

OCB

INFORMATION REPORT INFORMATION REPORT

CENTRAL INTELLIGENCE AGENCY

This material contains information affecting the National Defense of the United States within the meaning of the Espionage Laws, Title 18, U.S.C. Secs. 793 and 794, the transmission or revelation of which in any manner to an unauthorized person is prohibited by law.

S-E-C-R-E-T

COUNTRY USSR

REPORT

SUBJECT Soviet Technical Books **in the Field of Electronics**

DATE DISTR. 22 October 1958 25X1

NO. PAGES 2

REFERENCES RD

DATE OF INFO.

PLACE & DATE ACQ.

PROCESSING COPY

25X1

SOURCE EVALUATIONS ARE DEFINITIVE. APPRAISAL OF CONTENT IS TENTATIVE.

Soviet publications

25X1

- a. Signal (Signal), 403 pages, written by Igor Andreyevich Poletov and published by the Sovetskoye Radio Publishing House in Moscow in 1958. It is intended as a general reference on cybernetics for a wide reading public.
- b. Chastotnyye Metody v Avtomatike (Frequency Methods in Automation), 490 pages, represents a collection of foreign articles on the use of frequency methods in the solution of some problems of automation. It was published by the Foreign Literature Publishing House in Moscow in 1957.
- c. Teoriya Korrelyatsii i eye Primeneniye k Analizu Proizvodstva (The Correlation Theory and Its Application to Analysis of Productivity), 387 pages, written by Yakov Ilich Lukomskiy and published by the State Statistical Publishing House in Moscow in 1958. It expounds on the fundamentals of statistical studies utilized for a study of the inter-relationships of values, which are involved in the scientific analyses of experimental and computed data of various enterprises.
- d. Elektromekhanicheskiye Rele (Electromechanical Relays), 354 pages, written by Fayvel Aronovich Stupel, and published by the Kharkov State University Press, Kharkov, in 1956. It describes the fundamentals of the theory and engineering methods of computation, as well as the design of electromechanical relays and mechanisms having various principles of operation, e.g., electromagnetic, induction, magneto-electric, and others.
- e. Terminologiya Rele (Terminology for Relays), 42 pages, compiled by the Committee for Technical Terminology, Academy of Sciences, USSR, and published by the Academy of Sciences, USSR, in Moscow in 1958. It contains English, French, and German terminology, for recommended Soviet terminology.

Catalog

S-E-C-R-E-T

25X1

DATE	X ARMY	X NAVY	X AIR	X FBI	AEC				
Distribution by "#"									

INFORMATION REPORT

33 *W* *W*

S-E-C-R-E-T

-2-

25X1

- f. Elementy Elektroavtomatiki i Rascheta Samoletnykh Elektricheskikh Sistem (Elements of Electro-automation and Computation on Aircraft Electrical Systems), 118 pages, a collection of articles stemming from work done at the Moscow Order of Lenin Aviation Institute i/n S. Ordzhonikidze, and represents issue No. 85, of the "Trudy" of the Institute. It was published by the State Publishing House for the Defense Industry, in Moscow in 1957.
- g. Avtomatizatsiya Perevoda s Odnogo Yazyka na Drugoy (Automatic Translation of Languages), 34 pages, was written by D.Yu. Panov, A.A. Lyapunov, and I.S. Mukhin, and apparently represents the gist of a plenary meeting of the session on Scientific Problems of Industrial Automation. It was published by the Academy of Sciences in 1956.
- h. Opyty Avtomaticheskogo Perevoda na Elektrannoy Vychislitelnoy Mashine BESM (Experiment in Automatic Translation Utilizing the BESM Electronic Computer), 28 pages, written by I.S. Mukhin and published by the Academy of Sciences in Moscow in 1956. It briefly represents some of the procedures followed by D. Yu. Panov, et al., at the Institute of Precision Mechanics and Computer Engineering, beginning in January 1955 when they utilized BESM equipment to translate English into Russian.
- i. Certain Works in the Sphere of Computing Technique, 20 pages, is an English-language version of a pamphlet written by Academician S.A. Lebedev and published by the Academy of Sciences in 1956 on various aspects of the BESM installation. Some circuitry diagrams and photographs are given.

The above publications may be considered unclassified when detached from the covering report.

25X1

S-E-C-R-E-T

25X1

OCB

INFORMATION REPORT INFORMATION REPORT

CENTRAL INTELLIGENCE AGENCY

This material contains information affecting the National Defense of the United States within the meaning of the Espionage Laws, Title 18, U.S.C. Secs. 793 and 794, the transmission or revelation of which in any manner to an unauthorized person is prohibited by law.

S-E-C-R-E-T

25X1

COUNTRY	USSR	REPORT	[Redacted]
SUBJECT	Soviet Pamphlets on Electronic and Analytical Equipment	DATE DISTR.	15 May 1959
		NO. PAGES	3
		REFERENCES	RD

DATE OF INFO.	[Redacted]	25X1
PLACE & DATE ACQ.	[Redacted]	25X1

Reel # 717

SOURCE EVALUATIONS ARE DEFINITIVE. APPRAISAL OF CONTENT IS TENTATIVE.

brochures of Soviet equipment, and miscellaneous pamphlets

- ✓ 1. Automatic photo-colorimeter for nitrogen oxide determination (Russian).
2. Small brightness photometer (Russian, French, English, German).
- ✓ 3. Compressed air supply units for piezometric measuring equipment (Russian).
4. Vacuum diffraction spectrograph (Russian, English, French, German).
5. Quantometer, a photoelectric apparatus for emission spectral analysis (Russian, English, French, German).
- ✓ 6. Luminous spectrometer (Russian, French, English, German).
- ✓ 7. Diffraction spectrograph (Russian, French, English, German).
- ✓ 8. Equipment for measuring, registration, and automatic regulation of air pressure (Russian).
9. Electrochemical gas analyzer for sulphur gas (Russian).
- ✓ 10. Photoelectric steelometer (Russian, French, English, German).
- ✓ 11. Objective photoelectric color comparator (Russian, French, English, German).
12. Photometer for luminescent analysis (Russian, French, English, German).
13. Universal photometer (Russian, French, English, German).
- ✓ 14. High speed cinecamera (Russian, French, English, German).

STATE	X	ARMY	X	NAVY	X	AIR	X	FBI		AEC	X			
(Note: Washington distribution indicated by "X"; Field distribution by "#".)														

INFORMATION REPORT INFORMATION REPORT

S-E-C-R-E-T

-2-

25X1

25X1

15. Automatic gas analyzer (Russian).
16. Infra-red spectrometer (Russian, French, English, German).
17. Candle-power meter for measuring light flashes (Russian, French, English, German)
18. Quartz spectrograph (Russian, French, English, German).
19. Interferometer (Russian, French, English, German).
20. Electromagnetic valve (Russian).
21. Neutralization [acid/base] regulator (Russian).
22. Regulatory valve with pneumatic membrane and accessory mechanism (Russian).
23. Concentration meter for sulphuric acid and fuming sulphuric acid (Russian).
24. Magnetic gas analyzer (Russian).
25. Stem position regulator for use with a pneumatic valve (Russian).
26. Pneumatic indicator of liquid flow (Russian).
27. Slit liquid flow meter (Russian).
28. Visual spectral telescope (Russian, English, French, German).
29. Spectro-photometer (Russian, French, English, German).
30. High speed photo recording apparatus (Russian, French, English, German).
31. Spectrograph (Russian, English, French, German).
32. Spectroprojector (Russian, French, English, German).
33. Thermoconductometric gas analyzer (Russian).
34. Thermochemical gas analyzer (Russian).
35. Industrial [closed-circuit] TV equipment (Russian, French, English, German).
36. Used for determining hydrogen ion concentration with the aid of thick walled glass electrodes (Russian).
37. For investigating rapid exponential and non-exponential processes of decay and rise of luminescence and photo conductivity which are excited by periodical square light impulses produced by a mechanical light modulator.
38. developed by the Academy of Sciences, USSR (Russian, French, English, German).

S-E-C-R-E-T

S-E-C-R-E-T

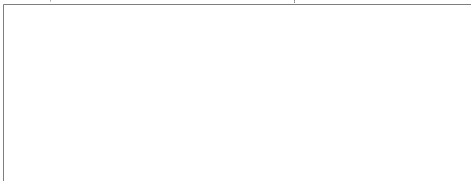
-3-



25X1

- ✓ 39. Diffraction gratings - Technical specifications for plane and concave gratings on aluminium layers, concave gratings on glass, and echelettes on metal specimens (Russian, German, French, English).
- 40. Spectroscopic investigations of the structure and composition of matter, 13 pages (English).
- 41. Regulator for programmed control of periodic chemical processes (Russian).
- 42. Short-wave solar radiation (Russian, French, English).
- 43. Self-limiting regulatory system for automatic search of a regime on the basis of given indicators, for use in the chemical industry (Russian).
- ✓ 44. The Peaceful Uses of Atomic Energy, a 31-page English-language propaganda pamphlet.
- 45. Automatic recording cloud ceilometer, developed by Dr. Ing. Frank Fruengel, GmbH Hamburg-Rissen (English).

The attachments may be treated as UNCLASSIFIED when detached from the covering report.



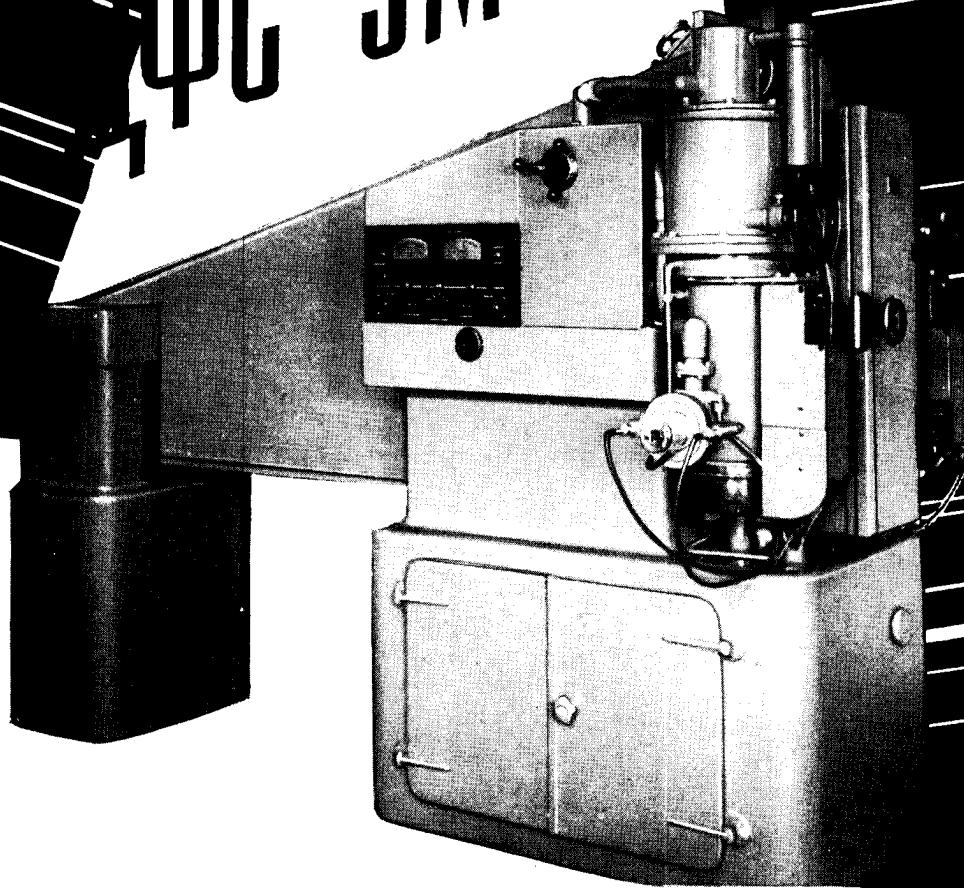
25X1

S-E-C-R-E-T



Вакуумная Электроника

ФЭС-5М



СЕКЦИЯ СССР НА ВСЕМИРНОЙ ВЫСТАВКЕ 1958 Г. В БРЮССЕЛЕ

SPECTROGRAPHE DE DIFFRACTION A VIDE

L'appareil est destiné à l'enregistrement photographique des spectres dans la région de l'ultra-violet à vide de 500 à 2000 Å. Il est construit d'après le schéma de l'incidence normale des rayons sur le réseau de diffraction.

L'élément dispersant: le réseau de diffraction de 100×60 mm, de 1200 traits per mm est porté sur une couche d'aluminium déposée sur le verre ayant le rayon de courbure de 3000 mm.

La dispersion linéaire inverse est près de 3 Å/mm.

Le châssis de 200×35 mm doit être chargé d'une pellicule souple.

Le châssis une fois chargé on peut photographier cinq spectres, de 3 mm de hauteur environ. Le châssis est réglé sans perturber le vide. La fente d'entrée est amovible et démontable.

La largeur de l'ouverture doit être réglée sous le microscope.

Entre la fente d'entrée et le corps du spectrographe est fixée une soupape à vide qui permet de régler la source et la fente sans perturber le vide dans l'appareil.

La source lumineuse est connectée à la bride de la fente à travers un joint étanche à vide démontable.

Le vide est de l'ordre de 10^{-5} mm Hg.

*Pour les conditions des livraisons adressez-vous:
« STANKOIMPORT », Moskva, G-200, Smolenskaja-
Sennaja ploschad, 32/34.*

VACUUM DIFFRACTION SPECTROGRAPH

The instrument is used for photographic recording of spectra in the region of vacuum ultra-violet from 500 to 2000 Å.

It is designed according to the diagram of normal beams incident on the diffraction grating.

The dispersion element is the grating of 1200 lines per mm, 100 × 60 mm; it is ruled on aluminium layer on glass (curvature radius – 3000 mm).

Reverse linear dispersion is about 3 Å/mm.

The adapter 200 × 35 mm is loaded with film on a flexible base.

5 spectra 3 mm high can be obtained without reloading the adapter and breaking the vacuum. The slit can be taken apart for cleaning. Its width is adjusted with a microscope.

Behind the slit a vacuum valve is provided whereby the light source and slit may be installed without breaking the vacuum.

The light source is connected to the flange of the slit through a dismountable vacuum seal.

Vacuum is about 10^{-5} mm of mercury.

*For delivery terms apply to "STANKOIMPORT",
Moskva, G-200, Smolenskaja- Sennaja ploschad, 32/34.*

VAKUUM-DIFFRAKTIONS-SPEKTROGRAPH

ДФС-5М

Das Gerät ДФС-5М ist für die photographische Registrierung von Spektren im Bereich von Vakuum-Ultraviolett von 500 Å bis 2000 Å bestimmt. Es ist nach dem Schema eines lotrechten Strahlen-Einfalls auf das Beugungsgitter aufgebaut.

Als Dispergator dient ein 100×60 mm = Gitter mit 1200 Teilchen/mm. Es ist auf einer Aluminiumschicht eingeschnitten, auf Glas aufgetragen war (Krümmungsradius 3000 mm).

Die reziproke lineare Dispersion beträgt ca. 3 Å/mm.

Die 200×35 -mm-Kassette wird mit einem Photomaterial auf biegsamer Unterlage geladen.

Ohne die Kassette umzuladen, kann man fünf Spektren mit einer Höhe von ca. 3 mm photographieren. Die Kassette wird mechanisch gesteuert, ohne daß das Vakuum gestört wird. Der Eintrittspalt ist einstellbar und zerlegbar. Die Öffnungsweite wird mit Hilfe eines Mikroskops eingestellt.

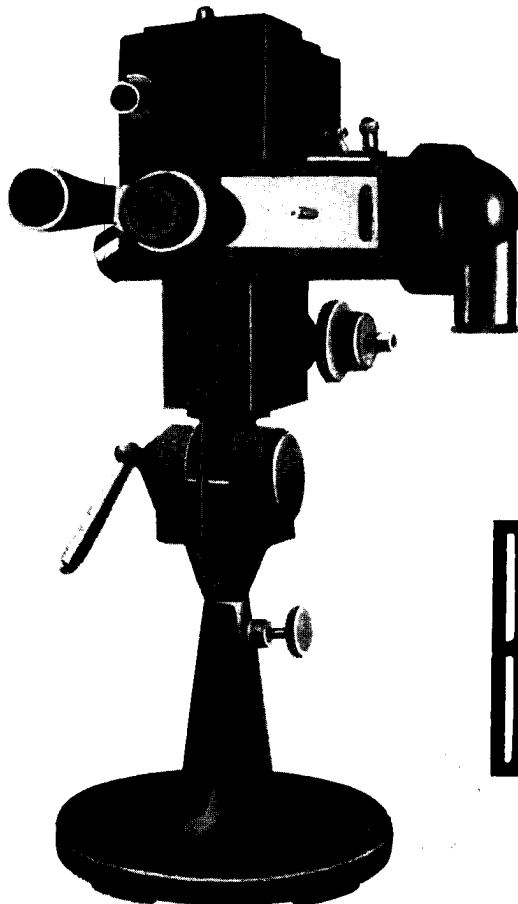
Zwischen dem Eintrittspalt und Spektrographen-Gehäuse ist ein Vakuum-Ventil angeordnet, das es gestattet, die Lichtquelle und den Spalt einzustellen, ohne daß das Vakuum im Spektrographen beeinträchtigt wird. Die Lichtquelle wird an einem Winkel des Spaltes durch die zerlegbare Vakuum-Dichtung verschlossen. Das Vakuum beträgt ca. 10^{-5} mm Q. S.

Alle Lieferungsanfragen sind an „STANKOIMPORT“,
Kosmonavtka, G-200, Smolenskaja-Sennaja Ploschad, 32/34 zu richten.

Sanitized Copy Approved for Release 2010/06/16 : CIA-RDP80T00246A045100330001-4

ФОТОМЕТР

МАЛЫХ ЯРКОСТЕЙ



ВФМ-57

СЕКЦИЯ СССР НА ВСЕМИРНОЙ ВЫСТАВКЕ 1958 Г. В БРЮССЕЛЕ

Sanitized Copy Approved for Release 2010/06/16 : CIA-RDP80T00246A045100330001-4

B(Д)M-57 est un photomètre visuel pour la détermination des brillances de différentes sources de lumière (luminophors, émanations lumineuses, matériaux, constitutions de lumière etc.).

L'instrument mesure: les brillances équivalentes (l'idée, proposée par le professeur A. A. Guerchun); brillances ordinaires (normalisées).

La construction de l'appareil prévoit les particularités de vue à l'adaptation aux brillances basses.

L'emmanchement télescopique permet de mesurer la brillance des naturels éclairages nocturnes, des installations de signalisations optiques, des couvertures de camouflage vertes etc.

Le champ photométrique – triple circulaire, dimension angulaire 28°. Pupille de sortie 8 mm. Le moindre champ visuel linéaire de l'instrument est 12 mm.

Trois vitesses de déplacement du coin photométrique assurent la même commodité et exactitude du mesurage de basses, moyennes et hautes brillances.

PHOTOMETRE DES BRILLANCES BASSES



Pour les conditions des livraisons
adressez-vous:
„STANKOIMPORT“, Moskva, G-200,
Smotenskaja-Sennaja Ploshchad, 32 34.

L'assortiment de l'appareil contient:
grand assemblage de filtres colorés pour la mesure
de la brillance des surfaces colorées;
filtres colorés neutres avec la courbe correctionnelle
spectrale de transmission;
l'absorbant tournant pour l'affaiblissement neutre de
la brillance mesurée;
les emmanchements pour le mesurage de la brillance
des luminophors et des surfaces écartées.

CARACTÉRISTIQUES TECHNIQUES

Limites de mesurage, apostilbes	
des brillances basses	0,0001–25
des brillances hautes	25–100 000
Erreur de mesurage, %	2–5
Encombrement, mm:	
de l'appareil	200 x 270 x 450
de l'installation alimentaire	170 x 170 x 230
de l'absorbant tournant	150 x 200 x 240
Poids, kg:	
de l'appareil	6
de l'installation alimentaire	3
de l'absorbant tournant	2,5

The BΦM-57 is a visual photometer designed for determining the brightness of various sources of light (luminiferous substances, luminiscent radiations, luminous materials, mixtures, etc).

The apparatus measures equivalent brightnesses (the notion was offered by prof. A. A. Gershun) and usual standard brightnesses.

The apparatus design has been made with taking into account specific features of vision in its adaptation to small brightnesses.

The telescopic headpiece makes it possible to measure the brightness of natural illumination of objects at night, the brightness of light signalling installations, the brightness of green masking covers, etc.

The photometric field is trippleannular; the angular size is 28°. The eye-piece outlet pupil is 8 mm in diameter. The least linear field of vision of the apparatus is 12 mm.

Three speeds at which photometric wedge is moving provide equal conveniences and accuracy of measurement of small, mean and large brightnesses.

The following accessories are attached to the apparatus:

- a numerous set of coloured light filters for measuring the brightness of coloured surfaces;
- neutral light filters with corrected curve of spectral passage;
- a rotary absorber for neutral weakening of the brightness measured;
- headpieces for measuring the brightness of luminiferous substances and of remoted surfaces.

SPECIFICATIONS

Limits of measurements, asb:	
weak brightnesses	0.0001-25
strong brightnesses	25-100,000
Measurement error, %	2-5
Overall size, mm:	
apparatus	200 x 270 x 450
feeding unit	170 x 170 x 230
rotary absorber	150 x 200 x 420
Weight, kg:	
apparatus	6
feeding unit	3
rotary absorber	2.5

PHOTOMETER OF SMALL BRIGHTNESSES

BΦM-57

For delivery terms apply to
"STANKOIMPORT", Moskva, G-200,
Smolenskaja-Sennaja Ploshchad, 32 34.

БФМ-57 ist ein Fotometer, der zu Helligkeitsmessungen verschiedenartiger Lichtquellen (Tageslichtlampen, lunimeszierende Materialien, leuchtende Zusammensetzungen und drg.) geeignet ist.

Das Gerät ergibt „equivalente Leuchtdichten“ (ein von Prof. Gerschun vorgeschlagener Begriff) und auch übliche (standartisierte) Leuchtdichten.

Die Ausführung des Gerätes trägt Rechnung der Seheigenschaften, die mit der Adaptation der Augen bei schwacher Beleuchtung verbunden ist.

Ein teleskopischer Ansatz dient zu Messungen der natürlichen Nachtbeleuchtungsstärken, der Helligkeit der Lichtsignalanlagen, des Leuchtens der grünen Tarnungsüberzüge und drg.

Das fotometrische Feld besteht aus drei Ringen, Winkegröße – 28°, Ausgangspupille 8 mm, das kleinste lineare Sehfeld beträgt 12 mm.

Drei Umstellungsgeschwindigkeiten des optischen Keilvorsatzes sichern das bequeme und genaue Messen der kleinen, mittleren und großen Lichtstärken.

FOTOMETER FÜR KLEINE LICHTSTÄRKEN

БФМ-57

Alle Lieferungsanfragen sind an
„STANKOIMPORT“, Moskau, G-200,
Smolenskaja Sennaja Ploshchad, 32/34,
zu richten.

Zum Gerät gehören:
ein voller Satz verschiedener **Farbenfilter** zur Messung
der Lichtstärke der farbigen Flächen;
ein **neutraler Lichtfilter** mit Korrigierung der Dis-
persiondurchlässigkeit;
ein **rotierender Lichtabsorber** zur neutralen Vermin-
derung der zu **bemessenen Leuchtdichte**;
Ansätze zum Messen der Lichtstärke der Tageslicht-
lampen und der entfernten Flächen.

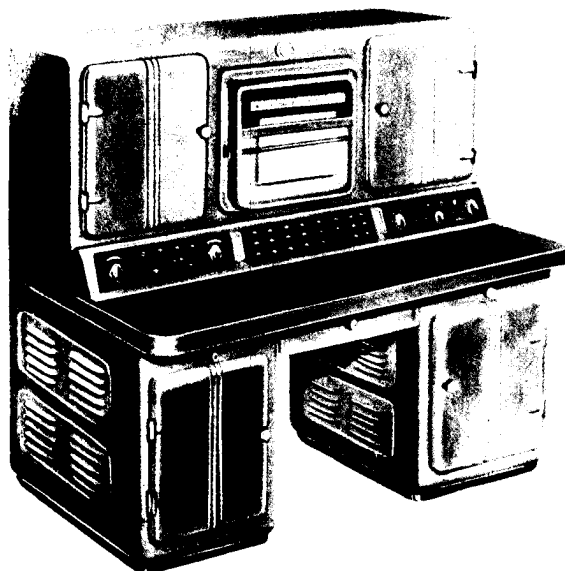
TECHNISCHE DATEN

Grenzen des Messbereiches, Apostilb:	
kleine Leuchtdichten	0,0001–25
große Leuchtdichten	25–100 000
Fehler in der Messung, %	2–5
Äußere Abmessungen, mm:	
des Gerätes	200 x 270 x 450
der Energiezufuhrvorrichtung	170 x 170 x 230
des rotierenden Absorbers	150 x 200 x 420
Gewicht, kg:	
des Gerätes	6
der Energiezufuhrvorrichtung	3
des rotierenden Absorbers	2,5



КВАНТОМЕТР
ДФС-10

СЕКЦИЯ СССР НА ВСЕМИРНОЙ ВЫСТАВКЕ 1958 г. В БРЮССЕЛЕ



INSTALLATION PHOTOELECTRIQUE POUR L'ANALYSE SPECTRAL D'EMISSION ДФС-10

L'installation photoélectrique ДФС-10 est destinée à l'analyse spectrale quantitative immédiate des métaux et des alliages.

L'équipement est réglable à 12 différents programmes analytiques et dans chaque programme on peut doser jusqu'à 11 éléments. La teneur en élément quelconque est déterminée par la mesure du logarithme du rapport du flux de la raie d'un élément analysé au flux de la raie de comparaison.

L'exécution de l'analyse est automatisée à partir du moment du branchement de la source lumineuse jusqu'à la lecture des résultats de mesure.

Pour le dosage de 11 éléments il ne faut que 2 minutes.

Encombrement, mm:	
appareil spectral	4500 x 1300 x 2500
dispositif électrique de mesure	1100 x 2100 x 2300
générateur	300 x 600 x 200
Poids, kg:	
appareil spectral	3000
dispositif électrique de mesure	500
générateur	160

Appareil spectral à réseau concave de diffraction (rayon de courbure 2 m, 1200 traits per mm, dispersion linéaire 4,2 Å/mm en ordre I-ier).

L'installation a 36 canaux récepteurs.

Lors du réglage de l'appareil, les fentes de sortie sont ajustées aux raies spectrales à analyser et sont fixées à demeure.

Comme récepteurs sont utilisées les cellules photoélectriques à vide à antimoine-césium ЦПБ-9 à fenêtre en uviol, qui assurent le fonctionnement de l'installation dans la région de 2200—5500 Å.

Le dispositif d'enregistrement et de réception a un canal de mesure: un voltmètre logarithmique à impulsion. Les condensateurs d'accumulation sont connectés en série à ce voltmètre à l'aide d'un schéma à relais. Le résultat de mesure est enregistré sur la bande du potentiomètre électronique ЭНН-09. Avec un réglage approprié de l'installation on peut enregistrer les résultats de l'analyse directement en pourcentage de la teneur en élément.

On utilise comme source lumineuse un générateur à commande électronique de l'impulsion d'amorçage ГЭУ-1. Le générateur assure une haute reproductibilité des conditions de la décharge dans les régimes de l'arc à courant alternatif, de l'étincelle de basse tension et du décharge d'impulsion.

*Pour les conditions des livraisons
adressez-vous: «STANKOIMPORT»,
Moskva, G-206, Smolenskaja-Semaja
ploubad, 32/34.*

PHOTOELECTRIC APPARATUS FOR EMISSION SPECTRAL ANALYSIS (QUANTOMETER) ДФС-10

The ДФС-10 photoelectric apparatus is designed for express quantitative spectral analysis of metals and alloys.

The installation is adjusted according to 12 various analytical programs, each program permitting determination of the contents of up to 11 elements. The content of an element is determined by measuring the logarithm of ratio of the radiant fluxes of lines of the element under investigation to the lines of the internal standard.

The analysis proceeds automatically from the moment the light source is switched on up to the time when results of measurements are obtained.

It takes 2 minutes to determine the composition of 11 elements.

The spectral instrument has a 2 m concave diffraction grating (1200 lines per mm, reverse linear dispersion 4.2 Å/mm in the first order).

The installation has 36 receiving channels. When adjusting the instrument, the exit slits are set on the spectral lines under study and are rigidly fixed.

Antimony-cesium ИСВ-9 vacuum photocells with windows of uvioi glass are used as

receivers, making it possible for the installation to operate in the region of 2200–5500 Å.

The receiving and recording arrangement has only one measuring channel — a logarithmic impulse voltmeter. The accumulator condensers are connected in series to it by means of a relay circuit. The result of the measurement is recorded on the strip of the ЭИИ-09 electronic potentiometer. Under a certain adjustment the results of analysis may be recorded directly in per cent composition of the element.

A ГЭВ-1 generator with electronic control is used as a source of light.

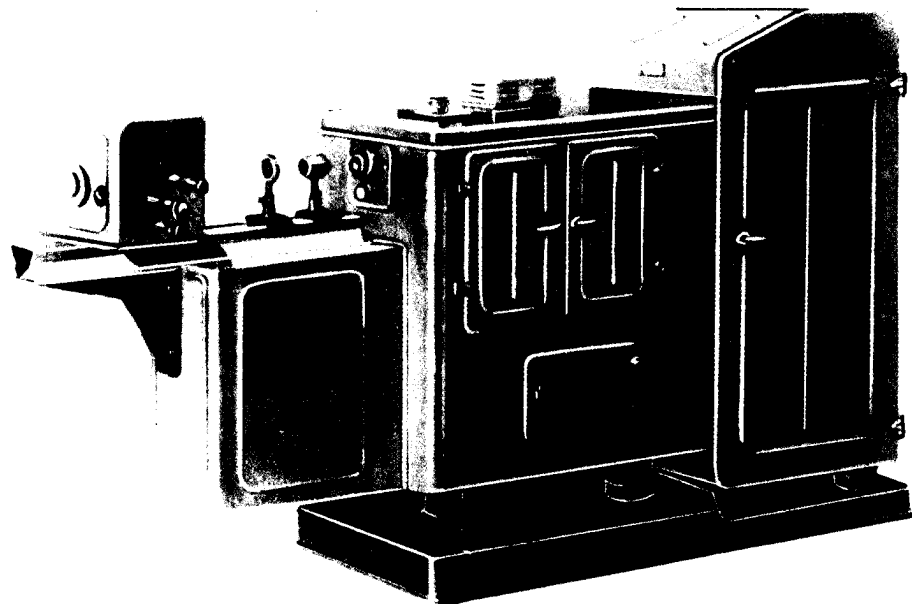
The generator ensures a high reproducibility of discharge conditions in a.c. arc, low-voltage spark and pulse discharge.

Overall dimensions, mm:

spectral instrument	4500 x 1300 x 2500
electromasuring arrangement	1100 x 2100 x 2300
generator	800 x 600 x 200

Weight, kg:

spectral instrument	3000
electromasuring arrangement	500
generator	160



For delivery terms apply to "STANKOIMPORT", Moskva, G-200, Smolenskaja-Sennaja ploshchad, 32/34.

PHOTOELEKTRISCHE ANLAGE FÜR DIE EMISSION SPEKTRALANALYSE ДФС-10

Die photoelektrische Anlage ДФС-10 ist für die quantitative Express-Spektalanalyse von Metallen und Legierungen bestimmt.

Die Anlage wird auf 12 verschiedene Analysenprogramme eingestellt, wobei in jedem Programm bis zu 11 Elemente bestimmt werden können. Das Gehalt an Elementen wird bestimmt durch die Messung vom Logarithmus des Verhältnisses der Licht-Ströme von der Linie des zu analysierenden Elements und der Linie des inneren Standards.

Der Analyseprozeß ist vom Augenblick der Einschaltung der Lichtquelle bis zum Empfang des Analyseergebnisses automatisiert.

Die Bestimmung des Gehalts an 11 Elementen erfolgt binnen 2 Minuten.

Das Spektralgerät besitzt ein konkaves Diffraktions-Gitter (Krümmungsradius 2 Meter, 1200 Str/mm, lineare Dispersion 4,2 Å/mm in der ersten Ordnung).

Die Anlage besitzt 36 Empfangskanäle.

Die Eintrittsspalten werden bei der Einstellung des Geräts auf die erforderlichen Spektrallinien eingestellt und starr festgelegt.

Als Empfänger werden Antimon-Cäsium-Photozellen ЦІВ-9 mit Uviol-Fenstern angewandt, die den Betrieb der Anlage im Bereich von 2200–5500 Å gewährleisten.

Die Empfangs-Registrieranlage besitzt einen Meßkanal – einen logarithmischen Impuls-Spannungsmesser. In Reihe damit werden durch eine Relais-schaltung Sammelkondensatoren geschaltet. Das Meßergebnis wird auf dem Papierstreifen des automatischen Potentiometers ЭПІІ-09 registriert.

Bei der entsprechenden Einstellung der Anlage können die Meßergebnisse der Analyse unmittelbar in Prozent-Gehalt des Elements registriert werden.

Als Lichtquelle wird in der Anlage ein Generator mit elektronischer Zündimpuls-Steuerung ГЭУ-І angewandt. Der Generator gewährleistet eine hohe Wiederholbarkeit der Entladungsverhältnisse in den Regimen: Wechselstromlichtbogen, Niederspannungsfunken und Impulsentladung.

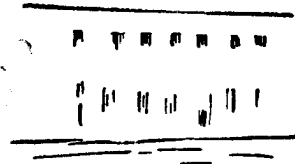
Außenmaße, mm:

Spektralapparat	4500 x 1300 x 2500
Elektrische Meßanlage	1100 x 2100 x 2300
Generator	800 x 600 x 200

Gewicht, kg:

Spektralapparat	3000
Elektrische Meßanlage	500
Generator	160

Alle Lieferungsanfragen sind an „STANKOIMPORT“,
Moskwa, G-200, Smolenskaja-Sennaja Ploschad, 32/34 zu
richten.



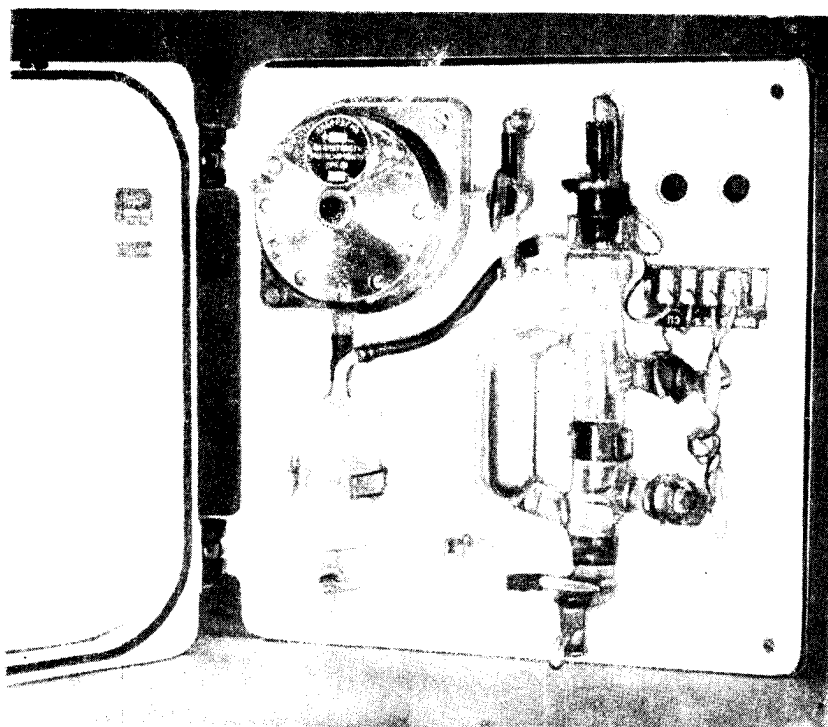
ЭЛЕКТРОХИМИЧЕСКИЙ ГАЗОАНАЛИЗАТОР

ТИПА

ЭХГ-3

НА СЕРНИСТЫЙ ГАЗ

ГОСУДАРСТВЕННЫЙ КОМИТЕТ
СОВЕТА МИНИСТРОВ СССР ПО ХИМИИ
МОСКВА



Внешний вид газоанализатора с открытой крышкой.

ЭЛЕКТРОХИМИЧЕСКИЙ ГАЗОАНАЛИЗАТОР типа ЭХГ-3 НА СЕРНИСТЫЙ ГАЗ

Электрoхимический газоанализатор типа ЭХГ-3 представляет собой стационарный автоматический прибор для непрерывного определения в газовых смесях малых концентраций сернистого ангидрида.

Действие газоанализатора основано на принципе электрохимической компенсации, заключающейся в непрерывном титровании сернистого ангидрида иодом, который электролитически выделяется из находящегося в приборе раствора, абсорбирующего анализируемый газ. При постоянном расходе газовой смеси показания прибора пропорциональны току электролиза; управление током производится по потенциалу платинового электрода, погруженного в раствор электролита.

В зависимости от условий применения, газоанализатор выпускается в трех модификациях:

Модификация газоанализатора	Назначение	Очистное устройство, поставляемое совместно с газоанализатором
ЭХГ-3А	Определение сернистого ангидрида в газовых смесях после контактирования (в производстве серной кислоты)	Две поглотительные колонки из нержавеющей стали. Предназначены для очистки газа от серного ангидрида поваренной солью
ЭХГ-3Б	Определение сернистого ангидрида в газовых смесях сероулавливающих установок ТЭЦ	Электрофильтр для очистки газа от мелкодисперсной золы
ЭХГ-3В	Определение сернистого ангидрида в газовых смесях прочих производств (см. п. 7 технической характеристики)	Очистные устройства не поставляются

В комплект поставки газоанализатора типа ЭХГ-3 входят:

- а) датчик;
- б) источник стабилизированного питания;
- в) вторичный прибор, изготовленный на базе электронного потенциометра типа ЭПД;
- г) очистное устройство (поставляется в зависимости от условий применения газоанализатора);
- д) водоструйный насос или воздушный эжектор (поставляется по отдельному заказу).

ТЕХНИЧЕСКАЯ ХАРАКТЕРИСТИКА

1. Шкала газоанализатора 0—0,5% SO₂ или 0—0,1% SO₂ (шкала 0—0,1% — для модификаций ЭХГ-3Б и ЭХГ-3В).
2. Основная погрешность показаний газоанализатора не превышает ±5% от максимального значения шкалы.
3. Питание от сети переменного тока напряжением 220 в ±10%; частота 50 гц ±1%; потребляемая мощность ~ 150 вт.
4. Время запаздывания начала реагирования и время установления показаний газоанализатора (без учета запаздывания за счет газоподводящей линии):

Модификация газоанализатора	Время запаздывания начала реагирования	Время установления показаний
ЭХГ-3А	6 мин.	10 мин.
ЭХГ-3Б	1,5 мин.	4 мин.
ЭХГ-3В	40 сек.	3 мин.

5. Расход газа для анализа 10—15 л/час.
6. Давление газа на входе в датчик должно быть постоянным и находится в пределах 800—1500 мм вод. ст.; температура газа от +10 до +40°С.
7. Газовая смесь не должна содержать компонентов и примесей, окисляющихся иодом.
8. Габариты и вес узлов прибора:

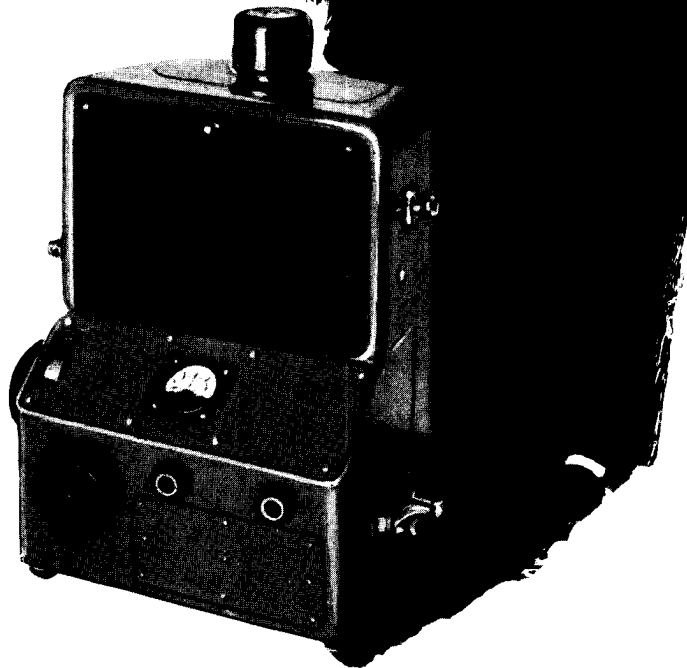
Узлы прибора	Длина мм	Ширина мм	Высота мм	Вес кг
Датчик	357	170	340	15
Стабилизированный источник питания	250	130	145	7,5
Вторичный прибор	420	292	506	29

Sanitized Copy Approved for Release 2010/06/16 : CIA-RDP80T00246A045100330001-4

ФОТОМЕТР

Для
ЛЮМИНЕСЦЕНТНОГО
анализа

ФМ-42



СЕКЦИЯ СССР НА ВСЕМИРНОЙ ВЫСТАВКЕ
1958 Г. В БРЮССЕЛЕ

Sanitized Copy Approved for Release 2010/06/16 : CIA-RDP80T00246A045100330001-4

PHOTOMÈTRE POUR L'ANALYSE LUMINESCENTE

ФМ-42 est un photomètre photoélectrique qui sert pour la détermination de la concentration des matières, capables de luminescer sous l'influence du rayonnement ultra-violet.

Les modèles recherchés – perles du diamètre de 3-4 mm ou disques du diamètre de 10-12 mm, d'épaisseur de 1-2 mm.

Le principe de fonctionnement est de mettre au même niveau les courants photoélectriques à force de changement de l'éclairage sur une des cellules photoélectriques à l'aide d'un diaphragme de fente variable.

L'instrument est construit d'après le schéma compensateur à deux rayons avec deux cellules photoélectriques. Les mesurages s'effectuent d'après la méthode de zéro. Au moment de mesurage les courants photoélectriques de deux cellules photoélectriques se compensent mutuellement et sur l'instrument-zéro sortant le signal est absent.

La source de lumière – la lampe à vapeur de mercure-quartz ЦВД-120. Le récepteur d'énergie lumineuse – cellules photoélectriques stibines-césium du type ЦИВ-3.

L'alimentation se produit du courant alternatif de 220 V, 50 hertz du réseau avec l'intermédiaire de la poulie d'alimentation.

La sensibilité de l'instrument est suffisante pour découvrir 10^{-10} g d'uranium dans le modèle, avec erreur de $\pm 5\%$.

Encombrement, mm:
du photomètre **250 × 300 × 400**
de la poulie d'alimentation **350 × 200 × 150**

Poids, kg:
du photomètre **20**
de la poulie d'alimentation **10**

Pour les conditions des livraisons adressez-vous : "STANKO-IMPORT", Moskva, G-200, Smolenskaja Sennaja ploshchad, 32/34.

PHOTOMETER FOR LUMINESCENT ANALYSIS

The ФМ-42 is a photoelectric photometer designed for determining the concentration of substances which are capable of glowing under the influence of ultra-violet rays.

The objects examined are pearls 3 to 4 mm in diameter or discs 10 to 12 mm in diameter, 1 to 2 mm thick.

The principle of action consists in levelling photoelectric currents by changing the intensity of illumination of one of photoelectric cells with the aid of an adjustable slot diaphragm.

The apparatus is designed according to a two-ray compensation diagram with two photoelectric cells. The measurements are made by zero method. At the moment of measurement photoelectric currents of two photoelectric cells are mutually compensated and no signal is received by the output zero-instrument.

The source of light is a ЦБД-120 mercury-quartz lamp. The energy of light is received by ЦЦБ-3 antimony-cesium photoelectric cells.

Power supply is provided from a. c. 220 V 50 cycles mains through a feeding unit.

The sensitivity of the apparatus is sufficient for revealing in the sample 10^{-10} g of uranium with an error of $\pm 5\%$.

Overall size, mm:

photometer **250 × 300 × 400**
feeding unit **350 × 200 × 150**

Weight, kg:

photometer **20**
feeding unit **10**

*For delivery terms apply to "STANKOIMPORT", Moskva,
G-200, Smolenskaja-Sennaja plashchad, 32/34.*

FOTOMETER FÜR LUMINESZENANALYSEN

ФМ-42 ist ein Fotometer zur Bestimmung der Konzentration des lumineszenzfähigen Stoffes unter Einwirkung der ultravioletten Bestrahlung.

Als Untersuchungsmuster werden Rundkörper (Perlen) von 3-4 mm oder Scheiben von 10-12 mm Durchmesser und 1-2 mm Stärke benützt.

Das Grundprinzip der Wirkung besteht in der Ausgleichung der Fotoströme mittels Beleuchtungsänderung einer der Fotoelemente mit Hilfe der regulierbaren Spaltblende.

Das Gerät ist entsprechend des zweistrahligten Kompensationsschemas mit zwei Fotoelementen aufgebaut. Die Messungen werden nach dem Nullverfahren geführt.

Ist die Ausgleichung der Fotoströme erreicht, so ist auf dem Ausgangsnullwerk kein Signal vorhanden.

Als Lichtquelle dient die Quecksilber-Quarzlampe (Höhensonne) Typ ЦБД-120. Als Empfänger der Lichtenergie sind zwei Antimon-Cäsium-Fotoelemente Typ ЦУБ-3 benützt.

Die Energiezufuhr geschieht vom Wechselstromnetz 220 V, 50 Hz.

Die Empfindlichkeit des Gerätes ist ausreichend, um in dem Muster das Vorhandensein von 10^{-10} g Uran mit Annäherung von $\pm 5\%$ festzustellen.

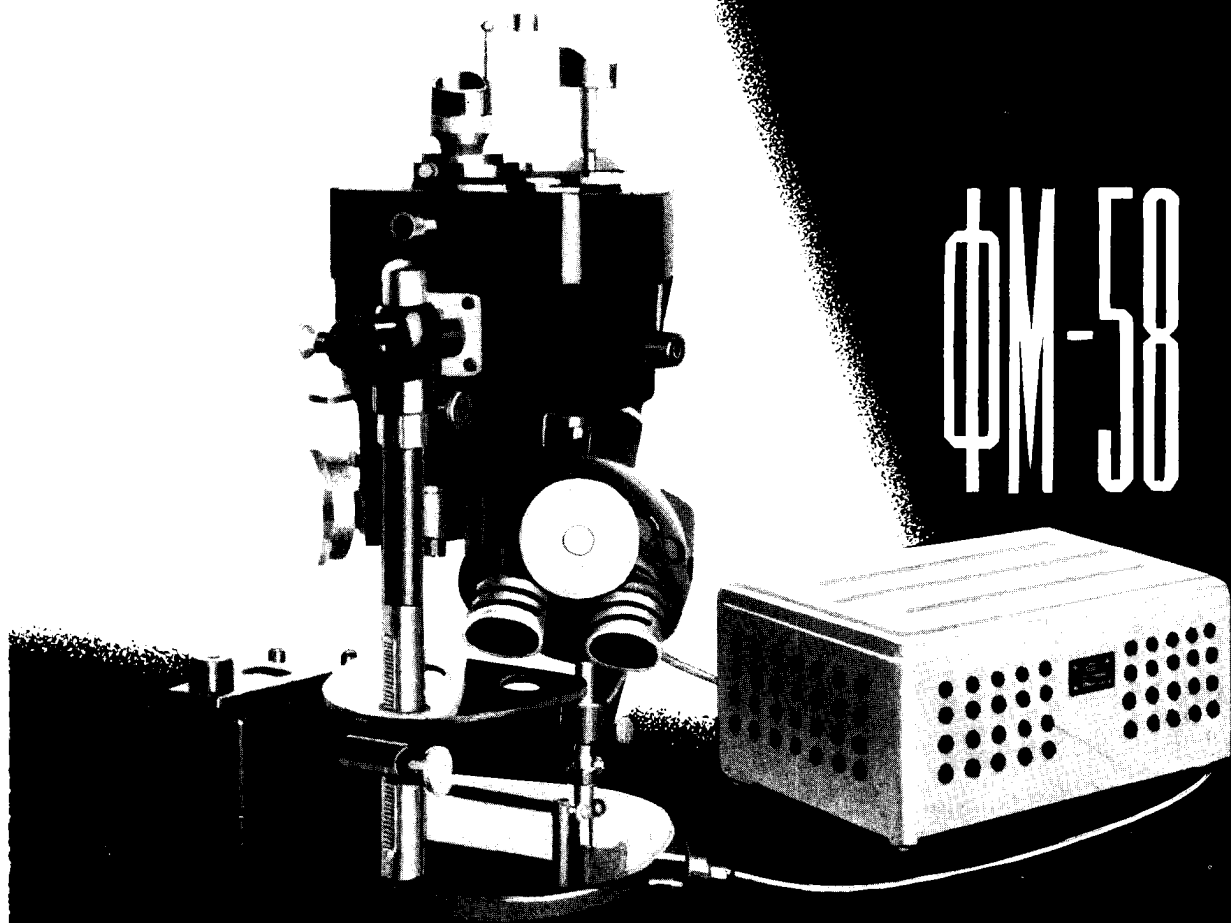
Äußere Abmessungen, mm:	
des Fotometers	250 × 300 × 400
der Energiespeisevorrichtung	350 × 200 × 150
Gewicht, kg:	
des Fotometers	20
der Energiespeisevorrichtung	10

Alle Lieferungsanfragen sind an „STANKOIMPORT“, Moskva,
G-200, Smolenskaja-Sennaja Ploschad, 32/34 zu richten.



ДИАЛЬНЬ
ОМЕТР

ФМ-58



ФМ-58 est un photomètre visuel-photoélectrique pour le mesurage du facteur de transmission (ou de la densité optique) des milieux durs et liquides transparents (non diffusés), pour le mesurage des coefficients de la brillance des modèles à lumière diffuse et leur éclat et pour les travaux spectrophotométriques. ФМ-58 peut être utilisé pour le mesurage des coefficients de réflexion et en qualité du microscope comparatif.

Le principe du fonctionnement de l'appareil — c'est mettre au même niveau les deux courants lumineux au moyen du changement de l'un d'eux à l'aide du diaphragme à l'ouverture alternative, l'image de laquelle se projette dans la pupille de l'observateur.

Les coefficients de la transparence et de la brillance plus de 0,1 se mesurent par la méthode photoélectrique (objective), moins 0,1 — visuellement.

Le passage des mesurages visuelles aux objectifs s'accomplit par un simple tour de la manivelle.

Deux cellules photoélectriques stibines-césium sont accouplées différenciellement.

Les mesurages par la méthode de zéro s'exécutent à l'aide des diaphragmes à l'ouverture alternative qui changent la grandeur des courants lumineux.

La tension stabilisée se maintient par un arrangement alimentaire spécial.

Les emmanchements microscopiques pour le mesurage des objets de petit format et l'utilisation du photomètre en qualité du microscope comparatif, les cuvettes pour les travaux colorimétriques, la petite table inclinée pour le mesurage des coefficients de brillance et de l'utilisation de l'appareil.

PHOTOMÈTRE UNIVERSEL

CARACTÉRISTIQUES TECHNIQUES

Limites du mesurage du coefficient de transparence, %	100-1
Le moindre coefficient mesuré de la brillance	0,01
La région des mesurages spectrophotométriques, m μ	400-720
L'erreur absolue des mesurages objectifs, %	1
... visuels, %	3-13
Encombrement	
de l'appareil	
d'arrangement alimentaire	315 x 210 x 141
Poids, kg:	
de l'appareil	14
d'arrangement alimentaire	9,5

Pour les conditions des livraisons adressez-vous: „STANKOIMPORT“, Moskva, G-200, Smolenskaja-Sennaja ploschad, 32/34.

UNIVERSAL PHOTOMETER

ΦM-58

The ΦM-58 is a visual photoelectric photometer for measuring transmission coefficients (or optical density) of solid and liquid transparent (not scattering) media, for measuring brightness coefficients of light-scattering specimens as well as for spectrophotometric analyses. The ΦM-58 can be used for measuring reflection factors and also as a comparing microscope.

The principle of action consists in levelling two light fluxes by changing one of them with the aid of a diaphragm with variable aperture whose image is projected to the observer's eye-pupil.

The transmission and brightness coefficients which are greater than 0.1 are measured photoelectrically, while those which are less than 0.1 — visually.

A change over from visual to photoelectrical measurements is effected by a simple turn of the lever.

Two antimony-cesium photoelectric cells are coupled to each other differentially.

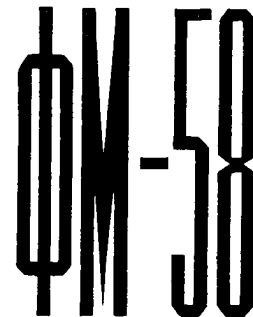
Measurements by zero method are made with the aid of diaphragms with variable aperture to change the value of light fluxes.

Stabilized voltage is maintained by a special feeding unit.

The microscopic headpieces for measuring small objects and for using the photometer as a comparing microscope, the baths for colorimetric analyses, and the tilting stage for measuring the brightness and brilliance coefficients give wider application to the apparatus.

SPECIFICATIONS

Limits of measurement of optical density, %	100-1
The least measurable brightness coefficient	0.01
Scope of spectrophotometric measurements, $m\mu$	400-720
Absolute error of photoelectric measurements, %	1
Relative error of visual measurements, %	1



apparatus	350 x 350 x 470
feeding unit	315 x 210 x 141
Weight, kg:	
apparatus	14
feeding unit	9.5

For delivery terms apply to
"STANKOIMPORT", Moskva, G-200,
Smolenskoja-Sennaja ploshchad,
32/34.

ФМ-58 ist ein kombiniertes visuelles und fotoelektrisches Fotometer, das zu Messungen des Lichtdurchlässigkeitswertes (oder optischer Dichte) der festen und flüssigen (nicht das Licht zerstreuernden) Körper, zur Bestimmung des Leuchtdichtheitswertes und der Brillanz der lichtausstrahlenden Muster (Lichtquellen), ebenso wie für spektralfotometrische Arbeiten geeignet ist. ФМ-58 kann auch für Messungen des Reflexionsvermögens und auch als Vergleichsmikroskop verwendet werden.

Das Grundprinzip der Wirkungsweise des Gerätes besteht in der optischen Ausglei- chung zweier Lichtströme, indem man eine mittels veränderlicher Öffnung versehene Blende verstellt, deren Abbildung in die Pupille des Beobachters projiziert wird.

Die Durchlässigkeits- und Leuchtdichtheitswerte, die größer als 0,1 sind, werden foto- elektrisch (objektiverweise) gemessen, welche unter 0,1 sind, visuell gemessen.

Der Übergang von visuellen Messungen zu objektiven Messungen geschieht mittels einfacher Umdrehung der Kurbel.

Zwei Antimon-Cäsium-Fotoelemente sind differentialartig verbunden.

Die Messungen nach Nullmethode werden mittels Blenden mit veränderlicher Öffnung, die die Größe der Lichtströme ändern, vollbracht.

Eine spezielle Energiequelle dient zur Zufuhr stabilisierter Spannung.

Mikroskopansätze, die für die Aufnahme kleiner Objekte und auch für Verwendung des Gerätes als Vergleichsmikroskop dienen sollen, Küvette für Farbmessungsarbeiten, geneigt einstellbare kleine Tische für Bestimmung der Leuchtdichten und Brillanzwerte erweitern das Anwendungsgebiet des Gerätes.

ФМ-58

UNIVERSALER FOTOMETER

TECHNISCHE DATEN

Meßgrenzen der Lichtdurchlässigkeitswerte, %	100 – 1
Kleinster Leuchtdichtemeßwert	0.01
Bereich der spektral-fotometrischen Messungen, <i>mμ</i>	400 – 720
Absoluter Fehler der fotoelektrischen Messun- gen, %	1
Relativer Fehler der visuellen Messungen, %	3 – 13
Außere Abmessungen, mm:	
des Gerätes	350 × 350 × 470
der Energiequelle	315 × 210 × 141
Gewicht, kg:	
des Gerätes	14
der Energiequelle	9,5

Alle Lieferungsanfragen sind an „STANKO-IMPORT“, Moskva, G-200, Smolenskaja-Sennaja Ploschad, 32/34 zu richten.

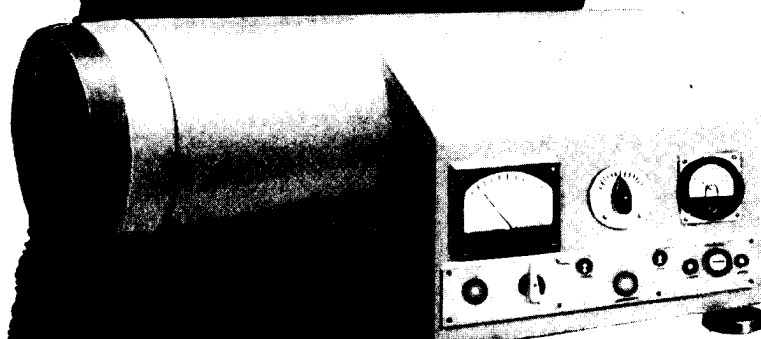
ABTEILUNG DER UDSSR AUF DER ALLGEMEINEN WELTAUSSTELLUNG IN BRÜSSEL 1958

Sanitized Copy Approved for Release 2010/06/16 : CIA-RDP80T00246A045100330001-4



Вечелер

ДЛЯ ИМПУЛЬСНЫХ ИСТОЧНИКОВ



СЕКЦИЯ СССР НА ВСЕМИРНОЙ ВЫСТАВКЕ 1958 Г. В БРЮССЕЛЕ

Sanitized Copy Approved for Release 2010/06/16 : CIA-RDP80T00246A045100330001-4

BOUGIEMETRE POUR LES SOURCES IMPULSIVES

Bougiemètre ИСМ-57 présente une combinaison du système optique télécentrique avec un photomètre photo-électrique.

Sert pour la détermination de bougiessecondes, c'est-à-dire du produit de la force de lumière des sources sur le temps de leur fonctionnement.

Peut être utilisé pour l'épreuve des lampes impulsives et lampes-flashes.

Qualités:

Indépendance des indications de la position de la source à mesurer;

indépendance du travail de l'appareil du moment de la mise en circuit de la source;

possibilité du travail dans un local à éclairage artificiel.



L'étendue des mesures, bougies	
à la durée du flash 0,02 sec.	100
à la durée du flash 50 μ sec	1×10^6
Dimension maximum du corps	
lumineux, mm	200
La tension du réseau électrique	
alimentaire du courant alternatif, V	220
Encombrement, mm	740 \times 395 \times 360
Poids, kg	20

Pour les conditions des livraisons adressez-vous: «STANKOIMPORT», Moskva, G-200, Smolenskaja-Sennaja Ploshchad, 32/34.

CANDLE POWER METER FOR MEASURING FLASHES OF LIGHT

The ICM-57 candle power meter is a combination of a telecentric optical system with a photoelectric photometer. It serves to determine candle-seconds, i. e., the product resulting from multiplication of the intensity of light of the sources by the time of their action.

It fits for testing flash lamps and bulbs.

Advantages:

Indications are independent of position of the source measured;

the apparatus operates independently of the moment at which the source is switched on;

the apparatus can operate in a room with artificial illumination.

The logo consists of the letters 'ICM-57' in a bold, stylized, sans-serif font. The letters are tall and narrow, with a slightly irregular, hand-drawn appearance. The 'I' and 'C' are connected at the top, and the 'M' is also connected to them. The '-57' is positioned to the right of the 'M', with a hyphen between them.

Measurement range, candles:

with 0.02 sec. duration of flash . . . **100**

with 50 μ sec. duration of flash . . . **1×10^6**

Maximum size of a flashing body

measured, mm. **200**

Voltage of a. c. feeding mains, V. 220

Overall size, mm. 740 \times 395 \times 360

Weight, kg. 20

For delivery terms apply to
"STANKOIMPORT", Moskva,
G-200, Smolenskaja-Sennaja
Ploshchad, 32/34.

LICHTKERZENMESSER FÜR IMPULSIONSBLITZQUELLEN

Der Lichtkerzenmesser ICM-57 ist ein kombiniertes Gerät, das aus einem telezentrischen Optiksystern und einem fotoelektrischen Fotometer besteht.

Das Gerät dient zur Bestimmung der „Kerzen-Sekunden“, d. h. des Produktes der Lichtquellenstärke und der Wirkungsdauer.

Das Gerät eignet sich zu Messungen der Impulslampen und einfachen Blitzlampen.

Vorzüge:

Unabhängigkeit der Anzeigen von der Lage der untersuchten Lichtquelle;

Unabhängigkeit der Wirkung des Gerätes von dem Einschaltmoment;

Möglichkeit, die Arbeiten im geschlossenen Raum bei künstlicher Beleuchtung zu erfüllen.



Bereich der Messungen, Kerzen:

bei einer Blitzdauer von 0,2 Sek.	100
bei einer Blitzdauer von 50 μ Sek.	1 · 10 ⁶
maximale Größe der untersuchten Lichtquelle, mm.....	200
Spannung im Speisewechselstromnetz, V	220
äußere Abmessungen, mm	740 × 395 × 360
Gewicht, kg	20

Alle Lieferungsanfragen sind an „STANKOIMPORT“, Moskva, G-200, Smolenskaja-Sennaja Ploschad, 32/34, zu richten.



ЭЛЕКТРОМАГНИТНЫЙ КЛАПАН ТИПА КЭ-2

МИНИСТЕРСТВО
ХИМИЧЕСКОЙ ПРОМЫШЛЕННОСТИ СССР
МОСКВА

ЭЛЕКТРОМАГНИТНЫЙ КЛАПАН типа КЭ-2

Электроманитный клапан типа КЭ-2 представляет собой воздухораспределительный трехходовой клапан с электромагнитным приводом, питаемым постоянным током.

Электроманитный клапан предназначен для включения или переключения пневматических линий с помощью постоянного тока и может быть использован для управления пневматическими клапанами в схемах блокировки, защиты и для различных переключений.

Принцип действия электромагнитного клапана заключается в том, что при отсутствии рабочего напряжения на обмотке электромагнита якорь закрывает один штуцер и открывает другой, соединенный с третьим штуцером. При подаче напряжения на обмотку электромагнита якорь перемещается, закрывая третий штуцер, и соединяет первый штуцер со вторым по внутренним каналам якоря.

Клапан работает на чистом и сухом воздухе при давлении до 2 кгс/см^2 .

ТЕХНИЧЕСКАЯ ХАРАКТЕРИСТИКА

1. Электромагнитный клапан питается от электросети постоянного напряжения $12 \text{ в} \pm 15\%$; $24 \text{ в} \pm 15\%$; $36 \text{ в} \pm 15\%$ (в зависимости от заказа).

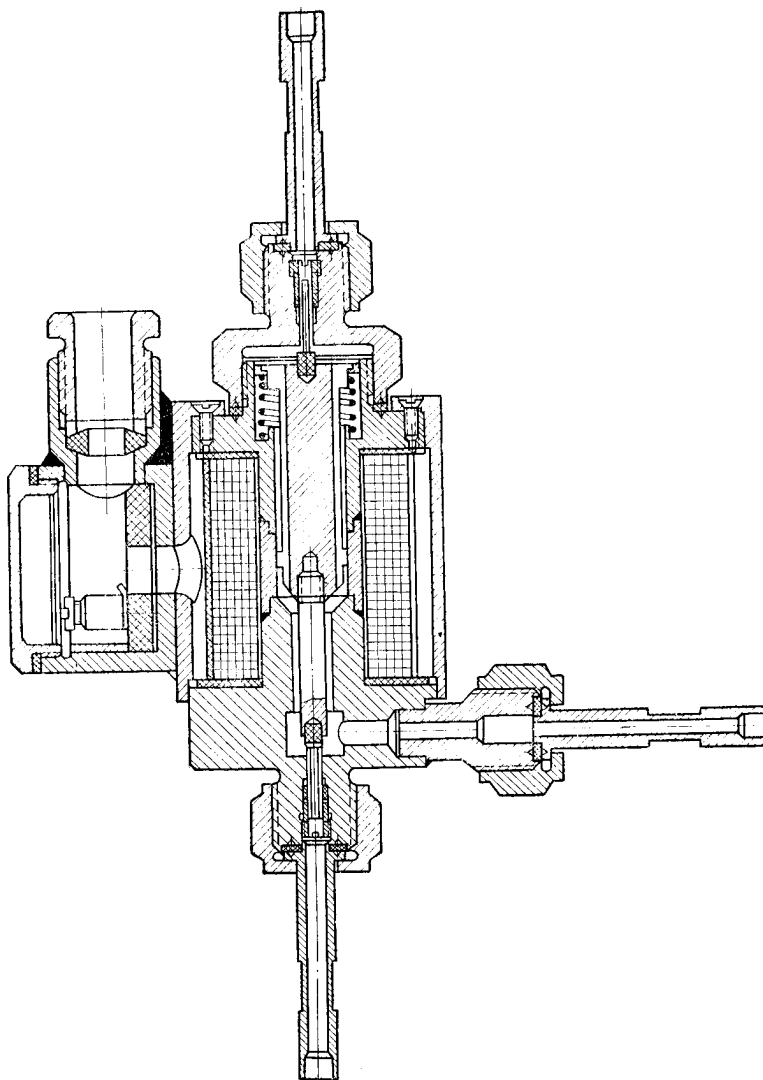
2. Потребляемая мощность 3 вт .

3. Расход воздуха, когда клапан находится под номинальным напряжением или обесточен, -- не менее 25 л/мин , при перепаде давлений на штуцере 2 кгс/см^2 .

4. Рабочий ход якоря 2 мм .

5. Габаритные размеры КЭ-2: $220 \times 150 \times 55 \text{ мм}$; вес $1,5 \text{ кг}$.

Электромагнитные клапаны изготавливаются Харьковским филиалом ОКБА (Харьков, 20, ул. Кибальчича, 10).

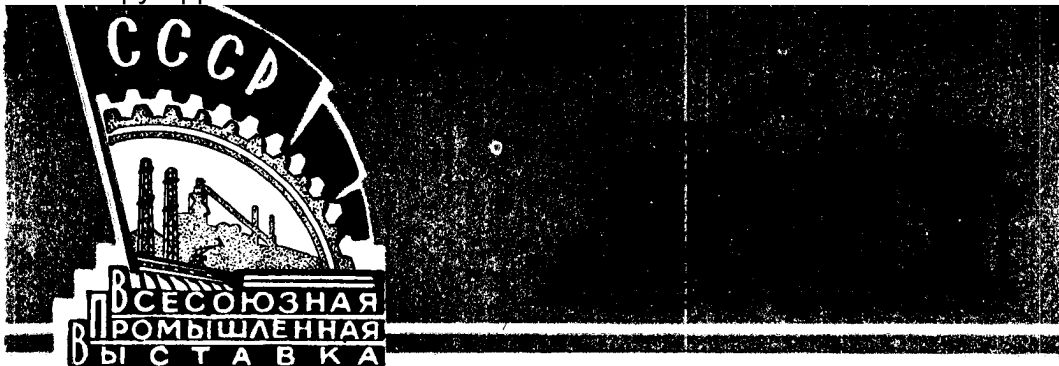


Sanitized Copy Approved for Release 2010/06/16 : CIA-RDP80T00246A045100330001-4

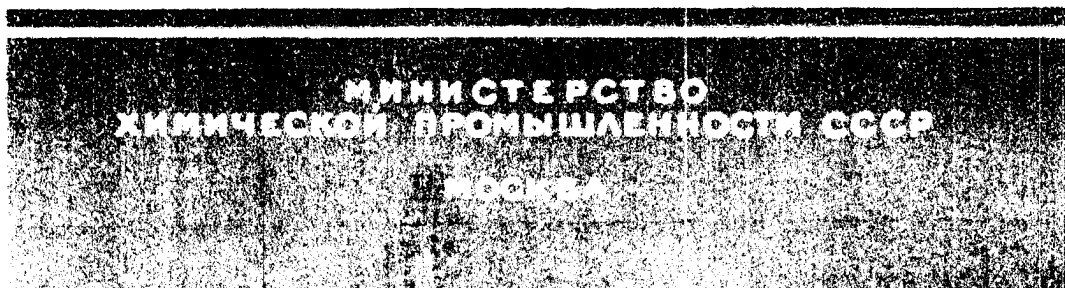
Г 04400 ГОСХИМИЗДАТ. № 60 Зак. 653

Типография Госхимиздата. Москва, 88, Угрешская

Sanitized Copy Approved for Release 2010/06/16 : CIA-RDP80T00246A045100330001-4

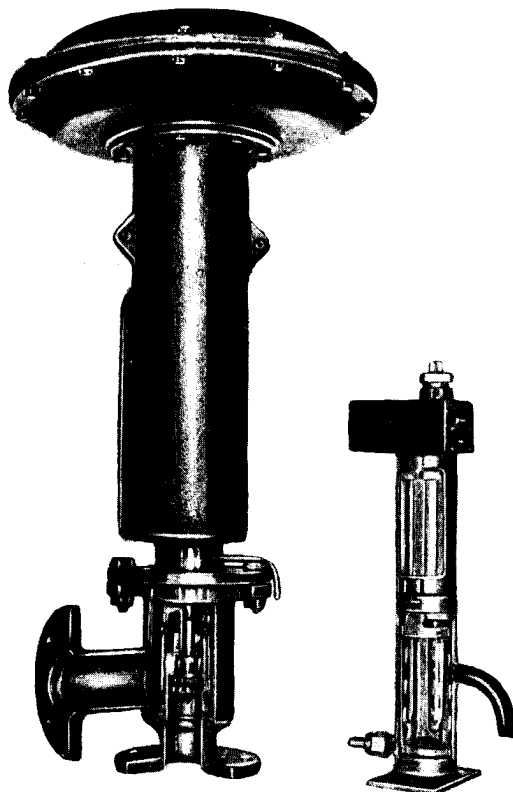


РЕГУЛЯТОР
НЕЙТРАЛИЗАЦИИ
ТИПА КРН



РЕГУЛЯТОР НЕЙТРАЛИЗАЦИИ типа КРН

Регулятор нейтрализации типа КРН представляет собой стационарный прибор для корректирования содержания избыточной кислоты или аммиака в щелоках из аппаратов ИТН (в процессе нейтрализации аммиака) при стабилизированном расходе аммиака и кислоты.



● Регулятор работает нормально только при отсутствии резких колебаний режима. В комплект регулятора входят:

- а) измерительная ячейка КРН;
- б) электронная приставка ЭПГ-2;
- в) электронный потенциометр ЭПД-32 на 45 мв;
- г) регулирующий клапан из нержавеющей стали с мембранным приводом типа КНМ-50с.

● Анализируемые щелока из аппарата ИТН проходят через ячейку

КРН. При этом на электродах ячейки возникает разность потенциалов, знак и величина которой зависят от того, насколько (и в какую сторону) изменилась величина кислотности щелоков или содержание аммиака в них. Разность потенциалов передается через электронную приставку на регулирующий электронный потенциометр типа ЭПД-32, который управляет мембранным клапаном на линии подачи кислоты

в аппарат ИТН (регулятором КРН можно регулировать процесс донейтрализации путем изменения подачи аммиака в донейтрализатор).

ИЗМЕРИТЕЛЬНАЯ ЯЧЕЙКА

- Измерительная ячейка КРН состоит из стального (ст. ЭЯ1-Т) корпуса, в котором находятся два платиновых электрода. Один электрод погружен в протекающие через ячейку анализируемые щелока, другой—в стеклянный сосуд со сравнительным раствором, содержащим требуемое количество кислоты или аммиака.
- Стеклянный сосуд со сравнительным раствором омывается протекающими через ячейку щелоками и отделен от последних стеклянным фильтром.
- При протекании через ячейку щелоков с таким же содержанием кислоты или аммиака, как в сравнительном растворе, потенциалы на обоих электродах будут равны и разность потенциалов на выходе из ячейки равна нулю. В противном случае возникает разность потенциалов, величина и знак которой зависят от того, насколько и в какую сторону кислотность измеряемых щелоков отклонилась от заданной величины.
- Для того чтобы уменьшить до минимума силу тока, отбираемую от ячейки, между ней и потенциометром включена электронная приставка ЭПГ-2. Благодаря включению приставки уменьшается поляризация электродов и предоставляется возможность использовать для регулирования стандартный потенциометр ЭПД-32 (нуль на середине шкалы).

ТЕХНИЧЕСКАЯ ХАРАКТЕРИСТИКА

- 1. Чувствительность регулятора устанавливается на месте (в зависимости от условий технологического процесса).
- 2. Расход щелоков через прибор равен 80 л/час.
- 3. Питание—от сети переменного тока напряжения 220 в $\pm 10\%$; частота 50 гц; потребляемая мощность 80 вт.
- 4. Давление сжатого воздуха для питания вторичного прибора составляет 2—6 кг/см². Воздух, подаваемый в регулятор, не должен содержать пыли, влаги и механических примесей.
- 5. Расход воздуха (максимальный) на регулятор 6—8 л/мин.
- 6. Длина соединительных труб от регулятора до исполнительного механизма—не более 50 м.

● 7. Габариты и вес узлов:

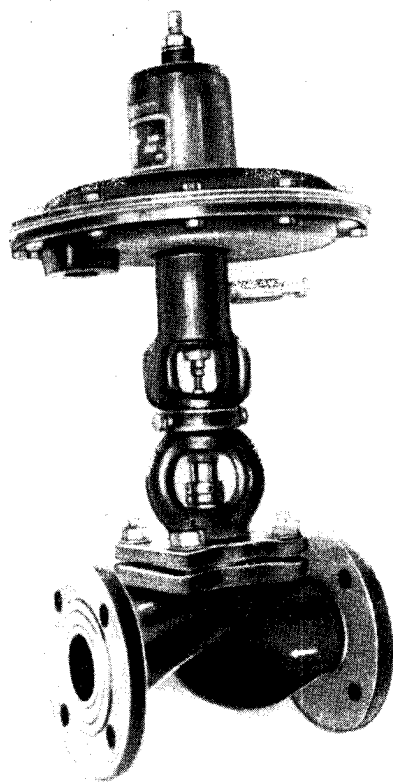
Узлы прибора	Диаметр мм	Длина мм	Ширина мм	Высота мм	Вес кг
Измерительная ячейка . . .	—	120	150	385	3
Электронная приставка . . .	—	275	130	216	6
Вторичный прибор ЭПД-32 . . .	—	420	292	506	35
Регулирующий клапан типа КНМ-50с	360	—	—	765	52



**КЛАПАН РЕГУЛИРУЮЩИЙ
С ПНЕВМАТИЧЕСКИМ МЕМБРАНЫМ
ИСПОЛНИТЕЛЬНЫМ МЕХАНИЗМОМ
ТИПА КРП**

**МИНИСТЕРСТВО
ХИМИЧЕСКОЙ ПРОМЫШЛЕННОСТИ СССР
МОСКВА**

Sanitized Copy Approved for Release 2010/06/16 : CIA-RDP80T00246A045100330001-4



Sanitized Copy Approved for Release 2010/06/16 : CIA-RDP80T00246A045100330001-4

КЛАПАН РЕГУЛИРУЮЩИЙ С ПНЕВМАТИЧЕСКИМ МЕМБРАННЫМ ИСПОЛНИТЕЛЬНЫМ МЕХАНИЗМОМ типа КРП

Регулирующий клапан типа КРП предназначен для использования в системах автоматического регулирования или автоматического управления потоками агрессивных жидкостей.

Принцип работы клапана основан на перемещении плунжера под воздействием давления воздуха в мембранном исполнительном механизме. Каждой величине давления воздуха соответствует определенное положение плунжера, а следовательно, определенная пропускная способность клапана.

Клапан типа КРП является односедельным. Все детали, соприкасающиеся с агрессивной жидкостью, изготовлены из стали ЭИ-629 и фторопласта-4.

Клапаны выпускаются с условным проходом $Dy-25; 50; 80$ и 100 мм в двух модификациях.

1. Нормально закрытый — «ВО» (воздух открывает).
2. Нормально открытый — «ВЗ» (воздух закрывает).

Клапаны типа КРП рассчитаны на условное давление среды $P_y = 2,5$ кгс/см².

ГАБАРИТНЫЕ РАЗМЕРЫ И ВЕС

D_n мм	Клапан «ВО»			Клапан «ВЗ»		
	Высота мм	Диаметр мм	Вес кг	Высота мм	Диаметр мм	Вес кг
25	485	230	18,6	440	230	17,7
50	590	290	31,0	510	290	31,0
80	815	360	51,5	680	360	51,0
100	860	360	61,5	725	360	61,0

Sanitized Copy Approved for Release 2010/06/16 : CIA-RDP80T00246A045100330001-4

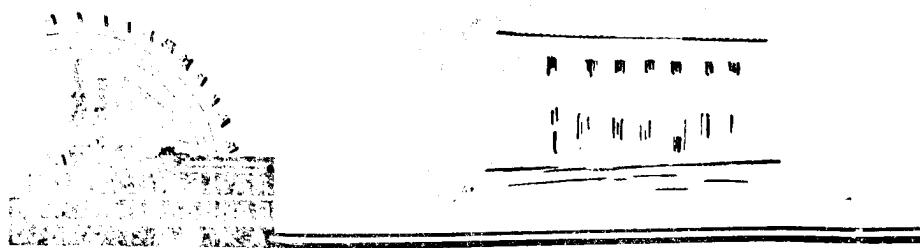
Т04400

ГОСХИМИЗДАТ. № 63

Зак. 651

Типография Госхимиздата. Москва, 88, Угрешская

Sanitized Copy Approved for Release 2010/06/16 : CIA-RDP80T00246A045100330001-4



ПОЗИЦИОНЕР ТИПА П-3

**ГОСУДАРСТВЕННЫЙ КОМИТЕТ
СОВЕТА МИНИСТРОВ СССР ПО ХИМИИ
МОСКВА**

ПОЗИЦИОНЕР типа П-3

Позиционер типа П-3 является пневматическим усилительным узлом, управляющим подачей сжатого воздуха в полость мембраны пневматического клапана для установки штока клапана в точном соответствии со значением входного сигнала регулятора.

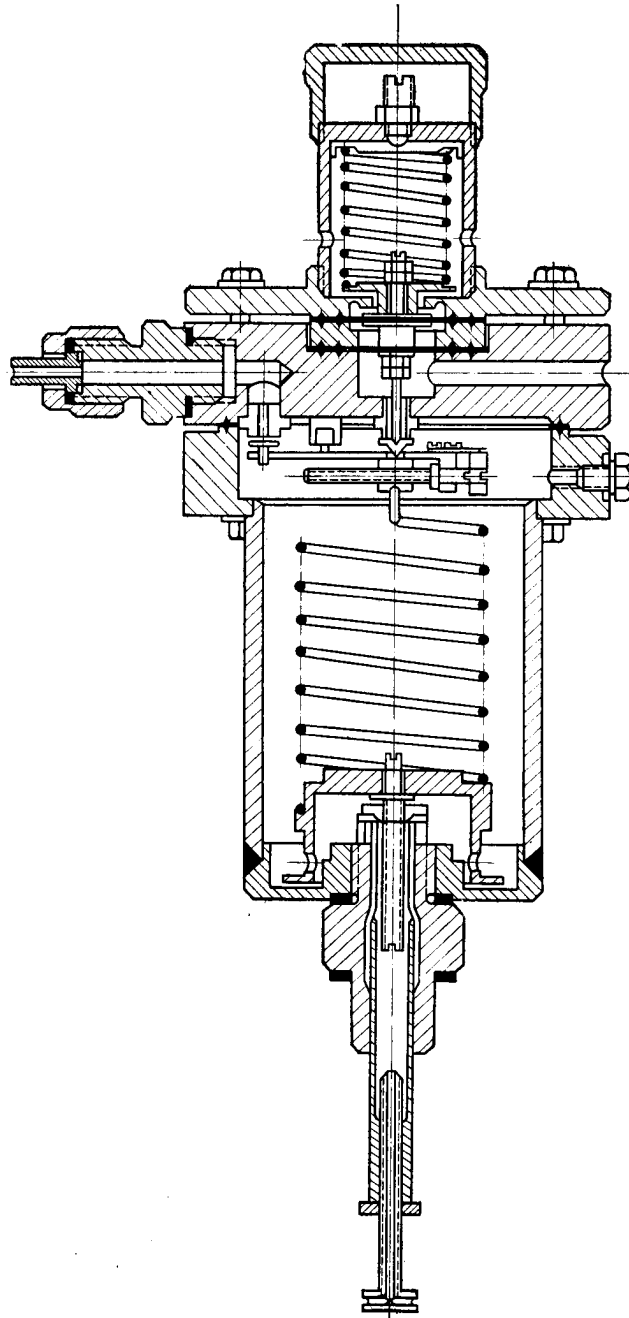
Позиционер предназначен для повышения быстрого действия мембранного клапана и усиления командного давления на мембрану пневматического клапана.

Позиционер устанавливается непосредственно на головке мембранного и исполнительного механизма клапана.

Позиционер П-3 является позиционером прямого действия, т. е. при повышении командного давления, давление в полости мембранного клапана также повышается и шток клапана перемещается вниз. При понижении командного давления шток клапана перемещается вверх. Определенному командному давлению соответствует определенное положение штока клапана.

ТЕХНИЧЕСКАЯ ХАРАКТЕРИСТИКА

1. Позиционер типа П-3 работает от стандартного пневматического регулятора.
2. Полному ходу штока клапана соответствует изменение командного давления от 0,2 до 1,0 *кгс/см²*.
3. Позиционер П-3 настраивается в зависимости от заказа на различный ход клапана в интервале от 7 до 34 *мм*.
4. Габариты: высота 300 *мм*; диаметр 165 *мм*; вес 5 *кг*.



Sanitized Copy Approved for Release 2010/06/16 : CIA-RDP80T00246A045100330001-4

К печати 10 VII 1958 г. Т04400 ГОСХИМИЗДАТ. № 62 Зак. 652
Типография Госхимиздата. Москва, 88, Угрешская

Sanitized Copy Approved for Release 2010/06/16 : CIA-RDP80T00246A045100330001-4

Самовосконт



А-12

STYLOSCOPE

СЛ-12

L'appareil est destiné à l'analyse spectrale visuelle qualitative et semi-quantitative rapide des aciers, des métaux non-ferreux et des alliages. L'appareil est utilisé pour le contrôle et le triage des matières premières aux dépôts, ateliers de dosage des minerais, aux fonderies, aux laboratoires pour les analyses immédiates, etc.

L'appareil comporte un instrument spectral, construit d'après le schéma d'auto-commutation, une platine porte-échantillon pour les échantillons à analyser et un générateur qui forment un seul ensemble.

La gamme de service de l'appareil est de 3700 Å à 7000 Å.

Grandissements foculaires: 20^x et 13,5^x.

Le dispositif optique comprend un coin photométrique pour le dosage des alliages de compositions semblables.

Il y a trois régimes stables de fonctionnement du générateur (régime à étincelles, régimes à arc (2 et 4A)).

Ensemble — 680×330×360 mm.

Poids — 48 kg.

Pour les conditions des livraisons adressez-vous:
«STANKOIMPORT», Moskva, G. 200, Sennaja Ploshhad, 32-34.

STEELSCOPE

СЛ-12

The instrument is designed for rapid visual qualitative and half-quantitative spectral analysis of steels, non-ferrous metals and alloys. It is used for sorting and checking of materials in storehouses and burden yards, in foundry express-laboratories, etc., and comprises a spectral apparatus built in accordance with autocollimation scheme, a stage for specimens and a generator, all being united in one unit.

Working range of instrument: 3700\AA – 7000\AA .

Magnification of eye-pieces: $20\times$ and $13.5\times$.

The ocular head has photometric wedge for sorting alloy grades similar in composition.

There are three constant operating conditions of the generator: a spark one and two arc ones (2 and 4A).

Overall dimensions – $680\times 330\times 360$ mm.

Weight – 48 kg.

Das Gerät ist für die schnelle visuelle qualitative und halbquantitative Spektralanalyse von Stählen, Buntmetallen und Legierungen. Es findet Anwendung für die Materialien-Kontrolle und -Sortierung in Lagern, Schichtböden, in Express-Labors der Gießereien und dgl.

Das Gerät besteht aus einem Spektral-Apparat, das nach dem Autokollimations-Schema aufgebaut ist, einem Probenisch und aus einem Generator, die konstruktiv zu einer Einheit zusammengeschlossen sind.

Arbeitsbericht des Gerätes 3700–7000 Å.

Vergrößerung der Okulare 20× und 13,5×.

In dem Okularkopf befindet sich ein photometrischer Keil für die Sortierung von Legierungsmarken, die nach ihrer Zusammensetzung einander nahe liegen.

Generator funktioniert in drei ständigen Arbeitsregimen: ein Funkenregime und zwei Lichtbogenregime (2 und 4A).

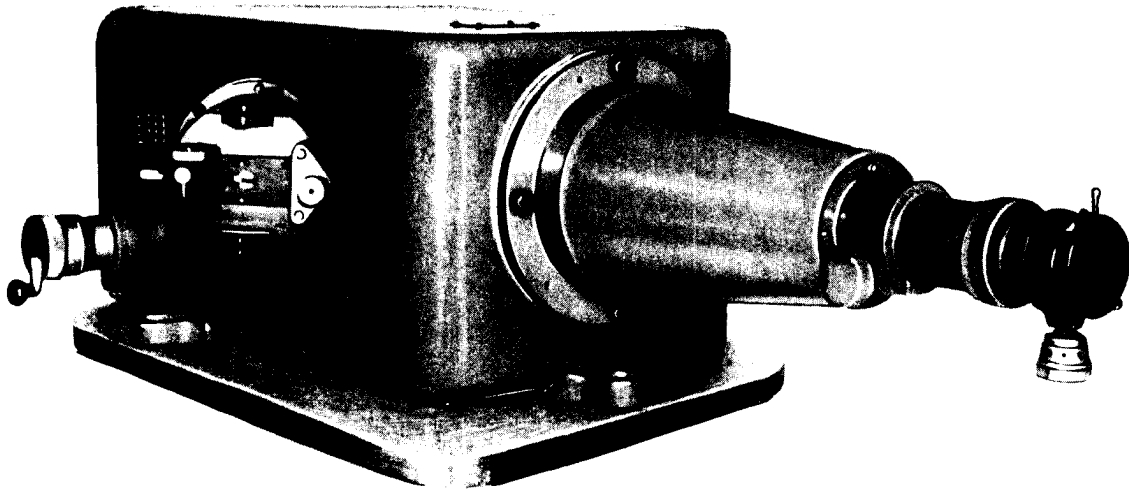
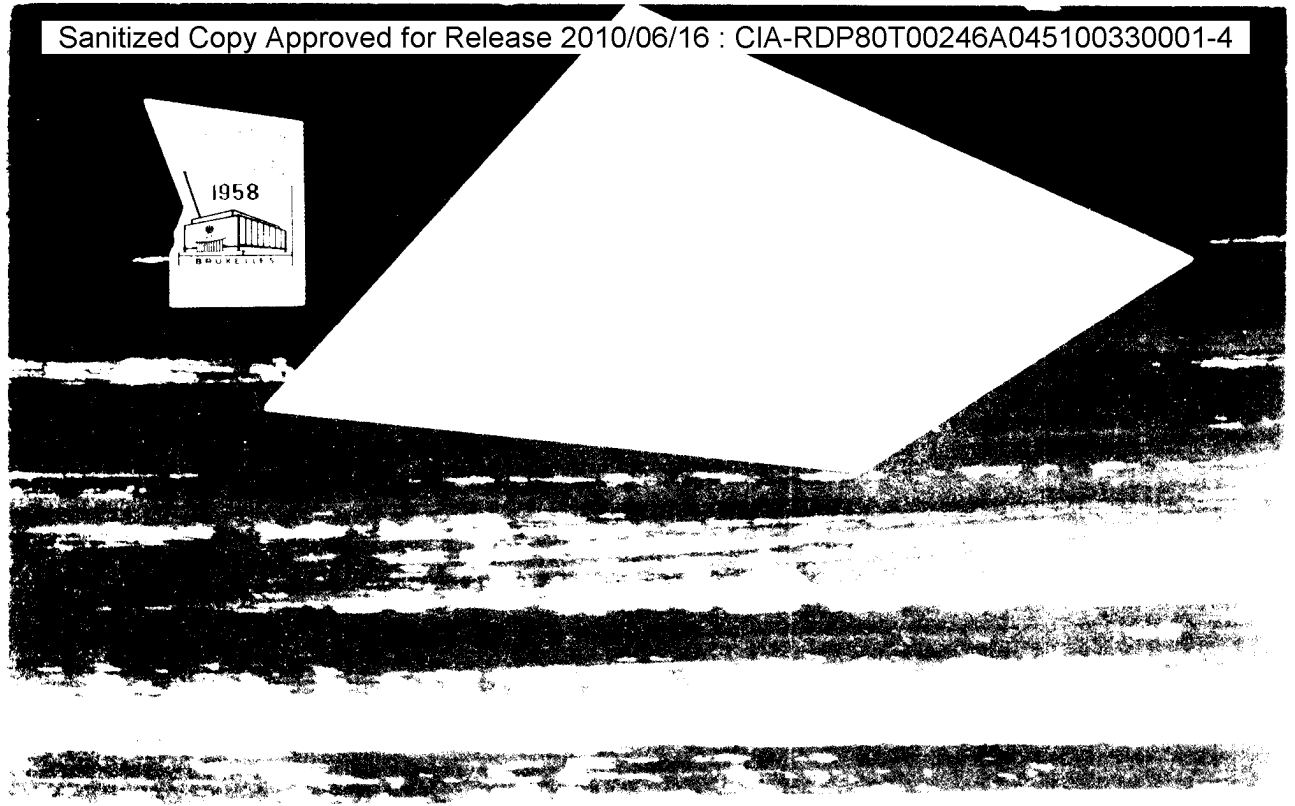
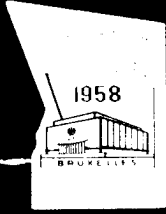
Äußere Abmessungen – 680×330×360 mm.

Gewicht – 48 kg.

STEELOSKOP

Alle Lieferungsanfragen sind an
„STANKOIMPORT“, Moskva, G-200,
Smolenskaja-Sennaja Ploschad,
32/34 zu richten.

ABTEILUNG DER UdSSR AUF DER ALLGEMEINEN Weltausstellung in BRUSSEL 1958



СП-48

СЕКЦИЯ СССР НА ВСЕМИРНОЙ ВЫСТАВКЕ 1958 Г. В БРЮССЕЛЕ

СП-48

SPECTROGRAPHE

Spectrographe lumineux СП-48 est spécialement conçu et construit pour les travaux effectués à l'occasion de l'Année Internationale Géophysique. Il est destiné pour les études des spectres de faibles luminescences dans la région visible de 3800 Å à 8000 Å (luminescence du ciel nocturne, des aurores boréales, etc.).

L'élément dispersant — une réplique prise du réseau de diffraction de 1200 traits/mm et fonctionnant en 1^{er} ordre.

La distance focale de l'objectif de caméra est 70 mm et l'ouverture relative est de 1:0,8. La dispersion linéaire inverse — 100 Å/mm.

Le spectre est photographié sur un film de 16 mm par parties, à quatre reprises.

L'appareil assure une haute qualité de l'image du spectre et un pouvoir résolvant élevé.

Pour les conditions des livraisons
adressez-vous:
V/O «STANKOIMPORT», Moskva G-200,
Smolenskaja-Sennaja ploschad, 32/34.

SPECTROGRAPH

СП-48

The large aperture spectrograph СП-48 was developed in connection with the International Geophysical year. It is designed for investigation of spectra of weak luminous objects in the region of 3800 to 8000 Å (light of the night sky, aurora, borealis, etc.).

The dispersion element is the replica of the diffraction grating of 1200 lines per mm operating in the first order.

The camera objective has a focal distance of 70 mm and a relative aperture of 1 : 0.8. Å reverse linear dispersion is 100 Å/mm.

The complete spectrum is photographed in 4 settings on a 16 mm film.

The instrument presents high quality of spectrum image and high resolving power.

For delivery terms apply to "STANKO-
IMPORT", Moskva G-200,
Smolenskaja-Sennaja ploschad, 32/34.

SPEKTROGRAPH СП-48

Für die Arbeiten, die mit der Durchführung des Internationalen Geophysischen Jahres verbunden sind, ist ein lichtstarker Spektrograph СП-48 entwickelt und hergestellt worden. Er ist für die Untersuchung der Spektren von schwach-leuchtenden Objekten im Bereich von 3800 bis zu 8000 Å (Leuchten des nächtlichen Himmels, des Nordlichts- usw.) bestimmt.

Das dispergierende Element ist eine Replik vom Beugungsgitter mit 1200 Str/mm, die in der 1. Ordnung arbeitet.

Das Objektiv der Kamera hat eine Brennweite 70 mm und eine relative Öffnung 1 : 0,8. Die reziproke lineare Dispersion beträgt cirka 100 Å/mm.

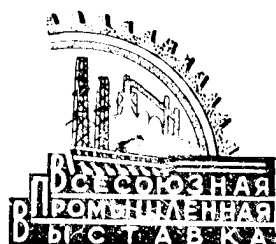
Das Spektrum wird in einzelnen Abschnitten auf einen 16 mm-Filmstreifen in 4 Einstellungen photographiert.

Das Gerät zeichnet sich durch gute Qualität der Spektrallinien und durch hohes Auflösungsvermögen aus.

Alle Lieferungsanfragen sind an V/O
„STANKOIMPORT“, Moskva, G-200,
Smolenskaja-Sennaja, Ploschad, 32/34
zu richten.

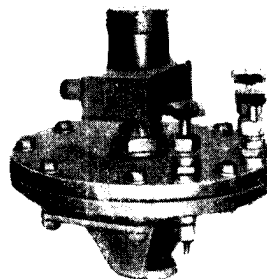
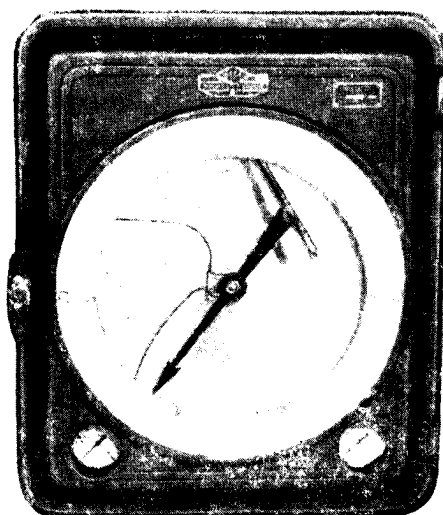
ABTEILUNG DER UdSSR AUF DER ALLGEMEINEN WELTAUSSTELLUNG IN BRÜSSEL 1958

Внешторгиздат. Заказ № Б-368-26



ИЗМЕРИТЕЛЬ ДАВЛЕНИЯ

ТИПА ЭМИД



ГОСУДАРСТВЕННЫЙ КОМИТЕТ
СОВЕТА МИНИСТРОВ СССР ПО ХИМИИ
МОСКВА

ИЗМЕРИТЕЛЬ ДАВЛЕНИЯ типа ЭМИД

Измеритель давления типа ЭМИД предназначен для измерения, записи и автоматического регулирования разности давлений (перепада), избыточного давления (напора) и разрежения (тяги).

Прибор может эффективно использоваться в системах с продувкой (в пьезометрических регуляторах плотности, концентрации, расхода и т. п.), так как вследствие малого переменного объема дифференциального манометра быстрое действие прибора достигается уже при небольшом количестве продуваемого через пьезометрические трубки воздуха или газа. Кроме того, при малом переменном объеме дифференциального манометра обеспечивается лучшая защита от агрессивного воздействия исследуемой среды.

Дифференциальный манометр с индукционным датчиком, имеющим условное обозначение «И», является искробезопасным прибором и может устанавливаться во взрывоопасных помещениях категории В-1А со средами: водородо-воздушной, эфира, пропана, ацетона, метана.

В комплект измерителя давления входят:

- а) мембранный дифференциальный манометр с индукционным датчиком, типа ДМ (конструкция ОКБА);
- б) электрический вторичный прибор типа ВЭП (конструкция ОКБА), изготовленный на базе электронного моста ЭМД или электронного потенциометра ЭПД.

Дифференциальный манометр состоит из двух металлических дисков, между которыми зажата резиновая мембрана, жесткий центр которой расположен между двумя измерительными пружинами. С жестким центром соединен железный сердечник, находящийся внутри индукционных катушек.

Плюсовое давление подается в верхнюю полость дифференциального манометра, минусовое — в нижнюю полость (через соответствующие запорные вентиля). Плюсовая и минусовая полости могут соединяться между собой посредством уравнительного вентиля.

При изменении перепада давления на мембране дифференциального манометра изменяется положение железного сердечника внутри индукционных катушек, в результате чего в диагонали измерительного моста возникает напряжение небаланса. Напряжение небаланса, усиленное через три каскада электронного усилителя напряжения и один каскад усилителя мощности, управляет реверсивным двигателем, который кинематически связан с движком реохорда.

Вращением движка реохорда измерительный мост приводится в равновесие, и в этот момент (равновесия) измеряемая величина отсчитывается по шкале прибора; запись измеренной величины непрерывно воспроизводится на дисковой диаграмме.

Измеритель давления выпускается в следующих модификациях:

- а) показывающий и записывающий;
- б) показывающий и записывающий, с трехпозиционным электрическим регулирующим устройством;
- в) показывающий и записывающий, с приспособлением для пневматического изохромного регулирования.

ТЕХНИЧЕСКАЯ ХАРАКТЕРИСТИКА

1. Основная погрешность в показаниях прибора при температуре $20^{\circ} \pm 5^{\circ}\text{C}$ составляет $\pm 2,0\%$ от предельного перепада давлений.

2. Условное статическое давление 1 кгс/см^2 .

3. Рабочий перепад давления: 10; 12,5; 16; 20; 25; 32; 40; 50; 60; 80; 100; 120; 160; 200; 250; 320 и 400 мм вод. ст. (определяется жесткостью измерительных пружин, установленных в дифференциальном манометре). По специальному заказу изготавливаются дифференциальные манометры и на другие рабочие перепады (в пределах от 10 до 400 мм вод. ст.).

Приборы, поставляемые комплектно с регуляторами плотности типа ПМ, имеют шкалу, градуированную в единицах плотности (г/см^3) или в единицах концентрации или крепости, а в комплекте со щелевыми регуляторами расхода — шкалу в весовых единицах расхода (т/час).

4. Рабочий ход мембраны дифференциального манометра составляет 1 мм. Переменный объем дифференциального манометра равен 30 см^3 .

5. Протяженность электрических проводов линии связи между индукционным датчиком дифференциального манометра и вторичным прибором — не более 500 мм. Линия связи — пятипроводная (в двух стальных трубах); сечение проводов $2,5 \text{ мм}^2$.

6. Питание прибора — от сети переменного тока напряжением 127 или 220 в (в зависимости от заказа); частота 50 гц; потребляемая мощность 60 вт.

7. Питание индукционного датчика на напряжении 0,65 в осуществляется от специального понижающего трансформатора, что предотвращает подачу сетевого напряжения на обмотки индукционного датчика. Этот трансформатор размещен внутри прибора ВЭП.

8. Минимальное время прохождения стрелкой вторичного прибора всей шкалы при мгновенном изменении давления составляет не более 30 сек. При работе регулятора в комплекте с пьезометрическими датчиками (типа КМ, ДПМ или РМ), через пьезометрические трубки которых продувается инертный газ в количестве 10—15 *см³/мин*, наименьшее время прохождения стрелкой всей шкалы составляет около 1¹/₂—3 мин.

9. Габариты и вес мембранного дифференциального манометра типа ДМ: диаметр 275 мм; высота 305 мм; вес 20 кг. Габариты и вес вторичного прибора типа ВЭП: 420×292×506 мм; вес от 29 до 35 кг (в зависимости от модификации).

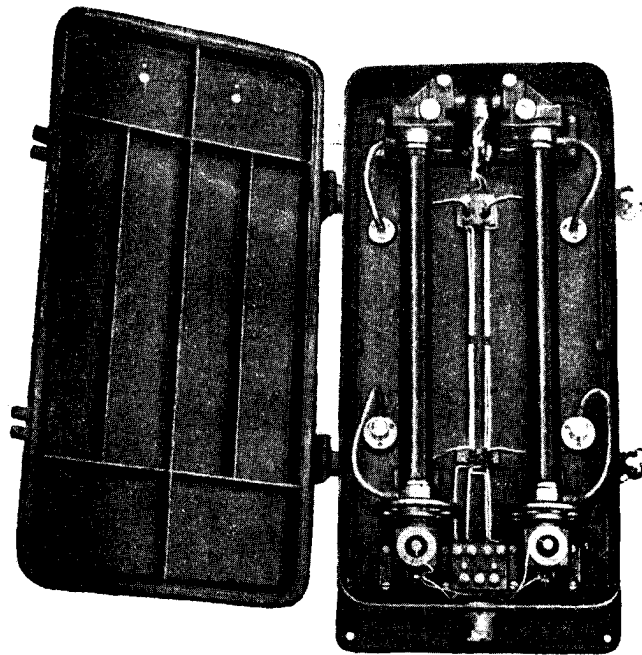
Измерители давления изготавливаются ОКБЛ (Москва, И-128, Сельскохозяйственная, 20).



**АВТОМАТИЧЕСКИЙ
ФОТОКОЛОРИМЕТР
ТИПА АФК-3
НА ОКИСЛЫ АЗОТА**

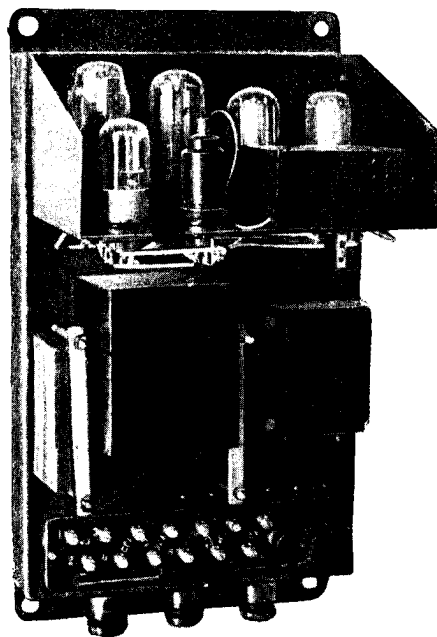
**МИНИСТЕРСТВО
ХИМИЧЕСКОЙ ПРОМЫШЛЕННОСТИ СССР
МОСКВА**

АВТОМАТИЧЕСКИЙ ФОТОКОЛОРИМЕТР
ТИПА АФК-3 НА ОКИСЛЫ АЗОТА



Автоматический фотоэлектрический колориметр типа АФК-3 предназначен для отдельного фотоколориметрического определения окислов азота (NO_2 и NO) в выхлопных газах башенного производства серной кислоты, производства азотной кислоты, а также и в других неокрашенных газах, если концентрация окислов азота не превышает 1%.

- Действие прибора основано на измерении прозрачности анализируемого газа при изменении в нем концентрации окрашенной двуокиси азота.
- Анализируемый газ должен проходить через первую колориметрическую трубку с такой скоростью, чтобы окись азота не успела окислиться до двуокиси азота, так как по первой трубке контролируется концентрация двуокиси азота. Наоборот,



Стабилизатор-усилитель.

через вторую трубку, по которой контролируется сумма NO_2 и NO , газ проходит с меньшей скоростью; за время прохождения (5—7 мин.) газа через окислительный объем окись азота успевает окислиться до двуокиси азота. Колориметрические трубки и трубки окислительного объема, во избежание образования в них тумана азотной кислоты, подогреваются до 50—60 С.

- Скорость проходящего газа через трубки поддерживается постоянной при помощи маностата.
- Лампа оптического устройства прибора питается от источника с постоянным стабилизированным напряжением на выходе.
- Шкала вторичного прибора градуируется в % NO_2 (в диапазоне от 0 до 1%). Максимальная погрешность в показаниях прибора составляет $\pm 10\%$ от диапазона шкалы.

- В комплект прибора входят:
 1. Датчик прибора.
 2. Стабилизатор-усилитель.
 3. Электрофильтры (2 шт.).
 4. Выпрямитель для электрофильтра.
 5. Окислительная емкость.
 6. Вторичный прибор.
- В качестве вторичного прибора могут быть использованы электронные потенциометры ЭПД, ЭПП-09 или указывающий милливольтметр МРЩПр-54 с соответствующим делителем на входе.

ОСНОВНЫЕ ТЕХНИЧЕСКИЕ ДАННЫЕ

- 1. Питание—от сети напряжением 220 в; потребляемая мощность 250 вт.
- 2. Давление на входе в прибор равно 100 мм вод. ст.
- 3. Расход газа составляет 75—100 л/час.
- 4. Постоянная времени прибора (без учета влияния газоподводящей линии и фильтра): для NO_2 равна 30 сек., для суммы ($\text{NO}_2 + \text{NO}$) составляет 5 мин.
- 5. Поступающая в прибор смесь должна быть очищена от пыли, тумана серной кислоты и влаги.

БЛОКИ ПИТАНИЯ ВОЗДУХОМ ТИПА БПВщ

Блоки типа БПВщ предназначены для питания пьезометрических измерителей сжатым воздухом и представляют собой соединение редуктора воздуха, манометра и контрольного стаканчика на одной общей панели. Давление воздуха перед редуктором может быть от 2 до 10 кг/см².

- Для предотвращения обратного переброса жидкости в воздушные коммуникации перед контрольными стаканчиками установлены ловушки.
- В зависимости от назначения блоки питания могут иметь разное число выходов:
- Блок БПВщ-1 один выход (один контрольный стаканчик).

Габариты блока: 250 × 158 × 160 мм; вес 3 кг.

- Блок БПВщ-2 два выхода (два контрольных стаканчика).

Габариты блока: 250 × 158 × 160 мм; вес 3,6 кг.

- Блок БПВщ-3 три выхода (три контрольных стаканчика).

Габариты блока: 250 × 230 × 160 мм; вес 4,8 кг.

- Производство блоков питания освоено на заводе треста «Химэнергоавтоматика» (г. Харьков, 20, ул. Кибальнича, 10).



БЛОКИ ПИТАНИЯ ВОЗДУХОМ ТИПА БПВЩ

БЛОКИ ПИТАНИЯ ВОЗДУХОМ ТИПА БПВщ

Блоки типа БПВщ предназначены для питания пьезометрических измерителей сжатым воздухом и представляют собой соединение редуктора воздуха, манометра и контрольного стаканчика на одной общей панели. Давление воздуха перед редуктором может быть от 2 до 10 $кг/см^2$.

- Для предотвращения обратного переброса жидкости в воздушные коммуникации перед контрольными стаканчиками установлены ловушки.

- В зависимости от назначения блоки питания могут иметь разное число выходов:

- Блок БПВщ-1—один выход (один контрольный стаканчик).

Габариты блока: 250×158×160 мм; вес 3 кг.

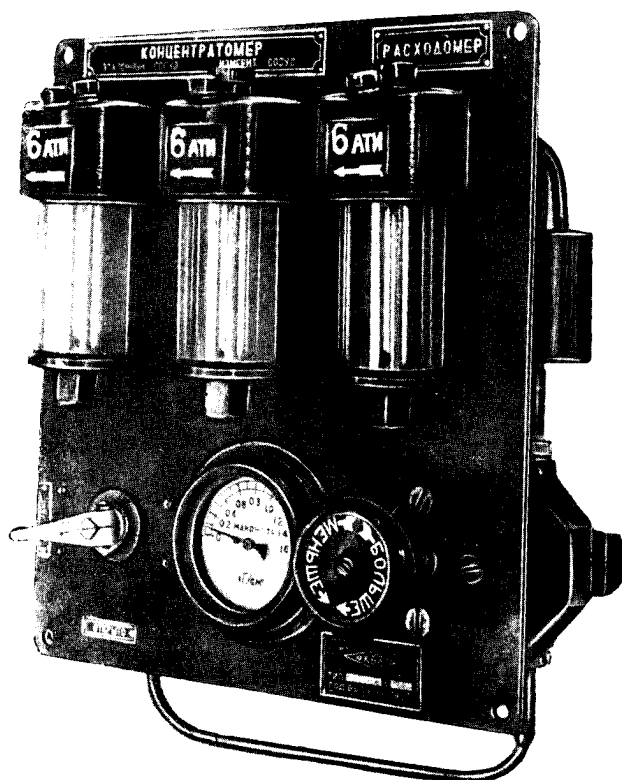
- Блок БПВщ-2—два выхода (два контрольных стаканчика).

Габариты блока: 250×158×160 мм; вес 3,6 кг.

- Блок БПВщ-3—три выхода (три контрольных стаканчика).

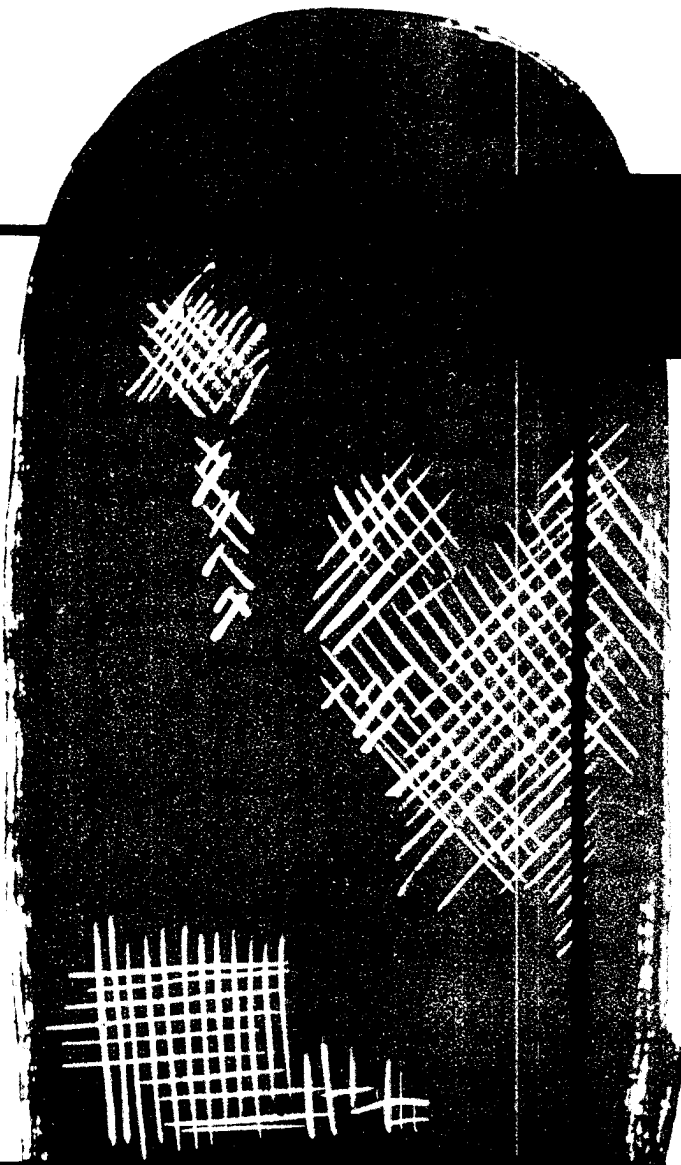
Габариты блока: 250×230×160 мм; вес 4,8 кг.

- Производство блоков питания освоено на заводе треста «Химэнергоавтоматика» (г. Харьков, 20, ул. Кибальчича, 10).



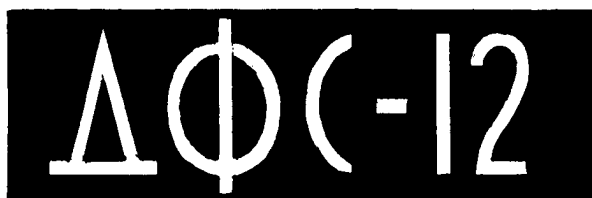
Sanitized Copy Approved for Release 2010/06/16 : CIA-RDP80T00246A045100330001-4

Sanitized Copy Approved for Release 2010/06/16 : CIA-RDP80T00246A045100330001-4



Сротоэлектрический
спектрометр.

СЕКЦИЯ СССР НА ВСЕМИРНОЙ ВЫСТАВКЕ 1958 Г. В БРЮССЕЛЕ



SPECTROMETRE LUMINEUX

L'appareil est destiné pour l'analyse des Raman-spectres des objets de grande dispersion. Il peut être aussi employé à l'étude d'autres faibles spectres, par exemple, de ceux de la luminescence, ainsi que pour l'analyse de structure qualitative et semi-quantitative de combinaisons organiques compliquées et de leurs mélanges.

Le dispositif principal de l'appareil est un monochromateur construit d'après un schéma original avec deux réseaux de diffraction (600 traits per mm, 150 × 140 mm).

Objectifs: miroirs paraboliques à distance focale de 800 mm.

Fentes de 40 mm de hauteur s'ouvrent dans les limites de 0 à 2 mm.

Pour développer le spectre il faut faire tourner de dispositif avec les réseaux de diffraction. La vitesse d'enregistrement des spectres peut être modifiée dans les limites de 0,25 à 25 Å/min.

Un photomultiplicateur à cathode à l'antimoine-césium (ФЭУ-17) sert de récepteur de l'énergie.

Le signal du photomultiplicateur est amplifié par l'amplificateur électronique à courant continu.

L'enregistrement du spectre s'effectue par l'appareil enregistreur automatique d'encombrement réduit.

Le spectromètre peut être utilisé aussi bien avec un monochromateur double qu'avec un monochromateur ordinaire, suivant le caractère de l'objet étudié.

L'appareil est fourni avec un illuminateur pour l'excitation des spectres de dispersion de combinaison dans les liquides.

*Pour les conditions des livraisons adressez-vous :
«STANKOIMPORT», Moskva, G-200, Smolenskaja-
Sennaja ploshchad, 32/34.*

LUMINOUS SPECTROMETER

ДФС-12

The instrument is designed for investigation of Raman-spectra of heterogeneous objects as well as for investigation of other weak spectra, f. e. luminescence. It may be also used for structural qualitative and semi-quantitative analysis of compound organic combinations and mixtures.

The main assembly of the instrument is the double monochromator of an original scheme with application of two diffraction gratings of 600 lines per mm. and the ruling area of 150×140 mm.

The objectives are parabolic mirrors with focal distance of 800 mm.

The slits of 40 mm. in height open within the range from 0 to 2 mm.

The scanning of spectrum is provided by turning the unit with the diffraction gratings. The gear box regulates spectra passing speed in slit plane within the range of 0.25 to 25 Å/min.

The energy receiver is a photomultiplier with an antimony-caesium cathode type ФЭУ-17

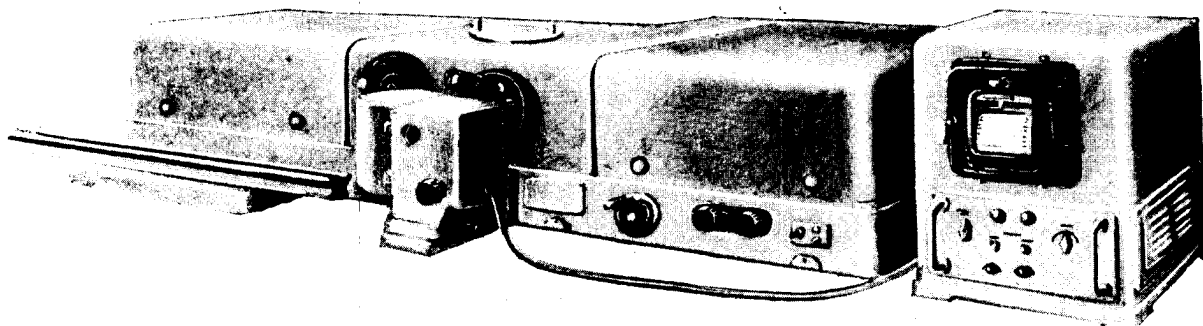
The signal from the multiplier is amplified by a d. c. electron amplifier.

The spectrum is recorded by a small-size recorder.

The instrument may be used with both the double monochromator and the ordinary one depending on the object under investigation study.

The instrument has a unit for excitation of Raman-spectra of liquids.

*For delivery terms apply to
"STANKOIMPORT", Moskva, G-200,
Smolenskaja-Sennaja ploshchad, 32/34.*



LICHTSTARKES SPEKTROMETER

ДФС-12

Das Gerät ist für die Untersuchung der Ramanspektren bestimmt. Es kann auch nicht nur für die Untersuchung schwacher Spektren (Lumineszenzspektren und dgl.), sondern auch für die qualitative und halbquantitative Strukturanalyse der zusammengesetzten organischen Verbindungen und ihrer Gemische verwendet werden.

Der Hauptteil des Geräts ist ein nach originalem Schema gebauter Doppelmonochromator mit zwei Beugungsgittern (600 Str/mm und mit einer Arbeitsfläche 150×140 mm).

Die Objektive stellen parabolische Spiegel mit 800 mm Brennweite dar. Spalte von 40 mm Höhe wird im Bereich von 0 bis 2 mm geöffnet.

Die Abwicklung des Spektrums erfolgt durch die Drehung der Einheit mit den Beugungsgittern. Die Geschwindigkeitswechselgetriebe ermöglichen die Steuerung des Spektraldurchgangs in der Ebene des Spaltes von 0,25 bis 25 Å/min.

Als Energie-Empfänger dient ein Photo-Vervielfacher mit Antimon-Cäsium-Kathode (ФЭУ - 17).

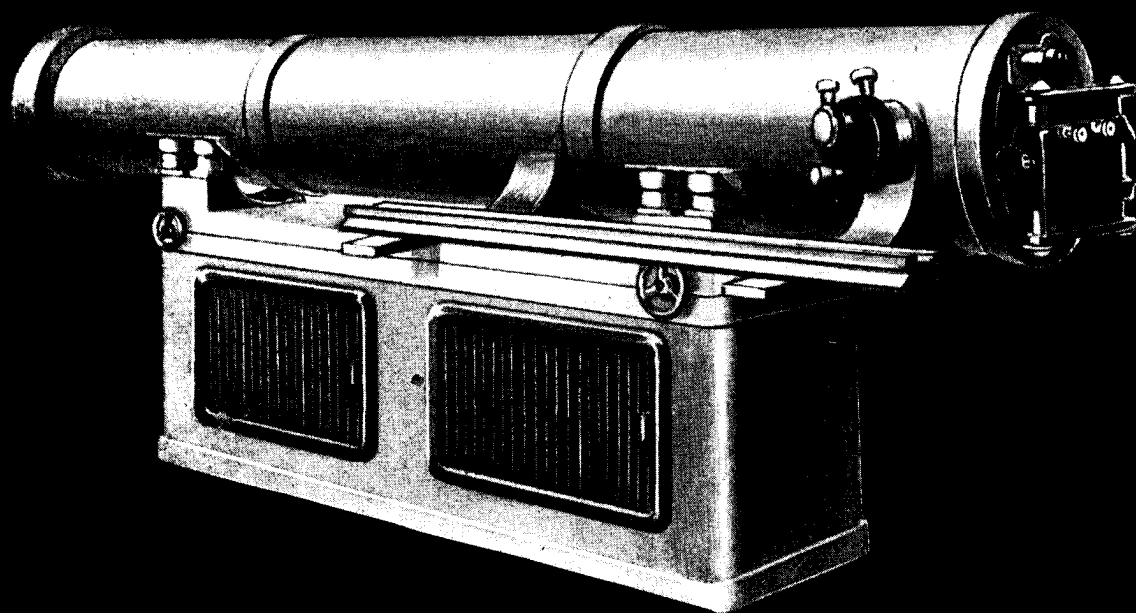
Das Signal aus dem Vervielfacher wird durch einen Gleichstrom-Elektronenverstärker verstärkt.

Die Aufzeichnung des Spektrums erfolgt auf einem Klein-Schreibgerät.

Das Gerät kann sowohl mit dem Doppelmonochromator wie auch mit einem Einfachen-Monochromator verwendet werden, je nach Beschaffenheit des zu untersuchenden Objekts.

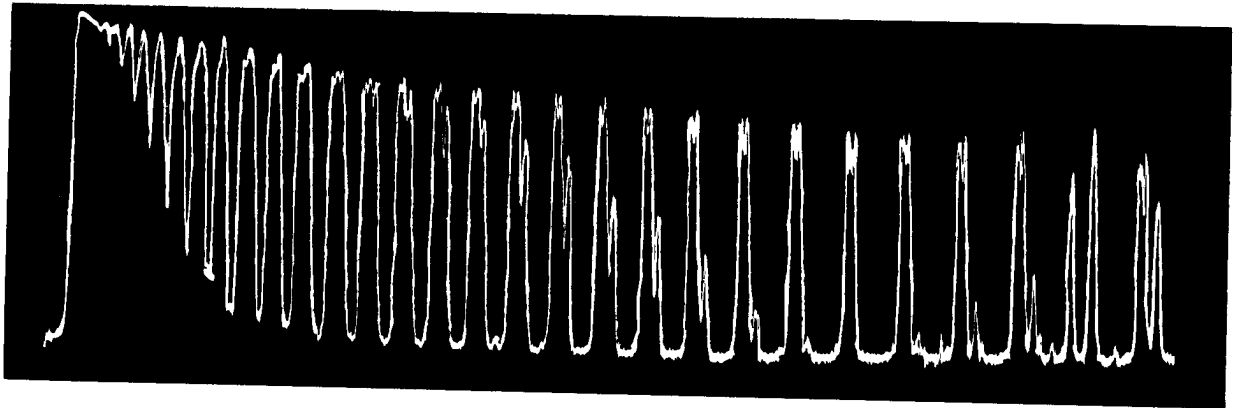
Dem Gerät wird ein Beleuchter für die Erregung von Ramanspektren in Flüssigkeiten beigegeben.

Alle Lieferungsanfragen sind an „STANKOIMPORT“
Moskva, G-200, Smolenskaja-Sennaja Ploschad, 32/34
zu richten.



СПЕКТРОГРАФ

ДФС-13



L'enregistrement de la raie CN 3883,4 – 3771,4 Å.

On voit une résolution nette des raies séparées par la distance de 0,05 Å.

The CN-band record at 3883,4 – 3771,4 Å.

The lines at a distance of 0.05 Å are permissible.

Aufschreiben des Bandes CN 3883,4 – 3771,4 Å.

Es ist eine gute Auflösung der Linien, die sich im Abstand von 0,05 Å befinden, ersichtlich.

SPECTROGRAPHE DE DIFFRACTION

L'appareil est destiné aux diverses études des spectres compliqués exigeant une grande dispersion dans la région du visible de 2000 à 10000 Å.

Le spectrographe assure une haute qualité de l'image des raies spectrales grâce au schéma d'auto-collimation vertical symétrique.

La distance focale de l'objectif à miroir – 4 m, l'ouverture relative – 1:42.

L'appareil travaille avec le réseau de diffraction de 1200 traits/mm ou bien de 600 traits/mm.

La dispersion linéaire inverse – 2 ou 4 Å/mm fonctionnant en 1^{er} ordre.

Le spectre est photographié par parties, chacune de 500 ou 1000 Å, sur un châssis plat de 9 x 24 cm.

Parallèlement à la prise de photographie du spectre l'échelle des longueurs d'ondes est imprimée sur la plaque.

Le spectrographe comprend:

système d'illumination achromatique à optique de quartz, à trois lentilles;

atténuateur à étages;

adaptateur pour la prise de photographie sur le film de 35 mm.

Pour les conditions des livraisons adressez-

vous: «STANKOIMPORT», Moskva,

G 200, Smolenskaja-Sennaja ploshchad, 32 34

ΔΦC-13

DIFFRACTION SPECTROGRAPH

The instrument is designed for investigation of complex spectra demanding large dispersion in regions from 2000 to 10000 Å.

Application of vertically symmetric autocollimation scheme provides high quality of the photographing of spectral lines.

Focal distance of mirror objective is 4 m, relative aperture—1:42.

The instrument operates in two variants: with diffraction grating of 1200 lines per mm or 600 lines per mm.

Reverse linear dispersion is correspondingly 2 or 4 Å/mm in the first order.

Spectrum is photographed by areas of 1000 or 500 Å on 9 x 24 cm flat plate.

Along with the photographing of spectrum imprinted on the plate is the wavelength scale.

The spectrograph ΔΦC-13 is equipped with:

quartz, achromatic three-lens illuminating system; stage attenuator;

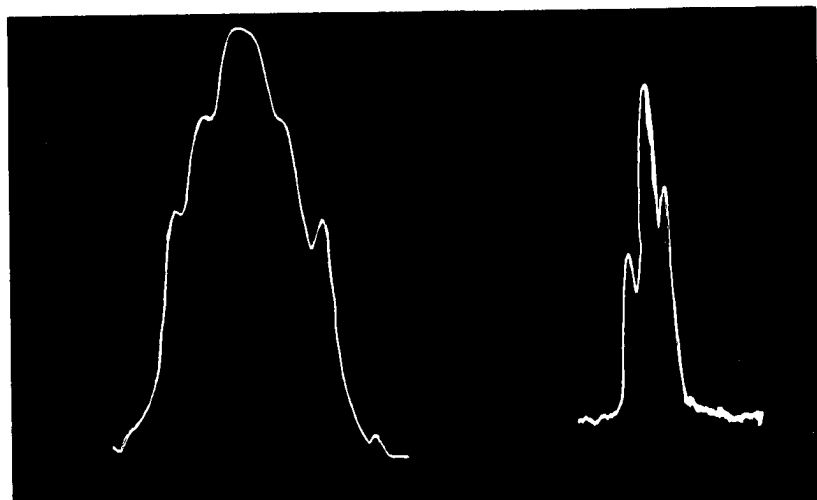
adapter for photographing on 35 mm film.

La structure fine des raies de mercure 3654 Å et 4358 Å

Fine structure of mercury lines 3654 Å and 4358 Å

Feinstruktur der Quecksilberlinien 3654 Å und 4358 Å

*For delivery terms apply to "STANKOIMPORT",
Moskva, G. 200, Smolenskaja Sennaja plosh-
chad, 32/34.*



USSR SECTION: BRUSSELS UNIVER-
SAL AND INTERNATIONAL EXHI-
BITION 1958

DIFFRAKTIONS-SPEKTROGRAPH

ДФС 13

Das Gerät ist für die Durchführung der verschiedenartigen spektroskopischen Untersuchungen im Spektralbereich von 2000 bis 10000 Å bestimmt. Es zeichnet sich durch gute Qualität der Spektrallinien-Abbildung aus. Das wurde durch die Verwendung des vertikal-symmetrischen Autokollimations-schemas erreicht.

Die Brennweite des Spiegel-Objektivs beträgt 4 Meter, Öffnung-Verhältnis — 1:2.

Das Gerät funktioniert in zwei Varianten: mit einem Beugungsgitter mit 1200 Str/mm und mit einem mit 600 Str/mm.

Die reziproke lineare Dispersion beträgt 2 oder 4 Å/mm in der 1. Ordnung.

Das Spektrum wird in Abschnitten von je 500 oder 1000 Å auf einer flachen 9 x 24-cm-Kassette photographiert.

Gleichzeitig mit dem Photographieren des Spektrums wird auf die Platte die Wellenlängenskala aufgedrückt.

Zum Spektrograph werden beigelegt:

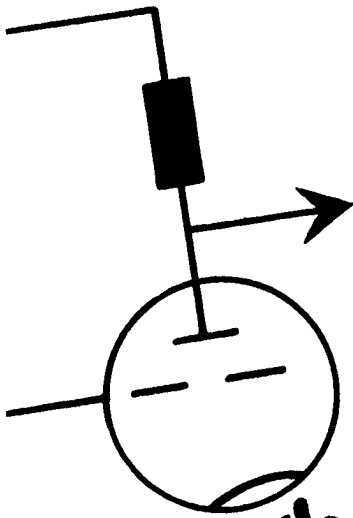
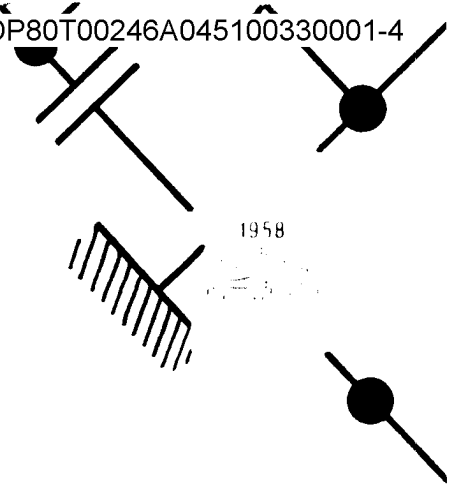
achromatisches, - Quarz - und Dreilinsen-Beleuchtungssystem

Stufenminderer

Adapter für das Photographieren auf einem 35-mm-Filmstreifen.

Alle Lieferungsanfragen sind an „STANKOIMPORT“,
Moskva, G-200, Smolenskaja-Sennaja Ploschad, 32/34 zu
richten.

ABTEILUNG DER UdSSR AUF DER ALLGEMEINEN WELTAUSSTELLUNG IN BRÜSSEL 1958



Фотозлектронный
стимулятор ФЭС-1

СЕКЦИЯ СССР НА ВСЕМИРНОЙ ВЫСТАВКЕ 1958 г. В БРЮССЕЛЕ

SECTION DE L'URSS À L'EXPOSITION UNIVERSELLE ET INTERNATIONALE DE BRUXELLES 1958

STYLOMETRE PHOTOELECTRIQUE

Le stylo-mètre photoélectrique ФЭС-1 est destiné à l'analyse spectrale quantitative des aciers faiblement alliés et d'autres alliages dans la région visible du spectre.

L'appareil mesure le rapport de l'intensité de la raie spectrale de l'élément analysé à l'intensité de la lumière non-décomposée qui est la fonction de la concentration de l'éléments dans l'alliage.

La détermination du taux de concentration de l'élément à doser s'effectue en deux ou trois minutes. Le dosage de plusieurs éléments s'effectue successivement.

L'appareil fonctionne dans la gamme des longueurs d'ondes de 3900 à 7000 Å, la dispersion linéaire dans la région de 4700 Å étant de 32 Å/mm.

Récepteurs d'émission: cellules photoélectriques à antimoine-césium du type CIIB-9.

Le placement de la raie à analyser sous la fente de sortie de l'appareil s'effectue manuellement, le contrôle est fait au moyen de l'oculaire.

Les résultats de l'analyse sont lus sur l'échelle de l'appareil de sortie ЭПБ-01, classe 0,5.

Le mesurage est automatisé du moment de branchement du générateur jusqu'à la lecture des résultats de l'analyse.

L'appareil complet comporte le générateur à arc et à étincelle à commande électronique de l'amorçage. Le générateur assure une haute reproductibilité des conditions de la décharge.

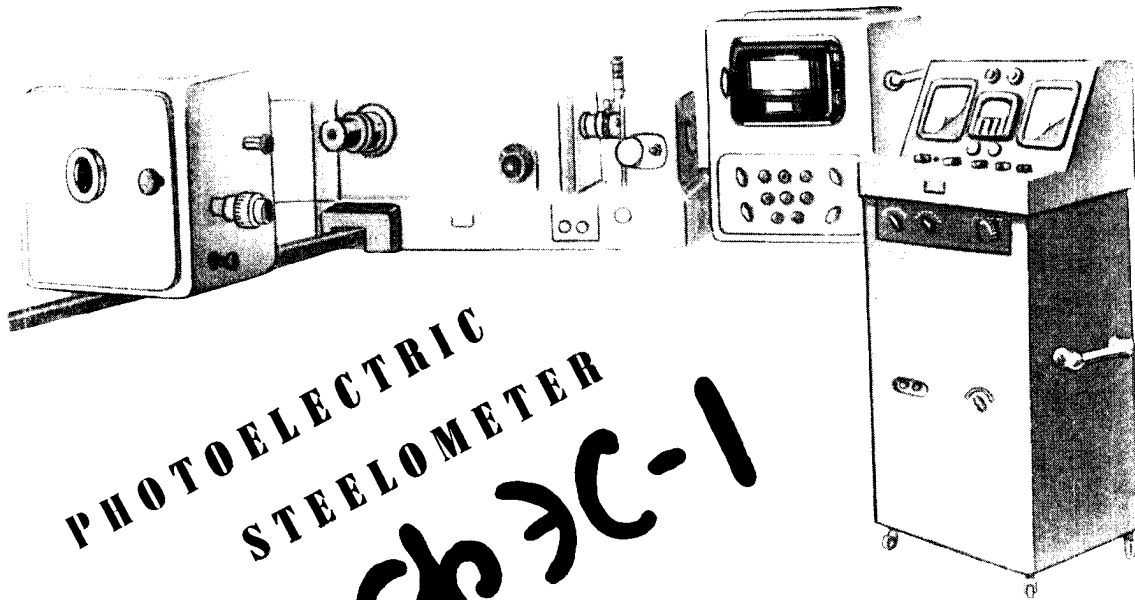
Encombrement de l'appareil en état de service, mm:

appareil spectral	1250×1700×400
dispositif électrique de mesure	516×420×1500
générateur à arc et à étincelle	650×450×1400

Poids, kg:

appareil spectral	218
dispositif électrique de mesure	100
générateur à arc et à étincelle	160

*Pour les conditions des livraisons
adressez-vous: «STANKOIMPORT»,
Moskva, G-200, Smolenskaja-Sennaja
Ploshchad, 32/34.*



The photoelectric steelometer is designed for carrying out quantitative spectral analysis of low-alloy steels and other alloys in the visible region of the spectrum.

By means of this instrument, the ratio is measured of the radiant fluxes of the spectral line of the element under study to the undecomposed light. This ratio is the function of the element concentration in the alloy.

The determination in percentage of the concentration of one element under study is done in two or three minutes. The analysis of several elements is carried out consecutively.

The instrument operates within a wavelength range of 3900 to 7000 Å with line dispersion in a region of 4700 Å = 32 Å/mm.

Radiation receivers — CHB-9 antimony-caesium photocells.

The adjustment of a desirable line against the exit slit of the instrument is carried out by hand. Control is accomplished through the eye-piece.

The result of the analysis is read from the scale of the CHB-01 exit instrument, class 0.5.

The measurement process is automatic from the moment when the generator is switched on until the analysis results are obtained.

The instrument set includes an arc and spark generator with an electronic igniting control. The generator ensures high reproduction of discharge conditions.

Overall dimensions of the instrument when in operation, mm:

spectral instrument	1250 x 1700 x 400
electromasuring arrangement	510 x 420 x 1500
arc and spark generator	650 x 450 x 1400

Weight, kg:

spectral instrument	218
electromasuring arrangement	100
arc and spark generator	160

The delivery terms apply to USSR-KOMMORTE, Moscow G-200, Smolensk, Kazanskaya plashina, 32-31

PHOTOELEKTRISCHE STEELOMETER $\Phi\Xi\text{C}-1$

Das photoelektrische Steelometer $\Phi\Xi\text{C}-1$ ist für die quantitative Spektralanalyse von schwachlegierten Stählen und anderen Legierungen im sichtbaren Spektralgebiet bestimmt.

Am Gerät wird das Verhältnis der Intensität der Spektrallinie des zu analysierenden Elements zu der Intensität des unzerlegten Lichts gemessen, das eine Funktion der Konzentration dieses Elements in der Legierung ist.

Die Bestimmung der Konzentration eines zu analysierenden Elements erfolgt binnen zwei bis drei Minuten. Die Analyse auf mehrere Elemente erfolgt der Reihe nach.

Das Gerät arbeitet im Wellenlängenbereich von 3900 bis zu 7000 Å mit einer linearen Dispersion im Gebiet 4700 Å — 32 μm .

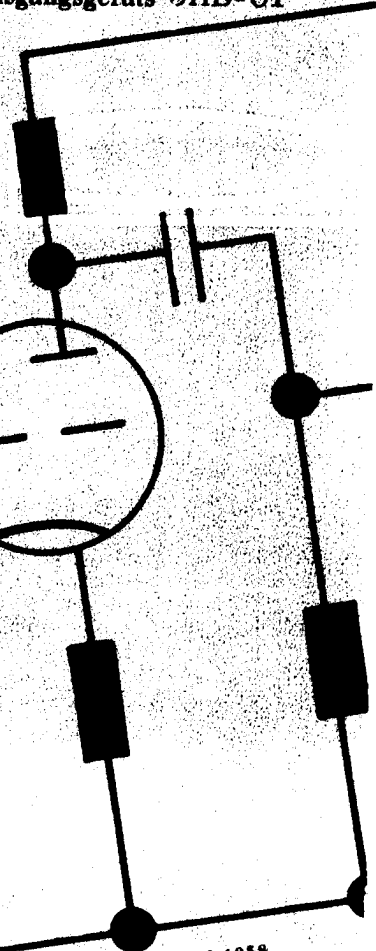
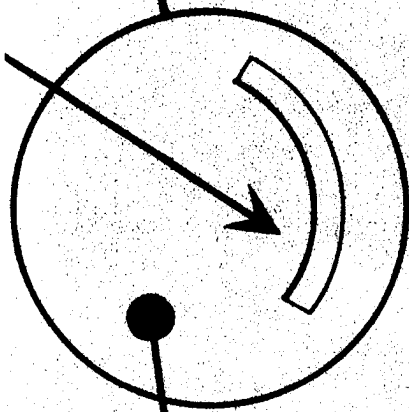
Als Strahlungsempfänger dienen Antimon-Cäsium-Photozellen Type CIIB-9.

Die Zuführung der erforderlichen Linse zum Ausgangspalt des Geräts erfolgt von Hand, die Kontrolle erfolgt durch das Okular.

Das Endergebnis der Analyse wird an der Skala des Ausgangsgeräts $\Xi\text{IIB-OI}$ der Klasse 0,5 abgelesen.

Der Meßvorgang ist von dem Augenblick der Einschaltung des Generators bis zum Empfang des Analyse-Ergebnisses automatisch.

Zum Gerätsatz gehört der Lichtbogen- und Funkengenerator mit elektronischer Steuerung der Zündung. Der Generator gewährleistet eine hohe Wiederholbarkeit der Entladungsverhältnisse.



Abmessungen des Geräts im Betriebszustand, mm:
das Spektralgerät

1250 x 1750 x 400
570 x 420 x 1300
530 x 450 x 1400

die elektronische Meßeinrichtung
der Lichtbogen- und Funkengenerator

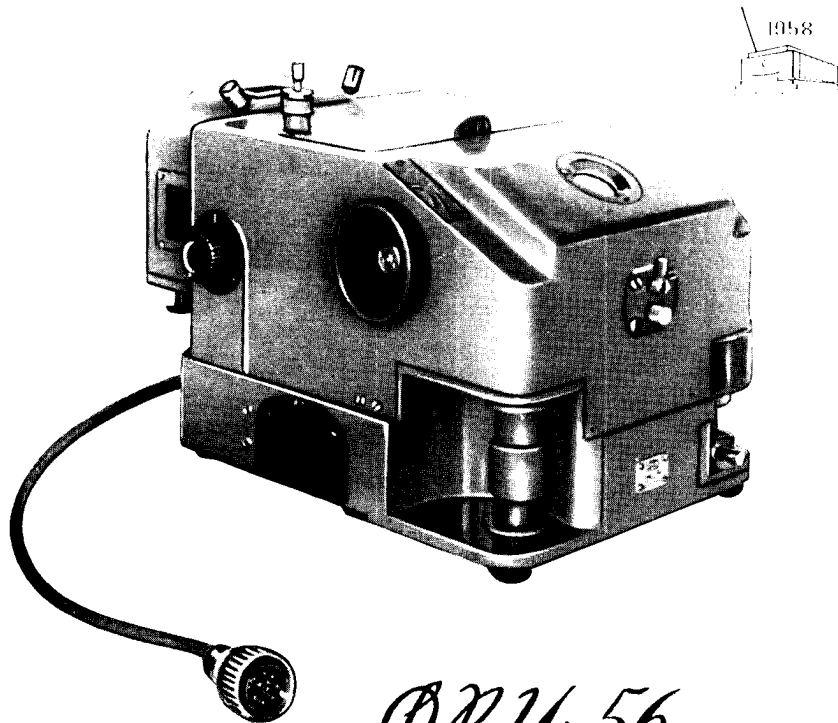
Gewicht, kg:
das Spektralgerät
die elektronische Meßeinrichtung
der Lichtbogen- und Funkengenerator

210
100
160

Alle Lieferungsanfragen sind an "STAN-KOIMPORT", Moskau, G-200, Srednenskaja-Sennaja Ploschad, 32/34, zu richten.

USSR SECTION: BRUSSELS UNIVERSAL AND INTERNATIONAL EXHIBITION 1958
ABTEILUNG DER UdSSR AUF DER ALLGEMEINEN WELTAUSSTELLUNG IN BRUSSEL 1958

Sanitized Copy Approved for Release 2010/06/16 : CIA-RDP80T00246A045100330001-4



DRU-56

**КОМПАРАТОР
ЦВЕТА**

СЕКЦИЯ СССР НА ВСЕМИРНОЙ ВЫСТАВКЕ 1958 г. В БРЮССЕЛЕ

Sanitized Copy Approved for Release 2010/06/16 : CIA-RDP80T00246A045100330001-4

OK 4-50

COMPARATEUR DE COULEUR

Cet appareil est un appareil photo-électrique à l'objectif pour la détermination de différences minces de couleur.

Pris en usage dans différentes branches de l'industrie : polygraphique, textile, de laques et de couleurs, céramique et d'autres pour comparer les couleurs de la production avec les étalons.

Se soumettent à l'épreuve les corps solides, la surface desquels diffuse la lumière.

La source de lumière est une lampe à incandescence à 12 V, 100 W.

Le récepteur de l'énergie lumineuse – deux cellules photo-électriques à selenium.

Renforceur des courants photo-électriques est un pont symétrique équilibré dans la diagonale duquel est mis en circuit un galvanomètre de la sensibilité de 10^{-6} A division.

Trois filtres colores sont assortis conformément à la caractéristique de l'œil.

L'appareil est trois fois plus sensible que l'œil humain.

Le diaphragme de mesure avec l'échelle circulaire sert pour la compensation de l'inclinaison de zéro-instrument.

L'alimentation se produit du réseau du courant alternatif (220 V, 50 hertz) avec l'intermédiaire de l'arrangement spécial alimentaire entrant dans l'assemblage de l'appareil.

Erreur relative des mesurages ne dépasse pas 3 %.

Encombrement, mm:

de l'appareil 386 x 275 x 248

de l'ensemble 386 x 275 x 311

Poids, kg 15

Pour les conditions des livraisons
adressez-vous:

« STANKOIMPORT », Moskva, G-200,
Smolenskaja-Sennaja Ploshchad, 32/34.

Sanitized Copy Approved for Release 2010/06/16 : CIA-RDP80T00246A045100330001-4
COLOUR

COMPARATOR

ФКЦ-56

The ФКЦ-56 is an objective photoelectric apparatus and is designed to determine small differences in colours.

It is used in printing, textile, varnish and dye-stuff, ceramic and other industries to compare colours of produced goods with standards.

To tests are subjected light scattering and transparent solids and liquids.

The source of light is a 12V filament lamp of 100W.

Two selenium photoelectric cells serve as the receiver of the energy of light.

The amplifier of photoelectric currents is a symmetrical balanced bridge in whose diagonal a ГМП galvanometer with a sensitivity of 10^{-6} A div. is included.

Three light filters are selected in conformity with colour characteristics of the eye.

The apparatus is three times more sensitive than a human eye.

The measuring diaphragm with a circular scale serves to compensate the deviation of the zero-instrument.

Power supply is received from A. C. 220V 50 cycles mains through a special feeding unit which is attached to the apparatus.

The relative error of measurements is not more than 3%.

Overall dimensions, mm:

Apparatus	386 x 275 x 248
Feeding unit	315 x 210 x 141
Weight, kg	15

For delivery terms apply to
"STANKOIMPORT", Moskva, G-200,
Smolenskaja-Sennaja Ploshchad, 32/34.

USSR Section : Brussels Universal and International Exhibition 1958

Внешторгиздат. Заказ № 13 229/26

Sanitized Copy Approved for Release 2010/06/16 : CIA-RDP80T00246A045100330001-4

Sanitized Copy Approved for Release 2010/06/16 : CIA-RDP80T00246A045100330001-4

OPTISCHER FARBENVERGLEICHAPPARAT KOMPARATOR

ОКН-56 ist ein fotoelektrisches Objektivgerät, das zum Erkennen kleiner Farbenabstufungen (Farbennuancen) benützt wird.

Es findet Verwendung in der Buchdruckindustrie, Textilindustrie, keramischen Industrie und anderen Industriegebieten und dient zum Vergleich der Farben der Werkserzeugnisse mit angegebenem Vergleichsmuster.

Es werden untersucht: feste Körper, deren Flächen das Licht zerstreuen, durchsichtige feste Körper und Flüssigkeiten.

Als Lichtquelle dient eine Glühlampe 12 V, 100 W.

Als Lichtempfänger – zwei Selenfotoelemente.

Als Verstärker des fotoelektrischen Stromes dient eine symmetrische balancierte Brückenschaltung mit einem Galvanometer (Mikroammperemesser) Typ I'MII von 10^{-6} A Teilung Empfindlichkeit auf der Diagonale.

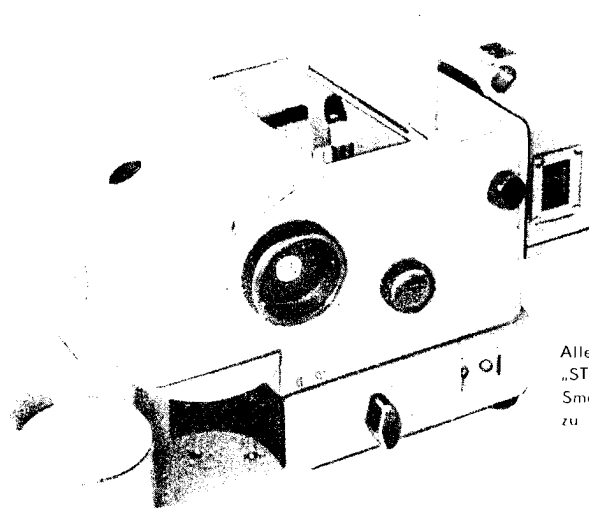
Drei Lichtfilter entsprechen der Charakteristik der drei Netzhautfarbempfangner des Auges.

Das Gerät ist dreimal mehr empfindlich als das Menschenauge.

Ein Meßdiaphragma (Blende) mit Ringskala dient zur Kompensierung der Null-Stellung des Gerätes.

Das Gerät wird gespeist vom Wechselstromnetz 220 V, 50 Hz mittels einer speziellen Speisevorrichtung, die als Teil dem kompletten Gerät zugehört.

Der relative Fehler der Messungen beträgt 3 %.

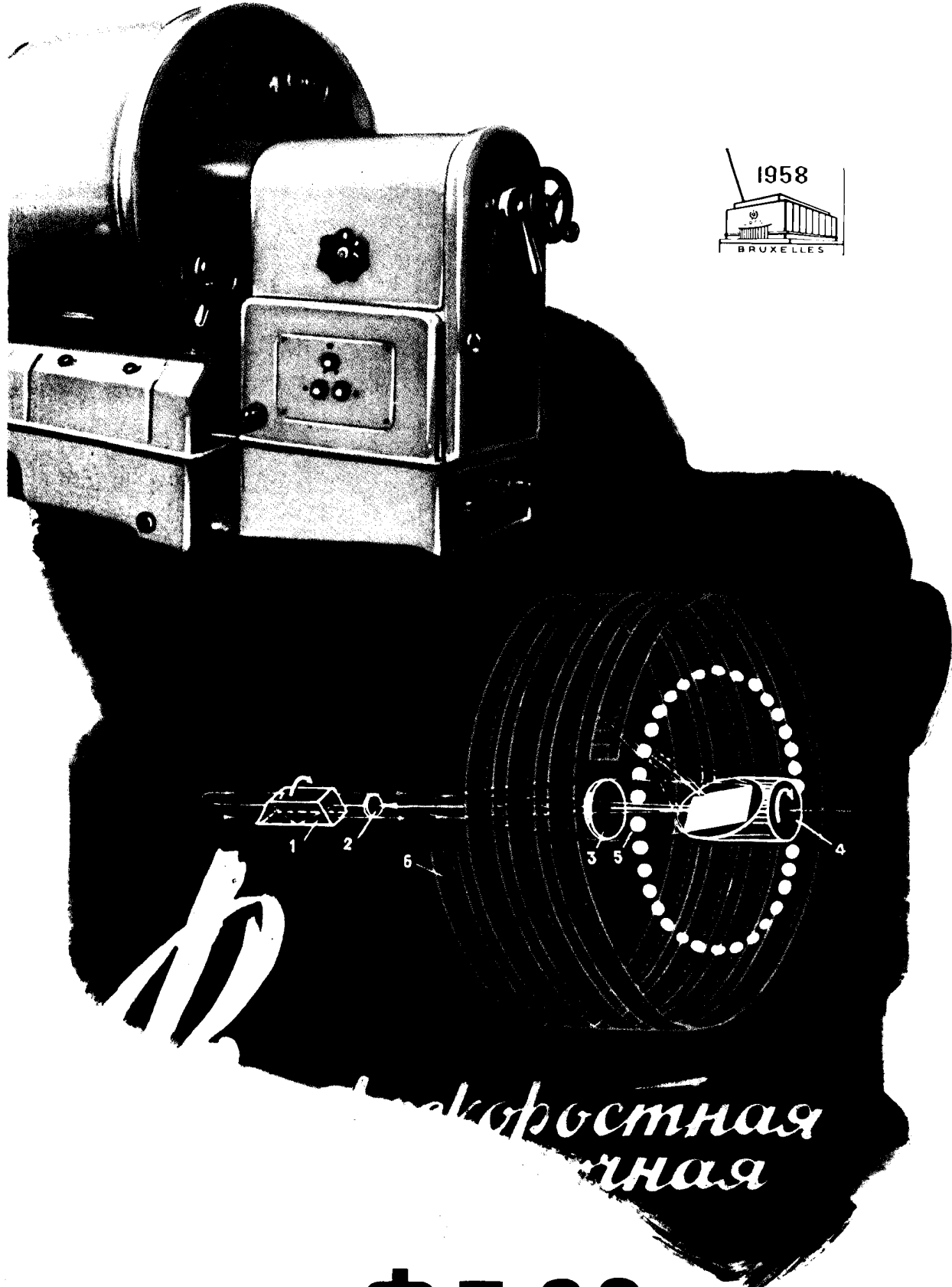


Außere Abmessungen mm:
des Apparates 386 × 275 × 248
der Energiespeisevorrichtung
315 × 210 × 141
Gewicht, kg 15

Alle Lieferungsanfragen sind an
„STANKOIMPORT“, Moskva, G 200,
Smolenskaja-Sennaja Ploshchad, 32/34
zu richten.

Abteilung der UdSSR auf der Allgemeinen Weltausstellung in Brüssel 1958

Sanitized Copy Approved for Release 2010/06/16 : CIA-RDP80T00246A045100330001-4



ФП-22

APPAREIL CINÉMATOGRAPHIQUE DE PRISES DE VUES EXTRA-RAPIDE

La caméra Φ II-22 est destinée à la cinématographie de processus se passant rapidement, avec obtention de films pour projection directe.

Principe du travail. L'objectif 2 et la lentille 3 portent l'image d'un objet éloigné à l'infini sur le miroir incliné 1 tournant rapidement. La lentille collective 3 forme la pupille du système de pénétration sur la couronne avec objectifs 5, l'obturateur optique assurant l'exposition successive des images est représenté à l'entrée de la couronne avec lentilles.

Les objectifs transportent l'image du miroir sur la pellicule 6 logée en spirale avec mouvement strictement synchronisé avec la rotation du miroir.

Le prisme rotatif 7 sert à la stabilisation de l'image en supprimant la déformation amenée par la rotation du miroir.

Alimentation électrique:

courant triphasé 220 V, 1,7 kVA;

courant continu 24 V, 4 A.

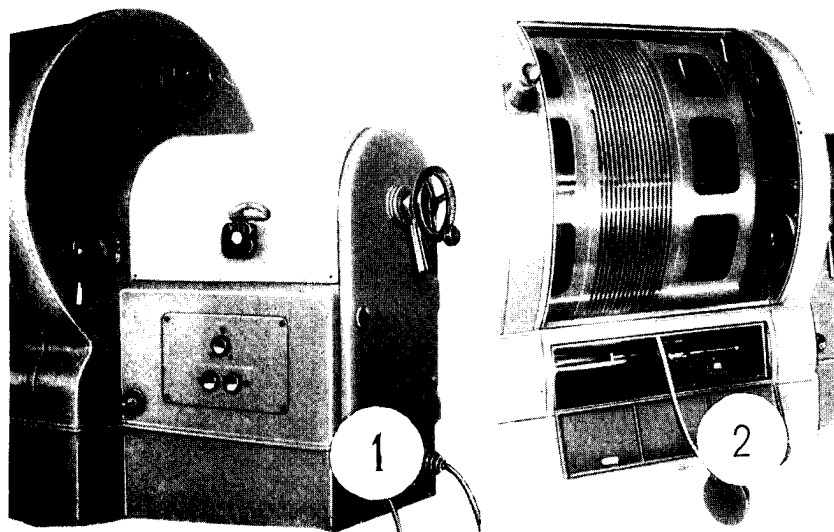
*Pour les conditions des livraisons adressez-vous:
«STANKOIMPORT», Moskou, G-200,
Smolenskaja-Sennaja Ploshchad 32/34.*

CARACTÉRISTIQUES TECHNIQUES

Fréquence des prises de vues, images/sec.....	5000 à 100000
Distance focale, mm	58
Ouverture relative	1 : 11,6
Capacité de résolution, lignes/mm	40
Grossissement du viseur	4,6*
Largeur de la pellicule, mm	8
Format des images, mm	3,6 × 4,8
Nombre d'images obtenues par prise de vues.....	7600
Longueur de la bobine, m	30
Durée de la démonstration d'un film complet, min.	8
Encombrement, m.....	1,7 × 0,7 × 0,8
Poids, kg	340

Complexe:

- Installations d'éclairage;
- Jeu de lentilles de mise au point pour cinématographie depuis l'infini jusqu'à moins d'un mètre;
- Fréquencemètre;
- Installation de développement;
- Appareil de projection cinématographique;
- Boîtes pour la pellicule et accessoires.



- 1 Organes de commande situés à droite de la caméra
 - 2 Pendant la mise en place de la pellicule et à la fin du filmage le tambour se trouve à gauche.
 - 3 La caméra étant prête pour le filmage, le tambour se trouve dans la position initiale.
 - 4 La caméra est pourvue de synchronisateurs.
-
- 1 All controls are arranged on the right-hand face of the camera.
 - 2 When loading and at the termination of photographing the drum is in the left-hand side position.
 - 3 The camera is ready for photographing, the drum is at the initial point.
 - 4 The camera is provided with a synchronization device.
-
- 1 Die Steuervorrichtungen sind am rechten Kameraende angebracht.
 - 2 Während der Ladung und am Ende der Aufnahme befindet sich die Trommel in der linken Stellung.
 - 3 Wenn die Kamera zur Filmaufnahme vorbereitet ist, dann befindet sich die Trommel in der Ausgangsstellung.
 - 4 Die Kamera ist mit einem Synchronisator versehen.

The $\Phi\Pi-22$ cinema camera is designed for the photographing of transient processes.

Operation principle. Lens 2 in conjunction with collective lens 3 gives the image of an object in the infinite distance on a sloping swiftly rotating mirror 4. Collective lens 3 gives the image of the system inlet pupil on a rim with lenses 5. The image of the inlet pupil skimming over the surface of the rim with lenses acts as an optical obturator and insures a successive exposure of frames.

The objective-lenses transfer the image from the mirror on to film 6; the latter is arranged spirally and runs in an evenly advancing motion in strict concurrence with the rotating mirror.

The rotating prism 1 serves as an image stabilizing safeguard against the turning effect of the rotating mirror.

Electricity feed:

220 V, 1.7 kVA three-phase current,
24 V, 4 A constant current.

*For delivery terms apply to
"STANKOIMPORT", Moskva, G-200,
Smolenskaja-Sennaja Ploshchad, 32/34.*

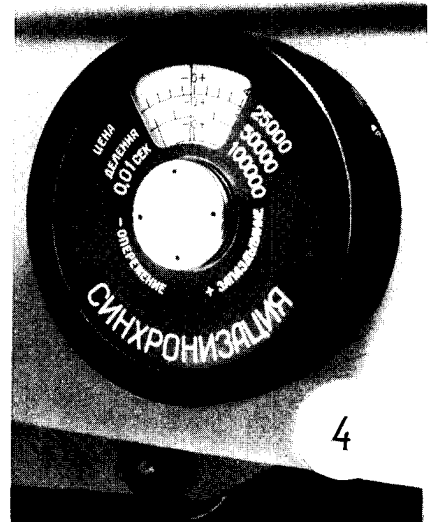
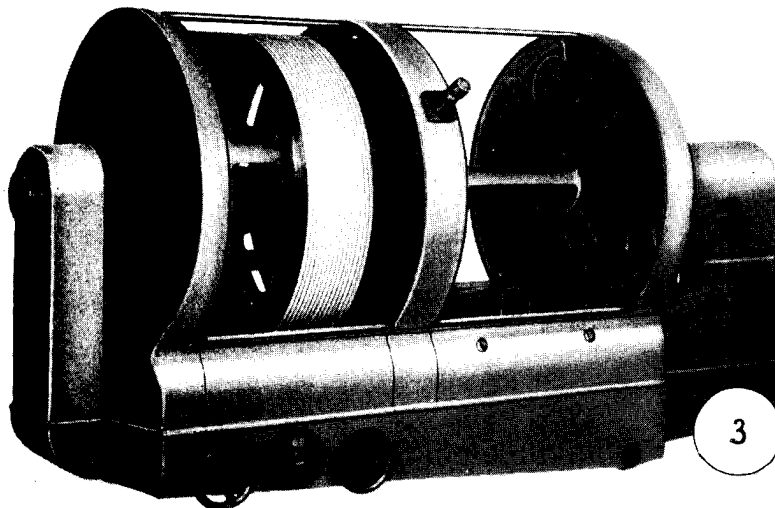
SPECIFICATIONS

Exposure frequency, frames per sec.	5000-100000
Focal length, mm	58
Relative aperture	1:11.6
Resolving capacity, line/mm	40
View-finder magnification	4.6*
Film width, mm	8
Frame size, mm	3.6 x 4.8
Number of frames obtained at one taking	7600
Magazine capacity, mm	30
Show time of one film, min.	8
Overall dimensions, m	1.7 x 0.7 x 0.8
Weight of camera, kg	340

Unit accessories:

lighting appliance;
set of lenses focussing at distances from less than a meter to infinity;
frequency meter;
developing equipment;
cinema projector;
film magazines and other accessories.

**SUPER-RAPID
CINEMA CAMERA**

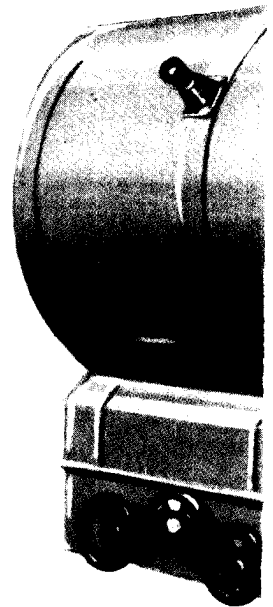


Φ Π-22

DIE HOCHGESCHWINDIGKEITS-KAMERA FÜR KINOFILMAUFNAHMEN

TECHNISCHE CHARAKTERISTIK

Aufnahme-Frequenz, Kader/Sek.	5000-100 000
Fokus-Abstand, mm	58
Relative Öffnung	1:11,6
Auflösungs-Vermögen	40
Visier-Vergrößerung	4,6*
Breite des Filmbandes, mm	8
Kader-Dimensionen, mm	3,6 × 4,8
Zahl der Kader, die bei einer Filmaufnahme erhalten werden	7600
Kassetten-Volumen, m	30
Demonstrationszeit des ganzen Films, Min.	8
Gabarit-Dimensionen, m	1,7 × 0,7 × 0,8
Gewicht, kg	340



Die Kamera Φ Π-22 ist für Kinofilmaufnahmen schnellablaufender Prozesse bestimmt mit der Erhaltung von Filmen, die für direkte Kino-projektion geeignet sind.

Wirkungsprinzip. Das Objektiv 2 zusammen mit der Kollektivlinse 3 bildet den unendlich entfernten Gegenstand auf dem geneigten, sich schnell drehenden Spiegel 4 ab. Die Kollektivlinse 3 bildet die Eingangsöffnung des Systems auf der Korona mit den Linsen 5 ab. Die gleitende Abbildung der Eingangsöffnung auf der Korona mit den Linsen stellt den optischen Obturator dar, welcher die aufeinanderfolgende Exponierung der Kader gewährleistet.

Die Objektive übertragen die Abbildung vom Spiegel auf den Film 6, der in Form einer Spirale angeordnet ist und sich bei der Filmaufnahme in strenger Übereinstimmung mit der Drehung des Spiegels gleichmäßig-fortschreitend bewegt.

Das sich drehende Prisma 1 dient zur Stabilisierung der Abbildung gegenüber Drehung, die durch Rotation des Spiegels hervorgerufen wird.

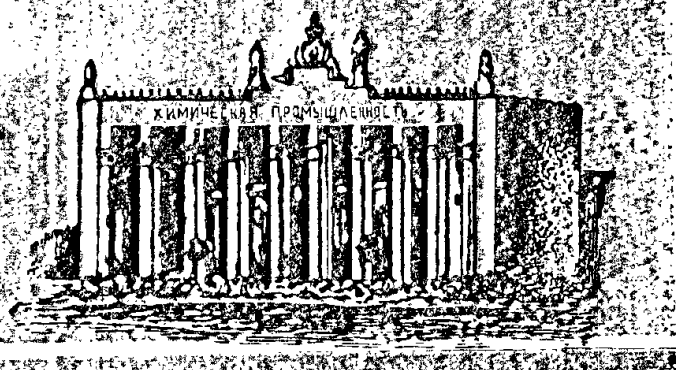
Elektrizitätsspeisung:

Dreiphasenstrom 220 V, 1,7 kVA
Gleichstrom 24 V, 4 A.

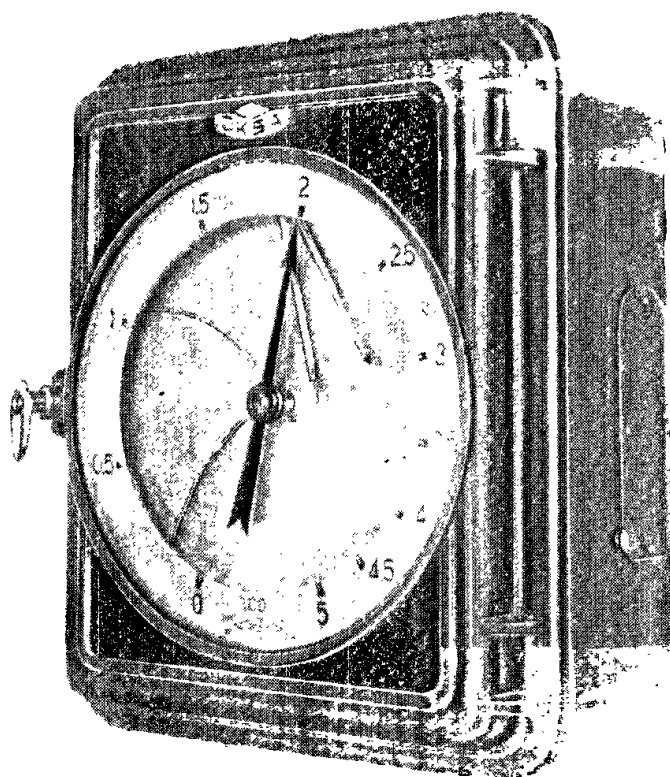
Apparatur-Bestandteile:

- Belichtungs-Einrichtungen;
- Fokussierende Linsen zur Aufnahme aus Entfernungen von unendlich bis zu Meter-Bruchteilen;
- Frequenzmesser;
- Entwicklungs-Einrichtungen;
- Kinoprojektor;
- Kassetten für den Film und Zubehör.

Alle Lieferungsanfragen sind an
„STANKOIMPORT“, Moskau, G-200,
Smolenskaja-Sennaja Ploschad, 32/34 zu richten.



ГАЗОАНАЛИЗАТОР ТИПА ГИП-5



МИНИСТЕРСТВО
ХИМИЧЕСКОЙ ПРОМЫШЛЕННОСТИ СССР
МОСКВА

ГАЗОАНАЛИЗАТОР типа ГИП-5

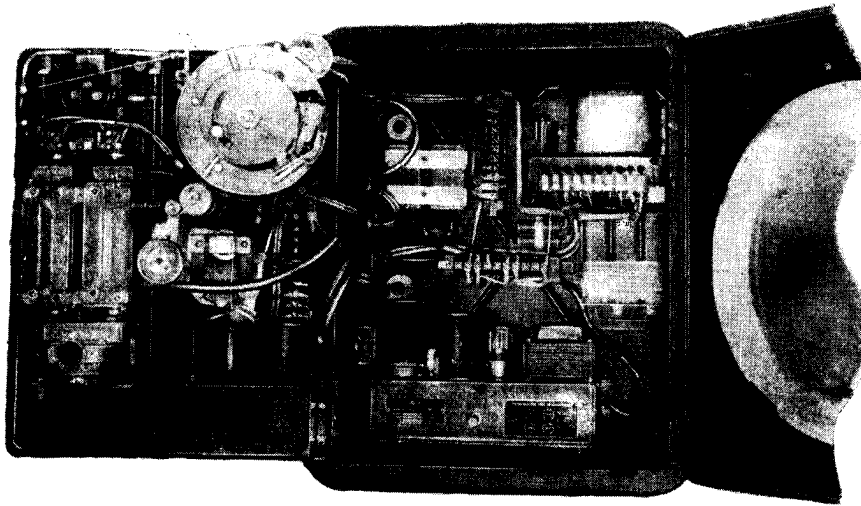
Газоанализатор типа ГИП-5 представляет собой стационарный автоматически действующий, показывающий и регистрирующий прибор для непрерывного определения содержания одного из компонентов в газовой смеси.

Работа газоанализатора основана на принципе избирательного поглощения инфракрасного излучения анализируемым компонентом. Газоанализатор ГИП-5 изготавливается на базе электронного потенциометра типа ЭПД.

Шкала прибора градуируется в объемных процентах анализируемого компонента. Измерения производятся компенсационным методом.

Источниками инфракрасного излучения являются две нихромовые спирали, накаливаемые электрическим током. Потоки инфракрасного излучения от нагретых спиралей прерываются обтюратором, проходят через две параллельно установленные кюветы и попадают в лучеприемник, в котором происходит сравнение энергий потоков. Через одну из кювет — измерительную — непрерывно продувается газовая смесь, содержащая анализируемый компонент, который поглощает часть потока инфракрасных лучей. Вторая кювета — сравнительная — заполнена воздухом, и проходящий через нее поток лучей остается неизменным. Поэтому в камеры лучеприемника попадают потоки, разность энергий которых зависит от концентрации анализируемого компонента в смеси.

Лучеприемник имеет две камеры, наполненные анализируемым компонентом и отделенные друг от друга мембранным конденсатором, мембрана которого воспринимает перепад давления между камерами. Прерываемые обтюратором потоки лучей периодически создают между камерами лучеприемника перепад давления. При равенстве энергий потоков сигнал от лучеприемника не поступает. При неравенстве потоков мембрана конденсатора начинает колебаться и возникает переменный ток, который поступает вначале на преобразовательный, а затем, после выпрямления, на окончательный усилитель и приводит во вращение реверсивный двигатель. С вала реверсивного двигателя движение через тягу передается на компенсационную заслонку, которая уменьшает площадь нижнего окна сравнительной кюветы до тех пор, пока потоки вновь не уравниваются. Таким образом, в приборе автоматически устанавливается равенство потоков и, следовательно, нуль сигнала. Каждой концентрации исследуемого газового компонента соответствует определенное положение компенсирующей заслонки и кинематически связанных с ней указывающей стрелки и записывающего пера прибора.



Внутренний вид газоанализатора типа ГИП-5.

Газоанализаторы типа ГИП-5 изготавливаются для анализа газов, указанных в таблице:

Анализируемый компонент	Газовая смесь, в которой определяется компонент	Показания шкалы прибора, %	Погрешность в показаниях прибора, % от диапазона шкалы
Оксид углерода (CO)	Воздух производственных помещений	0 - 0,05	± 10
Оксид углерода (CO)	Азото-водородная смесь в производстве синтетического аммиака	0 - 0,1	± 5
То же	Конвертированный газ в производстве синтетического аммиака	0 - 10	± 5
Двуокись углерода (CO ₂)	Азото-водородная смесь в производстве синтетического аммиака	0 - 0,1	± 5
То же	Газ после водной очистки в производстве синтетического аммиака	0 - 5	± 5
Метан	Циркуляционная азото-водородная смесь в производстве синтетического аммиака	0 - 10	± 5
»	То же	0 - 20	± 5
Аммиак	»	0 - 5	± 5
»	»	0 - 15	± 5
»	»	0 - 25	± 5

- В комплект газоанализатора ГИП-5 входят:
- а) датчик;
 - б) стабилизатор напряжения;
 - в) фильтр воздуха;
 - г) амортизатор (поставляется по отдельному заказу).

ТЕХНИЧЕСКАЯ ХАРАКТЕРИСТИКА

1. Питание — от сети переменного тока напряжением $220 \text{ в} \pm 10\%$; частота $50 \text{ гц} \pm 1\%$; потребляемая мощность $\sim 90 \text{ вт}$.
2. Постоянная времени прибора 30 сек.
3. Нормальная работа прибора обеспечивается при входе в прибор газа с давлением 100 мм вод. ст.; давление газа выше 190 и ниже 10 мм вод. ст. не допускается.
4. Температура анализируемой смеси должна быть в пределах от $+10$ до $+35^\circ\text{C}$.
5. Расход анализируемого газа $\sim 100 \text{ л/час}$.
6. Прибор предназначен для работы при температуре окружающего воздуха от $+10$ до $+35^\circ\text{C}$ и относительной влажности от 30 до 80%.
7. В газе, поступающем в прибор, не должно содержаться агрессивных веществ и пыли (сероводорода — не более $0,001 \text{ г/лм}^3$).
8. Габариты газоанализатора $420 \times 292 \times 506 \text{ мм}$, вес $\sim 45 \text{ кг}$. Габариты газоанализатора с амортизатором $554 \times 549 \times 292 \text{ мм}$, вес $\sim 55 \text{ кг}$.

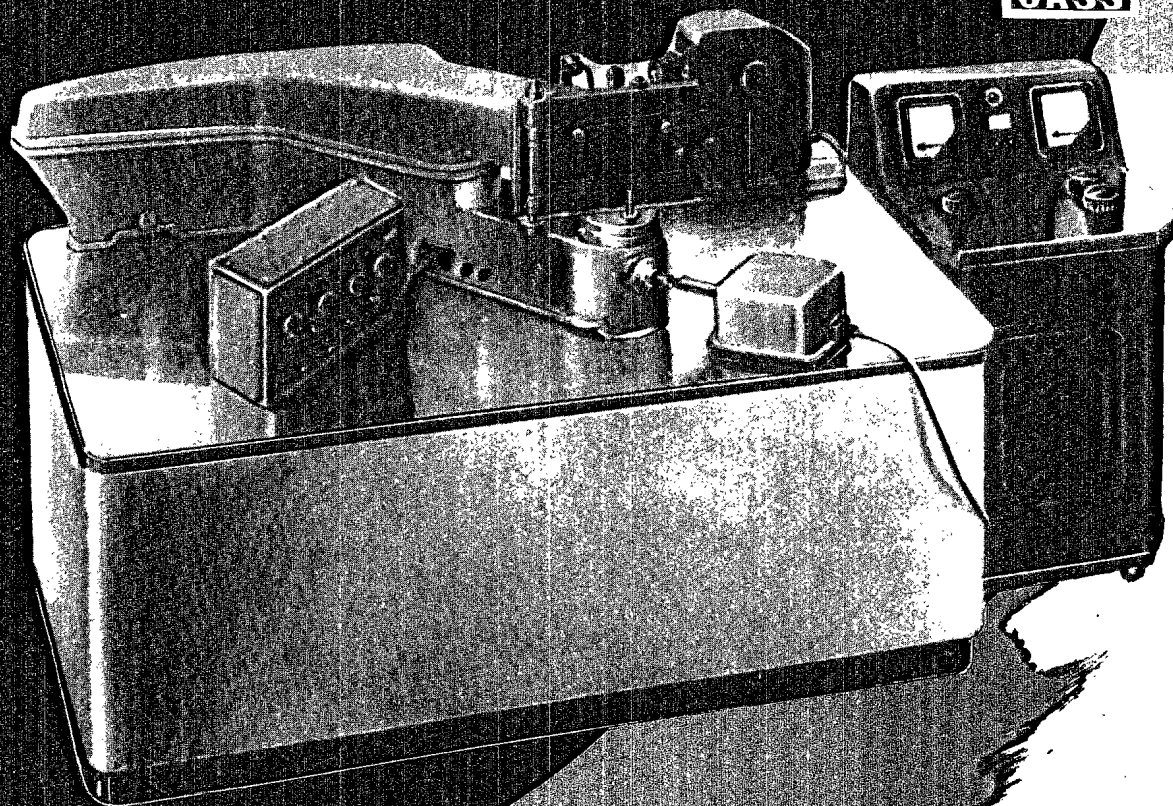
Г04400

ГОСХИМИЗДАТ

№ 356

Зак. 646

Типография Госхимиздата. Москва, 88, Угрешская



Кварцевый
спектрограф
ИСТ-28

СЕКЦИЯ СССР НА ВСЕМИРНОЙ ВЫСТАВКЕ 1958 Г. В БРЮССЕЛЕ

МОИ-28 SPECTROGRAPHE A QUARTZ

Le spectrographe à quartz de dispersion moyenne est largement répandu en Union Soviétique. Il est construit pour l'analyse spectrale industrielle, ainsi que pour la résolution de différents problèmes scientifiques spectroscopiques.

L'appareil assure une haute qualité de l'image des raies spectrales et un spectre plan grâce à l'application du miroir sphérique de collimation et de l'objectif double à quartz.

Le spectre entier dans l'intervalle entre 2000 Å et 6000 Å peut être photographié sur une plaque plane de 9x24 cm.

La dispersion du spectrographe est suffisante pour l'analyse spectrale des aciers à haute et à moyenne teneur en éléments alliés. Pour 2500 Å la dispersion linéaire inverse est de 9 Å/mm.

Le spectrographe complet comprend un système achromatique d'illumination à trois lentilles, un condenseur à quartz, un atténuateur à 3 et à 9 étages.

L'appareil résoud nettement le triplet 3100 Fe dont la distance entre les raies est de 0,34 Å (grossissement 20x)

The instrument provides sufficient resolution of iron triplet 3100 with distances between lines equal to 0.34 Å (magnification 20x)

Das Gerät löst das Eisentriplett 3100 Å mit einem Abstand zwischen den Linien 0,34 Å gut auf (20fache Vergrößerung)

QUARTZ SPECTROGRAPH



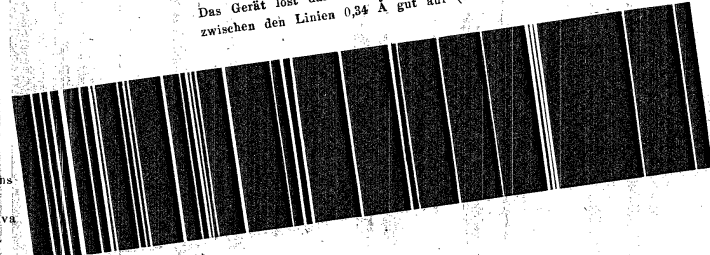
Partie du spectre du Fe dans la région spectrale de 3050 Å à deux poses (grossissement 10x)
The iron spectrum in the region 3050 Å, with two exposures (magnification 10x)
Spektrum des Eisens im Bereich von 3050 Å mit 2 Belichtungs-Zeiten (10fache Vergrößerung)

Widely used in the USSR the quartz spectrograph of a medium dispersion MOI-28 is designed for industrial spectrum analysis as well as for different scientific purposes.

Due to application of a spherical collimator mirror and double quartz camera objective the instrument displays high quality of spectral lines image and flat field of spectrum. The whole spectrum in the range from 2000 to 6000 Å is photographed on 9x24 cm plate.

Dispersion of spectrograph is sufficient to hold spectral analysis of medium alloy and alloy steels. For 2500 Å the reverse linear dispersion is 9 Å/mm.

The spectrograph set contains achromatic three-lense illuminating system, a quartz condenser, three- and nine-stage attenuators.



Pour les conditions des livraisons adressez-vous:
STANKOIMPORT, Moskva
G-200, Smolenskaja-Sennaja
ploschad, 32/34.

For delivery terms apply to
"STANKOIMPORT", Moskva
G-200, Smolenskaja-Sennaja
ploschad, 32/34

QUARZ-SPEKTROGRAPH ИСП-28

Der in der UdSSR weitverbreitete Quarz-Spektrograph der mittleren Dispersion ИСП-28 ist für die industrielle Spektralanalyse und für die Lösung verschiedener wissenschaftlichen spektroskopischer Aufgaben bestimmt.

Durch die Anwendung des sphärischen Kollimatorspiegels und des Doppel-Quarz-Objektivs wird im Gerät eine hohe Qualität des Spektrallinien-Bildes und ein flaches Spektrumfeld erreicht.

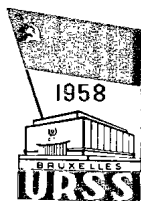
Das gesamte Spektrum im Bereich 2000—6000 Å wird auf eine 9×24 cm-Platte photographiert.

Die Dispersion des Spektrographen ist genügend, um Spektralanalysen von mittel-legierten und legierten Stählen durchzuführen. Für 2500 Å beträgt die reziproke lineare Dispersion 9 Å/mm.

Zum Spektrographensatz gehört ein Dreilinsen-Beleuchtungssystem, ein Quarz-Kondensator, drei- bzw. neunstufige Abschwächer.

Alle Lieferungsanfragen sind an
„STANKOIMPORT“, Moskva.
G-200, Smolenskaja-Sennaja
Ploschad, 32/34 zu richten

Духинской
интерактивный
спектрометр
УКС-14



СЕКЦИЯ СССР НА ВСЕМИРНОЙ ВЫСТАВКЕ 1958 Г. В БРЮССЕЛЕ

SPECTROMÈTRE INFRAROUGE

L'appareil est automatique, à deux faisceaux, destiné à l'enregistrement des spectres d'absorption des substances solides, liquides et gazeuses dans l'infrarouge. Les résultats sont exprimés directement en pourcentage de la transmission.

Il peut aussi fonctionner comme appareil à un faisceau.

Dans l'appareil est appliquée l'optique asphérique à miroir.

Le récepteur de radiation – un bolomètre métallique à vide à une sensibilité de seuil de $5 - 7 \cdot 10^{-10}$ W.

Un assortiment de prismes en verre, en fluorure de lithium, en chlorure de sodium et en bromure de potassium assure le fonctionnement dans l'intervalle compris entre 0,75 et 25 μ .

L'ouverture des fentes est automatique.

Le spectre est enregistré sur une bande de papier en échelle régulière des longueurs d'ondes ou en échelle des nombres d'ondes. En même temps des points de repère sont portés automatiquement sur la bande avec les intervalles de 5, 10 ou 20 divisions du tambour des longueurs d'ondes.

L'équipement complet comprend un assortiment de cuvettes de différentes épaisseurs pour les liquides et les gaz.

Pour les conditions
des livraisons adressez-vous:
"STANKOIMPORT", Moskva, G-200,
Smolenskaja-Sennaja ploshchad, 32/34.

INFRARED SPECTROMETER

Serial automatic double-beam spectrometer is designed for recording absorption spectra of solid, liquid and gaseous substances in the infrared region. Reads directly in transmission percentage.

The instrument can also operate by using the single-beam scheme.

Aspheric mirror optical elements are used.

Radiation receiver – a metallic vacuum bolometer whose threshold sensitivity is $5 - 7 \times 10^{-10}$ W.

A set of glass prisms, lithium fluoride, sodium chloride and potassium bromide enables operation in the range from 0.75 to 25 microns.

Slits are opened automatically.

Spectrum is recorded on a paper strip, a uniform scale of wave lengths or wave numbers being used.

Simultaneously with spectrum recording reference points are made on the strip at interval of 5, 10 or 20 divisions of the wave lengths reel.

A set of containers of various thickness is provided for liquids and gases.

For delivery terms apply to
"STANKOIMPORT", Moskva, G-200,
Smolenskaja-Sennaja ploshchad, 32/34.

Das serienmäßige automatische Zweistrahl-Spektrometer ist für die Registrierung der Absorptions-Spektren von festen, flüssigen und gasförmigen Stoffen im infraroten Gebiet bestimmt.

Die Ergebnisse werden unmittelbar in % der Durchlässigkeit abgelesen.

Das Gerät kann auf Wunsch auch nach dem Einstrahl-Schema arbeiten.

Im Gerät ist die asphärische Spiegeloptik angewandt. Als Strahlungsempfänger dient ein Metall-Vakuum-Bolometer mit einer Grenzempfindlichkeit $5 - 7 \cdot 10^{-10}$ Watt.

Die Arbeit im Bereich von 0,75 bis zu 25 μ wird durch einen Prismensatz aus Glas, LiF, NaCl und KBr gewährleistet.

Die Spalten werden automatisch geöffnet.

Das Spektrum wird auf einem Papierstreifen nach einer gleichmäßigen Wellenlängen-Skala bzw. nach einer Wellenzahlen-Skala aufgeschrieben.

Gleichzeitig mit der Registrierung des Spektrums werden auf dem Streifen Marken mit einem Intervall von 5,10 oder 20 Skalenteile der Wellenlängen-Trommel aufgezeichnet.

Das Gerät ist mit einem Satz Küvetten verschiedener Dicke für Flüssigkeiten und Gase ausgestattet.

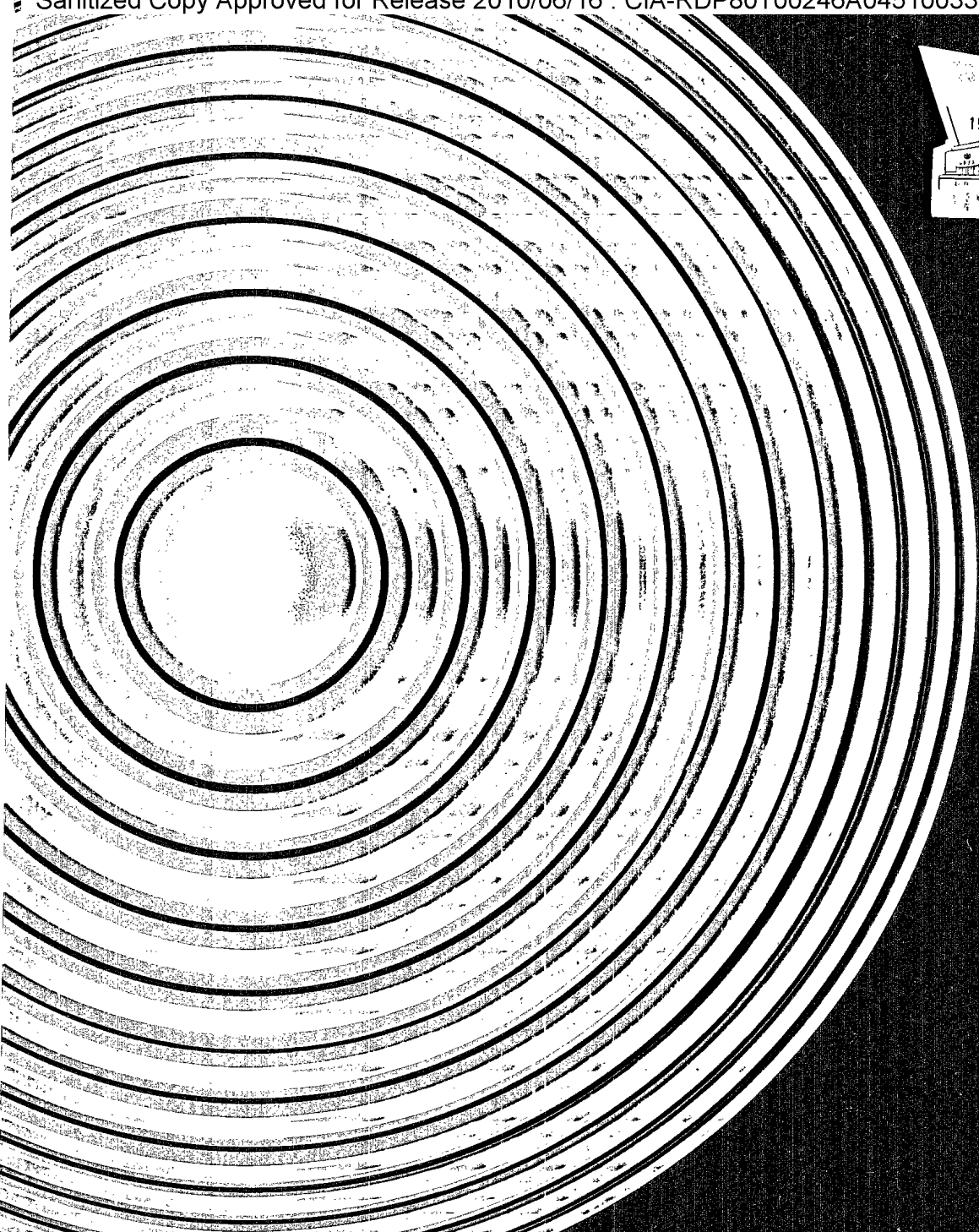
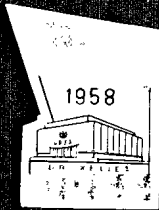
**INFRAROT-
SPEKTROMETER**

UKC-14

Alle Lieferungsanfragen sind an
„STANKOIMPORT“, Moskva, G-200,
Smolenskaja-Sennaja Ploschad, 32/34
zu richten.

ABTEILUNG DER UdSSR AUF DER ALLGEMEINEN WELTAUSSTELLUNG IN BRÜSSEL 1958

Внешторгиздат. Заказ № Б-361-26



Интерферометр
ИТ-28

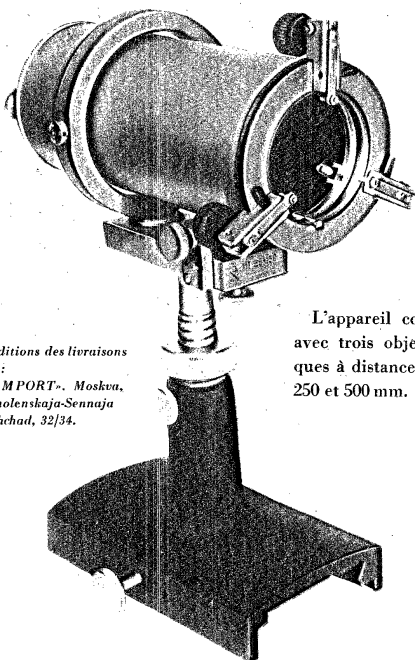
СЕКЦИЯ СССР НА ВСЕМИРНОЙ ВЫСТАВКЕ 1958 Г. В БРЮССЕЛЕ
SECTION DE L'URSS A L'EXPOSITION UNIVERSELLE ET INTERNATIONALE DE BRUXELLES 1958

INTERFEROMETRE
ИТ-28

L'interféromètre ИТ-28 à distance variable du type Fabry et Pérot, est destiné aux travaux différents comprenant les études de la fine structure et des contours des raies spectrales.

Il est possible de livrer l'appareil de deux types:
ИТ-28/30 — avec un assortiment de 18 anneaux de raccordement de 0,3 à 30 mm d'épaisseur;
ИТ-28/150 — avec un assortiment de 5 anneaux de raccordement de 40 à 150 mm d'épaisseur.

Chacun de ces deux types est muni de deux paires de coins à miroir: pour la région du spectre de 2000—3500 Å, une paire de coins en quartz cristallin avec un revêtement d'aluminium dont le coefficient de réflexion est de 80—84%; pour la région du spectre de 3200—5600 Å, une paire de coins en quartz fondu, avec un revêtement à plusieurs couches assurant le coefficient de réflexion de 87—92%.



Pour les conditions des livraisons adressez-vous:
„STANKOIMPORT“, Moskva,
G - 200, Smolenskaja-Sennaja
Ploshchad, 32/34.

L'appareil complet est livré avec trois objectifs achromatiques à distances focales de 150, 250 et 500 mm.

Interferometer ИТ-28 of the Fabry and Perot type with alternating distance is used for various kinds of work connected with study of fine structure and contours of spectral lines.

Two versions are manufactured:
ИТ-28/30 — with a set of 18 rings from 0.3 to 30 mm thick;
ИТ-28/150 — with a set of 5 rings from 40 to 150 mm thick.

Each of these has two couples of mirror wedges: for spectral region 2000—3500 Å — a couple of crystalline quartz with alluminium coating having reflection power of 80—84%; for spectral region 3200—5600 Å — a couple of fused quartz with multi-layer coating providing reflection power of 87—92%.

The set includes three achromatic objectives with focal distances of 150, 250 and 500 mm.

INTERFEROMETER
ИТ-28

For delivery terms apply to
„STANKOIMPORT“, Moskva,
G - 200, Smolenskaja-Sennaja
Ploshchad, 32/34.

Das Interferometer ИТ-28 nach Fabry und Perot mit variablem Abstand ist für die Durchführung von verschiedenen Arbeiten bestimmt, die mit der Untersuchung von Feinstruktur und von Umrisen der Spektrallinien verbunden sind.

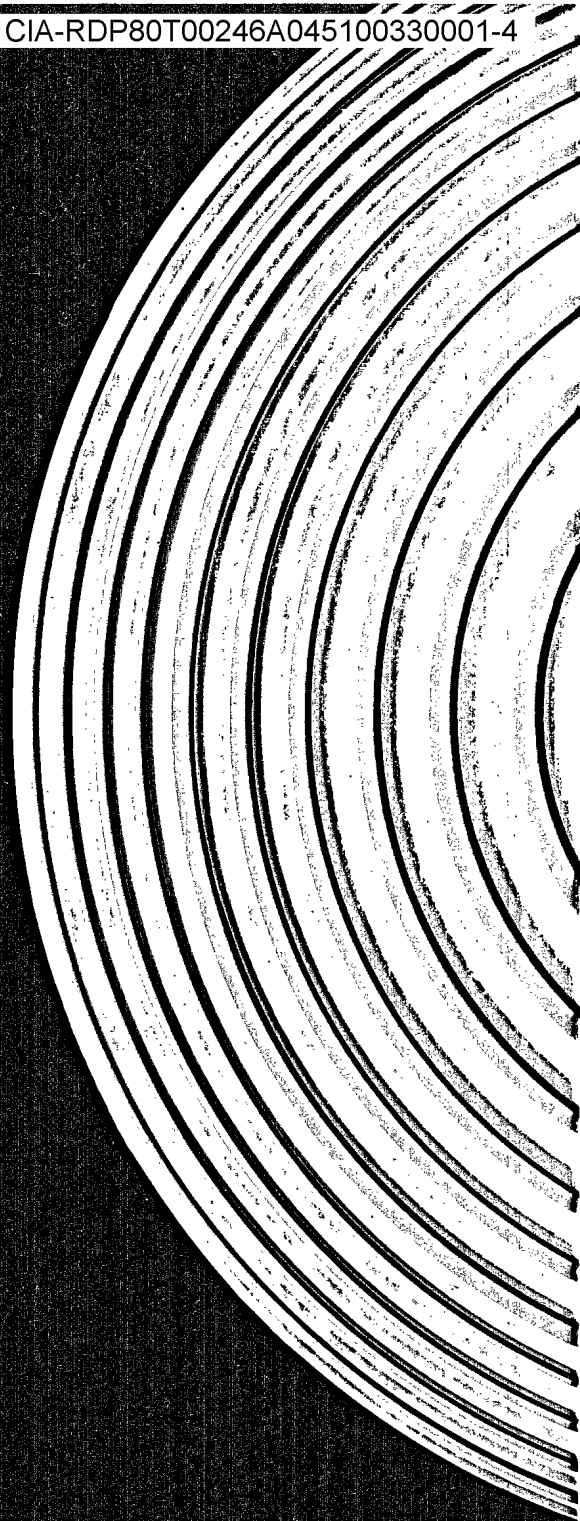
Das Gerät wird in zwei Varianten hergestellt:
ИТ-28/30 mit einem Satz aus 18 Zwischenringen von 0.3 bis 30 mm Dicke;
ИТ-28/150 mit einem Satz aus 5 Zwischenringen von 40 bis 150 mm Dicke.

Jedes Gerät besitzt zwei Paare Spiegelkeile: für den Spektralbereich 2000—3500 Å — ein Paar aus kristallinem Quarz mit Aluminium-Überzug mit einem Reflexionsfaktor 80—84%; für den Spektralbereich 3200—5600 Å — ein Paar aus geschmolzenem Quarz mit mehrschichtigem Belag, der einen Reflexionsindex von 87—92% gewährleistet.

Zu dem Gerätsatz gehören drei achromatische Objektive mit einer Brennweite von 120, 250 und 500 mm.

Alle Lieferungsanfragen sind an
„STANKOIMPORT“, Moskva,
G - 200, Smolenskaja-Sennaja
Ploshchad, 32/34, zu richten.

ИТ-28
INTERFEROMETER



USSR SECTION: BRUSSELS UNIVERSAL AND INTERNATIONAL EXHIBITION 1958

ABTEILUNG DER U.S.S.R. AUF DER ALLGEMEINEN WELTAUSSTELLUNG IN BRÜSSEL 1958

Внешторгиздат. Заказ № 5 326-26



КОНЦЕНТРАТОМЕР

ТИПА КСО-2

ДЛЯ СЕРНОЙ КИСЛОТЫ
И ОЛЕУМА

КОНЦЕНТРАТОМЕР ТИПА КСО-2 ДЛЯ СЕРНОЙ КИСЛОТЫ И ОЛЕУМА

Концентраномер типа КСО-2 представляет собою автоматический прибор для непрерывного определения концентрации серной кислоты и олеума, протекающих через датчик прибора. Концентраномер предназначен для установки на заводах, где может быть обеспечено непрерывное протекание серной кислоты или олеума через датчик прибора. Прибор может быть применен также для определения концентрации различных кислот и солей, электропроводность которых зависит от концентрации.

● Прибор состоит из следующих основных частей:

1. Датчика концентрамера.
2. Вторичного прибора—стандартного автоматического равновесного моста с несколько видоизмененной электрической схемой.

● Принцип действия концентрамера основан на измерении электропроводности серной кислоты. Измерение производится по схеме равновесного моста переменного тока, в котором двумя плечами моста служат электродные ячейки, установленные в датчике прибора. Шкала вторичного прибора градуирована в % H_2SO_4 или % SO_3 (своб.). Для кислоты датчик изготавливается из чугуна и состоит из корпуса, внутри которого вставляется стакан, в котором закреплены два измерительных электрода и сравнительная ячейка для температурной компенсации.

- Датчик для олеума изготавливается из обычной стали.
- Пределы шкалы и погрешность прибора приведены в следующей таблице:

№ п.п.	Наименование прибора	Пределы шкалы	Погрешность прибора
1	Концентраномер на моногидрат	95—99% H_2SO_4	$\pm 0,2\%$ H_2SO_4
2	Концентраномер на сульфилъную кислоту . . .	93—97% H_2SO_4	$\pm 0,3\%$ H_2SO_4
3	Концентраномер на олеум	15—25% SO_3	$\pm 0,5\%$ SO_3 своб.

Sanitized Copy Approved for Release 2010/06/16 : CIA-RDP80T00246A045100330001-4

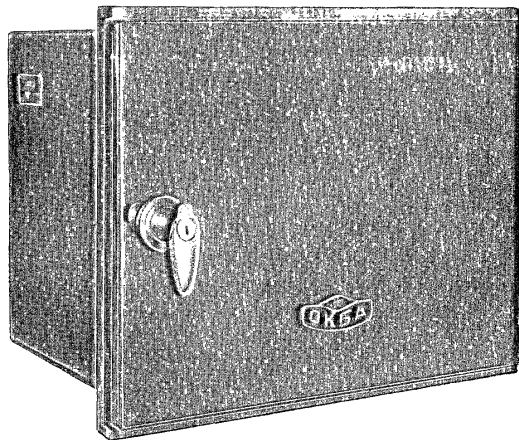
Sanitized Copy Approved for Release 2010/06/16 : CIA-RDP80T00246A045100330001-4



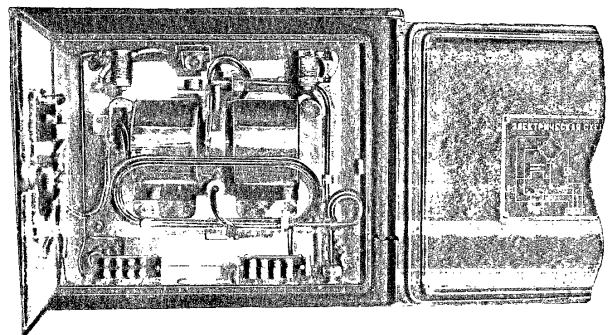
МАГНИТНЫЙ ГАЗОАНАЛИЗАТОР ТИПА МГК-2

МАГНИТНЫЙ ГАЗОАНАЛИЗАТОР ТИПА МГК-2

Магнитный газоанализатор типа МГК-2 предназначен для непрерывного определения или регулирования содержания кислорода в хвостовых газах при производстве азотной кислоты, серной кислоты и уксусной кислоты, а также для контроля обогащенного дутья, топочных газов и газов различных окислительных процессов.



- Работа газоанализатора основана на принципе термомагнитной конвекции. (Способность кислорода втягиваться в магнитное поле, т. е. его парамагнитные свойства с увеличением температуры уменьшаются. При наличии температурного перепада с двух сторон магнитного поля возникает термомагнитная конвекция.)
- Основной узел датчика—магнитный блок—помещен в корпус из антимагнитного материала. В магнитном блоке между полюсами постоянных магнитов эксцентрично установлена кольцевая камера (из стали ЭЯ1Т), по диаметру которой горизонтально расположена перемычка—стеклянная трубка. На трубке размещены две секции нагревательной обмотки из



Внешний вид датчика газоанализатора МГК-2 с открытой крышкой.

тонкой платиновой проволоки, которые включены как два плеча в неравновесный мост Уитстона и служат одновременно измерительными элементами.

- Анализируемая газовая смесь, поступающая в датчик, проходит кольцевую камеру. При наличии в газе кислорода, который обладает парамагнитными свойствами, в горизонтальной трубке возникает перепад давления. Повышенное давление создается со стороны, расположенной ближе к полюсам магнитов, так как в этом месте газ имеет более низкую температуру. При движении газа в трубке изменяется сопротивление рабочих плеч моста; это вызывает нарушение баланса моста, величина разбаланса пропорциональна термомагнитной конвекции, а следовательно, и концентрации кислорода в анализируемой смеси.

- В комплект прибора входят:
 - а) датчик прибора;
 - б) электронное реле типа ЭлР-1;
 - в) источник питания СИП-5;
 - г) панель подачи газа ППГ-2;
 - д) вторичный прибор типа ЭПД-12 или ЭПД-32 на 17 мм.

ТЕХНИЧЕСКАЯ ХАРАКТЕРИСТИКА

- 1. Шкалы газоанализатора:

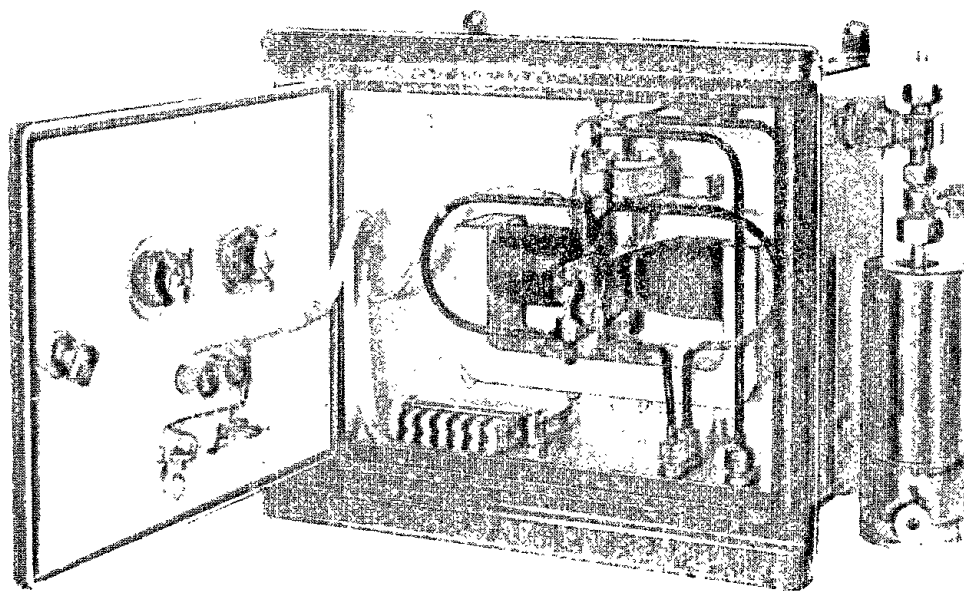
от 0 до 5% кислорода	от 20 до 35% кислорода
» 0 » 10% »	» 20 » 60% »
» 0 » 21% »	» 20 » 80% »

- 2. Погрешность в показаниях прибора составляет $\pm 5\%$ от диапазона шкалы.
- 3. Постоянная времени прибора равна 25 сек. (без учета запаздывания в газоподводящей линии и очистных устройствах).
- 4. Питание—от сети переменного тока напряжением 220 в $\pm 10\%$; потребляемая мощность 200 *вт*.
- 5. Подводимая к газоанализатору газовая смесь должна иметь давление не менее 50 мм вод. ст.; колебания допускаются не более $\pm 30\%$.
- 6. Прибор предназначен для работы в невзрывоопасных помещениях при температуре окружающего воздуха от $+10$ до $+35^\circ$ и относительной влажности до 80%.
- 7. Расход газа через прибор составляет 200 л/час.
- 8. Поступающая в прибор смесь не должна содержать пыли, смолы, серы и ее соединений, соединений железа, тумана щелочей и кислот.
- 9. Габариты и вес узлов прибора:

Узлы прибора	Длина мм	Ширина мм	Высота мм	Вес кг
Датчик	378	330	355	21
Электронное реле типа ЭлР-1	276	130	190	7
Источники питания типа СИП-5	416	165	285	16,6
Вторичный прибор типа ЭПД	420	292	506	29
Панель подачи газа с регулирующим ротаметром	130	185	235	3



МАГНИТНЫЙ ГАЗОАНАЛИЗАТОР ТИПА МГК-3



МАГНИТНЫЙ ГАЗОАНАЛИЗАТОР типа МГК-3

ТЕХНИЧЕСКАЯ ХАРАКТЕРИСТИКА

Магнитный газоанализатор типа МГК-3 предназначен для непрерывного определения или регулирования содержания кислорода в многокомпонентных газовых смесях. Содержание балластных компонентов и их физические свойства не вызывают дополнительной погрешности показаний.

Магнитный газоанализатор типа МГК-3 применяется в условиях, когда работа термоманитного газоанализатора невозможна, например в промышленности синтетического каучука и синтетического спирта.

Действие газоанализатора основано на использовании парамагнитных свойств кислорода, благодаря которым в магнитном поле истечению сравнительного газа, т. е. газа с постоянным содержанием кислорода, а следовательно, с постоянными магнитными свойствами, оказывается различное сопротивление, в зависимости от содержания кислорода в анализируемом газе. Это сопротивление вызывает перепад давления, служащий мерой содержания кислорода в анализируемом газе.

Основными узлами датчика являются магнитный блок с камерой, к которой подводятся анализируемый и сравнительный газы, и термоанемометр, служащий для измерения перепада давления. Два рабочих плеча термоанемометра, изготовленные из тонкой платиновой проволоки, включены в измерительный мост, разбаланс которого измеряется с помощью вторичного прибора.

Непосредственно к датчику укреплено нагревательное устройство, в котором продуваемый через датчик воздух нагревается до необходимой температуры. Продувка теплым воздухом обеспечивает термостатирование датчика и одновременно его взрывобезопасность.

В комплект прибора входят:

- датчик прибора;
- электронное реле типа ЭЛР-1;
- источник питания типа СИП-5;
- панель подготовки газа;
- реле давления типа РЛД;
- редуктор высокого давления типа РГ-1 и редуктор низкого давления типа РГ-2, скомпонованные вместе;
- вторичный прибор — электронный потенциометр типа ЭПД, ЭПП, ПС или ПСР.

1. Шкалы газоанализатора:

от 0 до 5%	кислорода
„ 0 до 10%	„
„ 0 до 21%	„
„ 15 до 45%	„

2. Основная погрешность в показаниях прибора составляет $\pm 5\%$ от максимального значения шкалы.

3. Запаздывание начала реагирования 20 сек., время установления показаний 1,5 мин. без учета влияния газоподводящей линии и очистных устройств.

4. Питание — от сети переменного тока напряжением $220 \text{ в} \pm 10\%$; потребляемая мощность 400 *ва*.

5. Прибор предназначен для работы в помещениях с температурой окружающего воздуха от $+10$ до $+35^\circ\text{C}$ и относительной влажностью до 80%.

6. Давление газовой смеси, подводимой к газоанализатору, должно быть не менее 50 *мм вод. ст.*; допускаются колебания давления не более $\pm 30\%$.

7. Для работы прибора необходим баллон со сравнительным газом (азот или воздух, в зависимости от диапазона шкалы).

8. Расход анализируемого газа через прибор составляет 200 *л/ч*; расход сравнительного газа через прибор равен 3 *л/ч*.

9. Анализируемая смесь, поступающая в прибор, не должна содержать пыли, смолы, серы, соединений железа, тумана щелочей и кислот.

10. Габариты и вес узлов прибора:

Узлы прибора	Ширина мм	Высота мм	Глубина мм	Вес кг
Датчик	378	355	330	22
Электронное реле	276	190	130	7
Источник питания	352	290	160	11
Вторичный прибор				
тип ЭПД	420	506	292	29
тип ПС или ПСР	330	287	404	22
Панель подготовки газа	365	250	200	8,5

Т 04400 ГОСХИМИЗДАТ. № 52 Зак. 618

Типография Госхимиздата. Москва, 88, Угрешская.



**ПНЕВМАТИЧЕСКИЙ
ИНДИКАТОР
РАСХОДА
ПИР-2М**

**МИНИСТЕРСТВО
ХИМИЧЕСКОЙ ПРОМЫШЛЕННОСТИ СССР**

МОСКВА

ПНЕВМАТИЧЕСКИЙ ИНДИКАТОР РАСХОДА
ПИР-2М

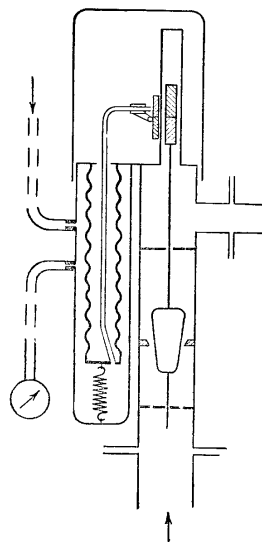
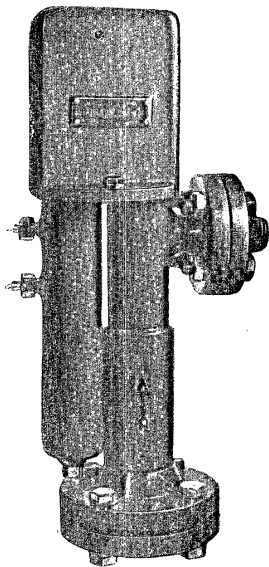


Схема прибора.

Прибор ПИР-2М представляет собой бесшкальный ротаметрический датчик с магнито-пневматическим преобразователем. Основным назначением прибора является дистанционное измерение и регулирование расхода жидкостей в условиях химических производств (при работе с агрессивными, горючими, токсичными жидкостями и растворителями во взрывоопасных условиях).

● Принцип действия прибора состоит в том, что поплавок ротаметра, передвигающийся на высоту, пропорциональную скорости протекающей жидкости, перемещает постоянный магнит, армированный в коррозионноустойчивый немагнитный материал. Этот магнит через трубку, выполненную из немагнитного материала, взаимодействует с постоянным магнитом следящей пневматической системы, перемещая его соответственно перемещению поплавка, благодаря чему регулируется выход воздуха из внутренней полости сильфонного сервомотора через систему «сопло-заслонка». При этом давление воздуха во внутренней полости сервомотора, передаваемое на вторичный прибор, пропорционально подъему поплавка и, следовательно, расходу протекающей через прибор жидкости.

● Прибор изготавливается без сальниковых и прокладочных соединений.

● В настоящее время выпускаются опытные партии приборов.

● Датчик ПИР-2М комплектуется следующими приборами:

- а) фильтром воздуха ФВ-2;
- б) редуктором давления воздуха МРДВ-1;
- в) контрольным манометром со шкалой $0-3 \text{ кг/см}^2$;
- г) вторичным сильфонным манометром МС-410 (МС-610, 04-МС-610 и т. п.) со шкалой $0-1 \text{ кг/см}^2$.

ТЕХНИЧЕСКАЯ ХАРАКТЕРИСТИКА

- 1. Максимальное рабочее давление равно 160 кг/см^2 .
- 2. Рабочая температура—до 120°C .
- 3. Шкала расхода—на 500, 1000, 1500, 2000, 3000 л/час (определяется размерами конусного поплавка).
- 4. Прибор может изготавливаться из стали 1Х18Н9Т или другого материала, не корродирующего в среде рабочей жидкости.
- 5. Питание—сжатым воздухом при давлении $1,1 \pm 0,1 \text{ кг/см}^2$; расход воздуха $0,25 \text{ нм}^3/\text{час}$.
- 6. Максимальное расстояние от датчика до вторичного прибора 100 м.
- 7. Погрешность в показаниях прибора составляет $\pm 2,5\%$.
- 8. Вес прибора 16 кг.

К печати 29/IV 1957 г. Т 01496. ГОСХИМИЗДАТ № 1-19. Зак. 310

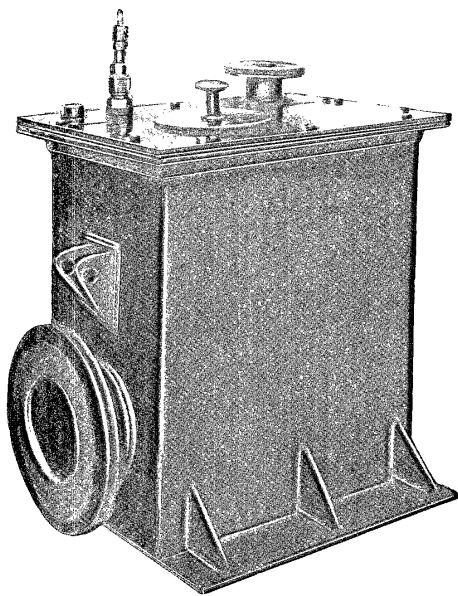
Типография Госхимиздата, Москва, 88, Угрешская



РАСХОДОМЕР ЩЕЛЕВОЙ ТИПА РМ

РАСХОДОМЕР ЩЕЛЕВОЙ ТИПА РМ

Щелевые расходомеры типа РМ являются стационарными приборами, предназначенными для автоматического непрерывного измерения и записи весового расхода жидкости, находящейся под атмосферным давлением, или для измерения, записи



Датчик расходомера типа РМ-3.

и автоматического регулирования весового расхода жидкости (в случае применения регулирующего вторичного прибора).

- Мгновенный весовой расход жидкости определяется по величине напора жидкости перед сливной щелью специального

профиля, обеспечивающего прямолинейную зависимость расхода от величины напора. Напор измеряется пьезометрической трубкой, через которую продувается воздух. Давление в пьезометрической трубке пропорционально напору, а следовательно, и весовому расходу жидкости.

- Вторичным прибором расходомера служит измеритель давления типа ЭМИД, состоящий из дифференциального манометра и электрического прибора, изготовленного на основе электронного потенциометра типа ЭПД или электронного моста ЭМД. Шкала прибора градуируется расчетным путем в $t/час$.
- Датчики расходомеров изготавливаются из различных материалов:

1. Расходомер типа РМ-1 из стали марки ЭЯ1Т.
2. Расходомер типа РМ-3 из свинца.

В комплект расходомера входят:

- а) датчик расходомера (один из указанных типов);
- б) измеритель давления (типа ЭМИД);
- в) блок питания воздухом (типа БПВЩ-1);
- г) фильтр для воздуха (типа ФВ-10).

ТЕХНИЧЕСКАЯ ХАРАКТЕРИСТИКА

- 1. Максимальный расчетный расход щелевых расходомеров типа РМ: 10, 20, 30 и 50 $m^3/час$.
- 2. Максимальные значения по шкале прибора: 8; 10; 12,5; 16; 20; 25; 30; 40; 50; 60 и 80 $t/час$ (при удельном весе измеряемой жидкости в пределах от 0,8 до 1,5).
- 3. Погрешность в показаниях прибора (с учетом погрешности ЭМИД) составляет $\pm 3,5\%$ от максимального значения шкалы.
- 4. Питание электроэнергией—от сети переменного тока напряжением 220 в; частота 50 гц; потребляемая мощность 50 вт.
- 5. Питание воздухом осуществляется от магистрали сжатого воздуха давлением 2—10 $кг/см^2$; расход воздуха составляет 0,6 $л/час$.
- 6. Максимальная длина импульсной трубки между датчиком и дифференциальным манометром 15 м; внутренний диаметр 4—8 мм.

- 7. Допустимая протяженность электрических проводов линии связи между дифференциальным манометром и вторичным прибором—не более 500 м. Линия связи—пятипроводная. Сечение провода—2,5 мм².
- 8. Измеряемый поток жидкости не должен иметь резких колебаний и пульсаций.
- 9. Жидкость не должна содержать механических примесей и в ней не должны образовываться осадки.
- 10. Габариты и вес датчика расходомера:

Датчики	Длина мм	Ширина мм	Высота мм	Вес кг
PM-1	580	280	800	55
PM-3	630	500	965	133

Специальный агент

СР Д-1



СЕКЦИЯ СССР НА ВСЕМИРНОЙ ВЫСТАВКЕ 1958 Г. В БРЮССЕЛЕ

СФД-1

SPECTROPHOTOMETRE

L'appareil est destiné à mesurer la transmission et la densité optique des substances liquides et solides dans la région du spectre de 220 à 1100 $m\mu$.

L'optique de l'appareil est construite d'après le schéma à un faisceau. L'objectif représente un monochromateur d'auto collimation à miroir parabolique hors d'axe (distance focale de 500 mm).

L'élément dispersant — une réplique prise du réseau de diffraction de 600 traits/mm, à concentration d'énergie de 1^{er} ordre, dans le domaine de 280 $m\mu$.

La dispersion linéaire inverse de l'appareil est près de 32 Å/mm.

Les mesures s'effectuent dans la région du spectre de 220 à 650 millimicrons au moyen d'une cellule photo-électrique à antimoine-césium; les mesures dans la région de 600 à 1100 millimicrons — au moyen d'une cellule photo-électrique à oxygène-césium. Le courant de la cellule photo-électrique est amplifié par l'application du schéma de courant continu.

Pour les études des liquides l'appareil est muni des cuvettes suivantes: cuvettes en quartz rectangulaires pour les couches de liquides de 10 mm d'épaisseur et cuvettes cylindriques démontables avec fenêtres en quartz ou en verre pour les couches de liquides de 5—100 mm d'épaisseur.

A l'aide d'intercalaires on peut avoir les couches de liquides de 0,05—1 mm d'épaisseur.

Les mesures des coefficients de transmission des substances liquides et solides s'effectuent au moyen d'un étalon dont le coefficient de la transmission est pris pour 100%. La valeur de la transmission de l'échantillon est lue sur l'échelle du potentiomètre à lecture directe.

Pour les conditions des livraisons adressez-vous:
„STANKOIMPORT“, Moskva, G-200, Smolenskaja-
Sennaja ploshchad, 32/34.

SPECTROPHOTOMETER СФД-1

The instrument is designed for measuring the permeability and optical density of liquid and solid substances in a spectral region of 220 to 1100 $m\mu$.

The optical part of the instrument is built according to one-beam diagram. Monochromator is the autocollimation one with a non-axial parabolic mirror as an objective (focal distance—500 mm).

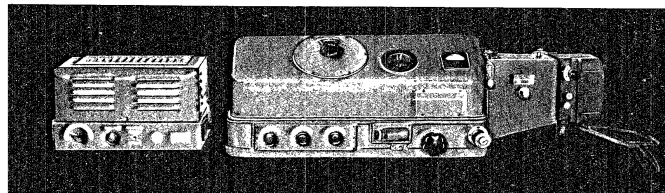
The dispersion element is the replica of the diffraction grating of 600 lines per mm with energy concentration of the first order in the region of 280 $m\mu$. The reverse linear dispersion is about 32 Å/mm.

Measurements in the spectral region of 220—650 $m\mu$ are made with the help of an antimony-cesium photocell; and measurements in the spectral region of 600—1100 $m\mu$ —with the help of an oxygen-cesium photocell. The photocell current is amplified by a D. C. amplifier.

For the investigation of liquids, the instrument set contains vessels: square quartz ones—for 10 mm thick layers of liquids—and split cylindrical ones with quartz and glass windows—for layers of liquids 5 to 100 mm thick. Layers of 0.05 to 1 mm can be obtained by means of inserts.

Measurements of permeability coefficients for liquid and solid substances are carried out in relation to the reference sample, the permeability coefficient of which is taken as 100%.

The value of permeability for a specimen is read on the scale of a reading potentiometer.



For delivery terms apply to "STANKO-
IMPORT", Moskva, G-200, Smolenskaja-
Sennaja ploshchad, 32/34.

SPEKTROPHOTOMETER

СФД-1

Das Gerät ist für die Messung der Durchlässigkeit und der optischen Dichte von flüssigen und festen Stoffen im Spektralbereich von 220 bis zu 1100 m μ bestimmt.

Der optische Teil des Geräts ist nach dem Einstrahl-System gebaut und hat als Objektiv einen Autokollimations-Monochromator mit Parabel-Spiegel (Brennweite 500 mm).

Das dispergierende Element ist ein Diffraktionsgitter 600 Str/mm mit Energie-Konzentration in der 1. Ordnung im Bereich von 280 m μ .

Die reziproke lineare Dispersion beträgt 32 Å/mm.

Die Messungen im Spektralbereich 220—650 m μ erfolgen mit einer Antimon-Cäsium-Photozelle, die Messungen im Spektralbereich 600—1100 m μ — mit einer Sauerstoff-Cäsium-Photozelle. Für die Verstärkung des Photozellenstroms wird das Gleichstromschema verwendet.

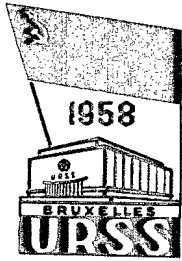
Für die Untersuchungen von Flüssigkeiten sind in den Gerätsatz folgende Küvetten eingeschlossen: rechteckige Quarz-Küvetten für die Flüssigkeitsschichten von 10 mm und zylindrische auseinandernehmbare mit Quarz- bzw. Glasfenstern für die 5 bis 100 mm starke Flüssigkeitsschichten. Mit Hilfe von Einlegestücken können Flüssigkeitsschichten von 0,05—1 mm Stärke erhalten werden.

Die Messungen der Durchlässigkeitsfaktoren von flüssigen und festen Stoffen werden im Vergleich mit einem Etalon ausgeführt, dessen Durchlässigkeit als 100% angenommen wird. Die Durchlässigkeitswerte der zu untersuchenden Probe werden auf der Skala des Ablese-Potentiometers abgelesen.

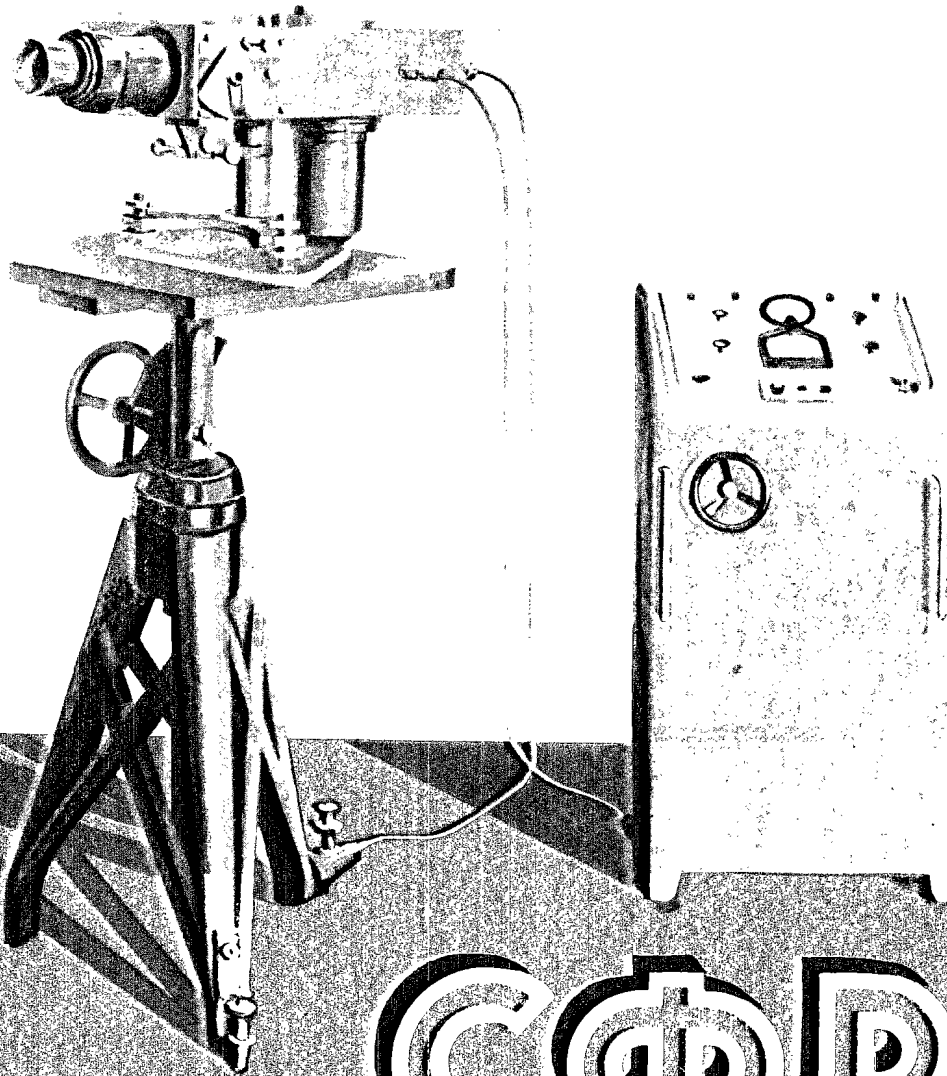
Alle Lieferungsanfragen sind an V/O „STANKO-IMPORT“, Moskva, G-200, Smolenskaja Sennaja Ploschad, 32/34 zu richten.

ABTEILUNG DER UdSSR AUF DER ALLGEMEINEN WELTAUSSTELLUNG IN BRÜSSEL 1958

Внешторгиздат. Заказ № Б. 376-26



С В Е Р Х С К О Р О С Т Н А Я Ф О Т О Р Е Г И С Т Р И Р У Ю Щ А Я У С Т А Н О В К А



СЕКЦИЯ СССР НА ВСЕМИРНОЙ ВЫСТАВКЕ 1958 Г. В БРЮССЕЛЕ



APPAREIL PHOTO-ENREGISTREUR EXTRA-RAPIDE

L'appareil COP est destiné à l'enregistrement de processus physique se passant rapidement.

L'appareil travaillant comme enregistreur photographique donne un développement continu du processus étudié.

L'appareil travaillant avec loupe coupant le temps d'exposition donne une série de photographies successives du processus étudié.

La largeur de l'interstice de l'enregistreur photographique est réglable.

L'obturateur électromagnétique est à double volet.

Le travail avec couplage de deux appareils est possible.

Alimentation électrique par courant alternatif 50 hertz, 220 volts.

Complexe:

- une caméra;
- un poste de commande;
- objectifs d'entrée de rechange;
- conduites électriques;
- pièces de rechange et accessoires.

Objectifs d'entrée de rechange

Objectifs	Distance focale, mm	Ouverture relative
« Industar-51 »	210	1: 4,5
Spécial	750	1: 5
Spécial	2000	1: 8

Capacité de résolution (dans le temps) de l'enregistreur photographique, sec.	10 ⁻⁸
Fréquence des photographies avec loupe coupant le temps d'exposition, images/sec.	25.10 ³ -25.10 ³
Largeur de la pellicule, mm	35
Longueur de la pellicule, m	1.6
Format des images, mm:	
enregistreur photographique (max)	25 x 375
avec loupe coupant le temps d'exposition	∅ 10 ou ∅ 5
Largeur de l'interstice, mm	0,01-0,5
Exposition depuis le début de l'événement, sec.	0,01; 0,03 et 0,1
Courant exigé, A	9
Encombrement (avec objectif F = 210 mm), mm:	
photo-enregistreur	670 x 415 x 345
loupe coupant le temps d'exposition	560 x 390 x 345
poste de commande	580 x 615 x 1200
Poids, kg:	
photo-enregistreur	42
loupe coupant le temps d'exposition	42
poste de commande	150

Pour les conditions des livraisons adressez-vous:
« STANKOIMPORT », Moskva, G-200,
Smolenskaja-Sennaja ploschad, 32/34.

Interchangeable inlet lenses

Lens	Focal distance, mm	Aperture ratio
"Industar-51"	210	1: 4,5
Special	750	1: 5
Special	2000	1: 8

Resolving power of photo-recorder (by of time), sec.	10 ⁻⁸
Exposure frequency in the "zeit-lupe" variant, frames per second	25.10 ³ -25.10 ³
Film width, mm	35
Length of film-strip, m	1.6
Frame dimensions, mm:	
in the photo-recorder variant (max)	25 x 375
in the "zeit-lupe" variant	∅ 10 or ∅ 5
Aperture width, mm	0.01-0.5
Exposure time from the moment of the phenomenon's initiation, sec.	0.01; 0.03 and 0.1
Current consumed, A	9
Overall dimensions (with f/210 mm lens), mm:	
of photo-recorder	670 x 415 x 345
of "zeit-lupe"	560 x 390 x 345
of control desk	580 x 615 x 1200
Weight, kg:	
of photo-recorder	42
of "zeit-lupe"	42
of control desk	150

For delivery terms apply to "STANKO-IMPORT", Moskva, G-200, Smolenskaja-Sennaja ploschad, 32/34.

SUPER-RAPID PHOTO-RECORDING APPARATUS

The COP apparatus is designed for the recording of transient physical processes.

When operating in the usual photo-recording apparatus routine, an incessant photo-sweep of the investigated process is obtained.

When operating in the "zeit-lupe" routine the apparatus gives a series of photo sequences of the process investigated.

The aperture width of the photo-recorder is controlled.

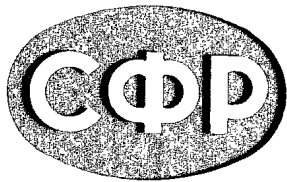
The two-petal shutter is electromagnetic.

A coupled operation of both the sets is quite possible.

Fed by 220 V, 50 hertz alternating current.

The unit:

- camera;
- control desk;
- interchangeable inlet lenses;
- wires;
- reserve parts and accessories.



DIE HOCHGESCHWINDIGKEITS- PHOTOREGISTRIERUNGS-ANLAGE

Die Anlage CΦP ist für Registrierung schnellablaufender physischer Prozesse bestimmt.

Bei Arbeit entsprechend dem Regime des Photoregistrators erhält man eine ununterbrochene Zerlegung des untersuchten Prozesses.

Bei Arbeit entsprechend dem Regime der Zeitlupe erhält man eine Reihe aufeinanderfolgender Photographien des untersuchten Prozesses.

Die Spaltbreite des Photoregistrators ist regulierbar.

Der Verschluss ist elektromagnetisch und besteht aus zwei Läppchen

Es ist eine gleichzeitige Kuppelung der Arbeit zweier Apparate möglich.

Elektrizitäts-Speisung: Wechselstrom 50 Hz, 220 V.

Auflösungs-Vermögen des Photoregistrators (in der Zeit), Sek.	10 ⁻⁸
Aufnahme-Frequenz in der Variante der Zeitlupe, Kader/Sek.	25.10 ³ -25.10 ⁵
Filmbreite, mm	35
Filmlänge, m	1,6
Kader-Dimensionen, mm:	
in der Variante des Photoregistrators (max.)	25 x 375
in der Variante der Zeitlupe	∅ 10 oder ∅ 5
Spaltbreite, mm	0,01-0,5
Expositions-dauer vom Momente der Injizierung der Erscheinung, Sek.	0,01; 0,03 und 0,1
Verbrauchsstrom, A	9
Gabarit-Dimensionen (mit einem Objekt F = 210), mm:	
Photoregistrator	670 x 415 x 345
Zeitlupe	560 x 390 x 345
Steuerpult	580 x 615 x 1200
Gewicht, kg:	
Photoregistrator	42
Zeitlupe	42
Steuerpult	150

Apparatur-Bestandteile:

- Kamera;
- Steuerpult;
- Auswechselbare Eingangs-Objektive;
- Elektrische Leitungsschnüre;
- Ersatzteile und Zubehör.

Auswechselbare Eingangs-Objektive

Benennung	Fokus-Abstand, mm	Relative Öffnung
„Industar-51“	210	1:4,5
Speziell	750	1:5
Speziell	2000	1:8

Alle Lieferungsanfragen sind an „STANKO-IMPORT“, Moskwa, G-200, Smolenskaja-Sennaja Ploschad, 32/34 zu richten.

ABTEILUNG DER UdSSR AUF DER ALLGEMEINEN WELTAUSSTELLUNG IN BRÜSSEL 1958

Внешторгиздат. Заказ № Б 377-26



Секция СССР на Всемирной выставке 1958 г. в Брюсселе

СПП-1

СЕКЦИЯ СССР НА ВСЕМИРНОЙ ВЫСТАВКЕ 1958 Г. В БРЮССЕЛЕ

SPECTROPROJECTEUR СПП-1

Le spectroprojecteur СПП-1 est un appareil auxiliaire pour effectuer l'analyse des images spectrales, obtenues à l'aide de différents spectrographes. C'est un dispositif de projection qui fournit une image agrandie des spectrogrammes sur un écran blanc de 300 mm de diamètre.

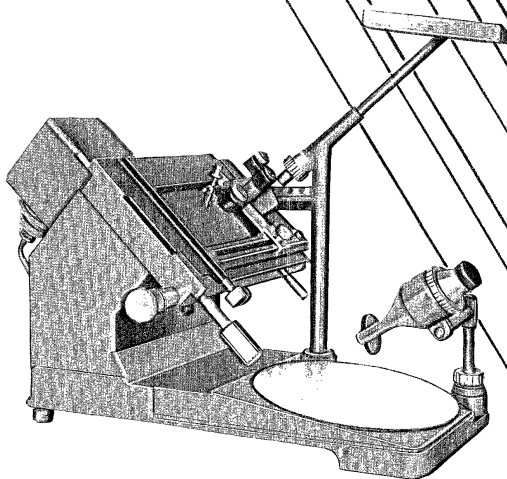
Application:
étude des spectres;
détermination approximative de la longueur d'ondes des raies spectrales;
photométrie visuelle des raies spectrales lors des analyses immédiates.
Les plaques de dimensions jusqu'à 130 x 240 mm et la pellicule de 35 mm sont fixées sur une platine porte-objet.

Les limites du déplacement de la platine sont transversal ± 55 mm, longitudinal ± 110 mm.

La précision de lecture du déplacement longitudinal est 0,1 mm.
Le grossissement du système est variable, 10^x et 20^x peut être réglé dans les limites de $\pm 1^{\circ}$.

L'appareil permet d'observer les spectres par plusieurs personnes simultanément dans un local non-obscurci.

L'appareil complet comprend un microscope à grossissement de 5^x et un dispositif pour les lectures des raies spectrales.



SPECTROPROJECTOR СПП-1

The СПП-1 is an auxiliary instrument for examining of spectra images obtained by different spectrographs. It represents a projecting device providing magnified image of spectrogram on the white screen of 300 mm in diameter.

The instrument design allows to carry out:
visual examination of spectra;
approximate determination of wavelength of spectral lines;
visual photometry of spectral lines during express analysis.
Plates up to 130 x 240 mm and 35 mm film are put on a stage which shifts ± 55 mm in transversal direction and ± 110 mm in longitudinal direction.

Accuracy of reading of longitudinal shifting is 0,1 mm.
Magnifications of system 10^x and 20^x are adjusted within $\pm 1^{\circ}$.

The instrument makes it possible to observe spectra in light premises by several persons simultaneously.

The set includes an attached 5^x microscope and a device for reading spectral lines.

SPKTROPROJEKTOR СПП-1

Spektroprojektor СПП-1 ist ein Hilfsgerät, das zur Durchsicht der mit den verschiedenen Spektrographen erhaltenen Spektrenbilder bestimmt ist.

Er stellt eine Projektionsanlage dar, die ein vergrößertes Spektrogramm-Bild auf einen weißen Schirm von 300 mm Durchmesser projiziert.

Das Gerät ermöglicht durchzuführen:
visuelle Untersuchung von Spektren;
angenäherte Bestimmung der Wellenlänge der Spektrallinien;
visuelle Photometrien der Spektrallinien bei Expressanalysen.

Die 130 x 240-mm-Platten und die 35-mm-Filme werden auf den Objektisch eingestellt.
Verstellbereich des Tisches: in Querrichtung ± 55 mm, in Längsrichtung ± 110 mm.

Ablesegenauigkeit der Längsverstellung - 0,1 mm.
Die Vergrößerung des Systems ist auswechselbar: 10^x und 20^x. Die Vergrößerung ist $\pm 1^{\circ}$ regulierbar.

Es ist möglich, die Spektren in einem nicht verdunkelten Raum gleichzeitig durch mehrere Personen zu beobachten.

Zu dem Gerätsatz gehört ein 5^x-Mikroskop und eine Vorrichtung zum Abzählen der Spektrallinien.

*Pour les conditions des livraisons adressez-vous:
«STANKOIMPORT» Moskva, G-200, Smolenskaja-Sennaja
ploschad, 32/34.*

*For delivery terms apply to "STANKOIMPORT".
Moskva, G-200, Smolenskaja-Sennaja ploschad. 32/34.*

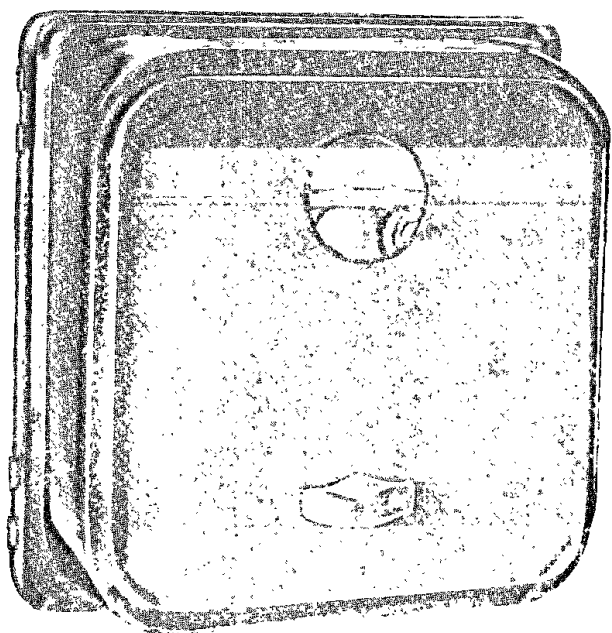
*Alle Lieferungsanfragen sind an „STANKOIMPORT“,
Moskva G-200, Smolenskaja-Sennaja Ploschad, 32/34
zu richten.*

ABTEILUNG DER UdSSR AUF DER ALLGEMEINEN WELTAUSSTELLUNG IN BRÜSSEL 1958

Внешторгиздат. Заказ № Б-373-26



ТЕРМОКОНДУКТОМЕТРИЧЕСКИЙ ГАЗОАНАЛИЗАТОР



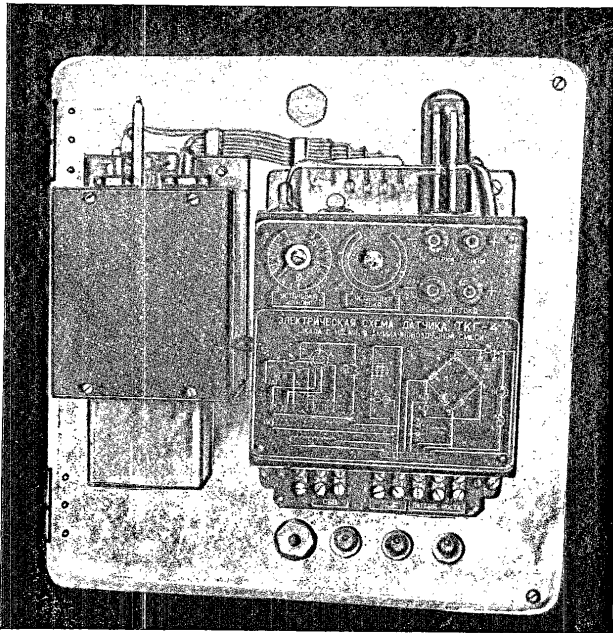
ТИПА

ТКГ-4

МИНИСТЕРСТВО
ХИМИЧЕСКОЙ ПРОМЫШЛЕННОСТИ СССР
МОСКВА

ТЕРМОКОНДУКТОМЕТРИЧЕСКИЙ ГАЗОАНАЛИЗАТОР ТИПА ТКГ-4

Термокондуктометрический газоанализатор типа ТКГ-4 представляет собой стационарный электрический автоматически действующий показывающий и регистрирующий прибор для непрерывного определения процентного содержания (по объему) одного компонента в газовой смеси путем сравнения теплопроводности анализируемого газа с теплопроводностью воздуха.



Внешний вид датчика газоанализатора (крышка снята).

● Сравнение теплопроводностей осуществляется с помощью неравновесного моста, два плеча которого помещены в рабочие газовые камеры, через которые протекает анализируемая

смесь, а два другие—в сравнительные камеры, заполненные воздухом. Все четыре плеча моста изготовлены из платины и нагреваются протекающим по ним током до определенной температуры. Для питания моста выпрямленным стабилизированным током используется источник питания типа СИП-4.

● При пропускании через газовые камеры анализируемой смеси, теплопроводность которой изменяется соответственно содержанию анализируемого компонента, электрическое сопротивление плеч моста, помещенных в эти камеры, меняется и в диагонали моста возникает ток, пропорциональный концентрации анализируемого компонента. В диагональ моста через делитель напряжения включен вторичный прибор типа ЭПД-12 или ЭПД-32.

● Постоянство расхода газа, проходящего через рабочие камеры прибора, обеспечивается регулятором расхода типа ППГ-1.

● Чтобы избежать влияния колебаний температуры окружающей среды на показания прибора, блок датчика с плечевыми элементами термостатирован. Нагрев происходит за счет индуктивных потерь в массе блока.

● Предусмотрена возможность продувки воздуха или азота через корпус датчика.

● Газоанализатор ТКГ-4 изготавливается в нормальном исполнении. Он применяется для анализа газов, указанных ниже.

Анализируемый компонент	Газовая смесь, в которой определяется компонент	Шкала прибора соответствует		Погрешность в показаниях прибора, % от диапазона шкалы
		содержанию компонента, %	диапазону прибора мВ	
Водород	Азото-водородная смесь производства синтетического аммиака	80—50	45	± 2
То же	Газ карбидных печей	0—20	17	± 5
»	Кислород электролитический	0—10	45	± 5
Кислород	Водород электролитический	0—10	17	± 5
Аммиак	Аммиачно-воздушная смесь	0—16	17	± 5
Сернистый ангидрид (SO ₂)	Печной газ в производстве серной кислоты	0—10	17	± 5
То же	То же	0—15	17	± 5

● В комплект газоанализатора типа ТКГ-4 входят:

- датчик;
- стабилизированный источник питания типа СИП-4;
- вторичный прибор—электронный потенциометр типа ЭПД;
- регулятор расхода типа ППГ-1;
- фильтр типа ФВ-10.

ТЕХНИЧЕСКАЯ ХАРАКТЕРИСТИКА

- с. 1. Питание—от сети переменного тока напряжением 220 в $\pm 10\%$; частота 50 *гц* $\pm 1\%$; потребляемая мощность 150 *вт*.
- 2. Нормальная работа прибора обеспечивается при наличии на датчике перепада давления 10—15 *мм вод. ст.*, а на комплекте прибора 1000—1500 *мм вод. ст.* (в зависимости от условий применения прибора).
- 3. Температура анализируемой смеси—от -10 до $+30^\circ\text{C}$.
- 4. Постоянная времени прибора 50—60 *сек*.
- 5. Расход анализируемого газа—не более 150 *л/час*.
- 6. Прибор предназначен для работы при температуре окружающего воздуха от -10 до $+35^\circ\text{C}$ и относительной влажности не выше 80%.
- 7. Габариты и вес узлов прибора:

Узлы прибора	Длина мм	Ширина мм	Высота мм	Вес кг
Датчик	330	193	340	18
Стабилизированный источник пита- ния типа СИП-4	352	160	290	16
Вторичный прибор типа ЭПД	420	292	506	29—35*
Регулятор расхода типа ППГ	130	185	235	3
Фильтр типа ФВ-10	138	88	245	2

* В зависимости от модификации.

К печати 29 IV 1957 г. Т01496 ГОСХИМИЗДАТ. № 1-1 Зак. 285
Типография Госхимиздата. Москва, 88, Угрейская



ТЕРМОХИМИЧЕСКИЙ ГАЗОАНАЛИЗАТОР

ТИПА

ТХГ-5

ТЕРМОХИМИЧЕСКИЙ ГАЗОАНАЛИЗАТОР типа ТХГ-5

Газоанализатор типа ТХГ-5 представляет собой стационарный электрический автоматически действующий прибор для непрерывного определения процентного (по объему) содержания горючих газов в присутствии (в избытке) кислорода и кислорода в присутствии (в избытке) горючих газов.

- Анализируемая газовая смесь не должна содержать агрессивных газов и паров, являющихся ядами для катализатора.
- Действие прибора основано на измерении теплового эффекта каталитической реакции, происходящей в слое катализатора. Тепловой эффект реакции измеряется с помощью двух чувствительных элементов, представляющих собой термометры сопротивления. Один чувствительный элемент—рабочий—расположен в камере, заполненной катализатором, другой—сравнительный—в камере, заполненной неактивной массой. Термометры включены в схему измерительного моста.
- Для повышения чувствительности газоанализатора и надежности работы катализатора измерительный блок датчика термостатирован. Температура измерительного блока $200 \pm 10^\circ\text{C}$.
- Постоянный расход анализируемой газовой смеси, проходящей через газоанализатор, поддерживается мембранным регулятором расхода.
- В комплект термохимического газоанализатора входят:
 - а) датчик;
 - б) вторичный прибор, изготовленный на базе автоматического уравновешенного моста типа ЭМД.

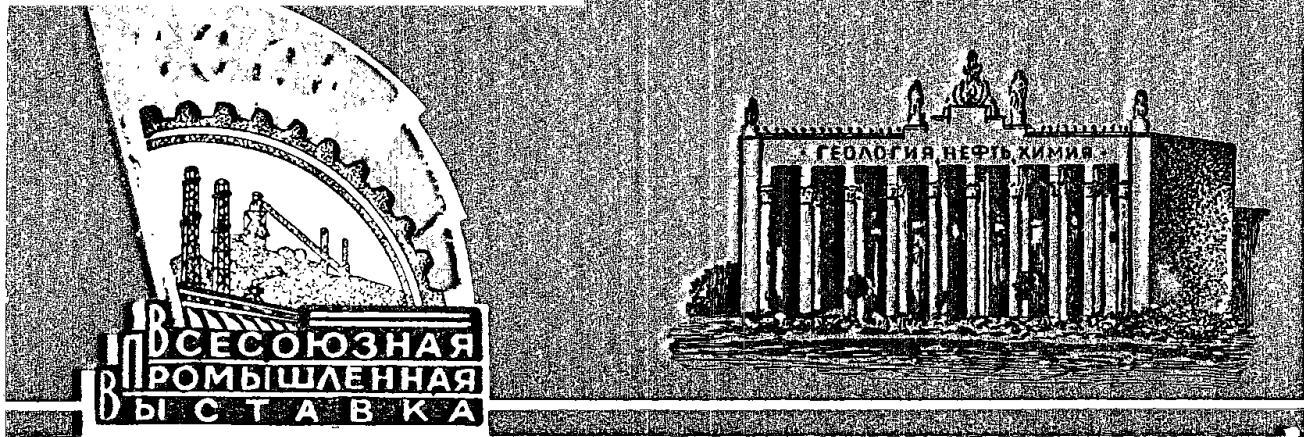
ТЕХНИЧЕСКАЯ ХАРАКТЕРИСТИКА

- 1. Термохимический газоанализатор типа ТХГ-5 изготавливается в нескольких модификациях для определения:
 - кислорода в электролитическом водороде (шкала прибора от 0 до 1,0% кислорода);
 - кислорода в генераторном газе (шкала прибора—от 0 до 0,5% и от 0 до 1,0% кислорода);
 - водорода в электролитическом кислороде (шкала прибора от 0 до 2% водорода).
- 2. Погрешность в показаниях прибора составляет $\pm 7\%$ от максимального значения шкалы.
- 3. Питание—от сети переменного тока напряжением $220 \text{ в} \pm 10\%$; частота $50 \text{ гц} \pm 1\%$; потребляемая мощность 100 в.
- 4. Постоянная времени (без учета влияния газозаборной линии) равна 2 мин.
- 5. Давление газовой смеси на входе в прибор должно быть от 400 до 1000 мм вод. ст.
- 6. Газоанализатор предназначен для работы при температуре окружающей среды от $+10$ до $+35^\circ\text{C}$ и относительной влажности среды до 80%.
- 7. Анализируемая газовая смесь не должна содержать сероводорода, сероорганических соединений, хлора и паров масла.
- 8. Расход газа через прибор составляет 60 л/час .
- 9. Габариты и вес газоанализатора:

Узлы газоанализатора	Длина мм	Ширина мм	Высота мм	Вес, кг
Датчик	312	258	235	15
Вторичный прибор ЭМД	420	292	506	29—35*
Регулятор температуры	422	166	288	
Регулятор расхода	275	130	236	

* В зависимости от модификации.

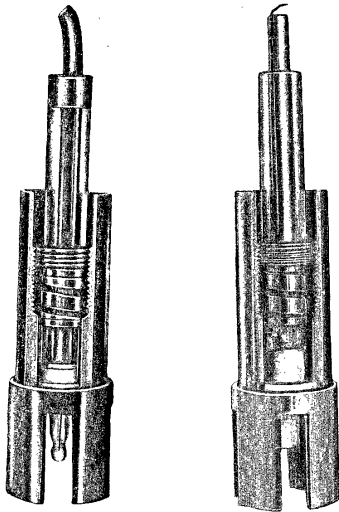
К печати 29.IV 1957 г. Т01496. ГОСХИМИЗДАТ. № 1—5. Зак. 289.
Типография Госхимиздата. Москва, 88, Угрешская



РН - МЕТР

МИНИСТЕРСТВО
ХИМИЧЕСКОЙ ПРОМЫШЛЕННОСТИ СССР
МОСКВА

pH-МЕТР



Прибор pH-метр предназначен для измерения и регулирования концентрации водородных ионов при помощи толстостенных стеклянных электродов. Этим прибором можно измерять величины pH растворов, находящихся в заземленных металлических баках.

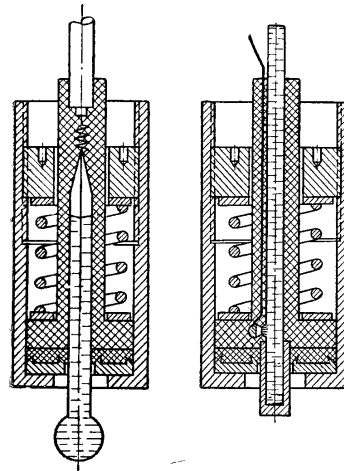
- В конструкции стеклянного электрода предусмотрено увеличение надежности изоляции места соединения электрода с высокочастотным кабелем, связывающим электрод со вторичным прибором.

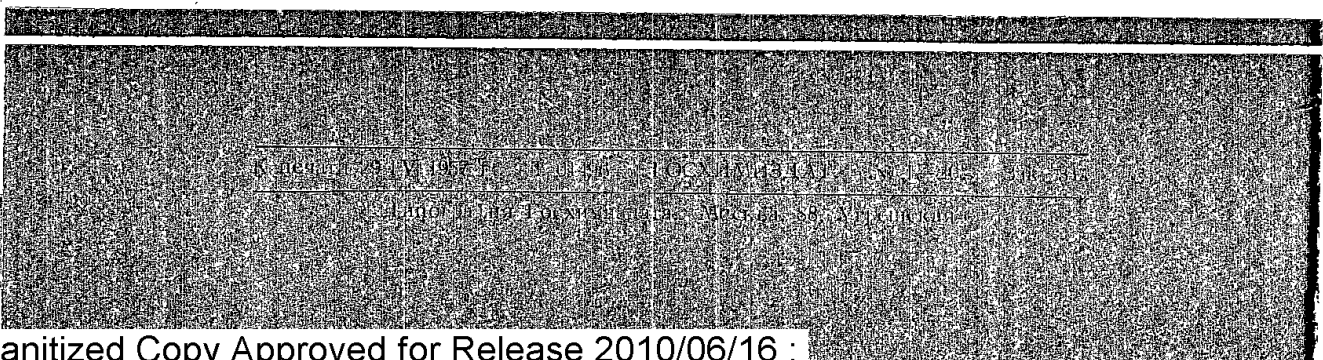
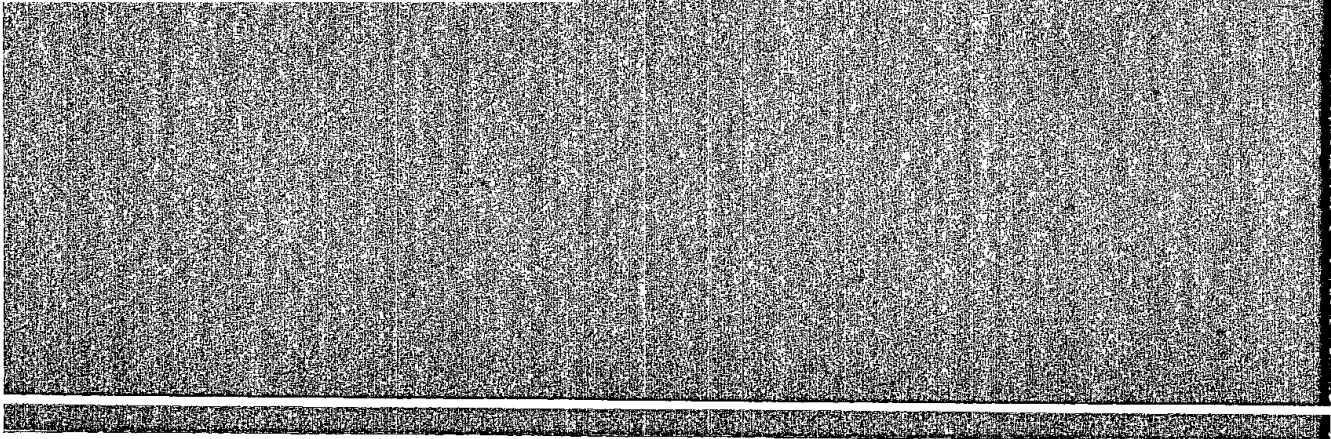
- Отличительной особенностью вторичного прибора pH-метра является схема включения оболочки кабеля, идущего от стеклянного электрода к электронному усилителю. Применение этой схемы позволяет в 1000 раз снизить требования к изоляции кабеля и тем самым увеличить эксплуатационную надежность прибора.

- Усилитель pH-метра собран на шасси усилителя электронного потенциометра типа ЭПД.

- Питание прибора осуществляется от сети переменного тока напряжением 220 в (с колебанием от плюс 10 до минус 15%).

- Погрешность и область измерения pH определяется качеством и характеристикой используемых стеклянных электродов.







СЕКЦИЯ СССР НА ВСЕМИРНОЙ ВЫСТАВКЕ
БРЮССЕЛЬ-1958
LA SECTION DE L'URSS A L'EXPOSITION
UNIVERSELLE ET INTERNATIONALE BRUXELLES
1958

Радио и телевидение

INSTALLATION INDUSTRIELLE DE TELEVISION
ПТХ-2М

APPAREILLAGE RADIO

INSTALLATION INDUSTRIELLE DE TELEVISION ПТТ-2М

L'installation industrielle de télévision ПТТ-2М est destinée à la surveillance à distance de différents travaux et à la surveillance des processus technologiques dangereux pour le personnel de service.

L'installation comprend le matériel suivant:

Des caméras d'émission de télévision (de 2 à 5), à têtes optiques, un bloc-canal, un dispositif récepteur vidéo, un bloc de commutation, un commutateur, des dispositifs mobiles de surveillance (de 1 à 5), un jeu d'objectifs, des combinés téléphoniques, un jeu de câbles, des câbles caméra-récepteur d'une longueur maximum de 1000 m et des câbles de liaison pour les dispositifs de surveillance mobile d'une longueur maximum de 800 m. Tout l'appareillage utilise principalement les lampes de la série miniature. Le vidicon type ЛИ-23 est utilisé comme tube émetteur.

Le récepteur vidéo utilise un tube à écran de 35 cm de diagonale et sert en même temps de source d'alimentation de la caméra. Le récepteur vidéo effectue la mise au point optique à distance de l'image, le réglage d'ouverture et le remplacement des objectifs de la caméra.

On prévoit la possibilité de fonctionnement de la caméra d'émission sans tête optique. Dans ce cas le remplacement des objectifs, la mise au point optique et le réglage du diaphragme se font à la main.

Le bloc-canal comprend un synchrégénérateur ainsi que les circuits d'amplification et de formation du signal de télévision.

La mise en service des caméras d'émission se fait à l'aide du commutateur. Les organes de commande du régime de fonctionnement des caméras d'émission sont également placés sur le commutateur.

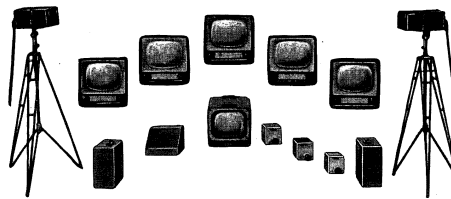
Une liaison téléphonique existe entre les caméras et le récepteur vidéo.

L'installation ПТТ-2М peut être simplifiée et transformée en installation de télévision industrielle à un seul canal (appelée ПТТ-1М). Dans ce cas on exclut le bloc de commutation et le commutateur, et la commande du régime de fonctionnement de la caméra d'émission se fait à partir du récepteur vidéo.

Caractéristiques techniques principales de l'installation

Analyse de l'image par points successifs, à 625 lignes pour 25 images par seconde.

Eclairage de fonctionnement à la photocathode du tube émetteur: 10 lux.



Jeu d'objectifs pour caméra:

l'objectif κ -25 (f = 100 mm, 1:2)

l'objectif ю-3 (f = 50 mm, 1:1,5)

l'objectif ПО 4-1 М (f = 35 mm, 1:2)

l'objectif „Юнирег” (f = 26 mm, 1:1,2)

Les objectifs κ -25 et ПО 4-1 М peuvent être employés dans la tête optique.

L'installation peut fonctionner pendant 23 heures consécutives.

Alimentation de l'installation: par un réseau à courant alternatif 220 V, 50 Hz. Puissance absorbée: 400 W maximum.

Poids de la caméra d'émission à tête optique: 6 kg environ; sans tête optique: 3 kg.

INDUSTRIAL TELEVISION EQUIPMENT TYPE ПТТ-2М

The type ПТТ-2М industrial television equipment is designed for remote observation of various operations and technological processes which are dangerous for the service personnel.

The following elements are included in the equipment: television pick-up cameras (2-5 pieces) with optical turrets, a channel unit, a video receiver unit, a commutation unit, a commutator, remote viewing units (1-5 pieces), a set of lenses, hand microtelephones, a set of cables, camera cables up to 1,000 metres long, and connecting cables up to 800 metres long for the remote viewing units.

Miniature lamps are predominant in most of the apparatus. The type ЛИ 23 vidicon is used as a pick-up tube.

The video receiver unit has a CRT with a screen diagonal of 35 cm. The optical focussing of the image, iris and a change of lenses for the camera are accomplished from the video receiver unit.

Means are provided for the operation of the pick-up camera without an optical turret. In such a case the lenses' change, focussing and iris are accomplished manually.

The channel unit contains a synchronizing generator and circuits for amplifying and forming the television signal.

The switching of the camera is accomplished through the commutator. The operating conditions for the pick-up tube are also controlled from the commutator.

Hand microtelephone communication is provided between the cameras and the video receiver unit.

Through simplification the ПТТ-2М equipment may be transformed into a single channel industrial television equipment (termed the ПТТ-1М equipment). The commutation unit and the commutator are absent in such a case, and the operating conditions for the pick-up camera are controlled from the video receiver unit.

ten. In diesem Fall werden die Auswechslung der Objektive, die Fokussierung und die Abblendung von Hand bewerkstelligt.

Das Kanalsystem enthält den Synchronisierungsgenerator sowie die Verstärker- und Formierungskreise des Bildsignals.

Die Umschaltung der Aufnahmekameras erfolgt durch Kommutator. Am Kommutator sind ebenfalls die Steuerorgane der Betriebsart der Aufnahmekameras angeordnet.

Zwischen den Kameras und der Videoempfangseinrichtung ist eine Mikrotelefonverbindung vorhanden.

Die Anlage ПТУ-2 М kann vereinfacht und in eine Einkanal-Industriefernsehanlage (genannt ПТУ-1 М) verwandelt werden. In diesem Fall wird die Kommutierungsgruppe und der Kommutator ausgeschlossen, die Betriebsartsteuerung der Aufnahmekamera wird von der Videoempfangseinrichtung bewerkstelligt.

Technische Grundkennwerte der Anlage

Bildzerlegung: Zeilensprungabtastung in 625 Zeilen bei 25 Bildwechseln pro Sekunde.

Betriebsbeleuchtungsstärke der Photokathode der Aufnahmeröhre: 10 lx.

Objektivsatz für jede Kamera:

ж-25 (F = 100 mm, 1:2)

ю-3 (F = 50 mm, 1:1,5)

РО 4-1 М (F = 35 mm, 1:2)

"Юпитер" (F = 26 mm, 1:1,2)

Im optischen Bildsucher können zwei Objektive:

ж-25 und РО 4-1 М benutzt werden.

Die Anlage kann ununterbrochen während 23 Stunden betrieben werden.

Stromversorgung: vom Wechselstromnetz 220 V, 50 Hz.

Leistungsaufnahme: höchstens 400 W.

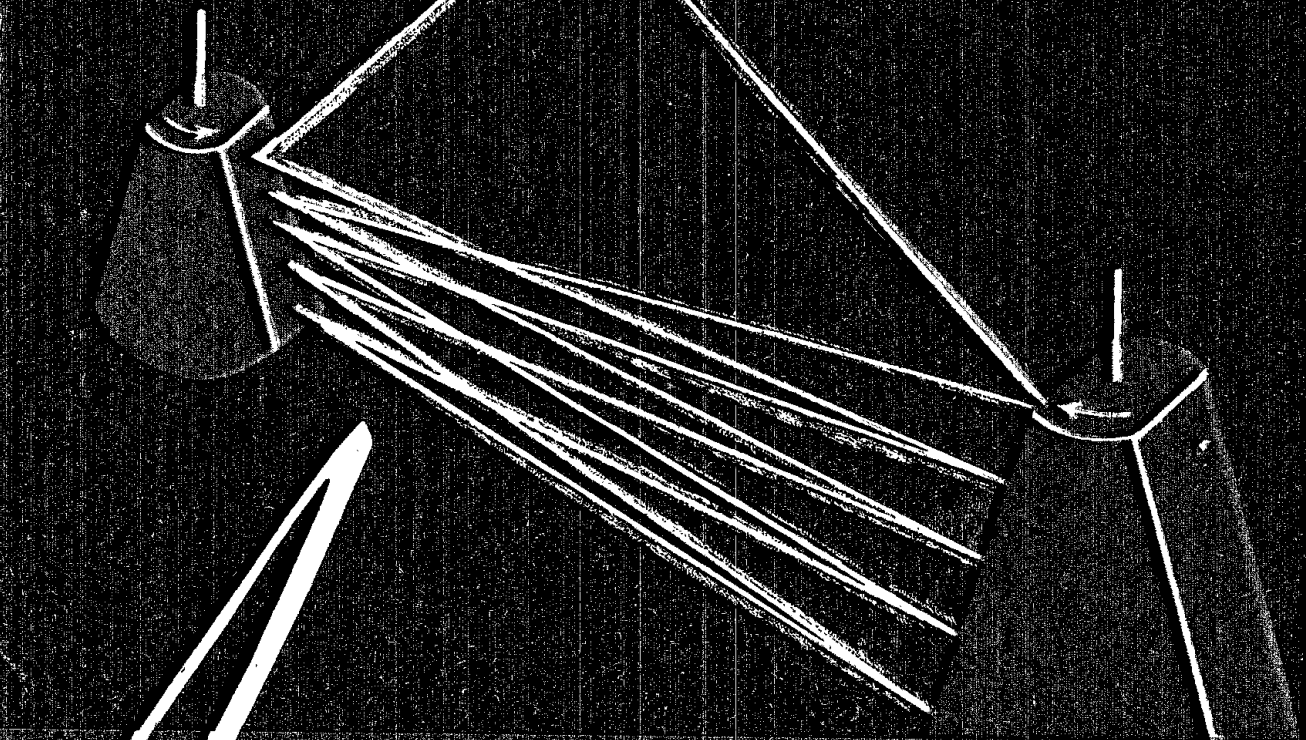
Gewicht der Aufnahmekamera mit optischem Bildsucher: ca. 6 kg,
ohne optischen Bildsucher: 3 kg.

Внешторгиздат. Заказ № Б 892-14

USSR SECTION: BRUSSELS UNIVERSAL
AND INTERNATIONAL EXHIBITION 1958

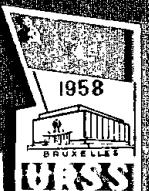
ABTEILUNG DER UdSSR AUF DER ALLGEMEINEN
WELTAUSSTELLUNG IN BRÜSSEL 1958

320000000
КАДРОВ В ГЕК



Цѣна **ВРЕМЕНИ**
С ОПТИЧЕСКИМ
РЕДУКТОРОМ

Section de l'URSS à l'Exposition Universelle et Internationale de Bruxelles 1958



LOUPE DE TEMPS FONCTIONNANT À HAUTE VITESSE

La loupe de temps, construite à l'Académie des Sciences de l'URSS, est destinée pour l'enregistrement photographique de haute vitesse des phénomènes autolumineux qui se passent rapidement (décharge d'étincelles, détonation d'explosifs etc.).

Le principe de fonctionnement du système optique de la loupe de temps est basé sur la réflexion multiple par miroirs des rayons de lumière, formant l'image d'un objet photographié. La réflexion multiple se produit par deux miroirs tournant à haute vitesse.

Le schéma de fonctionnement est représenté à la figure.

La fréquence de prise de vues est de 7 ÷ 32 millions de cadres par seconde.

La distance focale de la caméra peut être fixée par le changement de l'objectif d'entrée de la caméra: à la première variante 494 mm, à la seconde variante 943 mm.

Champ de vue:

à la première variante — de 1°22' à 41',

à la seconde variante — de 0°41' à 0°20',5.

Ouverture relative — de 1:44 à 1:88.

Dimension d'un cadre — de 5 mm à 12 mm (par diamètre).

Nombre de cadres enregistrés — 85 ÷ 416.

Dans la loupe de temps est utilisée la pellicule de 60 mm de largeur.

Les miroirs sont tournés par un moteur électrique asynchrone de puissance de 5 kW uni avec un réducteur-élévateur.

La loupe de temps est pourvue d'un système de commande automatique et de synchronisation et d'un dispositif qui engendre une impulsion de haute tension (5 kV) nécessaire pour commencer à enregistrer le phénomène examiné au moment correspondant à la position initiale des miroirs (au moment de commencement de prise de vues) ou à une avance en limites de 100 microsecondes. La production de cette impulsion s'effectue automatiquement au moment où les miroirs auront atteint la vitesse de rotation fixée d'avance (à la précision de $\pm 0,1\%$), ce qui exclut la nécessité de la mesure supplémentaire de la vitesse de l'enregistrement du phénomène.

Pour éviter la superposition répétitive de l'image du phénomène étudié dans la loupe de temps est utilisé un obturateur à explosion de rapide action, commençant à fonctionner quand un cycle de prise de vues est achevé.

La loupe de temps est un appareil de type laboratoire pourvu d'un pied de support permettant de faire la mise au point précise de l'objet à photographier.

Le châssis de l'appareil est fondu d'alliage d'aluminium.

Gabarits de l'appareil: longueur — 1085 mm, largeur — 1070 mm, hauteur — 1150 mm, poids — 250 kg.

HIGH-SPEED TIME MAGNIFYING GLASS

The time magnifying glass developed at the USSR Academy of Sciences is intended for high-speed photographic recording of fast self-luminous phenomena (spark discharge, detonation of explosives, etc.).

The principle of operating the optical system of the time magnifying glass is based on the multiple mirror reflection of light rays which build up the image of the object photographed. The multiple mirror reflection appears between two mirrors revolving at a high speed.

The key diagram is shown in the figure.

The frequency of photographing amounts to 7 ÷ 32 million pictures per second.

The focal length of the camera can be fixed by replacing the inlet lens of the camera: 494 mm in the first variant, and 943 mm in the second.

Field of vision:

For the first variant: from 1°22' to 41';

For the second variant: from 0°41' to 0°20',5.

Relative aperture: from 1:44 to 1:88.

Size of picture: from 5 mm to 12 mm (in diameter).

Number of pictures recorded: 85 ÷ 416.

A film 60 mm wide is used in the time magnifying glass.

The mirrors are revolved by a 5 kW asynchronous motor through a step-up reductor.

The time magnifying glass is equipped with a system of automatic control and synchronization and is supplied with an appliance producing a high tension impulse (5 kV) to initiate the phenomenon under examination at a moment corresponding to the initial position of the mirrors (at the beginning of photographing), or with a lead within 100 micro-seconds. The initiated impulse is produced automatically when the mirrors achieve the preset speed of rotation (to within $\pm 0,1\%$), which makes unnecessary any supplementary measurement of the speed of recording the phenomenon.

To preclude any repeated superimposition of the image of the phenomenon under investigation, the time magnifying glass is provided with a fast explosion shutter which operates at the end of a cycle of photographing.

The time magnifying glass is a laboratory type device supplied with a support which ensures accurate focusing.

The device housing is cast of an aluminium alloy.

Overall dimensions of the device: length — 1085 mm; width — 1070 mm; height — 1150 mm; weight — 250 kg.

ZEITLUPE FÜR HOCHGESCHWINDIGE FOTOAUFNAHMEN

Die von der Akademie der Wissenschaften der UdSSR entwickelte Zeitlupe ist für hochgeschwindige Fotoaufnahmen der mit großer Geschwindigkeit verlaufenden selbstleuchtenden Vorgänge (Funkenschlag, Detonation von Sprengstoffen und dgl.) bestimmt.

Die Wirkung des optischen Systems der Zeitlupe ist auf dem Prinzip der mehrmaligen Spiegelreflexion der Lichtstrahlungen aufgebaut, die den aufzunehmenden Gegenstand abbilden. Die mehrmalige Spiegelreflexion geschieht zwischen zwei Spiegeln, die sich mit großer Geschwindigkeit umdrehen.

Das prinzipielle Schema ist auf dem Bild dargestellt.

Die Frequenz der Aufnahmen beträgt 7 ÷ 32 Mil. Bilder je Sek.

Der Brennpunktstand der Kamera kann durch Umwechslung des Eingangsobjektives geändert werden: in der ersten Variante 494 mm, in der zweiten Variante 943 mm.

Das Blickfeld:

für die erste Variante — von 1° 22' bis 41'

für die zweite Variante — von 0° 41' bis 0° 20,5'.

Die relative Öffnung beträgt — von 1:44 bis 1:88.

Die Bildgröße ist von 5 mm bis 12 mm (im Durchmesser).

Die Zahl der registrierten Bilder 85 ÷ 416.

Für die Zeitlupe wird Fotorollfilm 60 mm breit verwendet.

Die Spiegel werden in Drehbewegung durch einen Asynchronmotor

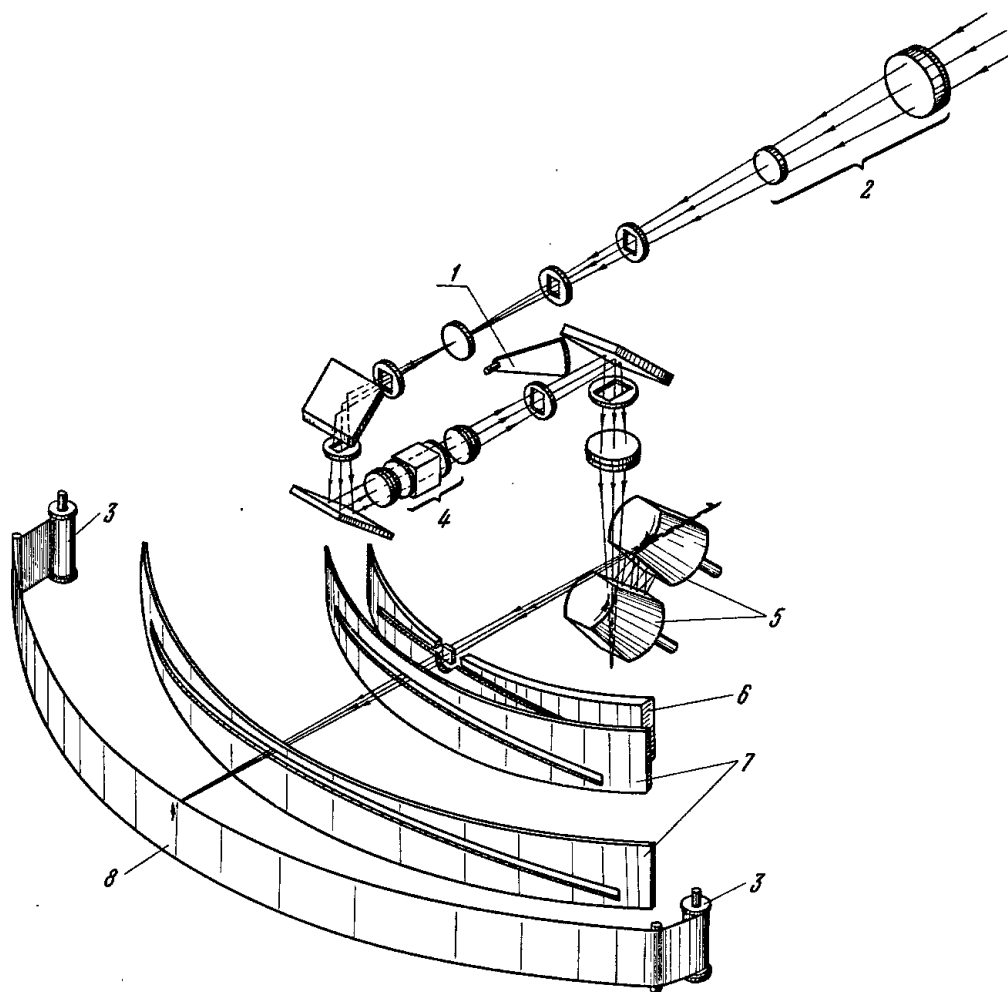
5 kW mittels eines Erhöhungsreduktors in Bewegung gesetzt.

Die Zeitlupe ist mit einem automatischen Steuerungs- und Synchronisierungssystem versehen, sowie mit einer Einrichtung, die einen hochgespannten Impuls (5 kV) erzeugt, um den Anfang des untersuchenden Vorganges exakt zum Moment hervorzurufen, der der Ausgangslage der Rotierung der Spiegel (Anfangsblick der Aufnahme) oder mit Vor-eilung in Grenzen von 100 Mikrosek. entspricht. Der Anfangsimpuls wird exakt zum Augenblick der Erreichung von den Spiegeln der angegebenen Drehgeschwindigkeit automatisch erzeugt (mit Genauigkeit von $\pm 0,1\%$). Damit wird die nachträgliche Messung der Registrierungsgeschwindigkeit des Vorganges nicht mehr nötig.

Um eine Aufeinandersetzung von Bildern des untersuchten Vorganges zu vermeiden, ist die Zeitlupe mit einem schnellwirkenden Explosionsver-schluß versehen, der nach der Vollendung jedes Aufnahmekreislaufes anspricht.

Die Zeitlupe ist ein Laborgerät. Sie ist mit einem Untergestell versehen, das eine exakte Einstellung des Apparates relativ des Aufnahme-gegenstandes gestattet.

Das Gehäuse des Apparates ist aus Aluminiumlegierung hergestellt. Abmessungen des Apparates: Länge 1085 mm; Breite 1070 mm; Höhe 1150 mm; Gewicht 250 kg.



1 – Obturateur mécanique; 2 – objectif; 3 – châssis aux bobines de pellicules;
4 – détails optiques de l'obturateur à explosion; 5 – miroirs tourants; 6 – len-
tille enchâssée; 7 – diaphragme; 8 – pellicule.

1 – mechanical shutter; 2 – lens; 3 – adapter; 4 – optical parts of the explosion
shutter; 5 – revolving mirrors; 6 – lens insert; 7 – blind; 8 – photographic film.

1. Mechanischer Verschluss; 2. Objektiv; 3. Kassette; 4. Optische Einzelheiten des
Explosionsverschlusses; 5. Rotierende Spiegel; 6. Linsenzusatzstück; 7. Blende;
8. Fotorollfilm.

Внешторгиздат. Заказ № Б-635-12

LOUPE DE TEMPS
FONCTIONNANT A HAUTE
VITESSE

HIGH-SPEED TIME
MAGNIFYING GLASS

HOCHGESCHWINDIGE
ZEITLUPE

USSR Section; Brussels Universal and International Exhibition 1958
Ausstellung der UdSSR auf der Allgemeinen Weltausstellung in Brüssel 1958

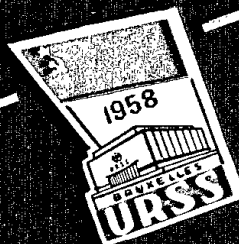
СПЕКТРОСКОПИЧЕСКИЕ

ИССЛЕДОВАНИЯ

СТРОЕНИЯ

И СОСТАВА

ВЕЩЕСТВА



USSR Section: Brussels Universal and
International Exhibition 1958

SPECTROSCOPIC INVESTIGATIONS OF THE STRUCTURE AND COMPOSITION OF MATTER

I. Introduction

Radiation, absorption and scattering of light by matter are closely associated with the processes taking place in atoms and molecules. Investigation of spectra is one of the principal methods of studying the structure of atoms and molecules, which yields most essential information on the structure of matter. At the same time spectral analysis, which is widely used to determine the composition of matter, is based on investigation of atomic and molecular spectra.

The study of the position and the intensity of lines in atomic emission and absorption spectra enables to obtain data on the structure of electron shells of atoms and certain information about atomic nuclei. It also throws some light on the state of the atoms studied, their ionization, velocity, gas temperature, etc.

Similarly, investigation of molecular spectra enables us to determine the structure of molecular electron shells, relative vibrations of nuclei, the rotation of the molecules as a whole, as well as a number of structural parameters of molecules. It also appears possible to obtain information on molecular interaction in solids, liquids and gases. The positions of spectral lines and bands is a characteristic feature of individual atoms and molecules. Owing to this, investigation of a spectrum can be used to determine the qualitative and quantitative composition of the object under test. Thus, spectral analysis of metals, alloys, ores and minerals, gases, organic mixtures and other objects becomes possible. Spectral analysis is, as a rule, much more sensitive than the chemical and many physical methods of analysis. This method is rapid, inexpensive and practically does not damage the object under study. These features of spectral analysis account for its wide use in scientific research and particularly for production control in industry.

Investigations of spectra and spectral analysis are being developed and employed by physicists, chemists, biologists, and engineers on a large scale all over the world. They form a traditional trend in Soviet science. Researches by D. Rozhdestvensky, L. Mandelstam, S. Vavilov, and G. Landsberg made a substantial contribution to theoretical and applied spectroscopy, and are being farther developed by their collaborators and pupils.

At present spectroscopic research is carried out in numerous Soviet research institutes, universities and industrial laboratories. Regularities of atomic and molecular spectra are being investigated theoretically and experimentally, the spectra of individual atoms and molecules and their spectroscopic characteristics being used for studying the structure of matter. Methods of spectral analysis are widely used in everyday production control in metallurgy, aircraft and motor car industries as well as in engineering and prospecting, practically superseding chemical methods.

The entire work in spectroscopy and spectral analysis in the U.S.S.R. is directed and coordinated by the U.S.S.R. Academy of Sciences Commission on Spectroscopy, which includes eminent scientists in spectroscopy and spectral analysis, from various research institutes, universities, and industrial enterprises. Biannual conferences on atomic and molecular spectroscopy and its applications are held in the U.S.S.R., the number of their participants increasing rapidly. The tenth Conference held in Lvov (1956) was attended by about 1500 participants, and about 350 papers were presented. A special periodical journal *Optics and Spectroscopy* is being published in the U.S.S.R. Papers on spectroscopy are also printed in many other physical and technical scientific publications.

Soviet industry manufactures various equipment for spectroscopic investigations and spectral analysis. Some specimens of this equipment are displayed at this Exhibition. Of certain interest are, among other exhibits, specimens of diffraction gratings and echelettes used in various Soviet spectroscopic instruments.

The exposition at the "Palais de Science" shows some works of Soviet scientists on spectroscopy and the new methods of spectral analysis. These are selected so as to give an idea of the diversity of spectroscopic investigations. The exposition does not include a rather large number of works embracing the various, now routine, methods of emission analysis, infrared absorption analysis, etc. The exposition is arranged on two stands, displaying the investigations of atomic and molecular spectra respectively. The exposition dealing with molecular spectra includes some works on application of molecular vibrations and rotations in the study of the structure and composition of matter. A considerable part in investigations conducted in the U.S.S.R. is played by the Raman effect discovered in 1928 by Raman (India) in liquids, and simultaneously by L. Mandelstam and G. Landsberg (U.S.S.R.) in crystals.

II. Atomic Spectroscopy

(Atomic Electron Shells)

I. Rozhdestvensky's "Hook" Method and its Applications.

Important results are obtained in the investigation of atomic spectra by studying abnormal light dispersion in the immediate vicinity of absorption lines, according to a method developed by D. Rozhdestvensky in 1912. Spectrograms obtained by Rozhdestvensky's method show hook-shaped interference fringes (Fig. 1), on both sides of the absorption line, whence the name,



Fig. 1. Abnormal dispersion in neon excited in a pulsed discharge.

the "hook" method. The linear distance between the apexes of the hooks is proportional to the square root of the product of concentration of the absorbing atoms and the value of the so-called oscillator strength (f) of the respective line. The latter value is related to the probability of the transition of the atom from one energy state to another and forms one of the most important atomic constants. Thus, the Rozhdestvensky method enables us to measure the probabilities of transitions of the atom or, if the latter are known, atom concentrations in the ground state and in the various excited states.

The Rozhdestvensky method was used at first to determine the relative and later on the absolute values of transition probabilities in atoms of alkali metals. In later studies Rozhdestvensky and his pupils extended the application of this method to more complicated spectra owing to improved experimental technique. At present the following elements have been investigated by the "hook" method: Li, Na, K, Rb, Cs, Cu, Ag, Ca, Sr, Ba, Zn, Cd, Hg, Al, Ga, In, Tl, Pb, Sc, Ti, V, Cr, Mn, Fe, Co, Ni, Mo, Ne.

The hook method is also used to study electric discharges in gases. The knowledge of atom concentrations in various energy states allows to draw conclusions about the extent to which conditions in gaseous-discharge plasma differ from thermodynamic equilibrium, as well as about the nature of atomic excitation, the degree of ionization, vapour density, etc.

In the U.S.S.R. considerable attention is also paid to theoretical calculations of f -values in atoms. This work is being carried out by V. Fok and his numerous pupils. The calculations are chiefly based on wave functions obtained by employing Fok's self-consistent field method taking exchange into account. At present the calculations of the f -values are performed by electronic computers.

References: D. Rozhdestvensky, *Investigations of Abnormal Dispersion in Metal Vapours*, "Classics of Science" Series, published by the U.S.S.R. Academy of Sciences, 1951; V. Fok, *Multielectron Problem of Quantum Mechanics and Atomic Structure*, collection of articles dedicated to the 30th anniversary of the October Socialist Revolution, part I, p. 255, the U.S.S.R. Academy of Sciences, Moscow-Leningrad, 1947.

2. Spectral Analysis Using the Vaporization Technique.

A number of branches of modern industry (atomic energy production, semiconductor devices, etc.) demand especially pure materials. Their analysis presents special difficulties, since conventional methods of chemical and spectral analysis fail, as a rule, to provide the required sensitivity of determination. On the other hand spectral analysis employing the vaporization technique makes it possible to obtain exceptionally high analytical sensitivities in a number of cases.

The method consists in separate vaporization of the impurities and in the excitation of their spectra, whereas impurity spectra free from the base spectrum interferences are obtained using the phenomenon of fractional vaporization.

Vaporization of impurities is performed by heating the sample to a given temperature in a special vaporizer. The impurities condense on the cooled surface of a copper or graphite electrode and are then burned in an arc or spark. The composition of the condensate is determined by using conventional methods of spectral analysis. The scheme of the vaporizer is shown in Fig. 2. Sublimation of impurities may be carried out either in air or in a vacuum. The method allows fairly accurate quantitative determination of trace impurities in low-volatile substances (1×10^{-9} — 1×10^{-7} g).

The method is used for the analysis of uranium, thorium, beryllium, aluminium, plutonium oxides and other metals as well. Thus, for example, a

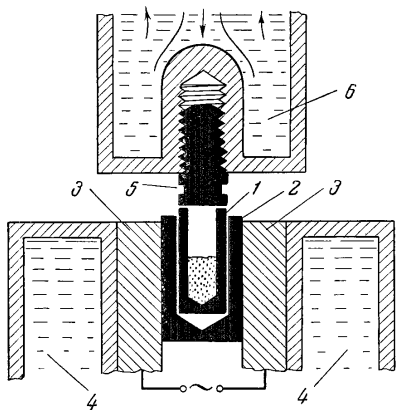


Fig. 2. Vaporizer diagram: 1—graphite container with sample; 2—graphite crucible; 3—current-heated graphite brushes; 4—water cooling; 5—graphite or copper condensate receiver; 6—water cooling of receiver.

quantitative determination of such impurities as 2×10^{-6} per cent Li, 1×10^{-5} per cent B, 2×10^{-5} per cent Mn, etc., is possible.

References: S. Mandelstam, N. Semyonov, Z. Turovtseva, *Journal of Analytical Chemistry*, 11, 9, 1956; A. Zaidel, N. Kaliteyevsky, L. Lipis, M. Chaika, Y. Belayev, *Journal of Analytical Chemistry*, 11, 21, 1956; Y. Belayev, A. Zaidel, *Journal of Analytical Chemistry*, 12, 30, 1957; A. Zaidel, N. Kaliteyevsky, L. Lipis, V. Tarakanov, *Optics and Spectroscopy*, 3, 16, 1957.

3. Powder Injection into Arc Discharge.

To attain a higher reproducibility in spectral analysis of ores and minerals, a new method of introducing the sample to be analysed into the discharge zone has been used. An air jet injects the pulverized sample under test into the discharge of an activated A. C. arc. A special apparatus, type ABP-2 (Fig. 3),

for a rapid semiquantitative analysis for the content of 10 to 15 elements in friable surface deposits has been developed. Plotting the concentrations of elements on maps makes their scattering haloes over ore outcroppings obvious. The apparatus makes it possible to obtain spectra of about 1,000 samples in 8 hours.

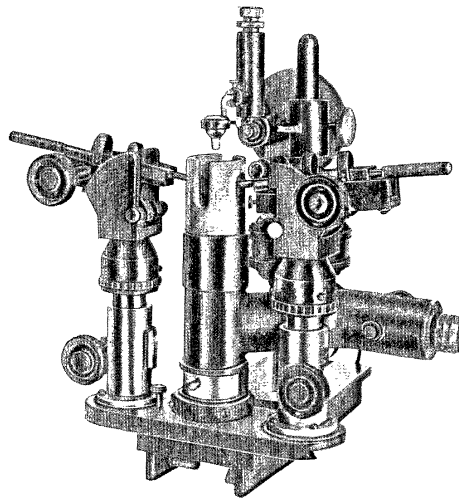


Fig. 3. External view of powder injector (ABP-2).

Air injection stabilizes the conditions of vaporization of the elements and of the excitation of their spectra, reducing to a minimum the role of processes on the electrodes, which considerably affect analysis results in the case when a sample is introduced into the crater of the carbon electrode by usual methods.

Application of the injection method reduces the error in quantitative determinations and increases analytical sensitivity for many elements. The method is widely used in the analysis of ores, rocks, minerals, slags, and other materials for content of all elements excited in a carbon arc.

References: A. Rusanov, V. Khitrov, *Plant Laboratory*, 28, 175, 1957; A. Rusanov, T. Tarasova, *Journal of Analytical Chemistry*, 10, 267, 1955; V. Nedler, *Plant Laboratory*, 21, 1056, 1955.

4. Spectral Determination of Gas Content in Metals by Means of Pulsed Discharge.

Analysis of gases in metals is an important technical problem. Until recently only chemical methods have been used for this purpose. Employment of spectral analysis often furnishes significant advantages. The method consists

in exciting the sample spectrum by means of a high- or low-voltage pulsed spark. Analysis of oxygen and nitrogen content in steels and hydrogen in titanium alloys (down to 0.005 per cent has been worked out). The sample is placed in an atmosphere free from the components to be determined. The spectra are photographed by means of a type ИСН-51 standard spectrograph with a $F=270$ mm camera. At present the methods are being worked out for determination of these gases in other metals (for instance, in Cr, Mo).

References: S. Mandelstam, O. Falkova, *Plant Laboratory*, 16, No. 4, 430, 1950; O. Falkova, *Bulletin of the U.S.S.R. Academy of Sciences*, Physical Series, 19, No. 2, 149, 1955; N. Gerasimova, T. Ivanova, N. Sventitsky, G. Startsev, K. Taganov, M. Trentovius, *Bulletin of the U.S.S.R. Academy of Sciences*, Physical Series, 19, No. 2, 147, 1955.

5. Hydrogen Content Determination in Metals by Isotopic Equilibration Method.

The method is based on isotopic equilibration of a certain amount of deuterium with the hydrogen contained in the metal sample under test. The composition of the gas surrounding the sample is established after equilibration by the relative intensity of hydrogen and deuterium lines. The spectrum is excited by means of an electrodeless high-frequency discharge. The method does not require the use of standards. The results are largely free from effects of other gases evolved when heating the sample. Two variations of the method have been developed: a more accurate one, with photoelectric recording of the $H\beta$ - and $D\beta$ -lines in a spectrum obtained by means of a monochromator with a diffraction grating, and a simplified semiquantitative one, where the relative intensity of the $H\alpha$ - and $D\alpha$ -lines is measured visually, using a polarization deuterometer (without a spectral instrument). In the former case the error approximates 6 per cent, while in the latter it is about 15 per cent. The method permits determination of hydrogen content of 0.1 cm^3 per each 100 g metal (zinc, aluminium, iron, titanium, zirconium and steel). Up to 20 analyses are carried out in 8 hours.

References: A. Zaidel, A. Petrov, *Journal of Applied Physics*, U.S.S.R., XXV, 14, 1955; A. Zaidel, A. Petrov, G. Veinberg, *Spectral Isotope Method of Hydrogen Determination in Metals*, Leningrad State University Publishing House, 1957.

III. Molecular Spectroscopy (Molecular Vibrations and Rotations)

I. Calculations of Vibrational Spectra.

Generally, $3N-6$ fundamental vibrations are possible in a molecule consisting of N atoms. The task of fundamental research is to determine the frequencies of normal vibrations and other spectral characteristics (intensity, polarization, etc.) from the known data on molecular structure. The inverse problem consists in the determination of molecular structure by spectral properties of the molecule. The principal results of investigations accomplished by

Soviet scientists on the theory of molecular vibrations are collected in a special monograph published in 1949. Simplified methods of calculating the frequencies and forms of molecular vibrations were developed and a theory of line intensity and line polarization in vibrational spectra has been evolved.

Of late the calculation of molecular vibrations has become much more rapid due to the use of electronic computers.

References: M. Volkenstein, M. Elyashevich, and B. Stepanov, *Molecular Vibrations*, State Publishing House of Theoretical and Technical Literature, 1949; L. Mayants, *Proceedings of the Physical Institute of the U.S.S.R. Academy of Sciences*, 5, 63, 1950; R. Podlovchenko, M. Sushchinsky, *Optics and Spectroscopy*, 2, 49, 1957.

2. Rotational Spectra of Molecules in the 0.5 mm Region.

The frequencies of rotational transitions of molecules are usually very low, and the corresponding lines lie in the far infrared region. For a long time our knowledge concerning this region was very limited due to the great experimental difficulties involved (low energy of thermal radiation in this spectral region, lack of sensitive detectors).

It is only recently, thanks to the progress in the techniques of detecting low intensity radiation, development of efficient diffraction gratings (echelettes), and to the use of aspheric optics, the construction of equipment for recording long-wave spectra proved to be possible. The method has come to be widely used for the solution of various structural problems of physics and chemistry.

A laboratory vacuum spectrometer (with a relative aperture ratio of 1 : 2) has been developed in the U.S.S.R. for automatic registration of infrared spectra within the $20\text{--}500 \mu$ range and upwards. It uses a silitite rod (for operation in the

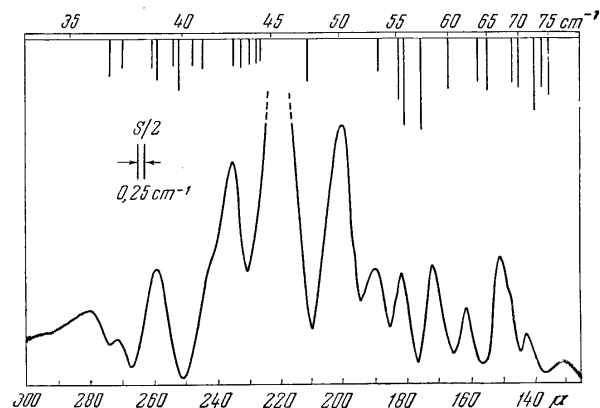


Fig. 4. Rotational spectrum of water in the $130\text{--}300 \mu$ region.

region up to 100μ) and a IIPK-4 quartz mercury tube (for the remaining region) as radiation sources, and an optical acoustic receiver as an indicator. The data are recorded by an automatic registering device. Four echelettes with a grating area of $250 \times 250 \text{ mm}^2$ are used.

Fig. 4 shows the recording of the rotational spectrum of atmospheric water in the $130\text{--}300 \mu$ range. Frequencies and intensities of the absorption bands are in good agreement with the theoretical predictions given in the upper part of the figure.

References: N. Yaroslavsky, B. Zheludov, A. Stanevich, *Optics and Spectroscopy*, 1, No. 4, 507, 1956; S. Luchin, *Journal of Applied Physics*, U.S.S.R., 16, 1115, 1945.

3. Investigation of Fine Structure of Absorption Bands in Infrared Region of Spectrum.

Investigation of the fine structure of vibrational-rotational spectra permits to determine important molecular constants. To resolve the fine structure of vibrational-rotational spectra of diatomic and more complex molecules, apparatus with high resolving power should be used. The general view of one such

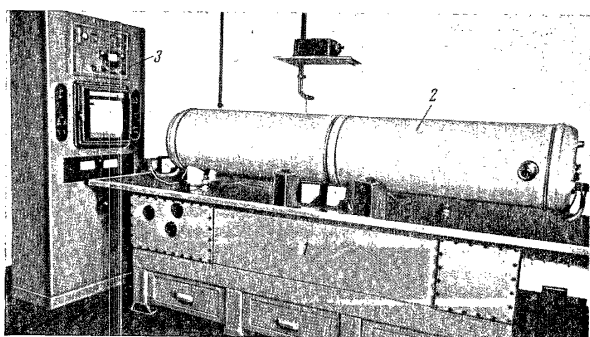


Fig. 5. Diffraction spectrometer for investigating the fine structure of vibrational-rotational spectra: 1 — spectrometer; 2 — gas cuvette; 3 — amplifier and register. device constructed in the U.S.S.R. is shown in Fig. 5. The grating (echelette) with an area of $150 \times 150 \text{ mm}^2$ (300 lines per 1 mm) ensures a resolution better than 0.1 cm^{-1} . Preliminary dispersion of the radiation is effected by means of a prismatic monochromator.

Radiation first enters a two-meter cuvette and owing to multiple reflection covers an optical path of 8 to 200 m. The light is detected by a lead sulphide resistor, and the amplified signal is recorded by an automatic registering device. Fig. 6 shows part of a CO spectrum obtained with the help of this spectrometer. The investigation of absorption spectra of water vapours allowed to

establish the dependence of absorbed energy of an individual rotational line or group of lines on vapour concentration, the path length and the pressure of foreign gases (He , N_2 , CO_2). The regularities thus established have been used for the study of molecular interaction and for the determination of gas humidity.

References: B. Neporent, K. Vasilevsky, N. Lapina, *Optics and Spectroscopy*, 3, 289, 1957.

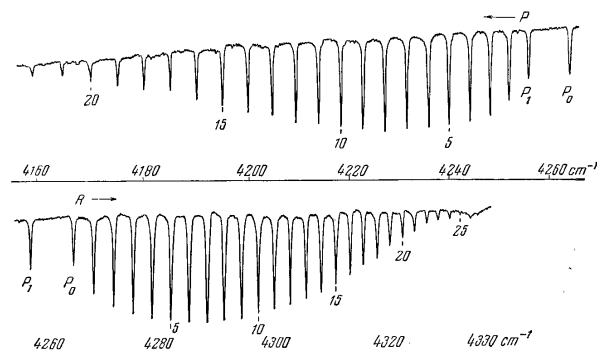


Fig. 6. Absorption spectrum for the first overtone of carbon dioxide at 300°K .

4. Resonance Raman Effect.

Raman line intensity rapidly increases as the exciting frequency approaches the absorption band of the scattering molecules. This effect compensates for light absorption in coloured media, thus enabling observation of the Raman spectrum in the hitherto inaccessible resonance region. Raman line intensities are tens and hundreds of thousands of times higher in the resonance region than far away from the absorption bands. It has been possible therefore to observe the Raman effect at exceedingly low concentrations (of the order of 10^{-4} per cent).

The above method made it possible to study the spectra of many intensely coloured aromatic compounds, to investigate the interaction of atomic groups in complex molecules and to study the effect solvents have on the vibrational spectra of solute at low concentrations.

References: P. Shorygin, *Reports of the U.S.S.R. Academy of Sciences*, 87, 101, 1952.

5. Investigation of Fog Transmission in the Infrared Region of Spectrum.

Due to the compensation of radiation absorption by water vapours, pure scattering and absorption effects caused by the drop phase of fogs could be observed. Instead of the expected absorption bands of liquid water ($\lambda = 2.95$

and 6.02μ) transparency bands were observed in the transmission spectrum of cooled fog, their position ($\lambda = 2.77$ and 5.80μ) unaffected by the size of the fog drops.

These bands are conditioned by selective scattering caused by the anomalous behaviour of the refraction index in the region of liquid water absorption bands, and by the true absorption of radiation by the drops. The position of the transparency bands coincides with the wave lengths, which correspond to the minima of the water refraction index. The use of so-called "integrating spheres" eliminated attenuation of the radiation beam caused by light scattering. Under these conditions practically no attenuation has been observed in the transmission spectrum of fog. Absorption bands have been observed whose position and shape are close to those of the absorption spectrum of liquid water film.

Similar phenomena have also been observed in the spectra of similar objects.

References: V. Malyshev, *Materials of the Xth U.S.S.R. Conference on Spectroscopy*, Lvov University Publishing House, 3, 1957.

6. Determination of Earth's Atmosphere Humidity.

The method of infrared absorption spectra has been used for the determination of humidity of the Earth's atmosphere at altitudes up to 20 km. A special altitude instrument (Fig. 7) was used, consisting of a vacuum monochromator with a diffraction grating, equipped with air-to-ground signal transmitters connected with the altitude, air temperature and humidity meters, etc.

References: B. Neporent, V. Belov, O. Dmitriyevsky, G. Zaitsev, W. Kastrov, M. Kiselyova, L. Kudryavtseva, I. Patalakhin, *Bulletin of the U.S.S.R. Academy of Sciences, Geophysical Series*, No. 4, 1957.

7. Study of Glass Structure by Method of Raman Spectra.

For the investigation of glass structure photoelectric recording of spectra, a powerful low-pressure spiral mercury tube (of the "Toronto" type), as well as double-monochromatization of radiation in a specially designed spectrometer have been used.

The examination of over 40 samples of two-component glass and molten quartz samples has shown that spectrum parameters do not always vary monotonously with composition of the objects. More or less abrupt changes of spectra point to structural changes in glass. It is supposed that during the formation of two-component glasses cations of monovalent metals are distributed in places of Si—O bond rupture non-statistically, and that regions of different structure are simultaneously present in glass (space net lattice, layers and chains of SiO_4 tetrahedrons).

References: Y. Bobovich, D. Gurevich, *Journal of Experimental and Theoretical Physics*, 27, 318, 1954; Y. Bobovich, T. Tulub, *Optics and Spectroscopy*, 2, 174, 1957; B. Kiselyov, *Optics and Spectroscopy*, 1, 597, 1956.

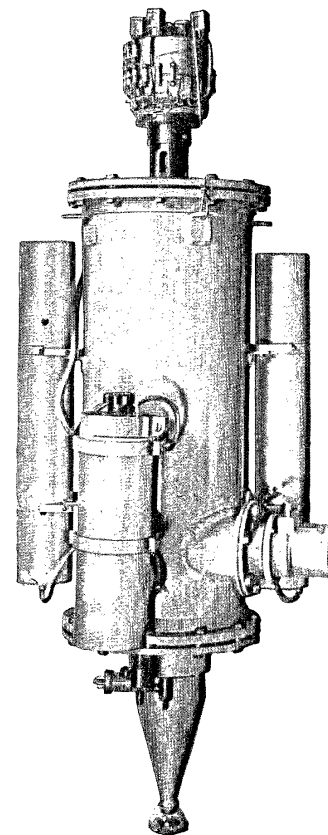


Fig. 7. Altitude infrared spectrometer for the determination of humidity in the Earth's atmosphere.

8. Combined Method of Determining the Individual Hydrocarbon Composition of Light Oil Fractions (Benzines).

The Raman method is widely used in the U.S.S.R. for the analysis of various organic mixtures. The method of determining the hydrocarbon composition of light fractions (benzines) has been developed more thoroughly. The original benzine is being successively decomposed by means of various physical and chemical processes into a series of simple fractions similar in their group hydrocarbon composition and molecular weight. These fractions are then investigated by using the Raman method. To decompose the original benzine the following methods are applied:

1. Precise rectification by means of distillation columns with an effective length of 40 to 50 theoretical plates enabling narrow fractions boiling within 4 to 5° to be obtained.
2. Chromatographic adsorption, allowing the separation of aromatic hydrocarbons with strong light scattering from paraffine and naphthene ones.
3. Dehydrogenation catalysis, by means of which cyclohexane naphthenes containing no heminal carbon atoms are separated from the paraffine-naphthene fractions by converting these into aromatic hydrocarbons.

Conditions have been worked out under which the results of hydrocarbon mixture analysis are reproduced with an accuracy of ± 10 per cent (relative). The analysis is performed without a large number of standards (two materials are sufficient: toluene and methyl cyclohexane). This required a previous investigation of Raman spectra of over 100 specially synthesized standards of individual hydrocarbons forming part of benzines. The intensities of the Raman lines of these compounds are given in uniform scale units. The method is widely used in the U.S.S.R. not only for benzine analysis, but also to determine the composition of various hydrocarbon reaction products.

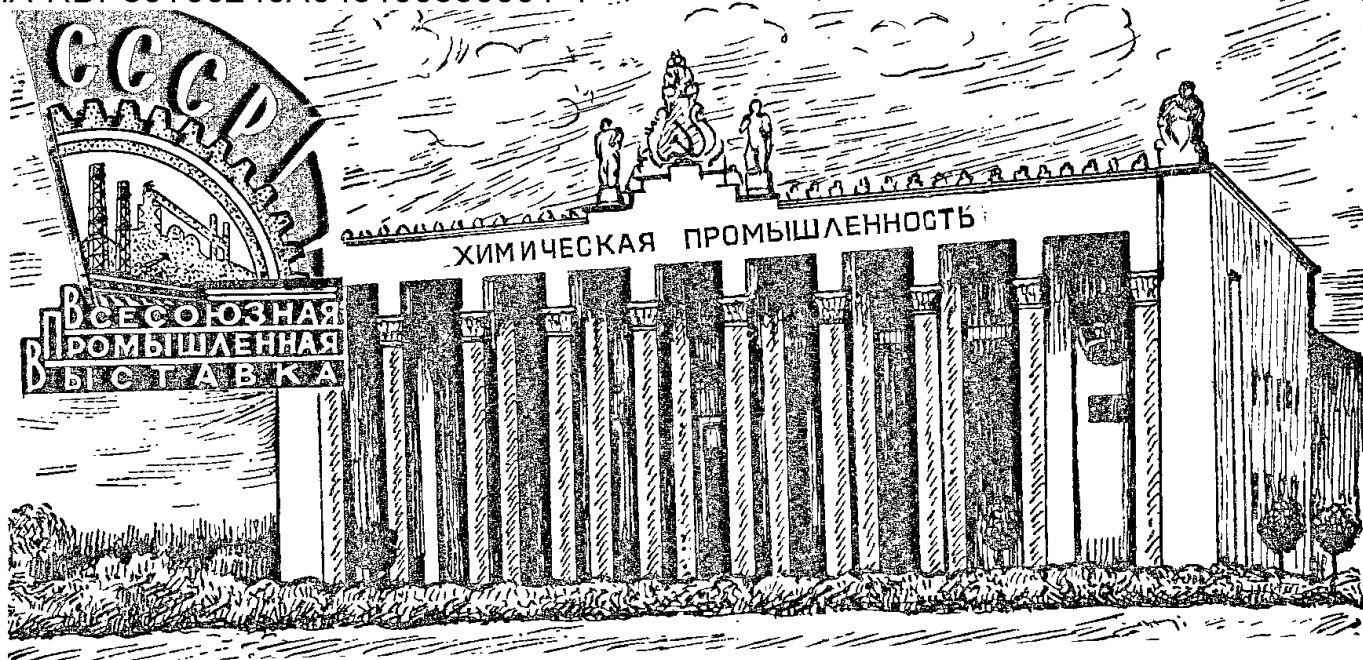
References: G. Landsberg, P. Bazhulin, M. Sushchinsky, *Principal Parameters in Raman Spectra of Hydrocarbons*, Moscow, 1956; G. Landsberg, B. Kazansky, P. Bazhulin, A. Liberman, Y. Mikhailova, A. Plate, Kh. Sterin, M. Sushchinsky, G. Tarasova, and S. Ukholin, *Determination of Individual Hydrocarbon Composition of Direct-Stilling Benzines by the Combined Method*, Moscow, 1958.

SPECTROSCOPIC

INVESTIGATIONS

OF THE STRUCTURE

AND COMPOSITION OF MATTER



**РЕГУЛЯТОР
ДЛЯ ПРОГРАММНОГО УПРАВЛЕНИЯ
ПЕРИОДИЧЕСКИМ
ХИМИЧЕСКИМ ПРОЦЕССОМ**

Регулятор для программного управления периодическим химическим процессом

НАЗНАЧЕНИЕ

Программный регулятор предназначен для автоматического управления химическим процессом, протекающим периодически по установленной временной программе и для записи отклонений от режима, принятого на задающем устройстве регулятора.

ПРИНЦИП ДЕЙСТВИЯ ПРОГРАММНОГО РЕГУЛЯТОРА

Блок-схема программного регулятора представлена на рис. 1. Параметр регулирования (выходная координата объекта регулирования) воспринимается датчиком, имеющим

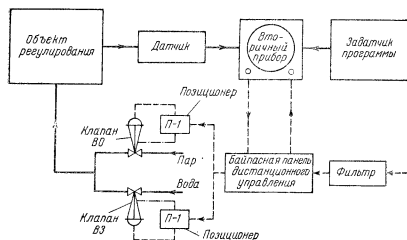


Рис. 1

электрический выход. Параметром регулирования может быть температура (датчик — термомпара, или термометр сопротивления), расход (датчик — дифманометр ДМ, или рота-

метр РЭД), рН среды (датчик — пара электродов). Электрическое напряжение от датчика сравнивается во вторичном приборе с напряжением, снимаемым с программного задачика.

Напряжение рассогласования, характеризующее отклонение от режима, принятого на программном задачике, измеряется электронным самопишущим и регулирующим прибором. Электрическая схема этого прибора выполнена таким образом, что «электрическому нулю» соответствует середина шкалы прибора. Задающая стрелка пневматического регулятора также установлена на центр шкалы.

При отклонении параметра регулирования от значения, заданного программным задачиком, пишущая стрелка регулирующего прибора отходит от задающей стрелки и, если, например, параметром регулирования является температура, пневматический регулятор приводит в действие клапаны для подачи воды или пара (в зависимости от знака отклонения) для того, чтобы параметр регулирования был приведен в соответствие с заданием. Паровой клапан, действующий по системе «воздух открывает» (ВО), и водяной клапан, действующий по системе «воздух закрывает» (ВЗ), могут быть настроены без взаимного перекрытия (настройка производится регулированием натяжения пружин у позиционера). Применение двух клапанов дает возможность программному регулятору работать на различных участках с экзотермическими и эндотермическими реакциями.

Кроме автоматического регулирования заданного параметра (по жесткой временной программе), программный регулятор может производить релейные переключения во времени для пуска и остановки мешалки, заполнения аппарата, переработки готового продукта и т. п. Для этого используют свободные электрические и пневматические контакты командного прибора КЭП-12.

Имеются три модификации программного регулятора, отличающиеся конструкцией программного задачика. Программный задачик потенциометрического типа позволяет производить набор требуемой ступенчатой программы изменения параметра регулирования.

Программный задачик кулачковый типа снабжен встроенным в прибор КЭП-12 цилиндрическим кулачком для плавного программного изменения параметра регулирования.

Программный задачик ленточного типа имеет для задания программы две дорожки с перфорацией.

Примерный вид требуемой по регламенту программы изменения температуры представлен на рис. 2, а реализация этой программы при помощи программного регулятора (запись отклонений от программы) иллюстрируется записью на диаграмме вторичного прибора (рис. 3).



Рис. 2

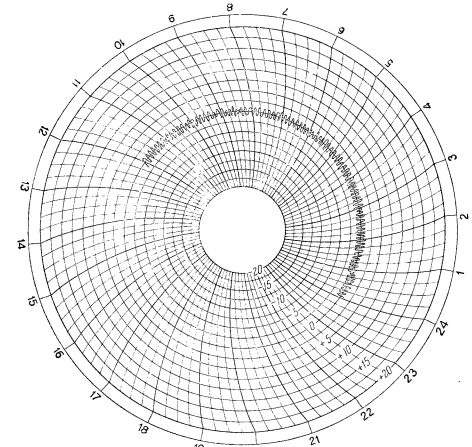


Рис. 3

КОМПЛЕКТНОСТЬ

В комплект программного регулятора входят:

1. Датчик с электрическим выходом (термопара, термометр сопротивления, дифманометр ДМ, ротаметр РЭД и т. п.).
2. Самопишущий и регулирующий электронный потенциометр ЭПД-32 или дифференциально-трансформаторный прибор ЭПИД-06 со встроенным пневматическим регулятором (электрическая схема приборов изменена).
3. Задатчик программы.
4. Прибор типа КЭП-12.
5. Мембранный пневматический клапан ВО для пара (при регулировании температуры).
6. Два мембранных пневматических клапана ВЗ для воды (при регулировании температуры).
7. Два позиционера П-1.
8. Панель дистанционного управления БПДУ.
9. Фильтр для воздуха.
10. Редуктор для воздуха.

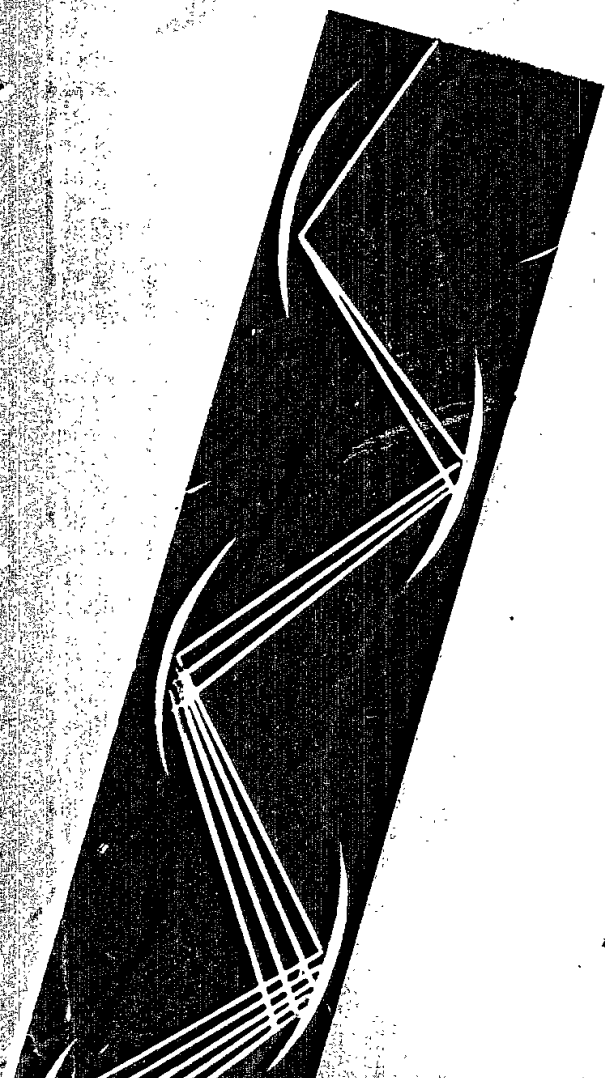
Г04400.

ГОСХИМИЗДАТ. № 51.

Зак. 597

Типография Госхимиздата. Москва, 88, Угрешская.

КОРОТКОВОЛНОВОЕ ИЗЛУЧЕНИЕ



Академия наук СССР

Первостепенный научный и практический интерес представляет исследование коротковолнового излучения Солнца. Как показали исследования последних лет, Солнце, помимо видимого света, испускает излучение, простирающееся в широкую область длин волн, начиная от рентгеновских лучей с длиной волны порядка нескольких стомиллионных долей сантиметра и кончая радиоволнами длиной в несколько метров.

Коротковолновое излучение Солнца (ультрафиолетовое и рентгеновское), а также радиоизлучение связано с физическими процессами, протекающими в малоизученных внешних слоях атмосферы Солнца (хромосфере и короне), и оказывает сильное влияние на атмосферу Земли.

Общая энергия коротковолнового излучения Солнца сравнительно невелика. Однако именно коротковолновое излучение оказывает чрезвычайно большое влияние на земную атмосферу. Объясняется это тем, что коротковолновое излучение обладает чрезвычайно высокой активностью и способно ионизировать молекулы воздуха, вызывая образование ионосферы — сильно ионизированных верхних слоев атмосферы.

Земная атмосфера почти полностью поглощает ультрафиолетовое излучение Солнца. Это поглощающее действие атмосферы предохраняет живые организмы от губительного для них коротковолнового излучения Солнца. В то же время оно делает невозможным исследование этого излучения с Земли. Поглощение молекулами воздуха настолько велико, что для наблюдения коротковолнового излучения необходимо вести наблюдения только за пределами земной атмосферы.

На втором советском искусственном спутнике был установлен прибор, состоящий из трех специальных фотоэлектронных умножителей, которые расположены под углом в 120° друг к другу. Каждый фотоумножитель последовательно переключался несколькими фильтрами из тонких металлических и органических пленок, что позволяло выделять различные диапазоны в рентгеновской области спектра Солнца и линию водорода в далекой ультрафиолетовой области. Электрические сигналы, даваемые фотоумножителем, который был направлен на Солнце, усиливались радиосхемами и передавались на Землю с помощью телеметрической системы.

Вследствие того, что спутник непрерывно изменял свою ориентацию относительно Солнца, а также часть времени находился в тени Земли, для экономии источников питания электрические цепи аппаратуры включались только при попадании Солнца в поле зрения одного из трех приемников света. Это включение осуществлялось с помощью фотосопротивлений, освещаемых Солнцем одновременно с фотоумножителями.

L'étude du rayonnement solaire ultra-violet à ondes courtes présente un très grand intérêt scientifique et pratique. Comme l'ont montré les recherches des dernières années, le Soleil émet en plus de la lumière visible, un rayonnement qui s'étend dans un très large domaine de longueurs d'ondes allant des rayons X à longueur d'onde de l'ordre de quelques cent-millionièmes de centimètre jusqu'aux ondes radioélectriques de plusieurs mètres de longueur.

Le rayonnement à ondes courtes du Soleil (ultra-violet et à rayons X) ainsi que le rayonnement radioélectrique sont en rapport avec les processus physiques se développant dans les couches supérieures peu étudiées du Soleil (chromosphère et couronne) et ont une très grande influence sur l'atmosphère terrestre.

L'énergie totale du rayonnement à ondes courtes du Soleil est relativement restreint. Cependant c'est le rayonnement à ondes courtes qui exerce une très grande influence sur l'atmosphère terrestre. Ceci s'explique par l'activité extrêmement élevée du rayonnement à ondes courtes et sa faculté de ioniser les molécules de l'air en formant la ionosphère — les couches supérieures fortement ionisées de l'atmosphère.

L'atmosphère terrestre absorbe presque complètement le rayonnement ultra-violet du Soleil. Cette action absorbante de l'atmosphère terrestre préserve les organismes vivants contre l'action néfaste du rayonnement solaire à ondes courtes cependant elle rend impossible l'étude de ce rayonnement à partir de la terre. L'absorption des molécules d'air est tellement importante qu'on ne peut observer le rayonnement à ondes courtes qu'au delà des limites de l'atmosphère terrestre.

Le deuxième spoutnik artificiel soviétique était doté d'un appareil comportant trois multiplicateurs photoélectroniques spéciaux disposés sous un angle de 120° par rapport l'un à l'autre. Chaque photomultiplicateur était successivement recouvert de plusieurs filtres composés des pellicules organiques et métalliques fixes ce qui permettait de dégager différentes bandes dans le domaine des rayons X du spectre solaire et la raie de l'hydrogène située dans la région extrême ultra-violet. Les signaux électriques du photomultiplicateur orienté vers le soleil étaient amplifiés par des schémas radioélectriques et envoyés à la terre au moyen d'un système téléométrique.

Etant donné que le spoutnik changeait continuellement son orientation par rapport au Soleil et que pendant un certain laps de temps il se trouvait à l'ombre de la terre, pour économiser les sources d'alimentation de l'appareillage, celles-ci n'étaient enclenchées que lorsque le Soleil entraînait dans le champ visuel de l'un des trois récepteurs de lumière. Cet enclenchement était réalisé au moyen de photorésistances, éclairées en même temps que les photomultiplicateurs.

Studies of the X-ray and short wave-length ultraviolet radiation of the Sun are of major scientific and practical interest. Research done in the last few years has shown that, along with visible light, the Sun emits radiations of a wide range of wave-lengths, beginning with X-rays having wave-lengths of a few hundred millions of a centimetre and ending with radio waves several metres long.

Solar ultraviolet and X-ray radiation as well as the radio waves emitted by the Sun are directly connected with the physical processes going on in the little known outer layers of the Sun's atmosphere (the chromosphere and the corona) and have a strong effect on the atmosphere of the Earth.

The total energy of solar short wave-length radiation is comparatively small. It is this short wave-length radiation, however, that has an exceedingly strong effect on the Earth's atmosphere; this is due to the fact that the activity of the X-ray and ultraviolet radiation is exceedingly high and this radiation is capable of ionizing air molecules. This results in the formation of the ionosphere — the heavily ionized upper layers of the atmosphere.

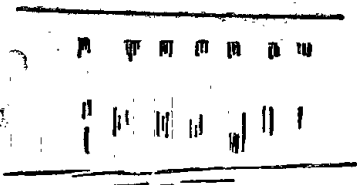
The Earth's atmosphere absorbs solar X-ray and ultraviolet radiation nearly completely. This absorbing effect of the Earth's atmosphere protects living organisms from solar short wave-length radiation which would be fatal to them. But at the same time it makes impossible the study of this radiation from the Earth. The absorption by air molecules is so great that for observing the X-ray and short wave-length ultraviolet radiation it is necessary to carry on observations beyond the limits of the terrestrial atmosphere.

Soviet Sputnik II carried three special photo-electron multipliers set at an angle of 120° to each other. Before each photo-multiplier rotates a disc with several filters consisting of thin metal and organic films. This made it possible to measure different bands in the X-ray region of the solar spectrum and the hydrogen line in the extreme ultraviolet region. Electric signals from the photo-multipliers, which were directed toward the Sun, were amplified by radio circuits and transmitted to the Earth by means of a telemetric system.

In view of the fact that the Sputnik continually changed its orientation with regard to the Sun and spent some time on the part of its orbit not illuminated by the Sun, to save on the power sources, the electrical circuits of the apparatuses were only switched on when the Sun entered the field of vision of one of the three light receptors. The switching on was done by means of resistor photo-cells illuminated by the Sun simultaneously with the photo-multipliers.

**RAYONNEMENT SOLAIRE
À ONDES COURTES**

**SHORT-WAVE RADIATION
OF THE SUN**



САМООРГАНИЗУЮЩАЯ СИСТЕМА РЕГУЛИРОВАНИЯ ДЛЯ АВТОМАТИЧЕСКОГО ПОИСКА РЕЖИМА ПО ЗАДАННОМУ ПРИЗНАКУ

ГОСУДАРСТВЕННЫЙ КОМИТЕТ
СОВЕТА МИНИСТРОВ СССР ПО ХИМИИ

МОСКВА

Самоорганизующаяся система регулирования для автоматического поиска режима по заданному признаку

НАЗНАЧЕНИЕ

Самоорганизующаяся система регулирования предназначена для непрерывного автоматического поиска режима, характеризующегося заданным признаком.

ПРИНЦИП ДЕЙСТВИЯ САМООРГАНИЗУЮЩЕЙСЯ СИСТЕМЫ РЕГУЛИРОВАНИЯ

Блок-схема самоорганизующейся системы регулирования представлена на рис. 1. Выходные координаты, характеризующие режим работы объекта регулирования, воспринимаются

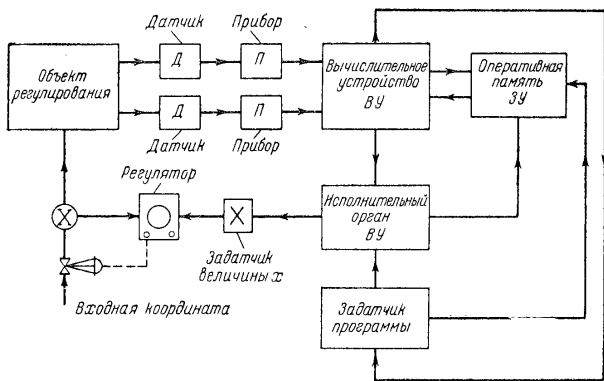


Рис. 1

датчиками Д и измеряются приборами П. Результаты измерений передаются в вычислительное устройство ВУ, где производится вычисление некоторого функционала, имеющего

характерный признак при искомом режиме. Значения этого функционала в дискретные моменты времени передаются в запоминающее устройство ЗУ (оперативная память). Вычислительное и запоминающее устройство периодически обмениваются информацией. С помощью элементов логического действия и реверсивного шагового двигателя вычислительное устройство производит дискретное изменение входных координат объекта регулирования. При этих изменениях учитывается история взаимодействия объекта регулирования с вычислительной машиной, управляющей ходом технологического процесса. Поиски требуемого режима, заданного в критериальной или дифференциальной форме, проводятся непрерывно, путем изменения входных координат объекта регулирования с учетом реакций объекта на прошлые изменения входных координат (учет памяти).

На рис. 2 представлена схема использования самоорганизующейся системы регулирования. Объектом регулирования

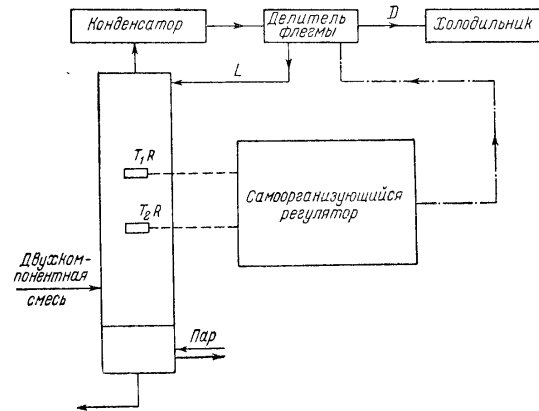


Рис. 2

является дистилляционная колонна непрерывного действия. В колонну непрерывно подается двухкомпонентная смесь. Сверху отбирается низкокипящая фракция этой смеси, а снизу (из куба колонны) — высококипящая. Разделительная способность колонны непрерывного действия зависит от скорости отбора Д. Наилучший режим колонны соответствует максимуму на кривой $\Delta T = f(D)$ (рис. 3).

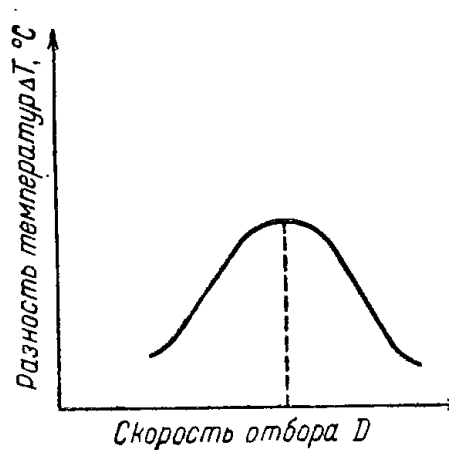


Рис. 3

Искомым признаком является отбор D , соответствующий максимуму на кривой $\Delta T = f(D)$. Температуры T_1 и T_2 (выходные координаты) воспринимаются двумя термометрами сопротивления R_{m_1} и R_{m_2} (рис. 2). Отбор D регулируется самоорганизующейся системой регулирования таким образом, чтобы всегда выполнялся заданный признак: максимум разности температур в двух измеряемых точках колонны.

Dr. Ing. FRANK FRÜNGEL GmbH. HAMBURG-RISSEN

Telephone: 81 21 51

Telex: 021 2883

Cable Code: Früngel-Hamburg

**CINEMATERIEL**

24, rue des Guillemins

LIÈGE — Téléphone : 52.30.91

AUTOMATIC RECORDING CLOUD CEILOMETER

Impuls light System Dr. Früngel

Introduction

The knowledge of the heights of clouds and their variations with time is of great importance for general meteorology and especially with respect to the requirements of the modern airport. The Früngel Ceilometer, for recording cloud heights, and the Früngel Transmissometer for recording visibility or transmission of light, are two modern instruments which meet present day needs for accurate and reliable instrumentation, designed for long time service without special maintenance. While there are various instruments used for measuring cloud heights, the Früngel Ceilometer is unexcelled for accurately recording heights down to 100 feet. It will measure and record such heights with an accuracy well within 5% error. This instrument has proven itself reliable and trouble-free and it was selected last year for installation at several Airports in Europe and North America. The present design is the result of many years of cooperative research and development at the German Meteorological Institute and the Impulse-Physics group working under Dr. Frank Früngel. It records cloud heights, automatically and accurately, during any part of the day or night.

Principles of Operation

The Ceilometer operates on the right-triangle principle of altitude measurements in which the impulse interception angle of the receiving unit is computed with a known distance between the impulse light transmitter and the receiver, to yield the cloud altitude. It is an excellent device for accurately measuring cloud heights up to 800 meters (2,600 ft.) and by increasing the base line one can measure cloud heights to unlimited altitudes.

The transmitter projects continuous light pulses, about 10 per second, with a pulse light duration of $1/\mu\text{s}$ time duration, with an essential peak duration of about $0.3/\mu\text{s}$. Because of the high luminosity of the pulsed light beam, the receiver needs only a small lens. The emitted spectrum is in the range of the medium ultraviolet to the invisible ultraviolet range. Infrared light is not used. The special peak brightness is $5 \cdot 10^7$ candles/cm². With the reflector surface having 2,000 cm², a peak light on the axis of about $2,000 \times 5 \cdot 10^7$ or 10^{11} cd. is obtained. The peak luminous density of the light pulses in the searchlight is considerably in excess over that of the daylight

and the beam of light, collimated by a ground glass parabolical reflector, projects a disc of light approx. the diameter of the reflector, onto the base of the cloud. The great luminous density of the sparked or impulsed light produces a sufficient reflected light intensity, even under daylight conditions and low reflectivity of the clouds, to permit measurements with the small diameter receiver lens. The reflected light pulses are transformed into sharp electrical pulses by the photocell receiver, with its optical lens system and iris. The pulses, amplified in a wideband amplifier, trigger constant amplitude and width pulses in a monostable multivibrator. Each of these pulses produce a contact in the high speed relay which in turn controls the recorder, which is connected by cable to the receiver. The receiver is moved in an arc, by a synchronous motor linked with the recorder unit, from a horizontal to a vertical position in the plane enclosing the axis of the searchlight beam and the base line. The recorder receives the pulses every time the moving point of interception of the optical axis of the transmitter light and the receiver, or every time the receiver sees the disc of light which the transmitter projects to the base of the cloud. The scribe on the recorder is linked by another synchronous motor connected with the receiver, and every make-contact of the high speed relay in the receiver triggers a spark imprint from the pen to the recording paper. There are scales on the receiver, graduated in degrees, as well as on the recorder, which can be read by the light of pulsed neon lamps. Also, there is a scale on the recorder graduated in height from 150 to 2,400 feet, matched to the length of the selected base line. A measurement is made every 30 seconds. The amplifier circuit is a wideband 3 tube amplifier and the output signal starts a monostable flip-flop, integrating three pulses in a time space corresponding to the pulse frequency of the transmitter. Sparks are made on the graph only when three signals are received; avoiding mistaken indications by single disturbing pulses. Flip-flop sensitivity can be regulated by a knob on the receiver. The length of the line on the graph depends upon the transparency of the cloud and it indicates a direct measurement of the angle range in which the reflected signal is received. The receiver continuously moves in the angular range and normally registers the scanning distance of 0 to 20 degrees and back within a total time of one minute. The receiver can be adjusted to move over a range 0 to 90 degrees, or any portion thereof. For meteorological research, the transmitter may be used with several receivers fixed on different base lengths, when specific measurements are desired. Caesium-Antimon cathode type photocells are used, good for the spectral range from blue to near ultra-violet. Problems due to thermal disturbance do not exist, as is found with the infra-red photocells. The light pulse of the trans-

mitter is strong enough to avoid the problem of noise, as it transmits a signal considerably higher than that found in continuous daylight. The lamp is demountable, and the housing may be removed and cleaned when it becomes blackened. Argon gas, filling the tube at an equalized pressure, increases the light output by a factor of 5. The lamp uses electrodes 5 mm. in diameter made from Tungsten. The projected light is concentrated in a beam of approx. 0.2° by means of a Mangin mirror. A trigger system is used to start the spark in the lamp.

The receiver has a filter to favor the high color temperature pulse light against the low color temperature daylight and it has an adjustable width slit to protect the photocell. No relay operated break-contacts are used at these cause considerable trouble; make the contacts are only used to guarantee almost an unlimited life of the relay serving the measuring circuit and the synchronization with the recorder.

The gas pressure in the spark chamber can be increased at will. At one atmosphere pressure in the spark chamber, a light output of 0.5×10^{11} candles is emitted. By increasing the Argon gas pressure to 5 atm. the power or light output increases to $2 \cdot 10^{11}$ candles.

When the transmitter is mounted on top of the roof of the airport tower it is fitted with a small diameter painted reflector plate which regularly sends a reflected beam of light back to the receiver for the fixed zero point (usually about 30°). The angle of the zero reflected light naturally will depend upon the height at which the transmitter is mounted.

The transmitter and the receiver are designed for outdoor use and protected against all types of weather. The transmitter has a heating coil to melt snow and ice from the sloping glass plate. It weighs 300 lbs. (135 kg) and is 51" high (1,300 mm) x 32" dia. (800 mm.). The base measures 32" (800 mm.) square. The receiver is very compact, and weighs 85 lbs. (32 kg.) with the scanning head. It is approx. 59" high x 8-1/2" dia. (1,300 x 220 mm.). It requires a base of 16" square (400 mm.). The recorder with accessories weighs approx. 24 lbs. (10-1/2 kg) and is 16" x 12" x 9-1/2" (400 x 320 x 240 mm.). The recorder base is 14" square (350 mm.). The power consumption of the transmitter is 1.0 kw. and the heating system requires 0.5 kw. extra. The power consumption of the receiver and recorder is 0.2 kw. The unit withstands a vibration factor during transportation of lg. The measuring accuracy for the lowest cloud is plus or minus 0.5° arc.

Based on a 250 ft. (75 m.) baseline between transmitter and receiver, the following data may be obtained:

Angle	Height of Cloud	Width of dispersion plus or minus
70°	830 ft. (250 m.)	20 ft. (6.25 m.)
90°	1,460 ft. (440 m.)	80 ft. (18 m.)
84.5°	2,640 ft. (800 m.)	165 ft. (50 m.)

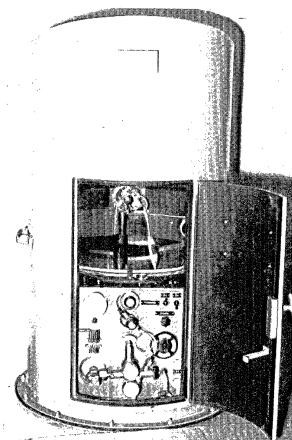
With a baseline of 250 ft. the range of automatic recording is more than 3,000 ft. (900 m.). For great heights, the instrument may be manually operated and a further increase in the operating range may be obtained by placing a small auxiliary oscilloscope in the receiver circuit. When the base is increased to 330 ft. (75 m.) cloud height measurements up to 4,000 ft. (1200 m.) may be recorded accurately, in small angle ranges.

Maintenance

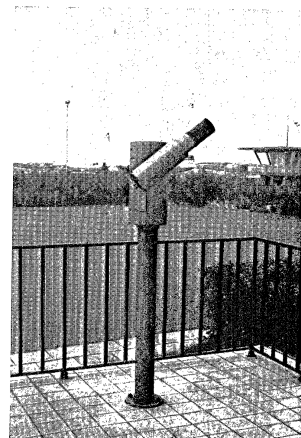
This instrument requires very little maintenance. The transmitter has two ranges, one for permanent working, that means for heights up to 1500 feet and an additional range with an 8 times higher energy per impuls for single measurements. After about 300 hours operation the spark lamp glass shell may become darkened and if it becomes too discolored, it should be removed and replaced with a spare glass housing. In order to guarantee a simple and quick exchange the used lamp can be replaced by a second one, so that the cleaning operation could be done afterwards. After about 300 hours of operation, the electrodes should be changed or adjusted, as the space-gap gets larger. This work takes about 6 minutes. The receiver tubes, photocell and other tubes are factory guaranteed and they usually last about 1 year or 10,000 hours. A complete manual of instructions with photographs is sent with each Ceilometer.

Advantages

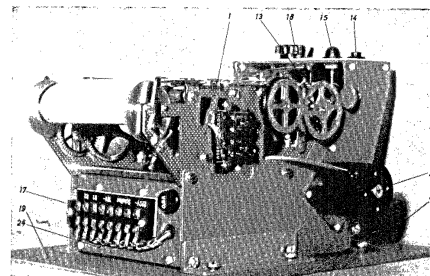
1. Impulsed light goes about three times higher than that of the rotating beam (normally approx. 3,000 ft. = 1,000 m.).
2. The instrument has 2 ranges of spot energy, the normal one for permanent working, sufficient for abt. 1500 ft. (500 m.) height, and a 8 time higher additional spot energy range for a max. load of the lamp for measuring great heights.
3. Small size receiver minimizes problems of weather such as snow, ice, rain, etc. A heating element melts ice and snow. Receiver opening is only 3" diameter, housing is 6" diameter.
4. Shows development of weather, indicates sinking clouds sharply on graph.
5. Impulsed light beams are projected at exact angles, always. They do not spread causing error to receiver.
6. As shot massive tungsten electrodes are used, they will not warp or change position, eliminating possibility of error which is caused by warpage of long electrodes in lamps of other ceilometers.
7. The impulsed light used is the most intensive spark light available. This permits the use of the transmitter for special work in meteorological research.
8. Several receivers can be used
 - a- to measure in the blue range,
 - b- to measure in the infra-red range for deeper penetration of the clouds,
 - c- additional receivers may be set at different base lines for measuring to various heights, up to approx. 10,000 m. at night.
9. Spark light can not extinguish suddenly, as f. i. an electric bulb, from which the filament is burnt. The employment of the impuls-lamp will be limited by the slow decreasing of the measuring range.



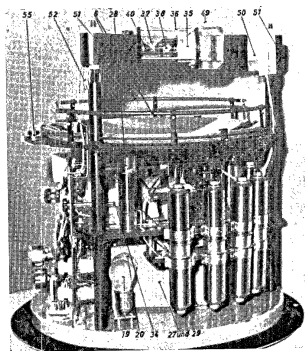
Shows the transmitter with sloping top and open door to permit viewing of the controls and gas pressure valves.



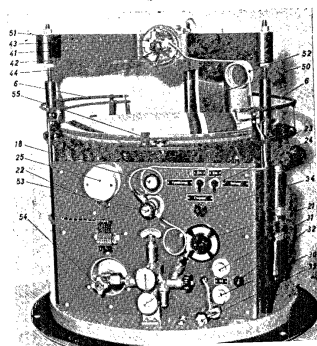
Shows the receiver mounted at the side of the post which contains the mechanical and electrical components. The post has a door to permit easy access to controls and adjustments.



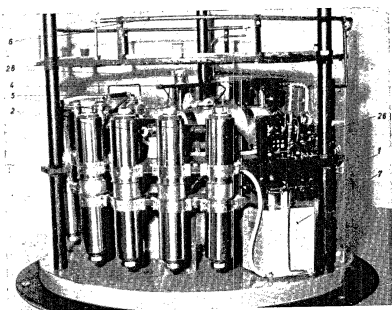
Shows the recorder with synchronizing button. (Built-in unit, also available as separate box)



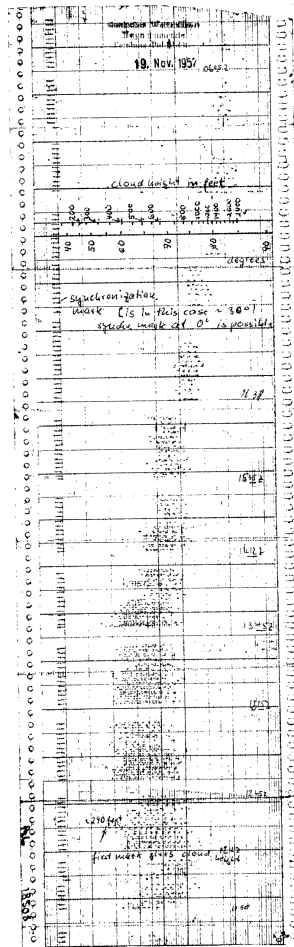
Side view of the transmitter showing the knobs of the gas valves at left and some of the condenser discharge batteries, with curved reflector (shown above) batteries and the spark chamber above the reflector.



Front view of the transmitter, showing Argon gas valves, controls and other components. The Argon gas line going to the spark chamber is visible as well as the curved reflector plate.



Showing the rear of the transmitter with balance of condenser batteries and other electrical components. The Argon Gas cylinder is partly visible in back of the cylinders, near the top.



Deutscher Wetterdienst - Flugwetterwarte Hamburg-Fuhlsbüttel

for the light is reflected and also refracted to a greater extent by dense clouds. The zero reference mark (000) at the extreme left is clearly shown. The graph is not marked in terms of height but a master scale is provided for such measurements. This is best, due to the fact that the receiver may be set for different base lines.

Sample graph obtained from Celiometer Recorder. This shows rising fog or mist in November at Hamburg Airport. The first mark in thick is quite thin. As the concentration increases and the penetration decreases, the penetration is indicative of the density of the clouds, the greater the density, the lower the penetration.

Schematic diagram Pulsed optical cloud height meter

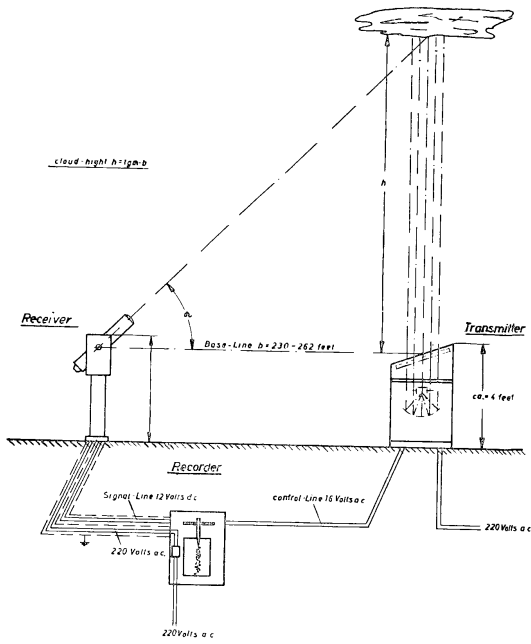


Fig. 11 - Schematic diagram showing pulsed optical cloud height meter.

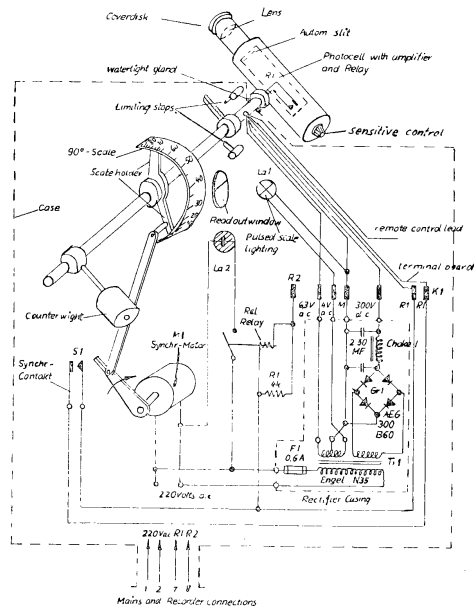


Fig. 12 - Layout and circuit diagram - receiver.

Fig. 14 - Circuit diagram amplifier - pulsed optical cloud height measuring set.

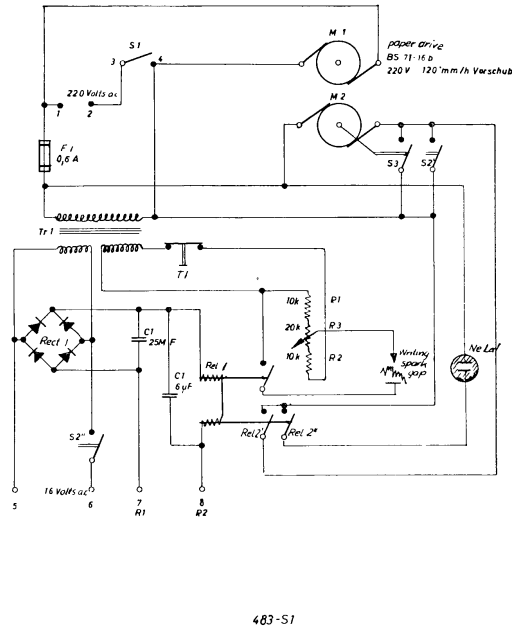
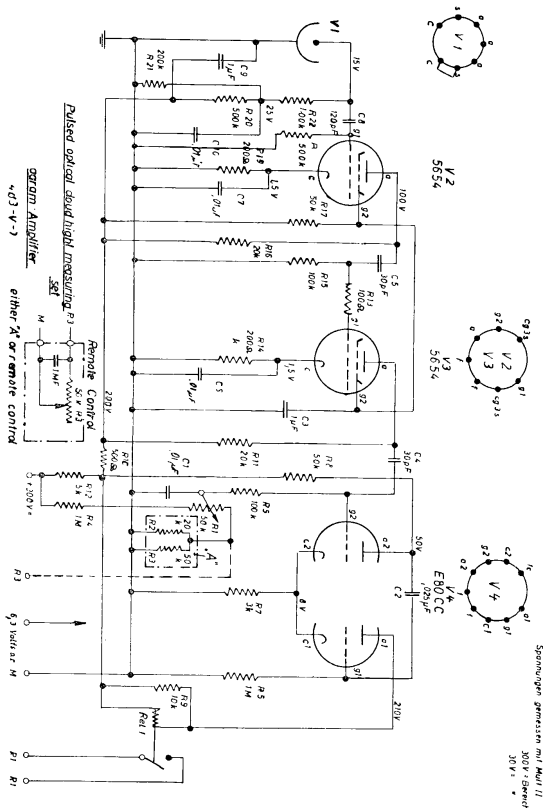


Fig. 13 - Circuit diagram recorder.



Уютная метр

СЕНТЯБРЬ СССР НА ВСЕМИРНОЙ ВЫСТАВКЕ 1958 Г. В БРЮССЕЛЕ



ULTRA-TAUMETRE

ULTRA-TAUMETER

L'appareil est destiné aux études des processus de la relaxation de la luminescence et de la photoconductibilité, qui s'écoulent en $5 \cdot 10^{-8}$ – 20 sec.

L'appareil se distingue par son caractère universel et permet d'effectuer les études suivantes:

- Etudes des processus rapides exponentiels et non-exponentiels du déclin et de la montée de la luminescence et de la photoconductibilité, processus excités par les impulsions périodiques rectangulaires de la lumière produites par le modulateur mécanique de lumière.

Durées des temps étudiés: de 10^{-7} à 10^{-1} sec.

- Etudes des processus de relaxation plus lents de la luminescence et de la photoconductibilité, excités par les impulsions **apériodiques** rectangulaires de la lumière, produites par le modulateur électrodynamique de lumière.

Durée des temps étudiés: de 10^{-3} à 20 sec. Ce genre d'études est surtout utile pour la démonstration et l'examen d'un nouveau phénomène de l'autostimulation.

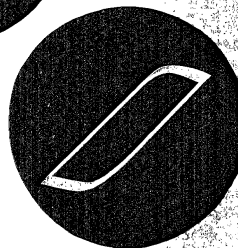
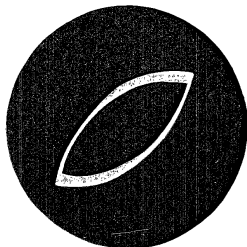
- Etudes des processus rapides exponentiels de la luminescence et de la photoconductibilité, excités par la **source lumineuse d'impulsion**. Une très grande intensité de l'excitation permet d'étudier les phénomènes faibles.

Durées des temps étudiés: de $5 \cdot 10^{-8}$ – 10^{-6} sec.

- Etudes de la photoconductibilité et du signe des porteurs par méthode de condensateur en régimes périodiques et d'impulsion.

L'appareil représente un ensemble des dispositifs optiques, mécaniques et électroniques, grâce auxquels l'appareil devient tout un laboratoire compact pour les études de la cinétique de la photoluminescence et de la photoconductibilité.

Encombrement – 3200 × 1500 × 1300 mm.
Poids – 500 kg.



The instrument is designed for the purpose of examining the process of relaxation of luminescence and photoconductivity which takes place during the period of time of $5 \cdot 10^{-8}$ to 20 sec.

Pour les conditions des livraisons adressez-vous: "STANKOIMPORT", Moskva, G-200, Smolenskaja-Sennaja-ploshchad, 32/34.

The instrument is distinctive for high universality and permits carrying out the following kinds of work:

- Investigation of rapid exponential and non-exponential processes of decay and rise of luminescence and photoconductivity, which are excited by periodical square light impulses produced by mechanical light modulator.

Periods of time under study: from 10^{-7} to 10^{-1} sec.

- Study of slower relaxation processes of luminescence and productivity excited by **aperiodical** square light impulses produced by **electrodynamical** light modulator. Periods of time under study: from 10^{-3} to 20 sec. (This kind of work is especially suitable for the demonstration and study of the new phenomenon called autostimulation.)

- Study of rapid exponential luminescence and photoconductivity processes excited by an **impulse** light source. The great intensity of excitation permits the study of rather weak phenomena.

Periods of time under study: $5 \cdot 10^{-10}$ – 10^{-6} sec.

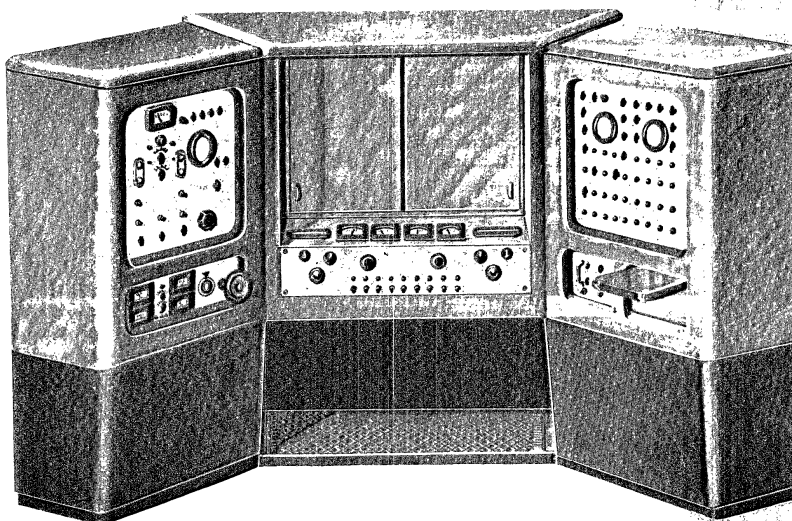
- Study of photoconductivity and carrier polarity of the **condenser method** in **periodical and single-pulse conditions of excitation**.

The instrument is a combination of optical, mechanical and electronic assemblies making it in fact a compact laboratory for the study of kinetics of photoluminescence and photoconductivity.

Overall dimensions of the instrument – 3200 × 1500 × 1300 mm.

Weight – 500 kg.

For delivery terms apply to "STANKOIMPORT", Moskva, G-200, Smolenskaja-Sennaja-ploshchad, 32/34.



ULTRA-TAUMETER

Das Gerät ist für Erforschung der Relaxation der Lumineszenz und der Photo-Leitfähigkeit, die innerhalb einer Dauer von $5 \cdot 10^{-8}$ bis zu 20 Sekunden verlaufen.

Das Gerät zeichnet sich durch große Vielseitigkeit aus, und ermöglicht die Durchführung folgender Arbeiten:

- Das Studium von schnellen Exponential- und Nicht-exponential-Prozessen der Ab- und Anklung der Lumineszenz und der Photoleitfähigkeit, die durch periodische von einem mechanischen Lichtmodulator erzeugte, rechteckige Lichtimpulse erregt werden.

Die zu untersuchenden Zeiten betragen von 10^{-7} bis zu 10^{-1} Sekunden.

- Das Studium von langsameren Relaxations-Prozessen der Lumineszenz und der Photo-Leitfähigkeit, die durch **aperiodische** von einem elektrodynamischen Lichtmodulator erzeugte rechteckige Lichtimpulse erregt werden.

Die zu untersuchenden Zeiten betragen von 10^{-3} bis zu 20 Sekunden. (Dieses Arbeitsverfahren ist besonders geeignet für die Vorführung und das Studium der neuen Erscheinung des „autostimulation“.)

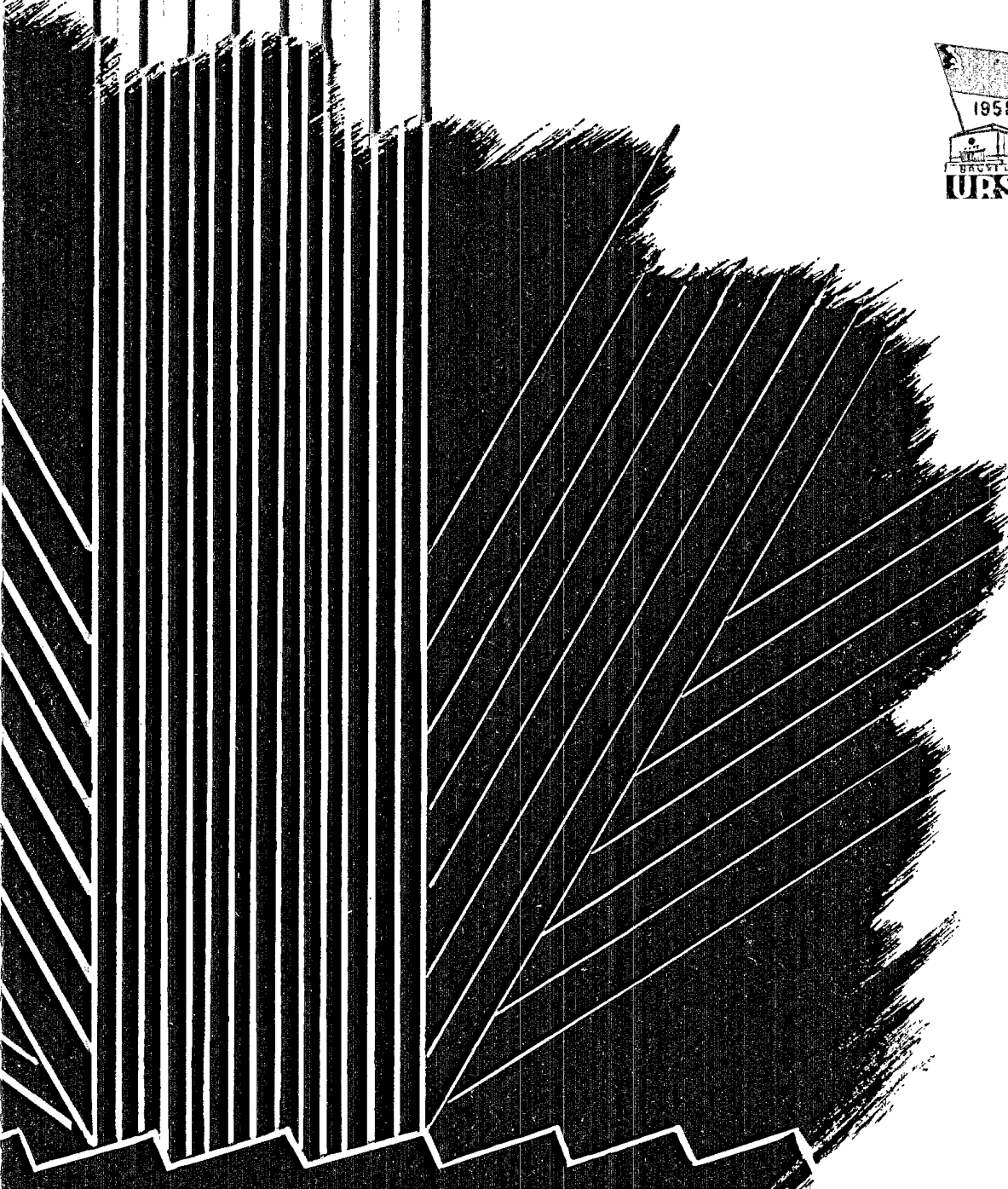
- Das Studium von raschen Exponential-Prozessen der Lumineszenz und der Photo-Leitfähigkeit, die von einer **Impuls**-Lichtquelle erregt werden. Die erhebliche Intensität der Erregung ermöglicht das Studium von schwachen Erscheinungen. Die zu untersuchenden Zeiten: $5 \cdot 10^{-8}$ – 10^{-6} Sekunden.

- Das Studium der Photo-Leitfähigkeit und der **Träger-Vorzeichen** nach der **Kondensator-Methode** bei **periodischen** und **einmaligen** Regimen.

Das Gerät stellt eine Kombination von optischen, mechanischen und elektronischen Einheiten dar, die es zu einem kompakten Laboratorium für das Studium der Photo-Lumineszenz- und der Photo-Leitfähigkeit-Kinetik macht.

Äußere Abmessungen des Gerätes: 3200x1500x1300mm
Gewicht: 500 kg

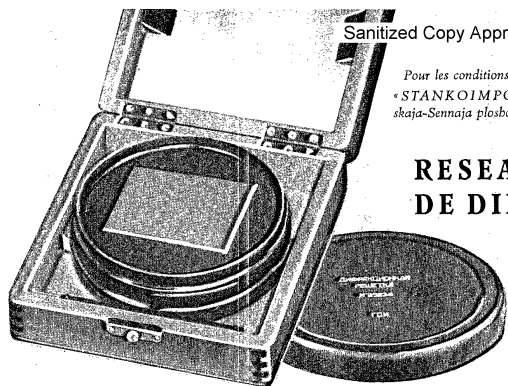
Alle Lieferungsanfragen sind an „STANKOIMPORT“,
Moskva, G-200, Smolenskaja-Sennaja Ploschad, 32/34,
zu richten.



Дифракционные
решетки

СЕКЦИЯ СССР НА ВСЕМИРНОЙ ВЫСТАВКЕ 1958 Г. В БРЮССЕЛЕ

Pour les conditions des livraisons adressez-vous :
« STANKOIMPORT », Moskva, G-200, Smolenskaja-Sennaja ploshchad, 32/34.



RESEAUX DE DIFFRACTION

Les réseaux de diffraction peuvent être utilisés dans les différents appareils spectraux destinés à l'analyse spectrale industrielle et aux recherches scientifiques.
En Union Soviétique on fabrique tous les principaux types de réseaux de diffraction utilisés en spectroscopie.

CARACTERISTIQUES ESSENTIELLES DES RESEAUX DE DIFFRACTION

Dimension du réseau, mm	Rayon de la courbure, mm	Nombre de traits per mm	Dimensions de la surface portant des traits, mm	Région spectrale, m μ	Dimension du réseau, mm	Rayon de la courbure, mm	Nombre de traits per mm	Dimensions de la surface portant des traits, mm	Région spectrale, m μ
Réseaux de diffraction plans sur les couches d'aluminium					Réseaux de diffraction concaves sur les couches d'aluminium				
30x30x10	—	1200	30x30	200-1000	Ø 90x15	1000	1200	60x50	200-1000
40x40x10	—	1200	40x40	200-1000	Ø 120x15	1995	1200	80x60	200-1000
Ø 90x15	—	1200	60x50	200-1000	Ø 150x20	2992	1200	100x60	200-1000
Ø 120x15	—	1200	80x70	200-4500	Ø 170x20	6020	1200	145x70	200-1000
Ø 150x20	—	1200	100x90	200-4500	Réseaux de diffraction concaves sur le verre				
150x110x20	—	300	100x100	600-4500	Ø 90x15	1000	1200	60x50	Régions des rayons X et ultra violet à vide
150x75x20	—	200	100x100	600-4500	20x35x10	1000	1200	18x32	
Ø 170x20	—	1200	120x60	200-1000	30x40x15	1995	1200	28x35	
Réseaux de diffraction plans sur les couches d'aluminium					Echelettes sur les plaques métalliques				
170x170x25	—	1200	120x110	200-4500	220x220x35	—	100, 50, 24, 12, 6, 4, 2	200x200	2,5-600 μ
		600	150x140	200-4500	270x270x40	—	50, 24, 12, 6, 4, 2	250x250	5-600 μ
		300			320x20x40	—	12, 6, 4, 2	300x300	20-600 μ

Dans les ordres bas du spectre les réseaux de diffraction donnent les raies spectrales pratiquement parfaites et possèdent un pouvoir séparateur proche à la valeur théorique. Le pouvoir séparateur des meilleurs réseaux dépasse 400 000.

Les réseaux de diffraction à 600 traits per mm ont une intensité relative des « fantômes » de Rowland dans le 1^{er} ordre de 0,01 à 0,2%. Les « fantômes » de Lyman ne sont pas observés.

Les réseaux plans et concaves sur les couches d'aluminium et les échelettes sont capables de concentrer de 40% à 80% de la lumière réfléchiée dans l'ordre du spectre donné.

L'angle d'inclinaison des gradins du réseau, qui détermine la position de la concentration maximum, varie dans les limites de 5° à 60°.

The diffraction gratings are used in various spectral instruments designed for industrial spectral analysis as well as for scientific purposes.
The Soviet Union manufactures all of the main types of diffraction gratings.

BRIEF CHARACTERISTICS OF THE GRATINGS

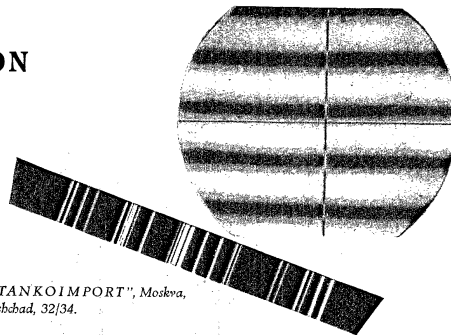
Dimensions of gratings, mm	Curvature radius, mm	Number of lines per mm	Dimensions of grating surface, mm	Spectral region of application, m μ	Dimensions of gratings, mm	Curvature radius, mm	Number of lines per mm	Dimensions of grating surface, mm	Spectral region of application, m μ
I. Plane gratings on aluminum layers					II. Concave gratings on aluminum layers				
30x30x10	—	1200	30x30	200-1000	Ø 90x15	1000	1200	60x50	200-1000
40x40x10	—	1200	40x40	200-1000	Ø 120x15	1995	1200	80x60	200-1000
Ø 90x15	—	1200	60x50	200-1000	Ø 150x20	2992	1200	100x60	200-1000
Ø 120x15	—	1200	80x70	200-4500	Ø 170x20	6020	1200	145x70	200-1000
Ø 150x20	—	1200	100x90	200-4500	III. Concave gratings on glass				
150x110x20	—	300	100x100	600-4500	Ø 90x15	1000	1200	60x50	X-ray and vacuum ultra-violet regions
150x75x20	—	200	100x100	600-4500	20x35x10	1000	1200	18x32	
Ø 170x20	—	1200	120x60	200-1000	30x40x15	1995	1200	28x35	
IV. Echelettes on metal specimens									
170x170x25	—	1200	120x110	200-4500	220x220x35	—	100, 50, 24, 12, 6, 4, 2	200x200	2,5-600 μ
		600	150x140	200-4500	270x270x40	—	50, 24, 12, 6, 4, 2	250x250	5-600 μ
		300			320x20x40	—	12, 6, 4, 2	300x300	20-600 μ

In low orders of spectrum the gratings present practically perfect spectral lines and have their resolving power approaching the theoretical value. The resolving power of best gratings exceeds 400 000.

With the gratings having 600 lines per mm, the relative intensity of Rowland's "ghosts" in the first order lies in the range of 0.01 to 0.2%. No Lyman's "ghosts" are observed.

Plane and concave gratings on aluminum layers and echelettes are capable of concentrating from 40% to 80% of reflected light (of a definite spectral interval) in a given spectrum order. The inclination angle of grating stages determining the position of the concentration maximum changes within a range of 5° to 60°.

DIFFRACTION GRATINGS



For delivery terms apply to "STANKOIMPORT", Moskva, G-200, Smolenskaja-Sennaja ploshchad, 32/34.

BEUGUNGSGITTER

Die Beugungsgitter werden in den verschiedensten Spektralgeräten angewandt, die so für die industrielle Spektralanalyse, wie für wissenschaftliche Zwecke bestimmt sind.

In der Sowjetunion werden alle bekannten Typen der Beugungsgitter hergestellt.
Die Hauptdaten der wichtigsten Typen der Gitter lauten wie folgt:

Größe des Gitters, mm	Krümmungsradius, mm	Strichzahl pro mm	Größe der mit Strichen bedeckten Fläche, mm	Spektral-Anwendungs-Gebiet m μ	Größe des Gitters, mm	Krümmungsradius, mm	Strichzahl pro mm	Größe der mit Strichen bedeckten Fläche, mm	Spektral-Anwendungs-Gebiet m μ
Flache Gitter auf Aluminiumschichten					Konkave Gitter auf Aluminiumschichten				
30×30×10	—	1200	30×30	200—1000	∅ 90×15	1000	1200	60×50	200—1000
		600			∅ 120×15	1995	1200		
40×40×10	—	1200	40×40	200—1000	∅ 150×20	2992	1200	100×60	200—1000
		600			∅ 170×20	6020	1200		
∅ 90×15	—	1200	60×50	200—1000	Konkave Gitter auf Glas				
		600			∅ 90×15	1000	1200	600	60×50
∅ 120×15	—	1200	80×70	200—4500	20×35×10	1000	1200	18×32	
		600	100×90	200—4500	30×40×15	1995	1200	28×35	
		300	100×100	600—4500	Escheletten auf Metallrohstücken				
150×110×20	—	300	120×60	200—1000	220×220×35	—	100, 50, 24, 12, 6, 4, 2	200×200	2,5—600 μ
		200			120×110	200—4500	270×270×40	—	50, 24, 12, 6, 4, 2
150×75×20	—	1200	150×140	200—4500	320×320×40	—	12, 6, 4, 2	300×300	20—600 μ
		600			170×170×25	—	1200	300	200
		300							
		200							

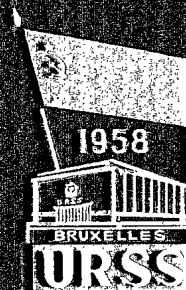
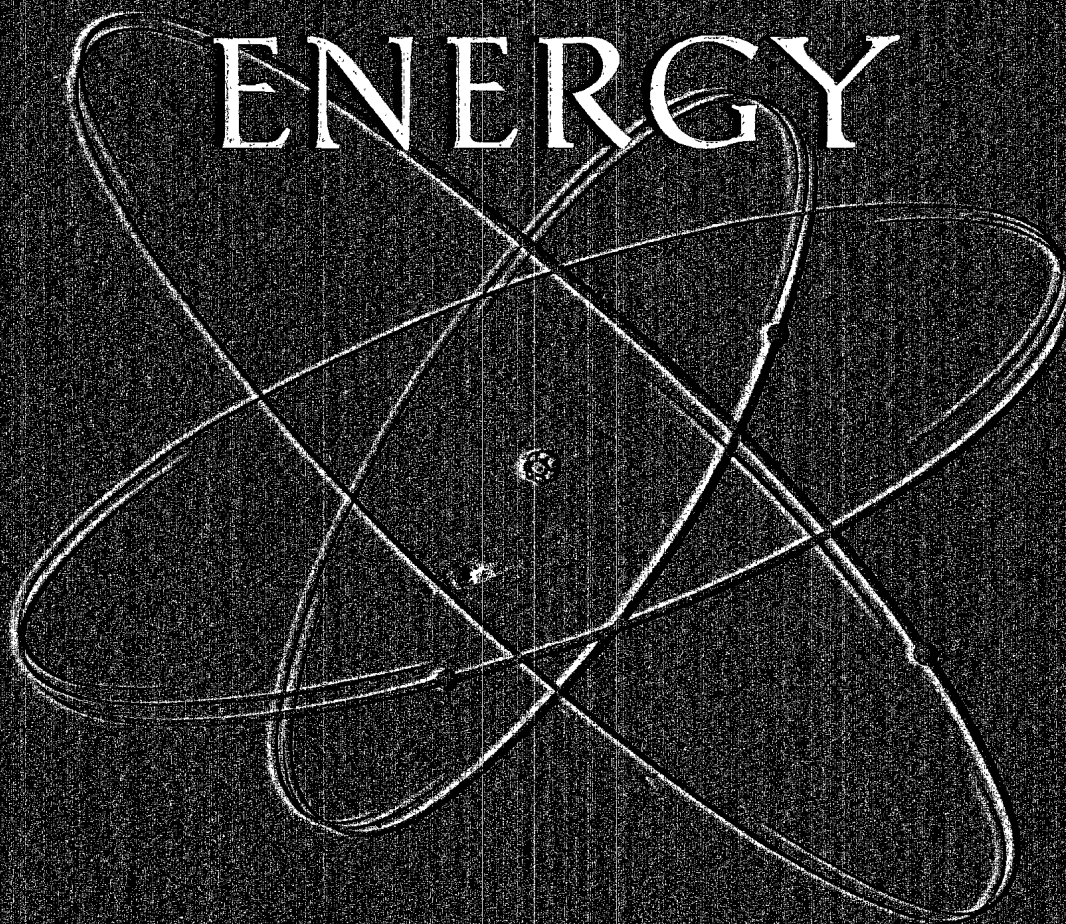
Bei niedrigen Spektral-Ordnungen ergeben die Gitter praktisch vollkommene Spektrallinien und besitzen dabei ein Auflösungsvermögen, das nahe an dem theoretischen Wert liegt. Das Auflösungsvermögen der besten Gitter übersteigt 400 000.

Bei den Gittern, die 600 Striche auf 1 mm aufweisen, beträgt die relative Intensität der Rowland's „Geister“ in der ersten Ordnung 0,01 bis 0,2%. Lyman's „Geister“ werden nicht beobachtet.

Flache und konkave Gitter auf Aluminiumschichten besitzen die Fähigkeit, von 40% bis zu 80% des reflektierten Lichtes (bestimmten Spektralbereich) in der gegebenen Ordnung des Spektrums zu konzentrieren. Der Beugungswinkel der Stufen des Gitters, der die Lage des Konzentrations-Maximums bestimmt, ändert sich im Bereich von 5° bis 60°.

Alle Lieferungsanfragen sind an
„STANKOIMPORT“ Moskva, G-200, Smolenskaja-Sennaja Ploschad, 32/34 zu richten.

THE PEACEFUL USES OF ATOMIC ENERGY



USSR SECTION AT THE UNIVERSAL AND
INTERNATIONAL EXHIBITION OF BRUSSELS 1958

**THE PEACEFUL USES
OF
ATOMIC
ENERGY**

ATOM SERVES THE PEOPLE

The inexhaustible resources of latent energy existing within the atom have for the first time been made to serve mankind.

Atomic energy opened up entirely new possibilities in the field of engineering and economics resulting in a growing power of man over nature. The emergence of this new power source has immeasurably increased the available power reserves. Nuclear radiations and radioactive isotopes have become a powerful means for developing research in every branch of science and engineering and a practical tool in biology, medicine, agriculture, and industrial production.

At present there is hardly a single person on the earth who would deny the intimate relationship between the further progress of mankind and the improvement of our knowledge of the atomic nucleus, the peaceful applications of atomic energy. Unfortunately, there are still some people, who see the main object of atomic engineering in accumulating ever increasing stock piles of atomic weapons. In the Soviet Union the achievements of nuclear physics and engineering have been made to

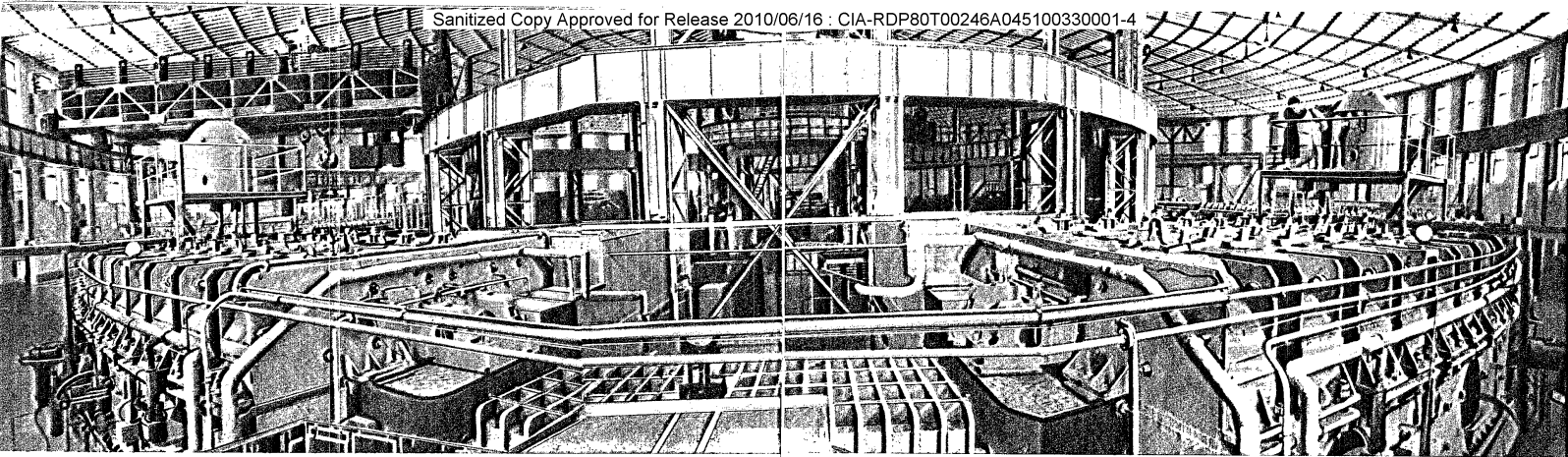
serve the people. There is hardly a branch of national economics that does not use to some extent or other new materials and methods brought into existence by the atomic age.

In the USSR National Pavilion you will be able to see how nuclear engineering is penetrating into the power industry and mechanical engineering, technology and transport, medicine and agriculture. The exhibits on show demonstrate the achievements of the USSR in peaceful uses of atomic energy.

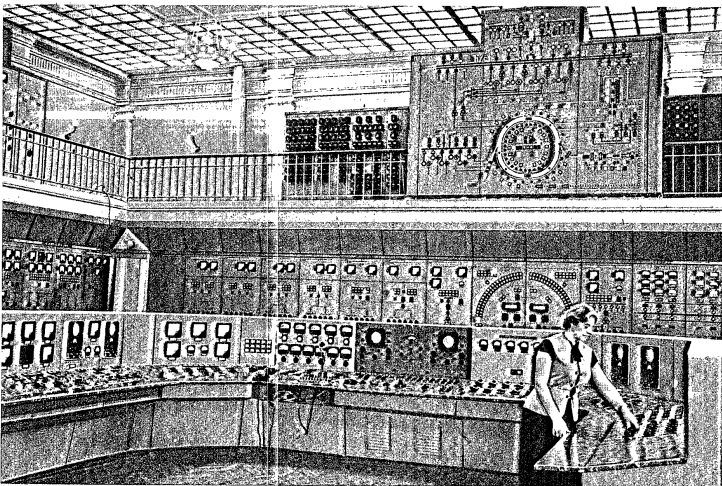
You will see a model of the world's most powerful 10 Bev accelerator, and the world's first atomic ice-breaker "Lenin" launched recently.

Here you will see how Soviet engineers are solving the problem of building the most reliable and economic atomic power station, and will learn what means are employed in the USSR to protect personnel handling radioactive materials.

The following pages will familiarize you very briefly with the main trends in the peaceful uses of atomic energy in the USSR.



Ring-shaped electromagnet of the world's largest charged-particle accelerator, the 10 Bev proton synchrotron



Control room of the 10 Bev proton synchrotron

ACCELERATORS

To investigate the structure of atomic nuclei, to master the laws governing nuclear forces and to find new paths toward the utilization of all power resources of atomic nuclei, huge experimental apparatus called charged particle accelerators are built. The world's largest particle accelerator operates at the Joint Nuclear Research Institute in the town of Dubna near Moscow. This gigantic atomic machine provides acceleration of protons with energies of up to 10 billion electron volts. The accelerator is of the proton synchrotron type. Its annular electromagnet is about 60 m in diameter and weighs about 36,000 t.

The vacuum maintained in the vacuum chamber located between the poles of the an-

nular electromagnet is equal to that attained in the best electronic devices. For this purpose 56 powerful high-vacuum pumps are installed along the chamber. For periodic acceleration of protons moving inside the chamber and for bringing their velocity close to that of light, the proton synchrotron uses a special system of high-frequency power source to supply the accelerating electrodes. Beams of accelerated protons and other particles obtained in the accelerator (neutrons, mesons, hyperons, and antiparticles) are widely employed for various nuclear experiments. Their results will enrich us with further knowledge which will help us in learning the laws of nature. The Joint Nuclear Research Institute has one more large

NUCLEAR POWER

After the world's first atomic power station was commissioned in the USSR on June 27, 1954, the attention of all mankind turned with a new vigour to the problems of nuclear power.

Use of nuclear energy for obtaining electricity will result in improved well-being and living standards of all nations of the world.

Development of nuclear power is of particular significance for the globe's areas poor in natural fuel. Construction of atomic power station in such areas will help to develop industries, irrigate arable land, increase the crops, raise the productivity of animal husbandry, improve the living and working conditions of the population.

The Soviet Union possesses huge reserves of natural fuel and hydropower, but the USSR also has areas remote from natural fuel re-

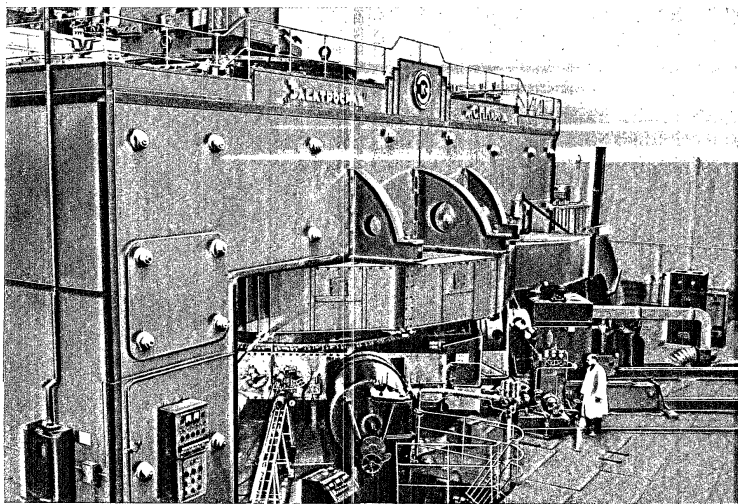
sources, where it is rational to construct atomic power stations right now.

In connection with this an extensive programme of building experimental atomic power stations has been undertaken in the Soviet Union. The total installed capacity of these stations is to reach 2—2.5 million kilowatts within the next few years.

Construction is already underway of a nuclear power giant, a commercial 420,000 kw atomic power station with nuclear reactors using ordinary water as a coolant and moderator.

It will use natural and slightly enriched uranium (containing an approximately double percentage of uranium-235 compared with natural uranium) as nuclear fuel.

Two circuits will be employed for cooling like at the world's first atomic power station.



The world's largest 680 Mev synchrocyclotron

accelerator, a synchrocyclotron, for acceleration of protons with energies of up to 680 Mev. It also ranks among the world's best accelerators of its type now in operation.

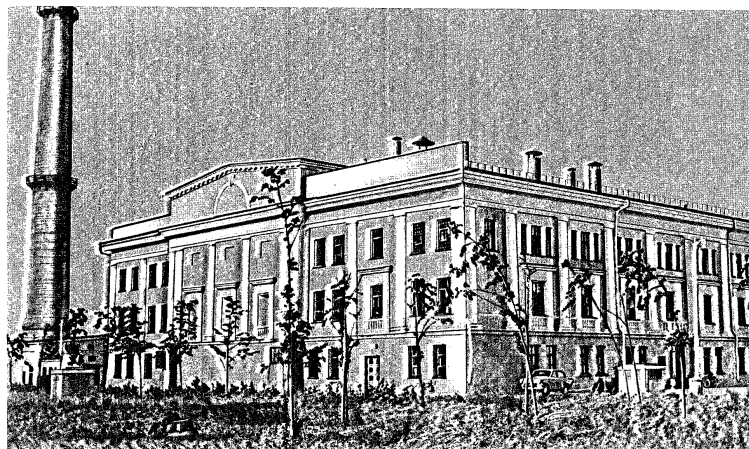
Models of these gigantic atomic machines are on show at the Exhibition.

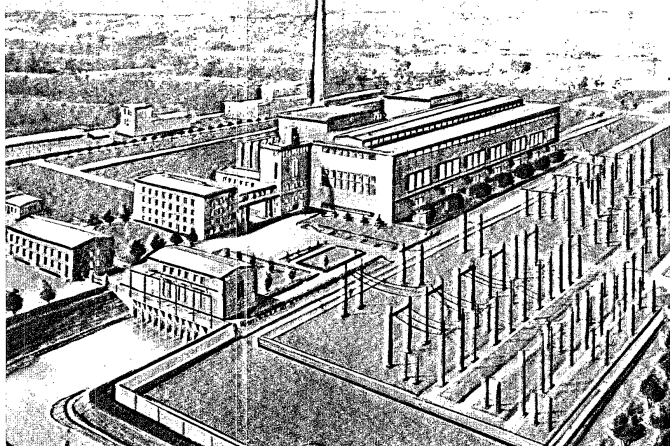
Exhibited in the Pavilion is also a scheme of a 50 Bev strong-focusing accelerator.

The powerful accelerators mentioned above have been handed over to the Joint Nuclear Research Institute, where scientists of many countries work on an equal footing with Soviet scientists. The doors of this Institute are wide open for all.

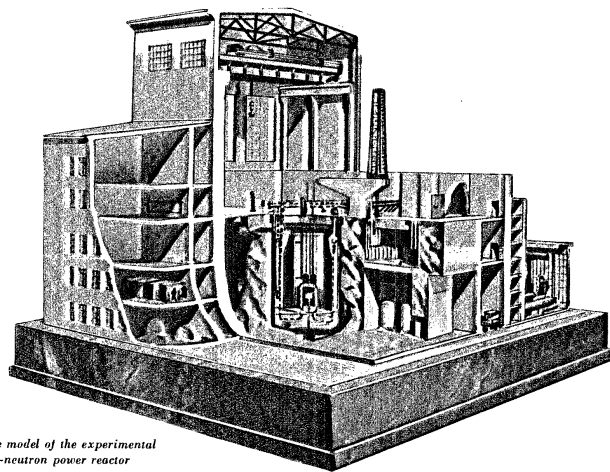
You will see here many exhibits reflecting the life of this Institute.

The building of the world's first commercial atomic power station in operation since June 27, 1954





The model of the experimental atomic power station with a capacity of 420,000 kilowatts now under construction



The model of the experimental fast-neutron power reactor

The model of a section (with a capacity of 210,000 kilowatts) of the experimental atomic power station with a total capacity of 420,000 kilowatts

When passing the core the water of the primary circuit will be heated to 275° C (at a pressure of 100 atmospheres in the system). The water will then be pumped through the pipes to special water-heating units, steam generators where the heat from the reactor is transferred to the secondary circuit for steam production. Each of the latter will provide 230 tons of steam per hour at a pressure of 32 atmospheres. The steam will be delivered from the secondary circuit to steam turbines which actuate the power generators. All working processes at the large atomic power station will be remotely controlled. Special care will be taken to safeguard the health of the operating personnel.

The cost of electric power supplied by the station will be the same as from a coal power station of equal capacity.

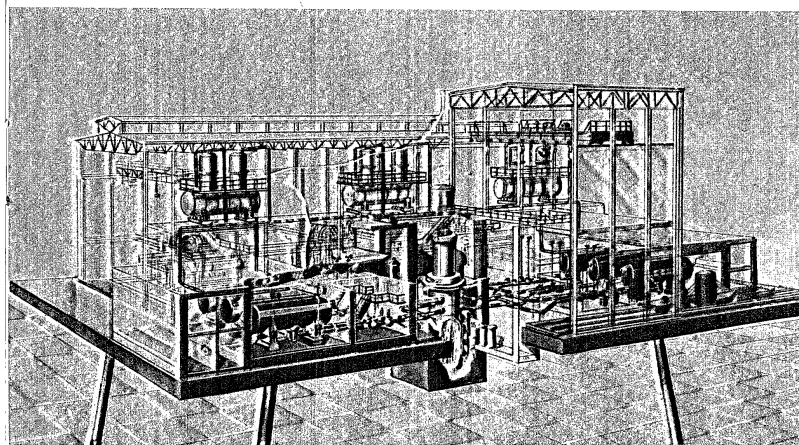
But will this be the best type of station economically? The answer can be obtained only by comparing several different reactor types. To do so, these should be built and sufficient operating experience accumulated.

Apart from the foregoing, the Soviet large-scale industrial experiment programme for atomic power stations includes the following power stations and reactors:

-An atomic power station with a heavy-water-moderated reactor using carbon dioxide as a coolant.

-An atomic power station with a graphite-moderated and water-cooled reactor which will in many respects resemble the world's first atomic power station operated since 1954.

-A graphite-moderated reactor using molten sodium as a coolant.



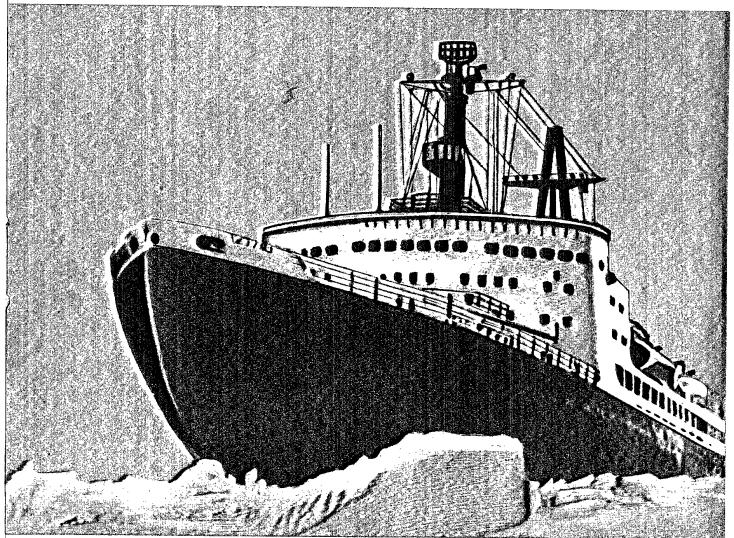
-A fast breeder reactor (with liquid sodium as a coolant) where fertile materials (uranium-238 and thorium), which cannot be called fuel in the nuclear sense of the word, will produce new fissionable materials (plutonium, uranium-233) in larger quantities than the original material used (uranium-235, plutonium, or uranium-233).

-A thermal thorium-breeder reactor with fuel suspension in heavy water, and, finally,

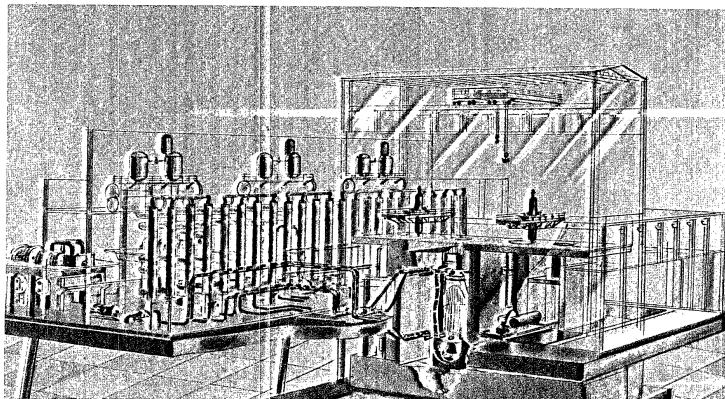
-A direct-cycle boiling water reactor which will be able to operate without an intermediate steam-generator. The steam from the reactor will be supplied directly to the turbogenerator.

Flow diagrams and models of these atomic power stations and reactors are shown at the Exhibition.

The atomic-powered ice-breaker "Lenin"



The model of the experimental atomic power station (graphite-moderated reactor)



ATOMIC ICE-BREAKER

The wide prospect of using nuclear power for water, land, and air transport are obvious.

An atomic-powered ice-breaker "Lenin" has been launched in the Soviet Union. The capacity of her main engines reaches 44,000 h.p., the displacement being 16,000 tons. A model of the ice-breaker is on view at the Exhibition.

The light weight of nuclear fuel with no need to replenish supplies for a long time will allow a considerably longer sailing at sea, without calling at ports. The spare room in the ice-

breaker due to the reduced fuel weight has been used for creating comfortable living and working conditions aboard for prolonged sailing in polar regions.

Atomic power plants are especially advantageous for longvoyage ships, whalers and vessels carrying bulk cargoes over long distances such as tankers.

The "Lenin" will pave the way for hundreds of new atomic-powered ships which are certain to appear on the seas and oceans in the near future.



RADIOACTIVE AND STABLE ISOTOPES

Use of new atomic power sources of tremendous energies is of great significance for the national economics as is shown in the respective sections of the Exhibition.

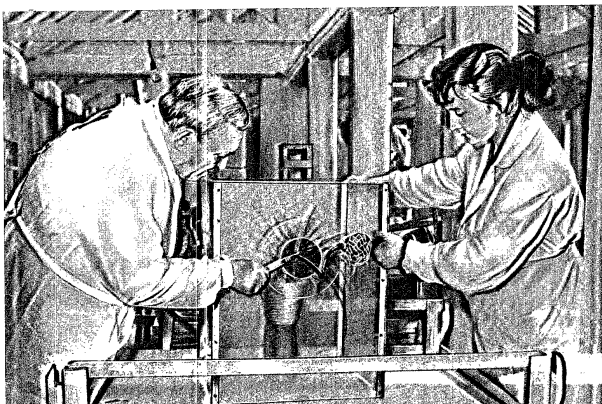
However, a part of no lesser importance is played by radioactive and stable isotopes obtained in nuclear reactors, accelerators, and special separation plants.

Isotopes are various forms of the same chemical element differing in their atomic weights and physical properties, but chemically identical. Almost every chemical element has a number of isotopes.

At present about 1000 artificial radioactive isotopes emitting alpha, beta and gamma rays, and about 300 stable (i.e. non-emitting) isotopes of all elements of Mendeleyev's periodic system are known. Production of artificial radioactive isotopes has now developed into a new industry,

Study of processes in a leaf by means of tracer atoms in natural conditions

Determination of the nature of carbon nutrition of plants by means of tracer atoms of carbon



Study of distribution of phosphorus in leaves of plants by the tracer atom method

while their application has become so wide and diverse that nowadays one can hardly name a branch where they are not or cannot be used.

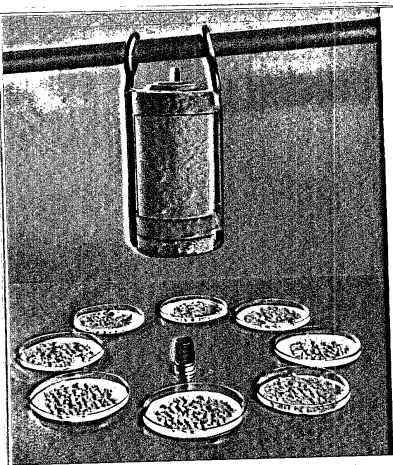
Thus, for instance, when used by doctors, radiations of radioactive isotopes are instrumental in diagnosing and curing diseases. When used by scientists, physiologists and biologists, they are helpful in penetrating into the deepest secrets of photosynthesis processes in plants and metabolism in animals. Biologists have succeeded in breeding new varieties of microorganisms producing large quantities of most valuable antibiotics (penicillin, biomyein, etc.). Conversely, other doses of the same radiations kill microorganisms when this is required for sterilization of tools, vessels, bandaging materials, food (detrichinization of pork, disinfection of grain, and sterilization of canned food), etc. Gamma radiation of radioactive co-

balt is used for killing silkworm cocoons, while irradiation of potatoes and other vegetables checks their germination in the case of prolonged storage.

When used by agronomists, radioactive isotopes make for expedient, highly effective use of fertilizers, and ensure an increased amount of green mass of fodder grass, etc.

The Exhibition displays some of the works, their results, and the equipment required.

Chromosomes of cells are known to convey hereditary properties to the future organism. Under ionizing radiation a number of chemical and structural changes take place in chromosomes, these changes often resulting in such changes of the embryo which either mutilate or kill it. These changes which are heritable are called mutations. But many newly arising mutations imply a number of useful properties.



Irradiation of plant seed

Organisms with such useful properties are selected by breeders and generate new forms of organisms.

Photosynthesis (carbon nutrition of plants) is one of the most complex biological processes. The tracer method has made and is making an inestimable contribution to its investigation. With the aid of this method many most important details of the successive stages of photosynthetic carbon transformations have been clarified.

It has been found that various products are formed in the course of photosynthesis, whose quantity varies with environment, plant type, and physiological condition of plants.

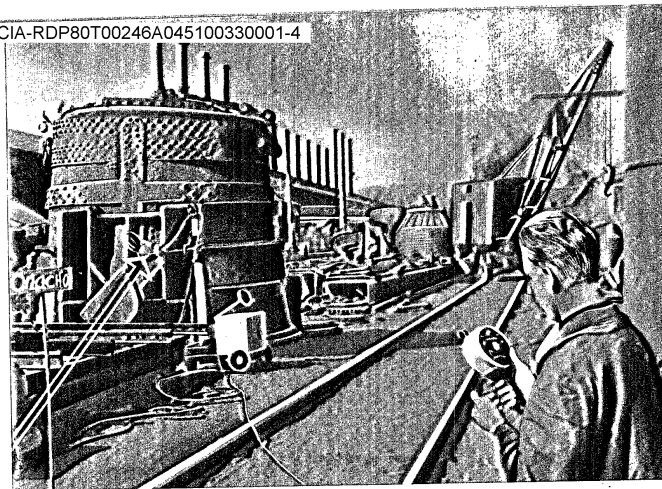
Thus, in the case of intensive nitrogen nutrition, with short-wave rays of the spectrum prevailing, the carbon of fast-growing plants assimilated during the course of photosynthesis, is increasingly converted to organic acids, amino acids, and proteins. Long-wave radiation makes for accumulation of carbons and reduces synthesis of proteins and organic acids.

It has been possible to show by means of carbon-14 that the movement of photosynthesis products is intensified in the case of nitrogen fertilization and watering, which creates better conditions for supply of assimilators to growing leaves and their accumulation in seed, fruit, and root crops.

By using P^{32} , a radioactive isotope of phosphorus, the most rational methods of introducing phosphoric fertilizers into various crops were devised, and practical recommendations given to agriculturists.

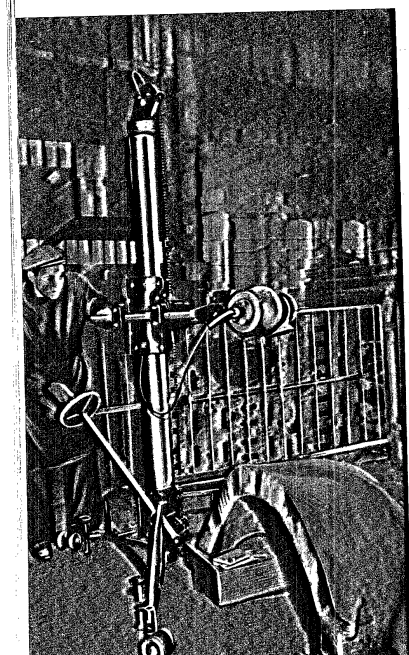
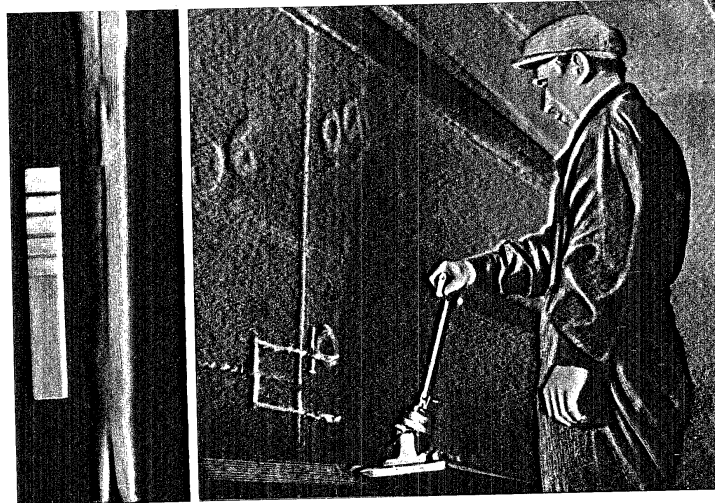
In the sphere of industrial production the application of radioactive isotopes is likewise exceptionally diverse. This diversity is so great

Gamma flaw detection in a steel casting in a shop

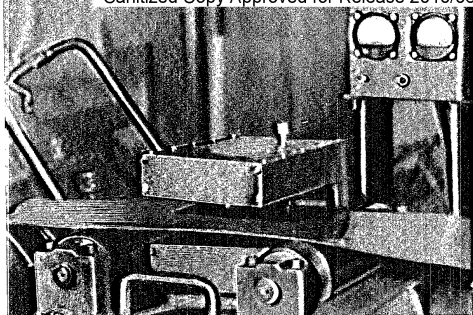


Gamma flaw detection in metallic articles in a works yard. The arrow indicates the radiation source

Gamma testing of the quality of welding. A gamma radiograph of a welding seam is shown on the left



Equipment for radioactive marking of steel

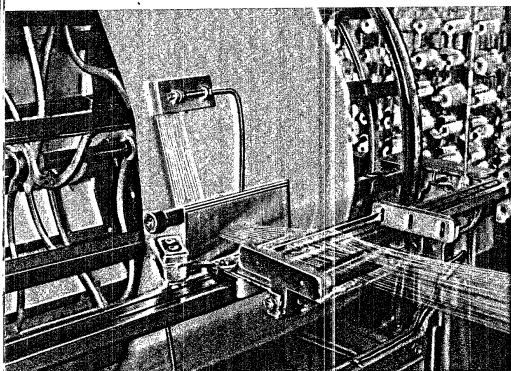


that it cannot possibly be fully reflected at an exhibition.

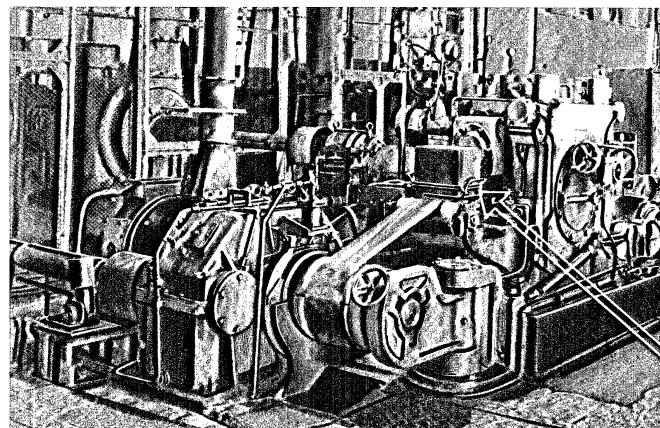
You can see, for instance, how with the aid of these isotopes the control of the pig-iron and steel production process can be made simpler and cheaper, how velocities of gas flows in blast

and welds permits detection of internal cracks, cavitations, and other defects. Wide application of gamma radiography has a great economic effect.

Use of radioactive isotopes makes marking of steel rolled stock easy and reliable.

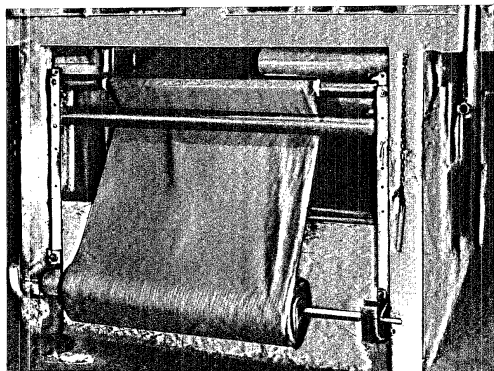


Radioactive isotopes enable easy and reliable removal of static electricity charges when making ribbons or yarn, which speeds up the process of production



Continuous testing of thickness of rolled steel band in industry

Contactless measurement of weight of cloth in industry



furnaces and hearth-bottom wear can be viewed, how changes can be traced in the structure of alloys depending on methods of casting, and distribution of alloy components and quality of welding checked, etc.

Isotopes enable the extent of wear of cutting tools or machine parts to be judged directly while in operation, without stopping or dismantling them.

Gamma radiography of metallic castings

All these possibilities have given rise to the production of a large number of various instruments now on view at the Exhibition.

Even this cursory review is sufficient to enable one to say, without running the risk of exaggeration, that in modern human society radioactive isotopes are becoming part of man's life, of production processes and of scientific research as rapidly and firmly as radio and television.

PRODUCTION OF RADIOACTIVE AND STABLE ISOTOPES

The ways of obtaining artificial radioactive isotopes are diverse. The most flexible and multi-purpose equipment for obtaining radioactive isotopes in greater quantities is a cyclotron. However, its low efficiency and complicated technological operations render it inferior to the nuclear reactor. To obtain radioactive isotopes, chemical compounds of various elements are enclosed in special cans and loaded into a reactor, where they are exposed to intensive bombardment with thermal or fast neutrons. This method is efficient and relatively cheap, its disadvantage being that it does not permit to produce all isotopes. The most common reaction is the capture of a neutron by a bombarded nucleus which is attended with emission of a gamma quantum, and results in the formation of a radioactive isotope of the same element. After decay it is transformed into a longer-lived isotope of another element, which can further be used. Bombardment of tellurium oxide with thermal neutrons may serve as an example. The isotope tellurium-131 thus produced which has a half-life of 24.8 minutes disintegrates into iodine-131 which has a half-life of 8 days.

Another reaction widely used for radioactive isotope production is where a nucleus captures a thermal neutron and emits a proton. Thus an isotope of another element is formed with an atomic number a unity less than the original one, as in the formation of carbon-14 from nitrogen.

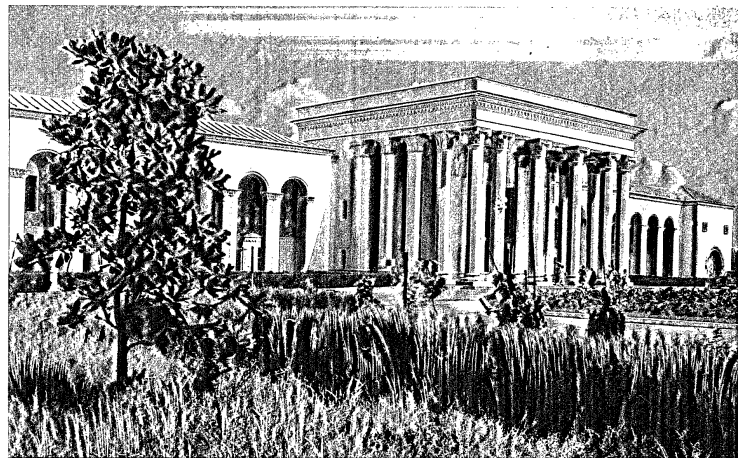
Reactions on fast neutrons are of the same type: they are used, for instance, for production of phosphorus-32 with great specific activity from sulphur.

Some isotopes are produced as a result of the reaction of a nucleus with a neutron when an alpha particle is emitted, as in the formation of phosphorus-32 from chlorine-35.

Of very great importance for radioisotope production is also use of fission products of the nuclei of some heavy elements. The most important isotopes thus obtained are cesium-137 and strontium-90.

No less important and promising is also production of unseparated fission products as powerful radiation sources. Irradiated chemicals are transformed into various compounds depending on their future use in special radio-chemical laboratories. In a number of cases it is necessary to use radiations of a radioisotope without introducing it into the composition of the material or organism to be tested. To do so, a wide variety of so-called radiation sources are produced in the Soviet Union, including sources of cobalt-60, iridium-192, thulium-170, cesium-137, strontium-90, etc. Geometrical dimensions and activities of the sources vary over a rather wide range. At present over 250 kinds of compounds among which there are many complex organic ones and over 100 radiation sources of various kinds are produced.

To obtain compounds enriched in stable isotopes, a number of physico-chemical separation methods are used: diffusion and thermodiffusion separation, ordinary and low-temperature rectification, electrolysis, electromagnetic separation, etc. The complex and perfect separation equipment available has enabled a wide variety of isotopes to be produced in considerable quantities.



Radon therapy clinic in Pyatigorsk

APPLICATION OF RADIOACTIVE ISOTOPES IN MEDICINE

Radioactive isotopes are used in medical institutions of the Soviet Union for treatment of various diseases. The isotopes which have found the widest application are cobalt-60, phosphorus-32, iodine-131, gold-198, sodium-24, strontium-90 and cesium-137.

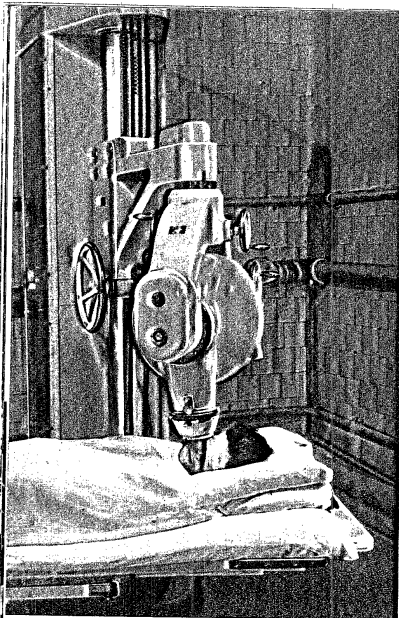
The basic principle of the Soviet health service, i.e. free medical aid for all, and centralized supply with radioactive isotopes of a large network of medical institutions, have enabled clinical observations of a large number of cases to be made in a comparatively short period.

Introduction of radioactive substances for treatment may be effected in several ways (for example, through the skin, the alimentary tract, respiration organs, blood) whose selection

depends on the nature of the disease and the kind of radiation employed.

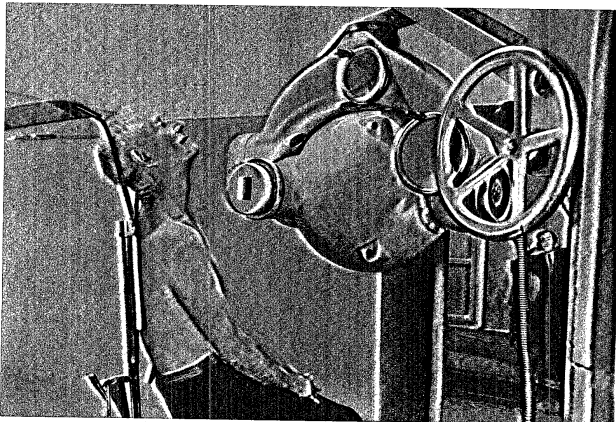
Radioactive cobalt has been used in the Soviet Union for treatment of a large number of cases of malignant tumours of mucous mouth cavity, tonsils, the upper jaw and the nasal cavity, larynx, gullet, lungs, mammary gland, and also of metastases into lymphatic nodes and bones, satisfactory immediate and long-term results being obtained.

For treatment of skin and lip cancer, and of other forms of neoplasms, radium, mesothorium and cobalt applicators are used, while for treatment of surface processes, both tumorous and inflammatory (capillary skin angiomas, eczemas, neurodermatites, etc.),



Treatment by means of radioactive cobalt (a type "FVT-Co-400" set)

Gamma irradiation of the thyroid gland by means of a type "FVT-Co-400" set



radioactive phosphorus (soluble sodium salt) is employed. For treatment of tumours use is also made of the intra-tissue method of radiumtherapy effected by means of radioactive preparations of cobalt-60, radium, and thorium, in platinum needle-shaped filters introduced to different depths into the tumour tissue.

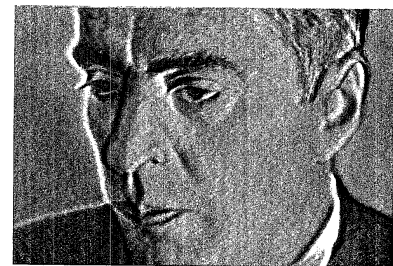
The intracavitary method of radiumtherapy is used for treatment of cancer of the internal organs. Of late internal irradiation methods have been used for treatment of various diseases, the radioactive substance being introduced into the body.

Apart from application for medical treatment, radioactive isotopes find wide use in diagnosing various diseases.

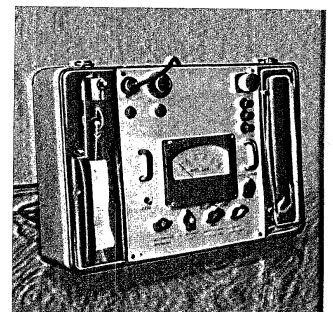
Radioactive iodine is used for determination of the functional state of the thyroid gland in the case of its diseases as well as for location of the brain tumours. Radioactive sodium is used for determination of blood flow rate and permeability of vessels, which also enables the state of the cardiovascular system to be tested. Concentration of various isotopes in tissues and organs is determined with the aid of special radiometers.



A case of cancer of the skin before treatment



Same after treatment



Clinical radiometer used for determination of blood circulation velocity by means of tracer atoms and determination of the outlines of radioactive substances deposited in the body

ATOMIC RAW MATERIALS

The principal atomic raw materials are uranium and thorium.

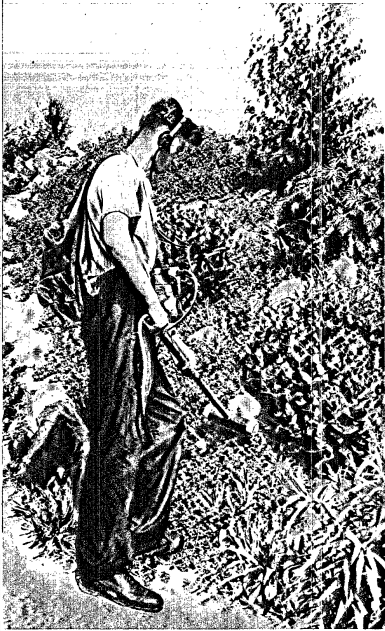
Natural uranium compounds vary in appearance and size of crystals. The Soviet Union has a large number of uranium and thorium deposits differing in ores and minerals. Some interesting samples of uranium minerals and commercial ores are represented at the Exhibition.

For prospecting, exploration, and mining of uranium deposits radiometric equipment is widely used, by means of which radiation of radioactive elements in ores and minerals is detected.

Also of interest are airborne instruments for complex aerial geophysical survey, permitting simultaneous automatic recording of gamma-radiation intensity in the troposphere, the magnetic field of the earth's surface, and the altitude of the aircraft, as well as instruments for gamma-survey from a moving vehicle.

Several types of instruments are made in the Soviet Union for land prospecting by gamma and beta-gamma radiation and radium emanation (radon): field gamma radiometers, field scintillation gamma radiometers, multi-purpose radiometers, different kinds of portable radiometers.

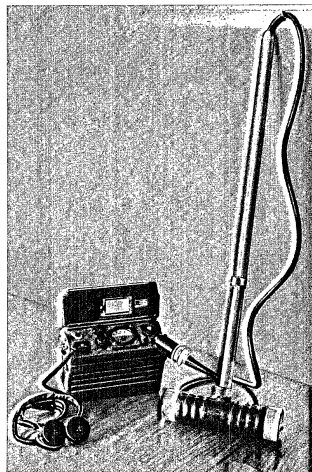
Some of the radiometers mentioned above are used for measurement of radioactivity in mine workings, for outlining ore bodies, and for evaluation of uranium content in ores.



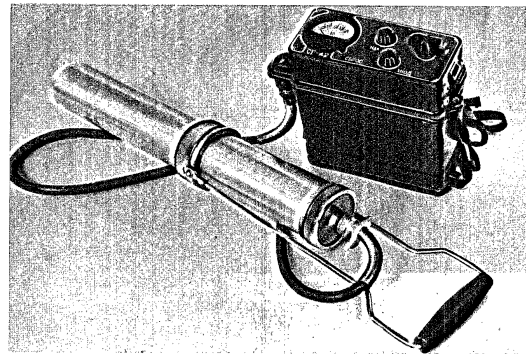
Work with a type "ПГР" radiometer in field



Work with a prospecting radiometer



A type "ПГР" field radiometer

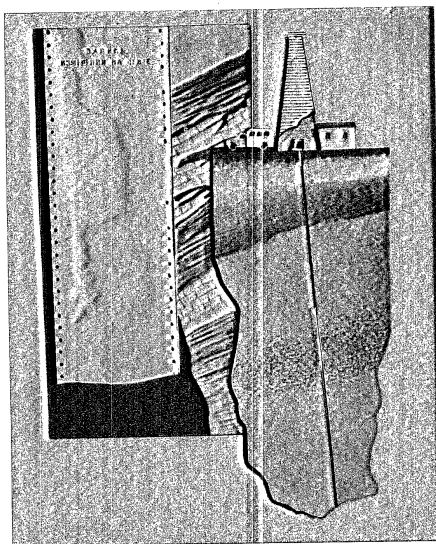


Prospecting radiometer with a scintillation counter

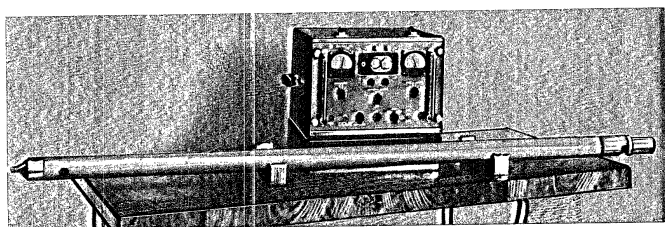
Radiometric equipment is also widely used for analysis of uranium ores in laboratory conditions. Represented at the Exhibition is similar analytical equipment for luminescence analysis based on the ability of some uranium minerals and compounds to emit light (lumi-

nescence) by giving off ultra-violet and blue rays, which is widely used for prospecting, exploration, and mining of uranium ores.

The method of radioactive logging has found wide application in geological investigations. This method is used not only for exploration of



Basic diagram of use of radioactive logging and type "PAPK" logging equipment.



Gamma survey of country from a helicopter

uranium deposits, but also for determination of oil and water-bearing strata, positioning of coal seams, and for solution of a number of other problems of well drilling.

For non-radioactive rocks neutron gamma logging is used (during well drilling the rock is irradiated with fast neutrons and a record is made of the gamma radiation excited as a result of capture of thermal neutrons by the rock, the characteristics of the radiation being dependent upon rock composition).

The Exhibition shows specimens of geophysical instruments for uranium ore exploration and radioactive logging instruments supplied by the Soviet Union to many countries of the world.

SAFETY

Wide application in our peaceful life of atomic energy in its various manifestations requires that special care be taken to protect the health of people dealing with radioactive materials and radiations.

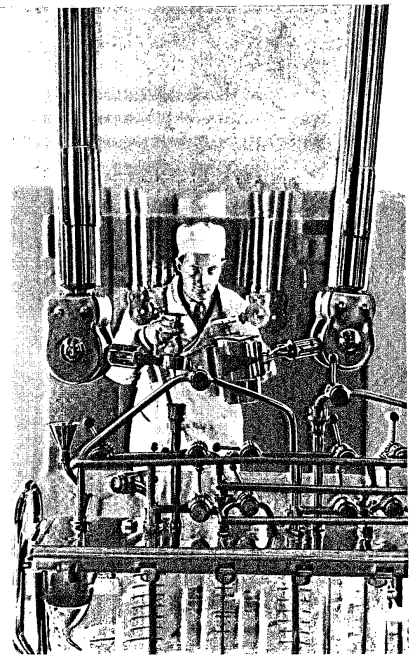
To prevent harmful effects of radioactive materials on personnel, it is highly important to avoid excessive radiations and radioactivity concentrations in the given medium. The great amount of experimental work carried out has made it possible to determine maximum permissible external radiation doses, air and water radioactivity standards and permissible contamination of surfaces, equipment, working clothes and hands.

In addition a series of protective measures have been developed for work with isotopes.

These measures ensure protection from external penetrating radiation, and prevent radioactive substances from getting inside the body and on the skin of personnel. They include special layouts and finishing of working space and use of special equipment and tools.

When handling highly radioactive materials, so-called "hot" cells are employed in which work is carried on with the aid of remote manipulators. The operator is separated from the active material by a thick shielding wall. Manipulators, these "extended arms" of the operator, have been brought to such a state of technical perfection that with their help the operator can easily perform the necessary work with various, at times very small-sized laboratory tools.

Great importance is attached to means of individual personnel protection (working clo-



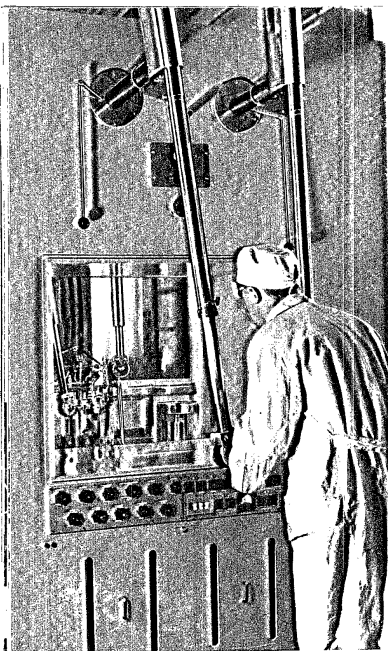
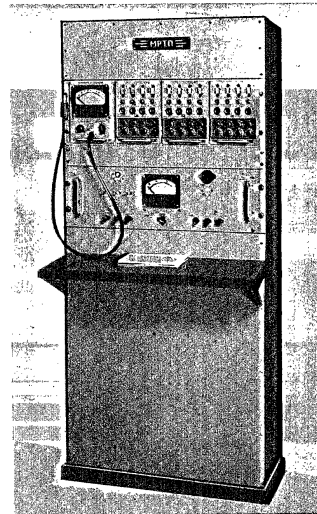
Operation by means of a remote-controlled manipulator in a "hot" cell

thes, special footwear, gloves, respirators, pneumatic suits, aprons, oversleeves, semi-robots, semi-overalls, etc.).

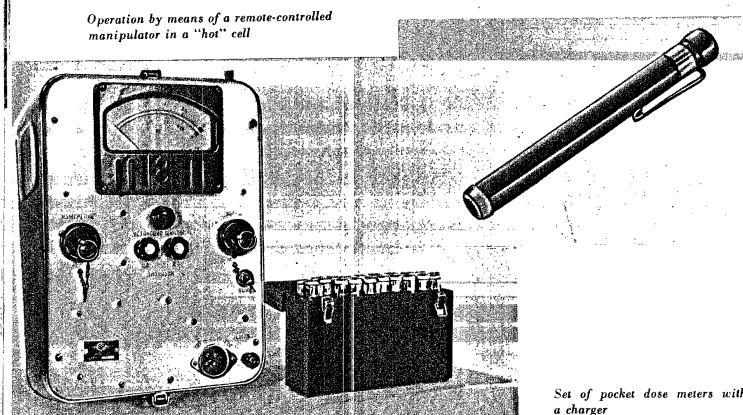
To protect the respiratory organs from radioactive dust and gases, use is made of valveless anti-dust respirators and pneumatic hoods, respectively.

When doing repair work under conditions of considerable radioactive contamination of the working rooms, special pneumatic suits are

A type "VCHT" remote-controlled signal measuring set for checking the level of radioactivity in premises



Operation by means of a remote-controlled manipulator in a "hot" cell



Set of pocket dose meters with a charger



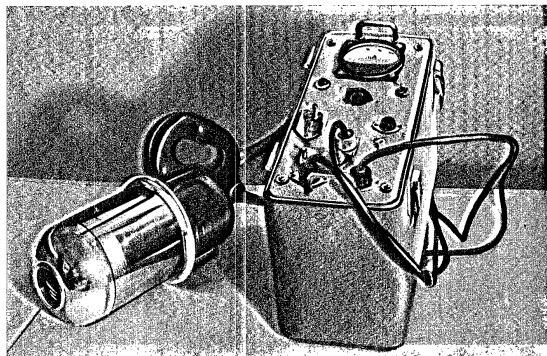
Safety clothing for work in premises with radioactive dust

Work in pneumatic safety clothes



Check of radioactive contamination of hands

A type "3D-2" electrical filter for sampling air with radioactive aerosols



worn, clean air being supplied under slight pressure.

The monitoring of the radiation level and immediate signalling in case the maximum permissible level is exceeded are effected by means of various stationary and mobile dosimetric apparatus provided with light and sound signalling devices.

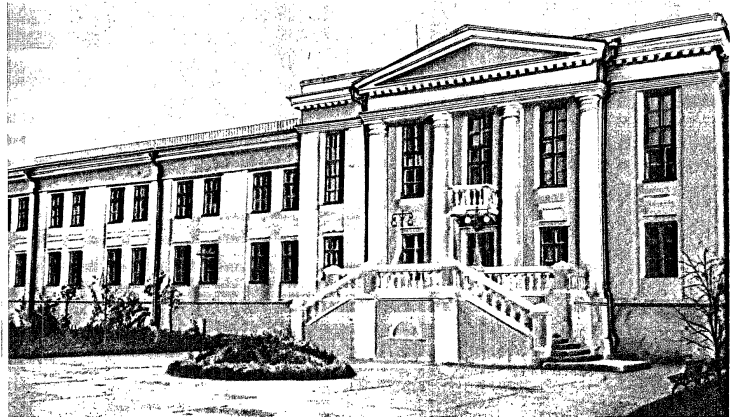
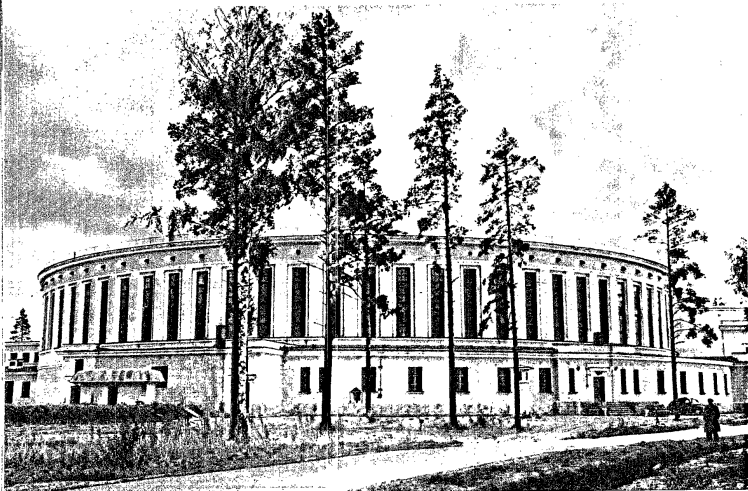
Protection control is implemented with the aid of portable microoentgen meters. Individual control of the dose received by the worker is now attained by using personal dosimeters which are made as condenser type double ionization chambers, and "fountain-pen" electrometers with individual counting.

**ACHIEVEMENTS OF SOVIET SCIENCE AND ENGINEERING
IN PEACEFUL USES OF ATOMIC ENERGY ARE AVAILABLE
FOR ALL**

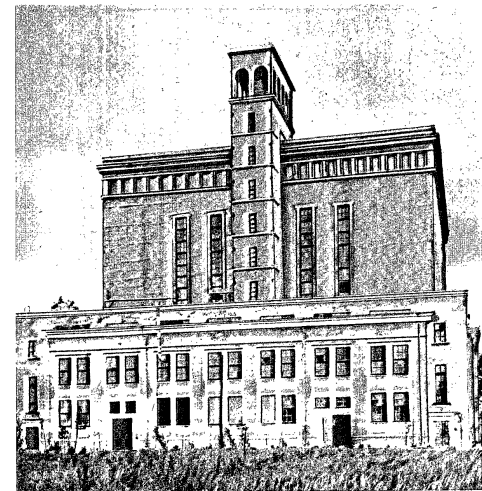
The Soviet Union persistently advocates the principle that the only reasonable uses of atomic energy are those pursuing peaceful objectives.

That is why the Soviet Union makes no secret of her work in this field, and, moreover, carries it out on the basis of international cooperation of scientists of various countries. She readily shares her experience and willingly and disinterestedly renders aid to countries which turn to her for assistance.

*The Joint Nuclear Research Institute (the town of Dubna, Moscow Region).
The building of the world's largest 10-Bev proton synchrotron*



The building of the synchrocyclotron control room



The building of the 680 Mev synchrocyclotron

The Soviet Union has supplied a number of states with nuclear reactors, particle accelerators, and various other equipment, necessary for research work in the field of nuclear physics.

Albania, Bulgaria, Hungary, the Democratic Republic of Viet-Nam, the German Democratic Republic, China, the Korean People's Democratic Republic, the Mongolian People's Republic, Poland, Rumania, the Soviet Union and Czechoslovakia, have united on an equal footing and set up a Joint Nuclear Research Institute in the town of Dubna, Moscow region, for cooperation in scientific research.

The Soviet Union has handed over to this Institute free of charge two world's largest accelerators, a 680 Mev synchrocyclotron and a 10 Bev proton synchrotron. There can be no doubt that such concerted team work where each can use his abilities and talent for the common welfare, will be of great benefit. The

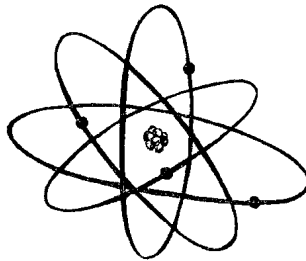
doors of this Institute are wide open for scientists of all countries.

The Soviet Union is a member of the International Atomic Energy Agency under the auspices of the UNO and actively participates in its work.

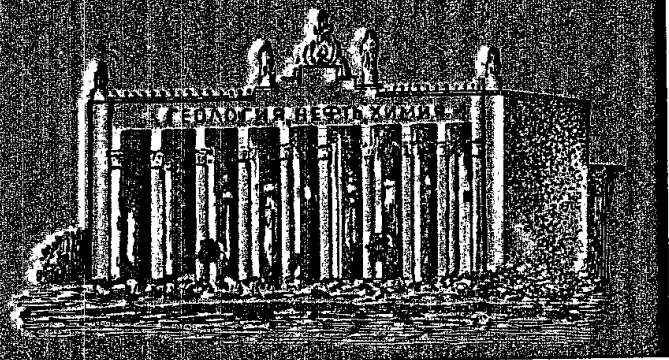
A large quantity of diverse preparations containing radioactive and stable isotopes are exported by the Soviet Union to other countries. Shipments of equipment and isotopes are made without any restrictions or special conditions.

The Soviet Union exchanges and publishes on an ever wider scale ample information in the sphere of peaceful uses of atomic energy.

Finally, the Soviet Union is carrying on intensive science popularization work, arranging exhibitions in many countries of the world which vividly display the possibilities and advantages of the introduction of nuclear engineering into various branches of economics, science, engineering, and medicine.



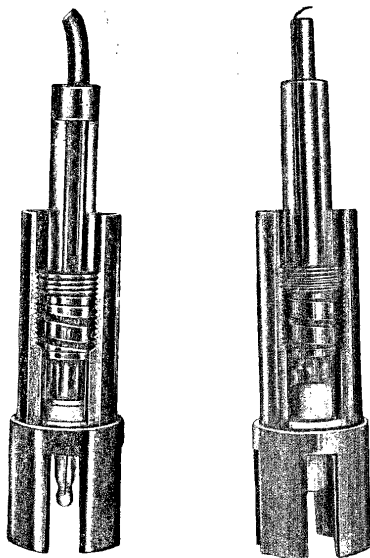
Внешторгиздат. Заказ № Б 510-45



РН-МЕТР

МИНИСТЕРСТВО
ХИМИЧЕСКОЙ ПРОМЫШЛЕННОСТИ СССР
МОСКВА

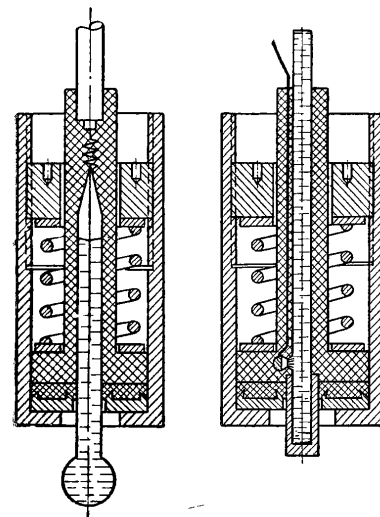
pH-МЕТР



Прибор pH-метр предназначен для измерения и регулирования концентрации водородных ионов при помощи толстостенных стеклянных электродов. Этим прибором можно измерять величины pH растворов, находящихся в заземленных металлических баках.

- В конструкции стеклянного электрода предусмотрено увеличение надежности изоляции места соединения электрода с высокочастотным кабелем, связывающим электрод со вторичным прибором.

- Отличительной особенностью вторичного прибора pH-метра является схема включения оболочки кабеля, идущего от стеклянного электрода к электронному усилителю. Применение этой схемы позволяет в 1000 раз снизить требования к изоляции кабеля и тем самым увеличить эксплуатационную надежность прибора.
- Усилитель pH-метра собран на шасси усилителя электронного потенциометра типа ЭПД.
- Питание прибора осуществляется от сети переменного тока напряжением 220 в (с колебанием от плюс 10 до минус 15%).
- Погрешность и область измерения pH определяется качеством и характеристикой используемых стеклянных электродов.



К печати 29 IV 1957 г. № 10496. ГОСХИМИЗДАТ № 1—26. 3 б. 3 и 1

Итография Госхимиздата, Москва, 88, Угличская

МОСКОВСКИЙ ОРДЕНА ЛЕНИНА
АВИАЦИОННЫЙ ИНСТИТУТ
ИМ. СЕРГО ОРДЖОНИКИДЗЕ

**ЭЛЕМЕНТЫ
ЭЛЕКТРОАВТОМАТИКИ
И РАСЧЕТА САМОЛЕТНЫХ
ЭЛЕКТРИЧЕСКИХ СИСТЕМ**

ОБОРОНГИЗ 1957

МИНИСТЕРСТВО ВЫСШЕГО ОБРАЗОВАНИЯ СССР
МОСКОВСКИЙ ОРДЕНА ЛЕНИНА
АВИАЦИОННЫЙ ИНСТИТУТ
имени СЕРГО ОРДЖОНИКИДЗЕ

ЭЛЕМЕНТЫ
ЭЛЕКТРОАВТОМАТИКИ
И РАСЧЕТА САМОЛЕТНЫХ
ЭЛЕКТРИЧЕСКИХ СИСТЕМ

Сборник статей под редакцией
докт. техн. наук профессора *Г. И. Атабекова*

ТРУДЫ ИНСТИТУТА

Выпуск 85

ГОСУДАРСТВЕННОЕ
ИЗДАТЕЛЬСТВО ОБОРОННОЙ ПРОМЫШЛЕННОСТИ
Москва, 1957

Сборник состоит из статей, посвященных методам расчета элементов электроавтоматики (дроссели насыщения, стабилизаторы, усилители, электромагниты, измерительные мосты), предназначенных для применения в наземных и самолетных условиях, а также некоторым вопросам расчета трехфазных электрических систем на самолете.

Зав. редакцией инж. *Е. В. Латынин*

ПРЕДИСЛОВИЕ

В настоящем сборнике содержатся статьи, посвященные теории и расчету некоторых элементов электроавтоматических устройств, а также вопросам, связанным с применением трехфазного тока на самолетах. Материалом указанных статей явились научно-исследовательские работы, выполненные в 1955 г. на кафедре «Теоретическая электротехника» МАИ.

В статье В. П. Щелкина дана теория дросселя насыщения с фазовым управлением, который может быть применен в различных фазочувствительных схемах. Н. Т. Кобяковой и В. А. Сартори приведен расчет ряда схем стабилизаторов напряжения переменного тока повышенной точности. В статьях А. К. Тер-Акопова показано существование оптимальной индуктивности, обеспечивающей минимальное время срабатывания электромагнитов; даны критический анализ существующих методов расчета динамических характеристик электромагнитов и рекомендации по практическому использованию их. Способ перехода от обыкновенного моста к двойному в измерительных схемах дан в статье С. Н. Иконникова. В статье В. Ф. Рахманова описана круговая диаграмма для определения коэффициентов усиления ламповых и полупроводниковых усилителей, работающих на линейном участке характеристики при любых видах нагрузки.

Ряд статей посвящен проблеме применения трехфазного тока на самолете, в том числе расчету несимметричных режимов при коротких замыканиях (статья В. Н. Истратова и В. В. Митюшиной), при обрывах фаз (статья А. В. Каменского и В. Г. Тер-Захаряна), а также некоторым вопросам релейной защиты самолетной электросети (статья А. Б. Тимофеева и В. Г. Тер-Захаряна) и применения электросистемы «два провода — корпус самолета» (статья А. А. Ковзана).

В процессе работы над рукописью учтены ценные замечания, сделанные доцентами Н. П. Удаловым, С. П. Колосовым и В. В. Андреевым, за что авторы сборника им весьма признательны.

Инж. В. П. ЩЕЛКИН

ДРОССЕЛЬ НАСЫЩЕНИЯ С ФАЗОВЫМ УПРАВЛЕНИЕМ

Дроссели насыщения получили в настоящее время широкое распространение в технике. Они являются составной частью магнитных усилителей, которые часто применяются в различного рода устройствах электроавтоматики.

У большинства магнитных усилителей входной величиной является постоянный ток, протекающий по управляющей обмотке усилителя, а выходной — переменный ток, протекающий по сопротивлению нагрузки.

Работа подобных усилителей и их расчет описаны в литературе.

У некоторых магнитных усилителей входной величиной является переменное напряжение, частота которого в несколько раз меньше частоты рабочего напряжения¹.

Созданы магнитные усилители, у которых в качестве входной величины используется выпрямленное переменное напряжение с частотой, равной частоте рабочего напряжения. Изменение тока нагрузки такого усилителя достигается изменением амплитуды импульсов, поступающих в управляющую обмотку.

В настоящей работе рассматривается дроссель насыщения, в управляющую обмотку которого подаются импульсы выпрямленного синусоидального напряжения с частотой, равной частоте рабочего напряжения. Изменение тока нагрузки, протекающего по рабочей обмотке дросселя и сопротивлению нагрузки, достигается сдвигом по фазе импульсов управляющего напряжения относительно рабочего напряжения.

Рассмотрим схему, изображенную на фиг. 1.

На этой схеме:

u — рабочее напряжение, $u = U_m \sin \omega t$;

u_y — управляющее напряжение, $u_y = U_{m_y} \sin(\omega t + \sigma)$, $0 \leq \sigma \leq 180^\circ$;

ω — число витков рабочей обмотки;

ω_y — число витков управляющей обмотки;

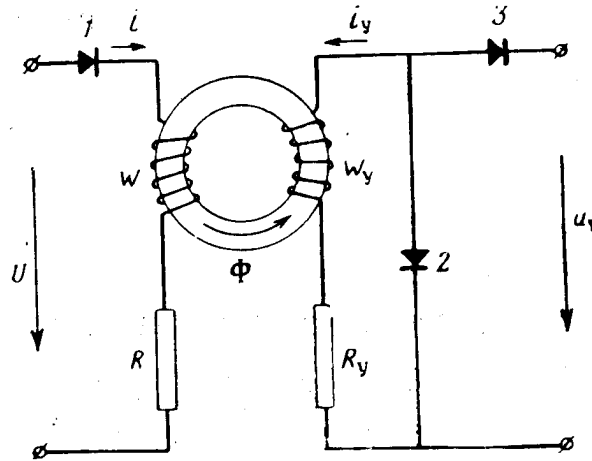
R — сопротивление нагрузки;

R_y — сопротивление управляющей цепи;

i — ток нагрузки;

i_y — управляющий ток.

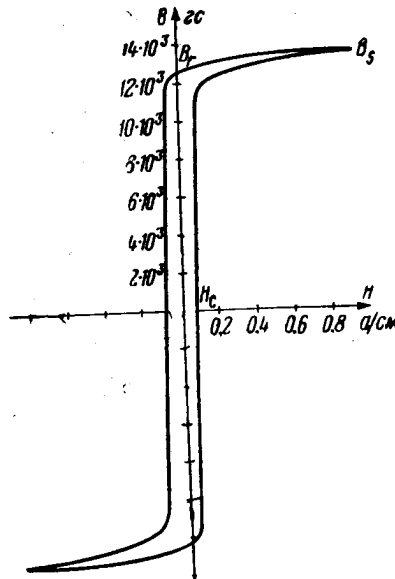
¹ М. А. Розенблат, Магнитные усилители, ГЭИ, 1955; F. K ü m m e l, Verstärkung kleiner Wechselstromspannungen mit dem magnetischen Verstärker, ETZ, 1954, Heft 11.



Фиг. 1. Схема дросселя насыщения с фазовым управлением.

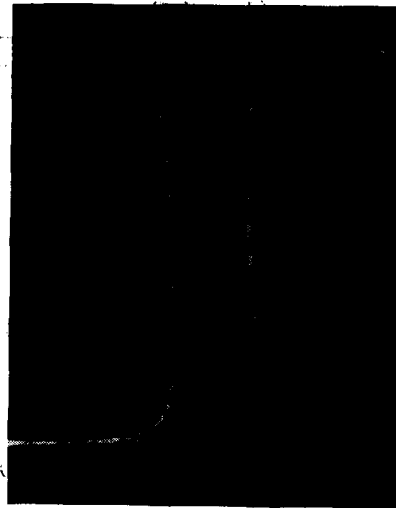
Сердечник дросселя изготовлен из магнитного материала, обладающего прямоугольной петлей гистерезиса. К таким материалам относятся сплавы 50НП, 65НП, Н34К29 (Мо-перминвар) и др.

Петля гистерезиса Мо-перминвара, снятая баллистическим методом, показана на фиг. 2.

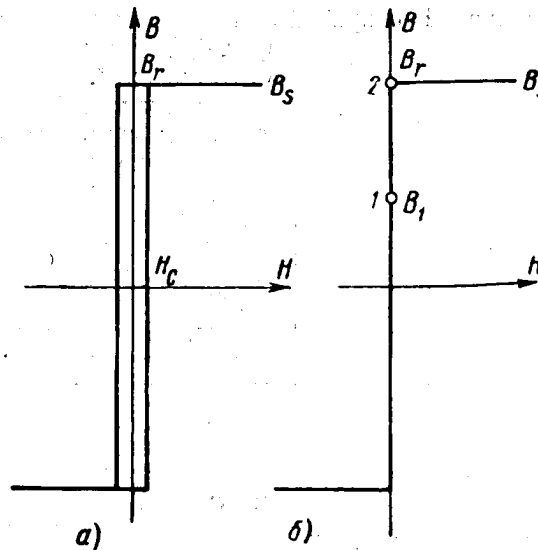


Фиг. 2. Петля гистерезиса сплава Н34К29 (Мо-перминвар), снятая баллистическим методом.

Петля гистерезиса сердечника, изготовленного из того же сплава, снятая на переменном токе, представлена на фиг. 3. Идеализован-



Фиг. 3. Динамическая петля гистерезиса сплава N34K29.



Фиг. 4. Идеализованные петли гистерезиса.

ная петля гистерезиса изображена на фиг. 4,а. При приближенных расчетах можно пренебречь величиной H_c и пользоваться кривой намагничивания сплава, изображенной на фиг. 4,б.

При анализе схемы, приведенной на фиг. 1, были приняты следующие допущения:

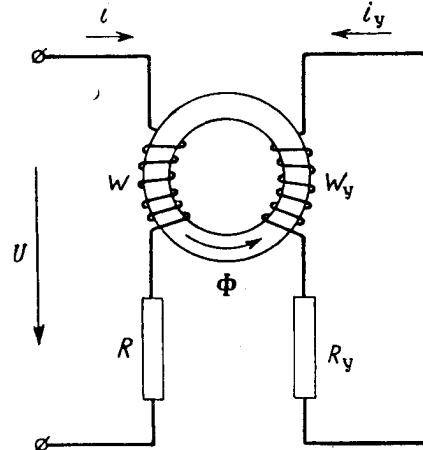
- 1) кривая намагничивания идеально прямоугольна, и коэрцитивная сила H_c равна нулю (фиг. 4,б);
- 2) вентили идеальны, т. е. их сопротивление в прямом направлении равно нулю, а в обратном направлении — бесконечно велико;
- 3) активные сопротивления обмоток малы и ими можно пренебречь;
- 4) сопротивления R и R_y — чисто активные;
- 5) потоки рассеяния обмоток отсутствуют.

Рассмотрим работу нелинейной схемы, изображенной на фиг. 1, в установившемся режиме. На различных интервалах времени этой схеме будут эквивалентны различные линейные схемы замещения. Границы интервалов времени определяются положением точки, соответствующей магнитному потоку в сердечнике, на кривой намагничивания и мгновенными значениями приложенных напряжений.

ПЕРВЫЙ ИНТЕРВАЛ

Границы интервала: $0 \leq \omega t \leq \varphi$, где φ — угол насыщения.

Предположим, что в начальный момент времени индукция в сердечнике равна B_1 и, следовательно, поток $\Phi_1 = B_1 s$, где s — площадь сечения сердечника. На кривой намагничивания, приведенной на фиг. 4,б, этому состоянию соответствует точка I .



Фиг. 5. Схема замещения для первого интервала.

Если к какой-либо из обмоток дросселя приложить переменное напряжение, то индукция в сердечнике будет изменяться по вертикальной части кривой намагничивания до тех пор, пока не наступит насыщение.

Обмотки намотаны таким образом, что положительным направлениям токов соответствует увеличение индукции в сердечнике.

На интервале $0 \leq \omega t \leq \varphi$ мгновенные значения напряжений u и u_y положительны. Поэтому вентиль I открыт, а вентиль 3 заперт. Вследствие этого на первом интервале схеме, показанной на фиг. 1, соответствует схема замещения, приведенная на фиг. 5.

Для этой схемы, пользуясь законами Кирхгофа и законом полного тока, можно написать систему уравнений, из которой определяются значения i , i_y и Φ :

$$\left. \begin{aligned} \omega \frac{d\Phi}{dt} + iR &= U_m \sin \omega t, \\ \omega_y \frac{d\Phi}{dt} + i_y R_y &= 0, \\ i\omega + i_y \omega_y &= 0. \end{aligned} \right\} \quad (1)$$

Начальное условие — при $t=0$ $\Phi = \Phi_1$.
Решая систему (1), получим

$$\Phi = \frac{U_m A}{\omega} (1 - \cos \omega t) + \Phi_1, \quad (2)$$

где

$$A = \frac{\omega R_y}{\omega^2 R_y + \omega_y^2 R};$$

$$i = \frac{U_m}{R + \frac{\omega^2}{\omega_y^2} R_y} \sin \omega t, \quad (3)$$

$$i_y = - \frac{U_m \frac{\omega}{\omega_y}}{R + \frac{\omega^2}{\omega_y^2} R_y} \sin \omega t. \quad (4)$$

Таким образом, на протяжении 1-го интервала схема работает как трансформатор напряжения. Амплитуда рабочего тока невелика, так как в практических схемах величина $\frac{\omega^2}{\omega_y^2} R_y$ выбирается достаточно большой.

При $\omega t = \varphi$ поток в сердечнике становится равным $\Phi_s = B_s s$, и сердечник, состояние которого в этот момент характеризуется точкой 2 на кривой намагничивания, насыщается. Угол насыщения φ определяется из равенства

$$\Phi_s = \frac{U_m A}{\omega} (1 - \cos \varphi) + \Phi_1,$$

откуда

$$\cos \varphi = 1 - \frac{(\Phi_s - \Phi_1) \omega}{U_m A}.$$

Таким образом, угол φ зависит от начальной величины потока Φ_1 .

ВТОРОЙ ИНТЕРВАЛ

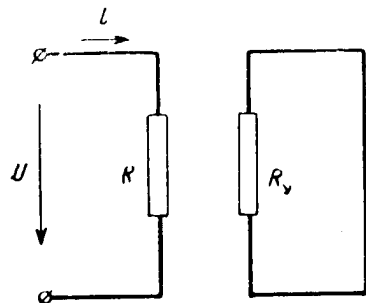
Границы интервала: $\varphi \leq \omega t \leq \pi - \sigma$.

Если сердечник обладает идеальной прямоугольной петлей гистерезиса, то теоретически при $\omega t = \varphi$ индуктивности управляющей и рабочей обмоток дросселя скачком меняются от бесконечно большой величины до нуля. Практически также можно считать, что когда сер-

дечник насыщен, индуктивные сопротивления обмоток дросселя пренебрежимо малы по сравнению с сопротивлениями R и R_y .

Учитывая сказанное выше, а также то, что в течение второго интервала вентиль 3 заперт, а вентиль 1 открыт, схему замещения можно представить так, как показано на фиг. 6.

Вследствие того, что в рассматриваемом интервале магнитной связи между обмотками нет, ток в управляющей обмотке равен нулю. Ток в рабочей обмотке определяется выражением



Фиг. 6. Схема замещения для второго интервала.

$$i = \frac{U_m}{R} \sin \omega t.$$

Точка, характеризующая состояние сердечника, лежит на верхнем горизонтальном участке кривой намагничивания.

Таким образом, во втором интервале.

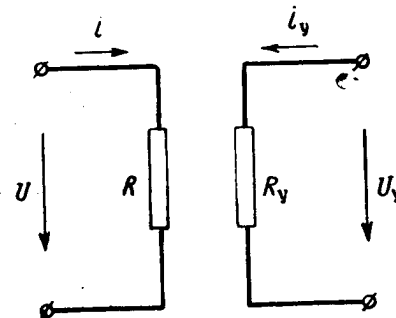
$$\begin{aligned} \Phi &= \Phi_s = \text{const}, \\ i &= \frac{U_m}{R} \sin \omega t, \\ i_y &= 0. \end{aligned} \quad (5)$$

Вследствие того, что в практических схемах $R \ll \frac{\omega^2}{\omega_y^2} R_y$, амплитуда рабочего тока во втором интервале значительно больше амплитуды рабочего тока в первом интервале.

ТРЕТИЙ ИНТЕРВАЛ

Границы интервала: $\pi - \sigma \leq \omega t \leq \theta$.

При $\omega t = \pi - \sigma$ управляющее напряжение U_y , уменьшаясь, проходит через нуль. В этот момент вентиль 3 отпирается, а вентиль 2 запирается. По управляющей обмотке начинает протекать ток. Схема замещения принимает вид, представленный на фиг. 7. Так как магнитной связи между обмотками и в этом интервале нет, токи определяются следующими выражениями:



Фиг. 7. Схема замещения для третьего интервала.

$$i = \frac{U_m}{R} \sin \omega t, \quad (6)$$

$$i_y = \frac{U_m y}{R_y} \sin (\omega t + \sigma). \quad (7)$$

Поток в сердечнике по-прежнему остается постоянным, равным Φ_s , так как точка, характеризующая состояние сердечника, лежит на верхнем горизонтальном участке кривой намагничивания.

Токи i и i_y , проходя по обмоткам, создают противоположно направленные ампер-витки намагничивания. Когда сумма этих ампер-витков станет равной нулю, состояние сердечника будет характеризоваться точкой 2 (см. фиг. 4,б). В этот момент времени $\omega t = \theta$ индуктивности обмоток скачком изменятся от нуля до бесконечно большой величины.

Угол θ определяется из условия

$$\sum a\omega = 0,$$

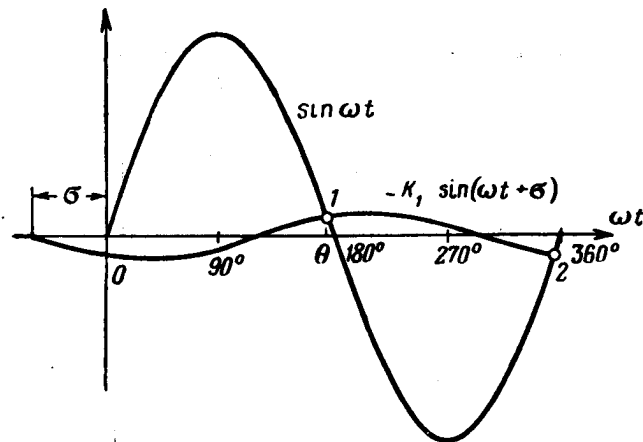
или

$$\frac{U_m \omega}{R} \sin \theta + \frac{U_m y \omega_y}{R_y} \sin (\theta + \sigma) = 0.$$

Обозначим

$$\frac{U_m y \omega}{U_m \omega_y} = K, \quad (8)$$

$$\frac{\omega_y^2 R}{\omega^2 R_y} = K_1. \quad (9)$$



Фиг. 8. График для определения угла θ .

Тогда предыдущее уравнение можно переписать так:

$$\sin \theta + K K_1 \sin (\theta + \sigma) = 0.$$

Обычно коэффициент K по величине несколько превышает единицу, а коэффициент K_1 значительно меньше единицы. Поэтому $K K_1 \ll 1$.

Угол θ можно определить как абсциссу точки пересечения двух синусоид: $\sin \omega t$ и $-K K_1 \sin(\omega t + \sigma)$ (фиг. 8). Таких точек пересечения две. Так как рассматривается положительный полупериод рабочего напряжения $U_m \sin \omega t$, то искомый угол определяется как абсцисса точки пересечения 1.

Решая уравнение, получим

$$\cos \theta = \frac{1 + KK_1 \cos \sigma}{\pm \sqrt{1 + 2KK_1 \cos \sigma + K^2 K_1^2}}.$$

Нетрудно заметить, что искомому углу θ соответствует знак «минус» перед радикалом, стоящим в знаменателе. Поэтому удобно записать это выражение так:

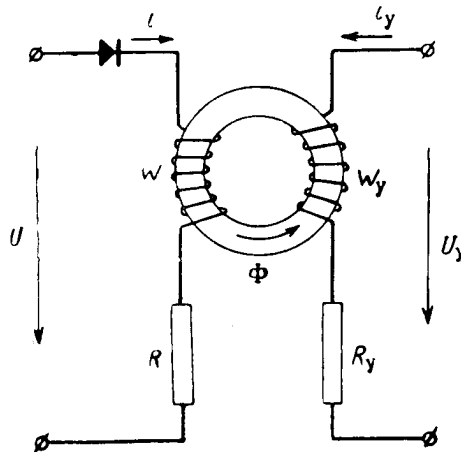
$$\cos \theta = - \frac{1 + KK_1 \cos \sigma}{|\sqrt{1 + 2KK_1 \cos \sigma + K^2 K_1^2}|}. \quad (10)$$

В приближенных расчетах можно полагать коэффициент K_1 равным нулю. Тогда $\cos \theta = -1$ и $\theta = \pi$.

ЧЕТВЕРТЫЙ ИНТЕРВАЛ

Границы: $\theta \leq \omega t \leq 2\pi - \sigma$.

Схема замещения для этого случая представлена на фиг. 9. Эта схема нелинейна: в ее рабочей цепи имеется вентиль. Вентиль остав-



Фиг. 9. Схема замещения для четвертого интервала.

лен в схеме потому, что в рассматриваемом интервале нельзя заранее предугадать направление тока в рабочей цепи, а следовательно, и режим работы вентиля.

Нелинейной схеме, представленной на фиг. 9, соответствуют две системы уравнений:

$$\left. \begin{aligned} \omega \frac{d\Phi}{dt} + iR &= U_m \sin \omega t, \\ \omega_y \frac{d\Phi}{dt} + i_y R_y &= U_{m_y} \sin (\omega t + \sigma), \\ i\omega + i_y \omega_y &= 0, \end{aligned} \right\} \quad (11)$$

$$\left. \begin{aligned} i &= 0; \\ \omega_y \frac{d\Phi}{dt} + i_y R_y &= U_{m,y} \sin(\omega t + \sigma), \\ i\omega + i_y \omega_y &= 0. \end{aligned} \right\} \quad (12)$$

Система (11) справедлива в том случае, если вентиль открыт, т. е. $i > 0$. Система (12) справедлива, если вентиль в рабочей цепи заперт.

Начальное условие — при $\omega t = \theta$ $\Phi = \Phi_s$.

Решая систему (11), получим

$$\begin{aligned} \Phi &= \Phi_s + \frac{U_{m,A}}{\omega} [\cos \theta - \cos \omega t] + \\ &+ \frac{U_{m,y} B}{\omega} [\cos(\theta + \sigma) - \cos(\omega t + \sigma)], \end{aligned} \quad (13)$$

где

$$A = \frac{\omega R_y}{\omega^2 R_y + \omega_y^2 R}, \quad B = \frac{\omega_y R}{\omega^2 R_y + \omega_y^2 R},$$

$$i = \frac{U_m}{R + \frac{\omega^2}{\omega_y^2} R_y} \sin \omega t - \frac{U_{m,y} \frac{\omega}{\omega_y}}{R + \frac{\omega^2}{\omega_y^2} R_y} \sin(\omega t + \sigma); \quad (14)$$

$$i_y = -\frac{U_m \frac{\omega_y}{\omega}}{R \frac{\omega_y^2}{\omega^2} + R_y} \sin \omega t + \frac{U_{m,y}}{R \frac{\omega_y^2}{\omega^2} + R_y} \sin(\omega t + \sigma). \quad (15)$$

Решения системы (11) описывают процессы, происходящие в схеме с открытым вентиляем 1. Вентиль открыт, если

$$i = \frac{U_m}{R + \frac{\omega^2}{\omega_y^2} R_y} \sin \omega t - \frac{U_{m,y} \frac{\omega}{\omega_y}}{R + \frac{\omega^2}{\omega_y^2} R_y} \sin(\omega t + \sigma) > 0.$$

Это неравенство выполняется на данном интервале при $\theta < \omega t < \gamma$, где γ определяется из соотношения

$$i(\gamma) = \frac{U_m}{R + \frac{\omega^2}{\omega_y^2} R_y} \sin \gamma - \frac{U_{m,y} \frac{\omega}{\omega_y}}{R + \frac{\omega^2}{\omega_y^2} R_y} \sin(\gamma + \sigma) = 0,$$

или

$$U_m \sin \gamma - U_{m,y} \frac{\omega}{\omega_y} \sin(\gamma + \sigma) = 0.$$

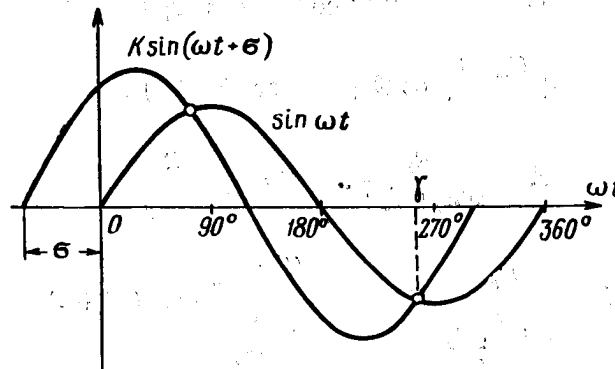
Определение угла γ сводится к нахождению точки пересечения двух синусоид, имеющих различные амплитуды и сдвинутых по фазе (Фиг. 10). Таких точек пересечения две. Так как рассматривается интервал $0 \leq \omega t \leq 2\pi - \sigma$, то искомая точка γ лежит во втором полупериоде. Угол γ можно определить из приведенного выше уравнения

$$\sin \gamma - K \sin(\gamma + \sigma) = 0,$$

$$\sin \gamma - K \sin \gamma \cos \sigma - K \sin \sigma \cos \gamma = 0,$$

откуда

$$\operatorname{tg} \gamma = \frac{K \sin \sigma}{1 - K \cos \sigma}.$$



Фиг. 10. График для определения угла γ .

Из этой формулы получим

$$\cos \gamma = \frac{1}{\pm \sqrt{1 + \operatorname{tg}^2 \gamma}} = \frac{1 - K \cos \sigma}{\pm \sqrt{1 - 2K \cos \sigma + K^2}}.$$

Двум знакам, стоящим перед радикалом, соответствуют две точки пересечения синусоид. Легко показать, что для определения точки γ , лежащей во втором полупериоде, необходимо брать отрицательное значение радикала. В дальнейшем удобно поменять знак числителя на обратный, а в знаменателе брать абсолютную величину радикала.

Тогда формулу можно записать так:

$$\cos \gamma = \frac{K \cos \sigma - 1}{|\sqrt{1 - 2K \cos \sigma + K^2}|}. \quad (16)$$

При $\omega t = \gamma$ ток i становится равным нулю и клапан в рабочей цепи запирается. Для последующего отрезка времени справедлива система (12):

$$\begin{aligned} i &= 0, \\ \omega_y \frac{d\Phi}{dt} + i_y R_y &= U_{m_y} \sin(\omega t + \sigma), \\ i\omega + i_y \omega_y &= 0. \end{aligned}$$

Начальное условие — при $\omega t = \gamma$

$$\Phi = \Phi_{\gamma} = \Phi_s + \frac{U_m A}{\omega} [\cos \theta - \cos \gamma] + \frac{U_m y B}{\omega} [\cos (\theta + \sigma) - \cos (\gamma + \sigma)].$$

Решениями данной системы являются

$$\left. \begin{aligned} \Phi &= \Phi_{\gamma} + \frac{U_m y}{\omega y} [\cos (\gamma + \sigma) - \cos (\omega t + \sigma)], \\ i &= 0, \\ i_y &= 0, \end{aligned} \right\} \quad (17)$$

Таким образом, в течение всего четвертого интервала поток уменьшается и точка, характеризующая состояние сердечника, перемещается вниз по вертикальной части кривой намагничивания (см. фиг. 4,б).

Если схема работает в установившемся режиме, то при $\omega t = 2\pi - \sigma$ $\Phi = \Phi_1$, т. е. поток принимает свое начальное значение.

ПЯТЫЙ ИНТЕРВАЛ

Границы: $2\pi - \sigma \leq \omega t < 2\pi$.

При $\omega t = 2\pi - \sigma$ управляющее напряжение $u_y = U_m y \sin (\omega t + \sigma)$ проходит через нуль и в дальнейшем становится положительным. Вентиль 3 (см. фиг. 1) в рабочей цепи запирается.

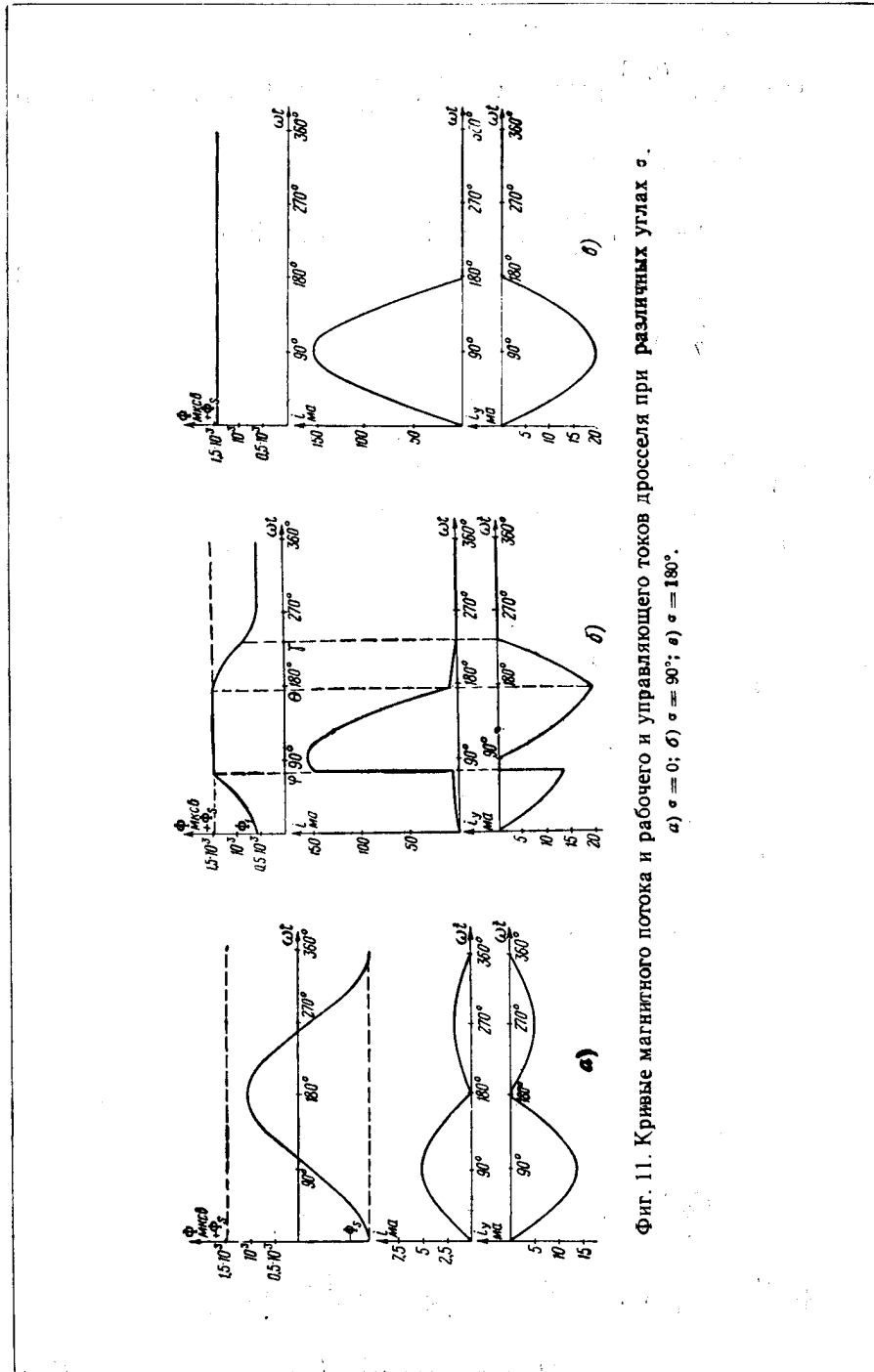
На протяжении всего интервала рабочее и управляющее напряжения приложены к запертым вентилям. Поэтому токи i и i_y равны нулю, а поток в сердечнике остается постоянным, равным Φ_1 .

При $\omega t > 2\pi$ цикл, описанный выше, повторяется.

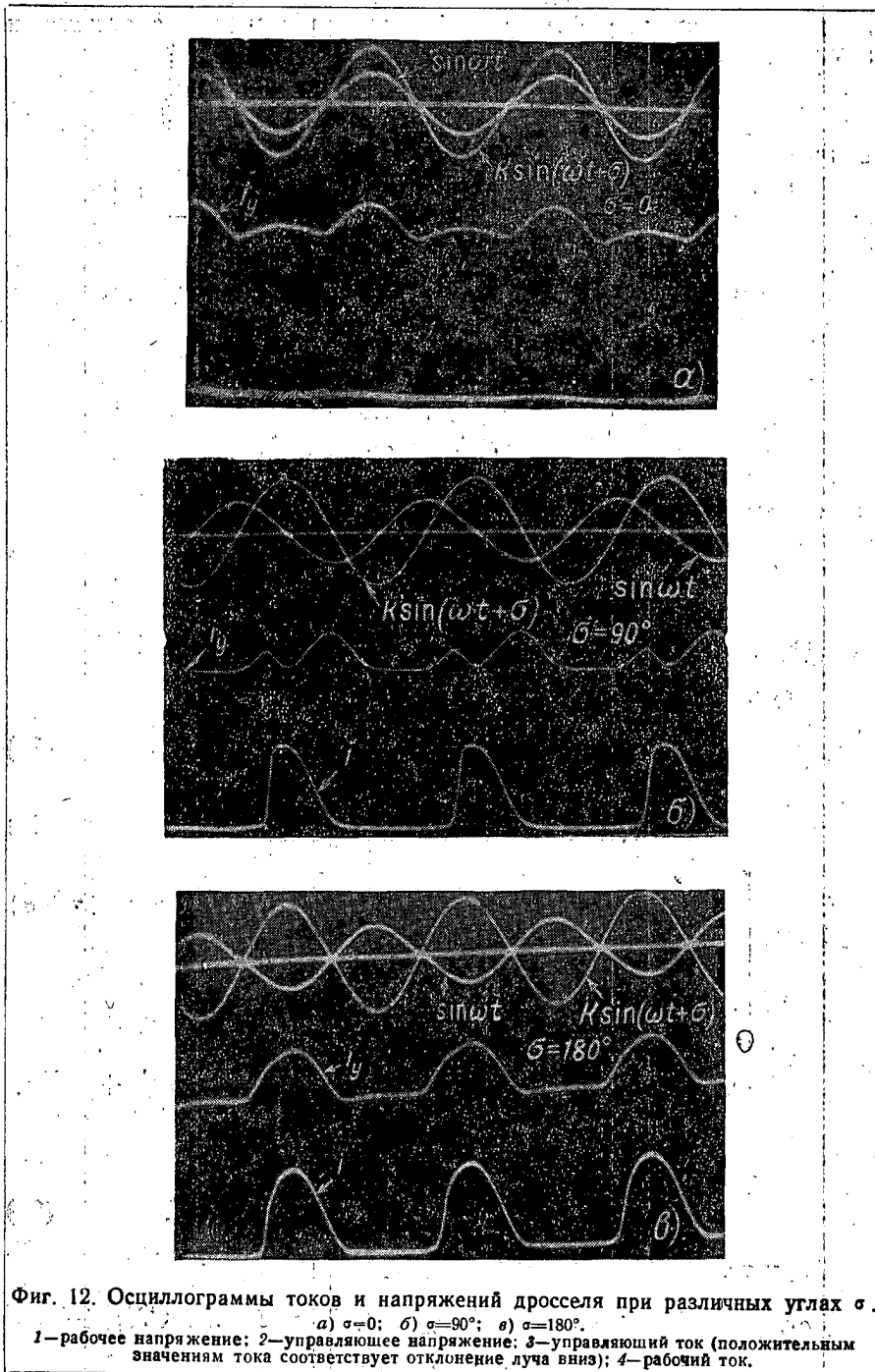
Пользуясь полученными выше выражениями, можно построить кривые изменения Φ , i , i_y в функции времени. Такие кривые, построенные для $\sigma = 0$, $\sigma = 90^\circ$, $\sigma = 180^\circ$ представлены на фиг. 11,а; б; в. На фиг. 12 показаны осциллограммы токов i и i_y , снятые на экспериментальной установке. Небольшие расхождения между теоретическими и экспериментальными кривыми объясняются тем, что кривые намагничивания сердечников не были идеально прямоугольными.

Выясним теперь зависимость между углом насыщения φ и фазой управляющего напряжения σ . Для этого рассмотрим изменение потока в сердечнике за период рабочего напряжения. В установившемся режиме это изменение в сумме должно равняться нулю:

$$\begin{aligned} & \frac{U_m A}{\omega} (1 - \cos \varphi) + \frac{U_m A}{\omega} [\cos \theta - \cos \gamma] + \\ & + \frac{U_m y B}{\omega} [\cos (\theta + \sigma) - \cos (\gamma + \sigma)] + \frac{U_m y}{\omega y} [\cos (\gamma + \sigma) - \\ & - \cos (2\pi - \sigma + \sigma)] = 0. \end{aligned} \quad (18)$$



Фиг. 11. Кривые магнитного потока и рабочего и управляющего токов дросселя при различных углах σ .
а) $\sigma = 0^\circ$; б) $\sigma = 90^\circ$; в) $\sigma = 180^\circ$.



Разделив выражение (18) на коэффициент при первом слагаемом и заметив, что

$$\frac{U_{myB}}{U_{mA}} = KK_1, \quad \frac{U_{my}}{U_{m\omega yA}} = K(K_1 + 1),$$

получим

$$1 - \cos \varphi + \cos \theta - \cos \gamma + KK_1 [\cos(\theta + \sigma) - \cos(\gamma + \sigma)] + \\ + K(K_1 + 1) [\cos(\gamma + \sigma) - 1] = 0.$$

Ранее мы получили зависимости между углами θ , γ и σ ,

$$\cos \theta = -\frac{1 + KK_1 \cos \sigma}{\sqrt{1 + 2KK_1 \cos \sigma + K^2 K_1^2}}, \\ \cos \gamma = \frac{K \cos \sigma - 1}{\sqrt{1 - 2K \cos \sigma + K^2}}.$$

Используя эти зависимости, а также уравнения

$$\sin \theta + KK_1 \sin(\theta + \sigma) = 0,$$

$$\sin \gamma - K \sin(\gamma + \sigma) = 0,$$

выразим углы $\theta + \sigma$ и $\gamma + \sigma$ через фазовый сдвиг σ .

Получим

$$\cos(\theta + \sigma) = -\frac{\cos \sigma + KK_1}{\sqrt{1 + 2KK_1 \cos \sigma + K^2 K_1^2}}, \quad (19)$$

$$\cos(\gamma + \sigma) = \frac{K - \cos \sigma}{\sqrt{1 - 2K \cos \sigma + K^2}}. \quad (20)$$

Подставляя выражения для $\cos \theta$, $\cos \gamma$, $\cos(\theta + \sigma)$ и $\cos(\gamma + \sigma)$ в уравнение (18), получим

$$\cos \varphi = -\frac{1 + KK_1 \cos \sigma}{\sqrt{1 + 2KK_1 \cos \sigma + K^2 K_1^2}} - \frac{K \cos \sigma - 1}{\sqrt{1 - 2K \cos \sigma + K^2}} - \\ - KM \frac{\cos \sigma + KK_1}{\sqrt{1 + 2KK_1 \cos \sigma + K^2 K_1^2}} + K \frac{K - \cos \sigma}{\sqrt{1 - 2K \cos \sigma + K^2}} + \\ + 1 - KK_1 - K = \sqrt{1 - 2K \cos \sigma + K^2} - \\ - \sqrt{1 + 2KK_1 \cos \sigma + K^2 K_1^2} + 1 - KK_1 - K. \quad (21)$$

Формула

$$\cos \varphi = \sqrt{1 - 2K \cos \sigma + K^2} - \sqrt{1 + 2KK_1 \cos \sigma + K^2 K_1^2} + \\ + 1 - KK_1 - K$$

определяет зависимость между углом насыщения φ и фазовым сдвигом управляющего напряжения σ . Из формулы вытекает, что если угол σ изменяется от нуля до π , угол насыщения изменяется от π до нуля. Необходимо заметить, что приведенная формула теряет силу при углах σ , близких к нулю. Точнее, формула справедлива, если

$$\sqrt{1-2K \cos \sigma + K^2} - \sqrt{1+2KK_1 \cos \sigma + K^2K_1^2} + 1 - KK_1 - K \geq -1,$$

Если

$$\sqrt{1-2K \cos \sigma + K^2} - \sqrt{1+2KK_1 \cos \sigma + K^2K_1^2} + 1 - KK_1 - K < -1,$$

то $\varphi = \pi$.

Формула неприменима при углах σ , близких к нулю, потому что при этих углах в конце четвертого интервала наступает насыщение сердечника. Поток, уменьшаясь под действием управляющего напряжения, достигает величины $-\Phi_s$. Индуктивности обмоток скачком уменьшаются до нуля и дальнейшее изменение потока прекращается. Для всех углов σ , при которых наступает подобное насыщение, угол φ равен π . Это явление насыщения наступает лишь при углах σ , близких к нулю, и учитывается наложением ограничения на применимость формулы (21).

В приближенных расчетах можно считать, что величина K_1 равна нулю. Тогда формула упрощается:

$$\cos \varphi = \sqrt{1-2K \cos \sigma + K^2} - K. \quad (22)$$

Данная формула справедлива для всех углов σ .

Из изложенного выше вытекает, что, изменяя угол сдвига фаз между рабочим и управляющим напряжением при постоянной амплитуде последнего, можно менять угол насыщения φ дросселя. Изменение угла насыщения φ влечет за собой изменение средней величины рабочего тока. Таким образом, изменяя фазовый сдвиг управляющего напряжения, можно менять среднюю величину рабочего тока, протекающего через сопротивление нагрузки.

Выясним зависимость между углом насыщения и средней величиной рабочего тока I .

Как было показано выше, мгновенные значения рабочего тока определяются на различных интервалах времени следующими формулами.

Первый интервал:

$$0 \leq \omega t \leq \varphi,$$

$$i = \frac{U_m}{R + \frac{\omega^2}{\omega_y^2} R_y} \sin \omega t + \frac{H_c l_{cp}}{\omega}.$$

Второй и третий интервалы:

$$\varphi \leq \omega t \leq \theta, \quad i = \frac{U_m}{R} \sin \omega t.$$

Четвертый интервал:

$$i = \frac{U_m}{R + \frac{w^2}{w_y^2} R_y} \sin \omega t - \frac{KU_m}{R + \frac{w^2}{w_y^2} R_y} \sin(\omega t + \sigma),$$

В формулу для рабочего тока, протекающего в течение первого интервала, введено слагаемое $\frac{H_c l_{cp}}{w}$, учитывающее намагничивающий ток, которым мы раньше пренебрегали. H_c — коэрцитивная сила $\left[\frac{a}{см} \right]$, l_{cp} — длина средней силовой линии сердечника [см].

Заметим, что

$$\frac{1}{R + \frac{w^2}{w_y^2} R_y} = \frac{1}{R \left(1 + \frac{1}{K_1} \right)} = \frac{K_1}{R(K_1 + 1)},$$

где

$$K_1 = \frac{w_y^2 R}{w^2 R_y}.$$

Средняя величина рабочего тока

$$\begin{aligned} I &= \frac{1}{2\pi} \int_0^{2\pi} i d(\omega t) = \frac{1}{2\pi} \int_0^{\varphi} \left[\frac{U_m K_1}{R(K_1 + 1)} \sin \omega t + \frac{H_c l_{cp}}{w} \right] d(\omega t) + \\ &+ \frac{1}{2\pi} \int_{\varphi}^{\theta} \frac{U_m}{R} \sin \omega t d(\omega t) + \frac{1}{2\pi} \int_{\theta}^{\gamma} \left[\frac{U_m K_1}{R(K_1 + 1)} \sin \omega t - \right. \\ &\left. - \frac{KU_m K_1}{R(K_1 + 1)} \sin(\omega t + \sigma) \right] d(\omega t) = \frac{1}{2\pi} \frac{U_m K_1}{R(K_1 + 1)} (1 - \cos \varphi) + \\ &+ \frac{1}{2\pi} \frac{H_c l_{cp}}{w} \varphi + \frac{1}{2\pi} \frac{U_m}{R} (\cos \varphi - \cos \theta) + \\ &+ \frac{1}{2\pi} \frac{U_m K_1}{R(K_1 + 1)} (\cos \theta - \cos \gamma) - \\ &- \frac{1}{2\pi} \frac{KU_m K_1}{R(K_1 + 1)} [\cos(\theta + \sigma) - \cos(\gamma + \sigma)]. \end{aligned}$$

Подставляя в это выражение выведенные ранее формулы для $\cos \varphi$, $\cos \theta$, $\cos \gamma$, $\cos(\theta + \sigma)$, $\cos(\gamma + \sigma)$, получим

$$I = \frac{1}{2\pi} \frac{U_m}{R} \left[\sqrt{1 - 2K \cos \sigma + K^2} + 1 - K \right] + \frac{1}{2\pi} \frac{H_c l_{cp}}{w} \varphi,$$

где φ определяется из формулы (21).

Таким образом, окончательно формулу можно записать так:

$$I = \frac{1}{2\pi} \frac{U_m}{R} [|\sqrt{1 - 2K \cos \sigma + K^2}| + 1 - K] + \frac{1}{2\pi} \frac{H_c l_{cp}}{w} \arccos [|\sqrt{1 - 2K \cos \sigma + K^2}| - |\sqrt{1 + 2KK_1 \cos \sigma + K^2 K_1^2}| + 1 - KK_1 - K]. \quad (23)$$

Формула (23) так же, как и формула (21), неверна, если

$$|\sqrt{1 - 2K \cos \sigma + K^2}| - |\sqrt{1 + 2KK_1 \cos \sigma + K^2 K_1^2}| + 1 - KK_1 - K < -1,$$

т. е. при углах σ , близких к нулю.

Для этих углов σ , как указывалось выше, $\varphi = \pi$ и средняя величина тока определяется по формуле

$$I = \frac{1}{\pi} \frac{U_m K K_1}{R(1 + K_1)} + \frac{H_c l_{cp}}{2w}. \quad (24)$$

На фиг. 13 представлена характеристика, построенная по формулам (23) и (24).

На этой же фигуре показаны точки, снятые экспериментально. Дроссель, который использовался при эксперименте, имел следующие характеристики.

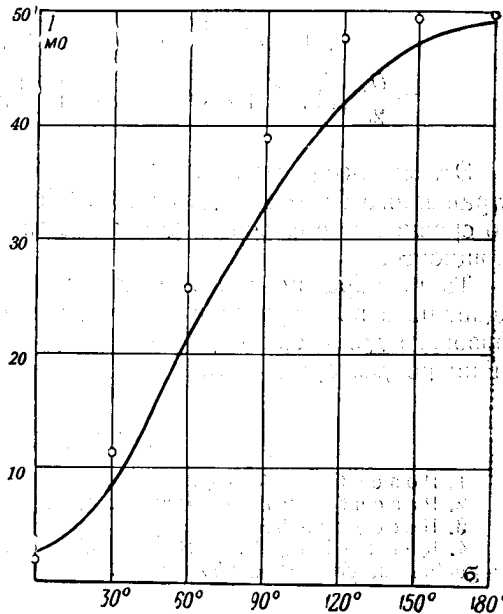
Сердечник дросселя был изготовлен из Мо-перминвара ($B_r = 12\,500$ гс, $B_s = 15\,000$ гс, $H_c = 0,18$ а/см). Толщина штамповок 0,1 мм. Площадь сечения сердечника 0,117 см².

Длина средней силовой линии $l_{cp} = 8,5$ см, $w = 1180$ витков, $w_y = 420$ витков, $R = 245$ ом, $R_y = 957$ ом.

Действующие значения рабочего и управляющего напряжений равнялись соответственно: $U = 27$ в, $U_y = 13$ в.

На основании изложенного можно сделать следующие выводы.

Угол насыщения дросселя с сердечником из материала, обладающего прямоугольной петлей гистерезиса, зависит от фазы переменного напряжения, приложенного к управляющей обмотке.



Фиг. 13. Зависимость средней величины рабочего тока дросселя от угла сдвига фаз между управляющим и рабочим напряжениями.

Сдвиг фазы управляющего напряжения относительно рабочего влияет на среднюю величину рабочего тока таким образом, что при нулевом сдвиге фаз ток минимален, а при сдвиге фаз, равном 180° , ток максимален.

Дроссель насыщения с фазовым управлением может применяться в схемах автоматического регулирования, в которых выходной величиной измерительного органа является напряжение, меняющееся по фазе с изменением регулируемой величины.

Два дросселя насыщения с фазовым управлением, соединенные в дифференциальную схему таким образом, что через сопротивление нагрузки проходит ток, равный разности токов дросселей, а углы сдвига фаз между рабочим и управляющими напряжениями в дросселях равны соответственно $\pi \pm \sigma$ и σ , обладают характеристикой, приближенно определяемой выражением

$$I = I_1 - I_2 = \frac{U_m}{2\pi R} [|\sqrt{1 - 2K \cos(\pi \pm \sigma) + K^2}| + 1 - K - |\sqrt{1 - 2K \cos \sigma + K^2}| - 1 + K] = \frac{U_m}{2\pi R} [|\sqrt{1 + 2K \cos \sigma + K^2}| - |\sqrt{1 - 2K \cos \sigma + K^2}|]. \quad (25)$$

Эта характеристика совпадает с характеристикой широко распространенного фазового детектора. Достоинством полученной схемы по сравнению с фазовым детектором является наличие усиления по мощности.

Таким образом, дроссели насыщения с фазовым управлением, соединенные по дифференциальной схеме, могут применяться вместо фазового детектора в различного рода устройствах электроавтоматики: регуляторах частоты, следящих системах и т. д.

ЛИТЕРАТУРА

1. Розенблат М. А., Магнитные усилители, ГЭИ, 1949.
2. Розенблат М. А., Магнитные усилители, ГЭИ, 1955.
3. Бессонов Л. А., Электрические цепи со сталью, ГЭИ, 1948.
4. K ü m m e l F., Verstärkung kleiner Wechselstromspannungen mit dem magnetischen Verstärker, ETZ, 1954, Heft 11.

Канд. техн. наук *Н. Т. КОБЯКОВА*, инж. *В. А. САРТОРИ*

СТАБИЛИЗАТОР НАПРЯЖЕНИЯ ПЕРЕМЕННОГО ТОКА ПОВЫШЕННОЙ ТОЧНОСТИ С ДРОССЕЛЕМ НАСЫЩЕНИЯ

При современном уровне техники вопрос автоматической регулировки напряжения переменного тока с большой степенью точности становится весьма актуальным. В каждом конкретном случае к стабилизатору напряжения переменного тока предъявляется ряд требований, от которых зависит выбор отдельных элементов схемы и их параметров. Ниже рассматривается схема регулятора напряжения, служащего для питания цепей накала электронных ламп. При изменении входного напряжения на $\pm 10\%$ действующее значение напряжения на выходе поддерживается с точностью $\pm 0,05\%$.

Прежде чем перейти к рассмотрению схемы стабилизатора, следует остановиться на выводе основных соотношений, необходимых для анализа работы стабилизатора.

Существует большое количество различных схем стабилизаторов, но все они в зависимости от способа включения измерительного органа (И. О.) могут быть разбиты на три основных типа:

а) на вход И. О. подается напряжение нагрузки (фиг. 1,а), т. е.

$$U_{изм} = U_{нагр};$$

б) на вход И. О. подается входное напряжение стабилизатора (фиг. 1,б), т. е.

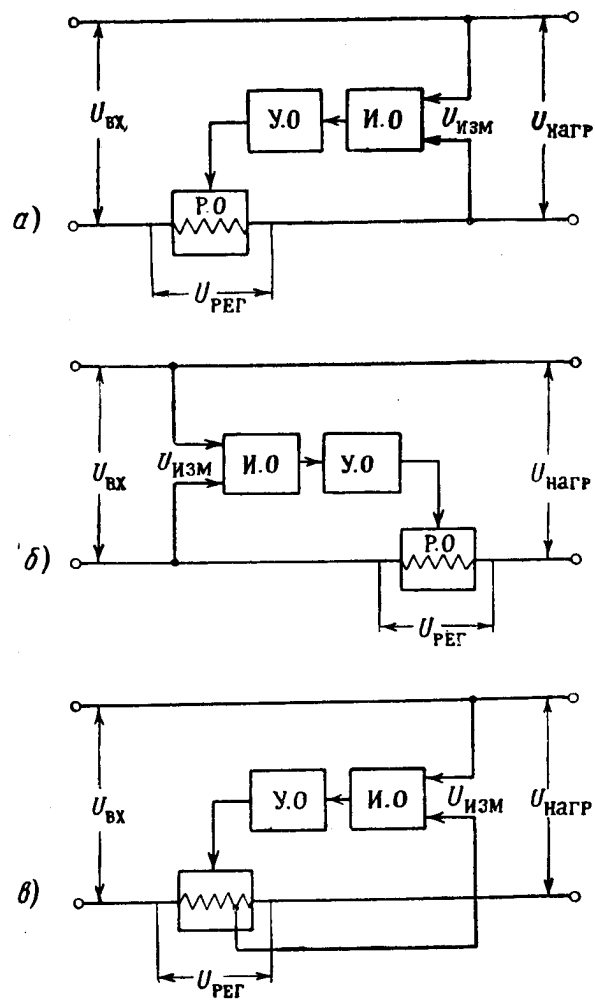
$$U_{изм} = U_{вх};$$

в) на вход И. О. подается сумма напряжений: напряжение на нагрузке и часть напряжения на регулирующем органе (Р. О.) (фиг. 1,в), т. е.

$$U_{изм} = |U_{нагр} + kU_{рег}|, \text{ где } 0 \leq k \leq 1.$$

Наибольшее распространение получили стабилизаторы, относящиеся к типу, приведенному на фиг. 1,а. Стабилизаторы типа, приведенного на фиг. 1,б, применяются реже, так как они, как будет показано ниже, обладают рядом существенных недостатков. Включение И. О. по схеме, изображенной на фиг. 1,в, до сих пор теоретически не исследовалось, но примененную инж. Евдокимовым схему стабилизатора¹ можно отнести к этому типу.

¹ Г. К. Евдокимов, Стабилизатор напряжения с дросселем насыщения, «Вестник электропромышленности», 1950, № 11.



Фиг. 1. Скелетные схемы регуляторов.
И. О.—измерительный орган, У. О.—усилительный орган, Р. О.—регулирующий орган.

ХАРАКТЕРИСТИКА СТАБИЛИЗАТОРА

Точность регулирования выходного напряжения стабилизатора при постоянном значении нагрузки характеризуется коэффициентом стабилизации K_c :

$$K_c = \frac{\Delta U_{\text{нагр}} \%}{\Delta U_{\text{вх}} \%} = \frac{\Delta U_{\text{нагр}} U_{\text{вх. ном}}}{\Delta U_{\text{вх}} U_{\text{нагр. ном}}}, \quad (1)$$

где $\Delta U_{\text{вх}} \%$ — изменение входного напряжения в % от его номинального значения;

$\Delta U_{\text{нагр}} \%$ — изменение напряжения на нагрузке в % от его номинального значения.

Обозначив $\frac{U_{\text{нагр. ном}}}{U_{\text{вх. ном}}} = \alpha$, получаем $K_c = \frac{\Delta U_{\text{нагр}}}{\Delta U_{\text{вх}}} \frac{1}{\alpha}$.

Очевидно, чем выше требуемая точность регулирования, тем меньше должен быть коэффициент стабилизации K_c . Для обеспечения заданного K_c при выбранной схеме необходимо, чтобы стабилизатор имел определенную величину коэффициента усиления

$$K_y = \frac{\Delta U_{\text{рег}}}{\Delta U_{\text{изм}}}, \quad (2)$$

где $\Delta U_{\text{рег}}$ и $\Delta U_{\text{изм}}$ — изменения напряжения соответственно на регулирующем органе и на входе измерительного органа. Требуемая величина K_y при заданном значении K_c зависит от способа включения измерительного органа.

Необходимые для сравнения различных схем стабилизаторов соотношения между K_y и K_c наиболее просто могут быть получены в том случае, когда напряжения на нагрузке и на Р. О. совпадают по фазе.

Для этого случая

$$U_{\text{вх}} = U_{\text{рег}} + U_{\text{нагр}}. \quad (3)$$

Найдем зависимость между $\Delta U_{\text{нагр}}$ и $\Delta U_{\text{вх}}$ при постоянной нагрузке для схемы, приведенной на фиг. 1, в. Соотношение между $\Delta U_{\text{нагр}}$ и $\Delta U_{\text{вх}}$ для стабилизаторов, выполненных по схемам, изображенным на фиг. 1, а и б, определяется из общего выражения путем приравнивания величины k нулю или единице.

Итак, для схемы, приведенной на фиг. 1, в, имеем

$$\Delta U_{\text{изм}} = \Delta U_{\text{вх}} - (1-k) \Delta U_{\text{рег}}. \quad (4)$$

Но, учитывая, что $\Delta U_{\text{изм}} = \frac{\Delta U_{\text{рег}}}{K_y}$, получаем

$$(1 - K_y k + K_y) \Delta U_{\text{рег}} = K_y \Delta U_{\text{вх}}. \quad (5)$$

Выражая $\Delta U_{\text{рег}}$ через $\Delta U_{\text{вх}}$ и $\Delta U_{\text{нагр}}$, получаем

$$\Delta U_{\text{рег}} = \Delta U_{\text{вх}} - \Delta U_{\text{нагр}}. \quad (6)$$

Подставляя (6) в (5), после некоторых преобразований получим

$$\frac{\Delta U_{\text{нагр}}}{\Delta U_{\text{вх}}} = \frac{1 - kK_y}{1 - kK_y + K_y}. \quad (7)$$

Таким образом, коэффициент стабилизации для схемы, приведенной на фиг. 1,в, запишется в следующем виде:

$$K_c = \frac{1 - kK_y}{1 - kK_y + K_y} \frac{1}{\alpha}. \quad (8)$$

Для стабилизатора, выполненного по схеме, показанной на фиг. 1,а ($k=0$), имеем

$$K_c = \frac{1}{(1 + K_y) \alpha}. \quad (9)$$

А для схемы стабилизатора, приведенной на фиг. 1,б ($k=1$), получим

$$K_c = (1 - K_y) \frac{1}{\alpha}. \quad (10)$$

Для схем 1,б и в имеется возможность получить $K_c=0$. Это будет для схемы 1,б при $K_y=1$, а для схемы 1,в — при $K_y = \frac{1}{k}$. Для схемы 1,а $K_c=0$ лишь при $K_y = \infty$.

Рассмотрим теперь зависимость между $\Delta U_{\text{нагр}}$ и изменением полной мощности нагрузки $S_{\text{нагр}}$. Предположим, что стабилизатор должен поддерживать постоянство $U_{\text{нагр}}$ с определенной степенью точности при m -кратном изменении $S_{\text{нагр}}$. Иными словами, предположим, что

$$S_{\text{нагр max}} = m S_{\text{нагр min}}.$$

Принимая фазовый угол нагрузки $\varphi_{\text{нагр}}$ и $U_{\text{нагр}}$ постоянными, получим

$$U_{\text{нагр}} = \frac{S_{\text{нагр max}}}{I_{\sim \text{max}}} = \frac{S_{\text{нагр min}}}{I_{\sim \text{min}}},$$

откуда

$$\Delta I_{\sim \text{max}} = I_{\sim \text{max}} - I_{\sim \text{min}} = \frac{S_{\text{нагр max}}}{U_{\text{нагр}}} \left(1 - \frac{1}{m}\right) = \frac{m-1}{m} I_{\sim \text{max}}. \quad (11)$$

Далее предположим, что $\frac{\Delta U_{\text{рег } \Delta I_{\sim}}}{\Delta I_{\sim}} = \text{const} = n$. Это предположение, как будет показано ниже, справедливо для рабочих участков характеристик дросселя насыщения. Тогда изменение напряжения на Р. О., вызванное m -кратным изменением нагрузки, можно выразить следующим образом:

$$\Delta U_{\text{рег } \Delta I_{\sim}} = n \Delta I_{\sim \text{max}} = \frac{n}{m} (m-1) I_{\sim \text{max}}. \quad (12)$$

Величина

$$\Delta U_{\text{рег } \Delta I_{\sim}} = \Delta U_{\text{рег}} + \Delta U_{\text{нагр}}, \quad (13)$$

где

$$\Delta U_{\text{рег}} = K_y \Delta U_{\text{изм}}. \quad (14)$$

В свою очередь

$$\Delta U_{\text{изм}} = (1 - k) (\Delta U_{\text{рег } \Delta I_{\sim}} - \Delta U_{\text{рег}}). \quad (15)$$

Подставляя выражения (13) и (14) в (15), после некоторых преобразований получим

$$\Delta U_{\text{нагр}} = \frac{\Delta U_{\text{рег } \Delta I_{\sim}}}{1 - K_y k + K_y}. \quad (16)$$

Принимая во внимание выражение (12), окончательно имеем

$$\Delta U_{\text{нагр}} = \frac{n(m-1)}{m(1 - K_y k + K_y)} I_{\sim \text{max}} \quad (17)$$

или

$$\Delta U_{\text{нагр } \Delta I_{\sim}} \% = \frac{100n(m-1)I_{\sim \text{max}}}{m(1 - K_y k + K_y)U_{\text{нагр}}}. \quad (18)$$

Для схемы, приведенной на фиг. 1,а, ($k=0$)

$$\Delta U_{\text{нагр}} = \frac{\Delta U_{\text{рег } \Delta I_{\sim}}}{1 + K_y}. \quad (19)$$

Для схемы, изображенной на фиг. 1,б, ($k=1$)

$$\Delta U_{\text{нагр}} = \Delta U_{\text{рег } \Delta I_{\sim}}. \quad (20)$$

Для схемы, показанной на фиг. 1,в, при

$$K = \frac{1}{K_y} \Delta U_{\text{нагр}} = \frac{\Delta U_{\text{рег } \Delta I_{\sim}}}{K_y}. \quad (21)$$

Равенства (19), (20) и (21) показывают, что влияние изменения нагрузки на $\Delta U_{\text{нагр}}$ в стабилизаторах, выполненных по приведенным на фиг. 1,а и 1,в схемам, практически одинаково.

Стабилизаторы, выполненные по схеме, показанной на фиг. 1,б, в этом случае не применимы, так как совсем не реагируют на изменение нагрузки.

Таким образом, способность стабилизатора поддерживать постоянство $U_{\text{нагр}}$ при изменениях нагрузки зависит в основном от K_y и n . Для того чтобы уменьшить влияние изменения нагрузки на $U_{\text{нагр}}$, следует K_y увеличивать, а n уменьшать за счет уменьшения крутизны характеристики дросселя.

Итак, изменения напряжения на нагрузке, вызванные изменениями напряжения сети и m -кратным изменением нагрузки, определяются следующим выражением:

$$\Delta U_{\text{нагр}} \% = \frac{1}{\alpha (1 - kK_y + K_y)} \left[(1 - kK_y) \Delta U_{\text{вх}} \% + \frac{100n(m-1)I_{\sim\text{max}}}{mU_{\text{вх. ном}}} \right]. \quad (22)$$

Количественная оценка влияния изменений входного напряжения, сопротивления нагрузки и неустойчивости K_y на $U_{\text{нагр}}$ для следующего частного случая, когда

$$U_{\text{вх. ном}} = 220 \text{ в}; \Delta U_{\text{вх}} = \pm 10\%; \alpha = 0,8;$$

$$K_y = 100; \Delta K_y = -20\%; I_{\sim\text{max}} = 2 \text{ а}; m = 2 \text{ и } n = 100 \frac{\text{в}}{\text{а}},$$

характеризуется приводимой ниже таблицей.

Тип стабилизатора	$\Delta U_{\text{вх}} = \pm 10\%;$ $S_{\text{нагр}} = \text{const}$		$\Delta U_{\text{вх}} = 0; \frac{S_{\text{max}}}{S_{\text{min}}} = 2$		$\Delta U_{\text{вх}} = \pm 10\%;$ $S_{\text{нагр}} = \text{const};$ $\Delta K_y = -20\%$	
	K_y	$\Delta U_{\text{нагр}} \%$	K_y	$\Delta U_{\text{нагр}} \%$	K_y	$\Delta U_{\text{нагр}} \%$
Схема (см. фиг. 1, а) $k=0$	100	$\pm 0,124$	100	0,564	80	$\pm 0,154$
Схема (см. фиг. 1, б) $k=1$	1	0	1	57,0	0,8	$\pm 2,5$
Схема (см. фиг. 1, в) $k=0,01$	100	0	100	0,570	80	$\pm 0,031$

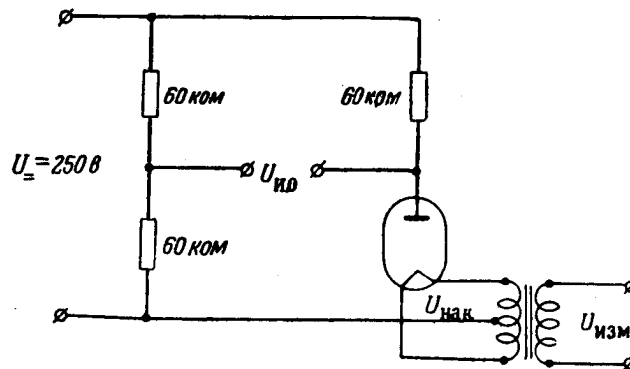
Из всего изложенного выше следует, что преимущества стабилизатора, выполненного по схеме, приведенной на фиг. 1, в, при $K = \frac{1}{K_y}$ сводятся к возможности получения (при сравнительно небольшой величине K_y) высокой точности стабилизации в случае изменения только $U_{\text{вх}}$. При этом зависимость выходного напряжения стабилизатора от колебаний тока нагрузки и неустойчивости K_y остается в этом случае практически такой же, как и для схемы, изображенной на фиг. 1, а.

ИЗМЕРИТЕЛЬНЫЙ ОРГАН (И. О.)

В процессе работы стабилизатора с дросселем насыщения форма кривой выходного напряжения ($U_{\text{нагр}}$) будет изменяться с изменением $U_{\text{вх}}$. Поэтому действующее и амплитудное значения $U_{\text{нагр}}$ бу-

дут поддерживаться с разной степенью точности, если, конечно, на выходе стабилизатора не имеется фильтра, исправляющего форму кривой выходного напряжения.

При включении И. О. по схеме, приведенной на фиг. 1, б, амплитудное значение может поддерживаться с точностью примерно до $n\%$ (n — изменение процентного содержания высших гармоник в $U_{\text{нагр}}$, обусловленное наличием дросселя насыщения). Действующее значение $U_{\text{нагр}}$ будет при этом поддерживаться с несколько большей точностью. Поэтому при необходимости поддерживать $U_{\text{нагр}}$ с большой степенью точности следует включать И. О. по схемам 1, а или 1, в, а тип И. О. выбирать исходя из того, какое значение необходимо поддерживать с большей степенью точности. При необходимости поддерживать постоянство амплитудного значения $U_{\text{нагр}}$ в качестве И. О.



Фиг. 2. Мостовая схема измерительного органа с диодом 4Ц1М.

могут быть применены различные схемы выпрямителей с RC-фильтром. При необходимости стабилизировать действующее значение $U_{\text{нагр}}$ можно применять различные варианты мостовых схем с нелинейными элементами, в качестве которых могут применяться, например, диоды прямого накала, работающие в режиме насыщения.

Они обладают способностью сильно изменять свое эквивалентное сопротивление $R_s = \frac{U_a}{I_a}$ при небольших изменениях напряжения накала¹.

На фиг. 2 изображена одна из возможных схем И. О. с диодом 4Ц1М в качестве нелинейного сопротивления моста. Коэффициент усиления И. О. может быть представлен в виде

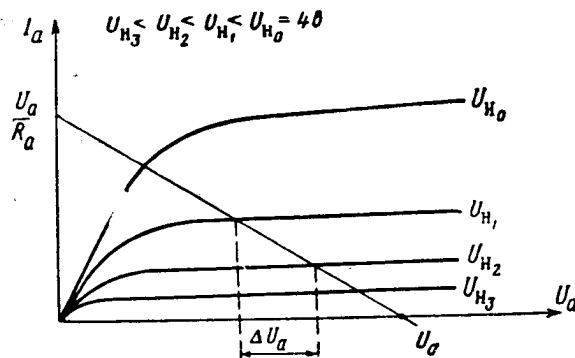
$$K_{\text{И. О.}} = K_m K_{\text{тр}}, \quad (23)$$

где $K_m = \frac{\Delta U_{\text{И. О.}}}{\Delta U_{\text{нак}}}$ — коэффициент усиления мостов с диодом;
 $K_{\text{тр}}$ — коэффициент трансформации трансформатора накала диода. (Обычно $K_{\text{тр}} \ll 1$).

¹ Г. Р. Герценберг, Высокостабильный диод в схемах регулирования и автоматики, «Вестник электропромышленности», 1948, № 5.

K_n определяется из характеристик диода, примерный вид которых представлен на фиг. 3, как отношение $\frac{\Delta U_a}{\Delta U_n}$, где ΔU_a и ΔU_n — изменения анодного напряжения и напряжения накала диода.

Для схемы И. О., приведенной на фиг. 2, при $U_n = 3,2 \div 3,4$ в $K_n = 250$. В заключение следует отметить, что к стабильности характеристик всех элементов И. О. предъявляются особенно высокие тре-



Фиг. 3. Вольтамперные характеристики диода 4Ц1М.

бования, так как они определяют собой стабильность работы всего стабилизатора. Эквивалентное сопротивление диода 4Ц1М очень мало зависит от окружающей температуры, но в точных стабилизаторах с этим изменением все же приходится считаться.

РЕГУЛИРУЮЩИЙ ОРГАН (Р. О)

В качестве регулирующего органа обычно применяется дроссель насыщения с подмагничиванием постоянным током или трансформатор с регулирующими лампами. Ниже рассматривается только дроссель насыщения.

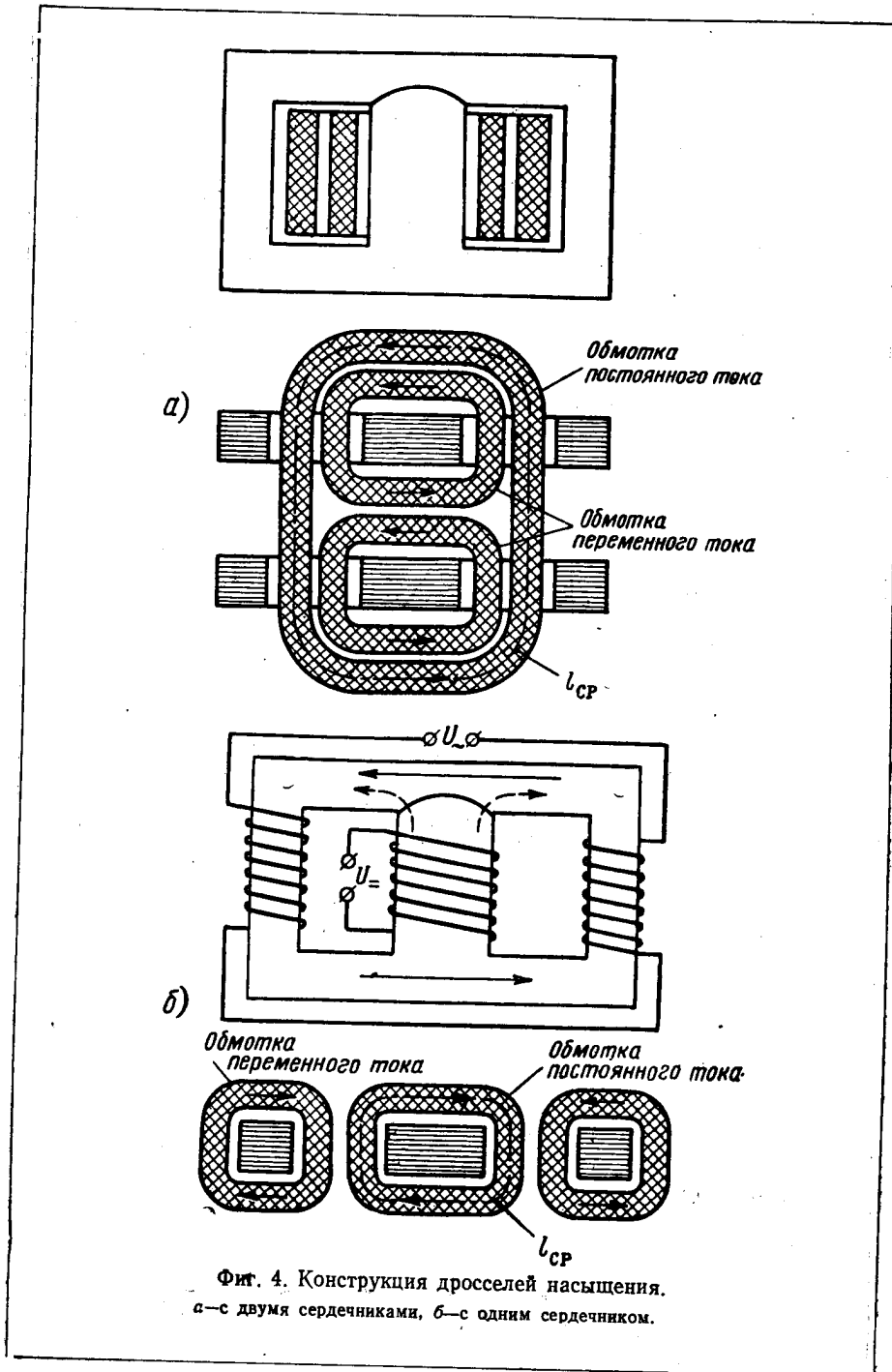
Для того чтобы в обмотке постоянного тока не наводилась переменная ЭДС, обусловленная током в обмотке переменного тока, дроссель выполняется либо с двойным сердечником, по схеме, приведенной на фиг. 4,а, либо с одним сердечником, по схеме, показанной на фиг. 4,б, в которой средний стержень является нерабочим и поэтому общий объем стали требуется больший, чем в предыдущей схеме. В дросселе, выполненном по схеме фиг. 4,а, может быть также применен тороидальный сердечник.

На фиг. 5 приведены графики зависимости действующего напряжения на обмотке переменного тока дросселя $U_{\text{пер}}$ от

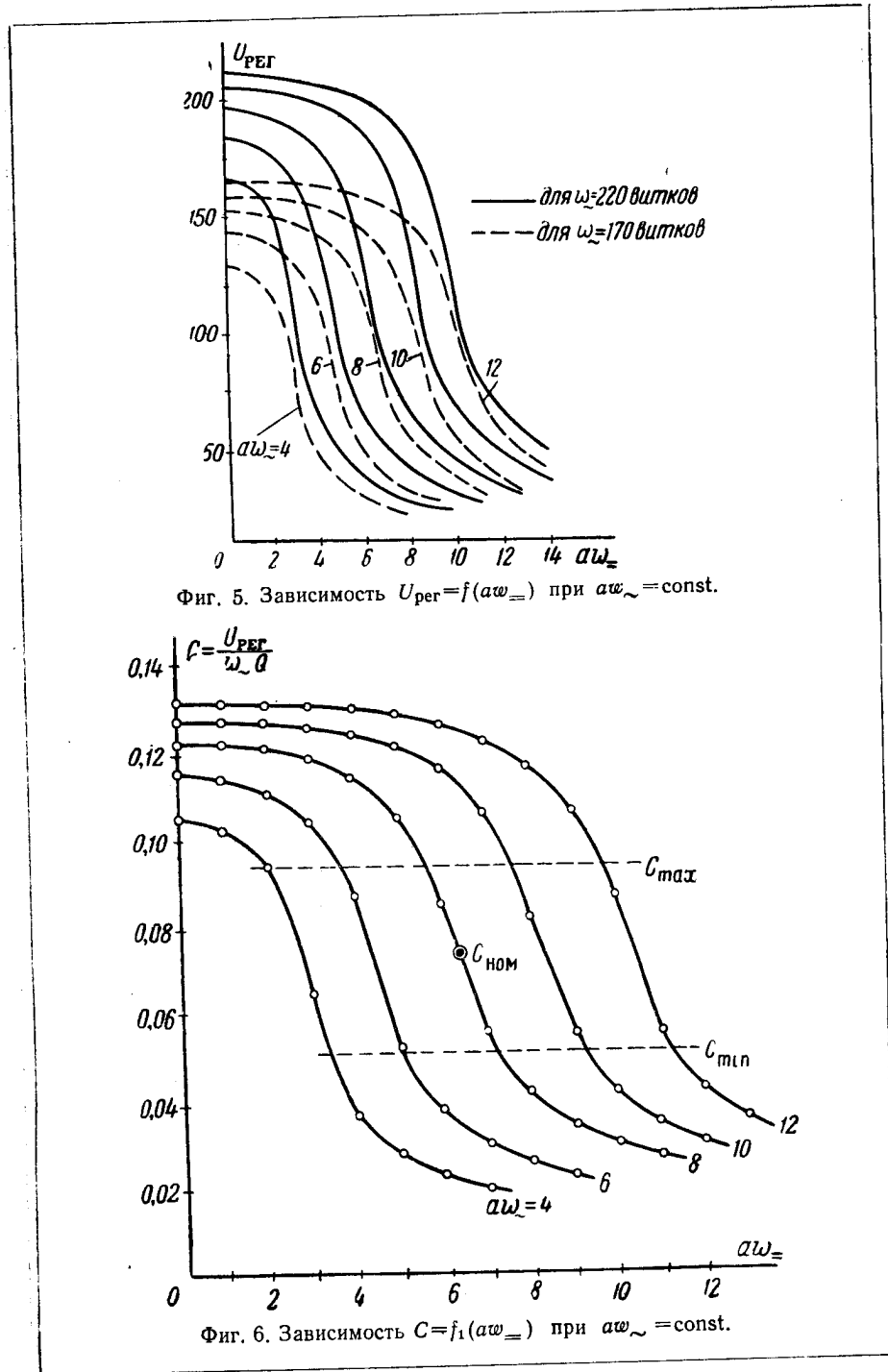
$$a\omega_{\sim} = \frac{I_{\sim} \omega_{\sim}}{I_{\sim}} \quad \text{при} \quad a\omega_{\sim} = \frac{I_{\sim} \omega_{\sim}}{I_{\sim}} = \text{const},$$

где I_{\sim} и ω_{\sim} — ток и число витков обмотки подмагничивания;

I_{\sim} и ω_{\sim} — действующее значение тока и число витков полуобмотки переменного тока;

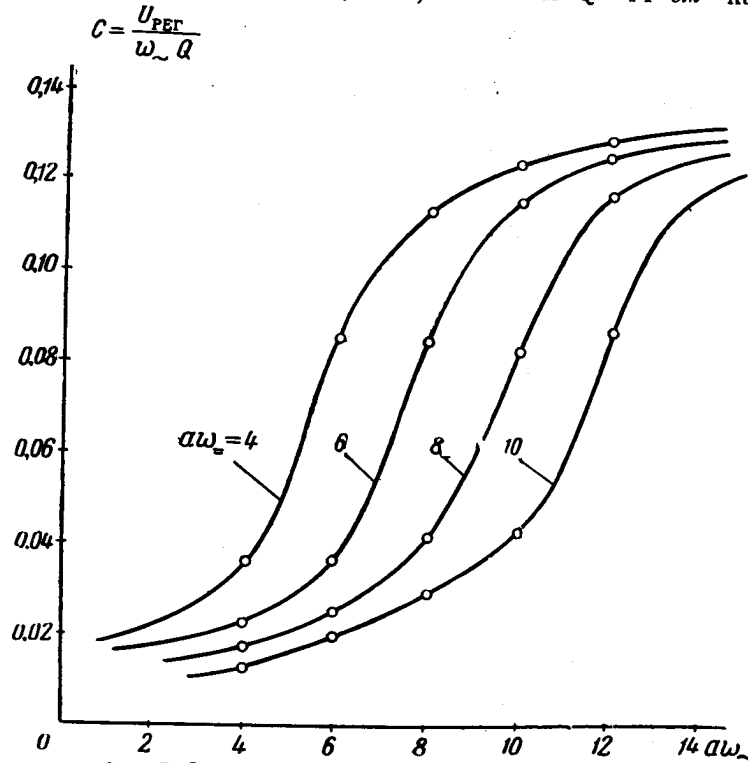


Фиг. 4. Конструкция дросселей насыщения.
а—с двумя сердечниками, б—с одним сердечником.



l_{\sim} и l_{\sim} — средние длины силовых линий магнитных полей постоянного и переменного токов (в см).

Действующие значения $U_{\text{пер}}$ и I_{\sim} замерялись с помощью электродинамических приборов. Графики сняты при частоте 50 гц для дросселя, собранного по схеме, приведенной на фиг. 4,а, на сердечниках из трансформаторной стали (Ш-35) сечением $Q=14 \text{ см}^2$ каждый.



Фиг. 7. Зависимость $C=f(a\omega_{\sim})$ при $a\omega_{\sim}=\text{const}$.

Кривые приведены для числа витков переменного тока $\omega_{\sim} = 170$ и $\omega_{\sim} = 220$ витков. Эти графики справедливы для данного сорта стали и частоты f_1 лишь при заданном сечении сердечника Q и определенном ω_{\sim} . Для того чтобы построить кривые, зависящие только от сорта стали и частоты f , необходимо разделить все ординаты графика (фиг. 5) на $Q\omega_{\sim}$.

$$C = \frac{U_{\text{пер}}}{Q\omega_{\sim}} \quad (24)$$

На фиг. 6 изображены кривые $C=f_1(a\omega_{\sim})$ при $a\omega_{\sim}=\text{const}$. При этом предполагалось, что величиной активного сопротивления обмоток и рассеянием можно пренебречь. Для того чтобы охарактеризовать зависимость $U_{\text{пер}}$ от $I_{\text{нагр}}$ на фиг. 7 построены графики $C=f_2(a\omega_{\sim})$ при $a\omega_{\sim}=\text{const}$.

Под коэффициентом усиления дросселя будем понимать

$$K_{др} = \frac{\Delta U_{рег}}{\Delta I_{\sim}}. \quad (25)$$

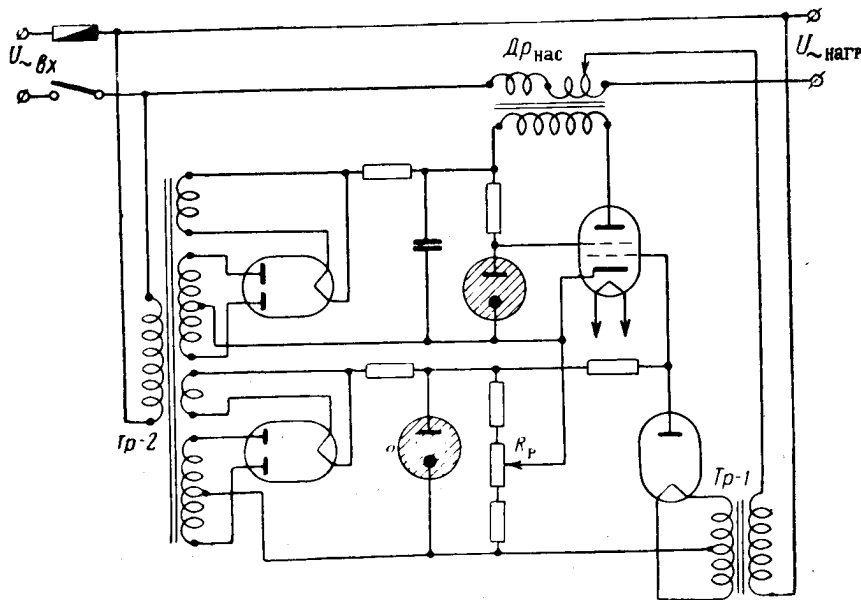
Тогда коэффициент усиления Р. О. запишется в виде

$$K_{р.о} = K_{др} S = \frac{\Delta C Q \omega_{\sim} S}{\Delta I_{\sim}}, \quad (26)$$

где S — крутизна лампы (см. фиг. 8).

СХЕМА СТАБИЛИЗАТОРОВ

На фиг. 8 приведена схема стабилизатора для питания цепей накала. Число включенных ламп не изменяется, поэтому можно считать, что величина нагрузки стабилизатора остается постоянной. При-



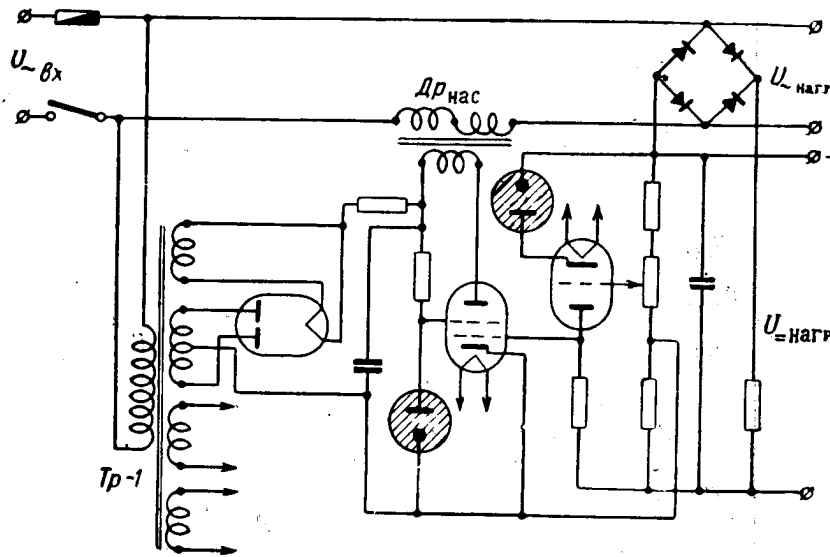
Фиг. 8. Схема стабилизатора для питания цепей накала.

мененный в качестве нелинейного элемента моста диод 4Ц1М поддерживает постоянство действующего значения выходного напряжения стабилизатора. В этом случае форма кривой выходного напряжения не имеет значения. В тех случаях, когда искажения формы кривой выходного напряжения стабилизатора не допустимы, на выход стабилизатора может быть включен фильтр.

Усилительный орган осуществлен с помощью электронных ламп, однако следует указать, что он с успехом может быть заменен одним каскадом магнитного усилителя.

Грубая регулировка $U_{\text{нагр}}$ в схеме, показанной на фиг. 8, осуществляется изменением числа витков первичной обмотки трансформатора Тр-1, точная — сопротивлением R_p .

При наладке схемы может оказаться, что диапазон изменения $U_{\text{вх}}$, в пределах которого происходит регулирование, несколько не соответствует желаемому. В этом случае следует изменить число витков обмотки переменного тока дросселя. При расчете стабилизаторов, работающих при различных токах нагрузки, необходимо выбрать значение α с таким расчетом, чтобы точки пересечения линий $U_{\text{пер max}}$ и $U_{\text{пер min}}$ с рабочими характеристиками дросселя нигде не выходили за их крутые участки. В некоторых случаях целесообразно часть $\alpha \omega =$



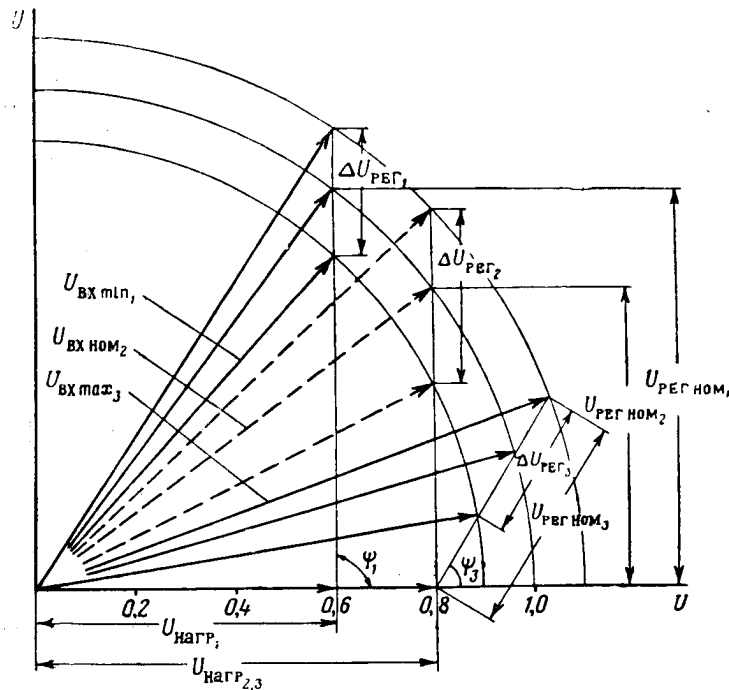
Фиг. 9. Схема стабилизатора с выпрямленным выходным напряжением.

создавать с помощью отдельной обмотки подмагничивания. Изготовленный по схеме, приведенной на фиг. 8, стабилизатор после тридцатиминутного прогрева в течение 8 часов работы в помещении поддерживал постоянство $U_{\text{нагр}}$ с точностью до $\pm 0,05\%$ при постоянных нагрузках от 100 до 400 вт и $\Delta U_{\text{вх}} = \pm 10\%$. При полном сбросе максимальной нагрузки $\Delta U_{\text{нагр}} = 0,8\%$. На фиг. 9 представлена схема стабилизатора с выпрямленным выходным напряжением. Такая схема, обеспечивая хорошее регулирование, обладает гораздо большим КПД, чем обычно применяемые электронные стабилизаторы напряжения постоянного тока, и может обеспечивать выходную мощность до 1 квт при использовании лишь одной регулирующей лампы. Кроме того, с выхода стабилизатора можно снять напряжение переменного тока, стабилизированное с несколько меньшей степенью точности.

РАСЧЕТ СТАБИЛИЗАТОРА

1. Определение $U_{\text{рег. ном}}$ и $\Delta U_{\text{рег. ном}}$

$U_{\text{рег. ном}}$ определяется с помощью векторной диаграммы. На фиг. 10 изображены векторные диаграммы для случая, когда фазовый угол $\psi = \varphi_{\text{рег}} - \varphi_{\text{нагр}}$ между напряжениями $U_{\text{рег}}$ и $U_{\text{нагр}}$ равен 90° , и для случая $\psi < 90^\circ$. Сначала рассмотрим случай $\psi = 90^\circ$. По оси абсцисс отложена величина $U_{\text{нагр}}$, которую для точных стабилизаторов можно считать неизменной. Для определения $U_{\text{рег. ном}}$ при номинальном напряжении на входе стабилизатора



Фиг. 10. Векторные диаграммы регулятора при различных α и ψ .

($U_{\text{вх. ном}}$) из конца вектора $U_{\text{нагр}}$ восстанавливается перпендикуляр до пересечения с окружностью, радиус которой равен $U_{\text{вх. ном}}$. Для того чтобы определить, насколько должно измениться $U_{\text{рег}}$ при заданном изменении напряжения на входе, следует провести новую окружность, радиус которой равнялся бы новому значению напряжения на входе, и тем же способом определить новое значение $U_{\text{рег}}$.

На фиг. 10 векторные диаграммы построены для двух значений $\alpha = \frac{U_{\text{нагр. ном}}}{U_{\text{вх. ном}}}$ ($\alpha_1 = 0,6$ и $\alpha_2 = 0,8$) и изменения входного напряжения $\pm 10\%$.

Из диаграммы видно, что с увеличением α :

1) уменьшается мощность дросселя $S_{\text{рег. ном}} = S_{\text{нагр}} \frac{U_{\text{рег. ном}}}{U_{\text{нагр}}}$;

2) увеличивается отношение $\frac{S_{\text{нагр}}}{S_{\text{вх. ном}}}$ стабилизатора.

Если не учитывать мощность, расходуемую на управление, то при $U_{\text{вх. ном}}$

$$\frac{S_{\text{нагр}}}{S_{\text{вх. ном}}} = \frac{U_{\text{нагр}}}{U_{\text{вх. ном}}} = \alpha, \quad (27)$$

а при повышении или понижении входного напряжения на $n\%$

$$\frac{S_{\text{нагр}}}{S_{\text{вх}}} = \frac{100\alpha}{100 \pm n\%};$$

3) увеличивается отношение $\frac{\Delta U_{\text{рег}}}{U_{\text{рег. ном}}}$, а при применении, например, трансформаторной стали дроссель будет работать в пределах прямолинейной части характеристики (см. фиг. 5) до тех пор, пока $\frac{\Delta U_{\text{рег}}}{U_{\text{рег. ном}}}$ не превышает $0,3 \div 0,4$.

Указанными выше соображениями следует руководствоваться при выборе значения α . Обычно α принимается равным $0,7 \div 0,8$ ¹.

Угол ψ равен 90° , если предположить, что сопротивление нагрузки чисто активное, а сопротивление дросселя чисто индуктивное. На самом деле угол ψ бывает равен примерно $60 \div 80^\circ$. На фиг. 10 приведено построение и для $\psi < 90^\circ$, из которого видно, что с уменьшением ψ отношение $\frac{\Delta U_{\text{рег}}}{U_{\text{рег. ном}}}$ увеличивается (изменение угла ψ в процессе регулирования невелико и им для упрощения можно пренебречь).

2. Определение мощности дросселя

Номинальный ток в обмотке дросселя

$$I_{\sim \text{ном}} = \frac{S_{\text{нагр. ном}}}{U_{\text{нагр}}} \quad (28)$$

Мощность дросселя при номинальном напряжении на входе стабилизатора и номинальном токе нагрузки

$$S_{\text{рег. ном}} = U_{\text{рег. ном}} I_{\sim \text{ном}} \quad (29)$$

¹ Чем в больших пределах меняется ток нагрузки, т. е. чем больше m , тем меньшей приходится принимать величину α .

3. Выбор объема сердечника

Значение коэффициента $C_{\text{ном}}$ при номинальном значении входного напряжения выбирается равным ординате средней точки прямолинейного участка кривой (см. фиг. 6), соответствующей выбранной величине

$$a\omega_{\sim\text{ном}} = \frac{a\omega_{\sim\text{max}} + a\omega_{\sim\text{min}}}{2}.$$

При выборе $a\omega_{\sim\text{ном}}$ следует учитывать, что с его увеличением:

- 1) объем стали сердечника дросселя уменьшается (см. ниже);
- 2) объем меди вследствие увеличения $a\omega_{\sim}$ и $a\omega_{=}$ увеличивается;
- 3) сопротивление обмотки подмагничивания, которая является анодной нагрузкой регулирующей лампы, увеличивается, в результате чего при данном $I_{=}$ сеточные смещения могут заходить в положительную область. Зная величину $C_{\text{ном}}$ и номинальную мощность $S_{\text{рег. ном}}$, можно определить необходимый объем стали каждого сердечника дросселя.

Подставляя в выражение (29) значение $U_{\text{рег. ном}}$ из (24) и учитывая, что $I_{\sim} = \frac{a\omega_{\sim} I_{=}}{\omega_{\sim}}$, получаем

$$S_{\text{рег. ном}} = C_{\text{ном}} \omega_{\sim} Q I_{=} a\omega_{\sim\text{ном}} = C_{\text{ном}} U_{\text{ст}} a\omega_{\sim\text{ном}},$$

откуда

$$U_{\text{ст}} = \frac{S_{\text{рег. ном}}}{C_{\text{ном}} a\omega_{\sim\text{ном}}}. \quad (30)$$

4. Расчет обмоток дросселя

Задавшись размерами штампа сердечника и определив l_{\sim} , находим $\omega_{\sim} = \frac{a\omega_{\sim\text{ном}} l_{\sim}}{I_{\sim\text{ном}}}$. Затем из кривых на фиг. 6 для $a\omega_{\sim\text{max}}$ и C_{min} определяем $a\omega_{=\text{max}}$. Величина тока $I_{=\text{max}}$ определяет тип и количество регулирующих ламп. Задавшись его величиной, находим

$$\omega_{=} = \frac{a\omega_{=\text{max}} I_{=}}{I_{=\text{max}}}. \quad (31)$$

Зная среднюю длину витка l_{cp} (см. фиг. 4,а) и диаметр провода обмотки подмагничивания, легко найти ее сопротивление. Если оно окажется таким, что нельзя будет получить выбранное $I_{=\text{max}}$, не заходя в область положительных сеточных смещений регулирующих

ламп, то следует уменьшить $a\omega_{\sim \text{ном}}$ или перейти на больший штамп сердечника.

5. Определение K_y и k

Определим необходимую величину K_y исходя из того, чтобы $\Delta U_{\text{нагр}}^{\text{доп}}/\%$, вызываемое m -кратным изменением нагрузки, не превышало заданного допустимого значения $\Delta U_{\text{нагр. доп}}^{\text{доп}}/\%$.

Для этого определим величину n :

$$n = \frac{\Delta U_{\text{рег}} \Delta I_{\sim}}{\Delta I_{\sim}} = \frac{\Delta C Q \omega_{\sim}}{\frac{\Delta a \omega_{\sim} I_{\sim}}{\omega_{\sim}}} = \frac{\Delta C Q \omega_{\sim}^2}{\Delta a \omega_{\sim} I_{\sim}}. \quad (32)$$

Значение $\frac{\Delta C}{\Delta a \omega_{\sim}}$ находится на фиг. 7 как тангенс наклона кривых на крутых участках характеристик. Подставляя полученное значение n в общее выражение (22), для случая $k = \frac{1}{K_y}$ по заданным значениям m , $I_{\sim \text{max}}$ и $\Delta U_{\text{нагр. доп}}^{\text{доп}}/\%$ определим необходимую величину K_y :

$$K_y = \frac{100n(m-1)I_{\sim \text{max}}}{\Delta U_{\text{нагр. доп}} m U_{\text{нагр}}}. \quad (33)$$

Определив из выражений (23) и (26) коэффициенты усиления И. О. и Р. О., находим коэффициент усиления У. О.

$$K_{y. \text{о}} = \frac{K_y}{K_{\text{и.о}} K_{\text{р.о}}}. \quad (34)$$

Затем находится относительное число витков $\omega_{\sim 0}$ обмотки переменного тока дросселя, напряжение с которых вместе с $U_{\text{нагр}}$ подается на вход И. О.:

$$k = \frac{1}{K_y} = \frac{\omega_{\sim 0}}{2\omega_{\sim}}. \quad (35)$$

При выполнении этого условия напряжение на выходе стабилизатора при изменении входного напряжения практически не будет меняться, так как коэффициент стабилизации K_s [см. уравнение (8)] при этом будет равен нулю.

При $I_{\sim} = \text{const}$ величина K_y в схеме, приведенной на фиг. 1, в, может быть выбрана произвольно исходя из схемных соображений. Для того чтобы $\Delta U_{\text{нагр}} = 0$, необходимо лишь выполнить условие (35).

Приведенный расчет позволяет достаточно просто производить расчет стабилизатора напряжения с необходимой для практических

целей точностью, а применение схемы, представленной на фиг. 1,б, позволяет при сравнительно малой величине коэффициента усиления получить достаточную стабильность выходного напряжения.

ЛИТЕРАТУРА

1. Евдокимов Г. К., Стабилизатор напряжения с дросселем насыщения, «Вестник электропромышленности», 1950, № 11.
 2. Герценберг Г. Р., Высокостабильный диод в схемах регулирования и автоматики, «Вестник электропромышленности», 1948, № 5.
 3. Терентьев Б. П., Электропитание радиоустройств, Гос. изд. литературы по вопросам связи и радио, 1948.
-

Канд. техн. наук *А. К. ТЕР-АКОПОВ*

ОПТИМАЛЬНАЯ ИНДУКТИВНОСТЬ И МИНИМАЛЬНОЕ ВРЕМЯ СРАБАТЫВАНИЯ ЭЛЕКТРОМАГНИТОВ ПОСТОЯННОГО ТОКА

Расчет электромагнита на минимальное время срабатывания имеет большое практическое значение. Проблема создания быстродействующих электромагнитов в качестве важных исполнительных механизмов представляет интерес для авиационных автоматических устройств (автопилоты, регуляторы двигателей, электромеханические установки), а также для многих других отраслей техники.

Несмотря на большое распространение и ответственность быстродействующих электромагнитов, в настоящее время нет каких-либо конкретных рекомендаций по выбору оптимальных параметров электромагнитов для получения минимального времени срабатывания. Более того, при расчете быстродействующих электромагнитов часто исходят из ошибочных исходных положений.

Настоящая статья посвящена выявлению оптимальных соотношений, обеспечивающих минимальное время срабатывания электромагнита.

Задача расчета быстродействующего электромагнита, в сущности, сводится к определению минимального времени срабатывания при заданных значениях потребляемой мощности из сети и производимой полезной работы. Решение этой задачи в общем виде должно быть произведено безотносительно к конкретной конструкции электромагнита, что не позволяет учесть ряд второстепенных факторов. Так, например, при неизвестной конструкции электромагнита нельзя учесть потоки рассеяния и выпучивания, насыщение стали, вихревые токи и т. д.

Поставленная задача в общем виде решалась при следующих допущениях:

- 1) потоки рассеяния и выпучивания отсутствуют;
- 2) магнитное сопротивление главного рабочего зазора пропорционально его длине;
- 3) магнитная система не насыщена;
- 4) отсутствуют вихревые токи и потери на гистерезис.

При известной конструкции электромагнита большинство из этих допущений можно учесть, а для быстродействующих электромагнитов они в значительной степени соответствуют действительности.

Динамика электромагнитов при этих допущениях полностью определяется следующими величинами:

U — приложенное напряжение сети;
 R — активное сопротивление катушки электромагнита;
 P_c — противодействующая сила;
 δ_0 — начальный зазор;
 w — число витков катушки электромагнита;
 S — площадь сечения якоря электромагнита;
 m — масса якоря и других подвижных частей.

Назовем эти величины *основными параметрами*, оперируя которыми, можно рассмотреть вопросы динамики без учета второстепенных факторов, связанных с конкретной конструкцией электромагнита.

В общем случае динамика электромагнитов, как известно, характеризуется следующими двумя уравнениями:

$$U = iR + \frac{d\psi}{dt}; \quad (1)$$

$$P_s = m \frac{d^2x}{dt^2} + P_c \left(x, \frac{dx}{dt} \right), \quad (2)$$

где x — путь, проходимый якорем электромагнита;

P_c — сила противодействия, зависящая в общем случае как от положения якоря, так и от его скорости;

P_s — тяговая сила электромагнита.

С учетом принятых допущений эти уравнения после ряда преобразований можно представить в следующем виде:

$$U = iR + L \frac{di}{dt} + i \frac{L}{\delta} \frac{dx}{dt}; \quad (3)$$

$$P_s = m \frac{d^2x}{dt^2} + P_c, \quad (4)$$

где индуктивность электромагнита

$$L = \frac{\mu_0 S}{\delta} w^2, \quad (5)$$

i , x , δ , L , P_s — текущие значения величин, соответствующих определенному моменту времени в процессе движения якоря.

При этом формулу тяговой силы электромагнита¹

$$P_s = -\frac{1}{2} i^2 \frac{dL}{d\delta} \quad (6)$$

весьма важно представить в таком виде:

$$P_s = \frac{1}{2} i^2 \frac{L}{\delta}. \quad (7)$$

¹ Принята абсолютная практическая рационализованная система.

Уравнения (3) и (4) характеризуют процесс как во время движения якоря, так и до его трогания. В последнем случае эти уравнения примут вид

$$U = iR + L_0 \frac{di}{dt}, \quad (8)$$

$$\frac{1}{2} I^2 \frac{L_0}{\delta_0} \leq P_c, \quad (9)$$

где L_0 — индуктивность при начальном зазоре.

В момент трогания якоря, очевидно, ток равен

$$i_{\text{тр}} = \sqrt{\frac{2P_c \delta_0}{L_0}}. \quad (10)$$

Формула (7) показывает, что тяговая сила электромагнита при прочих равных условиях пропорциональна индуктивности. Следовательно, с увеличением индуктивности при постоянной противодействующей силе [см. уравнение (4)] соответственно увеличивается «живая сила», т. е. уменьшается время срабатывания. С другой стороны, увеличение индуктивности, как известно, ведет к уменьшению величины тока в каждый данный момент времени. Уменьшение тока же соответственно ведет к уменьшению тяговой силы электромагнита, т. е. к увеличению времени срабатывания.

Таким образом, приходим к важному выводу, что существует некоторая *оптимальная индуктивность*, которая обеспечивает *минимальное время срабатывания*.

К этому же выводу можно прийти и другим путем. Уравнение (10) показывает, что с уменьшением индуктивности ток трогания увеличивается. Положим, что индуктивность уменьшена до такой величины, при которой ток трогания становится равным установившемуся значению. При этом, конечно, электромагнит не срабатывает, ибо тяговая сила его не превышает противодействующей силы, а равна ей, т. е. электромагнит находится «на пределе своих возможностей». Режим этот характеризуется равенством

$$P_c = \frac{1}{2} I_y^2 \frac{L_0}{\delta_0}, \quad (11)$$

где I_y — установившийся ток.

Этому режиму соответствует определенная *минимальная индуктивность*

$$L_0^{\min} = \frac{2P_c \delta_0}{I_y^2}. \quad (12)$$

При этой индуктивности электромагнит не сработает (время срабатывания $t_{\text{ср}} = \infty$).

Электромагнит начинает работать при малейшем увеличении этой индуктивности. Для любого конечного значения индуктивности время срабатывания, очевидно, будет иметь конечное значение.

Для случая
будем иметь

$$L_0 = \infty$$

$$t_{\text{ср}} = \infty.$$

Это видно из уравнения (8); при бесконечной индуктивности вполне можно пренебречь активным сопротивлением и тогда

$$i = \frac{U}{L_0} t. \quad (13)$$

Следовательно, для любого конечного промежутка времени ток равен нулю, т. е. равна нулю и тяговая сила электромагнита.

Итак, в функции

$$t_{\text{ср}} = f(L_0)$$

предельные значения аргумента L_0 обращают эту функцию в бесконечность. Естественно предположить, что функция проходит через некоторое минимальное значение. Индуктивность, при которой время срабатывания имеет минимальное значение, назовем *оптимальной*.

Для оптимального значения индуктивности начальный зазор электромагнита предполагаем заданным и вполне определенным. Следовательно, получаем оптимальные значения площади сечения якоря и числа витков катушки [см. формулу (5)]. В самом деле, из всех величин, входящих в исходные уравнения динамики (3 и 4), только две величины — тяговая сила электромагнита и индуктивность — включают в себя число витков и сечение якоря.

В формулу тяговой силы (7) и в выражение индуктивности (5) входят квадрат числа витков и первая степень сечения якоря.

При условии, что зазор является заданной величиной можно пользоваться уже известным понятием индуктивности и искать ее оптимальное значение.

Известно, что время срабатывания состоит из времени трогания и времени движения

$$t_{\text{ср}} = t_{\text{тр}} + t_{\text{дв}}. \quad (14)$$

Найдем оптимальное значение индуктивности для минимального времени трогания, т. е. времени с момента приложения напряжения к обмотке электромагнита до момента трогания якоря.

Из уравнения (8) легко можно вывести известную формулу времени трогания

$$t_{\text{тр}} = \frac{L_0}{R} \ln \frac{I_y}{I_y - i_{\text{тр}}}. \quad (15)$$

Пусть заданы:

- 1) потребляемая мощность из сети UI_y ;
- 2) полезная работа, производимая электромагнитом, $P_c \delta_0$;
- 3) напряжение сети U .

Требуется определить оптимальные условия для получения минимального времени трогания.

Подставим в уравнение (15) значение тока трогания (10):

$$t_{\text{тр}} = \frac{L_0}{R} \ln \frac{I_y}{I_y - \sqrt{\frac{2P_c \delta_0}{L_0}}}. \quad (16)$$

В последнем уравнении все величины, за исключением индуктивности, являются заданными. Из этой формулы видно, что с увеличением индуктивности время трогания, с одной стороны, увеличивается за счет увеличения постоянной времени, а с другой, — уменьшается за счет уменьшения тока трогания; при уменьшении индуктивности происходит обратное явление. Значит, действительно, можно искать оптимальную индуктивность.

В соответствии с выражением (12) при минимальной индуктивности получим

$$t_{\text{тр}} = \infty.$$

При $L_0^{\text{max}} = \infty$ получается неопределенность вида

$$t_{\text{тр}} = \infty \cdot 0.$$

Раскрывая неопределенность по правилу Лопиталья, приходим к ранее полученному результату

$$\lim_{L \rightarrow \infty} t_{\text{тр}} = \lim_{L \rightarrow \infty} \left(\frac{\sqrt{2P_c \delta_0}}{U} \sqrt{L} \right) = \infty.$$

Для нахождения оптимальной индуктивности возьмем первую производную от времени трогания (16) по индуктивности и приравняем ее нулю. Опуская все промежуточные операции, напишем полученное в итоге выражение оптимальной индуктивности:

$$L_{\text{тр}}^{\text{опт}} = \frac{3,92P_c \delta_0}{I_y^2}. \quad (17)$$

Сравнивая это выражение с выражением (12), нетрудно заметить, что:

$$L_{\text{тр}}^{\text{опт}} \approx 2L_0^{\text{min}}. \quad (18)$$

Если подставить выражение оптимальной индуктивности (17) в формулу (16), то получим искомую формулу минимального времени трогания:

$$t_{\text{тр}}^{\text{min}} = \frac{4,92P_c \delta_0}{I_y^2 R} = \frac{4,92P_c \delta_0}{UI_y}, \quad (19)$$

где P_c выражена в ньютонах, δ_0 — в метрах.

Формула (19) показывает, что минимальное время трогания зависит от подводимой мощности из сети и производимой электромаг-

нитом полезной работы. Для практических расчетов эту формулу удобнее представить в следующем виде:

$$t_{\text{тр}}^{\text{min}} = \frac{0,483 P_c \delta_0}{U I_y} \approx \frac{0,5 P_c \delta_0}{U I_y}, \quad (20)$$

где P_c выражена в $\kappa\Gamma$, δ_0 — в см .

Необходимо отметить, что при известной конструкции электромагнита можно найти оптимальную индуктивность и минимальное время трогания с учетом потоков рассеяния, выпучивания и магнитного сопротивления любых паразитных зазоров. Для этого необходимо рассчитать магнитную цепь при начальном зазоре электромагнита, найти необходимые коэффициенты, учитывающие все перечисленные факторы, и ввести их в исходное уравнение времени трогания (16). Затем аналогичным образом можно найти выражения оптимальной индуктивности и минимального времени трогания.

Если в электромагните имеется немагнитная прокладка, т. е. начальный воздушный зазор более рабочего хода якоря, то в формулах (17) и (19) должен фигурировать начальный зазор.

В практике могут встречаться случаи, когда противодействующая сила в зависимости от хода якоря изменяется по самым разнообразным законам. В этих случаях в формулы (17) и (19) следует подставлять начальные значения противодействующей силы; в последней можно учесть также и силы трения покоя.

Если имеется определенная конструкция электромагнита и задана площадь сечения якоря, то из выражений (5) и (17) легко найти оптимальное число витков для минимального времени трогания. Так, для электромагнита с одним рабочим зазором выражение для оптимального числа витков представится в следующем виде:

$$w_{\text{тр}}^{\text{опт}} = \frac{1,98 \delta_0}{I_y} \sqrt{\frac{P_c}{\mu_0 S}}. \quad (21)$$

Введением соответствующих расчетных коэффициентов здесь также можно учесть потоки рассеяния и выпучивания.

Если подставить выражение оптимальной индуктивности (17) в формулу тока трогания (10), то получим

$$i_{\text{тр}} = 0,714 I_y. \quad (22)$$

Это означает, что при правильно подобранном значении оптимальной индуктивности для минимального времени трогания ток трогания должен составлять 71% от установившегося тока.

Тяговую силу, развиваемую электромагнитом при начальном зазоре и при токе, равном установившемуся значению, назовем *максимальной тяговой силой* электромагнита:

$$P_s^{\text{max}} = \frac{1}{2} I_y^2 \frac{L_0}{\delta_0}. \quad (23)$$

Если подставим в это выражение значение оптимальной индуктивности (17), то получим

$$P_s^{\max} = 1,96P_c, \quad (24)$$

т. е. в оптимальном режиме противодействующая сила почти вдвое меньше максимальной тяговой силы электромагнита.

Определив оптимальные соотношения для минимального времени трогания, нетрудно заметить, что подобные соотношения можно получить и для времени движения, если иметь аналитическую зависимость времени движения от всех основных параметров. Эти оптимальные соотношения для минимального времени трогания и движения являются основой для разработки общих методов расчета быстродействующих электромагнитов. Но все эти вопросы вполне самостоятельны и требуют специального рассмотрения.

Наши выводы о существовании оптимальной индуктивности, в частности, оптимального числа витков [см. (21)], позволяют указать на следующее важное обстоятельство. Понятие «ампер-витки» никак не позволительно применять при рассмотрении вопросов динамики электромагнита, так как сила тока и число витков не всегда играют одинаковую роль. В самом деле, если с увеличением установившегося значения тока время срабатывания всегда уменьшается, то с увеличением числа витков за пределы оптимального значения время срабатывания увеличивается. Понятие «ампер-витки» имеет смысл лишь при рассмотрении статических тяговых характеристик.

Канд. техн. наук А. К. ТЕР-АКОПОВ

АНАЛИЗ МЕТОДОВ РАСЧЕТА ДИНАМИЧЕСКИХ ХАРАКТЕРИСТИК ЭЛЕКТРОМАГНИТНЫХ МЕХАНИЗМОВ

ОБЩИЕ СВЕДЕНИЯ И ЗАДАЧИ РАБОТЫ

Главным назначением любого электромагнитного механизма является создание движения, поэтому полную характеристику его могут дать именно динамические характеристики. Наиболее важной в практическом отношении характеристикой является время движения подвижного органа (якоря) электромагнита. Для полного определения динамического режима достаточно иметь две зависимости:

$$i=f(t) \text{ и } x=f(t),$$

где i — ток в обмотке электромагнита во время движения его якоря;
 x — путь, проходимый якорем.

По этим так называемым *главным динамическим характеристикам* можно найти любые другие зависимости, как, например,

$$P_s=f(t) \text{ и } \psi=f(t),$$

где P_s — тяговая сила электромагнита во время движения якоря;
 ψ — текущее значение потокосцепления.

Расчет динамических характеристик является наиболее трудной частью общего расчета всякого электромагнита. Основная трудность заключается в том, что в процессе движения якоря меняется магнитная проводимость воздушного зазора. Это ведет к изменению магнитного потока, поэтому, кроме ЭДС самоиндукции, появляется еще так называемая ЭДС движения. Последняя, в свою очередь, также ведет к изменению тока, отсюда — к изменению силы, развиваемой электромагнитом. От этой силы, собственно, зависит закон движения якоря, т. е., в конечном итоге, закон изменения магнитной проводимости воздушного зазора. Из сказанного видно, что электрические процессы в обмотке электромагнита находятся в тесной и сложной взаимосвязи с механическими процессами, обусловленными движением якоря.

В общем случае динамика электромагнитов характеризуется следующими двумя уравнениями:

$$U = iR + \frac{d\psi}{dt}; \quad (1)$$

$$P_0 = m \frac{d^2x}{dt^2} + P_c \left(x, \frac{dx}{dt} \right), \quad (2)$$

где R — сопротивление катушки электромагнита;
 m — масса подвижных частей электромагнита;
 P_0 — сила противодействия, зависящая в общем случае как от положения якоря, так и от его скорости.

Эти уравнения нелинейные и точное решение их обычными методами не представляется возможным. Если даже пренебречь силами противодействия (P_0), не говоря уже о пренебрежении такими второстепенными факторами, как рассеяние, насыщение стали, явление гистерезиса, вихревые токи и др., то все равно уравнения эти не поддаются точному решению.

В силу изложенного выше различные авторы предлагают самые разнообразные методы расчета динамических характеристик, отличающиеся друг от друга как подходом к разрешению задачи, так и принятыми допущениями. В зависимости от соответствия допущений, принятых в том или ином методе расчета, реальным условиям работы электромагнита, расчетные данные могут иногда в значительной степени отличаться от опытных. Тем не менее, большинством авторов не даются конкретные условия применения предлагаемых ими методов; иногда о некоторых допущениях вообще не упоминается. Более того, существуют методы расчета, которые никак нельзя рекомендовать для практического использования.

Однако в литературе почти отсутствуют какие-либо критические замечания по существующим методам расчета, нет конкретных указаний на условия, при которых целесообразно применение того или иного метода. Это в значительной степени затрудняет их практическое использование.

В настоящей работе сделана попытка проанализировать существующие методы расчета и определить, при каких условиях возможно практическое использование каждого из них. Кроме того, ставится задача выяснить, возможно ли аналитическое определение оптимальной индуктивности, существование которой установлено в предыдущей статье, на основании известных методов расчета времени движения якоря электромагнитных механизмов.

КРАТКИЙ ОБЗОР МЕТОДОВ РАСЧЕТА, ПРЕДЛОЖЕННЫХ РАЗЛИЧНЫМИ АВТОРАМИ

Несмотря на большое разнообразие методов расчета динамических характеристик, их можно разделить на две группы.

К первой группе отнесем методы, по которым весь процесс динамики разделяется на ряд участков; такое разделение позволяет на

каждом участке усреднить рассматриваемые величины. Каждый такой участок связывается, конечно, с предыдущим и последующим участком. Методы эти могут быть и графическими, и графо-аналитическими, и аналитическими. В них еще иногда используется принцип последовательных приближений. Отличительным признаком этих методов является то, что они не дают общих математических выражений для искомых динамических характеристик. Назовем эти методы *методами последовательных интервалов*.

Ко второй группе отнесем методы, дающие общие математические выражения динамических характеристик. Поскольку дифференциальные уравнения динамики, как мы уже говорили, не поддаются точному математическому решению, то эти методы основаны или на существенных допущениях или на некоторых упрощениях реального процесса динамики. Назовем эти методы *аналитическими*.

Методы последовательных интервалов ¹

Метод Шюлера ². Метод основан на том, что время движения разбивается на большое число интервалов, а величины скорости, воздушного промежутка, индукции и тока рассчитываются от одного интервала времени к другому. Изменения всех этих величин на протяжении какого-либо интервала времени учитываются в последующих интервалах.

При расчете пренебрегают потоками рассеяния, вихревыми токами и силами трения. Кроме того, Шюлер предполагает, что индукция одинакова по торцу сердечника.

Исходными являются следующие уравнения:

$$B_2 = \frac{et \cdot 10^8}{w_S} + B_1; \quad (3)$$

$$P_s = \left(\frac{B}{5000} \right)^2 S; \quad (4)$$

$$\Delta v = \frac{\Delta Pt \cdot 9,81}{G}, \quad (5)$$

где

B_1 и B_2 — индукции в начале и в конце интервала ($гс$);
 $e = U - i_{cp} R$ — средняя индуцированная ЭДС ($в$);
 G — общий вес подвижных частей ($кг$);
 $\Delta P = \frac{P_{s1} + P_{s2}}{2} - P_c$ — средний избыток развиваемой силы над силой противодействия;
 P_{s1} и P_{s2} — силы, развиваемые якорем в начале и в конце интервала времени ($кг$);
 Δv — приращение скорости за данный интервал времени.

¹ В описании всех методов расчета применены одни и те же буквенные обозначения. Системы единиц, принятые авторами методов расчета, сохранены.

² Schüler, Die Wirkungsgrad des Elektromagneten, ETZ, S. 613, 1913.

Уравнение (3) получено из закона электромагнитной индукции, выраженного в конечных разностях. Уравнение (4) — известная формула Максвелла. Шюлер пользуется своим методом для определения общего времени срабатывания, включая и время трогания.

Для первого интервала вначале принимается ЭДС самоиндукции, равная приложенному напряжению, поэтому уравнение (3) будет выглядеть так:

$$B_2 = \frac{Ut \cdot 10^8}{wS}.$$

Определив индукцию, находят затем требуемые ампер-витки для воздушного промежутка и для стали (если система насыщена). По найденным ампер-виткам определяют ток в конце первого интервала.

Средний ток в первом интервале

$$i_{\text{ср}} = \frac{0 + i_2}{2} = \frac{i_2}{2}.$$

Отсюда средняя индуктированная ЭДС

$$e = U - i_{\text{ср}}R.$$

По ней вторично находят индукцию, а затем и окончательное значение тока в конце первого интервала также посредством определения общих ампер-витков. Найденную величину тока наносят на координатную систему $i=f(t)$ и полученную точку соединяют с началом координат. Методом экстраполяции продолжая эту линию, находят среднее значение тока для второго интервала. Затем аналогичный расчет ведут для второго и последующих интервалов. После того как сила, развиваемая электромагнитом, превзойдет силы противодействия, расчет уже ведут для времени движения.

В общем случае предлагаемый Шюлером ход расчета следующий.

По найденному методом экстраполяции среднему току данного участка определяют

$$e = U - i_{\text{ср}}R.$$

Затем по уравнениям (3), (4) и (5) находят приращение скорости и конечную скорость на данном интервале:

$$v_2 = v_1 + \Delta v.$$

Путь, пройденный якорем:

$$\Delta x = v_{\text{ср}}t,$$

где

$$v_{\text{ср}} = \frac{v_1 + v_2}{2}.$$

Затем находят величину воздушного промежутка в конце этого интервала, а по ней и необходимые ампер-витки по воздуху и стали:

$$i\omega = i\omega_{\text{в}} + i\omega_{\text{ст}}.$$

По найденным ампер-виткам определяют ток в конце этого интервала. Дальнейший расчет проводят аналогичным образом.

Метод А. И. Москвитина¹. Как указывает автор, формула Максвелла, применяемая Шюлером для расчета тяговой силы, справедлива только при движении якоря по направлению силовых линий и только при небольших воздушных зазорах. Поэтому А. И. Москвитин предлагает пользоваться другой формулой:

$$P_s = \frac{1}{2} \frac{\Phi^2}{G} \frac{dG}{dx}, \quad (6)$$

где G — магнитная проводимость воздушных промежутков. Заметим, что формула эта справедлива для ненасыщенных систем. Очевидно, для пользования этой формулой необходимо предварительно построить зависимости

$$G = f(x) \quad \text{и} \quad \frac{dG}{dx} = \varphi(x).$$

В остальном расчет не отличается от метода, предложенного Шюлером.

Метод Б. С. Сотскова². Здесь, как указано автором, используется метод численного интегрирования. В отличие от предыдущих двух методов расчет ведется только для времени движения. Для определения тяговой силы Б. С. Сотсков использует ту же формулу, что и А. И. Москвитин, в следующей записи:

$$P_s = \frac{(0,4\pi i \omega)^2}{8\pi} \frac{dG}{dx}. \quad (7)$$

Однако, в отличие от А. И. Москвитина, Б. С. Сотсков предлагает начинать расчет с определения приращения силы тока при уже известном токе трогания.

Расчет ведется также по интервалам времени.

Для какого-либо интервала имеем

$$i_k = i_{k-1} + \left(\frac{di}{dt} \right)_{k-1} \Delta t_k; \quad (8)$$

$$\left(\frac{di}{dt} \right)_k = \frac{1}{L_k} \left\{ U - i_k \left[R + \left(\frac{dL}{dx} \right)_k \left(\frac{dx}{dt} \right)_{cp k} \right] \right\}, \quad (9)$$

где

$$L_k = L_{k-1} + \left(\frac{dL}{dx} \right)_{k-1} \Delta x_{k-1}. \quad (10)$$

В остальном ход расчета аналогичен предыдущим методам, за исключением того, что Б. С. Сотсков в уравнении динамики якоря

¹ А. И. Москвитин, Основные проблемы электрических машин прямолинейного движения, «Электричество», 1941, № 2.

² Б. С. Сотсков, Элементы автоматической и телемеханической аппаратуры, ГЭИ, 1950.

учитывает силы сопротивления, зависящие как от положения якоря, так и от его скорости. Нетрудно заметить, что расчет по этому методу требует предварительного определения зависимости $L=f(x)$.

Метод Н. Е. Лысова¹. Данный метод существенно отличается от методов, рассмотренных выше. Н. Е. Лысов разбивает процесс не по интервалам времени, а по интервалам хода якоря. Далее, уравнение динамики рассматривается не на основе баланса сил, а на основе энергетического баланса. В связи с этим в методе непосредственно не ведется расчет тяговой силы; здесь гораздо удобнее определять механическую работу, совершаемую якорем на каком-либо участке хода.

Расчетные уравнения представляются в конечных разностях:

$$U = i_{cp} R + \frac{\Delta \psi}{\Delta t}; \quad (11)$$

$$P_s \Delta x = \Delta \left(\frac{mv^2}{2} \right) + P_c \Delta x. \quad (12)$$

В общем случае Н. Е. Лысов предлагает решать эти уравнения графо-аналитическим методом последовательных приближений. Расчет ведется в следующей последовательности.

Ход якоря разбивается на ряд участков, число которых выбирается в зависимости от конкретных условий; во всяком случае, должно быть не менее 5—6 интервалов. Далее, для всех положений якоря расчетом магнитной цепи строится семейство кривых $\psi=f(l)$ (фиг. 1).

Значению тока трогания соответствует точка a на начальной кривой. Из этой точки проводится прямая ab до пересечения с соседней кривой так, чтобы потокосцепление увеличивалось. Как известно, заштрихованная площадь представляет собой механическую работу, совершаемую якорем на этом интервале. Зная эту работу, а также работу на преодоление противодействующих сил, нетрудно из уравнения (12) найти приращение скорости якоря на данном интервале.

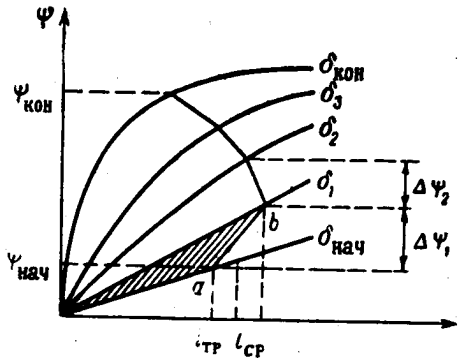
Для любого интервала имеем

$$v_{кон} = \sqrt{(P_s \Delta x - P_c \Delta x) \frac{2}{m} + v_{нач}^2}; \quad (13)$$

$$v_{cp} = \frac{v_{нач} + v_{кон}}{2}; \quad \Delta t = \frac{\Delta x}{v_{cp}},$$

где $v_{кон}$, $v_{нач}$, v_{cp} — соответственно конечная, начальная и средняя скорости на данном интервале.

¹ Н. Е. Лысов, Расчет электромагнитных механизмов, Оборонгиз, 1949; К расчету динамических характеристик электромагнитов, «Вестник электропромышленности», 1941, № 3.



Фиг. 1. К расчету динамических характеристик методом Н. Е. Лысова.

Найдя время движения, а по графику — приращение потоко-сцепления и средний ток на данном интервале (см. фиг. 1), следует проверить, удовлетворяется ли равенство (11). Если равенство это не удовлетворяется, то следует задаться другим направлением отрезка ab и весь расчет повторить сначала, и так до тех пор, пока равенство (11) не будет удовлетворено. При некотором навыке можно правильное направление этого отрезка выбрать, если не с первого, так обязательно со второго или третьего раза.

Метод этот является графо-аналитическим и пригоден при любой форме кривых $\psi = f(i)$. Если же эти кривые имеют прямолинейный характер, то Н. Е. Лысов предлагает вести расчет чисто аналитическим путем без каких-либо графических построений. При этом заштрихованная площадь на всех интервалах представляется в виде треугольника, и весь расчет сводится к простым алгебраическим действиям. При таком расчете учет потоков рассеяния Н. Е. Лысов предлагает производить введением соответствующих коэффициентов.

В остальном весь ход расчета аналогичен изложенному графо-аналитическому методу.

Метод Н. А. Лившица¹. Основой своего метода автор считает положение, что скорость и длительность перехода подвижных органов электромагнитного механизма определяются значением тяговых усилий, развиваемых потоком в рабочих воздушных зазорах, и тормозящих усилий, преодолеваемых подвижными органами механизма.

Для расчета необходимо иметь тяговую и нагрузочную характеристики электромагнита.

Для двух частных случаев относительного расположения этих характеристик Н. А. Лившиц дает аналитическое решение, которое мы рассмотрим в следующем разделе. Здесь же рассмотрим общее решение при любых законах изменения тяговой и нагрузочной характеристик. В этом случае для определения времени движения автор применяет метод двойного графического интегрирования. Как будет видно из дальнейшего рассмотрения, зависимости пути и скорости от времени строятся последовательно по отдельным точкам. Поэтому в некоторой степени мы имели право отнести этот метод к группе методов последовательных интервалов.

Тяговая характеристика строится из расчета магнитной цепи при постоянной силе тока, т. е. в статическом режиме. Вначале автор рассматривает случай, когда можно пренебречь влиянием сил трения. При этом уравнение движения представится так:

$$P_g - P_c = m \frac{d^2x}{dt^2}. \quad (14)$$

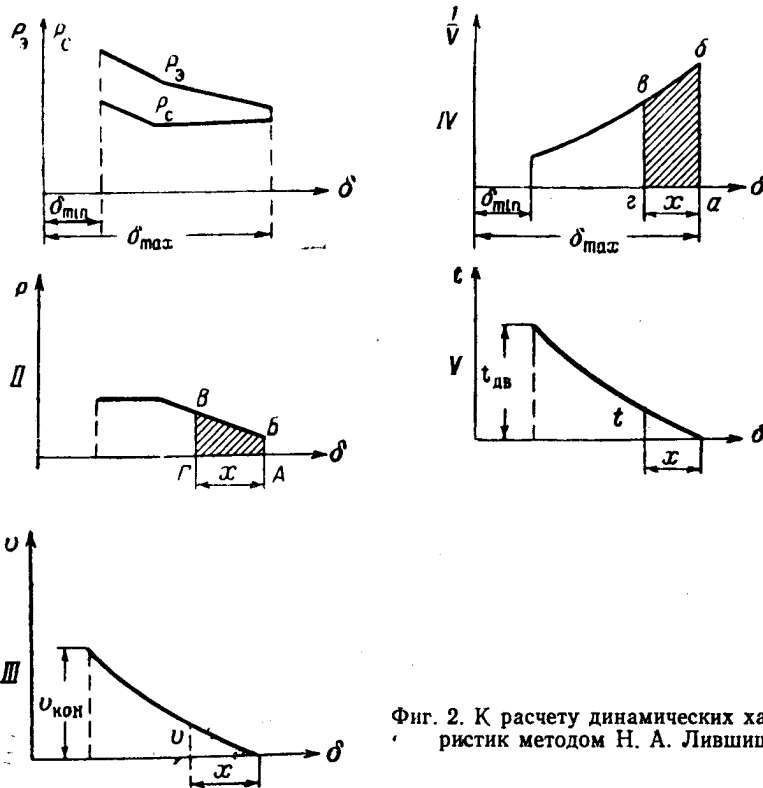
На фиг. 2 построены исходные характеристики: тяговая и нагрузочная. Ход расчета следующий.

¹ Н. А. Лившиц, К вопросу об анализе времени движения якоря электромагнитных механизмов, «Автоматика и телемеханика», 1939, № 2; Определение времени движения подвижных органов электромагнитного механизма при его срабатывании, «Известия электропромышленности слабого тока», 1940, № 11.

Из уравнения (14) непосредственно получаем

$$\frac{mv^2}{2} = \int_{\delta_{\max}}^{\delta} P d\delta = \text{площадь } ABVG,$$

где $P = P_s - P_c$ построена графически на основании исходных характеристик.



Фиг. 2. К расчету динамических характеристик методом Н. А. Лившица.

Из последнего уравнения графическим интегрированием определяют скорость для ряда последовательных положений якоря и строят кривую. Зависимость времени движения для каждого положения якоря определяют также графическим интегрированием по уравнению

$$t = \int_{\delta_{\max}}^{\delta} \frac{1}{v} d\delta = \text{площадь } abvg.$$

Но до этого предварительно надо построить зависимость

$$\frac{1}{v} = f(\delta).$$

Все графические построения в данной последовательности показаны на фиг. 2.

Для вращательного движения Н. А. Лившиц приводит аналогичное решение.

В случае, если надо учитывать еще силы трения, зависящие от скорости движения, автор предлагает в общем случае применять двойное графическое интегрирование по методу изоклин. Изложение этого метода в настоящей работе не представляется необходимым.

Аналитические методы

Метод Яссе¹. Автор пытается дать аналитическое решение динамики электромагнита, но так как такое решение, по мнению автора, представляет значительные трудности, то автор делает ряд допущений, которые сводятся в основном к следующему:

- 1) силы противодействия равны нулю;
- 2) магнитная система не насыщена;
- 3) индуктивность представляет линейную функцию от хода, т. е.

$$L = L_0 \sigma; \quad \frac{dL}{dx} = L_0 \frac{d\sigma}{dx} = L_0 \alpha,$$

где α — постоянная величина;

- 4) ток в процессе движения мало изменяет свою первоначальную величину;

- 5) скорость является сравнительно малой величиной.

Ток представляется в виде суммы

$$i = I + i',$$

где I — установившееся значение тока;

i' — переменное слагаемое тока.

Яссе считает i' и v (скорость) величинами малого порядка и при решении дифференциальных уравнений пренебрегает квадратами каждой из этих величин или их произведением.

Исходные уравнения Яссе приводит в следующем виде:

$$U = iR + L \frac{di}{dt} + iv \frac{dL}{dx}; \quad (15)$$

$$m \frac{dv}{dt} = \frac{1}{2} i^2 \frac{dL}{dx}. \quad (16)$$

Решив эти уравнения с указанным выше пренебрежением малых величин второго порядка, Яссе дает окончательные уравнения:

$$i' = \frac{I}{2} \left(e^{-\gamma t} \cos \gamma t - 1 + \frac{\eta}{\gamma} e^{-\gamma t} \sin \gamma t \right); \quad (17)$$

$$v = \frac{\gamma \sigma}{\alpha} \left(1 - e^{-\gamma t} \cos \gamma t + \frac{\gamma^2 - \gamma_1^2}{2\gamma_1} e^{-\gamma t} \sin \gamma t \right), \quad (18)$$

¹ Э. Яссе, Электромагниты, ГЭИ, 1934.

где η , γ , σ , α — постоянные величины, зависящие от основных параметров электромагнита.

Из уравнений (17) и (18) видно, что ток стремится уменьшиться до половины установившегося значения, а скорость якоря стремится к постоянной величине.

В заключение Яссе подчеркивает, что это решение справедливо для малых скоростей якоря и при малых изменениях первоначального значения тока.

Метод А. И. Москвитина¹. Автор подчеркивает важную роль переходных процессов в современных быстродействующих электромагнитах постоянного тока. Поэтому им «была поставлена цель дать достаточно простой аналитический метод исследования, позволяющий установить зависимости в общем виде».

Вначале автор рассматривает идеальный случай, делая следующие допущения:

- 1) активное сопротивление обмотки равно нулю;
- 2) противодействующие силы равны нулю;
- 3) магнитная система не насыщена;
- 4) потоки рассеяния отсутствуют;
- 5) явление выпучивания отсутствует;
- 6) вихревыми токами и явлением гистерезиса пренебрегает.

Ввиду того, что противодействующие силы приравнены к нулю, время трогания вовсе отсутствует и общее время срабатывания определяется лишь временем движения.

При этих допущениях А. И. Москвитин представляет уравнения динамики в следующем виде:

$$\frac{d\Phi}{dt} = k_1 U; \quad (19)$$

$$\frac{d^2 x}{dt^2} = k_2 \Phi^2; \quad (20)$$

$$i = k_3 (\delta_0 - x) \Phi, \quad (21)$$

где δ_0 — начальный зазор;
 k_1, k_2, k_3 — постоянные коэффициенты.

Уравнения эти после указания всех допущений, очевидно, не требуют пояснений; решение их осуществляется непосредственно:

$$\Phi = kt; \quad (22)$$

$$x = qt^4; \quad (23)$$

$$i = \alpha t - \beta t^5, \quad (24)$$

где k, q, α, β — постоянные коэффициенты.

¹ А. И. Москвитин, Уравнения процессов в электромагните с движущимся якорем. «Известия Академии Наук СССР», 1948, № 4; Электрические машины возвратно-поступательного движения, изд. АН СССР, 1950.

Полученные уравнения, собственно, и есть искомые динамические характеристики. Из этих уравнений можно получить наиболее интересные параметры, характеризующие динамический процесс.

Время срабатывания может быть определено как из уравнения (23), так и (24):

$$t_{\text{ср}} = \sqrt[4]{\frac{\delta_0}{q}} = \sqrt[4]{\frac{\alpha}{\beta}}. \quad (25)$$

Максимальное значение тока определяется из уравнения (24):

$$I_m = \frac{4}{5} \alpha \sqrt[4]{\frac{\alpha}{5\beta}}. \quad (26)$$

А. И. Москвитин подчеркивает, что указанные идеальные условия в значительной степени соответствуют действительности в *быстродействующих электромагнитах*. Тем не менее, после этого решения автор еще учитывает ряд факторов, которыми вначале пренебрегал.

Потоки рассеяния и явление выпучивания А. И. Москвитин учитывает введением в исходные уравнения соответствующего коэффициента. Идея заключается в том, что по картине поля, построенной для начального зазора, рассчитывается магнитная цепь с учетом потоков рассеяния и выпучивания, затем определяется действительный поток, проходящий через заданное сечение воздушного зазора. Отношение этого потока к потоку, рассчитанному без учета рассеяния и выпучивания, определит искомый поправочный коэффициент. Этот коэффициент вводится в исходные дифференциальные уравнения, которые уже будут правильно отражать реальную картину поля при начальном зазоре электромагнита.

Однако при движении якоря воздушный зазор будет уменьшаться; соответственно будет изменяться картина поля и найденный поправочный коэффициент. Очевидно, при притянutom якоря он будет примерно равен единице. Следовательно, надо было бы учесть изменение этого коэффициента во время хода якоря, что связано с большими трудностями. Оказывается, что такой учет практически и не требуется, ввиду, как удачно автор выразился, «одного счастливого обстоятельства». Оно заключается в том, что в течение большей части времени «якорь почти не движется», т. е. скорость движения весьма мала. Поэтому с достаточной точностью можно считать этот коэффициент постоянным:

$$\varphi = \text{const.}$$

С учетом этого коэффициента, а также активного сопротивления обмотки уравнения примут вид

$$\frac{d\Phi}{dt} = k_1 (U - iR); \quad (27)$$

$$\frac{d^2x}{dt^2} = \varphi^2 k_2 \Phi^2; \quad (28)$$

$$i = k\varphi (\delta_0 - x) \Phi. \quad (29)$$

Автор справедливо замечает, что «точное аналитическое решение этих уравнений невозможно».

Для разрешения этой задачи А. И. Москвитин предлагает считать в уравнении (27) активное падение напряжения постоянным за все время процесса, поставив вместо него среднее значение:

$$i_{\text{cp}} = \frac{1}{t_{\text{cp}}} \int_0^{t_{\text{cp}}} i dt.$$

Тогда уравнение (27) примет вид

$$\frac{d\Phi}{dt} = k_1 e_{\text{cp}}$$

и система уравнений динамики разрешится так же просто, как и уравнения (19), (20) и (21).

Все затруднение теперь заключается в определении средней ЭДС. Однако А. И. Москвитин обходит это затруднение искусственным путем. Дело в том, что максимальное значение тока, подсчитанное по уравнению (26) без учета активного сопротивления, в некоторых случаях в несколько раз превышает установившееся значение

$$I_y = \frac{U}{R},$$

что реально никак невозможно.

После ряда выкладок А. И. Москвитин приходит к выводу, что отношение максимального значения тока с учетом активного сопротивления к максимальному значению тока без учета последнего зависит от известных нам величин, т. е.

$$\frac{I'_m}{I_m} = C,$$

где I'_m — максимальное значение тока с учетом активного сопротивления обмотки;

I_m — то же, но без учета активного сопротивления, т. е. найденное по формуле (26);

$C = f\left(\frac{I'_m}{I_y}\right)$ — расчетный коэффициент.

Последнюю зависимость автор представляет в виде кривой (фиг. 3). Таким образом, найдя I'_m и I_y , мы по кривой определяем коэффициент C , а по нему и искомое амплитудное значение I'_m .

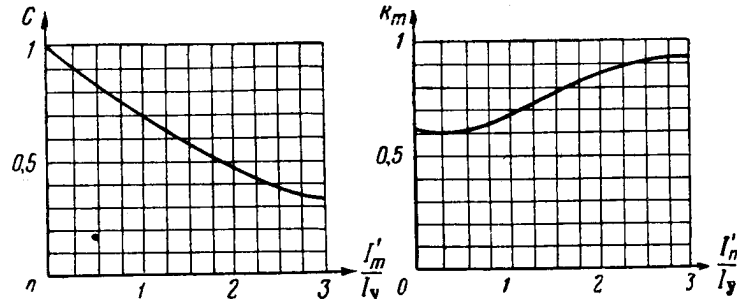
Средняя ЭДС определяется так:

$$e_{\text{cp}} = U - i_{\text{cp}} R = U - k_m I'_m R,$$

где $k_m = F\left(\frac{I'_m}{I_y}\right)$ определяется по кривой, также приведенной А. И. Москвитиным (см. фиг. 3).

Теперь можно легко решить уравнения (27) ÷ (29), подставив в первое уравнение вместо $(U - iR)$ найденное значение $e_{\text{ср}}$. Все расчетные уравнения получатся такими же, только вместо напряжения сети будет фигурировать $e_{\text{ср}}$.

В заключение автор подчеркивает, что данный метод является практически точным, когда противодействующие усилия много меньше притягивающего усилия электромагнита, и что в приложении к таким механизмам, как реле, пускатели и т. п., этот метод можно применять для ориентировочных расчетов.



Фиг. 3. Кривые расчетных коэффициентов для расчета динамических характеристик по аналитическому методу А. И. Москвитина.

Далее А. И. Москвитин приводит уравнение динамики с учетом противодействующих сил, хотя решение этого уравнения не дается.

Метод Б. С. Сотскова¹. Автор предлагает приближенный аналитический способ решения уравнений динамики с учетом всех основных параметров.

Для решения автор представляет уравнения в следующем виде:

$$U = iR + \omega \frac{d\Phi}{dt} 10^{-8}; \quad (30)$$

$$i = i_m + b\Phi(\delta_0 - x); \quad (31)$$

$$\frac{d^2x}{dt^2} + r \frac{dx}{dt} + kx + k_0 = m\Phi^2, \quad (32)$$

где δ_0 — начальный зазор;

b, r, k, k_0, m — постоянные величины, определяемые заданными параметрами;

$i_m \approx \text{const}$ — эквивалентное значение тока, учитывающего потерю намагничивающей силы в магнитопроводе.

Из последнего уравнения видно, что решение дается с учетом противодействующих сил, зависящих как от положения якоря, так и от его скорости.

¹ Б. С. Сотсков, Элементы автоматической и телемеханической аппаратуры, ГЭИ, 1950; К вопросу о приближенном решении уравнений движения для подвижной системы реле, «Автоматика и телемеханика», 1950, № 1; Зависимость времени движения подвижной системы реле от параметров реле, «Автоматика и телемеханика», 1951, № 4.

Б. С. Сотсков предлагает приближенное решение этих уравнений, полагая, что для периода движения

$$\Phi = \alpha + \beta t + \gamma t^2 + \varepsilon t^3 + \dots,$$

где $\alpha = \Phi_0$ при $t=0$.

Опуская все промежуточные операции, укажем, что подставив значение потока в уравнение (32), автор находит полное решение этого уравнения в виде

$$x = Qe^{-\rho t} \sin(\sqrt{k_1} t - \varphi) + A + Bt + Ct^2 + Dt^3, \quad (33)$$

где $Q, \rho, k_1, \varphi, A, B, C, D$ — постоянные величины.

Далее, подставив значения Φ и x в уравнение (31) и решив его совместно с уравнением (30), автор получает выражение для тока.

Б. С. Сотсков указывает, что для практики достаточно ограничить ряд потока двумя членами, тогда

$$x = Qe^{-\rho t} \sin(\sqrt{k_1} t - \varphi) + A + Bt + Ct^2, \quad (34)$$

$$i = i_m + b(\alpha + \beta t) \delta_0 - b(\alpha + \beta t) x. \quad (35)$$

Для облегчения практического использования этого метода Б. С. Сотсков совместно с И. Е. Декабрун и Н. А. Путиловой в работе «Зависимость времени движения подвижной системы реле от параметров реле» строит зависимости, дающие относительное перемещение якоря в функции времени, заданного в долях от значения постоянной времени. Тогда по заданным параметрам, используя эти кривые, можно найти значение времени движения для данного относительного перемещения якоря. Причем приводятся кривые как для случая отсутствия сил трения, так и с учетом их. Аналогичные кривые приводятся и для режима отпускания.

Метод Н. А. Лившица. В предыдущем разделе был изложен метод двойного графического интегрирования для решения задачи при любых законах изменения тяговой и нагрузочной характеристик.

Для частных случаев Н. А. Лившиц приводит аналитические решения уравнений динамики.

1) В случае, если

$$P = P_s - P_o = \text{const},$$

время движения определяется весьма просто уравнением

$$t = \sqrt{\frac{2m}{P} (\delta_{\max} - \delta_{\min})}. \quad (36)$$

2) В случае, если

$$P_s = \text{const} \text{ и } P_o = kx,$$

уравнение динамики также может быть решено аналитически.

После ряда преобразований решение представляется в следующем виде:

$$\cos \sqrt{\frac{k}{m} t_{\text{дв}}} = \frac{\frac{P_0}{k} - \delta_{\text{max}} + \delta_{\text{min}}}{\frac{P_0}{k}}. \quad (37)$$

Такую нагрузочную характеристику можно себе представить, например, в случае, если противодействующее усилие создается пружиной.

Н. А. Л и в ш и ц приводит также аналитические решения для этих же частных случаев, но с учетом сил трения, зависящих от скорости движения.

АНАЛИЗ МЕТОДОВ РАСЧЕТА И РЕКОМЕНДАЦИИ ПО ИХ ПРАКТИЧЕСКОМУ ИСПОЛЬЗОВАНИЮ

Методы последовательных интервалов

Сравнивая эти методы между собой, приходим к выводу, что первые три метода (Шюлер, Москвитин, Сотсков) основаны на одном и том же принципе: разделение процесса на ряд интервалов по времени, усреднение рассматриваемых величин в каждом интервале, представление уравнения динамики в виде баланса сил. На первый взгляд кажется, что между ними нет существенной разницы. На самом деле это не так, и неправильный выбор метода расчета в каждом конкретном случае может привести к существенным ошибкам.

Ш ю л е р пользуется формулой Максвелла для определения тяговой силы, которой, однако, не всегда можно пользоваться.

В книге М. А. Бабилова¹ дана общая формула Максвелла и весьма убедительно показано, что упрощенная формула (как у Шюлера) справедлива при условии равномерного распределения индукции по поверхности якоря, а также при условии, если вектор индукции нормален к поверхности якоря на всех его участках. Это, конечно, ограничивает область применения рассматриваемого метода.

Шюлер также не учитывает потоков рассеяния, хотя и в работе замечает, что учет потоков рассеяния вообще возможен, но, к сожалению, этим замечанием и ограничивается.

Метод Шюлера, однако, позволяет учитывать сопротивление стали. Резюмируя сказанное, можно сделать вывод, что методом Шюлера удобно пользоваться *при малых воздушных зазорах и при движении якоря вдоль силовых линий*. При малых зазорах во многих случаях, действительно, можно пренебрегать потоками рассеяния.

А. И. М о с к в и т и н вместо формулы Максвелла использует формулу, которая выводится из рассмотрения энергетического баланса электромагнита для ненасыщенных магнитных систем.

¹ М. А. Бабилов, Электрические аппараты, ч. I, ГЭИ, 1951.

Поэтому методом Москвитина А. И. удобно пользоваться *при больших зазорах, когда нельзя применить формулу Максвелла и при этом система не насыщена*. Однако при определении необходимых ампер-витков А. И. Москвитин учитывает также сопротивление стали, что противоречит применению указанной формулы для определения тяговой силы.

Правда, А. И. Москвитин в своей работе указывает, что приращение кинетической энергии при точных расчетах может быть определено графическим методом по характеристикам намагничивания, но, к сожалению, не излагает предлагаемый им метод. Тем не менее, А. И. Москвитин, несколько изменив методику Шюлера, значительно расширяет область применения своего метода, так как в большинстве случаев сопротивлением стали можно пренебречь.

Б. С. Сотсков применяет ту же формулу тяговой силы, что и А. И. Москвитин, но при этом, используя фактор отсутствия насыщения, расчет начинается сразу с определения приращения тока, принимая экспоненциальный закон его изменения до момента трогания. Этим он избегает предварительного определения средней индуктированной ЭДС для последующего нахождения потока, а затем и необходимых ампер-витков, как это делают Шюлер и Москвитин.

Метод Б. С. Сотскова позволяет также учитывать и потоки рассеяния, но при этом формулу тяговой силы (7) удобнее представить в следующем виде:

$$P_s = \frac{(0,4\pi i)^2}{8\pi} \frac{dL}{dx},$$

что даст возможность использовать уже построенную кривую $L = f(x)$, необходимую при расчете приращения силы тока.

Из всего сказанного видно, что методом Б. С. Сотскова так же, как и методом А. И. Москвитина, удобно пользоваться *при больших зазорах и отсутствии насыщения стали*.

Отметим, что И. И. Пеккер, пользуясь методом Б. С. Сотскова, построил динамические характеристики броневое электромагнита, вполне совпадающие с экспериментом¹.

Н. Е. Лысов разрешает проблему динамики наиболее строго. Метод этот применим во всех случаях: при любой степени насыщения стали, при любой величине воздушного зазора, при самых разнообразных формах полюсов, при любом характере механической нагрузки и т. д. Единственным допущением в методе является пренебрежение вихревыми токами. Следует отметить, что *ни один из известных методов расчета динамических характеристик не учитывает вихревых токов*.

Обычно считают, что метод Н. Е. Лысова громоздкий, ибо требует большого числа графических построений и основан на последовательных приближениях. По этому поводу следует сделать ряд замечаний.

¹ И. И. Пеккер, Применение методов численного интегрирования для расчета динамических характеристик электромагнитов постоянного тока, Труды МЭИ, вып. VII, ГЭИ, 1951.

Громоздкость графических построений заключается не в самом методе Н. Е. Лысова, а в построении зависимостей $\phi = f(t)$. Известно, что расчет магнитной цепи с учетом сопротивления стали и потоков рассеяния в общем виде встречает большие затруднения и известные методы расчета не отличаются ни строгостью, ни простотой расчета. Если же пренебрегать сопротивлением стали, то при этом всякие графические построения отпадают и задачу можно решить аналитическим путем, как предлагает Н. Е. Лысов.

Следует отметить, что в некоторых случаях строгий расчет динамических характеристик можно вести лишь методом Н. Е. Лысова. Например, в случае, если система насыщена и при этом нельзя применить формулу Максвелла, вряд ли уместно применение других методов. Что же касается применения Н. Е. Лысовым метода последовательных приближений, так это только увеличивает точность расчета.

Расчеты, проведенные нами методом Н. Е. Лысова, показали, что можно со второго раза найти правильное направление искомой переходной кривой $\phi = f(t)$. Здесь можно дать некоторые рекомендации. Если после первой попытки равенство (11) не соблюдается, то следует тут же прикинуть, какое значение $\Delta\phi$ дало бы соблюдение равенства (см. фиг. 1). Исходя из этого нужно и выбирать второй раз искомое направление. Следует заметить, что при изменении направления отрезка, изменится также и площадь треугольника, соответствующая механической работе, совершаемой якорем, т. е. соответственно изменится и отрезок времени. Поэтому предполагаемое равенство может нарушиться. Однако, как показывает практика, такое изменение во многих случаях не вносит существенной разницы в первоначальное значение найденного интервала времени.

Представляется весьма удобным даже при ненасыщенной магнитной цепи пользоваться все-таки графо-аналитическим методом, а не аналитическим. Это значительно облегчает работу, ибо после расчета двух-трех интервалов уже выявляется характер искомой переходной кривой тока и дальнейший расчет последующих интервалов большей частью не требует последовательных приближений; кроме этого, графическое построение увеличивает наглядность и сразу выявляет возможные ошибки при расчетах. В этом случае построение характеристик $\phi = f(t)$ в виде прямых не представляет никакого труда.

Метод Н. Е. Лысова имеет еще одно важное преимущество перед тремя рассмотренными выше методами, в которых процесс разбирается по временным интервалам. Ясно, что для определения числа интервалов и величины каждого из них необходимо предварительно хотя бы ориентировочно знать общее время движения. По методу же Н. Е. Лысова число интервалов и их величина могут быть с самого начала определены однозначно, так как полный ход якоря всегда известен.

Н. А. Лившиц при разработке своего метода берет за основу случай, когда в процессе движения ток в обмотке электромагнита остается неизменным, что встречается сравнительно редко. Правда, Н. А. Лившиц указывает, что в тех случаях, когда происходит из-

менение тока, «в частности, при значительных скоростях подвижных органов, необходимо определять скорость движения подвижных органов, пользуясь динамической тяговой характеристикой механизма, учитывающей изменение тока в обмотке в процессе перемещения подвижных органов».

К сожалению, автор ограничивается только этим указанием и не дает никаких пояснений, как же найти эти динамические характеристики, т. е. вопрос остается открытым. Сам автор излагает метод для случая, когда ток во время движения остается неизменным.

Здесь опять возникает вопрос, какой же принять расчетный ток во время движения подвижных органов. Н. А. Лившиц справедливо указывает, что тяговая характеристика определяется «величиной тока, протекающего через обмотку электромагнитного механизма во время перехода якоря от положения покоя до рабочего положения, и при разных значениях тока получаются разные времена движения». Автор не делает указаний относительно нахождения расчетного тока. Можно сделать предположение, что таким током в некоторых случаях может служить ток трогания.

Следует предостеречь от использования статических тяговых характеристик, построенных при токе, равном своему установившемуся значению. Такой подход может привести к грубым принципиальным ошибкам. Так, например, если увеличим напряжение, подаваемое на обмотку, но при этом последовательно включим такое добавочное сопротивление, чтобы установившийся ток остался неизменным, то статические характеристики в обоих случаях будут совершенно одинаковы. Поэтому времена движения, рассчитанные по методу Н. А. Лившица, также получаются одинаковыми. Нетрудно видеть, что на самом деле во втором случае время трогания и движения будет меньше, так как при этом подводимая мощность из сети больше, чем в первом случае.

Более того, установленное в предыдущей статье существование оптимальной индуктивности (в частности, оптимального числа витков) показывает, что в некоторых случаях чем выше тяговая характеристика, тем больше время срабатывания.

В самом деле, если будем увеличивать число витков за пределы оптимального значения, то при этом тяговые характеристики будут все выше и выше (считаем систему ненасыщенной). Однако времена срабатывания будут увеличиваться, что противоречит основному принципу метода Н. А. Лившица.

Поэтому использование статических тяговых характеристик, т. е. характеристик для установившегося значения тока, совершенно недопустимо при расчете методом Н. А. Лившица.

Из всего сказанного можно сделать вывод, что метод Н. А. Лившица можно применить в тех случаях, когда ток в обмотке электромагнита во все время движения якоря остается постоянным и когда известно значение этого тока, согласно которому и строится исходная тяговая характеристика расчетом магнитной цепи. Такой случай можно себе представить, если рассмотреть работу серьезного электромагнита (например, токовые реле). При движении якоря величина

тока все время остается неизменной и практически определяется внешней нагрузкой. При расчете же электромагнитов, включаемых на напряжение сети, использование метода Н. А. Лившица не имеет никакого смысла.

Аналитические методы

Яссе дает аналитическое решение уравнений динамики, но при этом делает ряд таких допущений, которые намного умаляют значение этого решения. Все эти допущения были показаны при изложении этого метода. Следует еще отметить, что в ходе решения дифференциальных уравнений Яссе отказывается даже от ранее им принятого допущения относительно индуктивности.

Сначала он принял, что индуктивность является линейной функцией от хода якоря, затем дал решение при $L \neq f(x)$ или $L = \text{const}$.

С таким допущением, конечно, нельзя согласиться. Если даже предположить, что остальные допущения для некоторых частных случаев могут быть справедливы, то при внимательном рассмотрении нетрудно заметить в них противоречия.

В самом деле, Яссе считает, что

$$\text{при } t=0 \quad i'=0, \quad i=I,$$

т. е. в начальный момент движения ток равен своему установившемуся значению. В то же время Яссе свое решение дает для случая, когда отсутствуют противодействующие силы. При этом, однако, время трогания совершенно отсутствует, а движение начинается при токе, равном нулю. Собственно, это наглядно видно из уравнения (16), приведенного самим же Яссе, ибо при $t=0 \quad v=0; \frac{dv}{dt} = 0$. Ясно,

что при таких противоречивых начальных условиях автор никак не мог получить решение, которое в какой-то мере могло соответствовать реальной картине.

Метод Яссе представляет интерес, лишь как попытка дать математическую трактовку динамики электромагнитов; практическое же использование этого метода для расчетов никак нельзя рекомендовать.

А. И. Москвитин разрешает проблему динамики электромагнитов постоянного тока весьма оригинальным способом. Заслуживает большого внимания удачное использование автором того положения, что большую часть времени якорь движется сравнительно медленно и поэтому картина поля (рассеяния, выпучивания) в течение этого времени практически одна и та же. Это позволяет расчет магнитной цепи производить только один раз — при начальном зазоре.

Кроме того, А. И. Москвитин предлагает весьма простой способ учета насыщения стали в конце хода якоря. Для большинства существующих электромагнитов действительно насыщение наступает только к концу хода. Такой учет поля рассеяния и насыщения стали значительно повышает точность расчета. Напомним, что Яссе не учитывал этих факторов. Если сравним методы Яссе и А. И. Москвитина,

то увидим, что, помимо этих факторов, у Яссе значительно больше допущений, чем у А. И. Москвитина.

В самом деле, общим допущением у них является отсутствие сил противодействия. Известно, что всю трудность решения уравнений динамики составляет главным образом наличие активного сопротивления обмотки. Эту трудность А. И. Москвитин обходит оригинальным путем, установив взаимосвязь между характерными величинами силы тока, о чем говорилось в изложении его метода. Правда, здесь некоторую погрешность вводит усреднение навсидимой в обмотке противо- ЭДС, но на конечный результат расчета это влияет, по-видимому, мало.

Яссе же эту трудность обходит весьма неудачно, приняв коэффициент самоиндукции постоянным за время движения, не говоря уже о грубых допущениях относительно изменения тока и скорости.

Таким образом, можно заключить, что метод А. И. Москвитина фактически *не учитывает только противодействующую силу*. Поэтому автор совершенно прав, указывая, что это решение в значительной степени отражает реальный процесс *в быстродействующих электромагнитах*, где сила, развиваемая электромагнитом, должна значительно превышать противодействующую силу.

Расчет методом А. И. Москвитина еще удобен тем, что сразу определяется общее время срабатывания.

Здесь только надо подчеркнуть, что само аналитическое решение дается без учета двух главных факторов: активного сопротивления обмотки и силы противодействия. Хотя и активное сопротивление в дальнейшем учитывается с достаточной для практических расчетов точностью, но оно непосредственно не фигурирует в полученных автором зависимостях, описывающих динамический процесс.

Несмотря на это, решение, предложенное А. И. Москвитиним, является важным шагом вперед в развитии расчетов динамических характеристик.

Б. С. Сотсков в отличие от предыдущих авторов дает аналитическое решение уравнений динамики с учетом всех основных факторов. В этом большая заслуга автора, несмотря на то, что решение это носит приближенный характер. Собственно, единственным допущением здесь является то, что Б. С. Сотсков заранее задается функцией потока от времени в виде степенного ряда. Экспериментальные данные показывают плавное возрастание потока в режиме динамики, поэтому такое представление потока в большой степени соответствует действительности. Насыщение стали учитывается введением постоянного слагаемого в уравнение магнитной цепи. Ценность этого слагаемого заключается в том, что в нем можно также учесть потери намагничивающей силы в паразитных зазорах, как например, в «воротничке» броневого соленоидного электромагнита.

Безусловно, пользуясь этим методом, можно также учесть и потоки рассеяния, выпучивания с достаточной для практических расчетов точностью. Учитывая еще компоненты противодействующих сил, зависящие от положения и скорости якоря, Б. С. Сотсков намного расширяет область применения своего метода.

Таким образом, метод Б. С. Сотскова учитывает все основные факторы, характеризующие динамический процесс, и поэтому он является наиболее общим из всех известных аналитических методов расчета.

Недостатком метода Б. С. Сотскова является сложность конечных уравнений, что несколько затрудняет их практическое использование. Поэтому автор, введя некоторые обобщенные параметры, дает зависимости относительного перемещения якоря в виде графиков, которые могут быть использованы для расчета времени движения электромагнитных механизмов.

Метод Б. С. Сотскова, хотя и приближенно, но в основном разрешает проблему аналитического расчета динамических характеристик.

Н. А. Л и в ш и ц дает аналитическое решение для двух частных случаев расположения тяговой и нагрузочной характеристик. Выше был приведен анализ общего метода двойного графического интегрирования; все те замечания в равной мере относятся и к этим аналитическим методам. К ним можно только добавить, что дополнительные допущения при аналитическом решении еще более сужают область применения данного метода.

Из рассмотрения различных методов видно, что советские ученые внесли наибольший вклад в разрешение проблемы динамики электромагнитов.

ВОЗМОЖНОСТИ АНАЛИТИЧЕСКОГО ОПРЕДЕЛЕНИЯ ОПТИМАЛЬНЫХ СООТНОШЕНИЙ ДЛЯ МИНИМАЛЬНОГО ВРЕМЕНИ ДВИЖЕНИЯ ПО СУЩЕСТВУЮЩИМ МЕТОДАМ РАСЧЕТА

Проведенный анализ методов расчета динамических характеристик электромагнитов позволяет перейти к важному вопросу, — возможно ли определение оптимальной индуктивности по существующим методам расчета времени движения.

Рассматривая с этой точки зрения методы расчета динамических характеристик, приходим к выводу, что методы последовательных интервалов не могут дать аналитического решения поставленной задачи. Выбрать оптимальные соотношения этими методами можно лишь серией расчетов и постепенным приближением к оптимальным параметрам.

Из рассмотренных аналитических методов заслуживают внимания методы Б. С. Сотскова и А. И. Москвитина.

Для получения зависимости времени движения от всех основных параметров, используя аналитический метод Б. С. Сотскова, следует в уравнение (34) подставить значение начального зазора, тогда получим

$$\delta_0 = Qe^{-pt} \sin(\sqrt{k_1}t - \varphi) + A + Bt + Ct^2. \quad (38)$$

Уравнение это необходимо решить относительно времени, т. е. надо получить функцию

$$t_{\text{дв}} = f(\delta_0, Q, p, k_1, \varphi, A, B, C). \quad (39)$$

Однако уравнение (38) является трансцендентным, и решение его в общем виде не представляется возможным.

Весьма заманчивым в этом отношении является аналитический метод А. И. Москвитина, дающий непосредственную зависимость времени срабатывания от основных параметров. Рассмотрим более подробно этот метод.

Метод А. И. Москвитина, как известно, дает аналитическое решение без учета противодействующей силы и активного сопротивления обмотки; последнее в дальнейшем учитывается но уже искусственным путем, с помощью расчетных коэффициентов, рекомендуемых автором.

Решим уравнения динамики по А. И. Москвитину, представив их в следующем виде:

$$\frac{d\Phi}{dt} = \frac{U}{w}, \quad (40)$$

$$m \frac{d^2x}{dt^2} = k\Phi^2, \quad (41)$$

где k — коэффициент, зависящий от конструкции электромагнита.

Определив выражение потока из уравнения (40), подставим его в уравнение (41). Тогда решение последнего уравнения осуществляется непосредственно, и в итоге получим

$$x = \frac{kU^2}{12mw^2} t^4. \quad (42)$$

Отсюда легко можно определить время срабатывания:

$$t_{cp} = \sqrt[4]{\frac{12\delta_0 m w^2}{kU^2}}. \quad (43)$$

Для электромагнита с одним рабочим зазором, например, броневое соленоидного

$$k = \frac{1}{2\mu_0 S}. \quad (44)$$

Для Ш-образного электромагнита

$$k = \frac{1}{\mu_0 S}, \quad (45)$$

где S — площадь сечения среднего сердечника.

Подставив значение этого коэффициента в формулу (43), получим соответственно для броневое электромагнита

$$t_{cp} = \sqrt[4]{\frac{24\mu_0 S \delta_0 m w^2}{U^2}}, \quad (46)$$

для Ш-образного

$$t_{cp} = \sqrt[4]{\frac{12\mu_0 S \delta_0 m w^2}{U^2}}. \quad (47)$$

Если принять во внимание, что для броневое

$$L_0 = \frac{\mu_0 S}{\delta_0} w^2, \quad (48)$$

а для Ш-образного

$$L_0 = \frac{\mu_0 S}{2\delta_0} w^2, \quad (49)$$

то формулы (46), (47) можно представить в общем виде

$$t_{cp} = \sqrt[4]{\frac{24L_0\delta_0^2 m}{U^2}}. \quad (50)$$

Это в сущности и есть аналитическая формула А. И. Москвитина, записанная в иной форме. Такая запись формулы удобна тем, что ее можно применить для любого электромагнита, подставив значение индуктивности, определяемое конструкцией электромагнита. Эта формула рекомендуется вообще для расчетов по методу А. И. Москвитина.

В предыдущей статье было показано, что при индуктивности, равной или меньшей некоторого минимального значения, электромагнит вообще не срабатывает. А по формуле А. И. Москвитина с уменьшением индуктивности время срабатывания все более уменьшается, и в пределе при $L_0=0$ имеем

$$t_{cp}=0,$$

что, конечно, не соответствует действительности.

Значит, пользуясь методом А. И. Москвитина, также нельзя найти оптимальные соотношения электромагнита. Это вполне понятно, ибо здесь мы пренебрегаем такими важными параметрами, как противодействующая сила и активное сопротивление обмотки.

Таким образом, все существующие методы расчета не позволяют решить поставленную задачу: найти аналитическим путем оптимальные соотношения для минимального времени движения.

ВЫВОДЫ

1. Графо-аналитический метод Н. Е. Лысова является наиболее общим не только из методов последовательных интервалов, но и вообще из всех известных методов расчета динамических характеристик электромагнитов. Он может быть применен практически при любых заданных условиях.

2. Методы последовательных интервалов А. И. Москвитина и Б. С. Сотскова могут быть применены при больших рабочих зазорах, когда можно пренебречь сопротивлением стали. При этом потоки рассеяния и выпучивания могут быть учтены.

3. Использование метода Шюлера целесообразно при малых рабочих зазорах, когда сопротивлением стали нельзя пренебречь. Движение якоря при этом должно быть вдоль силовых линий.

4. Методы расчета, предложенные Н. А. Лившицем, могут найти применение лишь для серийных электромагнитов, когда известно значение тока, протекающего по обмотке электромагнита, в процессе движения якоря.

5. Практическое использование метода Яссе не представляется возможным.

6. Аналитический метод А. И. Москвитина весьма удобен для расчета быстродействующих электромагнитов, когда тяговые усилия электромагнита намного превышают силы противодействия и когда магнитное сопротивление рабочего зазора пропорционально его длине. Магнитная система при начальном зазоре должна быть ненасыщенной.

7. Метод Б. С. Сотскова является наиболее общим из аналитических методов расчета. Практическое применение этого метода целесообразно при использовании приведенных автором графических зависимостей обобщенных параметров, особенно в случае, когда противодействующими усилиями нельзя пренебрегать, и поэтому аналитический метод А. И. Москвитина нельзя использовать.

8. Существующие методы расчета времени движения не позволяют получить необходимые аналитические выражения в общем виде для оптимальной индуктивности и минимального времени движения. Разрешение этой задачи возможно лишь при нахождении иного решения исходных дифференциальных уравнений динамики.

ЛИТЕРАТУРА

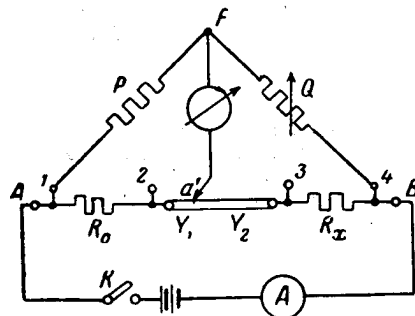
1. Бабиков М. А., Электрические аппараты, ч. 1, ГЭИ, 1951.
2. Лившиц Н. А., К вопросу об анализе времени движения якоря электромагнитных механизмов, «Автоматика и телемеханика», 1939, № 2.
3. Лившиц Н. А., Определение времени движения подвижных органов электромагнитного механизма при его срабатывании, «Известия электропромышленности слабого тока», 1940, № 11.
4. Лысов Н. Е., Расчет электромагнитных механизмов, Оборонгиз, 1949.
5. Лысов Н. Е., К расчету динамических характеристик электромагнитов, «Вестник электропромышленности», 1941, № 3.
6. Москвитин А. И., Основные проблемы электрических машин прямолинейного движения, «Электричество», 1941, № 2.
7. Москвитин А. И., Уравнения процессов в электромагните с движущимся якорем, «Известия Академии Наук СССР», 1948, № 4.
8. Москвитин А. И., Электрические машины возвратно-поступательного движения, Изд. АН СССР, 1950.
9. Пеккер И. И., Применение методов численного интегрирования для расчета динамических характеристик электромагнитов постоянного тока, Труды МЭИ, вып. VII, ГЭИ, 1951.
10. Сотсков Б. С., Элементы автоматической и телемеханической аппаратуры, ГЭИ, 1950.
11. Сотсков Б. С., К вопросу о приближенном решении уравнений движения для подвижной системы реле, «Автоматика и телемеханика», 1950, № 1.
12. Сотсков Б. С., Декабрун И. Е. и Путилова Н. А., Зависимость времени движения подвижной системы реле от параметров реле, «Автоматика и телемеханика», 1951, № 4.
13. Яссе Э., Электромагниты, ГЭИ, 1934.
14. Schüler, Die Wirkungsrad des Elektromagneten, ETZ, 1913, S. 613.

Канд. техн. наук С. Н. ИКОННИКОВ

О МЕТОДИЧЕСКОМ ПЕРЕХОДЕ ОТ ОБЫКНОВЕННОГО МОСТА К ДВОЙНОМУ МОСТУ (ТОМСОНА)

Для измерения малых сопротивлений при отсутствии двойного моста обычно рекомендуется метод двойного отсчета или, лучше сказать, метод двукратного измерения. Неизвестное сопротивление R_x включается по схеме, изображенной на фиг. 1.

В этом случае, кроме R_x , в схему входит неизвестное сопротивление Y , которое является сопротивлением участка, где присоединяются контакты от потенциальных зажимов 2 и 3 (А—1; А—2).



Фиг. 1. Схема измерения сопротивления по методу двукратного отсчета.

Присоединяя один раз гальванометр к зажиму 2, регулировкой сопротивления Q уравнивают мост, и тогда справедливо равенство

$$Q'R_0 = P(R_x + Y). \quad (1)$$

После этого гальванометр включают на зажим 3 и вновь уравнивают мост регулировкой сопротивления Q . Тогда справедливо равенство

$$Q''(R_0 + Y) = PR_x. \quad (2)$$

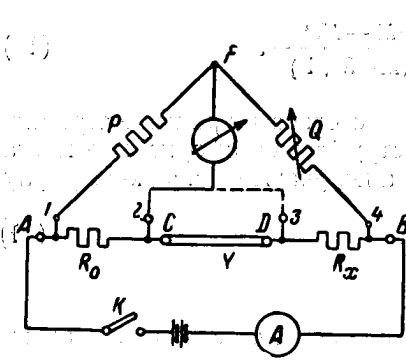
Решая совместно полученные два уравнения, определяют R_x и Y (А—2).

Следует обратить внимание на то, что уравнивание могло бы быть осуществлено при каком-то единственном значении Q , если бы, перемещая по соединительному проводнику CD конец проводника от гальванометра, мы бы нашли точку a' , делящую Y (фиг. 2) таким образом, чтобы было справедливо

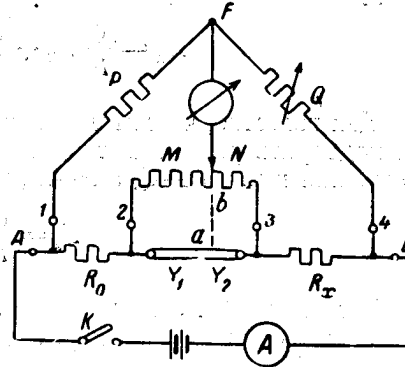
$$Q(R_0 + Y_1) = P(R_x + Y_2). \quad (3)$$

Но в этом случае нам необходимо было бы знать Y_1 и Y_2 , что могло бы вызывать еще большие затруднения, так как эти сопротивления, подобно Y , должны содержать еще переходное сопротивление в токовых зажимах C и D . Однако, присоединяя к потенци-

альным зажимам 2 и 3 некоторое сопротивление MbN (фиг. 3); можно отыскать на нем точку такого же потенциала, как потенциал точки F и точки a ; сопротивления M и N при этом становятся известными. Так как потенциалы точек b и a во время равновесия моста одинаковы, то перемычка ab может не существовать.



Фиг. 2. Принципиальная схема моста для измерения малых сопротивлений.



Фиг. 3. Схема двойного моста.

Из равновесия этого же моста с плечами Y_1 , Y_2 , M и N следует равенство

$$\frac{Y_1}{Y_2} = \frac{M}{N}, \quad (4)$$

а мост $AFBA$ позволяет написать

$$P \left(R_x + \frac{NY_2}{N+Y_2} \right) = Q \left(R_0 + \frac{MY_1}{M+Y_1} \right); \quad (5)$$

$$R_x = \frac{Q}{P} R_0 + \frac{Q}{P} \frac{MY_1}{M+Y_1} - \frac{NY_2}{N+Y_2}. \quad (5')$$

Из уравнения (4) можем получить

$$\frac{Y_1+Y_2}{Y_2} = \frac{M+N}{N}; \quad \frac{Y_1}{Y_2} = \frac{M+N}{N}; \quad Y_2 = Y \frac{N}{M+N}; \quad (6)$$

$$\frac{Y_1}{Y_1+Y_2} = \frac{M}{M+N}; \quad \frac{Y_1}{Y} = \frac{M}{M+N}; \quad Y_1 = Y \frac{M}{M+N}. \quad (7)$$

Значения Y_1 и Y_2 подставим в уравнение (5'):

$$R_x = \frac{Q}{P} R_0 + \frac{Q}{P} \frac{M \frac{MY}{M+N}}{M + \frac{MY}{M+N}} - \frac{N \frac{NY}{M+N}}{M + \frac{NY}{M+N}}. \quad (8)$$

Производя сокращение, получим

$$R_x = \frac{Q}{P} R_0 + \frac{Q}{P} \frac{MY}{M+N+Y} - \frac{NY}{M+N+Y}, \quad (9)$$

последние слагаемые уравнения (9) легко приводят к обычной форме уравнения с поправочным членом:

$$R_x = \frac{Q}{P} R_0 + Y \frac{QM - NP}{P(M+N+Y)}. \quad (10)$$

Как известно, числитель поправочного члена в практике стремятся подобрать равным нулю, т. е. конструктивное оформление моста позволяет иметь $N=Q$, а $P=M$, и тогда получают основную формулу

$$R_x = \frac{Q}{P} R_0. \quad (11)$$

Изложенный метод позволяет:

- 1) дать логический переход от простого моста к более сложному;
- 2) исключить необходимость решения трех контурных уравнений или необходимость преобразования треугольника сопротивлений в звезду;
- 3) наглядно уяснить укоренившийся в практике термин «двойного моста».

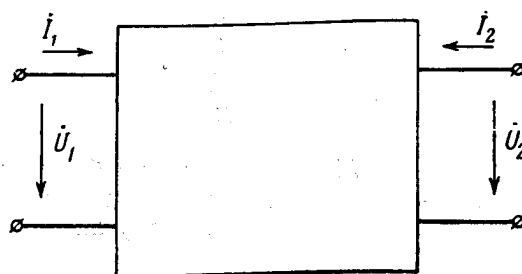
ЛИТЕРАТУРА

1. Карандеев К. Б., Методы электрических измерений, ГЭИ, 1952.
2. Шрамков Е. Г., Электрические и магнитные измерения, ГЭИ, 1937.

Инж. В. Ф. РАХМАНОВ

КРУГОВАЯ ДИАГРАММА ДЛЯ ОПРЕДЕЛЕНИЯ КОЭФФИЦИЕНТОВ УСИЛЕНИЯ ПО ТОКУ И НАПРЯЖЕНИЮ ПРИ КОМПЛЕКСНОЙ НАГРУЗКЕ

Усилитель, работающий при достаточно малых токах и напряжениях сигнала, может рассматриваться как активный линейный четырехполюсник. Такой четырехполюсник в самом общем виде показан на фиг. 1 с указанием положительных направлений токов и напряжений. Приведенные ниже рассуждения справедливы для любой схемы включения усилительного элемента, которым может быть, в част-



Фиг. 1. Четырехполюсник с указанием положительных направлений токов и напряжений.

ности, трехэлектродный прибор, работающий на линейном участке характеристики (трехэлектродная электронная лампа, полупроводниковый триод, многосеточная лампа, которая управляется одной из сеток, а потенциалы других остаются постоянными и т. д.).

Для усилителя с включенным на его выход сопротивлением нагрузки Z_n справедлива следующая система уравнений в параметрах Z :

$$\left. \begin{aligned} \dot{U}_1 &= Z_{11} \dot{I}_1 + Z_{12} \dot{I}_2, \\ 0 &= Z_{21} \dot{I}_1 + (Z_{22} + Z_n) \dot{I}_2. \end{aligned} \right\} \quad (1)$$

Используем эту систему уравнений для нахождения коэффициента усиления по току. Для этого определим токи \dot{I}_1 и \dot{I}_2 :

$$\dot{I}_1 = \frac{\begin{vmatrix} \dot{U}_1 & Z_{12} \\ 0 & Z_{22} + Z_H \end{vmatrix}}{\Delta_Z}; \quad (2)$$

$$\dot{I}_2 = \frac{\begin{vmatrix} Z_{11} & \dot{U}_1 \\ Z_{21} & 0 \end{vmatrix}}{\Delta_Z}. \quad (3)$$

Отношение тока \dot{I}_2 к току \dot{I}_1 равно коэффициенту усиления по току:

$$K_i = \frac{\begin{vmatrix} Z_{11} & \dot{U}_1 \\ Z_{21} & 0 \end{vmatrix}}{\begin{vmatrix} \dot{U}_1 & Z_{12} \\ 0 & Z_{22} + Z_H \end{vmatrix}} = -\frac{Z_{21}}{Z_{22} + Z_H}. \quad (4)$$

Разделив числитель и знаменатель выражения (4) на Z_{22} , получим

$$K_i = \frac{-\frac{Z_{21}}{Z_{22}}}{1 + \frac{Z_H}{Z_{22}}}. \quad (5)$$

Числитель выражения (5) представляет собой коэффициент усиления по току при $Z_H = 0$, который условимся обозначать $K_i(0)$. Следовательно, для любой схемы включения усилительного элемента

$$\frac{K_i}{K_i(0)} = \frac{1}{1 + \frac{Z_H}{Z_{22}}}. \quad (6)$$

Для определения коэффициента усиления по напряжению воспользуемся системой уравнений в параметрах Y

$$\left. \begin{aligned} \dot{I}_1 &= Y_{11}\dot{U}_1 - Y_{12}\dot{U}_2, \\ 0 &= -Y_{21}\dot{U}_1 + (Y_{22} + Y_H)\dot{U}_2. \end{aligned} \right\} \quad (7)$$

Решая систему уравнений (7) относительно \dot{U}_1 и \dot{U}_2 , получим

$$\dot{U}_1 = \frac{\begin{vmatrix} \dot{I}_1 & -Y_{12} \\ 0 & Y_{22} + Y_H \end{vmatrix}}{\Delta_Y}; \quad (8)$$

$$U_2 = \frac{\begin{vmatrix} Y_{11} & I_1 \\ -Y_{21} & 0 \end{vmatrix}}{\Delta_Y} \quad (9)$$

Отношение напряжения \dot{U}_2 к напряжению \dot{U}_1 является коэффициентом усиления по напряжению

$$K_V = \frac{\begin{vmatrix} Y_{11} & I_1 \\ -Y_{21} & 0 \end{vmatrix}}{\begin{vmatrix} I_1 & -Y_{12} \\ 0 & Y_{22} + Y_H \end{vmatrix}} = \frac{Y_{21}}{Y_{22} + Y_H} \quad (10)$$

Разделив числитель и знаменатель выражения (10) на Y_{22} , получим

$$K_V = \frac{\frac{Y_{21}}{Y_{22}}}{1 + \frac{Y_H}{Y_{22}}} \quad (11)$$

Числитель выражения (11), т. е. отношение $\frac{Y_{21}}{Y_{22}}$, представляет собой коэффициент усиления по напряжению при $Y_H=0$ (или при $Z_H=\infty$). Условимся обозначать его $K_V(\infty)$. Отношение коэффициента усиления по напряжению при конечной величине сопротивления нагрузки к коэффициенту усиления по напряжению при холостом ходе на выходе усилителя, т. е. при $Z_H=\infty$, можно записать в виде, совершенно аналогичном выражению (6):

$$\frac{K_V}{K_V(\infty)} = \frac{1}{1 + \frac{Y_H}{Y_{22}}} \quad (12)$$

Назовем отношение $\frac{K_I}{K_I(0)}$ — нормированным коэффициентом усиления по току, а отношение $\frac{K_V}{K_V(\infty)}$ — нормированным коэффициентом

усиления по напряжению. Будем обозначать их \bar{K}_I и \bar{K}_V . В силу того, что выражения (6) и (12) полностью аналогичны друг другу, они могут быть записаны в общем виде следующим образом:

$$\bar{W} = u + jv = \frac{1}{1 + M} = \frac{1}{1 + Me^{j\varphi}} \quad (13)$$

Следует заметить, что выражения для \bar{K}_I и \bar{K}_V приводятся к виду (13) независимо от выбранной системы параметров четырехполюсника.

Выражение (13) является аналитической функцией комплексного переменного $M = x + jy = Me^{j\varphi}$. Наглядной графической интерпрета-

$$U_2 = \frac{\begin{vmatrix} Y_{11} & I_1 \\ -Y_{21} & 0 \end{vmatrix}}{\Delta_Y} \quad (9)$$

Отношение напряжения \dot{U}_2 к напряжению \dot{U}_1 является коэффициентом усиления по напряжению

$$K_V = \frac{\begin{vmatrix} Y_{11} & I_1 \\ -Y_{21} & 0 \end{vmatrix}}{\begin{vmatrix} I_1 & -Y_{12} \\ 0 & Y_{22} + Y_H \end{vmatrix}} = \frac{Y_{21}}{Y_{22} + Y_H} \quad (10)$$

Разделив числитель и знаменатель выражения (10) на Y_{22} , получим

$$K_V = \frac{\frac{Y_{21}}{Y_{22}}}{1 + \frac{Y_H}{Y_{22}}} \quad (11)$$

Числитель выражения (11), т. е. отношение $\frac{Y_{21}}{Y_{22}}$, представляет собой коэффициент усиления по напряжению при $Y_H=0$ (или при $Z_H=\infty$). Условимся обозначать его $K_V(\infty)$. Отношение коэффициента усиления по напряжению при конечной величине сопротивления нагрузки к коэффициенту усиления по напряжению при холостом ходе на выходе усилителя, т. е. при $Z_H=\infty$, можно записать в виде, совершенно аналогичном выражению (6):

$$\frac{K_V}{K_V(\infty)} = \frac{1}{1 + \frac{Y_H}{Y_{22}}} \quad (12)$$

Назовем отношение $\frac{K_I}{K_I(0)}$ — нормированным коэффициентом усиления по току, а отношение $\frac{K_V}{K_V(\infty)}$ — нормированным коэффициентом

усиления по напряжению. Будем обозначать их \bar{K}_I и \bar{K}_V . В силу того, что выражения (6) и (12) полностью аналогичны друг другу, они могут быть записаны в общем виде следующим образом:

$$\bar{W} = u + jv = \frac{1}{1 + M} = \frac{1}{1 + Me^{j\varphi}} \quad (13)$$

Следует заметить, что выражения для \bar{K}_I и \bar{K}_V приводятся к виду (13) независимо от выбранной системы параметров четырехполюсника.

Выражение (13) является аналитической функцией комплексного переменного $M = x + jy = Me^{j\varphi}$. Наглядной графической интерпрета-

цией этого выражения является конформное отображение плоскости \dot{M} на плоскость \dot{W} . Всякая окружность, соответствующая постоянному модулю M , переходит в окружность на плоскости \dot{W} , уравнение которой записывается следующим образом:

$$u^2 + v^2 - 2 \frac{1}{1-M^2} u + \frac{1}{1-M^2} = 0, \quad (14)$$

а координаты центра и радиус определяются формулами

$$\left. \begin{aligned} u_{\text{ц}} &= \frac{1}{1-M^2}, \\ v_{\text{ц}} &= 0, \\ R &= \frac{M}{1-M^2}. \end{aligned} \right\} \quad (15)$$

Прямая в плоскости \dot{M} , соответствующая постоянному аргументу φ , переходит в плоскости \dot{W} в окружность, уравнение которой

$$u^2 + v^2 - u - v \operatorname{ctg} \varphi = 0. \quad (16)$$

Координаты центра и радиус этой окружности определяются формулами

$$\left. \begin{aligned} u_{\text{ц}} &= \frac{1}{2}, \\ v_{\text{ц}} &= \frac{1}{2} \operatorname{ctg} \varphi, \\ R &= \frac{1}{2 \sin \varphi}. \end{aligned} \right\} \quad (17)$$

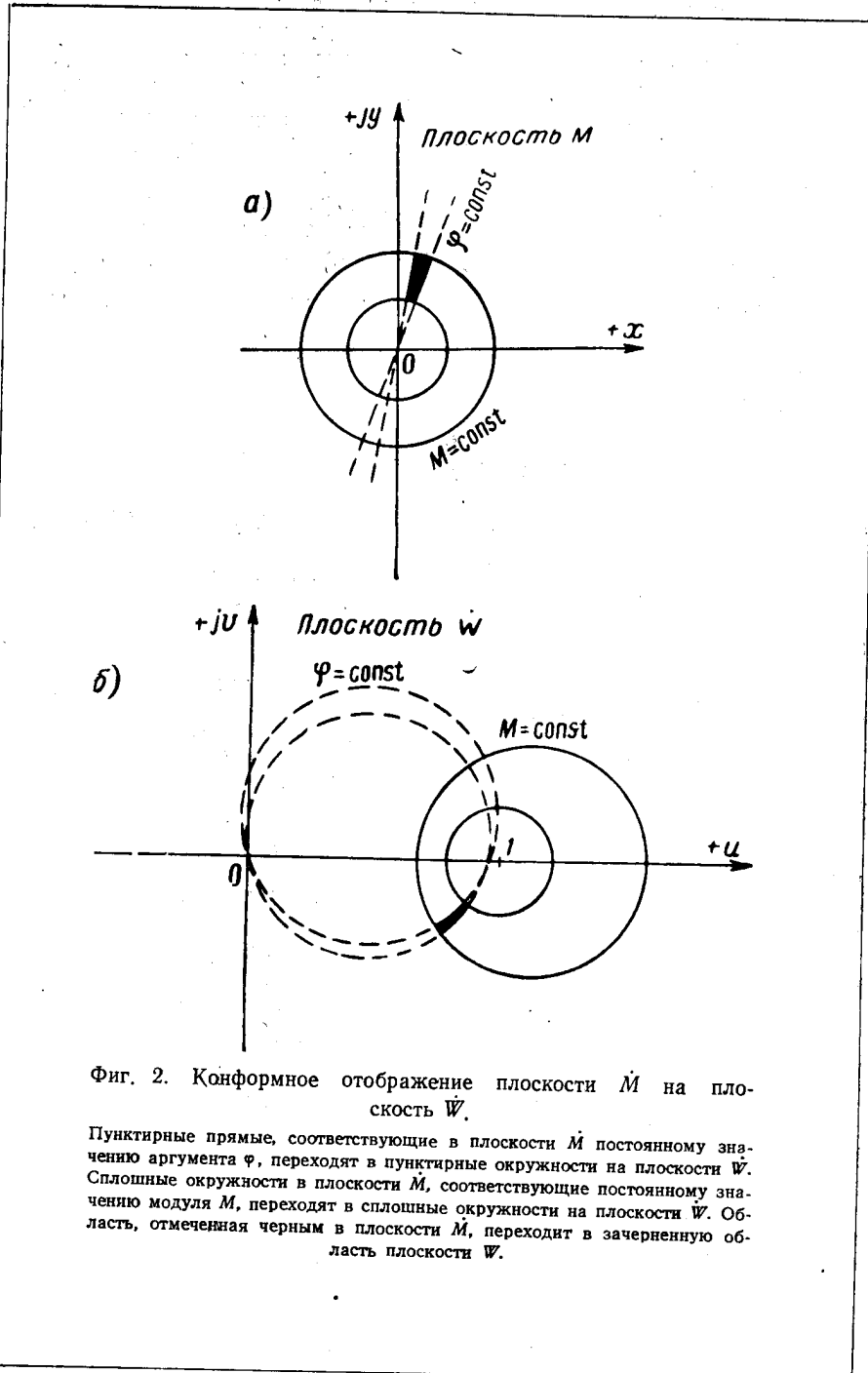
Уравнения (14) и (16) и формулы (15) и (17) получены обычными приемами аналитической геометрии¹.

На фиг. 2,а показаны линии постоянного модуля M и постоянного аргумента φ в плоскости \dot{M} . Соответствующие им окружности в плоскости \dot{W} показаны на фиг. 2,б. Показанные черным области переходят одна в другую. Точка 0 на плоскости \dot{M} переходит в 1 на плоскости \dot{W} , а точка ∞ переходит в 0.

Круговая диаграмма для определения коэффициентов усиления ламповых усилителей и усилителей с плоскостными полупроводниковыми триодами показана на фиг. 3. Она представляет собой два ортогональных семейства окружностей.

Окружности первого из этих семейств определяются уравнением (14). Каждая из окружностей этого семейства является линией постоянного модуля M . Второе семейство окружностей соответствует уравнению (16). Окружности этого семейства проведены при постоянных значениях угла φ .

¹ И. И. Привалов, Аналитическая геометрия, Гостехиздат, 1945.



Фиг. 2. Конформное отображение плоскости M на плоскость W .

Пунктирные прямые, соответствующие в плоскости M постоянному значению аргумента φ , переходят в пунктирные окружности на плоскости W . Сплошные окружности в плоскости M , соответствующие постоянному значению модуля M , переходят в сплошные окружности на плоскости W . Область, отмеченная черным в плоскости M , переходит в зачерненную область плоскости W .

Для расчета усилителей с точечными полупроводниковыми триодами можно использовать аналогичную диаграмму, включающую окружности, соответствующие отрицательным значениям угла φ .

Для определения нормированного коэффициента усиления на круговую диаграмму наносят точку, соответствующую значению \dot{M} , что производится с помощью вышеупомянутой сетки ортогональных окружностей. Значение нормированного коэффициента усиления находится непосредственно из круговой диаграммы. Оно дается координатами той же точки в прямоугольной системе координат.

Если отнести знак $K_i = K_i(0) \bar{K}_i$ к $K_i(0)$ и знак $K_V = K_V(\infty) \bar{K}_V$ к $K_V(\infty)$, то при условии, что φ не превышает $\pm \frac{\pi}{2}$, векторы \bar{K}_V и \bar{K}_i расположатся в правой полуплоскости \dot{W} .

Согласно определению

$$K_V = \frac{\dot{U}_2}{\dot{U}_1} = \frac{U_2 e^{j\alpha_2}}{U_1 e^{j\alpha_1}} = \frac{U_2}{U_1} e^{j(\alpha_2 - \alpha_1)},$$

где $(\alpha_2 - \alpha_1)$ — угол между действительной осью и вектором K_V ;

$$K_i = -\frac{\dot{I}_2}{\dot{I}_1} = \frac{I_2 e^{j\beta_2} e^{j\pi}}{I_1 e^{j\beta_1}} = \frac{I_2}{I_1} e^{j(\beta_2 - \beta_1 + \pi)},$$

$(\beta_2 - \beta_1 + \pi)$ — угол между действительной осью и вектором K_i .

Разность углов $(\alpha_2 - \alpha_1)$ и $(\beta_2 - \beta_1 + \pi)$ представляет собой угол между векторами K_V и K_i . Эта разность может быть записана следующим образом:

$$(\alpha_2 - \alpha_1) - (\beta_2 - \beta_1 + \pi) = (\alpha_2 - \beta_2) - (\alpha_1 - \beta_1) - \pi = \varphi_2 - \varphi_1 - \pi = \theta. \quad (18)$$

В последнем выражении $\varphi_2 = (\alpha_2 - \beta_2)$ — угол сдвига фаз между напряжением и током в сопротивлении нагрузки, $\varphi_1 = (\alpha_1 - \beta_1)$ — угол сдвига фаз между напряжением и током во входной цепи.

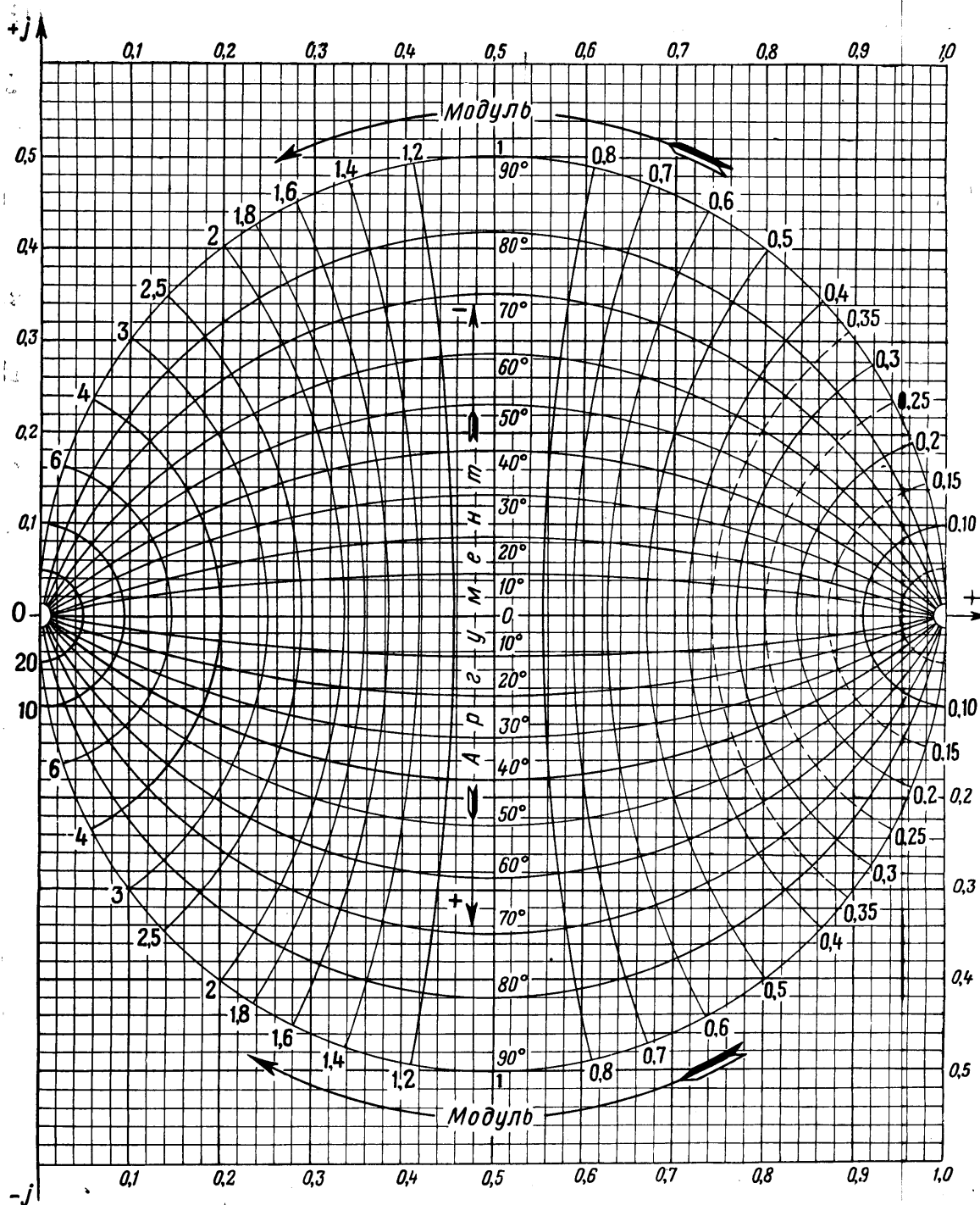
Для случая, когда комплексным характером параметров четырехполюсника можно пренебречь, разность $(\varphi_2 - \varphi_1)$ определяется непосредственным измерением угла между \bar{K}_V и \bar{K}_i на круговой диаграмме, так как при этом вектор \bar{K}_i (или \bar{K}_V) повернут относительно вектора K_i (или K_V) на угол π .

Пример. Найти коэффициенты усиления по току, напряжению и мощности, входное сопротивление и угол сдвига фаз между током и напряжением на входе полупроводникового усилителя, выполненного по схеме с заземленным эмиттером. Параметры полупроводникового триода, используемого в схеме, следующие:

$$r_s = 20 \text{ ом}, \quad r_r = 9 \cdot 10^5 \text{ ом},$$

$$r_{\text{осн}} = 200 \text{ ом}, \quad r_k = 1 \cdot 10^6 \text{ ом}.$$

Сопротивление нагрузки $Z_n = 34\,200 + j93\,397 \approx 10^5 e^{j70^\circ}$.



Фиг. 3. Круговая диаграмма для определения коэффициентов усиления по напряжению и по току.

Для схемы с заземленным эмиттером¹

$$Z_{11} = r_b + r_{оч} = 220 \text{ ом},$$

$$Z_{12} = r_b = 20 \text{ ом},$$

$$Z_{21} = r_b - r_c \approx -9 \cdot 10^5 \text{ ом},$$

$$Z_{22} = r_b + r_c - r_c = 1 \cdot 10^8 \text{ ом},$$

$$\Delta Z = Z_{11}Z_{22} - Z_{12}Z_{21} = 4 \cdot 10^7.$$

Определяем $K_i(0)$ и $K_V(\infty)$:

$$K_i(0) = -\frac{Z_{21}}{Z_{22}} = \frac{9 \cdot 10^5}{1 \cdot 10^8} = 9,$$

$$K_V(\infty) = \frac{Z_{21}}{Z_{11}} = \frac{-9 \cdot 10^5}{2,2 \cdot 10^2} = -4,09 \cdot 10^3.$$

Находим отношения $\frac{Z_{11}}{Z_{22}}$ и $\frac{Y_{11}}{Y_{22}}$:

$$\frac{Z_{11}}{Z_{22}} = \frac{10^5 e^{j70^\circ}}{10^8} = e^{j70^\circ}.$$

$$\frac{Y_{11}}{Y_{22}} = \frac{\Delta Z}{Z_{11}Z_{22}}; \quad \frac{\Delta Z}{Z_{11}Z_{22}} = \frac{4 \cdot 10^7}{2,2 \cdot 10^2 e^{j70^\circ}} = 1,82 e^{-j70^\circ}.$$

На круговой диаграмме находим точку с координатами $M=1$; $\varphi=70^\circ$, пользуясь дугами ортогональных окружностей. Прямоугольные координаты, соответствующие этой точке, дают значение нормализованного коэффициента усиления по току (фиг. 4):

$$\bar{K}_i = 0,5 - j0,35.$$

Аналогично для определения нормализованного коэффициента усиления по напряжению находим точку с координатами $M=1,82$; $\varphi=-70^\circ$. Читаем значение нормализованного коэффициента усиления по напряжению:

$$\bar{K}_V = 0,29 + j0,31.$$

Находим коэффициенты усиления по току K_i и напряжению K_V .

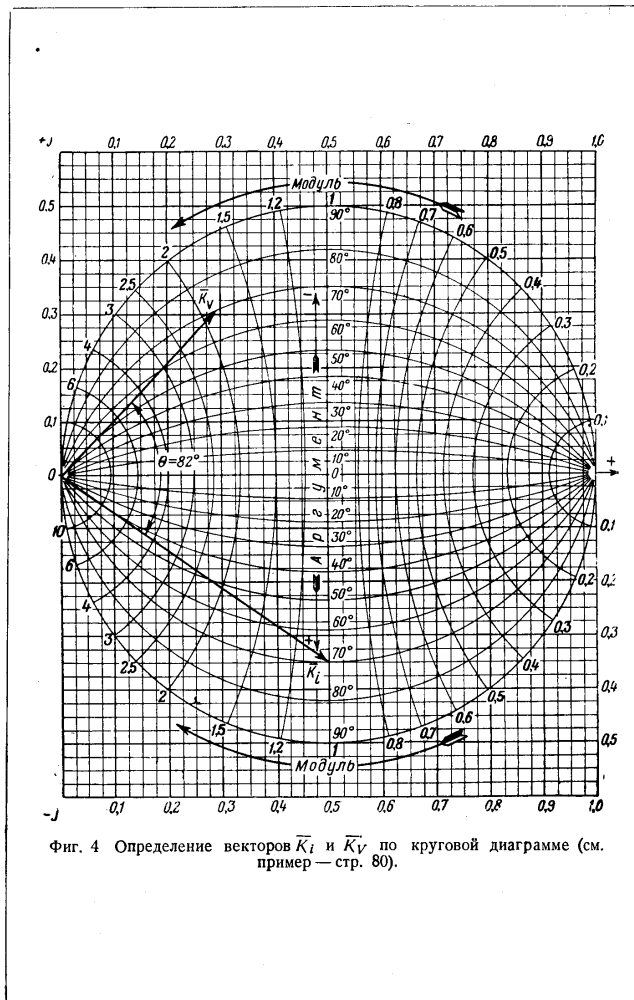
$$K_i = K_i(0) \cdot \bar{K}_i = 9(0,5 - j0,35) = 4,5 - j3,15,$$

$$K_V = K_V(\infty) \cdot \bar{K}_V = -4,09 \cdot 10^3(0,29 + j0,31) \approx -1190 - j1270.$$

Измеряем угол между \bar{K}_V и \bar{K}_i (определяем угол θ)

$$\theta = \varphi_2 - \varphi_1 \approx 82^\circ; \quad \varphi_1 \approx 12^\circ.$$

¹ Principles of Transistor Circuits, Edited by R. F. Shea, 1953.



Находим входное сопротивление:

$$Z_{вх} = Z_{11} + Z_{12}K_i = 220 + 20(4,5 - j3,15) = 310 - j63.$$

Находим коэффициент усиления по мощности:

$$K_p = |K_V| |K_I| \frac{\cos \varphi_2}{\cos \varphi_1} \approx 3300.$$

Круговую диаграмму, показанную на фиг. 3, можно назвать круговой диаграммой в полярных координатах, так как отношение \dot{M} при ее использовании приводится к виду $\dot{M} = Me^{j\varphi}$, где M и φ полярные координаты.

Является возможным построение аналогичной круговой диаграммы и для отношения \dot{M} , рассматриваемого в виде функции прямоугольных координат, т. е. $M = x + jy$.

Если известна зависимость отношения $\frac{Z_{11}}{Z_{22}}$ или $\frac{Y_{11}}{Y_{22}}$ от частоты, то круговая диаграмма может быть с успехом использована для нахождения амплитудно-фазовой характеристики усилителя, причем непосредственно по круговой диаграмме может быть определена вещественная часть частотной характеристики, зная которую можно рассчитать переходную характеристику усилителя¹.

В заключение следует заметить, что круговую диаграмму (см. фиг. 3) и формулы, приведенные выше, можно с успехом использовать и для расчета пассивных четырехполюсников.

¹ Основы автоматического регулирования, под ред. Солодовникова В. В., Машгиз, 1954.

Канд. техн. наук В. Н. ИСТРАТОВ, инж. В. В. МИТЮШИНА

К РАСЧЕТУ ТОКОВ КОРОТКОГО ЗАМЫКАНИЯ САМОЛЕТНОЙ ЭЛЕКТРИЧЕСКОЙ СИСТЕМЫ ТРЕХФАЗНОГО ТОКА

В связи с внедрением на самолетах систем трехфазного тока, а также усложнением схем коммутации и необходимостью защиты от коротких замыканий (к. з.) указанных систем расчет токов к. з. абсолютно необходим для правильного проектирования защиты от коротких замыканий элементов самолетной электрической системы и выбора аппаратуры.

В настоящей статье содержатся некоторые результаты расчетов токов к. з., которые были проведены авторами на кафедре «Теоретическая электротехника» МАИ для ряда простейших трехфазных систем. Недостаток экспериментальных данных, к сожалению, не позволил сделать окончательных выводов, однако и полученные результаты дают возможность провести некоторые сравнения и обобщения, позволяющие выполнить в известной степени задачу расчета токов к. з.

В простейшем случае изолированной нейтрали, отсутствия расщепления фаз и жгутовой прокладки рассмотрению подлежит прежде всего случай трехфазного, а также двухфазного к. з. В рассматриваемом случае вероятность того и другого вида к. з. следует считать примерно одинаковой; кроме того, двухфазное к. з. часто переходит в трехфазное вследствие образования дуги.

Это отличает самолетную электрическую систему трехфазного тока от наземной воздушной линии передачи, в которой степень вероятности трехфазного к. з. составляет не более 10% и приближает по характеру протекания процесса к. з. к кабельной сети, где почти все виды к. з. переходят в трехфазные.

Для проведения расчета и составления расчетной схемы замещения системы должны быть прежде всего заданы параметры элементов системы или даны рекомендации по оценке и выбору их.

Для генератора на основании характеристики холостого хода подсчитывают отношение короткого замыкания (ОКЗ) и коэффициенты спрямленной характеристики ЭДС, по которым определяют продольное индуктивное сопротивление генератора

$$x_d = \frac{E}{\text{ОКЗ} \cdot I_a}$$

и оценивают (если не заданы) переходное и сверхпереходное индуктивные сопротивления, например, в соответствии с выражениями, приводимыми в литературе¹, а также индуктивные сопротивления обратной и нулевой последовательностей.

Можно отметить, что значения аналогичных параметров для наземных турбогенераторов, наиболее близких по характеристикам к самолетным синхронным генераторам, отличаются от параметров последних, однако в ряде случаев они принимаются в качестве первого приближения.

При определении сопротивлений самолетной сети следует в отличие от наземных сетей учитывать активные сопротивления; при этом в случае применения проводов сечением $q > 25 \text{ мм}^2$ эти сопротивления значительно увеличиваются вследствие явления поверхностного эффекта.

Реактивные сопротивления прямой (обратной) и нулевой последовательностей должны определяться с учетом количества проводов в фазе, взаимного расположения фаз (в плоскости или треугольнике) и положения их относительно корпуса. В общем случае полные сопротивления, особенно нулевой последовательности, зависят также от материала и конструкции обшивки, т. е. последнюю следует учитывать при определении активных и реактивных сопротивлений.

В наиболее простых случаях жгутовой прокладки фаз и изолированных нейтралей благодаря взаимной компенсации влияния обшивки на каждую из фаз сопротивления прямой (обратной) последовательности определяются без учета влияния обшивки².

Сопротивления нагрузок, точное определение которых в режиме к. з. представляет большие трудности вследствие непостоянства сопротивлений асинхронных двигателей (зависимости их от напряжений и токов к. з.), учитывают при расчете токов к. з. в виде эквивалентных индуктивных или полных сопротивлений, в зависимости от $\cos \varphi$ нагрузки. Неучет нагрузок вообще приводит к большим ошибкам при расчете токов к. з.

Расчеты токов к. з. проводились для простых радиальных схем, питаемых несколькими генераторами на основании метода спрямленных характеристик, при этом расчетная схема приводилась к простейшему виду путем последовательного и параллельного сложения элементов, при котором эквивалентные ЭДС и сопротивления можно определять из простых выражений типа

$$\dot{E}_0(k) = \frac{\dot{E}_k Z_{k+1} + \dot{E}_{k+1} Z_k}{Z_k + Z_{k+1}},$$

$$Z_0(k) = \frac{Z_k Z_{k+1}}{Z_k + Z_{k+1}},$$

$$i_0(k) = \frac{\dot{E}_0(k)}{Z_{10}(k) + Z_{20}(k)},$$

¹ С. А. Ульянов, Короткие замыкания в электрических системах, ГЭИ, 1952.
² В. В. Андреев, Исследование сопротивлений прямой и нулевой последовательности самолетной трехфазной сети. Канд. диссертация, МАИ, 1949.

в которых нагрузочные ветви рассматриваются как генераторы с $E=0$. При этом следует иметь в виду, что вследствие учета нагрузок ЭДС $E_a(t)$ меняется в зависимости от места к. з.

В зависимости от типа и назначения релейной защиты для ее проектирования необходимо знание как установившихся, так и переходных и сверхпереходных токов к. з. в виде начальных значений, а также для любого момента времени.

Установившиеся значения тока к. з. должны определяться, как правило, с учетом регуляторов напряжения, причем в силу малых внешних сопротивлений цепи к. з. в радиальных самолетных электрических системах имеет место, как показывают расчеты, режим предельного возбуждения, в котором для расчета токов следует вводить генераторы их предельными ЭДС E_{np} и сопротивлениями

$$x_{np} = \frac{E_{np}}{OKZ I_{в, np}}$$

Определение начальных значений переходных и сверхпереходных токов I_0' и I_0'' производится по аналогии с определением установившихся токов путем замены ЭДС E_a и индуктивного сопротивления x_d генераторов, соответственно, значениями E_d' ; x_d' — для переходного режима и E_d'' ; x_d'' — для сверхпереходного режима.

Определение токов к. з. в любой момент времени целесообразно производить, принимая экспоненциальный закон изменения свободных составляющих переходного и сверхпереходного тока, тем более что действительный характер изменения их очень близок к принятому. Тогда без учета регуляторов напряжения соответствующие мгновенные значения будут

$$i'(t) = (I_0' - I_{уст}') e^{-\frac{t}{T_d'}} + I_{уст}'$$

$$i''(t) = (I_0'' - I_0') e^{-\frac{t}{T_d''}} + (I_0' - I_{уст}') e^{-\frac{t}{T_d'}} + I_{уст}'$$

где T_d' — постоянная времени обмотки возбуждения, определяющая затухание переходного тока;

T_d'' — постоянная времени успокоительной обмотки, определяющая затухание сверхпереходного тока.

Вообще говоря, постоянные времени T_d' и T_d'' зависят от удаленности к. з. и могут быть оценены по формулам вида

$$T_d' = T_b \frac{x_d' + x}{x_d + x}$$

(T_b — постоянная времени обмотки возбуждения при разомкнутой статорной обмотке). Однако при отсутствии данных относительно T_b или экспериментальных данных для T_d' и T_d'' целесообразно считать

последние постоянными, а также принимать в соответствии с данными типовых самолетных генераторов. Так в этой работе для проверки использовались данные из литературы¹.

Расчеты показали, что постоянные времена T_d' , T_d'' в зависимости от удаленности к. з. изменяются в небольших пределах, не более 10%.

При наличии регуляторов напряжения мгновенные значения токов к. з. содержат дополнительную составляющую² типа

$$(I_{уст. пр} - I_{уст}) \left[1 - \frac{T_d' e^{-\frac{t}{T_d'}} - T_{в. в} e^{-\frac{t}{T_{в. в}}}}{T_d' - T_{в. в}} \right],$$

где $I_{уст. пр}$ — установившееся значение тока к. з. при предельном возбуждении;

$T_{в. в}$ — постоянная времени обмотки возбуждения возбуждения.

Можно отметить, что при имеющем место соотношении $\frac{T_d'}{T_{в. в}} > 1$ значение выражения, заключенного в квадратных скобках, при $t \leq 1$ сек. не превышает примерно 10%, что означает, что в условиях быстродействующей защиты учет этой дополнительной составляющей практически не потребует (в силу медленного увеличения возбуждения).

Как показали расчеты, проведенные для сравнения по индуктивным и полным сопротивлениям (модулям и комплексным выражениям), их целесообразно вести, используя индуктивные сопротивления x (для менее точных расчетов) или модули z (в случае более точных расчетов), так как углы полных эквивалентных сопротивлений Z , уменьшаясь по мере удаления места к. з., сохраняют значения не менее 70°.

Для радиальной сети с питанием от двух генераторов (30 кВА, 120/208 В, $n_{ср} = 6000$ об/мин) получены установившиеся значения токов трехфазного к. з. (в относительных единицах) — от 3,22 до 3,00; то же для двухфазного к. з. — от 4,98 до 4,32 (в короткозамкнутых фазах); значения переходных токов трехфазных к. з. — от 5,95 до 5,62; то же для сверхпереходных токов — от 8,14 до 7,20; начальные значения переходных токов двухфазного к. з. — от 6,62 до 5,78; то же для сверхпереходных токов — от 7,44 до 6,24.

Здесь большие значения относятся к месту к. з., ближайшему к генераторам; меньшие — к наиболее удаленному.

Следует отметить специфическое распределение токов к. з. в самолетной электрической системе, при котором в силу малых сопротивлений участков магистральной линии, ток к. з. фактически не от-

¹ V. C. Holloway, Transient Characteristics of Aircraft A. C. Generators, Tr. AIEE, ч. II, 1954.

² С. А. Ульянов, Короткие замыкания в электрических системах, ГЭИ, 1952.

сасывается нагрузками, т. е. ток магистральной линии до места к. з. одинаков.

Естественно, что переходные и сверхпереходные режимы затухают, т. е. режим устанавливается значительно быстрее в самолетной электрической системе по сравнению с наземными в силу меньших соотношений $\frac{L}{r}$ (малых значений постоянных времени T_d' , T_d'' и т. п.). Как правило, время протекания переходного процесса достаточно для действия соответствующей релейной аппаратуры, а значения токов к. з., в том числе и установившиеся, позволяют спроектировать надежно действующую релейную защиту от коротких замыканий.

Очевидна также необходимость рассмотрения всех других возможных аварийных режимов трехфазной самолетной электрической системы с учетом результатов и достижений наземной электротехники.

ЛИТЕРАТУРА

1. Ульянов С. А., Короткие замыкания в электрических системах, ГЭИ, 1952.
2. Щедрин Н. Н., Ульянов С. А., Задачи по расчету коротких замыканий, ГЭИ, 1955.
3. Андреев В. В., Исследование сопротивлений прямой и нулевой последовательности самолетной трехфазной сети, Канд. диссертация, МАИ, 1949.
4. Кулебакин В. С., Морозовский В. Т., Синдеев И. М., Электропитание самолетов, Оборонгиз, 1956.
5. Ехпер D. W., Сингер G. H., Impedance Data for 400-Cycle Aircraft Distribution Systems, Tr. AIEE, ч. II, 1952.
6. Holloway V. C., Transient Characteristics of Aircraft A. C. Generators, Tr. AIEE, ч. II, 1954.
7. Истратов В. Н. и Каменский А. В., Расчет сопротивлений самолетных сетей трехфазного тока, Труды МАИ № 57, 1956.

Канд. техн. наук А. В. КАМЕНСКИЙ, инж. В. Г. ТЕР-ЗАХАРЯН

РАСЧЕТ НЕСИММЕТРИИ ТОКОВ ПРИ ОБРЫВАХ ФАЗЫ

В связи с наличием в низковольтных распределительных сетях трехфазного тока потребителей, не допускающих значительной асимметрии токов, представляет интерес методика расчета несимметрии, возникающей при обрыве фазы.

Согласно существующему стандарту на электрические машины трехфазная электрическая система считается симметричной по току, если отношение тока обратной последовательности к току прямой последовательности не превышает 5%. Очевидно, что при радиальном питании и изолированной нейтрали системы потеря одной из фаз на линии до потребителя ведет к стопроцентной несимметрии и расчет в данном случае не требуется. Потребитель при этом должен быть переключен на резервное питание. Однако всегда возможны обрывы фазы у одного из параллельно работающих генераторов или у соседних потребителей и на параллельных ответвлениях, что приведет к несимметрии токов, значительно отличающейся от стопроцентной.

Вопрос о переключении потребителя в этом случае должен быть решен после проведения соответствующих расчетов, позволяющих выявить величину возникающей несимметрии и определить ее допустимость.

Рассмотрению подлежат два режима работы:

- 1) нейтраль системы изолирована;
- 2) нейтраль системы заземлена.

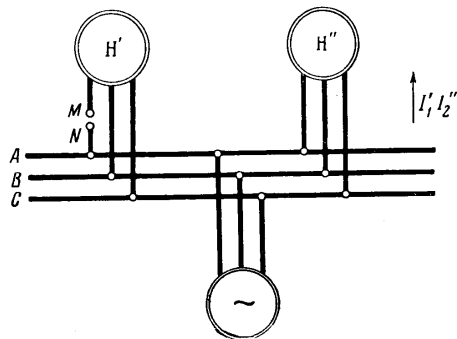
НЕЙТРАЛЬ СИСТЕМЫ ИЗОЛИРОВАНА

Обрыв фазы у потребителя

На фиг. 1 приведена схема сети, соответствующая обрыву фазы у параллельно включенного потребителя H' . Рассматривается появляющаяся несимметрия токов у потребителей H'' , количество которых может меняться. Переменным здесь является эквивалентное сопротивление прямой последовательности потребителей Z''_{1H} . Коэффициент несимметрии m_2 , определяемый как отношение тока обратной последовательности I''_{2H} к току прямой последовательности I''_{1H} для вторых потребителей, выводится в функции отношения сопро-

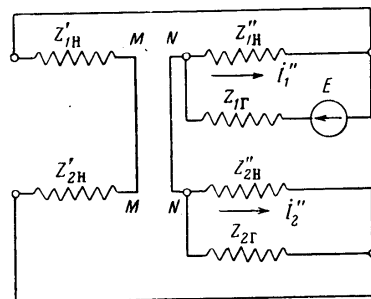
тивления прямой последовательности потребителя Н' к сопротивлению прямой последовательности потребителя Н'', т. е. в функции

$$K_H = \frac{Z'_{1H}}{Z''_{1H}}$$



Фиг. 1. Принципиальная схема сети с изолированной нейтралью (обрыв фазы у потребителя).

После этих предварительных замечаний переходим непосредственно к выводу расчетной формулы. На фиг. 2 приведена комплексная схема замещения сети, изображенной на фиг. 1.



Фиг. 2. Комплексная схема замещения сети, изображенной на фиг. 1.

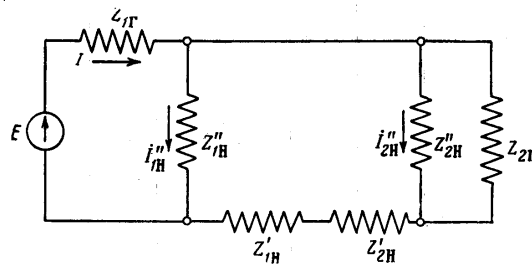
В целях наглядности и удобства расчета последняя схема может быть представлена так, как это показано на фиг. 3. Так как нас интересуют не абсолютные значения токов, а их отношения, то задачу можно решить в общем виде, не прибегая к числовым значениям.

Распределяя токи по ветвям схемы, изображенной на фиг. 3, и полагая сопротивление прямой последовательности потребителя H' равным сопротивлению обратной последовательности, т. е. $Z'_{1H} = Z'_{2H}$, получим

$$i'_{1H} = i \frac{2Z'_{1H} + \frac{Z'_{2H}Z_{2\Gamma}}{Z'_{2H} + Z_{2\Gamma}}}{Z'_{1H} + 2Z'_{1H} + \frac{Z'_{2H}Z_{2\Gamma}}{Z'_{2H} + Z_{2\Gamma}}}; \quad (1)$$

$$i'_{2H} = (i - i'_{1H}) \frac{Z_{2\Gamma}}{Z_{2\Gamma} + Z'_{2H}}, \quad (2)$$

где Z_H — сопротивление нагрузки; Z_Γ — сопротивление генератора.



Фиг. 3. Комплексная схема замещения сети, изображенной на фиг. 1, приведенная к удобному для расчета виду.

Коэффициент несимметрии определяется как отношение

$$m_2 = \frac{i'_{2H}}{i'_{1H}} = \frac{Z'_{1H}Z_{2\Gamma}}{2Z'_{1H}Z_{2\Gamma} + 2Z'_{1H}Z'_{2H} + Z_{2\Gamma}Z'_{2H}}$$

или, полагая

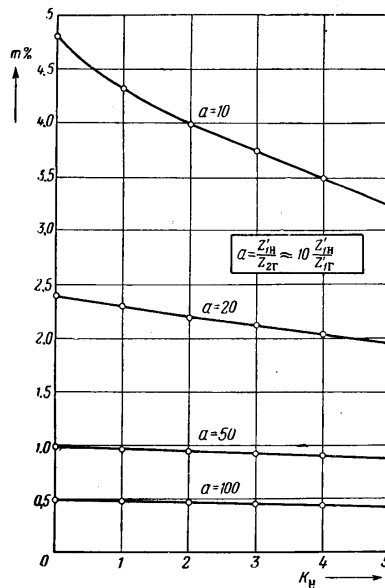
$$K_H = \frac{Z'_{1H}}{Z'_{2H}} \quad \text{и} \quad a = \frac{Z'_{2H}}{Z_{2\Gamma}},$$

имеем

$$m_2 = \frac{1}{1 + 2\left(K_H + \frac{Z'_{2H}}{Z_{2\Gamma}}\right)} = \frac{1}{1 + 2(K_H + a)},$$

что верно при $Z'_{1H} = Z'_{2H}$. При равенстве углов полных сопротивлений ветвей комплексной схемы замещения коэффициент m_2 действителен.

По этой формуле построены кривые фиг. 4. Из кривых следует, что при принятых отношениях $a=10, 20, 30...$ несимметрия не превышает 4%.



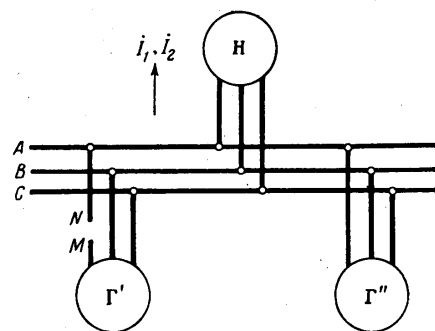
Фиг. 4. Кривые коэффициента несимметрии $m\%$ для сети с изолированной нейтралью (обрыв фазы у потребителя).

Обрыв фазы одного из параллельно работающих генераторов

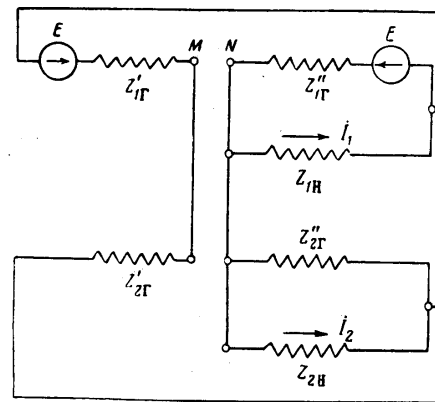
На фиг. 5 представлена схема сети, соответствующая обрыву фазы генератора Г'. Число неповрежденных генераторов изменяется от одного до трех. Случай отключения всех исправных генераторов не рассматривается, так как несимметрия при этом составит 100%.

Комплексные схемы замещения приведены на фиг. 6 и 7. Определяя аналогично предыдущему коэффициент несимметрии m_2 , получим

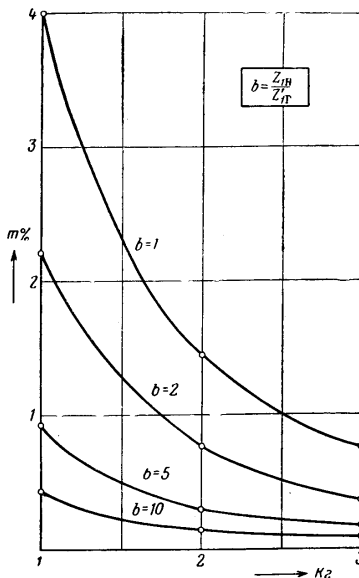
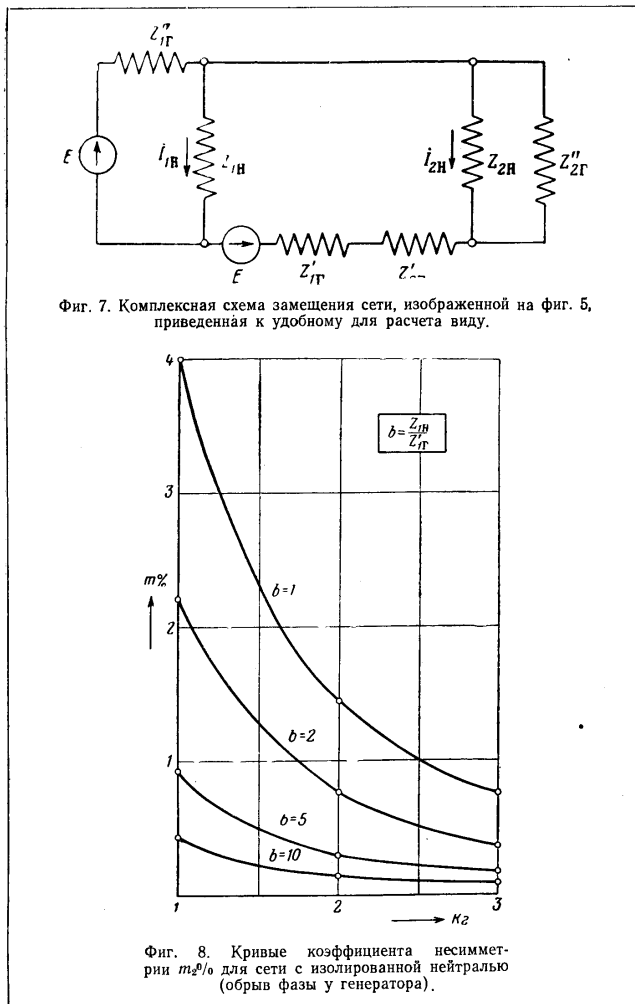
$$m_2 = \frac{-Z_{2Г}^* Z_{1Г}^*}{Z_{2Г}^* Z_{2Н} + (Z_{2Н} + Z_{2Г}) (Z_{1Г} + Z_{1Г} + Z_{2Г})}$$



Фиг. 5. Принципиальная схема сети с изолированной нейтралью (обрыв фазы у генератора).



Фиг. 6. Комплексная схема замещения сети, изображенной на фиг. 5.



или, обозначив $K_{\Gamma} = \frac{Z'_{1\Gamma}}{Z_{1\Gamma}}$ и полагая

$$Z_{1\Gamma} = 10Z_{2\Gamma}; Z_{1H} = Z_{2H} \text{ и } b = \frac{Z_{1H}}{Z_{1\Gamma}},$$

получим

$$m_2 = - \frac{1}{1 + 1,1K_{\Gamma} + 11K_{\Gamma}(1 + K_{\Gamma})b}.$$

По этой формуле построены кривые фиг. 8. Из этих кривых следует, что при принятых значениях $b = 1 \div 10$ обрыв фазы одного из параллельно работающих генераторов не вызывает значительной несимметрии токов у потребителя.

Однако в каждом конкретном случае следует рекомендовать проведение поверочных расчетов, учитывающих действительные сопротивления генераторов сети и нагрузки.

Максимальная несимметрия для первого случая соответствует 5%, а для второго 4% (см. фиг. 8), откуда следует, что при принятых отношениях сопротивлений, несимметрия не превышает допустимых стандартом величин.

НЕЙТРАЛЬ СИСТЕМЫ ЗАЕМЛЕНА

Схемы сети при заземленных нулевых точках и соответствующие комплексные схемы замещения приведены на фиг. 9, 10. На фиг. 11 приведены схемы, преобразованные к удобному для расчета виду.

Определяя коэффициент несимметрии для данного случая как отношение тока нулевой последовательности к току прямой последовательности, для схемы, приведенной на фиг. 9, получим

$$m_0 = \frac{Z'_{1H}}{\left[Z'_{1H} \left(1 + \frac{Z_0}{Z_2} \right) + Z_0 \right] \left(1 + \frac{Z'_{0H}}{Z_{0\Gamma}} \right)},$$

где

$$Z_2 = Z_{2H} + \frac{Z'_{2H} Z_{2\Gamma}}{Z_{2H} + Z_{2\Gamma}}$$

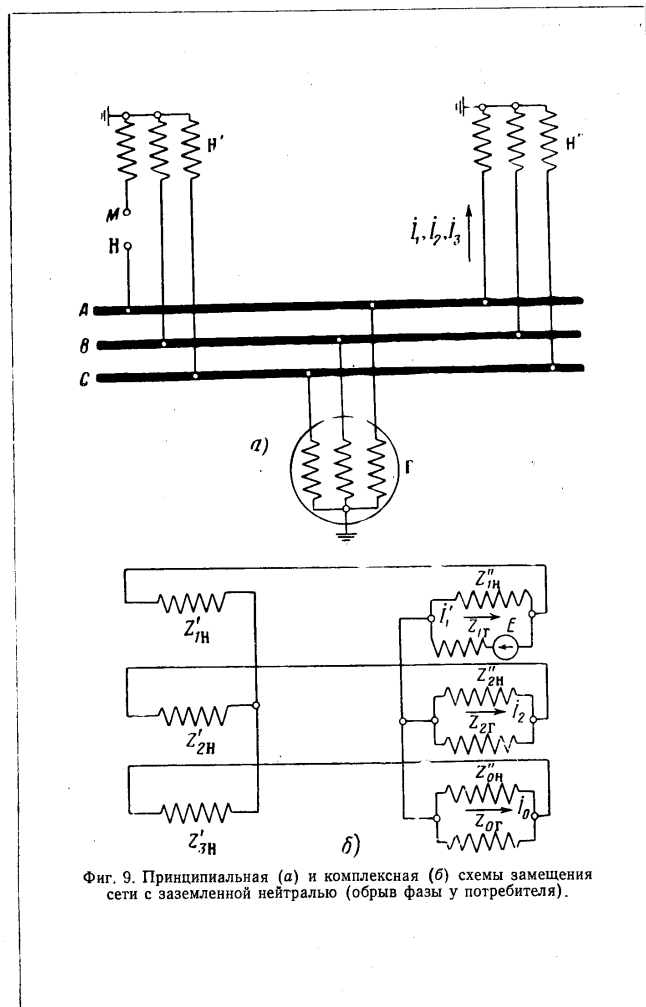
и

$$Z_0 = Z'_{0H} + \frac{Z'_{0H} Z_{0\Gamma}}{Z'_{0H} + Z_{0\Gamma}}.$$

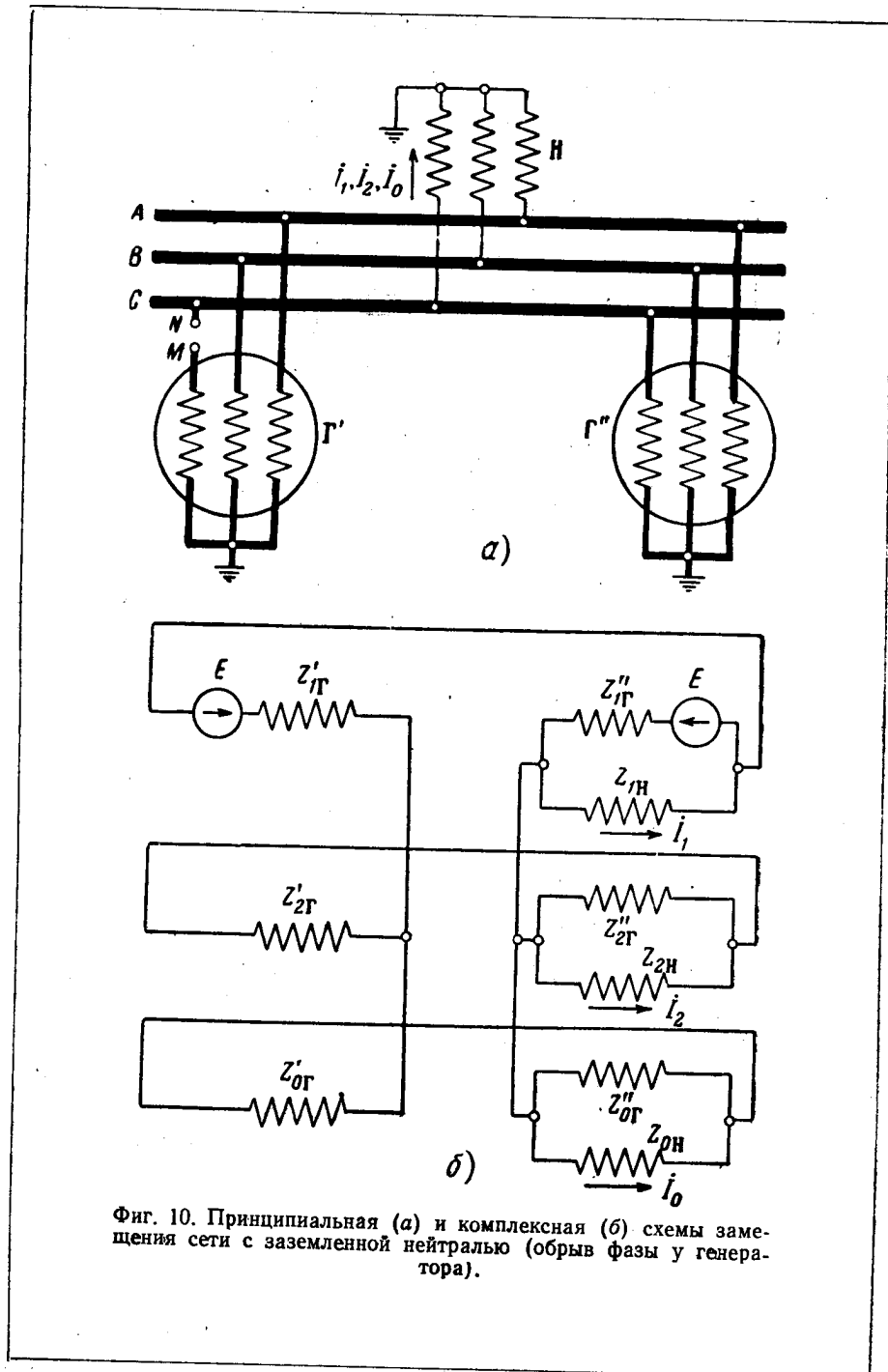
Аналогично для схемы, приведенной на фиг. 10, коэффициент несимметрии можно получить из следующего уравнения:

$$I_{1H} Z'_{1\Gamma} = I_{0H} \left[Z_{0H} + Z'_{0\Gamma} \left(\frac{Z'_{0\Gamma}}{Z_{0H}} + 1 \right) \right] + I'_{1\Gamma} (Z'_{1\Gamma} + Z'_{1\Gamma}),$$

$$m_0 = \frac{I_{1H} Z'_{1\Gamma} - I'_{1\Gamma} (Z'_{1\Gamma} + Z'_{1\Gamma})}{I_1 \left[Z_{0H} + Z'_{0\Gamma} \left(\frac{Z'_{0\Gamma}}{Z_{0H}} + 1 \right) \right]}.$$

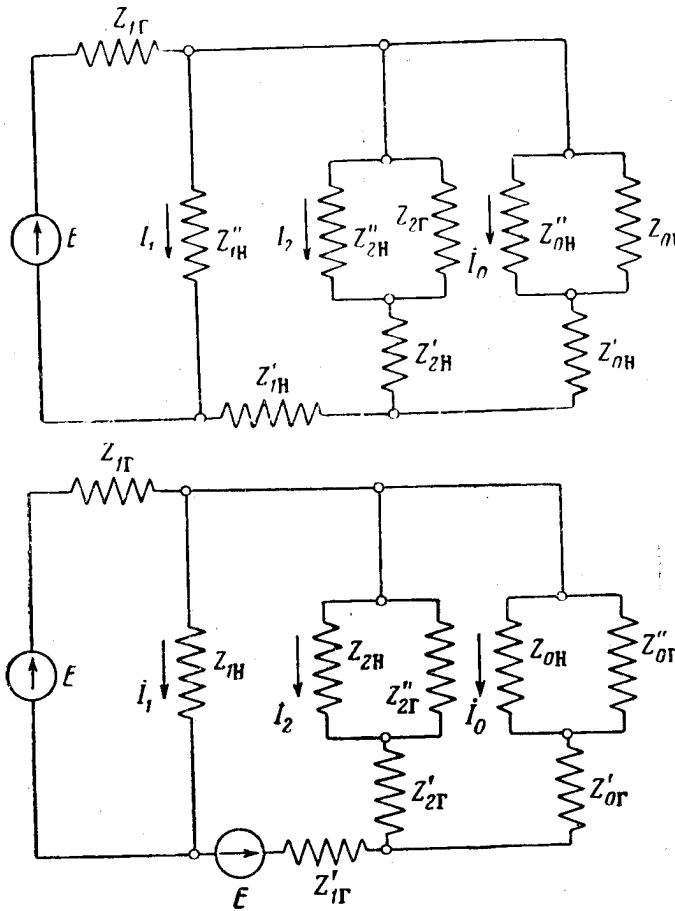


Фиг. 9. Принципиальная (а) и комплексная (б) схемы замещения сети с заземленной нейтралью (обрыв фазы у потребителя).



Фиг. 10. Принципиальная (а) и комплексная (б) схемы замещения сети с заземленной нейтралью (обрыв фазы у генератора).

Так как в этом случае коэффициент несимметрии зависит от многих параметров, то графическое представление получается более сложным, чем в предыдущем, и здесь не приводится. Полученные



Фиг. 11. Комплексные схемы замещения сети, изображенной на Фиг. 9 и 10, приведенные к удобному для расчета виду.

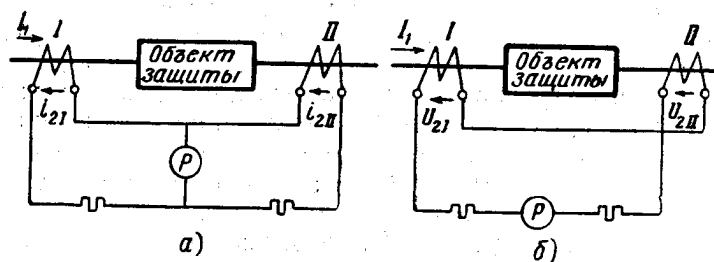
выше соотношения могут быть использованы при расчетах несимметрии токов низковольтных потребителей, вызванных обрывами фазы в различных точках сети.

Канд. техн. наук А. Б. ТИМОФЕЕВ, инж. В. Г. ТЕР-ЗАХАРЯН

ТРАНСФОРМАТОРЫ ТОКА ДЛЯ ПИТАНИЯ РЕЛЕ ДИФФЕРЕНЦИАЛЬНЫХ ЗАЩИТ САМОЛЕТНОЙ ЭЛЕКТРОСЕТИ

Для питания реле защиты самолетной электросети следует применять трансформаторы тока, значительно отличающиеся от измерительных.

Трансформаторы работают в комплекте с реле, которое имеет один определенный ток или напряжение срабатывания. Этому току (напряжению) срабатывания должен соответствовать определенный первичный ток. При этом несущественна погрешность в коэффициенте трансформации при токах, отличающихся от тока трогания защиты.



Фиг. 1. Принципиальные схемы включения трансформаторов.
а—схема на циркуляцию токов, б—схема на равновесие напряжений.

Следовательно, в отличие от трансформаторов тока, работающих на измерительные приборы, для которых важно, чтобы погрешность в коэффициенте трансформации не превосходила определенной заданной величины при изменении первичного тока в пределах измерения, для трансформаторов, работающих на реле, нужно обеспечить определенный коэффициент трансформации при одном токе и отстроить реле дифференциальной защиты от токов (напряжений) небаланса при сквозных токах коротких замыканий. Вес таких трансформаторов может быть получен в пределах нескольких десятков граммов.

Конструктивные данные трансформатора в значительной степени зависят от схемы включения их в защите. Основными схемами включения являются схемы, приведенные на фиг. 1, а и б.

Известно, что в наземных установках предпочтение обычно отдается схеме *a*. Для самолетных установок, видимо, следует предпочесть схему *b* (на равновесие напряжений).

В пользу такого решения можно привести следующие соображения.

1. При схеме *b* обрыв вспомогательных проводов не вызовет ложных срабатываний защиты в то время, как при схеме *a* обрыв вторичных проводов вызывает ложное срабатывание защиты. Правда, при схеме *b* замыкание между вторичными проводами может вызвать ложное срабатывание защиты, но при монтаже вторичных проводов на достаточном расстоянии друг от друга замыкание между проводами может произойти только через корпус самолета (в двух точках), что значительно менее вероятно, чем обрыв одного провода (при схеме *a*). Кроме того, такое двойное замыкание на корпус происходит обычно через два переходных сопротивления, что сильно снижает возможность ложной работы защиты.

2. Схема *a* вызывает необходимость применения реле со значительно меньшим сопротивлением обмотки, чем схема *b*. Имеющиеся в настоящее время чувствительные реле обычно имеют сравнительно большое сопротивление обмотки и поэтому применение таких реле целесообразнее в схеме *b*.

3. Сечение вспомогательных проводов защиты при схеме *a* должно быть больше, чем при схеме *b*.

4. При схеме *b* целесообразнее применение сравнительно больших магнитных индукций в сердечниках трансформаторов, чем при схеме *a*. Опасность ложного срабатывания защиты вследствие увеличения тока небаланса в случае переходных процессов при этом отсутствует благодаря весьма малой постоянной времени цепи защиты.

Как показал опыт, в случае применения насыщенных сердечников чувствительность малорелейных защит, питаемых от трансформаторов, используемых в качестве сумматоров, почти одинакова при различных видах повреждений (трех-, двух- и однофазные короткие замыкания) в то время, как в случае ненасыщенных сердечников, чувствительность защиты в зависимости от вида повреждения может изменяться в три раза.

Из сказанного выше следует, что для рассматриваемых трансформаторов можно допустить значительные индукции в сердечнике и получить благодаря этому выигрыш в весе. При этом для уменьшения токов небаланса при сквозных коротких замыканиях необходимо, чтобы каждые два трансформатора, включенные в схему дифференциальной защиты, были бы выполнены одинаково (одинаковый материал магнитопроводов, одинаковые размеры, одинаковые числа витков обмотки и т. п.). Существенным расчетным условием является мощность срабатывания реле. Чем меньше мощность срабатывания, тем меньше меди в обмотке (меньше сечение или число витков). Некоторое уменьшение сечения магнитопровода происходит также при увеличении первичного тока трогания защиты.

Проведенные расчеты показали, что при практически возможных токах короткого замыкания и использовании изготавливаемых ныне

высокочувствительных реле размеры сечения магнитопровода получаются настолько малыми, что механическая жесткость трансформатора оказывается недостаточной. В этих случаях приходится выбирать сечение магнитопровода, руководствуясь конструктивными соображениями.

При необходимости иметь первичную обмотку из нескольких витков можно соответствующее число раз пропустить провод линии через окно магнитопровода, но при этом необходимо увеличивать размеры окна. Как показали проверочные расчеты, при первичном токе трогания защиты порядка 100 а пропускать провод линии через окно трансформатора более одного раза нецелесообразно, так как магнитодвижущая сила получается достаточной для создания необходимого магнитного потока, а магнитопровод при этом имеет малый вес вследствие малого размера окна.

Задача инженерного расчета трансформатора не может быть решена однозначно, т. е. для исходных данных можно найти несколько трансформаторов с различными сечениями магнитопровода и различными числами витков вторичной обмотки.

Вследствие нелинейной зависимости между индукцией и напряженностью магнитного поля, а также большого количества причин, влияющих на расчет, наиболее простым и наглядным является расчет методом подбора наиболее выгодного варианта.

При этом нет необходимости стремиться к большой точности и вполне достаточно ограничиться двумя-тремя вариантами расчета поскольку расчетные величины, характеризующие работу трансформатора, полученные в результате расчета, могут отличаться от таковых у изготовленного трансформатора по следующим причинам:

а) кривая намагничивания материала сердечника может несколько отличаться от кривой намагничивания принятой при расчете;

б) могут ухудшаться магнитные свойства материала сердечника при резке и штамповке листов;

в) воздушные зазоры в стержнях могут отличаться от запроектированных.

При этом следует предусматривать первичный ток трогания защиты на 10÷15% меньшим, чем требуемый, так как этот ток можно увеличить при настройке защиты посредством изменения добавочных сопротивлений (увеличения — при схеме, изображенной на фиг. 1,а, и уменьшения — при схеме, приведенной на фиг. 1,б), а если первичный ток трогания защиты для изготовленных трансформаторов окажется больше заданного, то уменьшить его без переделки трансформаторов не всегда представляется возможным.

В заключение следует отметить, что вопросы расчета и конструирования рассмотренных трансформаторов для самолетной электросети еще мало исследованы и ввиду несомненной актуальности желательна дальнейшая разработка этих вопросов.

Инж. А. А. КОВЗАН

ОБ ИСПОЛЬЗОВАНИИ ОБШИВКИ САМОЛЕТА В КАЧЕСТВЕ ОДНОГО ИЗ ФАЗНЫХ ПРОВОДОВ

Теоретические предпосылки применения трехфазных сетей с использованием земли в качестве одного из фазных проводов относятся к двадцатым годам (О. Маур «Die Erde als Wechselstromleiter», E. T. Z., 1925).

Первые опыты до ДПЗ (два провода—земля) проводились преимущественно в электрификации сельского хозяйства уже в начале тридцатых годов.

Как показывает длительная эксплуатация таких систем, двухпроводные трехфазные сети (сети ДПЗ) имеют ряд несомненных преимуществ (экономия трети меди, части аппаратуры, повышение КПД электропередачи и т. д.), хотя и обладают некоторыми недостатками (асимметрия электропередачи, некоторое ухудшение условий безопасности и пр.).

При оборудовании самолетов системой трехфазного переменного тока система ДПЗ (в качестве «земли» может служить металлическая обшивка — корпус самолета) должна привлечь внимание благодаря своим серьезным преимуществам в весе перед обычной трехпроводной системой¹.

Такая двухпроводная система «два провода — корпус самолета» (ДПК) может найти практическое применение в авиации.

Асимметрия токов и напряжений при применении ДПК, как показывают расчеты, не выходит за пределы допустимой (т. е. 5%). Кроме того, как будет показано ниже, в некоторых случаях можно путем введения добавочных сопротивлений добиться того, что система ДПК будет практически аналогична обычной трехфазной транспонированной линии.

Следует отметить, что вся электрическая сеть, питающаяся от одних и тех же источников энергии (электрически связанная система), должна полностью переводиться на работу по схеме ДПК.

Выполнение каких-либо участков по трехпроводной схеме не оправдывается, так как эти участки будут все равно обладать особенностями сетей ДПК.

¹ В случае изолированной нейтрали генераторов.

**ЭЛЕКТРИЧЕСКИЙ РАСЧЕТ ЛИНИЙ В СИСТЕМЕ
«ДВА ПРОВОДА — КОРПУС САМОЛЕТА» (ДПК)**

Рассмотрим электрическую цепь, изображенную на фиг. 1¹.
Принимая поочередно сопротивление одной нагрузки равным нулю и одновременно сопротивление другой — бесконечности, т. е.

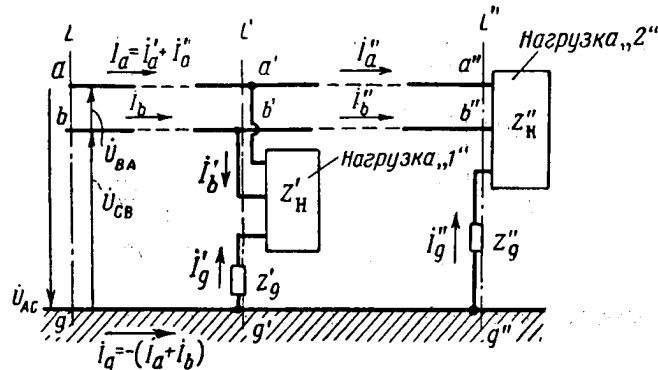
$$Z'_{HA} = Z'_{HB} = Z'_{HC} = 0, \quad Z''_{HA} = Z''_{HB} = Z''_{HC} = \infty \text{ и}$$

$$Z''_{HA} = Z''_{HB} = Z''_{HC} = 0, \quad Z'_{HA} = Z'_{HB} = Z'_{HC} = \infty,$$

можно напряжения в начале линии при токах, равных соответственно

$$i'_a, i'_b, i'_g \text{ и } i''_a, i''_b, i''_g$$

(фиг. 2, а, б), рассматривать как падения напряжений в линии.



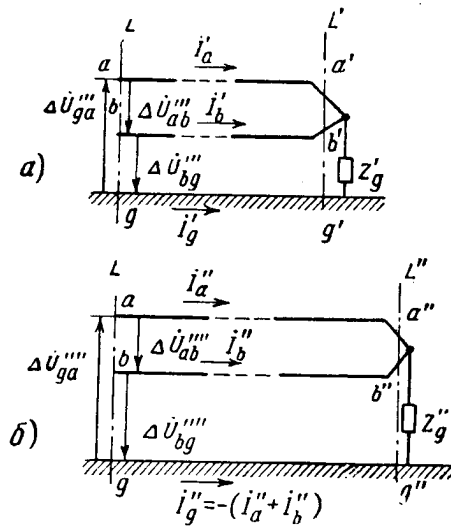
Фиг. 1. Принципиальная схема двухпроводной трехфазной системы ДПК с двумя нагрузками.

Проведем расчет последовательно для двух схем (фиг. 2, а и б). Очевидно наложение схем а и б друг на друга даст нам действительную картину распределения токов и падений напряжений в электрической линии, показанной на фиг. 1.

Схему, приведенную на фиг. 2, а, можно рассматривать как две петли „провод—корпус“, обтекаемые каждая токами i'_a и i'_b , связанные между собой взаимной индукцией и имеющие общее сопротивление Z'_g . Схему, изображенную на фиг. 2, б, можно рассматривать как две петли „провод—корпус“, обтекаемые токами i''_a и i''_b , связанные между собой взаимной индукцией и имеющие общее сопротивление Z''_g .

¹ Расчет цепи с одной нагрузкой в конце линии разными методами выполнен в работе Л. Е. Эбина и Ш. М. Алукера, Электрический расчет сетей с использованием земли в качестве одного из фазных проводов (Госэнергоиздат, М.—Л., 1949) и в работе В. В. Андреева, О двухпроводной трехфазной линии (Труды МАИ, № 57, 1956).

Если обозначим через Z'_{La} сопротивление самоиндукции петли „провод aa' —корпус“, через Z'_{Lb} —сопротивление самоиндукции петли „провод bb' —корпус“, через Z'_m —сопротивление взаимной индукции двух таких петель и соответственно через Z_{La} —сопротивление самоиндукции петли „провод aa'' —корпус“, через Z_{Lb} —сопротивление самоиндукции петли „провод bb'' —корпус“, через



Фиг. 2. Схема расчета линии ДПК с двумя нагрузками.

Z_m —сопротивление взаимной индукции двух петель, то падения напряжений в линии $\Delta U''_{ag}$, $\Delta U''_{bg}$, $\Delta U''_{ab}$ (см. фиг. 2,а) и $\Delta U'''_{ag}$, $\Delta U'''_{bg}$, $\Delta U'''_{ab}$ (см. фиг. 2,б) можно выразить так:

$$\left. \begin{aligned} \Delta U''_{ag} &= Z'_{La} i'_a + Z'_m i'_b - Z'_g i'_g, \\ \Delta U''_{bg} &= Z'_{Lb} i'_b + Z'_m i'_a - Z'_g i'_g, \\ \Delta U''_{ab} &= \Delta U''_{ag} - \Delta U''_{bg} = (Z'_{La} - Z_m) i'_a - (Z'_{La} - Z'_m) i'_b \end{aligned} \right\} (1)$$

и

$$\left. \begin{aligned} \Delta U'''_{ag} &= Z_{La} i''_a + Z_m i''_b - Z'_g i''_g, \\ \Delta U'''_{bg} &= Z_{Lb} i''_b + Z_m i''_a - Z'_g i''_g, \\ \Delta U'''_{cb} &= \Delta U'''_{ag} - \Delta U'''_{bg} = (Z_{La} - Z_m) i''_a - (Z_{Lb} - Z_m) i''_b. \end{aligned} \right\} (2)$$

Применяя принцип наложения для схем а и б фиг. 2 и приняв $Z'_{La} = Z'_{Lb} = Z'_L$, $Z_{La} = Z_{Lb} = Z_L$, получим линейные падения напря-

жений на участке LL'' для действительной схемы, приведенной на фиг. 1:

$$\left. \begin{aligned} \Delta \dot{U}_{ab} &= \Delta \dot{U}_{ab}'' + \Delta \dot{U}_{ab}' = \Delta \dot{U}_{ab}' + \Delta \dot{U}_{ab}'' \\ \Delta \dot{U}_{bg} &= \Delta \dot{U}_{bg}'' + \Delta \dot{U}_{bg}' = \Delta \dot{U}_{bg}' + \Delta \dot{U}_{bg}'' \\ \Delta \dot{U}_{ga} &= \Delta \dot{U}_{ga}'' + \Delta \dot{U}_{ga}' = \Delta \dot{U}_{ga}' + \Delta \dot{U}_{ga}'' \end{aligned} \right\} (3)$$

где

$$\left. \begin{aligned} \Delta \dot{U}_{ab}' &= (Z_L' - Z_m') (i_a - i_b), \\ \Delta \dot{U}_{bg}' &= Z_L' i_b + Z_m' i_a - Z_g' i_g', \\ \Delta \dot{U}_{ga}' &= -(Z_L' i_a + Z_m' i_b - Z_g' i_g'); \end{aligned} \right\} (4)$$

$$\left. \begin{aligned} \Delta \dot{U}_{ab}'' &= (Z_L'' - Z_m'') (i_a'' - i_b''), \\ \Delta \dot{U}_{bg}'' &= Z_L'' i_b'' + Z_m'' i_a'' - Z_g'' i_g'', \\ \Delta \dot{U}_{ga}'' &= -(Z_L'' i_a'' + Z_m'' i_b'' - Z_g'' i_g''); \end{aligned} \right\} (5)$$

$\Delta \dot{U}_{ab}'$, $\Delta \dot{U}_{bg}'$, $\Delta \dot{U}_{ga}'$ — линейные падения напряжений на участке линии LL' ;

$\Delta \dot{U}_{ab}''$, $\Delta \dot{U}_{bg}''$, $\Delta \dot{U}_{ga}''$ — то же на участке $L'L''$;

Z_L' и Z_m' — сопротивления само- и взаимной индукции петель „провод $a'a''$ — корпус“ и „провод $b'b''$ — корпус“; $Z_L = Z_L' + Z_L''$ и $Z_m = Z_m' + Z_m''$.

Разложим линейные падения напряжений $\Delta \dot{U}_{ab}$, $\Delta \dot{U}_{bg}$, $\Delta \dot{U}_{ga}$ в двухпроводной трехфазной линии (участок линии LL'') на симметричные составляющие:

$$\left. \begin{aligned} \Delta \dot{U}_{\lambda 1} &= \frac{1}{3} (\Delta \dot{U}_{ab} + a \Delta \dot{U}_{bg} + a^2 \Delta \dot{U}_{ga}), \\ \Delta \dot{U}_{\lambda 2} &= \frac{1}{3} (\Delta \dot{U}_{ab} + a^2 \Delta \dot{U}_{bg} + a \Delta \dot{U}_{ga}), \\ \Delta \dot{U}_{\lambda 0} &= \frac{1}{3} (\Delta \dot{U}_{ab} + \Delta \dot{U}_{bg} + \Delta \dot{U}_{ga}) = 0 \end{aligned} \right\} (6)$$

и, учитывая, что между линейными и фазными симметричными составляющими существуют зависимости

$$\left. \begin{aligned} \Delta \dot{U}_{\lambda 1} &= (1 - a^2) \Delta \dot{U}_{\phi 1}, \\ \Delta \dot{U}_{\lambda 2} &= (1 - a) \Delta \dot{U}_{\phi 2}, \end{aligned} \right\} (7)$$

где

$a = e^{j \frac{2\pi}{3}}$, $a^2 = e^{j \frac{4\pi}{3}}$ — фазные множители,

найдем симметричные составляющие фазных падений напряжений, выраженные через фазные токи:

$$\left. \begin{aligned} \Delta \dot{U}_{\phi 1} &= \frac{1}{3} [(Z'_L + aZ'_m - a^2Z'_g) i'_a + (aZ'_L + Z'_m - a^2Z'_g) i'_b + \\ &\quad + (Z_L + aZ_m - a^2Z_g) i''_a + (aZ_L + Z_m - a^2Z_g) i''_b], \\ \Delta \dot{U}_{\phi 2} &= \frac{1}{3} [(Z'_L + a^2Z'_m - aZ'_g) i'_a + (a^2Z'_L + Z'_m - aZ'_g) i'_b + \\ &\quad + (Z_L + a^2Z_m - aZ_g) i''_a + (a^2Z_L + Z_m - aZ_g) i''_b], \\ \Delta U_{\phi 0} &\neq 0. \end{aligned} \right\} (8)$$

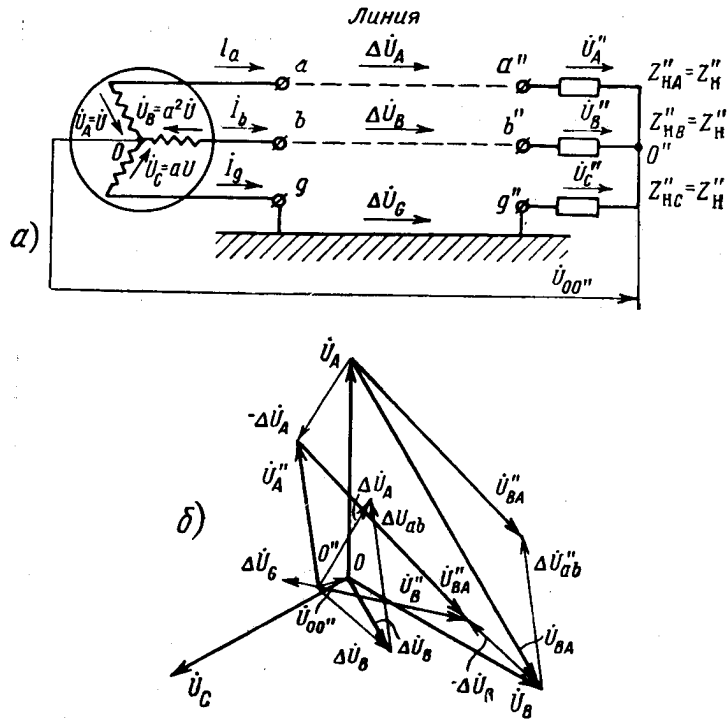
В общем случае двухпроводной трехфазной линии симметричная составляющая нулевой последовательности $\Delta \dot{U}_{\phi 0}$ не равна нулю. Поскольку величина $\Delta \dot{U}_{\phi 0}$ в дальнейшем не используется, метод нахождения ее в настоящей работе не рассматривается.

Найдем выражения фазных падений напряжений трехфазной системы на участке линии LL'' :

$$\left. \begin{aligned} \Delta \dot{U}_A &= (\Delta \dot{U}_{\phi 0} + \Delta \dot{U}_{\phi 1} + \Delta \dot{U}_{\phi 2}) = \frac{1}{3} [(2Z'_L - Z'_m + Z'_g) i'_a + \\ &\quad + (-Z'_L + 2Z'_m + Z'_g) i'_b + (2Z'_L - Z'_m + Z'_g) i''_a + \\ &\quad + (-Z_L + 2Z_m + Z_g) i''_b] + \Delta \dot{U}_{\phi 0}, \\ \Delta \dot{U}_B &= (\Delta \dot{U}_{\phi 0} + a^2 \Delta \dot{U}_{\phi 1} + a \Delta \dot{U}_{\phi 2}) = \frac{1}{3} [(-Z'_L + 2Z'_m + Z'_g) i'_a + \\ &\quad + (2Z'_L - Z'_m + Z'_g) i'_b + (-Z_L + 2Z_m + Z_g) i''_a + \\ &\quad + (2Z_L - Z_m + Z_g) i''_b] + \Delta \dot{U}_{\phi 0}, \\ \Delta \dot{U}_G &= (\Delta \dot{U}_{\phi 0} + a \Delta \dot{U}_{\phi 1} + a^2 \Delta \dot{U}_{\phi 2}) = \frac{1}{3} [(Z'_L + Z'_m + Z'_g) i'_g + \\ &\quad + (Z_L + Z_m + Z_g) i''_g] + \Delta \dot{U}_{\phi 0}. \end{aligned} \right\} (9)$$

Ввиду большой проводимости обшивки самолета, используемой в качестве фазного провода, падения напряжений в фазах будут различны. Легко убедиться, что падение напряжения ΔU_G будет меньше, чем ΔU_A и ΔU_B .

Считая напряжения, приложенные к линии, симметричными (симметричный трехфазный генератор с фазными напряжениями ($\dot{U}_A = \dot{U}$, $\dot{U}_B = a^2 \dot{U}$ и $\dot{U}_C = a \dot{U}$), можно определить линейные напряжения в конце линии на нагрузке „2“ (фиг. 3):



Фиг. 3. Определение фазных и линейных напряжений на нагрузке.

$$\left. \begin{aligned}
 \dot{U}_{BA}^* &= \dot{U}_{BA} + \Delta \dot{U}_{ab} = (a^2 - 1) \dot{U} + (Z'_L - Z'_m) (\dot{i}'_a - \dot{i}'_b) + \\
 &\quad + (Z'_L - Z'_m) (\dot{i}'_a - \dot{i}'_b), \\
 \dot{U}_{CB}^* &= \dot{U}_{cb} + \Delta \dot{U}_{bg} = a(1 - a) \dot{U} + (Z'_m + Z'_g) \dot{i}'_a + \\
 &\quad + (Z'_L + Z'_g) \dot{i}'_b + (Z'_m + Z'_g) \dot{i}'_a + (Z'_L + Z'_g) \dot{i}'_b, \\
 \dot{U}_{AC}^* &= \dot{U}_{AC} + \Delta \dot{U}_{ga} = (1 - a) \dot{U} - (Z'_L + Z'_g) \dot{i}'_a - \\
 &\quad - (Z'_m + Z'_g) \dot{i}'_b - (Z'_L + Z'_g) \dot{i}'_a - (Z'_m + Z'_g) \dot{i}'_b.
 \end{aligned} \right\} (10)$$

Если предположить симметричным приемник, включенный в „звезду“ в конце линии (т. е. $Z''_{HA} = Z''_{HB} = Z''_{HC} = Z''_H$), то можно установить фазные напряжения в конце линии по отношению к нейтральной точке, имея в виду, что для симметричной нагрузки

$$\dot{U}_A^* + \dot{U}_B^* + \dot{U}_C^* = 0,$$

$$\begin{aligned}
 \dot{U}'_A &= \frac{\dot{U}'_{AC} - \dot{U}'_{BA}}{3} = \dot{U} - \frac{1}{3} [(2Z'_L - Z'_m + Z'_g) \dot{i}'_a + \\
 &\quad + (-Z'_L + 2Z'_m + Z'_g) \dot{i}'_b + (2Z'_L - Z'_m + Z'_g) \dot{i}'_a + \\
 &\quad + (-Z'_L + 2Z'_m + Z'_g) \dot{i}'_b], \\
 \dot{U}'_B &= \frac{\dot{U}'_{BA} - \dot{U}'_{CB}}{3} = a^2 \dot{U} + \frac{1}{3} [(Z'_L - 2Z'_m - Z'_g) \dot{i}'_a + \\
 &\quad + (-2Z'_L + Z'_m - Z'_g) \dot{i}'_b + (Z'_L - 2Z'_m - Z'_g) \dot{i}'_a + \\
 &\quad + (-2Z'_L + Z'_m - Z'_g) \dot{i}'_b], \\
 \dot{U}'_C &= \frac{\dot{U}'_{CB} - \dot{U}'_{AC}}{3} = a \dot{U} - \frac{1}{3} [(Z'_L + Z'_m + 2Z'_g) \dot{i}'_g + \\
 &\quad + (Z'_L + Z'_m + 2Z'_g) \dot{i}'_g].
 \end{aligned} \tag{11}$$

Выражения (11) для случая симметричной нагрузки можно получить и другим способом, например, вычитая из фазных напряжений генератора соответствующие фазные падения напряжений (9) и учитывая напряжение $\Delta \dot{U}'_{\phi 0} = \dot{U}'_{00'}$ (см. фиг. 3).

Аналогичным образом определяются условия работы приемника электроэнергии «1» (см. фиг. 1). Не касаясь промежуточных расчетов, приведем окончательные выражения для падений напряжений в соответствующих фазах на участке линии LL' :

$$\begin{aligned}
 \Delta \dot{U}'_A &= \frac{1}{3} [(2Z'_L - Z'_m + Z'_g) \dot{i}'_a + (-Z'_L + 2Z'_m + Z'_g) \dot{i}'_b + \\
 &\quad + (2Z'_L - Z'_m) \dot{i}'_a + (-Z'_L + 2Z'_m) \dot{i}'_b] + \Delta \dot{U}'_{\phi 0}, \\
 \Delta \dot{U}'_B &= \frac{1}{3} [(-Z'_L + 2Z'_m + Z'_g) \dot{i}'_a + (2Z'_L - Z'_m + Z'_g) \dot{i}'_b + \\
 &\quad + (-Z'_L + 2Z'_m) \dot{i}'_a + (2Z'_L - Z'_m) \dot{i}'_b] + \Delta \dot{U}'_{\phi 0}, \\
 \Delta \dot{U}'_G &= \frac{1}{3} [(Z'_L + Z'_m + 2Z'_g) \dot{i}'_g + (Z'_L + Z'_m) \dot{i}'_g] + \Delta \dot{U}'_{\phi 0}.
 \end{aligned} \tag{12}$$

Если трехфазный генератор, присоединенный в начале линии, симметричный (см. стр. 107), то линейные напряжения на нагрузке «1» (фиг. 1) будут равны:

$$\begin{aligned}
 \dot{U}'_{BA} &= (a^2 - 1) \dot{U} + (Z'_L - Z'_m) (\dot{i}'_a - \dot{i}'_b), \\
 \dot{U}'_{CB} &= a(1 - a) \dot{U} + Z'_L \dot{i}'_b + Z'_m \dot{i}'_a - Z'_g \dot{i}'_g, \\
 \dot{U}'_{CA} &= (1 - a) \dot{U} - Z'_L \dot{i}'_a - Z'_m \dot{i}'_b + Z'_g \dot{i}'_g.
 \end{aligned} \tag{13}$$

Принимая нагрузку „1“ симметричной, включенной по схеме „звезда“ ($Z'_{HA} = Z'_{HB} = Z'_{HC} = Z'_H$), находим фазные напряжения \dot{U}'_A , \dot{U}'_B и \dot{U}'_C по отношению к нейтральной точке потребителя:

$$\left. \begin{aligned} \dot{U}'_A &= \frac{\dot{U}'_{AC} - \dot{U}'_{BA}}{3} = \dot{U} - \frac{1}{3} [(2Z'_L - Z'_m + Z'_g) i'_a + (-Z'_L + \\ &\quad + 2Z'_m + Z'_g) i'_b + (2Z'_L - Z'_m) i'_a + (-Z'_L + 2Z'_m) i'_b], \\ \dot{U}'_B &= \frac{\dot{U}'_{BA} - \dot{U}'_{CB}}{3} = a^2 \dot{U} - \frac{1}{3} [(-Z'_L + 2Z'_m + Z'_g) i'_a + \\ &\quad + (2Z'_L - Z'_m + Z'_g) i'_b + (-Z'_L + 2Z'_m) i'_a + (2Z'_L - Z'_m) i'_b], \\ \dot{U}'_C &= \frac{\dot{U}'_{CB} - \dot{U}'_{AC}}{3} = a \dot{U} - \frac{1}{3} [(Z'_L + Z'_m + 2Z'_g) i'_g + \\ &\quad + (Z'_L + Z'_m) i'_g]. \end{aligned} \right\} (14)$$

Интересны два частных случая, когда двухпроводная линия выполняется без добавочных сопротивлений Z'_g и Z'_g , введенных последовательно с фазами нагрузок, соединенными с обшивкой — корпусом самолета (см. фиг. 1), и второй, когда

$$Z'_g = Z'_L - 2Z'_m \text{ и } Z'_g = Z'_L - 2Z'_m^*.$$

Выражения, соответствующие первому случаю, получаются из предыдущих, положив $Z'_g = Z'_g = 0$.

Тогда падения напряжений в фазах линии для 1-го случая будут

$$\left. \begin{aligned} \Delta \dot{U}'_A &= \frac{1}{3} [(2Z'_L - Z'_m) i'_a + (-Z'_L + 2Z'_m) i'_b + (2Z'_L - Z'_m) i'_a + \\ &\quad + (-Z'_L + 2Z'_m) i'_b] + \Delta \dot{U}'_{\phi 0}, \\ \Delta \dot{U}'_B &= \frac{1}{3} [(-Z'_L + 2Z'_m) i'_a + (2Z'_L - Z'_m) i'_b + \\ &\quad + (-Z'_L + 2Z'_m) i'_a + (2Z'_L - Z'_m) i'_b] + \Delta \dot{U}'_{\phi 0}, \\ \Delta \dot{U}'_C &= \frac{1}{3} [(Z'_L + Z'_m) i'_g + (Z'_L + Z'_m) i'_g] + \Delta \dot{U}'_{\phi 0} \end{aligned} \right\} (15)$$

и

* Выбор дополнительных сопротивлений из условия возможной компенсации несимметрии двухпроводной линии будет показан далее при оценке несимметрии линии (см. стр. 115).

$$\left. \begin{aligned} \Delta \dot{U}'_A &= \frac{1}{3} [(2Z'_L - Z'_m) \dot{I}'_a + (-Z'_L + 2Z'_m) \dot{I}'_b] + \Delta \dot{U}'_{\phi 0}, \\ \Delta \dot{U}'_B &= \frac{1}{3} [(-Z'_L + 2Z'_m) \dot{I}'_a + (2Z'_L - Z'_m) \dot{I}'_b] + \Delta \dot{U}'_{\phi 0}, \\ \Delta \dot{U}'_G &= \frac{1}{3} (Z'_L + Z'_m) \dot{I}'_g + \Delta \dot{U}'_{\phi 0}. \end{aligned} \right\} (16)$$

Как и следовало ожидать, двухпроводная линия в данном случае оказывается несимметричной для обеих нагрузок. Следует заметить, что напряжения нулевой последовательности $\Delta \dot{U}'_{\phi 0}$ и $\Delta \dot{U}'_{\phi 0}$ в выражениях (15) и (16) не тождественны подобным величинам в выражениях (9) и (12).

Линейные и фазные напряжения при симметричных генераторе и нагрузках, соединенных «звездой», будут равными у потребителя «2»

$$\left. \begin{aligned} \dot{U}''_{BA} &= (a^2 - 1) \dot{U} + (Z'_L - Z'_m) (\dot{I}_a - \dot{I}_b) + (Z''_L - Z''_m) (\dot{I}''_a - \dot{I}''_b), \\ \dot{U}''_{CB} &= a(1 - a) \dot{U} + Z'_L \dot{I}_b + Z'_m \dot{I}_a + Z''_L \dot{I}''_b + Z''_m \dot{I}''_a, \\ \dot{U}''_{AC} &= (1 - a) \dot{U} - Z'_L \dot{I}_a - Z'_m \dot{I}_b - Z''_L \dot{I}''_a - Z''_m \dot{I}''_b; \end{aligned} \right\} (17)$$

$$\left. \begin{aligned} \dot{U}''_A &= \dot{U} - \frac{1}{3} [(2Z'_L - Z'_m) \dot{I}_a + (-Z'_L + 2Z'_m) \dot{I}_b + \\ &\quad + (2Z''_L - Z''_m) \dot{I}''_a + (-Z''_L + 2Z''_m) \dot{I}''_b], \\ \dot{U}''_B &= a^2 \dot{U} + \frac{1}{3} [(Z'_L - 2Z'_m) \dot{I}_a + (-2Z'_L + Z'_m) \dot{I}_b + \\ &\quad + (Z''_L - 2Z''_m) \dot{I}''_a + (-2Z''_L + Z''_m) \dot{I}''_b], \\ \dot{U}''_C &= a \dot{U} - \frac{1}{3} [(Z'_L + Z'_m) \dot{I}_g + (Z''_L + Z''_m) \dot{I}''_g] \end{aligned} \right\} (18)$$

и у потребителя «1»

$$\left. \begin{aligned} \dot{U}'_{BA} &= (a^2 - 1) \dot{U} + (Z'_L - Z'_m) (\dot{I}_a - \dot{I}_b), \\ \dot{U}'_{CB} &= a(1 - a) \dot{U} + Z'_L \dot{I}_b + Z'_m \dot{I}_a, \\ \dot{U}'_{AC} &= (1 - a) \dot{U} - Z'_L \dot{I}_a - Z'_m \dot{I}_b, \end{aligned} \right\} (19)$$

$$\left. \begin{aligned} \dot{U}'_A &= \dot{U} - \frac{1}{3} [(2Z'_L - Z'_m) \dot{I}_a + (-Z'_L + 2Z'_m) \dot{I}_b], \\ \dot{U}'_B &= a^2 \dot{U} - \frac{1}{3} [(-Z'_L + 2Z'_m) \dot{I}_a + (2Z'_L - Z'_m) \dot{I}_b], \\ \dot{U}'_C &= a \dot{U} - \frac{1}{3} (Z'_L + Z'_m) \dot{I}_g. \end{aligned} \right\} (20)$$

Интересно рассмотреть также выражения, аналогичные (15), (17), (18) и (16), (19), (20), для случая, когда последовательно с фазами нагрузок, соединенными с корпусом самолета, вводятся добавочные сопротивления $Z'_g = Z'_L - 2Z'_m$ и $Z''_g = Z'_L - 2Z'_m$ (см. фиг. 1).

Потери напряжений в каждой фазе линии ДПК будут равными для нагрузки „2“ в конце линии

$$\left. \begin{aligned} \Delta \dot{U}'_A &= (Z'_L - Z'_m) \dot{I}'_a + (Z'_L - Z'_m) \dot{I}'_a, \\ \Delta \dot{U}'_B &= (Z'_L - Z'_m) \dot{I}'_b + (Z'_L - Z'_m) \dot{I}'_b, \\ \Delta \dot{U}'_G &= (Z'_L - Z'_m) \dot{I}'_g + (Z'_L - Z'_m) \dot{I}'_g \end{aligned} \right\} \quad (21)$$

и для нагрузки „1“

$$\left. \begin{aligned} \Delta \dot{U}'_A &= (Z'_L - Z'_m) \dot{I}'_a + \frac{1}{3} [(2Z'_L - Z'_m) \dot{I}'_a + \\ &\quad + (-Z'_L + 2Z'_m) \dot{I}'_b] + \Delta \dot{U}'_{\phi 0}, \\ \Delta \dot{U}'_B &= (Z'_L - Z'_m) \dot{I}'_b + \frac{1}{3} [(-Z'_L + 2Z'_m) \dot{I}'_a + \\ &\quad + (2Z'_L - Z'_m) \dot{I}'_b] + \Delta \dot{U}'_{\phi 0}, \\ \Delta \dot{U}'_G &= (Z'_L - Z'_m) \dot{I}'_g + \frac{1}{3} (Z'_L + Z'_m) \dot{I}'_g + \Delta \dot{U}'_{\phi 0} \end{aligned} \right\} \quad (22)$$

или, принимая для симметричной нагрузки в конце линии симметричную систему токов \dot{I}'_a , \dot{I}'_b и \dot{I}'_g , выражение (22) можно преобразовать в (23):

$$\left. \begin{aligned} \Delta \dot{U}'_A &= (Z'_L - Z'_m) \dot{I}'_a + \frac{1}{3} (Z'_L - 2Z'_m) \dot{I}'_g + \Delta \dot{U}'_{\phi 0}, \\ \Delta \dot{U}'_B &= (Z'_L - Z'_m) \dot{I}'_b + \frac{1}{3} (Z'_L - 2Z'_m) \dot{I}'_g + \Delta \dot{U}'_{\phi 0}, \\ \Delta \dot{U}'_G &= (Z'_L - Z'_m) \dot{I}'_g - \frac{2}{3} (Z'_L - 2Z'_m) \dot{I}'_g + \Delta \dot{U}'_{\phi 0}. \end{aligned} \right\} \quad (23)$$

Из выражения (21) следует, что в случае введения дополнительных сопротивлений двухпроводная трехфазная линия для нагрузки в конце линии оказывается как бы транспонированной; симметричную составляющую нулевой последовательности $\Delta \dot{U}'_{\phi 0}$ в этом случае практически можно считать равной нулю. Потери напряжений в фазах характеризуются двумя системами токов, зависящих от нагрузок «1» и «2». Для нагрузки «1» (22) линия ДПК не транспонирована; таким образом, в данном случае введение дополнительных сопротивлений полностью не компенсирует несимметрию токов и напряжений.

Принимая генератор, включенный на двухпроводную линию, симметричным с фазными напряжениями $\dot{U}_A = \dot{U}$, $\dot{U}_B = a^2 \dot{U}$, $\dot{U}_C = a \dot{U}$, и учтя (21) и (23), легко получить выражения для линейных и фазных напряжений у потребителей¹.

Приведем выражения для фазных напряжений, принимая нагрузки симметричными и соединенными по схеме «звезда».

$$\left. \begin{aligned} \dot{U}'_A &= \dot{U} - (Z'_L - Z'_m) i'_a - (Z_L - Z_m) i'_g, \\ \dot{U}'_B &= a^2 \dot{U} - (Z'_L - Z'_m) i'_b - (Z_L - Z_m) i'_g, \\ \dot{U}'_C &= a \dot{U} - (Z'_L - Z'_m) i'_g - (Z_L - Z_m) i'_g \end{aligned} \right\} \quad (24)$$

и

$$\left. \begin{aligned} \dot{U}'_A &= \dot{U} - (Z'_L - Z'_m) i'_a - \frac{1}{3} (Z'_L - 2Z'_m) i'_g, \\ \dot{U}'_B &= a^2 \dot{U} - (Z'_L - Z'_m) i'_b - \frac{1}{3} (Z'_L - 2Z'_m) i'_g, \\ \dot{U}'_C &= a \dot{U} - (Z'_L - Z'_m) i'_g + \frac{2}{3} (Z'_L - 2Z'_m) i'_g. \end{aligned} \right\} \quad (25)$$

НЕСИММЕТРИЯ ДВУХПРОВОДНОЙ ТРЕХФАЗНОЙ СИСТЕМЫ

Для оценки асимметрии напряжений и токов двухпроводной системы с использованием обшивки самолета в качестве фазного провода выразим линейные напряжения (10) и (13) через их симметричные составляющие

$$\left. \begin{aligned} \dot{U}'_{\lambda 1} &= \frac{1}{3} (\dot{U}'_{BA} + a \dot{U}'_{CB} + a^2 \dot{U}'_{AC}) = (a^2 - 1) \dot{U} - \\ &- \frac{1}{3} (a^2 - 1) [(Z'_L + aZ'_m - a^2 Z'_g) i'_a + (aZ'_L + Z'_m - a^2 Z'_g) i'_b + \\ &+ (Z_L + aZ_m - a^2 Z_g) i'_a + (aZ_L + Z_m - a^2 Z_g) i'_b], \\ \dot{U}'_{\lambda 2} &= \frac{1}{3} (\dot{U}'_{BA} + a^2 \dot{U}'_{CB} + a \dot{U}'_{AC}) = \\ &= -\frac{1}{3} (a^2 - 1) [(aZ'_L + Z'_m - a^2 Z'_g) i'_a + \\ &+ (Z'_L + aZ'_m - a^2 Z'_g) i'_b + (aZ_L + Z_m - a^2 Z_g) i'_a + \\ &+ (Z_L + aZ_m - a^2 Z_g) i'_b], \\ \dot{U}'_{\lambda 0} &= \frac{1}{3} (\dot{U}'_{BA} + \dot{U}'_{CB} + \dot{U}'_{AC}) = 0; \end{aligned} \right\} \quad (26)$$

¹ Эти же выражения можно получить и другим путем, например, положив в выражениях (10), (11), (13) и (14)

$$Z'_g = Z'_L - 2Z'_m \text{ и } Z_g = Z_L - 2Z_m.$$

$$\begin{aligned}
 \dot{U}'_{a1} &= \frac{1}{3} (\dot{U}'_{BA} + a\dot{U}'_{CB} + a^2\dot{U}'_{AC}) = (a^2 - 1) \dot{U} - \\
 &- \frac{1}{3} (a^2 - 1) [(Z'_L + aZ'_m - a^2Z'_g) i'_a + (aZ'_L + Z'_m - a^2Z'_g) i'_b + \\
 &\quad + (Z'_L + aZ'_m) i'_a + (aZ'_L + Z'_m) i'_b], \\
 \dot{U}'_{a2} &= \frac{1}{3} (\dot{U}'_{BA} + a^2\dot{U}'_{CB} + a\dot{U}'_{AC}) = \\
 &= -\frac{1}{3} (a^2 - 1) [(aZ'_L + Z'_m - a^2Z'_g) i'_a + \\
 &\quad + (Z'_L + aZ'_m - a^2Z'_g) i'_b + (aZ'_L + Z'_m) i'_a + (Z'_L + aZ'_m) i'_b], \\
 \dot{U}'_{a0} &= \frac{1}{3} (\dot{U}'_{BA} + \dot{U}'_{CB} + \dot{U}'_{AC}) = 0.
 \end{aligned} \tag{27}$$

При выводе (26) и (27) использовано соотношение $\frac{1-a}{a^2-1} = a$.

Если сопротивления нагрузок для токов прямой и обратной последовательностей в трехфазной системе равны между собой ($Z'_{H1} = Z'_{H2}$ и $Z''_{H1} = Z''_{H2}$), то асимметрия токов и напряжений для нагрузки в конце линии будет

$$\epsilon'_U = \epsilon'_I = \left| \frac{\dot{U}'_{a2}}{\dot{U}'_{a1}} \right| = \left| \frac{(aZ'_L + Z'_m - a^2Z'_g) i'_a + (Z'_L + aZ'_m - a^2Z'_g) i'_b}{3\dot{U} - (Z'_L + aZ'_m - a^2Z'_g) i'_a - (aZ'_L + Z'_m - a^2Z'_g) i'_b} \right. \\
 \left. \frac{+ (aZ'_L + Z'_m - a^2Z'_g) i'_a + (Z'_L + aZ'_m - a^2Z'_g) i'_b}{-(Z'_L + aZ'_m - a^2Z'_g) i'_a - (aZ'_L + Z'_m - a^2Z'_g) i'_b} \right| \tag{28}$$

и для нагрузки „1“

$$\epsilon'_U = \epsilon'_I = \left| \frac{\dot{U}'_{a2}}{\dot{U}'_{a1}} \right| = \left| \frac{(aZ'_L + Z'_m) i'_a + (Z'_L + aZ'_m) i'_b + a^2Z'_g i'_g}{3\dot{U} - (Z'_L + aZ'_m) i'_a - (aZ'_L + Z'_m) i'_b - a^2Z'_g i'_g} \right|. \tag{29}$$

Для случая нагрузок с неравными сопротивлениями для токов прямой и обратной последовательностей ($Z'_{H1} \neq Z'_{H2}$, $Z''_{H1} \neq Z''_{H2}$) при определении асимметрии токов нужно пользоваться соотношением

$$\epsilon_I = \left| \frac{Z_{H1}}{Z_{H2}} \right| \epsilon_U, \tag{30}$$

где ϵ_I — асимметрия токов для соответствующей нагрузки;
 ϵ_U — асимметрия напряжений из выражений (28) или (29).

Абсолютное значение сопротивления обратной последовательности Z_{H2} при наличии электродвигателей в качестве нагрузок составляет примерно 0,25–0,5 от сопротивления прямой последовательности Z_{H1} . При статической нагрузке $Z_{H1} = Z_{H2}$.

Следовательно, для худшего случая асимметрия токов ϵ_I может быть больше асимметрии напряжений ϵ_U в 2—4 раза.

Положив в выражении (28) $Z'_g = Z'_L - 2Z'_m$ и $Z''_g = Z_L - 2Z_m$ и приняв с некоторым допущением (как это будет оговорено далее), что токи нагрузок образуют симметричные системы, найдем для этого случая, что

$$\epsilon'_U = \epsilon''_I \approx 0. \quad (31)$$

По аналогии выражение (29) принимает вид

$$\epsilon'_U = \epsilon'_I \approx \left| \frac{(Z'_L - 2Z'_m) I'_a}{3\dot{U} - 3(Z'_L - Z'_m) I'_a - (2Z'_L - Z'_m) I''_a} \right|. \quad (32)$$

Допущение, принятое при выводе выражения (31), необходимо, так как, несмотря на симметричность нагрузок, система токов не может быть симметричной, что следует из выражений (22), (23) и (32). Таким образом, введение дополнительных сопротивлений только частично компенсирует несимметрию двухпроводной линии. Практически можно считать для симметричной нагрузки, включенной в конце линии, при наличии симметричной другой нагрузки и включении дополнительных сопротивлений асимметрию токов и напряжений равными нулю (31).

Рассмотрим случай, когда двухпроводная линия выполнена без дополнительных сопротивлений, т. е. $Z'_g = Z''_g = 0$. Положив в выражениях (28) и (29) $Z'_g = Z''_g = 0$, найдем выражения асимметрии напряжений для двух нагрузок:

$$\begin{aligned} \epsilon''_U &= \left| \frac{(aZ'_L + Z'_m) I_a + (Z'_L + aZ'_m) I_b + (aZ''_L + Z''_m) I''_a + (Z''_L + aZ''_m) I''_b}{3\dot{U} - (Z'_L + aZ'_m) I_a - (aZ'_L + Z'_m) I_b - (Z''_L + aZ''_m) I''_a - (aZ''_L + Z''_m) I''_b} \right| \approx \\ &\approx \left| \frac{(Z'_L - 2Z'_m) I_a + (Z''_L - 2Z''_m) I''_a}{3\dot{U} - (2Z'_L - Z'_m) I_a - (2Z''_L - Z''_m) I''_a} \right|; \end{aligned} \quad (33)$$

$$\begin{aligned} \epsilon'_U &= \left| \frac{(aZ'_L + Z'_m) I_a + (Z'_L + aZ'_m) I_b}{3\dot{U} - (Z'_L + aZ'_m) I_a - (aZ'_L + Z'_m) I_b} \right| \approx \\ &\approx \left| \frac{(Z'_L - 2Z'_m) I_a}{3\dot{U} - (2Z'_L - Z'_m) I_a} \right|. \end{aligned} \quad (34)$$

Подобно предыдущему, можно получить выражения для несимметрии линии без дополнительных сопротивлений и для других нагрузочных условий, например для случая нагрузок с $Z'_{H1} \neq Z'_{H2}$ и $Z''_{H1} \neq Z''_{H2}$. Удобнее, оказывается, получить их, выражая симметричные составляющие соответствующих напряжений

через симметричные составляющие токов. Не приводя всех выводов, дадим окончательные формулы для асимметрии токов:

$$\varepsilon_I' = \left| \frac{Z_L' - 2Z_m'}{2Z_L' - Z_m' + 3Z_{H2}' + 3kZ_{ген2}'} \right|; \quad (35)$$

$$\varepsilon_I'' = \left| \frac{k_1(Z_L' - 2Z_m') + (Z_L - 2Z_m)}{k_2(2Z_L' - Z_m' + 3Z_{ген2}') + (2Z_L - Z_m + 3Z_{ген2}'' + 3Z_{H2}'')} \right|, \quad (36)$$

где $Z_{ген2}$ — сопротивление генератора токам обратной последовательности;

Z_{H2}' , Z_{H2}'' — сопротивления нагрузок токам обратной последовательности;

$$k = 1 - \frac{I_2''}{I_2'}; \quad k_1 = \frac{I_1'}{I_1''}; \quad k_2 = \frac{I_2'}{I_2''}; \quad I_2 = I_2' + I_2'';$$

I_1' , I_2' — токи прямой и обратной последовательностей нагрузки „1“;

I_1'' , I_2'' — то же для нагрузки „2“.

Отметим, что полученные выражения легко преобразуются в известные уже формулы (см. «Электрический расчет сетей с использованием земли в качестве одного из фазных проводов» Л. Е. Эбина и Ш. М. Алукера, а также «О двухпроводной трехфазной линии» В. В. Андреева) для частного случая одной нагрузки при значениях коэффициентов $k=1$, $k_1=k_2=0$.

Для определения асимметрии напряжений нужно воспользоваться выражением (30)

$$\varepsilon_V = \left| \frac{Z_{H2}}{Z_{H1}} \right| \varepsilon_I,$$

где Z_{H1} — сопротивление нагрузки токам прямой последовательности;

Z_{H2} — сопротивление нагрузки токам обратной последовательности.

Обобщая расчет для случая двух нагрузок, следует заметить, что не представляет особого труда рассчитать схему с большим числом нагрузок. Не производя множества расчетов, можно по аналогии с приведенными выше записать соответствующие выражения для более сложных случаев. Особенно хорошо это видно на примере выражений асимметрии напряжений (33) — (36).

Формулы для определения сопротивлений Z_L' , Z_L'' и взаимных индукций Z_m' , Z_m'' двух параллельных петель «провод—земля» приводятся в литературе. Для случая, когда в качестве фазного

провода используется обшивка самолета, соответствующие сопротивления можно определить по формулам¹:

$$Z_L = R + j29f10^{-4} \lg \frac{2h}{r_s} = R_{gL} + jX_{gL} \text{ [ом/км]}; \quad (37)$$

$$Z_m = j29f10^{-4} \lg \frac{D_{ab}}{d_{ab}} + R_{gm} + jX_{gm} \text{ [ом/км]}, \quad (38)$$

где R — активное сопротивление провода;
 f — частота переменного тока;
 h — высота подвеса проводов;
 d_{ab} — расстояние между проводами;
 D_{ab} — расстояние от одного провода до зеркального изображения другого провода относительно поверхности обшивки самолета;
 r_s — эквивалентный радиус провода. Для витых медных проводов величина определена в зависимости от радиуса провода r :
 для проводов, имеющих 7 прядей, $r_s = 0,724 r$,
 для проводов, имеющих 19 прядей, $r_s = 0,759 r$,
 для проводов, имеющих 37 прядей, $r_s = 0,767 r$,
 для проводов, имеющих 61 прядь, $r_s = 0,771 r$,
 для сплошных круглых проводов $r_s = 0,779 r$;

$R_{gL}, X_{gL}, R_{gm}, X_{gm}$ — сопротивления, учитывающие сопротивление обшивки и вихревые токи в ней. Эти сопротивления зависят от высоты подвеса проводов над обшивкой, расстояния между проводами, частоты переменного тока, толщины и проводимости обшивки и при частоте 400 гц могут быть определены из графиков, приведенных в работе В. В. Андреева «О двухпроводной трехфазной линии», в функции отношений

$$\frac{2h}{\rho_0} \text{ и } \frac{D_{ab}}{\rho_0},$$

где

$$\rho_0 = \frac{10^9}{2\pi\omega\Delta\gamma},$$

ω — угловая частота переменного тока;

Δ — толщина обшивки в см;

γ — проводимость обшивки в $\frac{1}{\text{ом см}}$.

При частотах, отличных от 400 гц, значения сопротивлений $R_{gL}, X_{gL}, R_{gm}, X_{gm}$, найденные по графикам для соответствующего ρ_0 , надо изменить пропорционально частоте. При очень

¹ В. В. Андреев, Исследование сопротивлений прямой и нулевой последовательностей самолетной трехфазной сети, канд. диссертация, МАИ, 1950.

большой проводимости обшивки $\frac{2h}{\rho_0}$ стремится к бесконечности, сопротивления $R_{gL}, X_{gL}, R_{gm}, X_{gm}$ стремятся к нулю. При малом расстоянии между проводами $D_{ab} \approx 2h$, а сопротивления R_{gL} и R_{gm} и соответственно X_{gL} и X_{gm} приблизительно равны.

ЗАКЛЮЧЕНИЕ

Внедрение в авиации трехфазной системы переменного тока открывает большие возможности перед использованием предлагаемой двухпроводной трехфазной системы, где в качестве одного из фазных проводов используется металлическая обшивка самолета. Такая двухпроводная система дает значительный выигрыш в весе, так как электрооборудование занимает большое место в общем оборудовании современного самолета.

Высокая проводимость обшивки самолета уменьшает потери электроэнергии в линии, повышая общий КПД электропередачи.

Несимметрия такой двухпроводной системы в общем не превышает допустимых 5%.

Введение дополнительных сопротивлений, как показывают расчеты, хотя и уменьшает несимметрию системы, но полностью компенсировать ее не может. Поэтому применение дополнительных сопротивлений может оказаться нецелесообразным.

Кроме того, при заземлении одной фазы линии создаются условия, когда изоляция фидеров и аппаратуры работает под линейным напряжением вместо фазного.

Увеличение диэлектрических потерь в изоляции и старение изоляции незначительны (в пределах норм). С другой стороны, к междуфазовой изоляции и изоляции между фазой и корпусом требования ОСТ одинаковы; следовательно, при переходе к двухпроводной системе дополнительные требования на изоляцию не налагаются.

Следует добавить, что применение двухпроводной системы ни в коем случае не затрудняет защиту сети. Новая система по сравнению с обычной трехфазной имеет даже некоторые преимущества, например, в том, что специальная защита от замыкания на массу в данном случае является излишней.

Таким образом, можно считать, что защита двухпроводной трехфазной системы в некоторых случаях проще обычной трехпроводной и, следовательно, выгоднее как в весовом, так и в экономическом отношении. Оборудование, предназначенное для работы в системах с изолированной нейтралью, может быть использовано в двухпроводных системах.

Проведение дальнейших исследований по использованию двухпроводной трехфазной системы переменного тока в авиации установит целесообразность применения подобной схемы и ответит на множество вопросов, связанных непосредственно с эксплуатацией такой системы.

ЛИТЕРАТУРА

1. Вагнер К. Ф., Эванс Р. Д., Метод симметричных составляющих, ОНТИ, 1936.
2. Марголин Н. Ф., Токи в земле, Госэнергоиздат, М.—Л., 1947.
3. Орешкинский П. С., Передача энергии трехфазным током по системе «Два провода—земля», изд. АН СССР, 1935.
4. Эбин Л. Е., Алукер Ш. М., Электрический расчет сетей с использованием земли в качестве одного из фазных проводов, Госэнергоиздат, М.—Л., 1949.
5. Эбин Л. Е., Алукер Ш. М., Руководство по электрическому расчету трехфазных сетей с использованием земли в качестве одного из проводов, Госэнергоиздат, 1940.
6. Симичев И. С., Система ДПЗ на вечной мерзлоте, Журнал «Электрические станции», 1947, № 2.
7. Олендорф Ф., Токи в земле, ГНТИ, 1932.
8. Андреев В. В., Исследование сопротивлений прямой и нулевой последовательностей самолетной трехфазной сети, канд. диссертация, МАИ, 1950.
9. Андреев В. В., О двухпроводной трехфазной линии, Труды МАИ № 57, 1956.
10. Зевеке Г. В., Ионкин П. А., Основы электротехники, Госэнергоиздат, 1955.

Замеченные опечатки

Стр.	Строка	Напечатано	Должно быть	По чьей вине
11	Фиг. 8	$-K_1 \sin(\omega t + \sigma)$	$-KK_1 \sin(\omega t + \sigma)$	авт.
17	Фиг. 12	$K \sin(\omega t + \sigma)$	$U_m \sin(\omega t)$	авт.
17	Фиг. 12	$\sin \omega t$	$U_{my} \sin(\omega t + \sigma)$	авт.
18	6 снизу	$-KM$	$-KK_1$	авт.
38	18 и 20 сверху	$U_{ст}$	$V_{ст}$	авт.
39	8 сверху	$\Delta U_{рег} \Delta I_{\sim}$	$\Delta U_{рег} \Delta I_{\sim}$	авт.
40	1 сверху	фиг. 1, б,	фиг. 1, в,	авт.
53	Формула (13)	$v_{нач}$	$v_{нач}^2$	авт.
91	Формула (1)	$\frac{Z_{2н}^* + Z_{2г}}{Z_{2н}^* + Z_{2г}}$	$\frac{Z_{2н}^* \cdot Z_{2г}}{Z_{2н}^* + Z_{2г}}$	авт.
105	8 сверху	i'_b	i_b	авт.
106	11 снизу	Z'_g	Z_g	авт.
108	4 сверху	\dot{U}_{BA}	\dot{U}'_{BA}	авт.
112	1 снизу	$Z'_g = Z_L - 2Z_m$	$Z_g = Z_L - 2Z_m$	авт.

СОДЕРЖАНИЕ

	Стр.
Предисловие	3
<i>В. П. Щелкин.</i> Дроссель насыщения с фазовым управлением	5
<i>Н. Т. Кобякова, В. А. Сартори.</i> Стабилизатор напряжения переменного тока повышенной точности с дросселем насыщения	23
<i>А. К. Тер-Акопов.</i> Оптимальная индуктивность и минимальное время срабатывания электромагнитов постоянного тока	41
<i>А. К. Тер-Акопов.</i> Анализ методов расчета динамических характеристик электромагнитных механизмов	48
<i>С. Н. Иконников.</i> О методическом переходе от обыкновенного моста к двойному мосту (Томсона)	72
<i>В. Ф. Рахманов.</i> Круговая диаграмма для определения коэффициентов усиления по току и напряжению при комплексной нагрузке	75
<i>В. Н. Истратов, В. В. Митюшина.</i> К расчету токов короткого замыкания самолетной электрической системы трехфазного тока	84
<i>А. В. Каменский, В. Г. Тер-Захарян.</i> Расчет несимметрии токов при обрывах фазы	89
<i>А. Б. Тимофеев, В. Г. Тер-Захарян.</i> Трансформаторы тока для питания реле дифференциальных защит самолетной электросети	99
<i>А. А. Ковзан.</i> Об использовании обшивки самолета в качестве одного из фазных проводов	102

3,

ЭЛЕМЕНТЫ
ЭЛЕКТРОАВТОМАТИКИ
И РАСЧЕТА САМОЛЕТНЫХ
ЭЛЕКТРИЧЕСКИХ СИСТЕМ

Редактор *Л. А. Кац*

Техн. редактор *В. П. Рожин*

T-01477

Подп. в печать 4/IV 1957 г.

Учетно-изд. л. 6,79

Формат бумаги $60 \times 92 \frac{1}{16} = 3,88$ бум. л.—7,65 печ. л. в т. ч. 1 вкл.

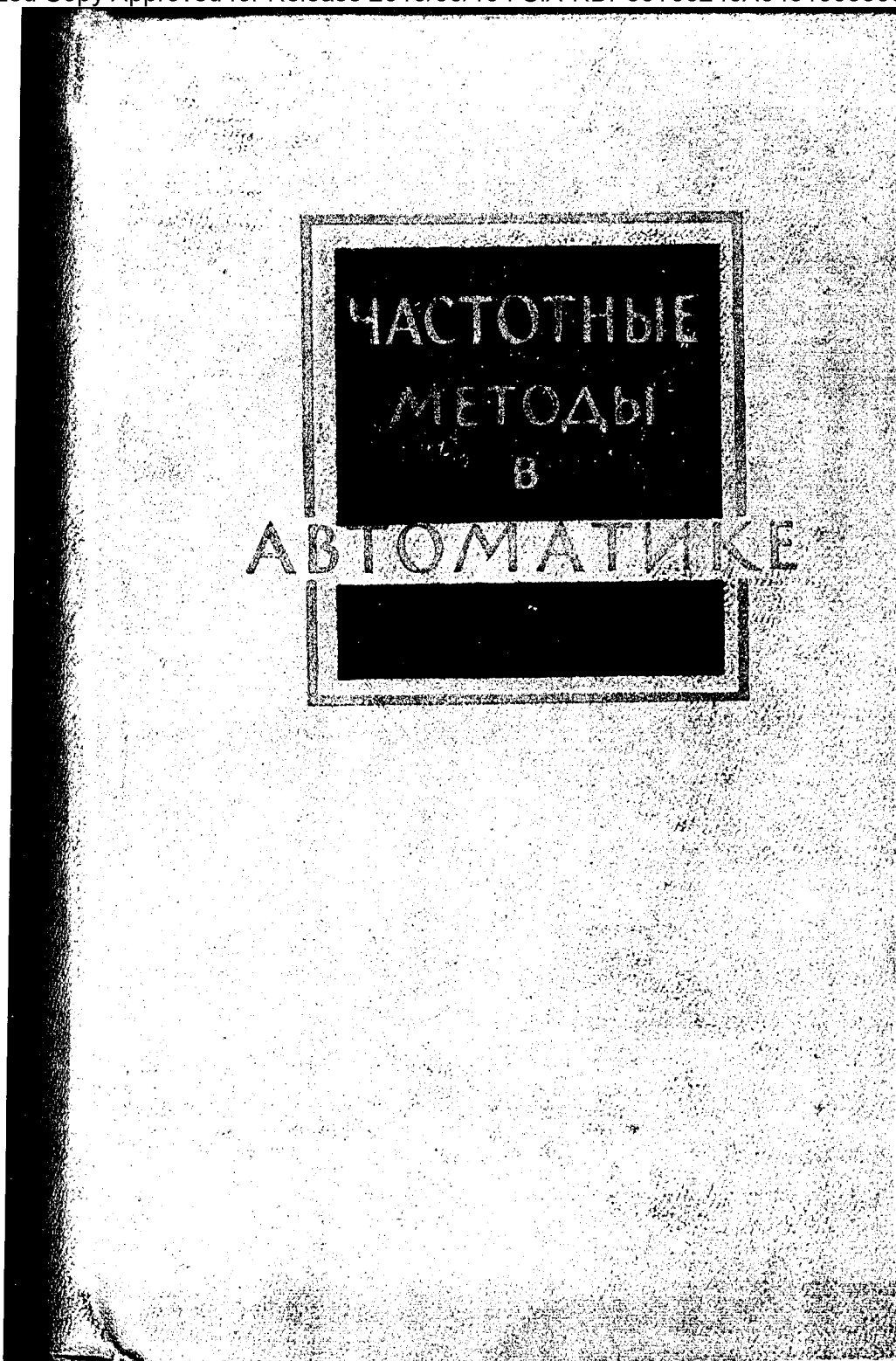
Цена 6 р. 75 к.

Тираж 4800.

Заказ 1848/8410.

Типография Оборонгиза

Цена 6 р. 75 к.



Sanitized Copy Approved for Release 2010/06/16 : CIA-RDP80T00246A045100330001-4

И * Л

*Издательство
иностранной
литературы*

*

Sanitized Copy Approved for Release 2010/06/16 : CIA-RDP80T00246A045100330001-4

ЧАСТОТНЫЕ МЕТОДЫ В АВТОМАТИКЕ

СБОРНИК СТАТЕЙ

Под редакцией

канд. техн. наук
В. А. КАРАБАНОВА

Вступительная статья

д-ра техн. наук
В. В. СОЛОДОВНИКОВА

ИЗДАТЕЛЬСТВО
ИНОСТРАННОЙ ЛИТЕРАТУРЫ

Москва, 1957

АННОТАЦИЯ

В сборник включены статьи из иностранной периодической литературы, посвященные применению частотных методов для решения некоторых задач автоматики. В нем рассмотрены способы приближенной оценки свойств переходных процессов в линейных системах автоматического регулирования по их частотным характеристикам, изложены результаты применения частотных методов к исследованию существенно нелинейных систем, а также импульсных систем, в состав которых входят цифровые вычислительные машины. Даны примеры применения частотных методов к анализу регулирования турбин в энергосистемах и движения самолета с автопилотом. Приведены также некоторые сведения об аппаратуре и методах экспериментальной оценки частотных характеристик.

Сборник предназначен для научных и инженерно-технических работников, специализирующихся в области автоматического регулирования. Он может служить также дополнительным пособием для студентов старших курсов и аспирантов ряда специальностей.

Редакция литературы по вопросам техники
Заведующий редакцией проф. А. Е. ШЕЙНДЛИН

ПРЕДИСЛОВИЕ

Настоящий сборник посвящен вопросам применения частотных методов к решению некоторых задач автоматики. Частотные методы, принятые „на вооружение“ специалистами в области автоматики примерно 20 лет назад, в настоящее время получили чрезвычайно широкое распространение и всеобщее признание. Это обусловливается в первую очередь богатством физического содержания частотных методов, их наглядностью, широкими возможностями замены аналитических расчетов графическими построениями и, наконец, тем, что применение частотных методов позволяет разработать сравнительно простые инженерные методы анализа и синтеза линейных систем автоматического регулирования.

Немалую роль в создании частотных методов сыграли работы советских ученых. Внедрение частотных методов в область автоматики было связано с разработкой частотного критерия устойчивости линейных динамических систем с отрицательной обратной связью. Эта задача была решена в работах Г. Найквиста (1932 г.), предложившего частотный критерий для анализа устойчивости усилителей с отрицательной обратной связью, и А. В. Михайлова (1938 г.), который первым применил этот критерий в теории автоматического регулирования и дал ему строгое математическое обоснование. С этого времени начинается быстрое развитие частотных методов как в нашей стране, так и за рубежом. Особенно много работ, в которых развиваются идеи частотного метода, было выполнено В. В. Солодовниковым, который примерно в это же время (1939 г.) начал разработку частотных методов анализа качества линейных систем автоматического регулирования, нашедшую завершение в его докторской диссертации (1948 г.). В. В. Солодовников разработал также практические приемы синтеза корректи-

рующих устройств систем автоматического регулирования, основанные на использовании частотного метода (1951—1953 гг.).

Для развития частотных методов в настоящее время весьма показательно то обстоятельство, что эти методы стали внедряться в такие области теории автоматического регулирования, как теория нелинейных систем и теория систем импульсного регулирования, в то время как вначале частотные методы применялись исключительно для исследования свойств линейных систем непрерывного действия.

Применительно к теории нелинейных (нелинеаризуемых) систем автоматического регулирования весьма плодотворным оказалось введение понятия эквивалентного комплексного передаточного коэффициента нелинейных элементов. Определение этого понятия основано на принципе гармонической линеаризации, впервые сформулированном и использованном для решения практических задач Н. М. Крыловым и Н. Н. Боголюбовым (1934—1937 гг.). Метод гармонической линеаризации, хотя и имеющий некоторые ограничения, позволяет эффективно исследовать поведение многих существенно нелинейных систем, причем эти исследования могут быть выполнены с использованием детально разработанных и хорошо известных частотных методов. Разработка такого рода методов исследования нелинейных систем проводилась целым рядом авторов как в нашей стране (Л. С. Гольдфарб, Е. П. Попов и др.), так и за рубежом.

Наряду с этим частотные методы с успехом были применены к исследованию часто встречающихся на практике систем импульсного регулирования. У нас разработкой методов исследования импульсных систем занимался Я. З. Цыпкин, который на основе дискретного преобразования Лапласа разработал методику, во многом совпадающую с частотным методом исследования систем непрерывного действия. В последнее время в этом направлении появилось довольно много работ и у иностранных авторов.

Таким образом, возможности частотных методов далеко не исчерпаны и они продолжают непрерывно развиваться и совершенствоваться.

Краткое изложение сущности частотного метода исследования систем автоматического регулирования дано во вступительной статье, написанной В. В. Солодовниковым.

В настоящем сборнике собраны переводы статей из американских журналов, в которых рассматриваются вопросы применения частотных методов к анализу нелинейных и импульсных систем автоматического регулирования, а также некоторые общие вопросы теории и результаты исследования конкретных систем. В большинстве своем представленные в сборнике статьи посвящены детальной разработке какого-либо частного вопроса или рассмотрению определенных практических задач. Поэтому они весьма показательны с точки зрения иллюстрации возможностей частотного метода, хотя и не содержат каких-либо новых, неизвестных советскому читателю принципиальных положений.

Материал сборника сгруппирован по тематическому признаку, и в соответствии с этим сборник разбит на четыре раздела.

В первом разделе рассматриваются некоторые общие вопросы применения частотных методов к анализу и синтезу систем автоматического регулирования.

В статье Ольденбургера содержится попытка сформулировать некоторые стандарты для представления частотных данных и установить расчетные критерии. Цель статьи — дать некоторые путеводные указания по использованию частотных методов для инженерно-технических работников, не имеющих глубоких познаний в специальных отделах математики, без которых применение частотных методов становится в известной мере формальным.

В статье Керби и Бомарьяжа читатель, знакомый с частотным методом анализа качества, разработанным В. В. Солодовниковым, пожалуй, не найдет для себя ничего нового, но определенный интерес в этой статье представляет подход к преобразованию Лапласа как к операции гармонического анализа. Авторы описывают целый ряд примеров и приводят графики, с помощью которых можно произвести оценку качества системы („относительной устойчивости“, по терминологии авторов).

В статье Чу рассматривается возможность оценки переходного процесса системы с помощью анализа распределения нулей и полюсов передаточной функции замкнутой системы, основанного на использовании метода корневого годографа Эванса. Статья представляет интерес с той точки зрения, что подход, предлагаемый автором, перебрасывает

мост между обычным частотным методом анализа переходных процессов систем и методом, основанным на анализе распределения нулей и полюсов передаточных функций.

Наконец, в статье Бьернсона излагается метод приближенной оценки свойств переходных процессов систем автоматического регулирования по частотным характеристикам разомкнутых систем. Читатель найдет в этой статье ряд полезных практических указаний о влиянии параметров логарифмических частотных характеристик разомкнутых систем на особенности и свойства переходных процессов.

Во второй раздел вошли статьи, в которых иллюстрируется применение частотного метода к анализу существенно нелинейных систем. Во всех этих статьях полученные результаты определены с помощью понятия эквивалентного коэффициента усиления нелинейного элемента (метод гармонической линеаризации).

В статье Принца указан прием построения частотных характеристик замкнутых систем автоматического регулирования, содержащих нелинейный элемент, который представляется эквивалентным передаточным коэффициентом.

В статье Левинсона делается попытка объяснить явление „скачкообразного“ резонанса, характерное для некоторых следящих систем с тахометрической стабилизацией и с нелинейным элементом типа насыщения. Для своих выводов автор использует частотный метод и понятие эквивалентного передаточного коэффициента.

Интересно отметить, что это, пожалуй, первое исследование „скачкообразного“ резонанса, основанное на использовании частотного метода. До этого „скачкообразный“ резонанс исследовался только с применением сравнительно сложных методов теории нелинейных колебаний.

В статье Честната на основе частотного метода проводится детальный анализ свойств систем автоматического регулирования, содержащих элементы с насыщением и зоной нечувствительности как в прямой цепи, так и в цепи обратной связи. Проектировщик следящих систем найдет в статье много полезных практических сведений.

В третьем разделе объединены статьи, в которых рассматриваются приложения частотного метода к анализу некоторых специальных систем автоматического регулирования, в частности импульсных систем. В статьях Лиц-

вилла и Ситтлера и Линвилла и Зальцера излагается одно из возможных применений частотных методов для анализа импульсных систем, в том числе систем, содержащих в качестве элемента цифровые вычислительные машины. В статье Зальцера рассматриваются частотные характеристики цифровой вычислительной машины, являющейся элементом системы автоматического регулирования и работающей в естественном времени.

Следует заметить, что первая статья очень конспективна и требует от читателя знания не только особенностей работы импульсных систем, но и элементов теории таких систем или, во всяком случае, хорошего знания свойств и основных теорем преобразования Лапласа.

В статье Ойя рассказывается о применении частотного метода для исследования системы регулирования турбин в энергосистеме Швеции. Статья представляет интерес не только потому, что частотные методы для анализа работы энергосистем стали применяться сравнительно недавно, но и в связи с тем, что в этой статье анализ работы энергосистемы с использованием частотного метода проводится с учетом экономических факторов работы системы.

В статье Сикорда рассматривается возможность применения частотного метода для анализа устойчивости движения такого сложного аэродинамического объекта, каким является самолет ¹⁾.

В четвертом разделе объединены статьи, в которых рассмотрены некоторые вопросы экспериментального определения частотных характеристик систем автоматического регулирования.

Статья Смита и Триpletта примыкает к статье Сикорда. В ней приводятся материалы об экспериментальных методах определения частотных характеристик самолета непосредственно в полете. Содержание этой статьи лишнее раз подчеркивает одно из преимуществ частотного метода, заключающееся в том, что исходные данные для анализа сложных объектов (частотные характеристики) могут быть определены экспериментально.

¹⁾ См. также Спирман Д., Построение и применение аэродинамических передаточных функций управляемых снарядов, *Вопросы ракетной техники*, 2 (38), 1 (1957).

В статье Мэтьюза рассмотрена задача расчета частотных характеристик линейной и эквивалентных характеристик нелинейной части системы по экспериментальным данным. Рассмотренная задача представляет определенный практический интерес. Очень часто при экспериментальном определении частотных характеристик результаты измерения искажаются присутствующими в системе нелинейными элементами, а по измеренным, т. е. искаженным, данным требуется определить истинные частотные характеристики линейной части системы. Решение такого рода задачи и рассмотрено в этой статье.

Наконец, в статье Сент-Клэра и др. дана классификация и рассмотрены характеристики некоторых типов генераторов синусоидальных колебаний, применяющихся при экспериментальном исследовании систем автоматического регулирования. Здесь же приводятся соображения о возможных способах построения таких генераторов.

В целом сборник представляет интерес для научных и инженерно-технических работников, специализирующихся в области автоматического регулирования, а также может послужить полезным пособием для студентов старших курсов и аспирантов ряда специальностей.

В. А. Карбанов.

ВСТУПИТЕЛЬНАЯ СТАТЬЯ

Частотный метод получил широкое практическое применение при расчете систем автоматического регулирования. Как известно, этот метод в теорию регулирования впервые ввел А. В. Михайлов, который в 1936—1938 гг. разработал, используя результаты Найквиста, основы частотного метода анализа устойчивости систем регулирования, изложенные наряду с другими его результатами в известной работе [1].

В основе частотного метода лежит математический аппарат преобразований Лапласа и Фурье. В общем случае линейной системы регулирования преобразование Лапласа $X(s)$ для регулируемой величины может быть представлено в виде

$$X(s) = \frac{H(s)}{1 + KW(s)}, \quad (1)$$

где через $KW(s)$ обозначена так называемая передаточная функция системы в разомкнутом состоянии, а через $H(s)$ — функция комплексного переменного s , зависящая в общем случае от характеристик воздействий, от начальных условий и передаточной функции объекта регулирования (см., например, [2]).

Частотный метод анализа и синтеза систем автоматического регулирования основан, во-первых, на изучении свойств зависимости, существующей при определенных условиях между показателями качества и вещественными функциями, получающимися из выражения (1) при замене в нем s на $j\omega$ (эти функции называются обобщенными частотными характеристиками), и, во-вторых, на использовании связи между обобщенными частотными характеристиками и частотными характеристиками разомкнутой системы, определяющими ее устойчивость,

Существенные особенности частотного метода заключаются в следующем:

а) основными исходными данными для него могут служить частотные характеристики, которые можно определить экспериментальным путем, не зная дифференциальных уравнений всей системы в целом или отдельных ее элементов;

б) он учитывает своеобразие систем автоматического регулирования, заключающееся в том, что их анализ в разомкнутом состоянии обычно проще, чем в замкнутом;

в) он позволяет производить полный анализ динамики и, в частности, анализ устойчивости, качества и переходных процессов, а также решать многие вопросы синтеза в системах любого порядка, как одноконтурных, так и многоконтурных, содержащих не только сосредоточенные, но и распределенные параметры;

г) он удобно обобщается для решения вопросов динамической точности или анализа и синтеза систем при непрерывно изменяющихся воздействиях и, таким образом, дает возможность провести полное исследование систем автоматического регулирования от начала до конца;

е) он в современном его состоянии, если иметь в виду применение логарифмических характеристик, основан на использовании специальных номограмм, таблиц и построений, устраняющих необходимость громоздких вычислений.

В настоящее время частотный метод анализа устойчивости и частотный метод качества представляют собою единый метод. Нужно, однако, заметить, что вначале две стороны вопроса — устойчивость и качество — рассматривались несколько в отрыве друг от друга. В работах А. В. Михайлова использован математический аппарат в виде теоремы Коши и принципа аргумента, ставящий частотный критерий устойчивости на прочное основание, причем для применения последнего используются частотные характеристики разомкнутой системы. В первых же работах по частотному методу анализа качества систем автоматического регулирования, наоборот, за основу были приняты частотные характеристики замкнутой системы и использован математический аппарат, позволяющий обосновать этот метод. Частотный критерий устойчивости, хотя и рекомендуется в этих работах, но исследование устойчивости рассматривается как самостоятельный этап и ничего не говорится о наличии связи

между обоими этапами — анализом устойчивости и последующим анализом качества.

Следующий шаг в развитии частотного метода заключался в преодолении указанного разрыва и в отчетливом выяснении и использовании того обстоятельства, что амплитудно-фазовая характеристика разомкнутой системы содержит исходные данные для анализа не только устойчивости, но и качества.

Таким образом было показано, что предварительный анализ качества можно производить не только по вещественной частотной характеристике замкнутой системы $P(\omega)$, но и по амплитудно-фазовой характеристике $KW(j\omega)$ разомкнутой системы и что процесс вычисления $P(\omega)$ может быть значительно упрощен при помощи специальных номограмм, если уже построены обычная или обратная амплитудно-фазовая характеристика или логарифмические характеристики, необходимые для анализа устойчивости.

В связи с рассматриваемым вопросом следует подчеркнуть еще одно очень существенное обстоятельство. Установление связи между свойствами частотных характеристик систем в замкнутом и в разомкнутом состояниях и свойствами частотных характеристик и качеством имеет важное значение не только для анализа систем автоматического регулирования, но и для синтеза и выбора параметров этих систем.

Трудность выбора параметров заключается в том, что в случае более или менее сложных систем нелегко установить непосредственную связь между показателями качества и параметрами системы. Но если нам удалось установить связь между свойствами вещественной частотной характеристики и качеством, то тем самым сделан шаг вперед в решении задачи выбора параметров, так как зависимость между параметрами системы и выражением для частотной характеристики всегда может быть представлена в явном виде.

Задача выбора параметров еще более упрощается, если установлена связь между свойствами переходного процесса и свойствами, скажем, амплитудной характеристики разомкнутой системы, поскольку зависимость последней от параметров часто имеет очень простой вид.

Как уже указывалось выше, одним из основных вопросов любого метода решения проблемы качества является

установление связи между выбранной косвенной характеристикой качества и показателями качества. В частотном методе анализа качества косвенной характеристикой качества служат обобщенные частотные характеристики, учитывающие как собственные динамические свойства системы, так и характеристики воздействия и начальных условий. Часто вместо обобщенных частотных характеристик достаточно рассмотреть одну из собственных частотных характеристик системы. Таковой в частотном методе обычно служит вещественная частотная характеристика $P(\omega)$.

Вещественная частотная характеристика связана с переходной функцией $x(t)$ при единичном ступенчатом воздействии известным интегральным соотношением

$$x(t) = \frac{2}{\pi} \int_0^{\infty} P(\omega) \frac{\sin t\omega}{\omega} d\omega. \quad (2)$$

Этот интеграл давно известен математикам, и некоторые его простейшие свойства рассматриваются в теории рядов Фурье. Однако постановка задачи об анализе качества, которую интеграл был призван решать в теории регулирования, существенно отличалась от предыдущих его применений¹⁾. В качестве основного здесь был поставлен вопрос не о том, как, зная $P(\omega)$, вычислить интеграл (2), а как, не вычисляя этого интеграла, по свойствам функции $P(\omega)$ судить о свойствах функции $x(t)$, характеризуемой показателями качества.

Анализ свойств интеграла (2) дал возможность установить такую связь; он позволил производить предварительную оценку качества по виду частотных характеристик в случае решения задачи анализа и выбирать форму частотных характеристик так, чтобы обеспечить удовлетворение требований, предъявляемых к системе, в случае решения задачи синтеза.

К числу этих свойств относятся: соотношения между начальными и конечными значениями переходной функции и частотных характеристик, достаточные условия для малых перерегулирований и для монотонности, необходимые условия

¹⁾ Своеобразие постановки проблемы качества, составляющее основу как частотного, так и всех других методов решения этой проблемы, уже отмечалось.

для монотонности, связь между параметрами частотных характеристик и показателями качества для типовых вещественных частотных характеристик (например, трапециoidalных), оценка нижнего предела для времени переходного процесса по интервалу положительности, оценка погрешности в определении переходного процесса в зависимости от погрешности в определении частотных характеристик и т. д.

Несмотря на то, что найденных соотношений не так много, они все же дают возможность составить ориентировочное представление о показателях качества и характере протекания переходного процесса непосредственно по виду частотных характеристик, без всяких дополнительных вычислений и построений, кроме тех, которые необходимы для построения самих характеристик.

Широкое применение, выходящее за рамки теории регулирования, получил частотный метод построения переходных процессов, обычно называемый методом трапециoidalных частотных характеристик.

Наиболее удобной и современной формой частотного метода является метод логарифмических характеристик; первоначальную же форму частотного метода, связанную с построением амплитудно-фазовых характеристик, на наш взгляд, можно считать уже несколько устаревшей.

Основные преимущества метода логарифмических характеристик заключаются в следующем:

- а) он устраняет необходимость почти всех расчетов;
- б) он дает единообразную методику анализа устойчивости многоконтурных систем;
- в) логарифмические характеристики являются весьма наглядным средством исследования, позволяющим судить по их виду о том, какие параметры системы наиболее существенны для устойчивости и для качества, для коэффициентов ошибок и ошибки при синусоидальном входном сигнале;
- г) логарифмические характеристики очень удобны для определения того, какие из параметров системы можно считать малыми не только с точки зрения устойчивости, но и качества;
- д) в случае одноконтурных систем асимптотическая логарифмическая амплитудная характеристика дает полное представление не только об устойчивости и качестве, но и о тех типовых звеньях, которые входят в состав системы, т. е. о ее структуре;

е) логарифмические характеристики очень удобны для синтеза корректирующих устройств;

ж) наконец, метод логарифмических характеристик; по крайней мере в случае минимально-фазовых систем, позволяет в принципе свести анализ к исследованию свойств вещественной функции $L(\omega)$ вещественного переменного ω , устраняя необходимость построения годографов.

Метод логарифмических частотных характеристик уже широко используется на практике. Здесь мы лишь вкратце остановимся на способе применения этого видоизменения частотного метода к синтезу корректирующих устройств.

Задача синтеза корректирующих устройств может быть сформулирована следующим образом. Задана структурная схема системы автоматического регулирования, состоящая из элементов с известными передаточными функциями и из элементов с неизвестными передаточными функциями. Требуется найти неизвестные передаточные функции корректирующих устройств и способ их реализации в виде схем так, чтобы обеспечить удовлетворение заданных условий качества при простейшем возможном виде искомым передаточных функций. Передаточная функция считается тем более простой, чем ниже порядок ее числителя и знаменателя.

Решая задачу синтеза корректирующих устройств, обычно следуют такому порядку операций:

а) определяются, если это возможно, оптимальные, т. е. совместимые с принятыми ограничениями условия протекания переходного процесса;

б) находятся логарифмические характеристики, соответствующие этим оптимальным условиям;

в) строятся желаемые логарифмические характеристики, представляющие собой разумный компромисс между стремлением приблизиться к условиям оптимальности и стремлением получить технически возможно более простое решение задачи, учитывающее свойства объекта и не нарушающее допустимых отклонений от условий оптимальности;

г) по найденным желаемым логарифмическим характеристикам и логарифмическим характеристикам неизменяемой части находятся логарифмические характеристики корректирующих устройств;

д) решается задача аппроксимации, т. е. находятся передаточные функции корректирующих устройств по найденным для них частотным характеристикам;

е) решается задача реализации полученных передаточных функций при помощи тех или иных технических средств (особенно хорошо разработана методика решения задачи реализации при помощи пассивных электрических контуров).

Подводя итоги приведенной выше краткой характеристике частотного метода, можно утверждать, что этот метод, основанный на наглядных физических представлениях и позволяющий в качестве исходных данных пользоваться не только дифференциальными уравнениями, но и экспериментально полученными частотными характеристиками, является в настоящее время основным средством анализа и синтеза сложных одноконтурных и многоконтурных систем автоматического регулирования, описываемых линейными уравнениями любого порядка и содержащих не только сосредоточенные, но и распределенные параметры.

Частотный метод дает единую общую методику решения всех основных задач теории автоматического регулирования, а именно: получения исходных данных, анализа устойчивости, анализа качества, построения переходных процессов, выбора параметров и решения вопроса о малости параметров, синтеза корректирующих устройств, анализа динамической точности и синтеза оптимальных систем.

Частотный метод удобен не только в теории непрерывного, но и в теории импульсного, прерывистого регулирования.

Необходимо отметить, что частотный метод за последнее время стал применяться для анализа не только линейных систем, но и систем с нелинейными и переменными параметрами. Так, например, сочетание частотных представлений с методом гармонического баланса привело к разработке практически удобного приближенного метода анализа автоколебаний [3], недавно разработаны основные положения частотного метода точного анализа релейных систем [4], показана возможность применения частотного метода к анализу систем с переменными параметрами на основе теоремы о свертке [5].

Поэтому утверждения, которые порой приходится слышать, что частотный метод полезен лишь для решения линейных задач, конечно, не соответствуют истине.

Первые работы, посвященные разработке частотного метода исследования динамики и расчета систем автоматического регулирования, принадлежат советским ученым. Им же принадлежат многие существенные результаты, полученные в этой области. Однако этот метод получил за последние годы существенное развитие за рубежом. Настоящий сборник дает об этом достаточно хорошее представление.

В. В. Солодовников.

ЛИТЕРАТУРА

1. Михайлов А. В., Метод гармонического анализа в теории регулирования, *Автоматика и телемеханика*, № 3 (1938).
2. Солодовников В. В., Метод частотных характеристик в теории регулирования, *Автоматика и телемеханика*, 8, № 2 (1947).
3. Гольдфарб Л. С., О некоторых нелинейностях в системах регулирования, *Автоматика и телемеханика*, 8, № 5 (1947)
4. Цыпкин Я. З., Частотный метод исследования периодических режимов релейных систем автоматического регулирования, сборник „Памяти А. А. Андропова“, М., 1955.
5. Солодовников В. В., Об одном применении операторного исчисления к динамическим системам с переменными параметрами, *Изв. АН СССР*, отд. техн. наук, № 12 (1945).

I. ОБЩАЯ ТЕОРИЯ

1. ИЗОБРАЖЕНИЕ ЧАСТОТНЫХ ХАРАКТЕРИСТИК, СТАНДАРТЫ И РАСЧЕТНЫЕ КРИТЕРИИ ¹⁾

Р. Ольденбургер

Введение

Задачей Комитета динамических систем, созданного при Американском обществе инженеров-механиков (ASME) в ноябре 1951 г., является разработка стандартов для представления частотных данных. Необходимость в таких стандартах стала очевидной ввиду широкого применения частотных методов для расчета систем автоматического регулирования. Из-за различного представления одних и тех же сведений, касающихся частотных характеристик, часто возникают недоразумения. Одни строят амплитудные характеристики в виде зависимости отношения амплитуд от частоты, другие — в виде зависимости логарифма отношения амплитуд от логарифма частоты. Между тем, одного взгляда на частотные характеристики элемента физической системы порой бывает достаточно, чтобы сделать важные заключения о динамических свойствах элемента, при условии, что используется привычная система координат. Поэтому для легкого чтения характеристик весьма желательно, чтобы они строились в стандартных координатах. Эти и другие соображения привели к созданию Комитета динамических систем, который после изучения всех имеющихся предложений вынес свои рекомендации, публикуемые в настоящей статье.

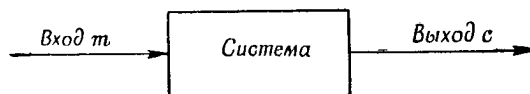
Чтобы облегчить понимание рекомендаций Комитета и способствовать применению частотных методов для расчета систем регулирования, в настоящей статье рассматриваются некоторые основные вопросы и даются правила расчета, позволяющие инженеру без специальной математической подготовки приступить к научному расчету систем автоматического регулирования. Большинство предлагаемых расчетных

¹⁾ Oldenburger R., *Trans. ASME*, 76, № 8, 1155—1169 (1954).

критериев в разное время уже было опубликовано в литературе по теории автоматического регулирования. Основная цель данной статьи состоит в том, чтобы дать специалистам некоторое представление о научных методах расчета систем автоматического регулирования.

Рекомендации относительно амплитудной и фазовой характеристик

Рассмотрим физическую систему, схематически изображенную на фиг. 1, с сигналом на входе t и выходным



Фиг. 1. Блок-схема физической системы с входом t и выходом c .

сигналом c . Входной сигнал t является переменной величиной, которая управляет величиной c на выходе. Примером может служить двигатель внутреннего сгорания, в котором дроссельная заслонка управляет числом оборотов двигателя. За переменную t можно принять положение заслонки ($см$), а за переменную c — скорость вращения вала двигателя (об/мин). Обычно вход и выход измеряются как отклонения от установившихся значений.

Если входной сигнал изменяется со временем синусоидально, т. е. $t = A \sin \omega t$ (A — амплитуда, ω — частота), то выход будет изменяться также по синусоидальному закону, если система линейна. Опыт и теория показывают, что всегда желательно при изучении системы считать ее линейной, если не доказано обратное.

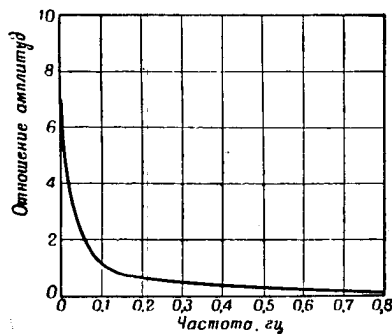
Величина $\omega = 2\pi f$, где f — частота колебания. Для установившихся колебаний выход линейной системы будет иметь вид $B \sin(\omega t + \varphi)$ (B — амплитуда, φ — фазовый угол). Для линейной системы при данной частоте f величина B пропорциональна амплитуде входа A . Частное B/A является отношением амплитуды выхода к амплитуде входа.

На фиг. 2 представлена амплитудная частотная характеристика двигателя внутреннего сгорания для случая нагрузки

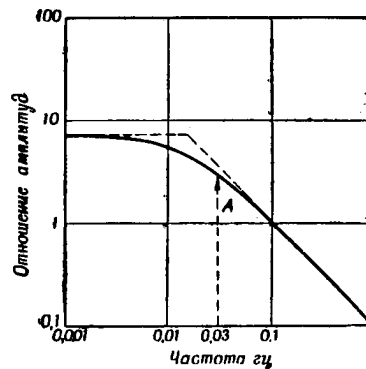
1. Изображение частотных характеристик и расчетные критерии 23

с малым демпфированием. Величина s выражена в единицах 100 об/мин, а вход m — в сантиметрах.

Читатель, знакомый с частотным методом, распознает в кривой на фиг. 2 характеристику одноемкостной системы. Воспроизводя данные, представленные на фиг. 2, в логарифмическом масштабе, мы получим кривую, изображенную на фиг. 3¹⁾. Заметим, что эту кривую можно приблизительно заменить двумя отрезками прямых — асимптот, как это показано пунктиром на фиг. 3. Ломаная линия, получаемая при помощи таких прямых, называется асимптотической



Фиг. 2. Амплитудная частотная характеристика двигателя Дизеля.



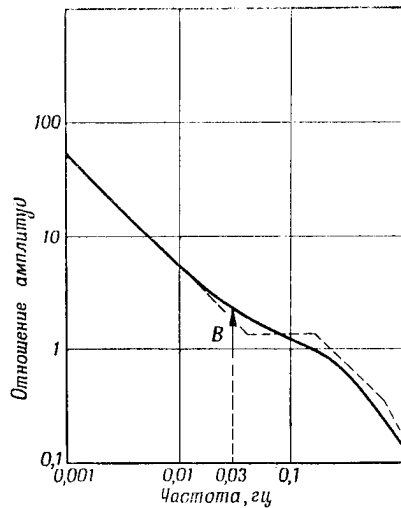
Фиг. 3. Логарифмическая амплитудная частотная характеристика двигателя Дизеля.

амплитудной характеристикой двигателя. Поскольку логарифмы действительных чисел при некотором основании пропорциональны логарифмам этих чисел при другом основании, можно при построении в логарифмических координатах выбирать произвольное основание логарифмов.

Предположим, что измеряется скорость двигателя и что разность между измеряемой и установленной величиной скорости, которую желательно поддерживать постоянной, приводит в действие регулятор, меняющий положение дроссельной заслонки. Можно рассматривать эту разность как вход регулятора, а положение заслонки — как его выход. Если вход изменяется по синусоидальному закону, то выход также

¹⁾ Такие кривые часто называют кривыми Боде.

будет синусоидальным (в предположении, что система линейна). Характеристика типового регулятора дается на фиг. 4. Ее можно также аппроксимировать асимптотической характеристикой, показанной пунктиром на фиг. 4 и составленной



Ф и г. 4. Логарифмическая амплитудная частотная характеристика регулятора двигателя.

из двух прямолинейных асимптот и двух промежуточных прямолинейных отрезков. Для построения этой асимптотической характеристики практически необходимо иметь ана-



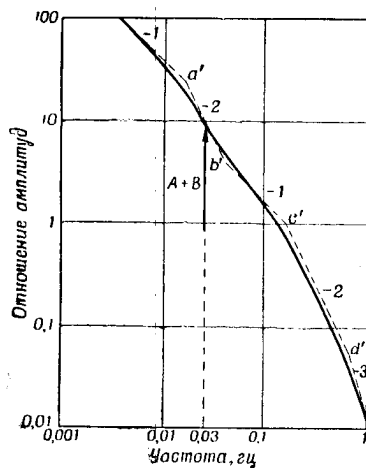
Ф и г. 5. Разомкнутая система регулятор — двигатель.

литическое выражение „передаточной функции“. Если специалист не знаком с этими функциями, ему, вероятно, придется иметь дело с готовыми характеристиками.

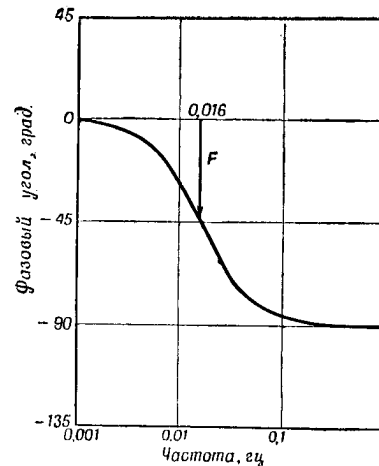
Предположим теперь, что регулятор соединен с двигателем и система разомкнута, как показано на фиг. 5. Здесь

1. Изображение частотных характеристик и расчетные критерии 25

элемент измерения скорости двигателя предполагается отключенным от входа регулятора. Если входной сигнал e (регулирующая переменная) регулятора изменяется синусоидально, то выход c (регулируемая переменная) двигателя, а именно скорость двигателя (отсчитываемая от установившегося значения скорости), практически будет изменяться тоже синусоидально. Можно показать, что амплитудная характеристика системы, состоящей из регулятора и двигателя, получается



Фиг. 6. Логарифмическая амплитудная частотная характеристика разомкнутой системы регулятор — двигатель.



Фиг. 7. Логарифмическая фазовая частотная характеристика двигателя Дизеля.

путем простого сложения ординат амплитудных характеристик, построенных в логарифмических координатах. Таким образом, если мы сложим векторы A и B на фиг. 3 и 4 при частоте $0,03$ гц, то получим вектор $A+B$ на фиг. 6, где представлена амплитудная характеристика разомкнутой системы регулятор — двигатель. Удобно принять горизонтальную линию с индексом 1 за ось абсцисс, так как логарифм единицы есть нуль. Векторы A , B и $A+B$ отсчитываются от этой оси.

Чтобы применять метод сложения ординат, необходимо выход регулятора и вход объекта регулирования (регулируемая система) измерять в одних и тех же единицах.

В нашем случае в одних и тех же единицах должны измеряться вход регулятора e и выход объекта регулирования c .

Если соответствующие ординаты асимптотических характеристик на фиг. 3 и 4 складываются, то на фиг. 6 получается ломаная линия. Это асимптотическая амплитудная характеристика разомкнутой системы регулятор — двигатель. Можно видеть, что результирующая ломаная линия мало отличается от точной кривой.

При увеличении отношений амплитуд системы регулятор — двигатель в K раз говорят об „увеличении усиления“ в K раз. В данной статье мы не будем рассматривать определение термина „усиление“ [5].

Повышение коэффициента усиления ¹⁾ системы регулятор — двигатель соответствует простому сдвигу амплитудной характеристики на фиг. 6 вверх. Коэффициент усиления в физических системах обычно легко регулируется. При применении логарифмического масштаба почти всегда можно приближенно заменить амплитудную характеристику асимптотической ломаной линией. Исходя из этих рассмотрений Комитетом динамических систем дается следующая рекомендация:

Рекомендация 1. Рекомендуется строить амплитудные характеристики в логарифмических координатах, причем по вертикальной оси откладывать отношение амплитуд, а по горизонтальной оси — частоту в герцах.

Это связано с тем, что обычно генераторы синусоидальных колебаний имеют градуировку в герцах и потребовались бы дополнительные расчеты для вычисления угловой частоты ω в радианах в секунду.

Пусть R есть отношение амплитуд. Напомним, что отношение амплитуд в децибелах есть $20 \lg R$. Использование логарифмической шкалы для отношения амплитуд исключает необходимость перевода значений отношения амплитуд в децибелы и, следовательно, экономит время при построении кривых по экспериментальным данным.

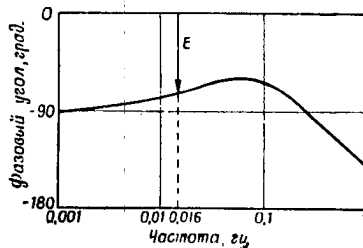
Для двигателя Дизеля фазовая характеристика показана на фиг. 7 с логарифмической шкалой для частоты и линейной шкалой для фазы. Как видно из графика, выход отстает

¹⁾ В нашей литературе более распространен термин „передаточный коэффициент“, — *Прим. ред.*

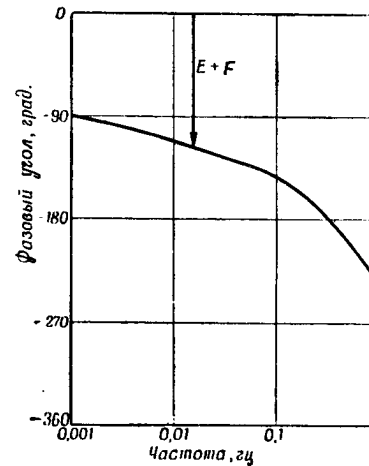
1. Изображение частотных характеристик и расчетные критерии 27

от входа. Регулятор двигателя имеет фазовую характеристику, показанную на фиг. 8. Кривая сначала изгибается вверх, а затем асимптотически приближается к прямой -180° .

Фазовый угол разомкнутой системы регулятор — двигатель при частоте f получается суммированием фазовых углов отдельных составляющих при той же частоте. Суммируя векторы E и F на фиг. 8 и 7, получают вектор $E + F$ на фиг. 9, на которой изображена фазовая характеристика всей системы регулирования скорости двигателя.



Ф и г. 8. Логарифмическая фазовая частотная характеристика регулятора.



Ф и г. 9. Логарифмическая фазовая частотная характеристика разомкнутой системы регулятор — двигатель.

Применение линейной шкалы в градусах для углов фазового сдвига между входом и выходом физической системы общепринято. Поэтому при построении фазовых характеристик рекомендуется следующее:

Рекомендация 2. Рекомендуется фазу в градусах откладывать по вертикальной оси в линейном масштабе, а частоту в герцах — по горизонтальной оси в логарифмическом масштабе.

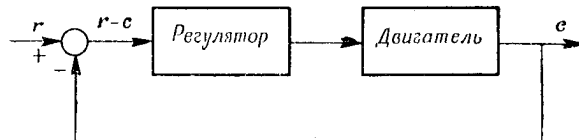
Согласно рекомендациям 1 и 2, для частоты на амплитудной и фазовой характеристиках удобно использовать одни и те же единицы.

Амплитудная и фазовая характеристики полностью характеризуют систему. Их можно использовать для вывода дифференциальных уравнений физической системы, для уста-

новления факторов, вызывающих расхождение теоретических и экспериментальных результатов, и для расчета систем автоматического регулирования.

Расчетные критерии, основанные на использовании амплитудной и фазовой характеристик

По амплитудной и фазовой характеристикам разомкнутой системы определяют характеристики замкнутой системы [1]. Обозначим через r скорость, которую необходимо поддерживать постоянной (заданный входной сигнал), а через



Ф и г. 10. Замкнутая система регулятор — двигатель.

c — действительную скорость двигателя (отсчитываемую от установившегося значения). Схема регулирования скорости двигателя изображена на фиг. 10. Для нормальной работы необходимо, чтобы система была устойчивой. По виду амплитудной и фазовой характеристик разомкнутой системы (см. фиг. 5) можно судить об ее устойчивости в замкнутом состоянии и вывести общие правила расчета такой системы.

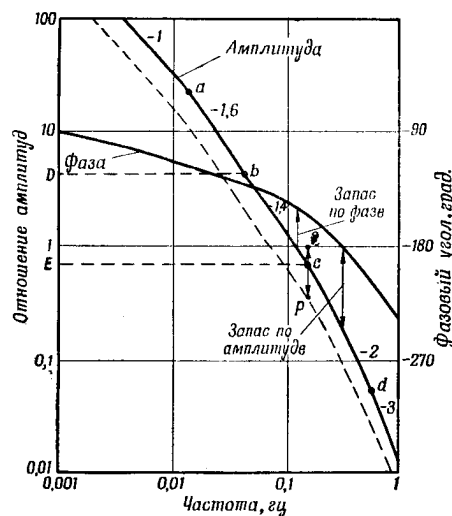
Амплитудные и фазовые кривые удобно строить на одном графике. Для рассматриваемого примера с двигателем такие кривые приведены на фиг. 11.

Первая работа, посвященная частотному методу, была опубликована Найквистом [2], который математически решил задачу об устойчивости линейных замкнутых систем, сведя задачу к изучению характеристик разомкнутой системы. Боде развил этот метод и ввел понятия усиления [3] и запаса устойчивости по фазе. Используя эти понятия, можно дать некоторые простые правила расчета.

Предположим, что к системе S приложено малое возмущение δ , которое, возможно, исчезает со временем. Если переходный процесс, вызванный этим возмущением, затухает со временем, то говорят, что система S устойчива. В противном случае система S неустойчива. Частота, при

1. Изображение частотных характеристик и расчетные критерии 29

которой отношение амплитуды на выходе разомкнутой системы к амплитуде на входе равно единице, называется „частотой среза“. Предположим, что имеется лишь одна частота среза. Назовем этот случай „нормальным“. Если мы из 180° вычтем отставание по фазе при частоте среза, то получим запас устойчивости по фазе; иначе говоря, запас устойчивости по фазе равен сумме 180° и фазового угла.



Фиг. 11. Логарифмические амплитудная и фазовая частотные характеристики разомкнутой системы регулятор — двигатель.

Фазовый угол считают отрицательным при отставании и положительным при опережении фазы. Если имеется несколько частот среза, то указанная сумма определяется для каждой частоты среза и наименьшая из них будет запасом устойчивости по фазе. Для рассматриваемого выше примера запас устойчивости по фазе показан на фиг. 11. Он равен приблизительно 30° . За ось абсцисс удобно принять горизонтальную прямую при значении отношения амплитуд (усиление по терминологии Боде), равном единице.

Точку, в которой отставание по фазе равно 180° (фазовый угол равен -180°), назовем „точкой перехода фазы“.

Предположим сначала, что имеется только одна такая точка. Считаем этот случай нормальным. Пусть R — значение отношения амплитуд в точке перехода фазы. Тогда запас устойчивости по амплитуде будет равен $1/R$, если R меньше единицы, и равен R , если R больше или равно единице. Запас устойчивости по амплитуде часто измеряется в децибелах. Если имеется более одной точки перехода фазы, то запас устойчивости по амплитуде определяется как наименьшее из значений относительной амплитуды $\{R\}$ при $R \geq 1$ или $\{1/R\}$ при $R \leq 1$ в точке перехода фазы.

На фиг. 11 запас устойчивости по амплитуде показан стрелкой, но он не равен длине соответствующего вектора. В нашем случае отношение амплитуд при частоте, соответствующей точке перехода фазы, приблизительно равно 0,2. Тогда запас устойчивости по амплитуде будет 1 : 0,2, т. е. 5. В децибелах это составит $20 \lg 5$, или 14 дб.

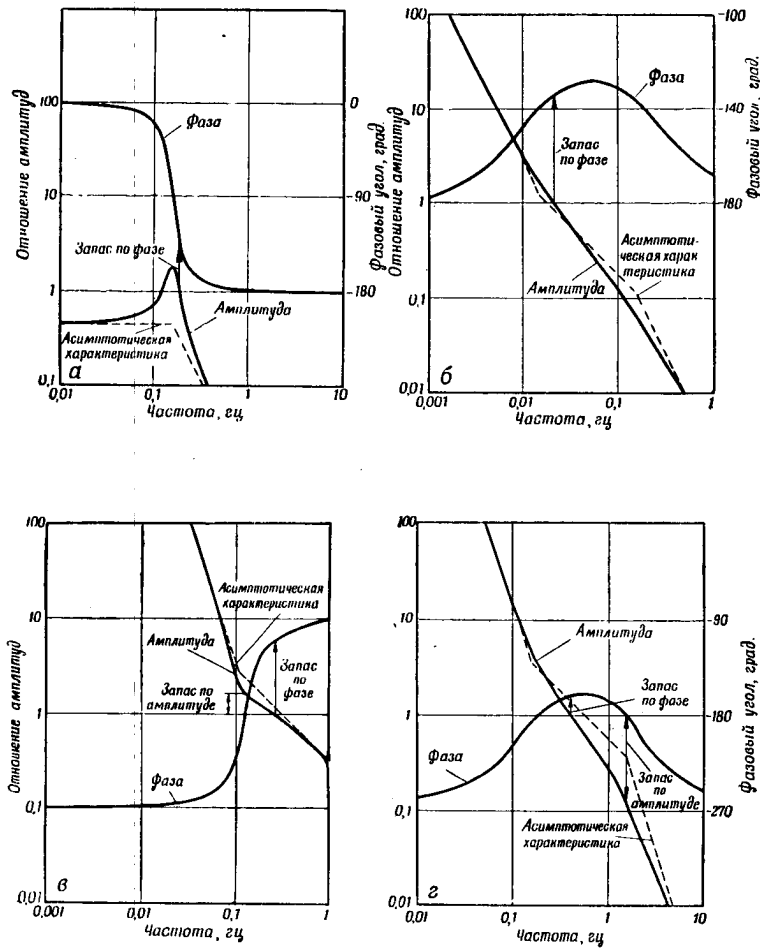
Отставание по фазе для разомкнутой системы может никогда не достигать 180° . В этом случае невозможно определить запас устойчивости по амплитуде. Точно так же отношение амплитуд может не достигать значения 1, и тогда невозможно определить запас по фазе. В случаях, когда запасы устойчивости определить нельзя, нет смысла пользоваться расчетными правилами.

Для дальнейшего изложения удобно распространить определение запаса устойчивости по амплитуде и по фазе на некоторые предельные случаи. Если отставание по фазе никогда не достигает 180° , но приближается к 180° , когда относительная амплитуда стремится к нулю, то можно сказать, что запас устойчивости по амплитуде равен бесконечности, и обозначить его символом ∞ . Аналогично, если отношение амплитуд никогда не достигает единицы, но приближается к 1, когда отставание по фазе стремится к нулю, то говорят, что запас по фазе равен 180° .

Из законов физики [4] и условий ограничения мощности известно, что при беспредельном возрастании частоты отношение амплитуд разомкнутой системы приближается к нулю или к конечному положительному значению.

На фиг. 12 показан запас устойчивости по амплитуде и по фазе для некоторых разомкнутых физических систем.

Обычно используется следующее правило расчета [5—9]:



Ф и г. 12. Запас устойчивости по амплитуде и фазе для некоторых физических систем.

Правило расчета 1. Запас устойчивости по фазе должен составлять по крайней мере 30° , а запас устойчивости по амплитуде — по крайней мере 2,5 (8 дб).

Это правило можно обосновать математически. С помощью теории функций комплексного переменного можно показать, что для разомкнутой системы следует избегать условий, при которых отставание по фазе составляет 180° , а отношение амплитуд равно единице. Согласно правилу 1, этого можно избежать при наличии определенного запаса устойчивости. Ниже мы вернемся к этому вопросу при рассмотрении амплитудно-фазовых характеристик.

Приведенное правило расчета должно рассматриваться как грубый критерий, позволяющий специалисту, не имеющему особой математической подготовки, строить частотные характеристики систем автоматического регулирования и избежать некоторых ошибок при проектировании.

Можно доказать, что вообще опасно допускать меньший запас устойчивости по фазе и по амплитуде. Так, например, при запасе устойчивости по фазе в 5° или при запасе по амплитуде, равном 1,05, система подходила бы так близко к границе неустойчивости, что решение было бы сильно колебательным; если бы при этом все или некоторые частотные характеристики системы были получены путем расчета, а не экспериментально, то даже незначительные факторы могли бы сделать систему полностью неустойчивой.

Может случиться, что, хотя предыдущее расчетное правило удовлетворено, замкнутая система все же будет слишком колебательной. В конце этой статьи будет дан пример, иллюстрирующий такой случай. Однако для большого числа физических систем приведенное правило обеспечивает достаточно хорошее протекание процесса регулирования. Правило определяет запас устойчивости, который часто бывает достаточным и лишь в редких случаях не приводит к цели. Чтобы быть уверенным в том, что система будет удовлетворительно работать, специалист может без особых трудностей и не углубляясь в теорию использовать готовый набор частотных характеристик, определяющих хорошую работу системы, и проследить за тем, чтобы характеристики, полученные им при решении данной конкретной задачи, совпадали или были близки к соответствующим стандартным характеристикам.

1. Изображение частотных характеристик и расчетные критерии 33

Расчет, основанный только на амплитудной частотной характеристике

Первоначально Боде [3] исследовал физические системы, которые могут быть описаны обыкновенными линейными дифференциальными уравнениями с постоянными коэффициентами. Допустим, что задана амплитудная характеристика W . Рассмотрим системы указанного типа с одинаковыми характеристиками W . Среди фазовых характеристик этих систем имеется одна, для которой отставание по фазе имеет наименьшее значение. Система с такой характеристикой называется минимально-фазовой системой. Боде показал, что для минимально-фазовых систем амплитудная характеристика определяет фазовую характеристику¹⁾. Поскольку иногда экспериментальным путем труднее получить фазовую характеристику, чем амплитудную, то желательно иметь расчетный критерий, основанный лишь на одной амплитудной характеристике.

Из теоремы Боде следует, что фазовый угол при данной частоте f для минимально-фазовой системы зависит от наклона амплитудной характеристики (построенной в логарифмических координатах) в окрестности частоты f . Если наклон достаточно постоянен в окрестности частоты f , то фазовый угол пропорционален наклону амплитудной характеристики. Действительно, фазовый угол равен приблизительно $90 m$ градусов для наклона m .

Для иллюстрации этого соотношения обратимся к амплитудной характеристике, показанной сплошной линией на фиг. 11. Часть кривой слева от точки a имеет наклон приблизительно -1 . Напомним, что линия с наклоном -1 опускается на одну единицу при смещении на одну единицу вправо, т. е. линия образует с горизонтальной осью угол

¹⁾ В области регулирования различных процессов запаздывание часто играет важную роль. Обыкновенные линейные дифференциальные уравнения перестают быть справедливыми для не минимально-фазовых систем. В этом случае необходимо иметь как амплитудную, так и фазовую характеристики. Однако хотя для этого случая теория расчета, приведенная в настоящем разделе, неприменима, основные положения, изложенные ниже, все же остаются справедливыми. Ольденбург и Сарториус [10] провели обширное математическое исследование запаздывания систем с сосредоточенными и распределенными параметрами. См. также работу Боде [3], гл. 13.

135°¹⁾. Согласно результатам Боде, фазовый угол должен быть приблизительно равен -90° для низких частот, что видно из графика фиг. 11. Для отрезка cd амплитудной характеристики наклон приблизительно равен -2 (линия опускается на две единицы при смещении на единицу вправо) и соответствующий фазовый угол должен быть около -180° . Справа от точки d наклон приблизительно равен -3 , соответствующий фазовый угол будет примерно равен -270° . Для высоких частот фазовая характеристика асимптотически стремится к линии, соответствующей -270° .

Наклон кривой от a до b приблизительно равен $-1,6$, а от b до c $-1,4$. Поэтому можно ожидать, что фазовый угол для частот, связанных с отрезком ac , будет равен -135° , что совпадает с действительным значением. Очевидно, имеется приближенная зависимость между наклонами -1 , -2 и -3 отрезков асимптотической кривой (см. фиг. 6) и действительными фазовыми углами.

Для частей амплитудной характеристики, где наклон m вдоль характеристики является постоянным целым числом, выражение $-90m$ градусов дает хорошее приближение для фазового угла. Вблизи концов отрезков приближение становится очень грубым; при коротких отрезках ломаной линии приближение также очень грубое.

Если частота среза находится внутри отрезка характеристики с постоянным наклоном -2 , то отставание фазы при частоте среза будет около 180° и запас по фазе будет очень малым.

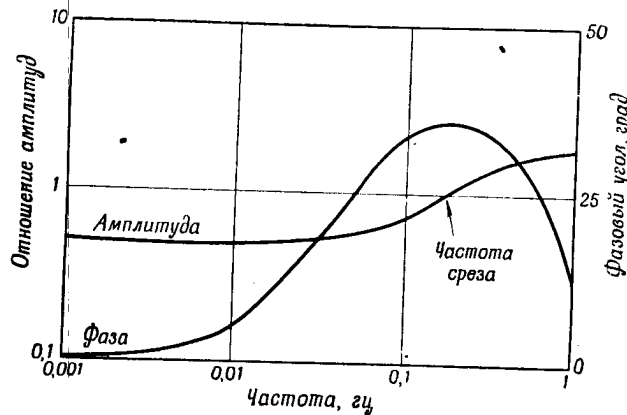
Правило расчета 2. Наклон амплитудной характеристики в логарифмических координатах при частоте среза и вблизи нее должен быть равен приблизительно -1 .

Наклон -1 выбран потому, что значение -1 больше „опасного“ значения -2 ; в то же время часто желательно иметь наклон, близкий к наклону -2 , с тем чтобы получить переходный процесс, который является достаточно колебательным. Пусть f — частота среза. Обычно принимают наклон характеристики близким к -1 в диапазоне частот от $\frac{1}{2}f$ до $2f$. Величина этого диапазона частот,

¹⁾ Этот наклон соответствует наклону -20 дб/декада. — Прим. ред.

1. Изображение частотных характеристик и расчетные критерии 35

вообще говоря, зависит от наклона кривой вдали от частоты среза. Расчетчик путем проверки сможет установить это. Пунктирная кривая на фиг. 11 получена путем смещения сплошной кривой вниз на расстояние, равное $\lg 2$. Поскольку пунктирная амплитудная характеристика на фиг. 11 имеет при частоте среза и вблизи нее наклон, равный приблизительно $-1,4$, (что для практических целей достаточно близко к наклону -1), то расчетное правило 2 удовлетворено.



Фиг. 13. Логарифмические амплитудная и фазовая частотные характеристики устойчивой системы.

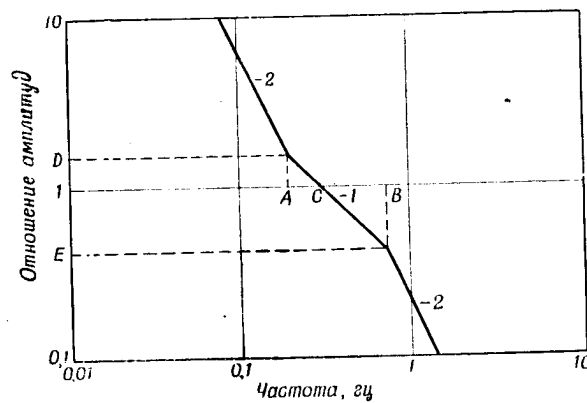
Для случая сплошной амплитудной характеристики это правило удовлетворяется не так хорошо, поскольку наклон этой кривой быстро стремится к -2 после того, как кривая проходит частоту среза. Однако сплошная амплитудная характеристика все еще описывает устойчивую систему.

На фиг. 13 приведены амплитудная и фазовая частотные характеристики для устойчивой разомкнутой системы, где амплитудная характеристика поднимается вверх при подходе к частоте среза. Такая характеристика вообще не типична. Запас устойчивости по фазе превышает 180° , и система имеет большой запас устойчивости, когда контур замыкается.

При вычерчивании частотных характеристик нет необходимости знать точные свойства этих характеристик при всех частотах. Поэтому для физической системы, переходный процесс в которой протекает одну или две минуты, нет

надобности знать характеристику при частоте 100 колебаний в минуту или 0,01 колебания в минуту. Действительно, полоса частот от 10 до 0,1 колебания в минуту будет в данном случае вполне достаточной.

Частотные характеристики должны строиться для области средних частот, и нет никакой необходимости строить их для очень низких частот. Там, где определяются запасы устойчивости по амплитуде и по фазе, эти характеристики должны быть известны для достаточно широкого диапазона



Фиг. 14. Часть асимптотической логарифмической амплитудной частотной характеристики.

частот. Общим случаем является тот, когда асимптотическая амплитудная характеристика в окрестности частоты среза составляется из трех отрезков с наклонами -2 , -1 и -2 соответственно, как показано на фиг. 14. Амплитудная кривая (сплошная или пунктирная) на фиг. 11 является примером, к которому применима фиг. 14. Чтобы обеспечить достаточные запасы устойчивости, автор рекомендует следующее правило¹⁾, основанное на математическом исследовании передаточных функций (см. Приложение):

Правило расчета 3. Рассмотрим разомкнутую систему, часть амплитудной характеристики которой

¹⁾ В работе [9] (гл. IV) рекомендуется отношение AC/AB , равное $1/8$.

1. Изображение частотных характеристик и расчетные критерии 37

аппроксимируется кривой, составленной из прямолинейных отрезков и соответствующей фиг. 14, причем отрезки с наклоном -2 длиннее среднего отрезка с наклоном -1 . Для этой системы желательно иметь длину отрезка AC , равную примерно $1/3 - 1/2$ от длины отрезка AB . Отношение относительной амплитуды, соответствующей точке D , к относительной амплитуде, соответствующей точке E , должно быть равно по крайней мере 5 (отношение частоты, соответствующей точке B , к частоте, соответствующей точке A , должно быть равно по крайней мере 5).

Легко показать, что отношение относительных амплитуд, соответствующих точкам D и E , равно отношению частот, соответствующих точкам B и A .

Согласно соотношению Боде между фазой и амплитудой, увеличение наклона амплитудной характеристики вызывает увеличение фазового угла. Если теперь один или оба отрезка характеристики, примыкающие к среднему отрезку с наклоном -1 , являются короткими отрезками с наклоном -2 , примыкающими к отрезкам с большим наклоном, то отношение частот, соответствующее точкам B и A , может иногда быть меньше 5 , например оно может быть равно 4 .

Для примера системы регулятор—двигатель на фиг. 11 средний отрезок bc сплошной кривой имеет наклон около $-1,4$, в то время как соседние отрезки ab и cd имеют наклоны, соответственно равные $-1,6$ и -2 .

Точки a , b и c , отмеченные на фиг. 11, являются точками сопряжения отрезков с различными наклонами. В строгом смысле термин „точка сопряжения“ (или „сопрягающая точка“) относится к точке, в которой пересекаются прямолинейные отрезки с различными наклонами. Точки сопряжения амплитудной характеристики на фиг. 11 могут быть выбраны не единственным образом ¹⁾. Для асимптотической амплитудной характеристики они являются единственными. На фиг. 6 точками сопряжения будут точки a' , b' , c' и d' . Грубо можно

¹⁾ Автор обратился одновременно к шести специалистам, предложив им выбрать точки сопряжения на этой характеристике, вычерченной в виде сплошной кривой. Все они указали точки, удивительно близкие к четырем точкам сопряжения, указанным на фиг. 11.

представить отрезок ab отрезком с наклоном, приблизительно равным -2 , а отрезок bc — отрезком с приблизительно наклоном -1 . Расстояние по вертикали от c до b равно $\lg 6,5$, откуда частное от деления относительных амплитуд, соответствующих DE на фиг. 11, равно $6,5$. В этом отношении правило расчета 3 удовлетворяется.

Отрезки $a'b'$, $b'c'$ и $c'd'$ асимптотической амплитудной характеристики для примера системы регулятор — двигатель имеют наклоны, соответственно равные -2 , -1 и -2 (см. фиг. 6). Расстояние по вертикали между b' и c' равно $\lg 4$, откуда соответствующее отношение получается равным 4 . Это не удовлетворяет правилу расчета 3. Однако отрезок $a'b'$ не очень велик по сравнению с $b'c'$, а отрезок асимптотической кривой слева от точки a' имеет наклон -1 . Отсюда следует, что частное от деления относительных амплитуд, соответствующих точкам b' и c' , может быть несколько меньше 5 (что требуется правилом расчета 3), и фактическое значение 4 будет вполне удовлетворительным.

Точка, соответствующая частоте среза на сплошной амплитудной характеристике (см. фиг. 11), расположена близко к правому концу среднего отрезка. Для асимптотической характеристики, представленной на фиг. 6, частота среза будет расположена на правом конце среднего отрезка. Таким образом, в этом отношении правило расчета 3 не соблюдается. Перемещая амплитудную характеристику вниз на расстояние $0,3$, как показано пунктирной кривой на фиг. 11 (за основание логарифмов принято число 10), мы сдвигаем тем самым точку среза так, чтобы правило расчета 3 удовлетворялось, а именно к середине или близко к середине среднего отрезка. Это справедливо как в случае использования амплитудной характеристики, так и для ее асимптотического приближения.

Как уже отмечалось, смещение амплитудной характеристики в логарифмических координатах вверх или вниз соответственно изменяет коэффициент усиления разомкнутой системы. Это одно из основных преимуществ построения амплитудной характеристики в логарифмическом масштабе. При увеличении коэффициента усиления вдвое каждое значение отношения амплитуд умножается на два, т. е. к ординатам характеристики прибавляется $\lg 2$, или $0,3$. Уменьшение

1. Изображение частотных характеристик и расчетные критерии 39

усиления вдвое соответствует сдвигу амплитудной характеристики вниз на расстояние 0,3, поскольку $\lg 0,5 = -0,3$. Как уже отмечалось в примере фиг. 11, мы уменьшали вдвое усиление при переходе от сплошной к пунктирной кривой. *Изменение коэффициента усиления не влияет на фазовую характеристику.*

На основании приведенного примера применения правила расчета 3 может показаться, что оно недостаточно точно. Это правило иллюстрировалось произвольно выбранным примером. Тем не менее приведенные здесь грубые методы дают хорошие количественные результаты. Желание иметь достаточный, но не чрезмерный запас по фазе оправдывает применение предыдущего правила. Чтобы помочь получить при расчете системы достаточный запас устойчивости по амплитуде, мы введем еще одно правило, которое основывается на изучении передаточных функций. Предполагается, что мы имеем обычный случай, когда амплитудная характеристика в районе частоты среза имеет отрицательный наклон.

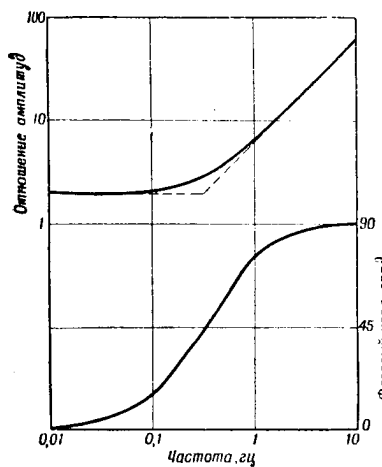
Пусть R — отношение амплитуд в точке P . *Запас устойчивости по амплитуде* в точке P по определению равен $1/R$, если $R < 1$, и R , если $R \geq 1$.

Правило расчета 4. Если после частоты среза вплоть до точки P наклон амплитудной характеристики больше —2 (т. е. кривая идет менее круто, чем при наклоне —2) и далее наклон уменьшается, то запас устойчивости по амплитуде в точке P должен быть равен 2 (6 дБ) или более.

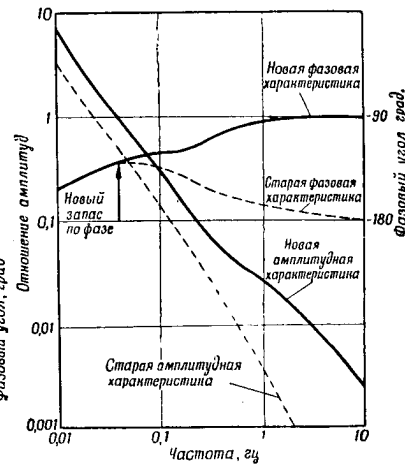
Точка P показана на пунктирной кривой фиг. 11. Запас устойчивости по амплитуде QP в точке P равен 3 (9 дБ). Для случая сплошной кривой этот запас устойчивости недостаточен.

На асимптотической характеристике точка P определяется единственным образом, и тогда правило расчета 4 особенно легко применить.

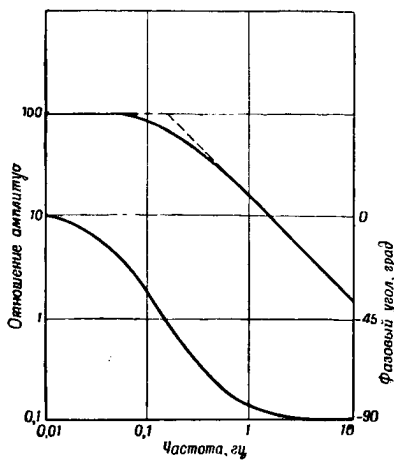
Как уже отмечалось, соотношение Боде между амплитудой и фазой означает, что при увеличении наклона амплитудной характеристики (уменьшении крутизны там, где кривая падающая) увеличивается фазовый угол. Таким образом, если частота среза, с которой связан запас по фазе, остается неизменной, а наклон амплитудной характеристики



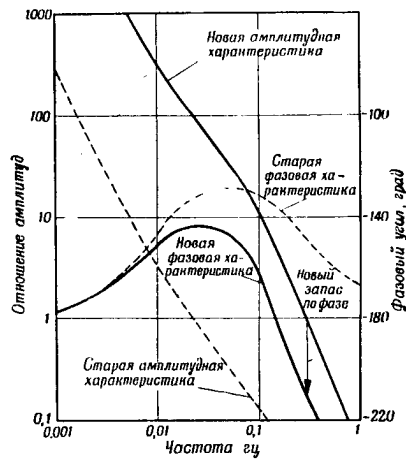
Фиг. 15. Логарифмические частотные характеристики опережающего звена.



Фиг. 16. Логарифмические частотные характеристики системы с учетом опережающего звена.



Фиг. 17. Логарифмические частотные характеристики запаздывающего звена.



Фиг. 18. Логарифмические частотные характеристики системы с учетом запаздывающего звена.

1. Изображение частотных характеристик и расчетные критерии 41

увеличивается, то запас по фазе увеличивается. Из того же соотношения Боде следует, что если наклон амплитудной характеристики никогда не становится меньше -2 (т. е. не становится круче в спадающих частях), то фазовый угол никогда не превышает 180° и запас устойчивости по амплитуде отсутствует. Любая часть амплитудной характеристики, наклон которой меньше -2 (т. е. где кривая идет круче), стремится увеличить отставание по фазе и уменьшить тем самым запас по фазе.

Можно сформулировать множество расчетных правил, подобных правилам, приведенным в этом разделе. Никакая комбинация этих правил не может заменить математического исследования конкретной задачи, но они могут помочь инженеру при грубом анализе задач регулирования.

Для более высоких частот элементы запаздывания имеют тенденцию уменьшать фазовые углы, в то время как элементы опережения имеют тенденцию увеличивать их.

Частотные характеристики для случая простого элемента опережения показаны на фиг. 15. Суммируя ординаты на фиг. 15 с ординатами характеристики системы, изображенной на фиг. 12, б, получим представленные на фиг. 16 характеристики системы с элементом опережения.

Для элемента запаздывания соответствующие характеристики даны на фиг. 17, а суммирование ординат фиг. 17 с ординатами характеристик фиг. 12, б дает в результате кривые, представленные на фиг. 18, для системы с элементом запаздывания. При более высоких частотах новая амплитудная характеристика приближается к старой.

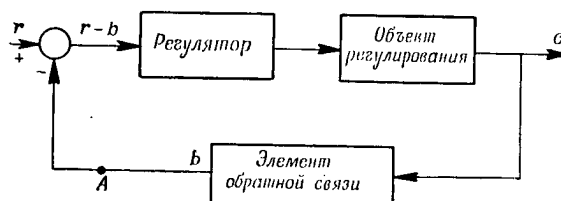
Система, характеристики которой представлены на фиг. 16, является устойчивой, тогда как система, соответствующая фиг. 18, неустойчива.

Системы с элементом в цепи обратной связи

Рассмотрим систему, изображенную на фиг. 19, содержащую элемент в цепи обратной связи, благодаря чему сигнал ошибки системы уже не будет равен $r - c$, как для системы, изображенной на фиг. 10.

Чтобы получить амплитудную характеристику системы, разомкнутой в точке А, мы должны сложить ординаты трех амплитудных характеристик для регулятора, объекта и эле-

мента в цепи обратной связи. Подобным же образом поступаем и с фазовыми характеристиками. Случай, представленный на фиг. 19, является общим случаем, поскольку должна измеряться регулируемая переменная c , а чувствительный элемент содержит, вообще говоря, запаздывающие и опережающие элементы. Регулятор реагирует на разность $r - b$ между входным сигналом r и выходом элемента обратной связи b . Например, при регулировании температуры с чувствительным



Фиг. 19. Система с элементом в цепи обратной связи.

элементом в виде термопары регулятор реагирует на разность между заданной температурой и температурой, измеряемой термопарой. Термопара обладает термическим запаздыванием.

Если характеристикой элемента обратной связи в диапазоне частот, в котором рассматривают устойчивость системы фиг. 19, можно пренебречь, то система, изображенная на фиг. 19, сводится к системе, представленной на фиг. 10. Это может иметь место в случае регулирования температуры при помощи термопары с малой постоянной времени.

Амплитудно-фазовые характеристики (диаграммы Найквиста)

Комитет не предлагает применять исключительно какой-либо один тип частотной характеристики. Тип кривой, которая наиболее отчетливо отражает влияние элементов на общую характеристику, зависит от положения и характера элементов и от критерия, положенного в основу выбора желаемой характеристики. Применение частотных методов для решения некоторых обычных задач хорошо известно. Однако их использование для решения сложных задач регу-

1. Изображение частотных характеристик и расчетные критерии 43

лирования и в особенности задач регулирования процессов совершенно не развито. Промышленные процессы подвергаются многочисленным возмущениям и взаимным влияниям, из которых лишь немногие определены количественно. Способы представления этих результатов зависят от инженера-практика. При публикации своих исследований он свободен в выборе и использовании типа характеристик, которые являются наиболее подходящими для его задачи. Однако, поскольку Комитет рекомендует, чтобы амплитудная и фазовая частотные характеристики были даны в логарифмических координатах, читатель, возможно, сочтет целесообразным использовать логарифмические амплитудно-фазовые графики, которые будут рассмотрены ниже. Поэтому Комитетом дается следующая рекомендация:

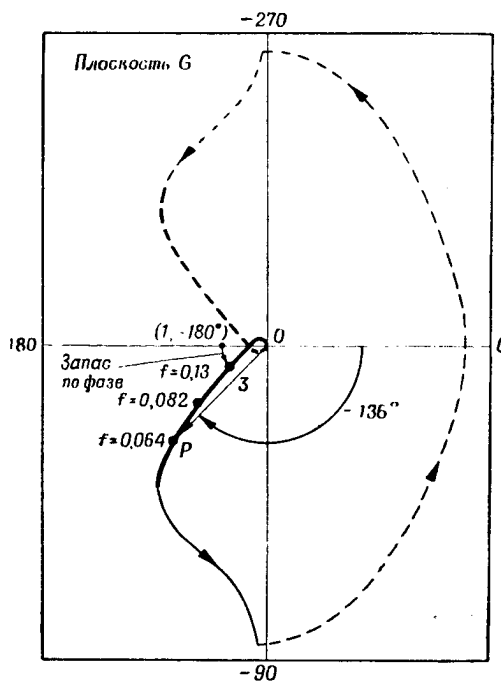
Рекомендация 3. Там, где это удобно, следует отдавать предпочтение логарифмическим амплитудно-фазовым характеристикам перед другими видами представления характеристик.

Амплитудно-фазовая характеристика обычно широко применяется при расчетах систем автоматического регулирования. Это хорошо известная диаграмма Найквиста, которая в одной кривой объединяет сведения, даваемые амплитудной и фазовой частотными характеристиками разомкнутой системы (в литературе и другие графики также называются амплитудно-фазовыми характеристиками, но мы предпочитаем сохранить эту терминологию для кривых Найквиста). Указанная характеристика получается вычерчиванием кривой в полярных координатах; векторный угол для частоты f является фазовым углом для разомкнутой системы, а величина радиуса-вектора является отношением амплитуд для разомкнутой системы при частоте f . Таким образом, частота является параметром кривой.

Характеристики разомкнутой системы регулятор — двигатель (см. фиг. 11), проведенные сплошной линией, объединяются в одну амплитудно-фазовую характеристику, представленную на фиг. 20. Так, например, когда частота $f = 0,064$, отношение амплитуд равно 3, а фазовый угол равен -136° , как это видно из фиг. 20 для точки P .

Части кривых, проведенные тонкой линией на фиг. 20 слева от вертикальной оси, искажены для того, чтобы показать, каким образом кривая асимптотически приближается

к вертикальной оси. Кривая же в окрестности начала координат увеличена в масштабе, чтобы сделать видимой малую петлю.



Фиг. 20. Амплитудно-фазовая характеристика разомкнутой системы регулятор — двигатель.

Для увеличения масштаба кривой в какой-либо части плоскости Комитет рекомендует следующее:

Рекомендация 4. Если можно улучшить удобочитаемость и точность отсчета амплитудно-фазовых характеристик путем частичного изменения масштаба по осям координат, рекомендуется производить такое изменение масштаба.

Математические основы

Хотя для математически неподготовленного читателя подробности настоящего раздела могут показаться неясными,

1. Изображение частотных характеристик и расчетные критерии 45

материал, изложенный в этом разделе, поможет ему составить общее представление о применяемой теории.

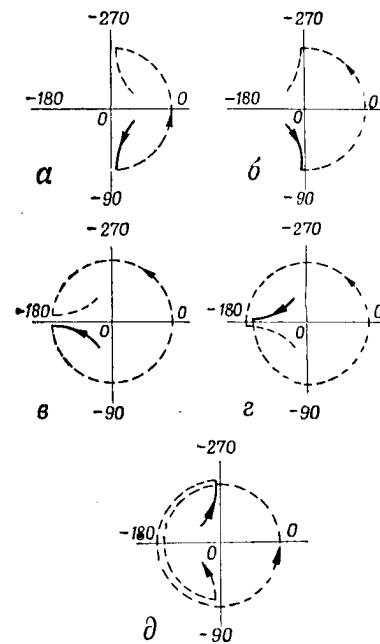
Зависимость выхода линейной системы от входа [11] описывается передаточной функцией $G(s)$, представляющей собой функцию комплексной переменной s , причем $G(s)$, по определению, родственна понятию импеданса в электротехнике. Можно показать, что амплитудно-фазовая характеристика является графиком этой функции при $s = j\omega$, где $j = \sqrt{-1}$ и $\omega = 2\pi f$, при изменении f от 0 до ∞ . Строя кривую, симметричную данной относительно горизонтальной оси, получим сопряженную ветвь, соответствующую $G(j\omega)$ при изменении ω от 0 до $-\infty$.

Для примера системы регулятор — двигатель сопряженная кривая показана пунктиром на фиг. 20 в виде ветви кривой, поднимающейся от начала координат вверх и затем асимптотически стремящейся к вертикальной оси, соответствующей фазовому углу -270° .

Для нормальных передаточных функций (например, для передаточной функции, аналитической всюду, кроме полюсов) теория показывает, что если величина радиуса-вектора увеличивается безгранично, когда ω стремится к нулю, то амплитудно-фазовая характеристика стремится асимптотически к вертикальным или горизонтальным полуосям, исходящим из начала координат. Если амплитудно-фазовая характеристика заканчивается в конечных точках, то она образует вместе с сопряженной ей ветвью замкнутую кривую. Конец амплитудно-фазовой характеристики, соответствующий $\omega = \infty$, является точкой в конечной области [4]. Если другой конец ($\omega = 0$) асимптотически стремится к одной из полуосей, исходящей из начала координат, то кривая, образованная амплитудно-фазовой характеристикой и ее сопряженной ветвью, замыкается проведенной надлежащим образом дугой. В физических системах характеристика часто стремится к полуоси -90° , иногда — к полуоси -180° и довольно редко — к полуоси -270° . Поэтому достаточно будет показать читателю, как замкнуть кривую, образованную амплитудно-фазовой характеристикой и ее сопряженной ветвью, в случаях когда полуоси -90° , -180° или -270° являются асимптотами.

Согласно теории функций комплексного переменного с учетом критерия устойчивости Найквиста, кривые должны

быть замкнуты определенным образом (замкнутая кривая должна обходиться один раз при обходе правой полуплоскости переменной s , скажем, по часовой стрелке), а в качестве дуг для замыкания кривых могут приниматься дуги



Фиг. 21. Способы замыкания амплитудно-фазовых характеристик.

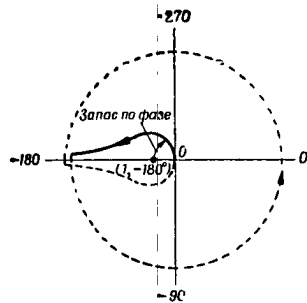
окружностей (они могут быть больше 360°) очень большого („бесконечного“) радиуса, как указано на фиг. 21. На этой фигуре в случаях *a* и *б* асимптотой является полуось -90° , а в случаях *в* и *г* — полуось -180° . Наконец, на фиг. 21, *д* показан способ замыкания в случае, когда асимптотой является полуось -270° (с фазовым углом -270°). На фиг. 21, *д* амплитудно-фазовая характеристика асимптотически стремится к полуоси -270° в третьем квадранте, а затем делает полтора оборота по часовой стрелке, замыкая кривую. Кривую, образованную амплитудно-фазовой характеристикой, ее сопряженной ветвью и замыкающей дугой, будем называть „замкнутой амплитудно-фазовой характеристикой“.

На фиг. 20 замыкание выполнено пунктирной дугой в правой полуплоскости. Замкнутую амплитудно-фазовую характеристику можно обходить в направлении против часовой стрелки. При обходе кривой мы не охватываем точку $(1, -180^\circ)$.

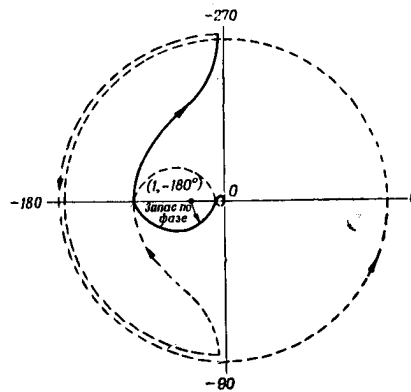
Амплитудно-фазовая характеристика, показанная сплошной кривой на фиг. 22, может быть замкнута, как показано пунктиром на том же рисунке. Обходя кривую в направлении стрелки, мы охватим точку $(1, -180^\circ)$ дважды. Замкнутая

1. Изображение частотных характеристик и расчётные критерии 4

амплитудно-фазовая характеристика для примера на фиг. 12,2 показана на фиг. 23. Полный охват точки $(1, -180^\circ)$ в этом случае равен нулю. Чтобы определить число охватов, проведем вектор из точки $(1, -180^\circ)$ в точку P амплитудно-фазовой характеристики. Пусть P пробегает всю замкнутую кривую. Вектор при этом повернется на угол $360^\circ \cdot n$, где n — число охватов амплитудно-фазовой характеристикой точки $(1, -180^\circ)$.



Фиг. 22. Замкнутая амплитудно-фазовая характеристика неустойчивой системы.



Фиг. 23. Замкнутая амплитудно-фазовая характеристика устойчивой системы (см. фиг. 12, 2).

Предположим, что рассматриваемая замкнутая линейная система является устойчивой в разомкнутом состоянии. Для такой системы имеет место следующий известный критерий устойчивости Найквиста:

Критерий устойчивости Найквиста¹⁾. Пусть S — замкнутая система, которая устойчива в разомкнутом состоянии. Система S будет устойчивой тогда и только тогда, когда при обходе замкнутой амплитудно-фазовой характеристики в одном направлении полный охват точки $(1, -180^\circ)$ равен нулю.

¹⁾ Частотный критерий устойчивости, введенный Найквистом и впервые использованный для анализа устойчивости систем автоматического регулирования А. В. Михайловым, в нашей литературе принято называть „критерием устойчивости Найквиста — Михайлова“. — Прим. ред.

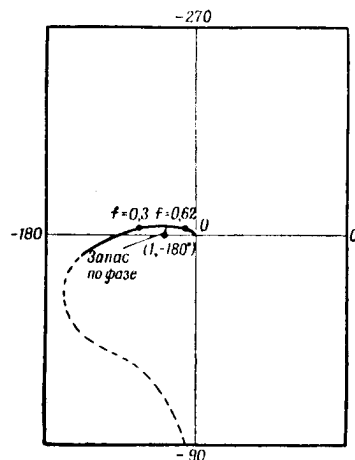
Случаи, представленные на фиг. 20 и 23, согласно этому критерию, соответствуют устойчивым системам, тогда как на фиг. 22 изображена характеристика неустойчивой системы. Система S будет неустойчива, если ее характеристика проходит через точку $(1, -180^\circ)$.

Критерий устойчивости Найквиста можно выразить и в более общем виде, распространив его на системы, неустойчивые в разомкнутом состоянии.

Расчет при помощи амплитудно-фазовых характеристик

Для практического использования критерий Найквиста можно представить в следующем виде:

Практический критерий устойчивости. Система устойчива, если она имеет положительный запас устойчивости по фазе. В противном случае она неустойчива.



Фиг. 24. Амплитудно-фазовая характеристика неустойчивой системы при большом коэффициенте усиления.

Увеличение коэффициента усиления разомкнутой системы в k раз соответствует простому умножению каждого радиуса-вектора на множитель k . Таким образом, повышая коэффициент усиления в 10 раз в примере регулятор — двигатель на фиг. 20, мы получим амплитудно-фазовую характеристику, изображенную на фиг. 24. Теперь характеристика охватывает точку $(1, -180^\circ)$ и, следовательно, замкнутая система неустойчива. Пунктирная часть кривой на фиг. 24

Этот критерий пригоден для системы S , которая является „абсолютно устойчивой“ [3], т. е. которая устойчива при некотором значении коэффициента усиления и неустойчива для больших значений или которая устойчива при всех значениях коэффициента усиления.

Увеличение коэффициента усиления разомкнутой системы в k раз соответствует простому умножению каждого радиуса-вектора на множитель k . Таким образом, повышая коэффициент усиления в 10 раз в примере регулятор — двигатель на фиг. 20, мы получим амплитудно-фазовую характеристику, изображенную на фиг. 24. Теперь характеристика охватывает точку $(1, -180^\circ)$ и, следовательно, замкнутая система неустойчива. Пунктирная часть кривой на фиг. 24

1. Изображение частотных характеристик и расчётные критерии 49

искажена здесь для того, чтобы показать, что кривая асимптотически стремится к вертикальной оси. Видно, что запас по фазе будет отрицательным.

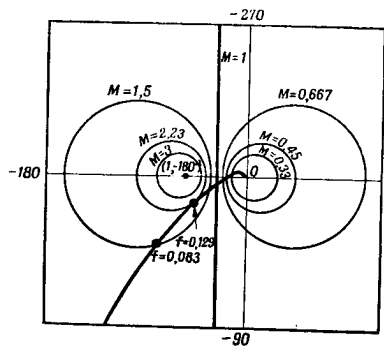
Амплитудные и фазовые круговые диаграммы

При помощи вспомогательных кривых частотные характеристики замкнутой системы, представленные на фиг. 10, можно получить непосредственно по амплитудно-фазовой характеристике разомкнутой системы.

Пусть M — отношение амплитуд замкнутой системы (см. фиг. 10, где r — входной сигнал, а s — выход). Можно показать, что кривые постоянных M („кривые M “) представляют собой окружности на плоскости амплитудно-фазовой характеристики, как показано на фиг. 25. Амплитудно-фазовая характеристика, представленная на фиг. 20, построена на той же диаграмме. При частоте $f=0,083$ характеристика пересекает кривую $M=1,5$. Следовательно, для замкнутой системы отношение амплитуды выхода s к амплитуде входа r при $f=0,083$ равно 1,5. Заметим, что „максимальное значение“ M_p для точек на амплитудно-фазовой характеристике равно 2,23. Можно также показать, что если максимальное значение M_p для амплитудно-фазовой характеристики будет очень большим, например будет равно 4, то характеристика переходного процесса замкнутой системы будет плохой. Амплитудно-фазовая характеристика, близкая к точке неустойчивости $(1, -180^\circ)$, дает высокие значения M_p .

Если $M > 1$, то окружность M имеет радиус

$$\frac{M}{M^2 - 1},$$



Фиг. 25. Круговые амплитудные характеристики M , наложенные на амплитудно-фазовую характеристику.

а центр ее расположен на горизонтальной оси на расстоянии

$$\frac{M^2}{M^2 - 1}$$

слева от начала координат. Если $M = 1$, то окружность сводится к прямой линии, отстоящей на $1/2$ слева от вертикальной оси.

Если $M < 1$, то окружность M имеет радиус

$$\frac{M}{1 - M^2},$$

а центр ее расположен на горизонтальной оси на расстоянии $\frac{M^2}{1 - M^2}$ справа от начала координат.

Обычно применяется следующее расчетное правило [12]. Если частота, при которой достигается максимум M , отлична от нуля, то эта частота называется резонансной частотой. Если при нулевой частоте $M = 1$, то значение частоты, при которой M впервые становится равной 0,707, называется шириной полосы пропускания.

Правила расчета 5. Максимальное отношение амплитуд M_p замкнутой системы должно быть меньше 2 и в среднем должно быть равно 1,3. Ширина полосы пропускания или резонансная частота (если она имеется) должны быть по возможности большими.

Математическое рассмотрение показывает, что низкая резонансная частота или малая ширина полосы пропускания характеризуют медленный переходной процесс. Ширина полосы пропускания (или резонансная частота) в правиле расчета 5 ограничивается требованиями подавления „шумов“ и другими условиями.

Кривые постоянных сдвигов фаз для частотной характеристики замкнутой системы являются окружностями и называются фазовыми круговыми диаграммами („кривые N “).

З а м е ч а н и е. Согласно правилу расчета 5, пунктирная амплитудная характеристика на фиг. 11 является более предпочтительной, чем характеристика, представленная сплошной кривой.

Величина M_p связана с запасом устойчивости по фазе (см. правило расчета 1). Предположим, что частота среза

1. Изображение частотных характеристик и расчетные критерии 51

расположена в точке, где запас по фазе равен 30° . Значение M в этой точке приблизительно равно 1,93, откуда M_p равно по крайней мере 1,93. Когда избыток фазы уменьшается, величина M при частоте среза увеличивается. Таким образом, если M_p меньше 1,9, то запас по фазе будет больше 30° . Аналогично, если $M_p = 1,3$, то запас по фазе будет более 45° .

Рекомендации относительно амплитудно-фазовых характеристик

Относительно амплитудно-фазовых характеристик Комитет дает следующие рекомендации:

Рекомендация 5. На амплитудно-фазовой характеристике рекомендуется давать достаточное число точек с отметкой частоты, указывающих скорость изменения амплитуды и фазы в зависимости от частоты. Использование амплитудных и фазовых круговых диаграмм на одном графике должно быть сведено к минимуму.

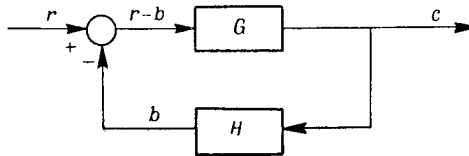
Рекомендация относительно минимального использования круговых диаграмм предлагается для того, чтобы избежать скопления на графике слишком большого числа кривых.

Обычно обозначают передаточную функцию разомкнутой системы в виде произведения KG , где K — коэффициент усиления системы. Часто бывает неудобно определять K само по себе и отделять его от передаточной функции, особенно при экспериментальных исследованиях. Поэтому предлагается следующее:

Рекомендация 6. Пусть G — передаточная функция разомкнутой системы. Плоскость амплитудно-фазовой характеристики при постоянном коэффициенте усиления K и множителе G , зависящем от частоты, должна обозначаться как плоскость G (см. фиг. 20), а не как плоскость KG .

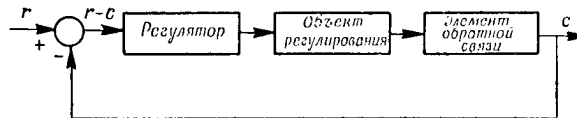
Предыдущее рассмотрение амплитудно-фазовой характеристики было ограничено характеристиками замкнутых систем без элементов в цепи обратной связи, зависящих от частоты. Пусть G обозначает передаточную функцию прямой части петли (регулятора и объекта на фиг. 19), а H — передаточную функцию элемента в цепи обратной связи. Тогда GH является передаточной функцией разомкнутой системы

(фиг. 26). Как уже отмечалось выше, амплитудная характеристика (в логарифмических координатах) разомкнутой системы получается суммированием ординат амплитудных характеристик прямой части системы и элемента в цепи обратной связи. Аналогично и для фазовой характеристики.



Фиг. 26. Система с элементами в прямой цепи и в цепи обратной связи.

При рассмотрении устойчивости можно представить систему так, как если бы все элементы находились в прямой цепи системы, как показано на фиг. 27. Но круговые диаграммы „ M “ и „ N “ в этом случае не определяют амплитудную и фазовую характеристики замкнутой системы, представленной на фиг. 19. Однако, используя приведенный выше



Фиг. 27. Система фиг. 19 с элементом обратной связи, перенесенным в прямую цепь.

критерий для максимума M , можно рассчитать элементы, рассматриваемые в задаче, так, чтобы обеспечить хорошую устойчивость для устройства, показанного на фиг. 27. Система, показанная на фиг. 19, будет тогда тоже устойчивой. Однако частотные характеристики двух систем не будут совпадать. Это может иметь значение там, где необходимо, чтобы выход c особенно точно следовал за входным сигналом r , как например при наводке зенитных орудий.

Таким образом, мы в дальнейшем концентрируем свое внимание на реакции системы c на входной сигнал r . Это обстоятельство имеет первостепенное значение при проектировании следящих систем. При обычных применениях регу-

1. Изображение частотных характеристик и расчетные критерии 53

лятора установка регулятора принимается постоянной или примерно постоянной и система может подвергаться лишь действию возмущений, приложенных не к регулятору. Такими часто являются возмущения нагрузки. Передаточная функция относительно нагрузки L и выхода c будет, вообще говоря, отличаться от передаточной функции, относящейся к переменным r и s . Однако проблема устойчивости остается той же самой. Таким образом, вместо исследования характеристик замкнутой системы при синусоидальных колебаниях нагрузки мы рассматриваем реакцию системы на синусоидальные колебания установки регулятора. Возможны, конечно, случаи, когда расчетчик пожелает непосредственно исследовать реакцию системы на колебания нагрузки.

Обратные характеристики Найквиста

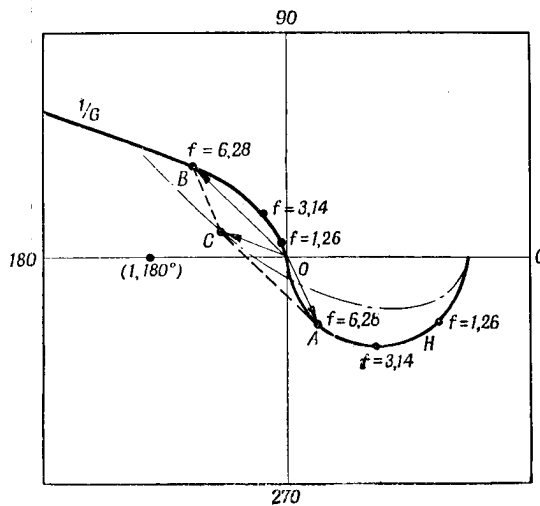
Амплитудно-фазовая характеристика является графиком передаточной функции разомкнутой системы G . Обратная характеристика Найквиста является графиком обратной передаточной функции $1/G$ в полярных координатах и получается следующим образом. Величина радиуса-вектора точки обратной характеристики Найквиста равна амплитуде входа разомкнутой системы, деленной на амплитуду выхода (отношение, обратное тому, которое применяется для амплитудно-фазовой характеристики), а фазовый угол равняется фазовому углу прямой амплитудно-фазовой характеристики с обратным знаком.

Для случая простой обратной связи системы на фиг. 10 обратная характеристика Найквиста для замкнутой системы (связывающая r и c) получается простым смещением точек обратной характеристики разомкнутой системы. Таким образом, обратная амплитудно-фазовая характеристика позволяет легко учесть характеристики элементов в цепи обратной связи. Эти соображения, а также особая наглядность обратных характеристик в низкочастотной области часто делают их удобными в работе. Обратная характеристика Найквиста для примера регулятор—двигатель (фиг. 20) показана на фиг. 28 сплошной кривой. Читатель может видеть, что кривая расположена выше критической точки $(1, 180^\circ)$, а не ниже, как в случае прямой характеристики Найквиста. На фиг. 28 стрелкой указан запас устойчивости по фазе.

1. Изображение частотных характеристик и расчетные критерии 55

обратной амплитудно-фазовой характеристикой замкнутой системы (см. фиг. 26).

Обратные круговые характеристики M являются концентрическими окружностями. Расстояние от начала координат до точки на обратной характеристике замкнутой системы равно просто $1/M$ для отношения амплитуд M , найденного



Фиг. 29. Обратная амплитудно-фазовая характеристика.

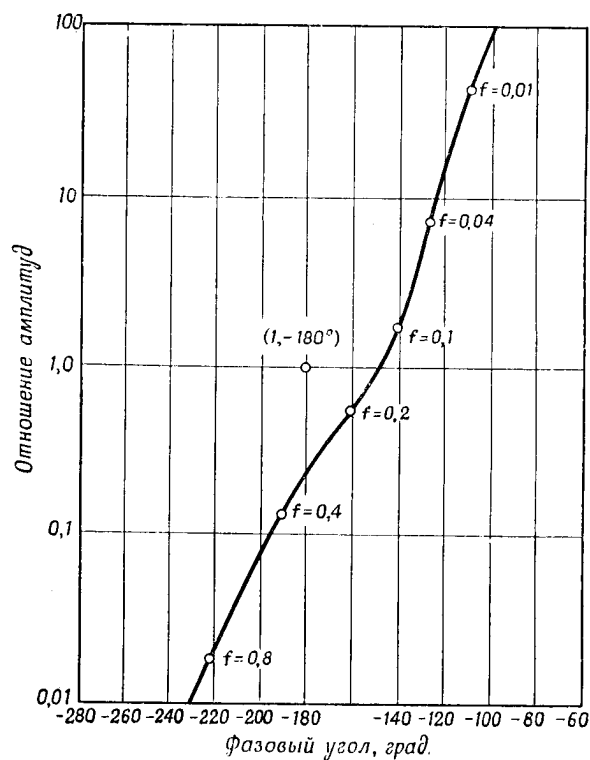
для прямой характеристики Найквиста (выходная амплитуда больше входной). Обратная передаточная функция для замкнутой системы равна просто сумме

$$\frac{1}{G} + H.$$

Рекомендация 7. Пусть G и H — передаточные функции соответственно прямой цепи и цепи обратной связи замкнутой системы. На графике обратной амплитудно-фазовой характеристики такой системы кривые для прямой и обратной частей системы должны обозначаться соответственно как $1/G$ и H (см. фиг. 29). Использование круговых характеристик M должно быть минимальным.

Логарифмическая амплитудно-фазовая характеристика

Если по вертикальной оси отложить отношение амплитуд в логарифмическом масштабе, а по горизонтальной оси —



Фиг. 30. Логарифмическая амплитудно-фазовая характеристика для случая системы регулятор — двигатель.

фазовые углы в линейном масштабе, то получим логарифмическую амплитудно-фазовую характеристику [9]. Применение такой характеристики дает следующие преимущества:

а) логарифмические графики амплитудных и фазовых частотных характеристик, рекомендуемые в этой статье, можно непосредственно преобразовать в логарифмическую амплитудно-фазовую характеристику при помощи циркуля;

1. Изображение частотных характеристик и расчетные критерии 57

б) произведения нескольких составляющих частотной характеристики можно заменить графическим суммированием, и, таким образом, легко выяснить влияние отдельных составляющих на полную характеристику системы;

в) логарифмическая амплитудно-фазовая характеристика позволяет изменять усиление разомкнутой системы простым переносом характеристики вверх или вниз. Изменение усиления не влияет на фазовый угол, как это имело место для амплитудной характеристики в логарифмических координатах. Обычно для изменения усиления системы перемещают вверх или вниз наложенные на логарифмическую амплитудно-фазовую характеристику шаблоны с круговыми диаграммами „М“ и „N“.

Логарифмическая амплитудно-фазовая характеристика имеет некоторые неудобства, заключающиеся в том, что круговые диаграммы „М“ не являются окружностями и, следовательно, для их построения необходимы особые приспособления или шаблоны. Для примера системы регулятор—двигатель, показанной на фиг. 11, логарифмическая амплитудно-фазовая характеристика приведена на фиг. 30.

Рекомендация 8. Рекомендуется вычерчивать логарифмическую амплитудно-фазовую характеристику, откладывая по вертикальной оси отношение амплитуд в логарифмическом масштабе, а по горизонтальной оси — фазовые углы (в градусах) в линейном масштабе.

Передаточные функции

Предположим, что физическая система S находится в покое до момента времени $t = 0$. Пусть в момент $t = 0$ к системе S прикладывается входной сигнал m ($m \equiv 0$ при $t < 0$). Обозначим выходной сигнал системы S через s и предположим, что существуют преобразования Лапласа от m и s .

Передаточной функцией системы S называется частное от деления преобразования Лапласа выходного сигнала на преобразование Лапласа входного сигнала при условии, что это частное единственно.

Рекомендация 9. Независимую переменную передаточной функции следует обозначать через s . Тогда передаточная функция $G(s)$ является функцией комплексного переменного s . При фактических расчетах s должно быть заменено на $j\omega$.

Рекомендации относительно представления данных измерений

При выборе формы представления данных измерения частотных характеристик рекомендуется учитывать общий уровень подготовки специалистов.

Поскольку требуется сосредоточить внимание на структурной схеме изучаемой системы, нежелательно перегружать ее излишними деталями. Если применяется относительно простая измерительная аппаратура, ее можно указать на схеме с помощью специальных обозначений, отличных от обозначений, применяемых для элементов системы, или с помощью пунктирных линий, указывающих реальные или эквивалентные цепи измерительной системы. Если измерительное оборудование сложно, можно повторить простую схему, указав на ней измерительную систему пунктирными линиями. Описание применяемой измерительной системы можно дать в табличной форме в тексте или на блок-схеме измерительной системы. Обычно эти данные включают следующее:

а) *входной импеданс* измерительной системы в механических или электрических единицах и его сопоставление с импедансом основной системы в точках измерения;

б) *чувствительность вход—выход* измерительной системы в целом или по отдельным ее элементам;

в) *эквивалентную схему измерительной системы, ее частотную характеристику* или ее *постоянную времени*, если она существует.

Следует обращать внимание на то, чтобы методы измерения, применяемые для получения частотных данных:

а) обеспечивали воспроизводимость данных;

б) по возможности не искажали „истинной“ частотной характеристики;

в) в случае искажения „истинной“ частотной характеристики позволяли ввести соответствующую поправку;

г) при невозможности ввести поправку допускали качественную оценку влияния методов измерения.

Следует сообщать такие сведения, которые дали бы возможность неподготовленному читателю составить себе качественное представление о значении измерительной системы, а специалисту позволили бы сделать и количественные оценки. Для этого недостаточно ограничиваться такими общими фор-

1. Изображение частотных характеристик и расчетные критерии 59

мулировками, как „станок для испытания проволоки на усталость типа вращающейся опоры“, „линейный акселерометр с широкой полосой пропускания“ или „характеристика системы не имеет заметного запаздывания“. Однако слишком подробные данные тоже не требуются. Надо руководствоваться соображениями здравого смысла, приспособивая описание к основным применениям системы; следует, по возможности, отмечать пределы применения приборов и влияние неконтролируемых факторов окружающей среды.

Дополнительные рекомендации

В частотные данные должны быть включены пределы для входных сигналов, при которых данные линейны, а там, где это возможно, следует указать экспериментальные или теоретические *оценки нелинейности*. Вообще если амплитудная и фазовая характеристики для больших и малых амплитуд синусоидального входного сигнала одинаковы, то процесс можно рассматривать как линейный, по крайней мере вблизи измеряемой точки. Если область линейности неизвестна, то необходимо дать подробное описание экспериментальной установки.

Случай, когда критерий расчета несправедлив

Может случиться, что применяемый обычно критерий устойчивости удовлетворяется, а система при этом практически неустойчива. Для иллюстрации такого случая рассмотрим передаточную функцию разомкнутой системы вида

$$G = \frac{0,486}{s^2 + 0,3s + 1}. \quad (1)$$

Пусть система имеет простую единичную обратную связь. Функция G характеризуется следующими величинами:

$$M_p = 1,34 \text{ при } \omega_r = 1,2;$$

запас устойчивости по фазе — 45° ;

запас устойчивости по амплитуде — ∞ .

Здесь ω_r равно $2\pi \times$ (резонансная частота). Корни уравнения

$$1 + G = 0 \quad (2)$$

равны

$$-0,15 \pm 1,21j. \quad (3)$$

Читатель, знакомый с методом исследования устойчивости по корням характеристического уравнения, сразу увидит, что переходный процесс будет сильно колебательным, так как коэффициент при мнимой части (1,21) намного больше вещественной части (0,15). При участии автора экспериментально установлено, что если вещественная часть мала по величине, например равна 0,1, а коэффициент при мнимой части примерно в 10 или более раз превышает вещественную часть, то система практически неустойчива. Это имеет место при регулировании скорости двигателя.

В настоящем примере нарушено правило расчета 2, поскольку наклон логарифмической амплитудной характеристики при частоте среза приблизительно равен — 4, как показано на фиг. 12, а. Однако *это правило было введено только для того, чтобы обеспечить достаточный запас устойчивости по фазе*, что и имеет место. Таким образом, *удовлетворение критерия устойчивости не обеспечивает достаточной устойчивости или достаточно хорошего переходного процесса*, но может помочь достичь их.

Роль частотного метода

Частотный метод получил широкое распространение в области автоматического регулирования отчасти потому, что многие инженеры чаще предпочитают иметь дело с графическими кривыми, а не с алгебраическими формулами и, как правило, плохо знакомы с простейшими методами [14] решения алгебраических уравнений высокого порядка. Там, где это возможно, желательно построить действительный переходный процесс или по крайней мере найти корни характеристического уравнения. По расположению этих корней можно судить о том, как будет протекать переходный процесс. При наличии моделирующего устройства задача определения переходного процесса значительно упрощается.

Широкое применение для решения алгебраических уравнений нашел в настоящее время метод корневого годографа Эванса [15]. Подобно частотному методу расчета, этот метод имеет тот недостаток, что необходимо графическое построение кривых. Все методы расчета систем автоматического регулирования, основанные на приближенных графических построениях, а также на применении моделей, имеют тот

1. Изображение частотных характеристик и расчетные критерии б1

недостаток, что результат можно получить только для конкретного численного случая. Точный и быстрый расчет часто можно выполнить лишь при помощи математических формул с буквенными обозначениями для некоторых констант. Но для этого требуется большое количество математических выкладок. В действительности оба метода — аналитический и численный — необходимы при анализе систем автоматического регулирования.

Для специалиста в области автоматического регулирования с хорошей математической подготовкой частотный метод предоставляет большие возможности в применении к нелинейным проблемам, где обычные алгебраические методы непригодны [16]. Примером могут служить задачи с насыщением, гистерезисом, переменным демпфированием и т. д. Использование частотного метода для анализа неопределимо в том отношении, что частотные кривые часто указывают на факторы, которыми нельзя пренебрегать в данной задаче.

Частотный метод особенно полезен для начинающих исследователей благодаря той связи, которая существует между частотными характеристиками и переходным процессом физической системы [7].

ПРИЛОЖЕНИЕ

Для пояснения использования передаточных функций в составлении расчетных правил рассмотрим передаточную функцию разомкнутой системы следующего вида:

$$\frac{k(s+a)}{s^2(s+1)}, \quad (4)$$

где a и k — вещественные положительные постоянные, причем $a < 1$. В этом случае асимптотическая амплитудная характеристика составлена из трех прямолинейных отрезков с наклонами -2 , -1 и -2 соответственно (см. фиг. 14).

Отношение амплитуд R для этой передаточной функции будет

$$R = \frac{k \sqrt{a^2 + \omega^2}}{\omega^2 \sqrt{1 + \omega^2}}. \quad (5)$$

Отношение p величины R при $\omega = a$ к величине R при $\omega = 1$ будет равно

$$p = \frac{2}{a(1+a^2)}. \quad (6)$$

Характеристическое уравнение замкнутой системы имеет вид

$$x^3 + x^2 + kx + ak = 0. \quad (7)$$

Алгебраический критерий устойчивости [6] требует, чтобы $1 > a$. Для значений a , близких к единице, например $a = 0,5$, корням предыдущего уравнения будет соответствовать слишком колебательный переходный процесс.

Для $a = 0,5$ имеем $p = 3,2$. Для $a = 0,2$ $p = 9,6$ и корни дают удовлетворительный переходный процесс. Поэтому правило расчета 3 рекомендует выбирать $p \geq 5$.

ЛИТЕРАТУРА

1. Guillemin E. A., *Communication Networks*, vol. 2, New York, 1935, p. 475.
2. Nyquist H., *Bell Syst. Tech. Journ.*, 11, 126 (1932).
3. Боде Г., Теория цепей и проектирование усилителей с обратной связью, Издательство иностранной литературы, 1948.
4. Oldenburger R., *Mathematical Engineering Analysis*, New York, 1950.
5. Brown G. S., Campbell D. P., *Principles of Servomechanisms*, New York, 1948.
6. Thaler G. J., Brown R. G., *Servomechanism Analysis*, New York, 1953.
7. Chestnut H., Mayer R. W., *Servomechanisms and Regulating System Design*, vol. I, New York, 1951.
8. Porter A., *Introduction to Servomechanisms*, London, 1950.
9. „Теория следящих систем“, под ред. Х. Джеймса, Н. Никольса и Р. Филлипса, Издательство иностранной литературы, 1953.
10. Ольденбург Р., Сарториус Г., Динамика автоматического регулирования, Гостехиздат, 1949.
11. Гарднер М., Бэрнс Дж., Переходные процессы в линейных системах, Гостехиздат, 1951.
12. Hall A. C., *Journ. Franklin Inst.*, 242, 287 (1946).
13. Марсу Н. Т., *Trans. AIEE*, 65, 521 (1946).
14. Oldenburger R., *Am. Math. Monthly*, 45, 335 (1948).
15. Evans W. R., *Trans. AIEE*, Part I, 69, 66 (1950).
16. Johnson E. C., *Trans. AIEE*, Part II, 71, 169 (1952).

2. ОТНОСИТЕЛЬНАЯ УСТОЙЧИВОСТЬ ЗАМКНУТЫХ СИСТЕМ ¹⁾

М. Керби и Д. Бомарьяж

Введение

Графическое изображение передаточных функций широко применяется при анализе и синтезе линейных систем автоматического регулирования [1—3]. График сравнительно простой передаточной функции указывает на свойства реакции системы и избавляет от необходимости решать дифференциальное уравнение, устанавливающее связь между выходным и входным сигналами и временем. Частота при построении является независимой переменной, а нанесенная на график функция показывает отношение выходного сигнала к входному при синусоидальном входном воздействии или связанное с ними преобразование.

Если система имеет сосредоточенные параметры, такой график дает три вида информации относительно реакции системы. В качестве передаточной функции он характеризует реакцию системы в установившемся режиме на периодические входные сигналы и работу системы в качестве составного элемента сложных систем. Кроме того, этот график, охватывая или не охватывая критическую точку, характеризует абсолютную устойчивость, или, иначе говоря, предельное поведение (возрастание или спад) конечного участка переходного процесса системы. Наконец, форма графика связана с начальным участком переходного процесса системы, и графики передаточной функции применяются для определения так называемой „относительной устойчивости“.

Понятие передаточной функции для установившихся режимов и определение абсолютной устойчивости являются общеизвестными; применение этой методики было распространено также за пределы систем с сосредоточенными

¹⁾ Kirby M. J., Beaumariage D. C., *Trans. AIEE*, Part II, 72, 22—43 (1953).

параметрами [4—6]. Но метод определения относительной устойчивости еще недостаточно разработан. Характеризующие систему величины обычно определены недостаточно четко, и толкование графиков может быть различным. Например, часто считают, что определение относительной устойчивости производится при помощи некоторого обобщения того же самого метода, посредством которого определяется абсолютная устойчивость. Однако неправомерное обобщение и развитие этого метода могут привести к ошибкам. Такой пример дается ниже, когда рассматривается приложение обычных методов к многоконтурной следящей системе.

В настоящей работе показано, как графики обычных передаточных функций определяют относительную устойчивость, поскольку они в то же время являются графиками преобразования Лапласа от импульсной переходной функции. Преобразование Лапласа оценивает временной процесс путем сравнения его с ядром преобразования e^{-st} — экспоненциальной функцией, свойства которой можно изменять, изменяя величину s .

Преобразование осуществляется путем интегрирования по времени; получение его легко представить как физический процесс, в котором имеют дело с произведениями, площадями и другими величинами, тесно связанными с временной реакцией системы. Хотя преобразование Лапласа для реакции определенной системы получают чаще с помощью операторного метода, чем посредством интегрирования, тем не менее график преобразования можно всегда интерпретировать при помощи процесса интегрирования, сравнения и измерения. Это дает возможность установить непосредственное соотношение между различными свойствами графика преобразования и существенными свойствами временной кривой, не используя промежуточные шаги и понятия, такие, как например плоскость комплексной частоты, расположение полюсов и нулей, фаза и другие данные, обычно связанные с анализом передаточных функций системы.

Не менее существенно и то обстоятельство, что применение преобразования Лапласа как средства оценки не ограничивается обычными графиками передаточных функций. Его можно использовать и для других целей:

1. В большинстве практических задач в качестве входных величин фигурируют не только импульсы; реакции си-

стемы на различные входные сигналы легко могут быть исследованы и сравнены.

2. Можно исследовать реакцию линейных систем, для которых понятие отношения преобразований выходного и входного сигналов недействительно, например для систем с элементами, изменяющимися во времени.

3. Необычными методами графического изображения преобразования можно воспользоваться для получения новых данных о реакции систем.

Общие черты метода

Ниже изложен метод определения относительной устойчивости, который имеет весьма широкое применение и лучше всего поддается обобщению. Однако, несмотря на его широкое практическое применение, он обычно не рассматривался в той последовательности, которая указана ниже.

1. Выбираем подходящую входную функцию и составляем дифференциальное уравнение, показывающее связь между выходным сигналом системы θ_o , входным сигналом θ_i и временем t .

2. Составляем преобразование Лапласа для каждого члена. Из преобразованных членов получаем уравнение, в котором преобразование Лапласа $u(s)$ выходного сигнала $\theta_o(t)$ является неизвестной функцией переменной Лапласа s .

3. Решаем уравнение относительно $u(s)$ как функции s .

4. Выбираем такой способ изменения s , при котором построенная кривая $u(s)$ становится чувствительной к представляющим интерес свойствам $\theta_o(t)$, и строим график для $u(s)$ как функции s . Обычный метод заключается в том, что действительная часть s (затухание) приравнивается нулю и изменяется мнимая часть (частота).

5. Отмечаем существенные свойства графика $u(s)$.

6. Определяем свойства функции времени $\theta_o(t)$, соответствующие свойствам $u(s)$, на основании опыта или путем непосредственного сопоставления с графиками преобразования знакомых функций.

Соотношения между специфическими свойствами графиков преобразования и временными кривыми рассматриваются ниже.

Удобство этого метода состоит в том, что преобразование $y(s)$, которое дает возможность лишь оценить временную реакцию $\theta_o(t)$, обычно значительно легче получается и выражается графически, чем сама величина $\theta_o(t)$. Однако здесь необходимо учитывать два ограничения. Во-первых, величина, нанесенная на график как преобразование Лапласа функции $y(s)$, должна быть равна интегралу

$$y(s) = \int_{t=0}^{t=\infty} e^{-st} \theta_o(t) dt, \quad (1)$$

хотя ее обычно получают другими способами (например, операторным методом или экспериментальным путем). Кроме того, этот интеграл должен сходиться, т. е. рассматриваемая система должна быть абсолютно устойчива; обычно это очень легко определить.

Указанные выше действия, состоящие из шести шагов, применяются в тех случаях, когда определяется относительная устойчивость системы с постоянными параметрами путем построения графика передаточной функции как функции частоты. Тогда первые четыре ступени не выделяются в том виде, как было указано; их существование даже ускользает от внимания. Это объясняется тем, что передаточная функция в этом случае равна преобразованию Лапласа для выходного сигнала, когда система возбуждается единичным входным импульсом и не имеет начального запаса энергии. Преобразование выходного сигнала (для системы с постоянными параметрами) равно произведению передаточной функции системы и преобразования входного сигнала. Передаточная функция системы в этом случае равняется передаточной функции замкнутой системы при $s = j\omega$, а преобразование единичного входного импульса равно единице. Иными словами, общепринятый способ построения графика передаточной функции в зависимости от частоты эквивалентен выбору входа в виде импульса, нахождению преобразования выходного сигнала системы, выбору такого способа изменения величины s , при котором s будет чисто мнимой величиной, и построению графика преобразования.

Роль преобразования Лапласа в обычных графиках была затемнена понятием передаточной функции частично потому,

что частотные характеристики при установившемся режиме обычно легче измеряются в лабораторных условиях, чем характеристики системы в переходном процессе. Кроме того, теория устойчивости многих систем автоматического регулирования в значительной мере заимствована из техники связи, которой чаще приходится иметь дело с характеристиками систем при установившемся режиме, чем с характеристиками в переходном процессе.

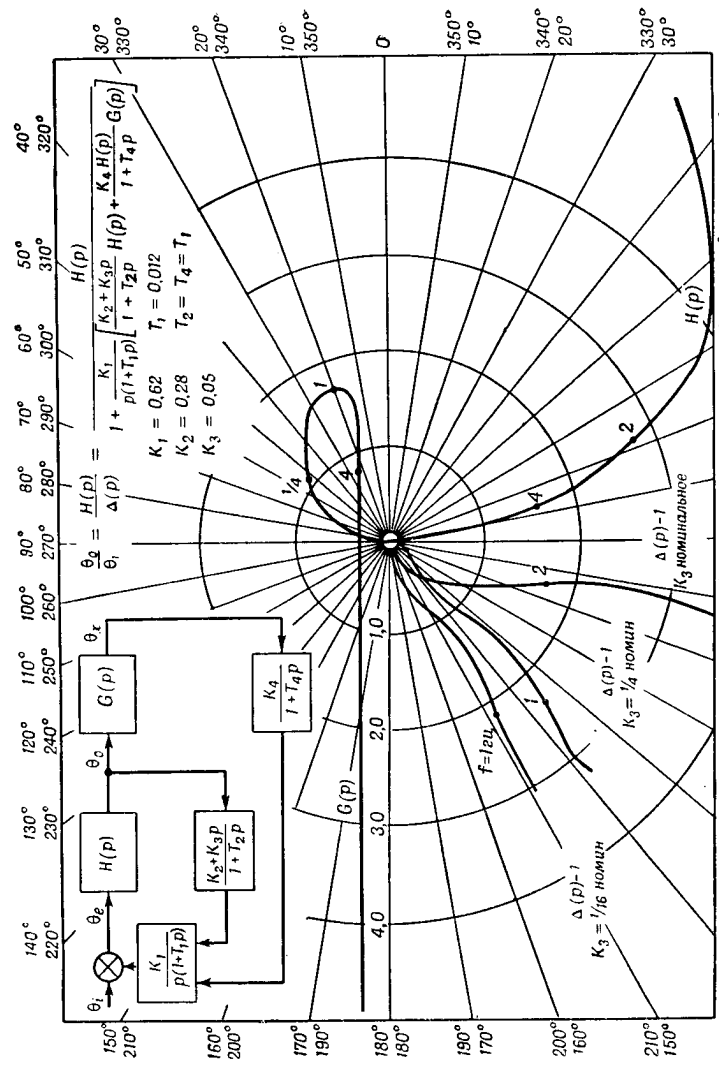
Ниже приводится случай, когда использование обычных передаточных функций, применяемых для определения абсолютной устойчивости, является неприемлемым для определения относительной устойчивости.

Анализ многоконтурной следящей системы

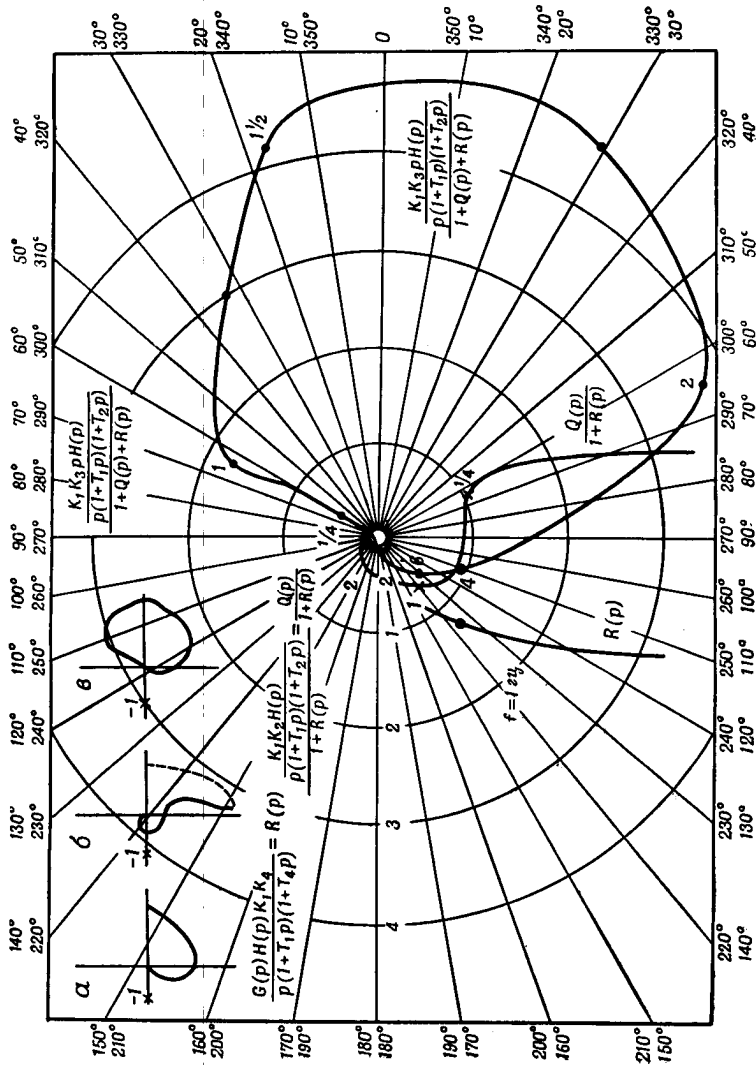
В простой следящей системе второго порядка с единичной обратной связью передаточная функция, применяемая для определения абсолютной устойчивости, может быть непосредственно использована для получения данных об относительной устойчивости системы. Обычно для этого при помощи, например, круговых диаграмм определяют максимум амплитудной характеристики и частоту, при которой он наблюдается, и далее находят характерные точки переходного процесса системы, которые единственным образом определяются для такой системы второго порядка [1]. На основе опыта и интуиции этот метод был распространен на системы более высокого порядка. Ниже приводится объяснение соотношения между амплитудно-фазовыми характеристиками и переходными процессами, исходя из измерительного свойства интеграла прямого преобразования Лапласа.

Для многих таких систем высшего порядка, особенно для многоконтурных следящих систем и систем с сочетанием авторегулирования по замкнутому и разомкнутому циклам, как например систем, описанных в работе [7], графики, построенные для определения абсолютной устойчивости, нельзя непосредственно использовать для получения данных об относительной устойчивости системы. В качестве примера рассмотрим многоконтурную следящую систему, изображенную на фиг. 1.

Для определения абсолютной устойчивости этой системы нужно найти определитель системы $\Delta(p)$, который является

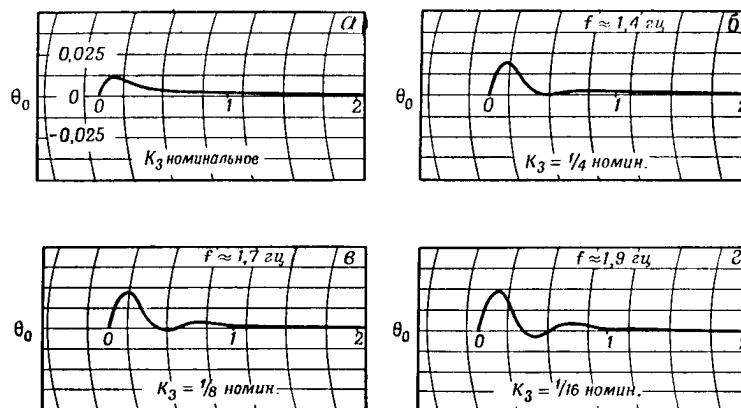


Фиг. 1. Частотные диаграммы для многоконтурной следящей системы.



Фиг. 2. Графики устойчивости многоконтурной следящей системы для определения запасов устойчивости по отношению к K_B .
 Графики а, б и в соответствуют действительным положительным частотам. Ни в одном из этих графиков нет охвата критической точки.

знаменателем передаточной функции [8]. Функция, построенная для определения устойчивости, равна $\Delta(p) - 1$, и поскольку для данного случая K_1 является общим множителем, этот единственный график $\Delta(p) - 1$ достаточен для определения абсолютной устойчивости. Система устойчива для

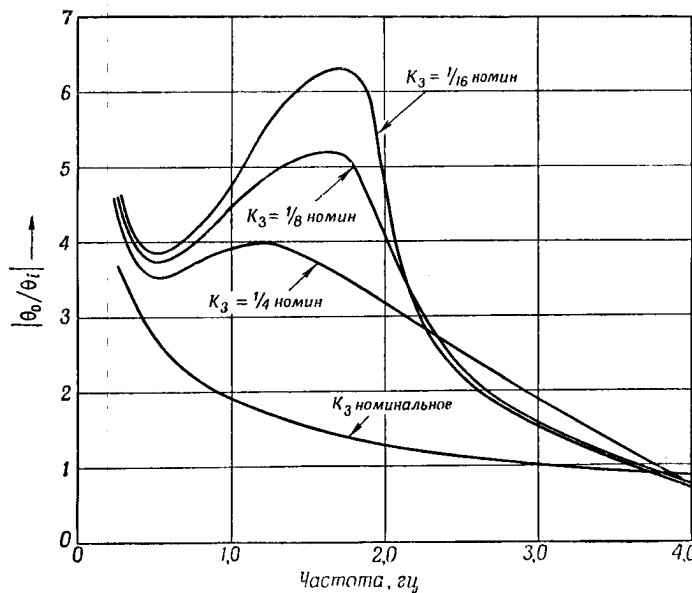


Фиг. 3. Характеристики $\theta_0(t)$ в переходном процессе для многоконтурной следящей системы при различных значениях K_3 .

выбранных величин передаточных коэффициентов, но, что более важно, запас устойчивости по усилению, определяемый из графика $\Delta(p) - 1$, является запасом только по отношению к K_1 . Если один из других передаточных коэффициентов, например K_3 , уменьшается, $\Delta(p) - 1$ изменяется, как показано на графике, а запас усиления по отношению к K_1 увеличивается. Чтобы найти влияние изменения K_3 на запас устойчивости по отношению к K_3 , нужно построить другие графики и использовать обобщенный критерий устойчивости, выведенный Боде [8] для многоконтурной системы.

Поскольку система была устойчива при нормальных величинах передаточных коэффициентов, обобщенный критерий устойчивости утверждает, что „общие числа обходов критической точки по часовой стрелке и против часовой стрелки должны быть равны друг другу в серии диаграмм Найквиста, построенных для отдельных контуров, получающихся при предположении, что все передаточные коэффициенты равны нулю,

и при последующем восстановлении их в любом порядке до нормальной величины". В частности, если значение K_3 будет восстановлено последним, конечный график покажет запас устойчивости по усилению по отношению к K_3 . Следуя этой



Фиг. 4. Графики преобразования Лапласа от $\theta_0(t)$ для импульсного входа $\theta_i(t)$.

процедуре, был составлен ряд графиков, показанных на фиг. 2, и из последнего графика видно, что при уменьшении K_3 система не будет менее устойчивой. Нужно отметить, что график для $\Delta(p) = 1$ на фиг. 1 при уменьшении значения K_3 „приближается к точке -1 “.

Как показывают графики переходных процессов системы, приведенные на фиг. 3, реакция приобретает более колебательный характер по мере уменьшения K_3 . На фиг. 3, а показано изменение θ_0 во времени, соответствующее импульсному входу θ_i при номинальных величинах передаточных коэффициентов, отмеченных на фиг. 1. Когда K_3 уменьшается, графики переходных процессов на фиг. 3, б, в и г характеризуют более колебательную систему, чего нельзя

было обнаружить на предыдущих диаграммах. Если, однако, построить график по преобразованию Лапласа величины θ_o для различных значений K_3 , как это сделано на фиг. 4, то можно убедиться в том, что этот график точно характеризует более колебательную реакцию при уменьшении K_3 . Так как истолкование и развитие графиков преобразования, подобных приведенному на фиг. 4, основаны на методе получения интеграла преобразования в уравнении (1), мы остановимся на этом методе подробнее.

Процесс измерения

Метод измерения свойств функции времени $\theta_o(t)$ при помощи преобразования $y(s)$ непосредственно вытекает из определения преобразования [уравнение (1)]. Функция времени $\theta_o(t)$ множится на ядро e^{-st} и произведение суммируется от нуля до бесконечности по оси времени. Переменная Лапласа s имеет, вообще говоря, комплексную форму с действительной частью σ и мнимой частью $j\omega$. Таким образом, интеграл (1) представляется суммой двух действительных интегралов, которые являются действительной и мнимой частями $y(s)$.

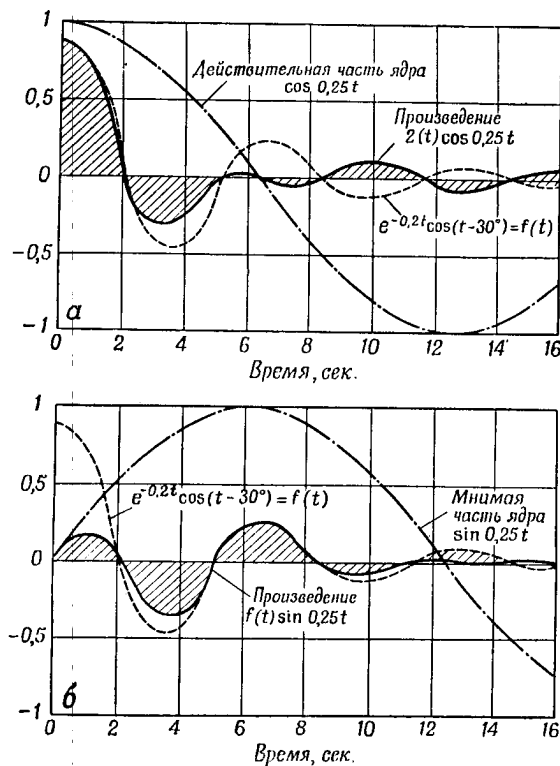
$$y(s) = \int_{t=0}^{t=\infty} e^{-\sigma t} \cos \omega t \theta_o(t) dt - j \int_{t=0}^{t=\infty} e^{-\sigma t} \sin \omega t \theta_o(t) dt. \quad (2)$$

На фиг. 5—8 показан способ, с помощью которого это умножение и суммирование сводится к сравнению $\theta_o(t)$ и ядра. Ступени образования интеграла преобразования для функции времени, характеризующей затухающие колебания, $\theta_o(t) = e^{-0,2t} \cos(t - 30^\circ)$, показаны для трех значений частоты ядра $j\omega$. Действительная часть σ от s принята равной нулю.

Когда частота ядра ω намного ниже частоты $\theta_o(t)$ (равной единице), обе части произведения $\theta_o(t) e^{-st}$ — действительная и мнимая — будут попеременно положительными и отрицательными, так что интегралы произведения от $t = 0$ до $t = \infty$ будут невелики (см. фиг. 5). Аналогично, когда ω намного превышает частоту $\theta_o(t)$, произведения будут снова

2. Относительная устойчивость замкнутых систем 73

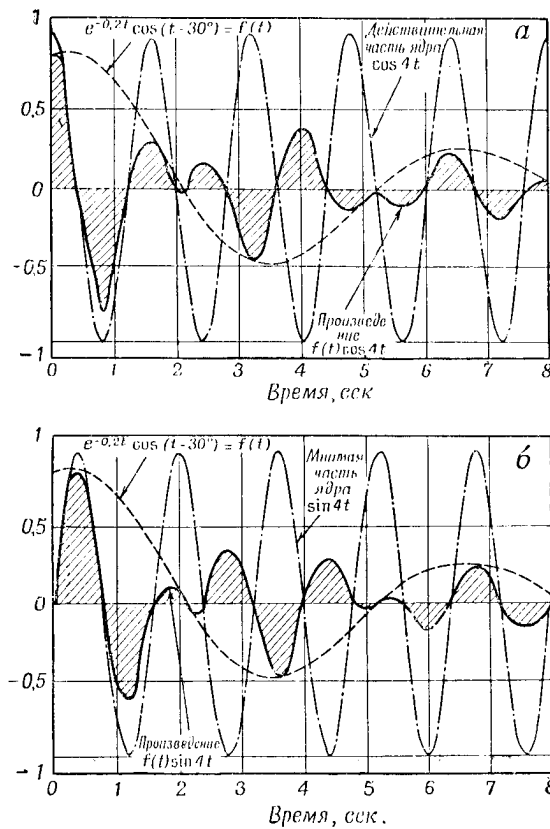
попеременно положительными и отрицательными и интегралы также будут невелики (см. фиг. 6). На фиг. 5—8 интеграл равен сумме заштрихованных площадей.



Фиг. 5. Образование действительной и мнимой частей преобразования Лапласа от $f(t)$ для частоты ядра, значительно меньшей, чем частота $f(t)$.

Когда частота ядра приближается к частоте $\theta_0(t)$, как показано на фиг. 7, действительная и мнимая части произведения $\theta_0(t) e^{-st}$ хотя и изменяются по величине в последовательные моменты времени, но в общем имеют некоторые не нулевые средние значения, так что интеграл (2) будет иметь довольно значительную величину. Относительные величины действительной и мнимой частей этого интеграла для

какого-либо значения ω зависят от того, будет ли фаза составляющей $\theta_o(t)$ (имеющая частоту ядра ω) приближаться эту составляющую к $\cos \omega t$ или $\sin \omega t$.



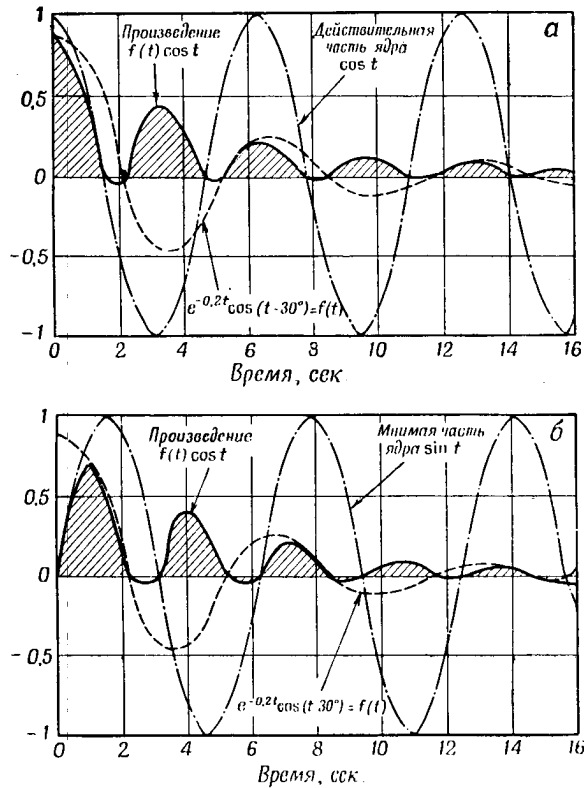
Фиг. 6. Образование действительной и мнимой частей преобразования Лапласа от $f(t)$ для частоты ядра, значительно более высокой, чем частота $f(t)$.

График преобразования $u(s)$ как функции ω дан на фиг. 8. Здесь показаны действительная и мнимая части и векторная сумма. Кривые были рассчитаны при помощи операторного уравнения для преобразования конкретной функции времени,

использованной на фиг. 5—7:

$$y(s) = \gamma \frac{s + a_0}{(s + a_1)^2 + \beta^2}, \quad (3)$$

где $\gamma = 0,866$, $a_0 = 0,777$, $a_1 = 0,2$ и $\beta = 1$.

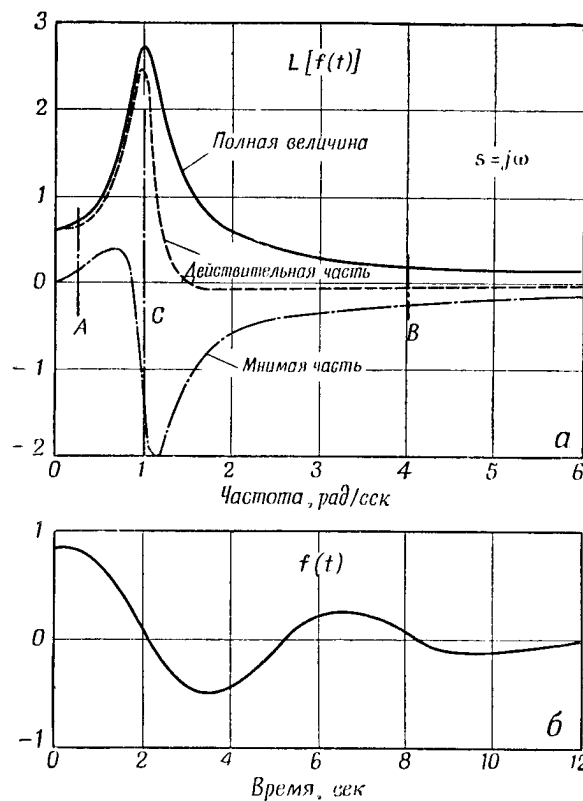


Фиг. 7. Образование действительных и мнимых частей преобразования Лапласа от $f(t)$ для частоты ядра, равной частоте $f(t)$.

Величины, соответствующие частотам, использованным для фиг. 5—7, обозначены на фиг. 8 точками А, В и С соответственно.

Процесс построения графика преобразования на фиг. 8, независимо от того, проводится ли он на основании уравнения

типа (3) или на основании экспериментальных или других данных, сводится к выбору величин для частоты ω ядра e^{-st} , получению произведений и суммированию их для каждой величины ω , как показано на фиг. 5—7. Ядро e^{-st} можно считать



Фиг. 8. Графики преобразования Лапласа от $f(t) = e^{-0,2t} \cos(t-30^\circ)$ и временная кривая.

Частоты А, В и С использованы соответственно в графиках фиг. 5, 6 и 7.

„испытательной“ функцией, а процесс изменения его частоты и построения графика преобразования можно рассматривать как процесс исследования по точкам временной функции. Когда выбрано значение для ω , соответствующее частоте

составляющей, присутствующей во временной функции $\theta_o(t)$, график преобразования для этой частоты будет иметь большую величину, указывая на присутствие соответствующей частотной составляющей в $\theta_o(t)$. Форма графика $u(s)$ и в особенности формы пиков и относительные размеры действительной и мнимой частей $u(s)$ имеют важное значение для определения характера временных кривых.

Интерпретация графиков преобразования

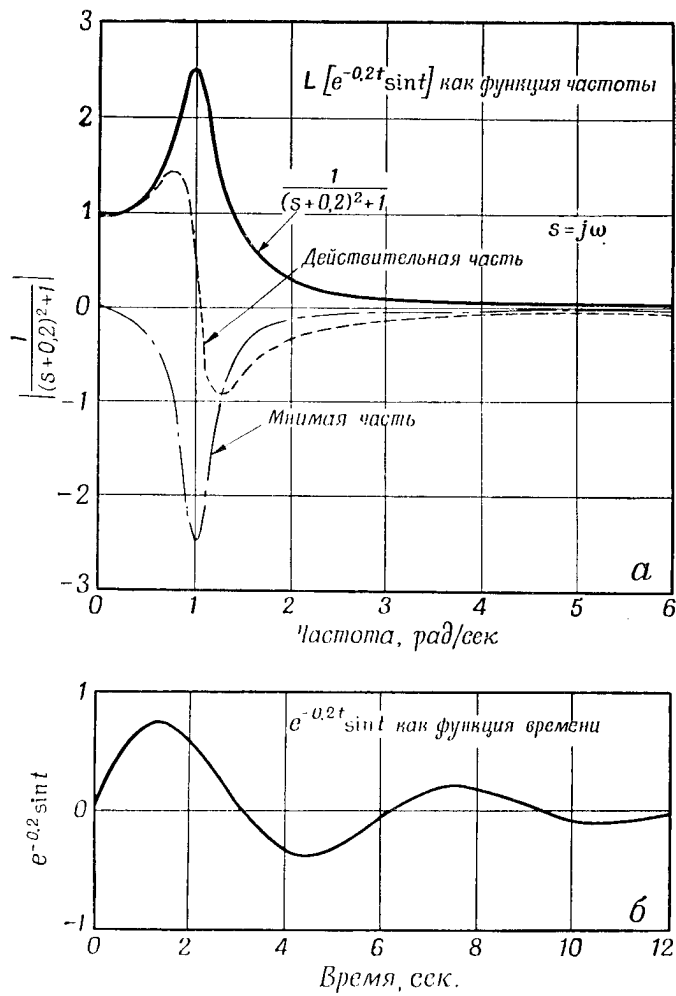
Шестой и последний шаг в общем методе интерпретации состоит в том, что отмечаются наиболее существенные свойства $u(s)$, определяются соответствующие свойства $\theta_o(t)$ и на бумаге либо в воображении строится временная функция, обладающая этими свойствами. В данном разделе определяются взаимоотношения между специфическими свойствами графиков преобразования и временными кривыми и иллюстрируются процессы аппроксимации $\theta_o(t)$.

Важными свойствами переходных процессов следящих систем обычно являются те, которые связаны с колебаниями; главная задача состоит в определении частоты, затухания и других свойств каждого существенного колебания. Соответствующие свойства преобразований связаны с пиками на графиках, изображающих зависимость $u(s)$ от s .

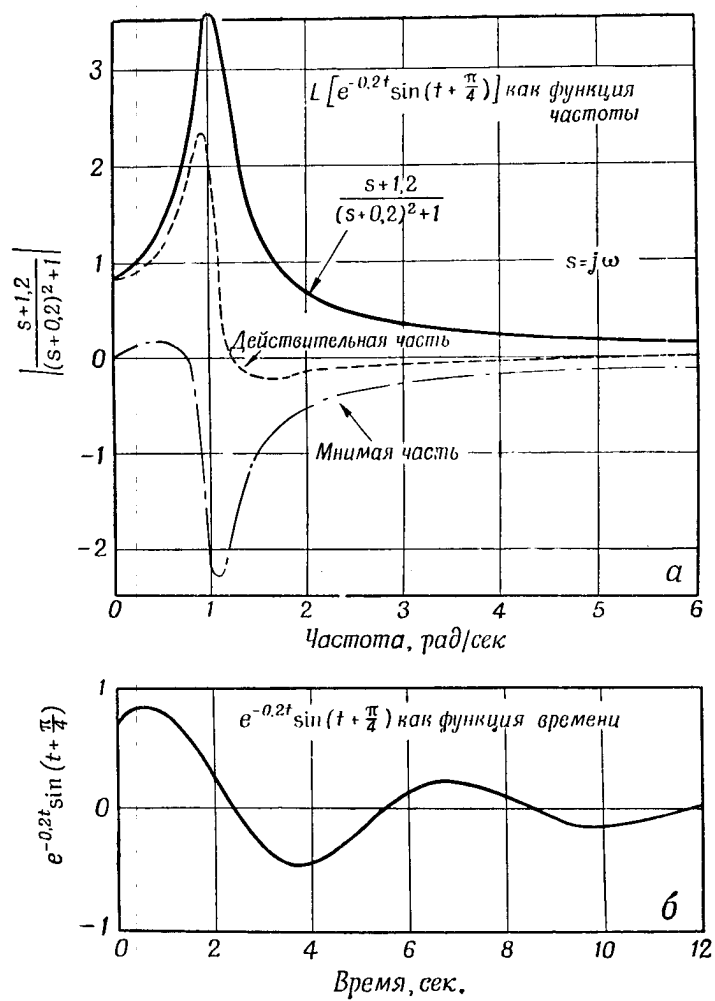
Первая ступень в интерпретации графика преобразования — вычисление величины отношения затухания к частоте основного колебания α/β ¹⁾. Последующие действия в процессе анализа будут зависеть от величины этого отношения.

Затухание колебания в $\theta_o(t)$ зависит от остроты соответствующего пика в $u(s)$. Общее численное уравнение, которое бы показывало соотношение ширины пика и скорости затухания $\theta_o(t)$, получить нелегко. На фиг. 9—17 приведены девять стандартных графиков характеристик реакции систем второго порядка с отношениями затухания к частоте, равными 0,2, 0,5 и 0,8, и тремя различными значениями фазового угла. Остроту пика можно определить по его высоте в точках возле его максимума при частотах не более чем на 20% выше или ниже частоты верхней точки. Когда

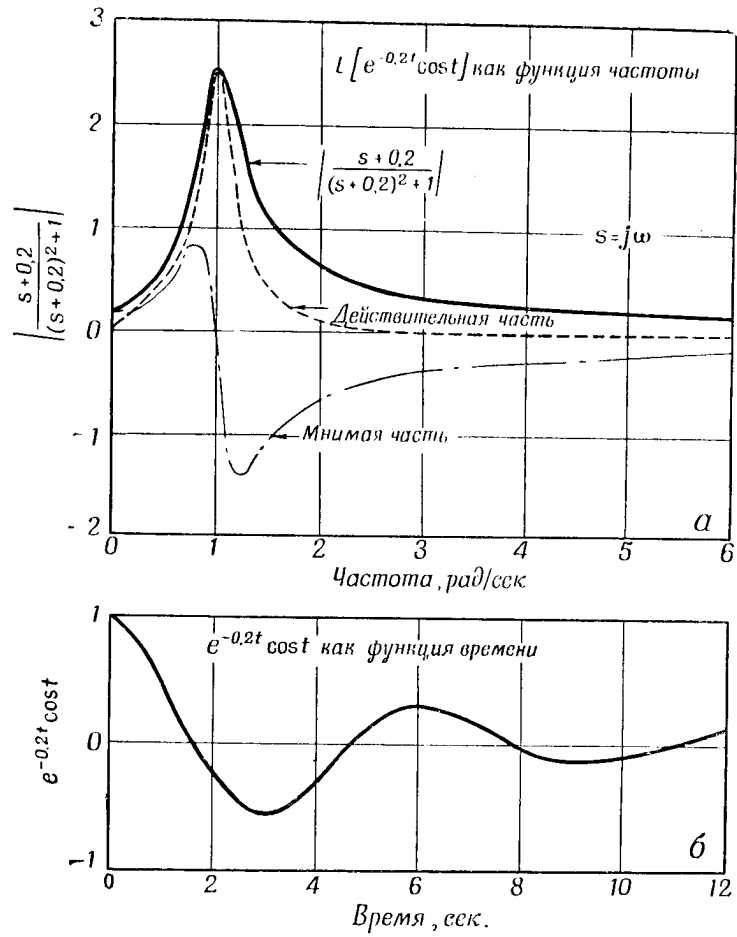
¹⁾ Рассматриваются колебания вида $Ae^{-\alpha t} \sin(\beta t + \varphi_0)$. — Прим. ред.



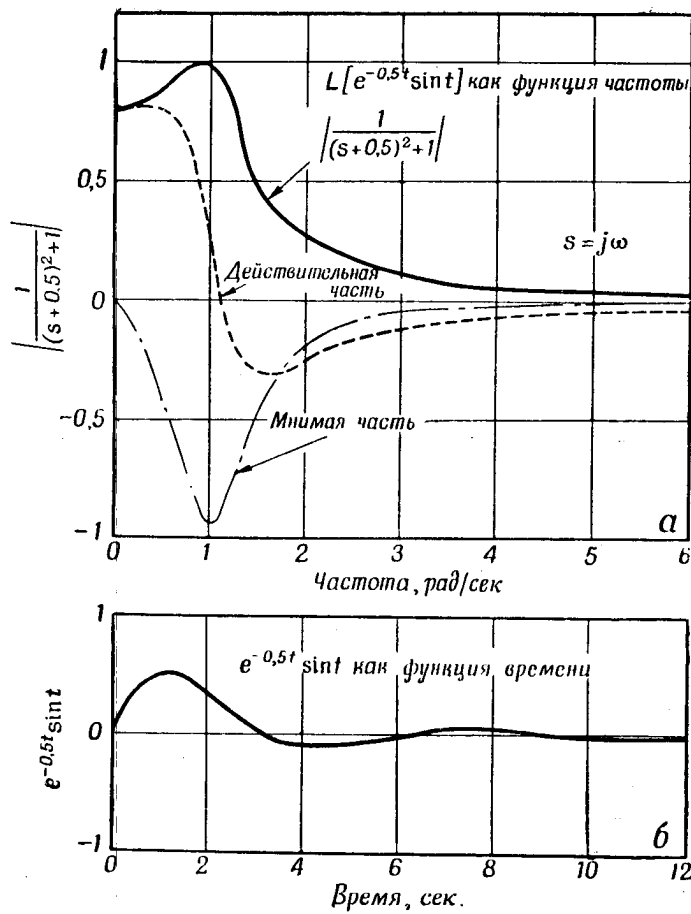
Фиг. 9. Преобразование Лапласа и временная кривая $e^{-0.2t} \sin t$.



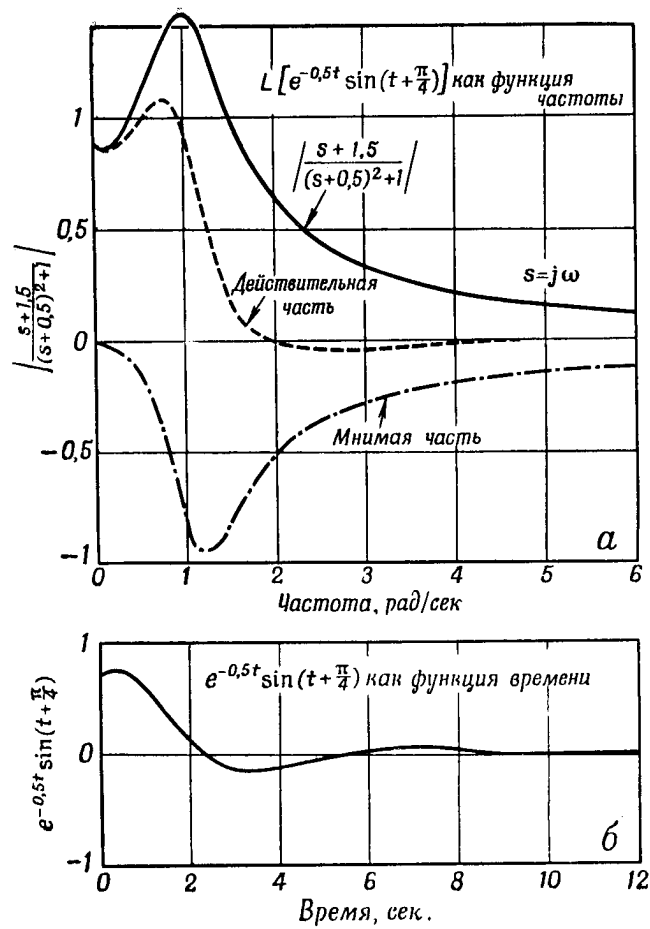
Ф и г. 10. Преобразование Лапласа и временная кривая $e^{-0,2t} \sin(t + \pi/4)$.



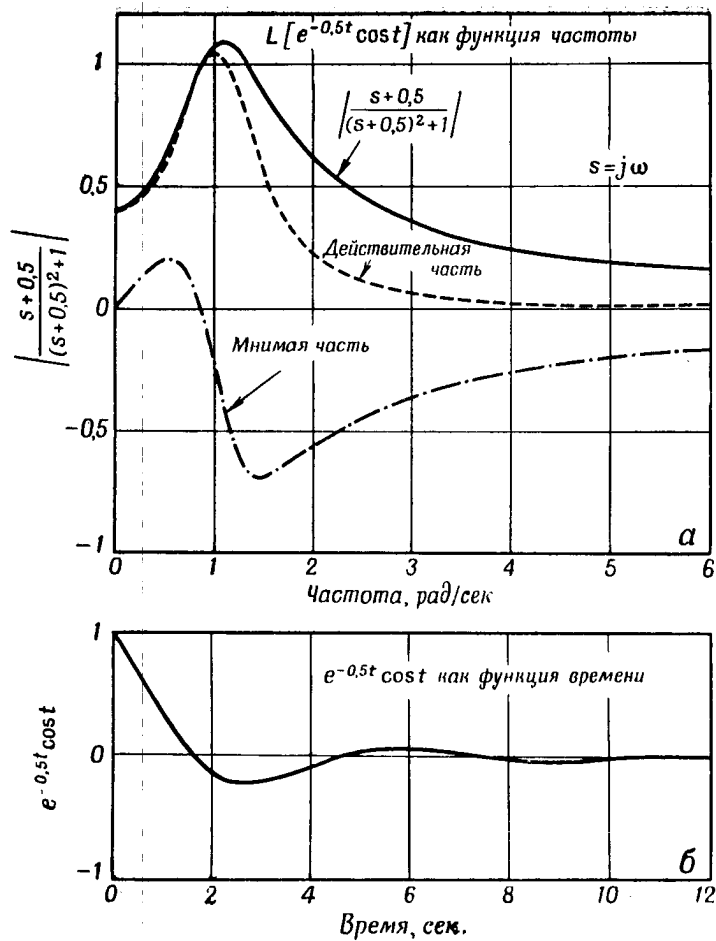
Фиг. 11. Преобразование Лапласа и временная кривая $e^{-0.2t} \cos t$.



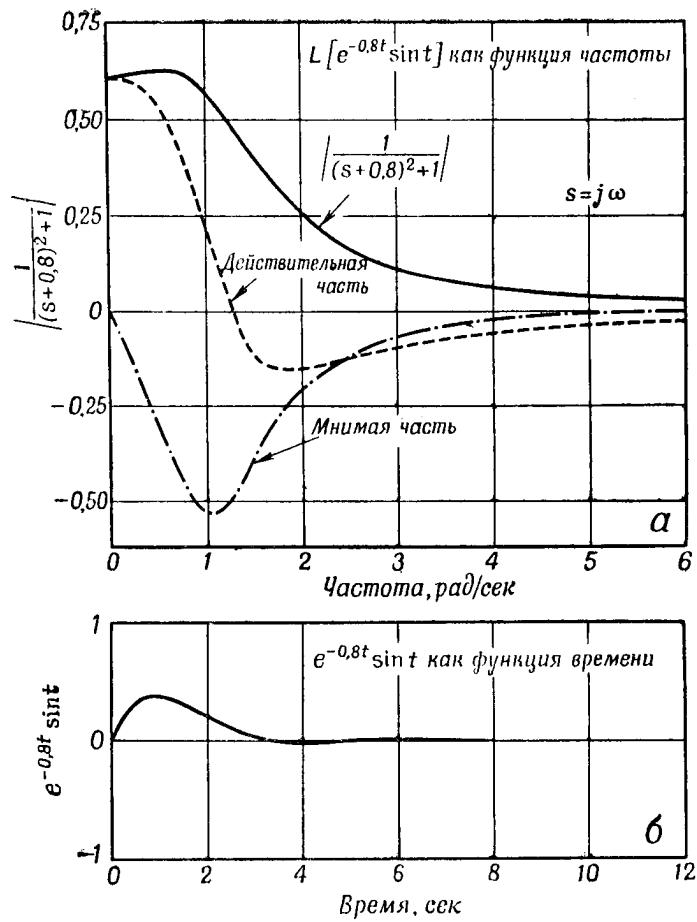
Фиг. 12. Преобразование Лапласа и временная кривая $e^{-0.5t} \sin t$.



Ф и г. 13. Преобразование Лапласа и временная кривая $e^{-0.5t} \sin(t + \pi/4)$.



Ф и г. 14. Преобразование Лапласа и временная кривая $e^{-0.5t} \cos t$.

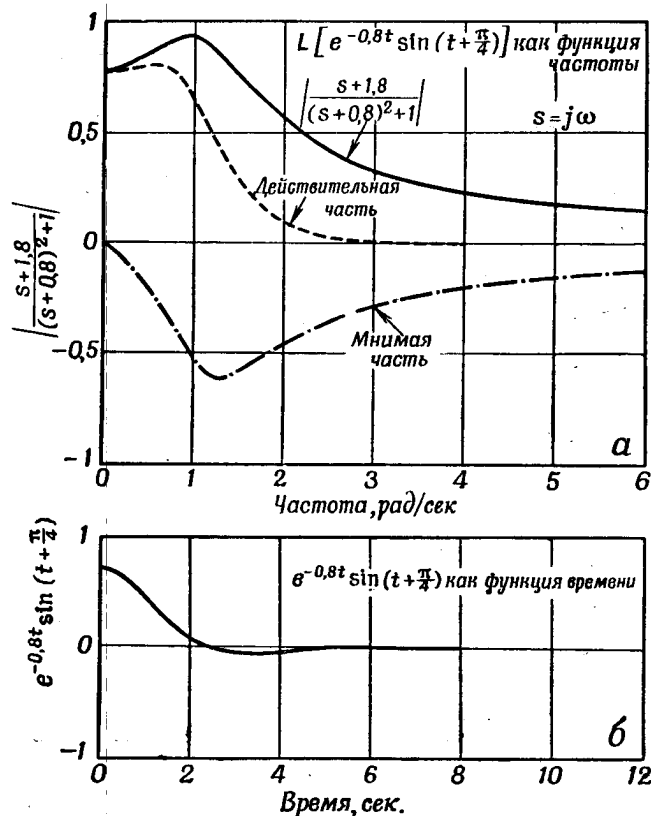


Фиг. 15. Преобразование Лапласа и временная кривая $e^{-0.8t} \sin t$.

2. Относительная устойчивость замкнутых систем

85

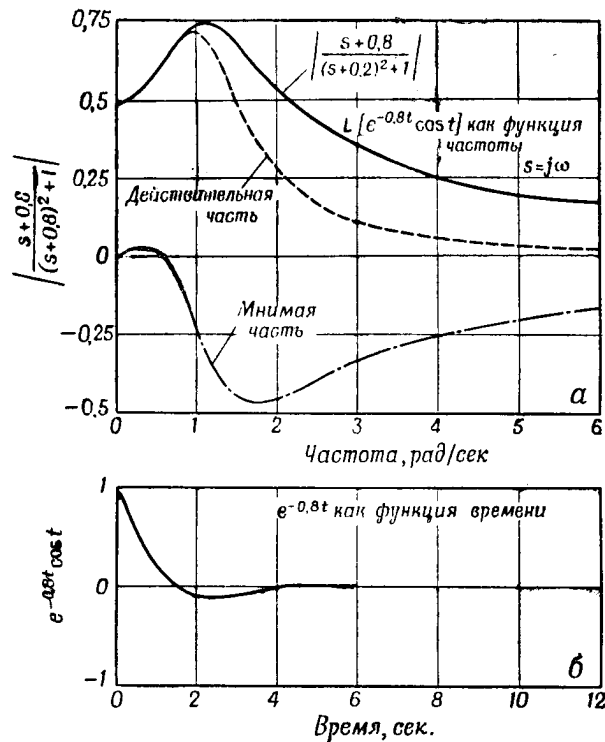
величина отношения затухания к частоте равна 0,2, величина $y(s)$ в точках, находящихся при частотах на 20% выше или ниже частоты наивысшей точки, составляет 60—80% максимального значения (фиг. 9—11). Когда величина отношения α/β равна 0,5 (фиг. 12—14), величина $y(s)$ в этих точках составляет 90—95% наибольшей величины, а когда $\alpha/\beta = 0,8$



Фиг. 16. Преобразование Лапласа и временная кривая $e^{-0.8t} \sin(t + \pi/4)$.

(фиг. 15—17), величина $y(s)$ составляет 98—99% наибольшей величины при частотах, расположенных на 20% выше или ниже частоты максимума.

При частотах, отклоняющихся более чем на 20% от частоты максимума, форма кривой возле нулевой частоты и вторичные пики на более высоких частотах (если таковые имеются) оказывают несоразмерно большое влияние на определяемую величину. Хотя указанные факторы и влияют на форму кривой (в смысле остроты пика) на графике $y(s)$, они не имеют отношения к скорости затухания основного колебания.



Ф и г. 17. Преобразование Лапласа и временная кривая $e^{-0.8t} \cos t$.

Когда величина отношения затухания к частоте $\alpha/\beta = 0,5$ или меньше, т. е. когда коэффициент затухания основного колебания в каждом цикле равен $e^{-\pi}$ или менее, существуют определенные и независимые соотношения между специфическими свойствами $y(s)$ и свойствами $\theta_0(t)$. Частота или

величина s , при которой достигается пик в $y(s)$, очень близка к соответствующей частоте колебания в $\theta_o(t)$. Величина $y(s)$ в точке максимума характеризует размах колебания; размах понимается здесь как произведение начальной амплитуды огибающей и постоянной времени ($1/\alpha$) [см. уравнение (8)]. Относительные величины действительной и мнимой частей $y(s)$ у его наивысшей точки точно указывают на близость $\theta_o(t)$ к косинусоиде или синусоиде, т. е. на фазу. Поэтому, когда установлено, что величина отношения затухания к частоте невелика, можно определить последовательно и в любом порядке размах колебания, его частоту и фазу.

Изучаемый график преобразования всегда полезно сравнить со стандартными графиками, данными на фиг. 9—17, чтобы подкрепить оценку $\theta_o(t)$. После определения свойств главного колебания нужно рассмотреть форму графика преобразования при частотах, близких к нулю и значительно превышающих частоту максимума. Различия между рассматриваемым графиком и стандартными графиками систем второго порядка, приведенными на фиг. 9—17, могут указывать на существование неколебательных составляющих или второстепенных колебаний в $\theta_o(t)$.

На практике для большинства следящих систем затухание нельзя считать незначительным. Затухание главного колебания происходит в течение небольшого числа периодов. Величина затухания $\alpha/\beta = 0,8$ является типичной. Когда величина этого отношения больше 0,5 или 0,6, соотношения между парами свойств $y(s)$ и $\theta_o(t)$ уже не являются более независимыми. Расположение максимальной величины (пика) $y(s)$ зависит не только от частоты колебания $\theta_o(t)$, но также и от фазы. Частота колебания значительно выше частоты максимума $y(s)$, если колебание близко к синусоиде, и ниже ее, если оно близко к косинусоиде. Однако фаза $\theta_o(t)$ не определяется с помощью простого измерения действительной и мнимой компонент у пика $y(s)$. Определение частоты и фазы нужно проводить одновременно путем сравнения формы главного пика изучаемого графика преобразования со стандартными графиками на фиг. 9—17. После определения частоты и фазы можно определить размах колебаний по амплитуде максимальной величины $y(s)$ при помощи уравнения (8). Последний этап в определении $\theta_o(t)$ — рассмотрение частей графиков преобразования при частотах, расположенных зна-

чительно выше или ниже частоты максимума, для исследования неколебательных составляющих и второстепенных колебаний.

Девять стандартных графиков, представленных на фиг. 9—17, показывают импульсные реакции системы второго порядка для значений α/β (отношение затухания к частоте), равных 0,2; 0,5 и 0,8, близкие к косинусоиде, синусоиде, а в одном случае—к промежуточной кривой между косинусоидой и синусоидой. В связи с этими графиками возникает ряд понятий и соотношений, способствующих интерпретации графиков преобразования. Каждый, кто пользуется этими стандартными графиками, может разработать свои собственные правила, приложимые к отдельным системам и к решению часто встречающихся задач. Каждый стандартный график является преобразованием временной кривой $\theta_o(t)$ следующей формы:

$$\theta_o(t) = Ae^{-\alpha t} \sin(\beta t + \varphi). \quad (4)$$

Преобразование уравнения (4) [9] равно

$$y(s) = \frac{A\beta}{(s + \alpha)^2 + \beta^2}, \quad (5)$$

когда $\varphi = 0$, и

$$y(s) = \frac{A\beta}{\sqrt{(a_0 - \alpha)^2 + \beta^2}} \frac{s + a_0}{(s + \alpha)^2 + \beta^2}, \quad (6)$$

когда $\varphi \neq 0$. Величина a_0 связана с фазовым углом φ соотношением

$$\varphi = \arctg \frac{\beta}{a_0 - \alpha}. \quad (7)$$

Соотношение между остротой пика и затуханием можно понять, рассмотрев способ измерения $\theta_o(t)$ при помощи преобразования. Величина $y(s)$ при любом (не обязательно экстремальном) значении $s = j\omega$ характеризует относительную величину колебания частоты ω , присутствующего в $\theta_o(t)$, т. е. показывает степень связи колебательного ядра e^{-st} с $\theta_o(t)$ при частоте ω . Когда величина отношения затухания к частоте велика, последовательные циклы $\theta_o(t)$ будут столь различны по форме, что преобладающая частота не так выделяется и близкие к ней частоты более выражены.

Отношение затухания к частоте (α/β) лучше отражает величину затухания, чем процент критического затухания,

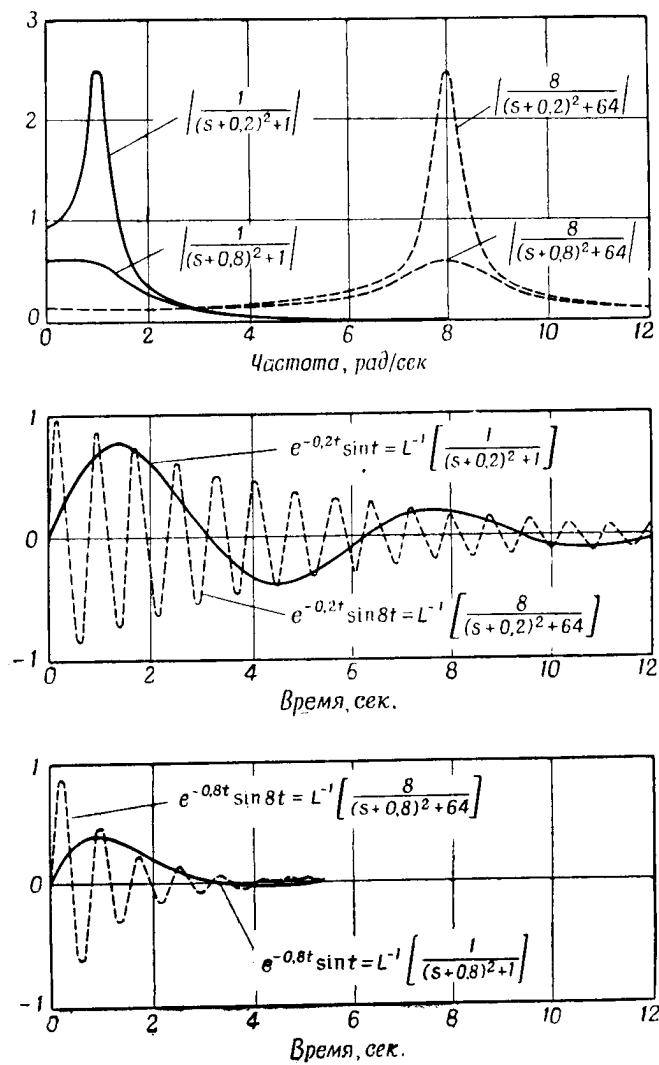
и кривые $y(s)$ удобнее строить в линейном, а не в логарифмическом масштабе по оси частот. Отношение α/β указывает на величину спада колебаний за каждый период и приложимо к колебаниям вообще, тогда как более обычный показатель — процент критического затухания — выражается через экспоненциальные функции.

Одним из наиболее существенных моментов в толковании графиков преобразования является соотношение между высотой кривой $y(s)$ и размахом соответствующего колебания $\theta_o(t)$. Весьма грубо можно считать, что высота кривых в графиках преобразования пропорциональна площадям соответствующих временных функций, поскольку преобразование Лапласа является интегралом; до некоторой степени правильно и обратное заключение, так как обратное преобразование также является интегралом. Приближительным уравнением, устанавливающим соотношение между высотой пика в графике $y(s)$ и размахом колебания $\theta_o(t)$, является уравнение (5). Затухание предполагается довольно незначительным ($\alpha^2 \ll \beta^2$).

Когда $s = j\beta$, уравнение (5) принимает следующую форму:

$$y \text{ пика} = \frac{A}{2\alpha j}. \quad (8)$$

Весьма существенно то, что частота затухающего колебания β не появляется в уравнении (8). Величина $1/\alpha$ является постоянной времени и служит мерой длительности спадающего колебания, а A — начальная амплитуда огибающей. Максимальное значение $y(s)$ поэтому показывает площадь под огибающей $\theta_o(t)$ и не зависит от частоты колебания внутри огибающей. Это иллюстрируется графиками преобразования фиг. 18, которые имеют одинаковую максимальную величину $y(s)$ для двух значений β (8 и 1), когда A , α и φ остаются одинаковыми. Однако высота пика на графике $y(s)$ до некоторой степени зависит от фазы для данной частоты и затухания, как это видно из сравнения графиков фиг. 9—17. Следует подчеркнуть, что максимальное значение $y(s)$ является мерой величины всей временной кривой $\theta_o(t)$ и не характеризует реакцию в какой-то определенный момент, например максимальное перерегулирование или первое изменение знака переходного колебания.



Фиг. 18. Сопоставление функций, для которых преобразование Лапласа дают сходные пики при различных частотах.

Как линейные, так и логарифмические частотные графики полезны, но отражают различные стороны явления. Если, например, два пика при различных частотах имеют одинаковую высоту и ширину при линейном масштабе, соответствующие колебания функций времени будут обладать одной и той же начальной амплитудой и длительностью огибающих, но различными частотами колебаний внутри этих сходных огибающих (см. фиг. 18). Если два пика при различных частотах имеют одинаковую высоту и форму при логарифмическом масштабе частот, то колебания будут иметь одинаковую величину отношения α/β , т. е. будут спадать за одинаковое количество циклов, но первоначальная амплитуда колебания для пика, расположенного на более высокой частоте, будет большей.

Приближение колебания к синусоиде или косинусоиде удобно выразить при помощи фазового угла, который равен нулю, когда колебание представляет собой синусоиду. При анализе передаточных функций вместе с графиком модуля передаточной функции строится и график фазового угла, поскольку первый один не может дать все необходимые сведения о системе.

Поскольку передаточная функция является также преобразованием для импульсной реакции, тангенс фазового угла есть отношение мнимой части преобразования к действительной. При построении графиков преобразования для оценки характеристик системы в переходном процессе понятие угла вносит ненужную ступень в интерпретацию графика. Более целесообразно строить графики действительной и мнимой частей непосредственно, вместе с общей амплитудой, как это сделано на фиг. 9—17. Действительная часть $u(s)$ при любой частоте показывает, в какой степени изменение во времени $\theta_o(t)$ содержит косинусоиду этой частоты, а мнимая часть преобразования указывает, в какой степени $\theta_o(t)$ связано с синусоидой этой частоты.

Апериодические составляющие и второстепенные колебания

При рассмотрении многих практических задач преобразование Лапласа не соответствует преобразованию для стандартной системы второго порядка. Тогда в $\theta_o(t)$, помимо

основного колебания, появятся другие существенные составляющие и характеристику системы нельзя уже будет приближенно представлять одной-единственной затухающей экспонентой. Различие между кривой преобразования для стандартной системы и действительной кривой $y(s)$ для рассматриваемой системы может быть наибольшим при $s = 0$ и при увеличении s может уменьшаться или же проявляться в виде ярко выраженного второстепенного пика.

В любом случае интерпретация кривой должна начинаться с оценки основного колебания и построения соответствующей части кривой преобразования. Это даст нулевую линию, от которой отсчитывается разность $\Delta y(s)$. Дальнейшее рассмотрение $\Delta y(s)$ имеет смысл в том случае, если кривая преобразования значительно отличается от стандартной кривой; в этом случае

1) определяем, является ли разность $\Delta y(s)$ наибольшей при $s = 0$ и уменьшается ли она по мере увеличения частоты, или же она больше при частотах более высоких, чем 0;

2) если $\Delta y(s)$ больше при более высоких частотах, определяем, является ли максимум относительно пологим в широком диапазоне частот или имеется ярко выраженный вторичный пик.

Если разность максимальна при $s = 0$ и уменьшается при больших значениях s , то имеется апериодическая асимметрия, помимо нормальной величины асимметрии, связанной с основным колебанием. Величина $y(s)$ при $s = 0$ является мерой среднего значения $\theta_o(t)$. Каждое колебание имеет определенную асимметрию, которая обычно возрастает по мере усиления затуханий (когда последовательные периоды существенно различаются). Асимметричная составляющая $\Delta\theta_o(t)$, имеющая место, помимо основных колебаний, может быть оценена, если допустить, что она близка к экспоненциальной форме с начальной амплитудой C , т. е.

$$\Delta\theta_o(t) = Ce^{-\gamma t}. \quad (9)$$

Преобразование этого приближения будет

$$\Delta y(s) = \frac{C}{s + \gamma}. \quad (10)$$

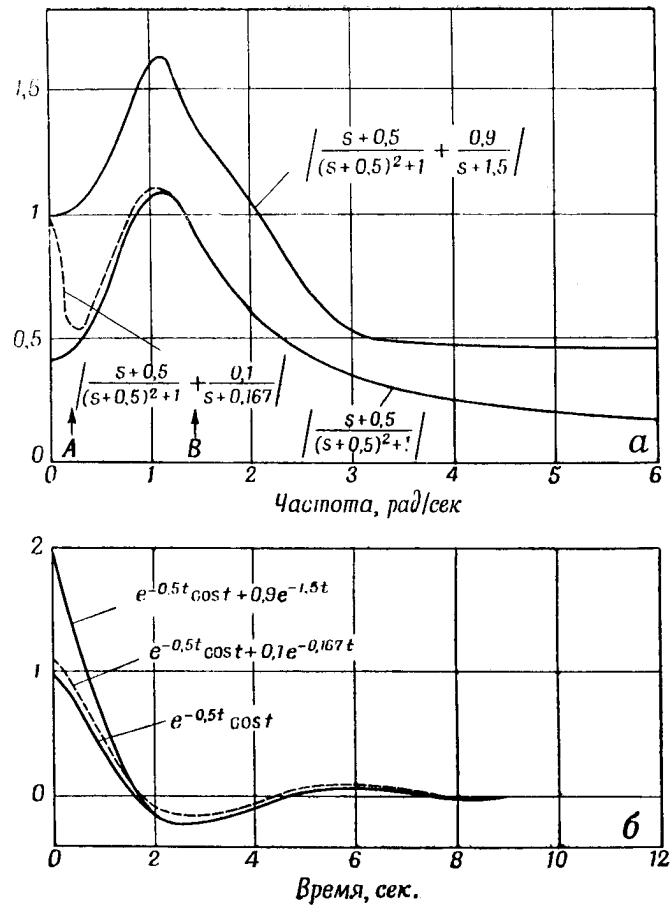
При $s = 0$ имеем $\Delta y(s) = C/\gamma$. Поскольку $1/\gamma$ является мерой продолжительности асимметричной составляющей, на-

чальная амплитуда $\Delta y(s)$ при $s=0$ показывает величину асимметрии $\Delta \theta_0(t)$ таким же образом, как высота пика $y(s)$ указывает на размах колебаний. Приближенные оценки амплитуды и постоянной времени могут быть получены отдельно, если определить значение $s=j\omega$, при котором $y(s)$ равно 0,7 от его значения при $s=0$. Эта величина приблизительно равна γ ; если амплитуда $\Delta y(s)$ прибавляется алгебраически, а не векторно к остатку $y(s)$, то амплитуда $\Delta y(s)$ при $s=j\gamma$ будет $|C/(\gamma + j\gamma)|$ или $(C/\sqrt{2}\gamma)$. После определения γ постоянная C может быть приближенно вычислена по значению $y(s)$ при $s=0$.

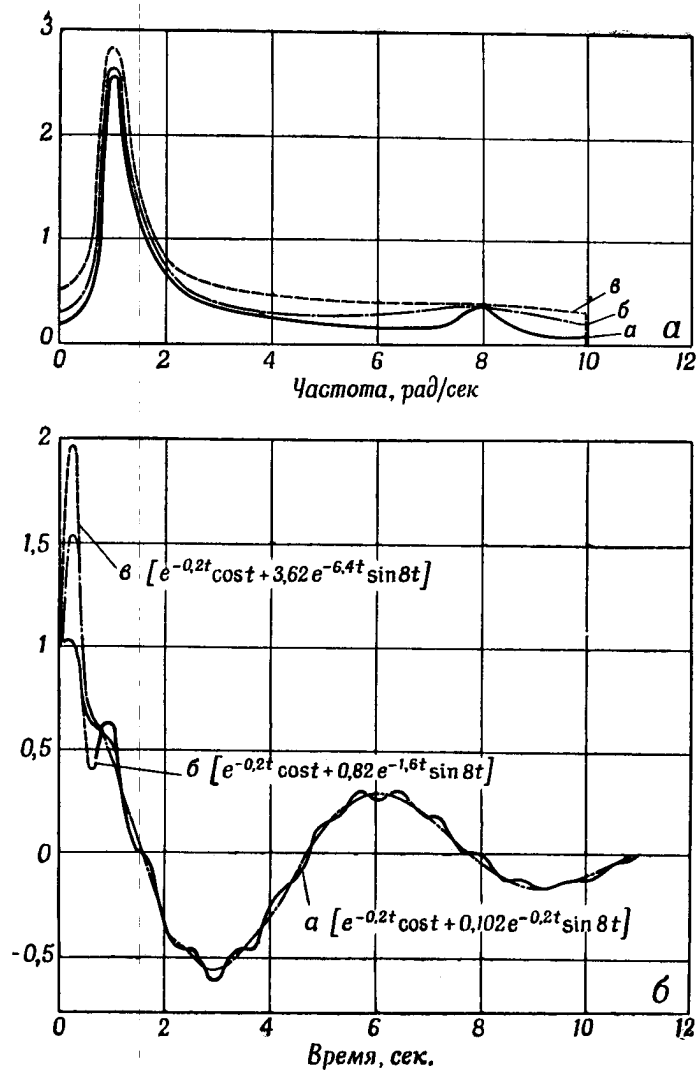
На фиг. 19 показаны две характеристики и их преобразования, имеющие те же основные колебания, но различные добавочные асимметричные составляющие. Постоянная времени одной больше, а другой меньше периода колебания. Величина обеих асимметричных составляющих одинакова; преобразования имеют одинаковые максимальные значения $\Delta y(s)$ при $s=0$, которые сделаны достаточно большими, чтобы проиллюстрировать предыдущее утверждение. Значения ω , при которых $\Delta y(s)$ в каждом случае составляет 0,7 своей величины при $s=0$, показаны в точках A и B для $\Delta \theta_0(t)$ с малой и большой постоянной времени соответственно. Видно, что эти величины ω действительно служат очень хорошим приближением γ .

Если максимальная разность $\Delta y(s)$ имеет место при частоте, отличной от нуля, но является пологой и не имеет ярко выраженного пика, временная характеристика, соответствующая ей, представляет собой импульс, продолжительность которого соответствует полупериоду частоты возле центра максимума $\Delta y(s)$. Размах колебания или произведение максимальных значений на продолжительность импульса дает нам высоту максимума $\Delta y(s)$ (фиг. 20—25).

Когда график преобразования содержит отчетливый вторичный пик, это соответствует отчетливому второстепенному колебанию. Проблема интерпретации состоит не в том, чтобы оценить параметры колебания с точностью в несколько процентов. Скорее она состоит в том, чтобы вообще установить наличие колебания и определить, на какой частоте оно наблюдается и каковы его амплитуда и продолжительность. Этим сведениям вполне достаточно, чтобы определить, например, какое количество энергии системы регулирования

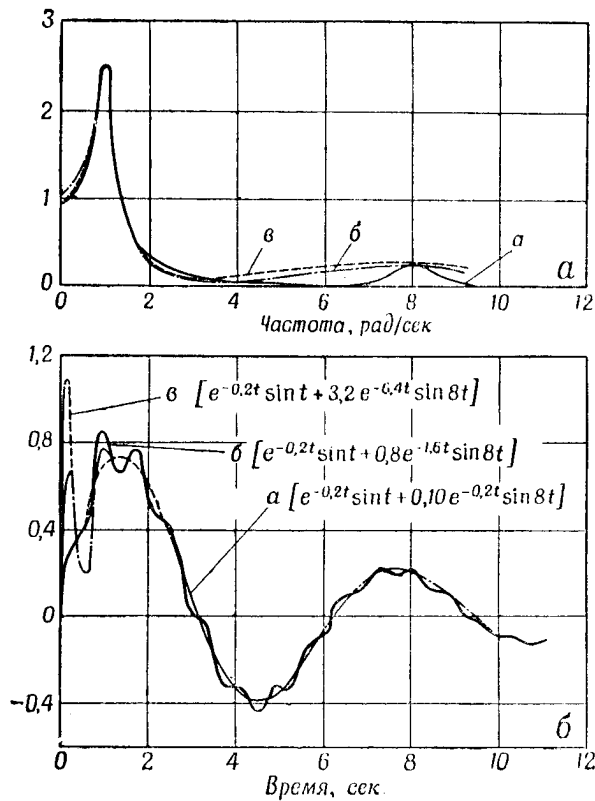


Фиг. 19. Примеры, иллюстрирующие результат добавления асимметрии к затухающей косинусоиде.



Ф и г. 20. Графики преобразования Лапласа (модули) и временные кривые $e^{-0.2t} \cos t$ с вторичными колебаниями, имеющими различное затухание.

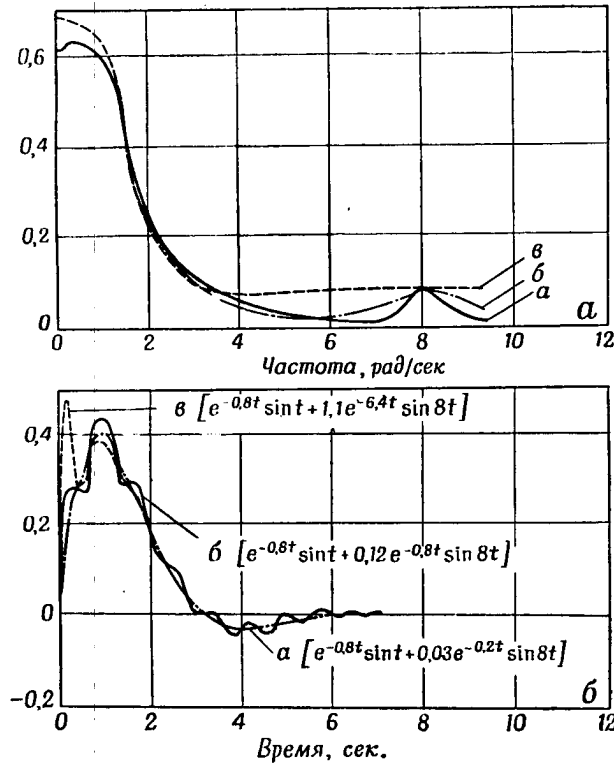
расходуется при колебаниях. Это очень важный вопрос, например, для некоторых пневматических устройств, где запас энергии системы регулирования ограничен.



Фиг. 21. Графики преобразования Лапласа (модули) и временные кривые $e^{-0,2t} \sin t$ с вторичными колебаниями, имеющими различное затухание.

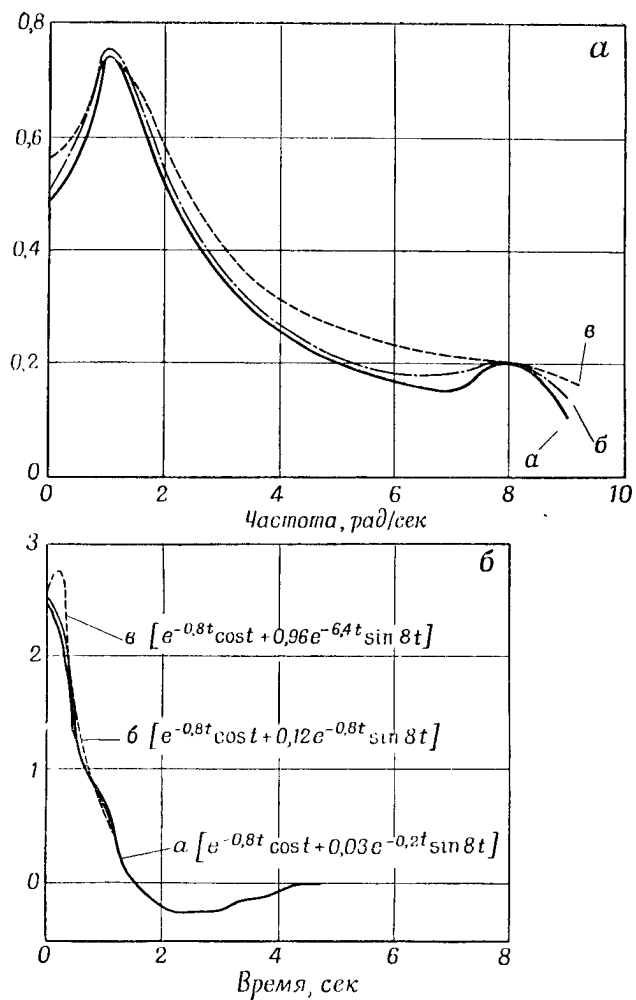
Принципы интерпретации вторичного пика обычно такие же, как и при больших пиках, которые соответствуют основному колебанию. Амплитуда и продолжительность, однако, легче определяются и более четко выражаются в долях амплитуды и продолжительности главных колебаний, а не

в абсолютных величинах. Следовательно, процесс определения величины превышения $\Delta y(s)$ над нулевой линией сводится к арифметическим действиям, вместо применения алгебры комплексных величин. Это не очень точный, но обычно вполне удовлетворительный метод.

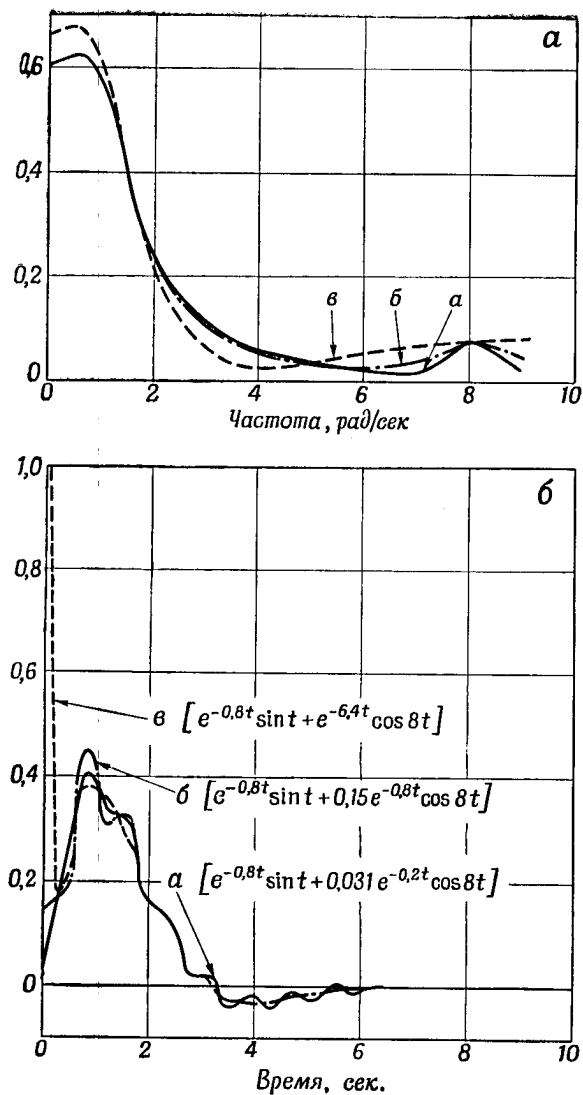


Фиг. 22. Графики преобразования Лапласа (модули) и временные кривые $e^{-0.8t} \sin t$ с вторичными колебаниями, имеющими различное затухание.

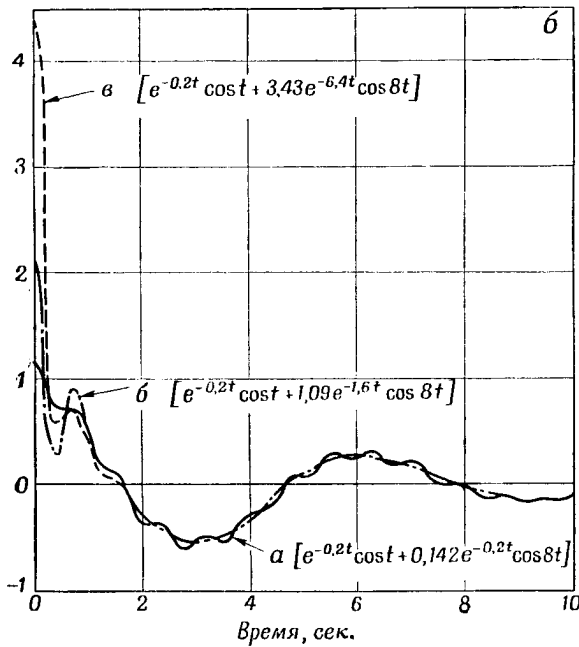
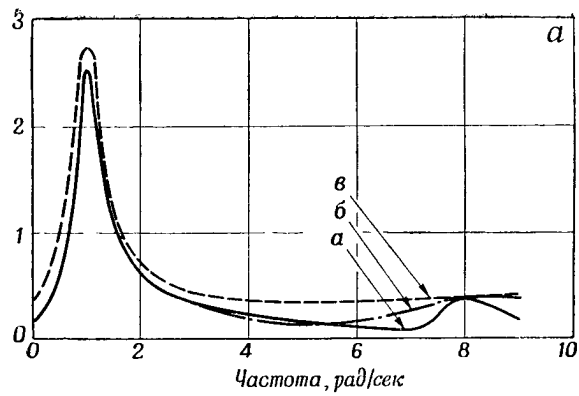
Если вторичный пик отчетливо выражен, то обычно второстепенные колебания имеют несколько периодов, так что частота, размах и затухание могут быть определены в отдельности. Фаза обычно не представляет интереса.



Фиг. 23. Графики преобразования Лапласа (модули) и временные кривые $e^{-0,8t} \cos t$ с вторичными колебаниями, имеющими различное затухание.



Ф и г. 24. Графики преобразования Лапласа (модули) и временные кривые $e^{-0,8t} \sin t$ с вторичными колебаниями; фазы вторичных колебаний отличны от фаз на фиг. 20—23.



Ф и г. 25. Графики преобразования Лапласа (модули) и временные кривые $e^{-0.2t} \cos t$ со вторичными колебаниями; фазы вторичных колебаний отличны от фаз на фиг. 20—23.

На фиг. 20—25 приведены графики преобразования и временные характеристики, в которых, кроме основных, присутствуют также небольшие второстепенные колебания. Амплитуда вторичных пиков составляет 10% основных пиков, а частота их на три октавы выше. Обычно невозможно интерпретировать вторичные пики, которые имеют меньшую относительную амплитуду или которые расположены ближе к основному пику, чем указано выше.

На графиках представлены четыре типа основных колебаний — синусоида и косинусоида с относительно большим затуханием ($\alpha = 0,8$) и малым затуханием ($\alpha = 0,2$). К каждому из этих колебаний на фиг. 20—23 прибавляется три второстепенных колебания, которые имеют различные величины затухания. Три вторичных пика на одинаковую величину превышают нулевую линию кривой преобразования.

Как и следовало ожидать, площади под огибающими второстепенных колебаний составляют около 10% от площадей основных колебаний. Наиболее пологие вторичные пики на графиках передаточной функции, не имеющие ярко выраженных максимумов, дают в соответствующей временной характеристике отдельные полупериоды или импульсы. По мере того как вторичные пики в кривых преобразования становятся более заметными, во временных характеристиках появляется большее число периодов; амплитуда огибающей уменьшается и длительность процесса увеличивается.

Все второстепенные колебания, приведенные на фиг. 20—23, являются синусоидальными. Для того чтобы показать, что фаза второстепенного колебания не влияет сильно на кривую преобразования, на фиг. 24 и 25 к тем же самым основным колебаниям, что и на фиг. 20—22, добавляются вторичные колебания в виде косинусоид.

Реакции на различные входные сигналы

При интерпретации кривых преобразования Лапласа, пожалуй, наиболее трудно установить тот факт, что кривая передаточной функции соответствует временной реакции на единичный импульс, поскольку при проведении экспериментов и при расчетах имеется тенденция связывать частотную характеристику с временной реакцией на единичный

ступенчатый входной сигнал. Это не наносит ущерба зависимости передаточной функции от ступенчатого входа, однако более удобной основой для сравнения и оценки реакций систем на различные действительные входные сигналы является импульсная реакция системы. Поэтому все кривые преобразований, описывавшиеся до сих пор в данной статье, являлись преобразованиями импульсных реакций. Когда исследуемая система принадлежит к большому классу систем с линейными сосредоточенными элементами, то преобразование от реакции на любой входной сигнал находится просто умножением преобразования от импульсной реакции на преобразование данного входного сигнала. Кривую произведения можно проанализировать методами, приведенными в этой статье, и таким путем оценить реакцию на интересующий нас входной сигнал.

Влияние входного сигнала или вынуждающей силы на линейную систему проявляется, во-первых, в возникновении установившегося или особого интегрального члена в реакции и, во-вторых, в перераспределении величин различных быстро-затухающих членов или дополнительных функций. Оба эти явления интересны при анализе кривой преобразования.

Большинство встречающихся на практике входных сигналов попадает в одну из следующих четырех групп:

- 1) ступенчатый скачок смещения, скорости, ускорения и т. д.;
- 2) экспоненциальные импульсы e^{-at} , где a — постоянная;
- 3) колебания, которые можно аппроксимировать синусоидами с затуханием или без него;
- 4) импульсы различной формы (прямоугольные, остроконечные и т. д.), продолжительность которых сравнима с постоянными времени системы.

Преобразования от первой группы сигналов имеют форму $[1/s^n]$, где n равно единице для ступенчатого скачка смещения, двум — для скачка скорости и так далее. Эти преобразования обращаются в бесконечность при $s = 0$, но сильно уменьшают любые высокие собственные частоты в импульсной реакции системы.

Преобразования второй группы имеют вид $[1/(s + a)]$. Такие функции иногда полезны при аппроксимации импульсов или постепенных подъемов (в комбинации со ступенькой).

Третья группа входных сигналов имеет преобразования вида $[1/(s^2 + a^2)]$, где a — частота колебания. Они обычно добавляют колебания с частотой a к импульсной реакции системы и ослабляют уже присутствующие колебания.

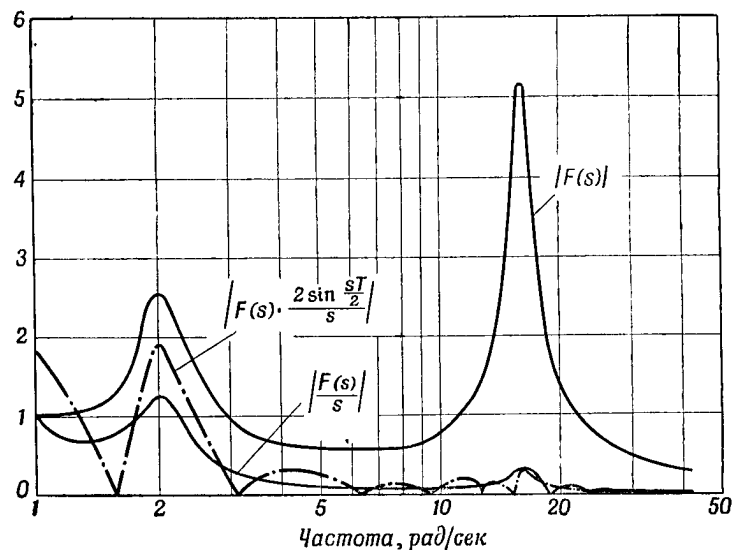
Преобразования импульсов заметной продолжительности обычно содержат трансцендентные члены вида e^{-sT} , где T — постоянная времени, связанная с продолжительностью импульса. В них входят также алгебраические члены, зависящие от формы импульса.

Кривая преобразования реакции системы на действительный входной сигнал обычно интерпретируется следующим образом: сперва исследуется установившаяся реакция, затем строится график ее преобразования и оцениваются размеры и вид пиков в общем преобразовании. Иногда установившийся член известен из граничных значений или из других условий. В любом случае, если преобразование от реакции равно бесконечности при $s = 0$, то имеются постоянные смещение, скорость или ускорение, которые остаются неопределенными. Если преобразование обращается в бесконечность при некоторой частоте, отличной от нуля, то в реакции имеются установившиеся колебания этой частоты. Если преобразование имеет трансцендентный член с периодически чередующимися при увеличении частоты минимумами и максимумами, то временная реакция обычно содержит импульс или серию импульсов, продолжительность которых просто связана с периодом изменения преобразования.

В качестве примера на фиг. 26 показаны преобразования различных реакций систем, описываемых дифференциальным уравнением 5-го порядка; соответствующие им временные реакции приведены на фиг. 27. Эти графики даны для иллюстрации различия между реакцией на единичный импульс и на единичное ступенчатое смещение. Реакция на импульс дает высокочастотные колебания, амплитуда которых значительно выше, чем у низкочастотных колебаний. В реакции на ступенчатое смещение амплитуда высокочастотных колебаний уменьшается до значений, значительно меньших, чем у низкочастотных.

Реакция системы на прямоугольный импульс и ее преобразование также показаны на графиках, чтобы дать краткий пример преобразования с трансцендентными функциями.

Здесь применимы те же общие правила интерпретации, как и раньше. Большие пики в преобразовании Лапласа¹⁾ указывают на наличие колебаний значительной величины во временной реакции, причем размах колебаний получается как произведение амплитуды на время, просуммированное

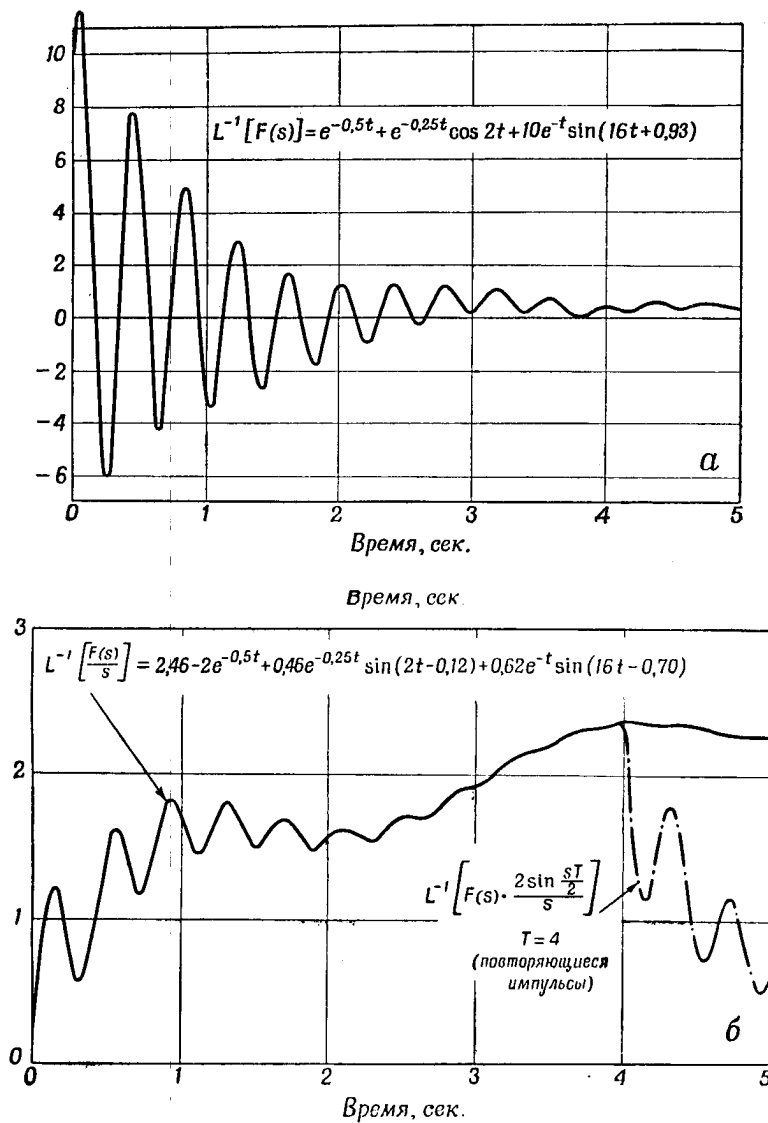


Фиг. 26. Графики преобразования Лапласа для различных входных сигналов в системе 5-го порядка.

по всему промежутку времени. Крутизна кривой характеризует быстроту затухания. Маленькие периодические пики на кривой преобразования Лапласа характеризуют наличие импульсов в реакции, а расстояние между их максимумами и минимумами на кривой преобразования дает постоянную времени или продолжительность импульса. Превосходное описание частотных компонент импульса дано Гольдманом [10].

Интерпретация кривых преобразования Лапласа указанным методом может и не дать точной формы каждого колебания в области около $t = 0$, но дает очень полезные сведения о временной реакции.

¹⁾ Имеются в виду пики на обобщенной амплитудной частотной характеристике. — Прим. ред.



Ф и г. 27. Временная кривая системы 5-го порядка, соответствующая фиг. 26.

а — импульсный вход; б — ступенчатый и повторяющийся прямоугольный импульсный входы.

Обобщение метода на системы с переменными коэффициентами

Для систем, отличных от систем с постоянными параметрами, т. е. для систем, у которых для получения преобразования от выходного сигнала необходимо решить уравнение, отличное от алгебраического, перемножение преобразований обычно незаконно. Для таких систем преобразование от реакции на желаемый входной сигнал необходимо искать, решая уравнение относительно $y(s)$ после подстановки преобразования входного сигнала в уравнение. Однако частоту, амплитуду и затухание колеблющихся членов можно, как и раньше, оценивать и сравнивать, исходя из кривых преобразования от выходного сигнала до тех пор, пока преобразование правомерно получается в виде интегрального уравнения (1).

К системам, для которых применение понятия передаточной функции становится туманным, но на которые можно распространить применение преобразования Лапласа, относятся системы, описываемые дифференциальными уравнениями с коэффициентами, линейно изменяющимися со временем. Такие системы имеют большое практическое значение; поэтому мы используем одну из них для того, чтобы проиллюстрировать, как из преобразований, которые не являются обычными отношениями двух рациональных многочленов от s , можно оценить относительную устойчивость.

Рассмотрим систему, в которой вход и выход связаны следующим соотношением:

$$(a_2 + b_2 t) \ddot{\theta}_o + (a_1 + b_1 t) \dot{\theta}_o + (a_0 + b_0 t) \theta_o = \theta_i. \quad (11)$$

В общем случае уравнение n -го порядка, в котором коэффициенты изменяются как r -я степень от t , можно преобразовать в дифференциальное уравнение порядка r относительно $y(s)$, в котором в коэффициентах уравнения относительно $y(s)$ появляется n -я степень от s . Каждый член уравнения преобразуется на основе положения, что умножение на t во временной области соответствует дифференцированию по s . Практически дифференциальное уравнение относительно s может быть решено только тогда, когда $r = 1$, так что применение преобразования Лапласа ограничивается уравнениями, подобными уравнению (11), в которых коэффициенты изменяются как $(a + bt)$.

Используя обычные методы [9, 11, 12], преобразуем уравнение (11) почленно в следующее уравнение:

$$\begin{aligned} & -a_2 s^2 y(s) - a_2 s \dot{\theta}_o(0) - a_2 \ddot{\theta}_o(0) - 2b_2 s y(s) - \\ & - b_2 s^2 y'(s) + b_2 \dot{\theta}_o(0) - a_1 \dot{\theta}_o(0) + a_1 s y(s) - b_1 y(s) - \\ & - b_1 s y'(s) + a_0 y(s) - b_0 y'(s) = L\{\theta_i(t)\} \end{aligned} \quad (12)$$

или

$$\begin{aligned} & -y'(s)[b_2 s^2 + b_1 s + b_0] + y(s)[a_2 s^2 + a_1 s + a_0 - 2b_2 s - b_1] = \\ & = L\{\theta_i(t)\} + \dot{\theta}_o(0)[a_2 s + a_1 - b_2] + \ddot{\theta}_o(0) a_2. \end{aligned} \quad (13)$$

Решение для $y(s)$ через преобразование Лапласа от $\theta_o(t)$ может быть представлено в обычном виде

$$y(s) = e^{+ \int \frac{Q(s)}{P(s)} ds} \left[C - \int e^{- \int \frac{Q(s)}{P(s)} ds} \frac{R(s)}{P(s)} ds \right], \quad (14)$$

где

$$\begin{aligned} P(s) &= (b_2 s^2 + b_1 s + b_0), \\ Q(s) &= (a_2 s + a_1 s^2 + a_0) - \frac{d[P(s)]}{ds}, \\ R(s) &= L\{\theta_i(t)\} + \dot{\theta}_o(0)[a_2 s + a_1 - b_2] + \ddot{\theta}_o(0) a_2, \end{aligned} \quad (15)$$

а C — произвольная постоянная.

Мы видим, что входной сигнал θ_i входит в выражение для $y(s)$ через интеграл частного решения и что преобразование Лапласа от θ_i не умножается на передаточную функцию системы, а появляется в $R(s)$ под знаком интеграла вместе с граничными условиями. Свободное решение такого уравнения, как (13), относительно $y(s)$ не обязательно является преобразованием от свободного решения уравнения (11) для $\theta_o(t)$. Свободное решение в уравнении (14) ведет себя так же, как и свободное решение любого другого дифференциального уравнения; оно применяется, когда необходимо удовлетворить граничным условиям, наложенным на y . Эти граничные условия часто получаются из граничных условий, налагаемых на $\theta_o(t)$ с помощью теоремы о начальном и конечном значениях.

Так как дифференциальное уравнение n -го порядка с изменяющимися во времени коэффициентами, подобно уравнению (11), можно преобразовать в дифференциальное уравнение первого порядка относительно $y(s)$, то, естественно,

возникает вопрос, как из единственного уравнения первого порядка относительно $y(s)$ получить преобразования от n составляющих свободного решения для θ_o . Процесс получения аналогичен процессу, когда уравнение (11) имеет постоянные коэффициенты, а уравнение (13) — алгебраическое, а не дифференциальное относительно $y(s)$. Граничные условия, наложенные на $\theta_o(t)$, видоизменяются и обычно входят в уравнение для $y(s)$ в виде n отдельных вынуждающих сил, каждая из которых дает частное решение для $y(s)$ в виде интеграла. Каждое решение является преобразованием от различных комбинаций $\theta_o(t)$. Когда некоторые граничные условия выпадают в процессе преобразований, как это имеет место в приводимом ниже примере, необходимо использовать свободное решение для y .

В виде примера рассмотрим уравнение (14) в случае, когда

$$a_2 = 0; \quad b_2 = 1; \quad a_1 = 1; \quad b_1 = 0,6; \quad a_0 = 0,3; \quad b_0 = 9,09.$$

Входной сигнал возьмем в виде $\theta_i(t) = a_0 + b_0(t)$ без начального накопления энергии. При этих условиях установившаяся часть выходного сигнала равна постоянной или единице. Если $Q(s)/P(s)$ разложить на множители и подставить в интеграл уравнения (14), то получим выражение для $y(s)$:

$$y(s) = \frac{1}{[s^2 + 0,6s + 9,09]^{1/2}} \left[C - \int \frac{9,09 + 0,3s}{(s^2 + 0,6s + 9,09)^{1/2}} \frac{10}{s^2} ds \right]. \quad (16)$$

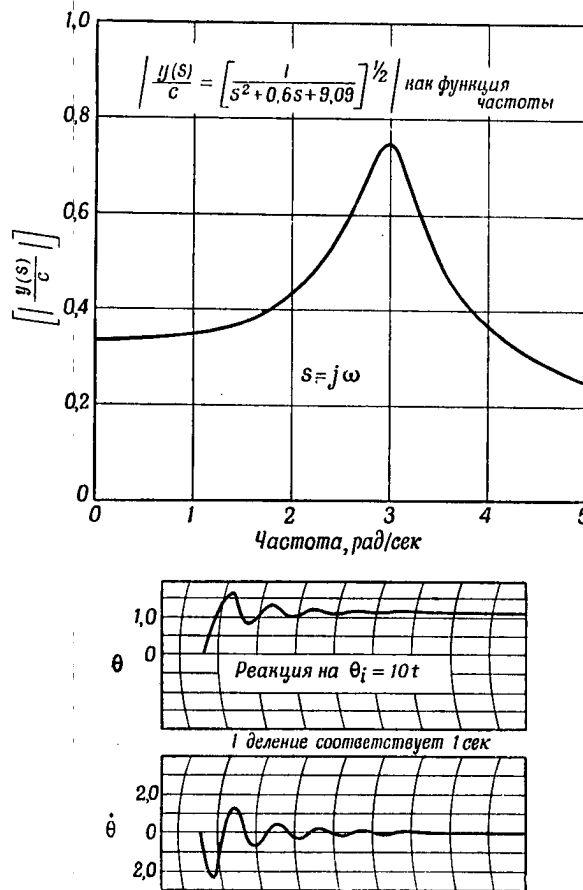
Таким образом,

$$y(s) = \frac{C}{[s^2 + 0,6s + 9,09]^{1/2}} + \frac{1}{s}. \quad (17)$$

На фиг. 28 показана кривая первого члена $y(s)$ уравнения (17). Наличие пика указывает на то, что временная реакция $\theta_o(t)$ содержит колебание, имеющее отношение затухание/частота (α/β), приблизительно равное 0,5, и частоту колебаний около 3 рад/сек. На фиг. 28 приведена также временная реакция, полученная с помощью электронной модели для той же системы при использовании уравнения (11). Форма огибающей и сами колебания временной реакции согла-

2. Относительная устойчивость замкнутых систем 109

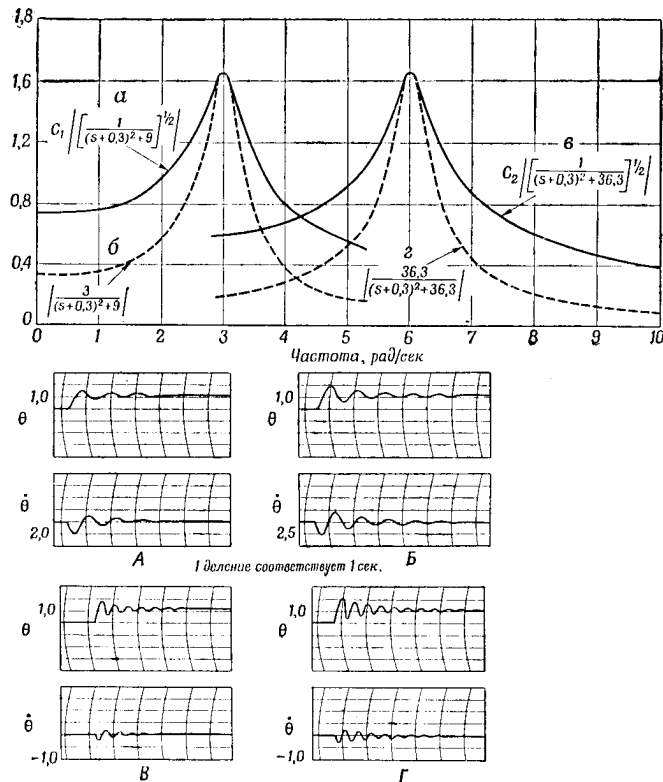
суются с оценкой, данной на основе кривой преобразования. Для удобства моделирования при получении временной



Фиг. 28. Графики преобразования и временная кривая, соответствующая уравнению (11) при $a_2 = 0$; $b_2 = 1$; $a_1 = 1$; $b_1 = 0,6$; $a_0 = 0,3$; $b_0 = 9,09$ и $\theta_i(t) = 10t$.

реакции, приведенной на фиг. 28, взят входной сигнал в виде простой наклонной функции, а не в виде $(a_0 + b_0 t)$. Такая замена не оказывает влияния на колебания.

Чтобы показать, как правила интерпретации кривых преобразования для систем с постоянными коэффициентами можно применить к системам, коэффициенты которых меняются



Фиг. 29. Сопоставление решений дифференциальных уравнений с графиками преобразования, имеющими одинаковые пики на различных частотах.

Наверху — графики преобразования Лапласа для решений дифференциальных уравнений $\alpha - \delta$.
Внизу — временные кривые.

A: $t\ddot{\theta} + (1+0,6t)\dot{\theta} + (0,3+9,09t)\theta = 10t$; B: $\ddot{\theta} + 0,6\dot{\theta} + 9,09\theta = 10$;
B: $t\ddot{\theta} + (1+0,6t)\dot{\theta} + (0,3+36,39t)\theta = 40t$; Г: $\ddot{\theta} + 0,6\dot{\theta} + 36,39\theta = 40$.

со временем, умножим значение b_0 в выражении предыдущего примера на 4. Кривая для $y(s)$ на основе уравнения (17) при увеличенном значении b_0 построена на фиг. 29. Форма

пика не изменяется, но пик сдвигается на частоту, равную удвоенному начальному значению. Из сравнения кривых преобразований можно заключить, что умножение b_0 на 4 не изменяет формы огибающей временной реакции, но удваивает частоту в пределах временной реакции. Это подтверждается кривыми временной реакции на фиг. 29, которые совершенно аналогичны кривым фиг. 18.

При изучении систем с переменными коэффициентами часто применяется метод фиксации коэффициентов. В данных примерах, если t становится очень большим, может оказаться вполне оправданным деление уравнения на t . Вместо уравнения (17) получаем тогда аппроксимацию преобразования в виде

$$y(s) = \frac{C}{s^2 + 0,6s + 9,09}. \quad (18)$$

Это преобразование, а также аппроксимация для случая, когда b_0 умножается на 4, построены на фиг. 29 для сравнения с точными преобразованиями. Сделанные по этим кривым оценки временной реакции показывают, что приближенная временная реакция затухает менее быстро, чем точная реакция, т. е. пик кривой, полученной на основе уравнения (18), более крутой. Временные реакции, полученные на модели для аппроксимации постоянными коэффициентами, подтвердили правильность такой оценки. Это показывает недостатки подобных аппроксимаций; кривая точного преобразования Лапласа от функции времени дает желаемую информацию о временной реакции.

Численные значения в примере уравнения (13) и фиг. 28 и 29 выбраны таким образом, чтобы получить вполне простое решение для $y(s)$ в замкнутом виде. В более сложных случаях потребуется система решений или же при решении дифференциального уравнения относительно $y(s)$ можно воспользоваться особым приемом, например предположить, что входной сигнал прикладывается в момент времени, отличный от $t=0$. Кроме того, метод получения преобразований от реакций систем, параметры которых не сосредоточены и не постоянны, по дифференциальным уравнениям не является единственным. Здесь кратко проиллюстрирован метод нахождения $y(s)$ по дифференциальному уравнению, так как он нигде не рассматривается подробно. Однако во многих превосходных книгах даются другие методы получения преобразований

Лапласа. До тех пор пока преобразование равно интегралу уравнения (1) и до тех пор пока он сходится, о свойствах реакции системы можно судить по кривой преобразования.

Иррациональные члены, появляющиеся, например, в уравнении (17), являются типичными элементами преобразований от реакций систем, имеющих меняющиеся со временем компоненты. Особый важный случай составляют преобразования от таких функций, как бесселевы, не имеющие демпфирующих членов под радикалом. Если построить преобразование $(1/\sqrt{s^2+1})$ от бесселевой функции нулевого порядка для чисто мнимых значений s , то кривая уходит в бесконечность при $s = 1$, что, очевидно, указывает на наличие незатухающих колебаний и на неустойчивость системы. Известно, что функция Бесселя $J_0(t)$ устойчива, конечна в начале координат и затухает приблизительно как синусоида с множителем $(1/\sqrt{t})$ после нескольких периодов. Однако огибающая бесселевой функции затухает так медленно, что интеграл по времени от нуля до бесконечности равен бесконечности, если не включить множитель сходимости (т. е. действительную часть) в экспоненту ядра преобразования. Это становится понятным, если вспомнить, что интеграл от нуля до бесконечности от $[1/(t+a)]$ равен бесконечности, в то время как интеграл от e^{-bt} конечен независимо от значений a и b (исключая нулевое значение). (Величина a включается в функцию $1/t$, чтобы он оставалась конечной при $t = 0$.)

Для интерпретации такого преобразования, как преобразование от функции Бесселя, можно построить преобразование, приняв, что s является комплексной величиной, имеющей, кроме изменяющейся мнимой части, малую фиксированную действительную часть σ . Кривая будет иметь максимальное значение при частоте колебаний. Для оценки размеров колебаний можно найти значение преобразования от незатухающей синусоиды той же частоты, например $[1/(s^2+1)]$, используя те же действительную и мнимую части, и сравнить преобразование синусоиды с исследуемым преобразованием. Абсолютная устойчивость всегда может быть увеличена другими путями [4,5]; применение малой действительной части в s справедливо для систем, о которых известно, что они находятся недалеко от границы области устойчивости.

Заключение

Преобразование Лапласа прочно заняло свое место как мощный способ решения трудоемких задач, но за ним всегда должен был следовать заключительный этап — обратное преобразование. Теперь, однако, само прямое преобразование можно рассматривать как заключительный этап. Оно образуется путем перемножения и суммирования, которые имеют физические аналоги и могут быть легко сделаны наглядными. Ядро преобразования можно считать стандартной или пробной функцией, частоту которой можно менять по своему усмотрению для анализа неизвестной реакции и выявления наличия в ней такой же частоты. Для оценки главных характеристик временной реакции требуется иметь лишь кривую преобразования. Это имеет важное практическое значение, так как обычно легче построить график преобразования, чем получить обратное преобразование, а затем строить его график.

Кроме того, область применения графиков преобразований можно значительно расширить, распространив ее на системы, для которых преобразование реакции системы может быть получено законным образом, но для которых обратное преобразование, являющееся функцией времени, получить не так легко.

В исследовательской работе, когда приходится использовать множество пробных решений для получения возможно более близкого приближения к желаемому результату, часто прибегают к помощи вычислительных устройств или моделей. Главная задача анализа с помощью простых графиков состоит тогда в том, чтобы задать направление для работы моделирующего устройства путем выделения областей, которые могут представить интерес, и проверить пригодность результатов, полученных на модели. Использование графиков преобразований Лапласа приводит к более широкому применению относительно дешевых математических моделей за счет сокращения количества более дорогостоящих физических моделирующих устройств.

Необходимо всегда помнить, что частота в графике частотных характеристик не является обязательно частотой установившегося синусоидального входного сигнала; это —

частота колебания ядра преобразования Лапласа, которая используется для анализа преобразуемой функции $\theta_o(t)$. Только для систем с постоянными коэффициентами или для систем, близких к ним, эта частота соответствует частоте установившегося испытательного входного сигнала. Переходная реакция системы является основной характеристикой системы, которая может использоваться в практических задачах и по важности равнозначна передаточной функции. Интерпретация кривой преобразования Лапласа не требует использования таких установившихся понятий, как запас устойчивости по фазе и амплитуде, для того чтобы определить реакцию системы на интересующий нас практически входной сигнал. Независимо от того, каким методом получена кривая, до тех пор пока величина площади под кривой равна интегралу уравнения (1), методы, описанные выше, можно применять для оценки временной реакции.

Это расширяет возможности практического использования метода, например, при испытаниях сложных электромеханических систем. Переходные процессы обычно легче получить, чем частотные реакции, особенно, когда система содержит много механических или аэродинамических элементов. Когда известен переходный процесс, то преобразование Лапласа как функцию от частоты можно определить методом, обратным методу данной статьи, а отсюда уже можно определить реакции на другие интересующие нас входные сигналы. Необходимо помнить, что если при экспериментальном исследовании использовался, например, ступенчатый входной сигнал, то получаемое при этом преобразование не является частотной реакцией в обычном смысле, если его не привести в соответствие с ней путем умножения на s .

Когда в переходном процессе имеется несколько периодов колебаний, то кривая преобразования совершенно ясно указывает на это, а также и на различия отдельных свойств реакции. Однако большинство сервомеханизмов на практике рассчитывается так, что основное колебание затухает только после двух или трех всплесков. Может показаться, что трудно определить, например, по относительно размытому пику кривой преобразования, представляет ли реакция затухающее колебание или сумму нескольких экспоненциальных

функций. Это не является принципиальным ограничением точности, а скорее указывает на то, что временную реакцию можно адекватно описать более чем одним способом. Метод аппроксимации реакций хорошо демпфированных сервомеханизмов использует то, что колебание, состоящее только из нескольких всплесков, можно представить двумя или тремя различными положительными и отрицательными экспоненциальными членами почти так же хорошо, как и затухающей синусоидой типа $e^{-at} \cos(\omega t + \varphi)$. Это распространяется также и на кривые преобразования Лапласа. Принципиально важным является то, что данная форма пика кривой преобразования связывается с определенной формой переходного процесса, причем не рассматривается формальное математическое выражение, описывающее переходный процесс.

Желательно отметить еще одну область возможного обобщения частотных графиков, хотя это и не соответствует строго методу, изложенному в этой статье. Найквист, а за ним и многие другие авторы при рассмотрении графиков кривых использовали комплексную частоту, т. е. как действительную часть σ , так и мнимую часть ω оператора s .

Возможны различные подобные обобщения обычных частотных кривых и, по-видимому, они в отдельных случаях обладают рядом преимуществ. Например, степень затухания отдельного колебания можно определить совершенно точно, приравняв мнимую часть s постоянной, равной частоте колебаний, и построив выражение для преобразования в виде функции от действительной части σ . Использовать положительные значения σ оказывается непродуктивно. При отрицательных значениях пик кривой наблюдается при значении σ , соответствующем значению затухания колебания. Эта операция сводится к исследованию расположения полюсов передаточной функции системы в левой полуплоскости комплексной частоты. Построенная при этом величина не соответствует полному интегралу, данному в уравнении (1), а определяет скорее только нижний предел отрицательных значений σ больших, чем вещественная часть любого полюса передаточной функции. Этот нижний предел, или передаточная функция системы, связан с временной характеристикой через интеграл обратного преобразования, а не через интеграл уравнения (1).

ЛИТЕРАТУРА

1. Brown G. S., Campbell D., Principles of Servomechanisms, New York, 1948.
2. Chestnut H., Mayer R. W., Servomechanisms and Regulating Systems Design, Vol. I, New York, 1951.
3. Ahrendt W. R., Taplin J. F., Automatic Feedback Control, New York, 1951.
4. Kirby M. J., *Trans. AIEE*, Part II, **69**, 1662 (1950).
5. Kirby M. J., Giulianelli R. M., *Trans. AIEE*, Part II, **70**, 1447 (1951).
6. Zadeh L. A., *Proc. IRE*, № 3, 291 (1950).
7. Moore J. R., *Proc. IRE*, № 3 (1951).
8. Бодэ Г., Теория цепей и проектирование усилителей с обратной связью, Издательство иностранной литературы, 1948.
9. Гарднер М., Бэрнс Дж., Переходные процессы в линейных системах, Гостехиздат, 1951.
10. Гольдман С., Гармонический анализ, модуляция и шумы, Издательство иностранной литературы, 1951.
11. Churchill R. V., Modern Operational Calculus, New York, 1941.
12. Goldman S., Transformation Calculus and Electrical Transients, New York, 1949.
13. McLachlan N. W., Modern Operational Calculus, London, 1948.
14. Harris H., Jr., Kirby M. J., Arx E. F., *Trans. AIEE*, Part II, **70**, 1452 (1951).

3. СВЯЗЬ МЕЖДУ ЧАСТОТНЫМИ ХАРАКТЕРИСТИКАМИ И ПЕРЕХОДНЫМ ПРОЦЕССОМ СИСТЕМЫ АВТОМАТИЧЕСКОГО РЕГУЛИРОВАНИЯ С ОБРАТНОЙ СВЯЗЬЮ¹⁾

И. Чу

Введение

Характеристики системы автоматического регулирования с обратной связью часто определяются по переходному процессу, возникающему при приложении ступенчатой функции на входе. Внезапное изменение входного сигнала или возмущения может соответствовать важным действительным условиям работы системы. Однако, когда система имеет характеристическое уравнение высокого порядка, синтез или расчет системы при таком воздействии математически трудно выполним. В связи с этим был развит и широко применяется частотный метод. Очевидно, очень желательно знать связь между частотной характеристикой и переходным процессом.

Математически такую связь можно установить при помощи преобразования Фурье. Но этот метод не раскрывает физической сущности явлений и поэтому приносит мало пользы при расчете систем автоматического регулирования.

Количественную связь между частотной характеристикой и переходным процессом обычно трудно сформулировать в общем виде²⁾.

Благодаря использованию метода корневого годографа, развитого Эвансом [1], стало возможным устанавливать характер такой связи, когда известна связь между амплитудно-фазовой характеристикой и корневым годографом. Точный характер этой связи будет зависеть от конкретного

¹⁾ С h u Y., *Trans. AIEE*, Part II, 72, 81—92 (1953).

²⁾ Здесь можно указать на метод анализа свойств переходного процесса, разработанный В. В. Солодовниковым, который основан на изучении свойств вещественной частотной характеристики и связи между ней и переходным процессом (см., например, „Основы автоматического регулирования“ под ред. В. В. Солодовникова, Машгиз, 1954).—*Прим. ред.*

вида системы. Поэтому, хотя в настоящей статье и не дается общего количественного соотношения, эта связь будет исследована для каждой системы в отдельности.

В настоящей статье мы ставим себе следующие задачи:

1. Показать, что частотные характеристики и переходные процессы связаны между собой конформным отображением благодаря соответствию между амплитудно-фазовой характеристикой и корневым годографом.

2. Указать в качестве примера [2—5,11] несколько известных методов определения переходных процессов по частотным характеристикам при помощи конформного отображения.

3. Выявить свойства переходного процесса на основе существующего M_p -критерия¹⁾ при синтезе систем частотным методом, исследуя расположение нулей и полюсов передаточной функции замкнутой системы, которая связана с передаточной функцией разомкнутой системы посредством корневого годографа.

4. Сравнить частотные характеристики и корневые годографы нескольких условно устойчивых систем.

5. Показать сходство методов синтеза в частотной и временной областях.

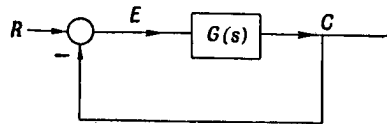
Связь между корневым годографом и амплитудно-фазовой характеристикой

На фиг. 1 показана упрощенная структурная схема системы автоматического регулирования с обратной связью, где $G(s)$ — передаточная функция разомкнутой системы, а s — комплексная переменная. Применяя частотный метод, можно, как показано на фиг. 2, построить на плоскости G амплитудно-фазовую характеристику. Эта характеристика является конформным отображением всей мнимой оси плоскости s на плоскость G . Амплитудно-фазовая характеристика на фиг. 2 и соответствующая ей мнимая ось на фиг. 3 указаны тонкой линией. Тонкая сплошная кривая представ-

¹⁾ Практический критерий, основанный на изучении свойств амплитудной частотной характеристики замкнутой системы, согласно которому максимальное значение этой характеристики не должно превышать 1,2—1,3. — *Прим. ред.*

3. Частотные характеристики и переходный процесс 119

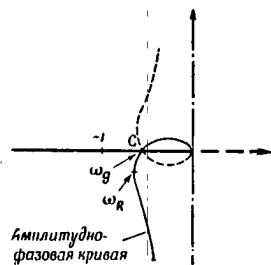
ляет отображение положительной части мнимой оси плоскости s , а тонкая пунктирная кривая — отображение отрицательной части мнимой оси. Амплитудно-фазовая характеристика симметрична относительно вещественной оси плоскости G .



Фиг. 1. Упрощенная структурная схема системы автоматического регулирования.

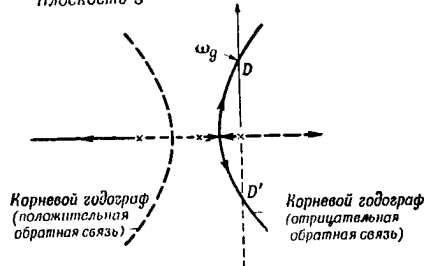
Для той же передаточной функции разомкнутой системы $G(s)$ можно методом корневого годографа построить на плоскости s корневой годограф, как показано на фиг. 3.

Плоскость G



Фиг. 2. Амплитудно-фазовая характеристика для $G(s) = \frac{K_v}{s(\tau_1 s + 1)(\tau_2 s + 1)}$.

Плоскость s



Фиг. 3. Корневой годограф для $G(s) = \frac{K}{s(\tau_1 s + 1)(\tau_2 s + 1)}$.

$$G(s) = \frac{K}{s(\tau_1 s + 1)(\tau_2 s + 1)}$$

Поскольку корневой годограф является геометрическим местом точек плоскости s , для которых функция $G(s)$ имеет фазовый угол, равный $\pm 180^\circ$, то эта кривая представляет собой конформное отображение отрицательной вещественной оси плоскости G на плоскость s . Отрицательная вещественная полуось на фиг. 2 и соответствующий ей корневой годограф на фиг. 3 показаны жирными сплошными линиями. В отличие от амплитудно-фазовой характеристики

конформное отображение положительной части вещественной оси плоскости G на плоскость s не является корневым годографом (однако оно будет корневым годографом, если система является регенеративной системой с обратной связью¹⁾). Это соответствие показано на фиг. 2 и 3 жирными пунктирными линиями. Корневой годограф симметричен относительно вещественной оси плоскости s .

Короче говоря, амплитудно-фазовая характеристика и корневой годограф представляют собой конформные отображения из плоскости s на плоскость G или обратно из плоскости G на плоскость s с помощью функции комплексного переменного $G(s)$. Это взаимное отображение и рассматривается как связь между амплитудно-фазовой характеристикой и корневым годографом.

Кривые на фиг. 2 и 3 построены для следующего вида передаточной функции разомкнутой системы $G(s)$:

$$G(s) = \frac{K_v}{s(\tau_1 s + 1)(\tau_2 s + 1)}.$$

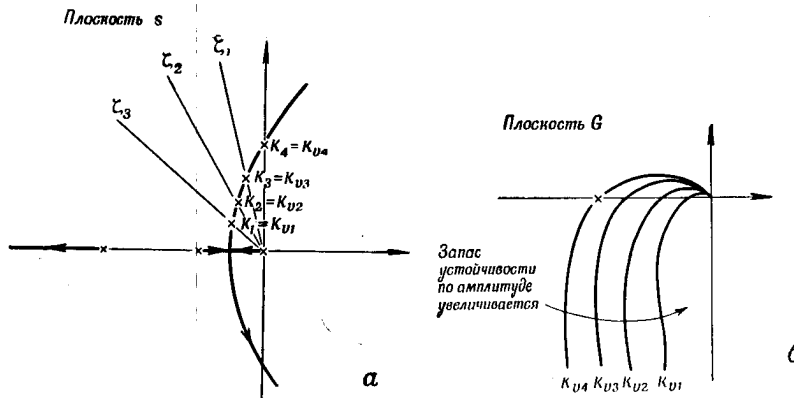
Следует заметить, что для данного значения K_v можно построить только одну амплитудно-фазовую характеристику. Однако соответствующий корневой годограф дает бесконечное число значений K_v . Чтобы учесть это обстоятельство, введем при построении корневого годографа вместо K_v новое обозначение K . Соотношение между K и K_v показано на фиг. 4, причем K является параметром, в то время как указанные значения K_v являются некоторыми заданными значениями.

Точка C на плоскости G (см. фиг. 2) соответствует точке D (а также точке D') на плоскости s (см. фиг. 3). Это следует из того факта, что точки C и D являются точками пересечения двух соответствующих пар кривых (жирной сплошной и тонкой сплошной). Следовательно, частота ω_D в точке C амплитудно-фазовой характеристики соответствует значению s в точке D на корневом годографе. Это очень важная пара точек. Из метода корневого годографа известно, что значение K в точке D есть максималь-

¹⁾ Т. е. будет фактически системой с положительной обратной связью. В этом случае критическая точка будет иметь координаты $(+1; j0)$. — *Прим. ред.*

ное значение коэффициента усиления системы, при превышении которого система будет неустойчивой. Это значение K можно назвать предельным усилением или пределом устойчивости.

Из теории метода частотных характеристик известно, что обратная величина абсолютного значения $G(j\omega)$ в точке C на отрицательной вещественной оси плоскости G является



Фиг. 4. Связь между коэффициентами K амплитудно-фазовой характеристики и корневого годографа.

a — корневого годограф на плоскости s .

$K_1 < K_2 < K_3 < K_4$

(K_4 — предельное усиление);

b — амплитудно-фазовая характеристика на плоскости G .

запасом устойчивости по амплитуде [6]. Легко можно показать, что запас устойчивости по амплитуде связан с предельным усилением соотношением

$$\frac{\text{Предельное усиление}}{\text{Запас устойчивости по амплитуде}} = K_v.$$

Так как при расчете в частотной области K_v и запас устойчивости по амплитуде известны, то предельное усиление системы определяется просто. Из предыдущего соотношения вытекает, что запас устойчивости по амплитуде обратно пропорционален K_v (предельное усиление постоянно для данной системы).

При увеличении запаса устойчивости по амплитуде (см. фиг. 4, б) система становится менее колебательной, потому

что при таком увеличении паре ближайших к мнимой оси корней на корневом годографе соответствует значительно больший коэффициент демпфирования (как это отмечено точками на корневом годографе фиг. 4, *a*) с соответствующим значением K .

Заметим, что система находится на границе устойчивости, когда значение K равно K_4 .

Вторая интересная особенность, вытекающая из соответствия между точкой C на плоскости G (см. фиг. 2) и точкой D на плоскости s (см. фиг. 3), состоит в том, что значение частоты ω_g на корневом годографе фиг. 3 характеризует максимально возможную полосу пропускания, которая получается, если расчет проводится в частотной области. Это происходит потому, что резонансная частота ω_R (см. фиг. 2) амплитудно-фазовой характеристики всегда меньше, чем частота ω_g (в точке C на фиг. 2), если элементы системы являются элементами минимально-фазового типа. Следовательно, как только построен корневой годограф при расчете системы методом корневого годографа, максимально возможная полоса пропускания немедленно определяется без перехода в частотную область. Таблица корневых годографов [7, 8] с соответствующими амплитудно-фазовыми характеристиками приведена на фиг. 5. На ней указана связь между корневыми годографами и амплитудно-фазовыми характеристиками. Все корневые годографы (кроме случаев P и C) имеют простую геометрическую форму: прямые линии, окружности, гиперболы (случаи M , H и O) и парабола (случай T). Заметим, что системы в случаях $Ж$ и O всегда неустойчивы.

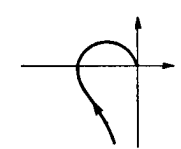
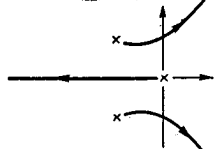
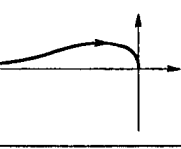
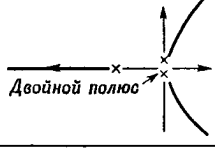
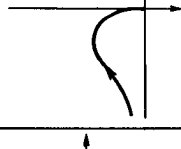
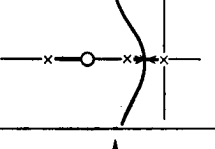
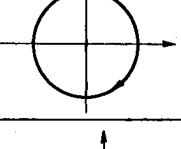
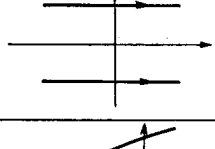
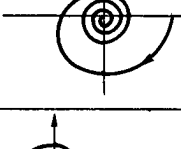
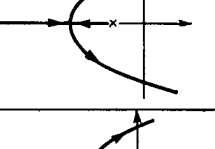
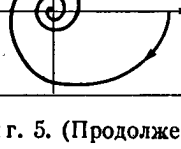
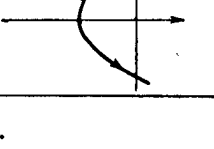
Поскольку амплитудно-фазовая характеристика является конформным отображением мнимой оси плоскости s на плоскость G с помощью передаточной функции $G(s)$, а корневой годограф является отображением отрицательной части вещественной оси плоскости G на плоскость s с помощью той же самой передаточной функции и поскольку переходный процесс можно построить, подобрав корни из корневого годографа, то можно сделать вывод, что частотная характеристика и переходный процесс связаны с помощью конформного отображения, как и должно было быть, учитывая известную связь, основанную на преобразовании Фурье.

Случай	$G(s)$	Амплитудно-фазовая характеристика в плоскости G	Корневой годограф в плоскости s
А	$\frac{1}{\tau s + 1}$		
Б	$\tau s + 1$		
В	$1 + \frac{1}{\tau s}$		
Г	$\frac{\tau s + 1}{\alpha \tau s + 1}$		
Д	$\frac{1}{s(\tau s + 1)}$		
Е	$\frac{1}{\tau^2 s^2 + 2\zeta \tau s + 1}$		

Фиг. 5. Таблица амплитудно-фазовых характеристик и соответствующих им корневых годографов.

Случай	$G(s)$	Амплитудно-фазовая характеристика в плоскости G	Корневой годограф в плоскости s
Ж	$\frac{1}{s^2}$		
З	$\frac{\tau_1 s + 1}{s(\tau_2 s + 1)}$		
И	$\frac{\tau_1 s + 1}{\tau_2^2 s^2 + 2\zeta\tau_2 s + 1}$ $\tau_2 > \zeta\tau_1$		
К	$\frac{\tau_1 s + 1}{\tau_2^2 s^2 + 2\zeta\tau_2 s + 1}$ $\tau_2 < \zeta\tau_1$		
Л	$\frac{\tau s + 1}{s^2}$		
М	$\frac{1}{s(\tau_1 s + 1)(\tau_2 s + 1)}$		

Фиг. 5. (Продолжение.)

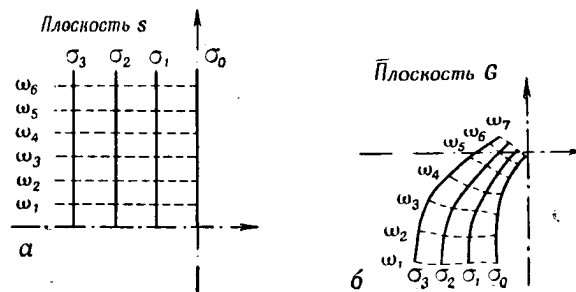
Случай	$G(s)$	Амплитудно-фазовая характеристика в плоскости G	Корневой годограф в плоскости s
H	$\frac{1}{s(\tau^2 s^2 + 2\zeta\tau s + 1)}$		
O	$\frac{1}{s^2(\tau s + 1)}$		 Двойной полюс
Π	$\frac{\tau_3 s + 1}{s(\tau_1 s + 1)(\tau_2 s + 1)}$		
p^*	e^{-Ts}		
C^*	$\frac{e^{-Ts}}{\tau s + 1}$		
T^*	$e^{-\sqrt{T}s}$		

Фиг. 5. (Продолжение).

В случаях, отмеченных звездочкой, показана лишь главная ветвь.

Соответствие между плоскостью s и плоскостью G для заданной функции $G(s)$

Амплитудно-фазовую характеристику можно обобщить посредством подстановки $s = \sigma + j\omega$ в передаточную функцию $G(s)$. Рассматривая σ как параметр, можно получить семейство кривых для различных постоянных σ . Амплитудно-фазовая характеристика является одной из этих кривых при $\sigma = 0$. Это семейство кривых, как показано на фиг. 6, является конформным отображением всех прямых, параллельных мнимой оси плоскости s , на плоскость G . Если прямые,



Фиг. 6. Семейства кривых постоянных σ и ω на плоскостях G и s .

параллельные вещественной оси плоскости s , конформно отобразить на плоскость G , то получится семейство кривых с постоянным ω , что также видно на фиг. 6. С помощью этих двух семейств кривых устанавливается взаимно-однозначное соответствие между плоскостями s и G .

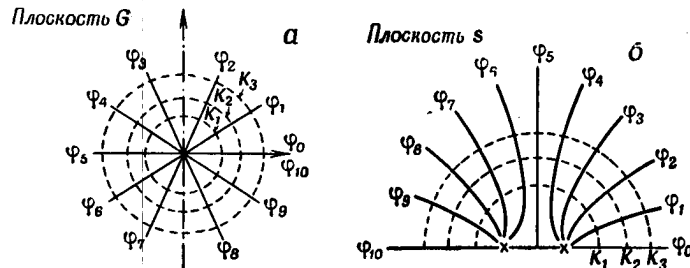
Аналогично корневой годограф можно обобщить с помощью соотношения $\arg G(s) = \varphi$. Для различных значений φ можно построить семейство кривых с постоянным φ (или фазовые кривые [10]). Корневой годограф является одной из кривых этого семейства при φ , равном $\pm 180^\circ$.

Это семейство показано на фиг. 7. Оно представляет собой конформное отображение всех радиальных линий, проходящих через начало координат плоскости G , на плоскость s . Если все концентрические окружности с центром в начале координат плоскости G конформно отобразить на плоскость s , то получится семейство кривых с постоянным r (или кривые усиления), изображенное на фиг. 7. С помощью

3. Частотные характеристики и переходный процесс 127

этих двух семейств кривых также устанавливается взаимно-однозначное соответствие между плоскостями s и G .

Это взаимно-однозначное соответствие между плоскостями s и G означает, что для каждой точки на плоскости G имеется соответствующая точка на плоскости s , и наоборот. Тогда любые сведения, содержащиеся на плоскости G , можно получить на плоскости s посредством отображения, и наоборот. Так как кривые на плоскости G дают сведения о частотной характеристике, а кривые на плоскости s дают сведения о переходном процессе, то очевидно, что любой



Фиг. 7. Семейство кривых постоянных φ (или кривые фазовых углов) и семейство кривых постоянного r (или кривые усиления) на плоскостях G и s .

из этих двух способов установления соответствия — с двумя семействами кривых при постоянных σ и ω или с двумя семействами кривых при постоянных усиления и фазе — можно использовать для нахождения искомой связи. Как указано в следующем разделе, имеется несколько методов [2—5, 11], использующих это соответствие для определения переходного процесса по частотной характеристике.

Очевидно, что если семейства амплитудных и фазовых кривых нанесены на плоскость s (фиг. 7, б), то амплитудно-фазовая характеристика является кривой на плоскости G , отвечающей значениям r и φ , которые соответствуют всем точкам мнимой оси плоскости s . Аналогично, если два семейства кривых с постоянными σ и ω (см. фиг. 6, б) нанесены на плоскость G , то корневой годограф является кривой в плоскости s , которая отвечает значениям σ и ω , соответствующим всем точкам отрицательной части вещественной оси плоскости G .

Определение переходного процесса по частотной характеристике

Амплитудно-фазовая характеристика, представляющая собой график передаточной функции разомкнутой системы, построенный в полярных координатах при $s = j\omega$ на плоскости G , является частотной характеристикой разомкнутой системы. Частотная характеристика замкнутой системы является графиком передаточной функции замкнутой системы при $s = j\omega$. Каждую из этих частотных характеристик можно построить либо в полярных координатах, либо как амплитудную и фазовую частотные диаграммы [9].

Для определения переходного процесса по амплитудно-фазовой характеристике (по частотной характеристике разомкнутой системы) или по частотной характеристике замкнутой системы были разработаны различные методы. Поскольку желаемые корни должны выбираться слева от мнимой оси плоскости s и поскольку амплитудно-фазовая характеристика является конформным отображением мнимой оси плоскости s , то сама амплитудно-фазовая характеристика не определяет значения желаемых корней. Таким образом, определение корней (а затем построение переходного процесса) с помощью соответствия между плоскостью s и плоскостью G потребует отображения не мнимой оси плоскости s , а некоторой линии или кривой слева от мнимой оси плоскости s , где расположены желаемые корни. В работах [2] и [3] было использовано соответствие, показанное на фиг. 6, а в работах [4, 5, 11] — конформное отображение с помощью передаточной функции разомкнутой системы линии постоянного коэффициента демпфирования плоскости s на плоскость G . Оба эти преобразования используют линии слева от мнимой оси плоскости s для непосредственного нахождения корней по соответствующим амплитудно-фазовым характеристикам. Возможность использования той или иной кривой левой полуплоскости s в значительной мере зависит от удобства вычерчивания соответствующих им амплитудно-фазовых характеристик на плоскости G .

В упомянутых выше методах использовалась амплитудно-фазовая характеристика (частотная характеристика разомкнутой системы). В то время как с помощью частотной характеристики разомкнутой системы нельзя непосредственно опре-

делить значения корней, частотная характеристика замкнутой системы позволяет найти их, потому что частотную характеристику замкнутой системы можно получить по передаточной функции замкнутой системы, а если последняя известна, можно определить переходный процесс.

Прекрасным примером применения частотной характеристики замкнутой системы является метод Флойда [6]. Этот метод позволяет найти импульсную характеристику системы по вещественной части частотной характеристики замкнутой системы (это же можно непосредственно проделать по амплитудно-фазовой характеристике с помощью номограммы¹⁾).

Возможны, по-видимому, и другие методы, использующие частотную характеристику замкнутой системы. Например, можно было бы попытаться найти множители первого порядка (такие, как $\tau s + 1$ или $1/(\tau s + 1)$) и множители второго порядка (такие, как $\tau^2 s^2 + 2\zeta\tau s + 1$ или $1/(\tau^2 s^2 + 2\zeta\tau s + 1)$) по частотной характеристике замкнутой системы методом последовательных приближений. После того как множители найдены, определяется передаточная функция замкнутой системы, и по ней уже строят переходный процесс. Хотя практичность предлагаемого метода будет зависеть от количества труда, необходимого для нахождения множителей для данной частотной характеристики замкнутой системы, все же такое соотношение между частотной характеристикой замкнутой системы и полюсами передаточной функции замкнутой системы показывает природу связи между частотной характеристикой и переходным процессом и, таким образом, обеспечивает лучшую интерпретацию частотной характеристики замкнутой системы. Напомним, что передаточная функция замкнутой системы связана с передаточной функцией разомкнутой системы посредством корневого годографа.

Исследование M_p -критерия частотным методом²⁾

Если известны нули и полюсы передаточной функции замкнутой системы, как это имеет место при синтезе

¹⁾ Метод трапециoidalных частотных характеристик, предложенный В. В. Солодовниковым, позволяет решить аналогичную задачу (см., например, «Основы автоматического регулирования», Машгиз, 1954). — *Прим. ред.*

²⁾ См. примечание на стр. 118. — *Прим. ред.*

методом корневого годографа, то частотную характеристику системы можно легко получить, если вычерчены как амплитудная, так и фазовая диаграммы [12].

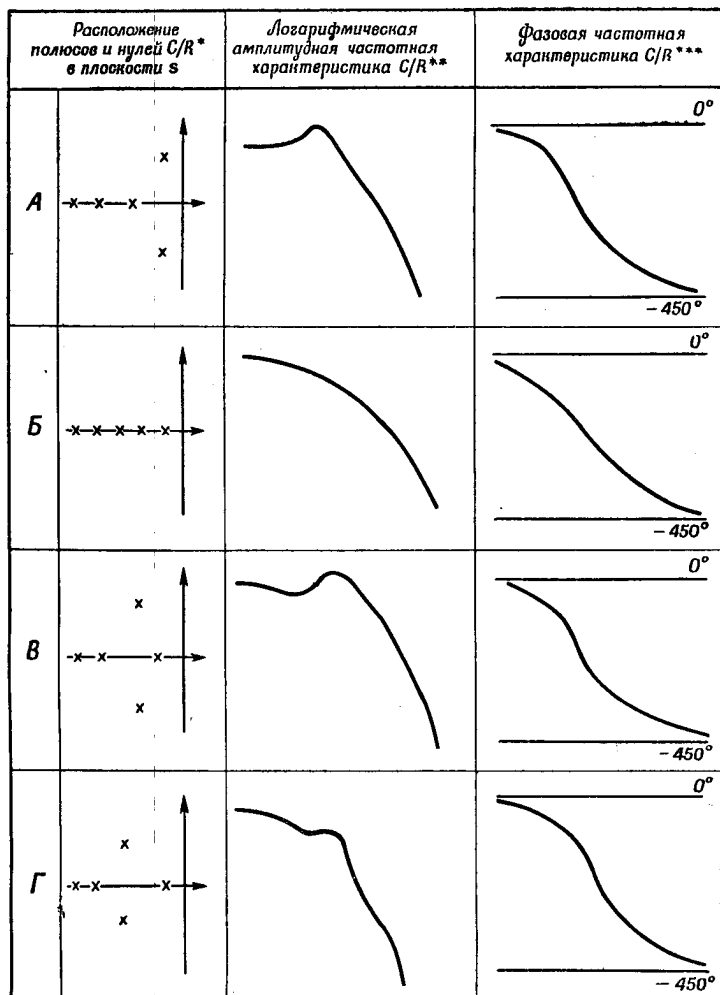
На фиг. 8 приведена таблица некоторых частотных характеристик замкнутой системы как амплитудных, так и фазовых с соответствующим им расположением нулей и полюсов на плоскости s .

Исследование случаев, представленных на фиг. 8, показывает связь между M_p -критерием частотного метода и расположением нулей и полюсов, которое определяет переходный процесс. В первых пяти случаях (фиг. 8, *A—D*) передаточные функции не содержат нулей. Амплитудная частотная характеристика в случае *A* имеет резонансный пик. В этом случае, который следует считать желательным типом частотной характеристики, имеется пара комплексно сопряженных полюсов, расположенных ближе других к мнимой оси в левой полуплоскости s .

Если такая пара корней отсутствует и все полюсы простые, то амплитудная характеристика имеет вид, показанный на фиг. 8, *B*. Если ближайшим к оси полюсом является простой полюс и, кроме того, имеется пара комплексно сопряженных полюсов, как в случаях *B* и *T*, то переходный процесс имеет большое время установления, обусловленное наличием простого полюса, ближайшего к мнимой оси. Различие в виде амплитудных характеристик в случаях *B* и *Г* зависит от относительного расположения простых и комплексных полюсов. Из теории частотного метода известно, что логарифмическая амплитудная частотная характеристика с крутым наклоном дает неудовлетворительный переходный процесс (случай *B*).

Если имеются две пары комплексно сопряженных полюсов, расположенных близко к мнимой оси, как это имеет место в случае *D*, то амплитудная частотная характеристика может иметь два резонансных пика.

В последних трех случаях (фиг. 8, *E—З*) передаточная функция имеет нуль. Амплитудная характеристика для случая *E* подобна характеристике для случая *A*, но их фазовые диаграммы отличаются. Случай с нулем характеризуется меньшей фазой при низких частотах. Случай *Ж* подобен случаю *B*, но в первом случае имеется нуль слева от мнимой оси плоскости s ; фазовые характеристики также заметно

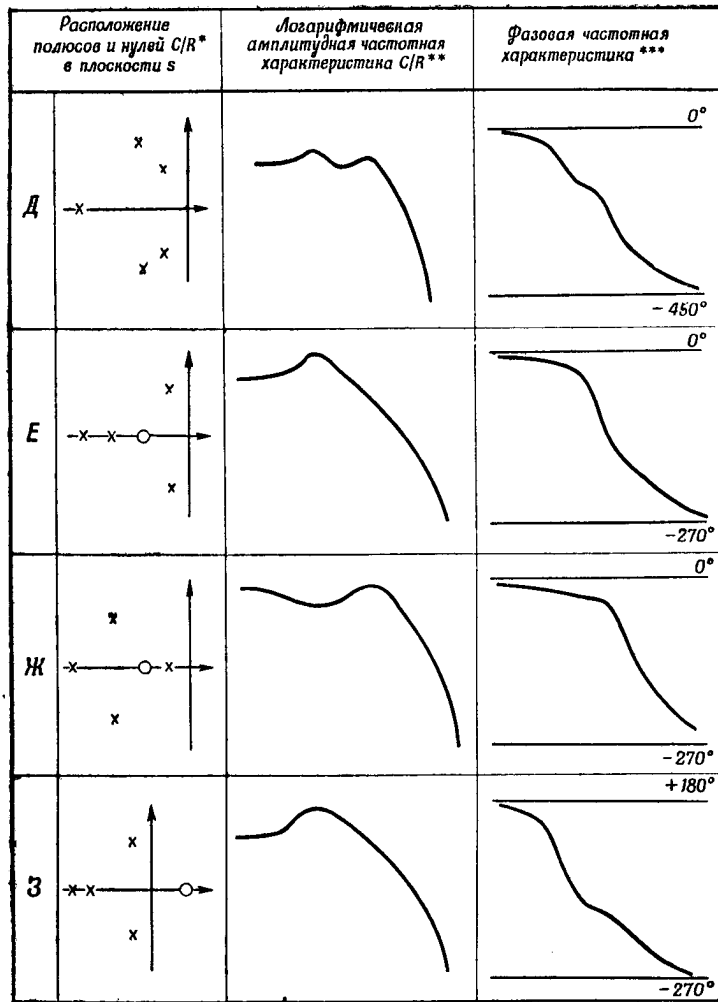


Фиг. 8. Таблица распределений полюсов и нулей C/R и соответствующие амплитудные и фазовые характеристики.

* C/R — передаточная функция замкнутой системы;

** Кривая $\lg |C/R|$ в функции $\lg \omega$;

*** Кривая фазовых углов в функции $\lg \omega$.



Фиг. 8. (Продолжение).

отличаются, и переходный процесс будет иметь большое время установления.

При очень близком расположении нуля к началу координат плоскости s пик на логарифмической амплитудной частотной характеристике становится очень большим. Это согласуется с выводом о том, что в случае, когда нуль расположен близко к началу координат, перерегулирование очень велико [10]. Известно также, что если нуль находится в правой полуплоскости s на некотором расстоянии от начала координат, то перерегулирование уменьшается. Об этом можно судить и по частотным характеристикам; действительно, логарифмическая амплитудная характеристика для случая Z подобна характеристике для случая E , но на фазовой характеристике в случае Z имеется большой положительный фазовый сдвиг, чего нет в случае E .

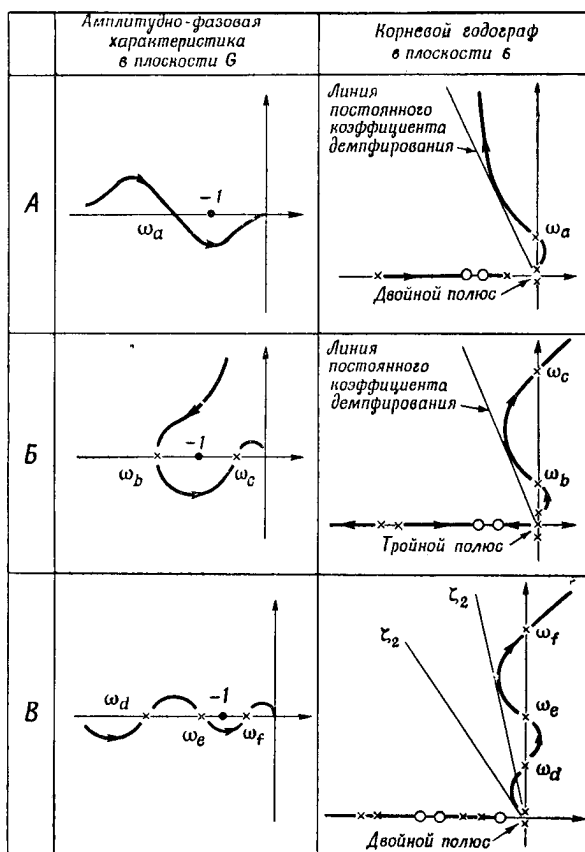
Из рассмотрения этих случаев можно сделать вывод, что частотная характеристика дает чрезвычайно ценные сведения относительно поведения системы в переходном процессе. Критерий M_p может обеспечить удовлетворительный переходный процесс, если, исследовав амплитудную и фазовую частотные характеристики замкнутой системы, мы установим наличие пары комплексно сопряженных корней, ближайших к мнимой оси, рассмотрев соответствующее им расположение нулей и полюсов передаточной функции замкнутой системы.

Сравнение частотного метода и метода корневых годографов для условно устойчивых систем

Условно устойчивой называется такая система, которая становится неустойчивой при достаточно малом коэффициенте усиления в противоположность абсолютно устойчивой системе, в которой не возникает неустойчивости при уменьшении коэффициента усиления системы. Примером условно устойчивой системы является позиционная следящая система с тахогенератором, включенным последовательно с пассивным RC -фильтром высоких частот в цепи обратной связи.

На фиг. 9 показаны амплитудно-фазовые характеристики для трех типов условно устойчивых систем и соответствующие им корневые годографы. Все амплитудно-фазовые характеристики построены для случая устойчивых систем.

Амплитудно-фазовой характеристике в случае *A* будет соответствовать неустойчивая система, если коэффициент усиления достаточно уменьшится, но эта система всегда



Фиг. 9. Характеристики условно устойчивых систем.

будет устойчива при увеличении коэффициента усиления. Это следует из соответствующего корневого годографа, так как только малая часть этой кривой находится в правой полуплоскости s , которая и представляет область неустойчивости при малых значениях коэффициента усиления.

Амплитудно-фазовой характеристике в случае *Б* будет соответствовать неустойчивая система при достаточно малых и достаточно больших коэффициентах усиления. Из соответствующего корневого годографа видно, что лишь одна часть кривой располагается в левой полуплоскости, являющейся областью устойчивости.

Амплитудно-фазовая характеристика в случае *В* соответствует точке корневого годографа между частотами ω_e и ω_r . Следовательно, эта система будет неустойчивой при достаточно малых или достаточно больших коэффициентах усиления. Система снова станет устойчивой при уменьшении коэффициента усиления до значения, которому соответствует точка на корневом годографе между началом координат и частотой ω_a (на плоскости s). Таким образом, по корневному годографу можно легко определить области значений коэффициента усиления, необходимого для устойчивой работы системы.

Амплитудно-фазовые характеристики пересекают отрицательную вещественную полуось плоскости G в точках, в которых значения частоты равны соответственно $\omega_a, \omega_b, \dots, \omega_r$.

При тех же значениях частоты корневые годографы пересекаются с положительной мнимой полуосью плоскости s .

Как было показано ранее, частота ω_a на плоскости s в случае *А* определяет минимальную полосу пропускания системы; в случае *Б* диапазон частот от ω_b до ω_e на плоскости s является допустимой полосой пропускания системы; в последнем случае *В* допустимая полоса пропускания устойчивой системы будет заключена в диапазонах частот от 0 до ω_a и от ω_e до ω_r на плоскости s . Таким образом, по корневному годографу можно также легко определить ширину полосы пропускания частот.

По корневым годографам можно не только определить область устойчивости условно устойчивых систем, но также и степень их устойчивости. Касательные, проведенные к частям корневого годографа в устойчивой области плоскости s , являются линиями постоянного коэффициента демпфирования. Они указывают максимально возможный коэффициент демпфирования для пары комплексно сопряженных корней, ближайших к мнимой оси этой устойчивой части кривой, как показано на фиг. 9.

Эквивалентный критерий для амплитудно-фазовой характеристики на плоскости G состоит в том, что амплитудно-фазовая характеристика должна проходить достаточно далеко от точки -1 . Поскольку установлена связь между амплитудно-фазовой характеристикой и корневым годографом, то параметры амплитудно-фазовой характеристики в плоскости G можно выразить через параметры переходного процесса.

Сходство синтеза в частотной и временной областях

Частотные методы синтеза систем автоматического регулирования являются мощным средством расчета, так как амплитудно-фазовую характеристику разомкнутой системы можно легко получить, зная амплитудно-фазовые характеристики отдельных элементов системы. На фиг. 10, *a* в качестве примера синтеза в частотной области приведена система автоматического регулирования. На фиг. 10, *б* и *в* показаны частотные характеристики (амплитудно-фазовые характеристики на плоскости G_1 и G_2 соответственно) регулятора и объекта регулирования. Умножая амплитуды и складывая фазовые углы амплитудно-фазовых характеристик двух элементов для различных частот, получаем частотную характеристику разомкнутой системы (амплитудно-фазовую характеристику на плоскости G), показанную на фиг. 10, *г*. Расположение амплитудно-фазовой характеристики относительно точки -1 на плоскости G определяет абсолютную устойчивость и степень устойчивости системы.

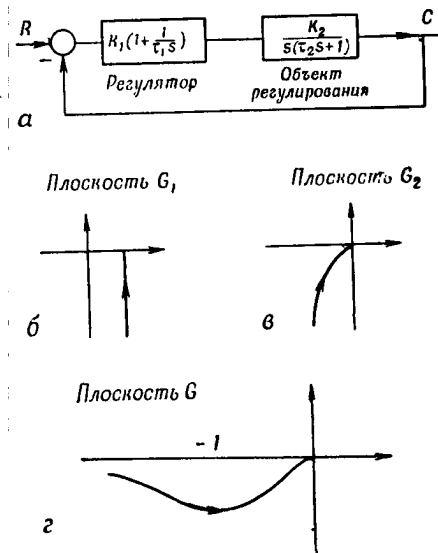
Итак, амплитудно-фазовая характеристика системы легко строится с помощью амплитудно-фазовых характеристик отдельных ее составляющих. Характеристики замкнутой системы можно определить на плоскости G с помощью двух семейств контуров¹⁾, не показанных на фиг. 10, *г*, а затем построить. Такой расчет упрощается при использовании логарифмических амплитудной и фазовой частотных характеристик [9].

Метод корневого годографа является методом синтеза систем во временной области. Выше установлена связь между амплитудно-фазовой характеристикой на плоскости G

¹⁾ Имеются в виду амплитудные и фазовые круговые диаграммы. — *Прим. ред.*

3. Частотные характеристики и переходный процесс 137

и корневым годографом на плоскости s . Способ синтеза систем методом корневых годографов, подобный методу амплитудно-фазовых характеристик, основан на введении в рассмотрение кривых постоянных фазовых углов [10].



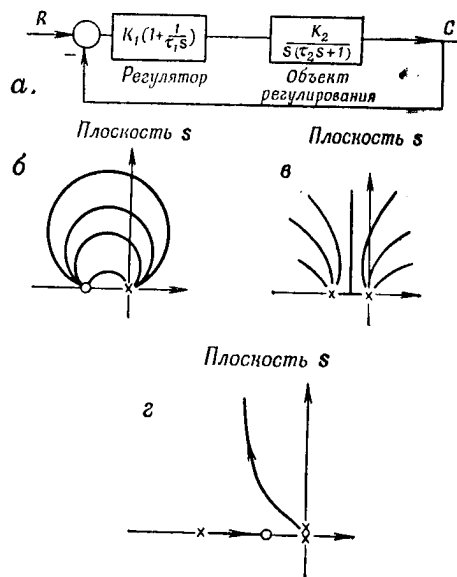
Ф и г. 10. Пример синтеза системы автоматического регулирования с обратной связью в частотной области.

a — структурная схема; $б$ — амплитудно-фазовая характеристика объекта регулирования; $в$ — амплитудно-фазовая характеристика разомкнутой системы; $г$ — амплитудно-фазовая характеристика регулятора.

На фиг. 11, a приведена система, подобная системе на фиг. 10, a . На фиг. 11, $б$ и $в$ соответственно показаны кривые фазовых углов регулятора и объекта регулирования. Путем наложения (не показанного на фиг. 11) этих двух кривых фазовых углов на той же плоскости s и установления точек, которые дают сумму $\pm 180^\circ$ для любых двух кривых фазовых углов, легко можно получить корневой годограф системы (фиг. 11, $г$). Поскольку корневой годограф дает все возможные корни характеристического

уравнения системы, желаемый переходный процесс можно получить, выбирая подходящие корни из корневого годографа.

Для того чтобы система была устойчивой, рассматривают лишь корни, расположенные в левой полуплоскости s . Корневой годограф можно также построить с помощью кривых фазовых углов отдельных элементов системы. Хотя синтез



Фиг. 11. Пример синтеза системы автоматического регулирования с обратной связью во временной области.

a — структурная схема; b — кривые фазовых углов регулятора; v — кривые фазовых углов объекта; z — корневой годограф системы регулирования.

во временной области содержит лишь суммирование углов на комплексной плоскости и не требует шаблонов, однако в настоящее время он, вообще говоря, требует большего графического труда, чем синтез в частотной области. Таким образом, показано сходство синтеза в частотной и временной областях.

Преимущество частотного метода состоит в том, что функция частоты является функцией комплексной переменной

и при этом можно применять конформное отображение. Метод корневого годографа позволяет применить конформное отображение для синтеза во временной области.

Заключение

Связь между частотной характеристикой и переходным процессом в общем виде трудно сформулировать, ибо она зависит от конкретного вида корневого годографа. Благодаря наличию связи между амплитудно-фазовой характеристикой на плоскости G и корневым годографом на плоскости s посредством конформного отображения можно определять частотную характеристику через параметры переходного процесса.

Характеристику системы автоматического регулирования можно найти, когда определены полюсы и нули передаточной функции замкнутой системы и соответствующие постоянные множители. Эти нули и полюсы, задающие, с одной стороны, передаточную функцию замкнутой системы, а с другой — переходный процесс, можно использовать как средство для выражения частотной характеристики через параметры переходного процесса, и наоборот.

Показано несколько условно устойчивых систем и рассмотрено соответствие между их амплитудно-фазовыми характеристиками и корневыми годографами. Показано также, что условная устойчивость, которая прежде объяснялась на основе критерия Найквиста на плоскости G , заключается в различном или, вернее, необычном пересечении корневым годографом мнимой оси плоскости s . Наконец, было показано, что синтез можно выполнить графически аналогичным образом как в частотной, так и во временной областях.

Таким образом, в настоящей работе показано, что связь между частотной характеристикой и переходным процессом легко устанавливается с помощью конформного отображения.

Какому из двух методов — частотному методу или методу корневого годографа — следует отдать предпочтение, зависит от усмотрения расчетчика, который учитывает удобство вычерчивания кривых и характер конкретной задачи. В некоторых случаях оба метода дополняют друг друга.

ЛИТЕРАТУРА

1. Evans W. R., *Trans. AIEE*, Part I, **63**, 66 (1950).
2. Profos P., *Sulzer Techn. Rev.*, **2**, 1 (1954).
3. Russel D. W., Weaver C. H., *Trans. AIEE*, Part II, **71**, 95 (1952).
4. Kusters N. L., Moore W. F., „A Generalization of the Frequency Response Method for the Study of Feedback Control Systems“, *Proceedings of the Conference on Automatic Control*, London, 1952.
5. Biernson G., *Trans. AIEE*, Part II, **72**, 53 (1953).
6. Brown G. S., Campbell D. P., *Principles of Servomechanisms*, New York, 1948, Chap. 11.
7. Yeh V. C. M., „The Study of Transients in Linear Feedback Systems by Conformal Mapping and the Generalized Root Locus Method“, *Диссертация Mass. Inst. Technol.*, 1952.
8. Chu Y., *Trans. AIEE*, Part II, **71**, 291 (1952).
9. Теория следящих систем, под ред. Х. Джеймса, Н. Никольса и Р. Филиппса, *Издательство иностранной литературы*, 1953.
10. Chu Y., *Trans. AIEE*, Part II, **71**, 330 (1952).
11. Леонард А., Относительное демпфирование как критерий устойчивости, а также как средство для нахождения корней полинома Гурвица, см. сборник „Автоматическое регулирование“, *Издательство иностранной литературы*, 1954, стр. 20.
12. Ward E. E., *Wireless Eng.*, **27**, № 320, 146 (1950).

4. ОПРЕДЕЛЕНИЕ ПЕРЕХОДНЫХ ПРОЦЕССОВ ПО ЧАСТОТНЫМ ХАРАКТЕРИСТИКАМ РАЗОМКНУТОЙ СИСТЕМЫ¹⁾²⁾

Дж. Бьернсон

Введение

Синтез системы автоматического регулирования с обратной связью обычно осуществляется с использованием ее частотных характеристик, однако поведение системы лучше всего оценивать по ее переходному процессу. Следовательно, необходимо установить по возможности непосредственную связь переходного процесса с частотными характеристиками. Хотя точные, достаточно общие математические соотношения между этими двумя характеристиками получить трудно, все же можно найти более или менее надежные приближения. В данной статье путем исследования и сравнения переходных процессов для большого числа типов передаточных функций систем автоматического регулирования выводятся некоторые наиболее важные из этих приближений.

Соотношения между реакцией системы на ступенчатое воздействие и частотными характеристиками систем автоматического регулирования рассматривались много раз, но, несмотря на это, между ними до сих пор не установлено достаточно полной связи. Обычно инженер знает, что быстрота реакции системы на ступенчатое воздействие некоторым образом связана с шириной полосы пропускания системы, однако при количественном определении этого соотношения он, вероятно, встретил бы большие трудности, исключая разве слишком идеализированные случаи.

Однако еще более важным, чем недостаток знания связи между реакцией на ступенчатое возмущение и частотными характеристиками, является полное игнорирование соотношений между частотными характеристиками и другими переходными процессами системы: импульсной переходной характеристикой, реакцией системы на вход с постоянной

¹⁾ Biernson G. A., *Trans. AIEE*, Part II, 74, 388—403 (1955).

²⁾ Рассматриваются логарифмические амплитудные характеристики, сокращенно ЛАХ. — *Прим. ред.*

скоростью, реакцией на скачок ускорения и т. д. Обычно эти переходные процессы гораздо легче определить по частотным характеристикам, чем реакцию на ступенчатое воздействие. Вероятно, главная причина пренебрежения этими переходными характеристиками состоит в том, что не в достаточной мере понимается их роль в общем поведении системы автоматического регулирования во времени.

В работе [1] было указано на значение различных переходных характеристик систем автоматического регулирования. Было показано, что для отдельных систем автоматического регулирования имеется лишь несколько переходных характеристик, которые являются важными, и если все они известны, то расчетчику достаточно использовать грубые приближения для каждой из них, чтобы легко и с приемлемой точностью вычислить кривую реакции системы на произвольный сигнал на входе. Следовательно, если расчетчик может достаточно точно оценить эти переходные характеристики по частотным характеристикам разомкнутой системы, то он сможет установить непосредственное влияние цепей, которые вводятся им в систему автоматического регулирования, на реакцию результирующей системы для любого входа. В данной статье делается попытка достичь этой цели.

Наиболее важные характеристики переходных процессов лучше всего можно оценить путем исследования частотных характеристик разомкнутой системы. Это обстоятельство следует считать весьма благоприятным, так как характеристики разомкнутого контура — это именно те характеристики, которые выбираются при расчете систем автоматического регулирования. Основным исключением является максимальное перерегулирование системы при ступенчатом возмущении, которое можно лучше всего установить по максимальным значениям (M_p) частотной характеристики замкнутой системы. Правила для оценки этих переходных характеристик устанавливаются в статье на основании эмпирических и интуитивных соображений путем исследования большого числа различных передаточных функций. Для этих правил автору удалось получить приемлемое теоретическое обоснование, слишком громоздкое, чтобы его можно было изложить в данной статье. Оно основывается на результатах, приведенных в статье [2], в которой показано, как полюсы

передаточной функции замкнутой системы связаны с частотными характеристиками системы в разомкнутом состоянии, и ему будет посвящена следующая статья. Большая часть материала в настоящей статье была опубликована первоначально в части IV работы [3].

Терминология

Термины и обозначения, используемые в этой статье, указаны на блок-схеме системы автоматического регулирования с обратной связью (фиг. 1), где X_i — входной сигнал, X_e — сигнал ошибки, X_u — сигнал обратной связи.

Сигнал обратной связи часто равняется выходному сигналу X_o . Наиболее важные передаточные функции системы автоматического регулирования таковы:

$G = \frac{X_u}{X_e}$ — передаточная функция контура;

$\frac{X_u}{X_i}$ — передаточная функция замкнутой системы по отношению к сигналу обратной связи;

$\frac{X_e}{X_i}$ — передаточная функция ошибки.

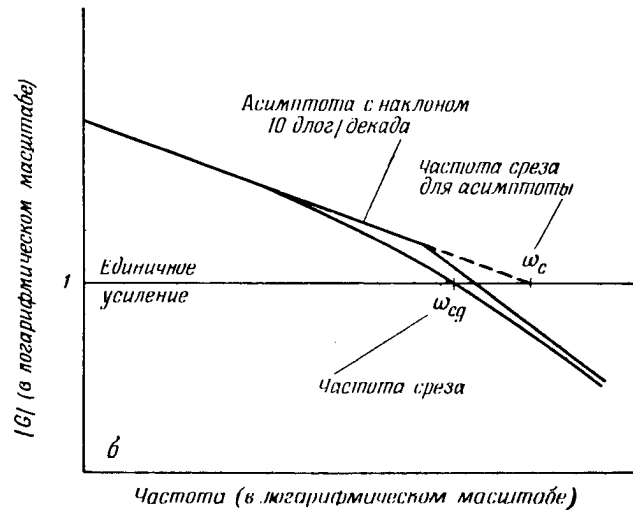
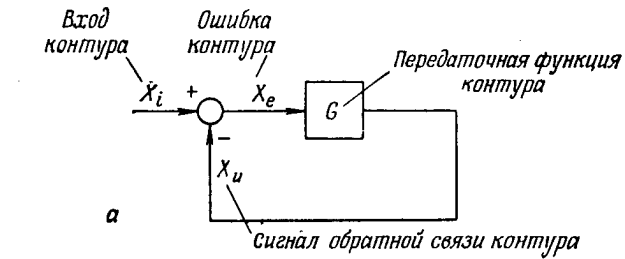
Передаточная функция контура G часто называется передаточной функцией разомкнутой системы. Модуль величины G при вещественных частотах называется амплитудной характеристикой, или усилением системы.

Наиболее важным динамическим параметром системы автоматического регулирования является ширина полосы пропускания; автор нашел, что прекрасным критерием для ширины полосы пропускания системы автоматического регулирования является частота среза, обозначаемая через ω_{cg} и определяемая как частота, при которой усиление системы равно единице. Частота среза имеет следующие преимущества перед другими критериями для ширины полосы пропускания: 1) ее легко можно измерить и вычислить, 2) она дает самое надежное указание относительно времени нарастания реакции системы на ступенчатое воздействие для всех типов систем автоматического регулирования, как это показано в настоящей статье.

Другим полезным параметром ширины полосы пропускания, связанным с частотой среза ω_{cg} , является частота среза

асимптоты, обозначаемая через ω_c . Частоты ω_c и ω_{cd} показаны на фиг. 1.

Вместо единицы децибел (дб) мы используем новую, более общую логарифмическую единицу — децилог (длог). Для



Фиг. 1. Пояснение принятой терминологии.

какого-либо безразмерного отношения A значение его в децилогах будет равно

$$A \text{ (в децилогах)} = 10 \log_{10} A. \quad (1)$$

Таким образом, отношению 10 соответствует 10 децилогов, что сокращенно записывается как 10 длог.

Элементы переходного процесса

В статье [1] было показано, что имеется три основных семейства характеристик, которые определяют общее поведение системы автоматического регулирования во времени; это составляющие переходного процесса $T_n(t)$, коэффициенты установившегося состояния c_n и коэффициенты начальных значений a_m . Эти характеристики появляются в выходной реакции системы на единичный скачок производной или интеграла от входного сигнала. Выходную реакцию системы автоматического регулирования на единичный скачок производной n -го порядка можно выразить следующим образом

$$x_u(t) = T_n(t) + c_n + c_{n-1} \frac{t}{1!} + c_{n-2} \frac{t^2}{2!} + \dots + c_1 \frac{t^{n-1}}{(n-1)!} + c_0 \frac{t^n}{n!}. \quad (2)$$

Переходную составляющую $T_n(t)$ назовем составляющей переходного процесса системы по n -й производной, а коэффициент c_n назовем коэффициентом установившегося состояния системы по n -й производной. Для единичного скачка m -кратного интеграла реакция на выходе будет

$$x_u(t) = T_{-m}(t) + a_{m-1}u_1(t) + a_{m-2}u_2(t) + \dots + a_0u_m(t), \quad (3)$$

где $u_1(t)$ — единичный импульс при $t = 0$; $u_2(t)$ — производная от единичного импульса; $u_3(t)$ — вторая производная от единичного импульса и т. д. Коэффициент a_m назовем коэффициентом начального значения m -й производной.

Коэффициенты установившегося значения и коэффициенты начальных значений связаны с начальными значениями соответствующих составляющих переходного процесса. Если система имеет конечную ширину полосы пропускания, то начальные значения реакции на ступенчатое возмущение и реакции на скачки производных должны равняться нулю¹⁾.

Полагая выражение (2) равным нулю при $t = 0$, получим

$$c_n = -T_n(0) \quad \text{для } n \geq 0. \quad (4)$$

¹⁾ Так как при конечной полосе пропускания системы порядок числителя передаточной функции меньше порядка знаменателя. — *Прим. ред.*

Можно также показать, что коэффициенты начальных значений должны быть равны

$$a_m = T_{-m}(0) \text{ для } m > 0. \quad (5)$$

Если система имеет конечную ширину полосы пропускания, то коэффициент начального значения реакции на ступенчатое возмущение a_0 равен нулю. Первые два члена разложения в уравнении (2) равны $[T_n(t) + c_n]$, и, как это видно из уравнения (4), их сумма определяет характеристику, имеющую начало в нуле. Это свойство очень удобно, и поэтому вместо составляющей переходного процесса $T_n(t)$ часто строится график суммы $[T_n(t) + c_n]$, который называется составным переходным процессом по n -й производной.

Скачок производной представляет собой низкочастотный вход, а для таких входов реакция на выходе следует за входом так близко, что часто необходимо рассматривать кривую изменения ошибки, а не выходную реакцию (переходный процесс). Ошибка системы связана с выходным сигналом соотношением

$$x_e(t) = x_i(t) - x_u(t). \quad (6)$$

Подставляя разложение для $x_u(t)$ [уравнение (2)] в выражение (6), получим следующее разложение для $x_e(t)$:

$$x_e(t) = T'_n(t) + c'_n + c'_{n-1} \frac{t}{1!} + c'_{n-2} \frac{t^2}{2!} + \dots \\ \dots + c'_1 \frac{t^{n-1}}{(n-1)!} + c'_0 \frac{t^n}{n!}, \quad (7)$$

где

$$T'_n(t) = -T_n(t), \quad (8)$$

$$c'_n = -c_n \text{ для } n > 0, \quad (9)$$

$$c'_0 = 1 - c_0. \quad (10)$$

Составляющая переходного процесса $T_n(t)$ является составляющей переходного процесса ошибки по n -й производной, а коэффициент c'_n является коэффициентом установившегося состояния ошибки, который обычно известен как коэффициент ошибки по n -й производной.

4. Определение переходных процессов по ЛАХ'ам 147

Для систем с простыми полюсами составляющие переходного процесса можно вычислить с помощью соотношения

$$T_n(t) = \sum_{i=1}^L \frac{K_{i_0}}{(s_i)^n} e^{s_i t}, \quad (11)$$

где s_i — один из L полюсов передаточной функции системы, а K_{i_0} — коэффициент реакции на ступенчатое воздействие для этого полюса, равный

$$K_{i_0} = \frac{(s - s_i) X_u}{s X_i} \Big|_{s=s_i}. \quad (12)$$

Независимо от порядка полюса системы составляющие переходного процесса по последовательным производным связаны соотношением

$$T_n(t) = \int T_{n-1}(t) dt. \quad (13)$$

Таким образом, если какая-либо составляющая переходного процесса определена с достаточной точностью, то все другие составляющие можно определить последовательным дифференцированием и интегрированием.

Коэффициенты установившегося состояния можно вычислить по формуле

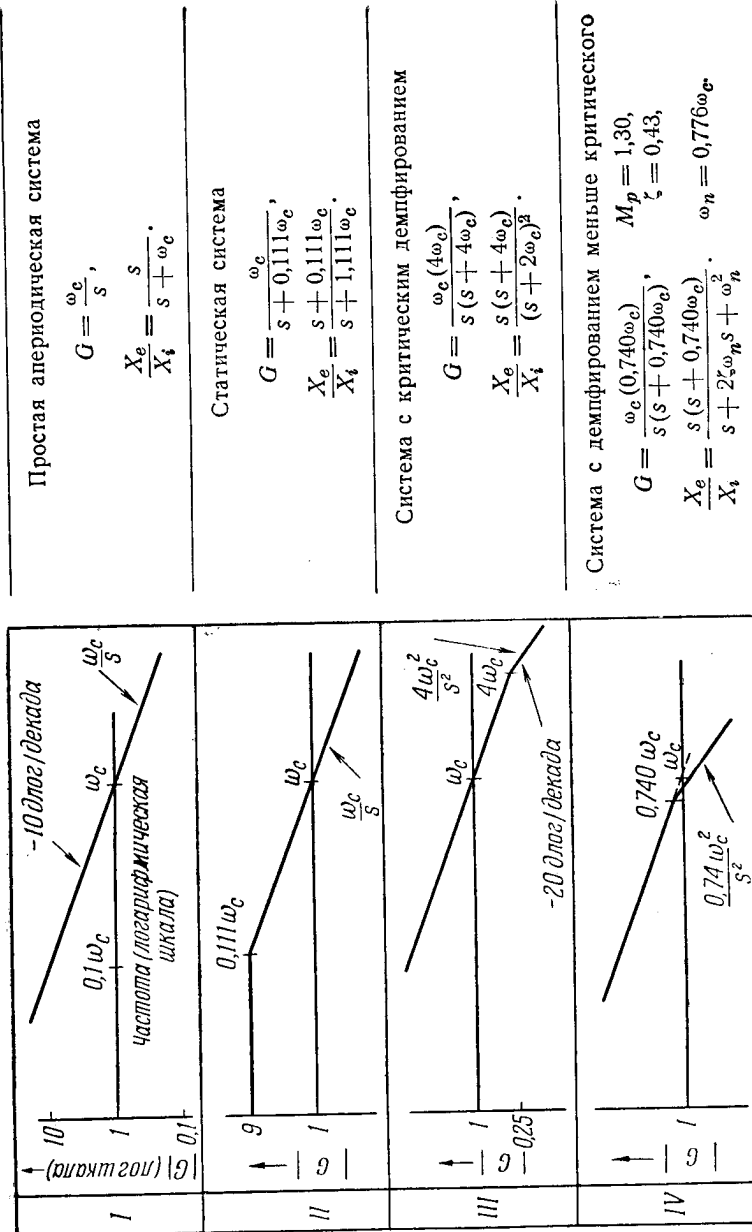
$$c_n = \frac{1}{n!} \frac{d^n}{ds^n} \frac{X_u}{X_i} \Big|_{s=0}. \quad (14)$$

Коэффициенты начальных значений можно определить, заменяя s в X_u/X_i на $1/\lambda$ и используя соотношение

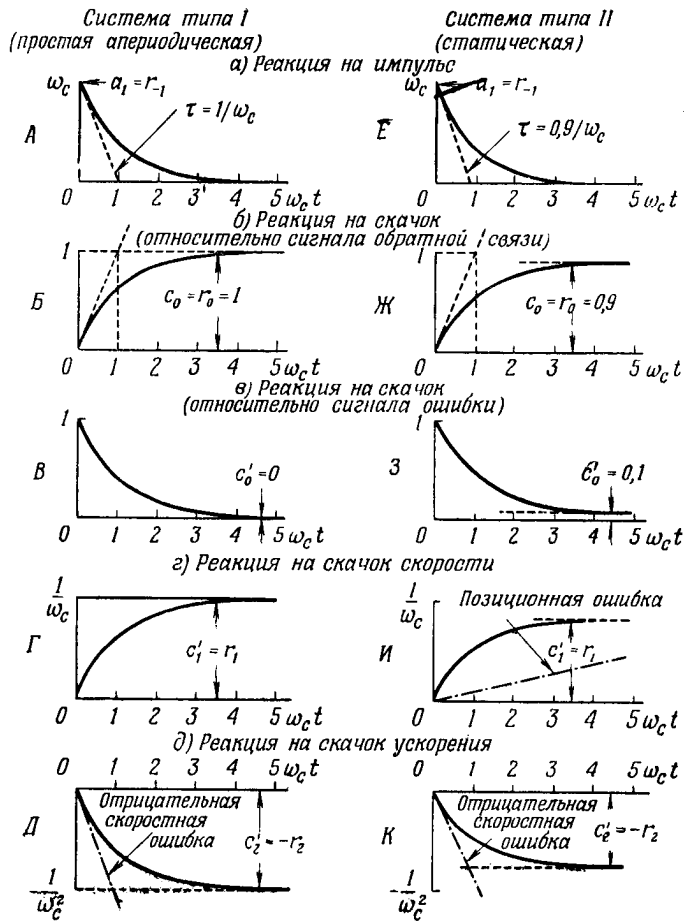
$$a_m = \frac{1}{m!} \frac{d^m}{d\lambda^m} \frac{X_u}{X_i} \left(\frac{1}{\lambda} \right) \Big|_{\lambda=0}. \quad (15)$$

С другой стороны, эти коэффициенты обычно гораздо легче вычислять с помощью разложения передаточной функции X_u/X_i по правилу деления многочленов, как показано в приложении к статье [1]. Если члены числителя и знаменателя расположить по возрастающим степеням s , то в результате деления получим коэффициенты установившегося состояния; если же расположить члены по убывающим степеням s , то при делении получим коэффициенты начальных значений.

Фиг. 2. Типы систем.



	<p>Система с малой величиной коэффициента ошибки по скорости</p> $G = \frac{\omega_c(s + 0,153\omega_c) \cdot 1,23\omega_c}{s(s + 0,0306\omega_c)(s + 1,23\omega_c)}, \quad \zeta = 0,522,$ $\frac{X_e}{X_i} = \frac{s(s + 0,0306\omega_c)(s + 1,23\omega_c)}{(s + 0,1750\omega_c)(s^2 + 2\zeta\omega_n s + \omega_n^2)}, \quad \omega_n = 1,033\omega_c.$
	<p>Астатическая система второго порядка (с нулевым коэффициентом ошибки по скорости)</p> $G = \frac{s^2(s + 3\omega_c)}{\omega_c(s + 0,333\omega_c) \cdot 3\omega_c},$ $\frac{X_e}{X_i} = \frac{s^2(s + 3\omega_c)}{(s + \omega_c)^3}.$
	<p>Астатическая система второго порядка (с малым коэффициентом ошибки по ускорению)</p> $G = \frac{1,050(s + 0,103)(s + 0,299) \cdot 3,08}{s^2(s + 0,02)(s + 3,08)},$ $\frac{X_e}{X_i} = \frac{s^2(s + 0,02)(s + 3,08)}{(s + 0,100)(s + 1,00)^3} \cdot \omega_c = 1,050.$



Фиг. 3. Составляющие переходного процесса систем I и II типов (систем первого порядка).

4. Определение переходных процессов по ЛАХ'ам 151

С помощью уравнения (15) можно показать, что отличный от нуля коэффициент начальных значений низшего порядка равен

$$a_m = s^m \frac{X_u}{X_t} \Big|_{s=\infty} = s^m G \Big|_{s=\infty}, \quad (16)$$

а с помощью уравнения (14) можно показать, что отличный от нуля коэффициент установившегося состояния низшего порядка равен

$$c_n = s^n \frac{X_u}{X_t} \Big|_{s=0}. \quad (17)$$

Предполагая, что при нулевой частоте система имеет усиление, большее нуля, можно утверждать, что коэффициент реакции на ступенчатое воздействие c_0 никогда не равен нулю и, следовательно, всегда является отличным от нуля коэффициентом установившегося состояния низшего порядка для выходной реакции. С другой стороны, для отличного от нуля коэффициента ошибки низшего порядка имеется аналогичное соотношение

$$c'_n = s^n \frac{X_e}{X_t} \Big|_{s=0} \quad (18)$$

и коэффициент ошибки на ступенчатое воздействие c'_0 часто равен нулю.

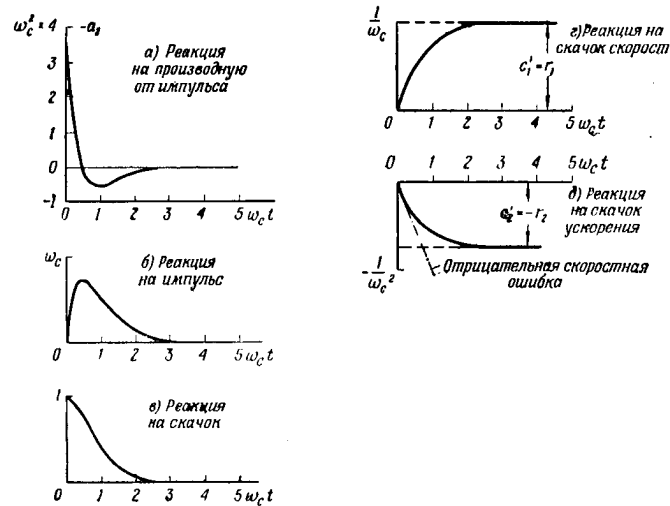
Роль составляющих переходного процесса, а также коэффициентов установившегося состояния и коэффициентов начальных значений не ограничивается тем, что они определяют реакцию на ограниченную группу входных сигналов; они описывают реакцию на любой вход. Как показано в статье [1], переходные процессы системы определяют ее общую временную характеристику в удобном для практики виде: имеется небольшая группа переходных процессов, которые являются важными для данной системы, и если все они известны, то каждый из них достаточно выразить лишь приближенно для установления реакции на любой входной сигнал.

Когда для этой группы переходных процессов используются составляющие переходного процесса, а также коэффициенты установившегося состояния и начальных значений, то реакция на любой вход может быть легко построена. Метод построения, развитый в статье [1], кратко изложен в Приложении I.

Группа переходных процессов, которая является важной для данной системы, неизменна независимо от входного сигнала, и ограничения, накладываемые на эту группу, зависят от характеристик передаточной функции системы. Так как настоящая статья посвящена исследованию переходных процессов отдельных систем, то указаны простые физические причины, по которым нет необходимости рассматривать другие переходные процессы. Дополнительные причины приведены в статье [1].

Исследования переходных процессов систем семи типов

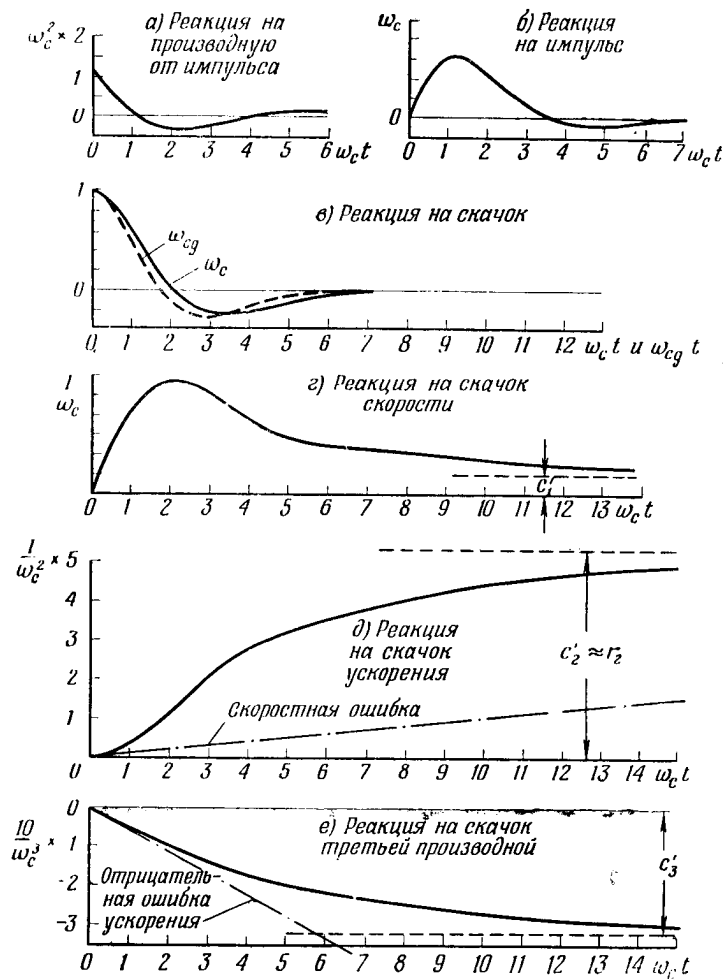
На фиг. 2 приведены семь различных типов передаточных функций системы. Переходные процессы вычислены



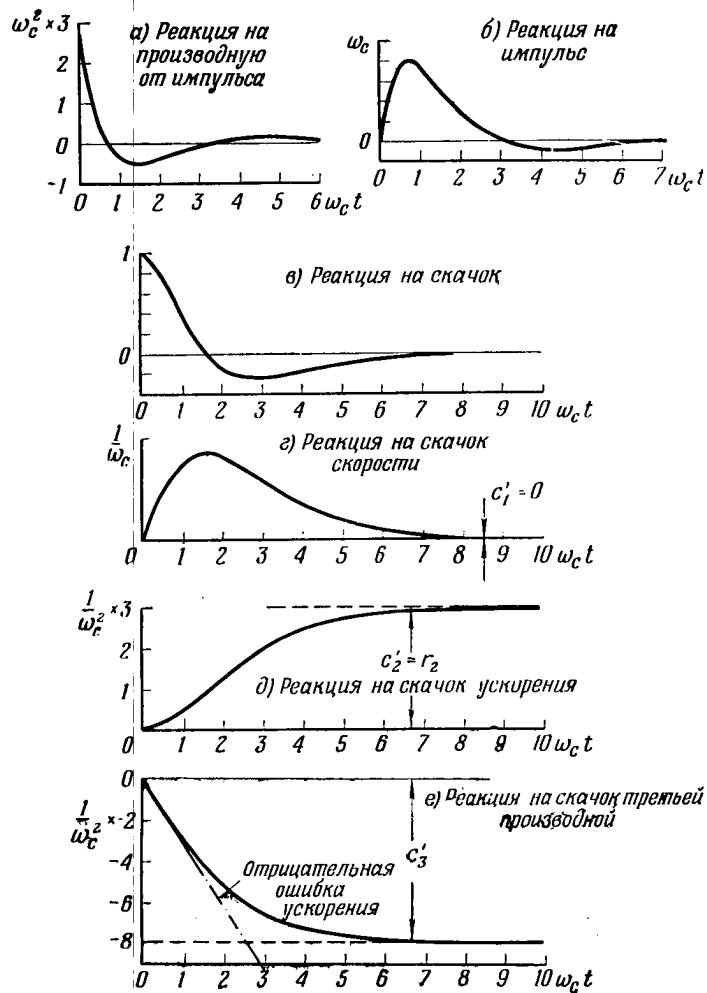
Фиг. 4. Составляющие переходного процесса системы III типа (системы с критическим демпфированием).

в Приложении II и построены на фиг. 3—8. На фиг. 2 даются также асимптотические амплитудные характеристики $|G(j\omega)|$ для каждой системы и выражения передаточных функций ошибки X_e/X_t .

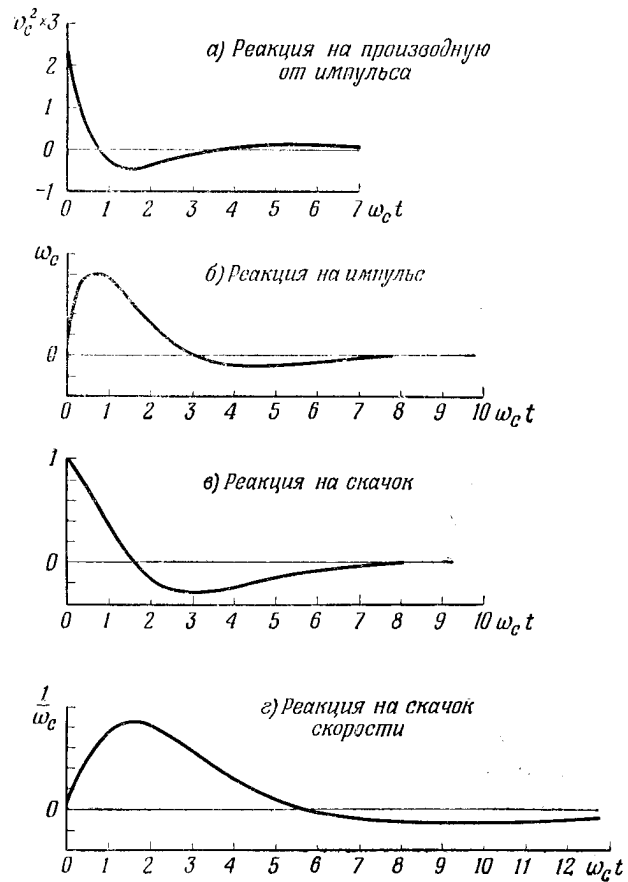
Системы выбраны так, чтобы показать влияние на переходные процессы следующих факторов: 1) интегрирующего



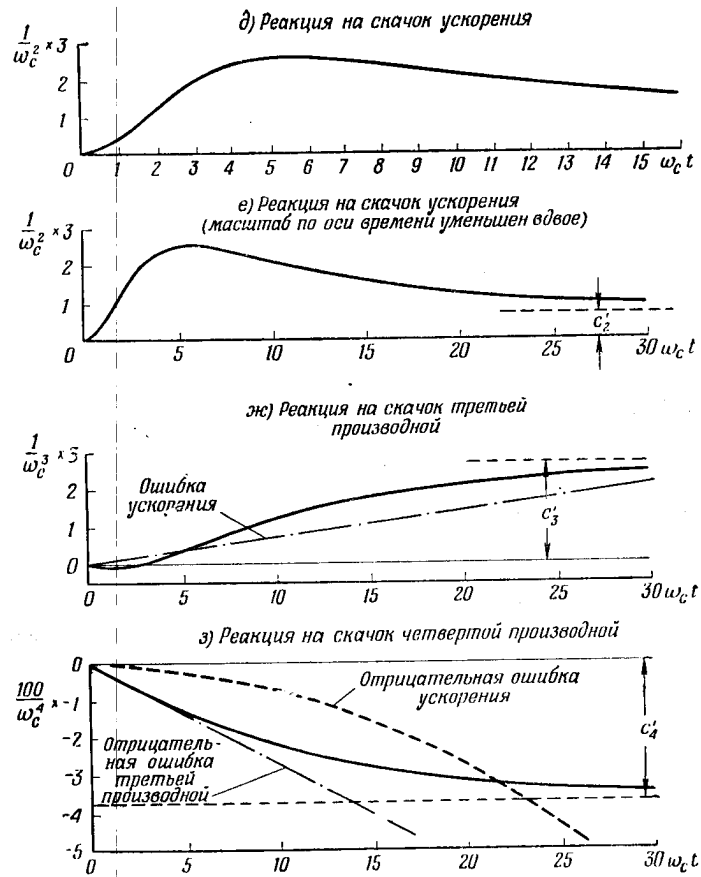
Фиг. 6. Составляющие переходного процесса системы V типа (системы с малым коэффициентом ошибки по скорости).



Фиг. 7. Составляющие переходного процесса системы VI типа (астатической системы второго порядка).



Фиг. 8. Составляющие переходного процесса системы VII типа (астатической системы второго порядка с малым коэффициентом ошибки по ускорению).



Фиг. 8. (Продолжение).

действия при низких частотах; 2) ширины полосы пропускания (частоты среза); 3) максимальных значений амплитудных частотных характеристик замкнутой системы.

Переходные характеристики строятся в безразмерных координатах в зависимости от асимптотической частоты среза ω_c . Для большинства систем частота ω_c почти равна действительной частоте среза ω_{cg} , но для системы IV типа $\omega_{cg} = 0,72 \omega_c$, а для системы V типа она равна $0,85 \omega_c$; следовательно, для этих двух систем некоторые переходные характеристики также безразмерны по отношению к ω_{cg} .

Кривые переходных характеристик выходной реакции строятся для скачкообразных изменений интегралов, а переходные процессы ошибки — для скачков производных и для ступенчатого возмущения на входе. Для систем I и II типов выходная реакция строится также и для ступенчатого воздействия на входе.

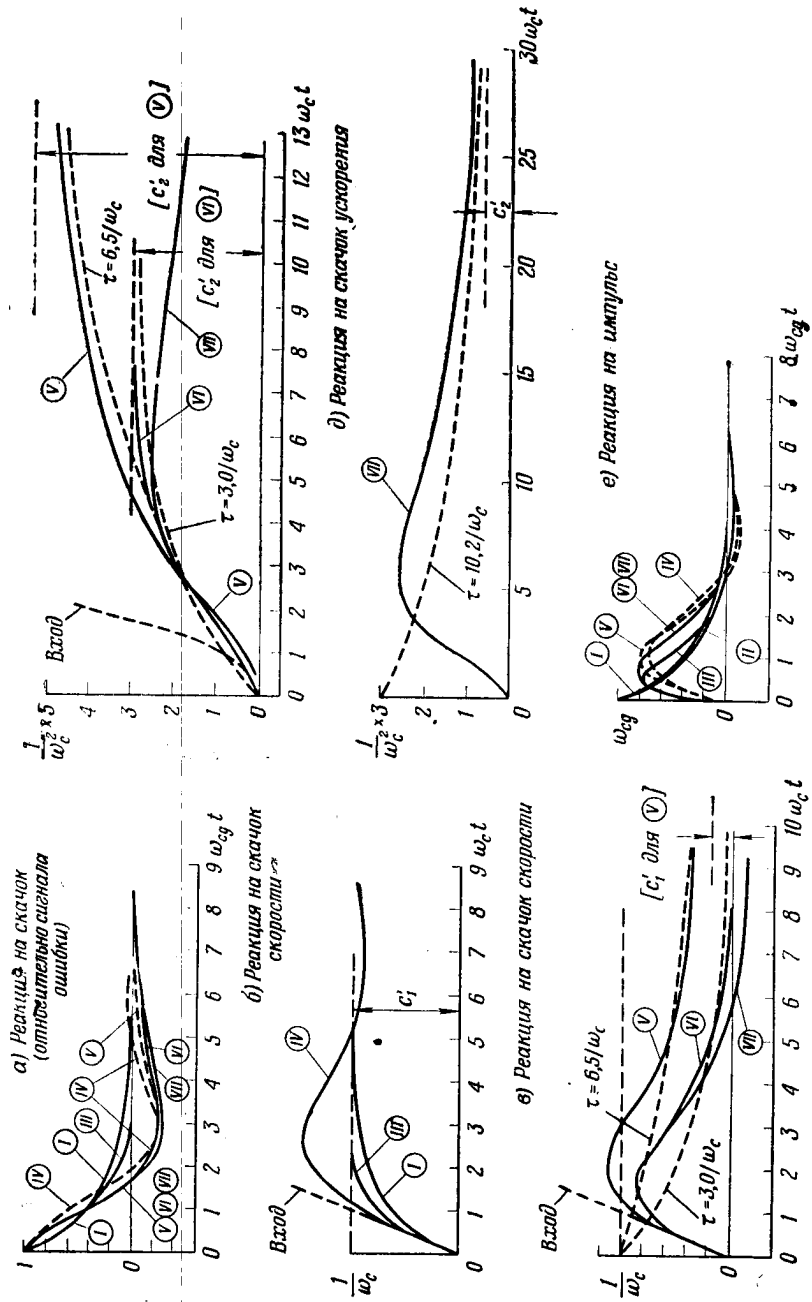
Реакция на ступенчатое возмущение

На фиг. 9, а показана кривая изменения ошибки системы при ступенчатом возмущении, построенная в безразмерном масштабе времени в зависимости от ω_{cg} для всех систем, за исключением системы II типа. Время, в течение которого ошибка системы при ступенчатом возмущении падает до 37% от своего начального значения (или, другими словами, время, в течение которого выходная реакция повышается примерно до 63% от своего конечного значения), называется временем нарастания при ступенчатом возмущении t_r . Из приведенного графика видно, что значение времени нарастания приближенно равно $1/\omega_{cg}$ для всех кривых. Действительные пределы значения t_r для различных переходных характеристик на фиг. 9, а равны

$$t_r = \frac{1,1}{\omega_{cg}} \pm 10\%. \quad (19)$$

Приближенная обратная зависимость между временем нарастания при ступенчатом воздействии и частотой среза является общей и очень надежной, и вследствие этого частоту среза можно рассматривать как хороший критерий ширины полосы пропускания.

Реакции на ступенчатое возмущение (фиг. 9, а) для систем IV, V и VI типов имеют 25%-ное, а система VII типа — 30%-ное



Фиг. 9. Сравнение соответственных составляющих переходного процесса.

перерегулирование. Если бы были построены кривые амплитудных частотных характеристик для замкнутых систем, то можно было бы показать, что эти характеристики для систем IV, V и VI типов имеют максимальное значение M_p , равное примерно 1,3, в то время как у системы VII типа имеется несколько большее значение. Таким образом, кривые доказывают хорошо известное соотношение, по которому значение M_p , по существу, определяет величину перерегулирования при ступенчатом воздействии.

В последние годы наблюдалась все возрастающая тенденция пренебрегать этим соотношением между M_p и максимальным перерегулированием, и вместо этого делались попытки выразить перерегулирование через коэффициент демпфирования так называемой доминирующей пары полюсов замкнутой системы. Часто предполагают, что для получения перерегулирования эта пара полюсов должна иметь коэффициент затухания меньше критического. Однако, как показано на фиг. 2, системы VI и VII типов не имеют таких полюсов; тем не менее реакция на ступенчатое возмущение имеет такое же перерегулирование, как и реакция системы IV типа, для которой коэффициент демпфирования равен 0,43. Таким образом, внешний вид кривых частотных характеристик может дать более верную информацию относительно максимального перерегулирования, чем значения полюсов передаточной функции замкнутой системы.

Все системы, кроме системы II типа, ведут себя при низких частотах как интегрирующее звено и, следовательно, имеют равный нулю коэффициент позиционной ошибки c'_0 .

Коэффициент позиционной ошибки c'_0 представляет собой конечное значение ошибки системы при ступенчатом возмущении и, как видно из фиг. 9, а, эти кривые для всех систем, кроме системы II типа, в установившемся состоянии стремятся к нулю. Система II типа имеет усиление при нулевой частоте, равное 9, так что ее коэффициент позиционной ошибки $c'_0 = 0,1$. Таким образом, кривая ошибки системы II типа при единичном скачке на входе стремится к конечному значению, равному 0,1 (фиг. 3, Ж).

Коэффициент установившегося состояния смещения выходной реакции для системы II типа равен

$$c_0 = 1 - c'_0 = 0,9. \quad (20)$$

4. *Определение переходных процессов по ЛАХ'ам* 161

Следовательно, выходная реакция системы II типа на единичное ступенчатое воздействие стремится к конечному значению 0,9 (фиг. 3, E). С другой стороны, для других систем коэффициент смещения выходной реакции c_0 равен единице. Следовательно, как показано на фиг. 3, B, для системы I типа выходные реакции на единичный скачок для всех систем, за исключением системы II типа, стремятся к конечному значению, равному единице, или, другими словами, выходной сигнал (сигнал обратной связи) равен в установившемся состоянии входному сигналу.

Реакция на скачок скорости

Для системы I типа коэффициент ошибки по скорости равен $1/\omega_c$. Следовательно, установившееся значение ошибки при единичном скачке скорости должно быть равно $1/\omega_c$. На фиг. 3, Г показана кривая изменения ошибки системы I типа при единичном скачке скорости входного сигнала, которая возрастает по экспоненциальному закону до конечного значения $1/\omega_c$.

Рассмотрим теперь, как изменится реакция системы при скачке скорости, когда для уменьшения коэффициента ошибки по скорости в систему вводится интегрирующая цепь. Рассмотрим систему VI типа, которая имеет коэффициент ошибки по скорости, равный нулю. Ее реакция на единичный скачок скорости приведена на фиг. 7, г. Сначала ошибка возрастает почти так же, как для системы I типа, как будто бы приближаясь к конечному значению $1/\omega_c$, но затем благодаря интегрирующему действию при низких частотах значение ошибки сводится к нулю. Однако интегрирующая цепь не дает заметного изменения максимального значения ошибки. Для любых систем максимальное значение ошибки при единичном скачке скорости является в основном функцией ширины полосы пропускания и равно приблизительно $1/\omega_{cg}$. Увеличивая в системе автоматического регулирования усиление при низких частотах, можно уменьшить ошибку после того, как она первоначально возникла, но, чтобы сохранить ошибку малой, с самого начала необходимо увеличивать ширину полосы пропускания.

Система V типа имеет интегрирующую цепь, которая не обеспечивает нулевого коэффициента ошибки по скорости,

но тем не менее этот коэффициент уменьшается благодаря наличию множителя, равного 5, до величины $1/5\omega_c$. На фиг. 6, з показана кривая изменения ошибки системы при единичном скачке скорости на входе. Максимальная ошибка близка к величине $1/\omega_{cg}$, которая несколько больше, чем $1/\omega_c$, поскольку для системы V типа ω_{cg} меньше, чем ω_c . Интегрирующая цепь понижает значение установившейся ошибки от $1/\omega_{cg}$ до значения $1/5\omega_c$. Сравнение этой характеристики с такой же характеристикой системы VI типа на фиг. 7, з показывает, что реакция на скачок скорости системы V типа, обладающая к тому же отличным от нуля конечным значением ошибки, имеет гораздо большее время установления. Причина этого состоит в том, что для системы V типа верхняя сопрягающая частота интегрирующей цепи, равная $0,153\omega_c$, вдвое меньше соответствующей сопрягающей частоты системы VI типа, равной $\omega_c/3$. Верхняя сопрягающая частота интегрирующей цепи определяет быстроту интегрирующего действия и, следовательно, время установления реакции на скачок скорости.

Система VII типа, как и система VI типа, имеет нулевой коэффициент ошибки по скорости, однако она в значительном интервале низких частот ведет себя как интегрирующее звено третьего порядка, что обуславливает относительно малое значение коэффициента ошибки по ускорению c'_2 . На фиг. 8, з приведена кривая изменения ошибки при единичном скачке скорости для этой системы. Несмотря на то, что ошибка в конечном счете достигает нуля, наличие у частотной характеристики участка с наклоном -30 длог/декада обуславливает появление длинного пологого участка кривой при отрицательных значениях ошибки, что увеличивает время установления.

Кривые изменения ошибки при скачке скорости для систем III и IV типов показаны на фиг. 4, з и 5, з. Кривая для системы III типа очень сходна с такой же кривой для системы I типа, а кривая изменения ошибки для системы IV типа имеет отчетливо выраженный колебательный характер, так как система обладает демпфированием ниже критического. Другое толкование характера реакции на скачок скорости для системы IV типа состоит в том, что ее максимальная ошибка почти равна $1/\omega_{cg}$, в то время как ее конечное значение равно $1/\omega_c$. Поскольку для системы IV типа ω_{cg} меньше,

4. Определение переходных процессов по ЛАХ'ам 163

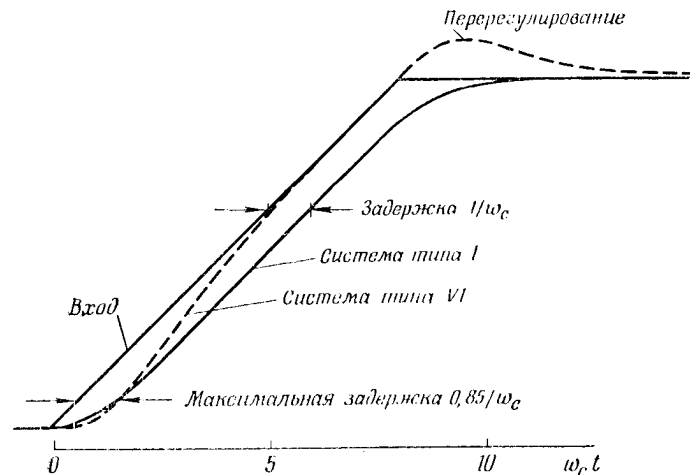
чем ω_c , ошибка должна переходить за величину $1/\omega_c$ до достижения установившегося значения.

Кривая изменения ошибки системы II типа при единичном скачке скорости показана на фиг. 3, И. Любопытная особенность этой характеристики состоит в том, что коэффициент ошибки по скорости c'_1 меньше, чем для системы I типа, несмотря на низкое усиление, которое имеет система II типа при низких частотах. Характеристики других систем, наоборот, показывают, что при большем усилении на низких частотах коэффициент ошибки по скорости меньше. Объяснение этого парадокса состоит в том, что кривая изменения ошибки при скачке скорости для системы II типа состоит из двух частей: экспоненциально возрастающей компоненты и установившейся составляющей позиционной ошибки, изменяющейся по линейному закону. Поскольку система II типа имеет установившуюся ошибку, пропорциональную величине входа, реакция на скачок скорости должна иметь установившуюся составляющую, равную коэффициенту ошибки по положению c'_0 , умноженному на величину скачка скорости на входе. Когда позиционная составляющая ошибки прибавляется к экспоненциальной составляющей ошибки по скорости, то полная ошибка при скачке скорости на входе будет всегда больше, чем для системы I типа, несмотря на то, что для системы II типа коэффициент ошибки по скорости меньше.

Желательно также исследовать выходную реакцию системы для скачка скорости на входе. На фиг. 10 показаны выходные реакции систем I и VI типов на „двойной“ скачок скорости, т. е. наклонный отрезок, начинающийся при $\omega_c t = 0$ и переходящий в горизонтальную прямую при $\omega_c t = 8$. Реакция на выходе системы I типа имеет по существу такую же форму, что и вход, только углы округляются и она запаздывает на время $1/\omega_c$. Для системы VI типа имеется начальное время запаздывания, почти равное $1/\omega_c$, но затем благодаря астатизму запаздывание сводится к нулю. Астатизм системы обуславливает сигнал, достаточно большой, чтобы компенсировать установившуюся ошибку по скорости, и, следовательно, в конце наклонного отрезка выходная реакция должна иметь перерегулирование и затем постепенно приближаться к установившемуся значению.

Для скачка скорости на входе мгновенное значение интервала времени, на которое выходная реакция запаз-

дывает относительно входа, равно мгновенному значению ошибки при единичном скачке скорости. Например, при приложении на входе системы VI типа единичного скачка скорости максимальная ошибка равна $0,85/\omega_c$ и, как показано на фиг. 10, максимальное время запаздывания выходной реакции на скачок скорости также равно $0,85/\omega_c$.



Фиг. 10. Реакция систем I и VI типов на скачок скорости.

Для доказательства этих общих соотношений рассмотрим вход с постоянной скоростью V

$$x_i(t) = Vt \quad (21)$$

и исследуем выходную реакцию, которую можно выразить следующим образом:

$$x_u(t) = x_i(t) - x_e(t) = Vt - x_e(t). \quad (22)$$

Вычислим $x_u(t)$ в данный момент времени t_2 и определим момент времени t_1 , когда вход, заданный уравнением (21), будет равен этому значению $x_u(t)$:

$$x_u(t_2) = Vt_2 - x_e(t_2) = x_i(t_1) = Vt_1. \quad (23)$$

Для интервала времени $(t_2 - t_1)$ мы получим

$$(t_2 - t_1) = \frac{x_e(t_2)}{V}. \quad (24)$$

Значение $(t_2 - t_1)$ представляет собой интервал времени, на который выходной сигнал в момент t_2 запаздывает относительно входа. Уравнение показывает, что время запаздывания для скачка скорости на входе (любой амплитуды) равно ошибке при единичном скачке скорости.

Поскольку максимальная ошибка при единичном скачке скорости приближенно равна времени нарастания при скачке, то максимальное время запаздывания, на которое реакция на скачок скорости отстает от входа, приближенно равно времени нарастания реакции на скачок. В дальнейшем это приближение можно будет распространить на случай произвольного входного сигнала. Выходная реакция системы I типа на произвольный вход равняется по существу входу, сдвинутому на время нарастания при ступенчатом воздействии, равное $1/\omega_c$, за исключением тех компонент, которые возрастают и убывают в интервале меньшем, чем время запаздывания. Для астатических систем второго порядка, таких, как система VI типа, реакция не такая простая, так как интегрирующее действие приводит к постоянному накоплению сигналов для компенсации установившихся ошибок, обусловленных составляющими с постоянной скоростью (иногда ускорения), и при этом имеет место перерегулирование всякий раз, когда эти компоненты исчезают. Из-за этого интегрирующего действия не имеется постоянного запаздывания между входом и выходной реакцией; тем не менее всякий раз, когда вход претерпевает излом, появляется запаздывание, которое по существу равно времени нарастания реакции на скачок.

Вообще характернейшей чертой реакции любой системы, являющейся фильтром низких частот, является время запаздывания, обратное пропорциональное ширине полосы пропускания. Запаздывание является неотъемлемым свойством и проявляется всякий раз, когда вход имеет резкие переходы. С другой стороны, можно выбрать низкочастотную часть передаточной функции так, чтобы при наличии составляющих постоянной скорости и, возможно, постоянного ускорения во входном сигнале путем изменения формы реакции

системы получить малые ошибки в течение интервалов времени, когда эти составляющие по существу постоянны.

Важно знать ограничения, свойственные методу уменьшения ошибки системы автоматического регулирования с помощью интегрирующего эффекта на низких частотах. Когда появляется составляющая постоянной скорости во входном сигнале, интегрирующая цепь должна „выждать“, пока появится ошибка, прежде чем начать ее корректировать; и после того, как эта составляющая исчезнет, цепь вызывает появление в реакции перерегулирования и пологого участка, которые появляются также при входном ступенчатом воздействии, как это видно на графиках реакций на скачок для систем V, VI и VII типов (см. фиг. 6, в, 7, в и 8, в).

Таким образом, двойное интегрирование может существенно уменьшить ошибку системы при входном сигнале, имеющем значительные скоростные составляющие, которые поддерживаются по существу постоянными в течение интервалов времени, намного превышающих время установления реакции на скачок скорости; но если скорость не является постоянной, двойное интегрирование может даже увеличить результирующую ошибку. Кроме того, в некоторых случаях, когда система следит за программным входом, бывает необходимо, чтобы форма выходной кривой была подобна форме входного сигнала хотя бы и с запаздыванием выхода относительно входа. Поскольку двойное интегрирование уменьшает установившуюся ошибку за счет искажения выходной кривой, то оно нежелательно там, где можно допустить запаздывание во времени.

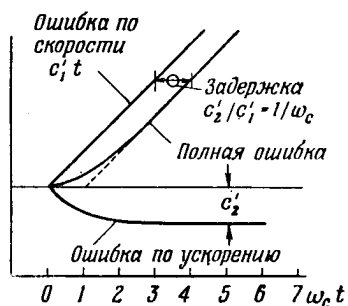
Реакция на скачок ускорения

Реакция на скачок ускорения является важной динамической характеристикой систем V, VI и VII типов, обладающих астатизмом и, следовательно, малыми установившимися ошибками по скорости. Однако для других систем установившиеся ошибки по скорости так велики по сравнению с составляющими ошибки по ускорению, что нет смысла принимать во внимание реакцию на скачок ускорения.

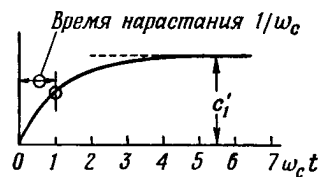
Для системы I типа сплошной кривой на фиг. 3, Д показано изменение ошибки по ускорению при единичном скачке

ускорения. Непривычно, что эта составляющая ошибки отрицательна (направлена в обратном направлении по отношению ко входу) и до момента времени $\omega_c t = 1$ она по амплитуде больше входного сигнала. С другой стороны, когда ускорение претерпевает скачок, скорость, а также и установившаяся составляющая ошибки должны возрастать по линейному закону, причем ошибка равна скорости, умноженной на коэффициент ошибки по скорости c'_1 . Эта положительная составляющая ошибки по скорости для удобства сравнения отложена на фиг. 3, Д в отрицательном направлении и обозначена как „отрицательная скоростьная ошибка“. Из кривой видно, что составляющая ошибки по скорости всегда больше по амплитуде, чем составляющая ошибки по ускорению. Это показывает, что ошибка по ускорению не имеет большого значения.

На фиг. 11, а показано, как составляющая ошибки по ускорению системы I типа на единичный скачок ускорения вычитается из возрастающей по линейному закону составляющей ошибки по скорости. Кривая результирующей ошибки является почти такой же, как для ошибки по скорости, но она запаздывает относительно нее на время c'_2/c'_1 , равное $1/\omega_c$. На фиг. 11, б показана реакция на скачок скорости, из которой видно, что время запаздывания составляющей ошибки по скорости, обусловленное ошибкой по ускорению, равно времени нарастания кривой ошибки при скачке скорости. Таким образом, составляющая ошибки по скорости запаздывает на время, равное времени нарастания ошибки при скачке скорости таким же



а



б

Фиг. 11. Влияние составляющей ошибки по ускорению для системы I типа.

а — ошибка системы при скачке ускорения; б — ошибка системы при скачке скорости.

образом, как и сама реакция при скачке скорости запаздывает на время, равное времени нарастания реакции на ступенчатое воздействие.

Обобщая этот результат, можно утверждать, что влияние составляющей ошибки по ускорению при произвольном входе в системе с монотонным изменением ошибки при скачке скорости (сюда относятся системы I, II и III типов) сводится по существу к задержке скоростной составляющей ошибки на время, равное времени нарастания ошибки при скачке скорости. Таким образом, составляющая ошибки по ускорению для этих систем, а фактически для всех систем с астатизмом первого порядка, не должна рассматриваться как основная составляющая ошибки, а лишь как поправка к скоростной составляющей ошибки. Если при вычислениях пренебрегают составляющей по ускорению, то при этом не должно заметно изменяться максимальное значение ошибки, а сдвигается лишь момент времени, при котором оно имеет место.

Реакции на скачок ускорения систем II и III типов являются по существу такими же, как и для системы I типа; несколько отличается только реакция для системы IV типа с демпфированием, меньшим критического. Однако составляющая ошибки по ускорению является совсем малой по сравнению с установившейся скоростной составляющей (за исключением коротких интервалов времени) и, следовательно, имеет малое значение. Из-за малого демпфирования для системы IV типа оценить ее реакцию более трудно, чем в других случаях. Всякий раз, когда прилагается возмущение, система IV типа имеет тенденцию колебаться. Однако, если система хорошо задемпфирована, эти колебания слабо влияют на реакцию, хотя и ограничивают точность, достигаемую при грубых приближениях.

Составляющая ошибки по ускорению имеет малое значение для астатических систем первого порядка. Однако когда для достижения малого значения коэффициента ошибки по скорости в значительном интервале частот ниже частоты среза используется интегрирование второго порядка, то установившиеся скоростные ошибки настолько уменьшаются, что по сравнению с ними ошибки по ускорению становятся значительными. Следовательно, для систем V, VI и VII типов реакция на скачок ускорения очень существенна.

4. *Определение переходных процессов по ЛАХ'ам* 169

Кривая изменения ошибки для системы VI типа при единичном скачке ускорения показана на фиг. 7, д. Она монотонно возрастает до конечного значения, равного коэффициенту ошибки по ускорению, и кривая, вдоль которой она возрастает, имеет по существу ту же форму, что и пологий участок реакции на скачок скорости на фиг. 7, г. Поскольку коэффициент ошибки по скорости для системы VI типа равен нулю, то в реакции на скачок ускорения отсутствуют составляющие ошибки по скорости.

Реакция системы V типа на скачок ускорения, приведенная на фиг. 6, д, имеет составляющую ошибки по скорости, изменяющуюся линейно, и составляющую ошибки по ускорению. Однако составляющая ошибки по скорости мала по сравнению с ошибкой по ускорению на достаточно большом интервале времени. Это указывает на то, что для системы V типа при произвольном входе составляющая ошибки по ускорению может быть заметной.

Реакция системы VII типа на скачок ускорения, приведенная на фиг. 8, д и е, показывает, как введение интегрирования третьего порядка обеспечивает малый коэффициент ошибки по ускорению. Ошибка по величине почти достигает значения коэффициента ошибки по ускорению, который имеют системы без интегрирования третьего порядка, но затем благодаря интегрирующему действию она медленно уменьшается до конечного значения, равного действительному коэффициенту ошибки по ускорению ϵ_2' . Время установления переходного процесса является функцией верхней сопрягающей частоты участка характеристики с тройным наклоном (0,103 на фиг. 2), которая для обеспечения достаточной устойчивости должна быть значительно меньше частоты среза. Следовательно, время переходного процесса для системы с участком интегрирования третьего порядка при скачке ускорения должно быть намного больше, чем время нарастания реакции на ступенчатое воздействие. В действительности это время при скачке ускорения для системы VII типа так велико, что для того, чтобы показать характер протекания процесса, график пришлось перестроить в меньшем масштабе времени.

Обычно в системах автоматического регулирования введение участка интегрирования третьего порядка мало используется для уменьшения установившихся ошибок. Если даже

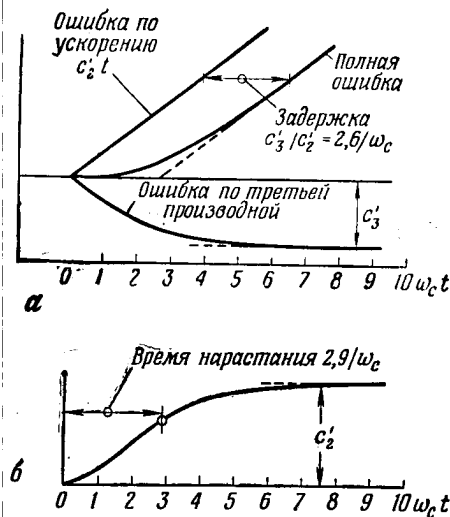
и можно уменьшить коэффициент ошибки по ускорению без увеличения ширины полосы пропускания, то это даст уменьшение ошибки только для составляющих по ускорению, сохраняемых в течение значительно большего времени по сравнению с временем нарастания реакции на скачок. И если скоростные составляющие входа системы автоматического регулирования часто остаются постоянными в течение времени, значительно большего, чем время нарастания реакции на скачок, то большая составляющая постоянного ускорения в течение длительного периода встречается очень редко. Если система автоматического регулирования и имеет характеристики с участком интегрирования третьего порядка, то обычно такой участок вводится не для уменьшения коэффициента ошибки по ускорению. Иногда, например, характеристика с участком интегрирования третьего порядка встречается потому, что в систему для увеличения ее жесткости при преодолении момента нагрузки добавляется интегрирующая цепь.

Важным недостатком такого рода систем является то, что у них к кривой изменения ошибки при скачке скорости добавляется пологий отрицательный участок, как это показано на фиг. 8, 2, который сильно затягивает время установления реакции на скачок скорости. Вследствие этого введение интегрирования третьего порядка часто приводит к увеличению низкочастотных ошибок.

Реакция на скачок третьей производной

Для астатических систем первого порядка реакции на скачки третьих производных значения не имеют и, следовательно, рассматриваться не будут. Реакции систем с участком интегрирования второго порядка (системы V и VI типов) на скачок третьей производной показаны на фиг. 6, e и 7, e. Из кривых видно, что составляющие ошибки по третьей производной малы по сравнению с установившимися составляющими ошибки по ускорению, изменяющимися по линейному закону, за исключением коротких промежутков времени, и, следовательно, ошибка по третьей производной должна рассматриваться лишь как поправка к ошибке по ускорению. На фиг. 12, a показано, как составляющая ошибки по третьей производной для реакции системы VI типа на скачок третьей

производной вычитается из составляющей ошибки по ускорению. Результирующая ошибка равна по существу ошибке по ускорению, запаздывающей на интервал времени c'_3/c'_2 , равный $2,6/\omega_c$. Это время запаздывания приблизительно равно времени нарастания кривой изменения ошибки при скачке ускорения, приведенной на фиг. 12, б, равному в свою



Фиг. 12. Влияние составляющей ошибки по третьей производной для системы VI типа.

а — ошибка системы при скачке третьей производной; б — ошибка системы при скачке ускорения.

очередь $2,9/\omega_c$. Таким образом, для систем с участком двойного интегрирования существенное влияние составляющей ошибки по третьей производной сводится к задержке составляющей ошибки по ускорению; время этого запаздывания приблизительно равно времени нарастания кривой ошибки при скачке ускорения.

Реакция системы V типа на скачок третьей производной, приведенная на фиг. 6, е, показывает только две из трех составляющих. Кроме построенных графиков переходной составляющей по третьей производной и линейно возрастающей составляющей по ускорению, реакция содержит также

параболическую составляющую ошибки по скорости, равную коэффициенту ошибки по скорости, умноженному на результирующую параболу скорости.

В случае системы VII типа наличие участка с интегрированием третьего порядка понижает коэффициент ошибки по ускорению настолько, что придает значительный вес реакции на скачок третьей производной. Эта реакция (фиг. 8, ж) состоит из двух частей: переходной составляющей по третьей производной и линейно возрастающей составляющей ошибки по ускорению.

Переходная составляющая возрастает очень медленно, по существу с таким же временем установления, как и у пологого участка реакции на скачок ускорения (фиг. 8, е).

Скачок четвертой производной

Система VII типа является единственной системой, для которой имеет значение реакция на скачок четвертой производной, и даже в этом случае она не очень существенна. Кривая изменения ошибки построена на фиг. 8, з и состоит из трех частей: переходной части, линейно изменяющейся ошибки по третьей производной и параболической части ошибки по ускорению. Две установившиеся составляющие ошибки положительны, но построены в отрицательном направлении. Для входа общего вида существенное влияние составляющей ошибки по четвертой производной сводится к задержке составляющей ошибки по третьей производной на интервал времени c'_4/c'_3 , равный $14/\omega_c$ и приблизительно соответствующий времени нарастания реакции на скачок третьей производной.

Реакция на импульс

Для определения реакции системы автоматического регулирования на высокочастотные сигналы на входе, такие, как сигналы шума, необходимо исследовать импульсную переходную характеристику, а по возможности — и реакции на первую производную от импульса, на вторую производную, на третью производную и т. д. Это соответствует реакциям на скачок интегралов от входа: импульс соответствует скачку первого интеграла, производная от импульса — скачку второго интеграла и т. д. Как было показано, наи-

4. Определение переходных процессов по ЛАХ'ам 173

больший порядок интегрирования в интервале частот ниже частоты среза (максимальный наклон характеристики) определяет число реакций на скачки производных, которые имеют значение. Аналогично наибольший порядок интегрирования в интервале частот выше частоты среза, т. е. максимальная крутизна амплитудной характеристики при высоких частотах, указывает, сколько реакций на скачки интегралов имеют значение.

Согласно уравнению (16), начальное значение импульсной переходной характеристики систем, которое представляется коэффициентом начальных значений a_1 , равно

$$a_1 = sG|_{s=\infty}. \quad (25)$$

Для систем I и II типов передаточная функция G асимптотически приближается при высоких частотах к ω_c/s и, следовательно, коэффициент начальных значений a_1 равен ω_c . Передаточные функции других систем при высоких частотах ведут себя как интегрирующее звено второго порядка и, следовательно, для них коэффициент начального значения a_1 равен нулю. Таким образом, импульсная переходная характеристика для систем I и II типов начинается при значении ω_c , как показано на фиг. 3, А и Е, тогда как для других систем импульсная переходная характеристика берет начало в нуле, как показано на фиг. 4—8 (б).

Импульсные характеристики всех систем на фиг. 9, е построены как функции относительного времени в зависимости от частоты среза ω_{cg} . Это сравнение показывает, что независимо от начальных значений импульсных характеристик их максимальные значения равны по существу ω_{cg} . Экстремальное значение импульсной переходной характеристики является количественной мерой величины возмущения системы шумом, так как импульс действует так же, как и сигнал шума на систему. Таким образом, системы с одинаковой ω_{cg} можно рассматривать как имеющие одинаковую ширину полосы пропускания, имея в виду их фильтрующее действие на сигналы шума, так как импульсные характеристики всех систем почти одинаковы.

Понятие ширины полосы можно рассматривать с двух сторон: система следует за частотными составляющими, лежащими в полосе пропускания, и фильтрует более высокочастотные составляющие. Время нарастания реакции на

скачок является хорошим показателем свойств системы по отношению к низкочастотным сигналам, так как оно показывает, как быстро система следует за низкочастотными составляющими входа, а максимальное значение импульсной переходной характеристики является хорошим показателем свойств системы по отношению к высокочастотным сигналам, так как оно показывает, как сильно возмущается система нежелательным высокочастотным шумом. Частота ω_{cg} , таким образом, выдерживает оба испытания как критерий для ширины полосы пропускания. Системы с одинаковой ω_{cg} имеют приблизительно одинаковое время нарастания и примерно одно и то же максимальное значение импульсной переходной характеристики.

Возможно, что максимальное значение импульсной характеристики системы будет намного больше ω_{cg} . Это имеет место, когда усиление системы очень слабо затухает при частотах выше ω_{cg} , что может быть вызвано нулями передаточной функции разомкнутой системы при высоких частотах. Как и следовало ожидать, такая система возбуждается шумом гораздо сильнее, чем система с такой же частотой среза и более быстрым затуханием при высоких частотах, но за входным сигналом она не следит сколь угодно точно. Максимальное значение импульсной характеристики у такой системы больше, однако время нарастания реакции на скачок не уменьшается.

Реакция на производную от импульса

Для систем I и II типов реакция на производную от импульса включает в себя импульс амплитуды a_1 . Хотя эта характеристика и имеет некоторое значение при определении реакции системы на обычные высокочастотные входы, все же она не столь важна по сравнению с составляющей, связанной с коэффициентом начального значения импульсной характеристики a_1 . Поэтому реакция на производную от импульса для этих двух систем не строилась.

Поскольку для систем III и IV типов коэффициент a_1 равен нулю, коэффициент начального значения реакции на производную от импульса для этих систем можно определить из выражения

$$a_2 = s^2 G \Big|_{s=\infty}. \quad (26)$$

Эту величину можно вычислить, определив по фиг. 2 уравнение для высокочастотной асимптоты. Например, высокочастотная асимптота системы V типа равна $1,23 \omega_0^2/s^2$, так что коэффициент начального значения a_2 должен равняться $1,23 \omega_0^2$. Это соответствует начальному значению реакции на производную от импульса, как показано на фиг. 6, а. Реакции на производные от импульсов систем III—VII типов приведены на фиг. 4—8 (а).

Реакция на вторую производную от импульса

Для систем I и II типов реакция на вторую производную не представляет интереса, однако для других систем она имеет некоторое значение. Для систем I и II типов реакция на вторую производную от импульса содержит импульс с амплитудой a_2 и производную от импульса с амплитудой a_1 , в то время как реакция на вторую производную от импульса для других систем содержит еще импульс амплитуды a_2 . Для произвольного высокочастотного входа системы часть переходного процесса, связанная с реакцией на вторую производную от импульса, обычно мала по сравнению с частью, обусловленной коэффициентом начального значения a_2 , но она все же может быть существенна.

Правила для приближенного определения переходных характеристик

Сравнивая переходные характеристики систем, можно вывести общие правила для приближенного определения переходных характеристик. В основе многих из этих правил лежит соотношение между частотными характеристиками разомкнутой системы и полюсами передаточной функции замкнутой системы.

Значения полюсов передаточной функции замкнутой системы

Поскольку значения полюсов передаточной функции замкнутой системы определяют постоянные времени различных составляющих переходного процесса, приближенное знание полюсов передаточной функции замкнутой системы необходимо для определения формы переходных характеристик.

Как показано в статье [2], указанные полюсы приближенно равны:

- 1) нулям передаточной функции разомкнутой системы, соответствующим частотам, меньшим частоты ω_c ;
- 2) полюсам передаточной функции разомкнутой системы, соответствующим частотам, большим частоты ω_c ;
- 3) полюсу, равному примерно $-\omega_c$.

Рассматриваемые правила аппроксимирования будут приложены лишь к реакции на ступенчатое воздействие и к реакции на скачки производных, поскольку эти характеристики обычно гораздо существеннее при расчете систем автоматического регулирования, чем реакции на скачки интегралов. Форма переходных составляющих этих реакций определяется в основном полюсами при средних и низких частотах. Однако при оценке формы переходных составляющих необходимо знать значение полюсов при средних частотах, так как по частоте среза и максимальному значению (M_p) амплитудной частотной характеристики замкнутой системы можно определить важнейшие свойства этих полюсов. Таким образом, остается оценить полюсы при низких частотах, а, как показано в статье [1], они приближенно равны нулям передаточной функции разомкнутой системы при низких частотах.

Действительные полюсы передаточной функции замкнутой системы при низких частотах несколько сдвигаются относительно нулей передаточной функции разомкнутой системы, и величина сдвига отдельного нуля приближенно равна обратной величине усиления системы $|G|$ при сопрягающей частоте этого нуля. Другими словами, чем больше усиление системы, тем меньше сдвиг. С другой стороны, можно получить хорошие приближения, пренебрегая сдвигами и предполагая, что полюсы передаточной функции замкнутой системы при низких частотах действительно равны нулям передаточной функции разомкнутой системы. Это допущение положено в основу вывода правил для приближенного определения переходных характеристик.

Вывод правил приближенного определения характеристик

В последующих разделах вновь исследуются переходные характеристики системы, чтобы найти способы их приближенного определения.

Реакция на скачок скорости

Ошибки систем при единичном скачке скорости сравниваются на фиг. 9, б и в. Для всех систем ошибка следует достаточно близко за входом до тех пор, пока не достигает максимального значения, примерно равного $1/\omega_{ср}$, после чего ошибка уменьшается до значения, равного коэффициенту ошибки по скорости. Для астатических систем, состоящих из одного интегрирующего звена, коэффициент ошибки по скорости равен $1/\omega_c$, и поскольку ω_c обычно очень близка к $\omega_{ср}$, конечное значение ошибки приближенно равно максимальному значению.

На фиг. 9, в показана кривая изменения ошибки на скачок скорости для систем с интегрирующим действием выше первого порядка. Для таких систем конечная величина ошибки значительно меньше максимальной величины. Поскольку система V типа имеет при низких частотах нуль с сопрягающей частотой $0,153 \omega_c$, то у передаточной функции замкнутой системы должен быть полюс с примерно такой же сопрягающей частотой. Этот полюс должен давать в переходном процессе составляющую с постоянной времени, равной примерно $1/0,153 \omega_c$, или $6,5/\omega_c$. На фиг. 9, в пунктирной линией от значения $1/\omega_c$ до конечного значения, равного для системы V типа коэффициенту ошибки по скорости c'_1 , построена кривая, представляющая собой экспоненту с этой постоянной времени. Из графика видно, что эта экспонента является достаточно хорошим приближением пологого участка реакции системы V типа на скачок скорости. Таким образом, значение нуля разомкнутой системы при низких частотах определяет количественно форму пологого участка реакции на скачок скорости.

Для систем VI и VII типов первый нуль ниже частоты среза имеет сопрягающую частоту, примерно равную $\omega_c/3$, и, следовательно, пологий участок кривой изменения ошибки системы на скачок скорости для каждой из них должен изменяться по экспоненте с постоянной времени $3/\omega_c$. Такая экспонента показана пунктирной кривой на фиг. 9, в, изменяющейся от $1/\omega_c$ до нуля. Реакция системы VI типа достаточно хорошо совпадает с этой экспонентой, тогда как реакция системы VII типа имеет дополнительную отрица-

тельную экспоненту, обуславливающую перерегулирование и пологий отрицательный участок кривой. Эта дополнительная экспонента соответствует полюсу замкнутой системы, полученному из нуля при низкой частоте с сопрягающей частотой 0,13 (или около 0,1 ω_c). Таким образом, интегрирующее действие третьего порядка вызывает появление в реакции на скачок скорости системы VII типа небольшой, но продолжительной отрицательной экспоненты с постоянной времени, примерно равной $1/0,1 \omega_c$, или $10/\omega_c$.

Реакция на скачок ускорения

На фиг. 9, 2 показаны переходные составляющие реакций систем V, VI и VII типов на единичный скачок ускорения. Переходная составляющая для системы V типа очень близка к экспоненте с постоянной времени $6,5/\omega_c$, возрастающей до конечного значения, равного коэффициенту ошибки по ускорению c'_2 . Аналогично для системы VI типа переходная составляющая близка к экспоненте с постоянной времени $3/\omega_c$, возрастающей до конечного значения, равного ее коэффициенту ошибки по ускорению. Для системы VII типа переходная составляющая возрастает по существу так же, как и для системы VI типа, и почти достигает значения $3/\omega_c^2$. Это выражение является коэффициентом ошибки по ускорению для системы VI типа и коэффициентом ошибки, который имела бы система VII типа, если бы у нее отсутствовало интегрирующее действие третьего порядка. Однако интегрирующее действие третьего порядка в конечном счете обуславливает переходную характеристику, изменяющуюся по экспоненте с постоянной времени, равной примерно $10/\omega_c$, до конечного значения, равного действительному коэффициенту ошибки по ускорению.

Реакция на скачок третьей производной

Переходная составляющая реакции системы VII типа на скачок третьей производной, приведенная на фиг. 8, ж, аппроксимируется экспонентой с постоянной времени $10/\omega_c$, возрастающей до конечного значения, равного коэффициенту ошибки по третьей производной.

Реакция на скачок

Кривые изменения ошибки системы при ступенчатом входе для всех систем, кроме системы II типа, показаны на фиг. 9, а. Все кривые ошибки, в том числе и для системы II типа, уменьшаются до 37% от своего начального значения за время, равное примерно $1/\omega_{ср}$. Величина перерегулирования является в основном функцией величины M_p , которая представляет собой максимальное значение амплитудной частотной характеристики замкнутой системы. При $M_p = 1,3$ перерегулирование составляет примерно 25%.

Для систем V, VI и VII типов с малым коэффициентом ошибки по скорости интегрирующее действие второго порядка добавляет к реакции на скачок отрицательную экспоненциальную составляющую с постоянной времени, примерно равной верхней сопрягающей частоте соответствующего участка характеристики. Эта экспонента определяет большую часть перерегулирования и добавляет пологую составляющую. Для системы VII типа интегрирование третьего порядка также добавляет к кривой изменения ошибки небольшой положительный пологий участок. Этот пологий участок имеет слишком малую амплитуду, чтобы его можно было заметить на графике; тем не менее он может иметь значение вследствие большой продолжительности. Поскольку верхняя сопрягающая частота системы с интегрированием третьего порядка приближенно равна $\omega_c/10$, то этот пологий участок имеет постоянную времени, примерно равную $10/\omega_c$, что в 10 раз больше, чем время нарастания реакции на ступенчатое воздействие.

Соотношение между последовательными переходными составляющими

Как видно из уравнения (13), составляющие переходного процесса при скачках производных связаны интегральной зависимостью. Поскольку коэффициент c_n равен $-T_n(0)$, это соотношение можно выразить в виде определенного интеграла

$$T_n(t) + c_n = \int_0^t T_{n-1}(t) dt \quad \text{для } n \geq 0. \quad (27)$$

Таким образом, интеграл от составляющей переходного процесса равен составляющей от скачка последующей производной более высокого порядка. Это соотношение можно использовать для объяснения значения характеристик некоторых переходных процессов.

Уравнение показывает, что ошибка системы при скачке скорости равняется интегралу от ошибки системы при скачке. Если конечное значение ошибки системы при скачке скорости намного меньше ее максимального значения, то интеграл от ошибки при скачке на входе должен быть приблизительно равен нулю или, другими словами, отрицательная площадь под кривой изменения ошибки при скачке на входе должна быть примерно равна положительной площади. Таким образом, чтобы получить малый коэффициент ошибки по скорости, реакция системы на скачок должна иметь перерегулирование. Для заданного времени нарастания реакции на скачок положительная площадь под кривой ошибки, по существу, фиксирована (за исключением систем с таким колебательным переходным процессом, когда имеются отрицательные перерегулирования); тем самым фиксируется и площадь перерегулирования, что справедливо для систем с малым коэффициентом ошибки по скорости. Поскольку площадь перерегулирования неизменна, то реакция системы с малым коэффициентом ошибки по скорости на скачок должна иметь либо большой выброс, либо медленно затухающий пологий участок. Длину этого участка можно устанавливать, меняя верхнюю сопрягающую частоту характеристики системы с интегрированием второго порядка. Однако если для уменьшения перерегулирования удлинять пологий участок реакции на скачок, то пологий участок реакции на скачок скорости также удлиняется, и проходит больше времени, пока не проявится влияние малого коэффициента ошибки по скорости.

Аналогично, если для уменьшения коэффициента ошибки по ускорению в системе используется интегрирование третьего порядка, то реакция на скачок скорости имеет перерегулирование и площадь перерегулирования должна примерно равняться положительной площади (как показано для системы VII типа на фиг. 8, 2). Таким образом, при интегрирующем действии третьего порядка кривая изменения ошибки при скачке скорости должна иметь либо большое перерегулиро-

4. Определение переходных процессов по ЛАХ'ам 181

вание, либо пологий участок кривой. Последний требует наличия такого же участка в реакции на скачок ускорения, что обуславливает большую задержку, прежде чем будут получены все выгоды от уменьшения коэффициента ошибки по ускорению.

Выводы

В статье рассмотрены некоторые важные частотные характеристики систем автоматического регулирования и показано их влияние на форму переходных процессов и реакций на более общие сигналы на входе. Кроме общего обсуждения этих вопросов, в статье дается несколько правил приближенного определения переходных характеристик с помощью исследования частотных кривых.

Краткое изложение правил приближенного определения переходных характеристик

Правила приближенного определения переходных характеристик систем автоматического регулирования для единичного скачка и скачков производных еще раз кратко сформулированы ниже.

Реакция на скачок

Время нарастания реакции на скачок, определяемое как время, необходимое для достижения уровня, составляющего 63% от конечного значения, приблизительно равно $1/\omega_{ср}$.

Максимальное перерегулирование определяется величиной M_p , причем для $M_p = 1,3$ перерегулирование составляет 25%.

Если система является астатической, то установившаяся ошибка будет равна нулю; для статической системы ошибка равна позиционному коэффициенту ошибки.

Интегрирование второго порядка обуславливает появление в кривой ошибки при скачке на входе отрицательного экспоненциального участка с постоянной времени, равной обратной величине верхней сопрягающей частоты участка с двойным наклоном.

Участок характеристики с тройным наклоном (интегрирование третьего порядка) добавляет небольшой, но более

пологий положительный экспоненциальный участок кривой с постоянной времени, равной обратной величине верхней сопрягающей частоты этого участка.

Реакция на скачок скорости

Для астатической системы первого порядка ошибка при скачке скорости на входе приблизительно следует линейному изменению входа до достижения максимальной величины $1/\omega_{cg}$ и после этого приближается к конечной величине, равной $1/\omega_c$.

Если система не является астатической, реакция по существу такая же, за исключением того, что имеется дополнительная линейная составляющая ошибки, равная позиционному коэффициенту ошибки, умноженному на значения входа постоянной скорости.

Для системы, имеющей участок частотной характеристики с двойным наклоном (интегрирование второго порядка), ошибка возрастает до максимальной величины $1/\omega_{cg}$ так же, как и для астатической системы первого порядка, но затем убывает по экспоненте до величины, соответствующей действительному значению коэффициента ошибки по скорости. Постоянная времени для этой экспоненты равна обратной величине верхней сопрягающей частоты участка характеристики с двойным наклоном.

Для системы с участком частотной характеристики с тройным наклоном кривая изменения ошибки такая же, как и для системы, имеющей участок характеристики с двойным наклоном, за исключением того, что имеется отрицательная пологая экспонента, постоянная времени которой равна обратной величине верхней сопрягающей частоты участка с тройным наклоном.

Реакция на скачок ускорения

Для системы, имеющей участок частотной характеристики с двойным наклоном, переходная составляющая по ускорению возрастает по экспоненте с постоянной времени, равной обратной величине сопрягающей частоты характеристики этого участка и конечным значением, равным коэффициенту ошибки по ускорению.

Для системы с частотной характеристикой, имеющей тройной наклон (интегрирование третьего порядка), переходная составляющая по ускорению сначала возрастает так, как если бы система имела характеристику с двойным наклоном, и почти достигает ошибки, равной значению коэффициента ошибки по ускорению, который имела бы эта система. Но затем ошибка убывает по экспоненте от этого максимального значения до действительной величины коэффициента ошибки по ускорению, и постоянная времени экспоненты равна обратной величине верхней сопрягающей частоты участка характеристики с тройным наклоном.

Реакция на скачок третьей производной

Для системы, имеющей участок частотной характеристики с тройным наклоном, переходная составляющая по третьей производной представляет собой возрастающую экспоненту с постоянной времени, равной обратной величине верхней сопрягающей частоты этого участка характеристики, и конечным значением, равным коэффициенту ошибки по третьей производной.

Когда эти правила применяются к системам, о которых говорится, что они имеют интегрирование второго или третьего порядков, то это не означает, что эти системы являются астатическими второго или третьего порядка. Система с интегрированием второго порядка является системой, которая имеет в разомкнутом состоянии амплитудную характеристику с наклоном — 20 длог/декада в значительном интервале частот ниже частоты среза, в то время как система с интегрированием третьего порядка имеет характеристику с наклоном — 30 длог/декада в значительной области частот.

В случае интегрирования второго порядка обычно имеется один нуль передаточной функции разомкнутой системы ниже ω_c и сопрягающая частота этого нуля представляет собой верхнюю сопрягающую частоту отрезка характеристики с наклоном — 20 длог/декада. Аналогично в случае интегрирования третьего порядка имеются два нуля передаточной функции разомкнутой системы ниже ω_c и наивысшая сопрягающая частота, соответствующая одному из этих нулей,

является верхней сопрягающей частотой отрезка с наклоном — 20 длог/декада, в то время как наименьшая сопрягающая частота равна верхней сопрягающей частоте асимптоты с наклоном — 30 длог/декада.

Степень общности приближенных вычислений

Выше предполагалось, что правила для приближенного определения переходных характеристик применимы в общем случае. Однако это фактически не было доказано. Тем не менее были представлены достаточные обоснования для их использования в практических расчетах. В конце концов хорошо известное соотношение между M_p и максимальным перерегулированием при скачке на входе никогда не доказывалось, хотя и было очень полезным при расчетах.

С другой стороны, в настоящее время автором подготовлена статья, посвященная общему методу последовательных приближений для расчета переходных процессов, основанному на использовании данных асимптотических частотных характеристик. Этот метод обосновывает рекомендации для получения приближений, связывающих частотные характеристики с переходными процессами, и поскольку метод допускает увеличение точности, расчетчик имеет готовые средства для определения достаточности приближений. Метод поможет обосновать приближения, представленные в этой статье, и, что более важно, обеспечит большую точность правил приближенного определения и более ясное толкование динамических принципов систем автоматического регулирования.

В конкретных случаях расчетчик может оценивать точность приближения путем вычислений или исследования с помощью моделирующих устройств. Тем не менее это не может заменить приближенного определения переходных характеристик, так как этот путь является слишком сложным, чтобы обеспечить достаточную основу для сознательного расчета. Если расчетная задача не является очень сложной и имеется на выбор большое количество вариантов ее решения, они, тем не менее, не устраняют необходимости ясного понимания основных динамических свойств системы, а такое понимание должно базироваться главным образом на надежном приближении.

ПРИЛОЖЕНИЕ I

Краткое изложение метода определения реакции на произвольный входной сигнал

В этом приложении кратко излагается метод определения реакции системы на произвольный входной сигнал, рассматриваемый в статье [1]. Метод основан на разложении реакции системы на составляющие, каждая из которых вызывается скачками производной и интеграла от входа.

Разложение выходной реакции на скачок n -й производной от входа, данное в уравнении (2), можно также представить в виде

$$x_u(t) = c_0 x_i(t) + c_1 D x_i(t) + c_2 D^2 x_i(t) + \dots \\ \dots + [c_n + T_n(t)] D^n x_i(t), \quad (28)$$

тогда как реакцию на скачок m -го интеграла, данную уравнением (3), можно записать следующим образом:

$$x_u(t) = a_0 x_i(t) + a_1 D^{-1} x_i(t) + \dots \\ \dots + a_{m-1} D^{-(m-1)} x_i(t) + T_{-m}(t) D^{-m} x_i(t). \quad (29)$$

Выражение $D^n x_i(t)$ обозначает n -ю производную от входного сигнала $x_i(t)$, а $D^{-m} x_i(t)$ — m -кратный интеграл.

Из этих выражений вытекает определенный порядок расчета реакции на произвольный входной сигнал. Прежде всего, входной сигнал представляют в виде суммы скачков производных и интегралов от функции на входе. Для этого производятся следующие операции:

1. Приближенно представляют скачками те участки входного сигнала, где он изменяется быстрее реакции системы на ступенчатое возмущение (скачок).
2. Высокочастотные колебания входного сигнала, возрастающие и убывающие значительно быстрее, чем реакция системы на скачок, заменяются осредненной кривой.
3. Устранив скачки и высокочастотные составляющие, дифференцируют остающуюся сглаженную кривую.
4. Представляют приближенно с помощью скачков те части продифференцированной кривой, которые изменяются более резко, чем составляющие переходного процесса для соответствующей производной, и вновь дифференцируют остающуюся сглаженную кривую.

5. Выполняют действия, указанные в п. 4, несколько раз, до тех пор пока не получат определенную производную, при которой кривая полностью заменяется приблизительно прямолинейными отрезками. Поскольку производная этого приближения целиком состоит из скачков, то этим приближением процесс заканчивается. Порядок производной, при котором достигается окончательная аппроксимация кривой ломаной линией, зависит только от передаточной функции системы и не зависит от входа (при условии, что вход не является таким простым, когда аппроксимация заканчивается производной более низкого порядка).

6. Высокочастотную составляющую входа, которая была отделена, необходимо несколько раз проинтегрировать и каждый последующий интеграл аппроксимировать скачками, чтобы выделить постоянные составляющие, длительность которых превышает время нарастания соответствующей составляющей переходного процесса системы.

7. Процесс интегрирования можно прекратить на определенном интеграле, порядок которого также зависит только от передаточной функции системы, и влиянием интегралов более высоких порядков можно пренебречь.

Затем строим кривую реакции, суммируя следующие три слагаемых:

- 1) кривые производных, умноженные на соответствующие коэффициенты установившихся состояний;
- 2) интегральные кривые, умноженные на соответствующие коэффициенты начальных значений;
- 3) соответствующие составляющие переходного процесса для каждого скачка производной или интеграла, пропорциональные величине этих скачков.

Нетрудно установить вполне удовлетворительную простую оценку точности вычисления реакции. Погрешность, вызванная аппроксимированием отдельных дифференциальных или интегральных кривых, не может быть больше произведения полученного от умножения максимального отклонения приближенной кривой от точной на соответствующий коэффициент остатка. Коэффициент остатка r_n для n -й производной равен

$$r_n = \int_0^{\infty} |T_{n-1}(t)| dt \quad (30)$$

и легко может быть вычислен посредством графического построения составляющей переходного процесса $T_n(t)$. Для m -кратного интеграла коэффициент остатка r_{-m} вычисляется аналогичным образом по $T_{-m}(t)$.

Важная черта этого метода состоит в том, что имеются определенные указания, где нужно остановить процесс дифференцирования и интегрирования, что зависит только от передаточной функции системы, и поэтому для данной системы имеется лишь фиксированная группа переходных процессов, которые необходимо знать. Ниже перечислены ограничения, накладываемые на эту группу.

Наивысшая дифференциальная составляющая переходного процесса, которую еще следует принимать во внимание, является первой составляющей, для которой сумма $[c_n + T_n(t)]$ не больше по амплитуде, чем составляющая $c_{n-1}t$, обусловленная коэффициентом установившегося состояния более низкого порядка.

Наивысшей интегральной составляющей переходного процесса будет составляющая, следующая за тем первым слагаемым переходного процесса, для которого начальное значение равно его максимальному значению по амплитуде.

Процесс дифференцирования прекращается при полной аппроксимации кривой наивысшей производной, осуществляемой шаг за шагом путем аппроксимирования прямолинейными отрезками кривой производной порядка на единицу ниже. У этой кривой производной более низкого порядка коэффициент ошибки обычно почти равен по величине коэффициенту остатка, так что точность вычисленной ошибки системы равна по существу точности прямолинейной аппроксимации.

ПРИЛОЖЕНИЕ II

Вычисление составляющих переходных процессов для семи систем

В этом приложении выводятся все коэффициенты установившихся состояний и начальных значений, все составляющие переходных процессов для реакций семи систем, приведенных выше,

Система	c'_0	c'_1	c'_2	c'_3	c'_4	a_1	a_2	a_3
I	0	$1/\omega_c$	$1/\omega_c^2$			ω_c	$-\omega_c^2$	
II	0,1	$0,81/\omega_c$	$-0,729\omega_c^2$			ω_c	$-1,111\omega_c^2$	
III	0	$1/\omega_c$	$-3/4\omega_c^2$			0	$4\omega_c^2$	$-16\omega_c^3$
IV	0	$1/\omega_c$	$0,346/\omega_c^2$			0	$0,74\omega_c^2$	$-0,547\omega_c^3$
V	0	$1/5\omega_c$	$5,35/\omega_c^2$	$-32/\omega_c^3$		0	$1,23\omega_c^2$	$-1,36\omega_c^3$
VI	0	0	$3/\omega_c^2$	$-8/\omega_c^3$		0	$3\omega_c^2$	$-8\omega_c^3$
VII	0	0	$0,69/\omega_c^2$	$26,6/\omega_c^3$	$-374/\omega_c^4$	0	$2,93\omega_c^2$	$-7,53\omega_c^3$

Таблица 1
КОЭФФИЦИЕНТЫ УСТАНОВИВШИХСЯ СОСТОЯНИЙ И НАЧАЛЬНЫХ ЗНАЧЕНИЙ ДЛЯ СЕМИ СИСТЕМ

Коэффициенты

Величины коэффициентов установившихся состояний и начальных значений для семи систем указаны в табл. 1. Они вычислены с помощью разложения передаточной функции системы по правилу деления многочленов, описанному в приложении I статьи [1].

Составляющие переходного процесса для систем V, VI и VII типов

Уравнения для составляющих переходного процесса систем V, VI и VII типов приведены ниже. Уравнения для других систем составлены в следующем разделе для иллюстрации процесса вычисления переходных составляющих.

СИСТЕМА V ТИПА

Выражение для переходных составляющих ошибки системы будет

$$T'_n(t) = K_{1n} e^{-0,175\omega_c t} + |2K_{2n}| e^{-\zeta\omega_n t} \sin(\omega_0 t + \varphi_n), \quad (31)$$

где

$$\zeta\omega_n = 0,54\omega_c, \quad (32)$$

$$\omega_0 = 0,88\omega_c. \quad (33)$$

Коэффициенты этого переходного процесса сведены в табл. 2. Эта система совпадает с системой, рассмотренной в работе [1]; вычисление составляющих переходного процесса приведено в приложении к работе [1].

Таблица 2

КОЭФФИЦИЕНТЫ СОСТАВЛЯЮЩИХ ПЕРЕХОДНОГО ПРОЦЕССА
ДЛЯ СИСТЕМЫ V ТИПА

n	K'_{1n}	$ 2K'_{2n} $	φ'_{2n}
-2	-0,00507 ω_c^2	2,88 ω_c^2	-57,3°
-1	0,0290 ω_c	2,79 ω_c	-178,8°
0	0,166	2,70	59,7°
1	0,949/ ω_c	2,61/ ω_c	-61,8°
2	-5,42/ ω_c^2	2,53/ ω_c^2	176,7°
3	30,9/ ω_c^3	2,46/ ω_c^3	55,2°

СИСТЕМА VI ТИПА

Выражение для переходных составляющих ошибки системы имеет вид

$$T'_n(t) = e^{-\omega_c t} (K_{1n} + K_{2n}t + K_{3n}t^2). \quad (34)$$

Коэффициенты указаны в табл. 3.

Таблица 3

КОЭФФИЦИЕНТЫ СОСТАВЛЯЮЩИХ ПЕРЕХОДНОГО ПРОЦЕССА
ДЛЯ СИСТЕМЫ VI ТИПА

n	K_{1n}	K_{2n}	K_{3n}
-2	$3\omega_c^2$	$-5\omega_c^3$	ω_c^4
-1	0	$3\omega_c^2$	$-\omega_c^3$
0	1	ω_c	$-\omega_c^2$
1	0	1	ω_c
2	$-3/\omega_c^2$	$-3/\omega_c$	-1
3	$8/\omega_c^3$	$5/\omega_c^2$	$1/\omega_c$

СИСТЕМА VII ТИПА

Выражение для переходных составляющих ошибки системы имеет вид

$$T'_n(t) = e^{-0,95\omega_c t} (K_{1n} + K_{2n}t + K_{3n}t^2) + K_{4n}e^{-0,95\omega_c t}. \quad (35)$$

Коэффициенты приведены в табл. 4.

Таблица 4

КОЭФФИЦИЕНТЫ СОСТАВЛЯЮЩИХ ПЕРЕХОДНОГО ПРОЦЕССА
ДЛЯ СИСТЕМЫ VII ТИПА

n	K_{1n}	K_{2n}	K_{3n}	K_{4n}
-2	$2,94\omega_c^2$	$-4,74\omega_c^3$	$0,929\omega_c^4$	$0,000311\omega_c^2$
-1	$0,00311\omega_c$	$2,93\omega_c^2$	$0,973\omega_c^3$	$-0,00311\omega_c$
0	0,967	$0,922\omega_c$	$-1,023\omega_c^2$	0,0327
1	$0,343/\omega_c$	1,293	$1,075\omega_c$	$-0,343/\omega_c$
2	$-4,34/\omega_c^2$	$-3,73/\omega_c$	-1,13	$3,66/\omega_c^2$
3	$11,20/\omega_c^3$	$6,51/\omega_c^2$	$1,18/\omega_c$	$-37,8/\omega_c^3$
4	$-21,8/\omega_c^4$	$-9,35\omega_c^3$	$-1,27/\omega_c^2$	$400/\omega_c^4$

4. Определение переходных процессов по ЛАХ'ам 191

Составляющие переходного процесса для систем I—IV типов

СИСТЕМА I ТИПА

Для системы I типа передаточная функция ошибки (см. фиг. 2) равна

$$\frac{X_e}{X_i} = \frac{s}{s + \omega_c}. \quad (36)$$

Для единичного скачка на входе $X_i(s) = 1/s$, а $X_e(s)$ равна позиционной составляющей переходного процесса $T'_0(s)$

$$X_e(s) = T'_0(s) = \frac{1}{s} \frac{X_e(s)}{X_i(s)} = \frac{1}{s + \omega_c}. \quad (37)$$

Обратное преобразование дает

$$x_e(t) = T'_0(t) = e^{-\omega_c t}. \quad (38)$$

Уравнение (11) показывает, что полюс первого порядка переходной составляющей при скачке n -й производной можно выразить следующим образом:

$$T'_n(t) = \sum_{l=1}^L K_{ln} e^{s_l t} = \sum_{l=1}^L \frac{K_{l0}}{s_l^n} e^{s_l t}, \quad (39)$$

где K_{l0} — коэффициент реакции на скачок для полюса s_l и K_{ln} — соответствующий коэффициент реакции на скачок n -й производной. Следовательно, общее уравнение для составляющей переходного процесса на скачок n -й производной системы I типа запишется в виде

$$T'_n(t) = \frac{e^{-\omega_c t}}{(-\omega_c)^n} = (-1)^n \frac{e^{-\omega_c t}}{\omega_c^n}. \quad (40)$$

СИСТЕМА II ТИПА

$$\frac{X_e}{X_i} = \frac{s + 0,11\omega_c}{s + 1,11\omega_c}. \quad (41)$$

Для единичного ступенчатого воздействия на входе

$$X(s) = \frac{c'_0}{s} + T'_0(s) = \frac{(s + 0,11\omega_c)}{s(s + 1,11\omega_c)} = \frac{0,1}{s} + \frac{0,9}{s + 1,11\omega_c}. \quad (42)$$

Обратное преобразование имеет вид

$$x_c(t) = c'_0 + T'_0(t) = 0,1 + 0,9e^{-1,11\omega_c t}. \quad (43)$$

Тогда общее уравнение для составляющей переходного процесса на скачок n -й производной запишется в виде

$$T'_n(t) = \frac{0,9e^{-1,11\omega_c t}}{(-1,11\omega_c)^n} = (-1)^n 0,9 \left[\frac{0,9}{\omega_c} \right]^n e^{-1,11\omega_c t}. \quad (44)$$

СИСТЕМА III ТИПА

$$\frac{X_e}{X_i} = \frac{s(s+4\omega_c)}{(s+2\omega_c)^2}. \quad (45)$$

Для единичного скачка на входе имеем

$$X_e(s) = T'_0(s) = \frac{(s+4\omega_c)}{(s+2\omega_c)^2} = \frac{1}{s+2\omega_c} + \frac{2\omega_c}{(s+2\omega_c)^2}. \quad (46)$$

Обратное преобразование таково:

$$T'_0(t) = (1 + 2\omega_c t) e^{-2\omega_c t}. \quad (47)$$

Поскольку эта система имеет полюс второго порядка, то уравнение (39) применять нельзя. Вместо него следует использовать более общее соотношение уравнения (13), которое применимо ко всем передаточным функциям и равно

$$T_n(t) = \int T_{n-1}(t) dt. \quad (48)$$

Таким образом, составляющие переходного процесса от скачков производных получаются интегрированием, а составляющие от скачков интегралов — дифференцированием. Для интегрирования используются следующие равенства:

$$\int e^{-u} du = -e^{-u}, \quad (49)$$

$$\int ue^{-u} du = -e^{-u}(1+u), \quad (50)$$

$$\int u^2 e^{-u} du = -e^{-u}(2+2u+u^2). \quad (51)$$

4. Определение переходных процессов по ЛАХ'ам 193

Уравнения (49) и (50) дают составляющие переходного процесса

$$T_1'(t) = \int T_0'(t) dt = \frac{-e^{-2\omega_c t}}{\omega_c} (1 + \omega_c t), \quad (52)$$

$$T_2'(t) = \int T_1'(t) dt = \frac{3}{4\omega_c^2} e^{-2\omega_c t} \left(1 + \frac{2\omega_c t}{3}\right), \quad (53)$$

Уравнение (51) необходимо при рассмотрении системы с полюсом третьего порядка. Для получения переходных составляющих на скачок интеграла выполняем дифференцирование:

$$T_{-1}(t) = -T_{-1}'(t) = -\frac{d}{dt} T_0'(t) = (2\omega_c)(2\omega_c t) e^{-2\omega_c t}, \quad (54)$$

$$T_{-2}(t) = \frac{d}{dt} T_{-1}(t) = (2\omega_c)^2 e^{-2\omega_c t} (1 - 2\omega_c t). \quad (55)$$

СИСТЕМА IV ТИПА

$$\frac{X_e}{X_t} = \frac{s(s + 0,74\omega_c)}{s^2 + 2\zeta\omega_n s + \omega_n^2}, \quad (56)$$

где $\zeta = 0,43$ и $\omega_n = 0,86\omega_c$. Для единичного скачка на входе $X_e(s)$ равна

$$X_e(s) = T_0'(s) = \frac{K_0}{(s + \zeta\omega_n - j\omega_0)} + \frac{K_0^*}{(s + \zeta\omega_n + j\omega_0)}, \quad (57)$$

где

$$K_0 = \frac{\omega_n}{2\omega_0} e^{j(\theta - 90^\circ)}, \quad (58)$$

$$\theta = \arccos \zeta, \quad (59)$$

$$\omega_0 = \omega_n \sqrt{1 - \zeta^2}. \quad (60)$$

Значение полюса в верхней полуплоскости

$$-\zeta\omega_n + j\omega_0 = -\omega_n e^{-j\theta}. \quad (61)$$

Следовательно, коэффициент K_n по уравнению (39) равен

$$K_n = \frac{K_0}{(-\zeta\omega_n + j\omega_0)^n} = \frac{K_0}{(-\omega_n e^{-j\theta})^n}. \quad (62)$$

Подставляя в уравнение (62) выражение для K_0 из уравнения (58), получим

$$K_n = \frac{(-1)^n (\omega_n)^{1-n}}{2\omega_0} e^{j[(n+1)\theta - 90^\circ]}. \quad (63)$$

Составляющая переходного процесса на скачок n -й производной равна

$$T'_n(s) = \frac{K_n}{s + \zeta\omega_n - j\omega_0} + \frac{K_n^*}{s + \zeta\omega_n + j\omega_0}. \quad (64)$$

Найдем обратное преобразование и, подставив выражение для K_n из уравнения (63), получим

$$T'_n(t) = \frac{(-1)^n (\omega_n)^{1-n}}{\omega_0} e^{-\zeta\omega_n t} \sin[\omega_0 t + (n+1)\theta]. \quad (65)$$

Подставляя значения параметров, будем иметь

$$T'_n(t) = \frac{(-1)^n 1,108}{(0,86\omega_c)^n} e^{-0,37\omega_c t} \sin[0,776\omega_c t + (n+1)64,5^\circ]. \quad (66)$$

ЛИТЕРАТУРА

1. Biernson G. A. *Trans. AIEE*, Part II, **74**, 227 (1955).
2. Biernson G. A. *Trans. AIEE*, Part II, **74**, 53 (1953).
3. Biernson G. A., "A Simple Method for Calculating the Time Response of a System to an Arbitrary Input", Servomechanisms Laboratory, Mass. Inst. Technol., Rep. № 7138-R-3, Jan. 20, 1954. ASTIA AD-25082.

**II. ПРИМЕНЕНИЕ
ЧАСТОТНЫХ МЕТОДОВ
К РАСЧЕТУ НЕЛИНЕЙНЫХ
СИСТЕМ**

5. ОБОБЩЕННЫЙ МЕТОД ОПРЕДЕЛЕНИЯ ЧАСТОТНЫХ ХАРАКТЕРИСТИК ЗАМКНУТЫХ НЕЛИНЕЙНЫХ СИСТЕМ¹⁾

Л. Принц

Введение

Частотный анализ нелинейных систем автоматического регулирования с обратной связью представляет собой результат развития известного метода расчета линейных систем. Поскольку поведение нелинейных систем не может описываться обыкновенными линейными дифференциальными уравнениями с постоянными коэффициентами, то для них не справедлив принцип суперпозиции и, следовательно, частотный анализ не имеет смысла. Однако с помощью частотного анализа можно получить важные сведения о нелинейных системах, как это показано в работах Кохенбургера [1, 2], Джонсона [3] и др.

Наиболее широко применяемый метод основан на представлении нелинейных элементов с помощью описывающих функций, которые при использовании частотной характеристики разомкнутой системы облегчают анализ и синтез.

Благодаря использованию описывающих функций можно определить, имеется ли опасность возникновения незатухающих колебаний или неустойчивости и как таких явлений можно избежать. К сожалению, не имеется общего метода определения частотной характеристики замкнутой системы; следовательно, для замкнутой системы можно, вообще говоря, получить количественные сведения лишь при рассмотрении устойчивости. Благодаря аппроксимированию нелинейных элементов с помощью эквивалентного коэффициента усиления можно так упростить математические уравнения, что удастся легко определить частотные характеристики многих замкнутых нелинейных систем.

Частотный метод является специальным методом, который полезен при изучении поведения некоторого класса нели-

¹⁾ Prince L. T., Jr., *Trans. AIEE*, Part II, 73, 217—224 (1954),

нейных систем, причем определение частотных характеристик замкнутой нелинейной системы может рассматриваться как дополнение к существующему частотному методу. Ценность метода можно установить лишь с помощью применения его к конкретным задачам.

Эквивалентный коэффициент усиления

Применение частотного метода к анализу нелинейных систем зависит, вообще говоря, от законности двух предположений: 1) гармониками высокого порядка на выходе нелинейного элемента можно пренебречь вследствие фильтрующего действия системы автоматического регулирования; 2) входной сигнал, подаваемый на нелинейный элемент, можно считать синусоидальным.

На основе этих предположений обычно представляют нелинейный элемент приближенной передаточной функцией для того, чтобы можно было применить операторный метод ко всей системе. Тогда многие методы линейного анализа можно распространить на нелинейные системы ¹⁾.

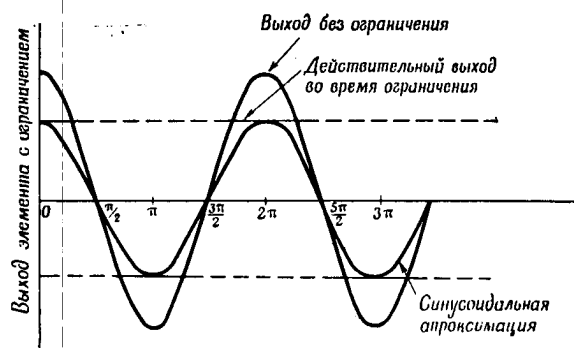
Понятие эквивалентного коэффициента усиления сходно с понятием описывающей функции, но обладает некоторыми преимуществами при анализе замкнутых систем. Описывающая функция определяется как комплексное отношение амплитуды основной гармоники на выходе к амплитуде синусоидального сигнала на входе нелинейного элемента. Комплексное отношение максимального значения выхода к максимальному значению входа нелинейного звена характеризует эквивалентный коэффициент усиления. Различие между терминами „нелинейный элемент“ и „нелинейное звено“ состоит в том, что первый содержит в себе второе плюс линейные элементы. Чтобы упростить нелинейные уравнения, обычно заменяют истинные характеристики нелинейного элемента кусочно-линейными. Это упрощение обычно проводят

¹⁾ Рассматриваемый здесь метод является методом эквивалентной или гармонической линеаризации. Основы этого метода заложены в работах Н. Н. Крылова и Н. И. Боголюбова и в дальнейшем развиты советскими учеными (см., например, Попов Е. П., Динамика систем автоматического регулирования, Гостехиздат, 1954). — *Прим. ред.*

5. Частотные характеристики замкнутых нелинейных систем 199

для того, чтобы определить простое аналитическое выражение, связывающее переменные нелинейной системы¹⁾.

При использовании частотного анализа определение эквивалентного коэффициента усиления указывает на существование фиктивной синусоиды на выходе нелинейного элемента, имеющей тот же период, что и синусоида на входе, и то же амплитудное значение, что и истинный выход нелинейного элемента, как это показано на фиг. 1. Строгое аналитическое обоснование этого определения в этой статье не дается,



Фиг. 1. Синусоидальные формы выходного сигнала нелинейного элемента с ограничением.

однако для проверки ошибки, вводимой такой аппроксимацией, были получены решения с помощью моделей. Показем, как можно определить частотную характеристику замкнутой нелинейной системы, представляя нелинейный элемент функцией эквивалентного коэффициента усиления.

ОБОЗНАЧЕНИЯ

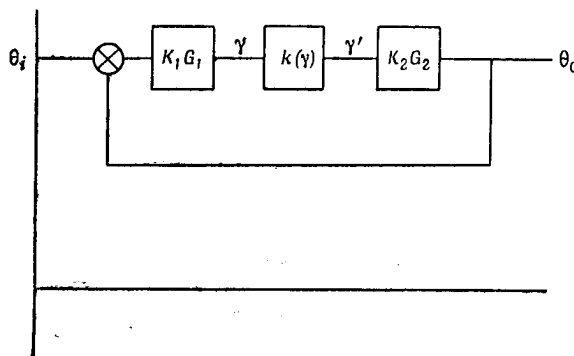
- K — коэффициент усиления;
- $G(s)$ — передаточная функция линейной части, зависящая от частоты;
- $k(\gamma)$ — эквивалентный коэффициент усиления;
- γ — вход нелинейного элемента;

¹⁾ В работе Я. З. Цыпкина (*Автоматика и телемеханика*, 17, № 4 (1956)) указан упрощенный метод вычисления эквивалентного коэффициента усиления нелинейного звена по его характеристике. — *Прим. ред.*

- γ' — выход нелинейного элемента;
 θ_i — управляющее воздействие;
 θ_o — выход системы;
 $\dot{\theta}_o$ — первая производная выхода системы;
 $\ddot{\theta}_o$ — вторая производная выхода системы;
 s — переменная Лапласа,
 $f[k(\gamma)]$ — функция эквивалентного коэффициента усиления;
 $\ddot{\theta}_o$ — максимально возможное значение $\ddot{\theta}_o$;
 ω — угловая частота [рад/сек];
 ω_{n0} — собственная частота системы при работе на линейном участке характеристики;
 ζ_0 — коэффициент демпфирования системы при работе на линейном участке характеристики;
 $u = \omega/\omega_{n0}$ — безразмерная угловая частота;
 T_m — постоянная времени двигателя.

Частотная характеристика замкнутой системы

На фиг. 2 представлена структурная схема системы автоматического регулирования с обратной связью, содер-



Фиг. 2. Структурная схема нелинейной системы автоматического регулирования с обратной связью.

жащей нелинейный элемент. Эквивалентный коэффициент усиления нелинейного элемента обозначен через $k(\gamma)$. По сравнению с описывающей функцией, которая может зави-

5. Частотные характеристики замкнутых нелинейных систем 201

сеть как от частоты, так и от амплитуды, эквивалентный коэффициент усиления зависит только от амплитуды входного сигнала, если статическая характеристика нелинейного элемента полностью описывает работу элемента, т. е. если в нелинейном элементе нет источников энергии (позднее мы на этом остановимся подробнее). Если на вход системы, показанной на фиг. 2, подается фиксированный синусоидальный сигнал θ_i , то после затухания всех переходных явлений координата γ будет установившейся синусоидальной величиной и, следовательно, величина $k(\gamma)$ будет постоянной. Эквивалентный коэффициент усиления определяется из следующего соотношения:

$$k(\gamma) = \frac{\gamma'}{\gamma}, \quad (1)$$

поскольку γ' можно определить из статической характеристики нелинейного элемента для всех значений γ . Однако если амплитуда или частота входного сигнала θ_i изменятся, от $k(\gamma)$ примет новое значение, так как на значение амплитуды γ влияют любые изменения θ_i . Следовательно, $k(\gamma)$ можно рассматривать как квазипостоянное усиление, а нелинейный элемент можно представить посредством функции, зависящей от амплитуды.

Частотная характеристика замкнутой системы, представленной на фиг. 2, определяется следующим уравнением:

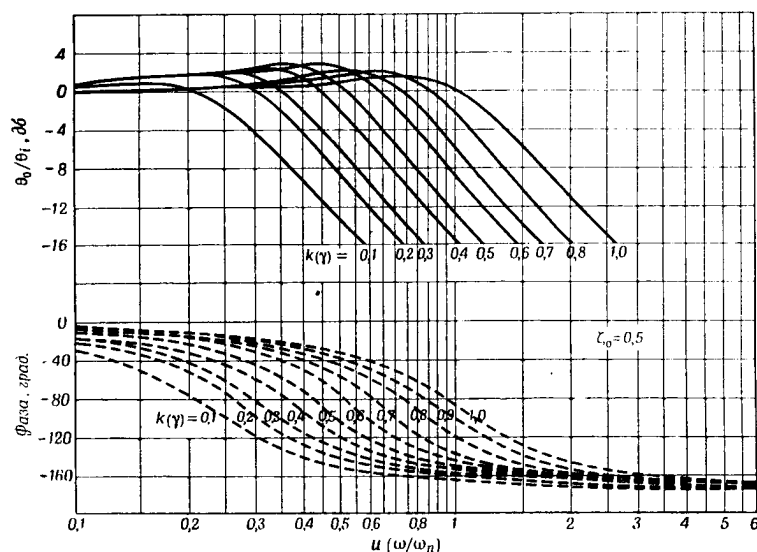
$$\frac{\theta_o}{\theta_i}(s) = \frac{k(\gamma) K_1 K_2 G_1(s) G_2(s)}{1 + k(\gamma) K_1 K_2 G_1(s) G_2(s)}, \quad (2)$$

где $s = j\omega$. Решение этого уравнения нельзя получить, так как θ_o и $k(\gamma)$ неизвестны. Поскольку $k(\gamma)$ при установившемся состоянии можно представить как квазипостоянное усиление, то можно построить семейство частотных характеристик замкнутой системы, считая $k(\gamma)$ параметром. Эти кривые можно определить по следующему уравнению:

$$\frac{\theta_o}{\theta_i}[s, k_n(\gamma)] = \frac{k_n(\gamma) K_1 K_2 G_1(s) G_2(s)}{1 + k_n(\gamma) K_1 K_2 G_1(s) G_2(s)}, \quad (3)$$

где индекс n указывает на значение $k(\gamma)$ для каждой отдельной кривой. Такие семейства амплитудных и фазовых характеристик замкнутой системы приведены на фиг. 3. Если построено достаточное число этих кривых в диапазоне

возможных значений $k(\gamma)$, то точки частотной характеристики замкнутой нелинейной системы для различных значений амплитуды и частоты входного сигнала θ_i будут находиться где-то на кривых постоянного $k(\gamma)$.



Фиг. 3. Частотные характеристики замкнутой системы с $k(\gamma)$ в качестве параметра.

Эти точки можно точно расположить, если величина $k(\gamma)$ определена как функция амплитуды и частоты управляющего воздействия при работе в замкнутом состоянии. Метод, приведенный в этой статье, позволяет определить $k(\gamma)$ по уравнениям (1) и (3) с использованием структурной схемы, изображенной на фиг. 2.

Аналитическое определение $k(\gamma)$ для замкнутой системы

Используя приближенную статическую характеристику нелинейного элемента, можно получить аналитическое выражение для $k(\gamma)$ как некоторую функцию γ , умноженную на постоянную. Обратно, γ можно определить как произведение постоянной на некоторую функцию от $k(\gamma)$; следова-

5. Частотные характеристики замкнутых нелинейных систем 203

тельно, либо $k(\gamma)$, либо γ можно рассматривать как независимую переменную. В этом ключ к решению нелинейного уравнения. Возможно получить два уравнения, в которых неизвестными будут только $k(\gamma)$ и γ . Принимая $k(\gamma)$ в качестве независимой переменной, можно одновременно решить эти два уравнения, исключить γ и определить, таким образом, $k(\gamma)$ как функцию амплитуды и частоты управляющего воздействия. Определив $k(\gamma)$, можно получить частотную характеристику замкнутой системы из семейства кривых, построенных по уравнению (3).

С помощью структурной схемы, изображенной на фиг. 2, можно определить передаточную функцию, связывающую преобразование Лапласа входа нелинейного элемента γ с преобразованием Лапласа управляющего воздействия θ_i на входе системы. Это уравнение легко получается из передаточной функции замкнутой системы $(\theta_o/\theta_i)(s)$ и будет иметь вид

$$\frac{\gamma}{\theta_i}(s) = \left[\frac{\theta_o}{\theta_i}(s) \right] \frac{1}{k(\gamma) K_2 G_2(s)}, \quad (4)$$

где $(\theta_o/\theta_i)(s)$ определяется из уравнения (2).

Разрешив уравнение (1) относительно γ и разделив обе части равенства на θ_i , получим другое уравнение, связывающее γ и θ_i :

$$\frac{\gamma}{\theta_i} = \left[\frac{\gamma'}{k(\gamma)} \right] \frac{1}{\theta_i}. \quad (5)$$

Это уравнение, полученное при рассмотрении статической характеристики нелинейного элемента, связывает величины, зависящие только от амплитуды. Таким образом, мы имеем два уравнения, связывающие две неизвестные величины $k(\gamma)$ и γ со значением θ_i на входе. Приравнявая эти уравнения, исключим γ и получим одно уравнение с одним неизвестным $k(\gamma)$:

$$\frac{1}{\theta_i} \left[\frac{\gamma'}{k(\gamma)} \right] = \frac{1}{k(\gamma) K_2 G_2(s)} \left[\frac{\theta_o}{\theta_i}(s) \right]. \quad (6)$$

Здесь удобно сделать следующую подстановку:

$$\frac{\gamma'}{k(\gamma)} = af[k(\gamma)], \quad (7)$$

где α — постоянная, а $f[k(\gamma)]$ — некоторая функция $k(\gamma)$, которую мы будем рассматривать как функцию эквивалентного коэффициента усиления.

Определим теперь значения $k(\gamma)$, которые удовлетворяют уравнению (6). Поскольку уравнение (6) содержит переменный коэффициент $k(\gamma)$, то целесообразно отделить функции, содержащие $k(\gamma)$, от обычных линейных передаточных функций, как это сделано в уравнении (8)

$$\left| \left[\frac{\alpha}{\theta_i} \right] K_2 G_2(s) \right| = \left| \left[\frac{\theta_o}{\theta_i} \right] (s) \frac{1}{k(\gamma) f[k(\gamma)]} \right|. \quad (8)$$

Можно предложить два метода решения этого нелинейного уравнения. Если значения $k(\gamma)$ заданы и пробное значение исправляется путем учета величины и знака результирующей ошибки, то могут быть определены значения $k(\gamma)$, которые удовлетворяют уравнению (8). Обычно этот метод приводит к сходящемуся решению, но расчет приходится повторять для каждого управляющего воздействия θ_i . Более полное решение можно получить графическим методом. Прежде всего, перепишем уравнение (8) так, чтобы можно было построить кривые для нескольких значений $k(\gamma)$

$$\left| \left[\frac{\alpha}{\theta_i} \right] K_2 G_2(s) \right| = \left| \frac{\theta_o}{\theta_i} [s, k_n(\gamma)] \frac{1}{k_n(\gamma) f[k_n(\gamma)]} \right|. \quad (9)$$

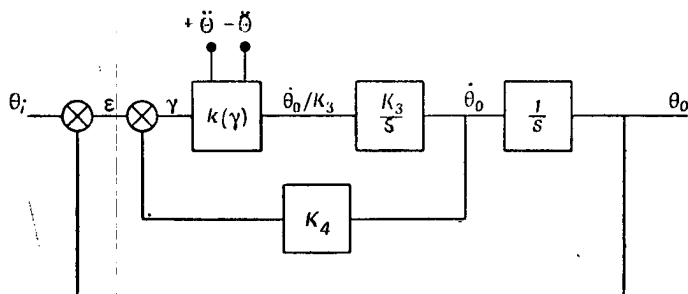
Построим теперь обе части уравнения (9) отдельно на одном и том же графике. Левая часть представляет собой обычную линейную частотную характеристику, а правая часть — семейство кривых, построенных в функции частоты, считая $k(\gamma)$ параметром. Произведение $k(\gamma) f[k(\gamma)]$ можно рассматривать как усиление, которое является постоянным при определенном значении $k(\gamma)$. Пересечения линейной кривой и кривых семейства, полученных из правой части уравнения (9), определяют значения $k(\gamma)$, удовлетворяющие уравнению, как функции амплитуды и частоты управляющего воздействия θ_i .

Эти значения $k(\gamma)$ можно применить для определения частотной характеристики замкнутой системы по семействам кривых, приведенных на фиг. 3. С помощью описанного выше метода будет проанализирована конкретная нелинейная система.

5. Частотные характеристики замкнутых нелинейных систем 205

Ограничение ускорения

На фиг. 4 приведена структурная схема системы с ограничением ускорения, из которой видно, что нелинейный элемент расположен в части системы, содержащей источник энергии. Выражение эквивалентного коэффициента усиления нелинейного элемента зависит от амплитуды только потому, что эта функция связывает вход и выход нелинейного элемента независимо от того, где нелинейный элемент расположен в системе.



Фиг. 4. Система с ограничением ускорения.

Используя уравнение (4), можно получить следующую функцию:

$$\frac{\gamma}{\theta_i}(s) = \left[\frac{\theta_o}{\theta_i}(s) \right] \frac{s^2}{K_3 k(\gamma)}, \quad (10)$$

где

$$\frac{\theta_o}{\theta_i}(s) = \frac{1}{\frac{s^2}{K_3 k(\gamma)} + K_4 s + 1}. \quad (11)$$

Статическая характеристика нелинейного элемента показана на фиг. 5. Из нее можно получить аналитическое выражение для $k(\gamma)$:

$$k(\gamma) = \begin{cases} 1 & \text{при } |\gamma| \leq |\ddot{\theta}_o|, \\ \frac{\ddot{\theta}_o}{\gamma} & \text{при } |\gamma| \geq |\ddot{\theta}_o|. \end{cases} \quad (12)$$

Применяя уравнения (5) и (7), получим

$$\frac{\gamma}{\theta_i} = \frac{\ddot{\theta}_o}{k(\gamma)} \frac{1}{\theta_i} \quad \text{при } |\gamma| \geq |\ddot{\theta}_o|, \quad (13)$$

где мы положили $\alpha = \ddot{\theta}_o$, а $f[k(\gamma)] = 1/k(\gamma)$. Уравнение (11) можно выразить в следующей безразмерной форме:

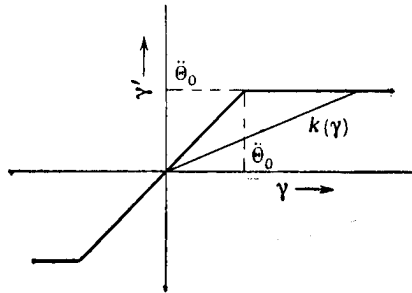
$$\frac{\theta_o}{\theta_i}(s) = \frac{1}{\frac{s^2}{\omega_{n0}^2 k(\gamma)} + \left[\frac{2\zeta_0}{\omega_{n0}} \frac{\sqrt{k(\gamma)}}{\sqrt{k(\gamma)}} \right] s + 1} = \frac{1}{1 - \frac{u^2}{k(\gamma)} + j[2\zeta_0]u}, \quad (14)$$

где

$$s = j\omega, \quad \omega_{n0}^2 = K_3, \quad \zeta_0 = \frac{K_4}{2} \sqrt{K_3}, \quad u = \frac{\omega}{\omega_{n0}};$$

ω_{n0} и ζ_0 обозначают соответственно собственную частоту и коэффициент демпфирования системы при $k(\gamma) = 1$. Заменяя уравнение (14) уравнением (10), получим уравнение, эквивалентное уравнению (4):

$$\frac{\gamma}{\theta_i}(s) = -\frac{u^2}{k(\gamma)} \frac{1}{1 - \frac{u^2}{k(\gamma)} + j[2\zeta_0]u}. \quad (15)$$



Фиг. 5. Статическая характеристика элемента с ограничением.

Теперь можно исключить γ ; приравняв правые части уравнений (13) и (15), получим уравнение

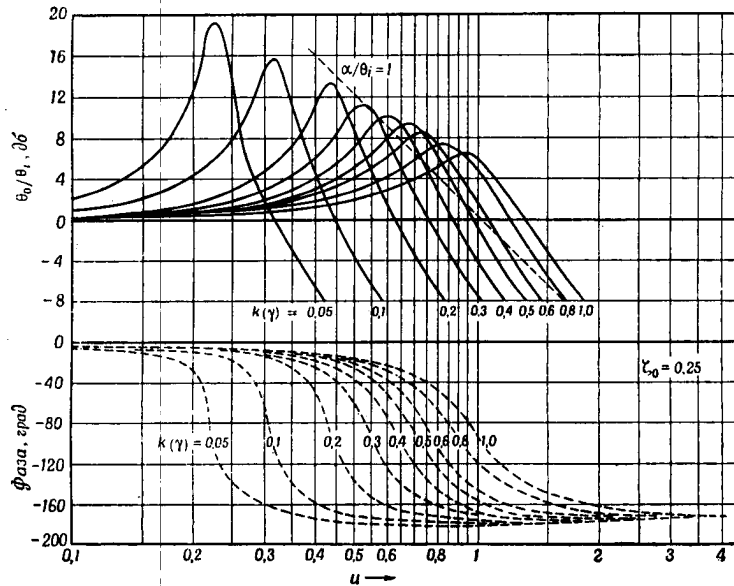
$$\left| \frac{\ddot{\theta}_o}{k(\gamma)} \frac{1}{\theta_i} \right| = \left| -\frac{u^2}{k(\gamma)} \frac{1}{1 - \frac{u^2}{k(\gamma)} + j[2\zeta_0]u} \right|, \quad (16)$$

5. Частотные характеристики замкнутых нелинейных систем 207

которое можно преобразовать к виду, подобному уравнению (8), а именно

$$\left| \frac{\ddot{\theta}_o}{\theta_i} \frac{1}{u^2} \right| = \left| \frac{1}{1 - \frac{u^2}{k(\gamma)} + j [2\zeta_0] u} \right|. \quad (17)$$

Для определения $k(\gamma)$ графическим методом необходимо построить семейство кривых при постоянных значениях $k(\gamma)$



Фиг. 6. Кривые постоянных $k(\gamma)$, построенные по уравнению (18).

как параметра. Аналитически это выражается уравнениями (9) и (18), т. е.

$$\left| \frac{\ddot{\theta}_o}{\theta_i} \frac{1}{u^2} \right| = \frac{1}{1 - \frac{u^2}{k_n(\gamma)} + j [2\zeta_0] u}. \quad (18)$$

На фиг. 6 приведено семейство амплитудных и фазовых кривых, построенных в функции u по правой части уравнения (18), причем $k(\gamma)$ принято за параметр. Частотная кривая, построенная по левой части уравнения (18), является

прямой линией с наклоном —12 дб на октаву, когда u откладывается в логарифмическом масштабе. Эта прямолинейная характеристика также показана на фиг. 6 для одного значения $\ddot{\Theta}_o/\theta_i$. Пересечение этой прямой с семейством кривых постоянного $k(\gamma)$ определяет значения $k(\gamma)$, которые удовлетворяют уравнению (18), когда u принимает различные значения, причем отношение $\ddot{\Theta}_o/\theta_i$ считается постоянным. Это отношение можно рассматривать как усиление которое определяет частоту среза линейных функций. Если это отношение изменяется, то определяется новое семейство значений $k(\gamma)$, которое удовлетворяет уравнению (18). Следовательно, $k(\gamma)$ является функцией ω и отношения $\ddot{\Theta}_o/\theta_i$.

Поскольку характеристику замкнутой системы можно определить, подставляя вычисленные значения $k(\gamma)$ в уравнении (3) при определенной частоте, то легко можно установить, как влияют на характеристику замкнутой системы амплитуда и частота управляющего воздействия θ_i . Если произведение $k(\gamma)f[k(\gamma)]$ равно 1, как это имеет место для нелинейного элемента на фиг. 5, то правая часть уравнения (18) будет такой же, как и уравнения (3). Следовательно, характеристика замкнутой системы и $k(\gamma)$ определяются одновременно. Для нелинейных элементов, у которых $k(\gamma)f[k(\gamma)]$ не равно единице, характеристика замкнутой системы и $k(\gamma)$ должны определяться отдельно.

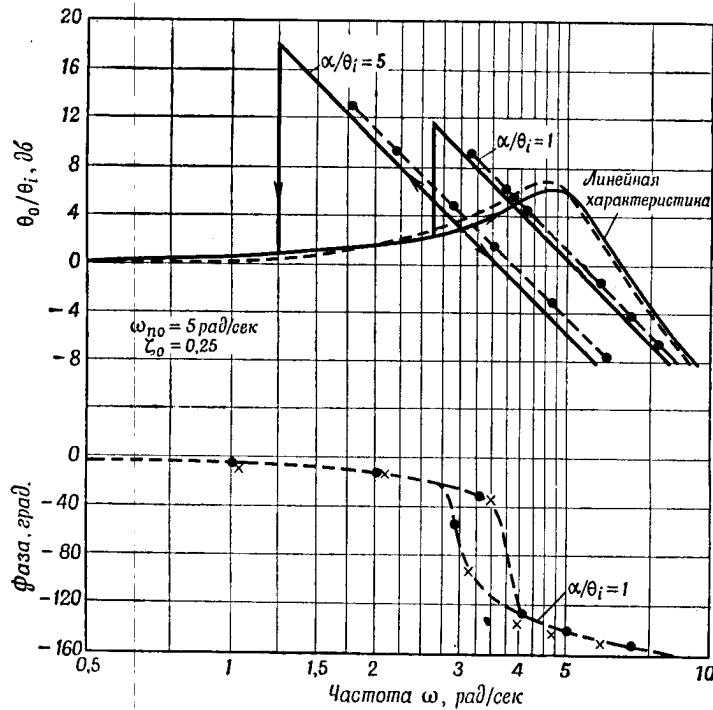
Пример нелинейного элемента, у которого функция эквивалентного коэффициента усиления $f[k(\gamma)]$ не равна единице, приведен в Приложении.

Сравнение аналитических и экспериментальных результатов

Логарифмические амплитудные частотные характеристики системы с ограничением ускорения и соответствующие значения $k(\gamma)$ для одного значения отношения $\ddot{\Theta}_o/\theta_i$ приведены на фиг. 6. Усиление нелинейного элемента равно единице до тех пор, пока $\gamma < \ddot{\Theta}_o$. Следовательно, частотная характеристика замкнутой системы будет совпадать с кривой $k(\gamma) = 1$, пока функция $(\ddot{\Theta}_o/\theta_i)(1/u^2)$ не пересечет кривую $k(\gamma) = 1$. В точке пересечения $\gamma = \ddot{\Theta}_o$. Эта точка является

5. Частотные характеристики замкнутых нелинейных систем 209

также пределом линейной работы системы. Частоту, соответствующую этой точке, обозначим через U_T . Значение $k(\gamma)$, которое удовлетворяет уравнению при $u > U_T$, определяется из пересечения графика линейной функции и кривых $k(\gamma)$. Поскольку для этой системы $k(\gamma) f[k(\gamma)] = 1$, то



Фиг. 7. Частотные характеристики замкнутой системы при наличии ограничения ускорения.

Верхний график: — расчетные данные; - - - данные, полученные с помощью модели; нижний график: ● — расчетные данные; X — данные, полученные с помощью модели.

частотная характеристика системы будет также располагаться вдоль кривой линейной функции, которая определяет значения $k(\gamma)$.

Из фиг. 6 видно, что $(\ddot{\theta}_0/\theta_i) (1/u^2)$ и кривые постоянных $k(\gamma)$ пересекаются при частотах, меньших частоты U_T .

Однако эти значения $k(\gamma)$ не реализуются до тех пор, пока не установится нелинейный режим работы, т. е. когда ускорение в системе будет ограничено. Это явление обычно называется гистерезисом характеристики или скачком частоты, и оно происходит вследствие наличия диапазона частот, в котором $k(\gamma)$ многозначно. Сдвиг фазы системы можно определить по семейству фазовых кривых, построенных по уравнению (14) с заменой $k(\gamma)$ на $k_n(\gamma)$. Используя значения $k(\gamma)$, определенные по уравнению (18), найдем значение фазы, нанося на график точки, полученные при пересечении кривых постоянного $k(\gamma)$ с определенными значениями u . Эти точки для одного значения $\ddot{\theta}_0/\theta_i$ также показаны на фиг. 6.

Чтобы определить ошибку, вводимую аппроксимацией нелинейного элемента эквивалентным коэффициентом усиления, на фиг. 7 приведены аналитическое решение и решение, полученное с помощью модели для нескольких различных значений $\ddot{\theta}_0/\theta_i$. Легко видеть, что в общем имеется совпадение между приближенным решением и точными результатами, полученными на модели. За время экспериментальной работы не было возможности получить столько же точек в нелинейной области для $u < U_T$, как и при анализе. Причина этого в данное время не выяснена. Значительные трудности встретились при определении этих точек с помощью модели, так как оказалось, что рассматриваемая область является полустойчивой. Можно видеть, что частотная характеристика в нелинейной области определяется с помощью характеристик отдельных элементов, расположенных справа от нелинейности на фиг. 4. Следовательно, не удивительно, что нелинейная характеристика показывает затухание 12 дБ на октаву, так как выход элемента с ограничением остается постоянным в течение работы в нелинейной зоне.

Краткое описание метода

Общий метод определения частотной характеристики замкнутой системы сводится к следующему:

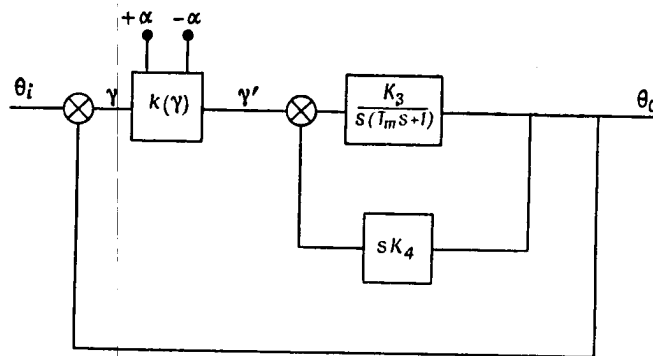
1. Уравнения системы с нелинейными элементами выражаются через передаточные функции и составляется структурная схема системы.

5. Частотные характеристики замкнутых нелинейных систем 211

2. Определяется передаточная функция замкнутой системы, связывающая вход нелинейного элемента γ с управляющим воздействием на входе системы θ_i .

3. Определяется аналитическое выражение, связывающее вход нелинейного элемента γ и эквивалентный коэффициент усиления $k(\gamma)$ нелинейного элемента, с использованием одной только статической характеристики. Обе части уравнения делят на θ_i .

4. Решая уравнения, полученные в пп. 2 и 3, совместно, исключаем γ . В результате получается нелинейное уравнение с одним неизвестным $k(\gamma)$.



Фиг. 8. Система с ограниченным рассогласованием.

5. Уравнение, полученное в п. 4, преобразуется так, чтобы линейные и нелинейные функции находились в разных частях равенства.

6. Заменяем $k(\gamma)$ через $k_n(\gamma)$, строим две функции и графически определяем значения $k(\gamma)$, которые удовлетворяют уравнению.

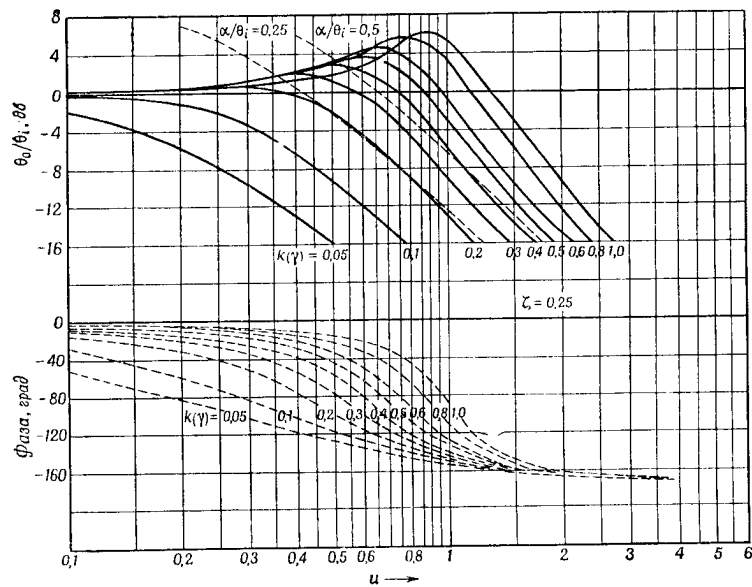
7. Строим логарифмические амплитудную и фазовую частотные характеристики замкнутой системы для тех же параметрических значений $k(\gamma)$, используемых в п. 6. Значения $k(\gamma)$ используются для определения частотной характеристики замкнутой системы.

Определим теперь, согласно указанной процедуре, шаг за шагом частотную характеристику замкнутой системы с ограниченным рассогласованием.

Шаг 1. См. фиг. 8.

Шаг 2.

$$\begin{aligned} \frac{\gamma}{\theta_i}(s) &= \frac{\theta_o}{\theta_i}(s) \frac{s(T_m s + 1)}{k(\gamma) K_B} = \\ &= \frac{1}{k(\gamma) K_B} \frac{s(T_m s + 1)}{\frac{T_m}{K_B k(\gamma)} s^2 + \frac{1 + K_B K_A}{K_B k(\gamma)} s + 1} \end{aligned} \quad (19)$$



Фиг. 9. Кривые постоянных $k(\gamma)$, построенные по уравнению (23)

Безразмерная форма этого уравнения будет иметь вид

$$\frac{\gamma}{\theta_i}(s) = \frac{1}{k(\gamma)} \frac{2\zeta_0 u \left(\frac{1}{2\zeta_0} u + 1 \right)}{1 - \frac{u^2}{k(\gamma)} + j \frac{2\zeta_0}{k(\gamma)} u}, \quad (20)$$

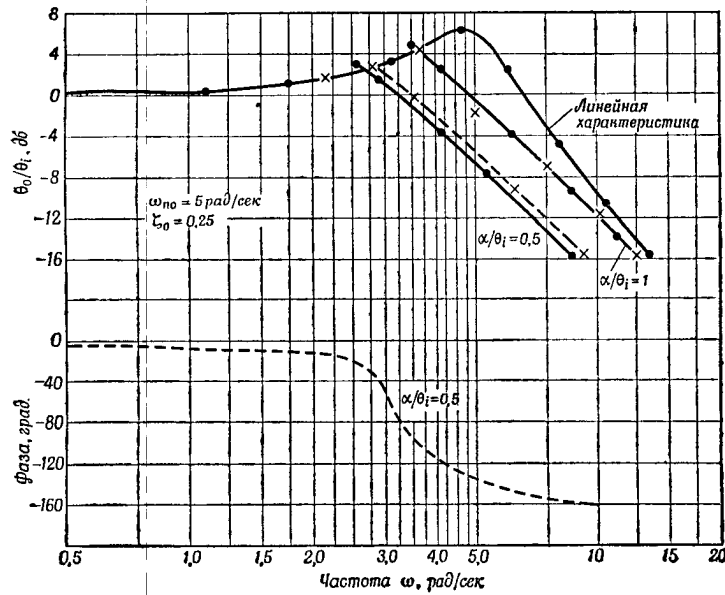
где

$$\zeta_0 = \frac{1 + K_B K_A}{2\sqrt{T_m K_B}}, \quad \omega_{n0} = \sqrt{\frac{K_B}{T_m}}, \quad s = j\omega \quad \text{и} \quad u = \frac{\omega}{\omega_{n0}}.$$

5. Частотные характеристики замкнутых нелинейных систем 213

Шаг 3. Поскольку элемент с ограничением имеет характеристику, приведенную на фиг. 5, то уравнение (13) можно записать, полагая $\alpha = \ddot{\theta}_0$, в виде

$$\frac{\gamma}{\theta_i} = \frac{\alpha}{k(\gamma)} \frac{1}{\theta_i} \text{ при } |\gamma| > |\alpha|. \quad (21)$$



Фиг. 10. Частотные характеристики замкнутой системы с ограниченным рассогласованием.

● — расчетные данные; × — данные, полученные с помощью модели.

Шаг 4.

$$\left| \left[\frac{\alpha}{\theta_i} \right] \frac{1}{k(\gamma)} \right| = \left| \frac{1}{k(\gamma)} \frac{2\zeta_0 u \left(\frac{1}{2\zeta_0} u + 1 \right)}{1 - \frac{u^2}{k(\gamma)} + j \frac{2\zeta_0}{k(\gamma)} u} \right|. \quad (22)$$

Шаг 5.

$$\left| \frac{\left[\frac{\alpha}{\theta_i} \right] \frac{1}{2\zeta_0}}{u \left(\frac{1}{2\zeta_0} u + 1 \right)} \right| = \left| \frac{1}{1 - \frac{u^2}{k_n(\gamma)} + j \frac{2\zeta_0}{k_n(\gamma)} u} \right|. \quad (23)$$

Шаги 6 и 7. Поскольку $k(\gamma)f[k(\gamma)] = 1$, эти два шага выполняются одновременно, потому что нелинейная часть уравнения (23) является также характеристикой замкнутой системы $(\theta_o/\theta_i)[s_1, k_n(\gamma)]$, считая $k_n(\gamma)$ параметром. Оба эти шага иллюстрируются фиг. 9. График зависимой от частоты левой части уравнения (23) можно построить с использованием шаблонов апериодического звена первого порядка и интегрирующего звена. Шаблон располагается при усилении $(\alpha/\theta_i)(1/2\epsilon_o)$. На фиг. 10 показаны аналитическое решение и решение, полученное с помощью модели.

Заключение

Представляя нелинейные элементы с помощью эквивалентного коэффициента усиления, можно получить точные количественные сведения относительно частотной характеристики замкнутой нелинейной системы. Точность метода зависит прежде всего от выполнения трех условий: 1) возможность пренебречь высшими гармониками на выходе системы; 2) допустимость кусочно-линейной аппроксимации характеристик нелинейного элемента и 3) синусоидальный вид сигнала, подаваемого на вход нелинейного элемента. Этим требованиям удовлетворяет достаточно широкий класс нелинейных систем. Обычно эквивалентный коэффициент усиления можно легко вычислить из геометрического построения. Эта функция является инвариантной по отношению к частоте, если динамические характеристики не зависят от производных входного сигнала.

Для большинства практических задач статические характеристики в равной степени описывают и динамический режим. Следовательно, функция эквивалентного коэффициента усиления остается инвариантной по отношению к частоте, даже если нелинейный элемент будет расположен во внутреннем контуре, содержащем источник энергии. Хотя этот метод был применен только к системам второго порядка, он применим также и к системам более высокого порядка, если они содержат также только один нелинейный элемент.

Другое преимущество аппроксимации с помощью эквивалентного коэффициента усиления состоит в том, что влияние изменений в частотной характеристике замкнутой системы, вызванных изменением амплитуды управляющего воздей-

5. Частотные характеристики замкнутых нелинейных систем 215

вия θ_i , легко определить, располагая шаблон, представляющий левую часть уравнения (23), в соответствии с различными значениями θ_i .

ПРИЛОЖЕНИЕ

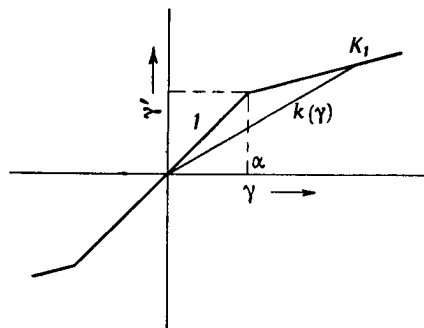
Аналитическое выражение для эквивалентного коэффициента усиления элемента с ограничением, которое использовалось в статье, дано уравнением (12). Более общее выражение было приведено в уравнении (1):

$$k(\gamma) = \frac{\gamma'}{\gamma},$$

которое можно записать в следующем виде:

$$\gamma = \frac{\gamma'}{k(\gamma)} = \alpha f[k(\gamma)].$$

Выражение $f[k(\gamma)]$ определяется как функция эквивалентного коэффициента усиления, равная $1/k(\gamma)$, для рассмотренного



Фиг. 11. Статическая характеристика нелинейного элемента.

типа элемента с ограничением. Это особый случай нелинейной функции, так как произведение $k(\gamma)f[k(\gamma)]$ в общем случае не равно единице. Поэтому мы покажем, как отклонение значения этого произведения от единицы влияет на решение нелинейного уравнения. Статическая характеристика нелинейного элемента более общего вида показана на фиг. 11. Определим функцию эквивалентного коэффициента усиления

для этого элемента. Из геометрического рассмотрения можно получить следующие соотношения:

$$\gamma' = \gamma [k(\gamma)] = \alpha + K_1(\gamma - \alpha) \quad \text{при } |\gamma| > |\alpha|, \quad (24)$$

$$\gamma [k(\gamma) - K_1] = \alpha(1 - K_1). \quad (25)$$

Следовательно,

$$\gamma = \frac{\alpha(1 - K_1)}{k(\gamma) - K_1} = \alpha f[k(\gamma)]. \quad (26)$$

В этом случае

$$f[k(\gamma)] = \frac{1 - K_1}{k(\gamma) - K_1}.$$

Если значение $f[k(\gamma)]$ в уравнении (9) заменить полученным выражением, то имеем

$$\left| \left[\frac{\alpha}{\theta_i} \right] K_2 G_2(s) \right| = \left| \frac{\theta_o}{\theta_i} [s, k_n(\gamma)] \left\{ \frac{k(\gamma) - K_1}{k(\gamma) [1 - K_1]} \right\} \right|. \quad (27)$$

В этом уравнении, как и в уравнении (9), только $k(\gamma)$ является неизвестной, когда заданы амплитуда и частота θ_i . Следовательно, значения $k(\gamma)$, удовлетворяющие уравнению, можно определить описанным выше методом. Левая часть уравнения остается неизменной, и единственное различие в правой части состоит в том, что кривые, рассчитанные с помощью функции $(\theta_o / \theta_i) [s, k_n(\gamma)]$, будут отличаться от каждой последующей коэффициентом усиления $[k(\gamma) - K_1] / [k(\gamma)(1 - K_1)]$, вычисленным при соответствующих значениях $k(\gamma)$.

Для построения частотной характеристики замкнутой системы необходимо прежде всего определить значения $k(\gamma)$ по уравнению (27). Эти значения затем применяют для построения характеристики замкнутой системы по $(\theta_o / \theta_i) [s, k_n(\gamma)]$, как это предлагает п. 7 краткого описания метода.

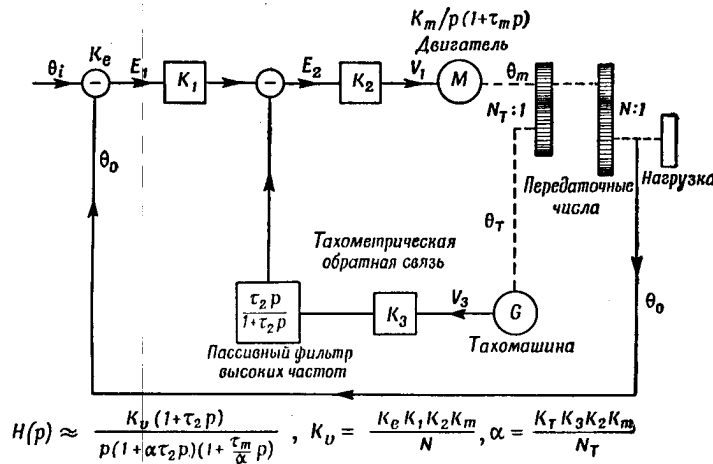
ЛИТЕРАТУРА

1. Kochenburger R., *Trans. AIEE*, Part II, **69**, 270 (1950).
2. Kochenburger R., *Trans. AIEE*, Part II, **72**, 180 (1953).
3. Johnson E. C., *Trans. AIEE*, Part II, **71**, 169 (1952).

**6. НЕКОТОРЫЕ ЯВЛЕНИЯ НАСЫЩЕНИЯ
В СЛЕДЯЩИХ СИСТЕМАХ С ТАХОМЕТРИЧЕСКОЙ
ОБРАТНОЙ СВЯЗЬЮ¹⁾**

Е. Левинсон

Тахометрическая обратная связь с пассивным фильтром высокой частоты хорошо известна [1] и широко применяется для стабилизации систем автоматического регулирования.



Ф и г. 1. Блок-схема следящей системы с тахометрической обратной связью.

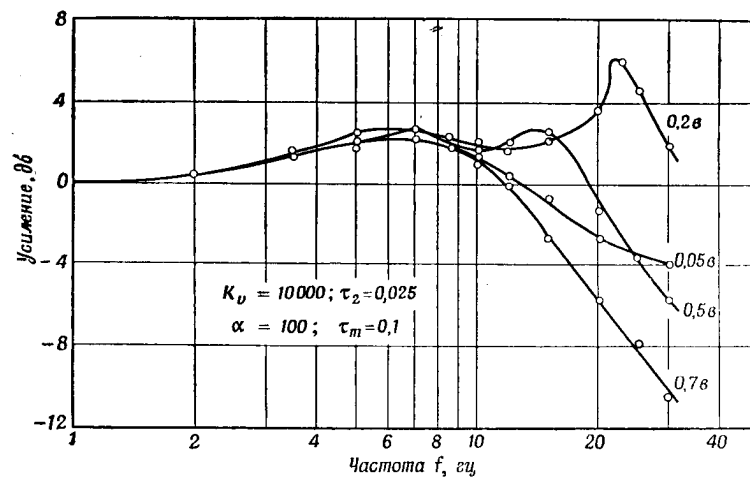
Большое преимущество такой стабилизации по сравнению с введением стабилизирующего контура²⁾ состоит в том, что можно получить высокий коэффициент передачи K_U (малое запаздывание по скорости) при узкой полосе пропускания системы (меньшей передаче шумов). На фиг. 1 приведена

¹⁾ Levinson E., *Trans. AIEE*, Part II, 72, 1—9 (1953).

²⁾ Подразумевается последовательный дифференцирующий контур.— *Прим. ред.*

схема типичной следящей системы, стабилизированной посредством тахометрической обратной связи с пассивным фильтром высоких частот.

При изучении поведения системы этого типа путем определения частотных характеристик замкнутой системы часто сталкиваются с некоторыми странными явлениями при изменении амплитуды входного сигнала. Типичное семейство



Фиг. 2. Типичное семейство частотных характеристик системы с насыщением и тахометрической обратной связью.

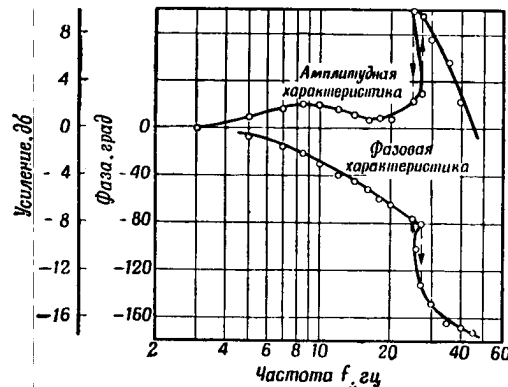
Показаны частотные характеристики при различных типовых сигналах на входе. $K_D = 10\ 000; \alpha = 100; \tau_2 = 0,025; \tau_m = 0,1$.

характеристик системы показано на фиг. 2. При малых сигналах поведение системы довольно хорошо согласуется с линейной теорией. При увеличении сигнала высота резонансных пиков увеличивается и они перемещаются в область более высоких частот, а в некоторые моменты может иметь место явление „скачкообразного“ резонанса¹⁾, как это видно на фиг. 3. Однако при еще большем увеличении входного сигнала высота пиков и частота среза уменьшаются.

¹⁾ Явление „скачкообразного“ резонанса, характерное только для нелинейных систем, изучается в теории нелинейных колебаний. См., например, Боголюбов Н., Митропольский Ю., Асимптотические методы в теории нелинейных колебаний, Гостехиздат, 1955. — Прим. ред.

6. Насыщение в системах с тахометрической обратной связью 219

Основная цель нашей статьи — дать приемлемое объяснение этого явления. При изложении будут использованы некоторые принципы и методы, которые могут быть полезны при анализе различных систем автоматического регулирования, имеющих элементы с насыщением. Кроме того, передаточная функция системы с тахометрической обратной связью и высокочастотным фильтром будет выражена в чрезвычайно удобной форме, в которой все сопрягающие точки



Фиг. 3. Явление скачкообразного резонанса на частотной характеристике.

легко связываются с параметрами системы. Передаточная функция подобного же вида получена Даукером несколько иным способом [2].

Порядок исследования

Первый основной принцип состоит в том, что эффект насыщения уменьшает среднее значение усиления цепи, в которой имеет место насыщение (типичная кривая насыщения и функция насыщения усиления показаны на фиг. 6). Эти изменения параметров усиления можно проследить на логарифмической амплитудной частотной характеристике, и они используются для исследования рассматриваемых характеристик замкнутой системы.

Для каждого условия насыщения получают различные логарифмические характеристики, причем здесь по-прежнему используется линейный анализ. Затем частотная характе-

ридика¹⁾ связывается с логарифмической характеристикой разомкнутой системы на том основании, что частота среза частотной характеристики замкнутой системы²⁾ близка к частоте среза логарифмической характеристики разомкнутой системы, а наклон последней в -12 дБ на октаву означает наличие пика в частотной характеристике замкнутой системы.

Для объяснения явления скачкообразного резонанса будет рассмотрен другой графо-аналитический метод, который можно использовать для получения частотной характеристики любой следящей системы с насыщением.

С помощью этого графо-аналитического метода показано, что это явление имеет место в простой одноконтурной следящей системе с насыщением в прямой цепи. Те же уравнения, как показано далее, справедливы при особых условиях и для случая тахометрической стабилизации.

Объяснение частотных характеристик на основе понятия „усиления“

Как показано в Приложении I и в соответствии с блок-схемой, изображенной на фиг. 1, приближенное выражение для передаточной функции разомкнутой системы с параллельным стабилизирующим устройством, состоящим из тахомашины, механически связанной с валом двигателя, и пассивного контура, имеет вид

$$H(p) = \frac{\theta_0}{E}(p) = \frac{K_v(1 + \tau_2 p)}{p(1 + \alpha \tau_2 p) \left[1 + \frac{\tau_m}{\alpha} p \right]}, \quad \alpha \tau_2 \gg \tau_m, \quad \alpha \gg 1, \quad (1)$$

где

$$K_v = \frac{K_e K_1 K_2 K_m}{N}$$

является коэффициентом усиления системы при разомкнутой внешней цепи, а

$$\alpha = \frac{K_T K_3 K_2 K_m}{N_T}$$

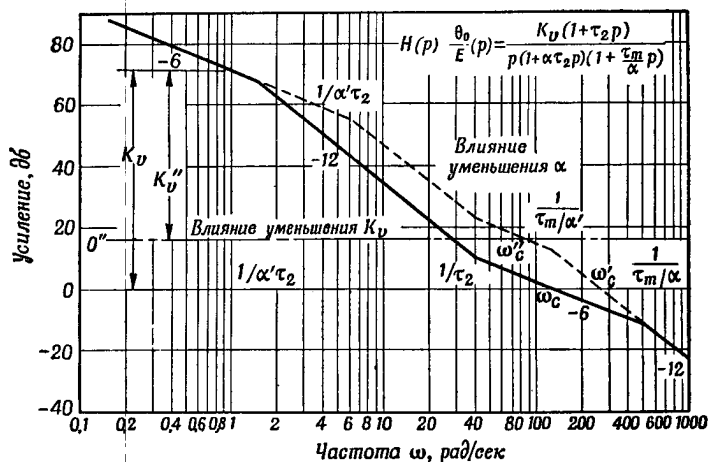
¹⁾ Имеется в виду частотная характеристика замкнутой системы.—
Прим. ред.

²⁾ Частота среза частотной характеристики замкнутой системы определяется как частота, при которой амплитудная частотная характеристика замкнутой системы становится равной единице.—
Прим. ред.

6. Насыщение в системах с тахометрической обратной связью 221

является коэффициентом усиления разомкнутой цепи тахометрической обратной связи. τ_2 — постоянная времени пассивного фильтра высокой частоты, а τ_m — постоянная времени двигателя и нагрузки.

Можно заметить, что точки сопряжения на логарифмической характеристике системы просто связаны с параметрами системы. Это хорошо видно на фиг. 4, где приведена характеристика, соответствующая передаточной функции (1).



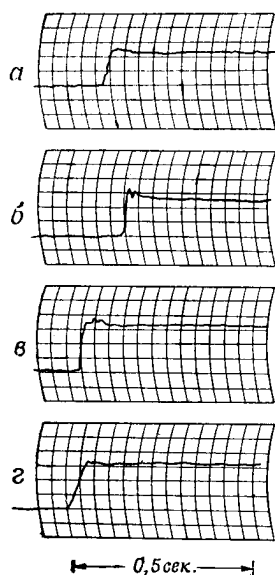
Фиг. 4. Логарифмическая частотная характеристика для стабилизированной тахометрической системы.

По этой характеристике мы будем теперь исследовать влияние насыщения.

Если в случае увеличения сигнала в цепи тахометрической обратной связи имеет место насыщение (например, в блоке K_3), в то время как основная часть системы работает еще в линейном режиме, то система будет вести себя так, как если бы имелось меньшее α при том же самом K_v . Уменьшение α сближает первую и третью точки сопряжения, как показано пунктирной линией на фиг. 4. Это приводит к новой более высокой по сравнению с ω_c частоте среза ω'_c , и характеристика при ω'_c имеет наклон -12 дБ на октаву.

Это означает, что амплитудная частотная характеристика¹⁾ будет иметь большой резонансный пик на более высокой частоте и более высокую частоту среза, как это видно на фиг. 2.

Если величина сигнала возрастает и далее, то основная часть системы также начинает насыщаться и система работает так, как будто K_p уменьшается. Вследствие этого горизонтальная линия 0 дб должна смещаться вверх, как показано на фиг. 4 штрих-пунктирной линией. Это приводит к другой частоте среза ω'_c , которая меньше ω_c , и характеристика при этой новой частоте среза может иметь наклон в -6 дб на октаву.



Фиг. 5. Экспериментальные кривые переходных процессов в системе с насыщением и тахометрической стабилизацией.

Скачок напряжения: $a - 0,07\sigma$;
 $b - 0,2\sigma$; $v - 0,5\sigma$; $z - 2\sigma$.
 Масштаб времени: 25 делений
 в 1 сек.

Следовательно, пики на кривой исчезают, и частота среза имеет место при более низких частотах. Это как раз соответствует поведению кривых на фиг. 2.

При более высоких амплитудах сигнала частота среза снова попадает на участок с наклоном характеристики в -12 дб на октаву, снова появляются высокие резонансные пики на низких частотах и частота среза замкнутой системы становится более низкой.

Если при увеличении сигнала насыщение сначала возникает в основной части системы, то, по-видимому, система не будет иметь вначале пика при высоких частотах, поскольку α начинает насыщаться в момент, когда характеристика системы попадает в область частоты среза ω_c . Следовательно, если желательно уменьшить обусловленный нелинейностью пик

¹⁾ Имеется в виду амплитудная частотная характеристика замкнутой системы.—Прим. ред.

6. Насыщение в системах с тахометрической обратной связью 223

при высоких частотах, величина K_v должна быть сделана такой, чтобы прямая цепь системы насыщалась одновременно или раньше цепи тахометрической обратной связи.

Влияние насыщения на переходные процессы при ступенчатых входных воздействиях различной амплитуды можно объяснить и предвидеть, исходя из частотных характеристик системы.

Так, например, в диапазоне сигналов, обуславливающих появление в характеристике пика при высоких частотах, могут быть получены более высокие быстродействие и колебательность, чем при сигналах, соответствующих линейной области. При еще более высоких амплитудах сигнала время переходного процесса системы будет увеличиваться и переходный процесс будет менее колебательным. Эти переходные процессы приведены на фиг. 5.

Этот основной принцип изменения усиления в зависимости от насыщения можно распространить на изучение влияния насыщения в любой системе автоматического регулирования, в частности на многоконтурные системы, в которых можно независимо друг от друга исследовать влияние каждого контура.

Применение понятия „усиление“ к одноконтурной системе

В качестве другого примера рассмотрим простую одноконтурную систему автоматического регулирования с последовательным стабилизирующим устройством.

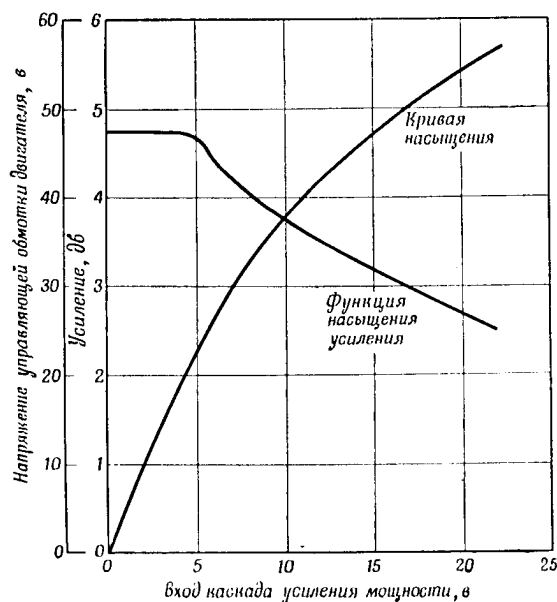
Передаточная функция системы такого типа будет иметь вид

$$H(p) = \frac{\theta_0}{E}(p) = \frac{K_v(1 + \tau_1 p)}{p(1 + \tau_m p)(1 + \tau_2 p)}. \quad (2)$$

На фиг. 7 построена логарифмическая амплитудная частотная характеристика, соответствующая передаточной функции (2), и приведена блок-схема системы. На фиг. 8 показано семейство частотных характеристик рассматриваемой системы при различных значениях входного сигнала. Можно заметить, что при увеличении амплитуды сигнала пик первоначально уменьшается, а частота среза также становится меньше. При дальнейшем увеличении амплитуды сигнала высота

пика начинает увеличиваться, а частота среза становится все меньше и меньше. Это явление можно пояснить, рассматривая логарифмическую характеристику на фиг. 7.

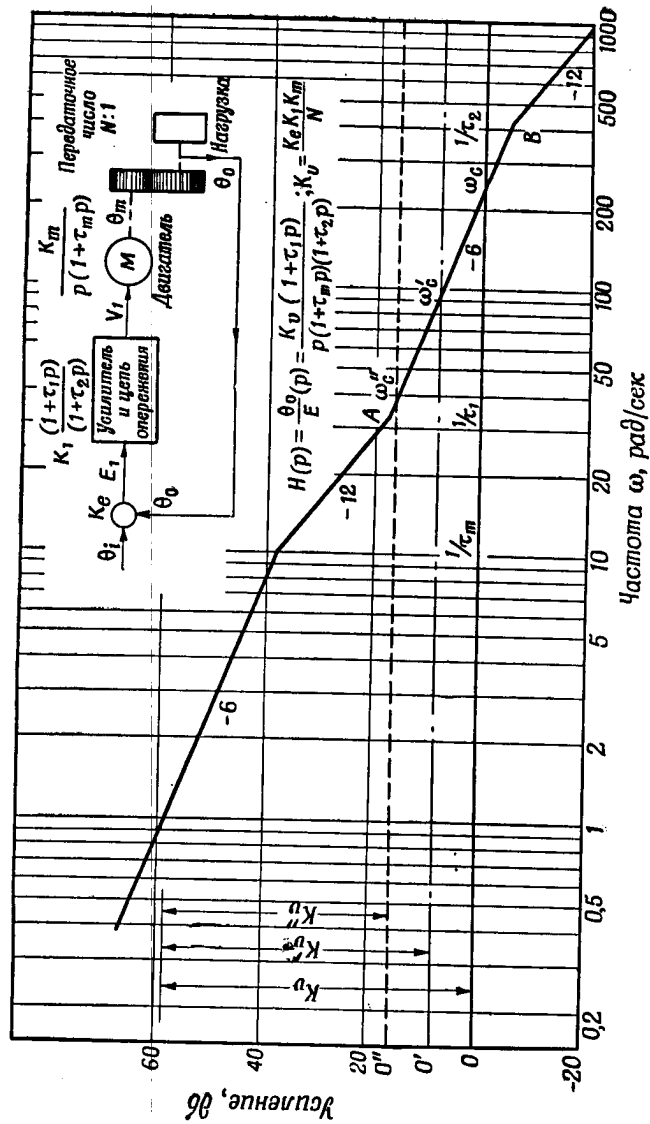
Единственным параметром усиления в одноконтурной системе является K_v , и, следовательно, при уменьшении K_v нулевая линия должна смещаться вверх. Когда амплитуда



Фиг. 6. Типичная кривая насыщения и функция насыщения усиления.

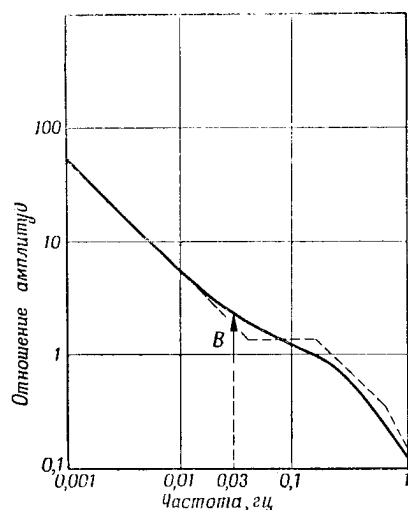
сигнала возрастает, действительное значение K_v уменьшается и частота среза соответствует новой более низкой частоте ω_c .

Начальное изменение пиков характеристики зависит от первоначального значения K_v в системе и от относительной близости расположения частоты среза к участку характеристики с наклоном -12 дБ на октаву. Если частота среза первоначально находится в точке В, то с увеличением амплитуды сигнала ожидаются сначала меньшие пики. При первоначальном расположении частоты среза в точке А с увеличе-



Ф и г. 7. Логарифмическая частотная характеристика и блок-схема простой замкнутой системы с целью опережения.

будет синусоидальным (в предположении, что система линейна). Характеристика типового регулятора дается на фиг. 4. Ее можно также аппроксимировать асимптотической характеристикой, показанной пунктиром на фиг. 4 и составленной



Ф и г. 4. Логарифмическая амплитудная частотная характеристика регулятора двигателя.

из двух прямолинейных асимптот и двух промежуточных прямолинейных отрезков. Для построения этой асимптотической характеристики практически необходимо иметь ана-



Ф и г. 5. Разомкнутая система регулятор — двигатель.

литическое выражение „передаточной функции“. Если специалист не знаком с этими функциями, ему, вероятно, придется иметь дело с готовыми характеристиками.

Предположим теперь, что регулятор соединен с двигателем и система разомкнута, как показано на фиг. 5. Здесь

1. ИЗОБРАЖЕНИЕ ЧАСТОТНЫХ ХАРАКТЕРИСТИК, СТАНДАРТЫ И РАСЧЕТНЫЕ КРИТЕРИИ ¹⁾

Р. Ольденбургер

Введение

Задачей Комитета динамических систем, созданного при американском обществе инженеров-механиков (ASME) в ноябре 1951 г., является разработка стандартов для представления частотных данных. Необходимость в таких стандартах стала очевидной ввиду широкого применения частотных методов для расчета систем автоматического регулирования. Из-за различного представления одних и тех же зависимостей, касающихся частотных характеристик, часто возникают недоразумения. Одни строят амплитудные характеристики в виде зависимости отношения амплитуд от частоты, другие — в виде зависимости логарифма отношения амплитуд от логарифма частоты. Между тем, одного взгляда на частотные характеристики элемента физической системы порой бывает достаточно, чтобы сделать важные заключения о динамических свойствах элемента, при условии, что используется привычная система координат. Поэтому для легкого представления характеристик весьма желательно, чтобы они строились в стандартных координатах. Эти и другие соображения привели к созданию Комитета динамических систем, который после изучения всех имеющихся предложений вынес свои рекомендации, публикуемые в настоящей статье.

Чтобы облегчить понимание рекомендаций Комитета и способствовать применению частотных методов для расчета систем регулирования, в настоящей статье рассматриваются некоторые основные вопросы и даются правила расчета, позволяющие инженеру без специальной математической подготовки приступить к научному расчету систем автоматического регулирования. Большинство предлагаемых расчетных

¹⁾ Oldenburger R., *Trans. ASME*, **76**, № 8, 1155—1169 (1954).

Тогда вместо p можно подставить $j\omega$, как это делается в линейном анализе, и мы получим в качестве частотной характеристики следующее выражение:

$$\frac{E}{\theta_i}(j\omega) = \frac{j\omega(1 + j\omega\tau_m)}{j\omega(1 + j\omega\tau_m) + K_v(|E|)}, \quad (7)$$

или

$$|E| = \frac{\omega \sqrt{1 + \omega^2 \tau_m^2} |\theta_i|}{\sqrt{[K_v(|E|) - \omega^2 \tau_m]^2 + \omega^2}} \quad (8)$$

и

$$\varphi_E = 90^\circ + \operatorname{arctg} \omega \tau_m - \operatorname{arctg} \frac{\omega}{[K_v(|E|) - \omega^2 \tau_m]}. \quad (9)$$

Поскольку $K_v(|E|)$ является функцией $|E|$, то непосредственно разрешить уравнение относительно $|E|$ или φ_E невозможно. Однако с помощью приведенного ниже графического метода удастся получить решение для $|E|$, а следовательно, и для φ_E , θ_o и φ_o .

Возводя в квадрат обе части уравнения (8) и приводя их к общему знаменателю, получим

$$\begin{aligned} [K_v(|E|) - \omega^2 \tau_m]^2 |E|^2 &= \\ &= \omega^2 (1 + \omega^2 \tau_m^2) |\theta_i|^2 - \omega^2 |E|^2. \end{aligned} \quad (10)$$

Извлекая из обеих частей квадратный корень, имеем

$$\begin{aligned} [K_v(|E|) - \omega^2 \tau_m] |E| &= \\ &= \pm \omega \sqrt{(1 + \omega^2 \tau_m^2) |\theta_i|^2 - |E|^2}, \end{aligned}$$

или

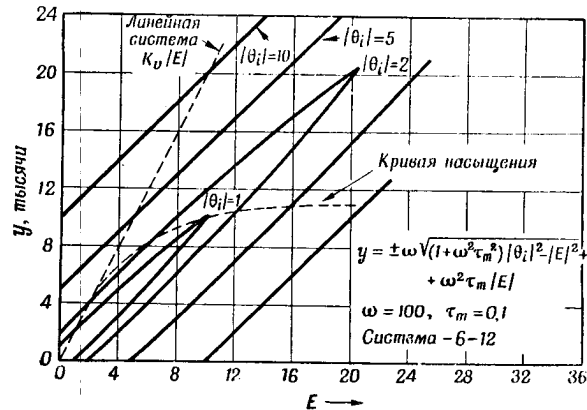
$$\begin{aligned} K_v(|E|) |E| &= \omega^2 \tau_m |E| \pm \\ &\pm \omega \sqrt{(1 + \omega^2 \tau_m^2) |\theta_i|^2 - |E|^2}. \end{aligned} \quad (11)$$

Левая часть уравнения (11) является кривой насыщения прямой цепи системы. Правая часть представляет собой семейство эллипсов. Пересечения этих эллипсов с кривыми насыщения дают величину $|E|$ при определенных ω и $|\theta_i|$. Каждый из параметров K_v , τ_m , ω и $|\theta_i|$ может изменяться и можно определить влияние их на величину $|E|$. На фиг. 9

6. Насыщение в системах с тахометрической обратной связью 229

построено типичное семейство эллипсов при фиксированных ω и τ_m и переменной величине входного сигнала $|\theta_i|$, а на фиг. 10 — типичное семейство эллипсов для фиксированных τ_m и $|\theta_i|$ и изменяющейся частоты ω .

Можно видеть, что увеличение $|\theta_i|$ должно увеличивать размеры эллипса, тогда как увеличение ω должно поворачивать контур эллипса по часовой стрелке при одновременном увеличении его размеров. В каждом случае строится типичная кривая насыщения и видно, что в первом квадранте может иметь место одно, два (касание) или три пересечения. Вследствие симметрии относительно начала координат



Фиг. 9. Типичное семейство эллипсов при фиксированных ω и τ_m и переменной амплитуде $|\theta_i|$.

те же пересечения имеются и в третьем квадранте и, следовательно, они дадут те же значения $|E|$. Это значит, что необходимо рассматривать только первый квадрант. Возможность трех пересечений приводит к явлению скачкообразного резонанса. В линейной системе эта возможность не может иметь места, что показывает прямая K_v на графике.

Амплитудная частотная характеристика для $|E|$, полученная по пересечениям на фиг. 10, построена на фиг. 11, причем имеет место явление скачкообразного резонанса. Три значения $|E|$ не могут одновременно существовать в физической системе. Какое именно значение имеет место,

д) решается задача аппроксимации, т. е. находятся передаточные функции корректирующих устройств по найденным для них частотным характеристикам;

е) решается задача реализации полученных передаточных функций при помощи тех или иных технических средств (особенно хорошо разработана методика решения задачи реализации при помощи пассивных электрических контуров).

Подводя итоги приведенной выше краткой характеристике частотного метода, можно утверждать, что этот метод, основанный на наглядных физических представлениях и позволяющий в качестве исходных данных пользоваться не только дифференциальными уравнениями, но и экспериментально полученными частотными характеристиками, является в настоящее время основным средством анализа и синтеза сложных одноконтурных и многоконтурных систем автоматического регулирования, описываемых линейными уравнениями любого порядка и содержащих не только сосредоточенные, но и распределенные параметры.

Частотный метод дает единую общую методику решения всех основных задач теории автоматического регулирования, а именно: получения исходных данных, анализа устойчивости, анализа качества, построения переходных процессов, выбора параметров и решения вопроса о малости параметров, синтеза корректирующих устройств, анализа динамической точности и синтеза оптимальных систем.

Частотный метод удобен не только в теории непрерывного, но и в теории импульсного, прерывистого регулирования.

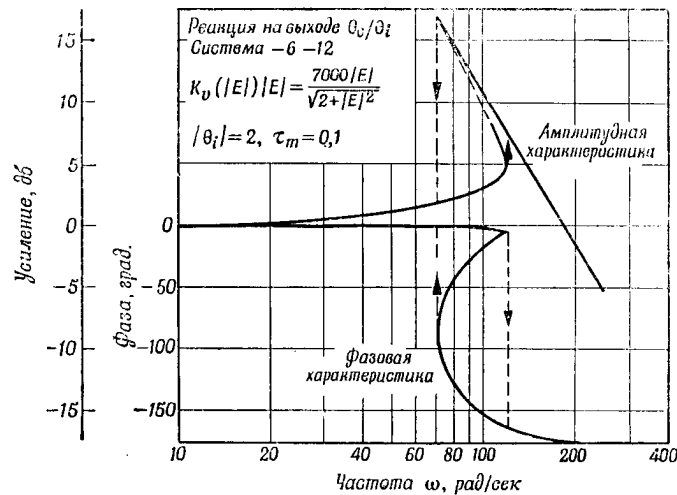
Необходимо отметить, что частотный метод за последнее время стал применяться для анализа не только линейных систем, но и систем с нелинейными и переменными параметрами. Так, например, сочетание частотных представлений с методом гармонического баланса привело к разработке практически удобного приближенного метода анализа автоколебаний [3], недавно разработаны основные положения частотного метода точного анализа релейных систем [4], показана возможность применения частотного метода к анализу систем с переменными параметрами на основе теоремы о свертке [5].

Поэтому утверждения, которые порой приходится слышать, что частотный метод полезен лишь для решения линейных задач, конечно, не соответствуют истине.

Подставляя уравнение (9) в уравнение (13), получим

$$\varphi_0 = - \operatorname{arctg} \frac{\omega}{(K_v(|E|) - \omega^2 \tau_m)}. \quad (14)$$

Амплитудная и фазовая выходные характеристики, соответствующие характеристике $|E|$ на фиг. 11, построены на фиг. 12. Можно видеть, что общий характер этих кривых идентичен экспериментальным кривым на фиг. 3.



Фиг. 12. Скачкообразный резонанс на амплитудной и фазовой частотных характеристиках для выхода.

Из уравнения (29) Приложения I видно, что для $\alpha = 0$ (отсутствие тахометрической обратной связи) передаточная функция системы с тахометрической обратной связью сводится к уравнению (3).

При быстром насыщении цепи тахометрической обратной связи α стремится к нулю, а передаточная функция — к уравнению (3) и появляется возможность скачкообразного резонанса.

Может вызвать удивление то обстоятельство, что это явление так редко встречается при изучении следящих систем. Одна из причин, бесспорно, состоит в том, что исследователи уделяли мало внимания изучению характеристик

6. Насыщение в системах с тахометрической обратной связью 233

следящих систем при очень больших сигналах на входе. Но, возможно, главнейшая причина состоит в том, что большинство следящих систем имеет цель опережения или, иначе говоря, они компенсированы таким образом, что передаточная функция разомкнутой системы обычно характеризуется прямой с наклоном, близким к -6 дБ на октаву. Мы покажем, что скачкообразный резонанс не может иметь места для систем этого идеального типа, если кривая насыщения имеет обычную форму. Для системы, представляемой частотной характеристикой с наклоном -6 дБ на октаву, сопрягающая частота при переходе к наклону -12 дБ стремится к бесконечности, или τ_m стремится к нулю. Для $\tau_m = 0$ уравнение (3) сводится к

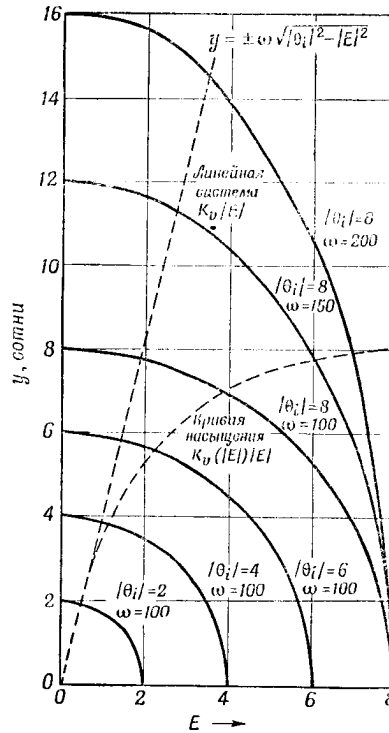
$$\frac{\theta_0}{E} = \frac{K_v(|E|)}{p},$$

а уравнение (11) при этом принимает вид

$$K_v(|E|)|E| = \pm \omega \sqrt{|\theta_i|^2 - |E|^2}. \quad (15)$$

В этом случае семейство эллипсов располагается, как показано на фиг. 13. Увеличение $|\theta_i|$ растягивает эллипсы по обеим осям, тогда как при увеличении ω эллипсы удлиняются лишь вдоль оси y .

Важным моментом, отмеченным на фиг. 13, является то обстоятельство, что для любой однозначной функции $K_v(|E|)|E|$ с положительным наклоном в первом квадранте всегда имеется



Фиг. 13. Семейство эллипсов для системы, описываемой логарифмической характеристикой в виде прямой с наклоном -6 дБ на октаву.

Правило расчета 1. Запас устойчивости по фазе должен составлять по крайней мере 30° , а запас устойчивости по амплитуде — по крайней мере 2,5 (8 дБ).

Это правило можно обосновать математически. С помощью теории функций комплексного переменного можно показать что для разомкнутой системы следует избегать условий, при которых отставание по фазе составляет 180° , а отношение амплитуд равно единице. Согласно правилу 1, этого можно избежать при наличии определенного запаса устойчивости. Ниже мы вернемся к этому вопросу при рассмотрении амплитудно-фазовых характеристик.

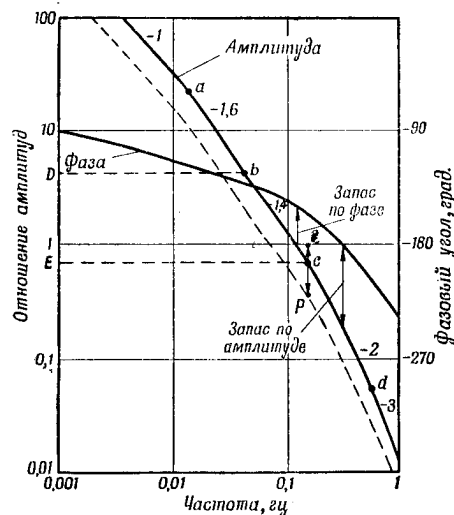
Приведенное правило расчета должно рассматриваться как грубый критерий, позволяющий специалисту, не имеющему особой математической подготовки, строить частотные характеристики систем автоматического регулирования и избежать некоторых ошибок при проектировании.

Можно доказать, что вообще опасно допускать меньший запас устойчивости по фазе и по амплитуде. Так, например при запасе устойчивости по фазе в 5° или при запасе по амплитуде, равном 1,05, система подходила бы так близко к границе неустойчивости, что решение было бы сильно колебательным; если бы при этом все или некоторые частотные характеристики системы были получены путем расчета а не экспериментально, то даже незначительные факторы могли бы сделать систему полностью неустойчивой.

Может случиться, что, хотя предыдущее расчетное правило удовлетворено, замкнутая система все же будет слишком колебательной. В конце этой статьи будет дан пример иллюстрирующий такой случай. Однако для большого числа физических систем приведенное правило обеспечивает достаточно хорошее протекание процесса регулирования. Правило определяет запас устойчивости, который часто бывает достаточным и лишь в редких случаях не приводит к цели. Чтобы быть уверенным в том, что система будет удовлетворительно работать, специалист может без особых трудностей и не углубляясь в теорию использовать готовый набор частотных характеристик, определяющих хорошую работу системы, и проследить за тем, чтобы характеристики, полученные им при решении данной конкретной задачи, совпадали или были близки к соответствующим стандартным характеристикам.

1. Изображение частотных характеристик и расчетные критерии 29

которой отношение амплитуды на выходе разомкнутой системы к амплитуде на входе равно единице, называется „частотой среза“. Предположим, что имеется лишь одна частота среза. Назовем этот случай „нормальным“. Если мы из 180° вычтем отставание по фазе при частоте среза, то получим запас устойчивости по фазе; иначе говоря, запас устойчивости по фазе равен сумме 180° и фазового угла.



Фиг. 11. Логарифмические амплитудная и фазовая частотные характеристики разомкнутой системы регулятор — двигатель.

Фазовый угол считают отрицательным при отставании и положительным при опережении фазы. Если имеется несколько частот среза, то указанная сумма определяется для каждой частоты среза и наименьшая из них будет запасом устойчивости по фазе. Для рассматриваемого выше примера запас устойчивости по фазе показан на фиг. 11. Он равен приблизительно 30° . За ось абсцисс удобно принять горизонтальную прямую при значении отношения амплитуд (усиление по терминологии Боде), равном единице.

Точку, в которой отставание по фазе равно 180° (фазовый угол равен -180°), назовем „точкой перехода фазы“.

Таблица 1

ФУНКЦИЯ $\alpha(\omega)$ и $\beta(\omega)$

Система	Передаточная функция разомкнутой системы	$\alpha(\omega)$	$\beta(\omega)$
— 6	$\frac{K_v}{p}$	0	ω
— 6—12	$\frac{K_v}{p(1 + \tau_1 p)}$	$-\omega^2 \tau_1$	ω
— 6—12—6—12	$\frac{K_v(1 + \tau_2 p)}{p(1 + \tau_1 p)(1 + \tau_3 p)}$	$\frac{-\omega^2 [(\tau_1 + \tau_3 - \tau_2) + \omega^2 \tau_1 \tau_2 \tau_3]}{1 + \omega^2 \tau_2^2}$	$\frac{\omega \{1 - \omega^2 [\tau_1 \tau_3 - \tau_2 (\tau_1 + \tau_3)]\}}{1 + \omega^2 \tau_2^2}$

6. Насыщение в системах с тахометрической обратной связью 237

как другие ее части являются линейными, то метод можно распространить на многоконтурные системы. Нужно получить частотное выражение, связывающее насыщающиеся величины с θ_i , например в случае тахометрической обратной связи $\theta_0/\theta_i(j\omega)$, и затем разложить полученное выражение аналогично уравнению (18) с тем, чтобы провести графическое построение.

Следует напомнить, что выше было сделано предположение о том, что искажения сигналов малы. В некоторых приложениях бывает невозможно сделать такое допущение.

Анализ переходных процессов простых следящих систем с насыщением

Некоторые свойства переходного процесса следящих систем с насыщением могут быть установлены из рассмотрения частотных характеристик. Однако этот метод сомнителен, так как для нелинейных систем принцип суперпозиции несправедлив. Хотя исследование переходных процессов следящих систем с насыщением и не является целью настоящей статьи, тем не менее можно указать некоторые возможные способы решения этой задачи для простых систем с насыщением $K_v(|E|)$ с амплитудными частотными характеристиками с наклоном -6 дБ на октаву, а также характеристиками типа $+6 -12$.

Для простой системы типа -6 характеристическое дифференциальное уравнение имеет вид

$$\frac{dE}{dt} + K_v(|E|)E = 0. \quad (21)$$

Это уравнение можно решить методом разделения переменных и получить решение для времени как функции ошибки E в виде

$$t = - \int \frac{dE}{K_v(|E|)E} + C. \quad (22)$$

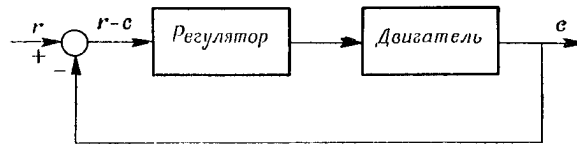
Интегрирование можно выполнить графически по обратной кривой насыщения.

Для нормальной кривой насыщения постоянная времени в случае насыщения больше постоянной времени в линейном

новления факторов, вызывающих расхождение теоретических и экспериментальных результатов, и для расчета систем автоматического регулирования.

Расчетные критерии, основанные на использовании амплитудной и фазовой характеристик

По амплитудной и фазовой характеристикам разомкнутой системы определяют характеристики замкнутой системы [1]. Обозначим через r скорость, которую необходимо поддерживать постоянной (заданный входной сигнал), а через



Фиг. 10. Замкнутая система регулятор—двигатель.

c — действительную скорость двигателя (отсчитываемую от установившегося значения). Схема регулирования скорости двигателя изображена на фиг. 10. Для нормальной работы необходимо, чтобы система была устойчивой. По виду амплитудной и фазовой характеристик разомкнутой системы (см. фиг. 5) можно судить об ее устойчивости в замкнутом состоянии и вывести общие правила расчета такой системы.

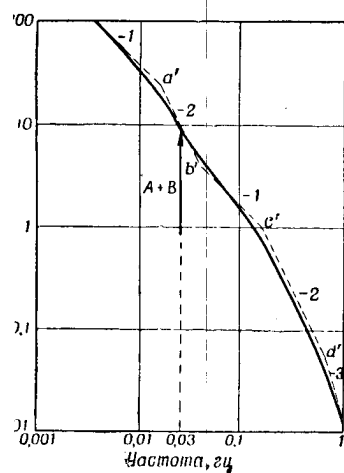
Амплитудные и фазовые кривые удобно строить на одном графике. Для рассматриваемого примера с двигателем такие кривые приведены на фиг. 11.

Первая работа, посвященная частотному методу, была опубликована Найквистом [2], который математически решил задачу об устойчивости линейных замкнутых систем, сведя задачу к изучению характеристик разомкнутой системы. Боде развил этот метод и ввел понятия усиления [3] и запаса устойчивости по фазе. Используя эти понятия, можно дать некоторые простые правила расчета.

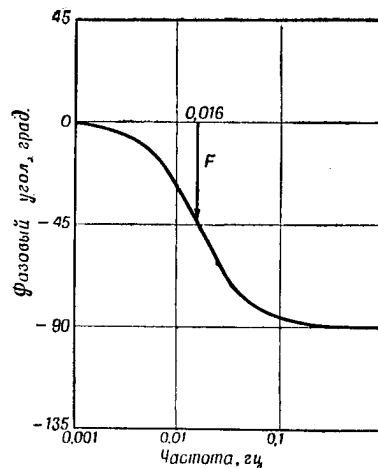
Предположим, что к системе S приложено малое возмущение δ , которое, возможно, исчезает со временем. Если переходный процесс, вызванный этим возмущением, затухает со временем, то говорят, что система S устойчива. В противном случае система S неустойчива. Частота, при

1. Изображение частотных характеристик и расчетные критерии 25

Элемент измерения скорости двигателя предполагается отключенным от входа регулятора. Если входной сигнал e (регулирующая переменная) регулятора изменяется синусоидально, то выход s (регулируемая переменная) двигателя, а именно скорость двигателя (отсчитываемая от установившегося значения скорости), практически будет изменяться тоже синусоидально. Можно показать, что амплитудная характеристика системы, состоящей из регулятора и двигателя, получается



Фиг. 6. Логарифмическая амплитудная частотная характеристика разомкнутой системы регулятор — двигатель.



Фиг. 7. Логарифмическая фазовая частотная характеристика двигателя Дизеля.

путем простого сложения ординат амплитудных характеристик, построенных в логарифмических координатах. Таким образом, если мы сложим векторы A и B на фиг. 3 и 4 при частоте $0,03$ гц, то получим вектор $A+B$ на фиг. 6, где представлена амплитудная характеристика разомкнутой системы регулятор — двигатель. Удобно принять горизонтальную линию с индексом 1 за ось абсцисс, так как логарифм единицы есть нуль. Векторы A , B и $A+B$ отсчитываются от этой оси.

Чтобы применять метод сложения ординат, необходимо вход регулятора и вход объекта регулирования (регулируемая система) измерять в одних и тех же единицах.

и фазу. Его можно записать в виде $\theta_o = B\theta_{im} \sin(\omega t - \varphi)$, где B и φ являются функциями ω (амплитудная и фазовая частотные характеристики следящей системы).

Следовательно, сигнал ошибки будет равен

$$E = \theta_i - \theta_o = \theta_{im} \sin \omega t - B\theta_{im} \sin(\omega t - \varphi) \quad (27)$$

или в векторной записи

$$E = \theta_{im} \angle 0^\circ - B\theta_{im} \angle -\varphi^\circ. \quad (28)$$

Величину E можно найти по частотной характеристике, приведенной на фиг. 14. При низких частотах B приблизительно равно единице, а $\varphi \simeq 0^\circ$, так что амплитуда сигнала ошибки очень мала. Однако при высоких частотах даже в хорошо стабилизированных системах автоматического регулирования B может быть около 1,4 и φ около 90° .

При определенной частоте напряжение ошибки должно быть равно корню квадратному из $[1 + (1,4)^2]$ или сигналу на входе, умноженному на 1,72. Таким образом, при высоких частотах действительный сигнал ошибки, который подается на усилитель, обычно много больше входного сигнала и может вызывать непредвиденные эффекты насыщения.

Заключение

В этой статье даются некоторые полезные методы анализа следящих систем с насыщением. Устройств, которые компенсировали или уменьшали бы влияние насыщения, предложено не было, хотя из анализа можно получить для этого необходимые рекомендации. Будет ли эффект насыщения вредным или нет, зависит от применения той или иной исследуемой следящей системы.

Изложенные здесь принципы ни в коем случае не следует рассматривать как естественные законы. Можно утверждать лишь то, что эти принципы применимы в нескольких простых случаях и, следовательно, могут использоваться наряду с другими аналитическими средствами при расчете нелинейных систем.

Наконец, вид полученной здесь передаточной функции тахометрически стабилизированной следящей системы очень удобен для анализа и синтеза таких систем и его область применения не ограничивается только вопросами насыщения.

6. Насыщение в системах с тахометрической обратной связью 241

ПРИЛОЖЕНИЕ I

Передаточная функция следящей системы с тахометрической обратной связью и пассивным фильтром высоких частот

Обращаясь к блок-схеме на фиг. 1, имеем

$$\theta_o = \frac{\theta_m}{N}, \quad \theta_T = \frac{\theta_m}{N_T},$$

$$\frac{\theta_m}{V_1} = \frac{K_m}{p(1 + \tau_m p)},$$

$$V_1 = K_2 E_2,$$

$$E_2 = K_1 E_1 - \frac{K_3 \tau_2 p}{1 + \tau_2 p} V_3,$$

$$V_3 = K_T p \theta_T,$$

$$E_1 = K_e (\theta_i - \theta_o) = K_e E.$$

Объединяя эти уравнения, получаем

$$\theta_o = \frac{K_m K_2}{N p (1 + \tau_m p)} \left[K_1 K_e E - K_3 K_T \frac{\tau_2 p^2}{1 + \tau_2 p} \frac{N \theta_o}{N_T} \right]$$

или

$$\theta_o \left[1 + \frac{K_m K_2 K_3 K_T}{N_T} \frac{\tau_2 p^2}{p(1 + \tau_m p)(1 + \tau_2 p)} \right] = \frac{K_m K_2 K_1 K_e}{N} \frac{E}{p(1 + \tau_m p)}.$$

Введем следующие обозначения:

$K_v = \frac{K_e K_1 K_2 K_m}{N}$ — полный коэффициент усиления разомкнутой системы;

$\alpha = \frac{K_T K_3 K_2 K_m}{N_T}$ — коэффициент усиления цепи тахометрической обратной связи. Тогда

$$\theta_o [p(1 + \tau_m p)(1 + \tau_2 p) + \alpha \tau_2 p^2] = K_v E (1 + \tau_2 p)$$

и

$$\frac{\theta_o}{E} = \frac{K_v (1 + \tau_2 p)}{p \{ \tau_m \tau_2 p^2 + [\tau_m + (1 + \alpha) \tau_2] p + 1 \}}. \quad (29)$$

Выражение в знаменателе можно преобразовать к виду

$$(1 + \tau_1 p)(1 + \tau_3 p),$$

где

$$\tau_1 \tau_3 = \tau_m \tau_2,$$

$$\tau_1 + \tau_3 = \tau_m + (1 + \alpha) \tau_2,$$

или

$$\frac{\tau_1}{\tau_2} \frac{\tau_3}{\tau_2} = \frac{\tau_m}{\tau_2} \quad (30)$$

и

$$\frac{\tau_1}{\tau_2} + \frac{\tau_3}{\tau_2} = \frac{\tau_m}{\tau_2} + 1 + \alpha. \quad (31)$$

Почти во всех практических системах τ_2 имеет тот же порядок, что и τ_m , в то время как $\alpha \gg 1$. Следовательно, $1 + \tau_m/\tau_2$ обычно мало по сравнению с α . Для устойчивости желательно также, чтобы τ_3 равнялось $1/8\tau_2$, тогда как при высоких коэффициентах усиления системы K_v величина τ_1 всегда больше τ_2 . Поэтому τ_3/τ_2 обычно значительно меньше τ_1/τ_2 .

Тогда в уравнении (31) $\tau_1/\tau_2 \simeq \alpha$, или $\tau_1 = \alpha\tau_2$ и, подставляя это в уравнение (30), получаем

$$\frac{\tau_3}{\tau_2} = \frac{\tau_m/\tau_2}{\alpha},$$

или

$$\tau_3 = \frac{\tau_m}{\alpha}.$$

Следовательно,

$$\frac{\theta_o}{E} \simeq \frac{K_v(1 + \tau_2 p)}{p(1 + \alpha\tau_2 p)\left(1 + \frac{\tau_m}{\alpha} p\right)}. \quad (32)$$

ПРИЛОЖЕНИЕ II

$\alpha(\omega)$ и $\beta(\omega)$ для системы типа — 6 — 12 — 6 — 12

$$\begin{aligned} C(j\omega) &= \frac{B(j\omega)}{A(j\omega)} = j\omega \frac{(1 + j\omega\tau_1)(1 + j\omega\tau_3)}{(1 + j\omega\tau_2)} = \\ &= j\omega \frac{[1 - \omega^2\tau_1\tau_3 + j\omega(\tau_1 + \tau_3)](1 - j\omega\tau_2)}{(1 + j\omega\tau_2)(1 - j\omega\tau_2)} = \\ &= j\omega \frac{\{1 - \omega^2[\tau_1\tau_3 - \tau_2(\tau_1 + \tau_3)] + j\omega[(\tau_1 + \tau_3 - \tau_2) + \omega^2\tau_1\tau_2\tau_3]\}}{1 + \omega^2\tau_2^2}, \end{aligned}$$

6. Насыщение в системах с тахометрической обратной связью 243

или

$$C(j\omega) = -\omega^2 \frac{[(\tau_1 + \tau_3 - \tau_2) + \omega^2 \tau_1 \tau_2 \tau_3]}{1 + \omega^2 \tau_2^2} + j\omega \frac{\{1 - \omega^2 [\tau_1 \tau_3 - \tau_2 (\tau_1 + \tau_3)]\}}{1 + \omega^2 \tau_2^2}. \quad (33)$$

Следовательно,

$$\alpha(\omega) = -\omega^2 \frac{[(\tau_1 + \tau_3 - \tau_2) + \omega^2 \tau_1 \tau_2 \tau_3]}{1 + \omega^2 \tau_2^2} \quad (34)$$

и

$$\beta(\omega) = \omega \frac{\{1 - \omega^2 [\tau_1 \tau_3 - \tau_2 (\tau_1 + \tau_3)]\}}{1 + \omega^2 \tau_2^2}. \quad (35)$$

ЛИТЕРАТУРА

1. Harris H., *Trans. AIEE*, **66**, 83 (1947).
2. Теория следящих систем, под ред. Х. Джеймса, Н. Никольса и Р. Филлипса, Издательство иностранной литературы, 1953.
3. Стокер Д., Нелинейные колебания в механических и электрических системах, Издательство иностранной литературы, 1952.
4. Rauscher M., *J. Appl. Mech.*, **5**, 169A (1938).
5. Ден Гартог Дж., Теория колебаний, Гостехиздат, 1942.
6. Minorsky N., *Non-Linear Mechanics*, Ann Arbor, Mich., 1947.
7. McDonald D., „Non-Linear Techniques for Improving Servo Performance“, *Proc. Nat. Electronic Conference*, vol. 6, 1950, p. 400.

7. ПРИБЛИЖЕННЫЙ ЧАСТОТНЫЙ МЕТОД ПРЕДСТАВЛЕНИЯ НАСЫЩЕНИЯ И ЗОНЫ НЕЧУВСТВИТЕЛЬНОСТИ¹⁾

Г. Честнат

Введение

Результаты работ Кохенбургера [1] и Джонсона [2], посвященных применению частотных методов для расчета нелинейных систем автоматического регулирования, можно применить в приближенном виде к задачам с нелинейностями типа насыщения и зоны нечувствительности. Можно показать, что каждая из этих нелинейностей может быть представлена с помощью „эквивалентного линейного“ коэффициента в качестве ее описывающей функции. Типичные примеры насыщения и зоны нечувствительности используются для иллюстрации этого физического приближения задачи.

В данной статье определяется изменение эквивалентного усиления для рассматриваемых типов нелинейностей при синусоидальных входных сигналах и приводится дополнительное соотношение между изменениями коэффициентов усиления, соответствующих насыщению и зоне нечувствительности. Проведено сравнение переходных характеристик при ступенчатом входном воздействии для следящих систем, имеющих либо насыщение, либо зону нечувствительности с результатами, полученными при использовании описывающей функции в форме эквивалентного линейного коэффициента. Как точный, так и приближенный методы анализа показывают, что, когда имеются насыщение или зона нечувствительности, в системе при некоторых условиях могут возникнуть незатухающие колебания.

Кроме получения качественных сведений о характеристике системы с насыщением или зоной нечувствительности, рассматриваемый в статье метод может быть использован для приближенных расчетов при наличии нелинейностей в более сложных многоконтурных системах.

¹⁾ Chestnut H., *Trans. ASME*, 76, № 8, 1345—1363 (1954).

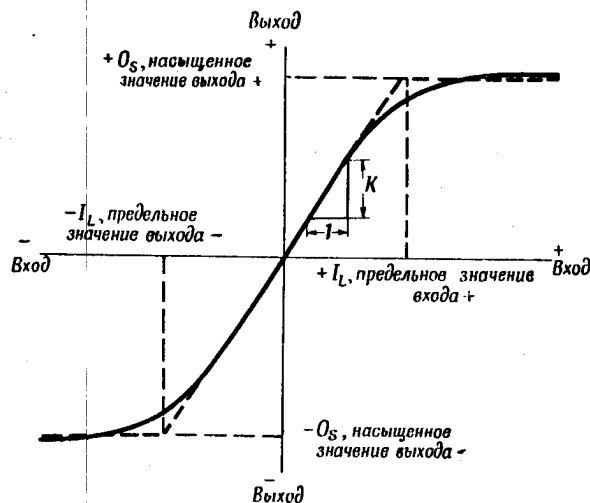
7. Частотный метод представления типовых нелинейностей 245

Типовые задачи с насыщением и зоной нечувствительности

Для раскрытия физической сущности явлений, происходящих при наличии насыщения и зоны нечувствительности, рассмотрим несколько характерных примеров с такими видами нелинейности.

Насыщение

Линейная зависимость между выходом и входом любого усилительного устройства (механического, электрического или гидравлического) справедлива в лучшем случае лишь



Фиг. 1. Характеристика элемента с насыщением.

— действительная кривая; - - - идеализированная кривая.

в ограниченной области значений входной функции. Для входных величин, превышающих некоторое предельное значение, выход перестает быть пропорциональным входу и действительное отношение выход/вход уменьшается. Это явление называется насыщением. На фиг. 1 сплошной кривой изображена характеристика насыщения. Вход и выход могут быть выражены в любых целесообразных единицах с тем

условием, чтобы между выходом и входом существовало однозначное численное соотношение. Чтобы показать влияние такой характеристики насыщения на описывающую функцию элемента, ниже будут рассмотрены примеры насыщения электронного усилителя по напряжению и насыщения двигателя постоянного тока по моменту.

Насыщение по напряжению в электронном усилителе

Для упрощения анализа насыщения электронного усилителя предположим, что действительная статическая характеристика, связывающая вход и выход усилителя, симметрична и что ее можно приближенно заменить идеализированной характеристикой насыщения, как показано на фиг. 1 пересекающимися пунктирными прямыми. Одна прямая проходит через начало координат и имеет тот же наклон K , что и действительная характеристика усилителя. Две другие линии имеют наклон, равный нулю, и значения $+O_S$ и $-O_S$ приближенно равны значениям выхода при насыщении соответственно для больших положительных и больших отрицательных значений входа.

Предельные значения входа $\pm I_L$ являются значениями, при которых линия с наклоном K пересекает прямые с нулевым наклоном и после которых значение выхода остается постоянным. Таким образом,

$$I_L = \frac{O_S}{K}. \quad (1)$$

Так как кривая, по предположению, симметрична, то предельные значения для положительных и отрицательных входов одинаковы и соответствующие положительные и отрицательные значения выхода при насыщении также равны.

Если ко входу усилителя, имеющего идеализированную характеристику (см. фиг. 1), приложено синусоидальное напряжение

$$i_a(t) = I_M \sin t, \quad (2)$$

где $I_M > I_L$, то выход $O_a(t)$ будет иметь форму, показанную на фиг. 2.

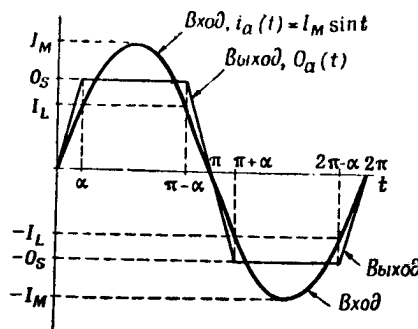
Для значений $-I_L < i_a < +I_L$ выход $O_a(t)$ будет пропорционален $i_a(t)$ с коэффициентом пропорциональности K .

7. Частотный метод представления типовых нелинейностей 247

Для значений $i_a > I_L$ выход будет постоянным и равным O_S (значение $O_a(t)$ при насыщении), для значений $i_a < -I_L$ выход также будет постоянным и равным $-O_S$. Угол α , при котором $O_a(t)$ перестает быть пропорциональным $i_a(t)$, равен

$$\alpha = \arcsin \frac{I_L}{I_M}. \quad (3)$$

Заметим, что для электронного усилителя форма кривых на фиг. 2 и значение α в интересующей нас области частот



Фиг. 2. Выход $O_a(t)$ элемента с насыщением как функция времени для входного сигнала $i_a(t) = I_M \sin t$, где $I_M > I_L$ (I_L — предельное значение входа).

$$\alpha = \arcsin (I_L/I_M).$$

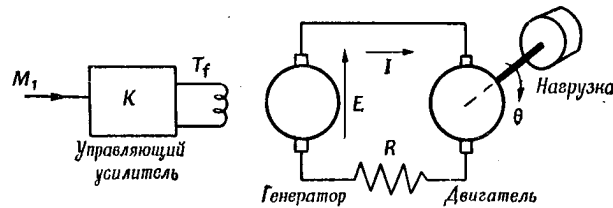
не зависят от частоты синусоидального входного сигнала. Для таких синусоидальных входов основная гармоника выхода O_{a1} находится в фазе с основной гармоникой входа I_M и отношение между ними является величиной, зависящей от α . Таким образом,

$$\frac{O_{a1}}{I_M} = G'(\alpha), \quad (4)$$

где $G'(\alpha)$ — описывающая функция усилителя с насыщением. Поскольку α и I_L/I_M связаны равенством (3), то очевидно, что отношение O_{a1}/I_M из уравнения (4) также можно выразить через I_L/I_M , а не через α .

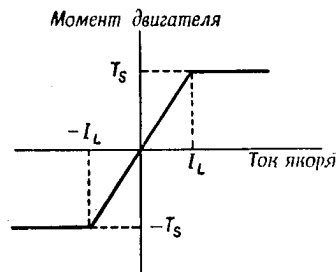
Насыщение момента двигателя

Явление насыщения момента двигателя в системе регулирования скорости двигателя постоянного тока, изображенной на фиг. 3, может оказать нежелательное влияние



Фиг. 3. Схема регулирования генератор — двигатель с насыщением двигателя по моменту.

на работу системы. Вследствие эффектов размагничивания, вызываемых неполной компенсацией обмоток и магнитным насыщением, момент двигателя, являющийся функцией тока якоря, может иметь идеализированную характеристику насыщения, показанную на фиг. 4,



Фиг. 4. Идеализированная характеристика насыщения момента двигателя в зависимости от тока якоря.

функцией максимального тока якоря I_M . Таким образом, основную гармонику момента двигателя можно выразить формулой

$$\text{Момент} = K_T' I, \quad (5)$$

7. Частотный метод представления типовых нелинейностей 249

где K'_T — коэффициент эффективного момента, принимающий следующие значения:

$$K'_T = K_T \text{ при } -I_L < I_M < +I_L,$$

$$K_T > K'_T > 0 \text{ при } I_M > I_L \text{ и } I_M < -I_L.$$

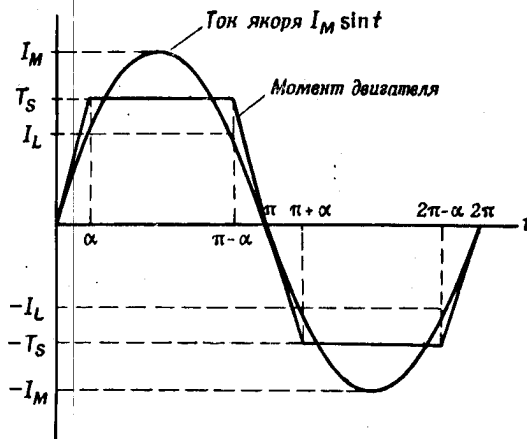
Для отдельных элементов системы, соответствующих фиг. 3, можно составить следующие уравнения:

$$KM_1 = (T_f s + 1) E, \quad (6)$$

$$E = RI + K_e s \theta, \quad (7)$$

$$K'_T I = Js^2 \theta, \quad (8)$$

где M_1 — входное напряжение управляющего усилителя; K — коэффициент усиления управляющего усилителя; T_f — постоянная времени управляющего усилителя.



Фиг. 5. Момент двигателя и ток якоря как функции времени при синусоидальном изменении тока якоря и наличии насыщения по моменту.

$$\alpha = \arcsin (I_L / I_M).$$

Исключая из последних уравнений E и I , имеем

$$\frac{\theta}{M_1} = \frac{K/K_e}{s(T_f s + 1)(T'_M s + 1)}, \quad (9)$$

где

$$T'_M = \frac{JR}{K'_T K_e}$$

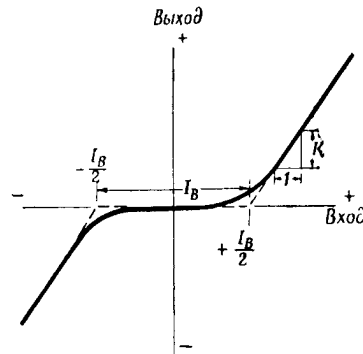
— эффективная постоянная времени двигателя.

Из этого выражения для постоянной времени двигателя следует, что T'_M увеличивается с усилением насыщения момента (т. е. при уменьшении значения K'_T). Нежелательные явления, возникающие при увеличении постоянной времени двигателя, хорошо известны.

Зона нечувствительности

Зона нечувствительности, или люфт, является другим видом нелинейности, также часто встречающимся в системах автоматического регулирования. Хотя обычно зона нечувствительности связана с несовершенством процесса изготовления, можно привести ряд примеров, когда для ограничения действия намеренно вводится зона нечувствительности.

Характеристика зоны нечувствительности представлена графически на фиг. 6. Пунктирной линией показана идеализированная кривая выхода элемента с зоной нечувствительности как функции входа. Сплошной кривой показана характеристика, более близкая к реальной, в которой наклон кривой вблизи $\pm I_B/2$ изменяется постепенно, а не скачком, как у идеализированной характеристики в точке разрыва $\pm I_B/2$. Для входных сигналов $I > I_B/2$ выход пропорционален входу с коэффициентом пропорциональности K , т. е.



Фиг. 6. Характеристика зоны нечувствительности.

— действительная кривая;
 - - - идеализированная кривая;
 I_B — ширина зоны нечувствительности.

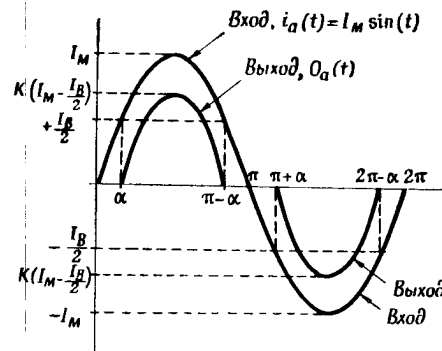
$$K = \frac{\text{Выход}}{I - I_B/2}. \quad (10)$$

7. Частотный метод представления типовых нелинейностей 251

Если к элементу, имеющему идеализированную характеристику, изображенную на фиг. 6, приложено синусоидальное возмущение

$$i_a(t) = I_M \sin t, \quad (11)$$

амплитуда которого $I_M > I_B/2$, то выход $O_a(t)$ будет иметь вид, показанный на фиг. 7. Для значений входа $-I_B/2 < i_a < I_B/2$, т. е. в пределах зоны нечувствительности,



Фиг. 7. Выход $O_a(t)$ элемента с зоной нечувствительности как функция времени для входа $i_a(t) = I_M \sin t$ при $I_M > I_B/2$.

$$\alpha = \arcsin(I_B/2I_M).$$

выходной сигнал будет равен нулю. Для значений $i_a > +I_B/2$ выход будет пропорционален $i_a - I_B/2$ с коэффициентом пропорциональности K из уравнения (10). Для $i_a < -I_B/2$ выход пропорционален $i_a + I_B/2$ также с коэффициентом пропорциональности K . Угол α , при котором $O_a(t)$ отлично от нуля, равен

$$\alpha = \arcsin \frac{I_B}{2I_M}. \quad (12)$$

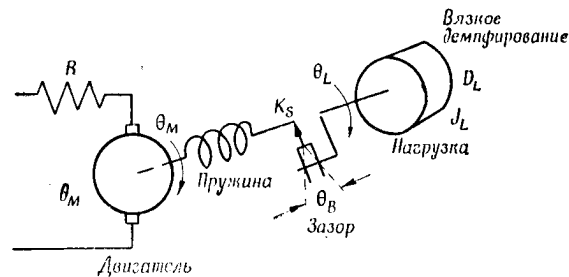
Заметим, что α не зависит от частоты синусоидального входа (см. фиг. 7) и что основная гармоника выхода находится в фазе с колебанием на входе. Амплитуда выхода является функцией отношения $I_B/2I_M$.

Следует заметить, что входной сигнал относится ко входу определенного элемента и в некоторых случаях может

представлять собой разностный сигнал или сигнал ошибки, а не исходный входной сигнал, приложенный к системе. Именно такой случай имеет место в механической системе при наличии люфта, рассматриваемой ниже.

Нагрузка, приводимая в движение от двигателя через упругий вал и элемент с люфтом

На фиг. 8 схематически показан двигатель постоянного тока, связанный через нелинейный элемент типа люфта и упругий вал с инерционной нагрузкой, имеющей вязкое



Фиг. 8. Двигатель постоянного тока, соединенный с инерционной нагрузкой, имеющей вязкое демпфирование, через упругий вал с зазором.

демпфирование. С помощью фиг. 9 и 10 можно получить описывающую функцию для этих элементов. Как видно из фиг. 9, разность между положениями оси двигателя θ_M и нагрузки θ_L определяет упругий момент.

При значениях $-\theta_B/2 < (\theta_M - \theta_L) < +\theta_B/2$ упругий момент отсутствует. Если угол $\theta_M - \theta_L$ изменяется по синусоидальному закону, то упругий момент будет изменяться так, как показано на фиг. 10, из которой можно также видеть, что основная гармоника упругого момента находится в фазе с входной угловой величиной $(\theta_M - \theta_L)$, причем коэффициент пропорциональности K'_s будет функцией отношения

$$\frac{\theta_B}{2(\theta_M - \theta_L)_{\max}}$$

7. Частотный метод представления типовых нелинейностей 253

K'_S по существу представляет собой эффективный коэффициент жесткости, значение которого является функцией указанного отношения амплитуд. Используя этот коэффициент, получаем уравнение для момента, действующего на нагрузку:

$$J_L s^2 \theta_L + D_L s \theta_L = K'_S (\theta_M - \theta_L). \quad (13)$$

Уравнение моментов, действующих на двигатель, имеет вид

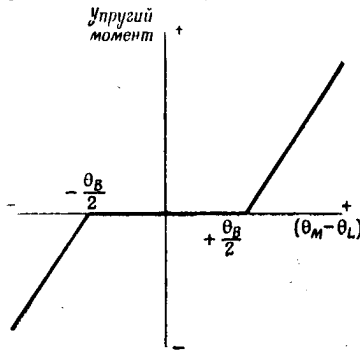
$$K_T I - K'_S (\theta_M - \theta_L) = J_M s^2 \theta_M, \quad (14)$$

а уравнение напряжений

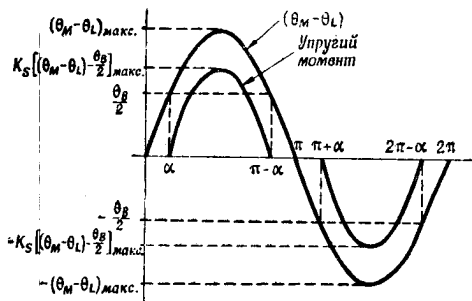
$$E = R I + K_e s \theta_M. \quad (15)$$

Объединяя уравнения (14) и (15), имеем

$$\frac{K_T E}{R} - K_S (\theta_M - \theta_L) = J_M s^2 \theta_M + \frac{K_e K_T}{R} s \theta_M. \quad (16)$$



Фиг. 9. Упругий момент, подведенный к нагрузке через пружину (см. фиг. 8), как функция разности положений валов двигателя и нагрузки при общем зазоре θ_B .

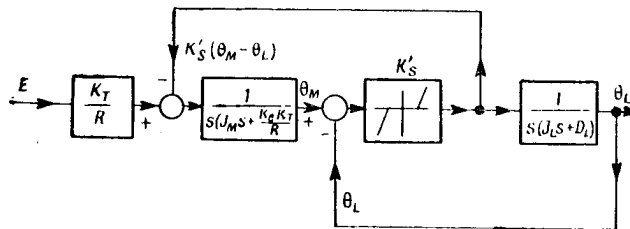


Фиг. 10. Упругий момент и разность угловых положений двигателя и нагрузки как функции времени при синусоидальном изменении разности угловых положений двигателя и нагрузки при общем зазоре θ_B .

$$\alpha = \arcsin \frac{\theta_B}{2} (\theta_M - \theta_L)_{\text{макс.}}$$

Уравнениям (13) и (16) соответствует блок-схема, приведенная на фиг. 11, где величина K'_S условно представлена

символом обратной связи. При аналитических расчетах можно считать, что K'_S является действительной линейной величиной,



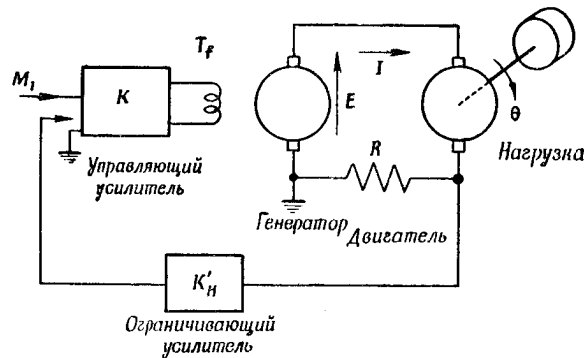
Фиг. 11. Блок-схема системы, представленной на фиг. 8.

зависящей от отношения

$$\frac{\theta_B}{2(\theta_M - \theta_L)_{\text{макс.}}}$$

Влияние ограничения тока

Несколько иной пример зоны нечувствительности приведен на фиг. 12, где показана система регулирования

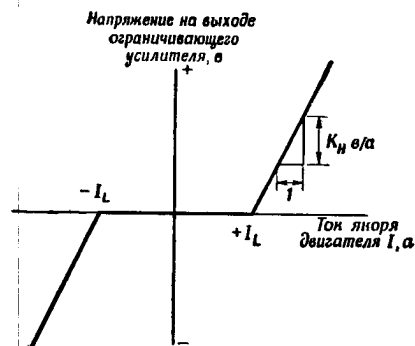


Фиг. 12. Система регулирования генератор—двигатель с ограничением обратной связи по току.
 K — отношение напряжения генератора к напряжению на входе управляющего усилителя.

генератор—двигатель, в которую введена цепь обратной связи для ограничения максимального тока двигателя. Сигнал с выхода ограничивающего усилителя подается в противофазе

7. Частотный метод представления типовых нелинейностей 255

на вход управляющего усилителя, ограничивая тем самым напряжение генератора и, следовательно, ток якоря. Этот пример ограничения тока по методу рассмотрения подобен примерам ограничения скорости, ускорения, давления или какого-либо другого подходящего параметра, причем соответствующая величина включается в цепь обратной связи через ограничивающий усилитель.



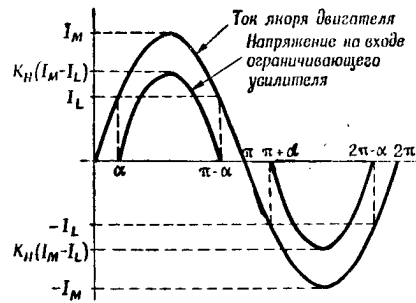
Фиг. 13. Характеристика ограничивающего усилителя в зависимости от тока якоря двигателя.

Характеристика выходного напряжения ограничивающего усилителя как функция тока якоря двигателя показана на фиг. 13. Для положительных значений тока, меньших I_L , где I_L — ток, при котором начинается ограничение, напряжение в обратной связи отсутствует. Для значений тока, превышающих I_L , напряжение обратной связи увеличивается пропорционально току, причем коэффициент пропорциональности равен K_H .

Для синусоидального изменения тока якоря с амплитудой, превышающей I_L , напряжение на выходе ограничивающего усилителя будет изменяться, как показано на фиг. 14. Основная составляющая напряжения на выходе ограничивающего усилителя находится в фазе с током якоря двигателя.

Амплитуда основной составляющей ограничивающего напряжения является функцией максимального значения тока якоря, причем коэффициент пропорциональности равен K'_H ,

где $0 < K'_H < K_H$. Таким образом, ограничивающее напряжение на выходе усилителя можно выразить в виде произведения $K'_H I$.



Фиг. 14. Ток якоря и напряжение на выходе ограничивающего усилителя как функции времени при синусоидальном изменении тока якоря.

$$\alpha = \arcsin(I_L/I_M).$$

Ограничение начинается с I_L .

Согласно фиг. 12, можно записать следующие уравнения: сигнал на входе управляющего усилителя

$$M_1 - K'_H I, \quad (17)$$

напряжение генератора

$$E = \frac{K(M_1 - K'_H I)}{(T_f s + 1)} = RI + K_e s \theta, \quad (18)$$

момент двигателя

$$K_T I = Js^2 \theta. \quad (19)$$

Объединяя эти уравнения, получим передаточную функцию системы в следующем виде:

$$\frac{\theta}{M_1} = \frac{K}{K_e s \left[T_f T_M s^2 + \left(T_M + T_f + \frac{K K'_H J}{K_e K_T} \right) s + 1 \right]}, \quad (20)$$

7. Частотный метод представления типовых нелинейностей 257

где

$$T_M = \frac{JR}{K_e K_T}$$

При $K'_H = 0$, т. е. в случае отсутствия ограничивающего действия, эта передаточная функция принимает обычное значение. При наличии ограничения и при $K'_H \simeq K_H$ система имеет большую постоянную времени вследствие наличия члена

$$\frac{KK'_H J s}{K_e K_T}$$

в знаменателе передаточной функции.

Определение коэффициента изменения усиления

Предыдущие примеры типичных задач с насыщением и с зоной нечувствительности были приведены, чтобы показать два общих типа нелинейных характеристик, для которых желательно установить отношение амплитуды основной гармоники выхода к амплитуде синусоидального сигнала на входе. Это отношение можно рассматривать как коэффициент изменения усиления, поскольку он указывает, во сколько раз усиление линейного элемента должно быть изменено, чтобы учесть влияние нелинейности.

Умножая линейный коэффициент элемента на коэффициент изменения усиления, можно найти значение „эквивалентного коэффициента усиления“ элемента, имеющего заданные входную амплитуду и область нелинейности.

Таким образом, если A является коэффициентом изменения усиления, а K — линейным коэффициентом элемента, то эквивалентный линейный коэффициент K' будет

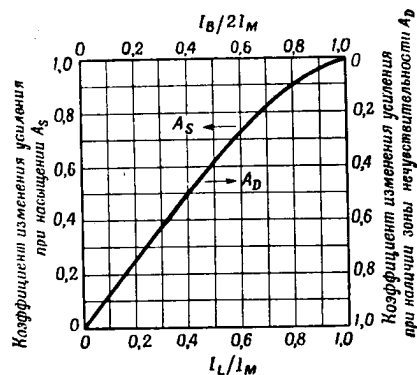
$$K' = AK, \quad (21)$$

где A является функцией амплитуды входного сигнала.

Коэффициент изменения усиления для элемента с насыщением A_s

Чтобы помочь установить взаимно дополняющие друг друга свойства коэффициентов усиления насыщения и зоны нечувствительности, повторим здесь вывод этого коэффи-

коэффициента усиления, данный Кохенбургером [3]. На фиг. 2 показан выход симметричного элемента с насыщением как функция времени при синусоидальном входе. Из графика видно, что между основными составляющими выхода и входа отсутствует фазовый сдвиг.



Фиг. 15. Коэффициенты изменения усиления для случая насыщения A_S и для случая зоны нечувствительности A_D в зависимости от отношения предельного значения входа к максимальному значению входа.

Если обозначить коэффициент усиления элемента в линейной области через K , значение входного сигнала, при котором имеет место насыщение, через I_L и максимальное значение входа через I_M , то выходной сигнал за один полупериод можно выразить следующим образом:

$$\begin{aligned} O_a(t) &= KI_M \sin t && \text{при } 0 < t < \alpha, \\ O_a(t) &= KI_M \sin \alpha && \text{при } \alpha < t < \pi - \alpha, \\ O_a(t) &= KI_M \sin t && \text{при } \pi - \alpha < t < \pi, \end{aligned} \quad (22)$$

где

$$\alpha = \arcsin \frac{I_L}{I_M}.$$

Учитывая свойство симметрии O_{a1} , основную гармонику выходного сигнала можно представить в виде

$$\begin{aligned} O_{a1} &= \frac{2KI_M}{\pi} \left(2 \int_0^{\alpha} \sin^2 t \, dt + \int_{\alpha}^{\pi-\alpha} \sin \alpha \sin t \, dt \right), \\ O_{a1} &= KI_M \left(\frac{2\alpha}{\pi} + \frac{\sin 2\alpha}{\pi} \right). \end{aligned} \quad (23)$$

Определяя коэффициент изменения усиления при насыщении A_S как отношение амплитуды первой гармоники на выходе элемента с насыщением к амплитуде первой гармоники на выходе того же элемента только с линейной характеристикой во всем диапазоне изменения входного сигнала,

7. Частотный метод представления типовых нелинейностей 259

получаем:

$$A_S = \frac{2\alpha}{\pi} + \frac{\sin 2\alpha}{\pi}. \quad (24)$$

На фиг. 15 представлена кривая A_S (левая ось ординат) как функция I_L/I_M и видно, как коэффициент усиления постепенно уменьшается при $I_L/I_M > 0,8$, тогда как при малых значениях $I_L/I_M < 0,3$ коэффициент усиления приблизительно пропорционален I_L/I_M . Как видно из фиг. 15 (правая ось ординат) и как будет показано в дальнейшем, коэффициенты изменения усиления при насыщении и в зоне нечувствительности отчасти сходны по своей природе.

Коэффициент изменения усиления для элемента с зоной нечувствительности A_D

На фиг. 7 приведен график выхода элемента с симметричной характеристикой типа зоны нечувствительности как функции времени при синусоидальном входе. Основная составляющая выхода также находится в фазе с основной составляющей входа, как и в случае насыщения.

Если обозначить коэффициент усиления элемента в линейной зоне через K , значение входа, при котором начинается линейное изменение, через $I_B/2$ и максимальное значение входа через I_M , то выход за один полупериод можно выразить следующим образом:

$$O_a(t) = KI_M (\sin t - \sin \alpha); \quad \alpha < t < \pi - \alpha,$$

где

$$\alpha = \arcsin \frac{I_B}{2I_M}. \quad (25)$$

Вне указанного интервала $O_a(t) = 0$.

Основную составляющую выходного сигнала можно представить в виде

$$\begin{aligned} O_{a1} &= \frac{2KI_M}{\pi} \int_0^{\pi-\alpha} (\sin t - \sin \alpha) \sin t dt = \\ &= KI_M \left(1 - \frac{2\alpha}{\pi} - \frac{\sin 2\alpha}{\pi} \right). \end{aligned} \quad (26)$$

Определяя коэффициент усиления для зоны нечувствительности как отношение амплитуды первой гармоники

выходного сигнала элемента с зоной нечувствительности к амплитуде первой гармоники выхода того же элемента, но с линейной характеристикой во всем диапазоне изменения входного сигнала, получим

$$A_D = 1 - \frac{2\alpha}{\pi} - \frac{\sin 2\alpha}{\pi}, \quad (27)$$

или

$$A_D = 1 - A_S \quad (28)$$

при соответствующем определении α в обоих случаях.

Уравнение (28) указывает дополнительное соотношение между коэффициентами изменения усиления при насыщении и наличии зоны нечувствительности, о которых упоминалось выше. Важно отметить, что условия „покоя“ для двух типов нелинейностей также дополняют друг друга. В то время как при бездействии для системы с насыщением $I_L/I_M \gg 1$ и $A_S = 1,0$, для элемента с зоной нечувствительности $I_B/2I_M \gg 1$ и $A_D = 0$. Если для больших значений входа $I_L/I_M \ll 1$ величина A_S в пределе достигает нуля и характеристика системы может сильно отклониться от линейной; то для $I_B/2I_M \ll 1$ величина A_D стремится к единице и характеристика системы стремится к линейному виду.

Короче говоря, для малых значений входных сигналов, подаваемых на элемент с насыщением, характеристика системы будет линейной; для больших же входных сигналов система работает при значительно уменьшенном усилении. Для малых входных сигналов, подаваемых в систему с зоной нечувствительности, система работает так, как если бы элемент с зоной нечувствительности имел пониженное усиление; для больших входных сигналов влияние зоны нечувствительности будет мало сказываться, и система при этом имеет тенденцию работать, как если бы элемент был линейным.

В двух последующих разделах иллюстрируется применение этих коэффициентов изменения усиления для получения эквивалентных коэффициентов усиления специальных систем.

Влияние насыщения на переходный процесс системы автоматического регулирования положения

Используя линейный анализ систем автоматического регулирования, мы покажем, что для получения желаемых характеристик системы в установившемся или переходном режиме

7. Частотный метод представления типовых нелинейностей 261

можно примерно с одинаковым успехом применять как метод последовательной стабилизации, так и метод параллельной стабилизации (обратной связью). Практически, когда система подвергается воздействию таких больших входных сигналов, какие наблюдаются в течение начального периода синхронизации, влияние насыщения в различных частях системы автоматического регулирования вызывает различную реакцию системы для этих двух методов стабилизации.

Используя для представления элемента с насыщением эквивалентный линейный коэффициент, полученный в предыдущем разделе, можно при помощи частотного метода анализа установить характер ожидаемого переходного процесса. Ниже мы сравним частотные и переходные характеристики систем с последовательной и параллельной стабилизацией. Мы будем считать, что эти системы условно устойчивы¹⁾ и имеют примерно одинаковые характеристики в разомкнутом состоянии для линейной части и одинаковые элементы с насыщением.

Для каждой системы проводится сравнение амплитудных и фазовых частотных характеристик разомкнутой системы и переходных характеристик замкнутой системы при различных начальных значениях отношения амплитуд a_0 , где

$$a_0 = \frac{\text{Предельное значение входного сигнала элемента с насыщением}}{\text{Максимальное значение входного сигнала элемента при линейном рассмотрении}}. \quad (29)$$

Предельное значение входного сигнала элемента с насыщением (числитель) равно величине I_L , показанной на фиг. 1. Максимальное значение входного сигнала элемента при линейном рассмотрении (знаменатель) есть максимальное значение входного сигнала указанного элемента, которое получается из анализа переходного процесса системы в предположении ее линейности. Этот линейный анализ выполняется в предположении отсутствия нелинейностей, и его можно проделать, используя преобразование Лапласа или другие методы решения линейных дифференциальных уравнений.

Для устойчивой системы максимальное значение сигнала на входе элемента является его начальным максимальным

¹⁾ Т. е. могут стать неустойчивыми при уменьшении коэффициента усиления прямой цепи. — *Прим. ред.*

значением. Поэтому величина a_0 называется также „начальным отношением амплитуд“. Таким образом, из линейного анализа поведения системы в предположении отсутствия насыщения в нелинейном элементе можно определить входной сигнал, подаваемый на элемент с насыщением, и установить значение знаменателя в выражении (29) для a_0 . Зная предельное значение входного сигнала элемента с насыщением I_L , можно получить численное значение a_0 .

Для определения приближенного значения эквивалентного коэффициента усиления при начальном значении отношения амплитуд a_0 применяется характеристика фиг. 15, показывающая изменение коэффициента усиления в зависимости от $I_L/I_M(a_0)$, считая a_0 независимой переменной. Несмотря на то, что входной сигнал, подаваемый на элемент с насыщением, не будет, вообще говоря, синусоидальным при ступенчатом воздействии на входе системы, все же входной сигнал элемента с насыщением при наличии незатухающих колебаний стремится изменяться приблизительно по синусоидальному закону и поэтому использование предположения о синусоидальном входе вполне оправдано. В случаях когда незатухающие колебания не возникают, предположение о синусоидальности входа позволяет определить значение коэффициента усиления элемента с насыщением, которое определяет начальное поведение, хорошо согласуемое с поведением реальной системы. После достижения перерегулирований, если они имеют место, значение I_L/I_M (и усиления) стремится к увеличению и поведение системы, определенное с помощью ее видоизмененной амплитудной частотной характеристики, соответствует в общем поведению реальной системы.

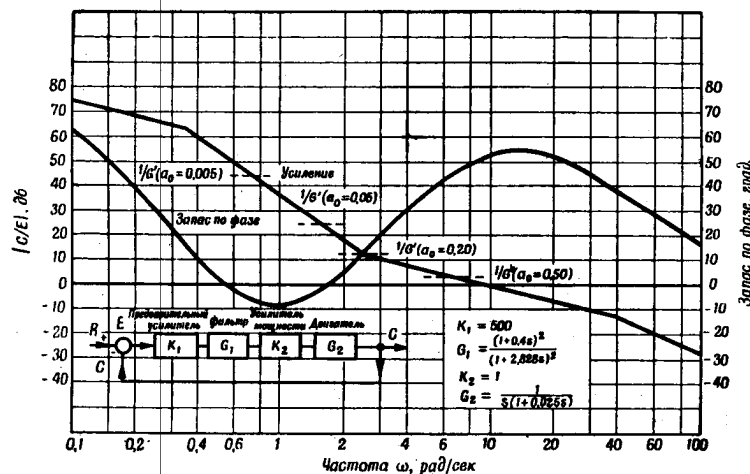
Появление незатухающих колебаний в нелинейных системах было путем рассмотрения сходящихся и расходящихся точек равновесия исследовано Кохенбургером [1] и здесь подробно не рассматривается.

Система с последовательной стабилизацией

На фиг. 16 приведены приближенные амплитудная и фазовая характеристики разомкнутой условно устойчивой системы, использующей последовательный стабилизирующий контур. Характеристики этой системы будут рассматриваться с учетом насыщения предварительного усилителя и усиления

7. Частотный метод представления типовых нелинейностей 263

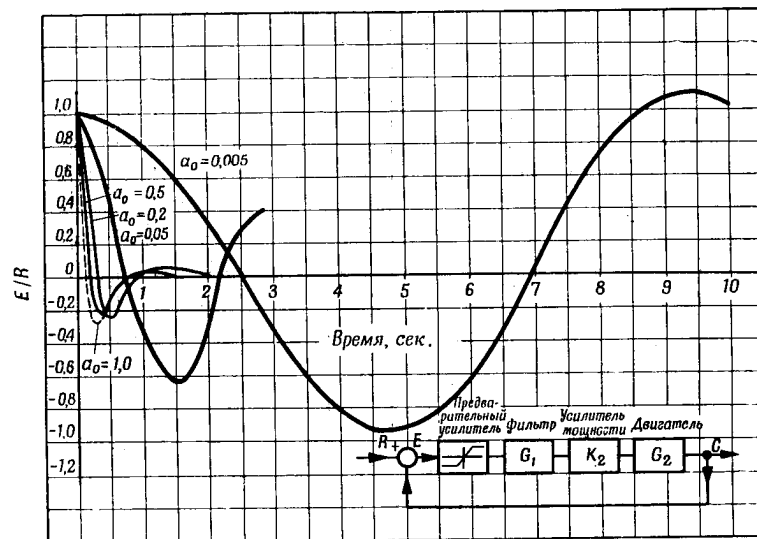
теля мощности. Горизонтальными пунктирными линиями показаны обратные значения коэффициента усиления усилителя, насыщающегося при начальных значениях отношения амплитуд $a_0 = 0,5; 0,2; 0,05$ и $0,005$. Эти линии по существу указывают величину, на которую уменьшается коэффициент усиления системы при отмеченном значении насыщения. Из



Фиг. 16. Амплитудная и фазовая частотные характеристики разомкнутой позиционной системы автоматического регулирования с последовательной стабилизацией (понижение усиления при различных начальных отношениях амплитуд a_0).

фиг. 16 видно, что для $a_0 = 0,005$ при $0 \text{ } \delta\text{б}$ имеет место отрицательный запас устойчивости по фазе, и следует ожидать, что работа системы будет неустойчивой в смысле появления больших незатухающих колебаний. Для $a_0 = 0,05$ система имеет небольшой отрицательный запас по фазе, так что в этом случае можно ожидать более устойчивой работы. Для значений $a_0 = 0,5$ и $0,2$ уменьшение коэффициента усиления недостаточно для получения неустойчивой системы. Однако в каждом случае быстродействие системы уменьшается из-за уменьшения частоты, при которой пунктирные линии пересекают амплитудные характеристики линейных элементов разомкнутой системы.

На фиг. 17 показаны переходные процессы в системе, представленной на фиг. 16, начинающиеся из одного и того же положения при различных значениях насыщения предварительного усилителя, вводимых различными значениями a_0 . Устойчивая работа имеет место для всех значений начальной относительной амплитуды a_0 , кроме $a_0 = 0,005$. Переходные



Фиг. 17. Переходные процессы при фиксированном входном сигнале и насыщении предварительного усилителя (последовательная стабилизация как на фиг. 16, при различных значениях a_0).

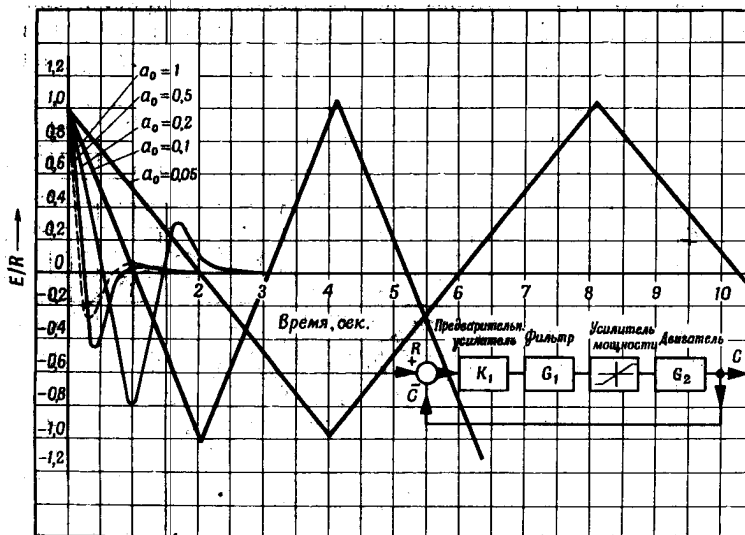
характеристики при больших значениях a_0 близки к характеристике в линейном случае, показанной пунктирной кривой ($a_0 = 1,0$); при $a_0 = 0,05$ характеристики сильно колебательны, а при $a_0 = 0,005$ система определенно неустойчива, так как возникают незатухающие колебания.

На фиг. 18 показаны переходные характеристики системы, представленной на фиг. 16, в случае насыщения усилителя мощности при тех же значениях a_0 . Заметим, что при начальных значениях относительной амплитуды a_0 , таких, как 0,05 и 0,1, система неустойчива. В случае насыщения пред-

7. Частотный метод представления типовых нелинейностей 265

варительного усилителя при этих значениях a_0 система устойчива (см. фиг. 17). Следовательно, даже при малых значениях насыщения усилителя мощности переходные характеристики ухудшаются и система делается вялой.

Для больших значений насыщения ($a_0 = 0,1$ или $0,05$) переходные характеристики имеют большие интервалы изменения с постоянной скоростью, соответствующей насыщению



Фиг. 18. Переходные процессы при фиксированном входном сигнале и насыщении усилителя мощности (последовательная стабилизация как на фиг. 16, при различных значениях a_0).

источника мощности. Этим данный случай отличается от случая насыщения предварительного усилителя, когда наличие стабилизирующего устройства между элементом с насыщением и регулируемой переменной предохраняет последнюю от заметного влияния источника насыщения.

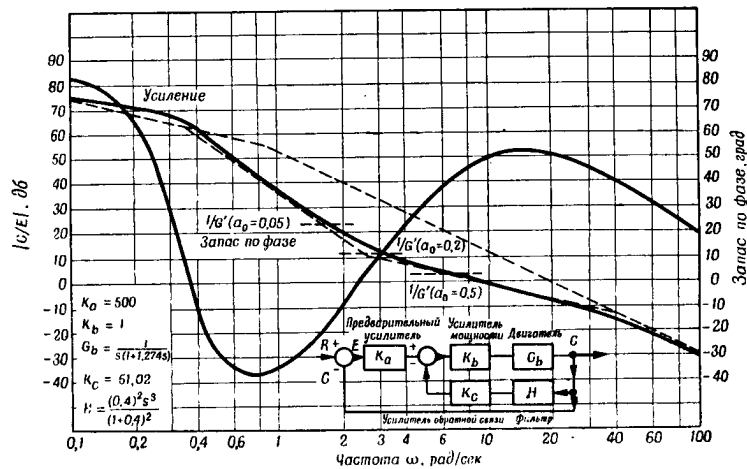
Переходные характеристики, показанные на фиг. 17 и 18, качественно совпадают с характеристиками, которые можно рассчитать по методу линейных приближений, используя

цему для элемента с насыщением эквивалентный линейный коэффициент усиления.

Насыщение предварительного усилителя и усилителя мощности дает фактически различные характеристики, хотя метод линейных приближений не указывает на это различие.

Системы с параллельной стабилизацией

На фиг. 19 показаны действительные и приближенные амплитудные и фазовые характеристики разомкнутой условно устойчивой позиционной системы регулирования, содержащей



Фиг. 19. Амплитудная и фазовая частотные характеристики разомкнутой системы автоматического регулирования со стабилизирующей обратной связью (понижение усиления предварительного усилителя при различных значениях a_0).

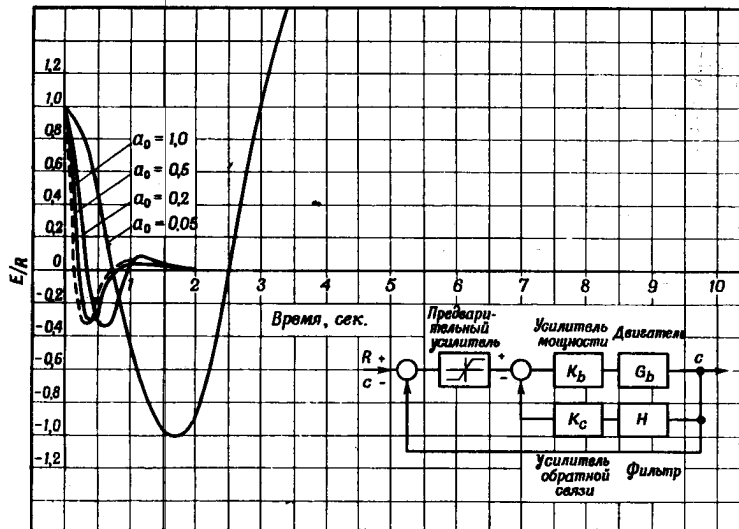
--- асимптотическая характеристика; ————— точная характеристика.

стабилизирующую обратную связь и имеющей амплитудную характеристику, подобную амплитудной характеристике разомкнутой системы с последовательной стабилизацией (см. фиг. 16).

Мы отдельно рассмотрим случаи насыщения предварительного усилителя, усилителя мощности и усилителя обратной связи.

7. Частотный метод представления типовых нелинейностей 267

Для случая насыщения предварительного усилителя горизонтальными пунктирными линиями отмечены начальные значения относительной амплитуды; значение $a_0 = 0,05$ должно давать неустойчивую работу, тогда как при $a_0 = 0,5$ и $0,2$ система будет устойчивой. Это подтверждается графиком фиг. 20, на котором показаны переходные процессы в рассматриваемой системе при насыщении предварительного усилителя для указанных значений a_0 .

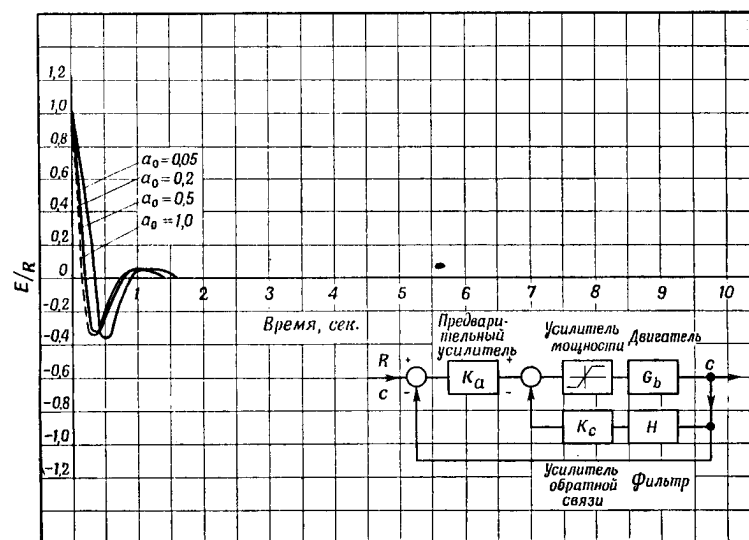


Ф и г. 20. Переходные процессы при фиксированном входном сигнале и насыщении предварительного усилителя (стабилизация обратной связи как на фиг. 19, при различных значениях a_0).

Обращаясь снова к фиг. 19 и рассматривая явление понижения коэффициента усиления источника мощности, следует заметить, что основным эффектом будет понижение усиления при низких частотах ($\omega < 1,0 \text{ рад/сек}$) и при высоких частотах ($\omega > 20 \text{ рад/сек}$). Хотя при этом низкочастотная часть характеристики системы понижается, избыток фазы при усилении 0 дб уменьшается незначительно и не изменяется частота, при которой это имеет место. Из этого, по-видимому, вытекает, что переходные характеристики

системы даже при достаточно большом значении насыщения усилителя мощности не должны заметно изменяться.

Из переходных характеристик системы (см. фиг. 19) с насыщением усилителя мощности, приведенных на фиг. 21, видно, что при начальном значении отношения амплитуд $a_0 = 0,05$ и более получается достаточно устойчивая работа системы.



Фиг. 21. Переходные процессы при фиксированном входном сигнале и насыщении усилителя мощности (стабилизация обратной связью как на фиг. 19, при различных значениях a_0).

Как уже отмечалось выше, при рассмотрении амплитудных частотных характеристик изменение усиления в усилителе мощности не влияет заметно на переходные характеристики системы.

Обращаясь опять к фиг. 19, рассмотрим теперь явление уменьшения усиления обратной связи при насыщении усилителя в цепи обратной связи. При уменьшении усиления обратной связи амплитудная характеристика разомкнутой системы приближается к характеристике разомкнутой системы без всякой стабилизирующей обратной связи. Так как ра-

Выводы относительно влияния насыщения на переходной процесс

Из рассмотренных данных, так же как и из изучения других систем с насыщением, можно сделать следующие выводы относительно работы систем в области насыщения.

Влияние насыщения можно свести к уменьшению коэффициента усиления элемента с насыщением. Коэффициент изменения усиления при насыщении, рассмотренный в предыдущем разделе, дает вполне удовлетворительную меру величины изменения коэффициента усиления, вызванного насыщением.

Для того чтобы явление насыщения могло вызвать незадаваемые колебания, необходимо, чтобы система была такой, в которой при линейном рассмотрении уменьшение коэффициента усиления насыщаемого элемента может привести к неустойчивой работе.

Поскольку для некоторых систем автоматического регулирования результат уменьшения усиления одного или нескольких элементов может оказывать небольшое вредное влияние или даже отчасти иметь благоприятное действие, то очевидно, что насыщение не всегда может вызывать неудовлетворительную работу системы.

Эти выводы хорошо согласуются с выводами, полученными Кохенбургером [3] и другими исследователями [4].

Учет влияния люфта при расчете механических систем автоматического регулирования

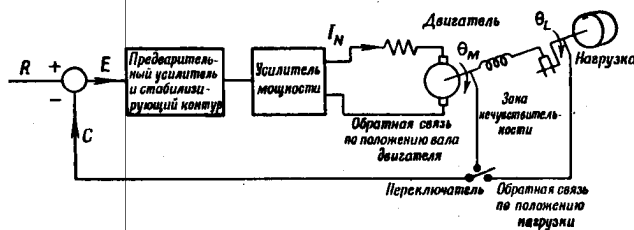
При изготовлении позиционных следящих систем трудно бывает получить совершенное соединение двигателя с нагрузкой. Обычно двигатель соединяется с нагрузкой через редуктор, в зубчатых зацеплениях которого существуют некоторые зазоры.

Было проведено большое число исследований [5, 6], в которых описывались характеристики двигателя и нагрузки при наличии люфта и при различных предположениях о типах характеристик нагрузки. В статье Томаса [7] содержится подробное исследование зоны нечувствительности, в котором рассматривается большое число различных примеров с нелинейностью типа люфта.

7. Частотный метод представления типовых нелинейностей 271

Представляет интерес рассмотреть эту задачу с учетом современных методов анализа систем автоматического регулирования. В этой работе основной упор будет сделан на использовании метода описывающей функции с получением эквивалентного коэффициента усиления для комбинации люфта с упругим элементом.

С помощью частотных методов будет исследовано влияние отношения момента инерции двигателя к моменту инерции нагрузки J_M/J_L , влияние механического резонанса (ζ),



Фиг. 23. Позиционная следящая система, имеющая инерционную нагрузку с вязким демпфированием, соединенную с двигателем через упругий вал с зазором.

коэффициента демпфирования самой нагрузки ζ и влияние места подключения обратной связи (к двигателю или нагрузке). Полученные таким образом качественные результаты будут сравнены с количественными результатами, найденными с помощью электронной модели.

Полезные сведения для расчета механических систем автоматического регулирования дает рассмотрение влияния люфта в типовой позиционной следящей системе (фиг. 23) при небольшом возмущающем воздействии в виде ступенчатой функции. Линейные элементы системы, предварительный усилитель, стабилизирующий контур и усилитель мощности показаны на фиг. 23 в виде структурной схемы. Двигатель и нагрузка показаны схематично и соответствуют таким же элементам, рассмотренным выше на фиг. 8. Структурная схема для этих элементов была получена на основании уравнений (13) и (16) и приведена на фиг. 11. Используя эквивалентный линейный коэффициент жесткости K'_S для представления комбинации упругого элемента и люфта и блок-схему на фиг. 11, можно выразить соотношение между

положением вала двигателя θ_M и входным напряжением двигателя I_N следующим образом:

$$\frac{s\theta_M}{I_N} = \frac{K_T/R (J_L s^2 + D_L s + K'_S)}{P_3 s^3 + P_2 s^2 + P_1 s + P_0}, \quad (30)$$

где

$$P_3 = J_M J_L; \quad P_2 = J_M D_L + J_L \frac{K_e K_T}{R};$$

$$P_1 = (J_M + J_L) K'_S + D_L \frac{K_e K_T}{R};$$

$$P_0 = K'_S \left(D_L + \frac{K_e K_T}{R} \right).$$

Таким же образом получим отношение положения вала нагрузки θ_L к входному напряжению двигателя I_N

$$\frac{s\theta_L}{I_N} = \frac{K'_S K_T / R}{P_3 s^3 + P_2 s^2 + P_1 s + P_0}. \quad (31)$$

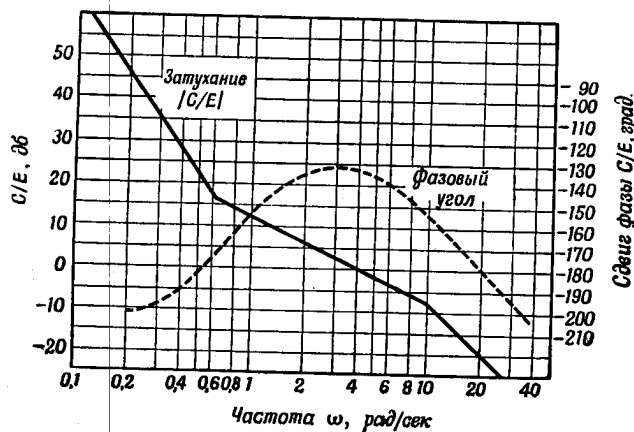
Эти уравнения сходны с уравнениями для тех же элементов без люфта с той лишь разницей, что коэффициент жесткости K'_S является функцией величины $\theta_B/2$ и максимальной разности амплитуд $(\theta_M - \theta_L)_{\max}$; K'_S может изменяться от нуля при $-\theta_B/2 < \theta_M - \theta_L < +\theta_B/2$ до K_S — действительного коэффициента жесткости вала при $|\theta_M - \theta_L| \gg \gg |\theta_B/2|$. Поскольку уравнения для положения оси двигателя и нагрузки записаны отдельно, мы имеем возможность путем сравнения установить, какую величину лучше принять в качестве регулируемой переменной для формирования сигнала обратной связи — положение вала двигателя или положение вала нагрузки (см. переключатель на фиг. 23).

Расчетные параметры системы

Сначала будем рассматривать систему как простую линейную систему и выразим расчетные конструктивные параметры через характеристики такой линейной системы. Предположим, что моменты инерции двигателя J_M и нагрузки J_L можно объединить в один эквивалентный момент инерции,

7. Частотный метод представления типовых нелинейностей 273

приведенный к валу двигателя. Коэффициент жесткости нагрузки предполагается равным бесконечности, а вязкое и сухое трение не принимается в расчет. Для этого случая



Фиг. 24. Приближенные амплитудная и фазовая частотные характеристики разомкнутой линеаризованной системы, соответствующей фиг. 23.

уравнения (30) и (31) становятся тождественными и содержат только одно простое звено с постоянной времени двигателя, т. е.

$$\frac{s\theta_M}{I_N} = \frac{s\theta_L}{I_N} = \frac{1}{K_e \left[\frac{(J_M + J_L)R}{K_e K_T} s + 1 \right]}, \quad (32)$$

где $K_e = 2,0$ в/(рад/сек),

$$\frac{K_T}{R} = 0,17 \text{ кгм/в},$$

$$J_M + J_L = 0,03 \text{ кгм} \cdot \text{сек}^2,$$

$T_T = [(J_M + J_L)R]/K_e K_T = 0,1$ сек. — полная эквивалентная постоянная времени двигателя.

Передаточная функция усилителя мощности равна

$$G_2 = \frac{20}{\left(\frac{s}{50} + 1\right)}. \quad (33)$$

Предварительный усилитель и стабилизирующие контуры были выбраны так, чтобы при принятых предположениях система была бы условно устойчивой. Их передаточная функция равна

$$G_1 = \frac{40 \left(\frac{s}{0,64} + 1 \right)^2}{\left(\frac{s}{0,064} + 1 \right)^2}. \quad (34)$$

На фиг. 24 показаны приближенные амплитудная и фазовая частотные характеристики разомкнутой линейной системы в окрестности 0 дб.

Введем следующие обозначения, связывающие параметры механической системы с параметрами характеристик:

$\omega_0 = \sqrt{\frac{K_S}{J_L}}$ — собственная частота нагрузки при фиксированном моменте инерции двигателя и без демпфирования (механическая резонансная частота);

$\omega_T = \frac{1}{T_T}$ — частота, соответствующая обратному значению эквивалентной постоянной времени двигателя;

$c = \frac{\omega_0}{\omega_T}$ — отношение частоты механического резонанса к частоте, соответствующей эквивалентной постоянной времени двигателя;

$\zeta = \frac{D_L}{2J_L\omega_0}$ — коэффициент демпфирования механической нагрузки при фиксированном моменте инерции двигателя и при отсутствии люфта;

$A_D = \frac{K'_S}{K_S}$ — коэффициент уменьшения постоянной жесткости, зависящий от величины $\frac{\theta_B}{2(\theta_M - \theta_L)_{\max}}$, как указано на фиг. 15.

В дальнейшем мы будем предполагать, что приведенный момент инерции двигателя и нагрузки $(J_M + J_L)$ постоянен и равен $0,03 \text{ кгм} \cdot \text{сек}^2$, хотя отношение J_M/J_L есть параметр,

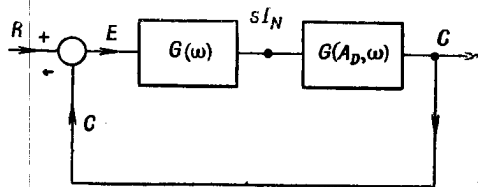
7. Частотный метод представления типовых нелинейностей 275

который при расчете систем автоматического регулирования можно изменять.

С учетом этих определений и предположений можно выразить механические расчетные параметры в безразмерной форме и аналитически исследовать влияние изменения этих отношений.

Основные соображения, касающиеся анализа устойчивости

Кохенбургер указал способ разделения нелинейной системы на линейную часть, характеризующуюся передаточной функцией, и нелинейную часть, характеризующуюся описывающей



Фиг. 25. Система фиг. 23, представленная в виде последовательного соединения передаточной и описывающей функций.

функцией (1). Рассматриваемая таким образом система представлена на фиг. 25. Для удобства двигатель включен в линейную часть. Передаточная функция замкнутой системы дается выражением

$$\frac{C}{R} = \frac{G(\omega) G'(A_D, \omega)}{1 + G(\omega) G'(A_D, \omega)}. \quad (35)$$

Условие потери устойчивости системой можно записать следующим образом:

$$G(\omega) = -\frac{1}{G'(A_D, \omega)}, \quad (36)$$

где

$$G(\omega) = \frac{800 (s/0,64 + 1)^2}{s (s/0,064 + 1)^2 (s/50 + 1)},$$

а $G'(A_D, \omega)$ — описывающая функция, определяемая уравнением (30) или (31), смотря по тому, используется ли

обратная связь по положению двигателя или по положению нагрузки; $G'(A_D, \omega)$ является функцией отношения амплитуд

$$\alpha = \frac{\theta_B}{2(\theta_M - \theta_L)_{\text{макс.}}} \quad (37)$$

и частоты ω .

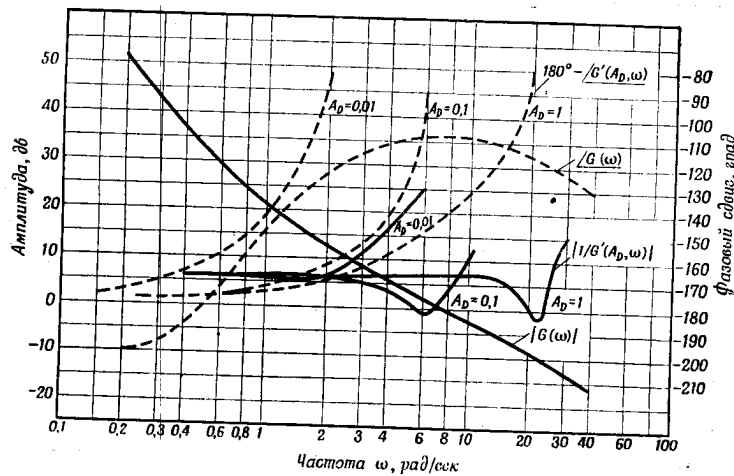
Чтобы облегчить применение частотных методов, строятся графики правой и левой частей уравнения (36) при различных задаваемых значениях A_D для всех значений параметров системы, которые представляют интерес. При тех частотах, при которых имеет место равенство амплитуд в уравнении (36), фиксируется имеющееся соотношение фаз. Если это фазовое соотношение таково, что запас устойчивости по фазе для замкнутой системы положителен, то система при этом значении A_D будет устойчивой и амплитуда колебаний будет стремиться к уменьшению, так что появляется меньшее значение A_D . Если же запас по фазе отрицателен, то система для этого значения A_D будет неустойчивой и амплитуда колебаний будет стремиться к увеличению, так что достигаются большие значения A_D .

В случае когда одновременно выполняются амплитудное и фазовое соотношения уравнения (36), колебания в системе являются незатухающими при соответствующих значениях коэффициента изменения усиления A_D и частоты незатухающих колебаний ω_S . По коэффициенту изменения усиления и кривой A_D как функции $I_B/2I_M$, приведенной на фиг. 15, можно определить значение отношения величины люфта к максимуму входа нелинейного элемента. Зная величину углового люфта для данной системы и определяя по кривой отношение $\theta_B/2(\theta_M - \theta_L)_{\text{макс.}}$, можно вычислить значение $(\theta_M - \theta_L)_{\text{макс.}}$. С помощью этих значений можно, используя обычные линейные методы, определить другие важные величины, например ошибку или положение нагрузки при частоте колебания ω_S .

Чтобы показать характер устойчивости этого процесса, обратимся к фиг. 26, на которой сплошными и пунктирными кривыми показаны соответственно амплитудные и фазовые характеристики для $G(\omega)$ и для $-1/G'(A_D, \omega)$ при трех различных значениях A_D , равных 1,0; 0,1 и 0,01. Используется обратная связь по положению нагрузки, так как при

7. Частотный метод представления типовых нелинейностей 277

этом наблюдается большая тенденция к возникновению незатухающих колебаний, чем при использовании обратной связи по положению вала двигателя. Для значения $A_D = 1$, соответствующего линейному случаю при отсутствии люфта, равенство амплитуд для уравнения (36) имеет место при $\omega = 3,7 \text{ рад/сек}$, хотя положительный запас по фазе существует до



Фиг. 26. Амплитудные и фазовые частотные характеристики и обратная отрицательная описывающая функция системы фиг. 25, показывающие влияние коэффициента изменения усиления элемента с зоной нечувствительности A_D при обратной связи по положению нагрузки.

$$I_M/I_L = 4; \zeta = 0,05; c = 2.$$

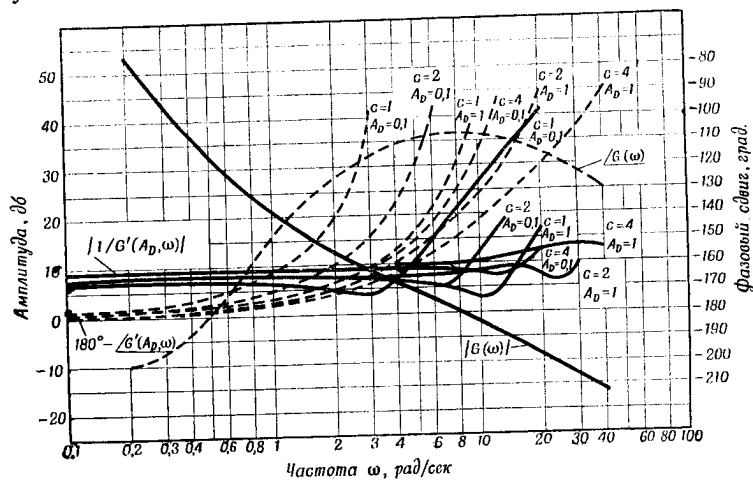
— амплитуда; - - - фаза.

$\omega = 15 \text{ рад/сек}$. Следовательно, указанные значения параметров системы и нагрузки обеспечивают устойчивую работу.

При $A_D = 0,1$ равенство амплитуд для уравнения (36) имеет место при $\omega = 7,0 \text{ рад/сек}$, хотя отрицательный запас по фазе существует для частот, превышающих $\omega = 5,5 \text{ рад/сек}$. Это значит, что для этого значения A_D , соответствующего $\theta_B/2(\theta_M - \theta_L)_{\text{макс.}} = 1,24$, система неустойчива и амплитуда колебаний увеличивается до тех пор, пока не будут достигнуты устойчивые незатухающие колебания. Амплитудные и фазовые характеристики при $A_D = 0,01$

также указывают на неустойчивую работу, так что существует определенная возможность для возникновения колебательного процесса, который будет усиливаться до получения значений $A_D > 0,1$.

Метод для определения возможности существования незатухающих колебаний в области $0,1 < A_D < 1,0$ используется в последующих разделах статьи. Этот диапазон



Фиг. 27. Амплитудные и фазовые частотные характеристики и обратная отрицательная описывающая функция системы фиг. 25, показывающие влияние относительной механической резонансной частоты c при обратной связи по положению нагрузки для $A_D = 0,1$ и $1,0$.

$$I_M/I_L = 1; \zeta = 0,1.$$

— амплитуда; - - - фаза.

значений A_D был выбран потому, что, по-видимому, при этом можно ограничить амплитуду незатухающих колебаний величиной, меньшей чем 1,25 величины люфта. Применение этого метода помогает получить качественное указание на относительное влияние различных конструктивных параметров, которые имеют принципиальный интерес, что и рассматривается ниже.

Влияние механической резонансной частоты нагрузки

На фиг. 27 показано влияние увеличения относительной резонансной частоты нагрузки c от 1 до 4 для значений

7. Частотный метод представления типовых нелинейностей 279

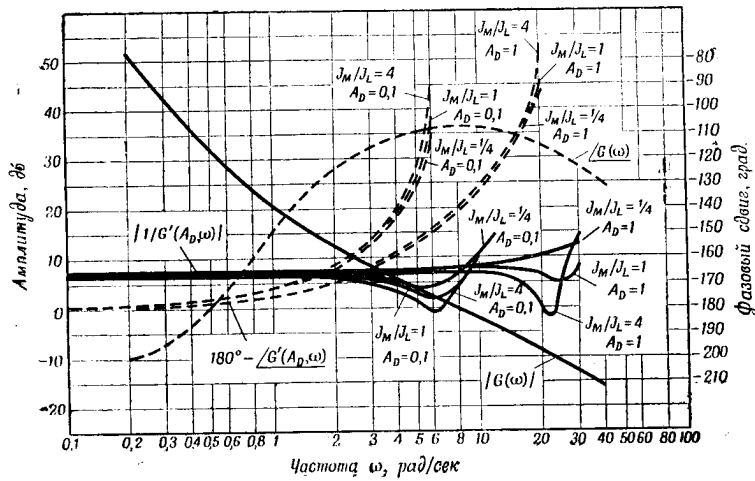
расчетных параметров $J_M/J_L = 1,0$ и $\zeta = 0,1$ в случае обратной связи по положению нагрузки. Изучение этих кривых показывает, что для $A_D = 0,1$ и $c = 1,0$ система неустойчива, тогда как для более высоких значений c при том же значении $A_D = 0,1$ и для всех значений c при $A_D = 1,0$ система устойчива. Кроме того, относительная степень устойчивости повышается с увеличением значения c , как на это указывает наибольшее значение запаса по фазе при равенстве амплитуд и более высокое отношение частот, соответствующих равенству амплитуд и нулевому запасу по фазе. Эти кривые указывают, что увеличение относительной резонансной частоты нагрузки уменьшает амплитуду незатухающих колебаний, вызванных наличием люфта. Очевидно, что значения $c \geq 2$ удовлетворяют рассматриваемой здесь системе.

Влияние отношения моментов инерции двигателя и нагрузки J_M/J_L

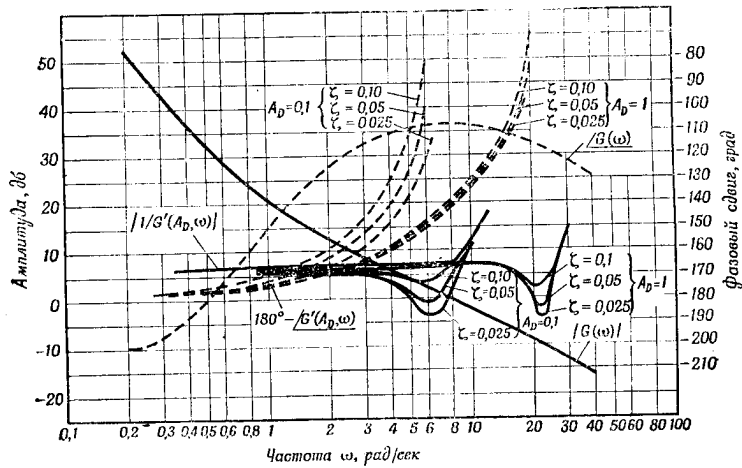
На фиг. 28 показано влияние изменения отношения моментов инерции J_M/J_L от $1/4$ до 4 на устойчивость системы с обратной связью по положению нагрузки. Видно, что при $A_D = 0,1$ и $J_M/J_L = 1/4$ система будет устойчивой. Однако для значений $J_M/J_L = 1$ и $J_M/J_L = 4$ и $A_D = 0,1$ система будет неустойчивой, причем неустойчивость возрастает при более высоких значениях J_M/J_L . При $A_D = 1,0$ все три значения J_M/J_L дают устойчивую систему, причем лучшие характеристики будут при более низком значении J_M/J_L . Следовательно, при $J_M/J_L = 1$ и $J_M/J_L = 4$ будут устанавливаться незатухающие колебания, амплитуда которых в 1,25 раза больше величины люфта. При использовании обратной связи по положению нагрузки и постоянном полном моменте инерции желательно, чтобы возможно большая часть полного момента инерции была сосредоточена в нагрузке, благодаря чему уменьшаются колебания, вызванные люфтом.

Влияние коэффициента демпфирования нагрузки ζ

На фиг. 29 показано влияние изменения коэффициента демпфирования нагрузки ζ в пределах от 0,025 до 0,10 для значений $A_D = 0,1$ и $A_D = 1,0$ на устойчивость системы с обратной связью по положению нагрузки. Очевидно, что при $\zeta = 0,1$ и $A_D = 0,1$ система будет находиться близко



Фиг. 28. Амплитудные и фазовые частотные характеристики и обратная отрицательная описывающая функция системы фиг. 25, показывающие влияние отношения моментов инерции I_M/I_L при обратной связи по положению нагрузки для $A_D = 0,1$ и $1,0$. $\zeta = 0,05$; $c = 2$. — амплитуда; - - - фаза.



Фиг. 29. Амплитудные и фазовые частотные характеристики и обратная отрицательная описывающая функция системы фиг. 25, показывающие влияние коэффициента демпфирования ζ при обратной связи по положению нагрузки для $A_D = 0,1$ и $1,0$. $I_M/I_L = 4$; $c = 2$. — амплитуда; - - - фаза.

7. Частотный метод представления типовых нелинейностей 281

к границе устойчивости; можно также видеть, что для значений $\zeta = 0,05$ и $0,025$ система будет определенно неустойчивой. При $A_D = 1,0$ система будет устойчивой для всех трех значений коэффициента демпфирования ζ , хотя в диапазоне частот, близких к механической резонансной частоте нагрузки, уменьшение коэффициента демпфирования вызывает уменьшение устойчивости системы. Таким образом, с точки зрения уменьшения амплитуды незатухающих колебаний, возникающих вследствие наличия люфта, желательнее иметь механический коэффициент демпфирования около $0,1$.

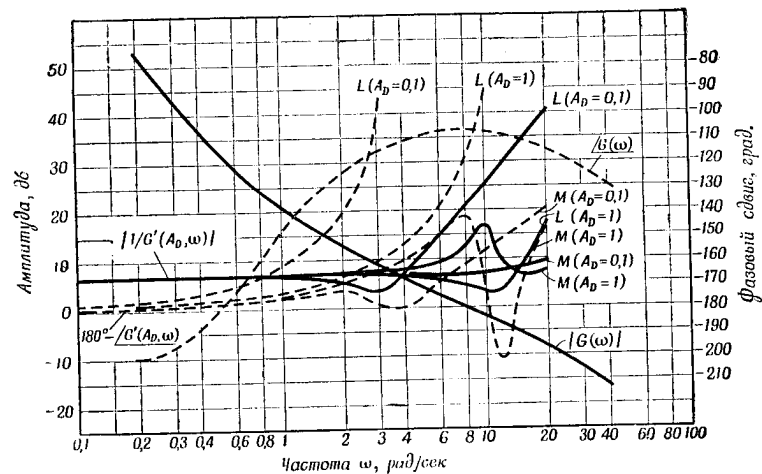
Так как вязкое трение характеризует потерю энергии при движении нагрузки с постоянной скоростью, то очевидно, что необходимо использовать ограниченное значение вязкого трения, и для подавления незатухающих колебаний, вызванных наличием люфта, следует искать другие средства.

Сравнение обратных связей по положению двигателя и нагрузки

На фиг. 30 показано сравнение устойчивости некоторой системы при использовании обратной связи либо по положению вала двигателя, либо по положению нагрузки (см. фиг. 23) при значениях $A_D = 0,10$ и $A_D = 1,0$. Для случая обратной связи по положению нагрузки система определенно неустойчива при $A_D = 0,1$ и устойчива при $A_D = 1,0$. Таким образом, при обратной связи по положению нагрузки следует ожидать незатухающих колебаний с амплитудой, превышающей в $1,25$ раза величину углового люфта. Для той же системы, но с обратной связью по положению двигателя для обоих значений A_D получается относительно высокоустойчивая работа. Этот результат подтверждается экспериментальными данными, полученными при исследовании систем автоматического регулирования в течение последних нескольких лет. Кроме того, этот результат показывает, что дополнительный фазовый сдвиг, возникающий при использовании обратной связи по положению нагрузки, а не по положению двигателя (уравнения (30) и (31)), может оказаться вредным в рассматриваемом диапазоне частот.

Различие между характеристиками системы при рассмотренных случаях обратной связи выражено слабо при низких частотах, т. е. при $\omega < 1,0$ рад/сек, тогда как при более

высоких частотах $\omega > 1$ результаты могут значительно отличаться. Исходя из этого, представляют особый интерес рекомендуемые и применяемые в Англии и других странах методы „распределенной интегральной обратной связи“. Эти результаты подтверждают давно установленный расчетный



Ф и г. 30. Амплитудные и фазовые частотные характеристики и обратная отрицательная описывающая функция системы фиг. 25, дающие сравнение обратных связей по положению вала двигателя и по положению нагрузки при $A_D = 0,1$ и $1,0$.

$$I_M/I_L = 1; \zeta = 0,1; c = 1,3.$$

M — Обратная связь по положению вала двигателя; L — обратная связь по положению нагрузки.
— амплитуда; - - - фаза.

критерий, заключающийся в том, что для получения наиболее устойчивой работы, когда это возможно, следует снимать позиционный сигнал, предназначенный для обратной связи, с двигателя, а не с нагрузки.

Результаты, полученные на модели

Схема системы регулирования по положению, представленная на фиг. 23, была исследована Р. Мэгьюзом на электронной модели. Ее характеристики были определены для

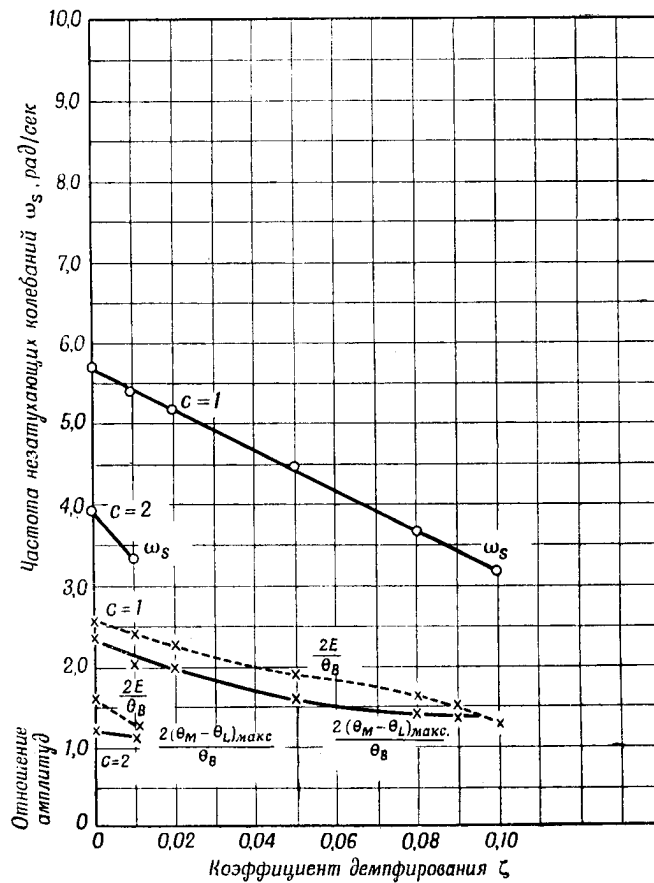
7. Частотный метод представления типовых нелинейностей 283

большого количества значений параметров с обратной связью как по положению двигателя, так и по положению нагрузки. На фиг. 31, 32 и 33 показано несколько важнейших характеристик системы с обратной связью по положению нагрузки в зависимости от механического коэффициента демпфирования для трех различных значений отношения моментов инерции двигателя и нагрузки J_M/J_L . Представленные характеристики включают частоту незатухающих колебаний ω_s , амплитуду отношения ошибки к величине люфта $2E/\theta_B$ и амплитуду отношения максимального угла поворота к величине углового люфта $2(\theta_M - \theta_L)_{\max.}/\theta_B$. Последние две характеристики являются просто отношением амплитуд. Полученные величины ошибки или угла скручивания прямо пропорциональны величине углового люфта θ_B . Из этого ясно, что влияние люфта может быть уменьшено путем его ограничения.

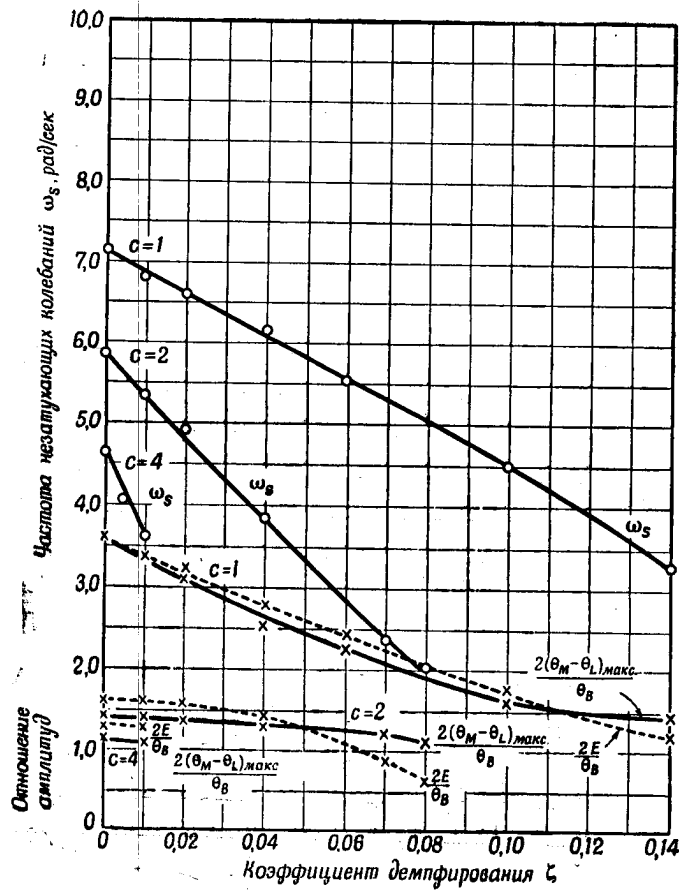
Из рассмотрения фиг. 31, для которой $J_M/J_L = 1/4$, т. е. основная часть момента инерции сосредоточена в нагрузке, можно видеть, что при значениях $c = 1$ и $c = 2$ существуют незатухающие колебания, тогда как при значении $c = 4$ незатухающие колебания отсутствуют. Частоты этих колебаний заключены в довольно узких пределах $\omega = 3-6$ рад/сек, т. е. в окрестности частоты среза амплитудной характеристики разомкнутой системы (см. фиг. 24).

Отношения „ошибка/величина люфта“ и „угол скручивания/величина люфта“ примерно равны и изменяются от 2,5 до $\sim 1,0$, постепенно уменьшаясь с увеличением значений коэффициента демпфирования ζ . При увеличении значений c амплитуды незатухающих колебаний становятся меньше и при $c \geq 4$ колебания совсем затухают.

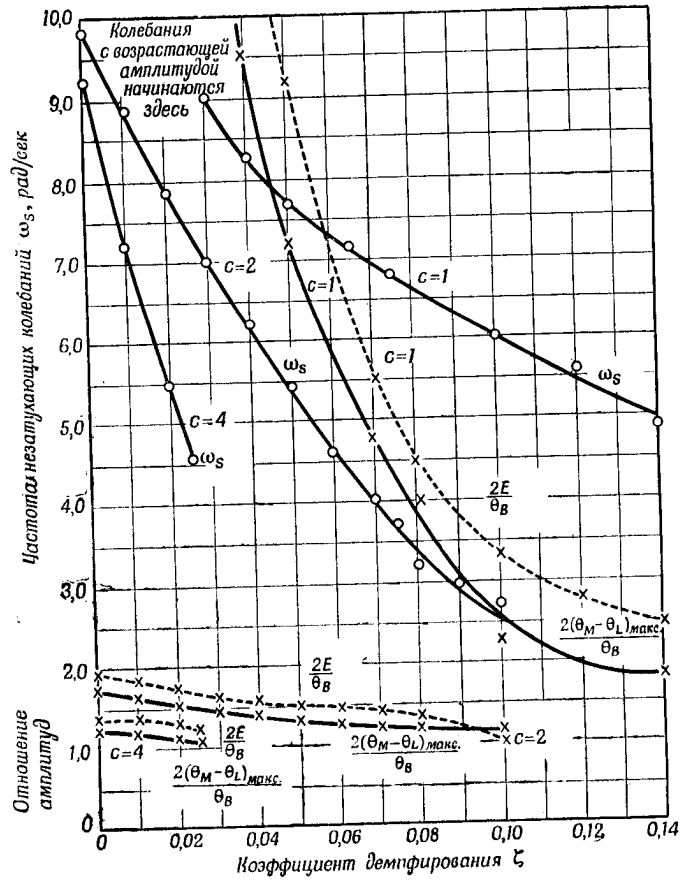
На фиг. 32 приведены данные, подобные тем, которые представлены на фиг. 31, но для случая $J_M/J_L = 1,0$. В этом случае, когда двигатель и нагрузка имеют одинаковые моменты инерции, незатухающие колебания существуют при $c = 1, 2$ и 4. Диапазон частот этих колебаний больше ($\omega = 2-7$), причем частота колебаний при низких коэффициентах демпфирования больше, чем частота среза амплитудной характеристики разомкнутой системы. При тех же самых значениях коэффициента демпфирования и c отношение ошибки и угла скручивания к величине люфта теперь больше при $J_M/J_L = 1$, чем при $J_M/J_L = 1/4$. Однако при



Ф и г. 31. Частота незатухающих колебаний ω_s , отношение разности углов между валами двигателя и нагрузки к величине углового люфта и отношение угла ошибки к величине углового люфта в зависимости от коэффициента демпфирования ζ при различных значениях относительной механической резонансной частоты c для $I_M/I_L = 1/4$.



Фиг. 32. Частота незатухающих колебаний ω_s , отношение разности углов между валами двигателя и нагрузки к величине углового люфта и отношение угла ошибки к величине углового люфта в зависимости от коэффициента демпфирования ζ при различных значениях относительной механической резонансной частоты нагрузки c для $I_M/I_L = 1$.



Фиг. 33. Частота незатухающих колебаний ω_s , отношение разности углов между валами двигателя и нагрузки к величине углового люфта и отношение угла ошибки к величине углового люфта в зависимости от коэффициента демпфирования ζ при различных значениях относительной механической резонансной частоты нагрузки c для $I_M/I_L = 4$.

7. Частотный метод представления типовых нелинейностей 287

$c \geq 2$ эти отношения меньше, чем 2,0, и представляют почти удвоенное значение самого люфта. Как и в предыдущем случае, увеличение коэффициента демпфирования нагрузки вызывает уменьшение амплитуды и частоты колебаний, а также ошибки и угла скручивания.

На фиг. 33, которой соответствует $J_M/J_L = 4$, приведены сильно улучшенные характеристики для тех же самых механических резонансных частот нагрузки и того же самого коэффициента демпфирования, как и в случае низких значений J_M/J_L . При $c = 1,0$ невозможно понизить коэффициент демпфирования ниже 0,03 без появления колебаний с возрастающей амплитудой. Даже при более высоких значениях ζ (до 0,1) получены гораздо более высокие значения отношений ошибки и угла скручивания к величине люфта. При $c \geq 2$, несмотря на то, что получаются незатухающие колебания довольно высокой частоты, отношение ошибки и угла скручивания к угловой величине люфта при $J_M/J_L = 4,0$ снова меньше, чем 2,0.

Сравнение результатов, полученных с помощью частотного анализа и с помощью модели

Сравнение данных, полученных на модели, с результатами частотного анализа с применением эквивалентного линейного коэффициента усиления позволяет сделать одинаковые качественные выводы относительно влияния параметров c , J_M/J_L и ζ на незатухающие колебания, вызванные наличием люфта. Однако обычно частота незатухающих колебаний ω_s , определяемая на основе метода линеаризации, на 10—40% выше частоты, определяемой с помощью модели. Между амплитудами колебаний, полученными этими двумя методами, имеется несколько лучшее совпадение.

Кривые, получаемые с помощью модели, показывают, что в большинстве случаев при появлении незатухающих колебаний разностный угол $\theta_M - \theta_L$ является основной гармоникой. Поэтому эти кривые дают независимую проверку предположения о том, что сигнал на входе элемента с люфтом имеет синусоидальную форму. Следует напомнить, что это допущение положено в основу определения понятия эквивалентного линейного коэффициента усиления при нашем методе анализа систем с зоной нечувствительности и люфтом

Исследования, проведенные как на модели, так и с помощью частотного метода анализа, показывают, что система с использованием обратной связи по положению вала двигателя обеспечивает устойчивую работу без появления незатухающих колебаний для рассмотренных случаев.

Краткое изложение результатов расчета, полученных при исследовании люфта

Результаты, полученные с помощью модели и частотным методом, указывают на следующие способы уменьшения незатухающих колебаний, вызванных наличием люфта: а) увеличение механической резонансной частоты нагрузки ω ; б) уменьшение отношения моментов инерции двигателя и нагрузки J_M/J_L ; в) увеличение коэффициента вязкого трения нагрузки ζ .

Применение обратной связи по положению двигателя, а не по положению нагрузки, также способствует ослаблению незатухающих колебаний, вызванных наличием люфта.

Несмотря на то, что результаты, приведенные в этой статье, относятся к случаю условно устойчивой позиционной системы регулирования, подобные же результаты были получены и являются справедливыми и для безусловно устойчивых систем. Из фиг. 26—29 видно, что появление незатухающих колебаний наблюдается в диапазоне частот вблизи частоты среза характеристики разомкнутой системы и не существуют при более низких частотах, соответствующих условно устойчивой работе системы.

Заключение

В статье показано, что имеется большое число нелинейных элементов с характеристиками типа насыщения или зоны нечувствительности и что для них в качестве описывающей функции можно использовать „эквивалентный линейный коэффициент усиления“. Показано, что эквивалентные коэффициенты усиления для элементов с насыщением и зоной нечувствительности имеют отчасти сходные свойства и относительно простой вид.

Хотя результаты, полученные с помощью этого приближенного метода анализа зоны нечувствительности и насыще-

7. Частотный метод представления типовых нелинейностей 289

ния, не дают точного решения, они все же позволяют получить при расчете систем автоматического регулирования качественные результаты. Благодаря этому метод особенно полезен при расчете более сложных систем, какие обычно и встречаются на практике. Вместе с тем он не снимает необходимости более точного аналитического исследования.

ЛИТЕРАТУРА

1. Kochenburger R. J., *Trans. AIEE*, 69, 270 (1950).
2. Johnson E. C., *Trans. AIEE*, 71 (1952).
3. Kochenburger R. J., *Trans. AIEE*, 72, 180 (1953).
4. Levinson E., *Trans. AIEE*, 72, 1—9 (1953), см. статью 6, настоящего сборника.
5. Tustin A., *Journ. IEE*, 94, Part II-A, 143 (1947).
6. Ливерсидж Дж., Влияние зазоров и упругих деформаций в замкнутых цепях систем автоматического регулирования на поведение этих систем, см. сборник „Автоматическое регулирование“, Издательство иностранной литературы, 1954, стр. 320.
7. Thomas C. H., *Trans. ASME*, 76, № 8, 1365 (1954).

**III. ПРИМЕНЕНИЕ
ЧАСТОТНЫХ МЕТОДОВ
К РАСЧЕТУ СПЕЦИАЛЬНЫХ
СИСТЕМ**

8. ПРИМЕНЕНИЕ ЧАСТОТНЫХ МЕТОДОВ К РАСЧЕТУ ИМПУЛЬСНЫХ СИСТЕМ¹⁾

В. К. Линвилл, Р. В. Ситтлер

Введение

Проблема анализа и синтеза импульсных систем на первый взгляд представляется весьма сложной. При рассмотрении импульсных систем инженеру приходится наряду с обычными элементами принимать во внимание действие элементов нового типа. Наличие в системе импульсных процессов и присутствие линейных фильтров, сигналы на входе и выходе которых имеют импульсный характер, заставляют его предположить, что перед ним—проблемы совершенно иной природы, чем те, с которыми ему до сих пор приходилось встречаться.

Первое впечатление о трудности решения проблемы еще более усиливается при обращении к литературе, посвященной импульсным системам. Чаще всего в ней усиленно подчеркивается различие между импульсными и обычными системами. Инженера смущает также использование новых необычных математических методов. Не удивительно, что при этом у него появляется убеждение об особой природе этих систем и о том, что они не могут быть исследованы применением обычных методов анализа.

Но это первое впечатление совершенно ошибочно. Установлено, что для анализа и синтеза подобных систем не требуется никаких новых методов. В действительности необходимо только расширить и приспособить хорошо изученные, привычные методы к этим системам. Сходство между импульсными и обычными системами очень велико, и оно более существенно, чем различия между ними.

В данной статье мы не ставим себе задачей подробно обосновать утверждение о сходстве обоих типов систем

¹⁾ Linvill W. K., Sittler R. W., *Convent. Record IRE*, Part I, 99—104 (1953).

в подавляющем большинстве признаков. В рамках настоящей статьи можно лишь высказать общие соображения, касающиеся исследования импульсных систем обычными методами, и проиллюстрировать их на ряде логически связанных примеров. Более полное и систематическое освещение этих проблем читатель найдет в работах [1] и [2].

Главные преимущества метода, основанного на использовании сходства между импульсными и непрерывными системами, обнаруживаются при синтезе систем. Синтез систем есть искусство достижения компромисса между желаемыми характеристиками некоторой системы и возможностью ее физической реализации. В то время как *задача анализа* всегда имеет единственное решение, *задача синтеза* часто может вовсе не иметь решения при заданных условиях или иметь несколько решений. Поэтому при синтезе систем обычно применяют метод эмпирического подбора, при котором последовательно рассматривают, анализируют и видоизменяют различные конкретные системы. Для успешного применения этого метода такой процесс следует проводить шаг за шагом, пользуясь на каждой ступени упрощенными методами приближенного анализа.

В этой статье мы начнем рассмотрение с анализа систем, а затем перейдем к их синтезу. Сначала будут описаны элементы импульсных систем. Затем будут рассмотрены методы оперирования со структурными схемами (блок-схемами) с целью сведения сложных блок-схем к простым эквивалентным цепям. Наконец, будет проведено исследование реакций последовательной цепи элементов импульсной системы по распределению полюсов и нулей в плоскостях s , e^{-sT} и $e^{-sT/2}$. В заключение будут рассмотрены вопросы синтеза импульсных систем.

Анализ систем

Элементы систем

Обычные методы анализа можно применять к импульсным системам вследствие их линейности, которая следует из линейности их элементов. Линейность же элементов системы обусловлена их конструкцией.

Мы рассматриваем три основных типа элементов. Первый тип включает обычные непрерывные линейные фильтры, кото-

рые являются основными узлами обычных систем непрерывного действия. Действие такого фильтра описывается обычной передаточной функцией $K(s)$. Ко второму типу относятся дискретные линейные фильтры, особенность которых состоит в приеме и передаче только специфических сигналов. Такой фильтр осуществляет линейное преобразование входной последовательности импульсов, следующих через равные интервалы, и дает на выходе также равноотстоящие во времени импульсы, синхронные с входными. Его действие описывается передаточной функцией $D(E^{-sT})$, которую обычно выражают в функции переменной e^{-sT} (T — интервал между импульсами), а не в функции переменной s . Третьим основным элементом импульсной системы, в котором впервые резко выражены свойства дискретности, является импульсное устройство, прерывающее через каждые T сек. сигнал на входе и передающее в последующие элементы системы только дискретные сигналы. Хотя такое устройство часто рассматривается как переключатель, лучше определять его как модулятор, несущая частота которого образуется единичными импульсами, следующими один за другим через интервалы в T сек. На выходе его получаются импульсы, равные по площади входным сигналам в дискретные моменты времени. Следовательно, описанное устройство — импульсный модулятор — является линейным, так как увеличение в несколько раз амплитуды входных сигналов ведет к увеличению во столько же раз амплитуды импульсов на выходе. Таким образом, все элементы импульсной системы являются линейными, как и сама система в целом.

Хотя модуляция (импульсная) линейна, однако это тип линейности с изменяющимися во времени параметрами. Сигналы, поступившие в разные моменты времени, вызывают различные реакции системы. Поэтому обычная передаточная функция не может, вообще говоря, применяться для описания работы импульсных систем. Однако процесс модуляции, дающий простое изменение амплитуды, хотя и является переменным во времени, но не имеет сложных динамических особенностей. Если считать, что *все модуляторы работают синхронно и синфазно с несущими импульсами, причем первый импульс находится в точке $t=0$* , то все импульсно-модулированные сигналы и импульсные фильтры будут иметь простые изображения и передаточные функции, рациональные

относительно e^{-sT} . Зависимость от времени характеристик импульсных систем сказывается только при модуляции, где она учитывается достаточно просто и, следовательно, не является важным фактором при анализе систем.

Преобразование структурной схемы

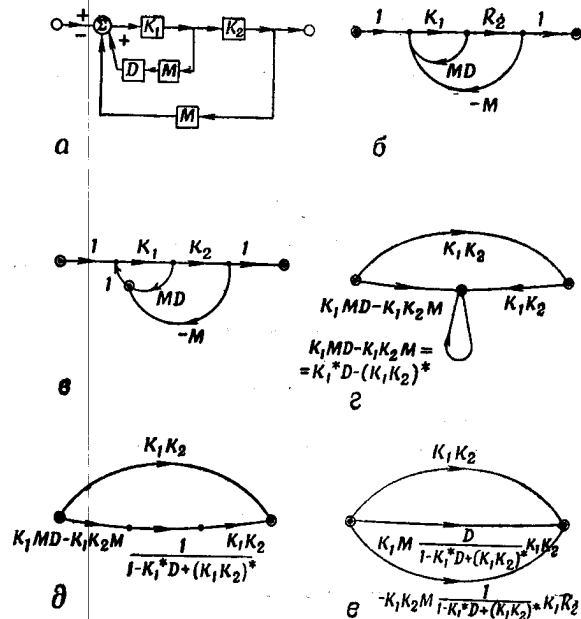
Первым шагом в проблеме анализа является преобразование данной системы в другую, что позволит применить более простые методы расчета. В первом примере мы покажем, как преобразовать достаточно сложную импульсную систему с обратной связью в простую эквивалентную систему без обратной связи, применяя обычные методы преобразования [3].

Рассмотрим блок-схему системы, изображенную на фиг. 1, а, которая на фиг. 1, б представлена в виде структурной схемы. Здесь M — импульсный модулятор, D — дискретный фильтр, K — непрерывный фильтр. Метод преобразования основывается на линейности системы и на возможности отыскания эквивалентной передаточной функции для замкнутой системы с одним контуром обратной связи, когда передаточные функции разомкнутой системы и обратной связи известны (обычная формула $K/(1-BK)$). В импульсных системах эта эквивалентность замкнутых и разомкнутых систем не совсем очевидна, так как система замыкается через модуляторы, действие которых не может быть описано обычными передаточными функциями. Однако если предварительно наложить на модулятор указанное выше требование синхронности, такая эквивалентность имеет место. *В результате любая импульсная система с обратной связью может быть заменена эквивалентной системой без обратной связи.*

Переходя далее к рассмотрению фиг. 1, в, мы используем метод исключения обратной связи, основанный на выборе узлов, в которых сигнал является чисто импульсным. Здесь необходимо выбрать только один такой узел. Мы закончим преобразование выбором наиболее подходящего узла в необходимой точке. Дальнейший выбор делать в нашем примере не имеет смысла. На фиг. 1, г сохранены только оставшиеся узлы, причем операции между ними записаны в символической форме. Эта новая диаграмма упрощается устранением петель вокруг узлов в порядке, обратном выбору этих узлов.

8. Частотный метод расчета импульсных систем 297

Единственный оставшийся узел, который мы должны рассмотреть в нашем примере (исключая узлы на входе и выходе системы), является узлом с чисто импульсным сигналом. Вследствие отмеченной выше синхронности действие этого замкнутого узла описывается эквивалентной передаточной



Фиг. 1. Преобразование структурной схемы импульсной системы с обратной связью в эквивалентную схему стандартного вида.

а — блок-схема; б — структурная схема; в — выбор остаточных узлов; г — нахождение передаточных функций; д — упрощение; е — стандартная схема.

функцией и он может быть заменен узлом с передаточной функцией $K/(1 - BK)$. Эти результаты представлены на фиг. 1, д и е.

Основным моментом этих преобразований явилось преобразование, позволившее нам заменить операцию над входными импульсами, осуществляемую непрерывным фильтром $K(s)$ и модулятором M , эквивалентной передаточной

функцией $K^*(s)$. Действительно, $K^*(s)$ является передаточной функцией дискретного фильтра, так как она преобразует импульсный сигнал в импульсный же, и поэтому ее можно записать в виде $K^*(e^{-sT})$. Функция K^* может быть определена следующим образом: на вход фильтра K подается импульс, определяется зависимость выходного сигнала от времени в виде суммы экспоненциальных членов и находится последовательность соответствующих им импульсов на выходе модулятора; к полученным выражениям применяется преобразование Лапласа, причем результат получается в форме ряда с коэффициентами вида e^{-sT} , после чего суммирование ряда дает решение в замкнутой форме.

В результате этого метода преобразования мы всегда можем, как показано на фиг. 1, е, получить эквивалентную систему, состоящую из нескольких стандартных последовательных цепей, соединенных параллельно. Каждая из этих цепей состоит из непрерывного фильтра, с которым последовательно соединены модулятор, импульсный фильтр и еще один непрерывный фильтр (имеются также „вырожденные“ системы). Следовательно, анализ реакции конкретной системы основан на анализе реакции основной последовательности элементов цепи K_a , M , D , K_b .

*Реакция последовательной цепи,
содержащей элементы K_a , M , D , K_b*

В качестве второго примера рассмотрим задачу нахождения реакции последовательно соединенных элементов K_a , M , D , K_b , представленных на фиг. 2, а ($K_a = 1$). Для этого мы используем преобразование Лапласа и анализ распределения полюсов и нулей на плоскости s .

Пусть в какой-то момент времени на вход системы подается единичный импульс $f_i(t)$. Мы можем считать, что такой сигнал возникает при подаче на вход фильтра с передаточной функцией $F_i(s)$ единичного импульса. Полная передаточная функция эквивалентного непрерывного фильтра, предшествующего модулятору, равна $F_i(s)K_a(s)$. Свяжем теперь с помощью передаточной функции эквивалентного дискретного фильтра $(F_i K_a)^*$ входной импульс с сигналом на выходе модулятора, т. е. рассмотрим преобразование сигнала модулятором. Передаточная функция системы, даю-

щей на выходе сигнал $f_0(t)$, будет равна $F_0 = (F_i K_a)^* DK_b$.
В нашем примере

$$\begin{aligned} (F_i K_a)^* &= \left(\frac{1}{s^2}\right)^* = 0 + Te^{-sT} + 2Te^{-2sT} + 3Te^{-3sT} + \dots = \\ &= \frac{Te^{-sT}}{(1 - e^{-sT})^2}; \\ F_0 = (F_i K_a)^* DK_b &= \frac{Te^{-sT}}{(1 - e^{-sT})^2} (1 - e^{-sT}) \frac{a}{s(s+a)} = \\ &= \frac{Te^{-sT}}{1 - e^{-sT}} \frac{a}{s(s+a)}. \end{aligned}$$

На фиг. 2, б дана картина распределения полюсов и нулей функции F_0 на плоскости s . Отметим две основные области этой диаграммы: центральную, полюсы которой соответствуют низкочастотной части реакции, и остальную, полюсы которой определяют высокочастотную часть реакции.

Диаграмма $F_0(s)$ на плоскости s характеризует реакцию с количественной стороны. В нашем примере эта диаграмма показывает, что выходной сигнал имеет линейно возрастающую компоненту (полюс второго порядка при $s=0$), возможную ступенчатую компоненту (искажаемую полюсом второго порядка), спадающую экспоненциально компоненту (полюс при $s=-a$) и пульсации постоянной амплитуды (простые полюсы при $s = \pm jn\Omega$, $n=1, 2, 3, \dots$, где Ω — частота повторения импульсов).

Так как диаграмма в общем случае имеет бесконечное число полюсов, мы не будем стремиться к полному количественному определению реакции, а постараемся подсчитать только ограниченное число компонент. Для приближенного расчета нам понадобится лишь применить преобразование Лапласа. Предположим, что нас интересуют лишь выходные сигналы в дискретные моменты времени. Мы найдем их, добавив к первоначальной системе второй синхронизованный модулятор. Применим наше основное преобразование, связав при помощи эквивалентного дискретного фильтра K_b^* импульсные сигналы, даваемые дискретным фильтром D , и выходные импульсы. Преобразование выходных импульсов,

следовательно, выражается функцией $F_0^* = (F_i K_a)^* DK_b^*$.
В примере системы фиг. 2, а мы имеем

$$\begin{aligned} F_0^* &= \left(\frac{1}{s^2}\right)^* (1 - e^{-sT}) \left[\frac{a}{s(s+a)}\right]^* = \\ &= \frac{T e^{-sT}}{(1 - e^{-sT})^2} (1 - e^{-sT}) \left[\frac{1}{s} - \frac{1}{s+a}\right]^* = \\ &= \frac{T e^{-sT}}{1 - e^{-sT}} \left[\frac{1}{1 - e^{-sT}} - \frac{1}{1 - e^{-aT} e^{-sT}}\right] = \\ &= \frac{T(1 - e^{-aT})(e^{-sT})^2}{(1 - e^{-sT})^2 (1 - e^{-aT} e^{-sT})}. \end{aligned}$$

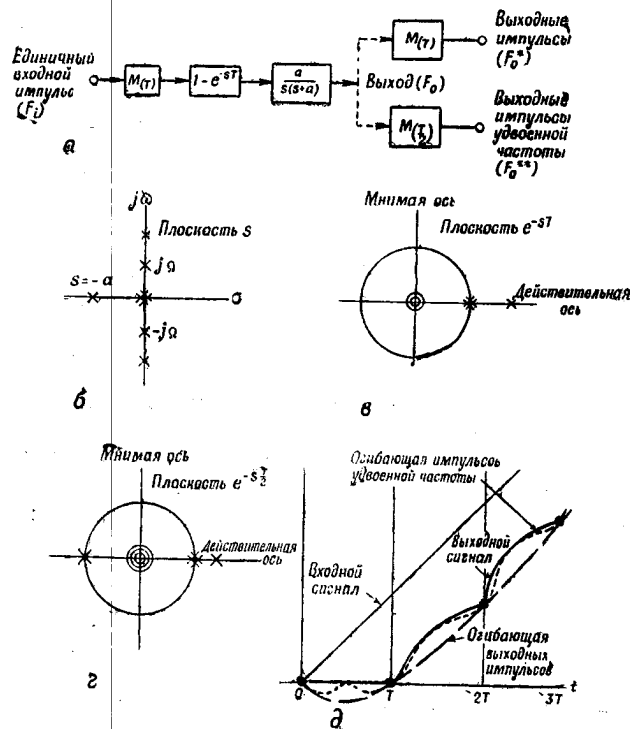
Функция F_0^* является рациональной функцией от e^{-sT} в отличие от функции F_0 , которая зависит не только от e^{-sT} , но и от s . Следовательно, F_0^* имеет ограниченное число полюсов в плоскости e^{-sT} , и каждому из них соответствует своя компонента реакции. Эти компоненты относятся только к низкочастотной части реакции, для которой и проводится расчет системы, поскольку поведение выходных сигналов в промежутках между импульсами не определено. Реакция может быть разложена на компоненты путем разложения F_0^* на простые дроби и приравнивания членов, содержащих функции e^{-sT} , членам, содержащим временные функции. *Изображение функции F_0^* на плоскости e^{-sT} дает возможность количественно оценить основную реакцию.*

Картина расположения нулей и полюсов функции F_0^* для нашего примера показана на фиг. 2, в. Полюс в точке $e^{-sT} = 1$ соответствует импульсным линейно возрастающей и ступенчатой компонентам, полюс $e^{-sT} = e^{aT}$ указывает на составляющую с экспоненциальным затуханием. Количественные вычисления дают

$$\begin{aligned} F_0^* &= \frac{T(1 - e^{-aT})(e^{-sT})^2}{(1 - e^{-sT})^2 (1 - e^{-aT} e^{-sT})} = \\ &= \frac{T e^{-sT}}{(1 - e^{-sT})^2} - \frac{T}{1 - e^{-aT}} \frac{1}{1 - e^{-sT}} + \\ &\quad + \frac{T}{1 - e^{-aT}} \frac{1}{1 - e^{-aT} e^{-sT}}. \end{aligned}$$

8. Частотный метод расчета импульсных систем 301

Член $\frac{Te^{-sT}}{(1-e^{-sT})^2} = 0 + Te^{-sT} + 2Te^{-2sT} + \dots$ представляет собой преобразование возрастающих импульсов. Аналогично



Ф и г. 2. Анализ переходного процесса системы, содержащей последовательно соединенные элементы.

a — блок-схема системы, возбуждаемой единичным импульсом; *б* — расположение полюсов и нулей преобразования Лапласа выходного сигнала на плоскости s ; *в* — расположение полюсов и нулей выходных импульсов на плоскости e^{-sT} ; *г* — расположение полюсов и нулей выходных импульсов удвоенной частоты на плоскости $e^{-sT/2}$; *д* — зависимость входного и выходного сигналов от времени.

$\frac{1}{1-e^{-sT}} = 1 + e^{-sT} + e^{-2sT} + \dots$ является преобразованием ступенчатой компоненты и $\frac{1}{1-e^{-aT}e^{-sT}} = 1 + e^{-aT}e^{-sT} + e^{-2aT}e^{-2sT} + \dots$ есть преобразование затуха-

хающих импульсов. Исследуя сигнал, мы можем найти огибающую импульсной реакции. В нашем примере $f_e(t) = t - \frac{T}{1 - e^{-aT}}(1 - e^{-at})$ для $t \geq 0$. Эта огибающая является приближенным представлением действительной реакции при пренебрежении пульсациями.

Для определения пульсаций методами преобразования мы предположим, что импульсы поступают на выход системы не с интервалами в T сек., а через каждые $T/2$ сек. Хотя фильтр D дает на выходе импульсы через каждые T сек., мы можем считать, что он удовлетворяет поставленному условию, полагая, что импульсы, поступающие на выход в моменты $T/2$, имеют амплитуду, равную 0. Применяя основное преобразование, мы сможем снова найти эквивалентный дискретный фильтр K_b^{**} , связывающий сигналы на выходе фильтра D с импульсами удвоенной частоты на выходе системы. Теперь, однако, аргументом K_b^{**} является переменная $e^{-s \frac{T}{2}}$ (вместо e^{-sT} для функции K_b^* и s — для функции K_b). Отсюда преобразование, соответствующее выходным импульсам удвоенной частоты, имеет вид:

$$F_0^{**} = (F_i K_a)^* D K_b^{**}.$$

В нашем примере

$$\begin{aligned} F_0^{**} &= \frac{T e^{-sT}}{1 - e^{-sT}} \frac{e^{-s \frac{T}{2}} \left(1 - e^{-a \frac{T}{2}}\right)}{\left(1 - e^{-s \frac{T}{2}}\right) \left(1 - e^{-a \frac{T}{2}} e^{-s \frac{T}{2}}\right)} = \\ &= \frac{T \left(1 - e^{-a \frac{T}{2}}\right) \left(e^{-s \frac{T}{2}}\right)^3}{\left(1 - e^{-s \frac{T}{2}}\right)^2 \left(1 + e^{-s \frac{T}{2}}\right) \left(1 - e^{-a \frac{T}{2}} e^{-s \frac{T}{2}}\right)}. \end{aligned}$$

Это выражение записано в виде рациональной функции от $e^{-s \frac{T}{2}}$. На фиг. 2, г показано распределение полюсов и нулей функции F_0^{**} на плоскости $e^{-s \frac{T}{2}}$. Сравнивая ее с фиг. 2, в, мы видим, что появляется новый полюс при

$e^{-s\frac{T}{2}} = -1$, обусловленный членом $\frac{1}{1 + e^{-s\frac{T}{2}}} = 1 -$

$-e^{-s\frac{T}{2}} + e^{-2s\frac{T}{2}} - e^{-3s\frac{T}{2}} + \dots$, который указывает на наличие в выходном сигнале колебательной компоненты. Это установившиеся пульсации. Следовательно, произведенное нами удвоение частоты импульсов выявило пульсации! *Изображение полюсов и нулей функции F_0^{**} на плоскости $e^{-s\frac{T}{2}}$ позволяет количественно охарактеризовать их и дать более точную оценку основной реакции.*

Если снова проделать все вычисления, можно определить величины этих компонент и получить огибающую импульсов удвоенной частоты на выходе, а также улучшить аппроксимацию действительной реакции. Огибающая получается в виде:

$$f_e = t - \frac{T}{1 - e^{-aT}} (1 - e^{-aT}) + \frac{T}{4} \frac{1 - e^{-a\frac{T}{2}}}{1 + e^{-a\frac{T}{2}}} \left(1 - \cos \frac{2\pi}{T} t\right).$$

На фиг. 2, д показаны входной и выходной сигналы и аппроксимирующие огибающие для обоих случаев. При необходимости аппроксимация может быть еще улучшена путем увеличения частоты выходных импульсов, так что применение преобразования Лапласа по существу дает полную картину реакции.

Синтез систем

Желаемые характеристики и условия их реализации

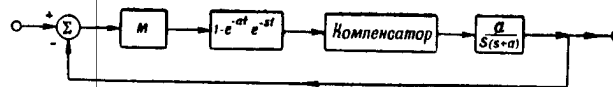
Решение задачи синтеза системы включает два аспекта. Первый касается желаемых характеристик системы, второй — возможности реализации той или иной системы. Оба аспекта тесно связаны и должны учитываться совместно. При расчете следящих систем сначала на основе общих соображений о мощности системы обычно выбирается определенный выходной элемент, который затем согласуется с системой при помощи компенсационных фильтров, чтобы обеспечить требуемые характеристики переходного процесса.

Мы прежде всего остановимся на желаемых характеристиках переходного процесса. В случае импульсных систем переходный процесс обычно считается приемлемым, если время переходного процесса мало и он быстро затухает, если возмущения в блоке прямой связи подавляются в достаточной степени, если ошибки при последующем сглаживании входных сигналов малы и если пульсация импульсов на выходе невелика. Большинство из указанных желаемых характеристик переходного процесса для конкретной системы можно определить при изучении упрощенной блок-схемы и исследовании распределения полюсов и нулей передаточной функции всей системы в плоскостях e^{-sT} и $e^{-sT/2}$. Насколько хорошо одноконтурная замкнутая система подавляет возмущения и следит за сглаженными входными сигналами, легко определить по коэффициентам ошибки [4].

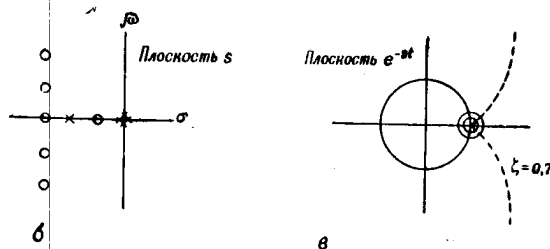
Рассмотрим далее, какие изменения характеристик реакции можно осуществить, используя компенсационные фильтры. Условия осуществимости непрерывного или импульсного компенсационного фильтра связаны с тем обстоятельством, что ни тот, ни другой фильтры не могут давать на выходе сигнал до их возбуждения. Влияние условий физической осуществимости данного фильтра на условия физической осуществимости передаточных характеристик всей системы зависит от выбора структурной схемы системы, но обычно передаточная функция одноконтурной замкнутой системы с обычным выходным элементом удовлетворяет этим условиям, если она имеет нуль в начале координат на плоскости e^{-sT} (этот нуль появляется при учете запаздывания в элементе мощности). Так же как и при расчете следящих систем непрерывного действия, мы можем получить только условия физической осуществимости отдельных фильтров, используемых для компенсации импульсных следящих систем, но можно показать, что осуществление компенсации в импульсных системах обычно сопряжено с меньшими трудностями, чем в непрерывных системах. Физическая причина этого состоит в том, что применение импульсов в системе сужает полосу частот до половины частоты повторения импульсов, и этот ограниченный диапазон частот обычно допускает применение легко осуществимых компенсационных фильтров.

Пример. Компенсация одноконтурной замкнутой следящей системы

Система, подлежащая компенсации, показана на фиг. 3, а. Эквивалентная система без обратной связи для нашей некомпенсированной системы представляется простой

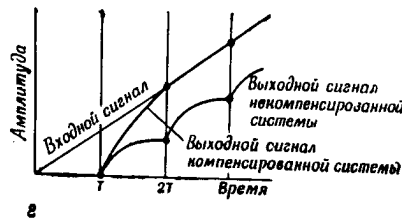


а



б

в



г

Фиг. 3. Расчет компенсатора для простой импульсной следящей системы.

а — блок-схема следящей системы, подлежащей компенсации; б — расположение полюсов и нулей импульсной передаточной функции некомпенсированной системы; в — расположение полюсов и нулей функции $1/(1+K^*)$ в плоскости e^{-st} , ограниченная контуром ζ ; г — сравнение переходных процессов компенсированной и некомпенсированной систем.

последовательностью элементов, изображенной на фиг. 2, переходный процесс которой уже определен. Он показан на фиг. 2, д. Задачей синтеза является устранение установившегося запаздывания в переходном процессе путем

добавления компенсирующего устройства в блок прямой связи. Коэффициенты ошибки по скорости при этом считаются равными нулю.

Систему без компенсации необходимо сначала подвергнуть анализу. Как показано в работах [1] и [2] и как легко проверить методом преобразования структурных схем (блок-схем), рассмотренным ранее, система с одним контуром отрицательной обратной связи, охватывающим блок модулятор — фильтр, может быть преобразована в эквивалентную систему следующего типа. Обозначим через DK передаточную функцию фильтра, состоящего из преобразователя дискретной части сигнала D и преобразователя непрерывной части K ; тогда эквивалентная система будет состоять из последовательной цепи элементов, соединенных в следующем порядке: модулятор, фильтр с передаточной функцией $1/(1+DK^*)$, фильтр с передаточной функцией DK . В нашем примере

$$D = (1 - e^{-aT}e^{-sT}); \quad K = \frac{a}{s(s+a)};$$

$$DK^* = \frac{(1 - e^{-aT}e^{-sT})(1 - e^{-aT})e^{-sT}}{(1 - e^{-sT})(1 - e^{-aT}e^{-sT})} = \frac{(1 - e^{-aT})e^{-sT}}{1 - e^{-sT}};$$

$$\frac{1}{1 + DK^*} = \frac{1 - e^{-sT}}{1 - e^{-aT}e^{-sT}}; \quad \frac{DK}{1 + DK^*} = \frac{(1 - e^{-sT})a}{s(s+a)}.$$

Так как эквивалентный фильтр $1/(1 + DK^*)$ передает сигналы в фильтр DK , то выходной сигнал фильтра $1/(1 + DK^*)$ является импульсной ошибкой системы. Коэффициенты ошибки являются соответствующими амплитудами сигналов ошибки при $t=0$ для серии единичных импульсов, последовательности линейно возрастающих импульсов и т. п., подаваемых на вход системы начиная с момента $t = -\infty$. *Метод определения коэффициентов ошибки совершенно аналогичен методу, используемому в обычных системах, исключая тот факт, что мы имеем дело с дискретными сигналами.*

В общем для импульсных систем с одной обратной связью k -й коэффициент ошибки можно представить выражением

$$\epsilon_k = \frac{1}{k!} \frac{d^k}{ds^k} \left[\frac{1}{1 + DK^*} \right]_{s=0, \text{ или } e^{-sT}=1}$$

Система с равным нулю коэффициентом ошибки по положению и нулевым коэффициентом ошибки по скорости будет иметь нулевую установившуюся ошибку при подаче на вход сигнала, состоящего из совокупности линейно возрастающих импульсов. В нашем примере для такого сигнала, преобразование которого равно $Te^{-sT}/(1 - e^{-sT})^2$, импульсную ошибку можно представить в виде

$$\frac{Te^{-sT}}{(1 - e^{-sT})^2} \frac{1}{1 + DK^*} = \frac{Te^{-sT}}{(1 - e^{-sT})(1 - e^{-sT}e^{-aT})}$$

Множитель $(1 - e^{-sT})$ в знаменателе указывает на наличие установившейся компоненты. Для устранения его, очевидно, необходимо добавить в числитель выражения $1/(1 + DK^*)$ член $(1 - e^{-sT})$. В обычных системах с одной обратной связью в этом случае говорят о необходимости дополнительного интегрирования. Здесь член $(1 - e^{-sT})$ аналогичен множителю s в обычных системах. Отметим, что

$$\lim_{T \rightarrow 0} \frac{1}{T} (1 - e^{-sT}) = s.$$

Проведя анализ системы без компенсации, вернемся к проблеме определения типа характеристик реакции, которые были бы желательны и осуществимы. Так как передаточная функция от входного импульса к импульсной ошибке имеет простой вид и описывает многие важные свойства переходного процесса, мы выберем такую передаточную функцию, которая была бы приемлемой и одновременно позволила бы установить физическую осуществимость компенсационного фильтра. В общем случае компенсационные фильтры могут быть включены или в основную цепь, или в обратную связь, а также во второстепенные цепи. В нашем примере мы предположим, что фильтр включен последовательно в основной контур системы.

Передаточная функция импульсного компенсатора, включенного последовательно в основной контур, является рациональной функцией от e^{-sT} и может быть представлена как частное двух многочленов от e^{-sT} в виде P/Q . Передаточная функция импульсной ошибки равна $1/(1 + P/Q) = Q/(P + Q)$. Нули этой передаточной функции соответствуют

полюсам выражения P/Q , полюсы этой передаточной функции соответствуют собственным частотам системы и функция $Q/(P+Q)$ может быть использована для определения коэффициентов ошибки. Возникает вопрос, возможна ли реализация такой системы? Для этого лучше всего рассмотреть выражение P/Q .

Наиболее простой путь выяснения осуществимости такой передаточной функции состоит в том, чтобы, идя обратно, найти по ней компенсирующий фильтр. Для случая непрерывной компенсации K_c может быть подсчитана следующим образом. Во-первых, из P/Q выделяем передаточную функцию дискретного фильтра. Далее находим передаточную функцию непрерывной части этого отношения. K_c является передаточной функцией, которая должна быть кратна передаточной функции фиксированной непрерывной части.

$$\frac{P}{Q} \frac{1}{D} = \left[K_c \frac{a}{s(s+a)} \right]^*$$

Данное выражение $\left[K_c \frac{a}{s(s+a)} \right]^*$ представляет возможность выбора величин $K_c \frac{a}{s(s+a)}$. Осуществимое значение K_c всегда можно выбрать, если $\left[K_c \frac{a}{s(s+a)} \right]^*$ имеет нуль в начале координат плоскости e^{-sT} . Обычно используется низкочастотный фильтр K_c , если он реализуем. Отметим, что некомпенсированная система K^*D имеет нуль в начале координат на плоскости e^{-sT} . Этот нуль соответствует тому факту, что импульсный переходный процесс начинается в нуле, так как число полюсов в его передаточной функции на две единицы больше, чем число нулей. Чтобы переместить этот нуль из начала координат, необходимо ввести в основную цепь блок, импульсный переходный процесс которого имеет величину, отличную от нуля при $t=0$, или предположить, что двигатель получает управляющий сигнал до момента $t=0$. Очевидно, это невозможно, так как компенсационный фильтр при этом должен был бы возбуждаться до подачи входного сигнала, чтобы «разогнать двигатель». (Детально условия физической осуществимости рассмотрены в статье [2].)

В нашем расчете мы выберем

$$\frac{1}{1 + DK_n^*} = \frac{Q}{P + Q} = \left(\frac{1 - e^{-sT}}{1} \right)^2.$$

При таком выборе имеем равную нулю ошибку по скорости и отсутствие полюсов у преобразования Лапласа. Отсюда мы найдем новое значение K , равное K_n , и определим, каким образом необходимо скомпенсировать K , чтобы получить K_n . Так как $Q/(P + Q)$ дано, мы сразу находим

$$P = 1 - (1 - e^{-sT})^2 = e^{-sT}(2 - e^{-sT})$$

и

$$\frac{P}{DQ} = \frac{e^{-sT}(2 - e^{-sT})}{(1 - e^{-aT}e^{-sT})(1 - e^{-sT})} = K_n^*;$$

$$K_n^* = \frac{Ae^{-sT}}{(1 - e^{-sT})^2} + B \left(\frac{1}{1 - e^{-sT}} - \frac{1}{1 - e^{-aT}e^{-sT}} \right),$$

или

$$K_n = \frac{A}{s^2} + B \left(\frac{1}{s} - \frac{1}{s + a} \right) = \frac{(A + aB)s + aA}{s^2(s + a)},$$

где

$$A = \frac{1}{T(1 - e^{-aT})}; \quad B = \frac{1 - 2e^{-aT}}{(1 - e^{-aT})^2}.$$

Следовательно,

$$K_c = \frac{K_n}{K} = \frac{s(s + a)}{a} \cdot K_n = \frac{(A/a + B)s + A}{s}.$$

Фильтр K_c интегрирует и усиливает входные сигналы и, очевидно, является физически осуществимым.

Для оценки полученных результатов отметим, что компенсация преобразует передаточную функцию системы M в выражение

$$\frac{DK_n}{1 + DK_n^*} = (1 - e^{-sT})(1 - e^{-aT}e^{-sT}) \frac{(A + aB)s + aA}{s^2(s + a)}.$$

Если на вход подан импульс, то преобразование выходного сигнала будет иметь вид:

$$F_0 = (F_i)^* \frac{DK_n}{1 + DK_n^*} = Te^{-sT}(1 - e^{-aT}e^{-sT}) \frac{(A + aB)s + aA}{s^2(s + a)}.$$

Изображение F_0 на плоскости s (см. фиг. 3, б) показывает, что оно имеет полюсы только при $s = 0$ и $s = -a$. Следовательно, пульсации отсутствуют. Компенсация привела

к исчезновению пульсаций. Целью расчета являлось получение системы с нулевой ошибкой по скорости, и мы нашли, что она равна нулю. Чтобы получить дополнительные количественные характеристики выходного сигнала, определим импульсы на выходе. Получим

$$F_0^* = (F_i) \frac{DK_n^*}{1 + DK_n^*} = \frac{T(e^{-sT})^2(2 - e^{-sT})}{(1 - e^{-sT})^2} = \frac{Te^{-sT}}{(1 - e^{-sT})^2} - Te^{-sT}.$$

Этот результат показывает, что первые два импульса равны нулю, а затем выходной сигнал прямо следует за входным. На фиг. 3, г изображены кривые переходных процессов компенсированной и некомпенсированной систем. Очевидно, что достигнуто значительное улучшение.

Отметим, что если мы по некоторым причинам пожелаем использовать другие полюсы, вместо единицы в знаменателе $1/(1 + DK_n^*)$, то изображенная на фиг. 3, в диаграмма $1/(1 + DK_n^*)$ на плоскости e^{-sT} дает характеристику затухания, соответствующую этим полюсам. На диаграмме можно изобразить контур ζ , показывающий относительное затухание указанными линиями (перенесенными из плоскости s). Для определения характера изменения затухания с увеличением воздействия можно также применить метод исследования расположения корней $1/(1 + DK_n^*)$ на плоскости e^{-sT} .

Возможно также компенсировать рассмотренную систему другим методом, используя компенсирующий дискретный фильтр. Однако это не даст улучшения характеристик пульсаций в системе.

ЛИТЕРАТУРА

1. Linvill W. K., Sampled-Data Control Systems Studied Through Comparison of Sampling with Amplitude Modulations, *Trans. AIEE*, 70, 1779 (1951).
2. Linvill W. K., Sittler R. W., Design of Sampled-Data System by Extension of Conventional Techniques, Project Whirlwind Report R-222, Mass. Inst. Technol., 1953.
3. Mason S. I., Some Properties of Signal Flow Graphs, *Proc. IRE* (в печати).
4. Теория следящих систем, под ред. Х. Джеймса, Н. Никольса Р. Филлипса, Издательство иностранной литературы, 1953 г.
5. Raggazzini I. R., Zadeh L. A., The Analysis of Sampled-Data Systems, *Trans. AIEE*, Part II, 71, 225 (1952).

**9. АНАЛИЗ СИСТЕМ АВТОМАТИЧЕСКОГО РЕГУЛИРОВАНИЯ,
СОДЕРЖАЩИХ ЦИФРОВЫЕ ВЫЧИСЛИТЕЛЬНЫЕ МАШИНЫ¹⁾**

В. К. Линвилл, Дж. М. Зальцер

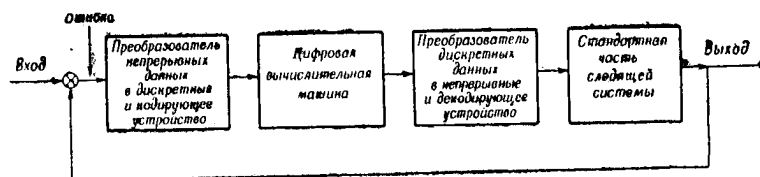
Введение

В настоящее время большинство исследований по цифровым вычислительным машинам непосредственно посвящено созданию устройств для численного решения научных и технических задач. Однако становится все более очевидным, что цифровые машины будут также играть важную роль в автоматическом регулировании. Вполне возможно, например, представить использование цифровых машин для регулирования автоматических процессов химического производства или для выполнения вычислений, связанных с использованием данных радиолокационной станции сопровождения в системе управления самолетом. Цифровая машина становится обычным элементом в замкнутых системах автоматического регулирования, в которых ее действие сравнимо с действием модели непрерывного действия или моделирующего фильтра. Можно считать, что точно так же, как моделирующее устройство дает решение дифференциального уравнения, цифровая машина позволяет получить решение разностного уравнения. Для применений важно, чтобы цифровая машина была достаточно быстродействующей, т. е. чтобы она могла завершать операции в реальные сроки.

Вопрос о преимуществах и недостатках применения цифровых вычислительных машин для целей управления выходит за рамки настоящей статьи. Можно лишь указать, что каждый случай такого применения должен рассматриваться с учетом особенностей конкретной задачи. Некоторые соображения, которые приходится принимать в расчет, станут ясными из дальнейшего изложения. Можно поставить следующую задачу: допустим, что принято решение использовать цифровую

¹⁾ Linvill W. K., Salzer J. M., *Proc. IRE*, 41, № 7, 901—906 (1953).

машину в системе регулирования; как проанализировать работу такой системы? Чтобы ответить на этот вопрос, необходимо прежде всего выяснить, в чем состоит основное различие между смешанной системой (в цепь которой включена цифровая машина) и обычной системой регулирования.



Фиг. 1. Простая следящая система, содержащая цифровую вычислительную машину.

Настоящая статья посвящена частному вопросу теории смешанных систем, связанному с неизбежным преобразованием сигнала к дискретному виду при прохождении через цифровую вычислительную машину. Действительно, любое устройство, действие которого дискретно во времени, может оперировать с данным входным сигналом лишь в течение определенного промежутка времени и тем самым преобразует непрерывные данные в дискретные. В вычислительных машинах современного вида арифметическое устройство используется для выполнения операций над последовательностью импульсов; в то же время последовательности импульсов служат для выполнения каждого шага требуемой операции.

Рассмотрим фиг. 1, на которой представлена блок-схема простой следящей системы, содержащей цифровую машину. Это лишь одна из большого числа возможных схем, в которых в качестве элемента системы регулирования может использоваться цифровая машина. Но анализ данной конкретной схемы позволит нам выяснить общие проблемы, с которыми мы сталкиваемся во всех подобных системах. Импульсно-кодирующее устройство преобразует непрерывные данные в дискретные, а затем представляет их в виде чисел для использования в цифровой вычислительной машине. Цифровая машина производит ряд операций над входным сигналом. Затем с выхода цифровой машины сигнал, который еще имеет дискретную числовую форму, поступает в преобразующее и декодирующее устройство. Последнее преобразует дискрет-

9. Анализ систем, содержащих цифровые машины 313

ный сигнал в непрерывный и подобный исходному для использования в обычных „непрерывных“ элементах системы.

При анализе системы, изображенной на фиг. 1, возникают три новые проблемы, связанные с описанием следующих процессов: 1) процесса получения дискретных данных на входе цифровой машины; 2) операций, производимых цифровой машиной над входным дискретным сигналом; 3) процесса получения непрерывного сигнала из дискретного на выходе цифровой машины. Остальная часть системы регулирования состоит из обычных элементов и производит операции уже над непрерывным сигналом. Обычно основную часть системы составляют стандартные элементы.

Анализ подобных смешанных систем значительно упрощается при описании работы всех ее элементов при помощи одинакового рода (либо частотных, либо временных) характеристик. В качестве общей основы для исследования в настоящей работе выбран частотный метод. К этому имелись следующие основания: 1) частотный метод оказался весьма эффективным при анализе обычных систем автоматического регулирования; в то же время известно, что система в значительной степени сохраняет свои обычные свойства даже тогда, когда в нее в качестве элемента введена цифровая вычислительная машина; 2) процессы получения дискретных данных из непрерывных и непрерывных из дискретных весьма удобно исследовать в частотной области; 3) операции, производимые цифровой машиной, часто являются линейными и, следовательно, могут быть легко рассмотрены в частотной области.

Вкратце результаты настоящей работы могут быть сформулированы следующим образом:

1. Показано, что процесс получения дискретных данных аналогичен процессу амплитудной модуляции.

2. Процесс получения непрерывных данных из дискретных аналогичен процессу сглаживания при амплитудной демодуляции.

3. Линейные операции, выполняемые цифровой машиной, могут быть описаны передаточными функциями.

Системы регулирования, содержащие цифровые машины, совершенно аналогичны обычным системам, но их можно представить себе состоящими из двух частей: в одной части распространяется обычный сигнал, а в другой — сигнал в виде огибающей,

полученной в результате детектирования амплитудно-модулированной несущей. Ту часть системы, в которой распространяется обычный сигнал, мы условимся называть стандартной, а часть, содержащую цифровую вычислительную машину, в которой распространяется амплитудно-модулированный сигнал,— нестандартной.

В настоящей статье будет проведен частотный анализ только нестандартной части системы регулирования. В предыдущей работе [1], посвященной системам прерывистого регулирования, был дан анализ системы, содержащей импульсный элемент, однако еще не был рассмотрен случай, когда система содержит цифровую вычислительную машину. Решение задачи анализа сложной системы, содержащей цифровую машину, получается сразу, как только определена передаточная функция цифровой машины. Цель настоящей работы состоит в том, чтобы в общих чертах показать, как в обычную систему регулирования может быть включена цифровая вычислительная машина. Более полное исследование цифровых машин и их свойств частотным методом будет дано одним из авторов в следующей статье¹⁾.

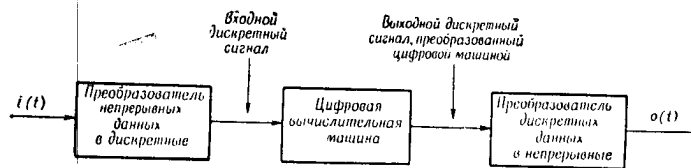
Эквивалентная схема

Рассмотрим фиг. 2, на которой показана нестандартная часть смешанной системы. Поскольку непосредственное описание дискретных сигналов на входе и выходе в частотной области невозможно, то следует пользоваться эквивалентными данными. В частности, если в схеме фиг. 2 мы заменим каждое дискретное значение импульсом, площадь которого равна соответствующему дискретному сигналу, то мы получим эквивалентную цепь, схема которой показана на фиг. 3. Эквивалентность схем фиг. 2 и 3 легко установить, если будет показано, что отношение $o(t)/i(t)$ в обоих случаях является одинаковым. В реальном устройстве сигнал $i(t)$ является дискретной последовательностью значений, следующих друг за другом через одинаковые интервалы времени. Эти значения можно кодировать для ввода в цифровую машину в виде последовательности чисел, записанных в двоичной системе исчисления. В эквивалентной цепи сигнал $i(t)$

¹⁾ См. следующую статью сборника.— *Прим. ред.*

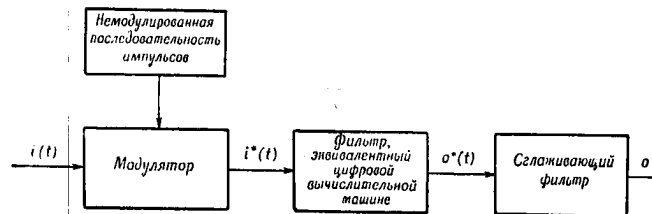
9. Анализ систем, содержащих цифровые машины 315

модулирует несущую в виде последовательности остроконечных импульсов единичной площади с частотой повторения $1/T$, где T — промежуток времени между двумя импульсами. Между площадями импульсов этой последовательности и дискретными кодированными числами, фактически поступающими на вход цифровой машины, имеется однозначное



Фиг. 2. Операции цифровой вычислительной машины над непрерывным входным сигналом.

соответствие. Эквивалентный дискретный фильтр производит операции над последовательностью импульсов точно так же, как цифровая машина производит операции над дискретной последовательностью чисел. Отсюда следует, что площади импульсов на выходе эквивалентного дискретного фильтра

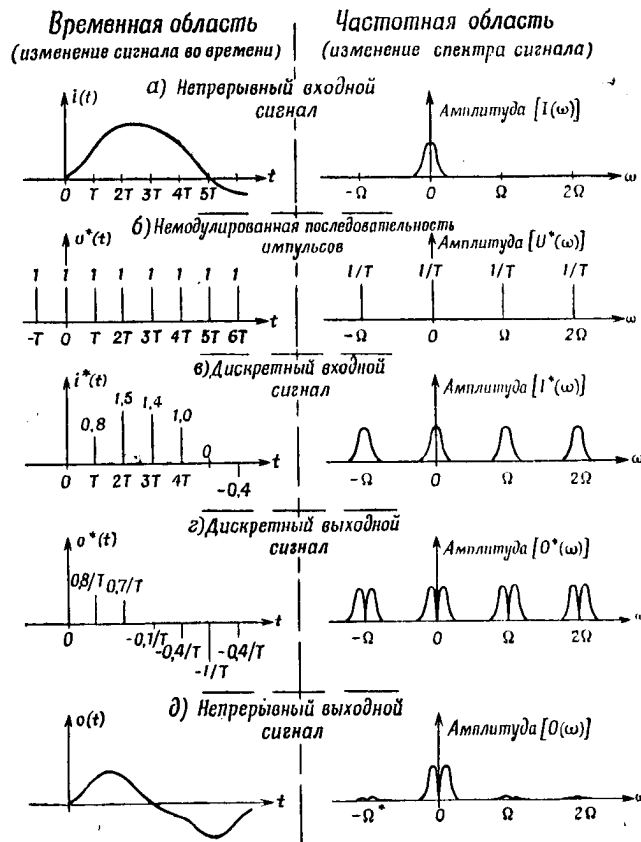


Фиг. 3. Эквивалентная схема к фиг. 2.

однозначно соответствуют числам на выходе цифровой вычислительной машины. Таким образом, демодулятор эквивалентной схемы на фиг. 3 имеет точно такую же импульсную переходную характеристику, как и истинный преобразователь дискретных данных в непрерывные. Теперь уже легко показать, что на выходе истинной и эквивалентной цепей мы получим эквивалентные сигналы.

Установив эквивалентность между действительной и эквивалентной цепями, рассмотрим последнюю более подробно. Обратимся к фиг. 4, на которой для частного случая

показана форма сигнала в различных точках эквивалентной цепи. Цифровая вычислительная машина в данном случае используется для дифференцирования входной функции. На



Фиг. 4. Вид сигнала в различных точках эквивалентной цепи фиг. 3.

фиг. 4 показано влияние каждого шага при последовательном преобразовании сигнала, как во временном, так и в частотном представлении.

Процесс получения дискретных данных

Как уже указывалось, процесс получения дискретных значений $i(t)$ аналогичен модуляции последовательности импульсов сигналом $i(t)$.

В результате разложения несущей последовательности единичных импульсов в ряд Фурье мы получим постоянную составляющую, основную гармонику с периодом T и бесчисленное количество гармоник с равными амплитудами. Для того чтобы выразить это в математической форме, мы обозначим через $u^*(t)$ последовательность единичных импульсов несущей частоты и запишем

$$u^*(t) = \sum_{k=-\infty}^{\infty} u(t - kT) = \frac{1}{T} \sum_{k=-\infty}^{\infty} e^{-jk\Omega T}, \quad (1)$$

где $u(t)$ — импульсная функция; T — период между импульсами; $\Omega = 2\pi/T$ — угловая частота, а e — основание натуральных логарифмов. Из уравнения (1) видно, что амплитуды всех гармоник равны $1/T$. На фиг. 4, а показан непрерывный входной сигнал $i(t)$, а на фиг. 4, б — несущая последовательности единичных импульсов $u^*(t)$. Процесс модуляции последовательности импульсов равносильен модуляции каждого члена ряда Фурье этой последовательности.

Составляющая, которая получится в результате модуляции постоянного члена ряда Фурье истинным сигналом $i(t)$, равна этому сигналу, умноженному на $1/T$. Составляющие, полученные в результате модуляции синусоидальных членов, дают бесконечно большое число побочных линий в спектре, соответствующих несущим частотам Ω , 2Ω , 3Ω , и т. д.

Назовем модулированные синусоиды побочными сигналами. Они возникают в процессе модуляции и появляются на выходе вместе с полезным сигналом. На фиг. 4, в показаны полезный сигнал и несколько побочных составляющих, возникающих вследствие модуляции синусоидальных гармоник. Изображение Лапласа дискретного сигнала $I^*(s)$ связано с изображением непрерывного сигнала при помощи следующей формулы:

$$I^*(s) = \frac{1}{T} \sum_{k=-\infty}^{+\infty} I(s + jk\Omega). \quad (2)$$

Процесс преобразования дискретных данных в непрерывные

Частотное рассмотрение процесса получения дискретного сигнала позволяет выяснить, как должен происходить обратный процесс получения непрерывного сигнала. Поэтому прежде чем заняться анализом работы цифровой вычислительной машины, нам следует рассмотреть вопрос о преобразовании дискретных данных в непрерывные.

Прежде всего мы выясним, как можно вновь получить первоначальный сигнал, если за процессом преобразования сигнала в дискретную форму непосредственно следует процесс получения непрерывного сигнала. Затем мы покажем, что те же рассуждения применимы и для случая преобразования дискретного сигнала в непрерывный на выходе цифровой вычислительной машины. Наконец, мы выясним, какое влияние оказывает цифровая машина на дискретный входной сигнал. Из рассмотрения фиг. 4, *a* и *b* и уравнения (2) ясно, что для того чтобы из дискретного сигнала вновь получить истинный непрерывный сигнал, необходимо просто отфильтровать побочные составляющие так, чтобы остался только полезный сигнал. Таким образом, если процесс получения дискретного сигнала подобен процессу модуляции, то обратный процесс подобен процессу сглаживания при демодуляции.

При фильтрации дискретного сигнала, точно так же как и при фильтрации демодулированного сигнала, невозможно избавиться от всех побочных гармоник. Напомним, что сглаживание при демодуляции становится более трудным для широкополосных сигналов, поскольку спектры сигнала и гармоник, подлежащих фильтрации, накладываются друг на друга в большей полосе частот. Это остается справедливым и для случая получения непрерывного сигнала из дискретного.

В то время как спектр полезного дискретного сигнала сконцентрирован вокруг нулевой частоты, спектры гармоник сконцентрированы вокруг частот Ω , 2Ω , 3Ω ... Низкочастотный фильтр может пропустить полезный сигнал и задержать побочные гармоники только в том случае, если имеется достаточное расстояние по оси частот между спектрами желаемого и нежелательного сигналов. Когда ширина спектра полезного сигнала приближается к величине $\Omega/2$, фильтрация становится невозможной вследствие того, что спектры желаемого

мого и мешающего сигналов накладываются друг на друга. Даже в том случае, когда спектры полезного и мешающего сигналов не накладываются друг на друга, необходимо выбрать оптимальное соотношение между фильтрующими свойствами фильтра и запаздыванием по фазе. Хорошая фильтрация может быть получена только ценой значительного запаздывания по фазе.

Цель устройства для получения непрерывного сигнала из дискретного на выходе цифровой машины состоит не в том, чтобы вновь получить непрерывную функцию, которая существовала до этого, а в получении новой непрерывной функции $o(t)$ из дискретного выходного сигнала $o^*(t)$ цифровой машины. Этот процесс показан на фиг. 4, г и д. Тем не менее основное требование остается пока тем же, что и раньше, т. е. устройство для получения непрерывного сигнала из дискретного сигнала $o^*(t)$ должно для фильтрации мешающих сигналов обладать характеристикой фильтра низких частот.

Операции, производимые цифровой вычислительной машиной

Наличие цифровой машины в системе не меняет природы дискретного сигнала, а только изменяет величину дискретных значений. Существенное преимущество цифровой машины состоит в том, что она производит действия над отдельными дискретными значениями с большой точностью.

Для того чтобы получить выходную последовательность $o^*(t)$, цифровая машина выполняет над входной последовательностью $i^*(t)$ совершенно определенную операцию. Связь между этими двумя последовательностями дается переходной характеристикой цифровой машины, действие которой определяется серией инструкций, имеющихся в ее заполняющем устройстве. Эта серия инструкций называется программой цифровой вычислительной машины. В том случае когда операции являются линейными, программу можно охарактеризовать передаточной функцией, причем имеется далеко идущая аналогия между составлением программ и расчетом линейных цепей.

Рассмотрим пример, изображенный на фиг. 4. В этом случае задачей цифровой машины является дифференцирование

входного сигнала. Напомним, что операция дифференцирования определяется соотношением

$$\frac{d}{dt} i(t) = \lim_{T \rightarrow 0} \frac{i(t) - i(t-T)}{T}. \quad (3)$$

Но поскольку цифровая машина оперирует над дискретными числами, то получить при ее помощи производную, строго говоря, нельзя. Временной интервал T между дискретными значениями сигнала должен оставаться конечным, ибо в противном случае мы не будем иметь дело с системой дискретного счета.

Цифровая машина может произвести операцию дифференцирования лишь приближенно, без перехода к пределу. В таком случае сигнал на выходе счетной машины можно записать в виде

$$o^*(t) = \frac{i^*(t) - i^*(t-T)}{T}, \quad (4)$$

где функция $i(t)$ в (3) заменена последовательностью импульсов $i^*(t)$. Это допустимо в силу того, что площадь каждого импульса $i^*(t)$ равна дискретному значению сигнала $i(t)$, взятому для того же момента времени.

Уравнение (4) аналогично линейному разностному уравнению, которое в общепринятых обозначениях можно записать в виде

$$o_j = \frac{l_j - l_{j-1}}{\Delta t}. \quad (5)$$

Согласно уже рассмотренной эквивалентной схеме, дискретные значения, входящие в (5), заменяются импульсами, которые имеют величину площади, равную дискретному значению.

Если преобразования Лапласа импульсных функций $o^*(t)$ и $i^*(t)$ соответственно равны $O^*(s)$ и $I^*(s)$, то преобразование Лапласа обеих частей (4) дает

$$O^*(s) = \frac{I^*(s) - I^*(s)e^{-sT}}{T}, \quad (6)$$

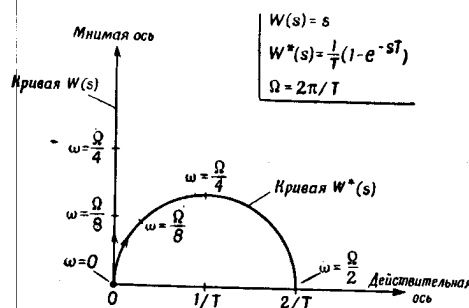
где множитель e^{-sT} соответствует запаздыванию T . По определению, передаточной функцией $W^*(s)$ программы вычислительной машины называется отношение преобразования

Лапласа выходной функции к преобразованию Лапласа входной функции. Таким образом, из (6) имеем

$$W^*(s) = \frac{O^*(s)}{I^*(s)} = \frac{1 - e^{-sT}}{T}. \quad (7)$$

Интересно исследовать точность этой приближенной формулы дифференцирования по сравнению с точным оператором дифференцирования

$$W(s) = s. \quad (8)$$



Фиг. 5. Сравнение истинной операции дифференцирования с разностной операцией, выполняемой цифровой машиной.

Кривые, соответствующие $W(s)$ и $W^*(s)$ при $s = j\omega$, показаны на фиг. 5, из рассмотрения которой сразу следуют хорошо известные факты. Кривая $W(s)$ совпадает с мнимой осью $j\omega$, в то время как кривая $W^*(s)$, определяемая уравнением $(1/T)(1 - e^{-j\omega T})$, есть окружность. При этом, как видно из фиг. 5, при возрастании частоты от $\omega = 0$ до $\omega = 1/T = \Omega/2$ точка, соответствующая этим частотам, перемещается по окружности по часовой стрелке. Из рассмотрения кривых фиг. 5 видно, что пока частота сигнала $i(t)$ мала по сравнению с частотой прерывания, формула (5) дает хорошую аппроксимацию формулы точного дифференцирования.

Для нулевой частоты, т. е. для случая $i(t) = \text{const}$, разностная формула совпадает с точной формулой дифференцирования. При любых других частотах операция, выполняемая вычислительной машиной, как по амплитуде, так и

по фазе, будет несколько отклоняться от истинной операции дифференцирования, однако при любой скорости изменения $i(t)$ частотный метод позволяет точно оценить эти ошибки. Например, если известен спектр входного сигнала, то может быть найдена среднеквадратичная ошибка программы.

Интересно рассмотреть, как в процессе предельного перехода (см. формулу (3)) кривая $W^*(s)$ переходит в кривую $W(s)$. Поскольку при $T \rightarrow 0$ центр и радиус окружности кривой $W^*(j\omega)$ (см. фиг. 5) стремятся к бесконечности как $1/T$, то ясно, что эта окружность превращается в прямую линию, совпадающую с мнимой осью. Из (7) легко получаем

$$\begin{aligned} \lim_{T \rightarrow 0} W^*(s) &= \lim_{T \rightarrow 0} \frac{1 - 1 + sT - 1/2(sT)^2 + \dots}{T} = \\ &= \lim_{T \rightarrow 0} \left(s - \frac{T}{2} s^2 + \dots \right) = s = W(s). \quad (9) \end{aligned}$$

Из рассмотрения фиг. 4, в и г видно влияние программы вычислительной машины. Например, если рассматривать преобразование сигнала во времени, то для площади импульса выходного сигнала в момент $4T$ имеем $\sigma^*(4T) = -0,4/T$; действительно, из формулы (4) мы при этом получаем

$$\frac{1}{T} [i^*(4T) - i^*(3T)] = \frac{1}{T} [1,0 - 1,4] = -\frac{0,4}{T}.$$

В частотной области операция дифференцирования сводится к умножению амплитуды входного сигнала на $|W^*(j\omega)|$. Отметим, что цифровая вычислительная машина производит действия как над полезным сигналом, так и над всеми гармониками, поступающими на вход вместе с полезным сигналом. Поэтому передаточная функция $W^*(s)$, соответствующая программе, имеет тот же период $j\Omega$, что и преобразование Лапласа входной функции $I^*(s)$. Периодичность $I^*(s)$ следует из уравнения (2), а периодичность $W^*(s)$ — из уравнения (7), ибо если учесть, что $\Omega T = 2\pi$, то получим

$$e^{-(s+jn\Omega)T} = e^{-jn\Omega T} e^{-sT} = e^{-sT}. \quad (10)$$

Этот результат можно было предвидеть, приняв во внимание, что сигнал на выходе цифровой машины, как и сигнал на входе, является дискретным и, следовательно, должен

также иметь преобразование Лапласа в виде периодической функции. Но если преобразования $O^*(s)$ и $I^*(s)$ являются периодическими функциями и если

$$O^*(s) = W^*(s) I^*(s), \quad (11)$$

то, как было показано раньше, $W^*(s)$ также должна быть периодической функцией.

Выводы

Прежде чем перейти к дальнейшему исследованию, полезно подвести итог тому, что было изложено. Для частного примера мы рассмотрели в частотной области схему эквивалентной цепи (см. фиг. 3). Мы показали, что для случая, когда операции являются линейными, цифровую вычислительную машину и устройство для преобразования дискретных данных в непрерывные можно охарактеризовать передаточными функциями. Кроме того, мы также показали, что действие устройства для получения дискретных данных аналогично действию модулятора. Для связи изображений сигналов на входе и выходе всей цепи служат следующие два уравнения:

$$I^*(s) = \frac{1}{T} \sum_{k=-\infty}^{\infty} I(s + jk\Omega) \quad (2)$$

и

$$O(s) = H(s) W^*(s) I^*(s), \quad (12)$$

где $H(s)$ — передаточная функция устройства для получения непрерывного сигнала из дискретного, которое является своего рода фильтром нижних частот. Мы использовали звездочку для того, чтобы отметить, что функции $W^*(s)$ и $I^*(s)$ являются периодическими с периодом, кратным $j\Omega$.

Для рассмотренного нами простого примера данное выше описание является исчерпывающим. Если цепь (или совокупность элементов), соответствующую схеме этого примера, включить в замкнутый контур, изображенный на фиг. 1, то анализ этой сложной системы можно провести методом, изложенным в работе [1]. Введение цифровой вычислительной машины не ставит никаких принципиально новых проблем, ибо цифровая машина описывается передаточной функцией точно так же, как и любой обычный элемент системы.

Однако полученная из выражения (7) передаточная функция цифровой машины, соответствующая определенной программе, не является рациональной функцией s , как это имеет место для обычных линейных систем. Тем не менее $W^*(s)$ является функцией от s и для любой частоты $s = j\omega$ имеет вполне определенную амплитуду и фазу. Следовательно, анализ такой системы не представляет затруднений.

Теперь остается обобщить результат, полученный для нашего примера, на общий случай цифровых вычислительных машин. Как уже указывалось, подробное рассмотрение такого обобщения будет сделано в следующей работе; здесь же будут приведены лишь некоторые предварительные соображения по этому вопросу.

Примерный расчет программы цифровой машины

Предположим, что разностная формула для получения первой производной, которая была рассмотрена в предыдущем примере, является недостаточно точной. Каким образом составить более точную программу цифровой машины? Конечно, в численном анализе имеется много формул для операции дифференцирования, и можно было бы, проанализировав их в частотной области, подобрать такую из них, которая удовлетворяла бы заданным требованиям точности.

Однако мы попытаемся получить приближенную формулу непосредственно в частотной области, основываясь на тех первоначальных сведениях, которые мы получили при разборе предыдущего примера. Ход рассуждений поясняется геометрическими построениями, приведенными на фиг. 6. На фиг. 6, a угол α есть фазовый угол, соответствующий функции $W^*(s)$, рассмотренной в предыдущем примере для некоторого значения частоты ω_1 . Этот угол по мере возрастания частоты уменьшается по сравнению с фазовым углом, соответствующим идеальному случаю, равным для всех частот $\pi/2$. Если бы мы могли каким-либо способом найти передаточную функцию, имеющую фазу $\pi/2 - \alpha$, то при умножении этой функции на функцию $W^*(s)$ мы получили бы передаточную функцию, имеющую необходимую фазу.

Известно, что треугольник, вписанный в окружность и имеющий одной своей стороной диаметр, является прямо-

угольным. Отсюда следует, что угол β (см. фиг. 6, а) равен $\pi/2 - \alpha$. На фиг. 6, б показана кривая

$$W_1^*(s) = k(1 + e^{+sT}), \quad (13)$$

которая имеет фазовый угол β . Таким образом, передаточная функция $W^*(s)W_1^*(s)$ будет для всех частот, представляющих интерес, иметь фазу $\pi/2$. Однако можно показать [2], что система с такой передаточной функцией является физически неосуществимой, поскольку в ней для получения выходного сигнала в данный момент времени требуется знание входного сигнала для будущих моментов времени, что для систем, работающих в реальном времени, является невозможным.

Физическая неосуществимость $W_1^*(s)$ связана с положительным показателем экспоненты e^{sT} , что соответствует случаю оператора упреждения во времени. Однако если попытаться использовать выражение

$$W_2^*(s) = k(1 + e^{-sT}), \quad (14)$$

то вместо фазового угла β мы получим угол $-\beta$. В этом легко убедиться, рассматривая фиг. 6, в. Для того чтобы теперь получить суммарную фазу, равную $\pi/2 = \alpha + \beta$, нам необходимо уже не умножить, а разделить $W^*(s)$ на $W_2^*(s)$. Это приводит к выражению

$$W_3^*(s) = \frac{W^*(s)}{W_2^*(s)} = \frac{2}{T} \frac{1 - e^{-sT}}{1 + e^{-sT}}. \quad (15)$$

В формуле (14) коэффициент k выбран равным $1/2$ для того, чтобы при $T \rightarrow 0$ передаточная функция $W_3^*(s)$ стремилась к s .

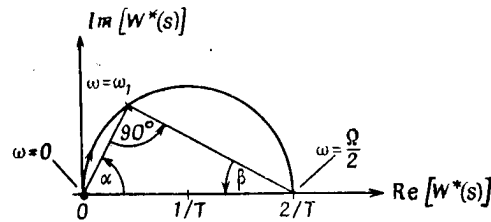
Нетрудно проверить, что $W_3^*(s)$ имеет идеальную фазу ($\pi/2$). Заменяя s через $j\omega$, мы получим выражение

$$W_3^*(j\omega) = j\omega \left[\frac{\operatorname{tg}(\omega T/2)}{\omega T/2} \right], \quad (16)$$

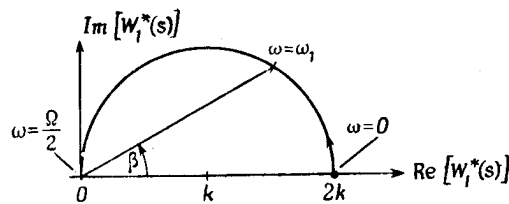
которое отличается от идеального оператора дифференцирования

$$W(j\omega) = j\omega \quad (17)$$

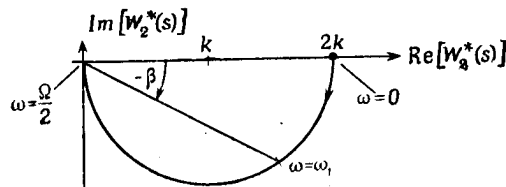
множителем в квадратных скобках. Этот множитель дает ошибку в амплитуде, но не влияет на фазу.



а) Кривая $W^*(s) = \frac{1}{T} (1 - e^{-sT})$



б) Кривая $W_1^*(s) = k(1 + e^{+sT})$



в) Кривая $W_2^*(s) = k(1 + e^{-sT})$

Фиг. 6. Последовательность операций для получения идеальной фазовой характеристики функции $W^*(s)$ для случая дифференцирования.

Однако в нашу задачу не входит исследование дифференциатора, имеющего передаточную функцию вида (15). Нас интересует вопрос о том, как можно реализовать передаточную функцию $W_3^*(s)$ в виде программы цифровой ма-

шины. По определению,

$$W_3^*(s) = O^*(s)/I^*(s),$$

откуда имеем

$$(1 + e^{-sT}) O^*(s) = \frac{2}{T}(1 - e^{-sT}) I^*(s), \quad (18)$$

или

$$O^*(s) = \frac{2}{T} [I^*(s) - I^*(s) e^{-sT}] - O^*(s) e^{-sT}. \quad (19)$$

Переходя в уравнении (19) к оригиналам, получим уравнение

$$o^*(t) = \frac{2}{T} [i^*(t) - i^*(t - T)] - o^*(t - T), \quad (20)$$

которое является линейным разностным уравнением. Это уравнение можно реализовать при помощи определенной временной программы, так как выходной сигнал в данный момент времени зависит только от значений входного и выходного сигналов в этот и в предшествующие ему моменты времени. Оказывается, что линейная программа, соответствующая разностному уравнению типа (20), всегда имеет передаточную функцию, рациональную относительно e^{-sT} . В этом состоит разница между передаточной функцией цифровой машины и передаточной функцией линейного фильтра, которая является рациональной функцией переменной s . Можно провести интересную параллель между фильтрами для дискретных сигналов, основанными на применении цифровых машин, и фильтрами для непрерывных сигналов. Но между этими двумя типами фильтров имеется и существенное различие, в связи с чем при исследовании дискретных фильтров возникает ряд новых проблем.

Один из возможных методов исследования сводится к введению новой комплексной переменной $z = e^{-sT}$ и решению задач анализа и синтеза применительно к функции переменной z . Требования, предъявляемые к передаточной функции как функции переменной s , можно выразить применительно к новой переменной z и обратно. Если дискретный фильтр не является частью системы с непрерывным сигналом, то все расчеты могут проводиться в плоскости z без перехода к прежней переменной s . Такой подход ранее

использовался в исчислении конечных разностей, как при применении операторных методов, так и при применении методов теории функций комплексного переменного.

Если дискретный фильтр является частью сложной системы с непрерывным сигналом, то необходимо установить переменные, которые являются общими как для дискретного фильтра, так и для части системы с непрерывным сигналом. Такими переменными являются время t и комплексная переменная s^1).

В настоящей работе на основании указанных выше соображений была выбрана область комплексной переменной s . Остается задача рассмотрения методов анализа и синтеза дискретных фильтров в частотной области. Оказывается, что многие результаты теории цепей и различные методы теории следящих систем могут быть использованы или распространены на случай дискретных фильтров. После того как введено понятие передаточной функции физически осуществимой программы цифровой вычислительной машины, появляется возможность использовать такие понятия, как амплитудная и фазовая характеристики, критерии устойчивости и спектр ошибки, частотные и логарифмические характеристики, импульсная переходная функция и ряд других. Эти вопросы будут рассмотрены в следующей статье.

ЛИТЕРАТУРА

1. Linvill W. K., *Trans. AIEE*, Part II, 70, 1779 (1951).
2. Salzer L. M., „Treatment of Digital Control Systems and Numerical processes in the Frequency Domain“, Диссертация, Mass. Inst. Technol., p. 148, 1951.
3. Теория следящих систем под ред. Х. Джеймса, Н. Никольса и Р. Филлипса, Издательство иностранной литературы, 1953, гл. V.

¹⁾ В действительности переменную z также можно использовать [3] в качестве основной переменной в том случае, когда в системе имеются одновременно участки с дискретным и непрерывным сигналом. Такое приближение допустимо только тогда, когда ширина спектра сигнала намного меньше половины частоты прерывания сигнала $\Omega/2$.

10. ЧАСТОТНЫЙ АНАЛИЗ ЦИФРОВЫХ ВЫЧИСЛИТЕЛЬНЫХ МАШИН, РАБОТАЮЩИХ В РЕАЛЬНОМ ВРЕМЕНИ¹⁾

Дж. М. Зальцер

Введение

В предыдущей статье [1]²⁾ были намечены методы анализа систем автоматического регулирования, в которых информация в какой-то точке преобразуется в дискретную форму. Было отмечено, что достаточно быстродействующая цифровая вычислительная машина могла бы совершать операции над дискретными данными и что для исследования этих операций вычислительной машины в случае их линейности можно применить частотный анализ. В статье был рассмотрен пример простого численного дифференцирования.

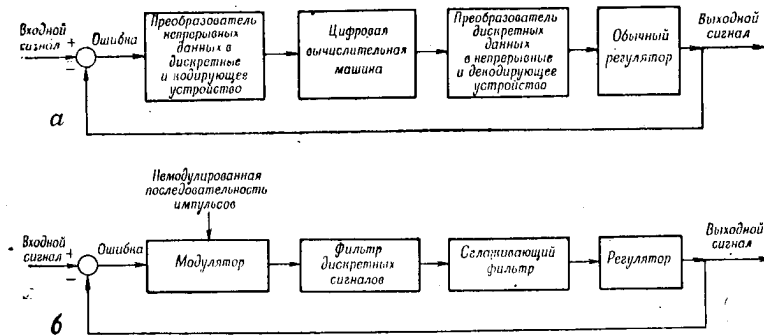
В настоящей статье делается попытка распространить частотный метод на анализ программ цифровых вычислительных машин. Если не считать докладов, зачитанных на конференциях в 1951 и 1952 гг. [2, 3], эта статья является первой опубликованной работой, обобщающей результаты исследований, проводившихся в течение нескольких лет [4]. Работа была выполнена, исходя из необходимости анализа систем автоматического регулирования, содержащих как цифровые, так и непрерывные элементы. Основная цель исследования состояла в том, чтобы показать, что линейная программа в реальном времени может быть описана разностным уравнением, преобразование которого дает передаточную функцию, характеризующую работу цифрового вычислительного устройства. Цифровую вычислительную машину с ее программой можно рассматривать как фильтр (дискретный фильтр), который можно анализировать при помощи частотного метода так же, как обычные (непрерывные) фильтры.

¹⁾ Salzer J. M., *Proc. IRE*, 42, № 2, 457—466 (1954).

²⁾ См. предыдущую статью сборника. — *Прим. ред.*

Пример

Рассмотрим снова пример, приведенный в предыдущей статье [1]. Блок-схема на фиг. 1, а представляет одну из ряда систем автоматического регулирования, управляемых цифровой машиной, или один из каналов сложной системы, или же простую систему, для которой применение специализированной цифровой машины может быть оправдано.



Фиг. 1. Эквивалентное представление простой системы автоматического регулирования, включающей цифровую вычислительную машину.

На входе и выходе цифровой машины данные должны подвергаться преобразованию. Входное устройство выполняет две функции: а) преобразует непрерывные физические величины в цифровые (например, напряжение — в двоичное число); б) преобразует непрерывную функцию в последовательность дискретных данных. Первый процесс вводит статистические ошибки главным образом из-за „квантования“ сигнала, которые могут быть проанализированы так же, как и другие шумовые эффекты. Второй процесс — превращение непрерывного сигнала в последовательность дискретных значений — влияет на динамические характеристики системы и часто является основным соображением, играющим роль при изучении замкнутых систем. В настоящей статье рассматривается в основном именно этот второй аспект систем, включающих цифровые и непрерывные элементы.

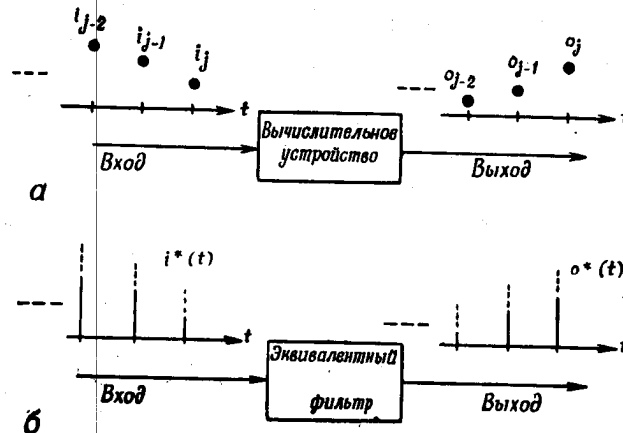
В эквивалентной схеме на фиг. 1, б мы пренебрегаем эффектом „квантования“ и дискретные значения функций

10. Частотный анализ цифровых вычислительных машин 331

заменяем импульсами, имеющими соответственные площади. Это эквивалентное представление, которое было рассмотрено в предыдущей статье [1], означает, что операция, выполняемая вычислительной машиной, может рассматриваться как операция над последовательностью импульсов, имеющих площади, равные действительным дискретным числовым значениям. Это взаимно однозначное соответствие между импульсами и дискретными данными сохраняется как в ходе вычислений, так и в конечном результате. Выбор такого представления позволяет применить частотный анализ и приспособить его к методам, уже используемым для анализа дискретных систем [5, 6].

Постановка задачи

В дальнейшем мы будем рассматривать только ту часть системы, которая соответствует цифровой вычислительной



Фиг. 2. Представление входной и выходной последовательностей сигналов.

а — дискретные данные; б — импульсы.

машине и ее программе. Введем следующие предположения которые весьма незначительно суживают область применения метода: а) данные являются дискретными и поступают с постоянной скоростью, т. е. время T между последовательными

дискретными значениями является постоянным; б) время задержки из-за процессов вычисления равно нулю или незначительно (это предположение часто оправдывается на практике; в противном случае это время может быть учтено при анализе); в) вычислительная машина устанавливает взаимно однозначное соответствие между входной и выходной величинами; г) вычислительная машина выполняет линейные операции, определенные ниже; д) вычислительная машина работает в реальном времени, используя только прошлую или настоящую информацию.

В результате этих ограничений мы имеем дело с задачей, иллюстрируемой фиг. 2,а. Входные и выходные сигналы вычислительной машины являются дискретной последовательностью чисел. Для получения выходного дискретного значения o_j для данного момента времени вычислительная машина может использовать входное дискретное значение i_j для данного момента и ограниченное число входных и выходных дискретных данных за предшествующие моменты времени. Таким образом:

$$\begin{aligned} o_j &= a_0 i_j + a_1 i_{j-1} + \dots + a_m i_{j-m} - b_1 o_{j-1} - \dots - b_n o_{j-n} = \\ &= \sum_{k=0}^m a_k i_{j-k} - \sum_{k=1}^n b_k o_{j-k}, \quad (1) \end{aligned}$$

где a_k и b_k — постоянные¹⁾.

Заменяя каждое дискретное значение импульсом с соответствующей площадью и обозначая последовательности импульсов соответственно через $i^*(t)$ и $o^*(t)$, мы можем записать уравнение (1) в виде

$$o^*(t) = \sum_{k=0}^m a_k i^*(t - kT) - \sum_{k=1}^n b_k o^*(t - kT). \quad (2)$$

Математические операции вида (1) или (2) вполне подходят для цифровой вычислительной машины, ее запоминающее устройство позволяет вводить произвольные идеальные задержки и запоминать постоянные, а ее арифметическое устройство — выполнять действия умножения и сложения.

¹⁾ Линейный анализ может быть распространен и на случай, когда a_k и b_k являются функциями времени, но этот вопрос выходит за рамки данной статьи.

Передаточная функция и программа

Передаточная функция фильтра может быть определена как отношение преобразований Лапласа выхода и входа системы. Если обозначить через $I^*(s)$ и $O^*(s)$ преобразования величин $i^*(t)$ и $o^*(t)$ соответственно, то из (2) получаем

$$O^*(s) = \sum_{k=0}^m a_k e^{-kTs} I^*(s) - \sum_{k=1}^n b_k e^{-kTs} O^*(s). \quad (3)$$

Отсюда получаем передаточную функцию

$$W^*(s) = \frac{O^*(s)}{I^*(s)} = \frac{\sum_{k=0}^m a_k e^{-kTs}}{1 + \sum_{k=1}^n b_k e^{-kTs}} = \frac{a_0 + a_1 e^{-Ts} + \dots + a_m e^{-mTs}}{1 + b_1 e^{-Ts} + \dots + b_n e^{-nTs}}. \quad (4)$$

При указанных ограничениях уравнение (1) соответствовало наиболее общей возможной программе, и это привело нас к передаточной функции вида (4). Кроме того, переходя к оригиналам, можно из (4) снова получить (1). Отсюда следует, что линейная программа в реальном времени может быть реализована в том и только в том случае, если ее передаточная функция может быть представлена в форме (4) (теорема реализуемости).

В то время как передаточная функция моделирующего (непрерывного) фильтра является рациональной функцией от s , для цифрового (дискретного) фильтра она рациональна относительно e^{-Ts} . Распространение обычного частотного метода на дискретные фильтры имеет большое значение при исследовании смешанных систем.

Ограничение полосы пропускания

Если непрерывный сигнал преобразуется в дискретную форму, то информация о поступающих данных, вообще говоря, теряется, так как через дискретные точки можно провести больше одной непрерывной кривой. Однако можно установить однозначное соответствие между непрерывной кривой и ее дискретной формой, если полоса пропускания формирующего устройства ограничена половиной частоты

повторения дискретных сигналов, или $1/2 T$ гц (теорема о дискретном представлении во времени) [7,8]¹⁾.

Действительно, при скорости прерывания, превышающей этот предел, можно восстановить сигнал только за счет значительной временной задержки. В разомкнутой системе, например в системе звукопередачи, допустимы временные задержки на несколько периодов наибольшей частоты сигнала; однако в замкнутой системе задержка даже в пределах части периода (10 или 20°) может оказаться критической. В системах автоматического регулирования, вероятно, потребуется, чтобы частота прерывания по крайней мере в 5 раз превышала наибольшую частоту сигнала. Несмотря на то, что теорема о дискретном представлении устанавливает верхний предел, она не должна рассматриваться как рабочее правило для дискретных систем автоматического регулирования.

Ограничение полосы пропускания импульсов зависит от вычислительной машины в целом, так как ее выходной сигнал имеет такую же дискретную форму, как и сигнал на входе. Чтобы показать это, отметим два свойства функции $W^*(s)$ — сопряженную симметрию и периодичность.

Так как каждый член функции $W^*(s)$ вида $a_k e^{-kTs}$ обладает сопряженной симметрией, т. е.

$$e^{-kT(\sigma \pm j\omega)} = e^{-kT\sigma} (\cos kT\omega \mp j \sin kT\omega), \quad (5)$$

то это свойство будет присуще и всей функции. Отсюда следует, что если $W^*(s)$ известна для положительных действительных значений частоты, то она известна и для отрицательных.

Кроме того, e^{-kTs} периодична в частотной области с интервалом повторения $j\Omega = j2\pi/T$. Действительно,

$$e^{-kT(s \pm jm\Omega)} = e^{-jkmT\Omega} e^{-kTs} = e^{-jkm2\pi} e^{-kTs} = e^{-kTs}. \quad (6)$$

Следовательно, $W^*(s)$ имеет ту же периодичность; она полностью определена, если известна в полосе плоскости s , ограниченной двумя прямыми $\pm j\Omega/2$. Однако по соображениям симметрии она *должна быть известна только в полосе*

¹⁾ Теорема сформулирована В. А. Котельниковым в 1933 г. (Котельников В. А., Материалы к 1-му Всесоюзному съезду по вопросам реконструкции связи, 1933). — *Прим. ред.*

10. Частотный анализ цифровых вычислительных машин 335

от 0 до $+j\Omega/2$, соответствующей тому же пределу полосы пропускания, который устанавливается теоремой о дискретном представлении. Например, программа в строгом смысле слова не может иметь характеристики фильтра нижних частот. Программа сглаживания имеет характеристики фильтра низких частот в пределах от 0 до $\Omega/2$ рад/сек, т. е. она имеет полосу пропускания, расположенную вблизи нулевой частоты. Следовательно, она будет иметь полосы пропускания также и при частотах, кратных Ω . Эти более высокие частоты могут быть срезаны непрерывными фильтрами.

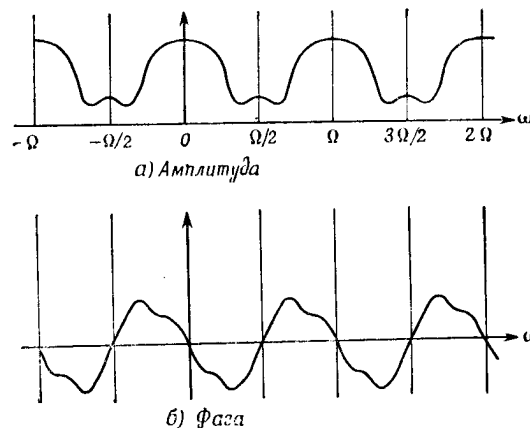
При использовании в системе цифровой вычислительной машины наибольшая возможная эффективная¹⁾ скорость формирования дискретного сигнала равна скорости, с которой машина может производить требуемые вычислительные операции. Таким образом, можно сказать, что более скоростная программа вычислительной машины соответствует большей полосе пропускания. Скорость программы зависит от нескольких факторов: скорости отдельного акта счета (скорости повторения импульсов), логической схемы, кода инструкций, способа вычислений и эффективности кодирования. Вычислительная работа часто может быть упрощена составлением такой программы, которая обеспечивает лишь минимально допустимую точность. Баланс между статистическими (округление) и динамическими (отсекание) ошибками [9,10] заслуживает такого же внимания, как и при проектировании непрерывных блоков. Пригодность вычислительного устройства для решения конкретной задачи регулирования определяется указанными выше соображениями.

Амплитуда, фаза и амплитудно-фазовая характеристика

Если мы в (4) заменим s на $j\omega$, то получим передаточную функцию $W^*(j\omega)$ программы в функции действительных частот. В нашем случае $W^*(j\omega)$ имеет тот же физический смысл, как и в непрерывном варианте. Синусоидальная входная последовательность сигнала с угловой частотой ω_1 ($\omega_1 < \Omega/2$)

¹⁾ Прибор, преобразующий входной сигнал, может иметь свою собственную скорость прерывания сигналов, которая обычно превышает скорость вычислений. Эффективная скорость формирования дискретных сигналов совпадает со скоростью, с которой программа использует входную информацию.

дает на выходе системы, управляемой программой $W^*(j\omega)$, синусоидальную последовательность той же самой частоты, но с амплитудой, увеличенной в $|W^*(j\omega)|$ раз, и фазой, сдвинутой на угол $\arg W^*(j\omega)$. Доказательство получается автоматически из уравнения (4). Отметим лишь, что входные и выходные функции являются дискретными (импульсными) и содержат в дополнение к частоте $\pm\omega_1$ также частоты $\pm\omega \pm \Omega$, $\pm\omega_1 \pm 2\Omega$ и т. д. той же интенсивности.



Фиг. 3. Возможные частотные характеристики программы.

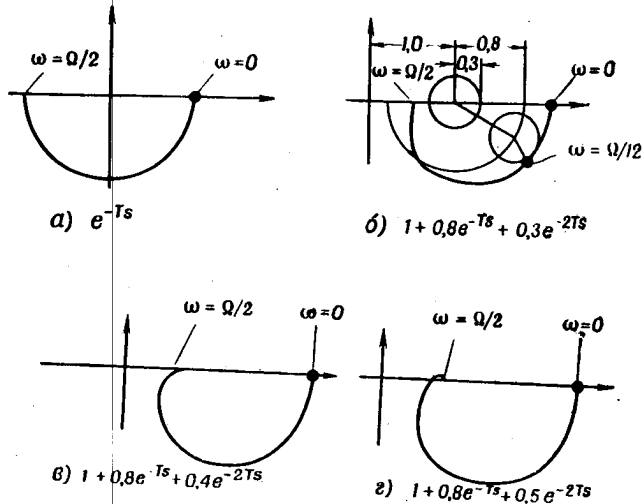
На фиг. 3 показаны возможные характеристики программы. Амплитуда является четной, а фаза — нечетной функцией частоты вследствие сопряженной симметрии, и обе кривые периодичны с периодом Ω рад/сек.

Амплитудные и фазовые характеристики могут быть объединены в одну полярную кривую (годограф), которую называют амплитудно-фазовой характеристикой. Она может быть довольно быстро построена, как это показано на фиг. 4. На фиг. 4,а изображена амплитудно-фазовая характеристика функции e^{-Ts} , у которой амплитуда равна 1, а фаза — $T\omega$. На графике изображен только участок кривой для частот $0 \leq \omega \leq \Omega/2$; распространив его на область $\Omega/2 \leq \omega \leq \Omega$, мы получим замкнутую окружность, которая в дальнейшем будет повторяться. Амплитудно-фазовые характеристики,

10. Частотный анализ цифровых вычислительных машин 337

изображенные на фиг. 4, соответствуют многочленам второй степени от e^{-Ts} . Пример построения, основанного на принципе вращения векторов, показан на фиг. 4, б для частоты $\omega = \Omega/12$ ($\omega T = 30^\circ$).

Для построения амплитудно-фазовой характеристики удобно пользоваться следующими правилами, приводимыми здесь без доказательства¹⁾:



Фиг. 4. Простые примеры амплитудно-фазовых характеристик программы.

- 1) амплитудно-фазовая характеристика повторяется для каждого интервала Ω ;
- 2) она обладает сопряженной симметрией относительно действительной оси;
- 3) при $\omega = 0$ и $\omega = \Omega/2$ она пересекает действительную ось под прямым углом (фиг. 4, а, б и з) или касается ее (фиг. 4, в);
- 4) согласно (4), амплитудно-фазовая характеристика может пересекать действительную ось $(m - 1)$ или $(n - 1)$ раз (в зависимости от того, какое из этих чисел больше) в промежутке $0 < \omega < \Omega/2$ (исключая его концы).

¹⁾ Доказательство см. [3], стр. 75, 76 и 265—267.

Программы упреждения

Для иллюстрации изложенного метода рассмотрим классические формулы для вычисления упреждения. Предположим, что цифровая вычислительная машина применяется для предсказания ближайшей величины входного дискретного сигнала; в идеальном случае задача состоит в решении разностного уравнения

$$o^*(t) = i^*(t + T). \quad (7)$$

Преобразование Лапласа этого уравнения дает передаточную функцию программы

$$W^*(s) = \frac{O^*(s)}{I^*(s)} = e^{Ts}, \quad (8)$$

которая не приводится к форме (4).

Классические формулы экстраполяции, основанные на аппроксимации многочленами, дают приближенное предсказание искомой величины. Преобразования Лапласа для четырех простейших выражений будут иметь вид

$$\begin{aligned} W_0^*(s) &= 1, \\ W_1^*(s) &= 2 - e^{-Ts}, \\ W_2^*(s) &= 3 - 3e^{-Ts} + e^{s-2T}, \\ W_3^*(s) &= 4 - 6e^{-Ts} + 4e^{-2Ts} - e^{-3Ts}. \end{aligned} \quad (9)$$

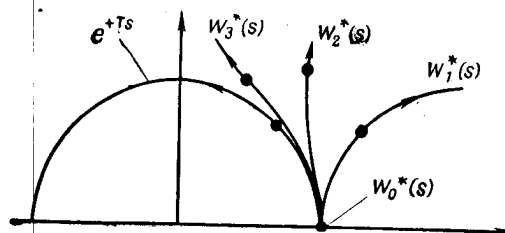
Для интерпретации этих выражений заменим $W_m^*(s)$ через $O_m^*(s)/I_m^*(s)$, умножим на $I_m^*(s)$ и выполним обратное преобразование Лапласа; в результате мы получим уравнения

$$\begin{aligned} o_0^*(t) &= i^*(t), \\ o_1^*(t) &= 2i^*(t) - i^*(t - T), \\ o_2^*(t) &= 3i^*(t) - 3i^*(t - T) + i^*(t - 2T), \\ o_3^*(t) &= 4i^*(t) - 6i^*(t - T) + 4i^*(t - 2T) - i^*(t - 3T). \end{aligned} \quad (10)$$

В каждом из этих уравнений выходные дискретные данные определяются только через входные, т. е. предыдущее упреждение не принимается во внимание при определении последующего. Это является результатом того, что знаменатель всех $W_m^*(s)$ в (9) равен 1.

10. Частотный анализ цифровых вычислительных машин 339

Амплитудно-фазовые характеристики, соответствующие формулам (9), показаны на фиг. 5. Легко видеть, что аппроксимация улучшается с возрастанием порядка формулы. Фактически характеристики можно сравнивать только при низких частотах; при $\omega = 0$ все они являются идеальными. При возрастании ω кривые отклоняются от идеальной. Можно заметить, что при всех значениях частоты амплитудно-фазовые характеристики остаются вне единичного круга,



Фиг. 5. Сравнение амплитудно-фазовых характеристик экстраполяционных формул с идеальной характеристикой.

соответствующего идеальной характеристике; это указывает на то, что характеристики могут быть улучшены для спектров определенных сигналов¹⁾.

Чтобы сравнить эти формулы при высоких частотах, рассмотрим случай, когда $\omega = \Omega/2$ рад/сек. Легко показать, что $W_m^*(j\Omega/2) = 2^{m+1} - 1$, т. е. при этой частоте усиление быстро растет с ростом порядка m упреждения. Поэтому при применении более сложной формулы упреждения подчеркиваются шумы, включая шумы квантования или округления.

Как и в непрерывном случае, решение задачи сводится к компромиссу между хорошим упреждением и хорошим сглаживанием.

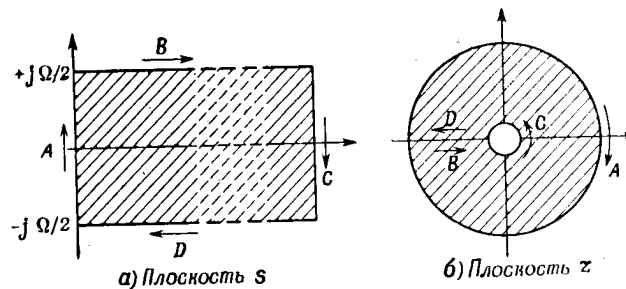
¹⁾ Можно показать, что эти формулы являются оптимальными в смысле среднеквадратичной ошибки *только* для сигнала с шириной полосы, стремящейся к нулю на нулевой частоте, что характеризует тривиальный случай.

**Комплексная переменная задержки.
Неограниченные весовые последовательности**

Для дальнейшего целесообразно ввести новую переменную z , определяемую соотношением

$$z = e^{-Ts}. \quad (11)$$

Такая замена не сводится только к изменению обозначений. Во-первых, z — комплексная функция комплексной переменной s и применение конформного отображения плоскостей обеих переменных позволит глубже изучить программы цифровых



Фиг. 6. Конформное отображение правой полуплоскости.

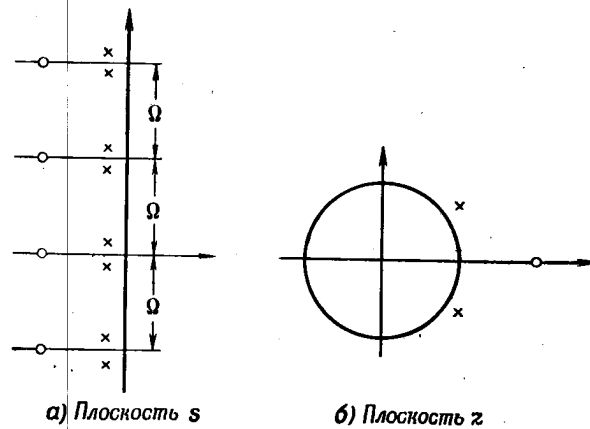
вычислительных машин и дискретных функций. Во-вторых, переменной z можно приписать определенный физический смысл.

На фиг. 4, *a* показано, что мнимая ось плоскости s является конформным отображением единичной окружности плоскости z , а правая половина плоскости s соответствует площади внутри этой окружности. Отметим, что на плоскости z окружность обходится один раз для каждого интервала $j\Omega$, который проходится на плоскости s . Чтобы найти конформное отображение *ограниченной* части правой полуплоскости s , достаточно отобразить контур полубесконечной полосы плоскости s (фиг. 6, *a*). Полученное отображение на фиг. 6, *б* исключает точку $z = 0$ на z -плоскости.

В качестве второго примера конформного отображения рассмотрим возможное расположение полюсов и нулей цифрового фильтра (или дискретной функции). Картина нулей и полюсов в плоскости s повторяется с интервалом повторения $j\Omega$ (фиг. 7, *a*), но имеет единственное конформное

10. Частотный анализ цифровых вычислительных машин 341

отображение в плоскости z (фиг. 7, б). Этим и объясняется, почему использование плоскости z для дискретных функций и фильтров часто является более предпочтительным. Для смешанных систем, описываемых рациональными функциями как от s , так и от z , только переменная s имеет определенный смысл. В таких случаях также использовали переменную z [11,6], но законность такой аппроксимации должна быть обоснована.



Фиг. 7. Конформное отображение части плоскости, содержащей нули и полюсы.

Переменную z можно рассматривать также как оператор, так как умножение на z соответствует задержке на T во временной области¹⁾. Этот разностный оператор аналогичен дифференциальному оператору p или D . Если рассматривать эти операторы как комплексные переменные, то появляется возможность широких обобщений.

Передаточная функция (4) программы в z -представлении будет иметь вид

$$W^*(z) = \frac{O^*(z)}{I^*(z)} = \frac{a_0 + a_1z + a_2z^2 + \dots + a_mz^m}{1 + b_1z + b_2z^2 + \dots + b_nz^n}, \quad (12)$$

¹⁾ В литературе такой оператор (E , p или z) обычно определяется как оператор опережения во времени и соответствует нашему $1/z$. В реальном времени опережение является нереализуемой операцией. Однако такой выбор ведет к упрощениям в математическом анализе.

откуда

$$(1 + b_1 z + \dots + b_n z^n) O^* = (a_0 + a_1 z + \dots + a_m z^m) I^*(z). \quad (13)$$

Числитель выражения (12) можно рассматривать как весовую последовательность, относящуюся к входному сигналу, а знаменатель относится к выходным дискретным данным.

Конечную весовую последовательность, относящуюся к предыдущим выходным сигналам, всегда можно представить в виде бесконечной весовой последовательности, относящейся только к входным дискретным данным; например, если

$$W^*(z) = \frac{1}{1 - bz}, \quad (14)$$

то

$$O^*(z) = I^*(z) + bzO^*(z), \quad (15)$$

или

$$O^*(z) = I^*(z) + bzI^*(z) + b^2 z^2 O^*(z). \quad (16)$$

Продолжая этот процесс, получим

$$O^*(z) = (1 + bz + b^2 z^2 + \dots) I^*(z), \quad (17)$$

что дает

$$W^*(z) = \frac{O^*(z)}{I^*(z)} = 1 + bz + b^2 z^2 + \dots \quad (18)$$

Это выражение можно прямо получить разложением в ряд из (14).

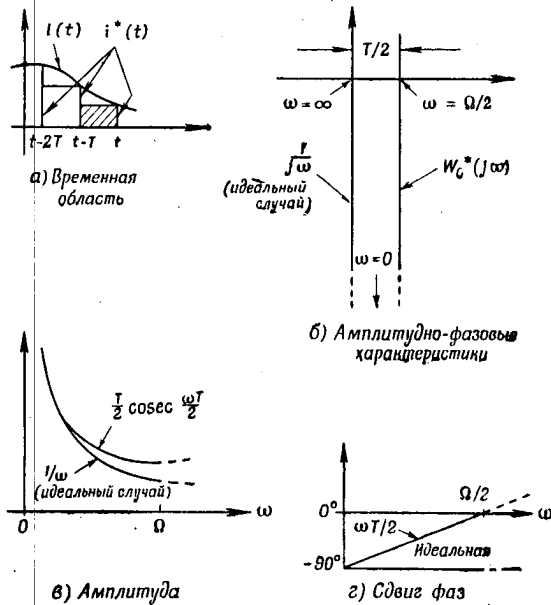
Левинсон (см. [12]) рассмотрел проблему оптимизации дискретных фильтров, используя неограниченную весовую последовательность, относящуюся только к входным дискретным данным. В практическом случае такая последовательность является ограниченной и соответствует передаточной функции в форме многочлена, представляющей собой частный случай рациональной функции (12).

Интегрирование

Если передаточная функция (12) имеет знаменатель, то ее можно путем деления представить в виде неограниченной весовой последовательности. Таким образом, знаменатель указывает на неограниченную (чего обычно нельзя достичь) память системы. Формулы интегрирования являются хорошим примером потому, что память идеального интегратора неограничена,

10. Частотный анализ цифровых вычислительных машин 343

Так как передаточная функция идеального интегратора равна $1/s$, что, очевидно, не соответствует форме (4), то программа только аппроксимирует эту величину. Интеграл от непрерывной кривой, т. е. площадь под ней, может быть аппроксимирован прямоугольниками (фиг. 8, а). Площадь



Фиг. 8. Характеристики при аппроксимации прямоугольниками.

каждого нового прямоугольника добавляется к уже накопленной площади. Таким образом, интеграл $o^*(t)$ вычисляется с помощью выражения

$$o^*(t) = T i^*(t) + o^*(t - T), \quad (19)$$

которое определяет передаточную функцию программы вида

$$W_0^*(z) = \frac{O^*(z)}{I^*(z)} = \frac{T}{1-z}. \quad (20)$$

Как и предполагалось, знаменатель $W_0^*(z)$ отличен от 1,

Подставляя $z = e^{-Ts}$ и заменяя $s = j\omega$, получаем

$$W_0^*(j\omega) = \frac{1}{1 + e^{-jT\omega}} = \frac{T}{2} \left(1 - j \operatorname{ctg} \frac{T\omega}{2} \right). \quad (21)$$

Полученная отсюда амплитудно-фазовая характеристика сравнивается на фиг. 8, б с идеальной. На фиг. 8, в и г сравниваются соответствующие амплитудные и фазовые характеристики. Если спектр входного сигнала известен, то можно определить среднеквадратичную ошибку интегрирования.

Несколько бóльшую ошибку, как показывает фиг. 9, а, дает так называемая интерполяционная формула трапеций. Этому случаю соответствует разностное уравнение

$$o^*(t) = (T/2) [i^*(t) + i^*(t-T)] + o^*(t-T) \quad (22)$$

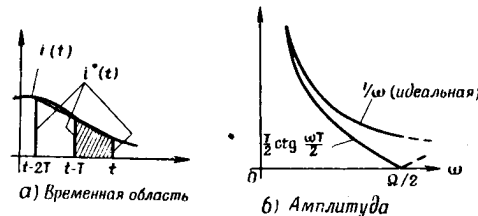
и передаточная функция

$$W_1^*(z) = \frac{T}{2} \frac{1+z}{1-z}, \quad (23)$$

откуда

$$W_1^*(j\omega) = \frac{T}{2} \frac{1 + e^{-jT\omega}}{1 - e^{-jT\omega}} = \frac{T}{2} \left(-j \operatorname{ctg} \frac{T\omega}{2} \right). \quad (24)$$

Сравнивая выражения (21) и (24), получаем, что последняя амплитудно-фазовая характеристика сдвинута влево на вели-



Фиг. 9. Характеристики при аппроксимации трапециями.

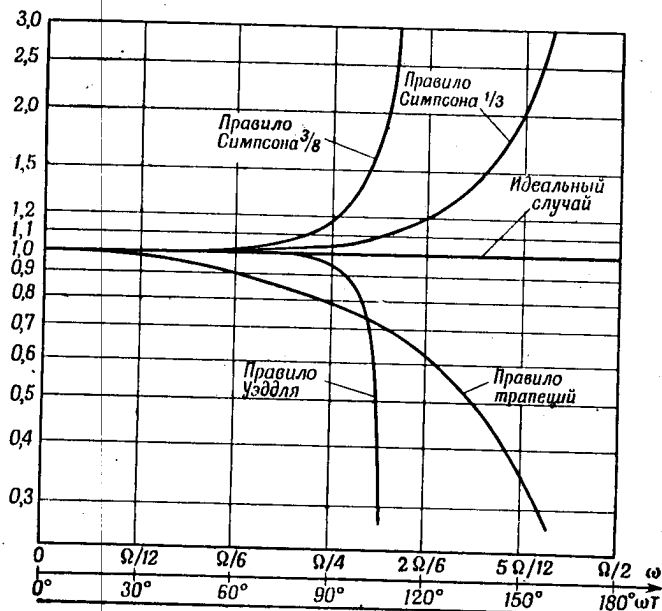
чину $T/2$ и поэтому совпадает с характеристикой идеального интегратора. Следовательно, метод трапеций дает такую же фазу, как и в идеальном случае, но имеет амплитудные характеристики, отличающиеся от идеальных. В этом его недостаток (фиг. 9, б).

Таким образом, оказывается, что формула интегрирования, дающая грубое приближение во временной области,

10. Частотный анализ цифровых вычислительных машин 345

имеет идеальную фазу. Другие классические правила интегрирования также имеют идеальные фазовые характеристики в обычных диапазонах частот¹⁾.

С учетом этого полное сравнение различных методов можно сделать по амплитудным характеристикам, показанным на фиг. 10. Фиг. 10 подтверждает некоторые известные



Фиг. 10. Сравнение формул приближенного интегрирования с идеальным случаем по амплитудным характеристикам.

свойства этих методов: превосходство правила „одной трети“ Симпсона над его же более сложным правилом „трех восьмых“, высокую точность правила Уэддл (до определенной частоты). Из фиг. 10 можно получить и другие результаты, которые обычно несущественны при численных расчетах, но важны при синтезе систем. Например, правило трапеций дает по сравнению со всеми другими методами

¹⁾ См. [3], стр. 233—234.

самые неточные результаты при низких частотах, но при высоких частотах оно уступает только правилу „одной третьей“ Симпсона. Однако фактически его можно предпочесть правилу Симпсона не только в силу его простоты, но и потому, что его амплитудная характеристика затухает на высоких частотах, нейтрализуя влияние шума.

Рассмотрим анализ ошибок на примере. Пусть синусоидальное колебание с частотой ω_0 рад/сек интегрируется с использованием шести дискретных значений в каждом периоде. Определим ошибку в определении амплитуды. При $\omega_0 = \Omega/6$ или $\omega T = 60^\circ$ (см. фиг. 10) ошибки таковы: для правила трапеций — 9,3%, для правила „одной третьей“ Симпсона +0,77%, для правила „трех восьмых“ Симпсона +1,12%, для правила Уэддля — 0,25%.

Ниже для справок приведены различные передаточные функции, соответствующие рассмотренным правилам.

Правило „одной третьей“ Симпсона:

$$W_2^*(z) = \frac{T}{3} \frac{1 + 4z + z^2}{1 - z^2}. \quad (25)$$

Правило „трех восьмых“ Симпсона:

$$W_3^*(z) = \frac{3T}{8} \frac{1 + 3z + 3z^2 + z^3}{1 - z^3}. \quad (26)$$

Правило Уэддля:

$$W_4^*(z) = \frac{3T}{10} \frac{1 + 5z + z^2 + 6z^3 + z^4 + 5z^5 + z^6}{1 - z^6} = \frac{3T}{10} \frac{1 + 6z + 6z^2 + 6z^3 + z^4}{1 + z - z^3 - z^4}. \quad (27)$$

Устойчивость

Следует иметь в виду, что цифровую вычислительную машину с определенной программой следует рассматривать как активный, а не как пассивный фильтр. Вопрос об устойчивости или сходимости всегда важен при численном анализе, а в приложении к системам регулирования он становится решающим.

Устойчивость программы может изучаться с использованием переменных s или z . В плоскости переменной s устойчивость определяется по отсутствию полюсов в правой полу-

10. Частотный анализ цифровых вычислительных машин 347

плоскости и кратных полюсов на мнимой оси. Числитель $W^*(s)$ ограничен при конечном s , так как он является целой функцией. Для доказательства отметим, что величина k -го члена

$$|a_k e^{-kTs}| = |a_k e^{-kT(\sigma + j\omega)}| = |a_k e^{-kT\sigma}| \quad (28)$$

ограничена для всех ограниченных значений σ . Поэтому необходимо исследовать только нули знаменателя.

Знаменатель $W^*(s)$ в (4) равен

$$Q^*(s) = 1 + b_1 e^{-Ts} + \dots + b_n e^{-nTs}. \quad (29)$$

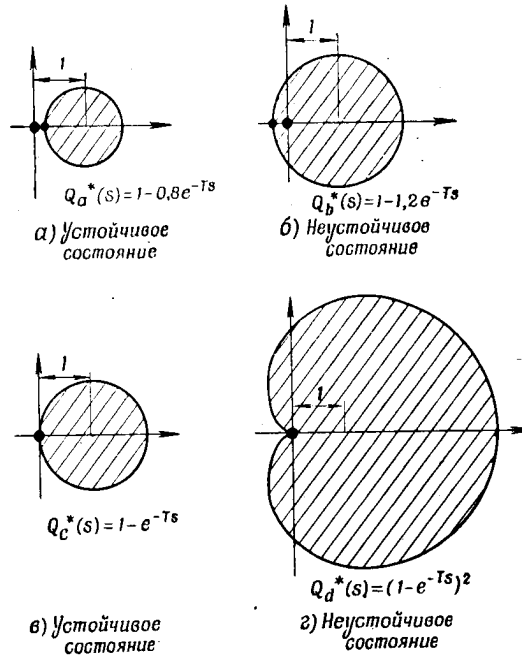
Если функция $Q^*(s)$ имеет нуль в правой полуплоскости, то она имеет нуль также и в заштрихованной полосе на фиг. 6, а, так как $Q^*(s)$ является периодической функцией. Если такой нуль имеется, то контур в плоскости Q^* , конформно отображающий контур $ABCD$ на фиг. 6, а, охватывает начало координат. Однако достаточно нанести только контур A (амплитудно-фазовая характеристика, соответствующая $Q^*(s)$), так как B и D всегда совпадают на плоскости Q^* (вследствие периодичности), а C дает бесконечно малую окружность около 1 (не затрагивая окрестности начала координат). Следовательно, если числитель и знаменатель $W^*(s)$ не имеют общих нулей, то для установления устойчивости реализуемой программы достаточно определить, охватывает ли амплитудно-фазовая характеристика, соответствующая знаменателю, начало координат.

Очень часто достаточно грубого наброска кривой Q^* , чтобы определить устойчивость. Об относительной устойчивости обычно принято судить по величине отклонения кривой от начала координат. Частота, соответствующая ближайшей к началу координат точке кривой, часто близка к резонансной частоте программы. На фиг. 11 приведено несколько простых примеров кривых Q^* и даны заключения об устойчивости.

Устойчивость может определяться также на плоскости переменной z . И в этом случае для исследования устойчивости достаточно рассмотреть только расположение нулей знаменателя (передаточной функции $Q^*(z)$). Так как правая полуплоскость переменной s отображается в круг единичного радиуса на плоскости z , то для устойчивости программы

корни уравнения $Q^*(z) = 0$ должны быть простыми и по абсолютной величине не меньше единицы.

Таким образом, для анализа устойчивости необходимо найти корни уравнения $Q^*(z) = 0$. Хотя при помощи билинейного преобразования [13] и применения критериев Рауса



Фиг. 11. Применение конформного отображения к изучению устойчивости.

или Гурвица можно определить абсолютную устойчивость, этот метод не позволяет находить относительную устойчивость.

Полезно также рассмотреть устойчивость во временной области. Если $W^*(z) = 1/Q^*(z)$, то для случаев фиг. 11, а, б и в мы получаем

$$W^*(z) = \frac{O^*(z)}{I^*(z)} = \frac{1}{1 - bz}. \quad (30)$$

Соответствующее разностное уравнение имеет вид

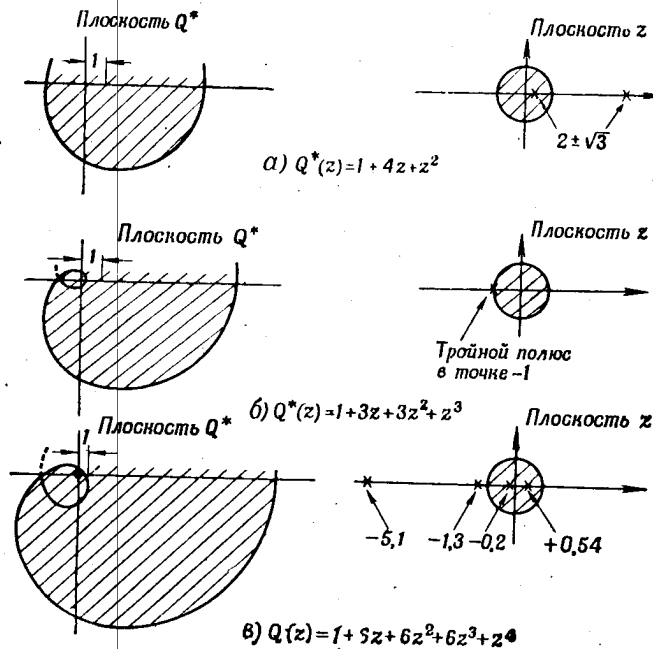
$$o^*(t) = i^*(t) + bo^*(t - T). \quad (31)$$

10. Частотный анализ цифровых вычислительных машин 349

Если рассматривать $t^*(t)$ как единичный достаточно малый импульс δ , действующий в момент прерывания сигнала, то $o^*(t)$ также будет равно δ ; второй выходной импульс уже будет равен $b\delta$, следующий за ним $b^2\delta$ и т. д. Очевидно, что если $b > 1$, то выходные импульсы постепенно возрастают, что соответствует неустойчивости. Так как корнем $Q(z) = 0$ является $1/b$, то указанный выше критерий устойчивости подтверждается.

Дифференцирование

В качестве иллюстрации определения условий устойчивости можно рассмотреть численное дифференцирование. Так



Фиг. 12. Доказательство неустойчивости „точных“ программ дифференцирования.

как передаточная функция идеального дифференциатора равна s , т. е. является обратной величиной передаточной функции идеального интегратора, то очевидно, что выражения,

обратные формулам интегрирования (20), (23), (25)—(27), должны соответствовать программам дифференцирования.

На основании предыдущей статьи [1] легко показать, что только две первые из пяти обратных программ будут устойчивы, так как первая из них имеет постоянный знаменатель, а вторая — простой полюс при $z = -1$.

Проверка устойчивости программ остальных трех дифференциаторов проведена методом конформного отображения и изучения расположения полюсов на плоскости z (фиг. 12). Мы рассмотрим только первый из них, изображенный на фиг. 12, *a*. Обычно нет надобности вычерчивать всю амплитудно-фазовую характеристику, так как легко вычислить точки ее пересечения с действительной осью. При частотах $\omega = 0$ и $\Omega/2$ $z = \pm 1$ и знаменатель $1 + 4z + z^2$ равен в этих точках соответственно $+6$ и -2 ; противоположные знаки указывают, что начало координат охватывается амплитудно-фазовой характеристикой. Рассматривая контур на плоскости z , мы увидим, что один из двух полюсов действительно нарушает условие устойчивости.

Программирование

При рассмотрении теоремы реализуемости в разделе „Передачная функция и программа“ мы уже иллюстрировали метод осуществления программы, заданной своей передачной функцией. В этом разделе мы рассмотрим другие методы реализации программ. Как и при расчете или синтезе систем, нам необходимо найти показатели, на основе которых можно оценить тот или иной метод. При реализации программ такими показателями являются число запоминающих регистров и требуемое время для вычислений. Но даже если два метода эквивалентны по памяти и времени, они могут отличаться в других важных деталях, таких, как накопление ошибок при округлении чисел. Определенное значение имеет установление связи между отдельными параметрами (постоянными) программы и такими показателями, как затухание, резонансная частота, расположение полюсов и нулей. Частично реализация программы включает в себя расчет и конструирование самого вычислительного прибора. Отсюда вытекают три метода реализации программ, которые мы рассмотрим и сравним.

10. Частотный анализ цифровых вычислительных машин 351

Прямое программирование

Исходя из выражения передаточной функции (4) и используя уравнение (3), можно перейти к уравнениям (2) и (1), являющимся двумя формами одного и того же разностного уравнения. Программа для решения уравнения (2) должна разрешить две различные задачи: арифметическую задачу и задачу „переадресовки“. Арифметическая задача заключается в воспроизведении действий умножения и сложения, указанных в уравнении (2). Так как это уравнение имеет $m + n + 1$ членов в правой части, необходимо произвести $m + n + 1$ умножений и $m + n$ сложений. Эти вычисления повторяются в каждом интервале между импульсами, но над новыми данными.

Второй задачей программы является приведение данных к настоящему моменту времени. Импульс, равный $i^*(t)$ в данный момент, становится равным $i^*(t - T)$ в последующий момент, и, следовательно, импульс $i^*(t)$ должен быть передан в ячейку памяти, соответствующую $i^*(t - T)$. Подобными же методами производится пересчет во времени и для других входных и выходных данных.

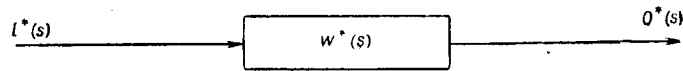
Большинство обычных вычислительных машин в состоянии осуществить операции, требующиеся для решения этих задач. Табличка, характеризующая требуемую емкость памяти прибора, является полезным инструментом, поскольку требуемое количество инструкций дает также указание о времени, необходимом для вычислений. Для рассматриваемого метода программирования подобная табличка приведена ниже для случая, когда все постоянные в уравнении (4) различны и отличны от нуля.

ЕМКОСТЬ ЗАПОМИНАЮЩИХ УСТРОЙСТВ ПРИ ПРЯМОМ ПРОГРАММИРОВАНИИ (УСЛОВНО)	
Количество постоянных	$m + n + 1$
Количество данных	$m + n$
Операции {	умножения $m + n + 1$
	сложения $m + n$
	передачи $m + n$
Полное количество операций	$3m + 3n + 1$
	<hr/>
Всего	$5m + 5n + 2$

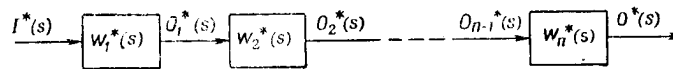
Отметим, что эта табличка является условной и не обязательно применима к конкретной вычислительной машине, так как число инструкций существенно зависит от логического и инструкционного кодов¹⁾.

Последовательное программирование

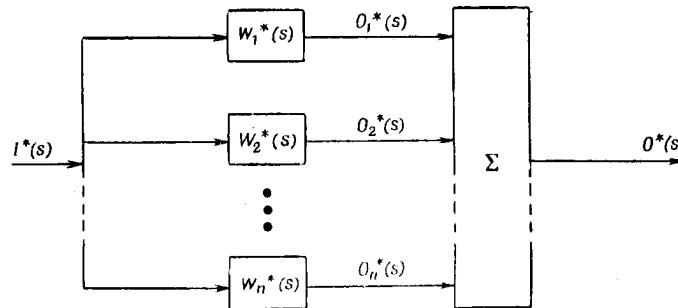
Воспользовавшись методами синтеза цепей, можно предложить другие средства реализации программ, исходя из



а) Прямое программирование



б) Последовательное программирование



в) Параллельное программирование

Фиг. 13. Сравнение методов реализации программ.

выражения для передаточной функции вида (4) или (12). Предположим, что числитель и знаменатель передаточной

¹⁾ Детали этой программы были разработаны в статье [3] стр. 273—292.

10. Частотный анализ цифровых вычислительных машин 353

функции имеют только действительные и простые нули и что $m \leq n - 1$. Эти ограничения, практически не уменьшающие общности наших рассуждений, позволяют нам представить уравнение (12) в виде

$$W^*(z) = \frac{a_0}{1 + d_1 z} \frac{1 + c_2 z}{1 + d_2 z} \cdots \frac{1 + c_n z}{1 + d_n z}, \quad (32)$$

где $-1/c_k$ и $-1/d_k$ дают нули соответственно числителя и знаменателя (некоторые из постоянных c могут равняться нулю).

Каждый множитель в (32), рассматриваемый отдельно, является передаточной функцией реализуемой программы, так что первоначальная программа (фиг. 13, а) заменяется последовательным рядом более простых программ (фиг. 13, б). Таким образом:

$$\frac{O^*(z)}{I^*(z)} = \frac{O_1^*(z)}{I^*(z)} \frac{O_2^*(z)}{O_1^*(z)} \cdots \frac{O_n^*(z)}{O_{n-1}^*(z)}. \quad (33)$$

Если приравнять соответствующие члены уравнений (32) и (33), произвести умножение и применить обратное преобразование Лапласа, мы получим следующую систему разностных уравнений

$$\left. \begin{aligned} o_1^*(t) &= a_0 i^*(t) && -d_1 o_1^*(t-T) \\ o_2^*(t) &= o_1^*(t) + c_2 o_1^*(t-T) && -d_2 o_2^*(t-T) \\ &\vdots && \vdots \\ o_n^*(t) &= o_{n-1}^*(t) + c_n o_{n-1}^*(t-T) && -d_n o_n^*(t-T) \end{aligned} \right\} \quad (34)$$

Простой подсчет членов этого уравнения дает следующую табличку:

ЕМКОСТЬ ЗАПОМИНАЮЩИХ УСТРОЙСТВ ПРИ ПОСЛЕДОВАТЕЛЬНОМ ПРОГРАММИРОВАНИИ (УСЛОВНО)	
Количество постоянных	$m + n + 1$
Количество данных	n
Операции {	умножения $m + n + 1$
	сложения $m + n$
	передачи n
Полное количество операций	$2m + 3n + 1$
Всего	$3m + 5n + 2$

Эта условная табличка показывает определенные преимущества последовательного программирования перед прямым. На практике метод последовательного программирования часто используется при численном анализе, но называется иначе. Например, при двойном интегрировании можно применить формулу для определения сразу двойного интеграла или дважды применить формулу квадратуры, используя две последовательные операции.

Параллельное программирование

Для характеристики этого метода запишем выражение $W^*(z)$ с учетом введенных выше ограничений в виде

$$W^*(z) = \frac{f_1}{1+d_1z} + \frac{f_2}{1+d_2z} + \dots + \frac{f_n}{1+d_nz}, \quad (35)$$

где f_k — вычеты $W^*(z)$ в полюсах $z_k = -1/d_k$. Как показывает фиг. 13, в, сумма передаточных функций реализуется путем приложения одного и того же входного сигнала ко всем блокам и суммирования их выходных сигналов. Математически этому будет соответствовать выражение ¹⁾

$$\frac{O^*(z)}{I^*(z)} = \frac{O_1^*(z)}{I^*(z)} + \frac{O_2^*(z)}{I^*(z)} + \dots + \frac{O_n^*(z)}{I^*(z)}. \quad (36)$$

Приравнивая соответствующие члены уравнений (35) и (36) и применяя обратное преобразование Лапласа, получим систему уравнений

$$\left. \begin{aligned} o_1^*(t) &= f_1 i^*(t) - d_1 o_1^*(t - T) \\ o_2^*(t) &= f_2 i^*(t) - d_2 o_2^*(t - T) \\ &\vdots \\ o_n^*(t) &= f_n i^*(t) - d_n o_n^*(t - T) \\ o^*(t) &= o_1^*(t) + o_2^*(t) + \dots + o_n^*(t) \end{aligned} \right\} \quad (37)$$

табличка для которой приведена ниже.

¹⁾ Выходные сигналы $O_1^*(z), \dots, O_n^*(z)$ не совпадают с таковыми в (33).

10. Частотный анализ цифровых вычислительных машин 355

ЕМКОСТЬ ЗАПОМИНАЮЩИХ УСТРОЙСТВ ПРИ ПАРАЛЛЕЛЬНОМ ПРОГРАММИРОВАНИИ (УСЛОВНО)	
Количество постоянных	$2n$
Количество данных	n
Операции	$\left\{ \begin{array}{l} \text{умножения } 2n \\ \text{сложения } 2n - 1 \\ \text{передачи } n \end{array} \right.$
Полное количество операций	$5n - 1$
Всего . . . $8n - 1$	

Параллельное программирование, по-видимому, занимает с точки зрения эффективности промежуточное положение между двумя другими методами. Оно редко используется в вычислительной работе. В самом деле, совершенно не ясно, каким образом соответствие между (2) и (37), которое было нами только что установлено, можно было бы получить обычными методами во временной области.

Сравнение методов программирования

Хотя число запоминающих регистров и необходимое время вычислений может служить оценкой методов программирования, многое зависит и от конкретных особенностей самого прибора и кода инструкций. Действительно, исследование методов программирования может повлиять на проектирование цифровых машин и разработку кода. Включение в код быстродействующей вычислительной машины так называемой „обменной“ (whirlwind) инструкции (которая передает результат из арифметического устройства в избранный запоминающий регистр и обратно) делает прямой метод программирования более эффективным, чем последовательный.

Прямой метод программирования более удобен в случае, когда некоторые постоянные в выражении $W^*(z)$ равны нулю, так как при последовательном методе преобразование „восстанавливает“ все постоянные, увеличивая число умножений. Кроме того, при прямом методе программирования можно свести к минимуму задержку, вносимую вычислением, так как все члены (2), исключая $a_0 i^*(t)$, могут быть подсчитаны до поступления на вход дискретного сигнала $i^*(t)$, что важно в применении к системам автоматического регулирования.

С другой стороны, последовательное программирование имеет то преимущество, что позволяет получать промежуточные результаты, что иногда также оказывается полезным.

Так, двойное интегрирование (в „прямом“ методе) определяет координату по измеренному ускорению, но два простых однократных интегрирования в „последовательном“ методе позволяют получать в качестве промежуточного результата также и скорость.

Параллельное программирование, помимо его академического интереса, может оказать влияние и на проектирование вычислительного устройства. Так как во всех уравнениях системы (37) фигурирует один и тот же входной сигнал $i^*(t)$, они могут быть решены совместно, если построены n соответствующих арифметических элементов. Такой вариант конструкции может показаться странным, но при этом можно получить выигрыш в скорости почти в n раз. В отдельных частных случаях для решения задач программирования может применяться совокупность этих методов.

При составлении или усовершенствовании программы экспериментальными методами могут обнаружиться специальные преимущества последовательной или параллельной схемы. Если проектирование основывается на изучении расположения полюсов и нулей передаточной функции, целесообразно использовать последовательный метод, при котором каждой постоянной в уравнении (34) непосредственно соответствует отдельный рассчитываемый параметр. Аналогично при параллельном программировании каждой постоянной соответствует или полюс, или вычет. Для получения желаемого результата отдельные постоянные должны подбираться.

Заключение

Цель статьи состояла в разработке метода анализа цифровых вычислительных машин, используемых в системах автоматического регулирования. При этом основное внимание уделено скорее подходу к проблеме, чем применению специальных методов. Рассмотренные методы могут найти применение и при решении обычных задач численного анализа, таких, как сходимость и резонанс численных решений. Хотя изложение не может претендовать на полноту, получены некоторые основные результаты, которые легко можно развить.

Трудность заключается в том, что область цифрового автоматического регулирования еще молода и ее проблемы и нужды не вполне ясно определены. Например, общая

10. Частотный анализ цифровых вычислительных машин 357

проблема аппроксимации желаемых характеристик с помощью передаточной функции реализуемой программы представляет большие трудности, а специальные частные случаи возникают лишь при конкретных применениях. Развитие техники реализации программ связано с разработкой новых элементов вычислительных устройств, конструктивных и логических схем.

Может возникнуть вопрос о целесообразности применения линейного анализа к цифровым вычислительным машинам, которые осуществляют разнообразные нелинейные операции. Однако даже при анализе нелинейных систем линейные методы анализа могут применяться и оказаться полезными в качестве вспомогательных методов. Распространение частотных методов анализа на цифровые вычислительные машины позволяет в принципе распространить существующие методы анализа следящих систем на смешанные системы, что во многих случаях может оказаться полезным.

ЛИТЕРАТУРА

1. Linvill W. K., Salzer J. M., *Proc. IRE*, 41, 901 (1953).
2. Linvill W. K., Salzer J. M., „Analysis of Digital Computers in Control Systems“, Wayne Computer Conf., Detroit, Mich., March 28, 1951.
3. Salzer J. M., „Frequency Analysis of Digital Computers Used in Control Systems“, Nat. Convention IRE, 1952.
4. Salzer J. M., „Treatment of Digital Control Systems and Numerical Processes in the Frequency Domain“, Диссертация, Mass. Inst. Technol., 1951.
5. Linvill W. K., *Trans. AIEE*, 70, 1779 (1951).
6. Ragazzini J. R., Zadeh L. A., *Trans. AIEE*, Part II, 71, 225 (1952).
7. Oliver B. M., Pierce J. R., Shannon C. E., *Proc. IRE*, 36, 1330 (1948).
8. Shannon C. E., *Proc. IRE*, 37, 11 (1949).
9. Salzer J. M., *Proc. NEC* (Chicago), 8, 621 (1952).
10. Salzer J. M., *Radio-Electronics Eng.*, Febr. 1953, p. 13—15, 30.
11. Tustin A., *Journ. IEE*, 94, Part 2A, 130 (1947).
12. Wiener N., *Extrapolation, Interpolation and Smoothing of Stationary Time Series*, New York, 1949, p. 129—148.
13. Madwed A., „Number Series Method of Solving Linear and Non-Linear Differential Equations“, Диссертация, Mass. Inst. Technol., 1950.

11. ПРИМЕНЕНИЕ ЧАСТОТНЫХ МЕТОДОВ ДЛЯ АНАЛИЗА
ПРОЦЕССА РЕГУЛИРОВАНИЯ ТУРБИН
В ЭНЕРГОСИСТЕМЕ ШВЕЦИИ¹⁾

В. Ойя

ОБОЗНАЧЕНИЯ

Функции времени

Указанные ниже величины относятся к синусоидальным отклонениям, наложенным на среднее значение. Они являются векторными величинами, но так как в этой статье они не используются в другом смысле, то знак вектора всюду опущен.

- a' — положение вспомогательного сервомотора;
- a — относительная величина открытия направляющего аппарата турбины;
- f — частота энергосистемы;
- P — мощность;
- P_T — выходная мощность турбины;
- P_s — возмущение мощности;
- P_2 — остаточное возмущение (после регулирования).

Передаточные функции

Эти функции являются функциями комплексного переменного, зависящими от частоты.

- F — передаточная функция элемента системы автоматического регулирования;
- $|F|$ — абсолютное значение F ;
- φ — фаза F ;
- Φ_k — передаточная функция разомкнутой системы с демпфированием;
- Φ_0 — передаточная функция разомкнутой системы без демпфирования.

Другие величины

- α_b — среднее значение величины открытия направляющего аппарата турбины;
- τ_0 — постоянная времени регулятора;

¹⁾ Oja V., *Trans. ASME*, 76, № 8, 1325—1331 (1954).

11. Анализ процесса регулирования турбин в энергосистеме 359

- τ_2 — время разгона воды;
 τ_c — постоянная времени стабилизирующей обратной связи;
 δ_p — постоянная жесткой обратной связи;
 δ_i — постоянная гибкой обратной связи;
 γ_r — эффективность регулирования;
 β_r — относительное усилие регулирования;
 ω — угловая частота возмущения;
 ρ — жесткость сети, *мгвт/гц*.

Энергосеть Швеции

Вся территория Швеции покрыта сетью электропередач, соединенных в единую систему; ее географическая протяженность составляет примерно 1400 км с севера на юг и 500 км с востока на запад. Общая мощность системы около 3500 мгвт (1\$49 г.). Энергия гидроэлектростанций составляет 97—98% всей энергии, производимой в течение нормального года.

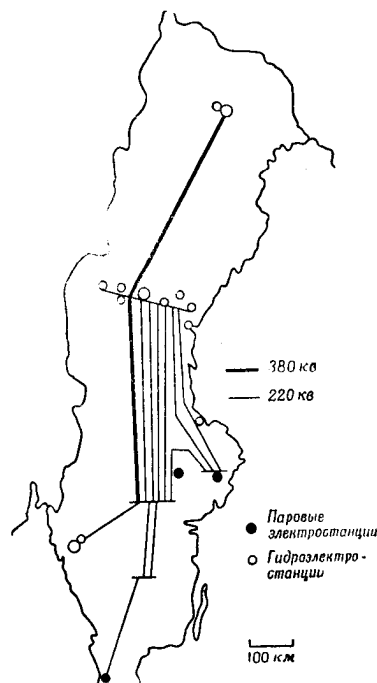
Около 85% энергии гидроэлектростанций производится в северной части страны, где проживает только одна шестая часть всего населения, а основные центры потребления расположены на юге. Электроэнергия передается на расстояние с помощью системы линий электропередачи на 380 и 220 кв, образующих основной костяк энергосети, как показано на фиг. 1. Потребление электроэнергии можно приблизительно классифицировать следующим образом: промышленность — 65%, бытовые нужды — 25% и железнодорожный транспорт — 10%.

Проектирование регуляторов частоты энергосистем осуществляется в Швеции Государственным управлением по энергетике. При данной конфигурации сети разрабатывается следующая схема [1]:

Одна мощная ведущая станция с достаточным водным резервом предназначена для слежения за большими, но медленными изменениями нагрузки. Она является астатической и сильно задемпфирована. Так как реакция ее является замедленной, то можно управлять нужным количеством агрегатов на этой станции, поддерживая хорошую полную эффективность и ограничивая износ.

Многочисленные вспомогательные станции работают в статическом режиме и при пониженном демпфировании. Падение

скорости вращения турбин на этих станциях может быть связано с изменением потребляемой мощности, ослаблением потока воды или его среднего значения, снижением уровня



Фиг. 1. Схема шведских линий электропередач на 380 и 220 кВ с указанием основных электростанций.

с помощью фиг. 2. В электронном регуляторе А создается напряжение, пропорциональное отклонению регулируемой величины (частоты). Это напряжение усиливается и преобразуется в постоянный ток. На изменение частоты регулятор реагирует с помощью настроенного LC-контура 1. Для настройки регулятора используется цепь 2. RC-цепочка 3 обеспечивает необходимое демпфирование. С помощью простых реостатов, расположенных на панели управления, можно легко

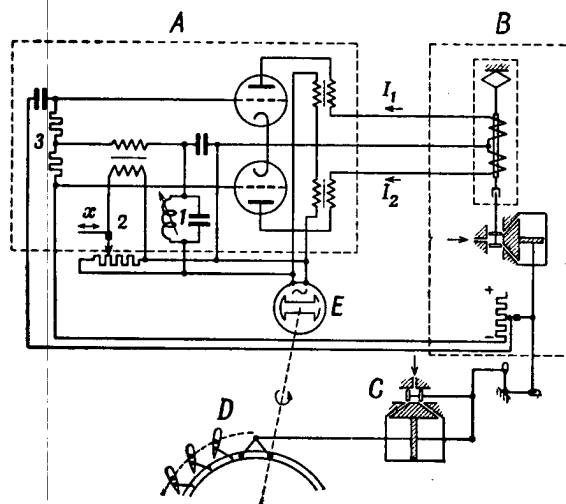
воды, понижением давления пара в паровых турбинах или изменением других параметров в зависимости от местных условий. Эти станции должны следить за быстрыми изменениями нагрузки, работая в узком диапазоне относительно их расчетной рабочей точки. Так как эта схема требует большой гибкости регуляторов, чтобы сочетать различные параметры в разных пропорциях и позволить гибкую регулировку демпфирования, то перед общим анализом будет дано краткое описание современного регулятора.

Работа регулятора

На всех станциях шведской энергосистемы, построенных в последнее время, используются электрические регуляторы турбин [2, 6]. Принцип их работы можно кратко пояснить

11. Анализ процесса регулирования турбин в энергосистеме 361

регулировать соотношения между этими различными величинами. Выходной ток регулятора приводит в действие электрогидравлическую систему *B*, которая превращает электриче-



Фиг. 2. Схема электрогидравлического турбинного регулятора.

A — электронный регулятор; *B* — механизм управления; *C* — главный сервомотор; *D* — механизм поворота лопаток; *E* — вспомогательный генератор переменного тока.

ский сигнал в механическое перемещение. Назначение главного сервопривода *C*, механизма поворота лопаток турбины *D* и вспомогательного генератора *E* не требует пояснений.

Замкнутая система

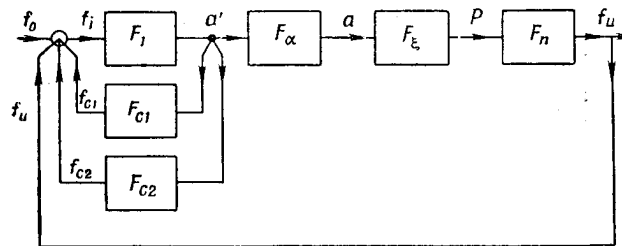
Если предположить, что во всей энергосистеме регулируется только один генератор и что регулятор воздействует только на частоту и степень открытия направляющего аппарата, то можно получить структурную схему, изображенную на фиг. 3. Схематически система регулирования показана на фиг. 4.

Первое звено на фиг. 3 представляет электрогидравлический регулятор турбины. Здесь F_1 — передаточная функция регулятора, связывающая отклонение частоты f_i с механическим выходом a' ; она может быть записана в виде

$$F_1 = \frac{a'}{f_i} = \frac{1}{j\omega\tau_0}. \quad (1)$$

Таким же образом можно представить передаточную функцию звена жесткой обратной связи

$$F_{c1} = \frac{f_{c1}}{a'} = \delta_p. \quad (2)$$



Фиг. 3. Структурная схема системы регулирования.

Выход этого звена для удобства выражен через отклонение частоты. Звено гибкой обратной связи, или демпфирующее звено, характеризуется передаточной функцией

$$F_{c2} = \frac{f_{c2}}{a'} = \delta_t \tau_c \frac{j\omega}{1 + j\omega\tau_c}, \quad (3)$$

где τ_c — постоянная времени цепочки RC .

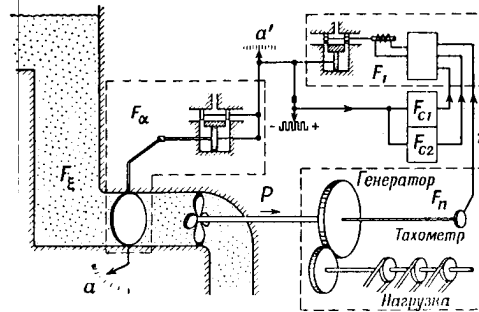
Звено F_{α} отражает свойства главного сервомотора, когда турбина находится под нормальным напором воды [3]. Следовательно, a — соответствующее открытие направляющего аппарата турбины, пропорциональное выходной мощности турбины P_{T_0} при нормальном напоре воды. В некоторых случаях эту передаточную функцию можно представить в виде аналитического соотношения или рассчитать методом последовательных приближений.

Влияние инерции воды F_{ξ} для среднего напора и небольших отклонений от средней величины можно выразить в виде

$$F_{\xi} = \frac{P}{P_{T_0}} = \frac{1 - j\omega\tau_2\alpha_b}{1 + j\omega 0,5\tau_2\alpha_b}, \quad (4)$$

11. Анализ процесса регулирования турбин в энергосистеме 363

где τ_2 — время разгона воды и α_b — средняя величина открытия направляющего аппарата турбины. Здесь P — дей-

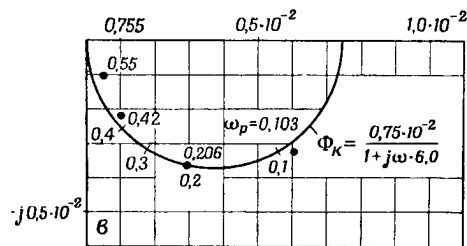
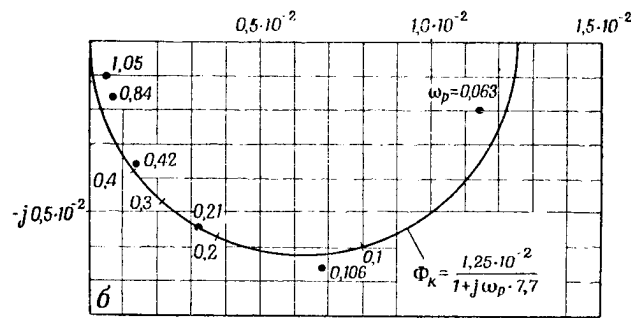
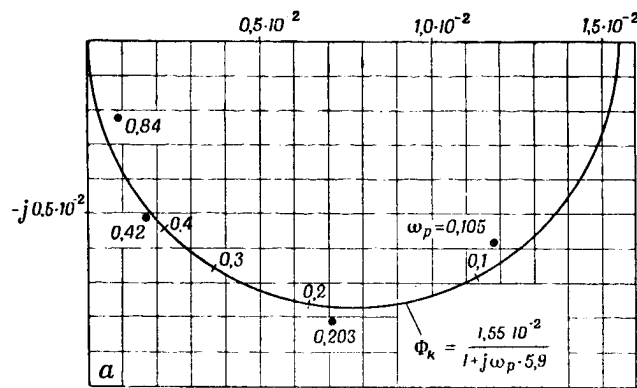


Ф и г. 4. Схема системы регулирования.

ствительная выходная мощность турбины. Последнее звено F_n относится к энергосистеме.

Передаточная функция сети

Эта функция [4] дает соотношение между небольшими изменениями частоты f и изменениями входной мощности P системы. Это соотношение было измерено экспериментально. На время этих измерений регуляторы всех турбин сети были выключены, а необходимая частота поддерживалась с помощью ручных регуляторов. Затем выход одного генератора на 50 мвт меняли по синусоидальному закону с различными частотами. Результирующие изменения частоты сети измерялись и сравнивались как по амплитуде, так и по фазе с меняющейся входной мощностью в специально спроектированной измерительной схеме. Для того чтобы отличить вынужденные изменения частоты сети от более значительных случайных изменений, возникающих в результате нормальных колебаний нагрузки, измеряемые величины умножались на синусоидальную функцию с частотой вынужденных колебаний мощности. Результат этого умножения интегрировался в течение длительного промежутка времени (около 10 мин.), чтобы исключить влияние случайных изменений.



Фиг. 5. Амплитудно-фазовые характеристики сети, $гц/мгвт$.
 а — воскресный день от 12.00 до 15.00; б — рабочий день от 00.00 до 04.00; в — рабочий день от 10.00 до 11.00.

11. Анализ процесса регулирования турбин в энергосистеме 365

Результаты этого исследования приведены в виде диаграмм Найквиста на фиг. 5 для различных условий нагрузки в сети в зависимости от времени суток. На тех же графиках вычерчены кривые, соответствующие уравнению

$$F_n = \frac{1}{\rho} \frac{1}{1 + j\omega\tau_n} \quad (5)$$

Уравнение (5) дает хорошее совпадение с экспериментальными значениями при низких частотах. На более высоких частотах возникает расхождение, обусловленное в основном реактивностью трансформаторов и высоковольтных линий. Но так как усиление на высоких частотах мало, то можно использовать это простое выражение для предварительного анализа всех наиболее важных частей сети.

Метод анализа

После определения всех элементов замкнутой системы можно разработать метод анализа, который даст сведения, необходимые для того, чтобы заставить генераторы на различных станциях работать согласно заранее заданной схеме.

Следуя основной идее частотного метода анализа, предположим, что возмущение выражается синусоидальной функцией. Зная передаточные функции элементов, можно легко определить частотную характеристику всей системы. С другой стороны, так как регулирование связано с износом и иногда с падением к. п. д., то анализ должен указать пути использования данных генераторов для регулирования наиболее благоприятным и экономически выгодным способом.

Исходя из этого, регулирование будет характеризоваться двумя величинами: эффективностью регулирования и суммарным регулирующим усилием.

Эффективность мы определим как отношение уменьшения амплитуды возмущения к ее первоначальной величине, т. е. как действительно полученное улучшение, выраженное в виде отношения. На фиг. 6 показана турбина с регулятором и сервомеханизмами с общей передаточной функцией F_T , присоединенная к сети с передаточной функцией F_n .

Запишем очевидное соотношение

$$P_2 = P_s - P_T; \quad (6)$$

действительное улучшение будет равно

$$|P_\eta| = |P_s| - |P_2|. \quad (7)$$

Согласно определению, эффективность регулирования

$$\eta_r = \frac{|P_\eta|}{|P_s|}. \quad (8)$$

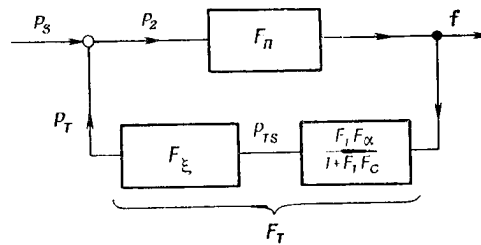
После подстановки передаточной функции разомкнутой системы

$$\Phi_k = \frac{P_T}{P_2}$$

получим, что

$$\eta_r = 1 - \left| \frac{1}{1 + \Phi_k} \right|. \quad (8a)$$

Таким путем можно определить кривую $\eta_r(\omega)$, которая показывает эффективность устройства для синусоидальных



Фиг. 6. Блок-схема с одной регулируемой машиной.

возмущений различных частот. Для эффективности регулирования получаем значение $\eta_r = 1$, если возмущения устранены полностью; $\eta_r = 0$, если амплитуда возмущений неизменна; $\eta_r < 0$, если регулирование увеличивает возмущение. Два последних случая имеют место, если мощность турбины P_T не совпадает по фазе с возмущением.

Второй характеристикой является регулирующее усилие, которое нужно приложить, чтобы получить эффективность регулирования, рассмотренную ранее. Регулирующее усилие определяется как отношение абсолютных величин изменения открывающего направляющего аппарата турбины к амплитуде возмущения. На фиг. 6 изменение открывающего направляющего аппарата представлено статической мощностью P_{Ts} .

11. Анализ процесса регулирования турбин в энергосистеме 367

Согласно определению, регулирующее усилие получается в форме

$$\beta_r = \frac{|P_{Tg}|}{|P_g|}. \quad (9)$$

Если подставить передаточную функцию, характеризующую изменение напора воды и ввести передаточную функцию разомкнутой системы, то получится следующее уравнение:

$$\beta_r = \left| \frac{1}{F_{\xi}} \frac{\Phi_k}{1 + \Phi_k} \right|. \quad (10)$$

Таким образом, можно построить кривую $\beta_r(\omega)$, которая дает амплитуду изменения открытия направляющего аппарата при синусоидальном возмущении единичной амплитуды.

На примере можно показать конкретное применение этих величин. При расчете были использованы данные генератора на 50 мвт с турбиной Каплана и напором воды около 30 м, исследованного экспериментально. Время разгона воды $\tau_2 = 1,5$ сек., а нагрузка в сети соответствовала воскресному дню.

Вычисление постоянных обратной связи

Предположим, что регулируется только один генератор и жесткая обратная связь F_{c1} разомкнута (см. фиг. 3). В этом случае получается астатическое регулирование.

Передаточная функция главной цепи без гибкой обратной связи F_{c2} запишется в виде

$$\Phi_0 = F_1 F_a F_{\xi} F_n, \quad (11)$$

которая при замыкании демпфирующей обратной связи F_{c2} станет равной

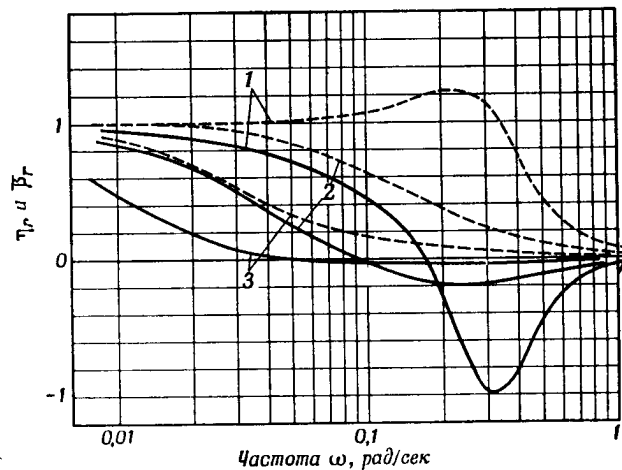
$$\Phi_k = \frac{\Phi_0}{1 + F_1 F_{c2}}, \quad (12)$$

где, исходя из уравнений (1) и (3),

$$F_1 F_{c2} = \frac{1}{\tau_0} \frac{\delta_t \tau_c}{1 + j\omega \tau_c}. \quad (13)$$

В этих уравнениях известны все величины, за исключением постоянных δ_t и τ_c . Предположим, что значения этих постоянных берутся из испытаний и обеспечивают достаточно

хорошее регулирование. Мы примем $\delta_t = 1\%$ и $\tau_o = 6,5$ сек. Затем определяем амплитудно-фазовую характеристику для разомкнутой системы Φ_k . Приняв эту кривую за основу и используя уравнения (8а) и (10), получаем кривые для эффективности и регулирующего усилия. Результаты показаны на фиг. 7 в виде кривых 1.



Фиг. 7. Эффективность регулирования η_r и регулирующее усилие β_r для системы регулирования с одним генератором без жесткой обратной связи.

— η_r ; - - - β_r .

Заметим, что эффективность регулирования становится отрицательной для $\omega > 0,17$, т. е. для частоты возмущения около 1,7 периода в минуту и достигает приблизительно величины -1 для $\omega = 0,3$, в то время как регулирующее усилие остается все еще высоким.

Согласно описанной схеме, астатически регулируемая ведущая станция должна быть сильно задемпфирована. Здесь будут описаны методы, которые дают возможность определить постоянные демпфирования, исходя из некоторых заранее известных условий для кривой эффективности $\eta_r(\omega)$. В качестве примера будет взято простое условие фиксации

11. Анализ процесса регулирования турбин в энергосистеме 369

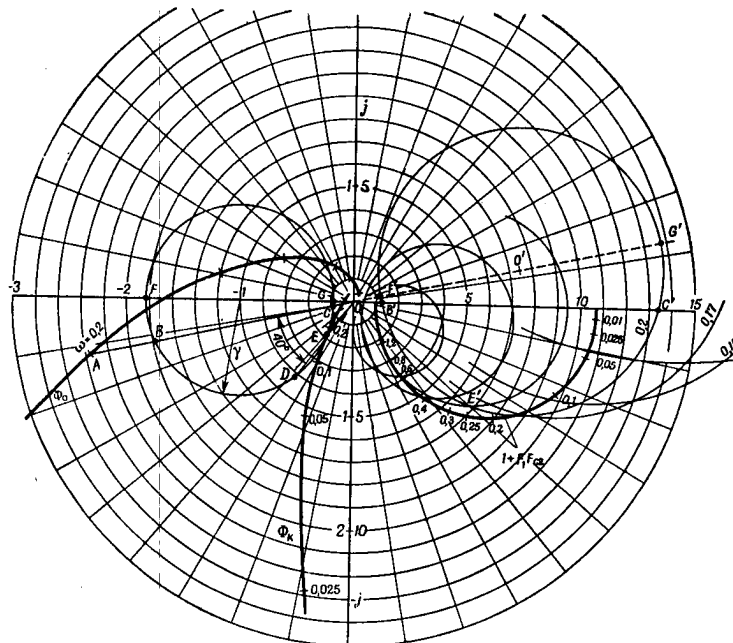
постоянной минимальной величины эффективности, равной $-0,2$. Согласно уравнению (8а), это означает, что

$$-0,2 \leq 1 - \left| \frac{1}{1 + \Phi_k} \right|, \quad (14)$$

т. е.

$$|1 + \Phi_k| \geq 0,83. \quad (14а)$$

Дальнейшие вычисления выполнены графически и показаны на фиг. 8.



Для любой точки известной кривой Φ_0 , например для $\omega_1 = 0,2$, можно найти определенные условия, которым функция $1 + F_1 F_{c2}$ должна удовлетворять при этой частоте. Если вектор $1 + F_1 F_{c2}$ при $\omega = 0,2$ имеет фазовый угол 0° , то вектор Φ_k будет иметь тот же угол, что и Φ_0 (в данном случае 168°) и амплитуду, которая определяется абсолютной величиной $|1 + F_1 F_{c2}|$. Для того чтобы соответствующая точка на кривой Φ_k не могла появиться внутри начерченного круга, абсолютная величина $|1 + F_1 F_{c2}|$ не должна лежать между 1,34 и 13,4. Эти ограничения получаются из отношений $\overline{AO}/\overline{BO}$ и $\overline{AO}/\overline{CO}$. Однако если вектор $1 + F_1 F_{c2}$ при $\omega = 0,2$ имеет, например, угол, равный 110° , то амплитуда не должна лежать внутри области, ограниченной в одном направлении величиной $6,2 |\overline{AO}/\overline{EO}|$. Таким способом получают окружность, внутри которой будут лежать те значения, которых функция $1 + F_1 F_{c2}$ для частоты $\omega = 0,2$ принять не может, если должно быть выполнено первоначально заданное требование.

Легко показать, что диаметр окружности равен

$$D = \overline{G'F'} = \frac{\overline{AO}}{\overline{GO}} - \frac{\overline{AO}}{\overline{FO}} = |\Phi_0| \left[\frac{1}{1-\gamma} - \frac{1}{1+\gamma} \right] = |\Phi_0| \frac{2\gamma}{1-\gamma^2}, \quad (15)$$

где γ — радиус первоначальной окружности (в этом случае 0,83).

Центр этой окружности лежит на продолжении \overline{OA} на расстоянии от начала, равном

$$OO' = \frac{\overline{AO}}{\overline{FO}} + \frac{D}{2} = |\Phi_0| \left[\frac{1}{1+\gamma} + \frac{\gamma}{1-\gamma^2} \right] = \frac{|\Phi_0|}{1-\gamma^2}. \quad (16)$$

Это построение для $\omega = 0,2$ можно повторить для различных частот и получить семейство окружностей для различных значений частоты. Функция обратной связи $1 + F_1 F_{c2}$ должна выбираться так, чтобы точка для частоты ω_n лежала снаружи соответствующей ей окружности. В данном случае функция обратной связи $1 + F_1 F_{c2}$ представляется, согласно уравнению (13), полуокружностью с центром на действительной оси, проходящей через точку $+1$ при $\omega = \infty$. Радиус полуокружности выбирается так,

11. Анализ процесса регулирования турбин в энергосистеме 371

чтобы он был наименьшим возможным радиусом, удовлетворяющим упомянутым выше требованиям. Эта полуокружность на фиг. 8 должна проходить в непосредственной близости от точки пересечения окружностей $\omega = 0,25$ и $\omega = 0,2$. Действительная полуокружность проведена с некоторым запасом. Затем для точки на полуокружности выбирается шкала частот, в данном случае для точки $\omega = 0,25$. Все последующие вычисления параметров функции $F_1 F_{c2}$ проделаны с помощью измерения координат точки кривой при $\omega = 0,25$ относительно точки $+1$, в данном случае равных $|F_1 F_{c2}| = 6,3$ и $\varphi = 49,5^\circ$. В результате вычисления по уравнению (13) мы получим $\delta_t = 5\%$ и $\tau_c = 4,7$ сек.

На фиг. 8 приведена кривая Φ_k , вычисленная для данного значения обратной связи. На фиг. 7 (кривые 2) показаны кривые эффективности и кривые регулирующих усилий.

Для сравнения аналогичные вычисления были проведены для случая еще большего демпфирования. В качестве начальной точки было принято $\eta_{r \text{ мин.}} = -0,04$. При этом получены постоянные обратной связи $\delta_t = 16,5\%$ и $\tau_c = 6$ сек. На фиг. 7 (кривые 3) показаны соответствующие кривые для η_r и β_r .

В предыдущих рассуждениях учитывалась только минимальная величина η_r . Можно задать более сложные условия и разрешить задачу подобным же образом.

Генератор с жесткой обратной связью можно исследовать аналогичным образом, включив обратную связь F_{c1} в первое звено. Затем по схеме фиг. 3 можно получить следующее соотношение:

$$F_{11} = \frac{F_1}{1 + F_1 F_{c1}} = \frac{1}{\delta_p} \frac{1}{1 + j\omega \frac{\tau_0}{\delta_p}}. \quad (17)$$

Характеристика разомкнутой системы принимает вид

$$\Phi'_0 = F_{11} F_a F_\xi F_n, \quad (18)$$

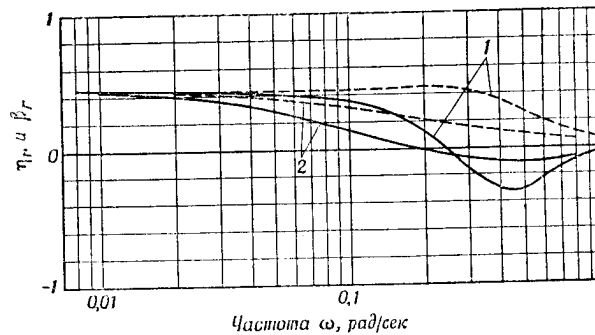
а когда включена гибкая обратная связь, имеем

$$\Phi_k = \frac{\Phi'_0}{1 + F_{11} F_{c2}}, \quad (19)$$

где, обращаясь к уравнениям (3) и (17),

$$F_{11}F_{c2} = \frac{\delta_t \tau_c}{\delta_p} \frac{j\omega}{1 + j\omega \left(\frac{\tau_0}{\delta_p} + \tau_c \right) + (j\omega)^2 \frac{\tau_0 \tau_c}{\delta_p}}. \quad (20)$$

Соответствующие кривые для эффективности и регулирующего усилия показаны на фиг. 9 при жесткой обратной



Фиг. 9. Эффективность регулирования η_r и регулирующее усилие β_r для одного генератора, имеющего 2% жесткой обратной связи.

— η_r ; - - - β_r .

связи в 2% и при демпфировании $\delta_t = 0$ (кривые 1) и $\delta_t = 3,2\%$, $\tau_c = 10$ сек. (кривые 2).

Регулирование системы

На практике в регулировании системы одновременно участвует несколько генераторов, как это показано на блок-схеме фиг. 10. Можно показать, что передаточная функция разомкнутой системы в этом случае будет равна

$$\Phi_{kt} = F_n F_{T1} + F_n F_{T2} + F_n F_{T3} + \dots, \quad (21)$$

где имеется в виду векторное суммирование. Эффективность всей системы можно подсчитать тем же способом, как это делалось раньше. На фиг. 11 представлены результаты такого подсчета для ведущей станции с двумя генераторами на 50 мвт и вспомогательной станции с тремя генераторами

11. Анализ процесса регулирования турбин в энергосистеме 373

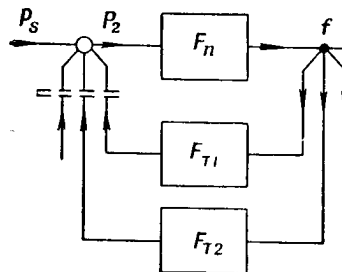
такого же типа. Постоянные обратной связи соответственно равны

$$\begin{aligned} \delta_p &= 0\%, & \delta_z &= 16,5\%, & \tau_e &= 6 \text{ сек.}; \\ \delta_p &= 2\%, & \delta_z &= 1,25\%, & \tau_e &= 4 \text{ сек.} \end{aligned}$$

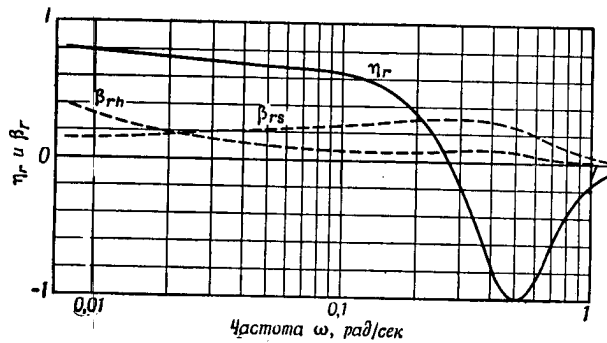
Здесь необходимо заметить, что ведущая станция для согласования с первоначальной схемой должна иметь более сильное демпфирование.

Можно также определить требования к каждому генератору в отношении максимального регулирующего усилия. Из рассмотрения блок-схемы на фиг. 12 следует, что частотная характеристика разомкнутой цепи Φ_{kn} для одного генератора при наличии в схеме ряда других генераторов будет равна

$$\Phi_{kn} = \frac{P_{Tns}}{P_{s2}} = \frac{F_n F_{Tn}}{1 + \Phi_{kt} - F_n F_{Tn}} \quad (22)$$



Фиг. 10. Блок-схема с несколькими регулирующими генераторами.



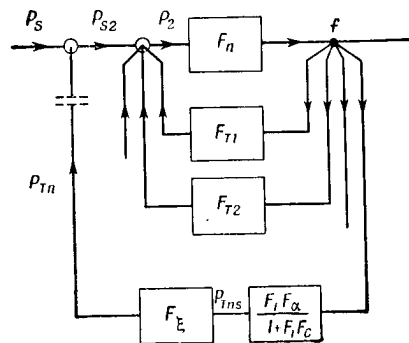
Фиг. 11. Коэффициент полезного действия регулирования всей энергосистемы и регулирующее усилие генераторов.

Регулирующее усилие определяется как соотношение между статической мощностью (изменение в открытии направляющего аппарата) и возмущением и записывается в следующей

форме:

$$\beta_{rn} = \left| \frac{P_{Tns}}{P_s} \right| = \left| \frac{1}{F_\xi} \frac{\Phi_{kn}}{1 + \Phi_{kn}} \right|. \quad (23)$$

Решение этих уравнений дает кривые η_r , β_{rh} (усилие ведущей станции) и β_{rs} (усилие вспомогательной станции), как это показано на фиг. 11. В данном примере эти общие кривые были получены при использовании только одного



Фиг. 12. Блок-схема с несколькими регулирующими генераторами.

типа генераторов для всех станций. Было выяснено, что реальный частотный диапазон мог бы быть перекрыт более эффективно и с меньшими регулирующими усилиями, если бы для различных диапазонов частоты использовались станции с подходящими характеристиками. В этом случае паровые станции можно было бы применять для регулирования быстрых возмущений малой длительности, а гидростанции с длительным временем разгона воды — для медленных, но существенных изменений нагрузки.

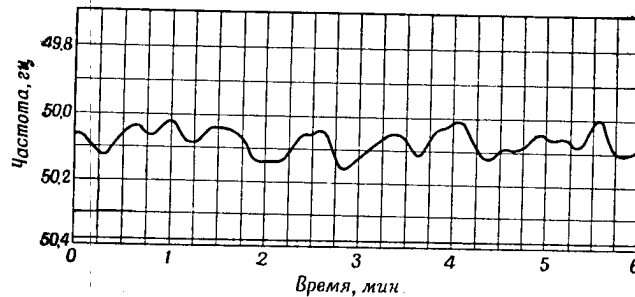
Возмущения в сети

Следующим шагом будет определение характеристики действительных возмущений, встречающихся в системе, путем определения их частотного состава, что необходимо для получения прямой связи между возмущениями и свойствами процесса регулирования. Практически это делается путем

11. Анализ процесса регулирования турбин в энергосистеме 375

наблюдения за изменениями частоты сети и вызывающими их возмущениями. Характерная запись такого рода показана на фиг. 13.

В большинстве случаев нет необходимости производить полный анализ кривой, а достаточно определить амплитуды и частоты преобладающих возмущений. Такие вычисления можно сделать, если известны соотношения между записанными изменениями частоты и величиной возмущений мощности, вызвавших эти изменения.



Фиг. 13. Изменения частоты в регулируемой сети.

Предположим, что приведенная запись относится к системе, выбранной нами в качестве примера. Соотношение между амплитудой возмущения P_s и соответствующим отклонением частоты f получается следующим образом. Согласно фиг. 6, для возмущения с частотой ω_n имеем

$$|P_2| |F_n| = |f| \quad (24)$$

и, по определению (уравнения (7) и (8)),

$$|P_2| = |P_s| (1 - \eta_r); \quad (25)$$

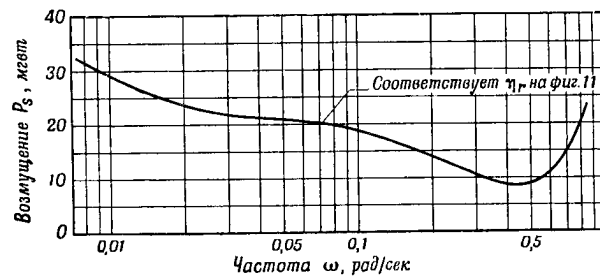
таким образом,

$$|P_s| = \frac{|f|}{|F_n| (1 - \eta_r)}. \quad (26)$$

Кривая $P_s(\omega)$ дана на фиг. 14 для $f = 0,1$ гц.

Оценка параметров вспомогательной станции должна производиться следующим образом. По записи изменений частоты сети находим амплитуду и частоту преобладающего возмущения. С помощью фиг. 14 определяем амплитуду

соответствующего возмущения мощности. Система регулирования вспомогательной станции должна быть достаточно быстродействующей, чтобы следить за действительной частотой возмущения (т. е. эффективность станции должна быть положительной и достаточно большой). Станция также должна



Фиг. 14. Амплитуда возмущения, дающая в регулируемой сети отклонение частоты 0,1 гц.

быть достаточно мощной, чтобы устранить это возмущение. После такой оценки расчет повторяется с учетом новой станции, присоединенной параллельно к другим.

Заключение

Из вышеизложенного ясно, что регулирующие свойства каждой станции системы должны быть согласованы с ее способностью самой стать объектом регулирования путем подбора соответствующей степени демпфирования.

Слишком малое демпфирование увеличит быстродействие без улучшения процесса регулирования в той же степени. Наконец, количество станций, используемых для регулирования, и их характеристики зависят от величины возмущений в сети.

Легко также видеть, что чувствительность регуляторов оказывает небольшое влияние, если их действие линейно и коэффициент усиления достаточно велик, чтобы преобладающую роль играла обратная связь. Другие характеристики обратной связи, отличные от классических, т. е. пропорциональных скорости открытия направляющего аппарата турбины, в этой статье не рассматриваются.

11. Анализ процесса регулирования турбин в энергосистеме 377

Как указывалось, в рассмотренном методе решения был сделан целый ряд допущений. Но так как регулируемые величины или медленно изменяются, или колеблются с небольшими амплитудами около среднего значения, то эти ошибки не будут выходить за известные пределы. Ввиду изменений характеристик сети со временем ни в коем случае нельзя ожидать высокой точности.

Хотя описанный метод расчета еще и не был использован в шведской энергосистеме в такой степени, чтобы можно было получить практические экспериментальные данные, но предполагают, что он даст хорошие результаты при проектировании систем регулирования с приемлемыми общими характеристиками при минимальных регулирующих усилиях. Данный метод также дает возможность приспособить регулирование к изменениям условий работы энергосистемы.

ЛИТЕРАТУРА

1. Lof B., „The Swedish Network, Actual and Planned Regulation System“, CIGRE Meeting, Rep. № 315, Part 2, Paris, 1950.
2. Hedstrom S. E., ASEA Journ., 25, 44 (1952).
3. Гардэ А., Электрогидравлический регулятор водяных турбин, см. сборник „Автоматическое регулирование“, Издательство иностранной литературы, 1954, стр. 463.
4. Almstrom K., Garde A., „Investigations by Frequency Analysis of the Swedish Network“, CIGRE Meeting, Rep. № 315, Part 3, Paris, 1950.
5. Oja V., ASEA Journ., 25, 57 (1952).
6. Hedstrom S. E., Johansson K. E., Linden G., Olsson B., ASEA Journ., 26, 52 (1953).

12. ПРИМЕНЕНИЕ ЧАСТОТНОГО МЕТОДА К АНАЛИЗУ УСТОЙЧИВОСТИ САМОЛЕТА С АВТОПИЛОТОМ ¹⁾

Ч. Л. Сикорд

Введение

В некоторых работах и, в частности, в работах [1] и [2], рассматривается графо-аналитический метод решения задачи устойчивости системы самолет—автопилот. Этот метод обычно называют частотным. Он основывается на принципах, широко используемых в электротехнике, однако лишь в течение последних лет эти принципы начали широко применяться к решению задачи устойчивости самолета. Частотный метод позволяет использовать для определения устойчивости экспериментальные характеристики автопилота и расчетные характеристики самолета. Возможность использования экспериментальных характеристик автопилота, позволяющая отказаться от применения сложных и неточных уравнений, описывающих работу автопилота, является одним из преимуществ этого метода по сравнению с методом непосредственного решения уравнений движения. Другое его преимущество состоит в том, что различные системы могут быть исследованы с меньшей затратой труда, чем при точном решении дифференциальных уравнений.

Возрастающее применение автопилотов требует от авиационного инженера понимания методов, используемых при расчете автопилотов, и границ их применимости. Цель настоящей работы состоит в том, чтобы представить в элементарной форме при минимальном использовании специальных разделов математики метод решения задач при анализе работы автопилота.

В статье приводятся основные уравнения движения самолета, а также некоторые упрощенные уравнения; подробно рассмотрен пример расчета частотной характеристики самолета, включающий ссылки на использованные аэродинамические данные.

¹⁾ Seacord Ch. L., Jr., *J. Aeronautical Sci.*, 17, № 8, 481—498 (1950). Сокращенный перевод.

12. Частотный анализ устойчивости самолета с автопилотом 379

Приведен метод для получения частотных характеристик автопилота, а также метод использования частотных характеристик самолета и автопилота для определения устойчивости системы. Указан также метод, позволяющий по частотным характеристикам получить кривую Найквиста, и рассматриваются критерии применения этой кривой.

В приложении А рассмотрен идеализированный пример, иллюстрирующий сходство результатов, полученных из частотных характеристик, из кривых Найквиста и из решения дифференциальных уравнений движения самолета с автопилотом.

ОБОЗНАЧЕНИЯ

- R_a — амплитуда колебаний самолета при единичном гармоническом воздействии на рули управления ¹⁾);
- R_p — амплитуда колебаний самолета, требуемая для того, чтобы автопилот вызвал колебания рулей управления единичной амплитуды ²⁾);
- ϵ — фазовый угол между δ и ψ ³⁾), град; $\epsilon > 0$, если $\psi > \delta$; индекс a относится к самолету, индекс p — к автопилоту;
- φ — угол крена, рад;
- ψ — угол рыскания, рад;
- β — угол скольжения, рад;
- δ — угол отклонения рулей управления, рад; индекс a относится к элеронам, e — к рулям высоты и r — к рулям направления;
- γ — угол между касательной к траектории центра тяжести и горизонтом, рад; индекс 0 относится к установившемуся состоянию, γ без индекса относится к приращению угла относительно значения γ_0 ;
- δ_0 — амплитуда синусоидальных колебаний рулей управления;

¹⁾ Имеется в виду гармоническое воздействие с амплитудой, равной единице. R_a — амплитудная частотная характеристика самолета. — *Прим. ред.*

²⁾ R_p — обратная амплитудная частотная характеристика автопилота. — *Прим. ред.*

³⁾ Для случая, когда δ и ψ изменяются по гармоническому закону. — *Прим. ред.*

$$r = \frac{d\psi}{dt} \text{ — скорость изменения угла рыскания, рад/сек;}$$

$$p = \frac{d\varphi}{dt} \text{ — скорость изменения угла крена, рад/сек;}$$

t — время, сек.;

ω — угловая частота колебаний, рад/сек;

ω_c — угловая частота, при которой $\varepsilon_a = \varepsilon_p$ и возникают автоколебания системы, рад/сек;

m — масса самолета, т. е. м.¹);

k_x — радиус инерции относительно главной продольной оси, м;

k_z — радиус инерции относительно главной вертикальной оси, м;

η — угол между продольной осью самолета и направлением полета, град; угол положителен, когда главная ось расположена над траекторией полета;

b — размах крыльев, м;

S — расчетная площадь крыльев, м²;

V — воздушная скорость, м/сек;

q — динамическое давление, кг/м²;

ρ — плотность воздуха, т. е. м./м³;

$$D = \frac{d}{dt} \text{ — оператор дифференцирования;}$$

$$j = \sqrt{-1} \text{ — мнимая единица;}$$

C_n — коэффициент момента рыскания;

C_l — коэффициент момента крена;

C_y — коэффициент боковой силы;

$$C_{n\beta}; C_{l\beta}; C_{y\beta} = \frac{\partial C_n}{\partial \beta}; \frac{\partial C_l}{\partial \beta}; \frac{\partial C_y}{\partial \beta};$$

$$C_{nr}; C_{lr}; C_{yr} = \frac{\partial C_n}{\partial (rb/2V)}; \frac{\partial C_l}{\partial (rb/2V)}; \frac{\partial C_y}{\partial (rb/2V)};$$

$$C_{np}; C_{lp}; C_{yp} = \frac{\partial C_n}{\partial (pb/2V)}; \frac{\partial C_l}{\partial (pb/2V)}; \frac{\partial C_y}{\partial (pb/2V)};$$

$$C_{n\delta}; C_{l\delta}; C_{y\delta} = \frac{\partial C_n}{\partial \delta}; \frac{\partial C_l}{\partial \delta}; \frac{\partial C_y}{\partial \delta};$$

$$C_{nD\delta}; C_{lD\delta}; C_{yD\delta} = \frac{\partial C_n}{\partial (D\delta b/2V)}; \frac{\partial C_l}{\partial (D\delta b/2V)}; \frac{\partial C_y}{\partial (D\delta b/2V)};$$

¹) Техническая единица массы 1 т. е. м. = 1 кг·сек²/м = 9,81 кг.—
Прим. ред.

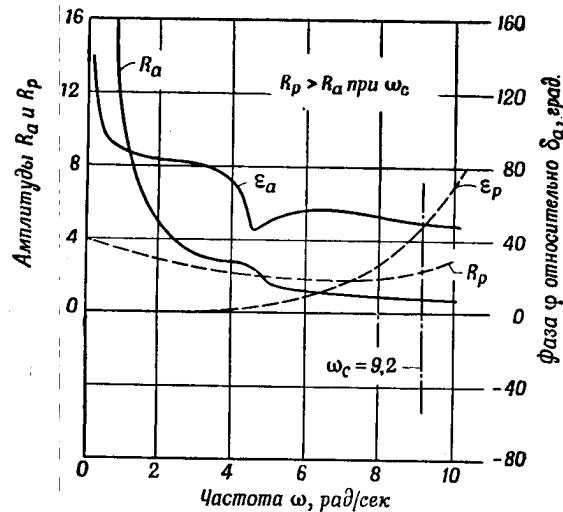
12. Частотный анализ устойчивости самолета с автопилотом 381

K — передаточное число автопилота (коэффициент усиления); при нулевой частоте $\delta = K\psi$, $K\varphi$;
 $P_a(j\omega)$ — передаточная функция самолета;
 $P_p(j\omega)$ — передаточная функция автопилота;
 $KG(j\omega) = P_a(j\omega)P_p(j\omega)$ (см. [2]).

Частотный метод

Общие соображения

Частотный метод для анализа системы самолет — автопилот предполагает использование четырех кривых (фиг. 1).



Фиг. 1. Частотные характеристики типового самолета и автопилота по углу крена.
 Из кривых видно, что система самолет — автопилот устойчива.

Эту систему кривых можно разбить на две пары. Одна пара кривых описывает связь между возмущающими колебаниями рулей (вход) и движением самолета (выход). Другая пара описывает связь между движением самолета (вход) и воздействием

автопилота на рули управления (выход). Для описания каждой характеристики (самолета или автопилота) необходимы две величины: отношение величины амплитуды перемещения самолета к величине амплитуды перемещения рулей управления и фазовый угол ϵ между перемещением самолета и перемещением рулей управления. Отношение величины, характеризующей перемещение самолета по амплитуде и фазе, к величине отклонения рулей управления зависит только от параметров самолета¹⁾. Отношение величины отклонения рулей, вызываемого автопилотом, к величине перемещения самолета как по амплитуде, так и по фазе зависит только от параметров автопилота. Амплитудная и фазовая характеристики обычно определяются для ряда значений частот от нуля до частоты, равной приблизительно тройному значению собственной частоты колебаний самолета. Обе кривые могут быть определены либо экспериментально, либо путем использования уравнений движения. Обычно характеристики самолета определяются расчетным путем, а характеристики автопилота — экспериментально.

Характеристики самолета определяются из решения уравнений движения [3, 4] для случая установившегося движения. Для их определения необходимо заменить оператор дифференцирования D частотной переменной $j\omega$. Это позволяет найти реакцию самолета на единичное возмущение как функцию частоты периодического возмущения.

Характеристики автопилота определяются путем осциллографирования отклонения рулей самолета или перемещения рулевых машин автопилота для различных частот и амплитуд. В том случае, когда перемещение рулей не является синусоидальным, необходимо воспользоваться приближенной заменой его синусоидальным колебанием. Способу определения эквивалентного синусоидального колебания посвящен специальный раздел настоящей работы. Две пары кривых на фиг. 1 были вычислены и сняты для некоторого типового самолета и автопилота. Они используются для определения устойчивости следующим образом: определяется пересечение ϵ -кривых. Точка пересечения опреде-

¹⁾ Здесь имеются в виду соотношения амплитуд и фаз для случая, когда входные и выходные координаты являются гармоническими функциями времени. — *Прим. ред.*

12. Частотный анализ устойчивости самолета с автопилотом 383

ляют критическую частоту, при которой могут существовать нейтрально-устойчивые колебания ¹⁾).

По значениям R при критическом значении частоты ω_c можно определить, будут ли существовать колебания. Если при критической частоте значения R равны, т. е. кривые R пересекаются в точке с частотой ω_c , то будут иметь место нейтрально-устойчивые колебания. Однако, если R_p больше R_a , как в данном примере, колебания будут затухающими, а если R_a больше R_p , то система будет неустойчивой. Существование устойчивых или неустойчивых колебаний — единственное заключение, которое можно сразу сделать из рассмотрения двух пар кривых. По определению [3, 4], частотные характеристики справедливы только при условии нулевого демпфирования, и, следовательно, если R -кривые не пересекаются при том же значении частоты, что и ϵ -кривые, то точка пересечения ϵ -кривых уже не будет определять частоту колебаний. Иными словами, если система самолет — автопилот является устойчивой или неустойчивой, то по основным частотным характеристикам можно установить этот факт, но сами частотные характеристики уже не определяют частоту (или демпфирование) колебаний.

Определение частотных характеристик самолета

Как упоминалось в предыдущем разделе, частотные характеристики самолета могут быть получены из уравнений движения самолета и определяют реакцию самолета на синусоидальное движение рулей управления. Движение самолета в общем случае описывается системой шести дифференциальных уравнений [5, 6]. Вследствие симметрии можно пренебречь связью между боковым движением (крен, рыскание и боковое скольжение) и продольным движением (тангаж и движение вдоль осей x и z) при условии, что колебания самолета малы. Следовательно, уравнения могут быть разделены на две группы, каждая из которых состоит из трех уравнений ²⁾. В некоторых специальных случаях, когда могут

¹⁾ Незатухающие колебания с произвольной постоянной амплитудой. Кривые, о которых здесь идет речь, называются амплитудными (R) и фазовыми (ϵ) характеристиками. — *Прим. ред.*

²⁾ Доказательство см., например, в книге В. Ведрова „Динамическая устойчивость самолета“. Оборонгиз, 1938. — *Прим. ред.*

быть сделаны упрощающие допущения, можно исключить одно или два уравнения.

Примеры таких упрощенных систем, а также полные уравнения продольного движения приводятся в Приложении Б. Полные уравнения бокового движения имеют следующий вид.

Уравнение для угла рыскания

$$\begin{aligned} & \left[\frac{m}{qSb} (k_z^2 \cos^2 \eta + k_x^2 \sin^2 \eta) D^2 - C_{nr} \frac{b}{2V} D \right] \psi + \\ & + \left[\frac{m}{qSb} (k_z^2 - k_x^2) \cos \eta \sin \eta D^2 - C_{np} \frac{b}{2V} D \right] \varphi - C_{n\beta} \beta = \\ & = C_{n\delta_a} \delta_a + \left(C_{n\delta_r} + C_{nD\delta_r} \frac{b}{2V} D \right) \delta_r. \quad (1) \end{aligned}$$

Уравнение для угла крена

$$\begin{aligned} & \left[\frac{m}{qSb} (k_x^2 \cos^2 \eta + k_z^2 \sin^2 \eta) D^2 - C_{lp} \frac{b}{2V} D \right] \varphi + \\ & + \left[\frac{m}{qSb} (k_z^2 - k_x^2) \cos \eta \sin \eta D^2 - C_{lr} D \frac{b}{2V} \right] \psi - C_{l\beta} \beta = \\ & = \left(C_{l\delta_a} + C_{lD\delta_a} \frac{b}{2V} D \right) \delta_a + C_{l\delta_r} \delta_r. \quad (2) \end{aligned}$$

Уравнение для угла скольжения

$$\begin{aligned} & \left[\frac{mVD}{qS} - \frac{W}{qS} \sin \gamma_0 - C_{yr} \frac{b}{2V} D \right] \psi - \\ & - \left[\frac{W \cos \gamma_0}{qS} + C_{yp} \frac{b}{2V} D \right] \varphi + \left[\frac{mVD}{qS} - C_{y\beta} \right] \beta = \\ & = C_{y\delta_a} \delta_a + \left(C_{y\delta_r} + C_{yD\delta_r} \frac{b}{2V} D \right) \delta_r. \quad (3) \end{aligned}$$

Уравнения, приведенные выше, записаны в обычной форме с безразмерными коэффициентами, но члены уравнений имеют размерность, поскольку оператор D означает дифференцирование по времени. Размерная форма уравнений была принята потому, что окончательные кривые должны сравниваться с опытными характеристиками автопилота, которые являются функциями времени или частоты. Примеры полностью безразмерных уравнений могут быть найдены в [7] и [8].

12. Частотный анализ устойчивости самолета с автопилотом 385

Таблица 1

Коэффициент	Численное значение	Способ определения
$C_{n\beta}$	0,2292	По данным продувки в аэродинамической трубе
$C_{l\beta}$	-0,1375	То же
$C_{y\beta}$	-0,9168	» »
C_{nr}	-0,699	Из уравнения $C_{nr} = C_{nr \text{ крыла}} + 2(1/b^2) C_{y\beta} \text{ оперения}$, см. [10]
C_{lr}	0,1719	Оценено по данным [11]
C_{nr}	-0,0040	Из уравнения $C_{nr} = -(C_L - C_{Da})/8$, см. [5]
$C_{l\dot{\beta}}$	-0,6429	Оценено по данным [11]
$C_{n\dot{\beta}}$	-0,0146	По данным продувки в аэродинамической трубе
$C_{l\dot{\alpha}}$	-0,0573	То же
$C_{lD\dot{\alpha}}$	0	Оценено по данным [12]
$C_{y\dot{\alpha}}$	0	По данным продувки в аэродинамической трубе
C_{yr}	0	Оценено по данным [13]
C_{yr}	0	То же
$C_{y\dot{\delta}_r} = C_{n\dot{\delta}_r} = C_{l\dot{\delta}_r} = C_{nD\dot{\delta}_r} = C_{yD\dot{\delta}_r}$	0	
m	90 кг·сек ² /м	
k_z	1,29 м	Вычислено
k_x	0,283 м	»
γ_0	0	
η	0°	Вычислено
b	3,51 м	
q	1775 кг/м ²	
V	179 м/сек	
S	1,95 м ²	

Необходимо отметить, что в уравнения входят коэффициенты двух типов, а именно массовые и аэродинамические коэффициенты. Коэффициенты, зависящие от масс, могут быть получены суммированием составляющих каждой части самолета или экспериментальным путем. Техника подобного эксперимента описана в работе [9].

Аэродинамические коэффициенты. Аэродинамические коэффициенты разделяются на статические и динамические. Статическими коэффициентами называются величины, зависящие от углового смещения самолета относительно потока, в то время как динамические коэффициенты связаны с угловой скоростью вращения самолета относительно воздуха.

При отсутствии достаточно полных данных летного эксперимента наиболее точным методом определения всех статических и некоторых динамических коэффициентов является метод продувки в аэродинамической трубе.

Коэффициенты, относящиеся к боковому движению, используемые в избранном нами примере, вместе с кратким указанием на их способ определения приводятся в табл. 1.

Расчет. Подстановка приведенных численных значений коэффициентов в уравнения (1) — (3) дает следующий результат:

$$\left. \begin{aligned} \text{(I)} \quad & (0,01231D^2 + 0,006859D)\psi + (0,00003935D)\varphi - \\ & - 0,2293\beta = -0,01146\delta_a \\ \text{(II)} \quad & (0,000591D^2 + 0,006308D)\varphi - (0,0001686D)\psi + \\ & + 0,1375\beta = -0,05730\delta_a \\ \text{(III)} \quad & 4,638D\psi + 0,2550\varphi + (4,638D + 0,9168)\beta = 0 \end{aligned} \right\} \quad (4)$$

Чтобы воспользоваться частотным методом, необходимо знать реакцию самолета на синусоидальное движение рулей с постоянной амплитудой и при различных частотах. Ее можно определить, заменив оператор D комплексной переменной частотой $j\omega$. Эта операция, описание которой можно найти в [3] и [4], подробно излагается ниже.

Система уравнений (I) — (III) является системой совместных уравнений с тремя неизвестными φ , β и ψ . Если необходимо определить φ , система решается относительно φ методом определителей. В нашем численном примере имеем:

$$\varphi = \frac{\Delta}{\Delta_0} \delta_a, \quad (5)$$

12. Частотный анализ устойчивости самолета с автопилотом 387

где

$$\Delta = \begin{vmatrix} 0,01231D^2 + 0,006859D & -0,01146 & -0,2292 \\ -0,0001686D & -0,05730 & 0,1375 \\ 4,638 & 0 & 4,638D + 0,9168 \end{vmatrix},$$

$$\Delta_0 = \begin{vmatrix} 0,01231D^2 + 0,006859D & 0,0003935D & -0,2292 \\ -0,0001686D & 0,0005914D^2 + 0,006308D & 0,1375 \\ 4,638D & -0,2550 & 4,638D + 0,9168 \end{vmatrix}.$$

Это выражение может быть упрощено и приведено к виду:

$$\frac{\varphi}{\delta_a} = \frac{10^{-8}(-327\,300D^3 - 247\,900D^2 - 6\,858\,000D)}{10^{-8}(3378D^5 + 38\,580D^4 + 90\,440D^3 + 720\,300D^2 + 23\,070D)}. \quad (6)$$

Замена $D = j\omega$ дает

$$\frac{\varphi}{\delta_a} = \frac{327\,300\omega^3 - 6\,858\,000 - j \cdot 247\,900\omega}{3378\omega^5 - 90\,440\omega^4 + 23\,070 + j(-38\,580\omega^3 + 720\,300\omega)}. \quad (7)$$

Выражение (7) как функция $j\omega$ может быть упрощено и приведено к виду

$$\frac{\varphi}{\delta_a} = A + jB,$$

где

$$A = \frac{110\,000\omega^6 - 4\,320\,000\omega^4 + 44\,930\,000\omega^2 - 1\,582\,000}{1141\omega^8 + 87\,740\omega^6 - 4\,724\,000\omega^4 - 51\,460\,000\omega^2 + 52\,310} \quad (8)$$

и

$$B = \frac{1\,179\,000\omega^5 - 4\,779\,000\omega^3 + 493\,400\,000\omega}{1141\omega^8 + 87\,740\omega^6 - 4\,724\,000\omega^4 + 51\,460\,000\omega^2 + 52\,310}. \quad (9)$$

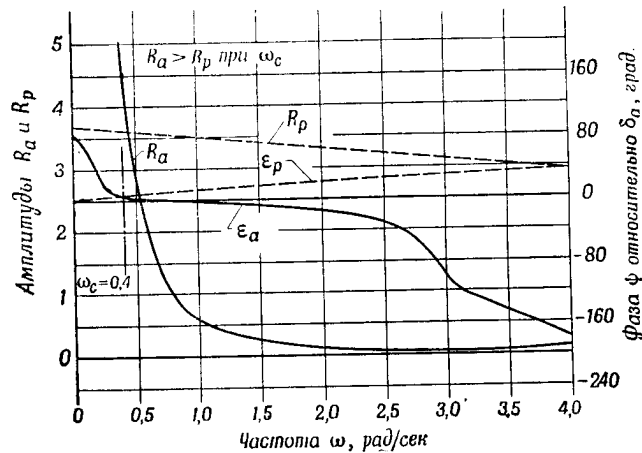
Величинами, которые необходимы для расчета, являются ϵ_a — фазовый угол между φ и δ_a и R_a — абсолютная величина отношения $\varphi_{\text{макс.}}/(\delta_0)_a$. Эти величины могут быть записаны в виде

$$\epsilon_a = \arctg \frac{B}{A} \quad \text{и} \quad R_a = \sqrt{A^2 + B^2}.$$

Подставляя величины A и B в выражения для ϵ_a и R_a , мы получим формулы для R_a и ϵ_a как функции частоты ω (частоты возмущающих колебаний рулей).

Графики, соответствующие формулам, приведены на фиг. 1. Форма кривых значительно меняется в зависимости от того,

какая величина рассматривается в качестве выходной. Как пример возможных изменений на фиг. 2 показаны частотные



Фиг. 2. Частотные характеристики типowego самолета и автопилота по углу рыскания.

характеристики по углу рыскания для того же самолета, который был рассмотрен в предыдущем примере.

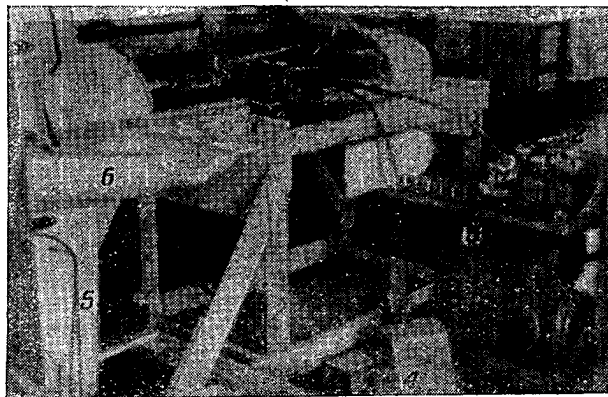
Определение частотных характеристик автопилота на экспериментальном стенде

Как уже указывалось, частотный анализ требует использования двух пар кривых. В предыдущем разделе был дан метод расчета характеристик самолета; в этом разделе дается метод преобразования экспериментальных данных, полученных для автопилотов, в частотные характеристики.

Испытываемый автопилот помещается на вибрационном столе, как показано на фиг. 3. Здесь же помещается записывающая аппаратура для записи положения стола и положения рулей управления, связанных с автопилотом. Блок-схема стенда показана на фиг. 4. Может оказаться, что невозможно расположить автопилот на вибрационном столе и в то же время приводить в движение рули управления самолета. В таком случае необходимо измерить перемещение

12. Частотный анализ устойчивости самолета с автопилотом 389

рулевых машинок без использования рычагов управления или построить модель системы управления. На фиг. 3 показаны рули управления, которые нагружены пружиной для создания аэродинамического шарнирного момента, так как характеристики рулевых машинок существенно меняются с изменением нагрузки на рули. Кривые перемещения стола и перемещения рулей записываются на одну пленку, как показано на фиг. 5.



Фиг. 3. Типовой стенд для снятия частотных характеристик автопилота.

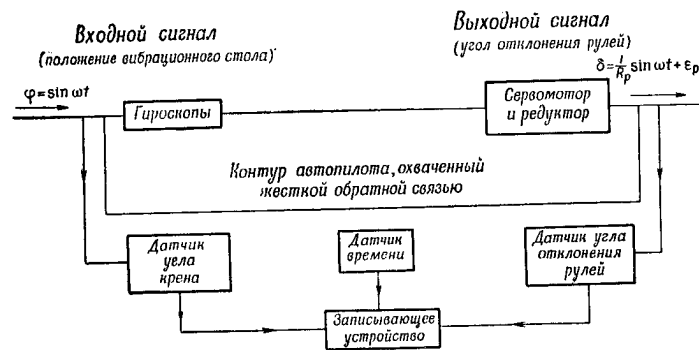
1 — записывающее устройство; 2 — автопилот; 3 — вибрационный стол; 4 — датчик времени; 5 — нагрузка; 6 — рулевой привод.

Следует иметь в виду, что не всегда можно использовать в качестве источника возмущений вибрационный стол. Если чувствительный элемент автопилота не реагирует на угловое отклонение, скорость или ускорение или если сигнал от датчика может быть получен более простыми средствами, то используются другие способы создания возмущений.

После того как произведена запись, мы располагаем тремя методами приближенной замены полученных сигналов эквивалентным синусоидальным сигналом (см. [1]):

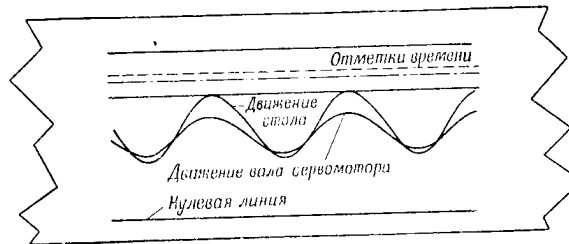
1. Если перемещение рулей представляет собой регулярные симметричные колебания (например, в случае прямоугольных или пилообразных колебаний), можно использовать

первую гармонику разложения в ряд Фурье этих колебаний и по записи перемещения рулей определить амплитуду и фазу этой гармоники.



Ф и г. 4. Блок-схема стенда для снятия частотных характеристик автопилота.

2. В случае когда не нужен точный ответ (т. е. когда имеется большой запас устойчивости), синусоидальный сиг-



Ф и г. 5. Образец осциллограммы, полученной при снятии частотной характеристики автопилота.

нал можно просто заменить самой функцией перемещения рулей, а фазу и амплитуду находить непосредственно по осциллограмме.

3. Если метод „1“ неприменим или если результаты, полученные методом „2“, не обладают требуемой точностью,

12. Частотный анализ устойчивости самолета с автопилотом 391

может быть применен графический метод, который в общих чертах описан в табл. 2.

Таблица 2

I. Величины, которые необходимо измерить или вычислить		
Обозначения	Описание	Единицы измерения
t_0	Период колебания стола на осциллограмме	см
F	Скорость движения ленты отметчика	см/сек
F_1	Калибровка отметчика	град/см
T	Период колебаний стола	сек.
$\omega = 2\pi/T$	Угловая частота	рад/сек
A_1	Площадь под кривой δ , вычисленная на протяжении половины периода	см ²
B_1	Площадь под кривой $\delta \cos \omega t$, вычисленная графическим методом	см ²

Последовательность определения величин:

t_0 , см; $t = (360/T)t_0$, град.; $\cos \omega t$; δ , см; $\delta \cos \omega t$, см;
 $t = 0$ в начале цикла колебаний стола.

II. Искомые величины R_p и ϵ_p

$$R_p = \frac{1}{\sqrt{A^2 + B^2}}, \quad \epsilon_p = \arctg \frac{B}{A},$$

где A и B определяются по формулам:

$$A = \frac{\omega A_1 F_1}{2Fa},$$

$$B = \frac{B_1 F_1}{\pi Fa},$$

— амплитуда колебаний стола.

В тех местах табл. 2, где сделаны ссылки на площадь под кривой, подразумевается площадь между кривой и нулевой линией, за исключением тех случаев, когда кривая не симметрична относительно нулевой линии. Это означает, что если записанная кривая симметрична, то интегрирование может быть проведено непосредственно, если же она несимметрична, то необходимо произвести некоторые операции

для нахождения оси симметрии. Операции эти сводятся к нахождению линии, параллельной нулевой оси, относительно которой кривая будет симметрична (по площади). Для этого необходимо определить площадь между записанной кривой и осью отсчета на протяжении одного периода и разделить ее на длину этого периода. В результате мы получим расстояние от линии отсчета до оси симметрии. Установлено, что точность и быстрота всех интегрирований и измерений возрастает с увеличением длины записи.

Перечисленные выше операции должны быть выполнены для записей, сделанных для ряда частот, а вычисленные кривые R_p и ϵ_p должны быть вычерчены как функции ω . Совокупность кривых, полученных экспериментально для одного вида автопилота с электрическим сервомотором, показана на фиг. 1 и 2.

Анализ частотных данных

Кривые R и ϵ в функции ω . В предыдущих разделах были рассмотрены методы получения частотных характеристик самолета и автопилота. Теперь мы покажем, как можно использовать эти кривые (см. фиг. 1 и 2) для определения устойчивости системы самолет—автопилот. Из фиг. 1 мы видим, что величина R_p при значении ω_c приблизительно в 3 раза больше R_a . Таким образом, амплитуда удовлетворяет условию устойчивости. Действительно, отношение R_p к R_a таково, что движение, вызванное возмущением, будет удовлетворительно демпфироваться.

Другой способ для оценки величины демпфирования состоит в рассмотрении величины запаса устойчивости по фазе $\epsilon_a - \epsilon_p$ в точке, где $R_p = R_a$.

В работе [14] в качестве значения, обеспечивающего удовлетворительное демпфирование, дается величина $+30^\circ$, которая выбирается в основном из экспериментальных соображений. Способ определения устойчивости на основании запаса по фазе становится более ясным при использовании кривых Найквиста, которые будут рассмотрены позднее.

На фиг. 2 имеем случай уже отмеченной неустойчивости. При значении ω_c величина R_p составляет приблизительно половину величины R_a , а запас по фазе в точке $R_a - R_p$ является отрицательным. Если бы самолет совершал полет

12. Частотный анализ устойчивости самолета с автопилотом 393

при этих условиях, то возникли бы колебания с возрастающей амплитудой. Колебания возрастали бы беспредельно, если бы в самолете или автопилоте не было некоторых нелинейностей, изменяющих частотные характеристики и тем самым ограничивающих амплитуду колебаний некоторой большой постоянной величиной.

Систему можно сделать устойчивой, если одну из четырех кривых изменить в нужном направлении. В рассматриваемом случае наклон характеристик самолета в критической области зависит главным образом от таких основных параметров, как инерция относительно оси рыскания и нагрузка на крыло; поэтому в данном случае нельзя найти какого-либо простого видоизменения параметров самолета, которое обеспечило бы устойчивость. Однако можно изменить характеристики автопилота. Средства, с помощью которых можно добиться таких изменений, будут рассмотрены ниже.

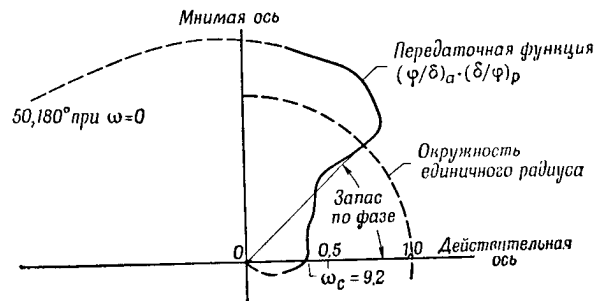
Диаграммы Найквиста. Частотные кривые, полученные экспериментальным путем, о которых говорилось выше, и кривые, полученные расчетным путем, могут быть объединены в одну кривую, вычерченную в полярных координатах. Тогда мы получим хорошо известную диаграмму Найквиста ¹⁾. Она дает в основном те же сведения, которые могут быть получены из отдельных частотных характеристик (амплитудной и фазовой), однако ее применение является более простым и удобным.

Для двух случаев, рассмотренных выше, кривые Найквиста изображены на фиг. 6 и 7. Эти кривые могут быть вычерчены в виде отношения амплитуд самолета и автопилота $R_a/R_p = (\psi/\delta)_a/(\psi/\delta)_p$, взятом при фазовом угле $(\epsilon_a - \epsilon_p)$. Таким образом, две кривые, характеризующие самолет и автопилот, могут быть объединены в одну.

В том случае когда система неустойчива, кривая охватывает точку (1,0), а частота, соответствующая точке пересечения кривой с действительной осью, является частотой колебаний системы, которая была определена нами раньше на основании четырех отдельных частотных кривых. Причина, по которой вместо точки -1 берется точка +1, изложена в Приложении А. Строгое доказательство этого

¹⁾ Амплитудно-фазовая характеристика разомкнутой системы. —
Прим. ред.

критерия (т. е. положения о том, что в случае нейтральной устойчивости кривая проходит через точку $(1,0)$) требует привлечения достаточно сложного математического аппарата и выходит за рамки этой статьи. Однако легко можно показать, что точка пересечения, найденная при использовании двух пар кривых, совпадает с точкой $(1,0)$ кривой Найквиста.



Фиг. 6. Кривая Найквиста для типового самолета с автопилотом по углу крена.
Из кривой видно, что система устойчива.

Функция ¹⁾ $(\psi/\delta)_a \cdot (\delta/\psi)_p$, как известно, носит название передаточной функции системы. Смысл произведения амплитуд R_a/R_p двух передаточных функций становится ясным при рассмотрении условий нейтральной устойчивости, исходя из отдельных частотных характеристик самолета и автопилота. Действительно, если $R_a = R_p$, то $R_a/R_p = 1$, а если $R_a > R_p$, то $R_a/R_p > 1$. Следовательно, точка, которая получается на отдельных частотных характеристиках самолета и автопилота (при условии их пересечения при одной частоте), будет соответствовать на кривой Найквиста точка $(1,0)$. Нетрудно видеть, что если $R_a > R_p$ при $\epsilon_a - \epsilon_p = 0$, то кривая Найквиста пересекает действительную ось вне отрезка $(0,1)$ и, следовательно, система неустойчива (см. фиг. 7); наоборот, если $R_a < R_p$ при $\epsilon_a - \epsilon_p = 0$, то кривая Найквиста

¹⁾ Эта функция в литературе [2] обозначается в виде $KG(j\omega)$. Она может быть также записана в виде

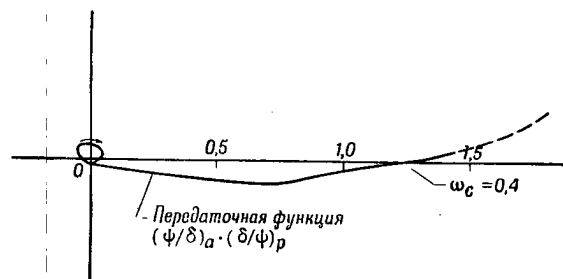
$$\frac{R_a}{R_p} e^{i(\epsilon_a - \epsilon_p)}.$$

12. Частотный анализ устойчивости самолета с автопилотом 395

пересекает действительную ось в пределах отрезка $(0, 1)$ и, следовательно, в этом случае система будет устойчивой (см. фиг. 6).

Введенное выше понятие „запас устойчивости по фазе“ становится ясным из рассмотрения фиг. 6. Запасом по фазе называется угол между действительной осью и лучом, проведенным из начала координат в точку пересечения кривой с окружностью единичного радиуса.

Метод Найквиста (частотный метод) и его различные разновидности достаточно полно рассмотрены в литературе



Фиг. 7. Кривая Найквиста для типового самолета с автопилотом по углу рыскания.

Из кривой видно, что система неустойчива.

[2, 14, 15]¹⁾, где можно найти различные примеры использования рассмотренных графиков. Основное преимущество метода Найквиста состоит в том, что он позволяет объединить в одну кривую любое число кривых, соответствующих отдельным звеньям системы. В том случае когда система состоит из большого числа отдельных звеньев, кривая Найквиста, как легко видеть, может быть получена в результате перемножения амплитудных характеристик отдельных звеньев и алгебраического сложения их фазовых углов. Иначе говоря, передаточная функция всей системы равна произведению передаточных функций отдельных звеньев, т. е. в нашем случае в обозначениях, принятых в работе [2]²⁾:

$$KG(j\omega)_{\text{системы}} = P_a(j\omega)P_p(j\omega).$$

¹⁾ См., например, „Основы автоматического регулирования“ под ред. В. В. Солодовникова, Машгиз, 1954. — Прим. ред.

²⁾ $P_a(j\omega)$ соответствует нашему $R_a = (\psi/\delta)_a$, а $P_p(j\omega)$ — нашему $1/R_p = 1/(\psi/\delta)_p$.

Следовательно, мы получаем возможность разложить нашу систему на отдельные звенья и выделить те из них, параметры которых необходимо подбирать. Нетрудно видеть, что на основании этого свойства можно заранее определять свойства системы самолет—автопилот при условии, что определенная часть системы (например, самолет) имеет заданные характеристики, и, таким образом, проектировать системы с желаемыми свойствами ¹⁾.

Заключение

Метод частотных характеристик как в форме отдельных амплитудно-частотной и фазо-частотной характеристик, так и в виде амплитудно-фазовых кривых Найквиста оказался удобным средством для анализа проблем устойчивости системы самолет—автопилот. Этот метод может применяться для определения характеристик существующих систем или для расчета новых систем и особенно полезен при изучении некоторых видов нелинейных систем. Результаты частотного анализа могут быть использованы для видоизменений конструкций самолета или автопилота с целью обеспечения удовлетворительной устойчивости. Однако точность указанного метода ограничивается рядом факторов, из которых основное значение имеют неточность аэродинамических данных для самолета, произвольные допущения, сделанные при выводе уравнений движения самолета, и существование нелинейностей в системе автопилота. Эти ошибки можно несколько уменьшить, если произвести вычисления, касающиеся самолета, для достаточно широких пределов изменения аэродинамических коэффициентов и при различных допущениях и если произвести тщательные испытания автопилота в достаточно широком диапазоне частот и амплитуд.

ПРИЛОЖЕНИЕ А

Сравнение частотного метода с точными математическими методами анализа устойчивости системы

Для иллюстрации положительных и отрицательных сторон частотного метода рассмотрим пример анализа устойчивости системы с одной степенью свободы. Примем следующие

¹⁾ Последний раздел статьи, посвященный обсуждению полученных результатов, при переводе опущен. — *Прим. ред.*

12. Частотный анализ устойчивости самолета с автопилотом 397

щий порядок расчета. Вначале мы получим частотные характеристики самолета и автопилота и на основании этих характеристик исследуем устойчивость системы рассмотренными выше методами. Затем для сравнения полученных результатов с точными мы составим и решим дифференциальные уравнения системы. Наконец, мы сравним частоту колебаний и коэффициент демпфирования системы, полученные на основании точного и приближенного решений.

Система выбрана, исходя из простоты получения точного решения описывающих ее дифференциальных уравнений.

Таблица 3

Обозначения	Описание	Величины
I_z	Момент инерции относительно оси z	2,9 кг·м·сек ²
$N_{\dot{\psi}}$	Производная от момента рыскания по угловой скорости рыскания	-14,5 кг·м·сек/рад
N_{ψ}	Производная от момента рыскания по углу рыскания	-0,24 кг·м/рад
N_{δ}	Производная от момента рыскания по углу отклонения рулей	-0,24 кг·м/рад
l	Постоянная времени системы управления	Переменная (сек.)
K	Передаточный коэффициент системы управления	0,6

В табл. 3 представлены величины, которые потребуются для расчета дополнительно к приведенным выше.

Уравнение

$$(I_z D^2 - N_{\dot{\psi}} D - N_{\psi}) \psi = N_{\delta} \delta \quad (A1)$$

определяет движение самолета с ограничениями по крену и скольжению, но свободного по углу рыскания.

Уравнение

$$\delta = K(1 - lD)\psi, \quad (A2)$$

использованное для автопилота, описывает безинерционную систему, в которой отклонение рулей пропорционально углу отклонения самолета, обладающую демпфированием, пропорциональным скорости отклонения самолета.

Для определения собственной частоты колебаний системы самолет — автопилот необходимо прежде всего решить характеристическое уравнение, соответствующее дифференциальному уравнению системы. Подставляя уравнение (A2) в уравнение (A1) и группируя подобные члены, получим

$$[I_z D^2 + (-N_{\dot{\psi}} + N_{\delta} K I) D - N_{\psi} - N_{\delta} K] \psi = 0. \quad (A3)$$

Как уже говорилось выше, частотные кривые были получены для случая нейтральной устойчивости. Поэтому величины $N_{\dot{\psi}}$ и $N_{\delta} I K$ были выбраны таким образом, чтобы полное демпфирование системы равнялось нулю. Это условие будет выполнено, если положить $N_{\dot{\psi}} = N_{\delta} I K$.

Учитывая это, уравнение (A3) можно записать в виде

$$[I_z D^2 + (-N_{\psi} - N_{\delta} K)] \psi = 0. \quad (A4)$$

Мы получили уравнение незатухающих колебаний, причем собственная частота системы равна

$$\omega_n = \sqrt{\frac{N_{\psi} + N_{\delta} K}{-I_z}}.$$

Теперь покажем, как та же самая величина ω_n может быть получена частотным методом. Первый шаг заключается в вычислении частотных характеристик самолета. Подробно метод вычисления частотных характеристик самолета изложен в работах [3] и [4], и поэтому мы здесь приведем лишь окончательный результат.

Частотная характеристика самолета, уравнение движения которого имеет вид (A1), может быть записана в виде

$$\left(\frac{\psi}{\delta}\right)_a = \frac{-N_{\delta}(I_z \omega^2 + N_{\psi})}{(I_z \omega^2 + N_{\psi})^2 + N_{\dot{\psi}}^2 \omega^2} + j \frac{N_{\delta} N_{\dot{\psi}} \omega}{(I_z \omega^2 + N_{\psi})^2 + N_{\dot{\psi}}^2 \omega^2},$$

т. е. в виде $A + jB$.

Частотная характеристика самолета определяется вектором, амплитуда и фаза которого зависят от частоты ω входного возмущения, являющегося гармонической функцией времени. Для амплитудно-частотной характеристики можно записать

$$R_a = \sqrt{A^2 + B^2}$$

12. Частотный анализ устойчивости самолета с автопилотом 399

или более подробно

$$R_a = \frac{N_\delta}{\sqrt{(I_z \omega^2 + N_\psi)^2 + N_\psi^2 \omega^2}} \quad (A5)$$

и для фазы

$$\epsilon_a = \arctg \frac{B}{A},$$

или

$$\epsilon_a = \arctg \frac{-N_\psi \omega}{I_z \omega^2 + N_\psi}. \quad (A6)$$

Эти кривые показаны на фиг. 8.

Частотная характеристика автопилота получается аналогичным образом, поэтому можно сразу записать на основании уравнения для автопилота

$$\delta = K(1 - ID)\psi,$$

$$\epsilon_p = \arctg l\omega, \quad (A7)$$

$$R_p = \frac{1}{K\sqrt{1 + l^2\omega^2}}. \quad (A8)$$

Эти кривые также представлены на фиг. 8.

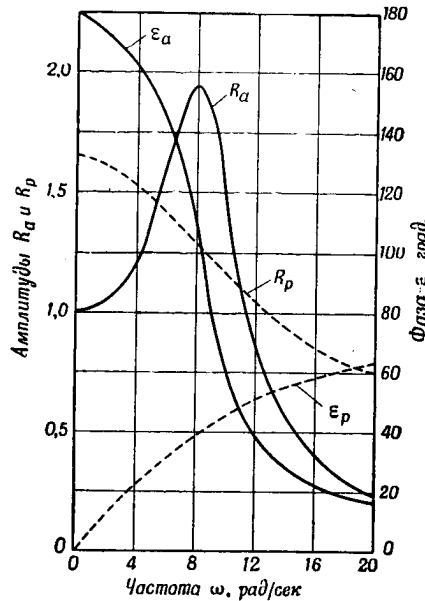
Условие возможности колебаний заключается в том, чтобы фазовые частотные характеристики самолета и автопилота пересеклись. При этом частота колебаний определяется из условия

$$\epsilon_a = \epsilon_p.$$

Для ее точного определения можно использовать уравнения (A6)

и (A7) и условие $N_\psi = IN_\delta K$. Тогда получим

$$\arctg l\omega = \arctg \frac{N_\psi \omega}{-I_z \omega^2 - N_\psi}, \quad \epsilon_a = \epsilon_p,$$



Фиг. 8. Частотные характеристики самолета, имеющего одну степень свободы — по углу рыскания.

Автопилот описывается уравнением $\delta = K(1 - ID)\psi$; $K=0,6$; $l=0,1$. Кривые соответствуют случаю нейтральной системы (границе устойчивости).

откуда для частоты собственных колебаний можем записать следующее выражение

$$\omega_c = \sqrt{\frac{N_\psi I + N_\psi}{-I_z}} = \sqrt{\frac{N_\psi + N_\delta K}{-I_z}}, \quad (\text{A9})$$

которое совпадает с результатом, полученным из дифференциального уравнения системы.

Теперь на основании уравнений (A5), (A8) и (A9) можно сравнить амплитудные значения при $\omega = \omega_c$. Нетрудно проверить, что после подстановки уравнения (A9) уравнение (A8) будет тождественно уравнению (A3) и, следовательно, условие нейтральной устойчивости будет выполнено, что и показано на фиг. 8.

Определение частоты демпфированных колебаний. Как уже указывалось ранее, точка пересечения фазовых частотных характеристик уже не будет определять частоту собственных колебаний системы в том случае, когда мы имеем устойчивую или неустойчивую систему, т. е. когда амплитудные и фазовые характеристики самолета и автопилота пересекаются при разных частотах. Однако собственная частота системы в этом случае может быть легко определена, если мы имеем частотные характеристики системы, вычисленные для заданного коэффициента демпфирования системы. В этом случае возможно следующее решение. Так же как и раньше, вначале пишутся уравнения движения системы, в которых условие нейтральной устойчивости заменяется условием $N_\psi > IN_\delta K$, соответствующим случаю устойчивого движения системы. Тогда уравнение системы можно записать в виде

$$(I_z D^2 - N_\psi D - N_\delta K) \psi = N_\delta K (1 - ID) \psi,$$

и частота демпфированных колебаний определится формулой

$$\omega_{nd} = \sqrt{\frac{-(-N_\psi + N_\delta K I)^2}{4I_z^2} - \left(\frac{N_\psi + N_\delta K}{I_z}\right)}.$$

Для проверки этого результата необходимо, как и раньше, получить частотные характеристики системы. Однако при замене оператора D демпфирующий член можно исключить.

12. Частотный анализ устойчивости самолета с автопилотом 401

Действительно, имеем

$$\left(\frac{\psi}{\delta}\right)_a = \frac{N_\delta}{I_z D^2 - N_\psi D - N_\psi},$$

причем вместо замены $D = j\omega$ мы используем подстановку $D = -u + j\omega$, где u равно сумме демпфирующих членов (т. е. коэффициенту при D , деленному на удвоенный коэффициент при D^2) в уравнении движения системы самолет — автопилот. Тогда частотная характеристика самолета может быть записана в виде

$$\left(\frac{\psi}{\delta}\right)_a = \frac{N_\delta}{I_z (-u + j\omega)^2 - N_\psi (-u + j\omega) - N_\psi},$$

откуда получаем амплитудную и фазовую частотные характеристики самолета:

$$R_a = \frac{N_\delta \sqrt{(I_z u^2 - I_z \omega^2 + N_\psi u - N_\psi)^2 + (2I_z u\omega + N_\psi \omega)^2}}{(I_z u^2 - I_z \omega^2 + N_\psi u - N_\psi)^2 + (2I_z u\omega + N_\psi \omega)^2},$$

$$\varepsilon_a = \arctg \frac{2I_z u\omega + N_\psi \omega}{I_z u^2 - I_z \omega^2 + N_\psi u + N_\psi}.$$

Совершенно аналогично частотные характеристики автопилота можно записать в виде

$$\left(\frac{\psi}{\delta}\right)_p = \frac{1 + lu}{K[(1 + lu)^2 + l^2 \omega^2]} + \frac{j l \omega}{K[(1 + lu)^2 + l^2 \omega^2]},$$

$$\varepsilon = \arctg \frac{l \omega}{1 + lu}.$$

Частотные характеристики самолета и автопилота для двух значений величины l представлены на фиг. 9 и 10. Как уже указывалось, собственная частота демпфированных колебаний определяется из условия $\varepsilon_p = \varepsilon_a$.

Таким образом, имеем

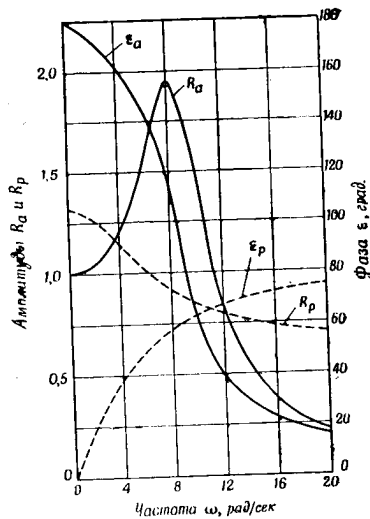
$$\arctg \frac{l \omega}{1 + lu} = \arctg \frac{2I_z u\omega + N_\psi \omega}{I_z u^2 - I_z \omega^2 + N_\psi u - N_\psi},$$

откуда находим

$$\omega_c = \sqrt{\frac{-(-N_\psi + lKN_\delta)^2}{4I_z^2} - \left(\frac{N_\psi + N_\delta K}{I_z}\right)}.$$

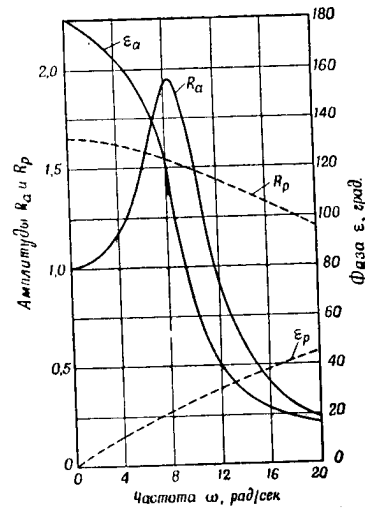
Из этого выражения видно, что ω_c совпадает с частотой ω_c , найденной на основании решения дифференциального уравнения системы.

На фиг. 8—10 представлены частотные характеристики самолета и автопилота для трех значений параметра l , а на фиг. 11 — соответствующие кривые Найквиста. Мы видим, что частоты, которые соответствуют точкам пересечения



Фиг. 9. Частотные характеристики самолета, имеющего одну степень свободы — по углу рыскания.

Автопилот описывается уравнением $\delta = K(1 - lD)\psi$; $K = 0,6$; $l = 0,2$. Кривые соответствуют случаю неустойчивой системы.



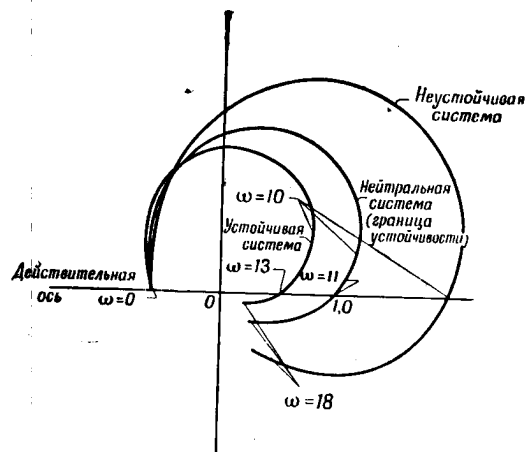
Фиг. 10. Частотные характеристики самолета, имеющего одну степень свободы — по углу рыскания.

Автопилот описывается уравнением $\delta = K(1 - lD)\psi$; $K = 0,6$; $l = 0,05$. Кривые соответствуют случаю устойчивой системы.

кривых Найквиста с действительной осью, совпадают с частотами, соответствующими точкам пересечения фазовых характеристик самолета и автопилота (см. фиг. 8—10). Следовательно, если использовать предложенный критерий устойчивости, то легко видеть, что оба метода (метод кривых Найквиста и метод частотных характеристик) совпадают.

12. Частотный анализ устойчивости самолета с автопилотом 403

Как уже указывалось, обычно в литературе при анализе устойчивости систем рассматривается точка $(-1,0)$, а не точка $(1,0)$, использованная в настоящей работе. Это различие объясняется выбором знаков в уравнениях самолета и автопилота. Обычно в электротехнике принято выражать коррекцию положительной ошибки отрицательной величиной.



Фиг. 11.5 Кривые Найквиста для случая устойчивой, нейтральной и неустойчивой системы самолет — автопилот.

Однако в соответствии с системой знаков, принятой в НАСА, коррекция положительной ошибки (соответствующей движению вверх) производится введением положительной величины (соответствующей отклонению руля вниз)¹⁾. Это обстоятельство и приводит к сдвигу критической точки на 180° . Однако понятно, что кривые Найквиста одинаково справедливы в обоих случаях и метод их использования в обоих случаях остается тем же самым.

¹⁾ Формулировка частотного критерия устойчивости зависит от принятой системы координат и в известной мере является произвольной. Однако в нашей литературе принято считать, что положительное отклонение системы компенсируется отрицательным регулирующим воздействием; это приводит к тому, что критической точкой является точка с координатами $(-1,0)$. — Прим. ред.

П Р И Л О Ж Е Н И Е Б
У р а в н е н и я д в и ж е н и я

При составлении уравнений продольного и бокового движения самолета были приняты следующие допущения:

1) отклонения самолета от установившейся траектории движения малы, т. е. приращение переменной значительно меньше единицы;

2) величинами приращений второго и более высокого порядка можно пренебречь;

3) суммарный эффект от нескольких воздействий равен сумме эффектов, вычисленных от каждого воздействия (принцип суперпозиции);

4) изменение аэродинамических сил и моментов, связанное с отклонением самолета от траектории установившегося движения, прямо пропорционально отклонению;

5) боковое и продольное движения самолета независимы друг от друга.

Уравнения продольного движения

Движение с тремя степенями свободы. При составлении этих уравнений были использованы следующие хорошо известные обозначения для параметров движения самолета: u — относительное приращение скорости самолета; θ — приращение угла касательной по отношению к горизонтальной оси и α — угол атаки. Обозначения производных C_D , C_L и C_m совпадают с общепринятыми¹⁾. Кроме того, могут быть также использованы дополнительные коэффициенты, необходимые для анализа беспилотного движения самолета. Эти коэффициенты могут быть использованы в уравнении при условии, что они выражены в тех же единицах, что и все остальные производные.

Уравнение силы лобового сопротивления

$$\begin{aligned} \frac{mV}{qS} Du = & -2(C_D + C_{D_u})u - (C_{D_\alpha} + C_{D_{D\alpha}} D \frac{c}{2V})\alpha - \\ & - (C_{D_{\delta_e}} + C_{D_{D\delta_e}} \frac{2}{2V} D)\delta_e - C_{D_{D\theta}} \frac{c}{2V} D\theta - \\ & - T_C \sin \alpha_0 \alpha - \frac{W}{qS} \cos \gamma_0 (\theta - \alpha). \end{aligned}$$

¹⁾ Обозначения основных координат и аэродинамических коэффициентов отличаются от принятых у нас. — *Прим. ред.*

12. Частотный анализ устойчивости самолета с автопилотом 405

Уравнение подъемной силы

$$\begin{aligned} \frac{mV}{qS} D\theta - \frac{mV}{qS} D\alpha = & (2C_L + C_{L_u}) u + \\ & + (C_{L_\alpha} + C_{L_{D\alpha}} \frac{c}{2V} D) \alpha + (C_{L_{\delta_e}} + C_{L_{D\delta_e}} \frac{c}{2V} D) \delta_e + \\ & + C_{LD\theta} \frac{c}{2V} D\theta + T_{C'} \cos \alpha_0 \alpha + \frac{W}{qS} \sin \gamma_0 (\theta - \alpha). \end{aligned}$$

Уравнение моментов

$$\begin{aligned} \frac{I_y}{qSc} D^2\theta = & 2C_{m_u} u + (C_{m_\alpha} + C_{m_{D\alpha}} \frac{c}{2V} D) \alpha + C_{m_u} u + \\ & + C_{m_{D\theta}} \frac{c}{2V} D\theta + (C_{m_{\delta_e}} + C_{m_{D\delta_e}} \frac{c}{2V} D) \delta_e + T_{C'} \frac{y}{c}. \end{aligned}$$

В этих уравнениях были использованы следующие коэффициенты:

m — масса самолета, т. е. м.;
 q — динамическое давление, кг/м²;
 ρ — плотность воздуха, т. е. м./м³;
 S — площадь крыла, м²;
 c — средняя аэродинамическая хорда, м;
 D — оператор дифференцирования d/dt ;
 V — воздушная скорость, м/сек;

$$u = \frac{\Delta V}{V};$$

C_D — коэффициент лобового сопротивления =
 = (сила лобового сопротивления) / qS ;

C_L — коэффициент подъемной силы =
 = (подъемная сила) / qS ;

C_m — коэффициент продольного момента =
 = (продольный момент) / qSc ;

$T_{C'}$ — коэффициент давления = (давление) / qS ;

α — угол атаки, рад;

θ — угол между осью самолета и горизонтом, рад;

γ — угол между касательной к направлению движения центра тяжести и горизонтом, рад; индекс 0

означает установившееся значение; γ без индекса означает отклонение от установившегося значения:

$$q = \frac{d\theta}{dt};$$

δ — угол отклонения рулей, рад;

u — кратчайшее расстояние от линии давления до центра тяжести, м;

$$C_{D\alpha}; C_{L\alpha}; C_{m\alpha} = \frac{\partial C_D}{\partial \alpha}; \frac{\partial C_L}{\partial \alpha}; \frac{\partial C_m}{\partial \alpha};$$

$$C_{D_{D\alpha}}; C_{L_{D\alpha}}; C_{m_{D\alpha}} = \frac{\partial C_D}{\partial (D\alpha c/2V)}; \frac{\partial C_L}{\partial (D\alpha c/2V)}; \frac{\partial C_m}{\partial (D\alpha c/2V)};$$

$$C_{Du}; C_{Lu}; C_{mu} = \frac{\partial C_D}{\partial u}; \frac{\partial C_L}{\partial u}; \frac{\partial C_m}{\partial u};$$

$$C_{D\theta}; C_{L\theta}; C_{m\theta} = \frac{\partial C_D}{\partial \theta}; \frac{\partial C_L}{\partial \theta}; \frac{\partial C_m}{\partial \theta};$$

$$C_{D_{D\theta}}; C_{L_{D\theta}}; C_{m_{D\theta}} = \frac{\partial C_D}{\partial (qc/2V)}; \frac{\partial C_L}{\partial (qc/2V)}; \frac{\partial C_m}{\partial (qc/2V)};$$

$$C_{D\delta}; C_{L\delta}; C_{m\delta} = \frac{\partial C_D}{\partial \delta}; \frac{\partial C_L}{\partial \delta}; \frac{\partial C_m}{\partial \delta};$$

$$C_{L_{D\delta}}; C_{m_{D\delta}} = \frac{\partial C_L}{\partial (D\delta c/2V)}; \frac{\partial C_m}{\partial (D\delta c/2V)}.$$

Уравнения системы с двумя степенями свободы. Движение, которое описывается тремя полученными уравнениями, состоит из двух движений: слабодемпфированного фугоидного движения с большим периодом, в котором угол атаки остается почти постоянным и меняются лишь скорость и угол γ , и сильнодемпфированного движения с малым периодом, в котором воздушная скорость остается постоянной, а меняются углы θ и α .

Поскольку автопилот реагирует лишь на изменение углов θ и α , то он будет действовать лишь в режиме короткопериодического движения самолета. Поэтому первое уравнение сил сопротивления при анализе системы можно опустить и рассматривать лишь уравнения относительно углов θ и α . Эти уравнения могут быть записаны в следующем виде.

12. Частотный анализ устойчивости самолета с автопилотом 407

Уравнение подъемной силы

$$\begin{aligned} \frac{mV}{qS} D\theta - \frac{mV}{qS} D\alpha - \left(C_{L\alpha} + C_{LD\alpha} \frac{c}{2V} D \right) \alpha - \\ - \left(C_{LD\theta} \frac{c}{2V} D \right) \theta - T_c \cos \alpha_0 \alpha - \frac{W}{qS} \sin \gamma_0 (\theta - \alpha) = \\ = \left(C_{L\delta_e} + C_{LD\delta_e} \frac{c}{2V} D \right) \delta_e. \end{aligned}$$

Уравнение продольных моментов

$$\begin{aligned} \frac{I_y}{qSc} D^2\theta - \left(C_{m\alpha} + C_{mD\alpha} \frac{c}{2V} D \right) \alpha - C_{mD\theta} \frac{c}{2V} D\theta - \\ - T_c \frac{y}{c} = \left(C_{m\delta_e} + C_{mD\delta_e} \frac{c}{2V} D \right) \delta_e. \end{aligned}$$

Уравнения бокового движения

Рыскание и боковое скольжение. Полные уравнения были приведены в начале работы. Для некоторых типов самолетов эти уравнения могут быть упрощены и сведены к двум или даже к одному уравнению, которые соответствуют системам с двумя и одной степенью свободы. Так, если самолет обладает симметрией относительно плоскостей (x, z) и (x, y) , то обычно можно считать, что движение крена не зависит от двух других движений. Это предположение можно подтвердить продувкой самолета в аэродинамической трубе и вычислением коэффициента $C_{I\beta}$, который в этом случае должен быть равен нулю. Таким образом, если не учитывать вращения самолета вокруг продольной оси, то можно записать следующие два уравнения:

Уравнение для угла рыскания

$$\begin{aligned} \frac{mk_z^2}{qSb} D^2\psi - C_{nr} \frac{b}{2V} D\psi - C_{n\beta} \beta = \\ = C_{n\delta_a} \delta_a + \left(C_{n\delta_r} + C_{nD\delta_r} \frac{b}{2V} D \right) \delta_r. \end{aligned}$$

Уравнение бокового скольжения

$$\left(\frac{mVD}{qS} - \frac{W}{qS} \sin \gamma_0 - C_{y_r} \frac{b}{2V} D\right) \psi + \left(\frac{mVD}{qS} - C_{y_\beta}\right) \beta = C_{y_{\delta_a}} \delta_a + \left(C_{y_{\delta_r}} + C_{y_{D\delta_r}} \frac{b}{2V} D\right) \delta_r.$$

Следует иметь в виду, что при сделанных допущениях справедливо лишь предположение о независимости рыскания и бокового скольжения от крена самолета. Однако обратное утверждение было бы совершенно неверно, поскольку уравнение для угла крена никак нельзя считать независимым от движения рыскания. Следовательно, допущение о независимости движений справедливо только в том случае, если имеется достаточно хорошая стабилизация самолета относительно продольной оси, которая ограничивает скорость крена до величины, практически независимой от момента рыскания.

Уравнение для угла крена. Основываясь на сделанных предположениях относительно симметрии самолета, уравнение для угла крена можно записать в следующей форме:

$$\left(\frac{mk_x^2 D^2}{qSb} - C_{l_p} \frac{b}{2V} D\right) \varphi = \left(C_{l_{\delta_a}} + C_{l_{D\delta_a}} \frac{b}{2V} D\right) \delta_a + C_{l_{D\delta_r}} \delta_r.$$

ЛИТЕРАТУРА

1. Greenberg H., „Frequency-Response Method for Determination of Dynamic Stability Characteristics of Airplanes with Automatic Control“, NACA, T. N. № 1229, March, 1947.
2. Hall A. C., The Analysis and Synthesis of Linear Servomechanisms, Mass. Inst. Technol., Cambridge, Mass., 1943.
3. Johnson, Mathematical and Physical Principles of Engineering, Analysis, New York, 1944, Chap. VII.
4. Карман Т. и Био М., Математические методы в инженерном деле, Гостехиздат, 1947, гл. IX.
5. Zimmerman Ch. H., „An Analysis of Lateral Stability in Power-off Flight with Charts for Use in Design“, NACA, Rep. № 589, 1937.
6. Jones T. M., Airplane Stability, vol. V of Aeronautical Theory, W. F. Durand, Ed., Berlin, 1934.
7. Sternfield L., „Effect of Product of Inertia on Lateral Stability“, NACA, T. N. № 1193, March, 1947.

12. Частотный анализ устойчивости самолета с автопилотом 409

8. Greenberg H., Sternfield L., „A Theoretical Investigation of Longitudinal Stability of Airplanes with Free Controls Including Effect of Friction in Control System“, NACA, Wartime Rep. L-430, Febr., 1944.
9. Soulé H. A., Miller M. P., „The Experimental Determination of the Moments of Inertia of Airplanes“, NACA, T. R. № 467, 1933.
10. Cotter W. E., Jr., „Summary and Analysis of Data on Damping in Yaw and Pitch for a Number of Airplane Models“, NACA, T. N. № 1080, 1946.
11. Pearson H. A., Jones R. T., „Theoretical Stability and Control Characteristics of Wings with Various Amounts of Taper and Twist“, NACA, Rep. № 635, 1938.
12. Cohen D., Jones R. T., „An Analysis of the Stability of an Airplane with Free Controls“, NACA, T. R. № 709, 1941.
13. Queijo M. J., Jaquet B. M., „Effect of Geometric Dihedral on the Low Speed Rolling Derivatives of Swept Wings“, NACA, T. N. № 1732, Oct. 1948.
14. Теория следящих систем, под ред. Х. Джеймса, Н. Никольса и Р. Филлипса, Издательство иностранной литературы, 1953.
15. Бодэ Г., Теория цепей и проектирование усилителей с обратной связью, Издательство иностранной литературы, 1948.

**IV. ЭКСПЕРИМЕНТАЛЬНЫЕ
МЕТОДЫ
ОПРЕДЕЛЕНИЯ
ЧАСТОТНЫХ
ХАРАКТЕРИСТИК**

13. ЭКСПЕРИМЕНТАЛЬНЫЕ МЕТОДЫ ОПРЕДЕЛЕНИЯ
ЧАСТОТНЫХ ХАРАКТЕРИСТИК САМОЛЕТА В ПОЛЕТЕ ¹⁾

Г. Смит, В. Триплетт

ОБОЗНАЧЕНИЯ

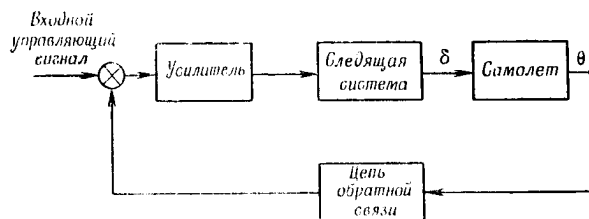
- I_x, I_y, I_z — моменты инерции относительно осей x, y и z ,
 $\text{кг} \cdot \text{м} \cdot \text{сек}^2$;
 I_{xz} — центробежный момент инерции, $\text{кг} \cdot \text{м} \cdot \text{сек}^2$;
 K — коэффициент усиления;
 L, M, N — моменты крена, тангажа и рыскания (см. фиг. 2),
 $\text{кг} \cdot \text{м}$;
 S — площадь крыльев, м^2 ;
 V — скорость, $\text{м}/\text{сек}$;
 X, Y, Z — силы, направленные вдоль осей, показанных на
фиг. 2, кг ;
 c — средняя аэродинамическая хорда, м ;
 m — масса, $\text{кг} \cdot \text{сек}^2/\text{м}$;
 q — угловая скорость тангажа, $\text{рад}/\text{сек}$;
 s — оператор Лапласа;
 t — время, сек.;
 $j = \sqrt{-1}$;
 α — угол атаки, рад ;
 δ — отклонение руля высоты, рад ;
 θ — угол тангажа, рад ;
 ρ — плотность воздуха, $\text{кг} \cdot \text{сек}^2/\text{м}^4$;
 ζ — коэффициент демпфирования;
 φ — фазовый угол, град ;
 ω — угловая частота, $\text{рад}/\text{сек}$;
 ω_n — собственная частота при отсутствии демпфирования,
 $\text{рад}/\text{сек}$;
 τ — постоянная времени, сек.

Точка, стоящая над величиной, обозначает производную по времени.

¹⁾ Smith G. A., Triplett W. C., *Trans. ASME*, 76, № 8, 1383—1390 (1954).

Введение

Необычайно широкий диапазон скоростей и высот, на которых летают современные самолеты, требует решения сложных задач, касающихся устойчивости и управления. Трудности, с которыми пришлось столкнуться пилоту при разрешении этих задач, в значительной степени обусловили то большое внимание, какое уделяется в настоящее время автоматическим самолетам-перехватчикам и управляемым снарядам, точное управление маневрами которых во время боевых действий зависит от поведения автопилотов. Для



Фиг. 1. Упрощенная блок-схема системы самолет — автопилот.

анализа этих сложных систем с учетом динамики самолета успешно применяются методы частотных характеристик.

Упрощенная скелетная схема типичного контура управления самолет — автопилот показана на фиг. 1. Для анализа такой замкнутой системы необходимо знать динамическое поведение каждого элемента системы или, более точно, передаточные функции, определяющие отношение выходного сигнала каждого элемента к входному сигналу. Передаточные функции самолета, например, определяют характер реакций самолета на движение рулей. К сожалению, передаточные функции самолета заметно изменяются при изменении высоты и скорости самолета, и определение характера этих изменений сопряжено с большими трудностями. Эта особенность является отличительной для задач, касающихся систем самолет — автопилот, по сравнению с задачей исследования обычной следящей системы и означает, что для удовлетворительного поведения всей замкнутой системы при изменении

13. Определение частотных характеристик самолета в полете 415

условий полета необходимо введение цепей для изменения усиления и стабилизации. Задача, кроме того, осложняется высокой точностью, требуемой от автопилотов, управляющих самолетом-истребителем во время боевых маневров. Эти условия противоположны существовавшим перед второй мировой войной, когда автопилоты предназначались главным образом в помощь пилоту на самолете, а пределы изменения скорости и высоты полета самолета были гораздо более ограниченны.

Задача аналогичного анализа существует и для некоторых пилотируемых самолетов, где необходимо изменять динамические характеристики самолета с помощью следящих устройств, таких, которые, например, демпфируют движения рыскания и тангажа так, чтобы летчик мог надлежащим образом управлять самолетом во всех условиях полета. Эта задача возникает в основном из-за малого демпфирования и высокой собственной частоты, неизбежных при полете с высокой скоростью на большой высоте. Разработка системы самолет — демпфер, приемлемой для пилотируемых полетов, по существу тождественна созданию замкнутой системы, аналогичной системе самолет — автопилот, рассмотренной выше, за исключением того, что сам пилот становится в этом случае звеном замкнутого контура. Из-за отсутствия удовлетворительной передаточной функции для описания реакции человека задача сводится к надлежащему подбору динамических характеристик демпфера с тем, чтобы с учетом поведения самолета результирующая система самолет — демпфер имела передаточную функцию, приемлемую для пилота согласно имеющимся экспериментальным данным.

Отсюда следует, что при решении задач, касающихся устойчивости и управления самолетом, необходимо иметь данные о частотной характеристике самолета. В работах [1] и [2] более подробно рассмотрены методы определения частотных характеристик самолетов и обсуждены некоторые достижения в этой области.

В настоящей статье главным образом описываются методы определения динамических характеристик самолета. Эти данные в форме передаточной функции могут вычисляться путем подстановки известных значений физических параметров в дифференциальные уравнения, описывающие движения самолета. В дополнение к обычным постоянным, таким, как масса и момент инерции, эти параметры состоят из членов,

устанавливающих зависимость между аэродинамическими силами и моментами, с одной стороны, и движениями самолета — с другой. Эти аэродинамические величины, выраженные в виде безразмерных производных коэффициентов, определяются на основе теоретического изучения и результатов испытания в аэродинамической трубе. В некоторых случаях это может быть выполнено с достаточной степенью точности для надежного предсказания основных свойств передаточной функции самолета. Однако обычно рекомендуется пользоваться результатами экспериментальных испытаний в условиях полета для получения достаточно полных и точных данных, необходимых для анализа. Это особенно справедливо для полетов со сверхзвуковыми скоростями, когда большие и незакономерные изменения этих параметров при изменении числа M не могут быть найдены с достаточной точностью, исходя из теории или из данных продувки в аэродинамической трубе.

Кроме того, имеются случаи, для которых обычно применяемые уравнения „твердого тела“ не определяют полностью движений самолета. Примером могут служить явления, связанные с упругой деформацией элементов конструкций самолета в условиях полета. Эти явления сопровождаются изменениями устойчивости траектории полета, обусловленными изгибанием и скручиванием фюзеляжа самолета под действием инерции и аэродинамических нагрузок. Возникновение дополнительных пиков на экспериментальной амплитудной частотной характеристике на частотах, превышающих величины, найденные для колебаний самолета, рассматриваемого как твердое тело, обычно показывает наличие и значимость колебаний конструкций, возникающих в системе. Движения, создаваемые в результате упругих деформаций, могут иметь важное значение при выборе подходящего автопилота и размещении его чувствительных элементов [3].

Зависимость поведения самолета от изменения числа M и упругих деформаций определяет те практические трудности, которые исключают возможность чисто теоретического решения задачи и требуют разработки методов летных испытаний для получения правильных передаточных функций самолета. Летные испытания имеют также дополнительное назначение, позволяя сравнивать экспериментальные и теоретически предсказанные значения, и способствуют, таким

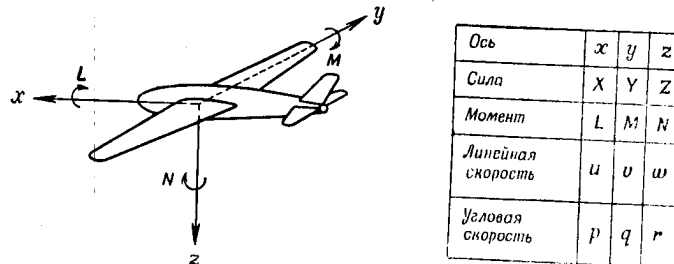
13. Определение частотных характеристик самолета в полете 417

образом, разработке более совершенных методов теоретического предсказания поведения системы.

После краткого обсуждения аэродинамической теории, положенной в основу уравнений движения самолета, будет рассмотрено определение частотной характеристики самолета на основании анализа летных данных.

Вывод теоретических передаточных функций

Самолет в условиях полета можно рассматривать как динамическую систему, имеющую шесть степеней свободы. Эта система может перемещаться и вращаться относительно любой из трех осей x , y и z , показанных на фиг. 2. Таким



Фиг. 2. Обозначение осей и определение основных величин.

образом, существует шесть уравнений, определяющих движения самолета, которые позволяют вычислить реакции на различные типы возмущающих сил.

Приравнявая силы и моменты инерции (левая часть) к аэродинамическим силам и моментам (правая часть), получаем следующие уравнения:

$$m(\dot{u} + qw - rv) = X, \quad (1)$$

$$m(\dot{v} + ru - pw) = Y, \quad (2)$$

$$m(\dot{w} + pv - qu) = Z, \quad (3)$$

$$I_x \dot{p} + (I_z - I_y)qr - I_{xz}(\dot{r} + pq) = L, \quad (4)$$

$$I_y \dot{q} + (I_x - I_z)rp - I_{xz}(r^2 - p^2) = M, \quad (5)$$

$$I_z \dot{r} + (I_y - I_x)pq - I_{xz}(\dot{p} - qr) = N. \quad (6)$$

Полный вывод этих уравнений можно найти в литературе (см., например, [4, 5]), и они лишь бегло рассмотрены в данной статье.

Очевидно, что система, описываемая приведенными выше уравнениями, является нелинейной и очень сложной. Однако по счастливому стечению обстоятельств движения самолета часто ограничены таким образом, что система уравнений для большинства практических задач распадается на две независимые линейные системы, каждая из которых имеет три степени свободы. Сначала предположим, что возмущающая сила приложена таким образом, что возмущение самолета происходит лишь в плоскости симметрии (плоскость (x, z)). Таким возмущением могут быть отклонение руля высоты, изменение мощности мотора или какое-либо атмосферное явление. Результирующие движения обычно называются продольными движениями самолета или движениями самолета по тангажу и состоят из ускорений вдоль осей x и z и вращений относительно оси y . Вследствие симметрии движения в любой другой плоскости отсутствуют.

Если же возмущение самолета осуществляется в результате перемещения элерона или руля поворота, то самолет будет поворачиваться относительно осей x и z (продольная и вертикальная оси) и перемещаться вдоль оси y (боковое скольжение). Движения этого типа обычно называются боковыми. Из уравнений, описывающих всю систему, следует, что боковые движения вызывают вторичные продольные движения в результате инерционной связи. Однако если принять, что движение системы ограничено небольшими перемещениями относительно горизонтальной плоскости, то члены, вызывающие возникновение этой связи, сравнительно малы и ими часто можно пренебречь. Тогда упрощенные уравнения для двух систем запишутся в следующем виде:

Система уравнений продольного движения

$$m(\dot{u} + qw) = X, \quad (7)$$

$$m(\dot{w} - qu) = Z, \quad (8)$$

$$I_y \dot{q} = M. \quad (9)$$

13. Определение частотных характеристик самолета в полете 419

Система уравнений бокового движения

$$m(\dot{v} + ru) = Y, \quad (10)$$

$$I_x \dot{p} - I_{xz} \dot{r} = L, \quad (11)$$

$$I_z \dot{r} - I_{xz} \dot{p} = N. \quad (12)$$

Эти уравнения определяют полные перемещения самолета при отсутствии взаимной связи между продольными и боковыми перемещениями.

Переходя к обсуждению продольных реакций, возникающих в результате отклонения руля высоты, можно сделать дальнейшее упрощение. Ограничив исследование небольшими возмущениями, нетрудно установить, что изменение скорости полета, сопровождающее отклонение средства управления, обычно очень мало и результирующая сила X , действующая на самолет, по существу равна нулю при отсутствии изменения тяги мотора. Таким образом, система с продольными реакциями преобразуется в систему с двумя степенями свободы, для которой переменная u может быть заменена постоянной скоростью V . Для упрощения дальнейшего исследования нормальная скорость w может быть записана как $V\alpha$, где α — угол атаки самолета, т. е. угол в плоскости симметрии между относительным вектором направления потока и осью самолета. Тогда уравнения, определяющие продольные реакции самолета, преобразуются к виду

$$mV(\dot{\alpha} - q) = Z, \quad (13)$$

$$I_y \dot{q} = M. \quad (14)$$

Для решения этих уравнений относительно координат q , α или нормального ускорения $V(\dot{\alpha} - q)$ необходимо выразить аэродинамические силы и моменты M и Z через величины α , q и отклонения руля высоты δ . Для небольших перемещений Z и M в основном являются линейными функциями α , $\dot{\alpha}$, q и δ и уравнения (13) и (14) могут быть записаны в виде

$$mV(\dot{\alpha} - q) = \frac{\partial Z}{\partial \alpha} \alpha + \frac{\partial Z}{\partial \delta} \delta, \quad (15)$$

$$I_y \dot{q} = \frac{\partial M}{\partial \alpha} \alpha + \frac{\partial M}{\partial \dot{\alpha}} \dot{\alpha} + \frac{\partial M}{\partial q} q + \frac{\partial M}{\partial \delta} \delta. \quad (16)$$

Согласно сделанному допущению о незначительной величине перемещений, частные производные являются постоянными величинами для данной скорости и высоты полета. В этих уравнениях α и δ являются приращениями углов и определяются как разность между мгновенными значениями и значениями, полученными для установившегося полета.

Тогда мы получаем два линейных дифференциальных уравнения с постоянными коэффициентами, которые можно совместно решить относительно α и q , выразив эти величины с помощью δ . Для придания уравнениям более простой формы уравнение силы делится на произведение mV , а уравнение момента на I_y ; введя соответствующие сокращенные обозначения, можно привести уравнения к следующему виду:

$$\dot{\alpha} - q = Z_\alpha \alpha + Z_\delta \delta, \quad (17)$$

$$\dot{q} = M_\alpha \alpha + M_\alpha \dot{\alpha} + M_q q + M_\delta \delta. \quad (18)$$

Исключив α из уравнений (17) и (18), получаем

$$\ddot{q} + b\dot{q} + kq = C_0 \delta + C_1 \dot{\delta}, \quad (19)$$

где

$$b = -Z_\alpha - M_q - M_\alpha, \quad (20)$$

$$k = Z_\alpha M_q - M_\alpha, \quad (21)$$

$$C_0 = Z_\delta M_\alpha - Z_\alpha M_\delta, \quad (22)$$

$$C_1 = Z_\delta M_\alpha + M_\delta. \quad (23)$$

Очевидно, что левая часть уравнения (19) имеет такую же форму, как и для хорошо известной системы, состоящей из пружины, массы и демпфера (или электрического контура $R-L-C$). Однако это уравнение можно выразить в более удобной форме путем применения преобразования Лапласа при нулевых начальных условиях и алгебраического решения уравнения относительно отношения преобразованных величин q и δ . Таким образом мы получим:

$$\frac{q}{\delta} = \frac{C_0 + C_1 s}{s^2 + bs + k}. \quad (24)$$

Выражение, стоящее справа, является передаточной функцией, которая определяет в операторной форме реакцию угловой скорости самолета по тангажу на отклонение руля высоты. При подстановке вместо s величины $j\omega$ можно непо-

13. Определение частотных характеристик самолета в полете 421

средственно определить частотную характеристику или с помощью обратного преобразования Лапласа вычислить время реакции для любого движения руля высоты. Существует аналогичная передаточная функция для α/δ , которая имеет такой же знаменатель, но отличные коэффициенты в числителе.

Вид передаточной функции позволяет установить, что система является колебательной и ее собственная частота и коэффициент демпфирования зависят от массы, инерции и аэродинамических характеристик самолета. Передаточная функция, определяемая уравнением (24), может быть записана в другой более обычной форме

$$\frac{q}{\delta} = \frac{K(1 + \tau s)}{1 + (2\zeta/\omega_n) s + s^2/\omega_n^2} \quad (25)$$

Колебательное движение, определенное системой уравнений продольного движения с двумя степенями свободы, называется „короткопериодическим движением“ и приобретает более существенное значение при увеличении скорости и высоты полета вследствие соответствующего увеличения собственной частоты и уменьшения коэффициента демпфирования. Для большинства самолетов наиболее типичные величины собственных частот лежат в пределах от 2 до 8 рад/сек при значениях коэффициента демпфирования не выше 0,2.

Если учесть при выводе третью степень свободы (уравнение для силы X), то передаточная функция для q/δ будет выражаться уравнением четвертого порядка и будет определять, помимо короткопериодических колебаний, колебания с большим периодом, называемые „фугоидными“. Несмотря на очень незначительное демпфирование, фугоидные колебания имеют такой продолжительный период (1—2 мин.), что они обычно не имеют существенного значения при рассмотрении динамических характеристик.

Аналогично может быть выведена и система уравнений бокового движения [6, 7]. В этом случае обычно необходимо рассматривать все три степени свободы, и при решении полученного уравнения относительно p , r и v для случаев движения руля поворота или элерона определяются три движения: колебательное движение („голландский шаг“), „спиральное“ движение первого порядка (часто расходящееся) и движение крена первого порядка (обычно сильно демпфированное). Однако обсуждение в последующих разделах

статьи будет ограничено рассмотрением только упрощенной продольной реакции, определяемой уравнением (24).

Очевидно, что реакцию самолета на любое движение руля высоты можно рассчитать, если в дополнение к постоянным, определяющим массу и инерцию, известны производные, характеризующие изменение аэродинамических сил и моментов при изменении α , $\dot{\alpha}$, q и δ . Как указывалось выше, эти величины часто не бывают известны с любой желаемой степенью точности, особенно для новых самолетов или самолетов необычной конструкции, предназначенных для полетов со звуковой или близкой к ней скоростью. Поэтому необходимо обращаться к методам испытательных полетов для определения динамического поведения, а затем на основании полученных результатов находить передаточные функции, необходимые для синтеза автоматических систем управления.

Определение частотных характеристик по данным летных испытаний

Наиболее простой метод определения частотной характеристики физической системы заключается в измерении установившейся реакции на синусоидальный входной сигнал для ряда различных частот. Несмотря на то, что этот метод с успехом используется для определения частотных характеристик летящего самолета [8], он во многих случаях неприменим. Этот метод требует установки оборудования для управления рулями по синусоидальному закону и, если данные должны определяться для большого числа частот, может потребовать недопустимо продолжительных летных испытаний.

В общем случае удобнее находить частотную характеристику путем анализа неуставившейся реакции самолета на произвольное отклонение рулей, обычно определяемое ступенчатой функцией или функцией, имеющей форму импульса. Реакция в неуставившемся режиме характеризуется непрерывным частотным спектром и может определяться при продолжительности полета порядка нескольких секунд. Можно считать, что неуставившееся колебание представляет собой сумму синусоидальных колебаний с различными амплитудами, перекрывающих весь частотный диапазон. Поэтому реакция линейной системы на неуставившийся входной сигнал может рассматриваться как реакция

13. Определение частотных характеристик самолета в полете А28

этой системы на сумму синусоидальных колебаний, образующих неустановившийся входной сигнал.

Метод преобразования данных переходного процесса системы, являющегося функцией времени, в данные, выраженные в функции частоты, основан на использовании интеграла Фурье, который в некоторых условиях позволяет преобразовать функцию времени $f(t)$ в комплексную функцию частоты $f[j\omega]$ ¹⁾.

Интеграл Фурье может быть записан следующим образом:

$$f[j\omega] = \int_0^{\infty} f(t) e^{-j\omega t} dt. \quad (26)$$

Этот интеграл должен вычисляться в пределах от времени, равного нулю, до времени, равного бесконечности, для каждой частоты ω , для которой желательно найти $f[j\omega]$. Интегрирование может быть выполнено только при условии, если известно поведение $f(t)$ на бесконечном интервале времени после возмущения системы. Это означает, что система должна достичь установившегося состояния в течение некоторого конечного периода времени T , так чтобы функция $f(t)$ могла быть выражена аналитически на интервале времени от T до бесконечности. Если при анализе летных данных входная величина достигает установившегося значения в течение сравнительно короткого периода времени, то выходной сигнал или одна из его производных также становится постоянной величиной в течение некоторого периода времени T при условии, что система является устойчивой и линейной. Таким образом, в случае реакции скорости тангажа на движение руля высоты в качестве входа интеграл Фурье можно представить в виде двух частей

$$q[j\omega] = \int_0^T q(t) e^{-j\omega t} dt + \int_T^{\infty} q_T e^{-j\omega t} dt, \quad (27)$$

где q_T — постоянное установившееся значение $q(t)$, соответствующее постоянной величине отклонения руля высоты $\delta(t)$

¹⁾ Круглые скобки мы будем использовать для обозначения функции времени, квадратные скобки — для обозначения функции частоты,

(в случае импульсного входного сигнала, для которого $\delta(t)$ точно становится равным нулю, q_T также должна равняться нулю). Второй интеграл может быть выражен аналитически, так что

$$q[j\omega] = \int_0^T q(t) e^{-j\omega t} dt + \frac{q_T}{j\omega} e^{-j\omega T}. \quad (28)$$

Для дальнейшего вычисления необходимо разложить величину $q[j\omega]$ на вещественную и мнимую составляющие

$$q[j\omega] = R + jI, \quad (29)$$

где

$$R = \int_0^T q(t) \cos \omega t dt - \frac{q_T}{\omega} \sin \omega T, \quad (30)$$

$$I = - \int_0^T q(t) \sin \omega t dt - \frac{q_T}{\omega} \cos \omega T. \quad (31)$$

Интегралы, включающие переходную часть величины $q(t)$, должны вычисляться с помощью одного из существующих методов численного интегрирования, что необходимо для получения R и I . После того как эти интегралы вычислены, можно выразить $q[j\omega]$ с помощью полярных координат

$$q[j\omega] = |q| e^{j\varphi_q}, \quad (32)$$

где $|q|$ — абсолютная величина $\sqrt{R^2 + I^2}$, а φ_q — фазовый угол $\arctg(I/R)$.

Это вычисление повторяется для входной величины $\delta(t)$, и затем частотная характеристика системы выражается как отношение двух функций частоты

$$\frac{q[j\omega]}{\delta[j\omega]} = \frac{|q| e^{j\varphi_q}}{|\delta| e^{j\varphi_\delta}} = \left| \frac{q}{\delta} \right| e^{j(\varphi_q - \varphi_\delta)} = \left| \frac{q}{\delta} \right| e^{j\varphi}. \quad (33)$$

Как указывалось выше, существует несколько практических методов численного интегрирования, которые могут использоваться при вычислении интеграла Фурье. Одним из наиболее известных является правило площадей Симпсона, которое применяется для определения площади, ограниченной

13. Определение частотных характеристик самолета в полете 425

кривой произведения функций $f(t) \cos \omega t$ и $f(t) \sin \omega t$, с помощью ряда параболических аппроксимаций. Для обеспечения достаточной точности при применении этого метода характер изменения функции $f(t)$ во времени должен определяться для достаточно малых интервалов времени Δt так, чтобы каждый период кривой, соответствующей произведению функций, мог определяться по меньшей мере в результате восьми измерений. Это означает, что произведение $\omega \Delta t$, где Δt — интервал времени между измерениями, не должно превышать 45° . Например, при исследовании короткопериодического продольного движения диапазон частот, представляющий интерес, простирается от нуля до 16 рад/сек , что потребует составления таблиц с интервалами $0,05 \text{ сек}$.

Дальнейшее развитие этого основного метода, сначала в работе Филона [9], а затем в работе Шумахера [10], позволяет использовать параболическую аппроксимацию непосредственно для функции времени $f(t)$, а не для произведения функций $f(t) \cos \omega t$ и $f(t) \sin \omega t$. Это обеспечивает более высокую точность в определении высокочастотных составляющих, и в общем случае возможно получение приемлемой точности при величинах $\omega \Delta t$, достигающих 120° , конечно, при условии, что используется достаточное число интервалов времени для надлежащего определения функции времени. Применение этого метода аналогично применению основного правила Симпсона, за исключением того, что коэффициенты $1/3$, $4/3$ и $2/3$ тогда становятся функциями $\omega \Delta t$, т. е. необходимо использовать различные ряды коэффициентов для каждой частоты, при которой должен выполняться анализ.

С успехом применялись и другие методы численного интегрирования, основанные на ступенчатой или треугольной аппроксимации [11, 12]. Кроме того, применимы и графические методы, разработанные Симансом, Блейзингемом и Клементсоном [12]. Однако численные методы, основанные на параболической аппроксимации, обеспечивают наиболее высокую точность. Такие вычисления можно легко выполнить с помощью счетных машин.

В любом случае вычисления являются сложными и требуют большой затраты времени. Вероятно, наибольшим достижением в этой области была бы разработка электронных анализаторов гармоник, которые очень быстро вычисляют частотную характеристику по кривой функции времени, исключая,

таким образом, необходимость табулирования данных. В Аэро-навигационной лаборатории НАСА разрабатывается устройство, в котором развертывающая система, аналогичная телевизионным системам развертки, преобразует кривую, заданную графиком и определяющую поведение системы в неустановившемся режиме, в напряжение, которое затем умножается на $\cos \omega t$ и $\sin \omega t$ и интегрируется для получения напряжения, пропорционального членам

$$\int_0^T f(t) \cos \omega t dt \quad \text{и} \quad \int_0^T f(t) \sin \omega t dt.$$

Для определения этих величин на каждой частоте требуется около 1 мин. В настоящее время эта экономия времени достигается за счет понижения точности, но можно ожидать, что дальнейшее усовершенствование прибора даст возможность повысить его точность.

Независимо от метода интегрирования, применяемого для получения частотной характеристики, геометрическая форма входной функции оказывает существенное влияние на диапазон частот, в пределах которого можно получить надежное преобразование, и поэтому требует к себе пристального внимания. Преобразование Фурье от входной величины характеризует возбуждение системы на любой частоте. Двумя предельными видами входных сигналов являются ступенчатый сигнал и импульсный сигнал.

Преобразование для единичного ступенчатого сигнала равно $1/\omega$, что обеспечивает бесконечное возбуждение для составляющей с частотой, равной нулю (установившееся состояние), за счет ослабления возбуждения на более высоких частотах. Наоборот, импульс (с продолжительностью во времени, равной нулю) имеет постоянную величину преобразования во всем спектре частот. Таким образом, для большинства применений импульс является оптимальным видом входной величины, тогда как ступенчатая входная величина дает хорошие результаты только на очень низких частотах.

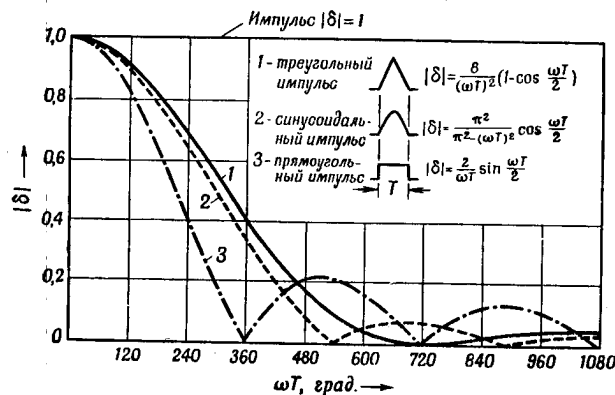
В реальных физических системах невозможно получить подлинный импульс, и наиболее точным приближением является треугольный входной сигнал с конечным основанием по оси времени T . Модуль преобразования треугольного

13. Определение частотных характеристик самолета в полете 427

импульса с единичной площадью определяется выражением

$$|\delta| = \frac{8}{(\omega T)^2} \left(1 - \cos \frac{\omega T}{2}\right). \quad (34)$$

Очевидно, что $|\delta|$ является функцией ωT и равняется нулю при $\omega T = 4\pi, 8\pi, \dots$. Для этих точек преобразование выходной величины также равно нулю, что делает частотную

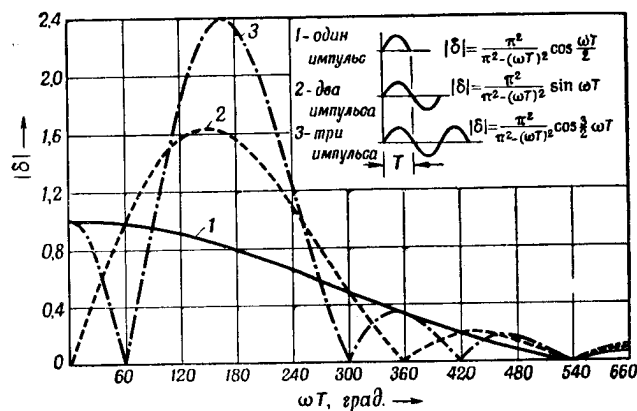


Фиг. 3. Амплитуды преобразования для импульсных входных сигналов с единичной площадью.

характеристику неопределенной. Для импульсов другой геометрической формы получаются до некоторой степени аналогичные преобразования, которые показаны на фиг. 3, где величины $|\delta|$ для треугольного и прямоугольного импульсов, а также для импульса в виде полуволны синусоиды, представлены в функции безразмерной частотной переменной ωT . В каждом случае площадь, ограниченная импульсом, и, следовательно, амплитуды, соответствующие нулевой частоте, равны единице. Три кривые имеют тождественную форму, приближаясь к нулю для величин ωT , лежащих в пределах от 2π до 4π , и являются типичными для любых импульсных входных сигналов произвольной формы.

Для получения наиболее точных результатов величина $|\delta|$ должна быть как можно больше в пределах полосы рассматриваемых частот, и поэтому желательно выбирать импульс с достаточно малым основанием так, чтобы при наиболее

высокой частоте ωT не превышала 360° . Это означает, что при использовании произвольного входного импульсного сигнала с основанием 1 сек. в общем случае можно определить частотную характеристику с достаточной степенью точности в полосе, ограниченной частотой порядка 6 рад/сек . Если бы основание импульса было уменьшено до $0,5 \text{ сек.}$, то максимальная надежная частота повысилась бы до величины порядка 12 рад/сек при условии, что не произойдет заметного уменьшения площади, ограниченной импульсом.



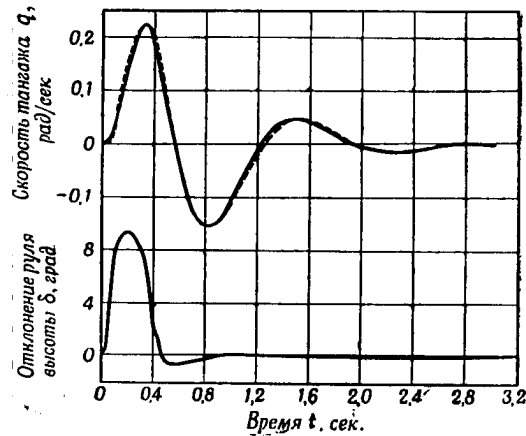
Фиг. 4. Амплитуды преобразования для последовательности синусоидальных импульсов.

Однако для любой физической системы существует максимальная скорость подачи входной величины. Следовательно, уменьшение T ниже некоторого предела означает уменьшение площади, ограниченной импульсом, и сопровождается уменьшением общей величины преобразования. Поэтому наиболее желательна такая входная величина, которая при достаточно большой площади импульса имела бы достаточно малое основание импульса. Однако для самолета движение рулей должно ограничиваться достаточно малыми величинами для получения линейной реакции системы.

Часто в тех случаях, когда необходимо получить точные данные в ограниченном диапазоне частот, может оказаться

13. Определение частотных характеристик самолета в полете 429

желательным использование последовательности импульсов, подобной изображенной на фиг. 4. На этом графике построены кривые, определяющие величины преобразований для одного, двух и трех импульсов. Площадь, ограниченная каждым



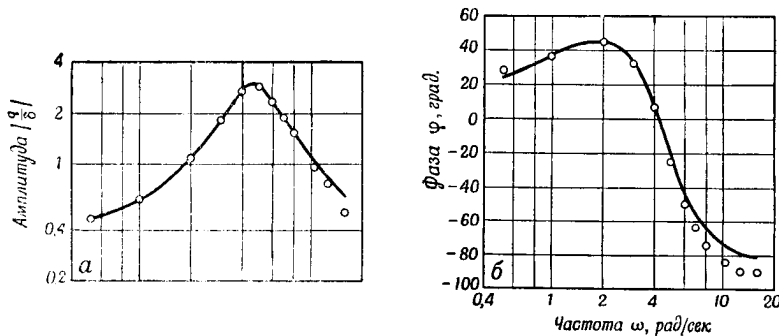
Фиг. 5. Типичная запись переходного процесса по скорости тангажа, полученная при летных испытаниях. Характеристика движения руля высоты импульсного типа.

— результаты летных испытаний; - - - результаты, полученные с помощью модели.

полупериодом, равна единице. Синусоидальные импульсы использованы в этом примере вследствие простоты вычисления, но полученные кривые типичны для любой произвольной формы импульса. Увеличение числа периодов увеличивает возбуждение на данной частоте, приближенно соответствующей $\omega T = \pi$, но уменьшает ширину рабочей полосы преобразования в диапазоне частот от 4 до 6 рад/сек, то желательно использовать последовательность из двух или трех импульсов с основанием порядка 0,6 сек.

На фиг. 5 приведена типичная запись реакции по скорости тангажа на импульсное движение руля высоты, сделанная в условиях полета. Кривые построены на основании фотографических записей, сделанных записывающими

испытательными приборами. Величины q и δ табулированы с интервалом 0,05 сек. и проанализированы путем аппроксимации функций времени семейством парабол. Для этого примера функция $q(t)$ достигает установившейся величины, равной



Фиг. 6. Частотные характеристики скорости тангажа при повороте руля высоты.

— вычислено по уравнению (40); \circ — экспериментальные данные.

нулю, приблизительно через 3,5 сек., а, следовательно, второй интеграл уравнения (27) равняется нулю. Если принять, например, что $\omega = 4$ рад/сек, то необходимо вычислить

$$R = \int_0^{3,5} q(t) \cos 4t dt = 0,0847 \quad (35)$$

и

$$I = - \int_0^{3,5} q(t) \sin 4t dt = -0,0777. \quad (36)$$

Это вычисление повторяется для $\delta(t)$, причем при $\omega = 4$ получаются величины $R = 0,0279$ и $I = -0,0331$. Тогда получаем

$$\frac{q[j\omega]}{\delta[j\omega]} = \frac{0,0847 - j0,0777}{0,0279 - j0,0331} = 2,65e^{j7,36^\circ}. \quad (37)$$

Эта операция численного интегрирования была выполнена для ряда частот в диапазоне от 0,5 до 16 рад/сек; полученная частотная характеристика показана на фиг. 6. На амплитудной характеристике имеется резонансный пик при

13. Определение частотных характеристик самолета в полете 431

частоте, близкой к $4,5 \text{ рад/сек}$, что типично для колебательных систем второго порядка. Положение этого пика обычно определяет диапазон частот, представляющий интерес для исследователя. Однако, как указывалось выше, точность, с которой могут вычисляться характеристики в пределах этого диапазона, зависит от характера входной величины и величины интервала времени, для которого составлены таблицы переходных процессов.

Преобразование частотных характеристик системы из графической формы в аналитическую

Во многих случаях можно ограничиться определением частотной характеристики в графической форме, но часто желательно или даже необходимо (например, для исследований с помощью моделирующих устройств) выразить эти данные в аналитической форме в виде передаточных функций, что особенно важно при исследовании сложных систем. Численные значения коэффициентов передаточной функции могут быть определены по графику частотной характеристики или непосредственно по параметрам переходного процесса. В любом случае возникает задача аппроксимации кривой, т. е. задача определения численного уравнения, наилучшим образом описывающего некоторую совокупность экспериментальных данных.

Методы аппроксимации кривых можно подразделить на два основных типа: аналитические и графические. Аналитические методы [13—17] часто основаны на использовании метода наименьших квадратов, и во всех случаях необходимо знать, или по меньшей мере заранее задаться видом передаточной функции, т. е. степенями многочленов, стоящих в числителе и знаменателе передаточной функции. Тогда задача сводится к определению (для данной последовательности экспериментальных значений и уравнения заданной формы) наилучших численных значений коэффициентов, при подстановке которых уравнение наиболее точно описывает экспериментальные данные. Подобные методы требуют выполнения трудоемких действий, особенно если принятая передаточная функция имеет порядок выше второго. Кроме того, необходимо пользоваться данными, определенными

с высокой степенью точности, не всегда достижимой при летных испытаниях.

При анализе данных, полученных при летных испытаниях, некоторые графические методы имеют практические преимущества по сравнению с упомянутыми выше аналитическими методами. Один из методов определения численных коэффициентов передаточной функции по графику частотной характеристики, с успехом применявшийся в Эймской аэронавигационной лаборатории, исключает значительную часть трудоемких вычислений, допускает возможность быстрой оценки данных (часто данные, полученные при летных испытаниях, бывают неполноценными из-за дополнительных возмущений, вызванных, например, резким ветром) и не требует априорного определения вида передаточной функции.

Этот метод заключается в аппроксимации графиков частотных характеристик с помощью ряда шаблонов частотных характеристик, определяющих функции комплексной переменной первого и второго порядков

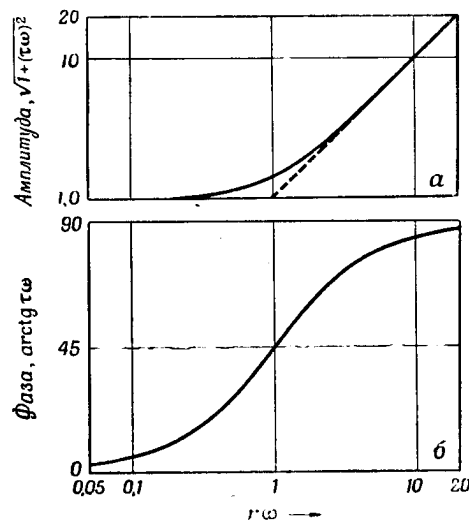
$$\lg [1 + j\omega\tau] \text{ и } \lg \left[1 + 2\zeta j \left(\frac{\omega}{\omega_n} \right) - \left(\frac{\omega}{\omega_n} \right)^2 \right],$$

где $\omega\tau$ и ω/ω_n — безразмерные частотные переменные и ζ — коэффициент демпфирования. Эти шаблоны, изготовленные из прозрачной пластмассы, были предложены К. Дрейпером (Лаборатория приборостроения Массачусетского технологического института) и описаны в работе [18]. Комплект шаблонов состоит из двух шаблонов, соответствующих кривым первого порядка (логарифмические амплитудная и фазовая характеристики), и шаблонов, соответствующих семейству кривых второго порядка с различными величинами параметра ζ . Образцы этих шаблонов показаны на фиг. 7 и 8. Величина, соответствующая единице на шкалах безразмерной частоты ($\omega\tau$ и ω/ω_n), называется точкой сопряжения.

В общем случае для линейных систем многочлены, образующие числитель и знаменатель передаточной функции, могут быть разложены на множители первого и второго порядков, имеющие указанную выше форму. Тогда кривая частотной характеристики может строиться путем надлежащего расположения необходимых шаблонов относительно оси частот и суммирования логарифмов амплитуд и фазовых углов.

13. Определение частотных характеристик самолета в полете 483

Возможно также решение обратной задачи нахождения членов первого и второго порядков, составляющих частотную характеристику, представленную в графической форме. При этом частотная характеристика вычерчивается на точной логарифмической шкале, и путем надлежащего комбинирования шаблонов обычно удается после нескольких попыток найти одну комбинацию, точно соответствующую



Фиг. 7. Частотная характеристика для члена первого порядка $1 + \tau j\omega$.

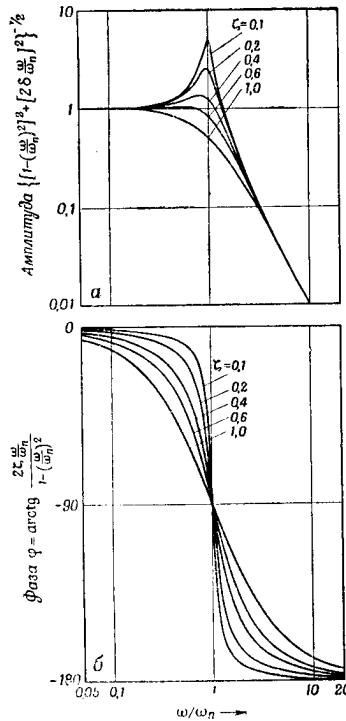
отношению амплитуд и фазовому углу. Эта частная комбинация позволяет найти форму уравнения, наиболее точно описывающего частотную характеристику в ее графическом представлении.

На фиг. 9 показано, каким образом для экспериментальной характеристики, изображенной на фиг. 6, подбирается множитель первого порядка, стоящий в числителе выражения, и множитель второго порядка, стоящий в знаменателе, при величине коэффициента демпфирования $\zeta = 0,3$. (Соответствующий подбор фазовой характеристики не показан.) Точка сопряжения кривой первого порядка получается при $\omega = 1,27$, откуда $\tau = 1/1,27 = 0,82$. Аналогичным образом

для точки сопряжения кривой второго порядка $\omega_n = 4,65$. Постоянная усиления K (амплитуда при частоте, равной нулю) составляет 0,460. Тогда передаточная функция запишется в виде

$$\frac{q}{\delta} = \frac{0,460(1 + 0,82s)}{1 + \frac{2 \times 0,3}{4,65}s + \left(\frac{s}{4,65}\right)^2} \quad (38)$$

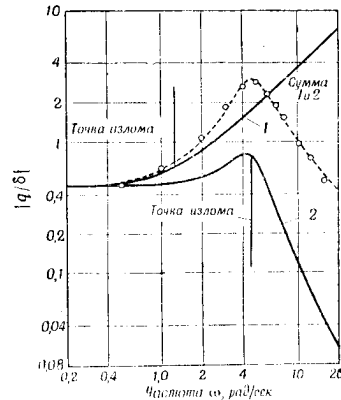
Численные значения передаточных функций, найденные с помощью шаблонов, могут оказаться достаточно



Фиг. 8. Частотная характеристика для члена второго порядка

$$\left[1 + \frac{2zeta}{\omega_n} j\omega - \left(\frac{\omega}{\omega_n}\right)^2\right]^{-1}$$

точными для многих применений; однако иногда желательно дальнейшее повышение точности этих данных, что достигается при помощи моделирующих устройств (с таблицами входных и выходных величин). При применении такого метода передаточная функция с численными величинами, найденными при помощи шаблонов, вводится в вычислительную машину. Запись истинного изменения входного возмущения во времени,

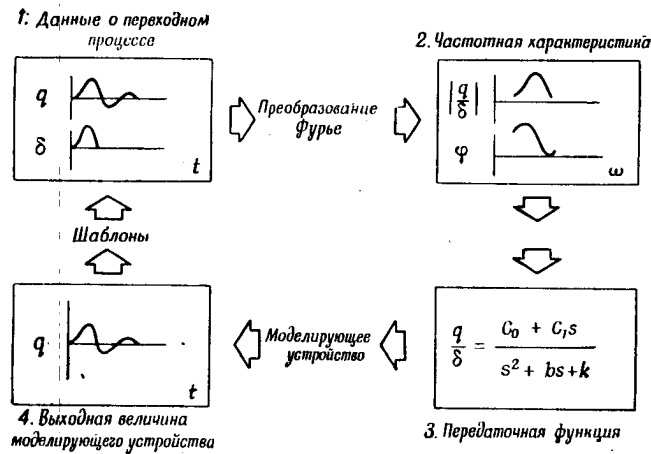


Фиг. 9. Подбор экспериментальной частотной характеристики с помощью шаблонов.

1 — член первого порядка; 2 — член второго порядка; \circ — экспериментальные данные.

13. Определение частотных характеристик самолета в полете 435

воздействующего на самолет при проведении летных испытаний, используется в качестве входной величины, тогда как выходная величина сравнивается с характеристикой, полученной в условиях полета. Точность совпадения записей позволяет судить о точности определения коэффициентов с помощью шаблонов. Если записи не совпадают, то величины коэффициентов, задаваемых в вычислительной машине, легко



Фиг. 10. Последовательность операций при преобразовании данных.

можно изменять до тех пор, пока не будет получено точное соответствие. Этот метод позволяет не только повысить точность определения коэффициентов, но и полностью проверить правильность всего анализа¹⁾.

На фиг. 10 в схематической форме показана последовательность операций, совершаемых над данными.

Для данных, использованных в предшествующем примере, передаточная функция, найденная с помощью шаблонов, была преобразована посредством вычислительной машины

¹⁾ При исследовании хорошо известных систем часто представляется возможным полностью исключить вычисление частотной характеристики и использовать моделирующее устройство для непосредственного определения передаточной функции, исходя из характера изменения переходных процессов, при условии наличия достаточно полных данных о передаточной функции.

таким образом, чтобы получить точное соответствие, показанное на фиг. 5. Соответствующая передаточная функция в окончательной форме определяется выражением

$$\frac{q}{\delta} = \frac{0,415(1 + 1,04s)}{1 + \frac{2 \times 0,356}{4,70}s + \left(\frac{s}{4,70}\right)^2}, \quad (39)$$

которое может быть записано так:

$$\frac{q}{\delta} = \frac{9,16 + 9,52s}{s^2 + 3,35s + 22,1}. \quad (40)$$

Частотная характеристика, вычисленная по этому уравнению, представлена в виде кривой на фиг. 6, причем получено очень хорошее совпадение с измеренной характеристикой на всех частотах, за исключением частот выше 8 *рад/сек*. Отклонения на более высоких частотах могут обуславливаться упругими деформациями самолета во время полета или недостаточно высокой точностью анализа переходного процесса с помощью преобразования Фурье. В любом случае расхождение не влияет на характер изменения кривых на фиг. 5. Изменение коэффициентов уравнений (38) и (39) показывает степень точности, достигаемую при использовании одних только шаблонов.

Очевидно, что передаточная функция, найденная экспериментально, имеет ту же форму, что и функция, полученная аналитически из теоретических уравнений движения (уравнение (24)), а, следовательно, допущения, сделанные при выводе уравнения (24), правильно описывают движения этого самолета.

Влияние изменения скорости и высоты на характеристику самолета

Передаточная функция, найденная по данным летных испытаний, справедлива только для тех условий полета (скорость и высота), при которых производились испытания. Однако может быть желательным определение характеристики самолета для иных рабочих условий, для которых не имеется экспериментальных летных данных. В некоторых случаях требуемые сведения могут быть получены путем экстраполяции коэффициентов передаточной функции.

13. Определение частотных характеристик самолета в полете 437

Согласно теории полета, аэродинамические силы и моменты, действующие на самолет, пропорциональны динамическому давлению $\rho V^2/2$, и таким образом Z и M обычно выражаются в следующем виде:

$$Z = -L = -C_L \frac{\rho V^2 S}{2}, \quad (41)$$

$$M = C_m \frac{\rho V^2 S c}{2}, \quad (42)$$

где L — полная подъемная сила, а C_L и C_m — безразмерные коэффициенты подъемной силы и момента. При использовании этих обозначений можно записать частные производные в уравнениях (17) и (18) в виде

$$\begin{aligned} Z_\alpha &= -\frac{\rho VS}{2m} C_{L\alpha}, & M_\delta &= \frac{\rho V^2 S c}{2I_y} C_{m\delta}, \\ Z_\delta &= -\frac{\rho VS}{2m} C_{L\delta}, & M_q &= \frac{\rho V S c^2}{4I_y} C_{mq}, \\ M_\alpha &= -\frac{\rho V^2 S c}{2I_y} C_{m\alpha}, & M_{\dot{\alpha}} &= \frac{\rho V S c^2}{4I_y} C_{m\dot{\alpha}}, \end{aligned}$$

где

$$C_{L\alpha} = \frac{\partial C_L}{\partial \alpha} \text{ и т. д.}$$

Тогда можно выразить коэффициенты C_0 и C_1 (уравнение (19), в котором мы пренебрегли всеми достаточно малыми членами) следующим образом:

$$b = \frac{\rho VS}{2} \left[\frac{C_{L\alpha}}{m} - (C_{mq} + C_{m\dot{\alpha}}) \frac{c^2}{2I_y} \right], \quad (43)$$

$$k = -\frac{\rho V^2 S c}{2I_y} C_{m\alpha}, \quad (44)$$

$$C_0 = \frac{\rho^2 V^2 S^2 c}{4mI_y} (C_{L\alpha} C_{m\alpha} - C_{L\delta} C_{mq}), \quad (45)$$

$$C_1 = \frac{\rho V^2 S c}{2I_y} C_{m\delta}. \quad (46)$$

Если производные $C_{L\alpha}$ и т. д. остаются постоянными во всей области изменения условий полета, представляющей интерес, то можно сразу записать, что наши коэффициенты удовлетворяют следующим формулам пропорциональности:

$$\begin{aligned} b &\sim \rho V, \\ k &\sim \rho V^2, \\ C_0 &\sim \rho^2 V^2, \\ C_1 &\sim \rho V^2. \end{aligned}$$

Поскольку плотность воздуха ρ является явной функцией высоты, эти соотношения позволяют экстраполировать коэффициенты экспериментальной передаточной функции и соответствующей частотной характеристики в данных пределах изменения скорости и высоты.

Существуют некоторые ограничения для применения этих соотношений, о которых следует упомянуть. Производные от аэродинамических коэффициентов, будучи по существу независимы от скорости, изменяются в зависимости от величины числа M . Характер этого изменения зависит от типа данного самолета, но обычно это изменение достигает ощутимой величины только для чисел M свыше 0,5. Однако в диапазоне сверхзвуковых скоростей изменения довольно велики и неупорядочены. Тогда, строго говоря, для высокоскоростных самолетов пропорциональность сохраняется только для изменений скорости и высоты, соответствующих постоянному числу M .

Существуют также другие ограничения, которые необходимо рассмотреть. При очень малых скоростях или на очень больших высотах самолет уравнивается лишь при больших углах атаки. При полетах с очень большими углами атаки могут возникнуть нежелательные изменения некоторых производных от аэродинамических коэффициентов. Кроме того, для самолетов со сравнительно гибкой конструкцией упругие деформации во время полета приведут к изменению фактической величины некоторых производных в соответствии с изменением динамического давления. Несмотря на эти ограничения, приведенные выше соотношения позволяют оценить характер изменений характеристик самолета при изменении скорости и высоты.

Заключение

Растущая потребность в данных, касающихся динамических характеристик самолета, вызвала к жизни совершенно новую область летных испытаний и анализа. Данные этого типа имеют чрезвычайно важное значение для расчета автопилотов и полностью автоматических систем управления.

Несмотря на то, что динамические характеристики можно во многих случаях вычислить, исходя из обычных производных от аэродинамических коэффициентов, полученных из

13. Определение частотных характеристик самолета в полете 43

испытаний в аэродинамической трубе, и теоретических исследований, для получения точных данных часто необходимо прибегать к методу летных испытаний, учитывая широкие пределы изменения рабочих условий для современных самолетов. Динамические характеристики могут быть представлены в виде частотных характеристик, но они приобретают наиболее важное значение, будучи выражены аналитически в форме передаточной функции. Применение различных графических средств совместно с электронными моделирующими устройствами позволяет произвести быстрое и достаточно точное определение передаточной функции самолета по данным переходного процесса, полученного при летных испытаниях.

ЛИТЕРАТУРА

1. Bolla W., *JAS*, 18, 569 (1951).
2. Millikan W. F., „Dynamic Stability and Control Research“, 3 Intern. Joint Conf. RAS—IAS, Brighton, England, Sept. 1951, Cornell Aero. Lab. Rep. № Cal-39.
3. Rea J. B., *JAS*, 18, 375 (1951).
4. Аэродинамика, под ред. В. Дюрэнда, том 5, Оборонгиз, 1940.
5. Perkins C. D., Hage R. E., *Airplane Performance Stability and Control*, New York, 1949.
6. Bretoi R. N., *Trans. ASME*, 74, 415 (1952).
7. Campbell J. P., McKinney M. O., „Summary of Methods for Calculating Dynamic Lateral Stability and Response and for Estimating Lateral Stability Derivatives“, NACA TN 2403, 1951.
8. Campbell G. F., Whitcomb D. W., Breuhaus W. O., „Dynamic Longitudinal Stability and Control Flight Tests of a B-25J Airplane—Forced Oscillation and Step Function Response Methods, Utilizing an A-12 Automatic Pilot“, Cornell Aero. Lab. Rep. № TB-405-F-3, Apr. 1947.
9. Filon L. N. G., *Proc. Roy. Soc. Edinburgh*, 49, 38 (1928—1929).
10. Schumacher L. E., „Methods of Analysing Transient Flight Data to Obtain Aircraft Frequency Response“, WADC Memo. Rep. № MCRFT-2268, Jan. 1950.
11. Seamans R. C., Jr., Bromberg E. G., Payne L. E., *JAS*, 15, 535 (1948).
12. Seamans R. C., Blasingame B. P., Clementson G. C., *JAS*, 17, 22 (1950).

13. Greenberg H., „A Survey of Methods for Determining Stability Parameters of an Airplane From Dynamic Flight Measurements“, NACA TN 2340, 1951.
14. Shinbrot M., „A Least Squares Curve Fitting Method with Applications to the Calculation of Stability Coefficients from Transient Response Data“, NACA TN 2341, 1951.
15. Donegan J. J., Pearson H. A., „Matrix Method of Determining the Longitudinal Stability Coefficients and Frequency Response of an Aircraft From Transient Flight Data“, NACA Rep. 1070, 1952.
16. Jones A. L., Briggs B. R., „A Survey of Stability Analysis Techniques for Automatically Controlled Aircraft“, NACA TN, 2275, 1951.
17. Eggleston J. M., Mathews C. W., „Application of Several Methods for Determining Transfer Functions and Frequency Response of an Aircraft From Flight Data“, NACA TN 2997, 1953.
18. Lees S., „Graphical Aids for the Graphical Representation of Functions of the Imaginary Argument“, MIT Instrumentation Lab. Eng. Memo. E-25, Febr. 1951.

14. МЕТОД ОЦЕНКИ НЕЛИНЕЙНЫХ СЛЕДЯЩИХ СИСТЕМ¹⁾

М. Мэтьюз

Проблема оценки следящей системы, рассматриваемая в этой статье, заключается в определении характеристик этой системы путем измерения ее реакции на испытательный входной сигнал. Например, для определения передаточной функции линейной системы могут быть использованы данные о реакции системы на синусоидальный, неустановившийся или случайный входной сигнал.

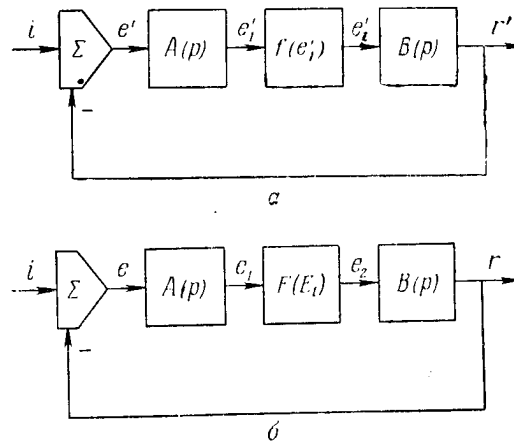
Задача оценки является одной из основных задач в области следящих систем, и необходимость улучшения способов ее решения, особенно для нелинейных следящих систем, давно признана всеми [1]. Методы оценки, основанные на чисто линейной теории, например на измерении переходного или синусоидального процесса, имеют значительные ограничения. Эти ограничения возникают из-за влияния нелинейностей, которые свойственны всем следящим системам. В самых идеальных случаях требуется тщательная регулировка амплитуды испытательных сигналов, чтобы обеспечить при испытаниях линейное действие следящей системы. Во многих практических случаях у следящей системы нельзя обнаружить диапазона линейного действия ввиду того, что влияние различных нелинейностей перекрывается. Это может иметь место при наличии зоны нечувствительности, сказывающейся при малых уровнях сигнала, и насыщения, влияющего при больших уровнях сигнала. В этом случае для получения характеристик линейной части системы почти невозможно истолковать данные на основе линейных методов оценки.

¹⁾ Mathews M. V., *Trans AIEE*, Part II, 74, 114—123 (1955).

В статье излагаются два метода, свободных от основных ограничений линейных методов оценки благодаря учету нелинейных характеристик, присущих следящим системам. Первый метод основан на использовании для оценки синусоидальных данных; второй метод подобен первому, но для оценки используются данные случайных сигналов.

Оценка по характеристикам синусоидальных сигналов

Способ оценки с помощью синусоидальных данных основан на использовании методов аналитического расчета реакции некоторых типов нелинейных следящих систем, изложенных Кохенбургером [2] и Джонсоном [3]. В этих работах



Фиг. 1. Структура исследуемых следящих систем.
a — нелинейная следящая система; *б* — эквивалентная квазилинейная система. Символ p означает оператор дифференцирования.

нелинейная следящая система аппроксимируется квазилинейной системой. Термин „квазилинейный элемент“ или „квазилинейная система“ в этой статье относится к линейному элементу или системе, параметры которых являются функциями входного сигнала элемента или системы.

Следящая система, характеристики которой необходимо определить, должна иметь структурную схему, подобную изображенной на фиг. 1, *a*, которая соответствует обычной

одноконтурной следящей системе. Прямая цепь контура состоит из двух линейных элементов с передаточными функциями $A(s)$ и $B(s)$ и одного безынерционного нелинейного элемента с характеристикой $f(e'_1)$. Выход нелинейного элемента e'_2 определяется уравнением

$$e'_2 = f(e'_1), \quad (1)$$

где e'_1 — вход нелинейного элемента.

Первый шаг в процессе определения характеристик системы заключается в определении наилучшей квазилинейной аппроксимации (фиг. 1, б) нелинейной следящей системы путем замены нелинейного элемента следящей системы его квазилинейной моделью. Квазилинейное представление характеристики $f(e'_1)$ нелинейного элемента является простой функцией усиления $F(E_1)$ [2, 3], которая определяется по $f(e'_1)$ с помощью уравнения

$$F(E_1) = \frac{2}{\pi E_1} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} f(E_1 \sin \theta) \sin \theta d\theta. \quad (2)$$

Таким образом, для квазилинейной системы

$$\frac{e_2}{e_1} = F(E_1), \quad (3)$$

где e_1 и e_2 — соответственно вход и выход квазилинейного элемента, а e_1 — синусоида с амплитудой E_1 .

Предположим, что квазилинейная система достаточно хорошо аппроксимирует нелинейную следящую систему, так что данные нелинейной системы могут быть использованы для определения характеристик элементов квазилинейной системы. Передаточные функции $A(s)$ и $B(s)$ и квазилинейная функция $F(E_1)$ оцениваются по данным вход — реакция, экспериментально определенным для следящей системы. Оценка следящей системы заканчивается использованием того факта, что $A(s)$ и $B(s)$ являются также передаточными функциями линейных элементов системы и определением функции $f(e'_1)$ по квазилинейной функции $F(E_1)$. Последнее выполняется путем решения уравнения (2), рассматриваемого

как интегральное уравнение. Это решение рассмотрено в Приложении I.

Принципиальное сграницение этого способа оценки состоит в необходимости введения предположения о том, что может быть найдена хорошая квазилинейная аппроксимация нелинейной следящей системы. Таким образом, эффективно оценить можно только те системы, которые допускают исследование методом Кохенбургера — Джонсона. Эти системы должны давать небольшие по величине нелинейные возмущения. Тем не менее мы будем считать, что эти системы могут иметь элементы, которые вносят существенно нелинейные функции, например резкое насыщение.

Экспериментальные данные получаются при возбуждении системы синусоидальным сигналом и измерении реакции системы как функции амплитуды и частоты входного сигнала. Мгновенное значение функции времени для синусоидального входного сигнала, приложенного к системе, указанной на фиг. 1, *a*, может быть записано в виде

$$i = I \cos \omega t, \quad (4)$$

где *I* — амплитуда входа, а ω — круговая частота. Для этого входа промежуточные координаты системы могут быть выражены в следующей форме:

$$\begin{aligned} e' &= E \cos(\omega t + \varepsilon) + \text{нелинейные возмущения,} \\ e'_1 &= E_1 \cos(\omega t + \varepsilon_1) + \text{нелинейные возмущения,} \\ e'_2 &= E_2 \cos(\omega t + \varepsilon_2) + \text{нелинейные возмущения,} \\ r' &= R \cos(\omega t + \rho) + \text{нелинейные возмущения.} \end{aligned} \quad (5)$$

В этих уравнениях строчные буквы обозначают мгновенные значения функций времени, прописные буквы — амплитуды основных гармоник этих функций, а греческие буквы — фазовые углы основных гармоник. Все другие составляющие координат, произведенные нелинейным элементом, объединены в члены, характеризующие нелинейные возмущения.

Если входной сигнал приложен к квазилинейной аппроксимирующей системе и сделано основное предположение, заключающееся в том, что основные гармоники сигналов в квазилинейной системе равны основным гармоникам сигналов следящей системы, то выражения для сигналов в ква-

в нелинейной системе могут быть записаны следующим образом:

$$\begin{aligned} e &= E \cos(\omega t + \varepsilon), \\ e_1 &= E_1 \cos(\omega t + \varepsilon_1), \\ e_2 &= E_2 \cos(\omega t + \varepsilon_2), \\ r &= R \cos(\omega t + \rho). \end{aligned} \quad (6)$$

Величины в уравнениях (6) определены точно так же, как и соответствующие величины в уравнениях (5).

Данные для оценки следящей системы представляются отношением амплитуды реакции системы к амплитуде сигнала ошибки R/E , измеренным в функции амплитуды сигнала ошибки E и частоты ω входного сигнала. Совокупность данных для типичной следящей системы, которая будет использована в качестве примера, указана на фиг. 2. Эти данные были получены путем моделирования следящей системы на электронном дифференциальном анализаторе. Прямая цепь контура моделируемой системы состояла из компенсационной цепи с передаточной функцией $A(s)$, элемента с насыщением с характеристикой $f(e_1')$, который характеризовал насыщение момента главного двигателя, и линейного элемента с передаточной функцией $B(s)$, представляющего линейные характеристики этого двигателя.

В некоторых следящих системах возможно непосредственное измерение R/E и E , в других удобно измерять либо реакцию, либо ошибку следящей системы как функцию амплитуды и частоты входного сигнала; в этом случае отношение R/E можно определить по измеренным данным с помощью любой из хорошо известных графических функций, например с помощью графиков логарифмических амплитудных частотных характеристик [4, 5]¹⁾.

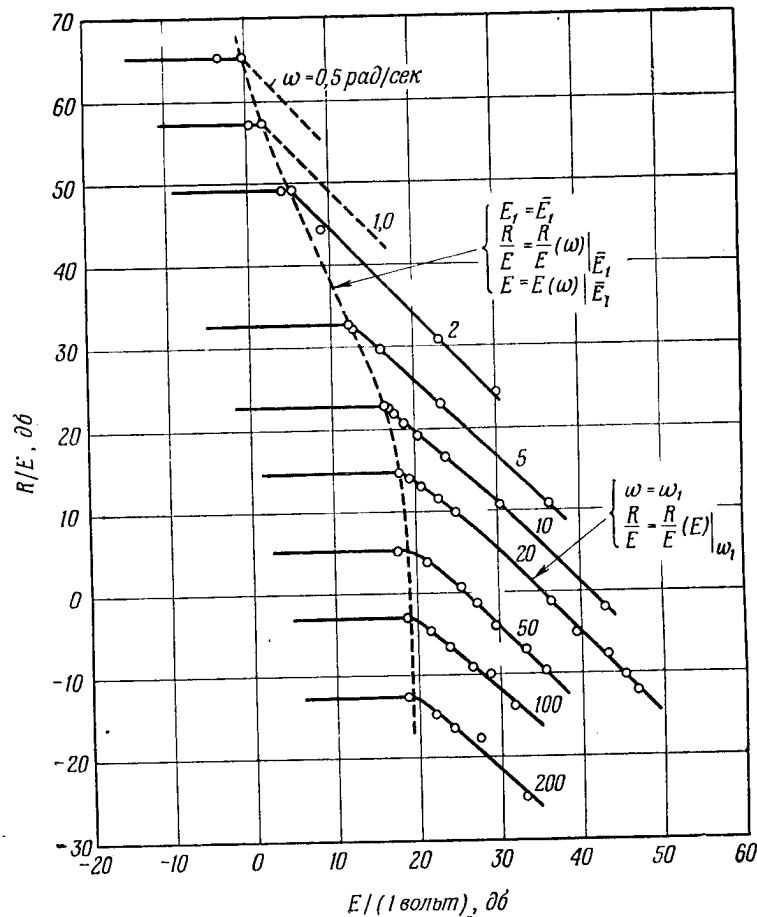
Характеристики системы как функции E и ω могут быть определены двумя уравнениями:

$$E_1 = |A(j\omega)| E, \quad (7)$$

$$\frac{R}{E} = |A(j\omega) B(j\omega) F| |A(j\omega)| E. \quad (8)$$

¹⁾ Имеются в виду номограммы для определения частотных характеристик замкнутых систем по частотным характеристикам разомкнутых систем. — Прим. ред.

Уравнение (7) определяет вход квазилинейного элемента E_1 как функцию $|A(j\omega)|$ и E , а уравнение (8) — отношение реакции системы к сигналу ошибки R/E как функцию $|A(j\omega)B(j\omega)|$ и $F(E_1)$.



Фиг. 2. Результаты опыта при синусоидальном входном сигнале.

Одна из кривых на фиг. 2, определяющих отношение R/E , может быть использована в уравнениях (7) и (8) для оценки квазилинейной функции усиления $F(E_1)$. Для выбранной

кривой частота ω постоянна и величина R/E поэтому может быть записана так:

$$\frac{R}{E}(E)|_{\omega_1}$$

Подстановка этой величины в уравнение (8) и решение его относительно $F(E_1)$ приводят к соотношению

$$F(E_1) = \frac{R/E(E)|_{\omega_1}}{|A(j\omega_1)B(j\omega_1)|} \quad (9)$$

Величина E_1 определяется соотношением

$$E_1 = |A(j\omega_1)|E. \quad (10)$$

Таким образом, функция $F(E_1)$ оценивается при помощи уравнений (9) и (10). Две произвольные постоянные $|A(j\omega_1)|$ и $|B(j\omega_1)|$, входящие в уравнения (9) и (10), не влияют на характеристику вход—выход системы, а служат лишь для определения масштабных коэффициентов внутренних сигналов в прямой цепи контура. Следовательно, они не могут быть оценены по данным вход—выход, а должны определяться произвольно или по результатам дополнительных измерений, сделанных над внутренними сигналами системы.

Численное определение функции $F(E_1)$ легко выполняется по табличным данным (см. табл. 1).

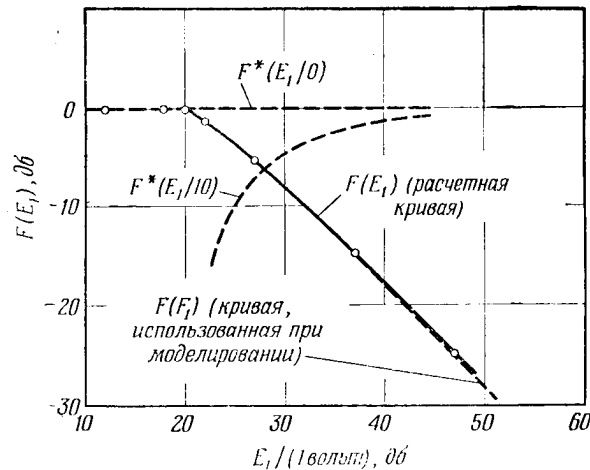
Таблица 1

ОЦЕНКА $F(E_1)$

$E/(1\sigma), \text{дб}$	$\frac{R}{E}(E) _{\omega_1}, \text{дб}$	$F(E_1), \text{дб}$	$E_1/(1\sigma), \text{дб}$
10	14,7	0	12
15	1,47	0	17
18	1,47	0	20
20	1,35	-1,2	22
25	9,6	-5,1	27
35	0,1	-14,6	37
40	-1,00	-24,7	47

Величины E и $R/E(E)|_{\omega_1}$, взятые из фиг. 2, табулированы в первых двух столбцах, величины $F(E_1)$ и E_1 , рассчитанные с помощью уравнений (9) и (10), представлены

в последних двух столбцах таблицы. Произвольные постоянные выбраны так, чтобы получить $F(E_1) = 0$ дб для малых значений E_1 и чтобы $F(E_1)$ начинала убывать при $E_1 = 20$ дб. Для этого принято $|A(j\omega_1)| = 2$ дб и $|B(j\omega_1)| = 12,7$ дб. Хорошее совпадение графика $F(E_1)$ на фиг. 3 с графиком квазилинейной функции усиления нелинейного элемента, использованного при моделировании,



Фиг. 3. Квазилинейная функция усиления.

характеризует точность этого способа оценки. Функция усиления $F(E_1)$ соответствует нелинейной характеристике с режимом насыщения, приведенной на фиг. 4. Определение $f(e'_1)$ по известной функции $F(E_1)$ рассматривается в Приложении I.

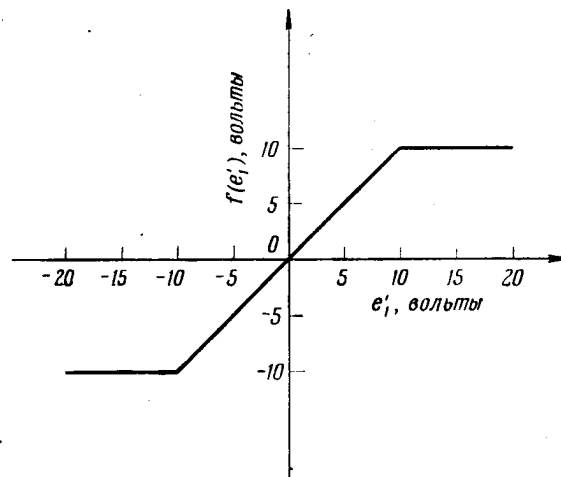
Линейные передаточные функции $A(j\omega)$ и $B(j\omega)$ могут быть оценены по одной из кривых R/E , для которой величина E_1 постоянна. Эта кривая не может быть определена непосредственно по экспериментальным данным, но может быть построена по кривым, для которых постоянна величина ω . Метод построения этой кривой будет описан ниже. Сейчас же предположим, что одна из таких кривых уже построена, как это показано пунктирной линией на фиг. 2. Координаты этой кривой можно задать параметрически

в функции ω . Таким образом, величина R/E на этой кривой определяется соотношением

$$\frac{R}{E} = \frac{R}{E}(\omega) |_{\bar{E}_1}, \quad (11)$$

а величина E на кривой — соотношением

$$E = E(\omega) |_{\bar{E}_1}. \quad (12)$$



Ф и г. 4. Нелинейная характеристика.

Подстановка выражения для E из уравнения (12) в уравнение (7) и решение последнего уравнения относительно $|A(j\omega)|$ дает

$$|A(j\omega)| = \frac{\bar{E}_1}{E(\omega) |_{\bar{E}_1}}, \quad (13)$$

а величина \bar{E}_1 может быть записана в виде

$$\bar{E}_1 = |A(j\omega_1)| E\omega_1 |_{\bar{E}_1}. \quad (14)$$

Комбинируя уравнения (13) и (14) для исключения \bar{E}_1 , получим соотношение

$$|A(j\omega)| = |A(j\omega_1)| \frac{E(\omega_1) |_{\bar{E}_1}}{E(\omega) |_{\bar{E}_1}}, \quad (15)$$

дающее окончательную оценку для $|A(j\omega)|$.

Подставляя данные из уравнения (11) в уравнение (8), можно получить выражение для $|A(j\omega)B(j\omega)|$:

$$|A(j\omega)B(j\omega)| = \frac{R/E(\omega)|_{\bar{E}_1}}{F(\bar{E}_1)}, \quad (16)$$

а для $F(\bar{E}_1)$ можно записать такое выражение:

$$F(\bar{E}_1) = \frac{R/E(\omega_1)|_{\bar{E}_1}}{|A(j\omega_1)B(j\omega_1)|}. \quad (17)$$

Исключая $F(\bar{E}_1)$ из (16) и (17), получаем соотношение для оценки $|A(j\omega)B(j\omega)|$:

$$|A(j\omega)B(j\omega)| = |A(j\omega_1)B(j\omega_1)| \frac{R/E(\omega)|_{\bar{E}_1}}{R/E(\omega_1)|_{\bar{E}_1}}. \quad (18)$$

Величина $|B(j\omega)|$ определяется из тождества

$$|B(j\omega)| = \frac{|A(j\omega)B(j\omega)|}{|A(j\omega)|}, \quad (19)$$

а величины $|A(j\omega)|$ и $|A(j\omega)B(j\omega)|$ — из уравнений (15) и (18).

Расчет величин $|A(j\omega)|$ и $|B(j\omega)|$ легко выполнить в табличной форме, как это иллюстрирует табл. 2.

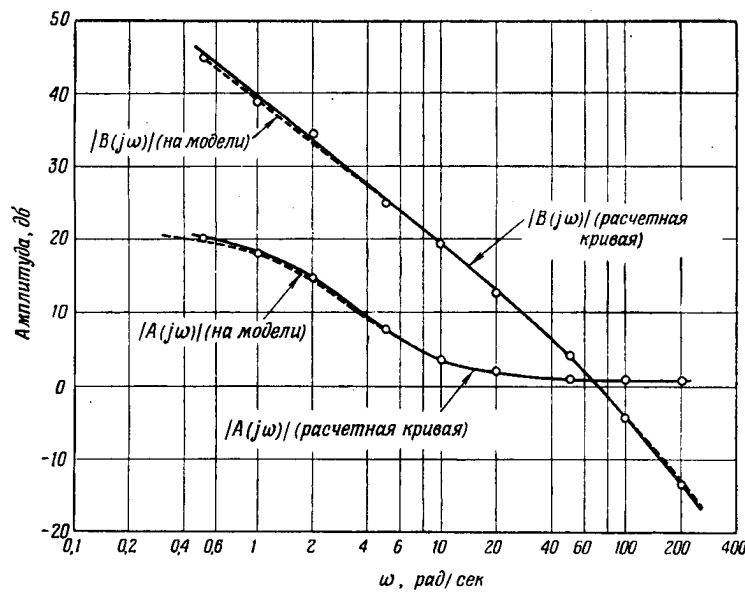
Таблица 2

ОЦЕНКА $|A(j\omega)|$ и $|B(j\omega)|$

ω , рад/сек	$\frac{R}{E}(\omega) _{\bar{E}_1}$, дб	$\frac{E(\omega) _{\bar{E}_1}}{\partial \sigma} / (1\sigma)$	$ A(j\omega) $, дб	$ A(j\omega)B(j\omega) $, дб	$ B(j\omega) $, дб
0,5	65,3	-0,3	20,3	65,3	45,0
1	57,0	2,0	18,0	57,0	39,0
2	49,2	5,3	14,7	49,2	34,5
5	32,8	12,2	7,8	32,8	25,0
10	22,9	16,5	3,5	22,9	19,4
20	14,9	17,9	2,1	14,9	12,8
50	5,3	18,9	1,1	5,3	4,2
100	-3,1	19,0	1,0	-3,1	-4,1
200	-12,6	19,2	0,8	-12,6	-13,4

Величины R/E и E , определенные по фиг. 2, табулируются как функции частоты ω в первых трех столбцах таблицы. Величины $|A(j\omega)|$, $|A(j\omega)B(j\omega)|$ и $|B(j\omega)|$, рас-

считанные по уравнениям (15), (18) и (19), даны в последних трех столбцах. Оценки амплитуд $|A(j\omega)|$ и $|B(j\omega)|$ графически представлены на фиг. 5. Эти амплитудные характеристики почти не отличаются от характеристик моделируемых передаточных функций, также приведенных на фиг. 5, что иллюстрирует точность данного метода оценки.



Фиг. 5. Амплитудные частотные характеристики линейных элементов.

Уравнения (15) и (19) определяют метод оценки $|A(j\omega)|$ и $|B(j\omega)|$. Для того чтобы получить передаточные функции $A(s)$ и $B(s)$ по амплитудным характеристикам, можно использовать метод Боде [6], и в дополнение к амплитудным найти также и фазовые характеристики. При этом придется делать некоторые предположения относительно фазовых характеристик передаточных функций $A(s)$ и $B(s)$ ¹⁾.

¹⁾ При использовании метода Боде приходится предполагать, что линейная часть изучаемой системы является минимально-фазовой. — Прим. ред.

Эти предложения могут быть проверены, так как общий фазовый сдвиг прямой цепи контура следящей системы может быть определен из экспериментальных данных.

Положение кривой, для которой E_1 постоянна, определяется построением $\ln(R/E)$ как функции $\ln E$ для нескольких значений ω . Такой логарифмический график использован на фиг. 2. Если прологарифмировать обе стороны уравнения (8) и подставить

$$E = e^{\ln E}, \quad (20)$$

то мы получим следующее соотношение:

$$\ln \frac{R}{E} = \ln |A(j\omega)B(j\omega)| + \ln F (e^{\ln E + \ln |A(j\omega)|}). \quad (21)$$

Из-за специфичной функциональной формы уравнения (21) кривые $\ln(R/E)$ как функции $\ln E$ для разных значений ω должны иметь одну и ту же форму и любая кривая может быть совмещена с любой другой кривой с помощью горизонтального и вертикального сдвига. Это подобие кривых иллюстрируется фиг. 2.

Если подобные точки на кривых постоянного значения ω соединить между собой, то результирующая кривая и будет одной из кривых постоянного значения E_1 . Другими словами, линия, вдоль которой одна из кривых постоянного значения ω может быть сдвинута для последовательного размещения ее со всеми другими кривыми постоянного значения ω , и есть линия постоянных значений E_1 . Линия $E = \bar{E}_1$ показана на фиг. 2.

Получение семейства подобных кривых является критерием применимости рассматриваемого способа оценки. Если результаты испытания следящей системы дают такое семейство кривых, то нелинейная следящая система может быть представлена нелинейной системой типа, указанной на фиг. 1, а, и система может быть аппроксимирована квазилинейной системой, изображенной на фиг. 1, б. Таким образом, путем изучения результатов испытания следящей системы может быть проверена законность основного предположения рассматриваемого способа оценки.

Оценка по характеристикам случайных сигналов

Метод оценки, при котором используются данные о случайных сигналах входа и выхода, подобен методу, в котором используются синусоидальные сигналы, только вместо возбуждения системы, как при испытании синусоидальным сигналом, и измерения основных гармоник реакции используют гауссов случайный входной сигнал и определяют среднее квадратичное значение реакции системы как функцию частотного состава и среднего квадратичного значения амплитуды входного сигнала. Использование метода случайных сигналов желательно в тех случаях, когда легче получить случайный испытательный сигнал или когда рабочий вход системы подобен гауссову случайному сигналу.

Метод случайных сигналов разработан на основе метода расчета средней квадратичной реакции некоторых типов нелинейных следящих систем, предложенного Бутоном [7]¹⁾. В методе Бутона нелинейный элемент аппроксимируется квазилинейным элементом, параметры которого являются функциями среднего квадратичного значения входного сигнала элемента. Выход нелинейного элемента e_2' опять-таки определяется выражением

$$e_2' = f(e_1'), \quad (22)$$

где e_1' — вход нелинейного элемента.

Квазилинейное представление функции $f(e_1')$ нелинейного элемента является простой функцией усиления $F(E_1)$ [7], которая определяется при помощи уравнения

$$F(E_1) = \frac{1}{\sqrt{2\pi} E_1^3} \int_{-\infty}^{\infty} e_1' f(e_1') \exp \left[-\frac{(e_1')^2}{2E_1^2} \right] de_1', \quad (23)$$

где E_1 — среднее квадратичное значение входа нелинейного элемента. Таким образом, для квазилинейной системы

$$\frac{e_2}{e_1} = F(E_1), \quad (24)$$

¹⁾ Аналогичный метод разработан И. Е. Казаковым. См., например, *Автоматика и телемеханика*, № 5 (1956).

где e_1 и e_2 — соответственно вход и выход квазилинейного элемента. При случайном входном сигнале средние квадратичные значения различных координат следящей системы могут быть определены следующими соотношениями:

$$\begin{aligned}\bar{i}^2 &= I^2, \\ \bar{e}^2 &= E^2, \\ \bar{e}_1^2 &= E_1^2, \\ \bar{r}^2 &= R^2,\end{aligned}\quad (25)$$

в которых строчными буквами обозначены мгновенные значения функций времени, а прописными — средние квадратичные значения.

Если случайный входной сигнал прикладывается к квазилинейной системе и принято основное предположение, заключающееся в том, что квазилинейная система хорошо аппроксимирует следящую систему, то соответствующие средние квадратичные значения координат в двух системах могут быть приравнены друг другу в соответствии с соотношениями

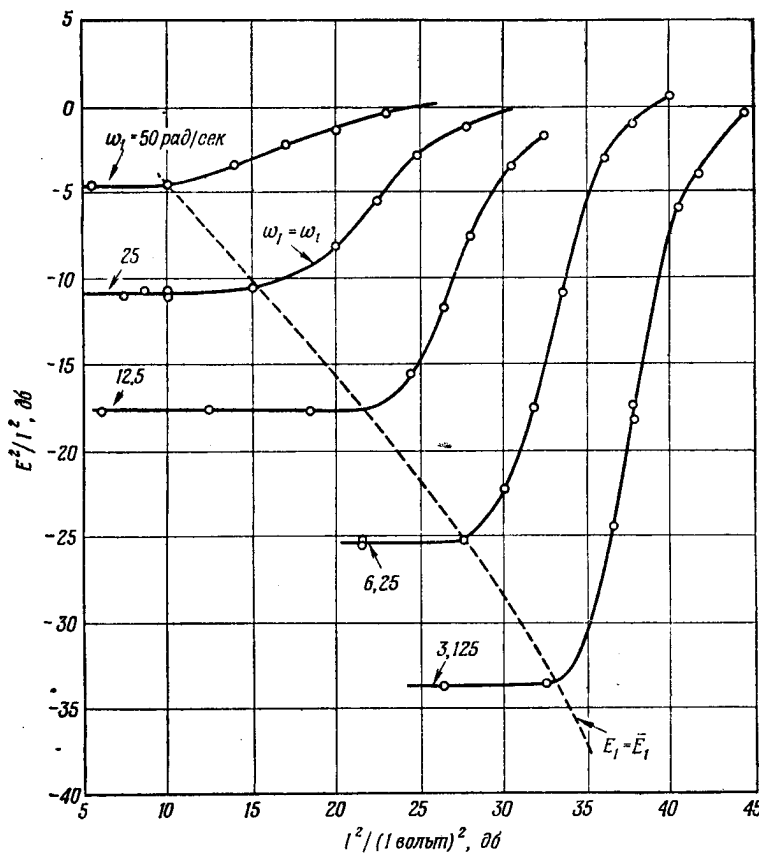
$$\begin{aligned}\bar{i}^2 &= I^2, \\ \bar{e}^2 &= E^2, \\ \bar{e}_1^2 &= E_1^2, \\ \bar{r}^2 &= R^2.\end{aligned}\quad (26)$$

Величиной для оценки следящей системы является отношение средней квадратичной ошибки системы к среднему квадратичному значению входного сигнала E^2/I^2 , измеренное как функция I^2 и ω_I . Частота ω_I характеризует полосу частот входного сигнала и точно определяется по спектру входного сигнала

$$\Phi_{II}(\omega) = \Phi_0 \frac{\omega_I^4}{\omega^4 + \omega_I^4}, \quad (27)$$

где Φ_0 — ордината функции спектральной плотности при нулевой частоте — имеет размерность $v^2/(\text{рад/сек})$, а энергия спектра поровну распределена между положительными и отрицательными значениями частоты. Спектр, определяемый

уравнением (27), был выбран потому, что он упрощает оценку системы. Рассмотрения одного спектра достаточно для иллюстрации способа.



Фиг. 6. Результаты опыта при случайном входном сигнале.

Совокупность экспериментальных данных для типичной следящей системы, которая была использована в качестве примера, приведена на фиг. 6 (та же следящая система была изучена в предыдущем разделе). Данные, приведенные на

фиг. 6, были получены при моделировании системы на электронном дифференциальном анализаторе.

Для проведения оценки необходимы следующие два основных статистических соотношения:

$$I^2 = \int_{-\infty}^{\infty} \Phi_{ii}(\omega) d\omega, \quad (28)$$

$$\Phi_{00}(\omega) = |H(j\omega)|^2 \Phi_{ii}(\omega). \quad (29)$$

Первое из них связывает среднее квадратичное значение I^2 случайного сигнала со спектральной плотностью $\Phi_{ii}(\omega)$ этого сигнала, а второе определяет спектр $\Phi_{00}(\omega)$ выходного сигнала линейной системы через передаточную функцию $|H(j\omega)|$ системы и спектр $\Phi_{ii}(\omega)$ входа.

Средние квадратичные характеристики квазилинейной системы могут быть рассчитаны с помощью уравнений (28) и (29). Среднее квадратичное значение входа может быть записано в следующей форме:

$$I^2 = \Phi_0 \int_{-\infty}^{\infty} \frac{\omega_I^4}{\omega^4 + \omega_I^4} d\omega. \quad (30)$$

Интеграл в уравнении (30) удобно вычислить с использованием таблиц интегралов, опубликованных в работе [8] (приложение С). Результат вычисления дает

$$I^2 = \Phi_0 \frac{\pi \omega_I}{\sqrt{2}}. \quad (31)$$

Средние квадратичные значения ошибки E^2 и входа нелинейного элемента E_1^2 определяются выражениями

$$E^2 = \Phi_0 \int_{-\infty}^{\infty} \frac{\omega_I^4}{\omega^4 + \omega_I^4} \left| \frac{1}{1 + A(j\omega)B(j\omega)F(E_1)} \right|^2 d\omega, \quad (32)$$

$$E_1^2 = \Phi_0 \int_{-\infty}^{\infty} \frac{\omega_I^4}{\omega^4 + \omega_I^4} \left| \frac{A(j\omega)}{1 + A(j\omega)B(j\omega)F(E_1)} \right|^2 d\omega. \quad (33)$$

Комбинируя уравнения (31) и (32), (31) и (33), получаем следующие соотношения для E^2/I^2 и E_1^2/I^2 :

$$\frac{\pi\omega_I E^2}{\sqrt{2}I^2} = \int_{-\infty}^{\infty} \frac{\omega_I^4}{\omega^4 + \omega_I^4} \left| \frac{1}{1 + A(j\omega)B(j\omega)F(E_1)} \right|^2 d\omega, \quad (34)$$

$$\frac{\pi\omega_I E_1^2}{\sqrt{2}I^2} = \int_{-\infty}^{\infty} \frac{\omega_I^4}{\omega^4 + \omega_I^4} \left| \frac{A(j\omega)}{1 + A(j\omega)B(j\omega)F(E_1)} \right|^2 d\omega. \quad (35)$$

Экспериментальная кривая E^2/I^2 , для которой E_1 постоянна, может быть использована совместно с уравнениями (34) и (35) для оценки линейных передаточных функций $A(j\omega)$ и $B(j\omega)$. Кривая постоянных E_1 ($E_1 = \bar{E}_1$), показанная пунктирной линией на фиг. 6, может быть определена с помощью метода, который будет описан позже. Для этой кривой величину E^2/I^2 можно записать так:

$$\frac{E^2}{I^2} = \frac{E^2}{I^2} (\omega_I) \Big|_{\bar{E}_1}, \quad (36)$$

а величину $1/I^2$ соответственно так:

$$\frac{1}{I^2} = \frac{1}{I^2} (\omega_I) \Big|_{\bar{E}_1}. \quad (37)$$

Обе эти величины являются функциями частоты ω_I . Подставляя (36) в (34), получим интегральное уравнение

$$\frac{\pi\omega_I}{\sqrt{2}} \left(\frac{E^2}{I^2} (\omega_I) \Big|_{\bar{E}_1} \right) = \int_{-\infty}^{\infty} \frac{\omega_I^4}{\omega^4 + \omega_I^4} \left| \frac{1}{1 + A(j\omega)B(j\omega)F(\bar{E}_1)} \right|^2 d\omega, \quad (38)$$

которое может быть решено для определения

$$A(j\omega)B(j\omega)F(\bar{E}_1).$$

Окончательная форма решения, как показано в Приложении II, может быть записана в виде

$$\left| \frac{1}{1 + A(j\omega)B(j\omega)F(\bar{E}_1)} \right|^2 = \operatorname{Re} \left[\frac{E^2}{I^2} (\omega e^{j\pi/4}) \Big|_{\bar{E}_1} \right] + \operatorname{Im} \left[\frac{E^2}{I^2} (\omega e^{j\pi/4}) \Big|_{\bar{E}_1} \right]. \quad (39)$$

Решение, выраженное уравнением (39), определено как функция величины

$$\frac{E^2}{I^2}(\omega_I) \Big|_{\bar{E}_1},$$

аргумент которой ω_I является комплексным. Так как экспериментальные данные могут быть получены только для действительных значений величины ω_I , то приходится применять аналитическое продолжение для расширения данных на комплексную плоскость. Детали этого приема изложены в Приложении II.

Чтобы найти $A(j\omega)B(j\omega)F(\bar{E}_1)$ из уравнения (39), необходимо определить передаточную функцию, удовлетворяющую соотношению

$$\left| \frac{1}{1 + A(j\omega)B(j\omega)F(\bar{E}_1)} \right| = |H_1(j\omega)|. \quad (40)$$

Функцию $H_1(j\omega)$ можно получить по методу Боде [6], определив фазовую характеристику по известной амплитудной характеристике или с помощью какого-либо другого стандартного метода. После того как будет определена функция $H_1(j\omega)$, уравнение (40) может быть решено относительно $A(j\omega)B(j\omega)$; при этом получается соотношение

$$A(j\omega)B(j\omega) = \frac{1 - H_1(j\omega)}{F(\bar{E}_1)H_1(j\omega)}. \quad (41)$$

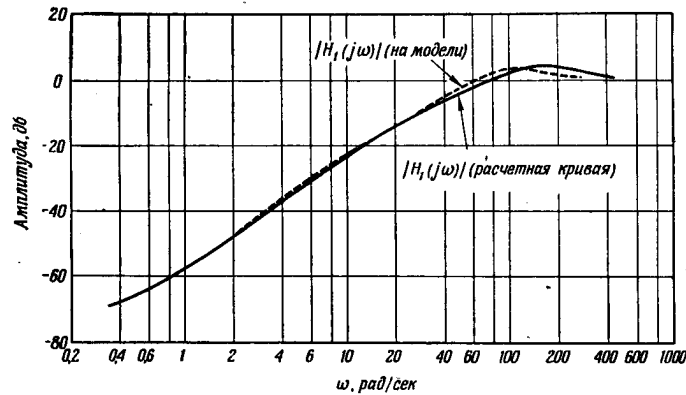
Постоянная $F(\bar{E}_1)$ является одним из производных масштабных коэффициентов, которые появляются в процессе оценки. Решение для получения $|H_1(j\omega)|$ по данным, приведенным на фиг. 6, выполнено в качестве примера в Приложении II. График $|H_1(j\omega)|$ показан на фиг. 7 и хорошо совпадает с графиком $|H_1(j\omega)|$, полученным на модели. Графическое решение уравнения (41) для определения $A(j\omega)B(j\omega)$ в предположении, что $F(\bar{E}_1) = 1$, приводит к выражению

$$A(j\omega)B(j\omega) = 995 \frac{0,077j\omega + 1}{(0,87j\omega + 1)(0,003j\omega + 1)}. \quad (42)$$

Передаточная функция, которая использовалась при моделировании, равна

$$A(j\omega)B(j\omega) = 1000 \frac{0,1j\omega + 1}{(j\omega + 1)(0,01j\omega + 1)}. \quad (43)$$

Сравнение уравнений (42) и (43) показывает, что все постоянные времени определены с приемлемой степенью точности, за исключением постоянной времени, равной 0,01 сек. Но так как наивысшая частота, для которой были определены экспериментальные данные, составляет 50 рад/сек, то ошибка не является неожиданной.



Ф и г. 7. Оценка $|H_1(j\omega)|$.

Уравнение (41) определяет произведение $A(j\omega)B(j\omega)$. Эти две передаточные функции можно разделить, используя данные, поставляемые уравнением (37). Подстановка уравнения (37) в уравнение (35) приводит к интегральному уравнению

$$\frac{\pi\omega_I}{\sqrt{2}} \left[\frac{1}{T^2}(\omega_I) \Big|_{\bar{E}_1} \right] = \int_{-\infty}^{\infty} \frac{\omega_I^4}{\omega^4 + \omega_I^4} \left| \frac{A(j\omega)}{1 + A(j\omega)B(j\omega)F(\bar{E}_1)} \right|^2 \frac{1}{\bar{E}_1^2} d\omega, \quad (44)$$

которое можно решить с помощью точно такого же приема, что и уравнение (39). При этом получают следующие соотношения:

$$\frac{1}{\bar{E}_1^2} \left| \frac{A(j\omega)}{1 + A(j\omega)B(j\omega)F(\bar{E}_1)} \right|^2 = \operatorname{Re} \left[\frac{1}{T^2}(\omega e^{j\pi/4}) \Big|_{\bar{E}_1} \right] + \operatorname{Im} \left[\frac{1}{T^2}(\omega e^{j\pi/4}) \Big|_{\bar{E}_1} \right], \quad (45)$$

или

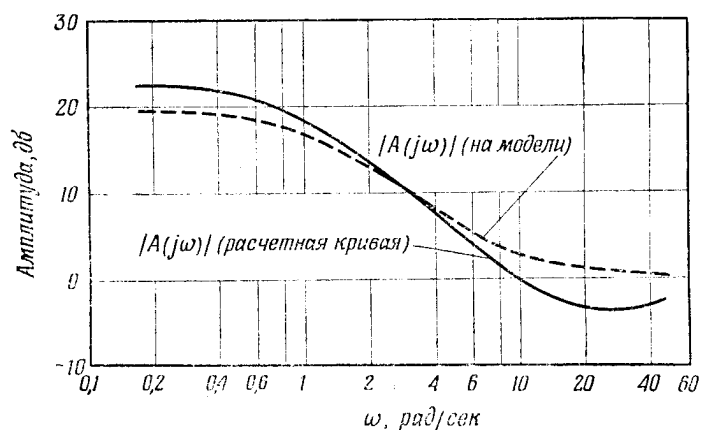
$$\frac{1}{\bar{E}_1^2} \left| \frac{A(j\omega)}{1 + A(j\omega)B(j\omega)F(\bar{E}_1)} \right|^2 = |H_2(j\omega)|^2. \quad (46)$$

Комбинируя уравнения (40) и (46), получим соотношения

$$A(j\omega) = \bar{E}_1 \frac{H_2(j\omega)}{H_1(j\omega)}, \quad (47)$$

$$B(j\omega) = \frac{1 - H_1(j\omega)}{\bar{E}_1 F(\bar{E}_1) H_2(j\omega)}, \quad (48)$$

завершив тем самым оценку линейной части квазилинейной системы. Величина \bar{E}_1 в этих двух соотношениях является второй постоянной — масштабным коэффициентом.

Фиг. 8. Оценка $|A(j\omega)|$.

Решение уравнений (45)—(47) для определения $|A(j\omega)|$ по данным, приведенным на фиг. 6, дает функцию, показанную на фиг. 8. Оценка для $|A(j\omega)|$ значительно отличается от величины $|A(j\omega)|$, использованной при моделировании, но передаточная функция, найденная при оценке, с несомненностью указывает на то, какие члены уравнения (42) должны быть связаны с $A(j\omega)$. Поэтому $A(j\omega)$ и $B(j\omega)$ могут быть

разделены и для них могут быть записаны следующие выражения:

$$A(j\omega) = 9,95 \frac{0,077 j\omega + 1}{0,87 j\omega + 1}, \quad (49)$$

$$B(j\omega) = \frac{100}{j\omega(0,003 j\omega + 1)}, \quad (50)$$

где второй масштабный коэффициент \bar{E}_1 выбран так, чтобы получить для $B(j\omega)$ на низких частотах выражение $100/s$.

Для оценки функции усиления квазилинейного элемента используются данные одной из кривых для постоянного значения ω . Для этого использована кривая на фиг. 6, для которой

$$\omega_I = \omega_1, \quad (51)$$

а величина E^2/I^2 , соответствующая этой кривой, обозначена следующим образом:

$$\frac{E^2}{I^2} = \frac{E^2}{I^2} (I^2) |_{\omega_1}, \quad (52)$$

Подстановка величины ω_1 и выражений для $A(j\omega)$ и $B(j\omega)$, полученных из уравнений (47) и (48), в уравнение (34) и оценка интеграла определяют E^2/I^2 как функцию $F(E_1)$. Обратная функция $F(E_1)$ как функция E^2/I^2 обозначена через $S_1(E^2/I^2)$:

$$F(E_1) = S_1\left(\frac{E^2}{I^2}\right). \quad (53)$$

Уравнения (52) и (53) при формальном решении дают соотношение для $F(E_1)$

$$F(E_1) = S_1\left[\frac{E^2}{I^2} (I^2) |_{\omega_1}\right], \quad (54)$$

определяющее $F(E_1)$ по экспериментальным данным, которые дают уравнение (52).

Фактическое решение удобно провести в табличной форме, что иллюстрируется табл. 3.

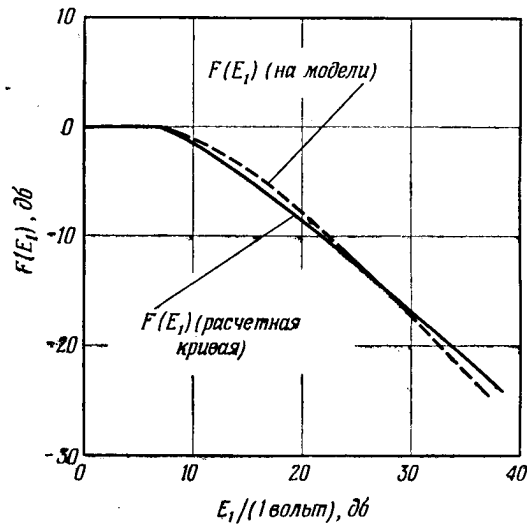
Величины $F(E_1)$ и E^2/I^2 , полученные из уравнения (53), представлены в первых двух столбцах. Экспериментальные данные (уравнение (52)) используются для определения величины I^2 в третьем столбце, которая соответствует величинам E^2/I^2 во втором столбце. Оценка завершается постановкой $A(j\omega)$, $B(j\omega)$ и ω_1 в уравнение (35) для определения

Таблица 3

ОЦЕНКА $F(E_1)$ ПО ХАРАКТЕРИСТИКАМ СЛУЧАЙНЫХ СИГНАЛОВ

$F(E_1)$, дб	E^2/P , дб	$P/(1 \sigma)^2$, дб	E_1^2/P^2 , дб	$E_1/(1 \sigma)$, дб
0	-9,8	17,2	-10,2	7,1
-0,9	-9,1	18,8	-9,5	9,2
-1,9	-8,3	19,8	-8,7	11,1
-6,0	-5,4	22,5	-5,6	16,9
-14,0	-1,2	27,3	-0,8	26,6
-20,	0,5	31,8	1,8	33,6

величины E_1^2/P^2 как функции $F(E_1)$, значения которой приведены в четвертом столбце. Затем может быть рассчитан вход E_1 как корень квадратный из произведения величин в третьем

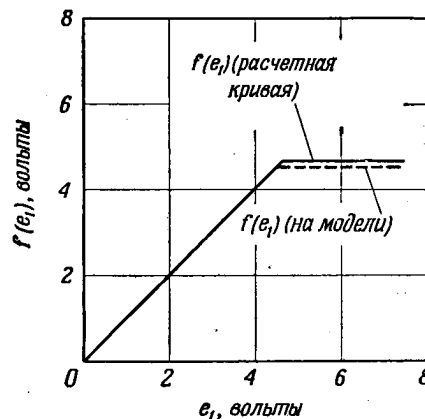


Фиг. 9. Квазилинейная функция усиления по данным испытания случайным входным сигналом.

и четвертом столбцах; результат расчета записывается в последнем столбце. Для $F(E_1)$ и функции усиления, использованной при моделировании, получено хорошее совпа-

дение; обе кривые показаны на фиг. 9. Нелинейная характеристика, полученная из $F(E_1)$ с помощью методов, изложенных в Приложении I, показана на фиг. 10 вместе с характеристикой, использованной при моделировании. Эти характеристики также хорошо совпадают.

Определением $F(E_1)$ заканчивается оценка квазилинейной системы, если не считать того, что необходимо еще определить положение линии постоянных значений E_1 . Определение положения этой линии не так просто, как для случая



Фиг. 10. Характеристика по данным испытания случайным входным сигналом.

синусоидальных сигналов, потому что кривые, характеризующие случайные данные, не подобны. Но все-таки некоторые способы определения линии постоянных значений E_1 будут рассмотрены.

В некоторых случаях влияние различных участков характеристики нелинейного элемента может быть замечено на каждой кривой E^2/I^2 как функции I^2 . Эти влияния дают простой способ определения кривой постоянных значений E_1 . В примере, показанном на фиг. 6, резкий излом характеристики нелинейного элемента приводит к появлению заметного изменения кривизны каждой из кривых E^2/I^2 и линия постоянных значений $E_1 = \bar{E}_1$ проводится через точки, в которых изменяется кривизна кривых.

Некоторые следящие системы имеют заметную область линейного действия, что может быть использовано для определения положения кривой постоянных значений E_1 . В области линейного действия функция $F(E_1)$ постоянна и существует семейство горизонтальных кривых E^2/I^2 (см. фиг. 6), так как

$$\frac{dF(E_1)}{dE_1} = 0. \quad (55)$$

В этой линейной области величина E^2/I^2 может быть оценена как функция ω_1 , и эти данные можно использовать при решении уравнения (38) для определения $A(j\omega)B(j\omega)$. Решение уравнения (38) для определения $A(j\omega)B(j\omega)$ возможно потому, что $F(E_1)$ постоянна в линейной области. Но E_1 не постоянна в этой области, и, следовательно, решение уравнения (44) для разделения $A(j\omega)$ и $B(j\omega)$ не может быть выполнено непосредственно. Вместо этого выражение $A(j\omega)B(j\omega)$ подставляют в уравнение (34) вместе с некоторым постоянным значением $F(E_1)$ и получают E^2/I^2 как функцию ω_1 . Постоянная величина $F(E_1)$ должна быть выбрана в нелинейной области. Функция E^2/I^2 может быть объединена с экспериментальными данными для определения линии постоянных значений E_1 .

Если нелинейный элемент не имеет области линейного действия и у него не существует резко выраженной характеристики, то для определения линии постоянных значений E_1 необходимо применять итерационный процесс. Итерация начинается с задания приближенного положения линии постоянных значений E_1 . Данные, взятые с этой линии, используются для оценки $A(j\omega)B(j\omega)$, и таким образом определяется положение второй линии постоянных значений E_1 . Данные, взятые со второй линии, используются для повторной оценки $A(j\omega)B(j\omega)$, что позволяет получить новую оценку положения подлинной линии. Этот процесс повторяется до тех пор, пока положение линии не станет вполне определенным.

Заключение

Методы оценки по характеристикам синусоидальных и случайных сигналов являются одним из примеров применения квазилинейных методов расчета нелинейных систем. Эти методы в сочетании с частным предположением отно-

сительны формы системы используются для того, чтобы получить специальный способ оценки ограниченного, но важного класса нелинейных следящих систем. Таким путем демонстрируется разумность и польза квазилинейного приближения для решения задач оценки, что является основным моментом в настоящем исследовании. Успешное применение квазилинейных методов для оценки следящих систем одного определенного класса указывает на то, что последующая работа с успехом может быть направлена на распространение изложенного метода применительно к другим видам следящих систем.

Способ, основанный на использовании характеристик синусоидальных сигналов, представляет собой непосредственное развитие существующих методов и легко понятен. Для получения требуемых данных можно использовать существующее оборудование для измерения амплитуды и фазы. Таким образом, этот способ может быть непосредственно применен к решению задач, представляющих практический интерес.

Практическая важность способа, основанного на изучении характеристик случайных сигналов, не так велика, как предыдущего способа, потому что этот способ более сложен как с расчетной, так и с экспериментальной стороны. Но изучение случайных сигналов представляет теоретический интерес и может привести к разработке важных методов оценки.

ПРИЛОЖЕНИЕ I

Оценка нелинейных характеристик

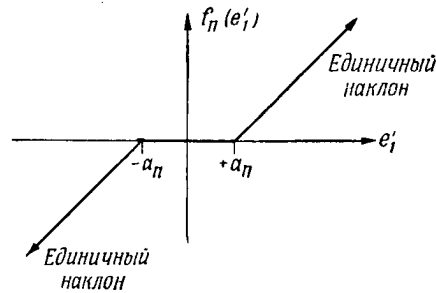
Методы оценки квазилинейной системы, приближенно представляющей нелинейную систему, уже были изложены. Для завершения оценки следящей системы необходимо по функции усиления квазилинейного элемента определить характеристику нелинейного элемента. В случае использования синусоидальных сигналов эти две характеристики связаны друг с другом уравнением (2), которое мы здесь заново выпишем:

$$F(E_1) = \frac{2}{\pi E_1} \int_{-\pi/2}^{\pi/2} f(E_1 \sin \theta) \sin \theta d\theta. \quad (56)$$

Если квазилинейная функция усиления $F(E_1)$ задана, а характеристика нелинейного элемента $f(e_1')$ неизвестна, то уравнение (56) является интегральным уравнением, для приближенного решения которого существует много способов. Простой и вполне удовлетворительный графический способ решения состоит в том, что функцию $F(E_1)$ аппроксимируют суммой

$$F(E_1) \simeq \sum_{n=1}^N \beta_n F_n(E_1). \quad (57)$$

Аппроксимация выполняется путем выбора подходящих коэффициентов β_n и функций $F_n(E_1)$. Функции $F_n(E_1)$ полу-



Фиг. 11. Аппроксимирующая функция.

чаются при подстановке известных аппроксимирующих функций $f_n(e_1')$ в уравнение (56) и вычислении полученного интеграла

$$F_n(E_1) = \frac{2}{\pi E_1} \int_{-\pi/2}^{\pi/2} f_n(E_1 \sin \theta) \sin \theta d\theta. \quad (58)$$

Характеристика $f(e_1')$ затем определяется из соотношения

$$f(e_1') \simeq \sum_{n=1}^N \beta_n f_n(e_1'). \quad (59)$$

Полубесконечная наклонная прямая, показанная на фиг. 11, является удобной аппроксимирующей функцией $f_n(e_1')$. Если

эту функцию подставить в уравнение (58), то получим следующий интеграл:

$$F_n(E_1) = F^*\left(\frac{E_1}{\alpha_n}\right) = \frac{4\alpha_n}{\pi E_1} \int_{\arcsin \alpha_n/E_1}^{\pi/2} \left(\frac{E_1}{\alpha_n} \sin \theta - 1\right) \sin \theta d\theta, \quad (60)$$

и, как это видно из уравнения (60), все функции $F_n(E_1)$ могут быть записаны в виде одной функции $F^*(E_1/\alpha_n)$ от аргумента E_1/α_n . Таким образом, $F(E_1)$ может быть переписана в виде суммы

$$F(E_1) \simeq \sum_{n=1}^N \beta_n F^*\left(\frac{E_1}{\alpha_n}\right) \quad (61)$$

и решение интегрального уравнения можно выполнить с помощью подстановки величин α_n и β_n , аппроксимирующих $F(E_1)$. Коэффициенты α_n и β_n можно выбрать, построив кривые $\ln F(E_1)$ как функцию $\ln E_1$ и $\ln F^*(E_1/\alpha_n)$ как функцию E_1/α_n . Эти два графика можно, наложив их друг на друга, перемещать до наилучшего совпадения и таким образом определить α_1 и β_1 . Затем вычерчивается функция $F(E_1) - \beta_1 F^*(E_1/\alpha_1)$ и процесс совмещения продолжается для определения α_2 и β_2 . Процесс совмещения может быть повторен столько раз, сколько необходимо для получения хорошей аппроксимации $F(E_1)$. Так как этот процесс приводит к кусочно-линейной аппроксимации $f(e_1')$, то любая функция $f(e_1')$ может быть аппроксимирована, если использовать достаточное число отрезков прямой.

Квазилинейная функция усиления, определенная в примере и показанная на фиг. 3, может быть аппроксимирована двумя членами суммы в уравнении (61). Эта сумма может быть записана в виде

$$F(E_1) \simeq F^*\left(\frac{E_1}{0}\right) + F^*\left(\frac{E_1}{10}\right). \quad (62)$$

Функции $F^*(E_1/0)$ и $F^*(E_1/10)$ также показаны на фиг. 3. Разность этих функций дает $F(E_1)$, хотя эта разность непосредственно не видна на графике из-за применения логарифмических координат.

Нелинейная характеристика $f(e_1')$, полученная при подстановке величин α_n и β_n из уравнения (62) в уравнение (59),

вычерчена на фиг. 4. В пределах точности построения эта характеристика совпадает с характеристикой, принятой при моделировании.

ПРИЛОЖЕНИЕ II

Решение интегральных уравнений, определяющих передаточную функцию

При оценке, основанной на использовании характеристик случайных сигналов, требуется решать интегральные уравнения (38) и (44). Эти уравнения имеют следующую форму:

$$\frac{\pi\omega_I}{\sqrt{2}} [R(\omega_I)] = \int_{-\infty}^{\infty} \frac{\omega_I^4}{\omega^4 + \omega_I^4} |H(j\omega)|^2 d\omega, \quad (63)$$

где неизвестная функция $H(j\omega)$ является передаточной функцией следящей системы, а $R(\omega_I)$ — экспериментально определенное отношение средних квадратичных значений напряжения.

Форма ядра уравнения (63) дает возможность применить для его решения преобразование Лапласа. Благодаря четности подынтегрального выражения уравнение (63) может быть переписано в виде

$$\frac{\pi\omega_I}{2\sqrt{2}} [R(\omega_I)] = \int_0^{\infty} \frac{\omega_I^4}{\omega^4 + \omega_I^4} |H(j\omega)|^2 d\omega. \quad (64)$$

После замены переменных $\omega_I = e^{-\tau}$ и $\omega = e^{-\sigma}$ уравнение принимает следующий вид:

$$\frac{\pi}{2\sqrt{2}} e^{-\tau} R(e^{-\tau}) = \int_0^{\infty} \frac{1}{1 + e^{4(\tau-\sigma)}} |H(je^{-\sigma})|^2 e^{-\sigma} d\sigma. \quad (65)$$

Ядро уравнения (65) является функцией разности $\tau - \sigma$ двух переменных. Следовательно, это интегральное уравнение можно решить с помощью двустороннего преобразования Лапласа с использованием теоремы свертки (9).

Используя эту теорему, решение уравнения (65) можно записать так:

$$|H(je^{-\sigma})|^2 e^{-\sigma} = \frac{1}{2\pi j} \int_{c-j\infty}^{c+j\infty} \frac{C(s)}{D(s)} e^{s\sigma} ds, \quad (66)$$

где $C(s)$ — преобразование левой части уравнения (65)

$$C(s) = \int_{-\infty}^{\infty} \frac{\pi}{2\sqrt{2}} e^{-\tau} R(e^{-\tau}) e^{-\tau s} d\tau, \quad (67)$$

а $D(s)$ — преобразование ядра уравнения (65)

$$D(s) = \int_{-\infty}^{\infty} \frac{1}{1+e^{4\sigma}} e^{-\sigma s} d\sigma. \quad (68)$$

Если преобразования $C(s)$ и $D(s)$ имеют общую полосу сходимости, то уравнение (66) определяет прямой метод нахождения решения интегрального уравнения (63), так как интегралы в уравнениях (66)—(68) могут быть найдены в численном виде. Но так как численный расчет утомителен, то здесь приводится более простой способ, включающий в себя аналитическое продолжение функции $R(\omega_l)$ на комплексные значения ω_l .

Интеграл в уравнении (68) можно вычислить с помощью метода контурных интегралов, что дает соотношение

$$D(s) = -\frac{\pi}{4 \sin(\pi/4) s}. \quad (69)$$

Величину $C(s)/D(s)$ можно теперь записать так:

$$\frac{C(s)}{D(s)} = -\frac{2}{j\pi} (e^{(j\pi/4)s} - e^{-(j\pi/4)s}) C(s). \quad (70)$$

Подстановка уравнения (70) в (66) дает интеграл, который можно вычислить, выразив результат через $R(\omega_l)$, что дает

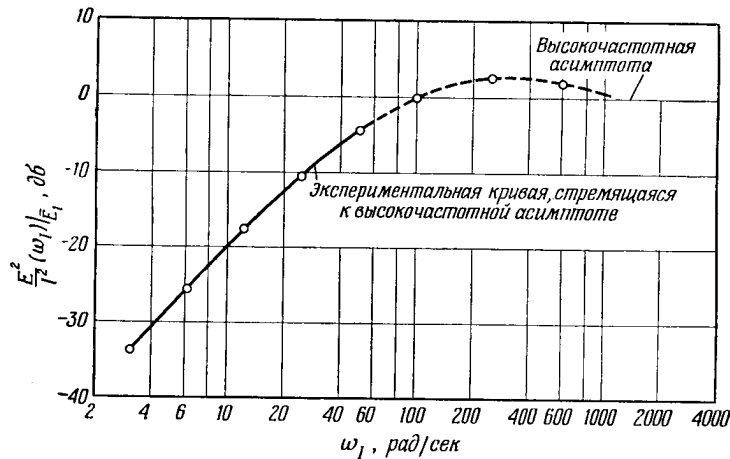
$$\begin{aligned} |H(je^{-\sigma})|^2 e^{-\sigma} &= \frac{j}{\sqrt{2}} [e^{-\sigma} e^{-j\pi/4} R(e^{-\sigma} e^{-j\pi/4}) - \\ &- e^{-\sigma} e^{j\pi/4} R(e^{-\sigma} e^{j\pi/4})]. \end{aligned} \quad (71)$$

После замены переменной $\omega = e^{-\sigma}$ перепишем уравнение (71) в такой форме:

$$|H(j\omega)|^2 = \operatorname{Re} [R(\omega e^{j\pi/4})] + \operatorname{Im} [R(\omega e^{j\pi/4})]. \quad (72)$$

Уравнение (72) определяет $|H(j\omega)|^2$ как функцию вещественной и мнимой частей $R(\omega e^{j\pi/4})$. Экспериментальные данные определяют $R(\omega_l)$ только для действительных значений ω_l . Поэтому необходимо использовать аналитическое

продолжение для расширения экспериментальных данных на комплексные значения ω_I . Это аналитическое продолжение удобно выполнить, аппроксимируя $R(\omega_I)$ аналитической функцией. Вполне определенная форма этой функции выбирается потому, что $H(j\omega)$ является рациональной функцией



Фиг. 12. Данные, полученные из кривой с постоянным E_1 .

$j\omega$ и, следовательно, $R(\omega_I)$ является рациональной функцией ω_I для положительных значений ω_I , как это показано в таблице интегралов в приложении С работы [8]. Поэтому форма

$$R(\omega_I) = \frac{b_0 + b_1\omega_I + b_2\omega_I^2 + \dots + b_n\omega_I^n}{a_0 + a_1\omega_I + a_2\omega_I^2 + \dots + a_m\omega_I^m} \quad (73)$$

является вполне подходящей. Коэффициенты a и b выбираются так, чтобы $R(\omega_I)$ совпадала с экспериментальными данными. Уравнение (73) можно разложить на ряд линейных алгебраических уравнений, которые можно решить для получения точек совпадения.

В качестве примера найдем решение уравнения (38), используя данные, представленные на фиг. 6. Отношение E^2/I^2 для постоянного значения E_1 как функции ω_I графически представлено на фиг. 12. Процесс приближения $R(\omega_I)$ к этой функции начинается с того, что уравнение (73)

14. Метод оценки нелинейных следящих систем 471

переписывается в следующей форме:

$$\begin{aligned} & \omega_I R(\omega_I) a_1 + \omega_I^2 R(\omega_I) a_2 + \omega_I^3 R(\omega_I) a_3 + \\ & + \omega_I^4 R(\omega_I) a_4 + \omega_I^5 R(\omega_I) a_5 - \omega_I^3 b_3 - \\ & - \omega_I^4 b_4 - \omega_I^5 b_5 = \omega_I^6 [1 - R(\omega_I)] + 10^{-6} \omega_I^2 - R(\omega_I). \end{aligned} \quad (74)$$

Уравнение (74) получено в предположении, что передаточная функция прямой цепи следящей системы стремится к нулю при высоких частотах и к 1000/s при низких частотах. Таким образом, из уравнения (34) следует

$$\frac{E^2}{I^2}(\omega_I) \rightarrow 1, \quad \text{когда } \omega_I \rightarrow \infty,$$

$$\frac{E^2}{I^2}(\omega_I) \rightarrow \frac{\omega_I^2}{10^6}, \quad \text{когда } \omega_I \rightarrow 0.$$

Величины ω_I и $R(\omega_I)$, взятые на фиг. 12 в точках, обозначенных кружками, подставляются в уравнение (74); в результате получаем 8 уравнений, которые можно записать в матричной форме:

$$\begin{bmatrix} 13,1 & 41 & 1,28 & 0,4 & 0,0125 & 3\,050 & 954 & 29,8 \\ 17,8 & 111 & 6,96 & 4,35 & 0,272 & 2\,240 & 1\,520 & 95,4 \\ 21,2 & 265 & 33,2 & 41,5 & 5,19 & 1\,950 & 2\,440 & 305 \\ 210 & 5\,250 & 1\,310 & 3\,280 & 820 & 15\,600 & 39\,000 & 9\,770 \\ 17,5 & 875 & 437 & 2\,180 & 1\,090 & 1\,250 & 6\,250 & 3\,120 \\ 97 & 9\,700 & 9\,700 & 97\,000 & 97\,000 & 10\,000 & 100\,000 & 100\,000 \\ 4,2 & 1\,060 & 2\,656 & 66\,400 & 166\,000 & 1\,560 & 39\,100 & 97\,700 \\ 0,918 & 551 & 3\,300 & 198\,000 & 1\,189\,728 & 2\,160 & 129\,600 & 778\,000 \end{bmatrix} \times$$

$$\times \begin{bmatrix} a_1 \\ a_{12} \\ 10^3 a_3 \\ 10^3 a_4 \\ 10^5 a_5 \\ -10^3 b_3 \\ -10^3 b_4 \\ -10^5 b_5 \end{bmatrix} = \begin{bmatrix} -4,1 \\ -2,81 \\ -1,68 \\ -8,31 \\ -0,387 \\ -0,93 \\ -1,72 \\ -24,7 \end{bmatrix} \quad (75)$$

В строки и столбцы уравнения (75) были введены различные множители, для того чтобы понизить порядок коэффициентов матрицы. Решение, полученное с использованием

метода Краута [10] и арифмометра, дало следующие результаты:

$$\begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \\ b_3 \\ b_4 \\ b_5 \end{bmatrix} = \begin{bmatrix} 3,27 \cdot 10^{-2} \\ 3,56 \cdot 10^{-4} \\ -49,6 \cdot 10^{-6} \\ 28,0 \cdot 10^{-8} \\ -24,1 \cdot 10^{-10} \\ 15,8 \cdot 10^{-6} \\ -28,2 \cdot 10^{-8} \\ -24,3 \cdot 10^{-10} \end{bmatrix} \quad (76)$$

Подстановка аналитической функции (73) с коэффициентами из уравнения (76) в уравнение (72) дает величину $|H_1(j\omega)|$, график которой приведен на фиг. 7, и завершает решение уравнения (38).

ЛИТЕРАТУРА

1. Gehmlich D. K., Van Valkenburg M. E., *Trans. AIEE*, Part II, **73**, 232 (1954).
2. Kochenburger R. J., *Trans. AIEE*, Part I, **69**, 270 (1950).
3. Johnson E. C., *Trans. AIEE*, Part II, **71**, 169 (1952).
4. Теория следящих систем, под ред. Х. Джеймса, Н. Никольса и Р. Филлипса, Издательство иностранной литературы, 1951.
5. Mathews M. V., «A Technique for the Evaluation of Nonlinear Servomechanisms», Диссертация, Mass. Inst. Technol., 1954, p. 13—14.
6. Боде Г., Теория цепей и проектирование усилителей с обратной связью, Издательство иностранной литературы, 1948.
7. Borton R. C., Jr., «Nonlinear Control Systems with Statistical Inputs», Dynamic Analysis and Control Laboratory, Rep. № 61, Mass. Inst. Technol., March 1, 1952.
8. Borton R. C., Jr., Mathews M. V., Seifert W. W., «Nonlinear Servomechanisms with Random Inputs», Dynamic Analysis and Control Laboratory, Rep. № 70, Mass. Inst. Technol., Aug. 20, 1953.
9. Ван-дер-Поль Б., Бреммер Г., Операционное исчисление, основанное на двустороннем преобразовании Лапласа, Издательство иностранной литературы, 1952.
10. Stout P., *Trans. AIEE*, **60**, 1235 (1941).

№ генератора	Применение в условиях и лабораториях	Регулировка частоты									Регулировка амплитуды			Регулировка уровня			Динамические характеристики		Предельные размеры устройства		Промышленные образцы			Способ задания частоты	Тип зависящего устройства	Регион	Где используется	Классификация по виду устройства	№ генератора				
		интервал, Гц			пере-рывная	кван-тизация	ступен-чатая	интервал, мВ/см			пере-рывная	кван-тизация	ступен-чатая	интервал, дБ/см			длина по оси частот	площадь по оси частот	инерт-ность, мс	габариты, г	вес, кг	затраты, руб.	стоимость, руб.							полная стоимость, руб.			
		мин.	макс.	мин.				макс.	мин.	макс.				мин.	макс.																		
1	+	+	5.0 · 10 ⁻²	12	+	+	+	0	1.0	+	+	+	7 · 10 ⁻²	2.1	+	+	+	+	1	3	25 × 25 × 38	13	4	112	125	Счетчик числа оборотов	Электронный	Бобли	Высокочастотное	—	1		
2	+	+	2.8 · 10 ⁻²	0.5	+	+	+	0	0.7	+	+	+	0	1.4	+	+	+	+	1	22	20 × 35 × 50	9	8	235	Система Метрон	Пневматический	Фоксборо	Химическая промышленность	—	2			
3	+	+	1.4 · 10 ⁻²	18	+	+	+	0	1.0	+	+	+	0	2.1	+	+	+	+	5	22	30 × 45 × 60	54	240	300	Привод Грама	Электронный	»	То же	—	3			
4	+	+	8.4 · 10 ⁻³	15	+	+	+	0	1.0	+	+	+	0	2.1	+	+	+	+	5	5	35 × 55 × 70	68	240	300	То же	»	»	»	—	4			
5	+	+	8.4 · 10 ⁻³	10	+	+	+	0.2	0.8	+	+	+	0.2	1.4	+	+	+	+	4	7	25 × 30 × 75	27	100	500	»	»	»	»	—	5			
6	+	+	2.3 · 10 ⁻²	15	+	+	+	3.5 · 10 ⁻³	0.35	+	+	+	0	1.4	+	+	+	+	9	10	15 × 25 × 50	11	40	370	500	Система Метрон	»	Фоксборо	»	—	6		
7	+	+	1.6 · 10 ⁻²	15	+	+	+	1.4 · 10 ⁻³	0.14	+	+	+	0.2	1.0	+	+	+	+	6	12	30 × 35 × 43	23	100	1000	Широкополосный фазовый	»	»	»	»	—	7		
8	+	+	8.4 · 10 ⁻³	15	+	+	+	0	0.5	+	+	+	0.14	1.0	+	+	+	+	6	12	25 × 45 × 55	23	100	1200	Редуктор	»	Гартман и Браун	Приборостроение	—	8			
9	+	+	5.6 · 10 ⁻³	0.5	+	+	+	0	0.4	+	+	+	0.2	1.0	+	+	+	+	6	12	20 × 35 × 55	32	240	1000	Двигатель постоянного тока	Пневматический	»	Химическая промышленность	+	9			
10	+	+	0.17	9	+	+	+	0	0.4	+	+	+	0.2	1.0	+	+	+	+	6	12	28 × 30 × 33	6.5	50	175	Двухфазный фазовый	»	»	»	»	—	10		
11	+	+	6.2 · 10 ⁻³	5.0 · 10 ⁻³	+	+	+	0	0.4	+	+	+	0.2	1.0	+	+	+	+	6	12	15 × 20 × 33	13	75	250	Редуктор	»	»	»	»	—	11		
12	+	+	1.0 · 10 ⁻²	3.3 · 10 ⁻³	+	+	+	0	0.4	+	+	+	0.2	1.0	+	+	+	+	6	12	15 × 20 × 38	10	100	300	То же	»	»	»	»	—	12		
13	+	+	5.6 · 10 ⁻³	12	+	+	+	3.5 · 10 ⁻⁴	0.5	+	+	+	7 · 10 ⁻⁴	1.3	+	+	+	+	4	7	25 × 30 × 55	38	100	3000	Система Метрон	Электронный	Мини-Комутал	»	—	13			
14	+	+	1.4 · 10 ⁻²	20	+	+	+	0	0.7	+	+	+	0	1.4	+	+	+	+	4	7	60 × 60 × 60	18	150	400	Привод Грама	»	Тейлор и Свартау	Нефтяная промышленность	—	14			
15	+	+	2.8 · 10 ⁻²	10	+	+	+	0	0.7	+	+	+	0	1.4	+	+	+	+	1	—	15 × 15 × 33	4.5	60	400	Двигатель постоянного тока	»	»	»	»	—	15		
16	+	+	2.8 · 10 ⁻²	10	+	+	+	1.7 · 10 ⁻³	0.17	+	+	+	0.2	1.0	+	+	+	+	180	—	30 × 45 × 60	23	60	400	Редуктор	Фототехнический	Тейлор	Приборостроение	—	16			
17	+	+	8.4 · 10 ⁻³	225	+	+	+	1.7 · 10 ⁻³	0.17	+	+	+	0.2	1.0	+	+	+	+	180	—	30 × 45 × 60	23	60	400	Звуковой генератор, двигатель постоянного тока	Электронный	»	»	»	—	17		
18	+	+	3.6 · 10 ⁻³	0.74	+	+	+	7 · 10 ⁻³	0.4	+	+	+	0	1.4	+	+	+	+	—	—	40 × 45 × 60	21	20	253	330	Редуктор	»	»	»	»	—	18	
19	+	+	0.2	35	+	+	+	0	3.5	+	+	+	0	210	+	+	+	+	—	—	35 × 40 × 40	11	30	50	100	Двигатель постоянного тока	Электронный	Нет	»	»	—	19	
20	+	+	0.17	250	+	+	+	0	3.5	+	+	+	0	16.5	+	+	+	+	—	—	10 × 15 × 25	6.5	20	50	100	То же	»	»	»	»	—	20	
21	+	+	100	2000	+	+	+	0	10.5	+	+	+	0	8.8	+	+	+	+	—	—	20 × 30 × 30	6.5	12	50	75	»	»	»	»	»	—	21	
22	+	+	40	40	+	+	+	0	40	+	+	+	0	8.8	+	+	+	+	—	—	15 × 15 × 15	4.5	12	50	1000	»	»	»	»	»	—	22	
23	+	+	5.0 · 10 ⁻²	20	+	+	+	0 см	2.5 см	+	+	+	0 см	2.5 см	+	+	+	+	—	—	15 × 30 × 50	50	12	50	1000	»	»	»	»	»	—	23	
24	+	+	8.3 · 10 ⁻³	18	+	+	+	0 см	2.5 см	+	+	+	0 см	2.5 см	+	+	+	+	—	—	20 × 45 × 48	85	12	50	3000	Привод Вилкеса	»	Нет	»	»	»	—	24

Примечание. Пустое означает «минимум», пустые клетки означают на отсутствие данных.

15. ГЕНЕРАТОРЫ СИНУСОИДАЛЬНЫХ КОЛЕБАНИЙ¹⁾
(Обзор пневматических, механических и электрических устройств
для получения частотных характеристик)

Д. Сент-Клэр, Л. Эррат, С. Гиллеспи

Введение

За последние несколько лет методы частотного анализа нашли широкое применение в автоматике. В данной статье собраны сведения о различных схемах и приборах, которые применяются или могут применяться в качестве источников синусоидальных колебаний, необходимых для экспериментального определения частотных характеристик промышленных систем. Авторы полагают, что эти сведения окажутся полезными всем, кто имеет дело с частотными методами исследования, и представят интерес также для широкого круга специалистов.

Настоящий обзор имеет в основном описательный характер, так как затрагивает лишь общие свойства генераторов. Другие свойства экспериментального оборудования, применяемого для определения динамических параметров, в котором генератор синусоидальных колебаний может быть лишь частью, упомянуты очень кратко. Авторы намеренно опустили здесь пространные обсуждения преимуществ, которые проявляются при применении частотных характеристик, поскольку большое количество используемых генераторов колебаний и интерес специалистов к частотным методам говорят сами за себя.

Вся рассматриваемая аппаратура подразделяется на пневматическую, механическую и электрическую в соответствии с видом ее основного выходного сигнала. Статья разбита на соответствующие разделы, посвященные отдельным типам устройств.

¹⁾ St. Clair D. W., Erath L. W., Gillespie S. L., *Trans. ASME*, 76, № 8, 1177—1183 (1954).

Пневматические устройства

Наиболее распространенный способ получения пневматических синусоидальных колебаний состоит в том, что вращательное движение сначала превращается в простое гармоническое движение, а затем уже преобразуется в пневматический сигнал. В одних случаях для возбуждения электропневматического реле используется электронный генератор. В других случаях для получения синусоидальных колебаний непосредственно применяется воздушный компрессор. Было изготовлено также пневматическое устройство, использующее принцип релаксационного генератора. Имеются некоторые сведения о самовозбуждающемся устройстве типа сопло — заслонка.

В таблицах (№№ 1 — 21) собраны данные о различных пневматических генераторах. Применяемые в таблице обозначения требуют некоторых пояснений. „Уровень“¹⁾ характеризует предельные возможности оборудования, включая амплитуду синусоидальных колебаний; „динамическая характеристика“ относится к фиксированной амплитуде в $0,07 \text{ кг/см}^2$, (1 фунт/дюйм^2); „плоская“ означает, что амплитуда выходного сигнала изменяется лишь в пределах $1 \div \sqrt{2}$ или $\sqrt{2}/2 \div 1$.

Обычно устройство состоит из электродвигателя, механизма для изменения скорости и редуктора для вращения вала с желаемой угловой частотой. Вращение оси преобразуется в простое гармоническое движение посредством эксцентрика (или шатунно-кривошипного механизма) и далее передается к пневматическому устройству, преобразующему перемещение в изменение давления воздуха. Уровень и амплитуда колебаний устанавливаются с помощью соответствующих рычагов.

Модели, которые сильно отличаются от обычного устройства, подробнее описываются ниже.

Некоторые дополнительные сведения, представляющие интерес, приводятся ниже под номерами, соответствующими номеру генератора в таблице.

¹⁾ Сигнал на выходе генератора может быть представлен в виде суммы постоянной составляющей („уровень“) и переменной синусоидальной составляющей, характеризующейся амплитудой. — *Прим. ред.*

1. См. [1]
4. Это устройство является взрывобезопасным. Оно приспособлено для генерирования электрических сигналов; когда это возможно, применяется электронный генератор.
5. Эта модель дана в дополнение к аналогичной модели под номером 13.
7. Для работ в полевых условиях может быть присоединен пневмодвигатель. Дополнительно разрабатывается устройство, основанное на принципе релаксационного генератора. Оно было испытано в полевых условиях.
8. Это новый образец аппаратуры. Запись фотографическая, но не требует проявления. Скорость ленты синхронизируется с частотой и изменяется шестью ступенями с помощью коробки передач. Вращение оси преобразуется в электрический сигнал посредством потенциометра, который управляет электропневматическим реле и самописцем без дополнительного усиления. Выход электропневматического реле питает трехтрубный клапан, который вводит колебания в процесс, не нарушая герметичности системы.
9. См. [2]. Это устройство имеет воздушную или азотную продувку. Запись ведется со скоростью, пропорциональной частоте. Фирма использует несколько подобных генераторов. Конструируется новый тип генератора для получения электрического выходного сигнала.
10. Это устройство используется для проверки приборов и регуляторов.
11. Это устройство используется для испытания образцов, оборудования и моделирующих устройств. Частота устанавливается пятнадцатью ступенями. Амплитуду можно регулировать пневматически.
12. Эта модель применяется для работы в полевых условиях. Частота устанавливается шестью ступенями, уровень регулируется пневматически. Трехкомпонентный регулятор является частью оборудования.
13. См. [3]. Этот генератор синусоидальных колебаний является частью аппаратуры для построения амплитудно-фазовой характеристики непосредственно в полярных координатах. Частота регулируется замкнутой системой автоматического регулирования, а уровень устанавливается пневматически. Это устройство имеет приспособление для получения электрического и механического выходных сигналов. Имеется

также прибор для измерения сдвига фаз между входным и выходным сигналами. Данные записываются двухканальным электронным самопишущим прибором и подаются на трубку осциллографа или автоматически вычерчиваются в полярных координатах.

14. Описанная модель создана на основе трех предшествующих моделей. Прибор позволяет получить механический выходной сигнал (от 0 до 38 мм) и электрический выходной сигнал (от 0 до 100 в) с такой же гибкостью регулирования, как и для воздуха. Максимальная механическая нагрузка устанавливается с помощью привода Грахама мощностью в $\frac{1}{4}$ л. с.

16. Эта модель находится в процессе разработки. Она будет взрывобезопасной. Частота будет изменяться 45 ступенями, отношение последующих ступеней 1:1,25.

17. Это устройство использует обычные средства для создания вращения оси со скоростью от 3 оборотов в час до 10 оборотов в секунду. С осью связан синусный потенциометр. Выход этого потенциометра или электронного генератора (от 2 до 225 гц) подается в ступенчатый аттенюатор (ступени 1 и 10 дб), а затем уже на усилитель. Выход усилителя приводит в действие электропневматический сервомотор.

18. Размеры и вес, указанные в таблице, включают размеры и вес пневматических усилителей и самопишущих приборов. Частота может устанавливаться ступенями 1:1,6. Уровень можно регулировать пневматически. Скорость ленты синхронизирована с частотой.

19. Генератор для специальных целей. Применяя непосредственно воздушный компрессор, можно менять амплитуду путем изменения объема, а уровень — смещением клапана. Утечка вокруг поршня отрицательно влияет при низких частотах.

20. Этот генератор представляет собой вращающийся пневматический клапан и действует как пневматический импульсный генератор. В современной конструкции при некоторых режимах нормальная работа нарушается, но эти нарушения устраняются дальнейшим совершенствованием конструкции.

21. Это пневматический генератор колебаний типа сирены, представляющий собой видоизменение прибора, описанного в работе [4]. Искажения незначительны при достаточно

тщательном расчете отверстий и использовании резонансной камеры. Амплитуда составляет приблизительно 80% от уровня давления.

Совсем недавно получены сведения о приборе, который состоит из вибрирующего стержня, возбуждаемого струей сжатого воздуха, направленного в центр стержня с помощью сопла соответствующей формы. Система начинает колебаться при частоте, равной собственной частоте стержня. Частоты между 800 и 2000 гц были получены в связи с тарировкой акселерометров.

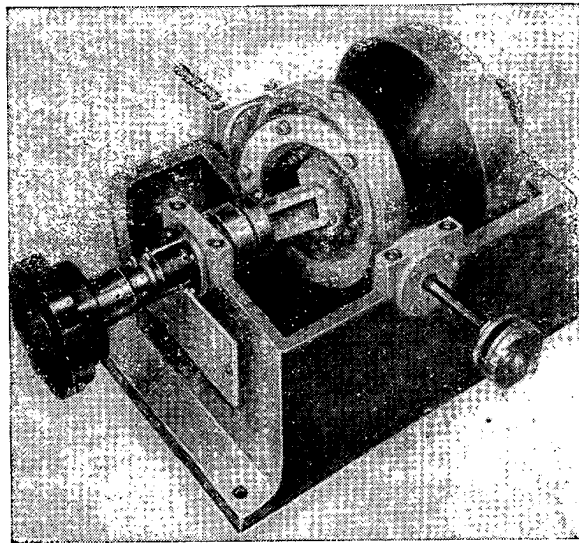
В общем, можно следующим образом суммировать дополнительные замечания, которые должны быть приняты во внимание. Наиболее желательные усовершенствования сводятся к созданию более гибкой и точной системы установки частоты, амплитуды и уровня. Первоначальные модели, изготовленные для лабораторных целей, для использования в полевых условиях делаются взрывобезопасными. Поскольку исследуемые системы автоматического регулирования являются замкнутыми, то для получения измерительных сигналов пригодно лишь устройство, снабженное пневматическим регулированием уровня (или электрическим регулированием, если регулятор электрический).

Кажущееся несоответствие между результатами, опубликованными ранее о некоторых из этих устройств, и данными, представленными в настоящей статье, объясняется тем, что в настоящей статье многие сведения были дополнительно уточнены.

Механические устройства

Существуют различные механизмы для создания механических колебаний. Если вращение с постоянной скоростью [5] нужно преобразовать в линейные гармонические колебания, то можно использовать следующие пять хорошо известных механизмов: гипоциклоида, кривошипно-шатунный механизм, кулисный механизм, диск, насаженный под углом на ось, и кулачковый механизм. Преобразование линейных колебаний в колебания вала выполняется обычно с помощью зубчатой рейки и шестерни. Однако часто конструкция беззубчатой рейки с колесом дает меньшую ошибку. Преобразование вращения с постоянной угловой скоростью в гармонические колебания

вала можно выполнить с помощью четырехзвенного рычажного механизма, ползунов и кривошипов и некоторых других устройств. Здесь будут описаны три типа механических генераторов синусоидальных колебаний. Первый применяется при изучении сервомеханизмов, где требуется малая мощность выходного сигнала; второй имеет среднюю мощность и конструируется сейчас для использования в нефтяной промышленности, а третий имеет большую мощность и применяется при исследовании



Ф и г. 1. Устройство, преобразующее вращение с постоянной скоростью в синусоидальные угловые колебания оси.

регуляторов скорости и двигателей. Их характеристики приведены в таблице под номерами 22, 23 и 24 соответственно. Дополнительные сведения следующие:

22. См. [6]. Это устройство показано на фиг. 1. Оно состоит из прецизионного шарикового подшипника с большим отверстием, укрепленного на оси так, что плоскость внутреннего кольца подшипника может наклоняться относительно плоскости, нормальной к оси входного вала, вращаю-

щегося с постоянной скоростью. Вращение входного вала вызывает круговое вращение подшипника, пропорциональное углу наклона, подобно вращению диска, насаженного под углом на ось. Внешнее кольцо подшипника удерживается от вращения с помощью цапфы, связывающей его с обоймой шарнира кардана. Обойма шарнира кардана поддерживается двумя выходными осями, проходящими через шариковые опоры.

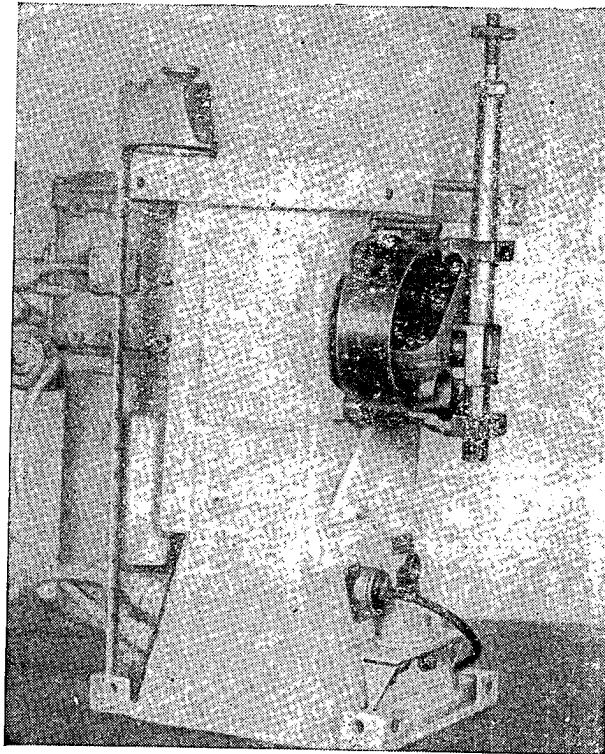
Таким образом, при вращении входной оси происходят колебания выходной оси на некоторый угол по обе стороны от среднего положения, равный углу наклона плоскости кольца подшипника к валу. Имеются регулировочная ручка и механизм для установки амплитуды во время работы.

Этот генератор может приводиться в движение любым двигателем постоянной скорости (крутящий момент от 0,0014—0,0057 кгм), и его выход присоединяется к механизмам, рассчитанным на вход в виде угловых колебаний оси. Для получения вращения с постоянной скоростью часто используется однофазный индукционный электродвигатель с электронным регулятором. Между двигателем и генератором синусоидальных колебаний имеется редуктор, позволяющий менять частоту колебания выходной оси. Так, в рассматриваемом устройстве частота меняется от 0,05 до 1,0 и от 1 до 25 гц. Выходной сигнал генератора подводится непосредственно к исследуемому сервомеханизму. По сообщению одного из потребителей, третья гармоника составляет 0,25% при 10° углового смещения.

23. Это устройство разработано недавно, и к моменту написания обзора о нем имелись ограниченные сведения. Устройство было сконструировано, чтобы получить усилие в 1,36 кг при максимальном ходе поршня и максимальной частоте. В приборе предусматривается также возможность получения электрического импульса шириной в 2° через каждые четверть периода гармонического движения.

24. Этот прибор, изображенный на фиг. 2а, был сконструирован для определения частотных характеристик регулятора скорости двигателя, работающего под нагрузкой. На фиг. 2б показана принципиальная схема устройства. Палец кривошипа находится на штоке поршня, который при нулевой амплитуде прижимается пружиной к неподвижному упору. Жидкость под давлением подается от

гидротрансмиссии типа Виккерс через электромагнитный клапан в цилиндр с поршнем. Под действием этого давления поршень перемещается до тех пор, пока не упрется в стопор



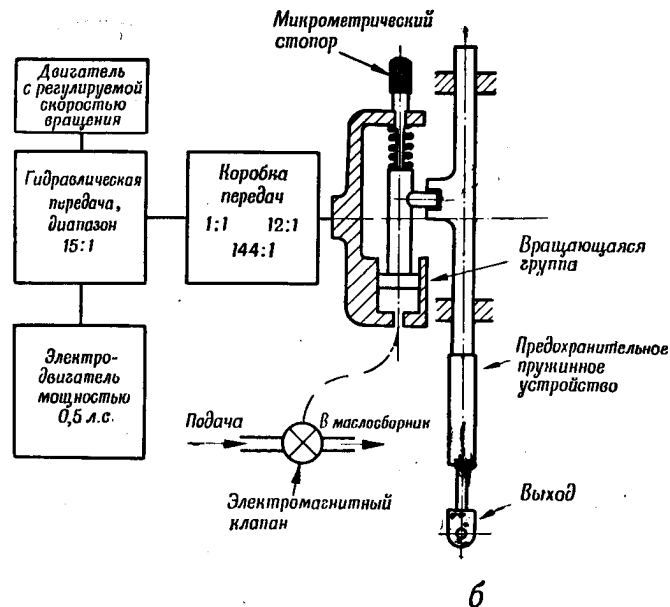
Фиг. 2а. Общий вид устройства для получения механических колебаний большой мощности.

микрометрического вала. При переключении электромагнитного клапана поршень под действием пружины возвращается обратно, что необходимо при эксплуатации установки.

Пружинное устройство предусмотрено для того, чтобы предохранить детали установки, если нагрузка будет превы-

шать расчетную нагрузку в 45 кг. Выходной стержень может поворачиваться на 360° для быстрой установки в нужном положении. Все узлы компактно размещены на сварной станине.

В полевых условиях этот генератор используется в соединении с записывающим устройством, усилителями, генераторами и т. д. и перевозится в подвижной испытательной

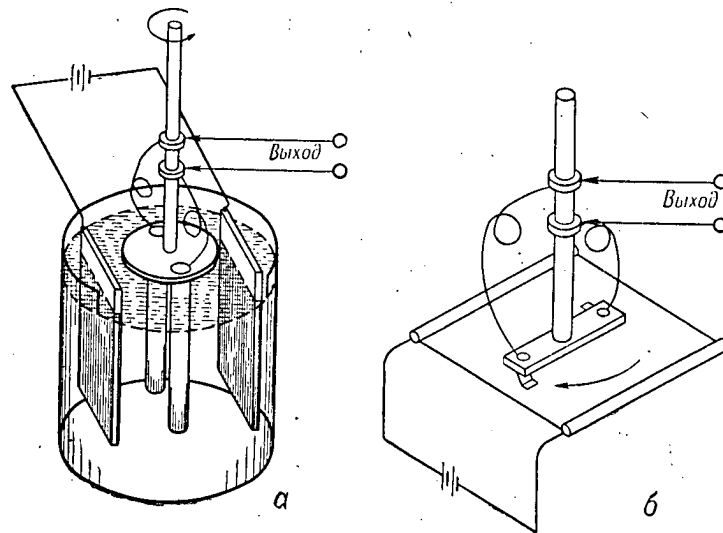


Ф и г. 26. Схема устройства, изображенного на фиг. 2а.

лаборатории, размещенной на грузовой автомашине. На приборной доске автомобиля имеется все необходимое для управления осциллографированием, для пуска и остановки двигателей привода, для регулирования частоты колебания сверх диапазона 15:1 изменения скорости трансмиссии. Оператор имеет телефонную связь с другим оператором, находящимся непосредственно у генератора, где имеется вторая группа приборов управления. Изменение величины амплитуды или переключение коробки передач должно производиться при выключенном оборудовании.

Электрические устройства

Электрические генераторы можно разделить на два основных типа: на генераторы, использующие дополнительные механизмы, и на генераторы, которые таковых не используют. Последние обычно относятся к разряду звуковых генераторов и из описанных в этом обзоре генераторов колебаний только генераторы этого типа выпускаются промышленностью в качестве стандартного оборудования.

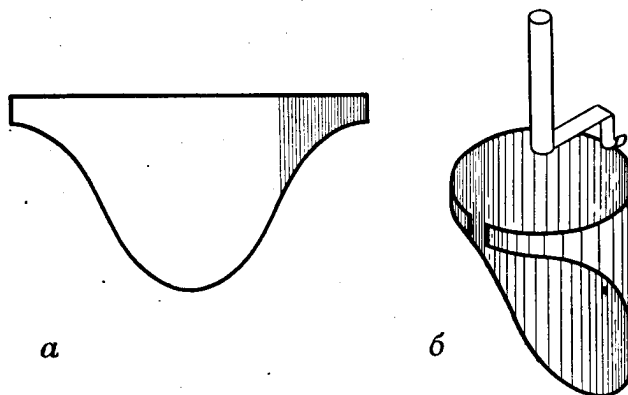


Фиг. 3. Два примера получения электрических синусоидальных колебаний.

Любой из вышеописанных методов получения механических синусоидальных колебаний можно использовать, чтобы привести в действие потенциометр или другое электрическое устройство. Однако наилучшим электромеханическим способом получения колебаний низкой частоты является способ, при котором два электрода или скользящих контакта вращаются в сосуде с жидкостью, имеющем плоскопараллельные эквипотенциальные поверхности, или по проводящей поверхности, содержащей прямые и параллельные эквипотенциальные линии. На фиг. 3 показаны такие устройства. Их возможности

лимитируются несовершенством эквипотенциальных плоскостей или прямых.

Другим электромеханическим устройством для получения колебаний синусоидальной или произвольной формы является профилированный потенциометр. Из изоляционного материала вырезается пластинка желаемой формы, как показано на фиг. 4, а, которая обматывается проволокой с большим сопротивлением и сгибается затем по окружности, образуя



Фиг. 4. Профилированный потенциометр.

потенциометр вида, показанного на фиг. 4, б. При вращении оси подвижного контакта на выходе потенциометра получается электрический сигнал, форма которого повторяет форму пластины.

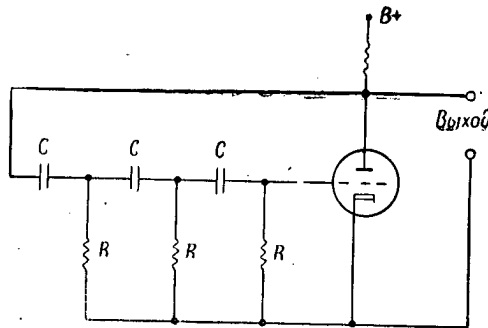
Все упомянутые электромеханические генераторы колебаний имеют стабильный выходной сигнал, не зависящий от частоты. И если в схеме нет механических или электрических реактивных элементов, то переходные ошибки отсутствуют. Однако их максимальная частота ограничена динамическими свойствами схемы. Если же применяются генерирующие элементы с проволочной обмоткой, то необходимо вводить электрические фильтры, чтобы сгладить скачки, появляющиеся в выходном напряжении.

Другим типом электрического генератора синусоидальных колебаний является звуковой генератор. Были созданы сотни различных типов звуковых генераторов. Некоторые

генераторы электрического типа управляются термически; имеются и такие, в которых используются характеристики ротационных электрических машин, таких, как машинные усилители. Однако в настоящее время наибольшее распространение получили ламповые генераторы и генераторы стандартных сигналов, которым и будет посвящена остальная часть статьи.

Ламповые генераторы

Одним из ранних видов лампового генератора низкой частоты является генератор с фазовым сдвигом. Его схема показана на фиг. 5. В схеме возникают колебания при частоте, для которой сдвиг фазы в контуре RC равен 180° ,



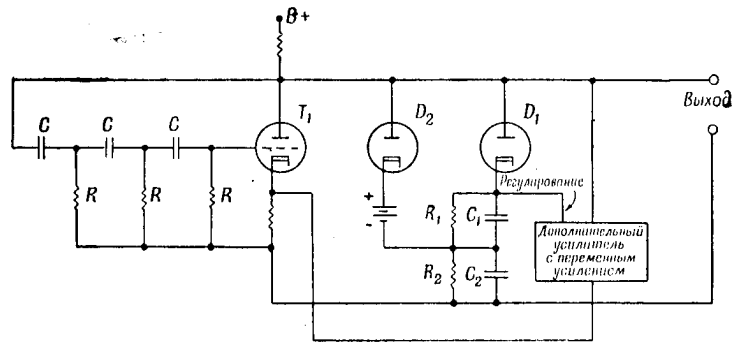
Фиг. 5. Упрощенная схема генератора.

и устанавливается такая амплитуда колебаний, при которой усиление лампы точно компенсирует потери в контуре. Частоту можно менять, изменяя сопротивления или емкости, или то и другое.

Первоначально предусматривались некоторые ручные способы регулирования усиления, и обычно требовалась настройка усиления при каждом изменении частоты.

За последние несколько лет в литературе появились схемы низкочастотных генераторов, имеющих автоматическое регулирование уровня. Обычно постоянная времени системы автоматического регулирования уровня должна быть значи-

тельно больше периода колебаний. Генераторы с частотой 1 гц и выше, использующие такую систему, являются вполне практичными. Однако в генераторах, предназначенных для использования в диапазоне ниже 1 гц, большое время стабилизации амплитуды являлось недостатком, и поэтому, прежде чем использовать такой генератор, необходимо было дальнейшее усовершенствование регулирования амплитуды.



Фиг. 6. Схема генератора синусоидальных колебаний с регулированием амплитуды сигнала.

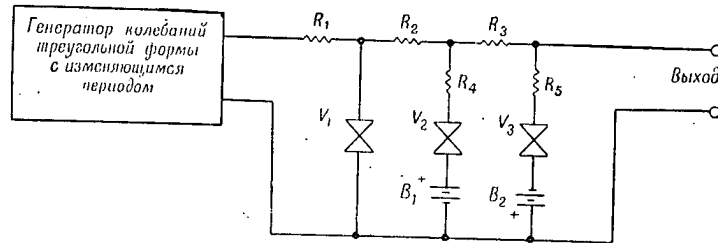
Выпущенный недавно промышленностью электронный низкочастотный генератор содержит устройство стабилизации уровня, которое сделало его удобным в работе и способствовало его широкому распространению в качестве генератора сигналов низкой частоты.

Система автоматического регулирования уровня устроена так, что ее постоянная времени при генерировании сигналов номинального уровня очень большая. Когда же амплитуда генерируемого сигнала отклоняется от заданного значения, постоянная времени системы становится малой. Это дает возможность генератору быстро подавить переходный процесс, который возникает при переключении или настройке.

На фиг. 6 показано видоизменение схемы фиг. 5, содержащее устройство регулирования амплитуды этого типа. Через высокоомный диодный детектор D_1 , сопротивление R_1 и конденсатор C_1 протекает постоянный ток. Напряжение на сопротивлении R_1 , обусловленное этим током, подается на дополнительный усилитель, усиление которого зависит

от величины этого напряжения. Дополнительный усилитель является широкополосным и включен в цепь отрицательной обратной связи генераторного каскада.

Если амплитуда генерируемого сигнала увеличивается, то возросшее напряжение на сопротивлении R_1 уменьшает коэффициент усиления дополнительного усилителя, что приводит к уменьшению амплитуды генерируемого сигнала до номинального значения. Диод D_2 при номинальном значении сигнала заперт. При увеличении амплитуды сигнала через



Фиг. 7. Вариант схемы генератора синусоидальных колебаний.

диод D_2 также течет постоянный ток, который создает напряжение на сопротивлении R_2 и конденсаторе C_2 (имеющих малую постоянную времени) и быстро уменьшает амплитуду сигнала. Очевидно, что можно использовать такую же схему для стабилизации амплитуды и в случае ее уменьшения. Для улучшения качества работы следует ввести дополнительно цепь, поглощающую избыток энергии в схеме. На фиг. 6 эти цепи не показаны. Рассмотренный тип генератора дает на выходе синусоидальные колебания весьма правильной формы.

Другой тип низкочастотного генератора сначала генерирует волну треугольной формы, которая затем при помощи нелинейной цепи превращается в синусоидальную волну. На фиг. 7 показана эта система. Генератором треугольной волны может быть любой из хорошо известных генераторов этого типа. Наиболее широко применяемым устройством является генератор (триггер) прямоугольной волны с последующим интегрированием ее с помощью контура RC . Нелинейная цепь для преобразования треугольной волны в синусоидальную, изображенная на фиг. 7, состоит из линейных сопро-

тивлений и термисторов (нелинейных сопротивлений). Сопротивление R_1 и термистор V_1 , например, предназначены для закругления пика треугольной волны, тогда как сопротивления и термисторы R_2 , R_4 , V_2 , R_3 , R_5 и V_3 в сочетании с соответствующим смещением (с помощью батарей B_1 и B_2) характеристик можно подобрать так, чтобы форма волны была возможно ближе к синусоидальной.

Преимущество этой системы состоит в том, что амплитуда остается постоянной во время переключения и настройки. Существенным недостатком является наличие высших гармоник в выходном сигнале, что иногда нежелательно.

Здесь были рассмотрены лишь упрощенные схемы генераторов с тем, чтобы просто пояснить их принцип работы.

25. Прибор для построения диаграмм Найквиста разработан фирмой „Plessey International Ltd“. Этот прибор автоматически пробегает весь диапазон частот, и результаты строятся в полярных координатах.

26. Массачусетский технологический институт имеет три генератора. В портативном приборе управления артиллерийским огнем типа М-8 низкочастотные синусоидальные колебания снимаются с потенциометра, приводимого во вращение двигателем, скорость которого можно изменять. Потенциометр можно питать либо постоянным током, чтобы получать синусоидальные низкочастотные колебания, либо переменным током, чтобы получить модулированную несущую частоту. Генераторы сигналов типа D-1 и D-2 генерируют колебания высокой частоты, модулированные низкой частотой вращения якоря синхронного двигателя в переменном поле.

Компании „Hewlett Packard“, „Krohn-Hite Instrument“ и „Southwestern Industrial Electronics“ являются тремя основными изготовителями электронных генераторов. Предел низкой частоты выпускаемых приборов составляет примерно 0,01 гц.

Генератор, показанный на фиг. 3, б, применялся в течение второй мировой войны фирмой „Bell Telephone“. О применении его в настоящее время сведений не имеется.

Потенциометр, показанный на фиг. 4, широко используется многими фирмами („Technology Instrument Corporation“, „Gamewell Company“, „Fairchild Camera and Instrument Corporation“ и др.).

Заключение

Имеются различные виды генераторов синусоидальных колебаний. Хотя одни более совершенны, чем другие, очевидно, что не может быть одного наилучшего типа прибора, который удовлетворял бы всем многочисленным требованиям.

Дополнительные сведения о большинстве перечисленных генераторов можно получить от организаций, которые их используют¹⁾, за исключением сведений о генераторах под номерами 3, 4, 5, 20 и 21, которые засекречены.

ЛИТЕРАТУРА

1. Eckman D. P., Moise J. C., „A Pneumatic Sine-Wave Generator for Process Control Study“, ISA, Tech. pap. 52-2-2, 1952.
2. Aikman A. R., „A Portable Instrument for Process Control Analysis“, *Instrument Practice*, № 5, 393 (1951).
3. Ehret R. J., Embree J. M., Hochschild E. F., Grogan E. C., „An Automatic Transfer Function Measuring and Recording System“, Proceedings of the AIEE Conference on Recording and Controlling Instruments, Philadelphia, Nov. 1952.
4. Meyer R. D., *Rev. Sci. Instr.*, 17, 199 (1946).
5. White C. F., „Signal Generators for Servo System Measurements“, Naval Res. Lab.
6. Ferguson F. H., „A Frequency Spectrum Analyser for the Study of Servo Mechanisms“, Naval Ordnance Lab.
7. „Harmonic Response Test Gear“, Vickers Armstrong Ltd, Newcastle, England.
8. Pearson E. B., Lingwood G. F., „A Technique for Rapid Determination of the Harmonic Response Characteristics of a Position Control Servo Mechanism“, *Instrument Practice*, № 3, 340 (1953).

¹⁾ Перечень организаций при переводе опущен. — *Прим. ред.*

СОДЕРЖАНИЕ

Предисловие	5
Вступительная статья	11
I. Общая теория	
1. Р. Ольденбургер. Изображение частотных характеристик, стандарты и расчетные критерии. (Перевод А. М. Егорова)	21
2. М. Керби, Д. Бомарьяж. Относительная устойчивость замкнутых систем. (Перевод А. М. Егорова и А. А. Брэдис)	63
3. И. Чу. Связь между частотными характеристиками и переходным процессом системы автоматического регулирования с обратной связью. (Перевод А. М. Егорова)	117
4. Дж. Бьернсон. Определение переходных процессов по частотным характеристикам разомкнутой системы. (Перевод А. М. Егорова)	141
II. Применение частотных методов к расчету нелинейных систем	
5. Л. Принц. Обобщенный метод определения частотных характеристик замкнутых нелинейных систем. (Перевод А. М. Егорова)	197
6. Е. Левинсон. Некоторые явления насыщения в следящих системах с тахометрической обратной связью. (Перевод А. М. Егорова)	217
7. Г. Честнат. Приближенный частотный метод представления насыщения и зоны нечувствительности. (Перевод А. М. Егорова)	244
III. Применение частотных методов к расчету специальных систем	
8. В. К. Линвилл, Р. В. Ситтлер. Применение частотных методов к расчету импульсных систем. (Перевод А. М. Баткова)	293

9. В. К. Линвилл, Дж. М. Зальцер. Анализ систем автоматического регулирования, содержащих цифровые вычислительные машины. (Перевод Ю. П. Леонова)	311
10. Дж. М. Зальцер. Частотный анализ цифровых вычислительных машин, работающих в реальном времени. (Перевод А. М. Баткова)	329
11. В. Ойя. Применение частотных методов для анализа процесса регулирования турбин в энергосистеме Швеции. (Перевод Е. Г. Ульянова)	358
12. Ч. Л. Сикорд. Применение частотного метода к анализу устойчивости самолета с автопилотом. (Перевод Ю. П. Леонова)	378
IV. Экспериментальные методы определения частотных характеристик	
13. Г. Смит, В. Триплетт. Экспериментальные методы определения частотных характеристик самолета в полете. (Перевод Н. В. Сытиной)	413
14. М. Мэтьюз. Метод оценки нелинейных следящих систем. (Перевод В. А. Карабанова)	441
15. Д. Сент-Клэр, Л. Эрат, С. Гиллеспи. Генераторы синусоидальных колебаний. (Обзор пневматических, механических и электрических устройств для получения частотных характеристик.) (Перевод А. М. Егорова)	473

**ЧАСТОТНЫЕ МЕТОДЫ
В АВТОМАТИКЕ**

Редактор *В. Я. ФРИДМАН*
Художник *Б. И. Фокин*
Технический редактор *С. В. Клименко*

Слано в производство 2/VIII 1957 г.
Подписано к печати 9/X 1957 г.
Бумага $84 \times 108 \frac{1}{32} = 7,9$ бум. л.
25,8 печ. л. в т/ч. 1 вкл.
Уч.-изд. л. 23,6. Изд. № 20/3119.
Цена 18 р. 50 к. Зак. 2356.

ИЗДАТЕЛЬСТВО
ИНОСТРАННОЙ ЛИТЕРАТУРЫ,
Москва, Ново-Алексеевская, 52.

Типография № 2 им. Евг. Соколовой
УПП Ленсовнархоза.
Ленинград, Измайловский пр., 29.

ИМЕЕТСЯ В ПРОДАЖЕ
КНИГА

Г. Ф. СТОРМ

**МАГНИТНЫЕ
УСИЛИТЕЛИ**

Перевод с английского
Под редакцией д-ра техн. наук
Б. С. СОТСКОВА

ИЗДАТЕЛЬСТВО
ИНОСТРАННОЙ ЛИТЕРАТУРЫ

Москва 1957

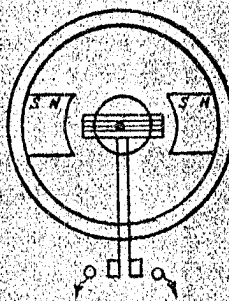
КНИГА ПРОДАЕТСЯ В ЦЕНТРАЛЬНЫХ
И СПЕЦИАЛИЗИРОВАННЫХ КНИЖНЫХ МАГАЗИНАХ

В случае отсутствия книги в местных магазинах следует
обращаться в отделы „Книга — почтой“.

Адреса основных отделов „КНИГА — ПОЧТОЙ“, высылаю-
щих книги наложенным платежом: Москва, Центр, ул. Ки-
рова, 6. Москва, К-9, ул. Петровка, 15.

Ф.А. СТУПЕЛЬ

ОСНОВЫ ТЕОРИИ,
ПРОЕКТИРОВАНИЯ И РАСЧЕТА



ИЗДАТЕЛЬСТВО
ХАРЬКОВСКОГО УНИВЕРСИТЕТА

Ф. А. СТУПЕЛЬ

ЭЛЕКТРОМЕХАНИЧЕСКИЕ РЕЛЕ

**ОСНОВЫ ТЕОРИИ,
ПРОЕКТИРОВАНИЯ И РАСЧЕТА**

УЧЕБНОЕ ПОСОБИЕ

Издание второе

ИЗДАТЕЛЬСТВО
ХАРЬКОВСКОГО ОРДЕНА ТРУДОВОГО КРАСНОГО ЗНАМЕНИ
ГОСУДАРСТВЕННОГО УНИВЕРСИТЕТА
им. А. М. ГОРЬКОГО

Харьков

1956

В книге излагаются основы теории и инженерные методы расчета и проектирования электро-механических реле и механизмов разных принципов действия: электромагнитных, индукционных, магнитоэлектрических и других. Рассмотрены вопросы расчета отдельных узлов реле: катушек, контактных систем, пружин и опор.

Книга составлена в виде учебного пособия для изучения теории реле, а также для курсового и дипломного проектирования студентов электротехнических специальностей. Книга может быть полезной инженерно-техническим работникам в их практической работе по проектированию и применению электрических реле и механизмов.

Ответственный редактор проф. *И. Л. Барг*

ПРЕДИСЛОВИЕ

Электромеханические реле являются одним из наиболее распространенных элементов устройств автоматики и телемеханики. Методы расчета реле базируются на теории ферромагнетизма, теории электромагнитного поля и на теории электромагнитных цепей.

Советскими физиками и металловедами — В. К. Аркадьевым, Н. С. Акуловым, А. С. Займовским, В. С. Меськиным и другими — решен ряд сложных задач по теории ферромагнетизма, выяснению структуры и свойств электротехнических сплавов. В результате их работ получены и освоены новые виды электроматериалов, применяемых в электрических цепях со сталью.

Вопросы электродинамики освещены в работах Я. И. Френкеля и И. Е. Тамма. Работы В. И. Коваленкова — «Теория электромагнитных цепей» и «Основы теории магнитных цепей», удостоенные Сталинской премии, дают возможность построения методики расчета сложных магнитных систем.

Вопросы теории электромагнитных систем освещены в работах А. Я. Буйлова, Б. С. Сотскова, Н. А. Лифшица. Теория индукционных систем и реле освещена в работах А. М. Бресслера, Б. К. Буля, Л. А. Гельбуха. Практические вопросы расчета кодовых и телефонных реле даны в работах М. Н. Витенберга.

В настоящей книге автор стремился к обобщению многочисленных работ по вопросам теории и методов расчета электрических реле и механизмов. Особое внимание уделено вопросам проектирования этих устройств.

В работе приняты такие формулы и соотношения, которые дают возможность анализировать при проектировании влияние отдельных параметров на работу реле; это позволяет проводить систематизацию расчетных материалов и накопление данных для получения оптимального решения поставленных задач. В книге приведен ряд примеров расчета и практических данных из опыта нашей промышленности.

Книга составлена в качестве учебного пособия для студентов электротехнических специальностей; автор надеется, что она будет полезной и для инженерно-технических работников в их практической работе по созданию новых, более совершенных конструкций электрических реле и механизмов.

Автор выражает глубокую признательность профессору Б. Ф. Ващура, доцентам А. П. Сукачеву и Н. А. Калужникову, давшим ряд ценных указаний при рецензировании настоящего пособия.

ПЕРЕЧЕНЬ ОБОЗНАЧЕНИЙ

1. Электрическая цепь

U — напряжение — действующее значение; U_m — напряжение — максимальное значение; u — напряжение — мгновенное значение; E — э. д. с. — действующее значение; e — э. д. с. — мгновенное значение; I — ток — действующее значение; I_m — ток — максимальное значение; i — ток — мгновенное значение;	R, r — сопротивление активное; X, x — сопротивление реактивное; Z — сопротивление полное; P — мощность активная; P_g — потери на гистерезис; P_v — потери на вихревые токи; P_n — суммарные потери на гистерезис и вихревые токи. Q — мощность реактивная; t — время; T — период; f — частота; ω — угловая скорость, электрическая; ρ — удельное сопротивление; γ — удельная проводимость; w — число витков.
--	---

2. Магнитная цепь

Φ — магнитный поток — максимальное значение; Φ_t — магнитный поток — мгновенное значение; B — магнитная индукция — максимальное значение; B_t — магнитная индукция — мгновенное значение; B_r — остаточная магнитная индукция; μ — магнитная проницаемость; μ_{max} — магнитная проницаемость — максимальное значение; μ_n — магнитная проницаемость — начальное значение; μ_r — магнитная проницаемость — обратимая.	μ_0 — магнитная проницаемость вакуума; R — магнитное сопротивление; G — магнитная проводимость; g — удельная магнитная проводимость; L — индуктивность; M — индуктивность, взаимная; σ — коэффициент рассеяния; H — напряженность магнитного поля; H_c — коэрцитивная сила; F — намагничивающая сила; W — магнитная энергия.
---	---

3. Механические и физические величины

W — энергия; A — работа; Q — сила — среднее значение; Q_t — сила — мгновенное значение; M — момент вращения — среднее значение;	M_t — момент вращения — мгновенное значение; j — момент инерции; v — скорость; ω — угловая скорость; R, r — радиус;
---	---

- D, d*** — диаметр;
n — скорость вращения (число оборотов в минуту);
t — температура;
 θ — превышение температуры;
V — объем;
m — масса;
g — ускорение силы тяжести;
G — вес;
 γ — удельный вес;
 μ — коэффициент трения;
- δ** — зазор;
 Δ — толщина диска, барабана;
 σ — напряжение нормальное;
 τ — напряжение касательное;
 $R_{изг}$ — допускаемое напряжение при изгибе;
 $R_{кр}$ — допускаемое напряжение при кручении;
 σ_b — предел прочности при растяжении.

Соотношения между практическими и абсолютными единицами измерений

Система	MKSA	CGSM
Поток, Φ	вебер (вб)	максвелл (мкс) $1 \text{ мкс} = 1 \cdot 10^{-8} \text{ вб}$
Индукция, B	вб/см ²	гаусс (гс) $1 \text{ гс} = 1 \cdot 10^{-8} \frac{\text{вб}}{\text{см}^2}$
Напряженность магнитного поля, H	ампер на сантиметр $\left(\frac{a}{\text{см}}\right)$	эрстед (эр) $1 \text{ эр} = 1,256 \frac{a}{\text{см}}$

ВВЕДЕНИЕ

ЗНАЧЕНИЕ АВТОМАТИЗАЦИИ ПРОЦЕССОВ

Развитие современной техники связано с широким применением автоматизации производственных процессов. Сущность автоматизации заключается в том, что соответствующие механизированные установки или агрегаты снабжаются такими приборами или устройствами, которые автоматически управляют работой механизмов по определенному заранее установленному режиму, без вмешательства человека.

Народнохозяйственное и политическое значение автоматизации чрезвычайно велико: при автоматизации процессов резко повышается производительность агрегатов и установок, улучшается качество продукции, достигаются такие темпы работы, которые немалымы, когда механизмами управляет человек. Автоматизация производства приводит к коренным изменениям в структуре рабочего состава на заводах, фабриках, электростанциях, приводит к созданию новых интеллектуально-технических рабочих специальностей, к уничтожению противоположности между физическим и умственным трудом. В исторических решениях XVIII съезда ВКП(б) были намечены меры по осуществлению автоматизации в важнейших отраслях промышленности и транспорта. В решениях XIX съезда КПСС также было уделено большое внимание вопросам автоматизации производственных процессов. В соответствии с директивами съезда, выпуск приборов управления и контроля должен увеличиться в текущей пятилетке в 2,7 раза, чтобы возможно было осуществить автоматизацию процессов во всех отраслях промышленности, в энергетике и на транспорте.

РАБОТЫ ДОРЕВОЛЮЦИОННЫХ И СОВЕТСКИХ УЧЕНЫХ И ИНЖЕНЕРОВ

Основные теоретические положения, принимаемые для построения устройств автоматики, заимствованы из электронизмерительной техники, области, развившейся значительно ранее автоматики. Приоритет в разработке теории и конструкций электронизмерительных приборов принадлежит русским ученым.

Первый измерительный прибор — электроскоп — был создан в 1751 г. Г. В. Рихманом и уже содержал основной принцип построения измерительного прибора и реле: взаимодействие двух

сил, одна из которых является функцией измеряемой величины, а другая — функцией положения подвижной части. М. В. Ломоносов, изучая электрические явления, дал идею использования сил взаимодействия электрических токов для построения электроизмерительных приборов. В начале XIX ст. русские ученые Э. Х. Ленц и Б. С. Якоби создали учение об электромагнитной индукции, дав, таким образом, предпосылки для построения индукционных приборов и реле.

В конце XIX ст. М. О. Доливо-Добровольский, развивая предложенную им идею применения трехфазного тока, разработал ряд новых измерительных приборов, в частности первый фазометр, основанный на принципе вращающегося магнитного поля. В 1913 г. М. О. Доливо-Добровольский, основываясь на классических работах А. Г. Столетова в области намагничивания стали, впервые разработал ферродинамические приборы.

Приоритет в области устройств автоматики принадлежит русским инженерам.

И. П. Ползунов применил в своей машине автоматическое парораспределение. К. Д. Фролов построил на Алтае завод с рядом механических полуавтоматических и автоматических станков. Современник Ломоносова, талантливый механик самоучка И. П. Кулибин изобрел и изготовил ряд физических приборов, а также уникальные часы с группой автоматически действующих фигур.

В первой половине XIX ст. П. Л. Шиллинг изобрел первый реально действующий телеграфный аппарат, телеграфное реле. Чиколевым разработан автоматический регулятор для дуговых печей, Константиновым — автоматический регулятор для электрических машин.

Однако работы русских ученых и изобретателей не получили должного применения в царской России. Осуществлению идей и изобретений русских ученых мешал не только низкий уровень промышленности в России, но и зависимость от крупных иностранных фирм. До революции отдельные реле изготовлялись только в небольших мастерских.

Великая Октябрьская социалистическая революция положила начало быстрому развитию отечественной электропромышленности. В годы осуществления довоенных пятилеток наша электротехническая промышленность, и в частности релестроение, обогатилась рядом крупных достижений. Производство реле в СССР было начато в 1928 г. на Харьковском электромеханическом заводе (ХЭМЗ). Уже к концу первой пятилетки было разработано и внедрено в производство большое количество реле для автоматического управления сложными электроприводами, например, для доменного подъема, блюмингов, бумагоделательных машин, а также защитных реле для энергосистем.

Дальнейшее развитие в релейном производстве ХЭМЗа проходило в области создания новых видов реле для защиты энергосистем и автоматизации производственных процессов. Ряд новых

конструкций создан инженерами завода А. М. Бресслером, Я. Н. Штрафуном, А. С. Тулиным, автором и другими.

Значительный прогресс в области релестроения в Советском Союзе определялся необходимостью автоматизации оборудования канала имени Москвы (1936—1937 гг.).

Теория устройств автоматики, теория релейной защиты обогатились работами ряда научно-исследовательских институтов и кафедр высших технических учебных заведений.

В области теории релейной защиты известны работы профессоров Н. Ф. Марголина, А. М. Федосеева, Г. И. Атабекова. В области расчета электромагнитных систем ряд положений разработан профессором А. Я. Буйловым, Б. С. Сотсковым, Б. К. Буль и другими. Б. К. Буль разработал методику расчета индукционных дисковых систем, исследовал работу таких систем.

Большое значение имеют работы Л. А. Гельбуха, который впервые дал методику расчета индукционных барабанных систем и методику аналитического исследования переходных процессов в этих системах. А. С. Тулин и В. Л. Фабрикант провели имеющие большое практическое значение работы по исследованию индукционных систем. На основе работ А. С. Тулина и А. Бресслера созданы совершенные модели дистанционных реле.

В настоящее время в СССР имеется ряд специализированных институтов и заводов, разрабатывающих и изготавлиющих большое число видов реле и устройств автоматики.

СТРУКТУРНАЯ СХЕМА АВТОМАТИЗИРОВАННОЙ УСТАНОВКИ

В общем виде всякую автоматизированную установку можно считать состоящей из следующих основных элементов:

1) механизмы или агрегаты, осуществляющие необходимые для данной установки операции или функции;

2) силовые аппараты, осуществляющие в механизме или агрегате необходимые манипуляции;

3) устройства автоматического управления, настроенные на определенный режим работы и воздействующие на силовые устройства.

Устройства автоматического управления исполняются в виде высокочувствительных приборов: реле, автоматических регуляторов, датчиков. В ряде случаев между этими приборами и силовыми аппаратами устанавливаются усилительные устройства.

Реле — прибор, настроенный на определенный режим работы механизма, агрегата или установки. Реле действует при определенных стадиях режима и действие его передается силовому аппарату.

Автоматический регулятор — прибор (или система приборов и усилителей), задачей которого является поддержание постоянным определенного состояния в автоматизированной установке или агрегате.

Датчики — специальные приборы, измеряющие определенные величины, характеризующие технологический процесс. Замеры датчиков передаются к измерительной части реле или регуляторов — непосредственно или через усилители.

ОБЩАЯ КЛАССИФИКАЦИЯ РЕЛЕ

Классификация реле производится по разным признакам: по назначению, по измеряемым величинам, по принципу действия, по способу включения.

1) Классификация по назначению

Различают реле:
управления — для автоматического управления работой электродвигателей;
защиты — для автоматической защиты электрических устройств;
автоматизации производственных процессов.

2) Классификация по измеряемой величине

Электрические — тока, напряжения, мощности, частоты, коэффициента мощности, сопротивления;
оптические — силы света, длины волны, прозрачности, мутности, цветности и другие;
тепловые — температуры, количества тепла;
акустические — силы звука и т. д.;
механические — силы, давления, скорости, ускорения, линейных перемещений, уровня, объема;
разных физических величин — времени, вязкости, плотности.

3) Классификация по принципу действия

Электромагнитные; индукционные; магнитоэлектрические; электродинамические; тепловые; механические.

4) Классификация по способу включения

Первичные — катушки реле включаются непосредственно в основную цепь защищаемого или управляемого объекта;
вторичные — катушки реле включаются во вторичные цепи измерительных трансформаторов тока и напряжения;
промежуточные — катушки реле включаются в цепь между основными реле защиты или управления и другими реле или аппаратами и получают питание от отдельного источника тока.

Реагирование на измеряемую величину — ток, напряжение, мощность и другие — может быть осуществлено по-разному. Различают:

реле максимального действия — срабатывающие при появлении или увеличении заданной величины;
реле минимального действия — срабатывающие при исчезновении или уменьшении заданной величины;
реле направленного действия — срабатывающие при изменении направления движения или знака заданной величины;
реле дифференциальные — срабатывающие при наличии разности значений нескольких величин;
балансные — в которых складываются или вычитаются силы, создаваемые двумя или несколькими элементами замера.

МЕТОДЫ РАСЧЕТА ЭЛЕКТРОМАГНИТНЫХ СИСТЕМ И ЭЛЕКТРИЧЕСКИЕ РЕЛЕ

Точный расчет электромагнитных систем и электрических реле связан с весьма громоздкими и трудоемкими вычислениями, что определяется в основном следующими причинами:

- 1) зависимость индукции от напряженности поля для цепи со сталью нелинейна;
- 2) значение индукции в рабочих зазорах магнитопровода распределяется неравномерно; по данным проведенных исследований при малых зазорах индукция имеет максимум у середины полюса, а по краям значение ее уменьшается, при больших зазорах получается обратная картина;
- 3) картина распределения поля в воздушных зазорах магнитопровода, а также поля утечки, весьма сложна;
- 4) закон распределения потоков в системе в значительной степени зависит от расположения катушки в магнитопроводе;
- 5) температура нагрева различных частей магнитопровода и катушки неодинакова.

Точный расчет электромагнитных систем затруднен еще тем, что принимаемые в расчетах данные о материалах и размерах получают иногда значительные отклонения в условиях производства. Так, данные о кривых намагничивания имеют отклонения в пределах $\pm (10 \div 15)\%$, данные о потерях на гистерезис и вихревые токи весьма неточны и в зависимости от условий обработки детали могут отклоняться от лабораторных в два-три раза, значения нерабочих зазоров — в стыках — могут изменяться в два-три раза, в условиях производства имеют отклонения значения коэффициентов заполнения меди и стали.

Существующие методы расчета электромагнитных систем и электрических реле как аналитические, так и графо-аналитические, построены с учетом ряда допущений:

- 1) поле в рабочих зазорах принимается однородным, с учетом среднего значения индукции;
- 2) упрощенно вычисляют значения проводимостей воздушных зазоров и путей утечки;
- 3) не учитывают влияния расположения катушки в магнитопроводе (по отношению к рабочему зазору);

4) потоки и токи в системах переменного тока принимают синусоидальными;

5) учитывают обычно среднестатистические данные кривых намагничивания.

Инженерные методы расчета должны учитывать возможности проектирования системы, т. е. разработки модели по заданным эксплуатационным параметрам, так и расчета системы, т. е. вычисления ее параметров по заданным геометрическим размерам.

Опубликованные в литературе аналитические методы расчета, в которых нелинейности магнитной цепи учтены определенной аппроксимацией кривых, содержат в себе сложные зависимости, не дающие возможности анализа влияния отдельных параметров на работу системы и установления оптимальных соотношений. Поэтому такие методы пригодны только для расчета системы определенных размеров.

В принятых в практике инженерных методах расчета пользуются как аналитическими соотношениями, так и графическими материалами — кривыми и векторными диаграммами. Такие методы расчета называют графо-аналитическими.

Получающиеся при этом соотношения, весьма простые по своей структуре, дают возможность проведения не только расчета, но и проектирования системы, установления оптимальных соотношений, позволяют конструктору определять, какие параметры системы требуется изменить в том случае, когда по технологическим причинам исполненная модель не соответствует проектным данным.

Принимая во внимание, что существующие методы расчета построены с учетом ряда допущений, а также учитывая возможные технологические отклонения, следует проводить проектирование с учетом некоторого запаса. Так, например, при расчете системы принимают значение превышения температуры несколько меньшим предельно допустимого (на 10°).

В данном пособии изложены инженерные графо-аналитические методы расчета, позволяющие решать вопросы как проектирования, так и расчета электромагнитных систем и реле. Такие методы также приняты и по методическим соображениям: при проведении расчета графо-аналитическими методами более наглядно выявляется физическая сущность происходящих процессов, что весьма важно для изучающих данную дисциплину.

В перечне литературы приведены статьи и книги, в которых освещены аналитические методы расчета.

ЧАСТЬ ПЕРВАЯ
МАГНИТНЫЕ СИСТЕМЫ

ГЛАВА I
МАГНИТОПРОВОД

**1-1. ОСНОВНЫЕ ВЕЛИЧИНЫ, ХАРАКТЕРИЗУЮЩИЕ СВОЙСТВА
МАТЕРИАЛОВ ДЛЯ МАГНИТОПРОВОДОВ**

Для изготовления магнитопроводов реле применяются разнообразные сорта сталей и сплавов, обладающих ферромагнитными свойствами. Выбор того или иного материала определяется, с одной стороны, назначением магнитной системы, а с другой — магнитными свойствами материала. К основным величинам, характеризующим свойства таких сталей и сплавов, относятся:

1) B_s — индукция насыщения — максимально возможное значение индукции в материале, при котором дальнейшее усиление поля не увеличивает намагниченности тела;

2) B_r — остаточная индукция, т. е. величина индукции при $H = 0$, получающаяся при размагничивании материала от максимального значения B_s ; при размагничивании от значения $B < B_s$ величина индукции при $H = 0$ уменьшается;

3) H_c — коэрцитивная сила — напряженность магнитного поля, необходимая для снижения остаточной индукции до нуля; значения B_r и H_c определяются по размагничивающему циклу кривой намагничивания. В данных заводов значения B_r и H_c приводятся для материалов, намагниченных до индукции насыщения B_s ;

4) μ — магнитная проницаемость; значение μ изменяется с изменением поля и определяется по начальной кривой намагничивания, как отношение $\frac{B}{H}$.

5) μ_n — начальная магнитная проницаемость — значение проницаемости в очень слабых магнитных полях, близких к нулю;

6) μ_{max} — максимальная магнитная проницаемость;

7) P_e и P_g — потери на гистерезис и вихревые токи;

8) точка Кюри — температура, при которой ферромагнитные материалы теряют свои магнитные свойства.

1-2. КЛАССИФИКАЦИЯ МАГНИТНЫХ МАТЕРИАЛОВ

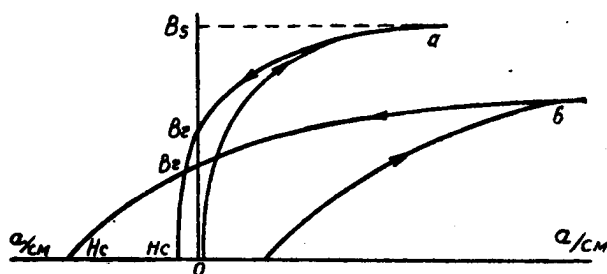
Промышленное значение получили магнитные материалы, обладающие крайними значениями магнитных свойств (фиг. 1,1).

Различают следующие основные виды магнитных материалов:

1) *магнитномягкие*, обладающие низкой коэрцитивной силой H_c и высоким значением магнитной проницаемости μ . Такие мате-

риалы применяются для изготовления магнитных систем разных видов реле;

2) *магнитотвердые*, обладающие высокими значениями коэрцитивной силы H_c ; значение μ у этих материалов во много раз



Фиг. 1, 1. Характеристики магнитных материалов:
а — магнитномягких; б — магнитотвердых.

меньше, чем у магнитномягких материалов. Магнитотвердые материалы применяются для изготовления постоянных магнитов. Особую группу представляют собой термомагнитные сплавы, обладающие низкой точкой Кюри (порядка $10 \div 100^\circ$). Эти сплавы применяются для температурной компенсации.

1-3. МАТЕРИАЛЫ ДЛЯ МАГНИТОПРОВОДОВ

В практике применяются различные магнитномягкие материалы, отличающиеся различными свойствами в области слабых, средних и сильных полей, различными значениями коэрцитивной силы, удельных потерь, условиями технологической обработки и стоимости.

Сталь конструкционная — разных профилей с содержанием углерода от 0,1 до 0,4%. Обладает хорошими технологическими свойствами в отношении штамповки и обработки режущими инструментами.

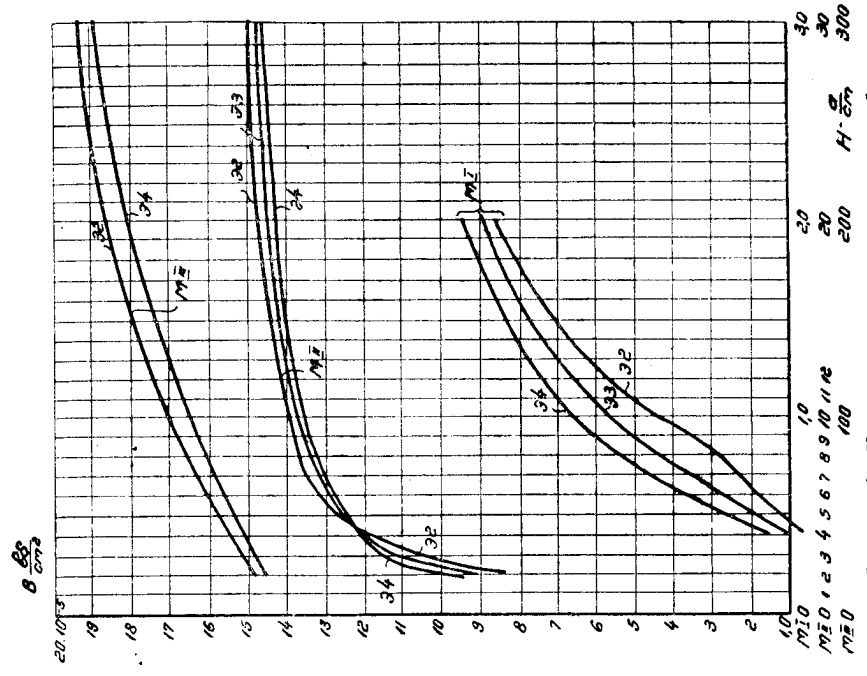
$$B_s \cong 23 \cdot 10^{-5} \frac{\text{вб}}{\text{см}^2}; H_c \text{ — до } 3 \frac{\text{а}}{\text{см}}; \gamma = 7,85 \frac{\Gamma}{\text{см}^2}.$$

Кривые намагничивания для сталей СТ-2 и СТ-3, выполненные по данным завода «Электросила», приведены на фигуре 1, 2.

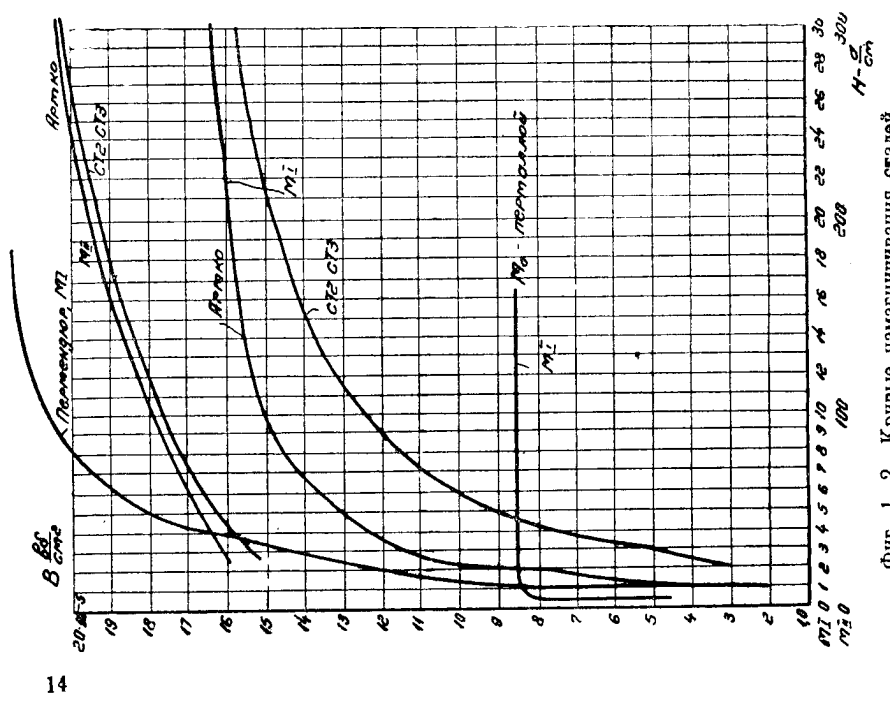
Кремнистая сталь — круглая, с добавлением кремния.

$$B_s = 20 \cdot 10^{-5} \frac{\text{вб}}{\text{см}^2}; B_r = 6 \cdot 10^{-5} \div 8 \cdot 10^{-5};$$

$$H_c = (0,4 \div 0,6) \frac{\text{а}}{\text{см}}; \gamma = 7,6.$$



Фиг. 1, 3. Кривые намагничивания листовой электротехнической стали.



Фиг. 1, 2. Кривые намагничивания сталей.

Железо Армко. Содержит примеси не более 0,14%, легко обрабатывается, изготавливается в виде листов и прутков. Выпускается двух марок:

для глубокой штамповки,

$$H_c = 1 \frac{a}{cm};$$

для магнитопроводов,

$$H_c = 0,8 \frac{a}{cm}.$$

Кривые намагничивания см. на фигуре 1, 2.

Электротехническая сталь. Изготавливается в виде листов толщиной от 0,35 мм до 2 мм, а для высоких частот от 0,05 до 0,2 мм. По степени легированности различают следующие марки электротехнической стали: Э1, Э2, Э3, Э4. Данные указанных марок см. в таблице 1. Кривые намагничивания показаны на фигуре 1, 3, кривые размагничивания — на фигуре 1, 4.

Сталь ХВП — холоднокатанная, высокой проницаемости, выпускается в виде листов. Листы и ленты этой стали обладают резко выраженной анизотропией: высокие магнитные свойства получают только в том случае, когда направление потока совпадает с направлением прокатки листа.

Средние данные сталей ХВП:

$$P_{10} = 0,84 \frac{вт}{кГ}; \quad P_{15} = 1,95 \frac{вт}{кГ};$$

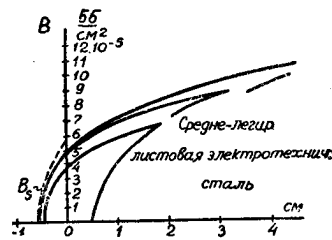
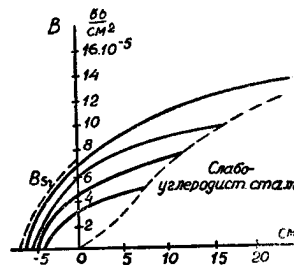
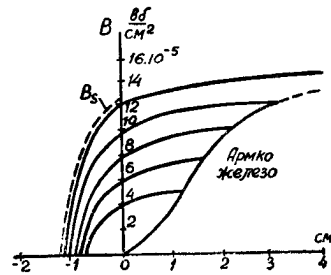
$$\text{Индукция при } H = 0,05 \frac{a}{cm}; \quad B = 500 \text{ гс} = 0,5 \cdot 10^{-5} \frac{вб}{cm^2}$$

$$H = 25 \frac{a}{cm}; \quad B = 17300 \text{ гс} = 17,3 \cdot 10^{-5} \frac{вб}{cm^2};$$

$$H = 50 \frac{a}{cm}; \quad B = 18400 \text{ гс} = 18,4 \cdot 10^{-5} \frac{вб}{cm^2};$$

$$H_c = 0,16 \text{ а/см.}$$

е) Пермаллой — сплав с содержанием никеля. Различают две основные группы пермаллоя: низконикелевые и высоконикелевые.



Фиг. 1, 4. Кривые размагничивания магнитномягких материалов.

Сплавы пермаллоя легируются молибденом, хромом и медью. Данные пермаллоя:

1) низконикелевый — с содержанием Ni — 45% ÷ 50%, легко обрабатывается резанием, штампуется. Изготавливается в виде холоднокатанных полос шириной 80—250 мм и толщиной 0,1—3 мм; прутков диаметром 8 ÷ 30 мм и проволоки диаметром 0,5 ÷ 3 мм.

$$B_s = 16 \cdot 10^{-5} \left(\text{в поле } 12 \div 13 \frac{a}{cm} \right); B_r = 10 \cdot 10^{-5} \frac{об}{cm^2};$$

$$H_c = 0,24 \frac{a}{cm}; \mu_{\text{макс}} = 25 \cdot 10^{-5} \frac{об}{a/cm} \left(\text{в поле } 0,16 \frac{a}{cm} \right).$$

$$\mu_n = 3,1 \cdot 10^{-5} \frac{об}{a/cm}; \gamma = 8,15; \rho = 0,45 \text{ ом } \frac{мм^2}{м};$$

2) высоконикелевый молибденовый пермаллой — с содержанием 78,5Ni + 3,8Mo + Fe легко обрабатывается штамповкой, резанием. Изготавливается в виде холоднокатанных полос шириной от 80 до 250 мм, толщиной от 0,1 до 3 мм; кованных и горячекатанных прутков диаметром от 8 до 30 мм; холоднокатанной проволоки диаметром от 0,5 до 3 мм.

$$B_s = 8,5 \cdot 10^{-5} \left(\text{в поле } 0,65 \frac{a}{cm} \right); B_r = 6 \cdot 10^{-5}; H_c = 0,024 \frac{a}{cm};$$

$$\mu_{\text{макс}} = 160 \cdot 10^{-5} \frac{об}{a/cm} \left(\text{в поле } 0,02 \frac{a}{cm} \right); \mu_n = 10 \cdot 10^{-5};$$

$$\rho = 0,55 \text{ ом } \frac{мм^2}{м}.$$

Пермендюр — сплав с содержанием Co — 49%, V — 1,7%, Fe. Обладает исключительно высокой индукцией насыщения B_s , достигаемой в сравнительно слабых полях. Изготавливается в виде холоднокатанных листов, толщиной от 0,2 до 2 мм; прутков диаметром от 8 до 30 мм.

$$B_s = 24 \cdot 10^{-5} \frac{об}{cm^2}; B_r = 13 \cdot 10^{-5} \frac{об}{cm^2}; H_c = 1,2 \frac{a}{cm}.$$

1-4. ВОПРОСЫ ТЕХНОЛОГИИ

Механическая обработка. Детали магнитопровода из стали, железа Армко, пермаллоя могут быть изготовлены штамповкой, ковкой, обработкой резами. Детали для шихтованных магнитных систем (жести) получают штамповкой.

Листовые электротехнические стали пригодны только для изготовления плоских деталей, так как обладают хрупкостью, возрастающей с увеличением содержания кремния. По данным ГОСТ листы марки Э1 и Э2 должны выдерживать без излома 10 перегибов, а марки Э4 — 1 перегиб.

Термическая обработка. В процессе изготовления деталей в них возникают местные изменения структуры, приводящие

Таблица 1

Сталь листовая электротехническая

Наименование	Марка	Толщина листов, мм	Магнитная индукция				Удельные потери при $f = 50$ вц, вт/кг не более		B_T , гс	H_c , а/см	Уд. вес γ	Сод. кремния
			не менее				P_{10}	P_{15}				
			B_{ac} при $a = 25$ см	B_{ac} при $H = 50$	B_{ac} при $H = 100$	B_{ac} при $H = 300$						
Слабо-легированная	Э1	2,0; 1,5	14700	15800	17000	19200	3,6	8,6	7000—10000	0,8	7,8	до 2%
	Э1	1,0	14700	15800	17000	19200	3,3	7,9				
	Э1А	0,42; 0,5	15000	16200	17500	19700	3,3	7,9				
	Э1АА	0,42; 0,5	14900	16000	17300	19600	3,3	7,9				
	Э1АБ	0,42; 0,5	15400	16400	17800	20200	2,85	7,0				
Э1ААБ	0,42; 0,5	15400	16400	17800	20200	2,85	7,0					
Средне-легированная	Э2	0,42; 0,5	14700	15800	17000	19200	2,6	6,4			7,75	до 3,3%
	Э2Б	0,42; 0,5	15300	16400	17600	20100	2,6	6,4				
Повышенно легированная	Э3	0,42; 0,5	14600	15700	16900	18900	2,3	5,6	5000—8000	0,52	7,65	до 3,85%
	Э3А	0,42; 0,5	14600	15700	16900	18900	2,0	4,7				
	Э3	0,3; 0,35	14600	15700	16900	18900	2,0	4,2				
	Э3А	0,3; 0,35	14600	15700	16900	18900	1,6	3,6				
Высоко-легированная	Э4	0,42; 0,5	14400	15500	16600	18700	1,8	3,9		0,32	7,55	до 5%
	Э4А	0,42; 0,5	14400	15500	16600	18700	1,6	3,6				
	Э4АА	0,42; 0,5	14400	15500	16600	18700	1,45	3,3				
	Э4	0,3; 0,35	14400	15500	16600	18700	1,45	3,4				
	Э4АА	0,3; 0,35	14400	15500	16600	18700	1,3	3,2				
Э4АА	0,3; 0,35	14400	15500	16600	18700	1,2	2,9					

к ухудшению магнитных свойств: уменьшается проницаемость, увеличивается коэрцитивная сила, повышается значение удельных потерь. Поэтому детали для магнитных систем, где указанные параметры играют важную роль, подвергают после механической обработки — термической. Для каждого вида материала применяется особый режим термической обработки. Ниже приводятся данные режимов обработки некоторых материалов.

Железо Армко — отжиг при 900° с последующим медленным охлаждением.

Листовая электротехническая сталь — отжиг производится или в сухом чистом песке или в печи с защитной газовой средой (водород, чистый азот и др.). Температура отжига $720-780^{\circ}$, выдержка — 1—1,5 часа, охлаждение в печи до $200-250^{\circ}$ со скоростью не выше $40-60^{\circ}$ в час, при температуре ниже $200-250^{\circ}$ детали могут остывать вне печи.

Пермаллой — отжиг в печи с защитной газовой средой (водород), температура 1100° , 3 часа, охлаждение до 200° со скоростью 50° в час.

Пермендюр — отжиг при 1000° , 50 часов, охлаждение 100° в час до 200° .

1-5. ОБЛАСТЬ ПРИМЕНЕНИЯ МАГНИТНЫХ МАТЕРИАЛОВ

При выборе материала для магнитной системы учитывают, с одной стороны, требования, предъявляемые к магнитной системе, с другой — качество материала и его стоимость.

Системы постоянного тока. Наиболее ходким материалом является железо Армко, марки Э. В простых конструкциях реле, например, в первичных реле управления, железо Армко может быть заменено более дешевым материалом — конструкционной сталью. В реле времени с магнитным демпфером сердечники выполняют из круглой кремнистой стали. Для высокочувствительных, быстродействующих реле применяют материалы с высокой начальной проницаемостью — пермаллой, реже пермендюр.

Особое значение придается выбору материалов для магнитных систем автоматических регуляторов. Точность работы регулятора определяется разницей между кривыми намагничивания и размагничивания системы. Для уменьшения этой разницы применяют материалы с весьма малой коэрцитивной силой: пермаллой, ХВП.

1-6. ИСПОЛНЕНИЕ ДЕТАЛЕЙ МАГНИТОПРОВОДОВ

Различают два основных вида магнитопроводов:

1) *сплошные*, изготавливаемые из полосового или круглого материала — конструкционной стали, железа Армко, круглой кремнистой стали;

2) *шихтованные* части магнитопровода представляют собой склепанные пакеты из листового материала толщиной $0,3-0,5$ мм — листовой электротехнической стали, пермаллоя, ХВП.

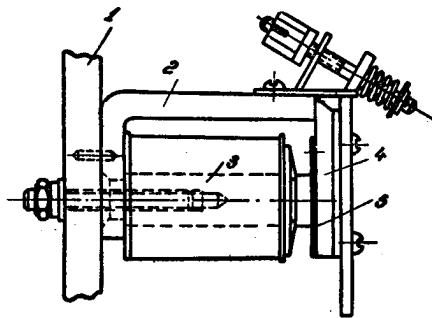
Как правило, реле постоянного тока имеют магнитопровод из сплошного материала, а реле переменного тока — шихтованный магнитопровод. Иногда, однако, отступают от этого общего правила. Так, магнитные системы постоянного тока исполняются с шихтованным магнитопроводом в тех случаях, когда требуется построить быстродействующее высокочувствительное реле или магнитную систему регулировочного реле, где необходимо применить материал с малыми значениями коэрцитивной силы и малыми потерями на вихревые токи (такие материалы изготавливают обычно в виде листов). Такое же положение имеет место при исполнении прямоходовых систем, когда изготовление шихтованной системы, состоящей из отдельных жестей, получаемых процессом штамповки, дешевле и проще магнитной системы из сплошного материала.

В отдельных случаях магнитопровод небольших реле переменного тока изготавливается из сплошного материала толщиной 2—3 мм (например, из кремнистой стали). В таких системах получаются повышенные потери на вихревые токи, но поверхность охлаждения катушки принимается несколько большей, чем у обычных исполнений, и значение температуры перегрева не выходит за пределы норм.

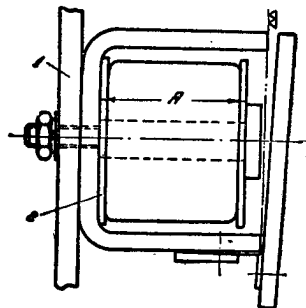
а) **Сплошной магнитопровод** состоит из следующих основных частей: сердечника, на который насаживается катушка, подвижного якоря (или плунжера) и ярма, на которое крепятся все части реле.

Сердечник обычно изготавливается из круглого материала; особое внимание уделяют креплению сердечника к ярму. Этот узел требует тщательного изготовления, так как при плохой пригонке деталей магнитное сопротивление в месте перехода получается значительным.

Наиболее рационально крепить детали наглухо легкой пресовой посадкой или расклепкой конца сердечника в скобе (фиг. 1, 5), однако, встречается и неглухое крепление сердечника (фиг. 1, 6).

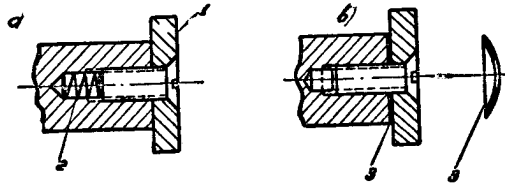


Фиг. 1, 5. Магнитопровод с расклепанным сердечником:
1 — плита; 2 — скоба, 3 — сердечник;
4 — ярмо; 5 — немагнитная прокладка.



Фиг. 1, 6. Магнитопровод:
1 — плита; 2 — шайба.

Полюсный наконечник. В системах постоянного тока с поворотным якорем рекомендуется применение полюсных наконечников, благодаря которым увеличивается проводимость рабочего воздушного зазора.



Фиг. 1, 7. Крепление полюсного наконечника: a — с витой пружиной; b — с пружинной шайбой. 1 — полюсный наконечник; 2 — витая пружина; 3 — пружинная шайба.

Если для смены катушки возможно снять все реле и отвинтить сердечник, последний вытаскивается заодно с полюсным наконечником (фиг. 1, 6). Если же по условиям эксплуатации съем реле с панели для смены катушки не-

удобен, ставят отдельный полюсный наконечник (фиг. 1, 7), который скрепляется с сердечником посредством винта с потайной головкой. Для предотвращения самоотвинчивания этого винта применяют пружины: плоские из фосфористой бронзы или винтовые из стальной проволоки.

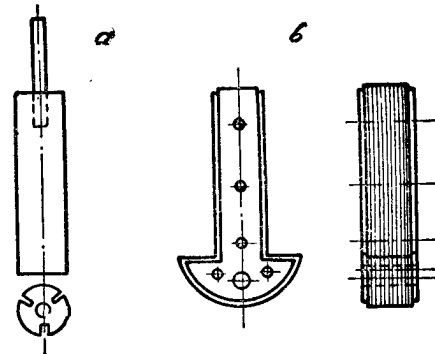
Якорь. В реле с поворотным якорем якорь выполняется из полосового материала; в прямоходовых — из круглого. В быстродействующих реле для уменьшения вихревых токов в круглом сердечнике делают прорезы (фиг. 1, 8 а).

В реле с поворотным якорем вращение якоря осуществляется реже на оси, чаще на призме (фиг. 1, 5), что дает лучшее конструктивное решение. В прямоходовых реле направляющей для поступательного движения якоря служит латунная трубка, помещаемая внутри катушки (фиг. 1, 9).

В реле с поперечным движением якоря (фиг. 1, 6) последний крепится на пружинах.

Немагнитная прокладка.

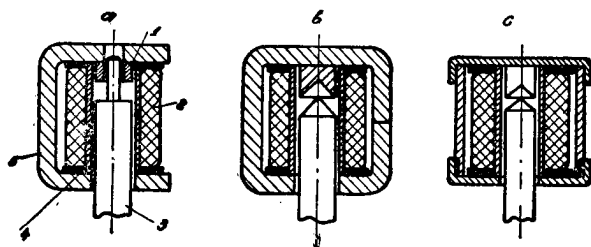
Для предохранения якоря от залипания вследствие остаточного потока в системе, необходимо выполнять магнитопровод так, чтобы при притянутаом якоря оставался некоторый немагнитный зазор. Величина немагнитных зазоров между отдельными частями магнитопровода: в стыках между сердечником и ярмом, якорем и сердечником, якорем и ярмом, определяемая неплотностью при-



Фиг. 1, 8. Якорь реле прямоходового типа: a — постоянного тока; b — переменного тока.

легания (приблизительно по 0,05 мм на стык), а также слоем цинкового покрытия (порядка 15—20 мк на деталь), считается достаточной для обычного исполнения реле постоянного тока.

В реле времени с магнитным демпфером, а также реле с повышенным коэффициентом возврата ставят дополнительные немагнитные прокладки из латуни или бронзы. В магнитопроводе с приставным полюсным наконечником немагнитная прокладка располагается между полюсным наконечником и сердечником.



Фиг. 1, 9. Магнитопровод прямоходовых систем:
1 — стоп; 2 — катушка; 3 — якорь; 4 — латунная трубка;
5 — скоба.

В других случаях (фиг. 1, 5) немагнитная прокладка в виде прямоугольной пластинки прикрепляется двумя винтами к якорю. Прокладка эта в реле времени в зависимости от требуемой выдержки времени бывает толщиной от 0,1 до 0,3 мм. Прокладки меньшей толщины (< 0,1 мм) исполняют в виде биметаллических пластинок — сталь — бронза, благодаря чему увеличивается стойкость прокладки и она не расклепывается при ударах якоря о сердечник.

В магнитопроводе (фиг. 1, 6) при притяннутом положении якоря должен быть оставлен небольшой зазор (порядка 0,1—0,2 мм) между якорем и полюсным наконечником.

Крепление катушки. Наиболее простое крепление катушки получается при наличии полюсного наконечника (фиг. 1, 6, 1, 7). Размер A обычно выдерживается точно, с допуском $\pm 0,3$ мм, размер же катушки по высоте может иметь отклонения порядка ± 2 мм. Для компенсации получающейся разницы в размерах ставят дополнительные шайбы из электрокартона. При исполнении без полюсного наконечника приходится ставить дополнительные пружинные шайбы из фосфористой бронзы (фиг. 1, 5), прижимающие катушку. Шайба вставляется в выточку сердечника и имеет вырезы для съема.

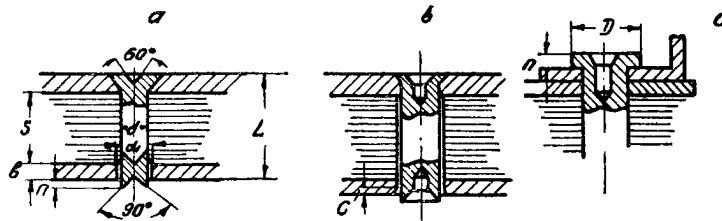
Ярмо. В реле с поворотным якорем ярмо исполняется в виде буквы Г (фиг. 1, 5) или в виде буквы П (фиг. 1, 6); в прямоходовых — в виде скоб (фиг. 1, 9) или в виде трубы.

Крепление магнитопровода. Скобка магнитопровода крепится к плите посредством винта, который завинчивается в сердечник (фиг. 1, 5) и закрепляется дополнительным штифтом от проворачивания.

Иногда сердечник магнитопровода для крепления заканчивается нарезанной частью (фиг. 1, 6). На фигуре 1, 27 показан способ крепления магнитопровода двумя винтами, завинчиваемыми в скобу. При исполнении открытых соленоидов наружные шайбы металлического каркаса делают утолщенными и к ним привинчивают угольники для крепления (фиг. 1, 21).

б) Шихтованный магнитопровод. Пакет. Пакет состоит из набора отдельных штампованных жестей толщиной 0,3—0,5 мм. Скрепление пакета осуществляется заклепками, реже винтами. Заклепки располагают так, чтобы было обеспечено плотное прилегание жестей пакета по всей его длине, обращая особое внимание на расположение заклепок у концов пакета, где жесты стремятся разойтись в виде веера. Заклепки следует располагать по одной линии пакета (средней) во избежание образования замкнутого контура, создающего дополнительные потери на вихревые токи.

Расстояние между отверстиями заклепок должно быть выдержано с допуском порядка $\pm 0,1$, $\pm 0,2$ мм. Диаметр отверстия под заклепку должен быть несколько больше диаметра заклепки (фиг. 1, 10).



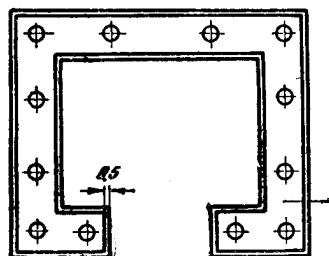
Фиг. 1, 10. Заклепочные соединения для магнитных систем.

Фиг.	d	d_1	D	v мин.	c	L	n
a	2	2,2	—	1	0,7		1
a	3	3,2	—	1	0,7		1
b, c	4	4,5	8	1,5	1	1	3
b, c	5	5,2	9	1,5	1,5	2	3

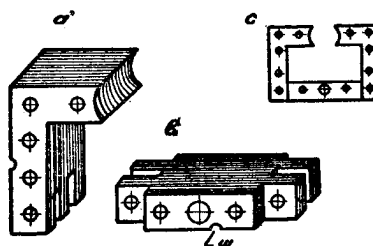
Допуск на отверстие под штифт в жестах — A_4 .

В ряде случаев для увеличения жесткости пакета наружные жести пакета штамуются из более толстого материала, чем основные (чаще всего толщиной 1 мм). Для наружных жестей требуется специальный штамп. Рекомендуется поэтому размеры наружной жести уменьшать по контуру на 0,5 мм (фиг. 1, 11), чтобы она не выступала из пакета в случае неточного совпадения ее размеров с основной.

Сложные пакеты. В конструкциях магнитопроводов сложной формы (фиг. 1, 12) пакеты исполняют разборными, чтобы можно



Фиг. 1, 11. Габариты наружных жестей.



Фиг. 1, 12. Конструкция сложного пакета:

a, b — отдельные пакеты; *c* — собранная система; *w* — штамповальный знак.

было надевать катушку. Отдельные собранные пакеты свинчиваются винтами. Для каждого вида пакетов требуется две жести — одна нормальной длины, другая укороченная.

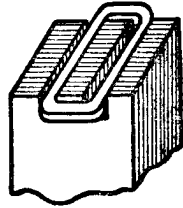
Набор пакета. Согласно требованиям ГОСТ 802—41 допуск по толщине для электротехнической стали может быть порядка $\pm 10\%$ и при сборке пакета строго по числу жестей могут получиться значительные отклонения в размерах толщины пакета. Рекомендуется поэтому в чертежах указывать не только число жестей, но и вес пакета. В производстве пакет собирают не по числу жестей, а по весу. При определении числа жестей и веса пакета следует учитывать коэффициент заполнения пакета, который зависит от способа изолирования жестей и колеблется в пределах 0,9—0,95.

Склепка пакета. Небольшие пакеты склепываются вручную — в ручном прессе. Пакеты для больших магнитных систем, например для контакторов, необходимо дополнительно подвергнуть прессовке в гидравлических прессах.

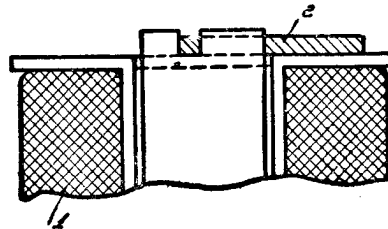
Штамповальный знак. При изготовлении штампа для жестей в нем предусматривается выштамповка так называемого знака, обычно в виде полукруга (фиг. 1, 12). Знак указывает направление штамповки при сборке пакета. Другое назначение знака — отличать жести, изготовленные одним штампом; допуски на размеры жестей и расстояния между отверстиями в них не могут быть выдержаны при изготовлении штампов с достаточно большой

точностью. Рекомендуется собирать пакет из жестей, изготовленных одним и тем же штампом; знак располагается в каждом отдельном штампе в разных местах.

Экран. Экраны, применяемые для устранения вибрации в магнитной системе, выштамповывают из полосового материала: латуни, алюминия, реже из меди. Для крепления экрана в магнитопроводе делаются прорезы. В простейшем случае экран укрепляется в торце магнитопровода и удерживается в нем благодаря плотной посадке (фиг. 1, 13 и 1, 14).



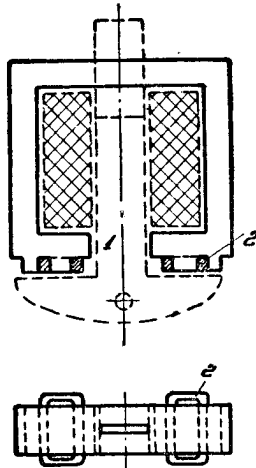
Фиг. 1, 13. Крепление экрана.



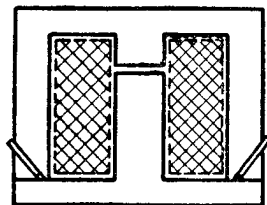
Фиг. 1, 14. Крепление экрана:
1 — катушка; 2 — экран.

Для уменьшения электрического сопротивления и увеличения поверхности охлаждения экрана (фиг. 1, 14) наружная часть его может иметь увеличенные размеры. Установка экрана в прямоходовых системах показана на фигурах 1, 15 и 1, 16, в E- и Ш-образных системах — на фигуре 1, 17.

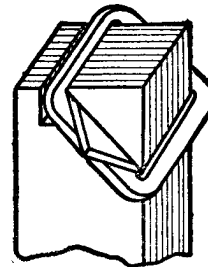
В последнем случае экран свободно входит в пазы магнитопровода и удерживается от выпадения благодаря тому, что его отгибают после вставления в



Фиг. 1, 15. Шихтованный магнитопровод прямоходовой системы:
1 — якорь; 2 — экран.



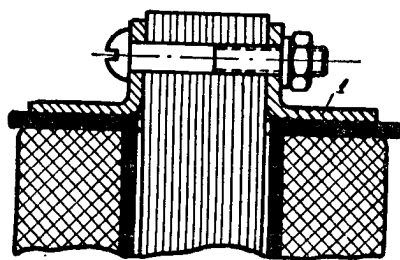
Фиг. 1, 16. Шихтованный магнитопровод прямоходовой системы.



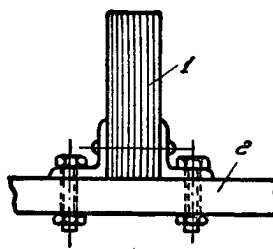
Фиг. 1, 17. Крепление экрана.

пазы; в других исполнениях для удержания витка отгибают крайнюю жесть.

Вращение якоря обычно осуществляется на осях. Якорь должен свободно ложиться на торцы сердечника и ярма, так как при малейшем перекосе якоря в притянутом положении могут образоваться большие воздушные зазоры, влекущие за собой возрастание намагничивающего тока и нагрева катушки. В реле с поворотным якорем отверстие в якоря делают несколько больше диаметра оси, в прямоходовых — внутренние размеры направляющей металлической гильзы несколько больше размеров якоря (плунжера), благодаря чему он свободно устанавливается в своем конечном положении так, что образуются минимальные воздушные зазоры между ним и торцом магнитопровода.



Фиг. 1, 18. Крепление катушки.



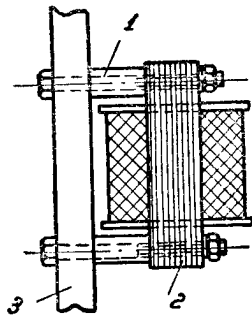
Фиг. 1, 19. Крепление шихтованного магнитопровода:
1 — магнитопровод; 2 — шайба.

Крепление катушки. В простейшем случае для небольших реле, если температура перегрева относительно невелика ($50-60^\circ$), катушка удерживается в требуемом положении экраном (фиг. 1, 14). В магнитных системах, где требуется частая смена катушки, последняя крепится специальными угольниками (фиг. 1, 18).

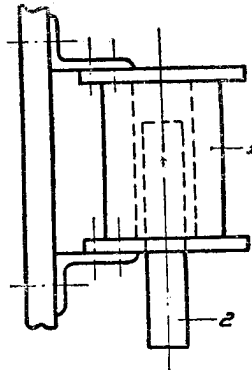
Крепление магнитопровода к основанию реле осуществляется в реле с поворотным якорем посредством дополнительных угольников, приклепываемых или привинчиваемых к ярму (фиг. 1, 19).

Магнитопровод прямоходовых систем крепится к основанию посредством винтов с надетыми на них распорными трубками (фиг. 1, 20). В некоторых случаях вместо трубок ставят дополнительные гайки. В прямоходовых системах без наружного стального магнитопровода — соленоидах — каркас катушки делается металлическим, шайбы его утолщаются, чтобы к ним можно было привинтить крепящие угольники (фиг. 1, 21).

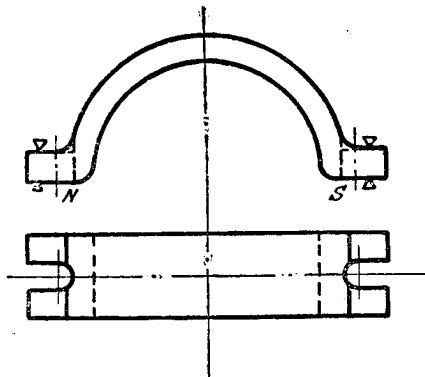
в) Магнитопровод с постоянными магнитами. При конструировании поляризованных реле с постоянными магнитами особое внимание должно быть уделено форме крепления этих магнитов. Магниты из вольфрамовой или хромистой стали, которые до закалки поддаются механической обработке (фрезеровке, сверлению), могут прикрепляться к остальным частям магнитопровода посредством винтов.



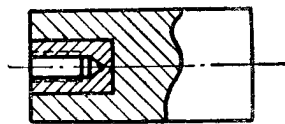
Фиг. 1, 20. Крепление шихтованного магнитопровода прямоходовой системы:
1 — трубка; 2 — магнитопровод; 3 — плита.



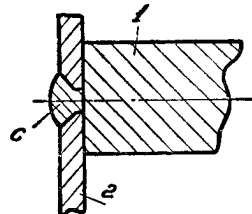
Фиг. 1, 21. Крепление соленоидов:
1 — катушка; 2 — якорь.



Фиг. 1, 22. Крепление постоянного магнита.



Фиг. 1, 23. Крепление постоянного магнита.

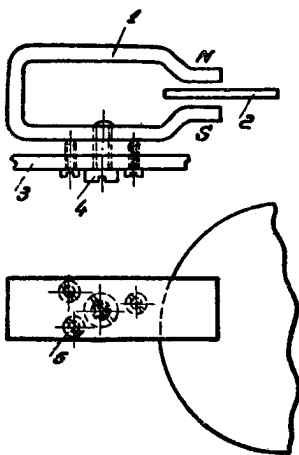


Фиг. 1, 24. Крепление постоянного магнита к стальной пластинке:
1 — магнит; 2 — пластинка.

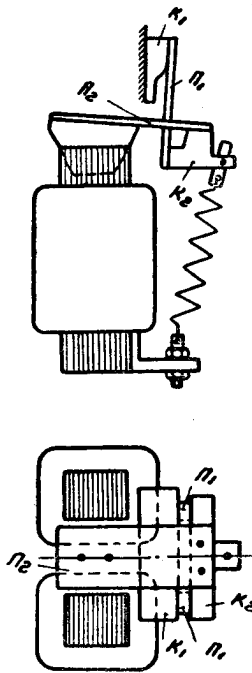
Хорошее конструктивное решение при материалах, не поддающихся механической обработке, получается исполнением магнита согласно фигуре 1, 22. При отливке такого магнита предусматривают открытые пазы для крепящих винтов. Другая форма крепления магнитов из высококоэрцитивных сплавов: в отливке постоянного магнита предусматривается канавка, которая затем шлифуется.

Магнит крепится к другой детали магнитопровода скобой из латуни.

Прямые магниты прямоугольного или круглого сечения (фиг. 1, 23) отливают со вставкой из магнитномягкого материала, в которой делается нарезка для ввинчивания крепящего винта. На наших заводах производят опыты по приварке постоянных магнитов к пластинкам из магнитномягких материалов (фиг. 1, 24): в пластинке просверливают отверстие с зенковкой и в этом месте приваривают деталь к магниту дуговой сваркой. При устройстве крепления тормозных постоянных магнитов необходимо предусматривать возможность регулирования положения магнита по отношению к диску. Для этой цели применяют четыре винта (фиг. 1, 25). Центральный винт является кре-

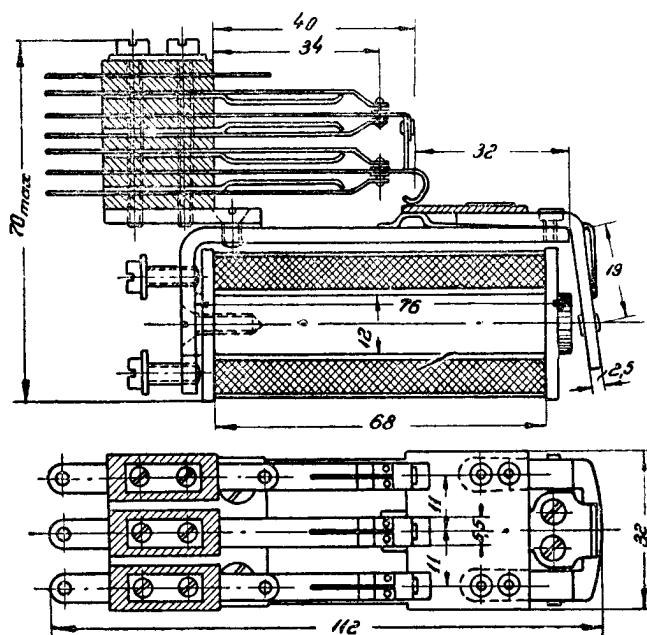


Фиг. 1, 25. Крепление постоянного магнита с центровкой его положения.



Фиг. 1, 26. Магнитопровод реле с поперечным движением якоря.

пещим, а три винта, ввинчиваемые в латунную пластинку, центрируют положение магнита.



Фиг. 1, 27. Конструкция кодового реле.

На фигуре 1, 26 приведен эскиз магнитопровода реле с поперечным движением якоря; на фигуре 1, 27 — конструкция кодового реле.

ГЛАВА 2

КАТУШКИ

2-1. ОСНОВНЫЕ ТРЕБОВАНИЯ И КЛАССИФИКАЦИЯ

Катушка является одним из основных элементов реле. Правильная и надежная работа реле, его конструктивные данные в значительной степени зависят от принятых размеров катушки и от ее исполнения. В процессе работы реле в катушке возникают различные напряжения:

механические — в катушках переменного тока витки подвергаются действию электромагнитных сил притяжения или отталкивания переменного значения; такое же явление имеет место в катушках реле постоянного тока при переходных процессах включения и отключения катушки;

термические — возникающие из-за перегревов, определяемых повышенным напряжением или током;

электрические — образующиеся из-за перенапряжений, возникающих в цепи катушки.

Требования, предъявляемые к катушке, кратко формулируются следующим образом:

катушка должна быть минимальных габаритов и удобна для производства;

катушка должна быть механически прочной;

катушка не должна перегреваться при всех возможных режимах сверх предельно-допустимой температуры;

катушка должна выдерживать определенное испытательное напряжение с целью проверки диэлектрической прочности изоляции.

при некоторых специальных режимах катушка должна отвечать требованиям влаго-, кислото- и маслостойкости, стойкости против воздействия разных химических реагентов.

В зависимости от конструктивного исполнения различают катушки:

каркасные, в которых намотка осуществляется на каркасе;

бескаркасные бандажированные — с намоткой непосредственно на съемном шаблоне; после намотки катушка бандажировается;

бескаркасные — с намоткой на сердечнике магнитной системы;

бескаркасные небандажированные.

2-2. МАТЕРИАЛЫ ДЛЯ НАМОТКИ КАТУШЕК

В качестве материала намотки применяют, как правило, медную проволоку. Алюминиевая проволока, ввиду ее большого удельного сопротивления, в релестроении не применяется. Обычно применяется проволока круглая, диаметром от 0,1 до 3 мм, квадратная — при требуемом сечении, превышающем 10 мм², и в некоторых случаях — полосовая медь. Проволоку диаметром меньше 0,1 мм применять избегают, так как она часто разрывается при намотке.

а) **Изоляция провода.** Намоточная проволока может иметь изоляцию разных видов.

Хлопчатобумажная представляет собой изоляцию относительно низкого качества, гигроскопичную. Толщина ее относительно велика. Проволока выпускается с одним слоем хлопчатобумажной изоляции (ПБО) и с двумя слоями (ПБД).

Шелковая изоляция является более совершенной, она обладает большей диэлектрической прочностью, чем хлопчатобумажная. Толщина шелковой изоляции меньше толщины хлопчатобумажной.

Выпускается проволока с одним слоем шелковой изоляции (ПШО) и с двумя (ПШД).

Эмалевая состоит из эмаль-лаков на масляной или глифталевой основе. Качество эмалевой изоляции выше, чем хлопчатобумажной и шелковой. Она тоньше их и обладает большей теплопроводностью и большей теплостойкостью. Пленка изоляции имеет хорошую влагостойкость и механическую прочность. К недостаткам эмалевой изоляции следует отнести наличие в ней ряда мелких, невидимых простым глазом, поврежденных точек, вследствие чего в намотке катушки возможны случаи образования короткозамкнутых витков. По ГОСТ 2773—44 допускается следующее число точек повреждения эмали на длине 15 м:

Марка	Диаметр проволоки от 0,05—0,14 мм	Диаметр проволоки от 0,15—0,35 мм
ПЭЛ-1	15	10
ПЭЛ-2	25	20

Комбинированная изоляция состоит из двух слоев разной изоляции. Существуют следующие марки комбинированной изоляции:

ПЭБО — слой эмали и слой хлопчатобумажной изоляции;

ПЭШО — слой эмали и слой шелковой изоляции;

ПЭТО — комбинированная эмалевая и теплостойкая изоляция.

Винифлекс является синтетической изоляцией, дающей тонкое покрытие, обладающее высокой механической прочностью. Проволока исполняется с одно-, двух- и трехслойным покрытием винифлекса.

Средняя радиальная толщина изоляции:

ПЭВ-1	ПЭВ-2	ПЭВ-3
0,02 мм	0,03 мм	0,05 мм

Провода со стеклянной и дельта-асбестовой изоляцией не применяются в релестроении.

Градации диаметров медной проволоки по ГОСТ 2773—44 приведена в таблице 2.

б) Выбор изоляции. При выборе изоляции намоточного провода принимается во внимание ряд факторов.

Механическая прочность. В отношении механической прочности все указанные виды изоляции удовлетворяют требованиям, предъявляемым к намоточной меди, за исключением эмалевой изоляции, механическая прочность которой недостаточна при относительно больших диаметрах проволоки.

Таблица 2

Проволока медная эмалированная¹

Диаметры, мм ²							
голой меди	с эмалевой изоляцией	голой меди	с эмалевой изоляцией	голой меди	с эмалевой изоляцией	голой меди	с эмалевой изоляцией
0,05	0,065	0,20	0,225	0,51	0,560	1,00	1,070
0,06	0,075	0,23	0,255	0,55	0,600	1,08	1,155
0,07	0,085	0,25	0,275	0,59	0,640	1,16	1,235
0,08	0,095	0,27	0,305	0,64	0,690	1,25	1,330
0,09	0,105	0,29	0,325	0,69	0,740	1,35	1,430
0,10	0,120	0,31	0,350	0,74	0,800	1,45	1,530
0,11	0,130	0,35	0,390	0,77	0,830	1,50	1,580
0,12	0,140	0,38	0,420	0,80	0,860	1,56	1,640
0,14	0,160	0,41	0,450	0,83	0,890	—	—
0,15	0,170	0,44	0,485	0,86	0,920	—	—
0,16	0,180	0,47	0,515	0,93	0,990	—	—
0,18	0,200	—	—	—	—	—	—

Примечания:

¹ Имеются следующие марки медной эмалированной проволоки: ПЭЛ-1 и ПЭЛ-2 — лакостойкая, изоляция на масляной основе, ПЭТ — та же проволока, но повышенной теплостойкости, изоляция на глифталевой основе, изготавливается диаметром от 0,33 мм и выше.

² Эмалированная проволока тоньше 0,05 мм изготавливается по особым техническим условиям.

Диэлектрическая прочность. Наилучшие данные по диэлектрической прочности имеет стеклянная изоляция, затем эмалевая и хромоксидная. Для получения необходимой диэлектрической прочности шелковой и хлопчатобумажной изоляции требуется применять двойную изоляцию (ПШД, ПБД).

Теплостойкость. Изоляция из шелка и хлопчатобумажная относятся к классу О и для повышения теплостойкости этой изоляции необходимо подвергнуть ее пропитке. Эмалевая и винифлексная изоляция относятся к классу А, хромоксидная — к классу В. Весьма высокой теплостойкостью обладает изоляция из стеклянной пряжи, но она гигроскопична, и ее приходится подвергать пропитке. При пропитке органическими лаками предельная допустимая температура для такой изоляции — 120°, а при пропитке кремнеорганическими смолами повышается до 175° и больше.

Химическая стойкость. Высокой химической стойкостью и стойкостью против большинства технических реагентов обладает

изоляция эмалевая, винифлексная и из стеклянной пряжи. Для повышения стойкости шелковой и хлопчатобумажной изоляции требуется дополнительная их пропитка.

Стоимость. Ниже приводятся сравнительные данные стоимости разных видов изоляции: Э—1; ШО—2,8; ШД—4,3; БД—4,0; ЭШО—3,8.

Толщина изоляции. Для получения высоких коэффициентов заполнения стремятся к применению изоляции с минимальной толщиной (см. таблицу 3).

Таблица 3

Толщина изоляции обмоточных проводов

Марка провода	Диаметр проволоки в мм								
	0,05 до 0,09	0,10 до 0,19	0,20 до 0,25	0,27 до 0,29	0,31 до 0,38	0,41 до 0,49	0,51 до 0,69	0,72 до 0,96	1,00 до 1,45
ПЭЛ1, ПЭЛ2, ПЭТ	0,015	0,020	0,025	0,03	0,04	0,045	0,05	0,06	0,07
ПЭЛ ШО	0,07	0,075	0,09	0,1	0,105	0,110	0,115	0,125	0,138
ПЭЛ БО	—	—	0,125	0,155	0,16	0,165	0,17	0,18	0,21
ПБД	—	—	0,19	0,22	0,22	0,22	0,22	0,22	0,27

Дефицитность. Наименее дефицитной является проволока с эмалевой изоляцией. В практике релестроения наиболее часто применяют проволоку с эмалевой изоляцией при намотке катушек проволокой до 0,9 мм. При больших диаметрах проволоки применяют изоляцию БД.

Катушки напряжения переменного тока часто исполняют с изоляцией ШД при желании осуществить намотку без бумажных прокладок. Для эмалевой изоляции прокладки в этом случае обязательны.

2-3. ВСПОМОГАТЕЛЬНЫЕ ИЗОЛЯЦИОННЫЕ МАТЕРИАЛЫ

При изготовлении каркасных и бескаркасных катушек применяют разного вида вспомогательные изоляционные материалы, служащие для изготовления каркасов, дополнительной их изоляции, изоляции выводов, наружных слоев намотки, бандажировки катушек, изоляции между слоями намотки. В таблице 4 приводятся краткие технические данные таких изоляционных материалов.

Таблица 4

Изоляционные материалы

Наименование	Толщина, мм		Пробивная напряженность кв/мм или пробивное напряжение, кв	Удельный вес	Примечания
	Размер	Допускаемое отклонение			
Электрокартон марки ЭВ (электропресспан)	0,1; 0,15; 0,2; 0,3 0,4; 0,5 1; 1,25 1,5; 1,75; 2; 2,5; 3	$\pm 0,015$ $\pm 0,3$	11 кв/мм 7,5 кв/мм	от 1,5 до 0,95	Вырабатывается из целлюлозы и льняного волокна. Рекомендуется пропитка лаком
	Гетинакс	0,5; 0,6 0,8; 1,0 1,2; 1,5 2 2,5; 3,0; 3,5	$+0,15$ $-0,03$ } $\pm 0,1$ $+0,12$ } $-0,15$ } $+0,15$ } $-0,2$ } $\pm 0,2$ }	15—23 кв 30—42 кв 40—55 кв	
Текстолит литой	0,5; 0,6; 0,8; 1,0 1,5 2 2,5; 3,0; 3,5	$\pm 0,1$ $\pm 0,15$ $\pm 0,2$ $\pm 0,25$	до 1 мм 3,5 кв/мм от 1 до 2 мм 3 кв/мм от 2 до 3 мм 2 кв/мм	1,3—1,4	Аналогичен гетинаксу, но с заменой бумаги текстильным материалом
ЛШ 1 Лако-ткань шелковая	0,1 0,15	$\pm 0,015$ $\pm 0,02$	2,1 кв 3,0 кв		Шелк, пропитанный лаком, светлым или черным
ЛХ 1 Лако-ткань хлопчатобумажная	0,15; 0,17; 0,20 0,24;	$\pm 0,02$ $\pm 0,03$	0,9—1,4 кв		Ткань, пропитанная лаком
Миканит прокладочный	0,5; 0,6 0,7; 0,8 0,9; 1,0 1,5; 2,0 2,5; 3,0		15—20 кв/мм		Слои щипаной слюды, склеенные изоляционными лаками с последующей опрессовкой

Продолжение таблицы 4

Наименование	Толщина, мм		Пробивная напряженность кв/мм или пробивное напряжение, кв	Удельный вес	Примечания
	Размер	Допускаемое отклонение			
Микалента	0,08 0,10 0,13 0,17	$\pm 0,01$ $\pm 0,01$ $\pm 0,02$ $\pm 0,03$ $-0,02$	от 20 кв/мм до 14 кв/мм		Гибкая лента, состоящая из одного слоя щипаной слюды, склеенной с бумагой
Микафоль	0,15 0,20	$+0,02$ $-0,03$ $+0,04$ $-0,03$ $\pm 0,05$	Не менее 13—16 кв/мм		Состоит из нескольких слоев щипаной слюды, наклеенных на бумагу
Бумага: конденсаторная пропиточная телефонная кабельная	От 0,007 до 0,0024 0,12 0,05 0,08 0,12 0,17	$\pm 0,01$ $\pm 5\%$ $\pm 0,005$ $\pm 0,007$ $\pm 0,01$	От 275 до 450 кв	0,97 0,65 0,82 0,7	
Лента: киперная тафтяная	0,45 0,25	$+0,02$ $+0,02$			Хлопчатобумажная
Линоксиновые трубки	Номинальный внутренний диаметр, мм 1,0; 1,5; 2,0; 2,5; 3; 3; 5; 4; 4,5; 5; 6; 7; 8; 9; 10; 11; 12	$\pm 10-30\%$	1000—1200 в		Плетеные чулки из хлопчатобумажной, шелковой или стеклянной пряжи, пропитанные светлым изоляционным лаком

В последнее время разработаны новые виды изоляционных материалов, весьма теплостойких:

Стеклолакоткань — лакоткань на комбинированной глифталевой и асфальто-масляной основе. Толщина изоляции 0,2—0,3 мм, класс изоляции В.

Стекломиканит — гибкий теплостойкий электроизоляционный материал толщиной 0,2—0,5 мм, состоящий из одного или

двух слоев тонкой стеклоткани и одного или нескольких слоев щипаной слюды, склеенных высокотеплостойким лаком.

Стеклослодяная лента — гибкая теплостойкая изоляция толщиной 0,18—0,25 мм.

Стеклотканевые трубки и стеклошнур.

Эти новые виды изоляционных материалов применяются для изоляции катушек, намотка которых исполнена из проволоки с теплостойкой изоляцией.

2-4. НАМОТОЧНЫЕ СТАНКИ

При выборе типа катушки и расчете ее необходимо учитывать конструкцию намоточного станка, на котором эта катушка изготавливается. К намоточным станкам предъявляются следующие требования:

1. Проволока должна укладываться точно: один ряд за другим. В противном случае в намотке образуются бугры, узлы и впадины, что приводит к уменьшению полезного объема катушки.

2. Скорость вращения станка и скорость подачи проволоки должны соответствовать размеру проволоки. Станок должен быть поэтому снабжен устройствами для регулировки числа оборотов шпинделя и подачи.

3. Натяжение проволоки при намотке должно быть определенным, не превышающим предельно допустимого. При повышенном натяжении возможны растягивание проволоки и частые разрывы ее.

4. Станок должен иметь приспособление для автоматической прокладки бумаги между слоями проволоки.

5. При всяком обрыве проволоки станок должен немедленно останавливаться.

6. Станок должен допускать возможность одновременной намотки нескольких катушек.

7. Станок должен быть снабжен счетчиком оборотов. Желательно устройство счетчика, обеспечивающее установку на определенное число витков, по достижении которых станок автоматически останавливается.

Существуют станки следующих типов.

Намоточные автоматы, полностью удовлетворяющие перечисленным выше требованиям. Они снабжены устройствами автоматической подачи проволоки с регулировкой подачи. Величина натяжения проволоки регулируется пружинами. Станки снабжаются соответствующими счетчиками, приспособлениями для прокладки бумаги, для остановки станка при обрыве проволоки или по достижении требуемого числа оборотов. На намоточном автомате можно проводить одновременную намотку 8—10 катушек. Наматывать катушки можно проволокой от 0,05 до 0,5—0,6 мм диаметром. Коэффициент заполнения катушки получается высоким, число ее витков точно соответствует заданному.

Полуавтоматы отличаются от автоматов тем, что они не оборудованы полностью всеми названными выше приспособлениями: не снабжены устройством для автоматической прокладки бумаги, для останковки станка при обрыве. На полуавтомате осуществляется одновременная намотка только одной катушки. При работе на полуавтоматах можно получить катушки такого же качества, как и на автоматах, но при условии наличия квалифицированных рабочих и максимального внимания их к процессу намотки.

Ручные намоточные станки. В таких станках натяжение проволоки и подача ее осуществляются от руки. Как правило, при работе на ручных станках получается «дикая намотка» с беспорядочной укладкой рядов и слоев. На ручных станках производится намотка только каркасных катушек. Следует учитывать более низкий коэффициент заполнения, получающийся при ручной намотке.

Тормозные станки. Изготовление токовых катушек с проволокой больше 0,9 мм в диаметре связано с некоторыми особенностями процесса: намотка осуществляется при небольшом числе оборотов, за укладкой проводов должен следить непосредственно рабочий, причем приходится часто останавливать станок, давать ему обратный ход для повторной подгонки проволоки, в некоторых случаях приходится по ходу намотки закреплять проволоку хлопчатобумажной лентой, так как в бескаркасных катушках с толстой проволокой не применяется прокладка бумаги. Для намотки таких катушек применяются специальные станки, называемые тормозными.

Тормозные станки отличаются от обычных ручных намоточных станков в основном большей мощностью двигателя, малым числом оборотов шпинделя, наличием тормозного и реверсирующего устройств.

2-5. КАРКАСНЫЕ КАТУШКИ

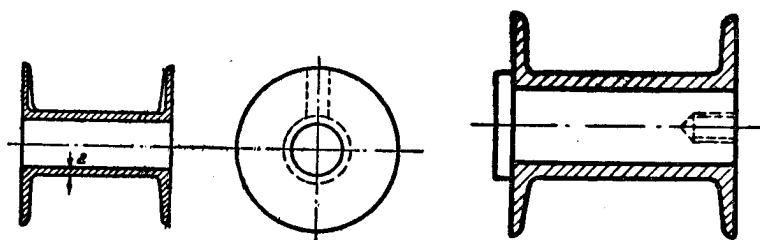
Намотка катушки осуществляется на заранее подготовленном каркасе. Различают следующие виды каркасов.

Прессованные каркасы из тенацита — порошка, состоящего из древесных опилок и бакелитовой смолы. Применение таких каркасов рационально при массовом выпуске изделия, когда стоимость прессформы для прессовки каркаса экономически оправдывается.

На фигуре 2, 1 показано примерное исполнение тенацитового каркаса. Отверстие для сердечника может быть круглым или прямоугольным. Толщина трубки и стенок каркаса, которую выбирают исходя из механических соображений, должна быть не менее 2 мм. Для облегчения съема прессованной детали с прессформы стенки каркаса имеют с внутренней стороны небольшой уклон ($\frac{1}{100}$ или $\frac{1}{50}$). В этих же стенках выпрессовываются углубления для прокладки выводов (см. пунктир).

Фарфоровые каркасы, имеющие конструкцию, аналогичную тенацитовым.

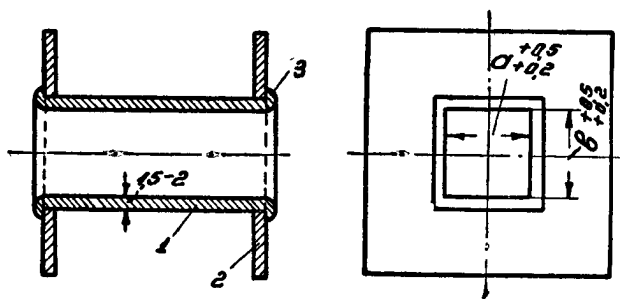
Для реле небольших размеров применяется конструкция тенацитового каркаса, опрессованного непосредственно на сердечнике (фиг. 2, 2).



Фиг. 2, 1. Тенацитовый каркас.

Фиг. 2, 2. Тенацитовый каркас, опрессованный на сердечнике.

Сборные каркасы из слоистой изоляции (фиг. 2, 3), состоящие из гетинаксовой трубки и шайб из гетинакса или текстолита. Трубка выполняется из гетинакса, изготовленного не на бакелитных, а на термореактивных неполимеризующихся лаках



Фиг. 2, 3. Сборный каркас.

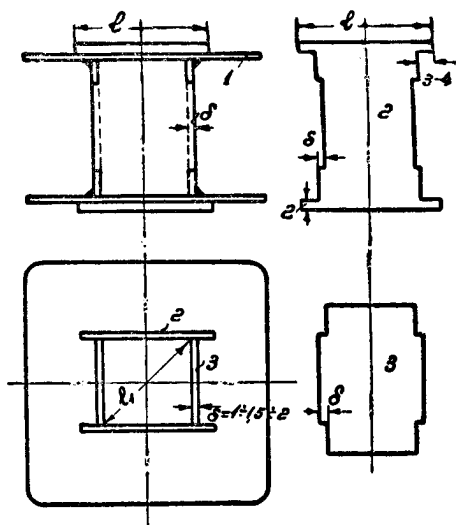
1 — изолит; 2 — текстолит; 3 — развальцевать.

(шеллак, глифталевый лак), благодаря чему трубка при нагреве под прессом может деформироваться.

Сборка каркаса осуществляется следующим образом: шайбы одеваются на трубку, и собранную конструкцию ставят в ручной пресс, где производится развальцовка трубки горячими пуансонами. В результате получается прочный, механически и электрически надежный каркас. Трубка выполняется так, чтобы внутренние размеры ее имели только положительный допуск порядка $+0,5$, $+0,2$ мм.

Удачная конструкция сборных каркасов, изготовляемых в последнее время на отечественных заводах, приведена на фигуре 2, 4. Каркас состоит из деталей, отштампованных из гетинакса толщиной 1—1,5 мм и для больших катушек — 2 мм. Деталь 2

имеет выступы, которыми удерживаются щеки 1. Размер l детали 2 должен быть несколько меньше размера l_1 по диагонали каркаса, благодаря чему осуществляется возможность надевания щек на деталь 2 по диагонали.

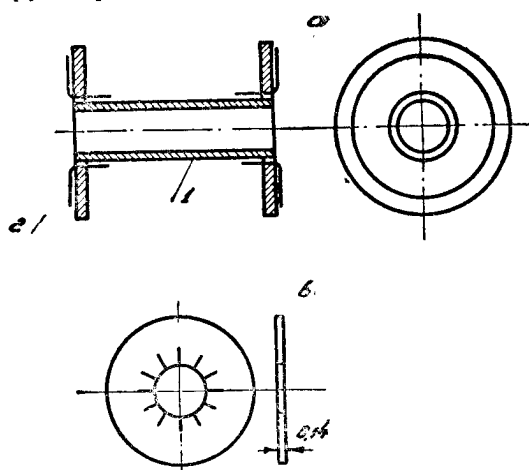


Фиг. 2, 4. Сборный каркас.

Наличие выступающих частей обеспечивает жесткость каркаса. После сборки каркаса углы между щеками и боковыми стенками заклеивают шеллачным лаком (или бакелитным с последующей запечкой).

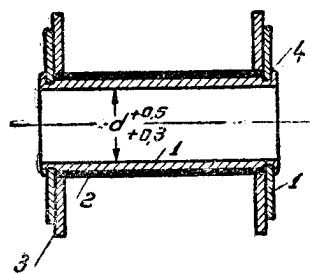
Клееные каркасы, в которых шайбы и трубки из электрокартона склеиваются посредством полотняных разрезных шайб или лент, расположенных по всей окружности каркаса (фиг. 2, 5, а). Клееные каркасы применяются только для редких, индивидуальных исполнений катушек. Эскиз заготовки шайбы показан на фигуре 2, 5, б.

Металлические каркасы, исполняемые в виде латунной трубки с насаженными по краям латунными шайбами. Концы трубки развальцовываются и шайбы припаиваются к ней (фиг. 2, 6).



Фиг. 2, 5. Клееный каркас: а — каркас; б — заготовка шайбы; 1 — прессшпан; 2 — полотно.

Изоляция катушки от трубки осуществляется посредством добавочных изоляционных шайб и изоля-



Фиг. 2, 6. Металлический каркас:

1 — латунь; 2 — лакополотно; 3 — гетинакс или текстолит; 4 — припаять.

ционных полос из лакополотна или микаполотна, наматываемых на металлическую трубку. Перед изолировкой трубки рекомендуется замазывать изоляционным клеем углы между изоляционными шайбами и трубкой.

Примерные размеры элементов каркаса следующие:

- толщина трубки и шайб — порядка 0,5 мм;
- толщина изоляционной шайбы — порядка 1 мм;
- толщина изоляции трубки — порядка 0,4—0,5 мм.

Обычно трубку для каркасов изготавливают из полосовой или листовой латуни (спайка шва не обязательна). Для некоторых исполнений катушек удается подбирать готовые латунные трубки.

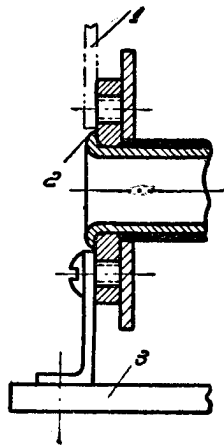
Металлические каркасы применяются в основном в конструкциях прямоходовых электромагнитных реле, в которых подвижный сердечник перемещается внутри катушки. Имеются конструкции соленоидных реле, в которых на металлическом каркасе крепятся конструктивные элементы реле: рычаги, контактные системы (фиг. 2, 7). В этих случаях наружные металлические шайбы каркаса должны иметь толщину, достаточную для крепления других деталей. Шайбы должны быть толщиной порядка 4 мм, трубка — толщиной 1 мм. Для катушек переменного тока трубка и шайбы металлического каркаса исполняются с разрезом в одном месте.

Намотка каркасных катушек может быть осуществлена на любом намоточном станке. На катушках постоянного тока не делается добавочных бумажных прокладок между слоями. В ряде случаев за 2—3 слоя до конца намотки ставится прокладка из телефонной бумаги или лакополотна, благодаря чему скрадываются неровности всей намотки. При намотке катушек переменного тока из проволоки с эмалевой изоляцией рекомендуется ставить через каждый слой проволоки прокладки из изоляционной бумаги.

Имеются катушки с двумя отдельными обмотками, разделенными шайбами. Преимуществом таких катушек является уменьшение напряжения между двумя соседними слоями. При изготовлении таких катушек производится намотка одной половины, другая заполняется временной деревянной колодкой. Наружная поверхность готовых катушек покрывается полосой лакополотна, наматываемой с некоторым нахлестом для склеивания концов.

Описание крепления выводов приводится в параграфе 2-7.

Готовая катушка снабжается табличкой, в которой указываются данные катушки — рабочее напряжение, число витков, диаметр проволоки, сопротивление намотки при 20°.



Фиг. 2, 7. Металлический каркас с утолщенными флянцами.

1 — деталь для крепления частей реле; 2 — припаять; 3 — плата.

2-6. БЕСКАРКАСНЫЕ КАТУШКИ

Изготовление катушки. Намотка бескаркасных катушек производится на временных разъемных металлических шаблонах, укрепляемых на валу намоточного станка. Наружные размеры шаблонов соответствуют внутренним размерам катушки. Различают два вида бескаркасных катушек: бандажированные и небандажированные. Наибольшее распространение имеют бандажированные катушки, обладающие после компаундировки значительной механической прочностью.

В магнитных системах постоянного тока в целях улучшения теплоотдачи исполняют иногда катушки, наматываемые непосредственно на сердечник или на железную трубку, которая одевается на сердечник магнитной системы.

В тех случаях, когда катушка в аппарате полностью защищена от внешних механических воздействий, например в реле с кожухом или в соленоидном реле с замкнутой магнитной системой, возможно применение небандажированных бескаркасных катушек.

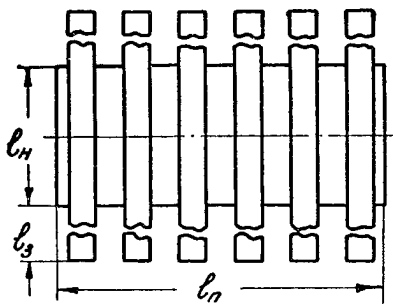
Как правило, процесс изготовления бескаркасных катушек состоит из трех стадий:

- 1) намотки катушки на станке;
- 2) бандажировки катушки после съема ее со станка;
- 3) компаундировки катушки.

Ниже коротко описываются первые две стадии изготовления катушки из различного обмоточного материала. Процесс компаундировки катушек рассматривается в параграфе 2-8.

Намотка и бандажировка бескаркасной катушки с эмалевой изоляцией. Процесс намотки состоит из следующих операций.

1. Заготавливается полоска из микаполотна, к которой приклеивают изоляционным клеящим лаком тафтяные ленты (фиг. 2, 8) с расстоянием между ними, равным их ширине. Длина полоски из микаполотна выбирается с таким расчетом, чтобы при наворачивании ее на шаблон получилась нахлест порядка 10—15 мм; $l_n = \pi D_{вн} + (10-15)$ мм, ширина ее l_n должна быть равна длине намотки. Длина тафтяных лент должна быть такой, чтобы ими



Фиг. 2, 8. Заготовка для бескаркасной катушки.

можно было укрепить всю намотку катушки.

2. Полоска микаполотна наматывается на съемный шаблон, закрепленный на валу станка (фиг. 2-9).

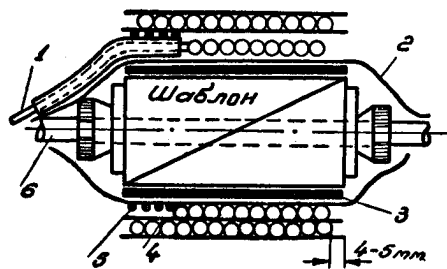
3. Укрепляется нитками у левого или правого края катушки вывод из нее (начало). Выводной конец и ленты закрепляют

на валике для предохранения их от болтания в процессе намотки.

4. Выполняется намотка катушки с прокладкой бумаги через каждый слой. Ширина бумажной прокладки должна быть больше длины намотки на 10 мм.

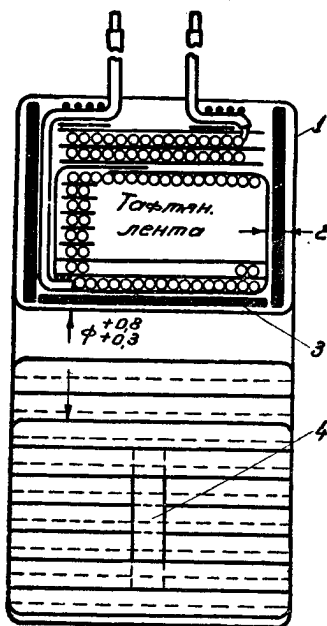
5. За два-три слоя до конца намотки станок останавливают и производят крепление катушки тафтяными лентами, предварительно приняв выступающие концы бумаги (фиг. 2, 10). Процесс крепления осуществляется посредством вспомогательного бандаж, наматываемого на левой половине катушки, с правой же стороны ее продолжают намотку и после того, как ленты закрепляются проволокой, снимают вспомогательный бандаж.

6. Закрепляются нитками на поверхности катушки выводы (фиг.2, 10).



Фиг. 2, 9. К описанию процесса намотки бескаркасной катушки;

1 — вывод; 2 — лента; 3 — трубка из лакоплетна; 4 — бумажная прокладка; 5 — нитка; 6 — валик.



Фиг. 2, 10. Разрез бескаркасной бандажированной катушки;

1 — бандаж из тафтяной ленты; 2 — шайба из микаплетна; 3 — трубка из микаплетна; 4 — табличка.

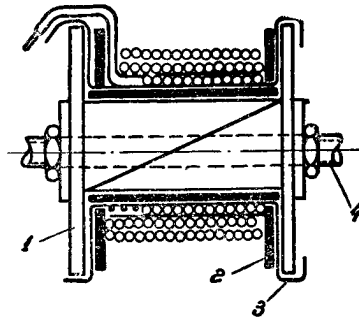
7. Катушка снимается с шаблона и опрессовывается в специальной оправке. При этом уминаются выступающие концы бумаги и катушка доводится до требуемых по чертежу размеров.

На торцах катушки накладываются шайбы из микаплетна и вся катушка бандажировается тафтяной лентой с перекрытием, равным $\frac{1}{3}$ ширины ленты. Разрез готовой катушки приведен на фигуре 2, 10.

После окончания вышеуказанных операций осуществляется процесс компаундировки катушки.

Намотка и бандажировка бескаркасной катушки из проволоки с шелковой или хлопчатобумажной изоляцией. Такие катушки исполняются без бумажных прокладок. Для формирования катушки и удержания слоев в процессе намотки шаблон снабжается

боковыми шайбами (фиг. 2, 11). На шаблон наматывается заранее заготовленная полоска микаполотна (фиг. 2, 8) и одеваются боковые шайбы из микаполотна. В остальном процесс намотки совпадает с таким же процессом для катушек из проволоки с эмалевой изоляцией (но без применения бумажных прокладок).

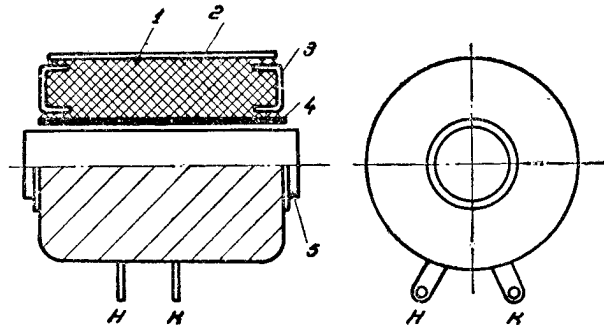


Фиг. 2, 11. Намотка бескаркасной катушки из проволоки с шелковой изоляцией:
1 — шаблон; 2 — шайба, 3 — ленты;
4 — валик станка.

Намотка и бандажировка катушки на сердечнике магнитной системы (фиг. 2, 12). Сердечник покрывается прочно скрепленным с ним тонким слоем изоляции — изолируется или миканизируется. По краям к сердечнику подклеивают разрезные полотняные шайбы (фиг. 2, 5). По окончании процесса намотки шайбы расправляются по торцу катушки, концы их загибают на наружную поверхность, которая затем бандажируется.

Намотка небандажированной бескаркасной катушки (фиг. 2, 13).

В конструкции этой катушки отсутствуют крепящие тафтяные ленты, торцевые шайбы и не применяется бандаж. Намотка катушки производится на металлической оправке. Основой катушки является трубка из бумаги, изготавливаемая на этой же оправке



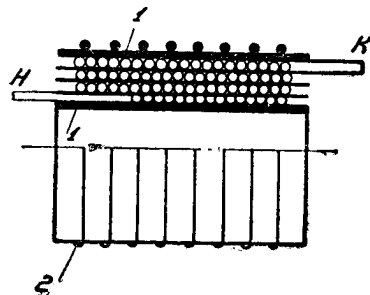
Фиг. 2, 12. Бескаркасная катушка с намоткой на сердечнике:
1 — прокладка микаполотна; 2 — бандаж, 3 — шайба полотняная;
4 — гильза; 5 — сердечник.

намоткой нескольких слоев (6—7) бумаги общей толщиной порядка 0,125 мм. В некоторых случаях применяют готовые трубки с толщиной стенок, равной 0,5 мм. Крепление катушки осуществляется с помощью бумажных прокладок, ширина которых берется такой, чтобы создавалась торцевая изоляция размером с каждой стороны по 2,5—3 мм. Сверх последнего ряда катушку обматывают одним слоем бумаги толщиной 0,125 мм, который закреп-

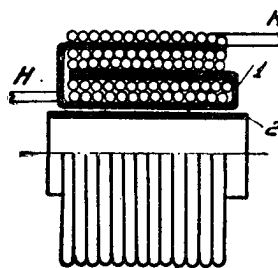
ляется затем тонкой ниткой. Выводы катушки — начало и конец — проходят через торцы катушки.

На автоматических намоточных станках можно осуществить одновременную намотку 8—10 описанных катушек. Длина оправки, трубки и бумажных прокладок соответствует суммарной длине всех катушек. По окончании процесса вся намотка разрезается специальными ножами на отдельные катушки, которые поступают затем в компаундировку.

Намотка бескаркасных токовых катушек. Как правило, для таких катушек применяется медь с хлопчатобумажной изоляцией. Намотка производится на тормозных станках. Основой катушки



Фиг. 2, 13. Бескаркасная небандажированная катушка:
1 — бумага толщиной 0,125 мм;
2 — нитка.



Фиг. 2, 14. Бескаркасная токовая катушка из проволоки с хлопчатобумажной изоляцией:
1 — лента; 2 — трубка.

является изоляционная трубка. Для крепления витков катушки и удержания порядка слоев намотка в процессе изготовления затягивается тафтяной лентой (фиг. 2, 14). Свободный конец ленты закрепляется ниткой. По окончании намотки катушка подвергается пропитке.

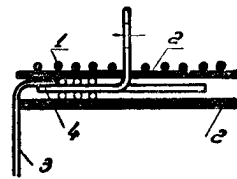
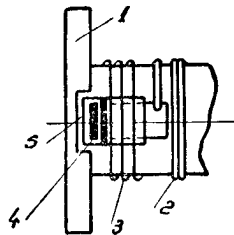
Маркировка катушек. Готовые катушки снабжаются табличкой, в которой указываются данные катушки: рабочее напряжение, число витков, диаметр проволоки, сопротивление при 20°. Табличка выполняется из прессшпана и укрепляется на катушке бандажом (фиг. 2, 10).

2-7. ВЫВОДЫ КАТУШЕК

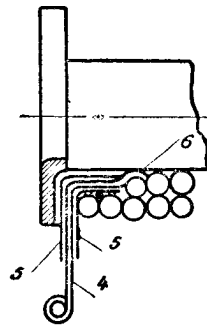
В практике применяются различные исполнения выводов катушек.

а) В каркасных катушках небольших размеров вывод осуществляется посредством флажков — латунных полосок толщиной 0,5—0,6 мм, шириной 4—6 мм, которые припаиваются концами к проводу намотки. Монтажный провод в свою очередь припаивается к флажкам. Флажки изолируются от намотки посредством двух полосок лакополотна; крепление флажков к намотке осуществляется тонкими нитками (фиг. 2, 15).

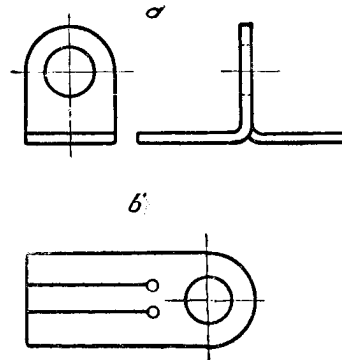
б) В каркасных катушках реле вывод может быть выполнен проводом намотки, присоединяемым к выводным зажимам, расположенным на цоколе. Изоляция выводов от намотки и от цоколя реле осуществляется посредством линооксиновых трубок. В случае, если намотка катушки производится тонкой проволокой, диаметром меньше 0,5 мм, необходимо применять специальные выводные провода большего диаметра.



Фиг. 2, 16. Крепление угольника в бескаркасной катушке:
1 — нитки; 2 — лакополотно;
3 — провод; 4 — припаять.



Фиг. 2, 15. Вывод флажком:
1 — каркас; 2 — намотка; 3 — нитки; 4 — флажок; 5 — лако-
ткань; 6 — припаять.



Фиг. 2, 17. Угольник для выводов:
b — заготовка.

В катушках с намоткой из весьма тонкой проволоки (диаметром 0,05—0,1 мм) в целях защиты наружных слоев от механических повреждений последние 2—3 слоя намотки производят более толстым проводом (порядка 0,6 мм), которым и осуществляется вывод конца катушки.

в) В бескаркасных катушках вывод осуществляется гибким проводом. Для этой цели применяют гибкие кабели, щеточные провода — жгуты с надетыми на них линооксиновыми трубками, или гибкие провода с кембриковой изоляцией. Провода начала и конца припаиваются к проводам намотки и укрепляются в катушке суровыми нитками.

г) В бескаркасных катушках применяется еще исполнение выводов с контактными угольниками. Угольники выштамповываются из латуни, устанавливаются на наружной поверхности

катушки (до ее бандажировки) и укрепляются посредством суровых ниток (фиг. 2, 16). Изоляция угольника (фиг. 2, 17) от намотки осуществляется посредством двух полосок из микаполотна или лакополотна. К контактному угольнику припаиваются провода начала и конца намотки, причем, если намотка производится проволокой диаметром меньше 0,15 мм, ставятся промежуточные провода диаметром не меньше 0,25 мм. Монтажный провод крепится к угольникам посредством винтов и гаек.

д) Промежуточные выводы катушек осуществляются проволокой намотки, выпускаемой из катушки в виде петли (без разрыва обмотки). Конец петли оголяется и к нему припаивается монтажный провод. Промежуточные выводы провода изолируются в намотке полосками из лакополотна.

2-8. ОТДЕЛКА КАТУШЕК

а) **Виды отделки.** В практике применяются различные виды отделки катушек, выбор их определяется режимом эксплуатации аппарата, а также и конструкцией катушки.

Каркасные катушки реле, снабженных закрытым пыленепроницаемым кожухом и предназначенных для установки в закрытых сухих помещениях, дополнительной отделке после намотки не подвергаются. Для улучшения внешнего вида таких катушек наружная поверхность их намотки обвертывается тонким слоем лакополотна, с небольшим нахлестом для склейки концов полотна. Дополнительная отделка обязательна для катушек реле, не имеющих кожуха, и для всех бескаркасных катушек.

Различают следующие виды отделки катушек:

Пропитка катушек пропиточными электроизоляционными лаками. При пропитке катушек улучшаются качества пористой и, в частности, волокнистой твердой изоляции элементов катушки: бумаги, картона, пряжи — шелковой, хлопчатобумажной и др. После пропитки поры в изоляции оказываются заполненными уже не воздухом, а высохшим лаком, благодаря чему увеличиваются пробивное напряжение и теплопроводность, уменьшается гигроскопичность и улучшаются механические свойства изоляции.

Пропитка катушек обычно осуществляется при определенном давлении, порядка 7 ат. Катушки с проволокой диаметром 0,9 мм и выше можно пропитывать и не под давлением.

Компаундировка катушек производится пропиточными компаундами при температуре 150—160° и при давлении порядка 7 ат. Жидкий компаунд при повышенной температуре и высоком давлении проникает в поры и капилляры изоляции обмоток, заполняет промежутки между проволоками, вследствие чего катушка получается в виде монолитной массы без пор. Создается надежное механически прочное сцепление между отдельными витками, что особенно важно для катушек переменного тока, а также для катушек постоянного тока, работающих при повторно-кратковременном

режиме с большой частотой включения. В этих случаях при отсутствии достаточного механического сцепления между отдельными витками могут появляться короткозамкнутые витки в результате истирания междувитковой изоляции, вызванного взаимным перемещением проволок при включении и отключении катушек.

Компаундированные катушки обладают высокой влагостойкостью и стойкостью против действия воды, паров, кислот и щелочей. Вследствие заполнения компаундом пор и капилляров повышается теплопроводность изоляции по толщине катушек.

Процесс компаундировки катушек является более рациональным, чем процесс пропитки, что определяется следующим:

1. Пропиточные лаки содержат большое количество летучих (порядка 50—60%), вследствие чего при просушке катушки в ней образуются поры. Пропиточные компаунды обладают весьма незначительной усадкой, содержат в себе незначительное количество летучих, поэтому компаундированная катушка получается монолитной.

2. Кислотность компаундов чрезвычайно низка, что исключает возможность разрушения хлопчатобумажной изоляции и окисления меди обмотки.

Пропитка катушек в горячих смолах, но без давления. Такой процесс применим только для бескаркасных катушек, не имеющих торцевого бандажа (фиг. 2, 13).

Покрытие лаками. Покровные лаки служат для образования на поверхности катушки механически прочной пленки, повышающей влагостойкость изоляции, стойкость против действия химически активных реагентов, уменьшающей возможность пристаивания к катушке пыли и грязи, улучшающей внешний вид катушки.

Особое значение имеет применение пигментированных покровных лаков (пигментированные эмали), содержащих в своем составе измельченный в тонкий порошок неорганический наполнитель — пигмент, придающий лаковой пленке определенную окраску и увеличивающий ее механическую прочность.

Покрытие лаками рекомендуется применять в следующих случаях:

1) для токовых катушек из проволоки с хлопчатобумажной изоляцией;

2) для катушек, предварительно подвергнутых пропитке в асфальтовых или асфальтомазных лаках;

3) для компаундированных катушек, работающих в особо тяжелых режимах в отношении влияния внешней среды. Эти катушки покрывают пигментированными лаками.

б) Процессы отделки. Ниже приводится описание нескольких примерных процессов отделки катушек.

Пропитка токовых катушек из проволоки с хлопчатобумажной изоляцией производится черными асфальтомазными лаками. Перед пропиткой катушки подвергаются сушке при температуре

105—110° в течение 3 часов, затем им дают охладиться до 60—70° и еще теплыми погружают на 4 часа в сосуд с лаком. После этого катушки раскладываются на 30 минут на сетке для стекания лака и затем снова подвергаются сушке при температуре 110—120° в течение 10 часов.

Пропитка катушек из проволоки с эмалевой изоляцией осуществляется под давлением в автоклаве черными асфальтомаляными лаками, не содержащими в себе в качестве растворителя бензола, разрушающе действующего на эмалевую изоляцию. Перед пропиткой катушки прогреваются в печи при температуре 65—80° в течение 3 часов и затем погружаются в автоклав. В автоклаве создается давление до 7 ат, которое выдерживается в течение 5 минут, затем давление снимается на 5 минут и снова повышается до 7 ат. Выдержка под давлением и без давления (или под вакуумом) повторяется три раза, после чего катушки выгружаются из автоклава, раскладываются на 30 минут на сетках для стекания лака и затем подвергаются окончательной сушке при температуре 120° в течение 36 часов. После пропитки катушки покрываются лаком.

Компаундировка катушек с проволокой из эмалевой изоляции. Каркасные катушки можно компаундировать только при снятых торцевых шайбах и без наружного бандаж. Катушки с бескаркасной намоткой могут подвергаться компаундировке в законченном виде. Пропиточный компаунд представляет собой битуминозную массу, получаемую из нефти. Температура размягчения компаунда 105—113°, при температуре 150° масса становится жидкой и однородной.

Весь процесс включает в себе следующие основные стадии.

1. Дополнительная бандажировка катушек хлопчатобумажной лентой (киперной).
2. Сушка в компаундировочном котле (без компаунда) в течение 3—4 часов, при температуре 110—120° при нормальном давлении, а затем в течение 30 минут при вакууме.
3. Тренировочный режим, заключающийся в том, что катушки подвергаются в котле с компаундом попеременно давлению в 7 ат в течение 5 минут, а затем вакууму (740—750 мм) в течение такого же времени, причем этот процесс повторяется пять раз.
4. После тренировочного режима катушки выгружаются из котла, и с них снимается временный бандаж с налипками компаунда.

На наших заводах разработан процесс компаундирования бандажированных катушек, при котором избегается необходимость накладки дополнительного временного бандаж, применяемого для устранения наливов компаунда и остатков примесей битума на поверхности готовой катушки.

В новом процессе битум загружается в варочный котел переплавленным, применена специальная конструкция каркаса для загрузки катушек. Этот каркас исполнен так, чтобы в процессе

передавливания битума из компаундировочного котла в варочный битум стекал с верхних слоев катушек, не попадая на нижние.

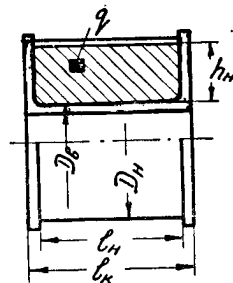
Покрытие покровными лаками для компаундированных катушек производится только при работе катушек в особых условиях.

Пропитка без давления в горячих смолах производится при температуре 160°. Такой пропитке можно подвергать бескаркасные небандажированные катушки из проволоки с эмалевой изоляцией.

Покраска катушек производится или пульверизатором, или процессом окунания. Как правило, применяют покровные лаки воздушной сушки, например, черные асфальтомаляные лаки, имеющие длительность сушки на воздухе около трех часов. Пигментированные лаки требуют печной сушки.

2-9. КОЭФФИЦИЕНТ ЗАПОЛНЕНИЯ

а) **Общие определения.** Коэффициентом заполнения катушки (f_k) называют отношение сечения меди катушки ко всему сечению катушки. Так, если катушка (фиг. 2, 18) длиной l_k , наружным диаметром D_n , внутренним D_o имеет ω витков с диаметром проволоки, равным d , то сечение меди катушки будет



$$q_m = \frac{\pi d^2}{4} \omega.$$

Сечение всей катушки

$$q_k = l_k \frac{(D_n - D_o)}{2}.$$

Коэффициент заполнения

$$f_k = \frac{q_m}{q_k} = \frac{\pi d^2}{2} \frac{\omega}{l_k (D_n - D_o)}. \quad (2-1)$$

Фиг. 2, 18. Сечение катушки.

При конструировании катушки стремятся к получению максимального значения f_k , которое зависит от следующих факторов: 1) изоляции катушки, 2) изоляции намотки, 3) формы сечения проволоки, 4) процесса намотки, 5) размеров проволоки.

Изоляция катушки. В каркасной катушке определенный объем ее занят изоляционными материалами каркаса, наружным бандажом; в бескаркасной — изоляцией катушки от магнитопровода, торцевой изоляцией, наружным бандажом.

Изоляция намотки. Определенный объем катушки занят изоляцией обмоточного провода — эмалевой, шелковой, хлопчатобумажной или другими. Кроме того, в ряде случаев между слоями проволоки прокладывают изоляционную бумагу, также занимающую определенный объем.

Форма проволоки. При намотке квадратной проволоки получается больший коэффициент заполнения, чем при намотке круг-

лой. В практике, однако, как правило, применяют круглую проволоку и только при сечениях проволоки, превышающих 10 мм^2 , ставят квадратную.

Процесс намотки. В процессе намотки в зависимости от совершенства намоточного станка или квалификации рабочего получается большая или меньшая плотность укладки проводов, что сильно сказывается на величине коэффициента заполнения. В зависимости от процесса намотки получают следующие виды намотки:

1) **рядовая**, когда витки в одном ряду расположены плотно друг к другу, а витки вышележащего ряда лежат точно над витками нижнего (фиг. 2, 19 б);

2) **шахматная**, когда витки вышележащего ряда расположены в промежутках между витками нижнего (фиг. 2, 19 а);

3) **дикая**, когда витки расположены рядами, но без точной укладки их.

Размеры проволоки. При малых сечениях проволоки величина ее изоляции относительно больше. Поэтому величина коэффициента

заполнения катушки тем меньше, чем меньше диаметр проволоки. Следует также учитывать и то обстоятельство, что при намотке катушек из тонкой проволоки нередки случаи обрыва ее. В связи с этим требуется спайка концов и изолировка их, что, в свою очередь, занимает дополнительно некоторое пространство в намотке. Поэтому в релестроении избегают применения проволоки диаметром меньше $0,1 \text{ мм}$. Но при исполнении катушек поляризованных и магнитоэлектрических реле приходится применять тонкую проволоку диаметром до $0,02 \text{ мм}$.

б) Коэффициент укладки. Наилучшие результаты в отношении достижения высокого значения величины коэффициента заполнения получают при шахматной намотке. В этом случае расстояние между серединами двух соседних слоев (фиг. 2, 19 а) равно

$$x = \sqrt{d_1^2 - \frac{d_1^2}{4}} = 0,86 d_1, \quad (2-2)$$

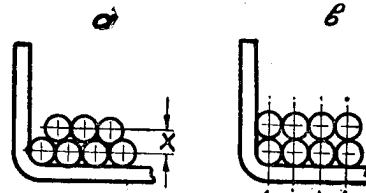
где d_1 — диаметр проволоки с изоляцией.

Число витков катушки, таким образом, при шахматной намотке получается

$$\omega_1 = \frac{h_n l_n}{0,86 d_1^2}, \quad (2-3)$$

в то время, как число витков при рядовой намотке составляет

$$\omega_2 = \frac{h_n l_n}{d_1^2}. \quad (2-4)$$



Фиг. 2, 19. Виды намоток:
а — шахматная; б — рядовая.

Отсюда

$$\frac{\omega_1}{\omega_2} = \frac{1}{0,86} = 1,16. \quad (2-5)$$

Таким образом, коэффициент заполнения при шахматной намотке должен быть на 16% выше, чем при рядовой. Однако в практике редко рассчитывают на выполнение шахматной намотки.

При рядовой намотке неизбежно попадание некоторых витков верхнего ряда в промежутки между витками нижнего, что увеличивает число витков катушки на 5—6%, но это увеличение аннулируется получающимися в процессе производства неизбежными неплотностями между витками одного ряда. Поэтому в практических расчетах учитывается рядовая намотка, причем полученное число витков умножают на коэффициент укладки k_y , значение которого колеблется в следующих пределах:

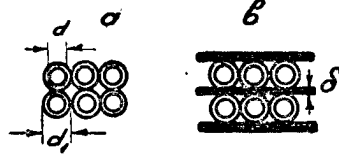
$k_y = 0,8—0,9$ для дикой намотки;

$k_y = 0,9—0,95$ для рядовой, выполненной на автоматах и полуавтоматах.

Меньшее значение k_y принимается для тонких проволок.

в) Коэффициент заполнения намотки — f_n . При определении его учитывают только пространство, занимаемое непосредственно обмоткой — площадка $h_n l_n$ (фиг. 2, 18). Следует принимать также во внимание значение коэффициента укладки k_y .

Значение f_n определяется из следующих соотношений: для намотки без прокладок (фиг. 2, 20, а)



Фиг. 2, 20. Сечение намотки: а — без прокладки; б — с прокладкой.

$$f_n = k_y \frac{\pi d^2}{4d_1^2}; \quad (2-6)$$

для намотки с прокладкой через слой (фиг. 2, 20, б)

$$f_n = k_y \frac{\pi d^2}{4d_1(d_1 + \delta)}, \quad (2-7)$$

где d — диаметр голой, а d_1 — диаметр изолированной проволоки; δ — толщина бумажной прокладки.

г) Коэффициент заполнения катушки. С учетом коэффициента заполнения намотки и коэффициента укладки коэффициент заполнения катушки (фиг. 2, 18) определится следующим образом:

$$f_k = f_n \frac{h_n l_n}{h_n l_n}, \quad (2-8)$$

и число витков катушки будет

$$\omega = f_k \frac{4h_n l_n}{\pi d^2}. \quad (2-9)$$

д) Практические данные. Обычно каждая магнитная система выполняется для разных пределов токов и напряжений рабочей

катушки. Так, например, реле с катушкой напряжения постоянного тока выпускаются для работы при напряжениях 24, 48, 110, 220 в.

При предварительных расчетах реле принимают во внимание какой-то средний коэффициент заполнения; а затем при окончательном расчете катушки для данного напряжения значение этого коэффициента уточняется.

Ввиду некоторой неопределенности значения коэффициента укладки k_u , зависящего от совершенства процесса намотки, для определения значений f_n и f_k в практике пользуются опытными данными. В таблицах 5 и 6 приведены опытные данные заводов «Красная заря» и ХЭМЗ.

Таблица 5

Опытные данные завода «Красная заря» для определения f_n проволоки с эмалевой изоляцией (намотка без прокладок)

Диаметр голой проволоки, мм	Коэффициент f_n	Диаметр голой проволоки, мм	Коэффициент f_n
0,05	0,25	0,18	0,66
0,10	0,48	0,20	0,685
0,11	0,49	0,23	0,73
0,12	0,51	0,25	0,76
0,14	0,56	0,27	0,786
0,15	0,58		

Таблица 6

Опытные данные ХЭМЗа для определения коэффициента заполнения катушки f_k для проволоки с эмалевой изоляцией (катушка каркасная)

Диаметр голой проволоки, мм	Намотка с ручной подачей				Намотка с автоматической подачей	
	Гильза цилиндрическая	Гильза прямоугольная	Гильза прямоугольная, прокладки через 1 слой	Гильза прямоугольная, прокладки через 2 слоя	Гильза цилиндрическая, прокладка через 1 слой	Гильза прямоугольная, прокладка через 1 слой
0,05	—	—	—	0,3	—	—
0,10	0,440	0,420	0,285	0,35	0,38	0,36
0,15	0,495	0,475	—	0,39	—	—
0,20	0,535	0,515	0,350	0,425	0,48	0,44
0,25	—	—	—	0,460	—	—
0,30	—	—	0,385	—	0,54	0,37
0,40	—	—	0,410	—	0,57	0,53

2-10. НАГРЕВ КАТУШЕК ПОСТОЯННОГО ТОКА

Нагрев катушки определяется потерей мощности при протекании тока I : $P = I^2 r$.

Предельная допустимая температура нагрева катушки ограничивается теплостойкостью изоляции. Последняя характеризуется той максимальной температурой, выше которой начинается разрушение материала или резкое ухудшение его свойств, как-то: резкое понижение электрической или механической прочности или прогрессивное старение материала. В катушках реле применяется изоляция классов *O* и *A*.

К классу *O* относятся органические изоляционные материалы: хлопок, шелк, бумага. К этому же классу относятся катушки с обмоткой из проволоки с эмалевой изоляцией, снабженные прокладками из непропитанной бумаги.

К классу *A* относятся органические материалы: хлопок, шерсть или бумага, пропитанные или погруженные в масло, а также эмалевая изоляция. По существующим нормам допускаются следующие предельные температуры: класс *O* — 90°, класс *A* — 105°.

Эти данные относятся к органическим изоляционным материалам, выдерживающим без доступа воздуха и влаги срок службы от 7 до 15 лет (в среднем считают 10 лет). При превышении указанных температур на 10—12° срок службы уменьшается вдвое.

В ряде аппаратов для катушек из проволоки с изоляцией класса *A* предельную температуру нагрева принимают равной 120°. Такая повышенная температура допускается для случаев, когда катушка подвергается компаундировке, при которой изоляция катушки и намотки пропитывается смолой, все промежутки между изоляцией также заполняются смолой и катушка становится монолитной. При таких условиях, если изоляция при повышенной температуре нагрева и становится хрупкой, катушка не выходит из строя, так как изоляция не разрушается механически. При применении в катушке изоляции из стеклянной пряжи и пропитке катушки специальными компаундами, например кремнеорганическими смолами, предельная допустимая температура нагрева повышается до 175° и выше.

При протекании тока I в течение достаточно большого промежутка времени в катушке устанавливается определенная температура, причем температура внутри катушки несколько выше, чем на ее поверхности. На фигуре 2, 21 приведена примерная кривая распределения температуры по сечению катушки. Максимальная температура получается приблизительно на расстоянии $\frac{1}{3}$ от начала катушки.

Существует несколько методов измерения температуры нагрева катушки. Наиболее точные данные получаются при определении температуры нагрева с помощью температурных детекторов (термопар), заложенных внутри катушки в разных местах

ее намотки. Этот метод применим только в лабораторных условиях. Из практических методов известны метод сопротивления и метод термометра.

Метод сопротивления основан на изменении удельного сопротивления меди с увеличением ее температуры. Это изменение происходит по линейному закону. Определяют сопротивление r_1 катушки в холодном состоянии при температуре окружающей среды ϑ_1 и r_2 — в горячем состоянии:

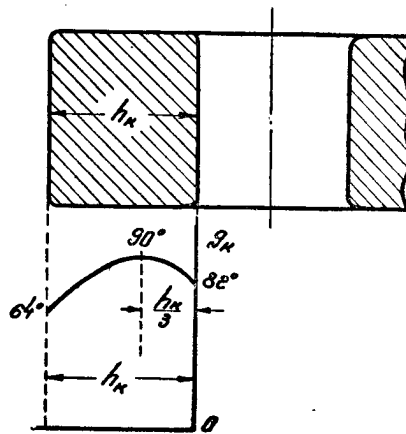
$$r_1 = r_0 (1 + \alpha \vartheta_1);$$

$$r_2 = r_0 (1 + \alpha \vartheta_2).$$

Здесь r_0 — сопротивление при температуре 20° .

По отношению сопротивлений r_1 и r_2 можно определить температуру ϑ_2 нагрева катушки. При методе сопротивления получают данные, относящиеся к средней температуре нагрева, несколько меньшей максимальной.

По методу термометра измеряют температуру на наружной поверхности катушки. Результаты измерения также получают заниженными по сравнению с максимальной температурой внутри катушки. На основании опытных данных установлена разница между измеренной и максимальной температурой катушки (см. табл. 7).



Фиг. 2, 21. Распределение температуры по сечению намотки.

Таблица 7

Допустимые значения нагрева и превышения температуры в $^\circ\text{C}$ (при $\vartheta_0 = 35^\circ$) катушек реле

Метод измерения	Класс О		Класс А		А компаундированная	
	нагрев	превышение температуры	нагрев	превышение температуры	нагрев	превышение температуры
	ϑ_k	θ	ϑ_k	θ	ϑ_k	θ
Термометром или термопарой . .	90	45	105	60	120	65
Сопротивления . .	90	55	105	70	120	85

При расчете катушки необходимо принимать во внимание еще значение температуры окружающей среды, учитывая наихудшие

условия. По нормам такая температура ϑ_0 принята равной 35° . Для работы аппаратов в специальных условиях, например на морских судах, в производствах с повышенной температурой окружающей среды, принимают $\vartheta_0 = 40^\circ$.

В таблице 7 приводятся данные допустимых значений нагрева катушки (ϑ_k) и превышения температуры θ с учетом $\vartheta_0 = 35^\circ$ для классов изоляции *O* и *A*.

При расчете катушек пользуются данными допустимых значений превышения температуры по методу сопротивления.

2-11. ТЕПЛОТДАЧА КАТУШЕК ПОСТОЯННОГО ТОКА

а) Общее уравнение. При установившемся нагреве все тепло, выделяемое током в катушке, переходит в окружающую среду. Переход тепла осуществляется различными путями: лучеиспусканием, конвекцией и теплопроводностью от внутренней поверхности катушки к стали магнитопровода. Законы передачи тепла для каждого из указанных видов различны. Величина теплоотдачи катушки определяется различными факторами.

Теплопередача с внутренней поверхности катушки. В системах постоянного тока магнитопровод не нагревается вихревыми токами и потерями на гистерезис. Поэтому в таких системах стремятся к уменьшению толщины изоляции между медью намотки и сталью, благодаря чему увеличивается коэффициент теплопроводности от внутренней поверхности катушки.

Наименее благоприятные условия получаются в этом отношении у каркасных катушек, в которых имеется относительно большой зазор с малой теплопроводностью, определяемый толщиной каркаса и наличием небольшого воздушного зазора между каркасом и сталью. Согласно опытным данным ХЭМЗа, наличие воздушного зазора в $0,25$ мм на сторону уменьшает теплопередачу катушки на 8% , а зазор в $0,5$ мм — на 11% .

Лучшие условия теплопередачи получаются в бескаркасных катушках, у которых толщина изоляции меньше, чем в каркасных. Еще более благоприятные условия теплопередачи имеют место в бескаркасных катушках, намотанных непосредственно на сердечнике магнитной системы. В таких катушках теплопередача повышается на 10% по сравнению с каркасной.

В системах переменного тока, где сердечник нагревается от вихревых токов и от потерь, определяемых гистерезисом, катушка, как правило, исполняется каркасной, чтобы устранить передачу тепла от сердечника к катушке. В индукционных реле часто применяют бескаркасные катушки, так как индукция в магнитных системах таких реле невелика и магнитопровод нагревается незначительно.

Влияние отделки на теплоотдачу. При пропитке катушек теплоотдача последних повышается на $3-4\%$, при компаундировке на $8-10\%$.

Влияние расположения катушки. На теплоотдачу катушки влияет еще ряд условий, определяемых расположением катушки в магнитной системе, наличием кожуха и другими причинами.

Точное определение количества тепла, отдаваемого катушкой, связано с весьма громоздкими расчетами. Поэтому в практических расчетах пользуются упрощенной формулой, которой учитывается только явление конвекции. Такое упрощение принято по следующим соображениям.

1. При относительно небольших температурах нагрева, допустимых для катушки (100—120°), количество тепла, отводимого лучеиспусканием, невелико, и этот процесс рассматривается совместно с конвекцией.

2. Точный учет количества тепла, отдаваемого теплопроводностью от внутренней поверхности катушки к магнитопроводу, весьма затруднителен, и этот процесс также рассматривается совместно с явлением конвекции, для чего в соответствующих соотношениях вводится величина внутренней поверхности катушки.

Таким образом, теплоотдачу лучеиспусканием, теплопроводностью и конвекцией объединяют и с достаточной точностью производят подсчет по уравнению

$$I^2 r = \mu S \theta, \quad (2-10)$$

где μ — коэффициент теплоотдачи, выраженный в ваттах на $см^2$ на градус;

S — поверхность охлаждения катушки (наружная и внутренняя) в $см^2$;

θ — превышение температуры катушки, представляющее собой разность между средней температурой намотки и температурой окружающей среды.

б) Поверхность охлаждения катушки. При определении поверхности охлаждения катушки учитывают, как правило, только внутреннюю и наружную ее поверхности. Теплоотдача с торцов катушки незначительна вследствие наличия у торцов относительно толстого слоя изоляции между медью намотки и окружающим воздухом. Кроме того, катушки располагаются обычно в магнитных системах так, что торцевые поверхности не участвуют в отдаче тепла конвекцией.

При исполнении катушек с металлическим каркасом следует учитывать теплоотдачу от меди намотки к металлу каркаса и с торцов катушки.

Имеются, однако, магнитные системы с развитыми по диаметру катушками. В этих случаях учитывают теплоотдачу торцевых поверхностей катушки.

Для круглых катушек

$$S = \pi (D_n + D_e) l_n \cdot 10^{-2} \text{ (см}^2\text{)}. \quad (2-11)$$

Так как средняя длина витка $l_{cp} = \frac{\pi(D_n + D_s)}{2}$ мм, можно написать

$$S = 2 l_{cp} l_k 10^{-2} \text{ (см}^2\text{)} \quad (2-12)$$

(l_{cp} , l_k , D_n и D_s принимаются в мм).

Это же соотношение (2—12) остается в силе и для прямоугольных катушек. При наличии металлических каркасов

$$S_k = S_n + S_s + S_m, \quad (2-13)$$

где S_n , S_s , S_m — соответственно поверхности катушки — наружная, внутренняя и торцевая.

в) Превышение температуры. Значение θ определяется как разность

$$\vartheta_k - \vartheta_n = \theta. \quad (2-14)$$

Величины допустимых превышений температуры приведены в таблице 7. При расчете катушки стремятся к получению значений θ , несколько меньших (на 5—10°), чем допустимые, что делается с учетом различных, возможных в производстве, отклонений в изготовлении катушки.

г) Коэффициент теплоотдачи, μ . Значение μ определяется на основании опытных данных. Так, установлено, что для матовой черной поверхности $\mu = 12 \cdot 10^{-4} \div 14 \cdot 10^{-4}$ (вт/см² °С).

При расчете катушек принимают несколько иные значения, в которых учитывается некоторая величина теплоотдачи от внутренней поверхности катушки. Значение μ находится в зависимости от превышения температуры, степени изоляции внутренней поверхности катушки от стали и от отделки катушки. Практически в отношении степени изоляции намотки от стали различают два исполнения:

1. Магнитные системы с хорошей теплопроводностью от меди катушки к стали. Сюда могут быть отнесены бескаркасные небандажированные катушки, бескаркасные с намоткой на сердечнике.

2. Магнитные системы с плохой теплопроводностью от меди к стали. Сюда могут быть отнесены каркасные, бескаркасные бандажированные, свободные катушки (без стали).

При компаундировке катушек получают лучшие условия теплопроводности от более нагретых слоев намотки к менее нагретым, поэтому значение μ повышается на 7—8%. В таблице 8 даны $\mu = f(\theta)$, полученные на основании ряда экспериментов с различными видами катушек.

В таблице 8 принято значение θ при температуре окружающей среды 35° ($\theta = \vartheta_k - 35$). Шкала А относится к катушкам с хорошей теплопроводностью от намотки к стали, шкала В — с плохой.

Пример. Определить допустимые потери (в вт) каркасной катушки из проволоки с эмалевой изоляцией.

Размеры катушки: $l_n = 120$ мм, $D_n = 60$ мм, $D_g = 20$ мм.
 Поверхность охлаждения 302 см².
 Принимаем: $\theta = 70^\circ$, по шкале В:

$$\mu = 10,99 \cdot 10^{-4}, P = I^2 r = 10,99 \cdot 10^{-4} \cdot 302 \cdot 70 = 23,1 \text{ вт}$$

Таблица 8
 Значения $\mu = f(\theta)$

θ°	Шкала А, вт/см ² °С	Шкала В, вт/см ² °С
40	$11,00 \cdot 10^{-4}$	$9,84 \cdot 10^{-4}$
45	$11,20 \cdot 10^{-4}$	$10,01 \cdot 10^{-4}$
50	$11,41 \cdot 10^{-4}$	$10,19 \cdot 10^{-4}$
55	$11,62 \cdot 10^{-4}$	$10,37 \cdot 10^{-4}$
60	$11,80 \cdot 10^{-4}$	$10,54 \cdot 10^{-4}$
65	$12,04 \cdot 10^{-4}$	$10,72 \cdot 10^{-4}$
70	$12,25 \cdot 10^{-4}$	$10,99 \cdot 10^{-4}$
75	$12,46 \cdot 10^{-4}$	$11,17 \cdot 10^{-4}$
80	$12,68 \cdot 10^{-4}$	$11,35 \cdot 10^{-4}$
85	$12,89 \cdot 10^{-4}$	$11,52 \cdot 10^{-4}$

д) **Определение требуемой поверхности охлаждения катушки.**
 Для быстрого расчета и проверки данных катушки на практике применяют несколько упрощенный метод, пользуясь опытными данными, в которых указывается потребное число квадратных сантиметров **наружной** поверхности катушки при условии, что превышение температуры ее не превосходит нормы.

В таблице 9 приведены опытные данные для каркасных катушек.

Таблица 9

Потребное число квадратных сантиметров наружной поверхности охлаждения каркасных катушек на 1 вт потерь

Отношение длины катушки к ее толщине	Изоляция проволоки		
	хлопчатобумажная, см ² /вт	хлопчатобумажная с пропиткой, см ² /вт	эмалевая см ² /вт
$l_n/h_n \geq 2$	14	12	10
$l_n/h_n \geq 1$	13	11	9
$l_n/h_n < 1$	12	10	

Пример. Определить допустимые потери (в *вт*) каркасной катушки из проволоки с эмалевой изоляцией с размерами, указанными в предыдущем примере:

$$l_w/h_k = \frac{12}{2} = 6; S_n = \pi \cdot 6 \cdot 12 = 226 \text{ см}^2.$$

Считая, что на 1 *вт* потерь требуется 10 *см*² наружной поверхности охлаждения, определяем, что для данной катушки допустимые полные потери равны

$$P_r = \frac{226}{10} = 22,6 \text{ вт}.$$

Результаты приближенного подсчета близки к данным предыдущего примера.

е) Влияние способа намотки на теплоотдачу катушки. М. Г. Кобленц провел исследования нагрева катушек контакторов постоянного тока при разных способах намотки катушки:

- 1) бескаркасной бандажированной;
- 2) бескаркасной, намотанной на стальной трубе (труба надевается на сердечник магнитопровода);
- 3) бескаркасной, намотанной непосредственно на сердечнике.

В результате проделанной работы установлены разные значения превышения температуры для указанных трех видов катушек. На основании результатов испытаний М. Г. Кобленц предлагает следующую формулу определения превышения температуры катушек:

$$P = \mu (S_n + \alpha S_v) \theta, \quad (2-15)$$

где S_n — наружная поверхность охлаждения;
 S_v — внутренняя поверхность охлаждения;
 α — коэффициент.

По данным исследования получены следующие результаты:
 при превышении температуры $\theta = 80^\circ$
 для бескаркасной бандажированной $\alpha = 0,9$;
 для катушки, намотанной на трубе, $\alpha = 1,7$;
 для катушки, намотанной на сердечнике, $\alpha = 2,4$.

Результаты исследования совпадают с данными таблицы 8 для случая бескаркасной бандажированной катушки. Так, по этой таблице по шкале А имеем при $\theta = 80^\circ$ и учете $S = S_n + S_v$ значение $\mu = 12,66 \cdot 10^{-4}$.

2-12. НАГРЕВ И ТЕПЛОТДАЧА КАТУШЕК ПЕРЕМЕННОГО ТОКА

Приведенные выше данные для определения превышения температуры катушек получены на основании анализа экспериментальных материалов исследования магнитных систем постоянного тока. При определении превышения температуры катушек переменного тока решение задачи несколько усложняется, так как в магнитной системе переменного тока сталь магнитопровода дополнительно подогревается потерями от вихревых токов и гистерезиса. Превышение температуры стали магнитопровода несколько ниже

превышения температуры катушки, поэтому для упрощения задачи считают, что магнитопровод не подогревает катушку, а с внутренней поверхности катушки нет отдачи тепла магнитопроводу.

Для расчета катушек переменного тока практически пользуются основным уравнением нагрева (2-10): $P = \mu S \theta$.

Принимают P как сумму всех потерь в магнитной системе: в омическом сопротивлении, на гистерезис и вихревые токи, а также в короткозамкнутых витках (экранах), если эти витки расположены непосредственно у катушки: $P_{\Sigma} = P_r + P_s + P_o + P_a$. Коэффициент теплоотдачи μ определяют по шкале B таблицы 8 (для катушек с плохой теплоотдачей от внутренней поверхности).

2-13. СОПРОТИВЛЕНИЕ НАМОТКИ

Общее выражение сопротивления

$$r = \rho \frac{l_{cp}}{q} \omega \cdot 10^{-3} \text{ ом}, \quad (2-16)$$

где l_{cp} — средняя длина витка в мм;
 q — сечение голой проволоки в мм².

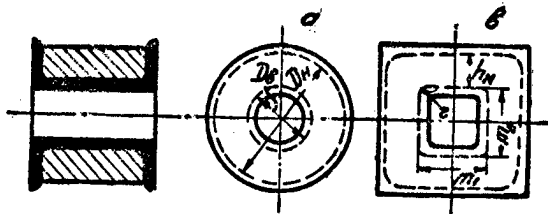
Для круглой катушки (фиг. 2, 22 а)

$$l_{cp} = \frac{\pi(D_n + D_o)}{2}, \quad (2-17)$$

где D_n и D_o приняты в мм.

Для прямоугольной катушки (фиг. 2, 22 б)

$$l_{cp} = 2(m_1 + m_2) + \pi(2r + h_n). \quad (2-18)$$



Фиг. 2, 22. Размеры намотки.

При расчете катушки следует принимать во внимание значение ρ при повышенной температуре. Так, если катушка исполняется с изоляцией класса А, следует учитывать ρ_{105} .

Температурный коэффициент меди в пределах нагрева от 10 до 100° принимается равным 0,0393. Для проводниковой меди пользуются значениями ρ , приведенными в таблице 10.

Таблица 10

Значения ρ			
Рабочая температура, °С	Сопротивление (удельное) ом. мм ² /м	Рабочая температура, °С	Сопротивление (удельное) ом. мм ² /м
20	$\rho_{20} = 0,01754$	90	$\rho_{90} = 0,02236$
35	$\rho_{35} = 0,01857$	105	$\rho_{105} = 0,02339$
40	$\rho_{40} = 0,01991$	120	$\rho_{120} = 0,02443$

Измеренное действительное сопротивление катушки не всегда совпадает с расчетным. Определяется это получающимися в производстве отклонениями в размерах проволоки (табл. 11), а также в числе витков.

Таблица 11

Допускаемые по ГОСТ отклонения в размерах проволоки

Диаметр голой меди, мм	От 0,05 до 0,08	От 0,09 до 0,25	От 0,25 до 0,69	От 0,74 до 1	От 1,15 до 1,56
Допускаемые отклонения, мм	$\pm 0,003$	$\pm 0,005$	$\pm 0,01$	$\pm 0,015$	$\pm 0,02$

Действительное сопротивление катушки получается равным

$$r_{\partial} = 4\rho \frac{l_{cp} (\omega \pm \Delta\omega)}{\pi (d \pm \Delta d)^2} 10^{-3} \text{ ом}, \quad (2-19)$$

где $\Delta\omega$ — величина отклонения в числе витков.

При определении большей величины сопротивления учитывают $+\Delta\omega$ и $-\Delta d$, при определении меньшей величины: $-\Delta\omega$ и $+\Delta d$.

Сопротивление в функции числа витков и габаритов. Подставляя в (2—16) значение q , определяемое из (2—9), получают выражение для сопротивления в функции числа витков и геометрических размеров катушки:

$$r = \rho \frac{l_{cp} \omega^2}{f_{\kappa} l_{\kappa} h_{\kappa}} 10^{-3} = k_r \omega^2, \quad (2-20)$$

здесь $k_r = \rho \frac{l_{cp}}{f_{\kappa} l_{\kappa} h_{\kappa}} 10^{-3}$ — постоянный коэффициент, зависящий

от материала проволоки и геометрических размеров катушки.

Выражение (2—20) применяется в том случае, когда необходимо определить сопротивление катушки при изменении числа витков. Следует учитывать и изменение коэффициента заполнения.

2-14. РАСЧЕТ КАТУШКИ ПРИ РАЗНЫХ РЕЖИМАХ РАБОТЫ

При расчете катушки учитывается режим работы аппарата. Различают следующие режимы работы:

ПРР — продолжительный (длительный) режим работы, когда катушка включена на длительный промежуток времени.

КРР — кратковременный режим работы, когда катушка включена на небольшой период, порядка нескольких секунд, и время паузы достаточно велико, чтобы катушка охладилась в течение этой паузы до температуры окружающей среды.

ПКРР — повторно-кратковременный (прерывистый) режим работы, при котором рабочий период t_e (период включения) регулярно (или нерегулярно) чередуется с периодами отключения. Общая продолжительность цикла $t_u = t_e + t_n$ обычно не превышает 10 минут.

РКЗ — режим короткого замыкания. Этот режим рассматривается только при расчете токовых катушек.

При нормальном режиме установки, где включен аппарат, в катушке протекает номинальный ток I_n . При аварийном режиме в катушке протекает ток короткого замыкания, который во много раз (50 ÷ 100) больше номинального. В таких случаях расчет катушки необходимо проводить с проверкой на термическую устойчивость.

ПРР. Исходное уравнение $P = \mu S \theta$. Значение P : для катушек постоянного тока $P = I^2 r$; для катушек переменного тока $P = I^2 r + P_n + P_s$, где P_n — мощность потерь на гистерезис и вихревые токи, P_s — мощность потерь в экране или роторе индукционной системы.

Для системы постоянного тока

$$I^2 r = \mu S \theta. \quad (2-10)$$

$$r = \rho \frac{l_{cp} \omega^2}{f_k l_k h_k} 10^{-3}. \quad (2-20)$$

Из (2—10) и (2—20) устанавливаем зависимость превышения температуры от н. с. катушки и ее геометрических размеров.

$$\theta = \rho \frac{l_{cp}}{10 f_k l_k h_k \mu S} (I \omega)^2, \quad (2-21)$$

где l_{cp} , l_k , h_k — в мм; S — см².

Для круглой катушки можно установить связь между l_k и S :

$$l_{cp} = \frac{\pi (D_n + D_s)}{2} (a); S = \pi (D_n + D_s) l_k \cdot 10^{-2} (см^2) (b).$$

Из (2—21) подстановкой a и b получаем

$$\theta = \frac{\rho}{20 \mu f_k l_k^2 h_k} (I \omega)^2 \quad (2-22)$$

КРР. При этом режиме работы катушка может выдержать значительно большие нагрузки, чем при ПРР.

При определении нагрева катушки, включенной в течение небольшого промежутка времени — порядка нескольких секунд — считают, что все тепло, выделенное в катушке, расходуется на нагрев меди, т. е. пренебрегают теплоотдачей во внешнюю среду. Уравнение нагрева катушки в этом случае будет

$$I^2 r t = c G \theta, \quad (2-23)$$

здесь t — в секундах;

c — удельная теплоемкость меди в $\text{дж}/\text{г} \text{ } ^\circ\text{C}$;

G — вес меди в граммах.

Преобразуем выражение (2-23) с тем, чтобы исключить из него величины, не влияющие на нагрев меди:

$$G = \gamma l_{\text{ср}} q w \cdot 10^{-3}, \quad (2-24)$$

где γ — удельный вес;

$l_{\text{ср}}$ — в мм;

q — в мм^2 .

Подставив (2-16) и (2-24) в (2-23), получаем:

$$I^2 r \frac{l_{\text{ср}}}{q} w t \cdot 10^{-3} = c \gamma l_{\text{ср}} q w \theta \cdot 10^{-3}.$$

Произведя соответствующие сокращения, получим

$$\frac{I^2}{q^2} = \frac{c \gamma \theta}{r t},$$

откуда плотность тока

$$I = \frac{I}{q} = \sqrt{\frac{c \gamma \theta}{r t}}. \quad (2-25)$$

При повышении температуры меди от θ_0 до θ_n сопротивление постепенно меняет свое значение. В расчетах принимают во внимание среднее значение ρ .

Таблица 12

Значение допустимой плотности тока для обмотки из меди
(при односекундной нагрузке)

Класс изоляции	О	А	Компаундированные катушки
$\theta_{np}, \text{ } ^\circ\text{C}$	90	105	120
$\theta = \theta_{np} - 35$	55	70	85
$\gamma - \text{г}/\text{мм}^3$	100	108	117

При этом приняты: $c = 0,39 \frac{\text{дж}}{\text{г} \text{ } ^\circ\text{C}}$; $\gamma_{\text{меди}} = 8,9 \frac{\text{г}}{\text{см}^3}$.

При длительности нагрузки t сек. надо значение I разделить на \sqrt{t} .

Примеры: 1. Определить сечение медной проволоки с эмалевой изоляцией для некомпандинированной катушки. $I = 0,3$ а; $t = 3$ сек; $\theta = 70^\circ$.

$$j_s = \frac{d_1}{\sqrt{3}} = \frac{108}{\sqrt{3}} = 62,2 \text{ а/мм}^2;$$

$$q = \frac{\pi d^2}{4} = \frac{I}{j} = \frac{0,3}{62,2} \text{ мм}^2,$$

$d = 0,079$, принимаем $d = 0,08$ мм.

2. Определить допустимую длительность включения t некомпандинированной катушки из проволоки с эмалевой изоляцией при токе $I = 0,5$ а; $d = 0,2$ мм; $\theta = 70^\circ$.

$$q = 0,0314 \text{ мм}^2; j = \frac{0,5}{0,0314} = 16 \text{ а/мм}^2;$$

$$t = \left(\frac{108}{16}\right)^2 = 45,5 \text{ сек.}$$

Приведенное в данном примере решение применяется для определения допустимой длительности включения катушки переменного тока при разомкнутом якоре, когда ток катушки в несколько раз больше тока при замкнутом якоре электромагнитной системы.

ПКРР. Повторно-кратковременный режим работы характеризуется величиной относительной продолжительности включения (ОПВ), т. е. отношением длительности рабочего цикла t_a к полному периоду цикла t_u : $\zeta = \frac{t_a}{t_u}$.

Для катушек, работающих при ПКРР, допускается по сравнению с продолжительным режимом работы (ПРР) значительное увеличение нагрузки. Обозначим: I — ток, протекающий через катушку; $\theta_{нк}$ и θ_n — установившиеся превышения температуры соответственно при ПКРР и ПРР с нагрузкой $P = I^2 r$.

$$\theta_{нк} = \theta_n \frac{1 - e^{-\frac{t_a}{T}}}{1 - e^{-\frac{t_u}{T}}} = \zeta' \theta_n, \quad (2-26)$$

где T — постоянная времени нагрева катушки.

Если t_u значительно меньше постоянной времени нагрева катушки, а следовательно, и $t_a \ll T$, то $\zeta' = \frac{t_a}{t_u} = \zeta$.

Следовательно,

$$\theta_{нк} = \zeta \theta_n. \quad (2-27)$$

Подставляя значение θ_n из (2-22), получаем

$$\theta_{нк} = \zeta \frac{P}{20 \mu f_n h_n l_n^2} (I \omega)^2 \quad (2-28)$$

Возникает вопрос о том, каково же максимальное время, при котором превышение температуры катушки достигает допустимого

значения. Для решения этой задачи воспользуемся общим уравнением нагрева

$$\theta = \theta_{\max} \left(1 - e^{-\frac{t_\theta}{T}} \right), \quad (2-29)$$

где T — постоянная времени нагрева катушки;

θ — допустимое значение превышения температуры;

θ_{\max} — максимальное значение превышения температуры, которое имело бы место в катушке при продолжительном ее включении.

Значение θ_{\max} может быть определено из уравнения нагрузки (2-10), соответствующего продолжительному включению

$$\theta_{\max} = \frac{P}{\mu S}. \quad (2-30)$$

Подставив (2-30) в (2-29), имеем

$$\theta = \frac{P}{\mu S} \left(1 - e^{-\frac{t_\theta}{T}} \right); \quad (2-31)$$

$$t_\theta = T \ln \frac{P/\mu S}{P/\mu S - \theta}. \quad (2-32)$$

Далее переходим к определению T .

Как известно, значение постоянной времени нагрева однородного тела равно отношению теплоемкости тела к количеству тепла, отдаваемого во внешнее пространство:

$$T = \frac{cG}{\mu S_\kappa}. \quad (2-33)$$

При нагреве катушки теплоемкость и теплоотдача непостоянны, так как тепло катушки в процессе ее нагрева частично передается стали магнитопровода. Поэтому при учете влияния отдачи тепла стали получается более сложная зависимость для определения T . Для приближенных вычислений можно пользоваться упрощенным выражением постоянной времени нагрева катушки:

$$T = \frac{c_1 G_1 + \beta c_2 G_2}{\mu S_\kappa}, \quad (2-34)$$

где $c_1 G_1$ — теплоемкость меди и изоляции катушки;

μ — коэффициент теплоотдачи;

S_κ — сумма внутренней и наружной поверхностей охлаждения катушки;

c_1, G_2 — теплоемкость стали магнитной системы;

β — коэффициент, учитывающий степень передачи тепла от катушки к стали.

Значение коэффициента β определяется условиями теплопередачи от катушки к стали и временем нагрева. По условиям теплопередачи можно рассматривать два крайних случая:

1) хорошая теплопередача от намотки к стали (бескаркасные небандажированные катушки, бескаркасные с намоткой на сердечнике);

2) плохая теплопередача от намотки к стали (каркасные, бескаркасные бандажированные катушки).

Влияние фактора времени на изменение β сильно сказывается в тех системах, где поверхность охлаждения катушки относительно мала. В этом отношении системы характеризуются коэффициентом

$$\gamma = \frac{c_2 G_2}{S_{\kappa}} \quad (2-35)$$

Так, по опытным данным, при $\gamma \leq 1,25$ значение β можно считать постоянным за все время нагрева катушки. При $\gamma \geq 5$ значение β следует считать в течение первого периода нагрева, равного $\frac{T}{3}$, уменьшенным вдвое, а затем после $t_e > \frac{T}{3}$ принимать полное значение β . В случае $1,25 < \gamma < 5$ изменение β принимают пропорциональным числу $5 - 1,25 = 3,75$. В таблице 13 указаны значения β , определенные на основании приведенных выше опытных данных.

Значения β

Таблица 13

Период нагрева	Хорошая теплопередача			Плохая теплопередача		
	$\gamma \leq 1,25$	$1,25 < \gamma < 5$	$\gamma \geq 5$	$\gamma \leq 1,25$	$1,25 < \gamma < 5$	$\gamma \geq 5$
$t_e < \frac{T}{3}$	0,55	$0,55 \left[0,5 + 0,5 \frac{5-\gamma}{3,75} \right]$	0,275	0,45	$0,45 \left[0,5 + 0,5 \frac{5-\gamma}{3,75} \right]$	0,225
$t_e > \frac{T}{3}$	0,55	0,55	0,55	0,45	0,45	0,45

Пример. Определить допустимую длительность включения t_e бескаркасной катушки с намоткой на сердечнике, рассчитанной для режима повторно-кратковременного включения $\zeta = 0,1$; изоляция проволоки — эмаль; потребляемая мощность — 130 *вт* (горячей катушки); вес: меди — 0,4 *кг*, изоляции — 0,07 *кг*; стали — 1,7 *кг*.

Поверхность охлаждения катушки — 200 *см*².

Принимаем: $\theta = 70^\circ$ по шкале А таблицы 8 $\mu = 12,19 \cdot 10^{-4}$,

$$\frac{P}{\mu S_{\kappa}} = \frac{130}{12,19 \cdot 10^{-4} \cdot 200} = 532^\circ;$$

$c_1 G_1$ — теплоемкость меди и изоляции:

$$(0,4 \cdot 390) + (0,07 \cdot 1540) = 264 \frac{\text{дж}}{^\circ\text{C}};$$

$c_2 G_2$ — теплоемкость стали равна $1,7 \cdot 495 = 840 \frac{\text{дж}}{^\circ\text{C}}$.

$$\gamma = \frac{c_2 G_2}{S_{\kappa}} = \frac{840}{200} = 4,2,$$

$$\beta = 0,55 \left[0,5 + 0,5 \frac{5-\gamma}{3,75} \right] = 0,334.$$

Постоянные времени катушки:

$$t < \frac{T}{3};$$

$$T_1 = \frac{264 + 0,334 \cdot 840}{12,19 \cdot 10^{-4} \cdot 200} = 2250 \text{ сек.}$$

$$t > \frac{T}{3};$$

$$T_2 = \frac{264 + 0,55 \cdot 840}{12,19 \cdot 10^{-4} \cdot 200} = 2850 \text{ сек.}$$

Допустимое значение t_0 согласно (2-32):

$$t_0 = T_1 \ln \frac{\frac{P}{\mu S}}{\frac{P}{\mu S} - \theta} = 2250 \ln \frac{532}{462} = 320 \text{ сек} = 5,34 \text{ мин.}$$

Полный период цикла t_u должен быть не менее

$$\frac{t_0}{\zeta} = \frac{5,34}{0,1} = 53,4 \text{ мин.}$$

Р К 3. При токе короткого замыкания нагрев катушки становится значительно больше допустимого при продолжительном режиме. В данном случае при длительности процесса короткого замыкания, не превышающей 1 секунду, допускается повышенная температура нагрева меди обмотки (до 200°). Свыше этой температуры начинается обугливание изоляции. Понятие термической устойчивости катушки, выражаемое определенным коэффициентом, характеризует допустимую в течение 1 секунды кратность тока в катушке при условии, что при этом предельный нагрев меди обмотки не превысит 200°. Термическая устойчивость, равная 50, обозначает, что катушка с номинальным током $I_n = 5a$ может выдержать в течение 1 секунды ток короткого замыкания, равный 250 a. Диаметр проволоки такой катушки определяют, исходя из условия, что все тепло, выделенное в обмотке, расходуется на нагрев меди до 200° (см. выражение (2-25)).

В данном случае

$$\rho_{ср} = 0,024; \quad \theta = 165°;$$

$$j = 155 \text{ а/мм}^2.$$

Пример. Определить диаметр медной проволоки токовой катушки реле: $I_n = 5 \text{ а}$; термическая устойчивость равна 50;

$$q = \frac{I_{кз}}{i} = \frac{250}{155} = 1,61 \text{ мм}^2; \quad d = 1,43 \text{ мм.}$$

Принимаем ближайший больший диаметр проволоки по таблице 2.

$$d = 1,45 \text{ мм.}$$

2-15. ОСНОВНЫЕ СООТНОШЕНИЯ ДЛЯ РАСЧЕТА

При расчете катушек стремятся к тому, чтобы выдержать два основных условия:

- 1) катушка должна обеспечить создание требуемых ампервитков магнитной системы при заданном напряжении или силе тока;
- 2) температура нагрева намотки катушки не должна превосходить предельно допустимой величины.

При расчете катушки требуется определить ряд данных: длину катушки, толщину намотки, количество витков и сечение проволоки. Для получения этих данных пользуются основными соотношениями, выведенными ранее (2-20), (2-21), (2-25) и (2-28), а также нижеследующими уравнениями.

Уравнение числа витков

$$\omega = f_{\kappa} \frac{l_{\kappa} \cdot h_{\kappa}}{q}. \quad (2-36)$$

Уравнения напряжения:

а) для постоянного тока

$$U = Ir = \rho \frac{l_{cp}}{q} I \omega \cdot 10^{-8} \quad (2-37)$$

б) для переменного тока

$$U = \sqrt{I^2 r^2 + E^2}, \quad (2-38)$$

где

$$E = \omega LI = \omega \omega^2 GI. \quad (2-39)$$

G — магнитная проводимость системы.

Уравнение потока (для систем переменного тока)

$$\Phi = \frac{E}{4,44 f \omega} \text{ вебер.} \quad (2-40)$$

На основе приведенных выше соотношений можно рассчитать катушку определенных заданных размеров. В том случае, если неизвестны габариты катушки, необходимы дополнительные данные, которые получаются при расчете всей магнитной системы в целом.

2-16. ДОПУСКИ И ОТКЛОНЕНИЯ

Точное соблюдение расчетных данных при намотке катушки не всегда удается. Вследствие различных отклонений в размерах проволоки и изоляции и неидентичности самого процесса намотки данные выпускаемых катушек, как правило, отличаются от расчетных. В производстве устанавливаются определенные допуски.

Допуск по сопротивлению. Допускаемое отклонение по величине сопротивления катушки находится в пределах от ± 3 до $\pm 5\%$. Точное соблюдение величины омического сопротивления важно для катушек постоянного тока, так как значение r характеризует потребляемую мощность катушки и, следовательно, ее

тепловой режим. В случае отклонения значения r от указанных допусков катушка доматывается или разматывается.

Допуск по числу витков. Точное соблюдение числа витков важно для катушек переменного тока, так как при определенном сечении стали магнитной системы число витков характеризует потребляемую мощность и число av системы в конечном положении якоря. Значения допускаемых отклонений по числу витков, принятые в практике, приведены в таблице 14.

Таблица 14

Значения допускаемых отклонений по числу витков катушек переменного тока.

Число витков катушки	Допускаемое отклонение
$w > 500$	$\pm 2\%$
$100 < w < 500$	$\pm 1\%$
$w < 100$	0

треннему диаметру колеблются в пределах $\pm (0,5 + 1,0)$ мм, по длине катушки в пределах $\pm (1 - 2)$ мм.

2-17. ИСПЫТАНИЕ И ПРОВЕРКА КАТУШЕК

Выпускаемые из производства катушки подвергаются поверочным испытаниям. Некоторые испытания проводятся только в специальных случаях, например, при типовых или приемо-сдаточных испытаниях, другие проводятся над каждой выпускаемой катушкой.

Тепловой режим катушек. Проверка нагрева катушек проводится в полностью собранном реле. Применяются два метода испытаний: 1) метод термометра или термопары, которые применяются для определения температуры катушки к наружной ее поверхности; 2) метод сопротивления, когда величина превышения температуры катушки определяется по величине изменения сопротивления.

Замер сопротивления катушки производится разными приборами: простым мостом — для сопротивлений, больше 1 ома, двойным мостом (если $r < 1$ ома), потенциометром, логометром или посредством амперметра и вольтметра. При измерении превышения температуры обмоток из медной проволоки методом сопротивления превышение это определяется по формуле

$$\theta = \vartheta_k - \vartheta_{oz} = \frac{r_z - r_x}{r_x} (235 + \vartheta_{ox} - \vartheta_{oz}), \quad (2-41)$$

где ϑ_x — температура нагретой обмотки;
 ϑ_{oz} — температура окружающей среды при нагретой катушке;
 ϑ_{ox} — температура окружающей среды при холодной катушке и начальная температура катушки;
 r_x — сопротивление обмотки при температуре ϑ_{ox} ;
 r_z — сопротивление обмотки при температуре ϑ_k .

Если испытание производится при температуре окружающей среды, отличной от $+35^{\circ}$, то величина превышения температуры, полученная по формуле (2-41), должна быть приведена к $+35^{\circ}$ путем умножения на коэффициент k .

Для катушек тока

$$k = \frac{270}{235 + \vartheta_{os}} \quad (2-42)$$

для катушек напряжения постоянного тока

$$k = \frac{235 + 2\vartheta + \vartheta_{os}}{270 + 2\vartheta} \quad (2-43)$$

Для катушек напряжения переменного тока применяется та же формула (2-42), что и для катушек тока. Объясняется это тем, что изменение сопротивления в процессе нагрева катушки напряжения переменного тока практически не влияет на величину ее тока. При испытании катушки переменного тока измерение сопротивления производится постоянным током, для чего в момент замера катушка переключается на прибор для измерения сопротивлений. Испытание токовых катушек производится при номинальном токе, а катушек напряжения — при номинальном напряжении.

Диэлектрическая прочность. В цепи магнитной системы в процессе ее эксплуатации возможны значительные перенапряжения, могущие привести к пробое изоляции катушки. Поэтому обязательна проверка диэлектрической прочности каждой катушки перед выпуском ее из производства. Проверка должна осуществляться над катушкой, встроенной в магнитную систему. По нормам испытательное напряжение значительно превосходит рабочее напряжение системы. Длительность испытания — 1 минута. Для ускорения процесса испытания его производят в течение 1 секунды, но при этом испытательное напряжение повышается.

В таблице 15 приведены данные напряжений для проверки на диэлектрическую прочность.

Таблица 15

Значение испытательных напряжений

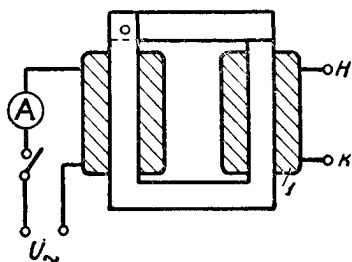
Рабочее напряжение, в	Испытательное напряжение, в	
	Одноминутное	Односекундное
До 24	250	2500
24—100	500	
101—500	2000	
501—600	$2U + 1000$	
Свыше 600	$2,25U + 2000$	

Проверка наличия короткозамкнутых витков. В процессе изготовления катушки возможны случаи оголения изоляции между двумя соседними витками, что приводит к образованию коротко-

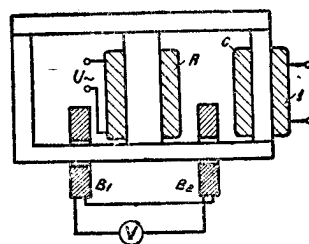
замкнутых витков. Применение таких катушек недопустимо в системах переменного тока, так как в короткозамкнутом витке индуцируется большой ток, виток сильно нагревается, портится изоляция катушки и последняя очень скоро выходит из строя.

В системах постоянного тока возможно применение катушек, имеющих короткозамкнутые витки, если только наличие этих витков не ведет к изменению омического сопротивления катушки. Однако для быстродействующих реле постоянного тока наличие короткозамкнутых витков в катушке вследствие их успокоительного действия также нежелательно.

Простейшее устройство для проверки наличия короткозамкнутых витков показано на фигуре 2, 23. На одном сердечнике магнитной системы насажена катушка возбуждения, включенная на



Фиг. 2, 23. Устройство для выявления к. з. витков:
I — испытуемая катушка.



Фиг. 2, 24. Устройство для выявления к. з. витков:
I — испытуемая катушка.

напряжение переменного тока, на другом — испытуемая катушка, концы которой остаются разомкнутыми. До установки испытуемой катушки проверяется ток катушки возбуждения. При насадке испытуемой катушки, не имеющей короткозамкнутых витков, значение этого тока не изменяется; при наличии же короткозамкнутых витков в испытуемой катушке величина тока I значительно повышается.

Схема другого устройства приведена на фигуре 2, 24. На среднем сердечнике Ш-образной системы насажена катушка возбуждения A , приключаемая к сети переменного тока, на яре системы насажены две балансные катушки B_1 и B_2 , включенные на вольтметр. Балансные катушки перед испытанием устанавливают на яре так, чтобы в них индуцировались одинаковые э. д. с.; при этом показания вольтметра равны нулю.

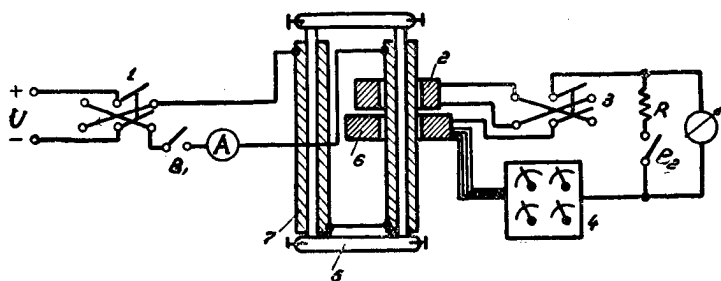
При насадке испытуемой катушки C , в случае наличия в ней короткозамкнутых витков, величины потоков в левой и правой частях магнитной системы уже не будут одинаковыми, и разность э. д. с. катушек B_1 и B_2 становится больше нуля, что показывает вольтметр.

При наличии высокочувствительного вольтметра в описанном устройстве можно выявить повреждение изоляции витков с отно-

сительно большим переходным сопротивлением. При наладке устройства необходимо следить, чтобы съемный якорь всегда устанавливался в одном и том же положении с минимальным зазором. В противном случае возможен небаланс между правой и левой частью и при отсутствии короткозамкнутых витков в испытуемой катушке.

В литературе приводится описание более совершенных современных устройств.

Проверка отсутствия обрывов в катушке производится на устройстве, описанном выше (фиг. 2, 23). Испытуемая катушка замыкается накоротко или через активное сопротивление. При наличии обрыва значение тока I катушки возбуждения остается



Фиг. 2, 25. Устройство для проверки числа витков:

1 — переключатель; 2 — испытуемая катушка; 3 — переключатель; 4 — декадный переключатель; 5 — магнитная система; 6 — стандартная катушка; 7 — катушка возбуждения.

тем же, каким оно было при отсутствии испытуемой катушки, при отсутствии обрыва значение I резко повышается.

Проверка числа витков катушки. Как правило, проверку числа витков катушки производят только периодически, выборочно. Устройства, с помощью которых осуществляется определение числа витков катушки, могут быть весьма точными. Так, имеются устройства с точностью замера до 1 витка на 100 000 витков катушки. Для практических целей выполняют устройства с точностью замера до 1 витка на 5000 витков катушки.

Действие устройства основано на сравнении э. д. с., индуцированных в искомой катушке с неизвестным числом витков и в стандартной — с известным числом витков. Если выполнять устройство так, что число потокосцеплений, созданных какой-то катушкой возбуждения, будет одинаковым для искомой и стандартной катушек, то индуцируемые в них э. д. с. не зависят от их габаритов, а только от числа витков.

Стандартная катушка выполняется с большим числом отпаек, подключаемых к декадному переключателю. В процессе проверки подбирают декадным переключателем такое число витков стандартной катушки, чтобы индуцируемая в ней э. д. с. была равна индуцируемой э. д. с. в искомой. При этом показания чувстви-

тельного гальванометра, включенного в цепь обеих катушек, будут равны нулю. Схема устройства показана на фигуре 2, 25.

Катушка возбуждения представляет собой распределенную обмотку, расположенную на обоих сердечниках магнитной системы. Катушка может быть включена на источник переменного или постоянного тока. В устройствах постоянного тока получаются более точные результаты. Стандартная и искомая катушки насаживаются на одном сердечнике одна под другой. Правильность соединения цепи обеих катушек проверяется предварительно посредством переключателя. Если катушки включены так, что их э. д. с. суммируются, получается большой бросок тока в цепи гальванометра.

Г Л А В А 3

РАСЧЕТ МАГНИТНОЙ ЦЕПИ СО СТАЛЬЮ

Задачей расчета магнитной цепи является вычисление магнитных потоков в разных частях магнитной системы, обусловленных наличием намагничивающей силы F , создаваемой током катушки. При проектировании системы приходится решать обратную задачу: определить н. с. катушки, требуемую для создания в системе магнитных потоков заданной величины.

Для расчета магнитных цепей применяются аналитические и графо-аналитические способы.

При проведении расчетов пользуются основными законами для магнитных цепей, которые совпадают по своей форме с основными законами электрических цепей.

3-1. ОСНОВНЫЕ ЗАКОНЫ МАГНИТНЫХ ЦЕПЕЙ

Магнитная цепь состоит из магнитопровода с насаженной на нем катушкой. Как правило, в магнитопроводе имеется несколько воздушных зазоров: $\delta_1, \delta_2 \dots$. При наличии магнитного потока в цепи получается падение магнитного потенциала в каждом элементе магнитопровода и в воздушных зазорах.

Для стали

$$Hl = \frac{\Phi}{\mu \sigma_c} l = F_c ; \quad (3-1)$$

для воздуха

$$H\delta = \frac{\Phi}{\mu_0 \sigma_a} \delta = F_a, \quad (3-2)$$

где μ — магнитная проницаемость стали;

μ_0 — магнитная проницаемость пустоты.

Первый закон Кирхгофа для магнитной цепи формулируется так же, как и для электрической: сумма магнитных потоков, сходящихся в каком-либо узле, равна нулю.

Второй закон Кирхгофа для магнитной цепи — линейный интеграл напряженности магнитного поля или сумма падений магнитных потенциалов для всякой замкнутой магнитной цепи равна сумме н. с. F тех витков или обмоток, через которые проходит рассматриваемая линия.

$$\oint H dl = \sum H_i l_i = H_1 l_1 + H_2 l_2 \dots + H_3 \delta_1 + H_4 \delta_2 = F. \quad (3-3)$$

Закон Ома для магнитных цепей — общий поток замкнутой магнитной цепи Φ равен намагничивающей силе катушки F , помноженной на магнитную проводимость всей цепи G . Подставляя (3-1) и (3-2) в (3-3), получаем

$$\Phi = \frac{F}{\frac{l_1}{\mu_1 g_1} + \frac{l_2}{\mu_2 g_2} + \dots + \frac{\delta_1}{\mu_0 g_3} + \frac{\delta_2}{\mu_0 g_4}}. \quad (3-4)$$

Обозначая

$$\frac{l_1}{\mu_1 g_1} + \frac{l_2}{\mu_2 g_2} = R_c; \quad \frac{\delta_1}{\mu_0 g_3} + \frac{\delta_2}{\mu_0 g_4} = R_e,$$

имеем из (3-4)

$$\Phi = \frac{F}{R_c + R_e} = \frac{F}{R} = FG, \quad (3-5)$$

где R , R_c и R_e — соответственно магнитные сопротивления всей магнитной цепи, стали и воздушных путей потока.

При проведении расчетов приняты

$$\Phi \text{ — в веберах; } G \text{ — } \frac{\text{об}}{\text{ас}}; \quad \mu_0 = 1,25 \cdot 10^{-8} \frac{\text{об}}{\text{а} \cdot \text{см}}.$$

3-2. СХЕМА ЗАМЕЩЕНИЯ МАГНИТНОЙ ЦЕПИ

При расчете сложных разветвленных магнитопроводов для удобства определения требуемых соотношений составляют схему замещения магнитной цепи, причем изображение схемы аналогично изображению электрической схемы.

В схеме замещения принимают следующие обозначения. Источник н. с. F обозначают как источник э. д. с. (батарея, генератор), магнитное сопротивление R — как электрическое. Величины и направления потоков магнитной цепи соответствуют величинам и направлениям токов электрической цепи.

Рассмотрим два примера построения схемы замещения.

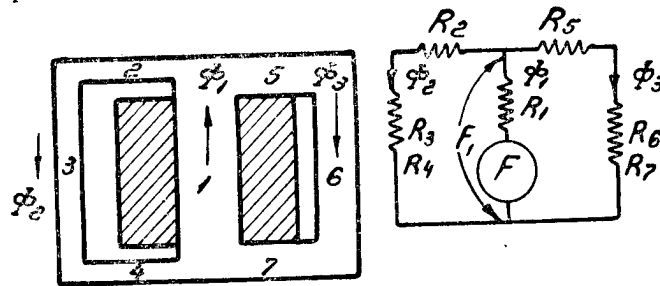
1. Замкнутая магнитная цепь без воздушных зазоров имеет две несимметричные ветви. Для упрощения схемы принимаем, что в магнитной системе отсутствуют потоки утечки. По аналогии с электрической цепью считаем, что под действием н. с. $F = I\omega$ создается поток Φ (фиг. 3, 1), протекающий в среднем сердечнике и разветвляющийся на два потока: Φ_2 и Φ_3 , причем значения этих потоков определяются одной и той же величиной н. с.:

$$F_1 = F - \Phi_1 R_1.$$

Тогда

$$\Phi_2 = \frac{F - \Phi_1 R_1}{R_2 + R_3 + R_4}; \quad \Phi_3 = \frac{F - \Phi_1 R_1}{R_5 + R_6 + R_7}.$$

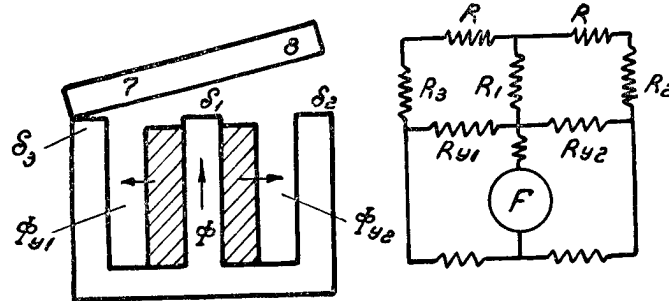
2. Магнитная система реле (фиг. 3, 2). В данной системе зазоры δ_1 и δ_2 рассматриваются как рабочие, а δ_3 — как нерабочий.



Фиг. 3, 1. Магнитная система.

В схеме замещения показаны отдельно сопротивления, соответствующие сопротивлениям воздушных зазоров, и отдельно сопротивления магнитопровода.

Потоки утечки Φ_{y1} и Φ_{y2} от среднего сердечника к крайним распределяются по всей длине сердечника, но в схеме замещения



Фиг. 3, 2. Магнитная система реле.

сопротивления путей потоков утечки показаны в виде активных сопротивлений, включенных за источником н. с., F . При этом следует учитывать приведенные значения магнитных сопротивлений (см. § 3-4).

По схеме замещения, пользуясь законами Ома и Кирхгофа для магнитных цепей, можно составить ряд уравнений, из которых определяются значения потоков в разных частях магнитной системы при заданной н. с. катушки.

Также решается и обратная задача: определение требуемой н. с. катушки при заданных значениях рабочих потоков в зазорах δ_1 и δ_2 .

Решение задачи несколько усложняется тем обстоятельством, что значение магнитного сопротивления стали не является постоянной величиной, а закон его изменения нелинейный. Поэтому для полного решения задачи приходится прибегать дополнительно к использованию кривых намагничивания материала. В некоторых случаях прибегают к решению задачи методом последовательных приближений (см. § 3-11).

3-3. МАГНИТНАЯ ПРОВОДИМОСТЬ ВОЗДУШНЫХ ЗАЗОРОВ

Методы вычисления проводимости. При расчете магнитной системы важно знать величины магнитных проводимостей воздушных зазоров: Рассмотрим вначале способы вычисления проводимостей воздушных зазоров между двумя эквипотенциальными поверхностями, которыми являются поверхности стальных деталей магнитопровода. Назовём эти поверхности магнитными полюсами.

Простейшим видом поля в воздушном зазоре является однородное поле, которое можно получить между двумя достаточной протяженности параллельными плоскостями — полюсами. В этом случае задача вычисления проводимости зазора решается просто. Однако однородное поле в воздушных зазорах магнитных систем электрических аппаратов и реле наблюдается весьма редко, так как поверхности полюсов имеют различную конфигурацию, полюсы часто располагаются непараллельно друг другу и имеют конечные размеры. Задача определения проводимости воздушных зазоров поэтому связана с рядом сложных вычислений.

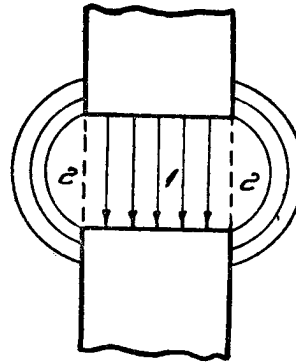
Поле между параллельными полюсами конечных размеров (фиг. 3, 3) можно рассматривать состоящим как бы из двух частей: однородного поля 1 и поля 2, определяемого выпучиванием линий индукции. Поток 2 называют краевым потоком.

В практических расчетах применяется несколько методов определения проводимости.

Аналитический метод. В простейшем случае, при наличии однородного поля между двумя параллельными плоскостями большой протяженности, проводимость воздушного зазора определяется как отношение площади q — сечения потока к длине δ — пути линии поля в воздухе.

$$G = \mu_0 \frac{q}{\delta} \frac{ab}{a}, \quad (3-6)$$

где q — см²; δ — см; $\mu_0 = 1,25 \cdot 10^{-8} \frac{ab}{a \cdot \text{см}}$.

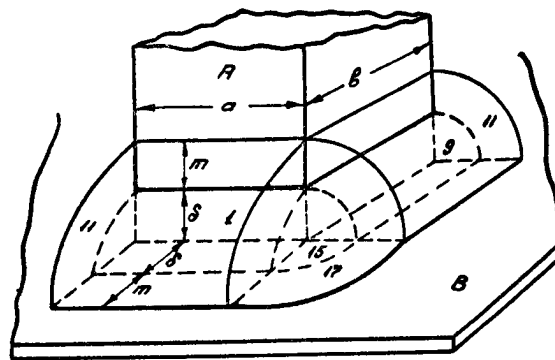


Фиг. 3,3. Поле между параллельными полюсами.

Для сложных случаев, например для вычисления проводимости поля между двумя параллельными цилиндрами или для поля между цилиндром и плоскостью, пользуются соотношениями, выведенными на основе теории поля.

В таблице 16 приведены соответствующие формулы, полученные на основе теории поля.

Метод разбивки поля на простые фигуры. Сущность этого метода заключается в том, что все поле в воздушном зазоре между полюсами разбивается посредством трубок индукции на



Фиг. 3, 4. Пример разбивки поля на простые фигуры:
1 — прямоугольный параллелепипед; 9 — четверть круглого цилиндра; 11 — четверть круглого полого цилиндра; 15 — половина шарового квадранта; 17 — половина квадранта шаровой оболочки.


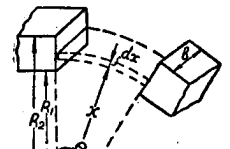
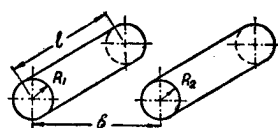
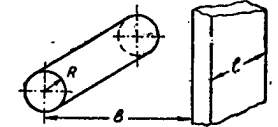
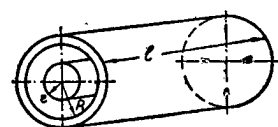

ряд потоков, причем разбивка производится так, чтобы эти трубки индукции имели форму простых геометрических фигур, например цилиндров, полуцилиндров и т. д.

Проводимость каждой трубки определяется как отношение среднего сечения трубки к средней ее длине. Полная проводимость равна сумме проводимостей отдельных трубок, если приложенная к концам разность магнитных потенциалов остается для всех трубок одинаковой. На фигуре 3, 4 приведен пример такой разбивки. Рассматривается поле между двумя полюсами A и B, причем размеры полюса B (в плане) в несколько раз больше размеров полюса A. Все поле разбивается на такие трубки индукции:

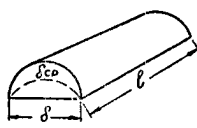
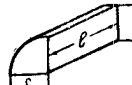
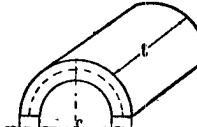


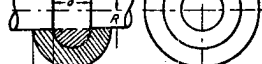
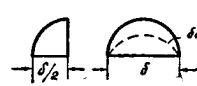
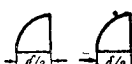


- 1 — прямоугольный параллелепипед размерами a , b и δ , причем δ — соответствует величине воздушного зазора;
- 9 — четверть круглого цилиндра с радиусом, равным δ ;
- 11 — четверть полого круглого цилиндра с внутренним радиусом, равным δ , и наружным, равным $\delta + m$, где m — некоторая произвольная величина, принимаемая обычно по опытным данным и равная $(1 + 2)\delta$;
- 15 — половина шарового квадранта;
- 17 — половина квадранта шаровой оболочки.

Таблица 16

Формулы для вычисления проводимостей воздушных путей

№	Эскиз	Проводимость $G \frac{\text{об}}{\text{ас}}$; $\mu_0 = 1,25 \cdot 10^{-8} \frac{\text{об}}{\text{ас}}$
1		<p>Параллельные цилиндры (без учета проводимости краевого потока)</p> $G_1 = \mu_0 \frac{ab}{\delta} \quad (1); \quad G_2 = \mu_0 \frac{\pi d^2}{4 \cdot \delta} \quad (2)$
3		<p>Непараллельные плоскости, расположенные под углом (без учета проводимости краевого потока)</p> $G_3 = \mu_0 \frac{b}{\theta} \int_{R_1}^{R_2} \frac{dx}{x} = \mu_0 \frac{b}{\theta} \ln \frac{R_2}{R_1} \quad (3)$
4		<p>Параллельные цилиндры.</p> <p>1) $G_4 = \mu_0 \frac{2\pi l}{\ln(k + \sqrt{k^2 - 1})} \quad (4)$</p> $R_1 \neq R_2; \quad k = \frac{b^2 - R_1^2 - R_2^2}{2R_1 R_2};$ $R_1 = R_2 = R; \quad k = \frac{b^2 - 2R^2}{2R^2};$ <p>2) при $b > 8R$; $G_{4a} = \mu_0 \frac{\pi l}{\ln b/R} \quad (4a)$</p>
5		<p>Цилиндр, параллельный плоскости.</p> $G_5 = \mu_0 \frac{2\pi l}{\ln \frac{b + \sqrt{b^2 - R^2}}{R}} \quad (5)$ <p>при $b > 4R$</p> $G_{5a} = \mu_0 \frac{2\pi l}{\ln 2b/R} \quad (5a)$
6		<p>Концентрические цилиндры.</p> $G_6 = \mu_0 \frac{2\pi l}{\ln R/r} \quad (6)$ <p>если $r > \delta$, где $\delta = R - r$, то</p> $G_{6a} = \mu_0 \frac{2\pi (r + \delta/2) l}{\delta} \quad (6a)$
7		<p>Цилиндр и концентрические поверхности (без учета проводимости краевого потока)</p> $G_7 = \mu_0 \frac{\theta b}{\ln R/r} \quad (7)$ <p>при $r > \delta$; $G_{7a} = \mu_0 \frac{(r + 0,5\delta) b \theta}{2\delta} \quad (7a)$</p>

(продолжение)

№	Эскиз	Проводимость $G \frac{\text{вб}}{\text{ав}}$; $\mu_0 = 1,25 \cdot 10^{-8} \frac{\text{вб}}{\text{а см}}$
8		<p>Половина цилиндра. Графически: $\delta_{cp} = 1,22\delta$;</p> $q_{cp} = \frac{\pi \delta^2 l}{8 \cdot 1,22 \cdot \delta} = 0,322\delta l$ $G_8 = \mu_0 0,26l \quad (8)$
9		<p>Четверть цилиндра</p> $G_9 = \mu_0 0,52l \quad (9)$
10		<p>Половина полого цилиндра</p> $\delta_{cp} = \pi \frac{\delta + m}{2}; \quad q_{cp} = ml.$ $G_{10} = \mu_0 \frac{2l}{\pi \left(\frac{\delta}{m} + 1 \right)} \quad (10)$
11		<p>Четверть полого цилиндра</p> $G_{11} = \mu_0 \frac{2l}{\pi \left(\frac{\delta}{m} + 0,5 \right)} \quad (11)$
12		<p>Тело вращения</p> $G_{12} = \mu_0 \cdot 0,26l = \mu_0 1,63r \quad (12)$
13		$G_{13} = \mu_0 \frac{2l}{\pi \left(\frac{\delta}{m} + 1 \right)} = \mu_0 \frac{4r}{\frac{\delta}{m} + 1} \quad (13)$
14		<p>Шаровой квадрант</p> $\delta_{cp} = 1,3\delta; \quad V = \frac{\pi}{3} \left(\frac{\delta}{3} \right)^3; \quad q = 0,1\delta^2;$ $G_{14} = \mu_0 \cdot 0,077\delta \quad (14)$
15		<p>Половина шарового квадранта</p> $G_{15} = \mu_0 \cdot 0,308\delta \quad (15)$
16		<p>Квадрант шаровой оболочки</p> $G_{16} = \mu_0 \frac{m}{4} \quad (16)$
17		<p>Половина квадранта шаровой оболочки</p> $G_{17} = \mu_0 \frac{m}{2} \quad (17)$

Номера поименованных форм соответствуют обозначениям таблицы 16. В этой же таблице приведены формулы для вычисления проводимостей различных простых фигур.

Расчетно-экспериментальные формулы. Некоторыми авторами на основании ряда расчетов и экспериментальных данных установлено численное влияние краевого эффекта на величину проводимости между плоскими прямоугольными и круглыми полюсами. А. Г. Сливинской на основе проведенных замеров предложены формулы для определения магнитной проводимости между полюсами различных форм — плоскими коническими, усеченными коническими. Б. К. Буль построил расчетные кривые для определения проводимости между плоскими полюсами прямоугольной формы. Шмидель построил ряд опытных кривых для определения проводимости зазора между плоскими полюсами.

Б. К. Буль и Шмидель определяют по формуле однородного поля $G = \mu_0 \frac{a_1 b_1}{\delta}$ проводимость зазора с учетом краевого эффекта; значения a_1 и b_1 являются размерами некоторого фиктивного полюса, отличающимися от действительных размеров a и b .

Графический метод. По этому методу посредством линий индукции и эквипотенциальных линий, построенных на основании данных о значении разности потенциалов между отдельными частями полюсов, поле разбивается на ряд отдельных трубок. Графически определяют среднее сечение и среднюю длину каждой трубки и по этим данным проводимость трубки. Общая проводимость воздушного зазора рассчитывается как сумма проводимостей отдельных трубок.

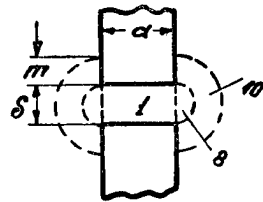
Сравнение методов и область их применения. Наиболее удобным для расчета является аналитический метод. Однако он применим для ограниченного числа случаев, когда для вывода формул можно пользоваться теорией поля.

Метод разбивки поля на простые фигуры дает решение задачи, в основном, для всех случаев определения проводимостей между полюсами разных форм. Этот метод дает решение в аналитической форме, поэтому его удобно применять для случаев, когда требуется вычислить ряд значений проводимости, а также производные проводимости по ходу якоря. Несмотря на приближенное решение задачи методом разбивки поля, результат получается в пределах точности, допустимой для практических расчетов.

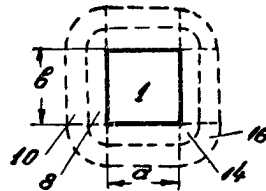
Расчетно-экспериментальные методы имеют ограниченную область применения, так как в них рассматриваются полюсы только определенной формы. Методы эти дают решение не в аналитической форме и неудобны, поэтому, при расчете проводимостей при меняющемся зазоре, производную проводимости приходится определять графическим дифференцированием, менее точным, чем аналитическое.

Графический метод применяется в тех случаях, когда картина поля весьма сложна, например при вычислении поля между криволинейными поверхностями.

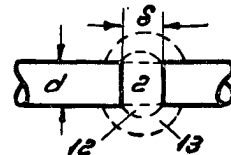
На фигурах 3, 5 до 3,12 приведены примеры вычисления проводимостей воздушных зазоров в различных частях магнитных систем реле. При вычислениях были применены формулы таблицы 16.



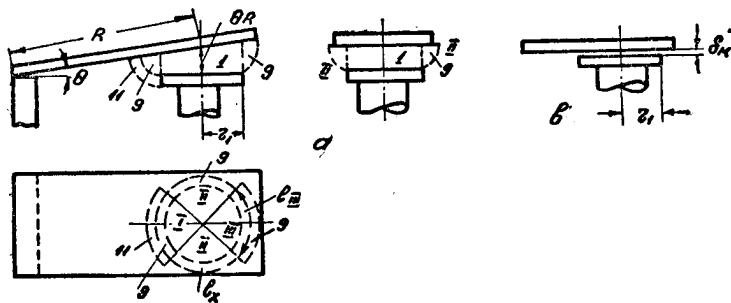
Фиг. 3, 5. Проводимость между прямоугольными полюсами:
 $G = G_1 + G_8 + G_{10} + 4G_{14} + 4G_{16}$;
 для G_8 и G_{10} : $l = 2a + 2b$; $m = (1 \div 2)\delta$.



Фиг. 3, 6. Проводимость между прямоугольным полюсом — плоскостью:
 $G = G_1 + G_9 + G_{11} + 4G_{15} + 4G_{17}$;
 для G_9 и G_{11} : $l = 2a + 2b$.

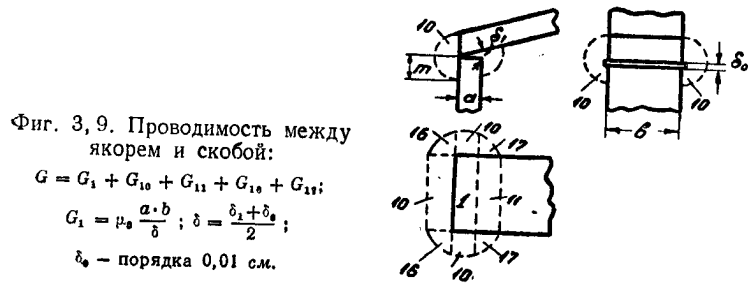


Фиг. 3, 7. Проводимость между двумя круглыми полюсами:
 $G = G_2 + G_{12} + G_{13}$.

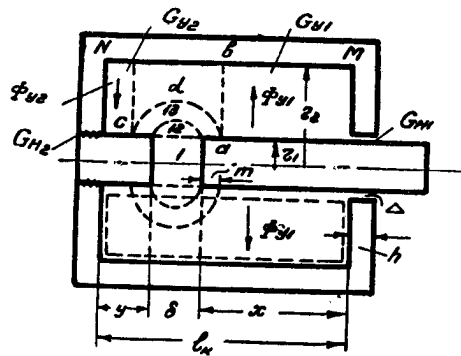


Фиг. 3, 8. Проводимость между ярком и полюсным наконечником:
 а — разомкнутое положение ярка:
 $G = G_1 + (G_9 + G_{11})_I + 2(G_9)_{II} + (G_9)_{III}$;
 $G_1 = \mu_0 \frac{\pi r_1^2}{\theta R}$; θ — рад;
 для G_9 и G_{11} вычисляются l_x в каждом квадранте
 б — замкнутое положение ярка:
 $G = \mu_0 \frac{\pi r_1^2}{\delta_k}$,

δ_k — при отсутствии немагнитной прокладки — порядка 0,01 см.



Фиг. 3, 9. Проводимость между якорем и скобой:
 $G = G_1 + G_{10} + G_{11} + G_{12} + G_{13};$
 $G_1 = \mu_0 \frac{a \cdot b}{\delta}; \delta = \frac{\delta_1 + \delta_2}{2};$
 δ_0 — порядка 0,01 см.



Фиг. 3, 10. Прямоходовой электромагнит цилиндрической формы:

$\delta < l_k; \delta$ — полный ход якоря.

а) Проводимости рабочего зазора:

$$G_p = \mu_0 \frac{\pi r_1^2}{\delta} + \mu_0 \cdot 1,63 r_1 + \mu_0 \cdot 4 \frac{r_1}{\delta/m + 1}.$$

Значение m определяется из условия равенства максимальной длины силовой линии adc двойного пути якоря к ярму — ab :

$$\pi \left(m + \frac{\delta}{2} \right) = 2 (r_2 - r_1).$$

б) Проводимости нерабочих зазоров:

$$G_{N1} = \mu_0 \frac{2\pi (r_1 + 0,5 \Delta)}{\Delta} h; \Delta - \text{порядка } 0,015 \text{ см};$$

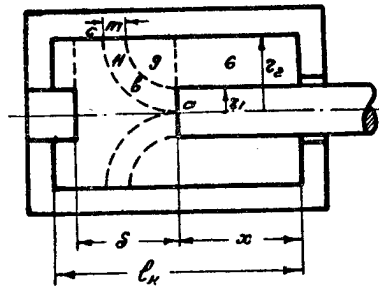
$$G_{N2} = \mu_0 \frac{2\pi r_1 \cdot h}{\Delta_0}; \Delta_0 - \text{порядка } 0,005 \text{ см}.$$

в) Проводимости утечки:

$$\text{удельная — на } 1 \text{ см длины } g = \mu_0 \frac{2\pi}{\ln r_2/r_1};$$

$$G_{y1} = g (x - m); G_{y2} = g (y - m).$$

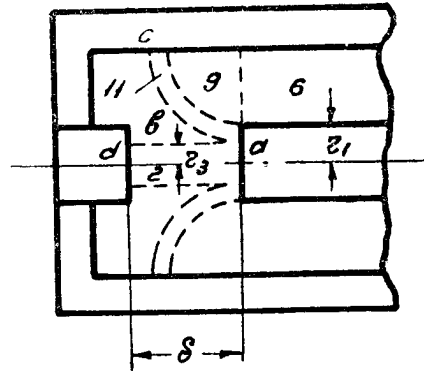
Приведенные значения проводимостей см. в параграфе 3-4.



Фиг. 3, 11. Прямоходовой электромагнит цилиндрической формы:
 $\delta > abc$; $G_{y0} = \mu_0 \cdot 4,7 \cdot r_1$
 (см. таблицу 16).

$$G_{y11} = \mu_0 \frac{2l}{\pi \left(\frac{\delta}{m} + 0,5 \right)}; l = \pi r_1; \delta = m = r_1.$$

Вычисление приведенных значений G_{y0} , G_{y0} и G_{y11} см. в параграфе 3-4.



Фиг. 3, 12. Прямоходовой электромагнит цилиндрической формы:

$$\delta = abc; G_1 = \mu_0 \frac{\pi r_1^2}{\delta}; G_{y0} = \mu_0 \cdot 4,7 \cdot r_1;$$

$$G_{y11} = \mu_0 \frac{2l}{\pi \left(\frac{\delta}{m} + 0,5 \right)};$$

$$l = \pi (r_1 + r_2); m = r_1 - r_2;$$

$$r_2 \text{ из условия } abc = \delta.$$

Вычисление приведенных значений G_{y0} , G_{y0} и G_{y11} см. в параграфе 3-4.

3-4. КОЭФФИЦИЕНТ РАССЕЯНИЯ

Учет распределения потоков. При расчете магнитной цепи необходимо иметь данные о распределении потоков по всему магнитопроводу. За основу принимают значение потока Φ_0 в рабочем воздушном зазоре, и все остальные значения потоков в разных частях магнитопровода выражают в функции Φ_0 , для чего пользуются значением коэффициента рассеяния σ . В сложных магнитопроводах приходится учитывать несколько значений σ . Рассмотрим для примера магнитную систему, приведенную на фигуре 3, 13. В плоскости N действует поток Φ_0 , а в плоскости M $\Phi_0 + \Phi_y$; таким образом,

$$\sigma = \frac{\Phi_0 + \Phi_y}{\Phi_0}. \quad (a)$$

Вычисление значений σ в функции потоков неудобно, так как неизвестны величины этих потоков. Поэтому уравнение (a) преобразовывают так, чтобы выразить σ в функции геометрических размеров магнитопровода.

Если не учитывать падения магнитного потенциала в стали, то значение Φ_0 определяется н. с. катушки F ; величина потока Φ_y определяется неполным значением н. с. F' , так как н. с. катушки меняется от нуля (плоскость M) до F (плоскость N).

$$\begin{aligned} \Phi_0 &= F \frac{G_1 \cdot G_2}{G_1 + G_2}; \\ \Phi_y &= F'G; F' < F. \end{aligned} \quad (b)$$

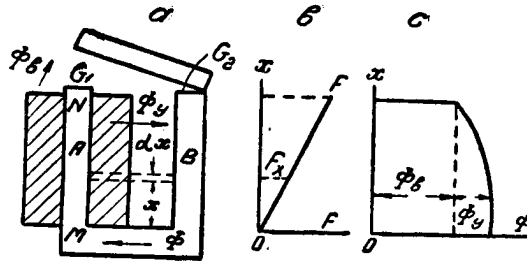
Удобнее пользоваться для вычисления Φ_y значением F , а не F' , для чего меняют соответственно значение G_y и пользуются значением G_{yn} — приведенного значения проводимости утечки.

$$\Phi_y = FG_{yn} \quad (c).$$

Из (a), (b) и (c):

$$\sigma = \frac{F \frac{G_1 G_2}{G_1 + G_2} + FG_{yn}}{\frac{G_1 G_2}{G_1 + G_2} F} = 1 + G_{yn} \frac{G_1 + G_2}{G_1 G_2} \quad (3-7)$$

Полученное значение σ назовем коэффициентом рассеяния по потоку.



Фиг. 3, 13. Распределение потоков в системе.

Учет распределения числа потокоцеплений. При проектировании магнитных систем переменного тока с катушкой параллельного включения приходится дополнительно учитывать коэффициент рассеяния по числу потокоцеплений.

Считая

$$U \cong E = \frac{\omega}{\sqrt{2}} \psi = \frac{\omega}{\sqrt{2}} (\psi_a + \psi_y), \quad (a)$$

где ψ , ψ_a , ψ_y — числа потокоцеплений, соответственно — полное, потока рабочего воздушного зазора и потока утечки.

Коэффициент рассеяния σ_c по числу потокоцеплений

$$\sigma_c = \frac{\psi_a + \psi_y}{\psi_a} = \frac{\omega F \frac{G_1 G_2}{G_1 + G_2} + \omega F G_{yc}}{\omega F \frac{G_1 G_2}{G_1 + G_2}}$$

или

$$\sigma_c = 1 + G_{yc} \frac{G_1 + G_2}{G_1 G_2}, \quad (3-8)$$

где G_{yc} — приведенное значение проводимости, G_y , по числу потокоцеплений.

Вычисление G_{yn} по потоку. Определим значение потока Φ_{yx} на высоте x (фиг. 3, 13). Обозначим: g — удельная магнитная проводимость потока утечки (на 1 см длины), вычисленную по

геометрическим размерам системы между сердечником A и ярмом B .
Н. с. на высоте катушки x равна

$$F_x = F \frac{x}{l} \quad (a);$$

$$\Phi_{yx} = \int_0^x F_x g dx = \frac{Fg}{l} \int_0^x x dx = Fg \frac{x^2}{2l} \quad (b)$$

при $x = l$;

$$\Phi_y = Fg \frac{l}{2} \quad (3-9)$$

значение $gl = G_y$ — проводимость потока утечки, вычисленная по геометрическим размерам.

Таким образом,

$$G_{yn} = \frac{gl}{2} = \frac{G_y}{2}. \quad (3-10)$$

Закон изменения потока утечки показан на фигуре 3, 13 *с*. Соотношение (3-10) применимо для ненасыщенных систем, когда максимальное значение индукции в магнитопроводе не превышает $11 \cdot 10^{-5} \div 12 \cdot 10^{-5} \frac{\text{вб}}{\text{см}^2}$. При большем значении индукции необходимо вводить поправки, учитывающие падение магнитного потенциала в сердечнике.

Согласно данным ряда проведенных расчетов, при которых принято среднее значение μ по всей длине сердечника, величина G_{yn} с учетом поправки

$$G_{yn} = \frac{G_y}{3}. \quad (3-11)$$

Вычисление G_{yc} по числу потокосцеплений. Рассмотрим магнитную систему, приведенную на фигуре 3, 13 *а*. Число потокосцеплений $d\psi$ потока утечки на участке dx , расположенном на расстоянии x от начала катушки, равно потоку $d\Phi_x$, умноженному на число охватываемых им витков w_x :

$$d\psi_x = d\Phi_x w_x = F_x g dx w_x$$

или

$$\psi_x = \int_0^x d\psi_x = Fwg \frac{x^2}{2l} = FwG_{y0}.$$

Таким образом, значение приведенной проводимости по числу потокосцеплений

$$G_{yc} = g \frac{x^2}{2l^2} \quad (3-12)$$

при $x = l$;

$$G_{yc} = \frac{gl}{2}. \quad (3-13)$$

В данном случае значение приведенной проводимости в три раза меньше вычисленной по геометрическим размерам магнитопровода: $G_y = gl$. Если учитывать падение магнитного потенциала в сердечнике, то можно принять

$$G_{yc} = \frac{2}{3} \frac{gl}{3}. \quad (3-14)$$

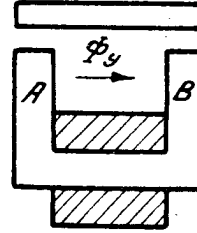
Пример — прямоходовая магнитная система (фиг. 3, 10). Определяют два коэффициента рассеяния:

σ_1 — для якоря; σ_2 — для стопа.

$$\sigma_1 = \frac{G_a + G_{yn1}}{G_a}; \quad \sigma_2 = \frac{G_a + G_{yn2}}{G_a}.$$

Таким образом, при заданном потоке Φ_a получаем: в плоскости M — $\sigma_1 \Phi_a$; в плоскости N — $\sigma_2 \Phi_a$. Учитывают удельную проводимость потока утечки g :

$$G_{yn1} = g \frac{(x-m)^2}{2l_h}; \quad G_{yn2} = g \frac{(y-m)^2}{2l_h}.$$



Фиг. 3, 14. Эскиз магнитной системы.

С учетом падения магнитного потенциала в стали

$$G_{yn1} = g \frac{(x-m)^2}{3l_h}; \quad G_{yn2} = g \frac{(y-m)^2}{3l_h}.$$

Пример — магнитопровод (фиг. 3, 14). В данном случае учитывают полное значение G_y , так как разность потенциалов между A и B , если не учитывать падения магнитного потенциала в стали, остается постоянной по всей длине.

3-5. МЕСТОПОЛОЖЕНИЕ МАКСИМАЛЬНОГО ЗНАЧЕНИЯ ПОТОКА В МАГНИТНОЙ СИСТЕМЕ

В зависимости от соотношения проводимостей воздушных зазоров системы меняется местоположение максимального значения потока в системе. В параграфе 3-3 при определении коэффициента рассеяния было принято, что максимальное значение потока имеет место у основания сердечника, т. е. у начала катушки (фиг. 3, 13).

Другая картина получается при наличии дополнительного воздушного зазора δ_3 между сердечником и ярмом (фиг. 3, 15). В этом случае вследствие падения магнитного потенциала в зазоре δ_3 часть потока утечки Φ_{y2} возвращается от ярма к сердечнику не у основания, по стали, а частично по воздуху. В такой системе плоскость ab является нейтралью и между точками a — b разность потенциалов равна нулю. Расстояние l_3 нейтрали от основания катушки зависит от соотношения проводимостей воздушных зазоров.

Для определения l_2 построим схему замещения (фиг. 3, 15 с), в которой

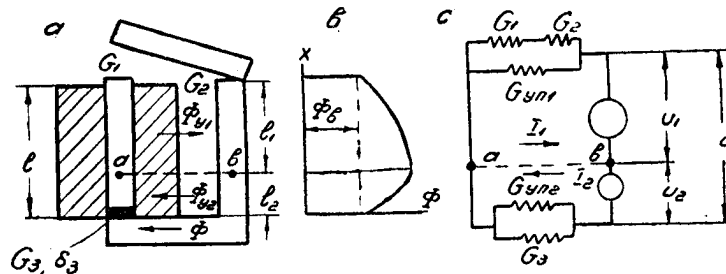
U — напряжение батарей, соответствующее н. с. катушки F .

U_1 — напряжение верхней части батареи, а U_2 — нижней.

$$U_1 = U \frac{l_1}{l}; \quad (3-15)$$

$$U_2 = U \frac{l_2}{l}, \quad (3-16)$$

gl_1 — gl_2 — соответственно приведенные проводимости утечки для верхней и нижней частей сердечника.



Фиг. 3, 15. Схема для определения местоположения максимального значения потока.

В нейтральном слое суммарные потоки верхней части магнитопровода должны быть равны суммарным потокам нижней части и направлены в разные стороны. Это условие выражается в схеме замещения.

$$I_1 = -I_2; \quad (3-17)$$

$$G = \frac{G_1 G_2}{G_1 + G_2}; \quad U_1 \left(G + \frac{gl_1}{2} \right) = U_2 \left(G_3 + \frac{gl_2}{2} \right). \quad (3-18)$$

Учитывая (3-16) и (3-17), получаем

$$U \frac{l_1}{2l} (2G + gl_1) = U \frac{l_2}{2l} (2G_3 + gl_2). \quad (3-19)$$

Так как $l = l_1 + l_2$ (3-19), решая (3-18) и (3-19) совместно, получаем

$$\frac{l_1}{l_2} = \frac{2G_3 + gl_2}{2G + gl_1}. \quad (3-20)$$

При зазоре $\delta_3 = 0$; $G_3 = \infty$, величина $l_2 = 0$. Этот случай соответствует системе, показанной на фигуре 3, 13. При наличии зазора δ_3 , с уменьшением δ нейтраль перемещается кверху.

При выводе (3-20) не принято во внимание падение магнитного потенциала в стали, что допустимо для ненасыщенных систем.

Влияние смещения нейтрали на работу реле.

а) В системах с поворотным якорем зазор δ_2 всегда мал сравнительно с зазором δ_1 , смещение нейтрали невелико, и в практических расчетах принимают, что максимум потока расположен у основания катушки.

б) В прямоходовых системах смещение нейтральной плоскости оказывает существенное влияние на величину силы на якорь: значение соленоидной силы зависит от потоков Φ_{VI} и Φ_{VII} , направленных в разные стороны (фиг. 3, 16). При уменьшении зазора δ нейтраль ab смещается к этому зазору, поток Φ_{VI} уменьшается, а Φ_{VII} увеличивается, в результате чего значение соленоидной силы резко падает. При величине $\delta_1 < 0,1l_k$ значением соленоидной силы пренебрегают.

Ниже приводится приближенная формула для определения значения l_2 .

$$l_2 = \frac{l(2G + gl_k)}{2(G + G_{н1} + gl_k)}, \quad (3-21)$$

где G — проводимость рабочего зазора.

3-6. ПАДЕНИЕ МАГНИТНОГО ПОТЕНЦИАЛА В СТАЛИ И В ЗАЗОРАХ

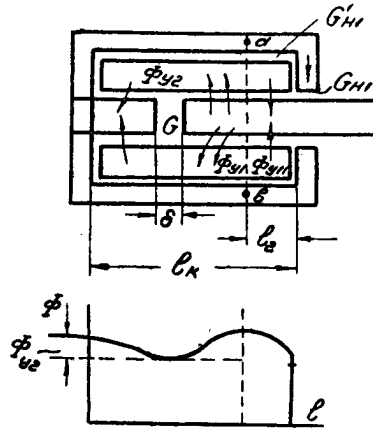
При расчете магнитной цепи вычисляют падение магнитного потенциала в рабочих и нерабочих зазорах, а также во всех частях магнитопровода.

а) **Падение в стали.** Для каждого элемента магнитопровода — скобы, сердечника, якоря — вычисляют индукцию B_c для этого элемента и находят по кривой намагничивания напряженность магнитного поля для данного материала, т. е. падение н. с. на 1 см длины стали. Падение н. с. на данном участке стального магнитопровода длиной l_1 :

$$F_{c1} = H_1 l_1,$$

где l_1 — в см.

В некоторых элементах магнитопровода значение индукции неодинаково по всей их длине. Так, в сердечнике электромагнита, изображенного на фигуре 3, 13, индукция B_{c1} в верхней части (у рабочего зазора) меньше значения индукции B_{c2} — внизу, что объясняется наличием потоков утечки. С достаточной для прак-



Фиг. 3, 16. Распределение потоков в прямоходовой системе.

тических целей точностью падение н. с. в сердечнике определяется по среднему значению напряженности поля:

$$F_c = \frac{H_1 + H_2}{2} l_c, \quad (3-22)$$

где H_1 — напряженность поля, соответствующая индукция B_{c1} , а H_2 соответствующая B_{c2} .

При определении H для систем переменного тока следует принимать во внимание амплитудное значение индукции в стали B . Далее нужно учитывать, что в шихтованных магнитных системах, применяемых, как правило, для электромагнитов переменного тока, сечение q детали не соответствует сечению стали, так как между отдельными листами шихтовки имеется изоляционная прослойка, необходимая для уменьшения потерь на гистерезис и вихревые токи. Каждый лист стали магнитопровода оклеивается с одной стороны бумагой толщиной 0,032 мм или покрывается слоем лака толщиной 0,02 мм. Покрытие слоем лака дает лучшее решение, так как пленка лака негигроскопична и более устойчива механически.

При определении значения индукции в стали шихтованного магнитопровода учитывают коэффициент заполнения шихтовки. Принимают $k = 0,95$ — для листов, покрытых лаком; $k_c = 0,90$ — для листов, оклеенных бумагой.

Таким образом,

$$B = \frac{\Phi}{k_c q}, \quad (3-23)$$

где Φ — амплитудное значение потока;

q — сечение детали в $см^2$;

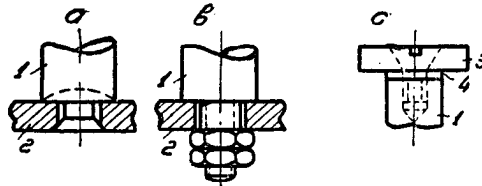
B — амплитудное значение индукции, $\frac{вб}{см^2}$;

б) Падение в воздушных зазорах. Помимо рабочего воздушного зазора между якорем и сердечником электромагнита, названного так потому, что момент на якоре создается потоком, протекающим в этом зазоре, имеются еще и нерабочие зазоры, где также теряется определенная часть н. с. Так, нерабочим считается зазор между якорем и ярмом. Кроме того, в магнитной системе имеется нерабочий зазор в стыке между сердечником и скобой (ярмом) (фиг. 3, 17), который определяется неточностью изготовления цапфы сердечника и отверстия скобы, а также толщиной покрытия цинком обеих деталей. Следует также учитывать нерабочий зазор между сердечником и полюсным наконечником (фиг. 3, 17 с).

На основании практических данных принимают следующие значения нерабочих зазоров в стыках двух неподвижных деталей:

1) при соединениях, исполненных по нормальному технологическому процессу, 0,05 мм;

- 2) при соединениях, исполненных с применением давления порядка 200 кг/см^2 , зазор считают практически равным нулю;
 3) конечный рабочий зазор между притянутым якорем и сердечником считают также равным $0,05 \text{ мм}$.
 Стальные детали магнитопровода (за исключением шихтованных магнитопроводов) обычно оцинковывают, причем слой цинка

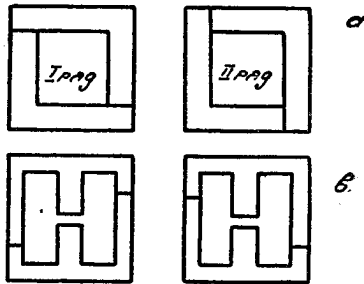


Фиг. 3, 17. Конструктивные элементы магнитопровода:
 1 — сердечник; 2 — ярмо; 3 — полюс; 4 — немагнитная прокладка.

в среднем бывает порядка 15 микрон. Таким образом, при сопряжении двух оцинкованных стальных деталей следует учитывать немагнитный слой толщиной $\delta = 0,05 \pm 2 \cdot 0,015 = 0,08 \text{ мм}$.

В некоторых случаях этот зазор искусственно увеличивают применением штифтов или прокладок из немагнитного материала (фиг. 3, 17 c).

Шихтованные магнитные системы для дроссельных катушек и вспомогательных трансформаторов исполняют так, чтобы уменьшить потери н. с. в стыках. На фигуре 3, 18 a приведен пример такого магнитопровода. Отдельные листы штампуются несимметричной формы. При наложении одного ряда листов на другой получается такой магнитопровод, что поток от одного листа к другому проходит в основном не по стыку между листами одного ряда, а через листы соседнего ряда. Ввиду большой проводимости зазора между плоскостями листов двух соседних рядов потери н. с. получаются незначительными. Если требуется исполнение дросселя с фиксированным воздушным зазором, магнитопровод исполняется так, как показано на фигуре 3, 18 b.



Фиг. 3, 18. Примеры шихтовки магнитных систем.

Падение н. с. в воздушном зазоре определяется по формуле

$$F = \frac{\Phi}{G}, \quad (3-24)$$

где Φ — поток в веберах;

G — проводимость зазора в $\frac{об}{ав}$.

3-7. КРИВЫЕ ПРОВОДИМОСТЕЙ

Для вычисления электромагнитных сил, действующих на якорь, и построения тяговой характеристики электромагнитного реле необходимо иметь данные проводимостей воздушных зазоров, изменяющихся по ходу якоря, а также значение производной этой проводимости для разных положений якоря.

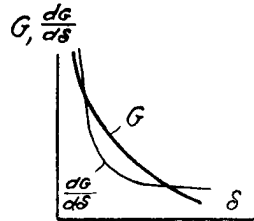
Определить значения $\frac{dG}{d\delta}$ или $\frac{dG}{d\theta}$ можно двумя путями: аналитическим или графическим. Наиболее удобное и точное решение задачи получается при аналитическом методе.

Рассмотрим для примера проводимость между двумя прямоугольными стержнями (фиг. 3, 5).

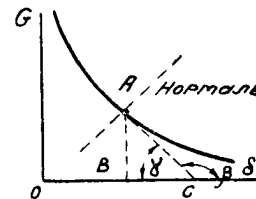
$$G = 1,25 \cdot 10^{-8} \left[\frac{ab}{\delta} + 0,26(2a + 2b) + 0,64 \frac{2a + 2b}{\delta/m + 1} + 0,306\delta + m \right];$$

$$\frac{dG}{d\delta} = 1,25 \cdot 10^{-8} \left[-\frac{ab}{\delta^2} - 0,64 \frac{2a + 2b}{m \left(\frac{\delta}{m} + 1 \right)^2} + 0,306 \right].$$

Для графического определения производной проводимости необходимо построить предварительно кривую $G = f_1(\delta)$ (фиг. 3, 19) или $G = \varphi_1(\theta)$ и затем провести операцию графического дифференцирования. Проводят касательные к отдельным точкам кривой $G = f_1(\delta)$ (фиг. 3, 20).



Фиг. 3, 19. Кривые проводимостей.



Фиг. 3, 20. Графическое определение производной проводимости.

Если m — масштаб оси ординат, а n — оси абсцисс, так что $G = \frac{AB}{m}$ и $\delta = \frac{OB}{n}$, то

$$\frac{dG}{d\delta} = \frac{n}{m} \frac{d(AB)}{d(BC)} = \frac{n}{m} \operatorname{tg} \beta = -\frac{n}{m} \operatorname{tg} \gamma.$$

Найдя значение $\frac{dG}{d\delta}$ для разных значений δ , строят кривую (фиг. 3, 19).

Для проведения касательной к кривой $G = f_1(\delta)$ находят вначале положение нормали к кривой в требуемой точке. Для этого

пользуются узкой полоской зеркала, которую ставят поперек кривой так, чтобы кривая и ее зеркальное отображение составляли одну плавную, без излома, кривую. Это положение зеркала определяет нормаль к кривой; касательная же к кривой будет перпендикулярна к нормали.

3-8. ПОТЕРИ В СТАЛИ МАГНИТОПРОВОДА

Полные потери складываются из потерь на гистерезис и вихревые токи: $P_n = P_a + P_v$. Величина потерь на вихревые токи зависит от индукции в системе, качества и толщины материала.

В ГОСТ 802—41 приводятся данные удельных потерь $\frac{вт}{кг}$ для разных марок листовой электротехнической стали при индукциях

$$B_1 = 1.10^{-4} \frac{вб}{см^2}$$

и

$$B_{1,5} = 1,5 \cdot 10^{-4} \frac{вб}{см^2}$$

при 50 гц.

При практических расчетах для индукции $B > 1.10^{-4} \frac{вб}{см^2}$ пользуются следующими упрощенными формулами:

$$P_a = \sigma_a \frac{f}{100} B^2 \left(\frac{вт}{кг} \right); \quad (3-25)$$

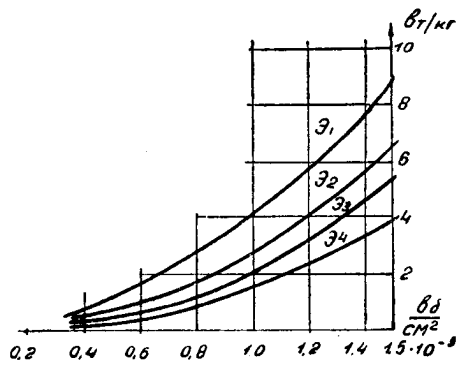
$$P_v = \sigma_v \left(\frac{f}{100} B \right)^2 \left(\frac{вт}{кг} \right). \quad (3-26)$$

Данные значений коэффициентов σ_a и σ_v для 50 гц приведены в таблице 17, кривые полных потерь для листовой электротехнической стали — на фигуре 3, 21.

Таблица 17

Значение коэффициентов σ_a и σ_v .

Марка стали	Э1							Э2	Э3	Э3а	Э4	Э4А	Э4АА
	2,0	1,5	1,0	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5	
Толщина материала, мм	2,0	1,5	1,0	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5
Коэффициент σ_a	4,4	4,4	4,4	4,4	3,6	3,2	2,85	3,0	2,6	2,4	2,4	2,4	2,4
" σ_v	89,6	50,5	22,4	5,6	3,2	2,9	2,3	1,2	1,2	1,2	1,2	1,0	1,0



Фиг. 3, 21. Кривые полных потерь для легированной стали 0,5 мм, 50 гц.

Следует учитывать, что данные, приведенные в ГОСТ и в таблице 17, относятся к отдельным листам стали. В собранном пакете магнитопровода потери получаются значительно большими (в 2,5—3 раза), что объясняется наклепом при штамповке, наличием заусениц, а также заклепок в пакете.

При термической обработке пакета потери снижаются.

3-9. ТОКИ И ПОТРЕБЛЯЕМАЯ МОЩНОСТЬ КАТУШЕК

При расчете магнитной системы необходимо вычислить значения тока в катушке и потребляемой катушкой мощности. Для систем постоянного тока эта задача решается весьма просто. Значительно более громоздкое решение получается для магнитных систем переменного тока.

В этом случае ток катушки равен

$$I = \sqrt{(I_{\mu a} + I_{\mu c})^2 + (I_n + I_a)^2}, \quad (3-27)$$

где $I_{\mu a}$ и $I_{\mu c}$ — составляющие тока намагничивания для воздушных зазоров и для стали;

I_n — составляющая тока катушки для компенсации потерь на гистерезис и вихревые токи в стали;

I_a , I_p — составляющие тока для компенсации потерь в короткозамкнутом витке электромагнитной системы — экране, или в роторе индукционной системы.

Значения указанных токов зависят от потоков системы, являющихся функцией э. д. с. катушки. Обычно в магнитных системах с катушкой напряжения электрическое сопротивление катушки в несколько раз меньше индуктивного. Поэтому в предварительных расчетах принимают, что все приложенное к катушке напряжение уравновешивается противодействующей э. д. с.

Так, если активное падение напряжения

$$Ir = 0,1U,$$

то индуктивное

$$Ix = U \sqrt{1 - 0,1^2} = 0,995U,$$

при

$$Ir = 0,2U; \quad E = 0,98U.$$

В тех случаях, когда к катушке приложено синусоидальное напряжение, изменение тока, как известно, происходит не по синусоиде. Однако для практических расчетов принимается допущение, что сила тока изменяется также на синусоиде.

Составляющая тока намагничивания для воздушных зазоров $I_{\mu a}$. Значение $I_{\mu a}$ может быть определено из соотношения

$$\sqrt{2} \cdot I_{\mu a} \omega G = \Phi, \quad (3-28)$$

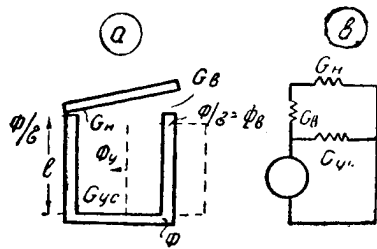
где $I_{\mu\phi}$ — действующее значение тока в амперах;
 Φ — амплитудное значение потока в веберах;
 G — проводимость воздушного зазора $\frac{\sigma\delta}{ab}$;

При наличии в магнитной системе нескольких воздушных зазоров и путей утечки значение G в (3-28) представляет собой полную приведенную для всех путей потока проводимость по воздуху. При этом учитывают приведенную по числу потокосцеплений проводимость утечки G_{yc} .

В ряде случаев для вычисления $I_{\mu\phi}$ удобнее пользоваться соотношением

$$I_{\mu\phi} = \frac{U}{\omega L} = \frac{U}{\omega w^2 G} \quad (3-29)$$

Примеры вычисления полной проводимости для разных форм магнитных систем приведены на фигурах 3, 22; 3, 23; 3, 24.

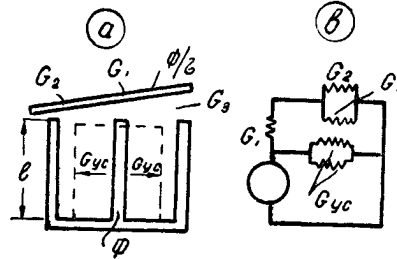


Фиг. 3, 22. Магнитная система переменного тока с поворотным якорем:

$$G = \frac{G_{\delta} \cdot G_{н}}{G_{\delta} + G_{н}} + G_{yc}; \quad \sigma = \frac{\Phi}{\Phi_{\delta}};$$

$$G_{yc} = \frac{gl}{3} \text{ — для ненасыщенных систем;}$$

$$G_{yc} = \frac{2}{3} \frac{gl}{3} \text{ — для насыщенных систем.}$$



Фиг. 3, 23. Магнитная система переменного тока с Ш-образным магнитопроводом:

$$G = \frac{G_{\delta 1} (G_{\delta 2} + G_{\delta 3})}{G_{\delta 1} + G_{\delta 2} + G_{\delta 3}} + 2G_{yc}$$

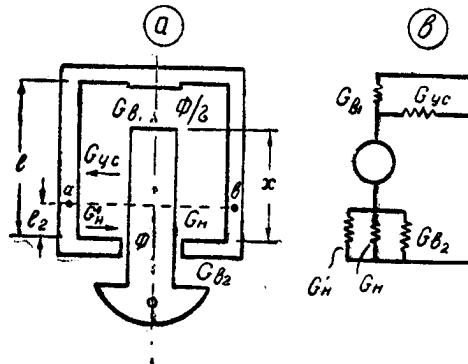
Фиг. 3, 24. Прямоходовая магнитная система переменного тока.

$$G_1 = G_{\delta 1} + G_{yc};$$

$$G_2 = G_{н} + G'_{н} + G_{\delta 2};$$

$$G = \frac{G_1 \cdot G_2}{G_1 + G_2}$$

ab — линия нейтрали для потоков утечки.



Составляющая $I_{\mu c}$. Эта составляющая определяется напряженностью поля H , которую находят по кривой намагничивания для данного значения индукции в стали B_c . При вычислении индукции B_c учитывается распределение потока в магнитопроводе.

Для данного участка магнитопровода $B_{ci} = \frac{\Phi}{k_c q_i}$.

Падение магнитного потенциала в магнитопроводе $F_c = \Sigma H_i l_i$, ток $I_{\mu c} = \frac{F_c}{\omega}$. (3-30)

Составляющая I_n . Вычисляется значение потерь в ваттах на гистерезис и вихревые токи, P_n .

$$I_n = \frac{P_n}{U}. \quad (3-31)$$

Составляющая I_s или I_p . Вычисляется мощность потерь в экране, P_s , или в роторе индукционной системы, P_p .

$$I_s = \frac{P_s}{U}. \quad (3-32)$$

Коэффициент мощности. Угол φ между вектором напряжения сети U и вектором тока катушки определяется из соотношения

$$\cos \varphi = \frac{P + P_n + P_s}{UI}, \quad (3-33)$$

где $P = I^2 r$ — активные потери в катушке.

Мощность, потребляемая токовыми катушками. Вычисляется кажущаяся мощность UI . Эту величину важно знать при включении катушки во вторичную цепь измерительных трансформаторов тока.

По данным расчета системы известны значения I , Φ , ω , r . Тогда

$$E = \frac{\omega \Phi \omega}{\sqrt{2}}; U_c = \sqrt{E^2 + (Ir)^2},$$

где U_c — падение напряжения на зажимах катушки. Тогда

$$U_c I = I \sqrt{E^2 + (Ir)^2}. \quad (3-34)$$

Данные для построения векторной диаграммы сложного магнитопровода. При построении векторных диаграмм сложных магнитопроводов, например систем реле переменного тока с экраном или тангенциальных индукционных систем, необходимо учитывать, что поток Φ основной части системы распределяется в системе на отдельные потоки. В этом случае падение магнитного потенциала и потери в стали зависят не от отдельных потоков, а от общего.

Рассмотрим для примера магнитную систему электромагнитного реле с экраном (фиг. 3, 25). Составляющие тока для потока Φ_1 : $I_{\mu s1}$, $I_{\mu c1}$ и I_{n1} .

2. Кривая представляет собой функцию потока в рабочем зазоре от н. с. всей катушки:

$$\Phi_e = \Phi_p + \Phi_k = f(F_p + F_n + F_c).$$

Ниже рассматриваются примеры построения таких кривых.

а) **Кривые намагничивания.** *Вариант 1.* Рассмотрим пример построения кривой

$$\Phi_e = \Phi_p + \Phi_k = f(F_n + F_c).$$

Поток распределяется в магнитопроводе неравномерно: при определенном значении потока в рабочем воздушном зазоре получаются разные значения потока в отдельных частях магнитопровода.

При построении кривой задаются определенным значением потока в рабочем зазоре, а значения потоков в других частях магнитопровода устанавливают в функции этого потока умножением на коэффициент рассеяния σ . Значение σ изменяется при ходе якоря; так, в разомкнутой системе $\sigma = 1,8 - 3$, в замкнутой — $\sigma = 1,05 - 1,1$.

При расчете электромагнита вычисляют значения сил или моментов для нескольких положений якоря: начального (разомкнутая система), конечного и промежуточного. Вследствие меняющегося при этом значения σ и величин нерабочих зазоров приходится строить несколько кривых намагничивания, соответственно принятым для расчета положениям якоря.

Методы вычисления проводимости воздушных зазоров приведены в параграфе 3-3, а падений н. с. в стали — в параграфе 3-6.

На фигуре 3, 26 приведена магнитная система с поворотным якорем и кривая намагничивания для нее:

$$\Phi_e = f(F_n + F_c).$$

К нерабочим относятся в рассматриваемой системе зазоры с проводимостями: G_2 и G_4 — неизменяющимися по ходу якоря, G_3 — изменяющейся по ходу якоря.

Для построения кривой задаются произвольно несколькими значениями потоков в рабочем воздушном зазоре. Пользуясь коэффициентом рассеяния σ , вычисляют значение потоков в разных частях магнитопровода при принятых значениях Φ_e . Все данные расчета сводят в таблицу 18.

Пользуясь кривой намагничивания, изображенной на фигуре 3, 26, можно решить следующие задачи.

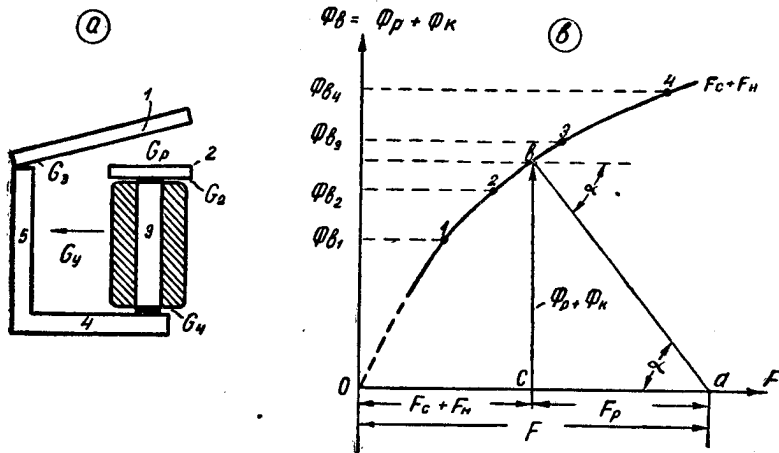
1. Определить значение потока Φ_p при заданном значении н. с. катушки и заданной проводимости рабочего зазора $G_p + G_k$.

Откладывают на оси абсцисс отрезок oa , соответствующий н. с. катушки, из точки a проводят прямую ab под углом α , причем $\operatorname{tg} \alpha = \frac{G_p + G_k}{p}$, (p — отношение масштабов оси ординат

и абсцисс; пример вычисления p см. стр. 111). Отрезок bc определяет собой величину потока в рабочем зазоре Φ_b .

Значение $\Phi_p = \Phi_s \frac{G_p}{G_p + G_k}$.

Падение н. с. в системе распределяется следующим образом: отрезок oc дает значение $F_c + F_n$, а отрезок ca — F_p . Угол α является функцией проводимости рабочего зазора.



Фиг. 3, 26. Магнитная система и кривая ее намагничивания.
Вариант $\Phi_s = f_1 (F_c + F_n)$.

Согласно фигуре 3, 26 имеем:

$$\operatorname{tg} \alpha = \frac{\Phi_s}{F_p}; \text{ но } \frac{\Phi_s}{\Phi_p} = G_p + G_k,$$

следовательно,

$$\operatorname{tg} \alpha = G_p + G_k,$$

или, с учетом масштабов,

$$\operatorname{tg} \alpha = \frac{G_p + G_k}{p}.$$

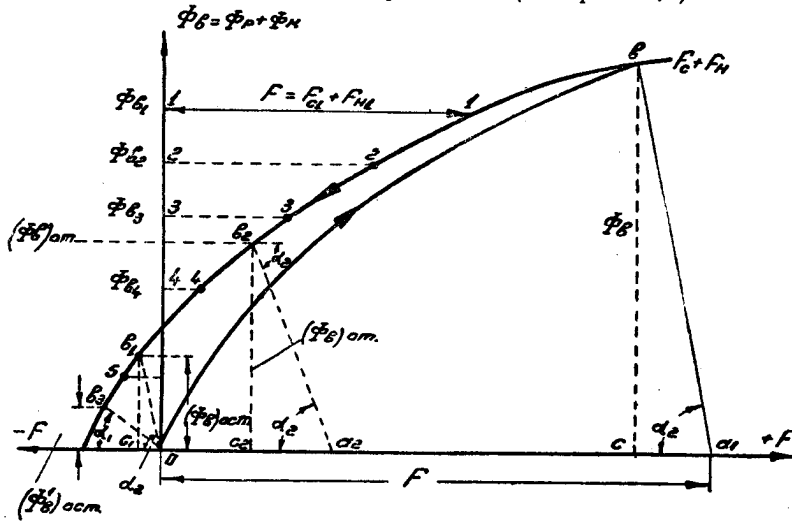
2. По заданному потоку Φ_s и проводимости $G_p + G_k$ найти значение требуемой н. с. катушки.

Из точки b кривой намагничивания, соответствующей значению Φ_s , проводят прямую ba под углом α , причем $\operatorname{tg} \alpha = \frac{G_p + G_k}{p}$. Полученный отрезок oa соответствует F катушки.

Вариант 2. Построение кривой $\Phi_p + \Phi_k = f(F_c + F_n + F_p)$. Построение проводим для магнитной системы, приведенной на фигуре 3, 26.

чивания. На фигуре 3,28 приведена кривая размагничивания. Задаются несколькими начальными значениями потока: Φ_{01}, Φ_{02} (точки 1, 2, 3, 4 ... кривой). Для точки 1 величина $F_c + F_n$ определяется отрезком 1—1.

При вычислении F_c для каждой части магнитопровода применяют кривые размагничивания для данного материала. Для каждого материала приводится ряд кривых, вид которых зависит от предельной величины индукции в стали, получающейся при намагничивании в прямом направлении (см. фиг. 1,4).



Фиг. 3, 28. Кривая размагничивания. Вариант $\Phi_0 = f_3 (F_c + F_n)$.

Следует обратить внимание на то, что для некоторых значений индукции в стали, когда кривая размагничивания материала располагается во втором квадранте, получается отрицательное значение H . Отрицательное значение n . с. Hl вычитают из величины n . с., необходимой для проведения потока в рабочих зазорах.

Все данные расчета сводят в таблицу, аналогичную таблице 18. Пользуясь кривой размагничивания, изображенной на фигуре 3,28, можно решить следующие задачи.

1. Найти n . с., при которой отпадает якорь. Задаваясь потоком $(\Phi_0)_{om}$, проводят прямую b_2a_2 под углом α_2 . Напомним, что в данном случае при построении кривой размагничивания угол α_2 определяется проводимостью рабочего зазора при замкнутом состоянии якоря. Отрезок oa_2 дает требуемое значение F_{om} .

2. Найти величину остаточного потока $(\Phi_0)_{ocm}$ в магнитной системе в сл. чае, если якорь не отпадает при $F = 0$. Для этого проводят прямую ob_1 под углом α_2 ; отрезок b_1c_1 дает значение $(\Phi_0)_{ocm}$.

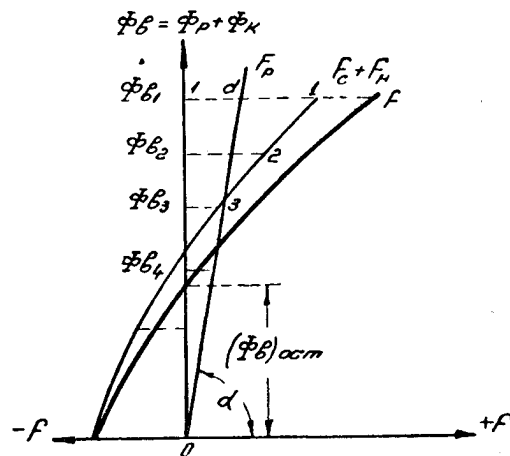
Таблица 18

Данные для построения кривой намагничивания

Точки	$\Phi_p + \Phi_k$, вб	Деталь	Φ_e , вб	Сече- ние, см ²	$B = \frac{\Phi}{q}$	$\frac{H, \text{ аэ}}{\text{см}}$	$F, \text{ аэ.}$	$\Sigma F, \text{ аэ}$
1	$\Phi_{p1} + \Phi_{k1} = \Phi_{e1}$	1	Φ_{e1}	q_1	B_1	H_1	$H_1 l_1$	ΣF_1
		2	Φ_{e1}	q_2	B_2	H_2	$H_2 l_2$	
		3	$\frac{\Phi_{e1}}{\sigma \Phi_{e1}}$	q_3	$\frac{B_3}{B_{s1}}$	$\frac{H_3}{H_{s1}}$	$\frac{H_3 + H_{s1}}{2} l_3$	
		4	$\sigma \Phi_{e1}$	q_4	B_4	H_4	$H_4 l_4$	
		5	$\frac{\sigma \Phi_{e1}}{\Phi_{e1}}$	q_5	$\frac{B_{s1}}{B_5}$	$\frac{H_{s1}}{H_5}$	$\frac{H_{s1} + H_5}{2} l_5$	
		нераб. зазор 2	Φ_{e1}				$\frac{\Phi_{e1}}{G_2}$	
		3	Φ_{e1}				Φ_{e1}/G_3	
		4	$\sigma \Phi_{e1}$				$\sigma \Phi_{e1}/G_4$	
2	Φ_{e2}	1	Φ_{e2}	q_1	B_{12}	H_{12}	$H_{12} l_1$	

и так далее

3. Найти величину остаточного потока в магнитной системе



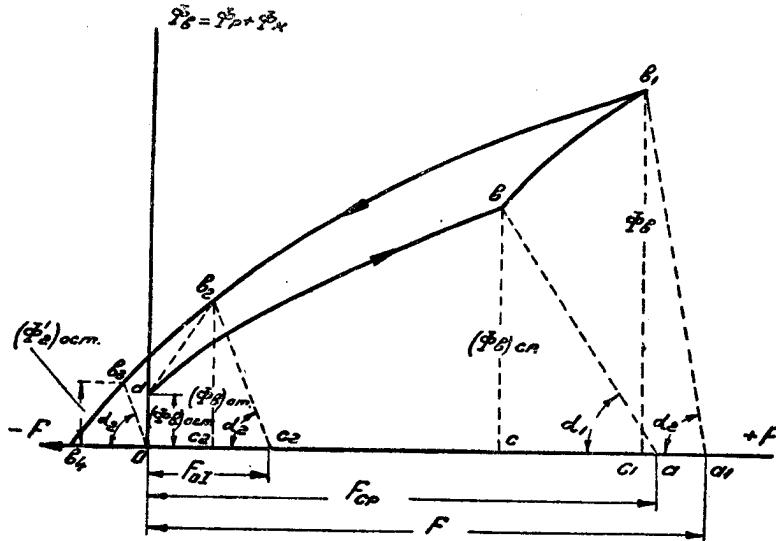
Фиг. 3, 29. Кривая размагничивания. Вариант $\Phi_e = f_s(F)$.

в случае, когда якорь отпадает. Для этого проводят прямую ob_3 под углом α_1 , соответствующим проводимости рабочего зазора при опущенном якоре.

Вариант 2. Строят кривую $(\Phi_p + \Phi_k) = f(F_c + F_n)$ и проводят прямую od под углом α (фиг. 3, 29). Для получения результирующей кривой суммируют для каждой выбранной точки требуемые значения F_p и $F_n \div F_c$.

3-11. КРИВЫЕ ПОЛНОГО ЦИКЛА РАБОТЫ ЭЛЕКТРОМАГНИТА

На фигуре 3, 30 приведены кривые полного цикла работы электромагнита. В магнитной системе при отключенной катушке и опущенном якоре имеется остаточный поток $(\Phi_e)_{ост}$. Кривая db соответствует намагничиванию системы при разомкнутом якоре.



Фиг. 3, 30. Кривые полного цикла работы электромагнита.

При увеличении н. с. до значения F_{cp} (отрезок oa) поток достигает величины $(\Phi_e)_{cp}$, при которой начинается движение якоря.

Отрезок $b-b_1$ кривой намагничивания определяется изменением потока при уменьшении зазора и увеличении н. с. катушки до значения F (отрезок oa_1).

Угол α_1 соответствует проводимости $G_p + G_k$ при опущенном якоре, угол α_2 — при притянutom. При снижении н. с. катушки поток спадает по кривой b_1b_2 . Точке b_2 соответствует значение потока $(\Phi_e)_{от}$, при котором якорь отпадает. В дальнейшем, при снижении н. с. до нуля размагничивание происходит по линии b_2d . Если якорь системы не отпадает, спадение потока происходит по кривой $b_1-b_2-b_3$.

3-12. РАСПРЕДЕЛЕНИЕ ПОТОКОВ В МАГНИТОПРОВОДЕ

Пользуясь кривыми намагничивания, можно решить задачу определения значений потоков в разных ветвях замкнутого магнитопровода. Рассмотрим два примера.

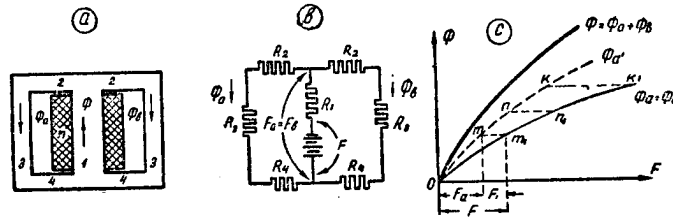
Магнитная система с симметричными ветвями (фиг. 3, 31). Такие системы применяются при исполнении различного рода дросселей

и вспомогательных трансформаторов. Магнитопровод выполняется шихтованным так, что в системе практически отсутствуют воздушные зазоры.

Согласно схеме замещения $F_a = F_\sigma = F - \Phi R_1$, или

$$F = \Phi R_1 + \Phi_a (R_2 + R_3 + R_4).$$

Для решения задачи строят кривую намагничивания. Вначале строят кривую $\Phi'_a = f(F_a)$, для чего задаются рядом значений потока Φ_a , определяют индукции в частях 2, 3, 4 и по кривой



Фиг. 3, 31. Магнитная система и кривые ее намагничивания.

намагничивания для данного материала находят значения напряженности поля. Тогда

$$F_a = H_2 l_2 + H_3 l_3 + H_4 l_4.$$

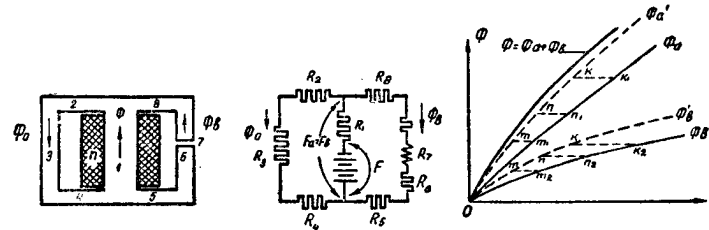
В каждой из полученных точек (m, n, k) добавляют значение

$$F_1 = H_1 l_1.$$

Величина H_1 соответствует напряженности поля при индукции $B_1 = \frac{2\Phi_1}{q_1}$.

Полученные точки $m_1, n_1 \dots$ соответствуют окончательной кривой для потоков Φ_a и Φ_σ .

Магнитная система с несимметричными ветвями. Такие системы (фиг. 3,32) применяются в тех случаях, когда желательно иметь



Фиг. 3, 32. Несимметричная магнитная система.

практически постоянное значение потока в одной ветви — Φ_σ , независящее от изменения н. с. Это достигается введением небольшого воздушного зазора. Решение задачи осуществляется аналогично предыдущему.

Строят кривые $\Phi'_a = f_1(F_a)$; $\Phi'_s = f_2(F_s)$, задавая ряд произвольных значений этих потоков.

В точках m, n, k добавляют значение $F_1 = H_1 l$. При нахождении H_1 учитывают, что

$$\Phi'_1 = \Phi'_{a1} + \Phi'_{s1}; \quad B = \frac{\Phi'_{a1} + \Phi'_{s1}}{q_1}.$$

Через полученные точки $m_1 n_1 \dots m_2 n_2$ проводят кривые, которые соответствуют потокам Φ_a и Φ_s . Суммируя ординаты этих потоков, получают точки кривой Φ .

3-13. МЕТОД ПОСЛЕДОВАТЕЛЬНЫХ ПРИБЛИЖЕНИЙ

При расчете магнитной системы в случае наличия в магнитопроводе относительно большого воздушного зазора значение потока в этом зазоре можно вычислить методом последовательных приближений, избегая таким путем применения более громоздкого графического метода.

Сущность метода постепенных приближений заключается в следующем. Вначале определяют значение магнитного потока, не учитывая падения н. с. в стали; далее находят значение индукции в стали и падение н. с. F_c . После этого корректируют полученные предварительные данные. Для пояснения сказанного рассмотрим пример.

Пример. Требуется определить значение потока в рабочем зазоре. Данные системы: $F = 2000$ ав; зазор: $q = 2$ см², $\delta = 0,25$ см; магнитопровод: железо Армо, $q = 2$ см², $l = 50$ см.

Первое приближение.

$$\Phi = FG = 2000 \cdot 1,25 \cdot 10^{-8} \frac{2}{0,25} = 2 \cdot 10^{-4} \text{ вебер.}$$

Корректировка: индукция в стали $B = \frac{2 \cdot 10^{-4}}{2} = 1 \cdot 10^{-4} \frac{\text{вб}}{\text{см}^2}$. Железо

Армо: $H = 2,4 \frac{\text{а}}{\text{см}}$.

$$F_c = 2,4 \cdot 50 = 120 \text{ ав.}$$

Таким образом, для потока $\Phi = 2 \cdot 10^{-4}$ требуется 2120 ав.

Второе приближение.

Уменьшают значение потока до $\Phi = 1,9 \cdot 10^{-4}$ вб.

$$B_c = 0,95 \cdot 10^{-4}; \quad H_c = 2,1; \quad F_c = 2,1 \cdot 50 = 105.$$

$$F_s = \frac{\Phi \delta}{\mu_0 q} = \frac{1,9 \cdot 10^{-4} \cdot 0,25}{1,25 \cdot 10^{-8} \cdot 2} = 1900.$$

$$F = F_c + F_s = 105 + 1900 = 2005 \text{ ав.}$$

Таким образом, при значении $F = 2000$ ав рабочий поток $\Phi \approx 1,9 \cdot 10^{-4}$ вб.

ГЛАВА 4

РАСЧЕТ ЦЕПЕЙ С ПОСТОЯННЫМИ МАГНИТАМИ

4.1. МАТЕРИАЛЫ ДЛЯ ПОСТОЯННЫХ МАГНИТОВ

Постоянные магниты применяются в разнообразных устройствах и приборах: реле, электроизмерительных приборах, электрических машинах; постоянные магниты применяются также для построения компасов, тахометров, спидометров, магнитных сепараторов и других устройств.

Разнообразие требований, предъявляемых к магнитам в указанных устройствах, обусловило создание различных высококоэрцитивных сплавов, отличающихся по составу, технологическим и магнитным свойствам.

В СССР проведены многочисленные работы по созданию новых сплавов для магнитов и исследованию их свойств. Известны работы А. С. Займовского, Б. Г. Лившица, Я. М. Довгалевского и других.

Различают следующие основные группы сплавов для постоянных магнитов:

- 1) мартенситные стали;
- 2) сплавы на основе железо — никель — алюминий, пластически недеформируемые;
- 3) сплавы пластически деформируемые;
- 4) сплавы на основе серебро или платина.

1. Мартенситные стали (см. таблицу 19). Лучшие магнитные свойства получают при структурном состоянии сталей, называемом дисперсионным твердением мартенсита, т. е. металл должен состоять из мартенсита с включением высокодисперсных карбидов:

а) углеродистая сталь (лучшими свойствами обладает сталь марки У-10, применяемая для магнитов малых размеров, неотвественного назначения); б) вольфрамовая сталь; в) хромистая сталь, более дешевая и менее дефицитная, чем вольфрамовая; г) кобальтовая сталь.

Таблица 19

Мартенситные стали

Марки	Содержание добавок к железу, %					$\frac{B_r}{\text{вб}} \frac{\text{вб}}{\text{см}^2}$	$\frac{H_c}{a} \frac{a}{\text{см}}$	$\frac{\gamma}{g} \frac{g}{\text{см}^3}$	$\frac{\rho}{\text{ом} \cdot \text{мм}^2}{m}$
	С	Cr	W	Co	Mo				
ЕХЗ	1,0	3,3	—	—	—	$0,95 \cdot 10^{-4}$	48	7,8	0,29
ЕВ6	0,75	0,5	6	—	—	$1,0 \cdot 10^{-4}$	44	8,1	0,30
ЕК5	1,0	6	—	6	—	$0,8 \cdot 10^{-4}$	72		
ЕК15	1,0	9	—	15	1,5	$0,75 \cdot 10^{-4}$	120		
ЕК30	0,8	5,5	6	30	—	$0,75 \cdot 10^{-4}$	160	8,1	0,27

Перечисленные виды сталей поддаются механической обработке, ковке, обработке резцом. После механической магниты подвергают

термической обработке: нормализации, подогреву, закалке; режим термической обработки должен быть строго выдержан, в противном случае получаются магниты с меньшей магнитной энергией.

Стали обладают склонностью к структурному старению: коэрцитивная сила магнита со временем снижается; во избежание старения во время эксплуатации применяются процессы структурной стабилизации.

Мартенситные стали редко применяются в современных конструкциях реле, так как углеродистые, вольфрамовые и хромистые стали обладают относительно небольшим значением магнитной энергии, а кобальтовые дороги.

2. Сплавы на основе железо—никель—алюминий (см. таблицу 20). В современных конструкциях реле применяют почти исключительно только указанные сплавы вследствие их относительно невысокой стоимости при большом значении относительной величины

Сплавы на основе Fe — Ni — Al

Таблица 20

Марка	Содержание добавок к железу, %					$\frac{B_r}{\text{вб}} \frac{\text{см}^2}{\text{см}^2}$	$\frac{H_c}{\text{а}} \frac{\text{см}}{\text{см}}$	$\frac{\gamma}{\text{г}} \frac{\text{см}^3}{\text{см}^3}$	μ_r при B_0	Способ изготовления	
	Ni	Al	Co	Cu	Si						
АН1 (ални 1)	22	11	—	—	—	$0,7 \cdot 10^{-4}$	200			Литье	
АН2 (ални 2)	24,5	13	—	3,5	—	$0,6 \cdot 10^{-4}$	345	6,9			
АН3 (ални 3)	23,5	15,5	—	4	—	$0,5 \cdot 10^{-4}$	400				
АНК (алниси)	33	13,5	—	—	1	$0,4 \cdot 10^{-4}$	600				
АНКО1 алнико 12	18	10	12	6	—	$0,68 \cdot 10^{-4}$	400	7,1		Металло- керамика Металло- керамика	
АНКО2 алнико 15	20	9	15	4	—	$0,75 \cdot 10^{-4}$	480	7,3			
АНКО3 алнико 18	19	10	18	3	—	$0,9 \cdot 10^{-4}$	520				
АНКО4 магнико *	13,5	9	24	3	—	$12 \cdot 10^{-4}$	400	7,3			
Ални12—23	23	12	—	—	—	$0,6 \cdot 10^{-4}$	280	6,7			
Ални13—25	25	13	—	—	—	$0,5 \cdot 10^{-4}$	400	6,6			
Алнико 9—20—15	20	9	15	4	—	$0,65 \cdot 10^{-4}$	465	6,8			
Магнико Ални— 15—24	14	6	24	3	—	$1,1 \cdot 10^{-4}$	360	7,1	6		
Алнико 9—20—15	24	15	—	4	—	$0,33 \cdot 10^{-4}$	360	5,3			Металло- пластика
Алнико 9—20—15	20	9	15	—	—	$0,44 \cdot 10^{-4}$	400	5,7			

* Примечание. Для магнитов АНКО3 и АНКО4 указанные магнитные свойства относятся к направлению приложенного магнитного поля при термомангнитной обработке.

магнитной энергии. Для получения различных магнитных свойств к сплавам добавляют дополнительно кремний, медь, кобальт.

Указанные сплавы нековки, обладают высокой твердостью и детали из них выполняются поэтому литыми с последующей обработкой требуемых поверхностей шлифовкой.

Готовые изделия перед намагничиванием должны подвергаться термической обработке. По данным ГОСТ 4402—48 предусматриваются следующие режимы: для сплава АН1 — закалка при 1200° в кипящей воде, АН2 — такая же закалка и отпуск при 550°, АН3 — нормализация при 1000°, АНК — при 1200°, АНКО1 — 1250°; АНКО2 — 1300°; АНКО3 — закалка при 1300° в магнитном поле (скорость охлаждения $5 \frac{\text{град}}{\text{сек}}$ до 500°) и отпуск при 600°.

В последние годы разработаны новые процессы изготовления постоянных магнитов на основе железо — никель — алюминий.

а) Металлокерамический магнит — порошки из соответствующих металлов или лигатур спрессовываются в прессформах под высоким давлением (порядка $10 \frac{\text{Т}}{\text{см}^2}$) и затем спекаются при определенной температуре (1100 ÷ 1300°). Металлокерамические магниты обладают по сравнению с литыми большей механической прочностью, спрессованное изделие не требует дополнительной механической обработки. Магнитная энергия таких магнитов меньше, чем у литых (примерно на 10%), что объясняется наличием пор в спрессованной детали.

б) Металлопластический магнит — порошок, полученный из литых заготовок сплавов ални или алнико, смешивается с твердым диэлектриком, например с бакелитной смолой, и прессуется под давлением. Полученные детали нагревают до 120° и выдерживают при такой температуре в течение нескольких часов для полимеризации смолы.

Магнитная энергия прессмагнита составляет всего 40% магнитной энергии литого по объему и 50% — по весу. Механическая прочность прессмагнитов значительно выше, чем литых. Так, предел прочности на разрыв составляет: у литых $3-6 \frac{\text{кг}}{\text{мм}^2}$, у прессмагнитов — $12 \frac{\text{кг}}{\text{мм}^2}$. Размеры прессмагнитов получаются точными.

3. Пластически деформируемые сплавы. Необходимость получения лент и проволоки для выполнения подвесных магнитов электроизмерительных приборов, компасных стрелок, спидометров и других устройств привела к созданию высококоэрцитивных сплавов, допускающих механическую обработку ковкой, волочением, резцовую.

К таким сплавам относятся:

викаллой I — $52\text{Co} + 9,5\text{V} + \text{Fe}$; $B_r = 0,9 \cdot 10^{-4} \frac{\text{вб}}{\text{см}^2}$;

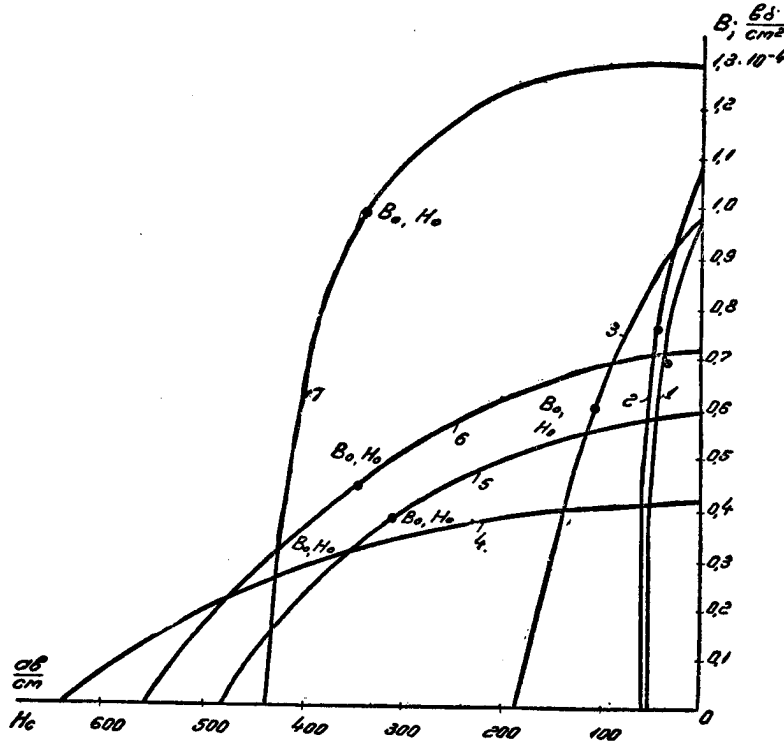
$H_e = 240 \frac{\text{а}}{\text{см}}$;

виккаллоу II — $52\text{Co} + 13\text{V} + \text{Fe}$; $B_r = 0,9 \cdot 10^{-4}$; $H_c = 320$;
 кунифе — сплав $\text{Fe} - \text{Cu} - \text{Ni}$;
 кунико — сплав $\text{Cu} - \text{Ni} - \text{Co}$.

4. Сплавы на основе — серебро или платина:

силманал — $\text{Ag} - \text{Mn} - \text{Al}$; $B_r = 0,5 \cdot 10^{-4}$; $H_c = 4800$;
 сплав $\text{Pt} - \text{Co}$; $B_r = 4 \cdot 10^{-4}$; $H_c = 3200$.

Перечисленные сплавы применяются только в малогабаритных измерительных приборах, магнитометрах, вибрационных приборах.



Фиг. 4, 1. Кривые размагничивания магнитотвердых материалов:
 1 — хромистая сталь; 2 — вольфрамовая сталь; 3 — кобальтовая сталь; 4 — альниси;
 5 — альни; 6 — альнико; 7 — магнитко.

На фигуре 4, 1 изображены кривые размагничивания разных материалов. Приведенные в таблицах 19 и 20 значения для B_r и H_c относятся к максимальным возможным в данном материале.

Практически, учитывая отклонения в технологии изготовления магнита и в процессе намагничивания, получают несколько сниженные (на 10—15%) значения параметров магнита.

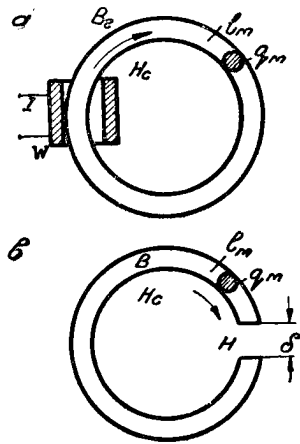
Намагничивание магнита следует производить так, чтобы получить индукцию насыщения B_s . Следует учесть, что магнитная проницаемость материалов для постоянных магнитов низка и

поэтому при намагничивании магнитов необходимо иметь источник с большим значением напряженности поля. Так, для сталей требуется для намагничивания до индукции насыщения значение $H = 450 \div 500 \frac{a}{см}$; для высококоэрцитивных сплавов значение H принимают равным $6H_c$.

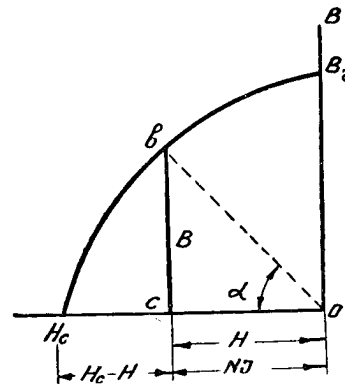
Вопросы крепления и отделки постоянных магнитов рассматриваются в главе I.

4-2. МАГНИТНЫЕ ПОТОКИ ОТДЕЛЬНОГО МАГНИТА

а) **Характеристические данные постоянных магнитов.** Постоянный магнит представляет собой кусок стали или какого-либо другого твердого сплава, который, будучи намагничен, устойчиво сохраняет запасенную часть магнитной энергии. Назначение постоянных магнитов — служить источником магнитного поля, не меняющимся заметно ни со временем, ни под влиянием таких факторов, как сотрясения, изменения температуры, наличие внешних магнитных полей.



Фиг. 4, 2. Кольцевой магнит; а — замкнутый; б — разрезанный.



Фиг. 4, 3. Кривая размагничивания.

В рестроении постоянные магниты применяются в качестве элементов поляризованных магнитных систем и магнито-электрических реле или в качестве тормозных магнитов индукционных реле.

Для определения магнитных характеристик готовых магнитов пользуются кривой размагничивания, снятой для материала, из которого изготовлен магнит. Кривая эта снимается для кольца (фиг. 4, 2, а), которое намагничивается вначале до индукции насыщения B_s . Характерными точками кривой являются: остаточная индукция в кольце B_r и коэрцитивная сила H_c (фиг. 4-3).

Для использования энергии магнита необходимо изготовить его с воздушным зазором. При разрезании кольца и образовании в нем зазора δ (фиг. 4, 2, *b*) часть н. с. магнита затрачивается на проведение потока в этом зазоре, вследствие чего индукция в магните падает от B_r до B (фиг. 4, 3). Составляющая н. с., затрачиваемая постоянным магнитом на проведение потока в воздушном зазоре, называется свободной н. с. Явление снижения индукции магнита в данном случае аналогично явлению, которое имело бы место, если бы пропустить в катушке, надетой на тороид, размагничивающий ток обратного знака, пропорциональный величине H . Величина этой размагничивающей силы H , а также значение индукции B вычисляются на основании следующих соображений:

1) при замкнутом кольце вся н. с. магнита $F = H l_m$, индукция в магните B_r ;

2) при разрезании кольца часть н. с. магнита теряется в зазоре δ :

$$F = F_c + F_g,$$

где F_g — составляющая н. с. магнита, приходящаяся на воздушный зазор.

Из предыдущего известно, что

$$F_g \cdot G_g = \Phi_g \quad \text{или} \quad H l_m \cdot G_g = \Phi_g,$$

где G_g — магнитная проводимость воздушного зазора.

Предположим сначала, что в системе магнита нет потоков рассеяния. В этом случае можно считать, что весь поток магнита проходит через воздушный зазор. Тогда

$$B q_m = H l_m G_g;$$

или

$$\frac{B}{H} = \frac{l_m}{q_m} G_g. \quad (4-1)$$

С другой стороны, на основании диаграммы, приведенной на фигуре 4, 3,

$$\frac{B}{H} = \operatorname{tg} \alpha.$$

Таким образом,

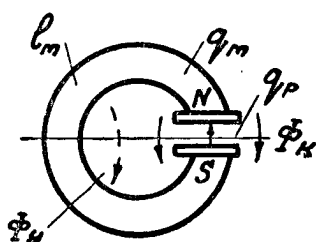
$$\frac{l_m}{q_m} G_g = \operatorname{tg} \alpha, \quad (4-2)$$

где $G_g = \mu_0 \frac{q_g}{\delta}$ — магнитная проводимость воздушного зазора.

Имея данные о материале магнита (в виде кривой размагничивания), размеры магнита l_m и q_m , размеры зазора q_g и δ можно вычислить, пользуясь (4-2), величину потока в зазоре. Для этого следует провести на диаграмме (фиг. 4, 3) прямую ob под углом α . Отрезок bc дает величину индукции B магнита. Поток в воздушном зазоре равен $\Phi_g = B q_m$.

В дальнейших расчетах будут приняты во внимание потоки рассеяния, значением которых в предварительных рассуждениях мы пренебрегли.

б) Учет потока рассеяния магнита. Рассмотрим более общий случай (фиг. 4, 4), когда площадь сечения рабочего зазора $q_p \neq q_m$ не равен площади сечения магнита. Далее будем учитывать потоки: рабочий, в зазоре, Φ_p , краевой Φ_k и утечки Φ_y .



Фиг. 4, 4. Эскиз магнита с полюсными башмаками.

Тогда коэффициент рассеяния

$$\sigma = \frac{\Phi_p + \Phi_k + \Phi_y}{\Phi_p} = \frac{G_p + G_k + G_{yn}}{G_p}. \quad (4-3)$$

Определение Φ_p производится следующим образом (фиг. 4, 3): проводят прямую ab под углом α , причем

$$\operatorname{tg} \alpha = \frac{G_p \sigma l_m}{p q_m}, \quad (4-4)$$

где p — коэффициент масштабов B и H .

Полученное значение B характеризует индукцию в среднем сечении магнита.

Поток в среднем сечении магнита $\Phi_m = Bq_m$;

Поток, рабочий, в зазоре

$$\Phi_p = \frac{Bq_m}{\sigma}; \quad (4-5)$$

Индукция в рабочем зазоре $B_p = \frac{\Phi_p}{q_p}$.

в) Прямолинейные магниты. Выражение (4-4) дает решение задачи для магнита замкнутой, например подковообразной формы, где проводимости G могут быть вычислены с достаточной для практических целей точностью.

В настоящее время широко применяются постоянные магниты из высококоэрцитивных сплавов: альни, альниси и других, выполняемых, как правило, в виде прямых или немного изогнутых стержней прямоугольного или круглого сечения. Задача вычисления проводимостей в воздухе для этих магнитов весьма трудна, и в таких случаях вопрос об определении значения потока магнита решается иначе, чем было показано выше.

Размагничивающее действие воздушных зазоров, определяемое на фигуре 4,3 отрезком $oc = H$, пропорционально интенсивности намагничивания J :

$$H = NJ,$$

где N — коэффициент размагничивания, зависящий от размеров магнита и данных воздушного зазора. Величина $N = 0$ в образце, имеющем форму замкнутого кольца.

Опытным путем найдены значения N в функции отношения длины стержня к его диаметру: l_m/d_m ; этими же данными для N

пользуются и при расчете магнитов прямоугольного сечения с площадью q_m , определяя диаметр d_m эквивалентного круглого сечения.

$$d_m = \sqrt{\frac{4q_m}{\pi}}. \quad (4-6)$$

Из диаграммы фигуры 4,3 имеем

$$\operatorname{tg} \alpha = \frac{B}{H} = \frac{B}{NJ} \cong \frac{4\pi J}{NJ} = \frac{4\pi}{N} (a).$$

Таким образом, зная N (таблица 21) можно определить угол α . Полученное после проведения прямой ab значение B относится к среднему сечению магнита.

Значения N

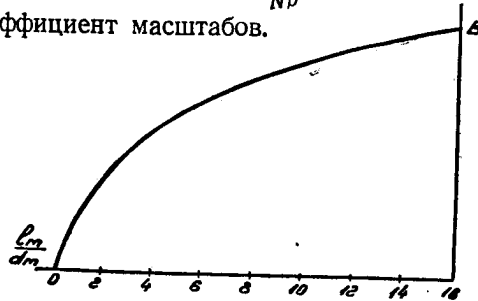
Таблица 21

l_m/d_m	1,35	1,4	1,5	1,6	1,7	1,85	2,0	2,1	2,25	2,4	2,7	3,0	3,4	3,8	4,3	5,0	6,0	7,5	10	16
N	2,0	1,9	1,8	1,7	1,6	1,5	1,4	1,3	1,2	1,1	1,0	0,9	0,8	0,7	0,6	0,5	0,4	0,3	0,2	0,1

Для практического пользования таблицей выражение (a) преобразовывается следующим образом:

$$\operatorname{tg} \alpha = \frac{10 \cdot 12,56}{Nr}, \quad (4-7)$$

здесь r — коэффициент масштабов.



Фиг. 4, 5. Кривая $B = f\left(\frac{l_m}{d_m}\right)$;

Пример: $l_m/d_m = 2,0$; $N = 1,4$.

Масштабы:

$$B = 1 \cdot 10^{-4} \frac{86}{\text{см}^2} \text{ в } 1 \text{ см}; \quad H = 100 \frac{a}{\text{см}} \text{ в } 1 \text{ см};$$

$$P = \frac{1 \cdot 10^{-4} \cdot 10^8}{100} = 100; \quad \operatorname{tg} \alpha = \frac{10 \cdot 12,56}{100 \cdot 1,4} \cong 0,9;$$

$$\alpha = 42^\circ.$$

Пользуясь соотношением (4-7), можно построить кривую $B = f\left[\frac{l_m}{d_m}\right]$, удобную для расчета. На фигуре 4, 5 построена примерная кривая для определенного материала.

Явление старения определяется рядом причин.

а) **Структурное старение.** Материал магнита после закалки или отливки имеет неравномерную структуру, но со временем эта неравномерность переходит в более равновесное — стабильное состояние. При этом в металле исчезают внутренние напряжения. Одновременно уменьшаются значения B_r и H_c .

б) **Механическое старение** происходит вследствие ударов, толчков, вибраций и влияния высоких температур, которые ослабляют поток магнита.

в) **Магнитное старение** определяется влиянием внешних магнитных полей.

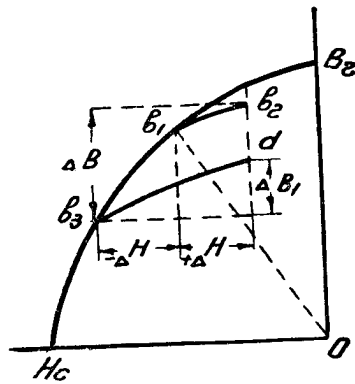
Всякий магнит перед установкой его в приборе должен быть подвергнут дополнительному процессу — стабилизации, после которого увеличивается сопротивляемость магнита уменьшению потока.

Структурная стабилизация заключается в дополнительной термической обработке, которая проводится до намагничивания магнита. Так, например, процесс стабилизации магнитных сталей заключается в кипячении закаленного магнита в воде в течение 4 часов. Такое кипячение приводит к некоторому отпуску и к удалению закалочных напряжений, сплавы на основе Fe—Ni—Al не требуют структурной стабилизации.

Механическая стабилизация заключается в том, что намагниченный магнит подвергается перед установкой в приборе ударам, толчкам, вибрациям при условиях, соответствующих режиму работы прибора в установке.

Магнитная стабилизация заключается в том, что намагниченный магнит подвергают действию внешних полей переменного знака после чего магнит становится более устойчивым к воздействию внешних полей, к температурным и механическим воздействиям. Сущность явления магнитной стабилизации заключается в следующем.

Допустим, что в намагниченном магните точка b_1 соответствует индукции B_1 в среднем сечении магнита (фиг. 4, 8). По



Фиг. 4, 8. Диаграмма, поясняющая действие магнитной стабилизации.

действием внешнего поля, напряженностью $+\Delta H$, точка b_1 переходит по кривой возврата в b_2 , при поле $-\Delta H$ — точка b_1 опускается до b_3 . Таким образом, индукция B меняется при $\pm \Delta H$ на величину ΔB .

Если же магнит предварительно подвергнуть действию $\pm \Delta H$, то после этого, под действием такого же поля, значение индукции будет меняться только на величину $\Delta B_1 < \Delta B$. Д.

стабилизации магнита его подвергают воздействию постепенно убывающего до нуля переменного поля, напряженность которого должна быть такого же порядка, как и поля, воздействующие на магнит в процессе его работы в приборе. При расчете постоянного магнита учитывается снижение его потока на 10—15%, происходящее в процессе стабилизации.

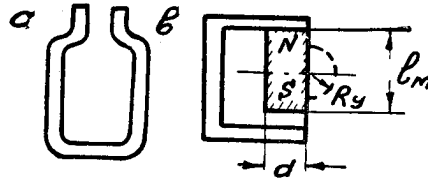
4.6. ВОПРОСЫ РАСЧЕТА МАГНИТНЫХ ЦЕПЕЙ С ПОСТОЯННЫМИ МАГНИТАМИ

Расчет магнитной цепи с постоянными магнитами сводится к решению двух задач:

- 1) в заданной системе определить величину потока в рабочем воздушном зазоре;
- 2) найти размеры постоянного магнита, который создает в рабочем зазоре определенной магнитной системы поток заданной величины.

Более трудной является вторая задача — проектирование магнита, при котором необходимо разрешить следующие вопросы: выбор материала; оптимальное решение; выбор значения коэффициента рассеяния.

Выбор материала зависит от намеченных данных рабочего зазора и конструкции магнитной системы. Рассмотрим для сравнения два материала (1 и 2 фиг. 4, б), обладающие разными значениями B_r и H_c . Для материала с меньшим значением H_c лучшие условия получаются при большем значении α , а при большем H_c — с меньшим углом α . При соблюдении этого условия получается решение, близкое к оптимальному, т. е. дающее минимальный вес магнита.



Фиг. 4, 9. Магнит:
а — из низкокоэрцитивного сплава; б — из высококоэрцитивного сплава.

Учитывая, что $\operatorname{tg} \alpha = \frac{l_m}{d} G_0$, получаем следующие практические выводы:

1) в системах, где длина магнита относительно велика, а рабочий зазор мал, лучше применять низкокоэрцитивные сплавы (материал 1);

2) в системах с большим зазором лучше применять высококоэрцитивные сплавы (материал 2).

Для одной и той же магнитной цепи можно создать условия наилучшего использования обеих групп материалов. Так, тормозной магнит счетчика или индукционного реле может быть изготовлен:

- 1) из низкокоэрцитивного сплава — хромистой или вольфрамовой стали — относительно большой длины (фиг. 4, 9 а);
- 2) из высококоэрцитивного сплава, например альни; в этом случае магнит должен быть коротким (фиг. 4, 9 б).

Оптимальное решение. Различают два вида магнитных систем с постоянными магнитами.

1) Без арматуры — к ним относят такие системы, в которых воздушный зазор магнита не меняется при встройке магнита в прибор. В качестве примера можно привести магнит для счетчика (фиг. 4, 9а).

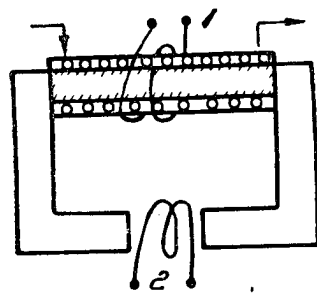
2) С арматурой, в которых воздушный зазор магнита меняется при встройке в прибор. Так, если магнит, приведенный на фигуре 4, 9б, намагничивается в собранном виде, то такую систему можно назвать системой без арматуры. Если же магнит намагничивается отдельно, а затем его прикрепляют к магнитопроводу, то получается система с арматурой.

Оптимальное решение для систем без арматуры находится просто; значительно сложнее решение этой задачи для систем с арматурой. Подробнее этот вопрос рассматривается ниже.

Выбор значения коэффициента рассеяния. В цепях с постоянными магнитами при данной величине н. с. $F = H_c I_m$ значение потока в рабочей части зазора Φ_p находится в прямой зависимости от коэффициента рассеяния σ :

$$\Phi_p = \frac{Bq_m}{\sigma}.$$

В этом отношении цепи с постоянными магнитами значительно отличаются от цепей с магнитномягкими материалами. Объясняется это тем, что магнитотвердые материалы обладают малым значением магнитной проницаемости μ , и поэтому потоки утечки создают в системе значительные падения магнитных потенциалов.



Фиг. 4, 10. Электромагнитная модель постоянного магнита.

В системе с магнитномягкими материалами при заданной н. с. F величина Φ_p в ненасыщенной системе зависит от коэффициента рассеяния в очень малой степени.

По аналогии с другим источником энергии, например электрической, цепь с постоянным магнитом можно рассматривать как генератор с большим внутренним сопротивлением, а цепь с магнитномягкими материалами — с малым. Поэтому при расчете постоянного магнита весьма важно знать точное значение σ , определение которого представляет значительные трудности. При проектировании постоянного магнита необходимо задаваться значением σ , величина которого колеблется в исполненных системах в пределах 2—5.

Для уточнения вопроса о значении σ исполняют из магнитномягкого материала модель системы таких же размеров, как и система с постоянным магнитом (фиг. 4, 10).

На часть магнитопровода, заменяющую постоянный магнит, насаживается катушка, намагничивающая сила которой выбирается так, чтобы получить необходимое значение Φ_p в рабочей части зазора. Определяют флюксметром значения потока в средней части катушки Φ и в рабочей части зазора.

Тогда

$$\sigma = \Phi / \Phi_p.$$

4-7. КОЭФФИЦИЕНТ РАССЕЯНИЯ СИСТЕМЫ С ПОСТОЯННЫМ МАГНИТОМ

При расчете систем с постоянными магнитами необходимо точно учитывать значение коэффициента рассеяния. Проводимости, рабочие и утечки, определяют различными методами, например методом разбивки поля на ряд простых фигур. Однако для некоторых форм магнитов эти методы не дают требуемой точности. Поэтому приходится весьма часто применять поправочный коэффициент β , на который умножают вычисленное значение σ . В ряде случаев для уточнения величины σ выполняют модели системы (фиг. 4, 10).

Ниже приводятся несколько примеров определения коэффициента рассеяния.

Система с подковообразным магнитом (фиг. 4, 11). Магнит M намагничивается отдельно, чаще всего с пристроенными полюсными башмаками, затем, при встройке магнита в прибор, проводимость между полюсными башмаками увеличивается благодаря наличию сердечника c . Необходимо вычислить проводимость утечки G_{yn} , краевую G_k , проводимость между башмаками G_1 — три отсутствия сердечника и G_p — при наличии его.

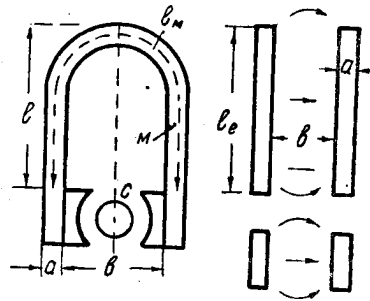
Для определения индукции в отдельном магните пользуются значением

$$G_0 = G_1 + G_{yn} + G_k.$$

После встройки магнита в прибор

$$G = G_p \cdot \sigma; \quad \sigma = \frac{G_p + G_{yn} + G_k}{G_p}.$$

Вычисление G_1 , G_p и G_k производится обычными методами. Для вычисления G_{yn} подковообразного магнита применяется существенный способ: подковообразный магнит заменяется двумя параллельными плоскостями такой длины, чтобы проводимость между ними по воздуху была равна проводимости подковообраз-



Фиг. 4, 11. Подковообразный магнит.

ного магнита. Длина рабочей части магнита $l_m = 2 \left(l_1 - \frac{b}{2} - a \right)$.
Эквивалентная длина плоскости

$$l_e = \left(l_1 - \frac{b}{2} - a \right) + 1,2(a + b). \quad (4-9)$$

Поле между плоскостями разбивают на ряд простых фигур, вычисляют значение G_y . Учитывая, что разность магнитных потенциалов изменяется равномерно по длине (это положение дает не вполне точное решение), получают $G_{yn} = G_{y/2}$.

Система с прямолинейным магнитом (фиг. 4, 9 б). Проводимость утечки по воздуху между полюсами магнита определяется следующим образом. Площадь потока равна $S = \pi d l_m$. Средняя длина силовой линии упрощенно принимается $l_c = \pi R_y = \frac{\pi l_m}{4}$, тогда

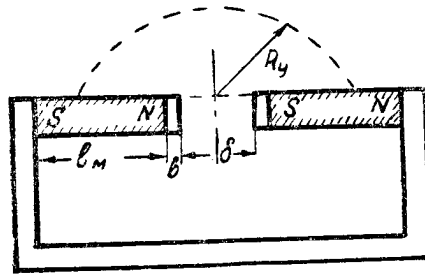
$$G_{yn} = \frac{1}{2} \mu_0 \frac{2\pi d l_m}{\pi l_m} = \mu_0 d.$$

$$\sigma = \frac{G_p + G_k + G_{yn}}{G_p}.$$

Система с двумя прямолинейными магнитами (фиг. 4, 12). Магниты снабжены двумя башмаками из магнитномягкого материала.

$$G_p = \mu_0 \frac{\pi d^2}{4\delta};$$

$$G_k = \mu_0 \left[0,26l + 4 \frac{d}{2(\delta/m + 1)} \right] = \mu_0 \left[1,63 \frac{d}{2} + 4 \frac{a}{2(\delta/b + 1)} \right].$$



Фиг. 4, 12. Система с двумя прямолинейными последовательно включенными магнитами.

Проводимость G_y определяем из следующих условий: площадь потока $S = \pi d l_m$; средняя длина силовой линии $l_c = \pi \left(\frac{l_m}{2} + b + \frac{\delta}{2} \right)$.

$$G_{yn} = \mu_0 \frac{2d l_m}{l_m + 2b + \delta}.$$

Учитывая, что н. с. определяется двойной длиной магнита: $F = H 2l_m$, при определении G_{yn} не нужно делить значение G_y на два. Численный пример:

$$\delta = 0,515 \text{ см}; \quad d = 2,56 \text{ см}; \quad b = 0,2 \text{ см}; \quad l_m = 3 \text{ см}.$$

$$\left. \begin{aligned} G_p &= \mu_0 \cdot 9,9; \\ G_k &= \mu_0 \cdot 3,5; \\ G_{yn} &= \mu_0 \cdot 3,94; \end{aligned} \right| \delta = \frac{9,9 + 3,5 + 3,94}{9,9} = 1,75.$$

Значение коэффициента рассеяния, полученное по данным измерения на модели, $\sigma_1 = 2,27$. Таким образом, в расчет надо вводить поправочный коэффициент

$$\beta = \frac{\sigma_1}{\sigma} = \frac{2,27}{1,75} = 1,3.$$

4-8. РАСЧЕТ МАГНИТНОЙ ЦЕПИ С ПОСТОЯННЫМИ МАГНИТАМИ

Предполагаются заданными: материал магнита, его размеры, размеры всей магнитной цепи. Необходимо определить индукцию в рабочем зазоре.

Ниже рассматривается два случая: а) система без арматуры — сюда будем относить случаи, когда магнит намагничивается в собранном с арматурой виде и при встройке в прибор проводимость рабочего зазора не меняется; б) система с арматурой — магнит намагничивается отдельно и при встройке в прибор меняется проводимость рабочего зазора.

В дальнейшем пренебрегаем величиной магнитного сопротивления деталей систем из магнитномягких материалов, величиной весьма малой сравнительно с сопротивлением магнита и воздушных путей.

а) Система без арматуры. Определяют значения проводимостей: G_p ; G_k ; G_{yn} ;

$$\sigma = \frac{G_p + G_k + G_{yn}}{G_p}.$$

Пользуясь диаграммой, данной на фигуре 4,3, устанавливают значение индукции B в среднем сечении магнита.

$$\operatorname{tg} \alpha = \frac{l_m}{q_m} \frac{G_p \sigma}{p}, \quad (4-4)$$

здесь p — коэффициент масштабов B и H .

Тогда

$$B_p = \frac{B q_m}{q_p \sigma}. \quad (4-10)$$

б) Система с арматурой. Определяют проводимости: 1) G_1 — сех воздушных путей для магнита в том виде, в каком производится его намагничивание. 2) G_p , G_k , G_{yn} — для собранной арматурой системы.

Коэффициент рассеяния при наличии арматуры

$$\sigma = \frac{G_p + G_k + G_{yn}}{G_p}.$$

остроения:

1) На диаграмме, приведенной на фигуре 4,7, проводят прямо под углом α_1 , причем

$$\operatorname{tg} \alpha_1 = \frac{l_m}{q_m} \frac{G_1}{p}.$$

Отрезок b_1c_1 — величина индукции B_1 магнита при отсутствии арматуры.

2) Проводят кривую возврата b_1d и прямую ob_2 под углом α_2 , причем

$$\operatorname{tg} \alpha_2 = \frac{l_m \sigma \cdot G_p}{q_m p}.$$

Отрезок b_3c_2 — величина индукции B_2 в магните в собранной системе.

$$B_p = \frac{B_2 q_m}{q_p \sigma}.$$

в) Система с арматурой и прямолинейными магнитами. Несколько иначе, чем указано в пункте б, решается задача для систем с прямолинейными магнитами.

Для примера рассмотрим систему, изображенную на фигуре 4, 12. Магниты намагничиваются отдельно. По отношению l_m/d_m определяют значение угла α_1 :

$$\operatorname{tg} \alpha_1 = \frac{12,56}{1000 N p}. \quad (4-7)$$

В точке B_1 (фиг. 4, 7) проводят кривую возврата; затем проводят прямую ob_2 под углом α_2 , причем

$$\operatorname{tg} \alpha_2 = 2 \frac{l_m \sigma \cdot G_p}{q_m p}.$$

Значение $\operatorname{tg} \alpha_2$ увеличено вдвое, так как в системе имеется два последовательно включенных постоянных магнита.

4-9. ПРОЕКТИРОВАНИЕ ПОСТОЯННЫХ МАГНИТОВ

В данном случае требуется установить материал и размер магнита по заданным параметрам прибора: проводимости рабочего зазора в собранной системе, индукции в рабочей части зазора, B_p .

а) Выбор материала. Выбирают такой материал, при котором в заданных условиях работы прибора, получается оптимальное решение, т. е. параметры магнита равны соответственно B_0 и H_0 (или близки к ним). На фигуре 4, 1 показаны характеристики разных материалов и отмечены точки B_0 и H_0 . Рассмотрим, пользуясь этими характеристиками, область применения трех наиболее ходких материалов при следующих условиях работы прибора

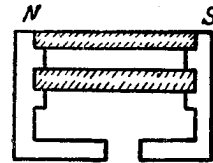
1) относительно малое значение рабочего зазора δ , большое значение B_p . В данном случае получается относительно большое значение G_p , а следовательно и угла α : лучше применить сплав альни.

2) Относительно большое значение δ , малое значение I лучше применить сплав альниси.

3) относительно большое значение δ , большое значение B_p : лучше применить сплав **магнико**.

Нами указано только общее решение задачи. В каждом отдельном случае необходимо дополнительно учитывать ряд условий: конструктивных, технологических и экономических.

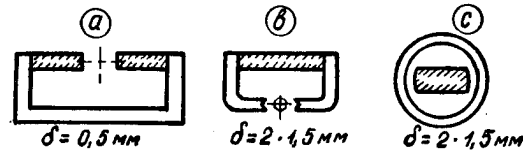
б) **Отношение l_m/q_m** . Как указывалось выше, это отношение принимают из условия получения максимального значения свободной магнитной энергии. Однако такое решение не всегда получается удачным. Так, при необходимости получения весьма больших значений потоков размеры магнита по сечению получаются громоздкими. В таких случаях выгоднее применить систему с параллельными магнитами (фиг. 4, 13). Необходимо располагать магниты на некотором расстоянии друг от друга, так как параллельные магниты размагничивают друг друга.



Фиг. 4, 13. Система с двумя параллельно включенными магнитами.

При необходимости получения высоких значений н. с. F в рабочем зазоре рациональнее применять систему с последовательно включенными магнитами (фиг. 4, 12).

в) **Магнитная цепь без арматуры**. Заданы: B_p ; q_p ; δ . Некоторая неопределенность в решении поставленной задачи получается вследствие отсутствия данных о величине коэффициента рассеяния системы. На основании опытных данных задаются значением σ , которое уточняется в проверочном расчете. Иногда приходится проводить несколько вариантов расчета.



Фиг. 4, 14. Формы систем с магнитами.

На фигуре 4, 14 приведены эскизы и данные δ для некоторых видов систем.

При значениях, указанных на эскизе зазоров δ и пределах индукции B от $(0,1 - 0,4) 10^4 \frac{об}{см^2}$, получаются следующие пределы изменения σ :

$$\begin{aligned} \sigma_a &= 1,55 \div 1,87; \\ \sigma_b &= 2,60 \div 2,88; \\ \sigma_c &= 1,75 \div 2,06. \end{aligned}$$

При увеличении δ увеличивается σ .

Размеры магнита могут быть определены из следующих двух соотношений, в которых индукция в магните и напряженность поля принимаются равными B_0 и H_0 :

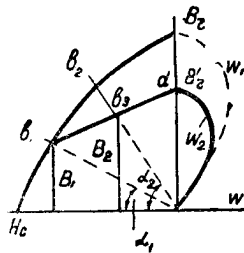
из уравнения потока

$$\Phi_m = B_0 q_m = B_p q_p \sigma; \quad q_m = \frac{B_p q_p \sigma}{B_0}; \quad (4-11)$$

из уравнения н. с.

$$H_0 l_M = \frac{B_p \delta}{\mu_0 l_M}; \quad l_M = \frac{B_p \delta}{\mu_0 H_0}. \quad (4-12)$$

г) **Магнитная цепь с арматурой.** Рассмотрим работу магнита: проводимость воздушных путей без арматуры примем равной G_1 , а с арматурой G_2 . Тогда, согласно диаграмме, данной на фигуре 4, 15, индукция магнита в собранной системе будет B_2 . Изменение проводимости G_2 , а следовательно и угла α_2 приводит в этом случае к изменению индукции по кривой возврата $b_1 d$, и максимальное значение получают равным B_2' . Кривая, характеризующая закон изменения магнитной энергии, будет уже не W_1 , а W_2 и, таким образом, условия оптимума для данного материала изменяются.



Фиг. 4, 15. Диаграмма работы магнита.

В системах с арматурой условия оптимума в определенной степени зависят от соотношения G_1 и G_2 .

Введем понятие об относительном изменении проводимости воздушных путей

$$\eta = \frac{G_2 - G_1}{G_2}. \quad (4-13)$$

Учитывая, что углы α_1 и α_2 связаны с проводимостями соотношением

$$\frac{\operatorname{tg} \alpha_2}{\operatorname{tg} \alpha_1} = \frac{G_2}{G_1},$$

получаем из (4-13)

$$\operatorname{tg} \alpha_2 = \frac{1}{1 - \eta} \operatorname{tg} \alpha_1. \quad (4-14)$$

Пользуясь введенным понятием, можно найти путем графических построений условия для оптимального соотношения размеров магнита и воздушных зазоров.

Задаваясь определенным значением η строят кривую $BH = f(\operatorname{tg} \alpha_1)$ при $\eta = \text{const}$.

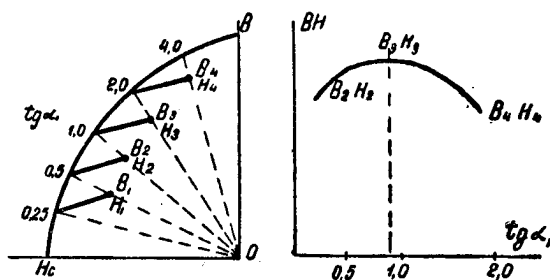
Для примера приведена кривая при $\eta = 0,5$ (фиг. 4, 16); задаемся $\operatorname{tg} \alpha_1$, равным 0,25, 0,5, 1,0 и 2,0. Соответственно получаем значения

$$\operatorname{tg} \alpha_2 = \frac{1}{1 - 0,5} \operatorname{tg} \alpha_1, \text{ т. е. } 0,5, 1,0; 2,0; 4,0.$$

Свободная магнитная энергия в каждом случае пропорциональна BH . Из кривой $BH = f(\operatorname{tg} \alpha_1)$ видно, что оптимальное решение при $\eta = 0,5$ получается в случае, когда $\operatorname{tg} \alpha_1 = 1,0$.

Таким образом, в случае, когда относительное изменение проводимостей G_2 — при наличии арматуры и G_1 — при отсутствии ее, равно $\eta = 0,5$, надо выбрать магнит так, чтобы тангенс начального угла был равен 1.

Если магнит намагничивается в замкнутом магнитопроводе, но без сердечника, например в системе, приведенной на фигуре 4, 11, то отношение l_m/q_m находят из выражения $\operatorname{tg} \alpha_1 = \frac{l_m}{q_m} G_1/p$, где G_1 — проводимость путей потока по воздуху.



Фиг. 4, 16. Кривые для определения оптимального угла.

В случае прямолинейного или слегка изогнутого магнита, намагничиваемого отдельно или слегка изогнутого магнита, намагничиваемого отдельно без сердечника или без остальных частей магнитопровода (фиг. 4, 12), отношение l_m/d_m определяется из выражения (4-7). Для данного материала магнита строится кривая $\operatorname{tg} \alpha_1 = f(\eta)$, из которой определяют значение $\operatorname{tg} \alpha_1$ при заданном значении η (фиг. 4, 17).

Основные данные и соотношения для расчета. При выбранном материале строится кривая $\operatorname{tg} \alpha_1 = f(\eta)$. Из предыдущего имеем:

$$\operatorname{tg} \alpha_1 = \frac{l_m \cdot G_1}{q_m \cdot p} \quad (4-4),$$

или

$$\operatorname{tg} \alpha_1 = \frac{12,56}{1000Np} \quad (4-7);$$

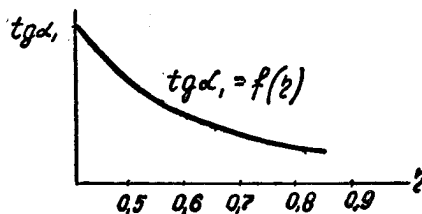
$$\operatorname{tg} \alpha_2 = \frac{1}{1-\eta} \operatorname{tg} \alpha_1 \quad (4-14);$$

$$\operatorname{tg} \alpha_2 = \frac{l_m}{q_m} \mu_0 \frac{q_p \sigma}{\delta p} \quad (4-15) \text{ для одиночного магнита;}$$

$\operatorname{tg} \alpha_2 = 2 \frac{l_m}{q_m} \mu_0 \frac{q_p \sigma}{\delta p} \quad (4-16)$ — для двух последовательно включенных магнитов.

После проведения расчета применяется следующее соотношение для проверки:

$$B_p = \frac{B q_m}{q_p \sigma}. \quad (4-10)$$



Фиг. 4, 17. Кривая зависимости $\operatorname{tg} \alpha_1$ от относительного изменения проводимости.

Ход расчета. Приведенные выше соотношения не позволяют решить задачу определения размеров магнита. Некоторыми величинами приходится поэтому задаваться в предварительном расчете, уточняя их в проверочном. После этого может возникнуть необходимость проведения других вариантов расчета. Количество требуемых вариантов расчета зависит от того, насколько удачно приняты предварительные данные. В ходе расчета заданными являются обычно величины: B_p , q_p , иногда l_m .

При расчете задаются значением σ , варьируют значение η , отношение q_m/q_p (изменением размеров полюсных башмаков), а также величиной δ — рабочего воздушного зазора.

ЧАСТЬ ВТОРАЯ
ЭЛЕКТРОМАГНИТНЫЕ СИСТЕМЫ И РЕЛЕ

ГЛАВА 5
ВОПРОСЫ ПРОЕКТИРОВАНИЯ ЭЛЕКТРОМАГНИТОВ

5-1. ЗАДАЧИ РАСЧЕТА

Электромагниты применяются в качестве приводного или управляющего устройства в целом ряде механизмов, электрических аппаратов и реле: в подъемных и тормозных электромагнитах, в приводах для включения или отключения коммутационных аппаратов, в приводах задвижек, золотников, в электромагнитных контакторах, в первичных и вторичных реле, в автоматических регуляторах и других.

Электромагнит должен иметь такую форму, при которой получается наиболее рациональное сочетание его конструкции с остальными элементами устройства: контактной системой или исполнительным механизмом; тяговая характеристика электромагнита должна соответствовать требованиям устройства. Задача расчета состоит в том, чтобы найти по заданным условиям работы механизма конструктивные размеры электромагнита и определить его параметры — данные магнитопровода и катушки, а также тяговую характеристику. Расчету электромагнита предшествует построение кинематической схемы устройства, на основании которой строятся характеристики противодействующих сил и моментов:

$$Q_n = Q(\delta) \text{ или } M_n = M(\theta).$$

Из этих характеристик получают основные расчетные данные для выбора формы и расчета электромагнита.

Процесс расчета имеет следующие стадии:

- 1) построение кинематической схемы и характеристики исполнительного механизма;
- 2) выбор формы магнитной системы;
- 3) предварительный расчет для определения основных размеров и параметров электромагнита;
- 4) окончательный расчет для уточнения данных предварительного расчета.

При проведении предварительного расчета применяют упрощенные соотношения, некоторыми значениями параметров, например индукцией B в стали, в рабочем зазоре, падением магнитного потенциала в нерабочих зазорах и стали и другими, задаются. После проведения окончательного расчета выявляется рациональность принятой формы электромагнита и принятых параметров.

Если полученное решение не является оптимальным, проводят варианты расчета с измененными параметрами. Число вариантов расчета зависит от того, насколько удачно выбраны параметры для определения габаритов системы.

5-2. ОБЩАЯ КЛАССИФИКАЦИЯ ЭЛЕКТРОМАГНИТНЫХ СИСТЕМ

1) По конструктивному исполнению системы характеризуются перемещением подвижного якоря:

системы с поворотным якорем (фиг. 8, 1), в которых якорь совершает поворотное по отношению к неподвижному полюсу движение;

системы с поперечным движением якоря (фиг. 9, 1), в которых якорь перемещается так, что в своем движении он пересекает среднюю линию, проведенную между неподвижными полюсами;

системы прямоходовые — с поступательным движением якоря (фиг. 10, 1).

Приведенные наименования систем не узаконены стандартами (ГОСТ); применяются и другие наименования: клапанные, с вращающимся якорем, соленоидные и другие.

2) По способу включения катушек и по роду тока различают два вида электромагнитных систем.

Системы $F = \text{пост.}$, в этих системах установившееся значение н. с. катушки не меняется по ходу якоря, а значение потока Φ меняется.

К системам $F = \text{пост.}$ относятся системы с токовой катушкой постоянного и переменного тока и с катушкой напряжения постоянного тока.

Системы $\psi = \text{пост.}$ В системах с катушкой напряжения переменного тока, если не учитывать активной составляющей падения напряжения, весьма малой сравнительно с индуктивной, приложенное напряжение U компенсируется противодействующей э. д. с., определяемой числом потокосцеплений $\psi : U \approx E = \frac{\omega \psi \Phi}{\sqrt{2}}$. При

заданном напряжении U и постоянной частоте $\psi = \Phi \omega$ остается постоянным по ходу якоря; значение тока I , а следовательно и F меняется по ходу якоря.

5-3. ФАКТОРЫ, ВЛИЯЮЩИЕ НА ВЫБОР ФОРМЫ ЭЛЕКТРОМАГНИТА

Ниже приводится ряд условий работы основного устройства, определяющих выбор формы электромагнита.

1. Электромагнитная характеристика должна быть такой, чтобы в любом положении якоря электромагнитная сила или момент были больше силы или момента противодействующих (механических) сил или моментов устройства.

2. Ход или угол поворота якоря. В зависимости от условий работы устройства требуется больший или меньший ход или угол поворота якоря. В быстродействующих реле ход якоря

бывает порядка $0,3 \div 1$ мм, в обыкновенных $2 \div 5$ мм, в контактах $10 \div 20$ мм, в тормозных магнитах $30 \div 150$ мм.

3. Особенности исполнительного механизма. В реле прямого действия его исполнительный механизм действует непосредственно на устройство автоматики. В реле косвенного действия исполнительным механизмом является контактная система. Существует большое разнообразие форм контактных систем, отличающихся по числу контактов, ходу контактов, по разрывной мощности. Особенности исполнительных механизмов в некоторой степени определяют форму магнитных систем: необходимо выполнять эту систему с учетом конструктивного оформления всего устройства в целом.

4. Род тока. Электромагнитные характеристики систем постоянного тока несколько отличны от характеристик систем переменного тока. Поэтому в ряде случаев выбор формы магнитной системы определяется и родом тока.

5. Коэффициент возврата, представляющий собой отношение величины тока или напряжения отпуска якоря к величине тока или напряжения втягивания ($k_a = \frac{I_0}{I_a}$; $k_a = \frac{U_0}{U_a}$), зависит от степени расхождения характеристик электромагнитной и приводящего устройства.

Требования, предъявляемые к значению k_a , различны: в одних случаях достаточен $k_a = 0,1$; в защитных реле требуется $k_a = 0,8 \div 0,85$; в регуляторах — k_a должен быть порядка 0,99.

6. Время срабатывания системы также определяет выбор формы магнитной системы. В быстродействующих реле время срабатывания должно быть порядка $0,003 \div 0,05$ сек, в обыкновенных — $0,15$ сек, в замедленных — $0,2 \div 20$ сек. Имеются устройства, где требуется еще большая выдержка времени: порядка нескольких минут.

7. Потребляемая мощность. Значение потребляемой мощности катушек реле и электромагнитов колеблется в исполненных системах от сотых долей ватт до нескольких сот ватт. Специальные требования в отношении малого потребления мощности предъявляются к вторичным реле, включаемым в цепи вторичных обмоток измерительных трансформаторов тока и напряжения. Такое же требование относится и к некоторым видам реле автоматики.

8. Степень уравновешенности подвижных частей системы. При работе устройства в условиях тряски и вибраций, например при установке на кораблях и других подвижных установках, подвижная часть магнитной системы должна быть уравновешена.

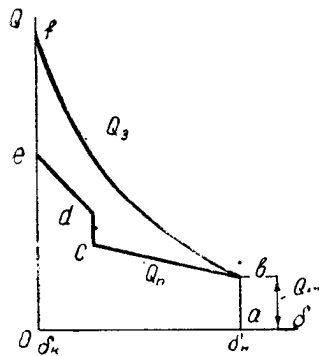
9. Производственные условия. Всякая магнитная система должна быть удобной для производства, части ее должны быть, по возможности, простыми, удобными для изготовления и сборки.

10. Эксплуатационные условия. Устройство должно быть удобно для монтажа, доступно осмотру в эксплуатации, отдельные части должны быть легко сменяемы.

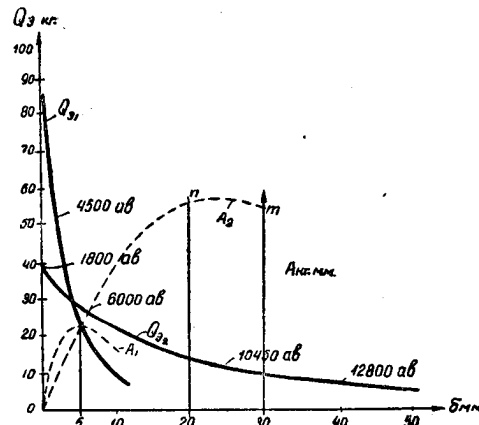
Ниже приводится ряд примеров, показывающих влияние условий работы устройства на выбор формы магнитной системы.

5.4. УСЛОВНАЯ РАБОТОСПОСОБНОСТЬ ЭЛЕКТРОМАГНИТА

Электромагнит всегда проектируется так, чтобы электромагнитные Q_3 в процессе движения якоря были больше противодей-



Фиг. 5, 1. Тяговая характеристика электромагнита.



Фиг. 5, 2. Тяговые характеристики электромагнитов с поворотным якорем:

$Q_{31} - A_1$ — постоянного тока: $F = \text{пост.}$
 $Q_{32} - A_2$ — переменного тока: $\psi = \text{пост.}$

ствующих сил Q_n (фиг. 5,1). Работа A_3 , которая может быть получена от электромагнита в процессе движения якоря, выражается

$$\text{площадью } oabfo : A_3 = \int_0^{\delta_n} Q_3 d\delta.$$

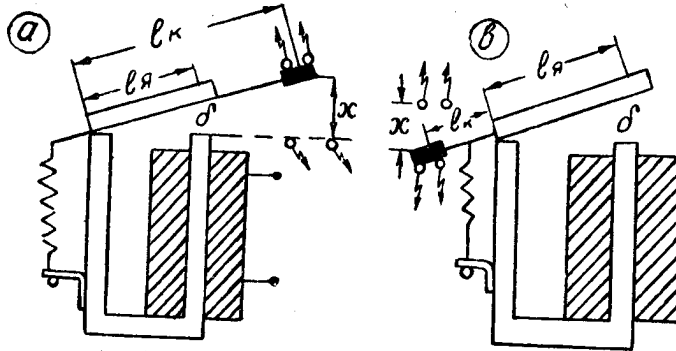
Эта работа получается большей по своему значению, чем требуемая для перемещения якоря (площадь $oabcdeo$): Поэтому электромагниты характеризуют не по полной работоспособности A_3 , а по условной A , представляющей собой произведение начальной силы Q_{3n} на ход δ_n . Для каждой формы электромагнита существуют оптимальные условия получения минимального размера и веса при определенных величинах Q_{3n} и δ_n . Значение A изменяется по ходу якоря.

Рассмотрим два примера.

1) Электромагнит с поворотным якорем.

На фигуре 5,2 приведены характеристики системы определенных размеров, но снятые при параллельном включении катушек постоянного и переменного тока.

Соответствующие кривые условной работоспособности A_1 и A_2 . Из кривых видно, что системы переменного тока обладают большим значением A_2 при относительно больших значениях δ_n , а системы постоянного тока A_1 при относительно меньших δ_n .

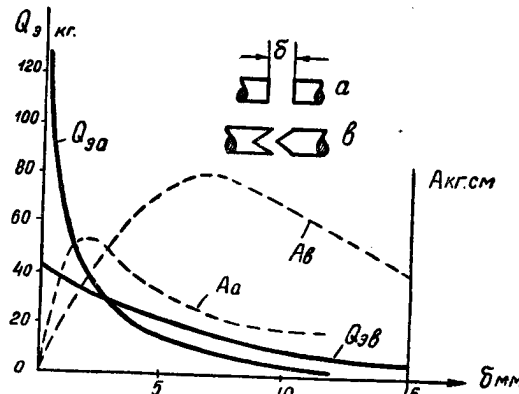


Фиг. 5, 3. Магнитные системы реле:
а — постоянного тока; б — переменного тока.

Из полученных данных следуют определенные практические выводы:

а) При построении систем постоянного тока необходимо стремиться к тому, чтобы ход якоря δ_n был минимальным при заданном ходе X контактов (фиг. 5, 3 а). Система должна быть выполнена так, чтобы $l_я$ было меньше $l_к$. Это правило не всегда выполняется. Так, в кодовом реле КДР1 по конструктивным соображениям принято $l_я > l_к$. Однако при исполнении этого же реле с меньшим потреблением энергии — тип КДР2 — принято $l_я < l_к$.

б) При построении систем переменного тока рекомендуется принимать $l_я > l_к$ (фиг. 5, 3 б).



Фиг. 5, 4. Характеристики систем с плоским и коническим стопом.

На фигуре 5, 4 приведены характеристики двух электромагнитов прямоходовой системы постоянного тока одинаковых размеров и с одними и теми же данными катушки, но выполненных с разными формами стопов: плоским и коническим. Условная работа электромагнита с коническим стопом больше при больших величинах хода, конечное значение силы и условная работа при малых величинах хода выше у системы с плоским стопом.

5-5. УДЕЛЬНЫЙ РАСХОД МАТЕРИАЛА

Для сравнительной оценки разных систем в отношении расхода материала при заданных условиях работы электромагнита вводится понятие об удельном расходе материала

$$m = \frac{\text{вес электромагнита}}{\text{условная работа}}$$

Для каждой формы электромагнита имеются определенные величины сил и хода, при которых вес рационально сконструированного электромагнита получается минимальным.

Для удобства выбора формы электромагнита вводится понятие о конструктивном факторе (к. ф.), представляющем собой отношение

$$\text{к. ф.} = \frac{\sqrt{Q_{эн}}}{\delta_n} = \frac{\sqrt{\text{силы в кг}}}{\text{ход в см.}} \quad (5-1)$$

Выбор формулы (5-1) определяется на основании следующих соображений:

- а) длина электромагнита пропорциональна требуемой величине ход — δ ;
- б) поперечное сечение стали электромагнита определяется величиной требуемой начальной силы $Q_{эн}$.

Каждой форме электромагнита соответствует определенная зона величин к. ф., при которых эта система выполняется с оптимальными данными по расходу материала.

На основании ряда произведенных расчетов получены практические данные для выбора формы электромагнита, приведенные на таблице 23.

Таблица 23

Значения к. ф. для разных систем

Форма электромагнита	Пределы значений к.ф. для систем $F = \text{пост.}$
С поворотным якорем	2,6—26
Прямоходовая с плоским стопом	90—16
» с коническим стопом	16—1,85
» без стопа	ниже 0,2

Пример. 1) Для системы $F = \text{пост.}$ $Q_{эн} = 16 \text{ кг}$, $\delta_n = 0,2 \text{ см}$;

к. ф. $= \frac{\sqrt{16}}{0,2} = 20$; лучше применить прямоходовой электромагнит с плоским стопом.

2) Для той же системы при $\delta_n = 0,5 \text{ см}$

к. ф. $= \frac{\sqrt{16}}{0,5} = 8$; лучше применить конический стоп.

5-6. ДАННЫЕ ДЛЯ ПРОЕКТИРОВАНИЯ ЭЛЕКТРОМАГНИТА

Для расчета электромагнита по заданным размерам магнитной системы пользуются основными соотношениями, принятыми для расчета электрической и магнитной цепи. В случае, когда требуется спроектировать магнитную систему, т. е. определить ее размеры и параметры катушки по заданным условиям работы устройства, необходимо задаваться некоторыми параметрами и величинами: значениями индукции в рабочих зазорах и в стали, коэффициентов рассеяния, отношением длины катушки к толщине намотки и другими.

Принятые значения перечисленных данных определяют собой экономичность конструкции, т. е. создают условия получения модели минимальных размеров и веса. Выбор таких данных основывается на результатах расчета целого ряда электромагнитов. При отсутствии соответствующих данных приходится прорабатывать ряд вариантов с целью получения оптимального решения.

Следует учитывать разницу в условиях решения задачи для систем постоянного и переменного тока. Определяется эта разница тем, что при одном и том же значении потока в рабочем зазоре сила при постоянном токе вдвое больше силы при переменном токе.

Ниже приводятся общие соображения по выбору определенных величин для проектирования магнитной системы электромагнита.

Индукция в рабочем воздушном зазоре. Согласно принятому ранее методу расчета, весь поток в рабочем воздушном зазоре разбивается на два потока: основной, называемый нами рабочим, и краевой $\Phi_p = \Phi_r + \Phi_k$.

Как будет показано в дальнейшем, основная величина силы на якоре определяется потоком Φ_r , поэтому в предварительных расчетах для их упрощения принимают во внимание действие сил, определяемых только основным, рабочим потоком.

В предварительном расчете приходится задаваться значением индукции $B_p = \frac{\Phi_p}{q}$. Оптимальное значение B_p колеблется в боль-

ших пределах: от $0,06 \cdot 10^{-4}$ до $1 \cdot 10^{-4} \frac{\text{вб}}{\text{см}^2}$. Выбор B_p определяется соотношением между заданной тяговой силой Q_n при отпущенном якоре и величиной хода: с увеличением силы необходимо принимать большее значение индукции.

Опытные и расчетные данные оптимальных значений B_p удобно систематизировать в функции к. ф.

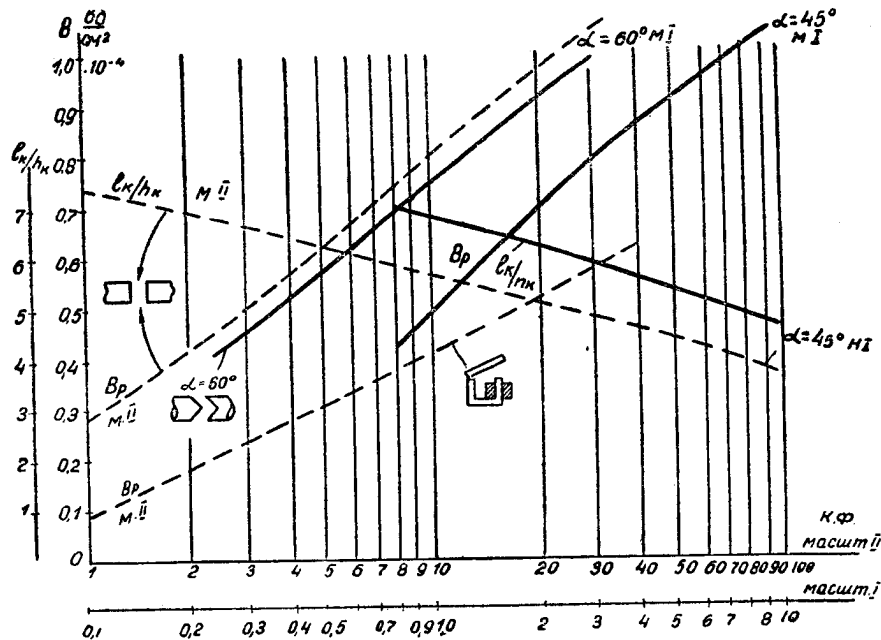
На фигуре 5, 5 приведены кривые $B_p = f(\text{к. ф.})$ для некоторых видов электромагнитов постоянного тока. Эти же кривые применимы и для систем переменного тока, при этом вводится понятие об эквивалентном значении конструктивного фактора:

$$\text{к. ф.} = \frac{\sqrt{2Q_n}}{\delta} \quad (5-2)$$

Это эквивалентное значение получается из условия: при одном и том же максимальном значении индукции B сила на якоре при постоянном токе вдвое больше силы при переменном.

Отношение длины катушки к толщине намотки.

Площадь сечения меди катушки $q_k = l_k h_k$ принимают в зависимости от требуемого значения н. с. F . Значение q_k может быть



Фиг. 5, 5. Кривые зависимости $B_p = f(\kappa, \Phi)$ для магнитных систем:
масштаб I — сплошные линии;
масштаб II — пунктирные.

получено при разных отношениях $\frac{l_k}{h_k}$. При изменении этого отношения получаются разные условия в отношении расхода меди и стали:

- 1) увеличение $\frac{l_k}{h_k}$ приводит к уменьшению расхода меди, так как при этом уменьшается объем меди и увеличивается поверхность охлаждения катушки;
- 2) увеличение $\frac{l_k}{h_k}$ приводит к увеличению расхода стали, так как при этом увеличивается длина магнитопровода.

Минимальный вес всей конструкции достигается для различных видов электромагнитов и условий их работы разными путями.

Практикой установлено следующее: при относительно большом ходе якоря и малом значении силы принимается большее значе-

ние $\frac{l_{\kappa}}{h_{\kappa}}$, при увеличении силы и уменьшении хода — это отношение уменьшается.

Таким образом, и отношение $\frac{l_{\kappa}}{h_{\kappa}}$ можно фиксировать в функции значения к. ф. Значение $\frac{l_{\kappa}}{h_{\kappa}}$ колеблется в практически выполненных конструкциях в пределах 1 ÷ 8.

Падение магнитного потенциала в стали и нерабочих зазорах

Значение $F_c + F_n$ составляет определенный процент от общего значения F — н. с. катушки. Рассматривая систему при разомкнутом положении системы — якорь отпущен — можно отметить следующее:

а) большое удельное значение F_n приводит к увеличению расхода меди, малое значение F_n , получающееся при увеличенных размерах стали в местах перехода от якоря к ярму, связано с увеличением расхода стали;

б) такое же положение имеет место при рассмотрении F_c . В рационально исполненных конструкциях

$$F_n = 0,05 \div 0,15 F, \quad F_c = 0,1 \div 0,2 F.$$

Следовательно,

$$F_c + F_n = 0,15 \div 0,35 F. \quad (5-3)$$

Коэффициент рассеяния. Для определения размеров магнитопровода требуются данные о величине полного потока в системе $\Phi = \sigma \Phi_p$. В предварительном расчете задаются значением σ , которое представляет собой отношение

$$\sigma = \frac{G_p + G_{\kappa} + G_{yn}}{G_p}.$$

Значение σ колеблется в пределах 1,3 ÷ 2,5; меньшие значения получаются в электромагнитах с малым ходом якоря, большие — с большим.

5-7. УСТАНОВЛЕНИЕ ОПТИМАЛЬНОГО ЗНАЧЕНИЯ ПРИНЯТЫХ ДАННЫХ

Системы постоянного тока. Как указывалось в параграфе 5-4, условная работоспособность электромагнита выражается как произведение величины начальной силы на ход якоря:

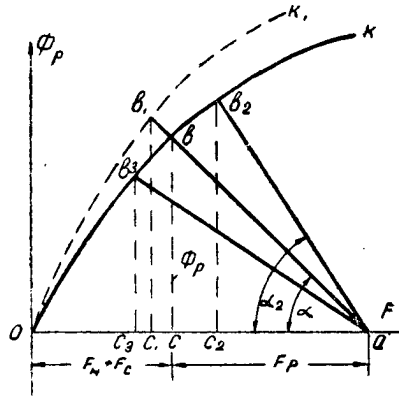
$$A = Q_{\text{ан}} \delta.$$

Для данной формы электромагнита значение A зависит от величины принятых при предварительном расчете данных: B_c — индукции в стали и B_p — индукции в рабочем зазоре. Таким обра-

зом, выбор значения этих данных определяет собой экономичность конструкции, т. е. получение максимальной работоспособности при минимальном расходе стали и меди.

Проверку соответствия принятых данных оптимальным проводят, пользуясь кривой намагничивания системы.

На фигуре 5, 6 кривая K соответствует потерям н. с. в стали и нерабочих зазорах, отрезок oa — полной н. с. катушки, oc — н. с. $F_c + F_n$ при данном потоке Φ_p , ac — oc , приходящейся на рабочий зазор.



Фиг. 5, 6. Диаграмма для определения оптимальной проводимости рабочего зазора.

Угол α соответствует проводимости рабочего зазора $\text{tg } \alpha = \frac{G_p}{p}$ (p — отношение масштабов).

Можно показать, что площадь Δabc пропорциональна работе A , совершаемой при перемещении якоря на пути δ :

$$\text{площадь } \Delta abc = \frac{1}{2} F_p \Phi_p, \quad (5-4)$$

с другой стороны,

$$A = Q_p \delta = 4,06 \cdot 10^8 \frac{\Phi_p^2}{q}. \quad (5-5)$$

Далее,

$$\Phi_p \frac{\delta}{\mu_0 q} = F_p. \quad (5-6)$$

Следовательно,

$$A = c \Phi_p F_p. \quad (5-7)$$

Сопоставляя (5-4) и (5-7), видим, что площадь $\Delta abc = c_1 A$.

Рассмотрим влияние принятых значений индукций в стали и в воздухе B_p на величину площади Δabc .

1. Характер кривой K , определяющей собой потери $F_c + F_n$, зависит от принятого значения B_c . При уменьшении B_c получается кривая K_1 (пунктирная); площадь треугольника, характеризующего A электромагнита, будет уже иной (см. ab_1c_1), размеры магнитопровода увеличиваются.

2. Изменение индукции B_p связано с изменением сечения рабочего потока, т. е. с изменением проводимости G_p .

При уменьшении B_p необходимо увеличивать G_p . Пользуясь кривой K_1 , мы устанавливаем, что увеличение G_p приводит к изменению площади треугольника, так как новому углу α_2 соответствует треугольник ab_2c_2 . Варьируя значение G_p , а следовательно, и B_p , можно получить максимальное значение работоспособности A электромагнита при заданной н. с. системы (oa).

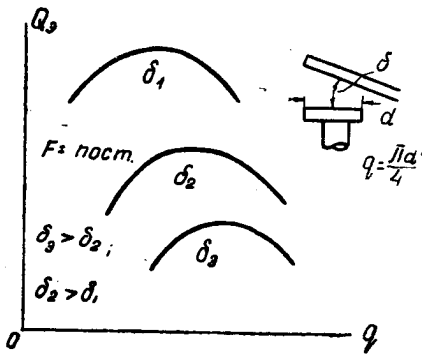
Практически для определения оптимального значения G_p , а следовательно, и B_p , поступают следующим образом. Откладывают на оси абсцисс отрезок, соответствующий принятой н. с.

катушки, проводят ряд прямых $ab, ab_2, ab_3 \dots$, определяют значения площадей треугольников abc, ab_2c_2 и останавливаются на выборе проводимости, соответствующей максимальному значению площади треугольника.

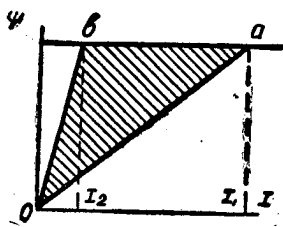
Изменение G_p весьма удобно осуществляется в магнитных системах постоянного тока с поворотным якорем: для этого достаточно изменить диаметр полюсного наконечника, укрепляемого на сердечнике. При этом согласно формуле $Q_s = 4,06 \cdot 10^8 \cdot B^2 q$ при заданной величине Q_s изменяется значение индукции B_p .

В системах переменного тока, а также соленоидных системах для изменения G_p приходится изменять сечение сердечника. Для иллюстрации влияния принятого значения G_p , а следовательно B_p , приводим кривые, показывающие изменения величины силы, действующей на якорь при изменении сечения полюсного наконечника. Все кривые (фиг. 5, 7) сняты при 2000 ампервитках катушки.

Каждой величине зазора δ соответствует оптимальное значение размеров полюсного наконечника.



Фиг. 5, 7. Кривые зависимости силы от площади полюса.



Фиг. 5, 8. Диаграмма работы системы ψ -пост.

Системы переменного тока

В магнитных системах механическая работа перемещения якоря от положения δ_n до δ_k совершается за счет магнитной энергии системы W .

В системах с постоянным числом потокосцеплений ψ эта работа определяется площадью треугольника oab (фиг. 5, 8):

$$\text{площадь } oab = 0,5\psi(I_1 - I_2),$$

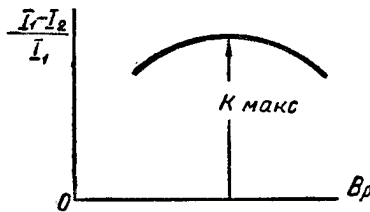
где I_1 — значение тока при зазоре δ_n ;

I_2 — значение тока при зазоре δ_k .

При увеличении разности этих токов увеличивается и работоспособность электромагнита, причем эта разность в определенной степени зависит от принятого значения индукции B_p при разомкнутом якоре.

При построении зависимости $\frac{I_1 - I_2}{I_1} = f(B_p)$ получается кривая, в которой имеется ясно выраженное максимальное значение $k_{\text{макс}} = \frac{I_1 - I_2}{I_1}$ (фиг. 5, 9). В рационально выполненных конструкциях значение $k_{\text{макс}}$ бывает порядка 0,9.

Практически для определения оптимального значения поступают следующим образом. Подсчитывают значение k для разных величин индукции B_p , строят кривую $k = f(B_p)$ и принимают значение B_p , соответствующее $k_{\text{макс}}$. При накоплении ряда опытных и расчетных данных строят кривые $B_p = f(Q_{\text{ан}}, \delta_n)$, и в этом случае несколько громоздкая работа по построению кривой k отпадает.



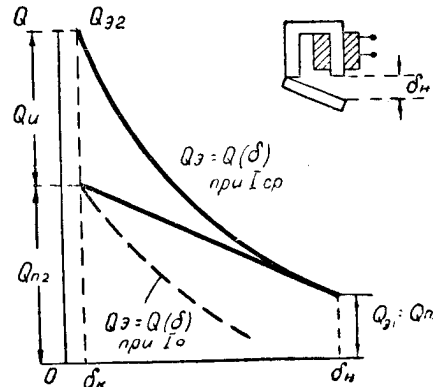
Фиг. 5, 9. Кривая $k = f(B_p)$.

которой происходит отпуск якоря, к величине, при которой происходит его втягивание.

Коэффициент возврата по току равен $k_a = \frac{I_{om}}{I_{cp}}$.

Коэффициент возврата по напряжению равен $k_a = \frac{U_{om}}{U_{cp}}$.

Значение коэффициента возврата всегда меньше единицы и колеблется для разных исполнений электромагнитов в пределах от $k_a = 0,1$ до $k_a = 0,99$. Для вычисления k_a необходимо иметь характеристики противодействующих механических сил и тяговой электромагнитной. Для примера рассмотрим магнитную систему с поворотным якорем. Характеристики $Q_s = Q(\delta)$ и $Q_n = Q(\delta)$ приведены на фигуре 5, 10. При начальном зазоре δ_n имеем следующие значения сил на якоре:



Фиг. 5, 10. Диаграммы для определения коэффициента возврата.

1) противодействующие: вес якоря и начальное натяжение пружины;

2) электромагнитная: при каком-то токе срабатывания I_{cp} создается сила такой величины, что она превышает Q_{n1} .

Обычно принимают, что началу движения якоря соответствует равенство $Q_{a1} = Q_{n1}$.

При конечном зазоре δ_k имеют место следующие значения сил на якоре:

1) механические противодействующие силы: добавляется сила растяжения пружины при ходе от δ_n до δ_k

$$Q_{n2} = Q_s + Q_{n1} + \Delta Q_n.$$

2) электромагнитные: электромагнитная сила при том же токе I_{cp} и зазоре δ_k будет большей и равна $Q_{\partial 2}$.

Якорь реле будет отходить обратно в том случае, когда ток катушки снизится до значения I_{om} , при котором электромагнитная сила $Q_{\partial 2}$ становится меньше силы Q_{n2} .

Вычислим значение коэффициента возврата по току: $k_a = \frac{I_{om}}{I_{cp}}$. При зазоре δ_k имеем: $Q_{\partial 2} = k_1 I_{cp}^2$; $Q_{\partial 2} = k_1 I_{om}^2$; отсюда

$$I_{cp} = \sqrt{\frac{Q_{\partial 2}}{k_1}}; I_{om} = \sqrt{\frac{Q_{\partial 2}}{k_1}} = \sqrt{\frac{Q_{n2}}{k_1}}.$$

Тогда

$$k_a = \sqrt{\frac{Q_{n2}}{Q_{\partial 2}}} = \sqrt{\frac{Q_n + Q_{n1} + \Delta Q_n}{Q_n + Q_{n1} + \Delta Q_n + Q_u}}, \quad (5-8)$$

где $Q_u = Q_{\partial 2} - Q_{n2}$ — избыточная сила на электромагните в конечном притянутом положении якоря.

Преобразуем (5-8), добавляя в числитель $+Q_u$ и $-Q_u$,

$$k_a = \sqrt{\frac{Q_{n2} + Q_u - Q_u}{Q_{n2} + Q_u}} = \sqrt{1 - \frac{Q_u}{Q_{n2} + Q_u}}. \quad (5-9)$$

Таким образом, значение k_a тем выше, чем меньше значение Q_u и чем больше Q_{n2} .

На основании (5-9) устанавливают следующее положение: высокий коэффициент возврата получается при условии совпадения электромагнитной и механической характеристик электромагнита. Практически такое совпадение на всей части характеристики неосуществимо. Однако можно выполнить реле так, чтобы имело место совпадение на определенном участке характеристик; тогда якорь реле должен иметь движение только на этом участке.

Следует отметить, что (5-9) дает приближенное значение. Дело в том, что при выводе этого выражения значение k_1 принято одинаковым для сил $Q_{\partial 2}$ и Q_{n2} . Практически же коэффициент этот неодинаков для обеих сил, что объясняется наличием коэрцитивной силы в стали магнитопровода.

Для получения точного значения k_a необходимо строить кривые намагничивания и размагничивания системы.

Вычисляют значения Φ_{cp} и Φ_{om} — потоков в рабочем зазоре, при которых происходит срабатывание якоря (втягивание) и отпущение его. Пользуясь кривой полного цикла работы электромагнита (фиг. 3, 30), определяют значение коэффициента возврата

$$k_a = \frac{F_{om}}{F_{cp}} = \frac{oc_2}{oa}.$$

Изменение коэффициента возврата в определенной системе осуществляется изменением конечного зазора δ_k или постановкой пружины с другим углом жесткости j .

ГЛАВА 6

ВЫЧИСЛЕНИЕ СИЛ И МОМЕНТОВ

6-1. ИСХОДНЫЕ ПОЛОЖЕНИЯ

В электромагнитных системах механическая работа перемещения якоря создается за счет изменения подведенной из сети магнитной энергии. Выражения для сил и моментов на якоре определяются соответственно:

$$Q = \frac{dW}{d\delta}; \quad (6-1)$$

$$M = \frac{dW}{d\theta}. \quad (6-2)$$

Здесь δ и θ — параметры, определяющие величину хода или угла поворота якоря.

В зависимости от характера силы, создаваемой на якоре, различают:

Силу притяжения поверхностей двух тел из магнитного материала, когда силы продольного тяжения силовых трубок, проходящих между этими поверхностями, вызывают притяжение подвижной части к неподвижной, например якоря к сердечнику.

Соленоидную силу, наблюдающуюся в прямоходовых соленоидных системах, где на якоре создается сила, определяемая взаимодействием потока утечки от якоря к наружному ярму с током, протекающим в катушке. В соленоидных системах перемещение якоря, при наличии стопа в системе, определяется обеими видами сил: притяжения якоря к стопу и соленоидной.

6-2. ВЛИЯНИЕ КРАЕВОГО ПОТОКА

Значения сил и моментов электромагнитов пропорциональны величине $\frac{dG_a}{d\delta}$ или $\frac{dG_o}{d\theta}$, где G_o — проводимость воздушных рабочих зазоров: $G_o = G_p + G_k$.

На основании ряда подсчетов установлено, что с изменением хода якоря проводимость краевого потока G_k менее резко изменяется с изменением хода, чем рабочая проводимость G_p , т. е.

$$\frac{dG_k}{d\delta} < \frac{dG_p}{d\delta}.$$

Это значит, что сила на якоре, создаваемая краевым потоком, незначительна по сравнению с силой, создаваемой рабочим потоком. Рассмотрим для примера поле между якорем и сердечником цилиндрической формы.

Примем $r = 2$ см; $\delta = 0,2$ см.

$$G_e = G_p + G_k = G_p + (G_{12} + G_{13}) = 78,5 \cdot 10^{-8} + 13,8 \cdot 10^{-8} = 92,3 \cdot 10^{-8} \text{ вб/ав.}$$

$$\frac{dG_p}{d\delta} = -393 \cdot 10^{-8} \text{ вб/ав; } \frac{dG_k}{d\delta} = 4 \cdot 10^{-8} \text{ вб/ав.}$$

$$\frac{dG_e}{d\delta} = -389 \cdot 10^{-8} \text{ вб/ав.}$$

Таким образом, в данном примере получаем следующее: величина краевого потока равна 17,5% рабочего, производная проводимости краевого потока равна лишь 1% производной проводимости рабочего. Для некоторых форм магнитных систем эта цифра повышается до 5—10%.

6-3. ФОРМУЛА МАКСВЕЛЛА

Сила притяжения поверхностей двух тел из магнитного материала, между которыми протекает поток Φ , выражается по Максвеллу

$$Q = \frac{\Phi^2}{8\pi q} \text{ дин,} \quad (6-3)$$

где Φ — в мкс, q — см².

Формула Максвелла справедлива для случая однородного поля. При рассмотрении поля между якорем и сердечником электромагнитной системы можно приближенно считать поле рабочего потока однородным. Это поле и учитывается в (6-3), а краевое поле относят к потокам рассеяния.

Формула Максвелла дает в ряде случаев только приближенное решение с ошибкой порядка 5—15%, поэтому она применяется, как правило, в предварительных расчетах систем с поворотным якорем и прямоходовых. При проведении точных расчетов формула применима для вычисления сил в конечном положении якоря, когда зазор δ невелик.

Применяя формулу, необходимо учитывать род тока.

Система постоянного тока. Преобразуя (6-3), получаем

$$Q = 4,06 \cdot 10^8 \frac{\Phi^2}{q} \text{ кг,} \quad (6-4)$$

где Φ — в веберах, q — см². Часто пользуются соотношением

$$Q = 4,06 \cdot 10^8 B_p^2 q \text{ кг,} \quad (6-5)$$

где B_p — $\frac{\text{вб}}{\text{см}^2}$.

Система переменного тока. Рассмотрим систему, где поток меняется по закону

$$\Phi_i = \Phi \sin \omega t.$$

В таких системах учитывается мгновенное значение силы Q_t и среднее значение силы за период Q .

$$Q_t = 4,06 \cdot 10^8 \frac{\Phi_t^2}{q} = 4,06 \cdot 10^8 \frac{\Phi^2 \sin^2 \omega t}{q}; \quad (6-6)$$

$$Q = \frac{1}{T} \int_0^T Q_t \cdot dt = \frac{1}{T} 4,06 \cdot 10^8 \frac{\Phi^2}{q} \int_0^T \sin^2 \omega t dt,$$

или

$$Q = 2,03 \cdot 10^8 \frac{\Phi_p^2}{q} \text{ кг}; \quad (6-7)$$

$$Q = 2,03 \cdot 10^8 B_p^2 q \text{ кг}. \quad (6-8)$$

При расчете систем переменного тока с катушкой напряжения $\psi = \text{пост.}$ значение Φ_p определяют из значения Φ — потока системы.

Если

$$\Phi = \frac{V \sqrt{2} U}{\omega \sigma}; \quad \sigma = \frac{G_p + G_k + G_{yc}}{G_p}; \quad \Phi_p = \frac{\Phi}{\sigma},$$

где G_{yc} — приведенное по числу потокосцеплений значение проводимости утечки, тогда

$$Q = 2,03 \cdot 10^8 \frac{\Phi^2}{\sigma^2 q}. \quad (6-9)$$

6-4. ВЫРАЖЕНИЕ СИЛЫ В ФУНКЦИИ ПРОИЗВОДНОЙ ПРОВОДИМОСТИ

Формула Максвелла не всегда удобна для расчетов, например при вычислении моментов в системах с поперечным движением якоря. Кроме того, эта формула в ряде случаев не дает точных результатов расчета.

Более точные результаты получаются при пользовании формулой, вывод которой произведен на основе учета баланса магнитной энергии системы (см. 6-1 и 6-2). Преобразования этих формул приводят к виду

$$Q = \frac{1}{2} I^2 \frac{dL}{d\delta} \frac{\partial \mathcal{E}}{\partial \delta}; \quad (6-10)$$

$$M = \frac{1}{2} I^2 \frac{dL}{d\theta} \frac{\partial \mathcal{E}}{\partial \theta}. \quad (6-11)$$

Значение $L = \frac{\omega^2}{R}$, где $R = R_s + R_c$.

При наличии воздушных зазоров в системе $R_c \ll R_s$. Поэтому при практических расчетах электромагнитов принимают

$$L = \frac{\omega^2}{R_s} = \omega^2 G \text{ генри}, \quad (6-12)$$

где G — проводимость воздушных путей потоков в системе в $\frac{об}{а}$.
Из (6-10), (6-11) и (6-12) имеем

$$Q = \frac{1}{2} I^2 \omega^2 \frac{dG}{d\delta} \frac{\partial \lambda}{\partial \delta} = 5,1 I^2 \omega^2 \frac{dG}{d\delta} \text{ кГ.} \quad (6-13)$$

$$M = 5,1 I^2 \omega^2 \frac{dG}{d\theta} \text{ кГ см.} \quad (6-14)$$

При пользовании соотношениями (6-13) и (6-14) необходимо учитывать особенности систем постоянного и переменного тока.

Системы постоянного тока. Произведение $I\omega$ представляют собой часть н. с. катушки, затрачиваемой в рабочем зазоре, F_p . Таким образом, соотношения (6-13) и (6-14) получают следующий вид:

$$Q = 5,1 F_p^2 \frac{dG}{d\delta} \text{ кГ.} \quad (6-15)$$

$$M = 5,1 F_p^2 \frac{dG}{d\theta} \text{ кГ.} \quad (6-16)$$

Задачу определения F_p решают графически, пользуясь кривой намагничивания системы: $\Phi_a = f(F_c + F_n)$ (фиг. 3,26). В точке a , соответствующей полному значению F катушки, проводят прямую ab , под углом α , причем $\text{tg } \alpha = \frac{G_a}{p}$ (p — коэффициент, учитывающий масштабы Φ и F). Отрезок ca на оси абсцисс определяет значение F_p . Это значение меняется по ходу якоря, когда меняется проводимость воздушного зазора G_a . При расчете электромагнита принимают во внимание наиболее неблагоприятные условия, т. е. минимально возможное значение н. с. катушки. Этому условию соответствует случай включения катушки на пониженное напряжение сети, например $U = 0,85U_n$, причем предполагается, что сопротивление катушки соответствует максимальному значению превышения температуры

$$F_{\text{мин}} = \frac{0,85U_n \omega}{r_{\text{гор}}} \quad (6-17)$$

Системы переменного тока. $F = \text{пост.}$ Для системы с токовой катушкой действительны соотношения (6-13) и (6-14), но при этом следует учитывать, что значения Q и M получаются средними за период, F — эффективное значение н. с.

Системы переменного тока. $\psi = \text{пост.}$ В таких системах принимают

$$U \cong E = \frac{\omega \Phi \omega}{\sqrt{2}} \quad (6-18)$$

и

$$I = \frac{U}{\omega L} \quad (6-19)$$

Из (6-13), (6-14), (6-18) и (6-19)

$$Q = 0,5 \frac{U^2}{\omega^2 L^2} \frac{dL}{d\delta} = 0,5 \frac{U^2}{\omega^2 \omega^2 G^2} \frac{dG}{d\delta} \partial \mathcal{K} / \text{см},$$

или, учитывая $G_p = G/\sigma$, имеем:

$$Q = 5,1 \frac{U^2}{\omega^2 \omega^2 \sigma G_p^2} \frac{dG_p}{d\delta} \text{кГ}. \quad (6-20)$$

$$M = 5,1 \frac{U^2}{\omega^2 \omega^2 \sigma G_p^2} \frac{dG_p}{d\theta} \text{кГсм}. \quad (6-21)$$

При вычислении Q и M следует учитывать $U = 0,85U_n$.

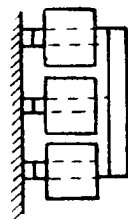
6-5. ТРЕХФАЗНЫЕ ЭЛЕКТРОМАГНИТЫ

Существуют электромагниты, снабженные тремя системами, которые действуют на общий якорь. Катушки этих систем включаются на три фазы сети переменного тока (фиг. 6,1). Мгновенное значение силы на якоре с учетом действия одной катушки:



$$Q_{ti} = \frac{1}{2} i^2 \frac{dL}{d\delta}. \quad (6-10)$$

Примем $i = I \sin \omega t$. В трехфазной симметричной системе, при условии, что потоки отдельных фаз системы не влияют друг на друга:



$$Q = 0,5 \left[I_1^2 \sin^2 \omega t + I_2^2 \sin^2 \left(\omega t + \frac{2}{3} \pi \right) + I_3^2 \sin^2 \left(\omega t + \frac{4}{3} \pi \right) \right] \frac{dL}{d\delta}.$$

После проведения соответствующих преобразований получаем

$$Q = \frac{3}{4} I^2 \frac{dL}{d\delta} \partial \mathcal{K} / \text{см}. \quad (6-22)$$

Фиг. 6, 1.
Строенная магнитная система переменного тока.

Таким образом, при наличии тройной системы с катушками, включенными в симметричной трехфазной сети, общее значение средней силы на якоре в 1,5 раза больше средней силы одной магнитной системы.

Заметим, что сила Q строенной системы остается постоянной во времени и, следовательно, отсутствуют явления вибрации, получающиеся в однофазной системе.

6-6. СОЛЕНОИДНАЯ СИЛА

В прямоходовых системах сила на якоре, определяемая изменением магнитной энергии системы по ходу якоря, рассматривается для удобства расчета состоящей из двух сил (фиг. 6, 2):

а) Q_1 — силы притяжения двух поверхностей, определяемой потоком Φ_v ;

б) Q_2 — соленоидной силы, определяемой взаимодействием потока утечки Φ_v от якоря к наружному кожуху с током, протекающим в катушке.

Вычисление значения Q_2 .

$$Q_2 = \frac{dW_v}{dx} = \frac{1}{2} I \frac{d\psi_v}{dx} dj / \text{см}. \quad (6-23)$$

Число потокоцеплений потока утечки ψ_v :

$$\psi_v = \frac{\omega}{l_k} x B_v y x = \frac{\omega}{l_k} B_v y x^2, \quad (6-24)$$

где B_v — индукция потока, выходящего нормально к оси якоря, y — периметр якоря. Из (6-23) и (6-24)

$$Q = \frac{\omega x y}{l_k} I B_v.$$

Учитывая, что напряженность поля катушки

$$H = \frac{I \omega}{l_k}; \quad \Phi_v = B_v y x,$$

имеем

$$Q_2 = 10,2 H \Phi_v \text{ кг}. \quad (6-25)$$

Для систем $\psi = \text{пост.}$ следует учитывать, что Q_2 — среднее значение силы за период, Φ_v — амплитудное значение потока. В этом случае

$$Q_2 = \frac{10,2}{\sqrt{2}} H \Phi_v. \quad (6-26)$$

При вычислении соленоидной силы необходимо принимать во внимание смещение нейтрали $a-b$, которое по мере увеличения x и уменьшения δ перемещается в сторону рабочего зазора, что приводит к уменьшению Φ_v , а следовательно и Q_2 . При значениях $\delta < 0,1 l_k$ величину Q_2 практически можно не учитывать.

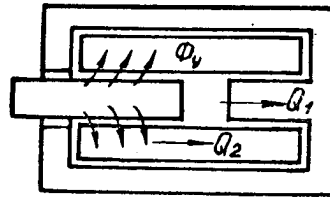
6.7. ВЛИЯНИЕ РАСПОЛОЖЕНИЯ КАТУШКИ В СИСТЕМЕ

Величина тяговой силы электромагнита в значительной степени зависит от расположения катушки в магнитной системе. Возможны следующие варианты расположения катушки при отпущенном якоре:

а) средняя линия катушки совпадает со средней линией начального воздушного зазора между подвижным якорем и неподвижным сердечником;

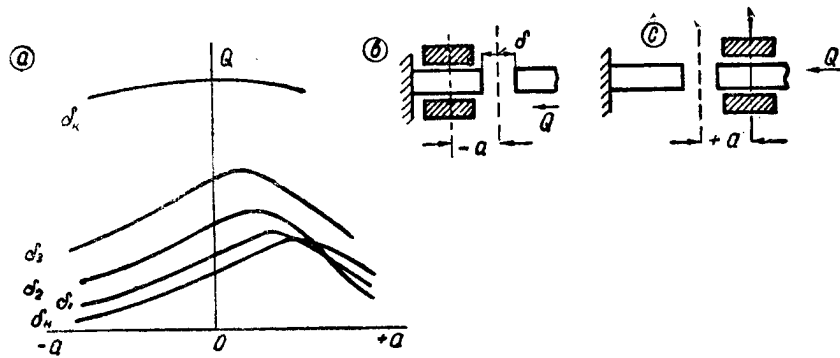
б) средняя линия смещена влево (фиг. 6,3 в);

в) средняя линия смещена вправо (фиг. 6,3 с).



Фиг. 6, 2. Силы на якоре прямоходовой системы.

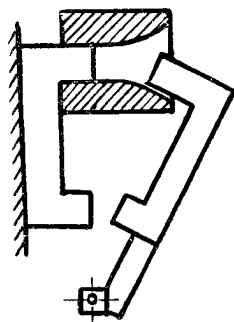
А. П. Назаренко, при исследовании магнитных систем контакторов, установил влияние расположения катушки на величину тяговой силы электромагнитов с поворотным якорем. На фигуре 6,3 приведены полученные А. П. Назаренко кривые, из которых



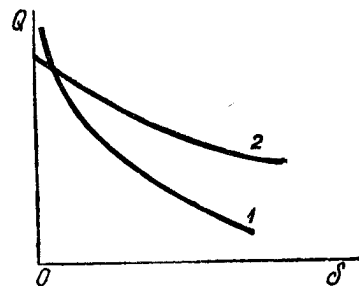
Фиг. 6, 3. Кривые зависимости силы на якоре от расположения катушки.

видно, что наиболее благоприятные условия в отношении создания начальной силы на якоре, получаются при положительном значении a .

А. П. Назаренко предложена соответствующая конструкция для электромагнитов, контакторов и реле переменного тока



Фиг. 6, 4. Магнитная система.



Фиг. 6, 5. Тяговые характеристики электромагнитов:
1 - якорь вне катушки; 2 - якорь в катушке.

(фиг. 6,4). По сравнению с ранее применявшейся для контакторов переменного тока магнитной системой E-образного типа с утопленным расположением катушки, А. П. Назаренко получил значительную экономию меди и стали.

ГЛАВА 7

РАСЧЕТ ЭКРАНА ДЛЯ УСТРАНЕНИЯ ВИБРАЦИЙ ЯКОРЯ

7-1. ВИБРАЦИИ ЯКОРЯ

При включении катушек электромагнитов на напряжение переменного тока получающееся на якоре мгновенное значение силы Q_t не является величиной постоянной, а изменяется вместе с изменением по времени величины потока в системе.

При наличии потока $\Phi_t = \Phi \sin \omega t$ мгновенное значение силы получается равным

$$Q_t = 4,06 \cdot 10^8 \frac{\Phi_t^2}{q} = 4,06 \cdot 10^8 \frac{\Phi^2}{q} \sin^2 \omega t. \quad (7-1)$$

Заменяя в (7-1) $\sin^2 \omega t = 0,5(1 - \cos 2\omega t)$,
имеем

$$Q_t = 2,03 \cdot 10^8 \frac{\Phi^2}{q} - 2,03 \cdot 10^8 \frac{\Phi^2}{q} \cos 2\omega t. \quad (7-2)$$

Таким образом, можно считать, что мгновенное значение силы состоит как бы из двух частей:

1) постоянной составляющей $Q = 2,03 \cdot 10^8 \frac{\Phi^2}{q}$; (7-3)

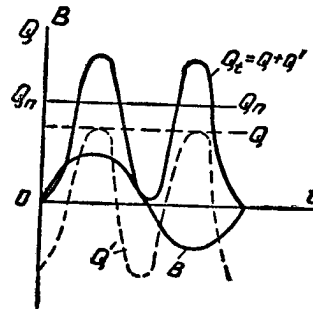
2) переменной составляющей $Q' = 2,03 \cdot 10^8 \frac{\Phi^2}{q} \cos 2\omega t$. (7-4)

Полное значение силы Q_t меняется с двойной частотой от нуля (при $\omega t = 0$) до двойного значения (при $\omega t = 0,5\pi$) (фиг. 7, 1).

Если противодействующая — механическая сила — в электромагните равна Q_n , то в моменты, когда $Q_n > Q_t$, якорь начинает отходить от сердечника и снова возвращается в момент, когда $Q_t > Q_n$.

Таким образом, якорь в течение одного периода стремится отойти от сердечника и два раза снова притянуться к нему, т. е. при переменном токе с частотой 50 гц якорь имеет 100 колебаний в секунду. Эти колебания недопустимы, так как они создают условия, приводящие к износу магнитной системы и к подгоранию контактов.

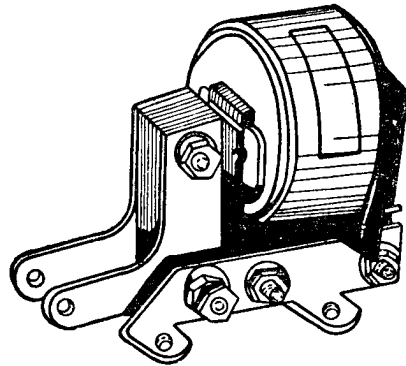
Кроме того, вибрации якоря вызывают шум и гудение, неприятно действующие на обслуживающий персонал. Поэтому в электромагнитах переменного тока применяют, как правило, специальные меры к устранению вибраций.



Фиг. 7, 1. Кривые $Q = t(f)$ электромагнита переменного тока.

7-2. СПОСОБЫ УСТРАНЕНИЯ ВИБРАЦИЙ

Для устранения вибраций необходимо, чтобы при переменном токе электромагнитная сила притяжения якоря не падала ниже силы Q_n противодействующего устройства. Для этого магнитную систему исполняют так, чтобы притяжение якоря происходило от действия двух потоков, существующих одновременно, но сдвинутых в пространстве и во времени. Тогда при переходе значения одного потока через нуль другой имеет какую-то определенную величину, благодаря чему значение электромагнитной силы не снижается до нуля.



Фиг. 7,2. Магнитопровод с экраном.

Для устранения вибраций в практике применяют два способа.

1. Магнитная система реле состоит из двух отдельных магнитопроводов; катушка одного магнитопровода подключается непосредственно к сети, а второго — через емкость. В этом случае поток во втором магнитопроводе сдвинут по фазе по отношению к потоку первого магнитопровода.

2. В сердечнике или в якоре магнитной системы у рабочего зазора ставится короткозамкнутое кольцо — экран (фиг. 7, 2). Экран охватывает только часть сечения сердечника или якоря и поток, пронизывающий эту экранированную часть, отстает по фазе от потока, пронизывающего неэкранированную часть.

В релейной практике применяется, как правило, второй способ устранения вибраций. Первый способ, дающий более громоздкое конструктивное решение, применяется редко.

Следует отметить, что в электромагнитах с тремя катушками, включенными в цепь трехфазного тока, явление вибрации не имеет места. Объясняется это тем, что общая сила притяжения трех потоков, сдвинутых по фазе на 120° , пропорциональна Φ^2 и имеет поэтому постоянную величину, не зависящую от времени.

7-3. ПУЛЬСАЦИИ СИЛЫ

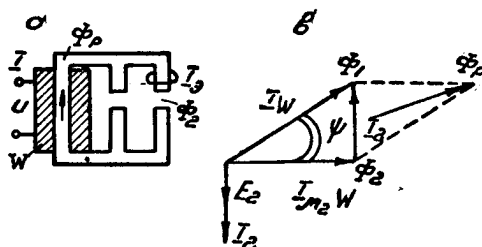
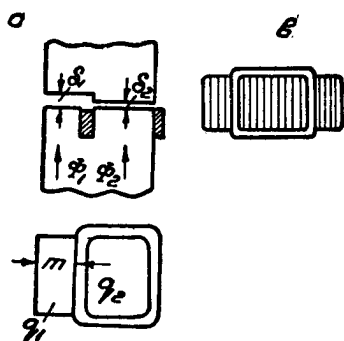
Ниже рассматриваются системы с экраном. На фигуре 7, 3 представлена та часть магнитопровода, где происходит стык между якорем и сердечником. В сердечнике установлен экран в виде короткозамкнутого витка из латуни или меди.

Примем следующие обозначения:

а) неэкранированная часть — поток Φ_1 , площадь сечения q_1 , зазор δ_1 ;

б) экранированная часть — соответственно, Φ_2, q_2, δ_2 .

Упрощенная векторная диаграмма без учета углов потерь для потоков Φ_1 и Φ_2 приведена на фигуре 7, 4. Намагничивающая сила катушки магнитной системы $F = Iw$ определяется как геометрическая сумма $F_2 = I_{\mu_1}w$ для потока Φ_2 и $F'_2 = I_2$, где I_2 ток экрана. Индуктивное сопротивление экрана весьма невелико и его значение не учитывается. Поток Φ_2 сдвинут по фазе по отношению к потоку Φ_1 — угол сдвига ψ .



Фиг. 7, 3. Варианты расположения экранов в магнитопроводе. Фиг. 7, 4. Схема замещения и векторная диаграмма.

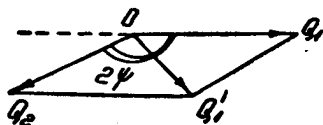
При наличии потоков Φ_1 и Φ_2 мгновенные значения силы на якоре, согласно (7-2):

$$Q_{i1} = 2,03 \cdot 10^8 \frac{\Phi_1^2}{q_1} - 2,03 \cdot 10^8 \frac{\Phi_1^2}{q_1} \cos 2\omega t. \quad (7-5)$$

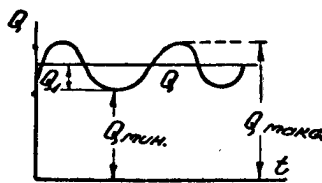
$$Q_{i2} = 2,03 \cdot 10^8 \frac{\Phi_2^2}{q_2} - 2,03 \cdot 10^8 \frac{\Phi_2^2}{q_2} \cos (2\omega t - 2\psi). \quad (7-6)$$

Или, с учетом обозначений (7-3):

$$Q_{i1} = Q_1 - Q_1 \cos 2\omega t \quad (7-7); \quad Q_{i2} = Q_2 - Q_2 \cos (2\omega t - 2\psi). \quad (7-8)$$



Фиг. 7, 5. Векторная диаграмма.



Фиг. 7, 6. Силы на якоре системы с экраном.

Амплитудное значение суммарной переменной составляющей может быть определено из векторной диаграммы (фиг. 7, 5):

$$Q' = \sqrt{Q_1^2 + Q_2^2 + 2Q_1Q_2 \cos 2\psi}. \quad (7-9)$$

Суммарное значение постоянной составляющей $Q = Q_1 + Q_2$. Представляет интерес рассмотрение двух сил (фиг. 7, 6): максимальной $Q_{i\max} = Q + Q'$; минимальной $Q_{i\min} = Q - Q'$.

Пульсация силы, т. е. отношение амплитудного значения переменной составляющей Q' к постоянной

$$p = \frac{Q'}{Q}. \quad (7-10)$$

Пульсации в системе отсутствуют при $Q' = 0$. Установить условия $Q' = 0$ можно пользуясь диаграммой, приведенной на фигуре 7, 5. Амплитудное значение $Q' = 0$ получается при наличии двух условий: 1) $Q_1 = Q_2$; 2) $2\psi = \pi$.

Таким образом, для устранения пульсации сил в электромагнитной системе необходимо выполнить следующие условия:

- 1) угол сдвига между потоками ψ должен быть равен 90° ;
- 2) постоянные составляющие сил должны быть равны

$$Q_1 = Q_2. \quad (7-11)$$

Как будет показано ниже, в системах с экраном выполнить первое условие практически невозможно. В существующих системах значение $\psi = 50^\circ \div 80^\circ$. При несоблюдении первого условия минимальное значение пульсации получается при наличии второго условия.

Ниже рассматриваются две задачи: расчет системы с экраном и проектирование экрана.

7-4. РАСЧЕТ СИСТЕМЫ С ЭКРАНОМ

Задачей расчета является вычисление минимального значения электромагнитной силы $Q_{\text{мин}}$ на якоре в протянутом его положении. Предполагаются известными все данные электромагнита и экрана: рабочее напряжение катушки U_n , число витков катушки w , коэффициент рассеяния магнитной системы при протянутом якоре — σ , размеры полюса q_1 и q_2 (фиг. 7, 3), омическое сопротивление экрана. Основные соотношения для расчета можно получить из векторных диаграмм.

Распределение потоков. Рассмотрим схему замещения магнитной системы с расщепленным полюсом (фиг. 7, 4 а). Система состоит из двух частей: дроссельной катушки с воздушным зазором δ_1 и трансформатора с зазором δ_2 и вторичной короткозамкнутой обмоткой (экраном). Общий поток сердечника Φ_p разделяется у расщепленного полюса на две части: Φ_1 и Φ_2 .

При построении упрощенной диаграммы (фиг. 7, 4 б):

- 1) не учитываются потери н. с. катушки в стали магнитопровода;
- 2) углы потерь для потоков Φ_1 и Φ_2 принимаются равными, поэтому при построении диаграммы эти углы не учитываются;
- 3) не учитывается реактивное сопротивление экрана, которое во много раз меньше омического — по экспериментальным данным $r_{\text{э}}/X_{\text{э}}$ получается порядка 10.

Значение потоков Φ_1 и Φ_2 определяется н. с. катушки системы

$$F = I\omega = I_{\mu_1}\omega + I_\vartheta.$$

Обозначим $G_1 = \mu_0 \frac{q_1}{\delta_1}$; $G_2 = \mu_0 \frac{q_2}{\delta_2}$ — магнитные проводимости полюса у неэкранированной и экранированной части его.

Тогда $\Phi_1 = \sqrt{2} I\omega G_1$; $\Phi_2 \sqrt{2} I_{\mu_2}\omega G_2 = \sqrt{2} I\omega G_2 \cos \psi$;

$$\frac{\Phi_1}{\Phi_2} = \frac{G_1}{G_2 \cos \psi} = \frac{q_1 \delta_2}{\delta_1 q_2 \cos \psi}.$$

Обозначив

$$\alpha = q_1/q_2; \quad (7-11)$$

$$\beta = \delta_1/\delta_2, \quad (7-12)$$

получаем

$$\frac{\Phi_1}{\Phi_2} = \frac{\alpha}{\beta \cos \psi}. \quad (7-13)$$

В обычных исполнениях электромагнитов $\delta_1 = \delta_2$; тогда

$$\Phi_1/\Phi_2 = \frac{\alpha}{\cos \psi} = c. \quad (7-14)$$

Значение угла ψ определяется следующим образом: из фигуры 7, 4 б

$$\operatorname{tg} \psi = \frac{I_\vartheta}{I_{\mu_1}\omega} = \frac{E_\vartheta}{r_\vartheta I_{\mu_1}\omega}. \quad (7-15)$$

При числе витков экрана, равном 1,

$$E_\vartheta = \frac{\omega}{\sqrt{2}} \Phi_2; \quad \Phi_2 = \sqrt{2} I_{\mu_2}\omega G_2.$$

Подставляя значения E_ϑ и Φ_2 в (7-15), получаем

$$\operatorname{tg} \psi = \frac{\omega G_2}{r_\vartheta}. \quad (7-16)$$

Определив отношение потоков Φ_1 и Φ_2 (7-14) и значение угла ψ , можно найти величины этих потоков в функции результирующего потока Φ_p , определяемого н. с. катушки.

Из диаграммы, приведенной на фигуре 7, 4 б, имеем

$$\Phi_p^2 = \Phi_1^2 + \Phi_2^2 - 2\Phi_1\Phi_2 \cos(\pi - \psi) = \Phi_1^2 + \Phi_2^2 + 2\Phi_1\Phi_2 \cos \psi;$$

учитывая (7-14), получаем

$$\Phi_2 = \frac{\Phi_p}{\sqrt{1+c^2+2c \cos \psi}}; \quad (7-17)$$

$$\Phi_1 = c\Phi_2. \quad (7-18)$$

Силы на якоре. Минимальное значение электромагнитной силы на якоре в притяннутом его положении, $Q_{\text{л.мин}}$ определяется как

сумма постоянных составляющих сил от потоков Φ_1 и Φ_2 и переменной составляющей.

$$Q_{l\min} = Q_1 + Q_2 - Q'.$$

С учетом (7-9) получаем

$$Q_{l\min} = Q_1 + Q_2 - \sqrt{Q_1^2 + Q_2^2 + 2Q_1Q_2 \cos 2\psi}. \quad (7-19)$$

Значение $Q_{l\min}$ должно быть больше $Q_{нк}$ — значения противодействующих механических сил на якоре в притянутом его положении. В противном случае будет наблюдаться явление вибрации якоря.

Расчет следует проводить с учетом падения напряжения в сети, т. е. при $U = 0,85 U_n$.

Потери в экране. Активные потери в экране $P = I_s^2 r_s$.

Учитывая

$$I_s = \frac{E_s}{r_s} = \frac{\omega \Phi_2}{\sqrt{2} r_s},$$

получаем:

$$P = \frac{\omega^2 \Phi_2^2}{2 r_s}. \quad (7-20)$$

При вычислении активных потерь и превышения температуры экрана следует учитывать условие $U = U_n$.

Значение тока в рабочей катушке системы определяется построением полной векторной диаграммы с учетом падения н. с. в стали магнитопровода и потерь на гистерезис и вихревые токи (см. главу 3).

7-5. ПРОЕКТИРОВАНИЕ СИСТЕМЫ С ЭКРАНОМ

Задачей проектирования является установление размеров q_1 и q_2 расщепленного полюса, а также размеров экрана с тем, чтобы минимальное значение электромагнитной силы на якоре в притянутом его положении, $Q_{l\min}$, было больше $Q_{нк}$ — конечного значения противодействующей механической силы. При этом принимают определенное значение коэффициента запаса κ_n .

$$Q_{l\min} = \kappa_n Q_{нк}.$$

При выводе основных соотношений для расчета принимают ряд допущений, которые приводят к тому, что расчетные значения сил получаются меньшими, чем действительные. Поэтому значения κ_n принимают небольшим, порядка 1,1.

Общий расчет и проектирование электромагнита проводится и условия получения требуемой электромагнитной силы при заданных значениях начальной силы противодействующих устройств $Q_{нк}$ и начального хода δ_n . При этом влияние экрана не учитывается. Расчет проводится по среднему значению силы. Статиче

ская электромагнитная характеристика также строится без учета влияния экрана.

В конечном положении якоря на работе электромагнита сказывается переменное во времени значение электромагнитной силы. В данном случае учитывается действие экрана.

При проектировании экрана учитываются данные электромагнита, полученные при проведении предварительного расчета:

$Q_{\text{нн}}$ — конечное значение противодействующих сил;

q — площадь полюса;

ω — число витков катушки;

σ — коэффициент рассеяния магнитной системы в конечном положении якоря.

Выбор значений α и β

Расчетные данные для экрана определяются двумя основными соотношениями:

$$\alpha = \frac{q_1}{q_2} \quad (7-11); \quad \beta = \frac{b_1}{b_2} \quad (7-12).$$

Предельные значения α и β ограничиваются практическими условиями.

В определенной конструкции уменьшение значения α достигается за счет уменьшения q_1 . При этом увеличивается величина B_1 индукции в неэкранированной части полюса, что приводит к резкому увеличению н. с. катушки.

Предельное значение B_1 принимается порядка $1,6 \cdot 10^{-4}$. Следует учитывать, что при малом значении α уменьшается размер m (фиг. 7, 3), что создает трудности при штамповке.

В практических условиях $\alpha > (0,415 \div 0,425)$. Как указывалось выше, учитывая технологические условия, принимают значение $\beta = 1$.

МЕТОДЫ ПРОЕКТИРОВАНИЯ

В литературных источниках приводятся разные методы проектирования экрана. Наиболее рациональное решение задачи получается при методе анализа работы системы, предложенном А. В. Шорыгиным. Пользуясь этим методом, Г. В. Могилевский получил в своей диссертационной работе весьма простые соотношения для расчета. Ниже приводятся основные положения, принятые Г. В. Могилевским при решении задачи.

Исходное уравнение

$$Q_{\text{л. мин}} = Q_1 + Q_2 - \sqrt{Q_1^2 + Q_2^2 + 2Q_1Q_2 \cos 2\psi}. \quad (7-19)$$

Примем следующие обозначения:

$$f = \frac{Q_{\text{л. мин}}}{Q_c}, \quad (7-21)$$

где

$$Q_{\text{ср}} = 2,03 \cdot 10^8 \frac{\Phi_p^2}{q_p} \quad (7-22)$$

— среднее значение силы при отсутствии экрана;

$$q_0 = q_1 + q_2. \quad (7-23)$$

$$\tau = (1 + \alpha)^2 r_0^2. \quad (7-24)$$

$$m = \alpha \left(\frac{\omega \mu_0 q_0}{\delta_2} \right)^2. \quad (7-25)$$

Разделив все члены уравнения (7-19) на Q_{cp} и произведя соответствующие преобразования, получаем с учетом принятых обозначений

$$f = f(\alpha, \beta, m, \tau). \quad (7-26)$$

Задача расчета сводится к отысканию зависимости между значениями α , β и r_0 , при которой величина f имеет максимальное значение.

При $r_0 = 0$ и $r_0 = \infty$ значение $f = 0$. Таким образом, имеется оптимальная величина сопротивления экрана, при которой получается значение f_{max} .

Учитывая, что в зависимости (7-26) имеется одна переменная τ , берем производную $\frac{\partial f}{\partial \tau} = 0$.

Решая полученное уравнение, получаем следующие соотношения для расчета (с учетом $\beta = 1$):

$$\alpha = \frac{2 - f}{4f}. \quad (7-27)$$

$$r_0 = \frac{\omega \mu_0 q_0}{\delta_2} \cdot \frac{4f}{(3f + 2)^2} \sqrt{4 - f^2}. \quad (7-28)$$

Экран выполняется в виде кольца с неравномерным сечением. При конструировании экрана учитывают условия его охлаждения. Обычно экран выполняют из материала с относительно большим удельным сопротивлением — латуни или алюминия.

В части, расположенной в пазу, принимают меньшее сечение экрана с целью уменьшения ширины паза. В части вне паза увеличивают площадь кольца, благодаря чему увеличивается поверхность охлаждения экрана.

Значение превышения температуры Θ для экрана принимается различным. Если экран расположен вблизи катушки, значение Θ должно быть не больше 70° , если удалён от неё — допускается значение Θ порядка 100 — 120° .

Методика проведения расчета поясняется приводимым ниже примером.

7-6. ПРИМЕР РАСЧЕТА ЭКРАНА (фиг. 7, 7).

1. Заданные величины: $U_n = 220$ в; $\omega = 4000$; $\sigma = 1,05$;
 $q = 1,8 \times 1,2 = 2,16$ см²; $Q_{nk} = 1,8$ кг;

2. Исполнение полюса: учитывая паз шириной 2 мм, получаем

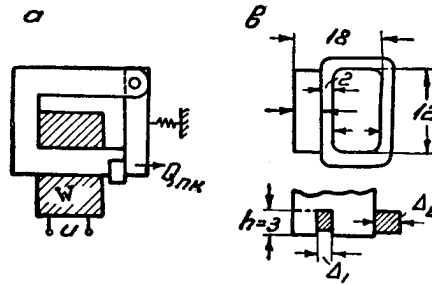
$$q_0 = 1,6 \times 1,2 = 1,92$$
 см².

3. Значение результирующего потока Φ_p : при определении $Q_{l \text{ мин}}$ учитывается напряжение $U = 0,85U_n$, при определении потерь в экране $U = U_n$.

$$\Phi_p = \frac{\sqrt{2} \cdot 0,85U_n}{\sigma \omega \omega} = \frac{\sqrt{2} \cdot 0,85 \cdot 220}{1,05 \cdot 314 \cdot 4000} = 2 \cdot 10^{-4} \text{ об.}$$

4. Среднее значение силы при отсутствии экрана:
из (7-22)

$$Q_{cp} = 2,03 \cdot 10^8 \frac{\Phi^2}{q_0} = 2,03 \cdot 10^8 \frac{4 \cdot 10^{-8}}{1,92} = 4,22 \text{ кг.}$$



Фиг. 7. Эскизы к расчету экрана.

5. Коэффициент f :
из (7-21)

$$f = \frac{k_n Q_{l \text{ мин}}}{Q_{cp}} = \frac{1,1 \cdot 1,8}{4,22} = 0,47.$$

6. Коэффициент α :
из (7-27)

$$\alpha = \frac{2-f}{4f} = \frac{2-0,47}{4 \cdot 0,47} = 0,81; \quad q_1 = 0,858; \quad q_2 = 1,062.$$

7. Сопротивление экрана:

$$r_s = \frac{\omega \mu_0 q_0}{\delta_2} \frac{4f}{(3f+2)^2} \sqrt{4-f^2} = 4,78 \cdot 10^{-4} \text{ ом.}$$

8. Угол ψ :
из (7-16)

$$\operatorname{tg} \psi = \frac{\omega G_2}{r_s} = \frac{314 \cdot 1,25 \cdot 10^{-8} \cdot 1,062}{4,78 \cdot 10^{-4} \cdot 5 \cdot 10^{-3}} = 1,755.$$

$$\psi = 60^\circ 20'; \quad \cos \psi = 0,495; \quad \cos 2\psi = -0,51.$$

9. Поток:
из (7-17)

$$\Phi_p = \sqrt{\Phi_3^2 (1 + c^2 + 2c \cos \psi)};$$

из (7-14)

$$c = \frac{\alpha}{\cos \psi} = \frac{0,81}{0,495} = 1,63;$$

$$\Phi_2 = 0,87 \cdot 10^{-4} \text{ об}; \quad \Phi_1 = c\Phi_2 = 1,415 \cdot 10^{-4} \text{ об.}$$

10. Силы:

$$Q_1 = 2,03 \cdot 10^8 \frac{1,415^2 \cdot 10^{-8}}{0,858} = 4,75 \text{ кг.}$$

$$Q_2 = 1,445 \text{ кг}$$

$$Q' = \sqrt{Q_1^2 + Q_2^2 - 2Q_1Q_2 \cos 2\psi} = 4,21 \text{ кг.}$$

$$Q_{\text{мин}} = 4,75 + 1,445 - 4,21 = 1,985 \text{ кг.}$$

По условию требуется $Q_{\text{т мин}} = 1,1 \cdot 1,8 = 1,98 \text{ кг.}$

11. Активные потери в экране:
из (7-20)

$$P = \frac{\omega^2 \Phi_2^2}{2 \cdot 0,85^2 \cdot r_3} = 1,17 \text{ вт.}$$

12. Индукция в неэкранированной части полюса:

$$B_1 = \frac{\Phi_1}{q_1} = \frac{1,415 \cdot 10^{-4}}{0,858} = 1,65 \cdot 10^{-4} \frac{\text{вб}}{\text{см}^2}.$$

ГЛАВА 8

МАГНИТНЫЕ СИСТЕМЫ С ПОВОРОТНЫМ ЯКОРЕМ

8-1. ВОПРОСЫ ПРОЕКТИРОВАНИЯ СИСТЕМ ПОСТОЯННОГО ТОКА

Формы магнитных систем. В практике применяются разнообразные формы магнитных систем, на фигуре 8, 1 приведены наиболее распространенные из них.

а) Система реле кодового типа; сердечник круглого сечения, без полюсной надставки. Такое решение принято из условия осуществления легкой и быстрой замены катушки.

б) Система реле кодового типа, но с Ш-образным магнитопроводом, сердечник круглого сечения. Такие системы применяют в замедленно действующих реле. Благодаря принятой форме магнитопровода увеличивается значение потока в системе при конечном положении якоря, что приводит к увеличению времени отпуска реле.

в) Система телефонного реле плоского типа, обладающая высокой чувствительностью, так как в этой системе сила на якоре создается не только потоком рабочего зазора, но и потоком рассеяния между сердечником и якорем.

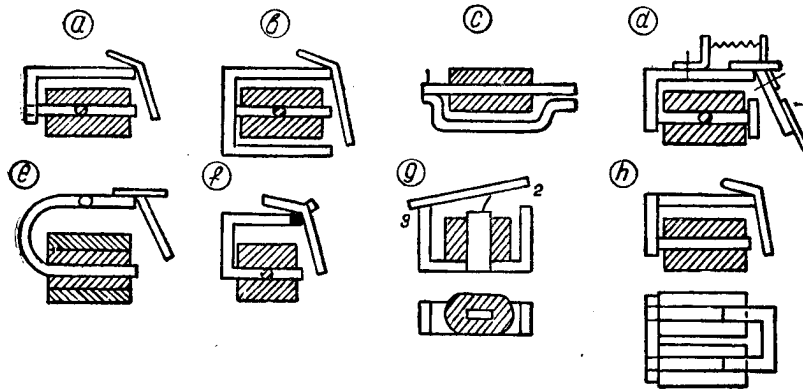
г) Система, принятая для различных реле управления и промежуточных.

е) Система, принятая для замедленно действующих реле управления. Особенностью её является отсутствие нерабочего зазора между сердечником и ярмом — магнитопровод выполнен цельным из железа Армко круглого сечения.

ф) Система, принятая для высокочувствительных и быстродействующих реле. Магнитопровод её выполняется шихтованным из

листового материала с высоким значением магнитной проницаемости в слабых полях, например из пермаллоя. Якорь вращается на призме и удерживается на ярме специальной пружиной.

г) Система, принятая для промежуточных реле постоянного и переменного тока. Магнитопровод цельный, выштампованный из одной полосы кремнистой стали.



Фиг. 8. 1. Формы систем постоянного тока с поворотным якорем.

н) Сдвоенная система, принятая для быстродействующих реле, имеющих контакты большой разрывной мощности.

В практике применяются и другие формы систем, приспособленных к условиям работы механизма.

8-2. ИСХОДНЫЕ ДАННЫЕ И СООТНОШЕНИЯ ДЛЯ СИСТЕМ ПОСТОЯННОГО ТОКА

Исходными для проектирования магнитной системы принимают данные исполнительной части реле: контактной системы, приводного механизма. Эти данные сводятся к двум основным величинам: начальной силе Q_n , или моменту M_n , и ходу механизма x , или углу поворота θ . Необходимо установить наиболее выгодные условия работы магнитной системы, т. е. подобрать такое значение хода якоря δ , при котором получается наибольшее значение условной работоспособности A (см. главу 5). Как правило, в системах постоянного тока значение δ принимают меньше x (см. фиг. 5, 3).

Для проектирования системы пользуются следующими основными соотношениями:

$$\text{сила } Q = 4,06 \cdot 10^8 B_p^2 g \text{ кГ (6-5),}$$

н. с. катушки F складывается из падения магнитного потенциала в стали и нерабочих зазорах $F_c + F_n$ и в рабочем зазоре F_p

$$F = F_p + F_c + F_n = \frac{B_p \delta}{\mu_0} + F_c + F_n \quad (8-1)$$

напряжение

$$U = Ir = I\omega p \frac{l_{cp}}{q} 10^{-3} \quad (8-2)$$

превышение температуры обмотки

$$\theta = \rho \frac{1}{20\mu f_n h_n l_n^2} (I\omega)^2, \quad (2,22)$$

где l_n и h_n в мм.

Пользуясь приведенными соотношениями, можно провести расчет магнитной системы готовой формы с определенными габаритами.

При проектировании же системы, когда нет данных о размерах ее отдельных частей, необходимо задаваться некоторыми значениями параметров, причем удачный выбор этих значений предопределяет экономичность конструкции. Поэтому вначале проводится предварительный расчет системы, данные которого уточняют в окончательном расчете.

В предварительном расчете приходится задаваться следующими значениями параметров: индукцией в рабочем воздушном зазоре B_p (принимается по значению к. ф. (фиг. 5, 5); коэффициентом рассеяния; падением магнитного потенциала в стали и нерабочих зазорах и др.

8-3. ПРЕДВАРИТЕЛЬНЫЙ РАСЧЕТ СИСТЕМЫ ПОСТОЯННОГО ТОКА

Исходные данные. Заданными являются следующие данные: напряжение сети U_n , характеристика противодействующих сил $Q_n = f(\delta)$ или $M_n = \varphi(\delta)$.

На фигуре 8, 2 показаны два вида характеристик.

В случае (а) предварительный расчет проводится по значениям Q_{n1} и δ_{n1} , в поверочном расчете уточняется значение Q_n для точек 2 и 3.

В случае (б) расчет по точке 1 привел бы к неверным результатам, ибо в точке 2 сила Q_{n2} получается меньше Q_{n1} . Поэтому в данном случае расчет проводится по точке 2: Q_{n2} , δ_{n2} . Подобрав наиболее подходящие для магнитной системы значения Q_{n1} и δ_{n1} (см. главу 5), перестраивают механическую характеристику с тем, чтобы на оси абсцисс были отложены значения δ_{n1} .

Диаметр полюсного наконечника. Установив значение конструктивного фактора $\frac{\sqrt{Q_n}}{\delta}$ определяют по кривой, изображенной на фигуре 5, 5, величину индукции B_p — в рабочем зазоре. Из формулы $Q = 4,06 \cdot 10^8 B_p^2 q$ определяют значение q и диаметр полюсного наконечника.

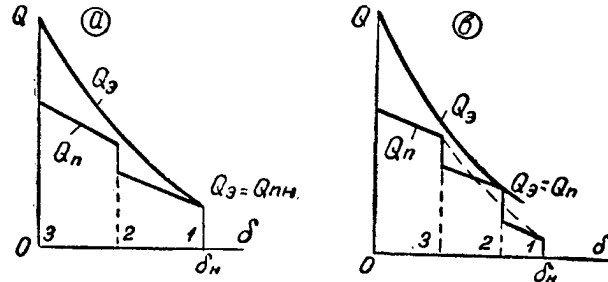
Диаметр сердечника. Задаются значением σ — коэффициента рассеяния и величиной B_c индукции в стали. Значение B_c при-

нимают порядка $(1,2 \div 1,4) 10^{-4} \frac{\text{сб}}{\text{см}^2}$ для мощных магнитных систем и порядка $(0,4 \div 0,7) 10^{-4} \frac{\text{сб}}{\text{см}^2}$ для релейных систем.

Поток у основания сердечника

$$\begin{aligned} \Phi &= \sigma \Phi_p = \sigma B_p q; \\ q_c &= \frac{\sigma B_p q}{B_c}. \end{aligned} \quad (8-3)$$

Полученное значение q_c относится к сердечнику и к яру, сечение якоря может быть взято несколько меньшим.



Фиг. 8, 2. Тяговые характеристики электромагнитов.

Н. с. катушки. $F_p = \frac{B_p \delta}{\mu_0}$.

Задаются величиной $F_c + F_n = \alpha F$ (α принимается порядка $0,15 \div 0,35$, см. 5-6).

Тогда

$$F = F_p + \alpha F = \frac{B_p \delta}{\mu_0 (1 - \alpha)}. \quad (8-4)$$

Размеры катушки. Задаются значением $k = \frac{l_k}{h_k}$ — отношением длины катушки к толщине намотки, коэффициентом заполнения катушки f_k (приблизительно 0,5) и допускаемой температурой превышения (см. главу 2). Подставляя в уравнение (2-21) $I \omega = \frac{F}{0,85}$; ρ — соответствующее температуре $\vartheta_0 + \theta$, принятое значение f_k , коэффициент теплоотдачи μ (см. главу 2), заменив $l_k = h_k k$, получают из (2,22)

$$h = \sqrt[3]{\frac{F^2}{20 \cdot 0,85^2 \mu f_k \theta k^2}}. \quad (8-5)$$

При проведении расчета следует принять во внимание то, что напряжение U , подведенное к катушке, может изменяться в процессе эксплуатации в пределах от $0,85 U_n$ до U_n . Поэтому при вычислении силы учитывают $U = 0,85 U_n$, а при вычислении температуры превышения учитывают полное U_n . Таким образом, значение F из (8-4) относится к случаю $U = 0,85 U_n$.

Диаметр проволоки.

Из выражения (8-2) следует:

$$q_m = \frac{\rho l_{cp} F \cdot 10^{-3}}{U}. \quad (8-6)$$

Здесь q_m — площадь сечения проволоки в $мм^2$;
 l_{cp} — средняя длина витка в $мм$, которая может быть вычислена по размерам сердечника и катушки.

Диаметр голой проволоки $d = \sqrt{\frac{4q_m}{\pi}}$. Полученное значение может не соответствовать шкале нормальных диаметров по ГОСТ. В таких случаях принимают ближайшее значение по шкале (см. главу 2).

8-4. ПОВЕРОЧНЫЙ РАСЧЕТ СИСТЕМЫ ПОСТОЯННОГО ТОКА

В поверочном расчете выясняют правильность принятых в предварительном расчете данных, уточняют заданные соотношения, определяют оптимальное значение величины B_p . При несоответствии полученных результатов с заданными и оптимальными производят новый вариант расчета с измененными параметрами.

Катушка. Принимают по условиям эксплуатации или производства конструктивную форму катушки — каркасную или бескаркасную, и материал изоляции проволоки.

Определяют по изоляции катушки размеры намотки l_n и h_n , уточняют значения коэффициента заполнения, среднего значения витка.

Число витков катушки

$$\omega = f_n \frac{4l_n h_n}{\pi d^2}. \quad (8-7)$$

н. с. $F = \frac{U}{r} \omega$; значение r определяют с учетом ρ (горячее).

Температура превышения уточняется по полученным уточненным данным катушки (см. 2-21). Если температура превышения получается выше допустимой, необходимо пересчитать катушку с ближайшим по шкале меньшим диаметром проволоки.

Проводимость воздушных зазоров. Вычисляют проводимости: G_p — рабочего потока, G_n — краевого, G_n — нерабочих, G_{yn} — приведенного значения потока утечки.

Вычисляют значение коэффициента рассеяния

$$\sigma = 1 + G_{yn} \frac{G_p + G_n + G_n}{(G_p + G_n) G_n}. \quad (8-8)$$

Построение кривой намагничивания. Строится кривая $F = f(F_r + F_n)$ для начального положения якоря (фиг. 3, 26), на оси абсцисс откладывается значение F , соответствующее $0,85I\omega$, где $I\omega$ — н. с.,

соответствующая полному напряжению сети, в точке a проводится прямая под углом α . Значение α определяется из условия

$$\operatorname{tg} \alpha = \frac{G_p + G_{\kappa}}{p},$$

где p — коэффициент, представляющий собой отношение масштабов Φ и F .

Полученное из диаграммы значение F_p учитывается при вычислении силы по уточненной формуле

$$Q_s = 5,1 F_p^2 \frac{dG}{d\delta}. \quad (6-13)$$

Если полученное значение Q_s больше или меньше требуемого, производят изменения в конструкции магнитной системы.

В первую очередь проверяют, получены ли оптимальные данные системы (см. параграф 5-7) и для того чтобы получить их, меняют диаметр полюсного наконечника.

Построение тяговой характеристики. После того как получены окончательные данные расчета, необходимо построить тяговую характеристику $Q_s = f(\delta)$. Вычисляют значение Q_s для нескольких (трех, четырех) точек хода якоря, для чего строят соответствующие кривые намагничивания. Для каждого положения якоря вычисляют σ , F_p и $\frac{dG}{d\delta}$.

8-5. ПРИМЕР РАСЧЕТА СИСТЕМЫ ПОСТОЯННОГО ТОКА

Требуется разработать систему для следующих условий:

начальный момент $M = 2 \text{ кг} \cdot \text{см}$;

угол поворота якоря $\alpha = 9^\circ$;

номинальное напряжение $U_n = 110$ вольт. Это напряжение в процессе эксплуатации может снижаться до $U = 0,85U_n$.

а) Предварительный расчет. В предварительном расчете необходимо задаться некоторыми величинами:

B_p — индукцией в рабочем воздушном зазоре;

B_c — индукцией в стали;

$F_c + F_n$ — потерей м. д. с. в стали и нерабочих зазорах при опущенном якоре;

$k = \frac{l_{\kappa}}{h_{\kappa}}$ — отношением длины катушки к толщине намотки.

Все эти величины зависят от силы $Q_{\text{ан}}$ и хода якоря δ_n , отнесенным к среднему радиусу R_c (фиг. 8, 3).

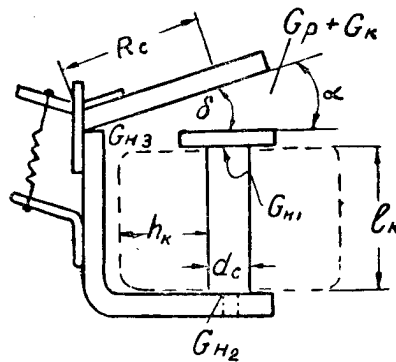
При заданном моменте $M_{\text{ан}}$, принимая различные значения δ_n , можно получить различные значения R_c и, таким образом, подобрать наиболее выгодное соотношение между Q и δ , соответствующее лучшим показателям работоспособности системы клапанного типа (см. параграф 5-4).

Примем $\delta_n = 0,4 \text{ см}$.

Так как зазор $\delta = R_c \cdot \alpha$; $Q_{\text{ан}} \cdot R_c = M_{\text{ан}}$, то

$$Q_{эн} = \frac{M_{эн}}{R_c} = \frac{M_{эн} \alpha}{\delta} = \frac{2 \cdot 9}{57,4 \cdot 0,4} = 0,78 \text{ кг}.$$

При таком соотношении $Q_{эн}$ и δ_n задаются значением $B_p = 0,16 \cdot 10^{-4}$.



1. Пользуясь формулой $Q = 4,06 \cdot 10^8 B_p^2 q$ определяем поверхность q_{nn} — полюсного наконечника:

$$q_{nn} = \frac{0,78}{4,06 \cdot 0,16^2} = 7,6 \text{ см}^2.$$

Принимаем $d_{nn} = 3,0 \text{ см}$, тогда $q_{nn} = 7,06 \text{ см}^2$.

Вследствие уменьшения значения q_{nn} с 7,6 до 7,06 увеличивается индукция B_p до $0,1665 \cdot 10^{-4}$.

Фиг. 8, 3. Эскиз магнитной системы.

2. Для определения сечения сердечника задаемся значением $B_c = 1,1 \cdot 10^{-4}$; материал — слабоуглеродистая сталь; коэффициент рассеяния системы $\sigma = 2$.

Площадь сечения сердечника

$$q_c = \frac{B_p q_{nn} \cdot \sigma}{B_c} = \frac{0,1665 \cdot 10^{-4} \cdot 7,06 \cdot 2}{1,1 \cdot 10^{-4}} = 2,14 \text{ см}^2.$$

Принимаем $d = 1,7 \text{ см}$; тогда $q_c = 2,26 \text{ см}^2$.

Ярмо-скобу, а также якорь выполняем из полосового материала $5 \times 40 \text{ мм}$.

$$q_a = 2,0 \text{ см}^2.$$

3. Для определения н. с. системы задаемся падением магнитного потенциала в стали и нерабочих зазорах в разомкнутой системе. Учитывая, что в данной конструкции имеется три нерабочих зазора, принимаем несколько повышенное значение потерь

$$F_c + F_n = 0,3F,$$

$$F = \frac{B_p \cdot \delta}{\mu_0 (1 - \alpha)} = \frac{0,1665 \cdot 0,4 \cdot 10^{-4}}{1,25 \cdot 0,7 \cdot 10^{-8}} = 760.$$

При вычислении моментов на якоре учитывают н. с., получающиеся при пониженном напряжении $U = 0,85U_n$. Этому напряжению будет соответствовать н. с. F_1 .

При вычислении превышения температуры учитывают F , соответствующую полному напряжению

$$F = \frac{F_1}{0,85} = \frac{760}{0,85} = 895.$$

4. Размеры катушки находим так: принимаем материал намотки — медь с эмалевой изоляцией, превышение температуры $\theta = 60^\circ$,

катушка бандажированная, коэффициент теплоотдачи по шкале А (табл. 8).

$$\mu = 11,92 \cdot 10^{-4}.$$

Отношение длины катушки к её высоте

$$k = \frac{l_k}{h_k} = 6.$$

$$h_k = \sqrt[3]{\frac{23,39 \cdot 10^{-3} \cdot 895^2}{20 \cdot 60 \cdot 11,92 \cdot 10^{-4} \cdot 0,45 \cdot 30}} = 9,32 \text{ мм.}$$

Принимаем $h_k = 10 \text{ мм.}$

5. На основании полученных данных выполняется эскиз магнитной системы (фиг. 8, 3).

В данной конструкции величина $R_c = \frac{\delta}{a} = \frac{4 \cdot 57,4}{9} = 25,6 \text{ мм.}$

б) Поверочный расчет. 1. Проводимости воздушных зазоров вычисляются обычно для трех положений якоря: начального, промежуточного и конечного — при притяннутом якоря.

При вычислении проводимостей рабочего зазора при притяннутом якоря (0°) величина δ учитывается как сумма: величины минимального возможного зазора $0,05 \text{ мм}$ и толщины цинкового покрытия на полюсном наконечнике и якоря $0,03 \text{ мм}$ (считая по 15 микрон на сторону).

Такая же величина зазора учитывается при вычислении проводимостей нерабочих зазоров — в стыках между полюсным наконечником и сердечником, между сердечником и ярмом. Для других положений якоря отдельно вычисляем проводимости рабочего зазора: G_p — рабочую и G_k — краевую.

2. Коэффициенты рассеяния определяются по проводимостям. Полный коэффициент рассеяния

$$\sigma = \frac{G_p + G_k + G_{yn}}{G_p}.$$

Коэффициент рассеяния рабочего зазора

$$\sigma_k = \frac{G_p + G_k}{G_p}.$$

Ниже приведена сводная таблица данных проводимостей и коэффициентов рассеяния, обозначения соответствуют обозначениям фигуры 8, 3.

Положение якоря	$10^8 G_p$	$10^8 G_k$	$10^8 G_{н1}$	$10^8 G_{н2}$	$10^8 G_{н3}$	$10^8 G_{yn}$	σ	σ_k
9°	22,06	11,35	353	1000	112	12	2,06	1,51
3°	66,18	15,2	353	1000	206	12	1,41	1,23
0°	1100		353	1000	385	12		

3. Для построения кривой намагничивания (для начального положения якоря) задаемся значениями рабочего потока Φ_p :

$$1 \cdot 10^{-4}, 1,5 \cdot 10^{-4}, 1,7 \cdot 10^{-4} \text{ вебер.}$$

Методика подсчета $F_c + F_n$ приведена в главе 3.

Строим кривую намагничивания $\Phi_p = f(F)$.

Отношение масштабов Φ и F равно $20 \cdot 10^{-8}$;

$$\operatorname{tg} \alpha = \frac{G_p}{p} = \frac{22,06}{20} = 1,1; \quad \alpha = 49^\circ.$$

Получаем

$$\Phi_p = 1,16 \cdot 10^{-4} \text{ вб}; \quad Q_{эн} = \frac{4,06 \cdot 10^8 \cdot 1,16^2 \cdot 10^{-8}}{7,06} = 0,735 \text{ кГ.}$$

Несколько сниженное по сравнению с расчетным (0,78 кГ) значение силы получилось из-за повышенного значения $F_c + F_n$.

Проверка оптимальности принятого значения индукции B_p производится по кривой намагничивания (см. главу 5).

$$\text{Площадь треугольника } ab_1c_1 = \frac{11,6 \cdot 510}{2} = 2850;$$

$$\text{» » } ab_2c_2 = 3070;$$

$$\text{» » } ab_3c_3 = 3000.$$

Оптимальные условия соответствуют потоку $\Phi_p = 1,28 \cdot 10^{-4} \text{ вб}$, который получается при угле $\alpha_2 = 52^\circ$; $\operatorname{tg} 52^\circ = 1,28$:

$$G_p = 1,28 \cdot 20 \cdot 10^{-8} = 25,6 \cdot 10^{-8} \text{ вб/ав.}$$

Этому условию соответствует полюсный наконечник диаметром 32 мм.

Таким образом, надо принять

$$d_{пн} = 32 \text{ мм}; \quad q_{пн} = 8,04 \text{ см}^2.$$

При этом условии сила на якоре

$$Q_{эн} = \frac{1,28^2 \cdot 10^{-8}}{8,04} \cdot 4,06 \cdot 10^8 = 0,825 \text{ кГ.}$$

Получается некоторый запас в силе, благодаря чему возможно снизить величину н. с. системы.

4. Определение числа витков и диаметра проволоки катушки. Размеры катушки: $D_n = 38 \text{ мм}$; $D_e = 18 \text{ мм}$; $l_n = 60 \text{ мм}$. Средняя длина витка

$$l_{cp} = \frac{\pi(38 + 18)}{2} = 88 \text{ мм.}$$

Сечение q_m меди катушки определяем из уравнения напряжения

$$U_n = Ir = \rho \frac{l_{cp}}{q} I \omega \cdot 10^{-3};$$

$$q_m = \frac{\rho l_{cp} I \omega \cdot 10^{-3}}{U_n} = \frac{23,39 \cdot 10^{-8} \cdot 88 \cdot 10^{-3} \cdot 895}{110} = 16,7 \cdot 10^{-3} \text{ мм}^2.$$

Наиболее близкий диаметр проволоки

$d = 0,15 \text{ мм}$; $q_m = 17,6 \cdot 10^{-8} \text{ мм}^2$,
 d_1 проволоки с изоляцией равно $0,17 \text{ мм}$.

Коэффициент заполнения катушки f_n :

$$f_n = k_v \frac{q_m}{d_1(d_1 + \delta)} = 0,9 \frac{17,6 \cdot 10^{-8}}{0,17 \cdot 0,20} = 0,463.$$

Задаваясь размерами изоляции катушки, получаем размеры катушки

$l_n = 56 \text{ мм}$, $h_n = 8 \text{ мм}$.

Тогда

$$\omega = f_n \frac{l_{cp} \cdot \omega}{q_m} = 0,463 \cdot \frac{56,8}{17,6 \cdot 10^{-8}} = 11800;$$

$$r = \rho \frac{l_{cp} \cdot 10^{-8}}{q_m} \omega = 1370 \text{ ом};$$

$$\text{ток } I = \frac{U}{r} = \frac{110}{1370} = 0,08 \text{ а};$$

$$F = 0,08 \cdot 11800 = 950 \text{ ав}.$$

Потери в катушке

$$P = \frac{U^2}{r} = \frac{110^2}{1370} = 8,8 \text{ вт}.$$

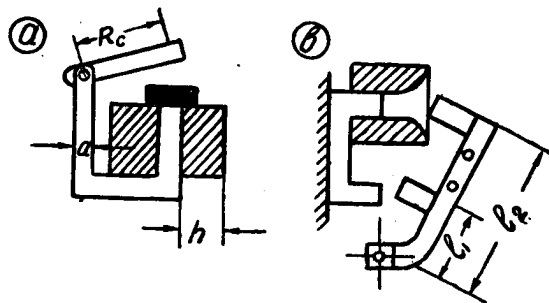
5. Для получения данных $Q_s = f(\delta)$ требуется построить кривые намагничивания для промежуточного и конечного положений якоря.

6. Для вычисления коэффициента возврата строят кривую размагничивания и определяют н. с. отпуска (метод вычисления коэффициента возврата см. в главе 5).

8-6. ВОПРОСЫ ПРОЕКТИРОВАНИЯ СИСТЕМ ПЕРЕМЕННОГО ТОКА

Формы магнитных систем. Ниже рассматриваются наиболее распространенные системы с поворотным якорем.

1) Фиг. 8, 4 а. Обычная форма, применяемая для исполнения разных реле управления. Магнитопровод шихтованный из листов электротехнической стали, ось вращения якоря располагается против ярма, экран — на сердечнике.



Фиг. 8, 4. Формы систем переменного тока с поворотным якорем.

2) Фиг. 8, 1 г. Форма исполнения для промежуточных

реле с контактами малой разрывной мощности. Особенность системы в том, что она выполнена из листовой кремнистой стали

2 мм. Такая же форма применяется и для систем постоянного тока, в этом случае экран удаляется.

3) Фиг. 8, 4 *b* — система с П-образным якорем и ярмом. Применяется для исполнения реле управления и контакторов переменного тока.

4) Для многоамперных контакторов переменного тока применяется система с Е-образным якорем и ярмом.

Вопросы исполнения магнитных систем. 1) Сечение магнитопровода. По условиям работы катушки магнитопровод лучше всего выполнять с квадратным сечением. Однако часто предпочитают делать его прямоугольным, так как при таком исполнении якорь в притянутом состоянии обладает большей устойчивостью в отношении вибраций при возникающих в конструкции боковых составляющих моментов пружин.

2) Зазоры. Во избежание залипания якоря необходимо выполнить магнитную систему так, чтобы при притянутом якоря оставались нерабочие воздушные зазоры порядка 0,2—0,5 мм: в системе, изображенной на фигуре 8, 4 *a*, этот зазор предусматривается у основания якоря; в системе, приведенной на фигуре 8, 1 *g* — у конца якоря — зазор 3, в Е-образной системе — в средней части якоря. В месте соприкосновения якоря с экранированным полюсом зазор должен быть доведен до минимума, для чего места соприкосновения отшлифовывают.

Исходные данные. Исходными являются следующие данные: рабочее напряжение, начальное значение силы или момента, начальный ход якоря или угол поворота. Необходимо учитывать, что максимальное значение условной работоспособности *A* в системах переменного тока получается при больших значениях хода якоря, чем у систем постоянного тока. Рекомендуется выполнять связь между электромагнитом и исполнительным механизмом так, чтобы угол поворота якоря получился максимально возможным (фиг. 5, 3).

Ход расчета. 1) Строится диаграмма механических сил

$$Q_n = f(\delta) \quad \text{или} \quad M_n = \varphi(\theta).$$

2) Полученная диаграмма перестраивается по отношению к силам или моментам на якоря: получают диаграмму, приведенную к якорю.

3) Проводят предварительные расчеты системы и устанавливают основные размеры ее.

4) Проводят окончательный расчет для уточнения данных предварительного. Если получаются неудовлетворительные результаты, пересчитывают систему, задавшись другими значениями параметров.

8-7. ОСНОВНЫЕ СООТНОШЕНИЯ ДЛЯ ПРЕДВАРИТЕЛЬНОГО РАСЧЕТА

Уравнение силы. В предварительных расчетах удобнее пользоваться формулой Максвелла (6-8)

$$Q_s = 2,03 \cdot 10^8 B_p^2 q_p,$$

где B_p — амплитудное значение индукции в рабочей части воздушного зазора;

q_p — сечение рабочей части зазора, $см^2$.

Оптимальное значение B_p колеблется в больших пределах. Для ориентировки в выборе B_p можно пользоваться кривыми, изображенными на фигуре 5, 5, но при этом следует учитывать, что величина к. ф. должна определяться для систем переменного тока из условия $\frac{\sqrt{2Q}}{\delta}$. В системе, приведенной на фигуре 8, 4 а имеется один рабочий зазор и сила определяется величиной магнитного потока в этом зазоре. Сила в нерабочем зазоре не учитывается. В других, более сложных магнитных системах, например Ш- и Е-образных, имеется два или три рабочих зазора и сила определяется величиной потока во всех этих зазорах.

В этих случаях также применима формула (6-8), где значения B_p и q_p относят к одному зазору. Так, в П-образной системе с осью вращения вне полюса при заданном моменте можно определить следующим образом приведенное значение Q (фиг. 8, 4b):

$$M = Q_1 l_1 + Q_2 l_2.$$

Принимая одинаковыми сечения стали полюсов и не учитывая потоков рассеяния, можно считать $Q_1 = Q_2 = 2,03 \cdot 10^8 B_p^2 q$. Тогда приведенное значение Q :

$$Q = \frac{M}{l_1 + l_2}.$$

В Ш-образной системе (фиг. 8, 1g) не учитывают в разомкнутом положении якоря силу у полюса 2.

Коэффициент рассеяния. В предварительном расчете необходимо задаваться значением коэффициента рассеяния σ , причем краевые потоки рабочего зазора также относят к потокам рассеяния. Значение σ колеблется в пределах 1,5 ÷ 3, меньшее значение получается в системах с малым углом поворота якоря, большее — с большим углом поворота (порядка 12—15°).

Отношение $k = \frac{l_k}{h_k}$ принимается в зависимости от величины к. ф.: большему значению к. ф. соответствует меньшая величина отношения.

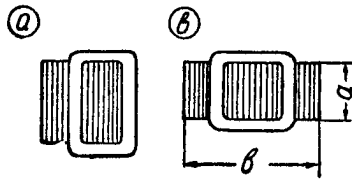
8-8. ПРЕДВАРИТЕЛЬНЫЙ РАСЧЕТ СИСТЕМЫ ПЕРЕМЕННОГО ТОКА

Схема расчета. а) Заданные величины: U_n , начальное значение момента противодействующих сил M_{nn} , θ_n .

б) Для предварительного расчета необходимо определить значение силы $Q_{эн} = \frac{M_{nn}}{R_c}$ (фиг. 8, 4 а). Величина R_c определяется из конструктивных соображений. Учитывая, что в системах переменного тока условная работоспособность $A = Q\delta$ получается большей при относительно больших зазорах, выгоднее увеличивать

$\delta = R_c\theta$ и уменьшать $Q_{эн}$. Принимают $R_c = (1,2 - 1,5) R_k$ (фиг. 5, 3).

в) Площадь сечения рабочей части воздушного зазора. Задавшись значением B_p , вычисляют q из (6-8):



Фиг. 8, 5. Размещение экрана.

$$q = \frac{Q_{эн}}{2,03 \cdot 10^8 B_p^2}.$$

Полное сечение q_c сердечника получается несколько меньшим. С учетом k_c коэффициента заполнения пакета (порядка 0,9—0,95)

$$q_c = k_c q \text{ кв. см.}$$

г) Поток в системе. Учитывается полное значение потока, для чего задаются σ — коэффициентом рассеяния.

$$\Phi = \sigma k_c q B_p \text{ вебер.}$$

д) Число витков катушки. Принимают, что при разомкнутом якоре $E \cong (0,97 \div 0,98) U$; далее учитывают, что требуемое значение силы на якоре должно иметь место при пониженном напряжении

$$U = 0,85 U_n.$$

Получаем

$$E = 0,98 \cdot 0,85 U_n = \frac{\omega \Phi w}{\sqrt{2}};$$

$$w = \frac{0,98 \cdot 0,85 \sqrt{2} U_n}{\omega \Phi}.$$

е) Размеры сердечника в плане. Площадь сечения сердечника $q = ab$. В электромагнитах небольшой мощности стремятся к квадратному сечению, благодаря чему уменьшается средняя длина витка. Экран размещается согласно изображению на фигурах 8, 5 а и 8, 5 б. В мощных электромагнитах, в целях большей устойчивости якоря при притянutom его положении, принимают $b > a$.

ж) Катушка. В электромагнитах переменного тока применяются каркасные катушки, при тяжелых условиях работы (повторно-кратковременный режим, условия среды) применяют бескаркасные катушки. Размер h катушки определяется из следующих условий:

принимая сечение ярма равным сечению сердечника, получают $c = R_c - 2a$ (фиг. 8, 4), далее, учитывая необходимость свободной посадки катушки, размер h принимают несколько меньшим размера c (на 0,2—0,3 см).

В зависимости от значения к. ф. $= \frac{\sqrt{2Q}}{\delta}$ принимают отношение l_n/h_n .

8-9. ПОВЕРОЧНЫЙ РАСЧЕТ СИСТЕМЫ ПЕРЕМЕННОГО ТОКА

Катушка. Устанавливаются размеры намотки l_n и h_n , которые получаются из размеров катушки l_k и h_k , за вычетом изоляции. Диаметр проволоки определяют тремя этапами:

1) Задаются коэффициентом заполнения намотки f_n , в среднем порядка 0,5 (см. расчет катушек). Из соотношения $\omega = f_n \frac{4l_n \cdot h_n}{\pi d^2}$ определяют d голой меди.

2) Проверяется значение f_n . При исполнении катушки с прокладкой бумаги через слой, учитывая d_1 — диаметр проволоки с изоляцией

$$f_n = k_v \frac{\pi d^2}{4d_1(d_1 + \delta)}$$

3) С уточненным значением f_n снова вычисляется диаметр проволоки. Принимается значение d , ближайшее по ГОСТ.

Проводимости воздушных зазоров. Вычисляются проводимость $G_s = G_p + G_n$, нерабочего зазора G_n и приведенная по числу потосцеплений проводимость утечки G_{yc} . Далее определяется производная $\frac{dG_s}{d\theta}$. Для вычисления σ — коэффициента рассеяния, пользуются схемой замещения. Вычисления проводимостей и σ проводят для нескольких точек, например, θ_1 , θ_2 и $\theta = 0$; данные расчета сводят в таблицу. При $\theta = 0$ обычно считают величину рабочего зазора 0,005 см; нерабочего 0,02 см.

Значение момента в начальном положении якоря. Для θ_n

$$M_{эн} = \frac{5,1 \cdot 0,85^2 U^2}{\omega^2 \omega^2 \sigma G_s} \frac{dG_s}{d\theta} \text{ кг см.}$$

Если полученное значение $M_{эн}$ отличается от заданного, меняют число витков катушки с тем, чтобы приблизиться к значению $M_{эн}$. При изменении ω вновь определяют диаметр проволоки. Необходимо стремиться к полному заполнению окна катушки, чтобы не ухудшать условий ее охлаждения.

Значение момента при угле θ_2 . При проверке этого значения $M_{э2}$ должен быть больше $M_{н2}$. В противном случае необходимо увеличивать значение $M_{эн}$.

Значение силы при $\theta \cong 0$ (конечное положение якоря). В этом случае удобнее пользоваться формулой Максвелла. Без учета действия экрана среднее значение силы:

$$Q_{э0} = 2,03 \cdot 10^8 \frac{\Phi^2}{\sigma^2 q'} \text{ кг.}$$

В данном случае $q' < q$, так как в магнитопроводе предусматривается вырез (порядка 2 мм) для укладки экрана.

Расчет экрана (см. главу 7).

Токи в катушке. Расчет проводится при U_n . Вычисляются токи при разомкнутом и замкнутом положениях якоря (см. главу 3).

Нагрев катушки. Из уравнения нагрева $P = \mu S \theta$, где P — полные потери $P = P_k + P_n + P_s$, определяют температуру превышения θ .

Проверка эффективности использования магнитной системы проводится из соотношения

$$k = \frac{I_n - I_0}{I_n},$$

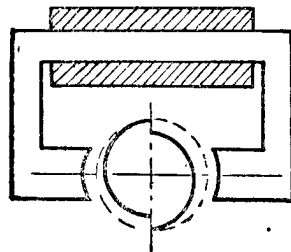
где I_n и I_0 — соответственно значения токов при разомкнутом и замкнутом положениях якоря.

ГЛАВА 9

МАГНИТНЫЕ СИСТЕМЫ С ПОПЕРЕЧНЫМ ДВИЖЕНИЕМ ЯКОРЯ

9-1. ФОРМЫ МАГНИТНЫХ СИСТЕМ

Системы с поперечным движением якоря отличаются тем, что якорь в них движется в поперечном направлении к средней линии между полюсами. Применяются такие системы в тех случаях, когда требуется получить высокое значение коэффициента возврата, например при проектировании автоматических регуляторов или специальных реле.



Фиг. 9, 1. Эскиз системы.

Особенностью рассматриваемых систем является то, что в них рабочий воздушный зазор между якорем и полюсом имеет небольшое значение при всем ходе якоря. В этом случае закон изменения профиля якоря определяет собой закон изменения момента. Изменяя профиль якоря можно получить любые требуемые виды характеристик. Системы с поперечным движением якоря применяют предпочтительно в устройствах постоянного тока, так как наличие в системе постоянного зазора между полюсами и якорем затрудняет условия устранения вибраций, возникающих в системах переменного тока. Практическое применение получили следующие формы магнитных систем.

а) Системы со сплошным якорем (фиг. 9, 1), применяющиеся при углах поворота якоря до $100-110^\circ$. Для получения требуемых тяговых характеристик якорю придают различные профили: арифметической спирали, гиперболы, параболы и другие.

б) Системы с выступающим якорем (фиг. 9, 5), которые применяются при углах поворота порядка $25-40^\circ$.

в) Системы с втягивающимся якорем (фиг. 9, 8), применяющиеся при небольших углах поворота якоря ($10-15^\circ$), но в случаях, когда требуется получение тяговой характеристики особого вида.

9-2. ОСНОВНЫЕ СООТНОШЕНИЯ И ДАННЫЕ

Рассматриваются только системы $F = \text{пост.}$

Моменты на якоре: $M = 5,1 F_p^2 \frac{dG}{d\theta}$ кГ см. Полная проводимость рабочих воздушных зазоров

$$G_s = G_p + G_n; \quad \frac{dG_s}{d\theta} = \frac{dG_p}{d\theta} + \frac{dG_n}{d\theta}.$$

В системах с поперечным движением якоря значение $\frac{dG_n}{d\theta}$ оказывает влияние на величину момента только в крайних положениях якоря, когда выступающая его часть находится вблизи полюса. В рабочем диапазоне положения якоря

$$\frac{dG_n}{d\theta} \ll \frac{dG_p}{d\theta}.$$

Поэтому в практических расчетах учитывают только $\frac{dG_p}{d\theta}$.

Индукция B_p . Значение B_p меняется с ходом якоря. В среднем для начального положения якоря индукция в сечении полюса принимается порядка $(0,3 \div 0,4) 10^{-4} \frac{\text{вб}}{\text{см}^2}$.

Полюсная дуга α (фиг. 9, 2). Значение α принимается: в системах со сплошным якорем порядка $45-70^\circ$, с выступающим якорем $60-65^\circ$.

Рабочий угол поворота выбирается из условия устранения влияния краевого потока, сказывающегося при крайних положениях якоря: в системах со сплошным якорем порядка $100-110^\circ$, с выступающим $25 \div 40^\circ$.

Значения минимальных воздушных зазоров $\delta_{\text{мин}}$ принимаются из конструктивных условий порядка $0,3 \div 0,5$ мм.

9-3. СИСТЕМЫ СО СПЛОШНЫМ ЯКОРЕМ

В таких системах якорь очерчен по специальному профилю (фиг. 9, 2) радиусом $R_x = f(\theta)$, где θ угол отсчета от оси симметрии якоря.

Обозначим: R_0 — минимальный радиус, R_1 — максимальный. Зазор между полюсом, очерченным по окружности радиусом R , и якорем: $\delta_x = R - R_x = R - f(\theta)$. Минимальный зазор получается при R_1 : $\delta_0 = R - R_1$. Определим проводимость между полюсом и якорем. Выделим элементарный объем воздушного промежутка

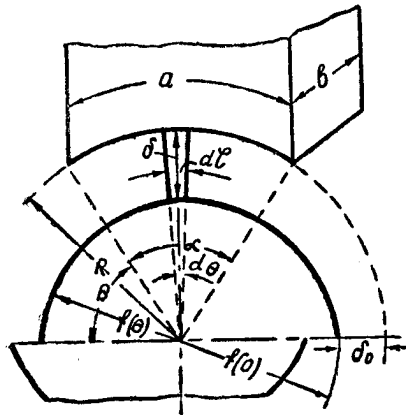
шириной в средней части $d\zeta$, длиной δ :

$$d\zeta = \left[f(\theta) + \frac{\delta}{2} \right] d\theta = \left[f(\theta) + \frac{R - f(\theta)}{2} \right] d\theta = 0,5 [R + f(\theta)] d\theta.$$

Элементарная проводимость зазора

$$dG_p = \mu_0 \frac{d\zeta b}{\delta}.$$

Полная проводимость между полюсом и якорем без учета проводимостей крайних потоков



$$G_p = 0,5\mu_0 b \int_{\theta}^{\theta+\alpha} \frac{R + f(\theta)}{R - f(\theta)} d\theta. \quad (9-1)$$

Для вычисления моментов необходимо знать производную проводимости по углу. Учитывая два последовательно включенных воздушных промежутка, имеем

$$\frac{dG_p}{d\theta} = 0,25\mu_0 b \frac{d}{d\theta} \int_{\theta}^{\theta+\alpha} \frac{R + f(\theta)}{R - f(\theta)} d\theta =$$

Фиг. 9, 2. Профиль полюса и якоря.

$$= 0,25\mu_0 b \left[\frac{R + f(\theta + \alpha)}{R - f(\theta + \alpha)} - \frac{R + f(\theta)}{R - f(\theta)} \right] =$$

$$= 0,5\mu_0 b R \frac{f(\theta + \alpha) - f(\theta)}{[R - f(\theta + \alpha)][R - f(\theta)]} \quad (9-2)$$

Выбор профиля якоря, т. е. закона изменения R_x , производится в зависимости от требуемого закона изменения тяговой характеристики $M_o = M(\theta)$.

$$M_o = 5,1 \cdot F_p^2 \frac{dG}{d\theta} = 2,55\mu_0 F_p^2 b R \frac{f(\theta + \alpha) - f(\theta)}{[R - f(\theta + \alpha)][R - f(\theta)]}. \quad (9-3)$$

Обозначим: $A = 2,55F_p^2 \mu_0 b R$,

тогда

$$M_o = A \frac{f(\theta + \alpha) - f(\theta)}{[R - f(\theta + \alpha)][R - f(\theta)]}. \quad (9-4)$$

Из (9-4)

$$f(\theta + \alpha) = \frac{A f(\theta) + R^2 M_o - R M_o f(\theta)}{A + R M_o - M_o f(\theta)}. \quad (9-5)$$

При $\theta = \pi$ получается минимальный зазор: $R - R_1$.

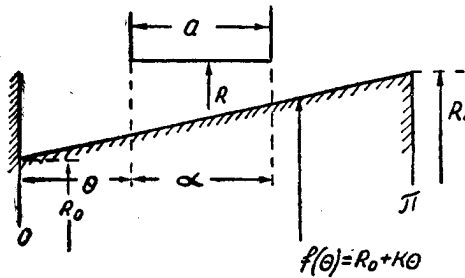
Для получения большей наглядности рассматривают профили полюса и якоря в развертке (фиг. 9, 3), где показана развертка якоря, очерченного по арифметической спирали.

$$R_x = f(\theta) = R_0 + k\theta. \quad (9-6)$$

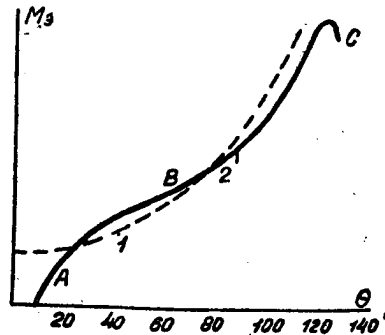
Подставляя (9-6) в (9-2), получаем:

$$\frac{dG_p}{d\theta} = 0,5\mu_0 b R \frac{k\theta}{[R - R_0 - k(\theta + \alpha)][R - R_0 - k(\theta)]} \quad (9-7)$$

Момент M_p изменяется по гиперболическому закону (см. пунктир на фиг. 9, 4); сплошной показана экспериментальная кривая, по-



Фиг. 9, 3. Развертка полюса и якоря.

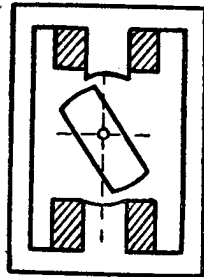


Фиг. 9, 4. Тяговые характеристики системы $F = \text{пост}$:
1 — без учета изменения F_p и действия краевого потока; 2 — экспериментальная.

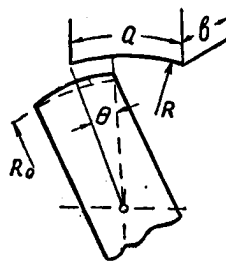
лученная в Харьковском политехническом институте инженером К. И. Диденко при проведении им исследования системы. Разница между расчетной и экспериментальной кривыми определяется следующими факторами: участки A и C действием краевого потока, участок B — изменением величины н. с. F_p , приходящейся на зазоры, по ходу якоря.

9-4. СИСТЕМЫ С ВЫСТУПАЮЩИМ ЯКОРЕМ

В таких системах допускается малый угол поворота. Профиль якоря может быть очерчен по окружности (фиг. 9, 5 и 9, 6) в том



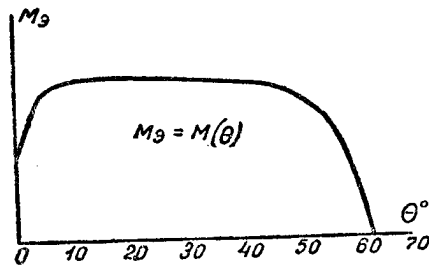
Фиг. 9, 5. Эскиз системы с выступающими полюсами.



Фиг. 9, 6. Профиль полюса и якоря.

случае, когда требуется получить приблизительно постоянное значение момента при рабочем угле поворота (фиг. 9, 7). Если же

требуется увеличение момента по ходу якоря, то профиль якоря очерчивается не по окружности, концентричной с профилем полюса, а по другой так, чтобы при втягивании якоря уменьшался воздушный зазор.



Фиг. 9, 7. Примерная кривая моментов системы с выступающими полюсами, $F = \text{const}$.

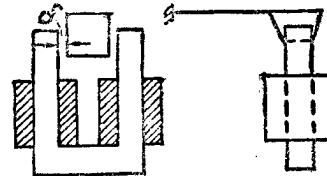
Для якоря, очерченного по окружности, концентричной с окружностью полюса, с учетом двух последовательно включенных зазоров

$$G_p = 0,25\mu_0 b \int_0^\theta \frac{R + R_0}{R - R_0} d\theta;$$

$$\frac{dG_p}{d\theta} = 0,25\mu_0 b \frac{R + R_0}{R - R_0}. \quad (9-8)$$

9-5. СИСТЕМЫ СО ВТЯЖНЫМ ЯКОРЕМ

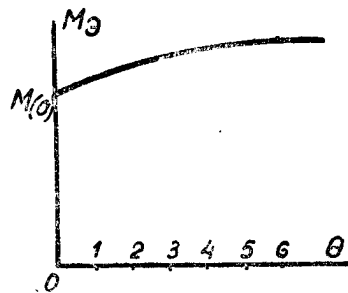
В таких системах якорь подвешивается на пружине (фиг. 9, 8). Рабочий угол поворота якоря должен быть взят таким, чтобы якорь не занимал крайних положений против полюса. Профиль якоря выбирается в зависимости от требуемой характеристики.



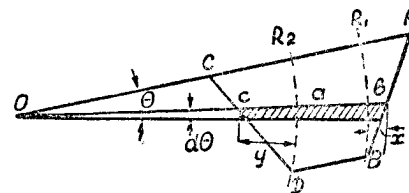
Фиг. 9, 8. Эскиз системы со втяжным якорем.

На фигуре 9, 9 приведена примерная кривая моментов для электромагнита астатического угольного регулятора, на фигуре 9, 10 — профиль якоря. Если якорь не занимает крайних положений, то можно не учитывать составляющей момента, определяемой изменением проводимости краевого поля.

Профиль якоря выполняется так: линия AB принимается прямой; закон изменения кривой CD определяется из требуемого закона изменения момента. На заданной по условиям работы регулятора кривой моментов устанавливаются значения M_0 , M_1 и т. д.



Фиг. 9, 9. Примерная кривая моментов системы со втяжным якорем, $F = \text{const}$.



Фиг. 9, 10. Примерный профиль якоря.

Рассмотрим элементарную трубку поля между якорем и полюсом, площадкой dS и зазором 2δ . Площадка dS может быть разбита на три части:

a — определяемую радиусами R_1 и R_2 ;

b — определяемую радиусом R_1 и линией AB ;

c — определяемую радиусом R_2 и кривой cD ;

$$dG_1 = \mu_0 \frac{dS_1}{2\delta}; \quad dS_1 = dS_a + dS_b + dS_c;$$

$$dS_1 = \left| \frac{R_1^2 - R_2^2}{2} \right| d\theta + R_1 x_1 d\theta + R_2 y_1 d\theta;$$

$$\frac{dG_1}{d\theta} = 0,5 \frac{\mu_0}{\delta} \left(\frac{R_1^2}{2} - \frac{R_2^2}{2} + R_1 x_1 + R_2 y_1 \right). \quad (9-9)$$

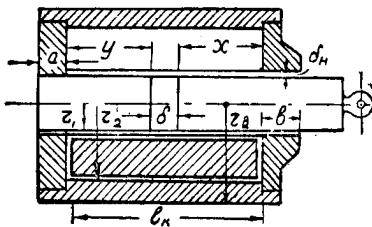
При заданном значении x_1 определяют y_1 .

Г Л А В А 10

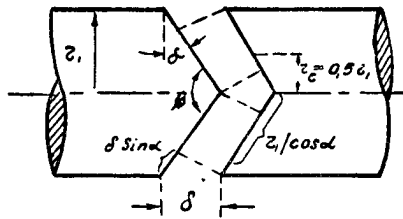
ПРЯМОХОДОВЫЕ МАГНИТНЫЕ СИСТЕМЫ

10-1. ФОРМЫ МАГНИТНЫХ СИСТЕМ ПОСТОЯННОГО ТОКА

Прямоходовые системы постоянного тока исполняются, как правило, в виде соленоидов. Поэтому такие системы часто называют соленоидными. В устарелых конструкциях автоматических регуляторов соленоидные системы выполнялись с открытым магнитопроводом. Практическое значение имеют в настоящее время



Фиг. 10, 1. Прямоходовая система.
Плоский стоп.



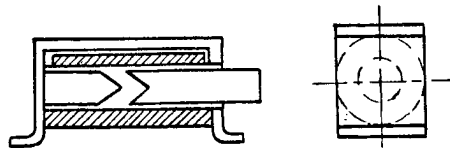
Фиг. 10, 2. Конический стоп.

системы с замкнутым магнитопроводом. По своему назначению прямоходовые системы могут быть разбиты на две основные группы.

а) **Короткоходовые системы** — предусматриваются для получения больших значений сил при относительно малом ходе якоря. Такие системы применяются при больших значениях конструктивного фактора $\frac{\sqrt{Q}}{\delta}$ в пределах 4—90. Короткоходовые системы всегда снабжаются стопом. На фигуре 10,1 приведен эскиз системы с плоским стопом, на фигуре 10,2 показано исполнение конического стопа.

Для исполнения реле небольших габаритов применяются упрощенные формы систем, изображенных на фигуре 10, 3.

б) **Длинноходовые системы** — предусматриваются для получения относительно больших ходов (фиг. 10, 4). Такие системы применяются при малых значениях конструктивного фактора. В длинноходовых системах сила на якоре определяется взаимодействием

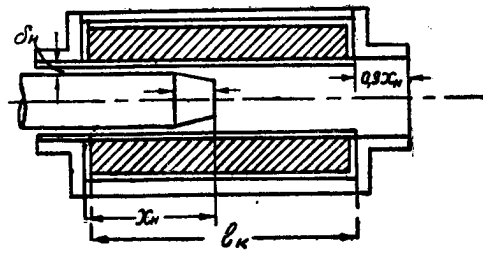


Фиг. 10, 3. Прямоходовая система упрощенного вида.

потока утечки с током в катушке: $Q = cI\Phi$. Поэтому в таких системах возможно получить любое желательное значение силы при повышении значения тока в катушке (индукция достигает своего максимального значения при насыщении). Это свойство

системы используется в установках, работающих при режиме кратковременной нагрузки, когда можно пропускать через катушку ток большой силы без опасения перегрева сверх допустимой по нормам температуры.

Другая особенность рассматриваемой системы: в определенной части хода якоря соленоидная сила является постоянной и не зависит от расстояния между входящим концом якоря и передним краем катушки. Поэтому можно выполнить соленоидную систему с любой величиной хода якоря, удлиняя соответственно катушку. Так, имеются исполнения длинноходовых соленоидных систем с величиной хода до 200 мм. Такая величина хода недостижима в системах поворотного типа.



Фиг. 10, 4. Длинноходовая система постоянного тока.

Значение хода якоря в длинноходовых системах бывает порядка $0,2 \div 0,6 l_k$.

10-2. СИЛЫ НА ЯКОРЕ

В короткоходовых системах учитывается только сила притяжения двух поверхностей — якоря и стопа. Соленоидная сила в таких системах относительно мала и ею пренебрегают.

Исходное уравнение для вычисления сил на якоре

$$Q = 5,1 F_p^2 \frac{dG_s}{db}, \quad (6-15)$$

где F_p — н. с. катушки за вычетом падения магнитного потенциала в стали и нерабочих зазорах,

G_s — проводимость рабочего зазора.

Упрощая (6-15), не учитывая влияния краевого потока, получаем формулу Максвелла, более удобную для расчетов, причем эта формула приводится к виду, в котором учитывается индукция в якоре, в конце его, у рабочего зазора.

Плоский стоп (фиг. 10, 1). Формула Максвелла

$$Q = 4,06 \cdot 10^8 B^2 \pi r_1^3 \text{ кг.} \quad (10-1)$$

В данном случае значения индукции в якоре B и B_p в рабочем зазоре совпадают.

Конический стоп (фиг. 10, 2). Для данного случая формулу Максвелла получаем, преобразуя (6-15). Проводимость рабочего зазора, без учета краевой:

$$G_p = \mu_0 \frac{\pi r_1^2}{\delta \cos^2 \alpha} - \mu_0 \frac{\pi \delta \sin^2 \alpha}{2}. \quad (10-2)$$

Второй член правой части (10-2) относительно мал. Поэтому примем:

$$G_p = \mu_0 \frac{\pi r_1^2}{\delta \cos^2 \alpha}. \quad (10-3)$$

Тогда

$$\frac{dG_p}{d\delta} = -\mu_0 \frac{\pi r_1^2}{\delta^2 \cos^2 \alpha}, \quad (10-4)$$

далее

$$F_p G_p = \Phi_p = B_p q_p. \quad (10-5)$$

Из (10-3), (10-4), (10-5) и (6-15) имеем

$$Q = \frac{5,1 \cdot 10^8 B_p^2 \cdot q_p^2 \cos^2 \alpha}{\mu_0 \pi r_1^2}. \quad (10-6)$$

Приведем (10-6) к индукции B в якоре. Поток в якоре, у рабочего зазора, равен потоку рабочего зазора

$$\Phi = \Phi_p = B_p q_p = B \pi r_1^2. \quad (10-7)$$

Из (10-6) и (10-7) имеем

$$Q = 4,06 \cdot 10^8 B^2 \pi r_1^2 \cos^2 \alpha. \quad (10-8)$$

10-3. ВЫБОР ФОРМЫ И ДЛИНЫ СТОПА

Форма стопа. Выбор формы стопа определяется значением конструктивного фактора для данного устройства. При большом значении к. ф., т. е. при больших значениях сил и малых — хода, рациональнее применять плоский стоп, при больших значениях хода — конический.

Практические данные: плоский стоп применяется при значении к. ф. в пределах $16 \div 90$, конический стоп с углом $\alpha = 45^\circ$ при к. ф. $= 4 \div 16$; с углом $\alpha = 60^\circ$ при к. ф. $= 1,85 \div 4$.

Длина стопа. Короткоходовые системы всегда снабжаются стопом, так как по условиям конструкции длина катушки системы в несколько раз больше величины рабочего зазора δ (фиг. 10,1). В рационально исполненных системах $\delta = 0,05 - 0,1 l_k$ и в редких случаях, в маломощных системах, $\delta = 0,12 \div 0,15 l_k$. Начальное значение x_n — величины вхождения якоря в катушку, принимается порядка $l_k/3$. Таким образом, длина стопа y получается в пределах

$$y = l_k - \frac{l_k}{3} - (0,05 \div 0,1) l_k. \quad (10-9)$$

10-4. ПРЕДВАРИТЕЛЬНЫЙ РАСЧЕТ КОРОТКОХОДОВОЙ СИСТЕМЫ ПОСТОЯННОГО ТОКА

Исходными являются следующие данные: рабочее напряжение U , рабочий ход δ , начальная сила на якоре Q_n , режим работы катушки — длительный, повторно-кратковременный.

При проведении предварительного расчета не учитывают соленоидной силы, а принимают во внимание только силу притяжения якоря к стопу.

Некоторыми значениями параметров, определяющих собой габариты системы, приходится задаваться: индукцией B в якоре рабочего зазора, падением магнитного потенциала в стали и нерабочих зазорах — $\sum H_i l_i$, температурой превышения катушки, отношением длины намотки к толщине ее. Правильный выбор указанных параметров определяет экономичность конструкции.

Индукция B . Значение B выбирается в зависимости от конструктивного фактора и от режима работы катушки: чем больше значение к. ф., тем большее значение B (см. кривые фиг. 5, 5), при кратковременном режиме работы принимают большее значение, чем при длительном.

Падение магнитного потенциала в стали и нерабочих зазорах. В рационально построенных конструкциях падение магнитного потенциала в стали магнитопровода составляет при начальном положении якоря 10—20% от величины н. с. катушки, а в нерабочих зазорах — 5—10%.

Следовательно,

$$F_n + F_c = (0,15 \div 0,3) F.$$

Обычно принимают значение индукции в стали кожуха равным 80% индукции в якоре.

Нерабочий зазор δ_n (фиг. 10, 1) определяется толщиной латунной направляющей трубки (порядка 0,5 мм) и зазором между трубкой и якорем. Этот зазор принимают порядка 0,125 мм для случая, когда внутренняя поверхность трубки не обработана, и 0,05 мм для трубки с внутренней обработанной поверхностью.

Для уменьшения магнитного сопротивления нерабочего зазора рекомендуется увеличивать размеры a и b (фиг. 10, 1).

Отношение $k = \frac{l_k}{h_k}$ бывает в пределах 5 ÷ 8 и выбирается также в зависимости от значения к. ф. (фиг. 5, 5). Толщина намотки h принимается из условия $m = \frac{r_2}{r_1} = 2 \div 2,5$, тогда

$$h = r_1(m - 1) \quad (10-10).$$

Схема расчета.

1) Уравнение силы.

$$Q = 4,06 \cdot 10^8 B^2 \pi r_1^2 \text{ кг} \quad (10-8)$$

При заданном значении Q и выбранном значении B определяют радиус якоря r_1 .

2) Уравнение магнитной цепи.

$$F = F_p + F_c + F_n = \frac{B\delta}{\mu_0} + (0,15 \div 0,3) F \quad (10-11)$$

Полученное значение н. с. катушки должно иметь место при напряжении $U = 0,85U_n$.

Полное значение н. с. катушки, которое надо учитывать при расчетах теплового режима, получается равным $\frac{F}{0,85}$.

3) Уравнение нагрева. Для круглых катушек

$$\theta = \rho_{20} \frac{(F/0,85)^2}{20 f_{\kappa \mu} h_{\kappa} l_{\kappa}^2} \quad (10-12)$$

Из (10-12) определяют l_{κ} и h_{κ} . Для этого требуется принять значения θ , $\frac{l_{\kappa}}{h_{\kappa}}$, f_{κ} — порядка 0,5. По принятому значению θ определяют ρ и μ (см. главу 2), значение F — из (10-11). Из вычисленного значения h_{κ} определяют радиус кожуха $r_2 = h_{\kappa} + r_1$; из l_{κ} определяют величину y (10-9). Значение x_n принимается равным $\frac{l_{\kappa}}{3}$.

4) Уравнение электрической цепи:

$$U = Ir = I\rho \frac{l_{cp}}{q} \omega = \frac{4\rho(r_1 + r_2)}{d^2} I\omega 10^{-3}, \quad (10-13)$$

$I\omega$ — соответствует полному значению н. с. катушки. Из (10-13) определяется диаметр проволоки намотки, а затем вычисляется число витков катушки.

10-5. ПРОВЕРОЧНЫЙ РАСЧЕТ КОРОТКОХОДОВОЙ СИСТЕМЫ ПОСТОЯННОГО ТОКА

На основании данных предварительного расчета устанавливаются основные размеры системы. Задачей проверочного расчета является уточнение параметров системы, проверка системы в отношении оптимального решения, построение тяговой характеристики.

После уточнения данных системы иногда приходится проводить расчет варианта с измененными параметрами. Число пересчетов,

необходимых для получения удовлетворительных результатов, зависит от того, насколько удачно выбраны в предварительном расчете значения параметров.

Схема расчета. 1) Уточняют данные катушки — коэффициент заполнения f_k , число витков w , сопротивление обмотки, температуру превышения.

2) Проводят расчет магнитной цепи. Для этого вычисляют проводимости всех путей в воздухе, определяют коэффициенты рассеяния и строят кривую намагничивания

$$\Phi_a = f(F_c + F_n).$$

3) Пользуясь кривой намагничивания, устанавливают рациональность принятых данных, дают им принятые значения. Оптимальное решение задачи см. в параграфе 5-7. Если полученное решение не дает удовлетворительных результатов, меняют значение индукции B в стали y рабочего зазора, что влечет за собой изменение площади сечения якоря, меняют также размеры немагнитных зазоров. В некоторых случаях приходится проверять вопрос о рациональности принятой формы стопа — плоской или конической.

4) После полного уточнения всех данных системы строится тяговая характеристика по нескольким характерным точкам, например начальной, промежуточной и конечной. При этом, ввиду изменения по ходу якоря коэффициента рассеяния, приходится строить несколько кривых намагничивания.

При вычислении сил учитывают:

Силу Q_1 — притяжения якоря к стопу

$$Q_1 = 5,1 F_p^2 \frac{dG_a}{d\delta},$$

где $\frac{dG_a}{d\delta}$ — производная проводимости с учетом изменения как рабочего, так и краевого поля;

F_p — определяется из кривой намагничивания.

Силу Q_2 — соленоидную определяют из соотношения

$$Q_2 = 10,2 H \Phi_y, \quad (6-25)$$

где

$$H = \frac{Iw}{l_k}.$$

При вычислении Φ_y необходимо учитывать смещение нейтрали (фиг. 3, 16), которая приближается к рабочему зазору по мере уменьшения последнего. В результате с уменьшением δ резко падает значение потока $\Phi_y = \Phi_{y1} - \Phi_{y11}$, а следовательно, и значение соленоидной силы. Характеристика соленоидной силы имеет вид кривой с максимумом, соответствующим определенной величине вхождения якоря в катушку. Величиной Q_2 при ходе якоря $\delta < 0,1l_k$ практически можно пренебречь.

10-6. РАСЧЕТ ДЛИННОХОДОВОЙ СИСТЕМЫ ПОСТОЯННОГО ТОКА

Общие соображения. Сила на якоре в длинноходовых системах возникает в результате взаимодействия потоков Φ_y с токами в катушке. Катушка в таких системах получается длинной. Поэтому каркасом для нее принимают необходимую для направления движения якоря латунную трубку с боковыми стальными фланцами (фиг. 10, 4). Входящий конец якоря обычно выполняют суживающимся, с углом сужения порядка $8-9^\circ$, максимальная длина суженной части принимается порядка $\frac{x_n}{3}$.

Исходные данные: рабочее напряжение U , рабочий ход x_p , начальная сила на якоре Q_n , режим работы катушки.

Принимаемые величины:

1) Индукция B в начальном положении якоря — у перехода от фланца к якору — принимается высокой — порядка 90% индукции насыщения материала.

2) Значение падения магнитного потенциала в стали и нерабочем зазоре принимают в предварительном расчете, исходя из следующих соображений.

Падение магнитного потенциала в нерабочем зазоре принимается порядка 10% от н. с., требуемой для проведения потока Φ_y , т. е. $F_n = 0,1F_y$. Для вычисления падения магнитного потенциала в стали находят по кривой намагничивания материала для заданной индукции B значение H_c . Учитывая, что индукция в начале якоря равна B , а в конце $B = 0$, можно считать

$$F_{c1} = \frac{H_c x}{2}.$$

Потери в наружном ярме принимают равными 60% потерь в якоре, так как индукция в ярме обычно меньше индукции в якоре. Следовательно,

$$F_c = \frac{H_c x}{2} + 0,6 \frac{H_c x}{2} = 0,8H_c x.$$

3) Намагничивающая сила для потока утечки (с учетом только воздушных путей)

$$F_{y0} = \frac{\Phi_y}{G_{yn}} = B \pi r_1^2 \frac{2(r_2 - r_1) \cdot 2}{\mu_0 \pi (r_1 + r_2) x} = \frac{4B r_1^2 (r_2 - r_1)}{\mu_0 (r_2 + r_1) x}.$$

Полная н. с. для потока утечки

$$F_y = \frac{I \omega x}{y} = 8,8 \frac{B r_1^2 (r_2 - r_1)}{x \mu_0 (r_2 + r_1)} + 0,8H_c x. \quad (10-14)$$

4) Толщина намотки

$$h = r_2 - r_1, \text{ причем } m = \frac{r_2}{r_1} = 2 \div 2,5.$$

Соотношения для расчета.

1. *Уравнение силы.* Уравнение $Q_2 = 10,2H\Phi_y$ (6-25) преобразуем к виду, более удобному для проведения предварительных расчетов. Если r_1 — радиус якоря, B — индукция в якоре в начальном его положении, то

$$B = \frac{\Phi_y}{\pi r_1^2}. \quad (10-15)$$

Далее

$$H = \frac{I\omega}{l_k}. \quad (10-16)$$

Следует дополнительно учесть уменьшение силы за счет потерь магнитной энергии, затрачиваемой на намагничивание входящей в магнитную систему части якоря. Это снижение силы учитывается определенным коэффициентом k . Значение k для сталей хорошего качества получается порядка 0,85. Таким образом, уравнение (6-25) получается в виде

$$Q_2 = 10,2 \cdot 0,85 \frac{I\omega}{l_k} B \pi r_1^2 = 27,2 \frac{I\omega}{l_k} B r_1^2. \quad (10-17)$$

2. *Уравнение нагрева.* Из (10-12), учитывая $h = r_2 - r_1$,

$$\theta = \rho \frac{(F/0,85)^2}{20\mu f_k (r_2 - r_1) l_k^2}. \quad (10-18)$$

3. *Уравнение электрической цепи.*

$$U = \frac{4\rho(r_1 + r_2)}{d^2} F \cdot 10^{-3}, \quad (10-13)$$

Схема предварительного расчета.

1. *Вычисление r_1 .* Из (10-17) и (10-18) имеем

$$\frac{I\omega}{l_k} = \frac{Q_2}{27,2 B r_1^2}; \quad \frac{I\omega}{l_k} = \sqrt{\frac{20\mu f_k (r_2 - r_1) \theta 0,85^2}{\rho}}.$$

Следовательно,

$$\frac{Q_2}{27,2 B r_1^2} = \sqrt{\frac{20\mu f_k (r_2 - r_1) \theta 0,85^2}{\rho}}.$$

Учитывая, что

$$r_2 - r_1 = r_1 \left(\frac{r_2}{r_1} - 1 \right) = r_1 (m - 1),$$

получаем

$$r_1^5 = \frac{Q_2^2 \rho}{27,2^2 B^2 20\mu f_k (m - 1) \theta 0,85^2}.$$

2. *Вычисление $\frac{I\omega}{l_k}$* производится из выражения (10-18).

3. *Вычисление x_n* производится из (10-14).

4. *Длина катушки $l_k = x_n + x_p$.*

Проверочный расчет. На основании данных предварительного расчета устанавливаются основные размеры магнитной системы. Задачей проверочного расчета является уточнение параметров и введение изменений в величинах этих параметров в том случае, если принятые значения не привели к удовлетворительным результатам. Расчет производится по приведенным ранее соотношениям.

10-7. ФОРМЫ МАГНИТНЫХ СИСТЕМ ПЕРЕМЕННОГО ТОКА

В практике получили распространение короткоходовые системы разных форм. На фигуре 10, 5 приведены наиболее ходкие из них: *a* — с Ш-образным ярком и ярмом; *b* — с Т-образным ярком и Е-образным ярмом; *c*, *d* — соленоидные. Формы *a* и *b* дают лучшее решение в отношении получения минимального веса конструкции при заданной условной работоспособности. Однако в реле-строении применяют предпочтительно соленоидные системы (форма *c*), более удобные в конструктивном отношении. При необходимости получения больших сил на якоре применяется форма *d*.

Соотношение сечений. Сечение ярма принимается меньшим, чем сечение ярка и стопа. Отношение размеров *a* и *b* принимают порядка $1,5 \div 2$.

Размещение экрана. Экран располагают в месте наиболее удобном для шлифовки торцов.

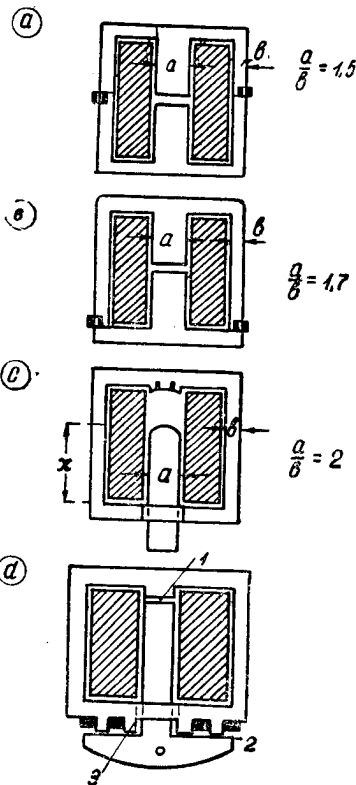
Конечный зазор. Для предохранения ярка от залипания необходимо иметь конечный зазор в системе; в местах же крепления экрана зазор должен быть минимальным. В системе формы *c* конечным зазором является промежуток между ярком и ярмом.

10-8. СИЛЫ НА ЯКОРЕ

В магнитных системах $\phi = \text{const}$. поток системы определяется приложенным напряжением и числом витков катушки.

Считая $U \cong E$, получаем

$$\Phi = \frac{\sqrt{2}U}{\omega\psi}$$



Фиг. 10, 5. Прямоходовые системы переменного тока.

Значение рабочего потока равно $\Phi_p = \frac{\Phi}{\sigma}$; при проведении расчета по формуле Максвелла краевой поток рабочего зазора относят к потокам рассеяния.

При вычислении сил на якоре необходимо учитывать конструктивные особенности магнитной системы.

Форма с. Задаваясь значением индукции B_p в рабочей части зазора, определяют силу

$$Q = 2,03 \cdot 10^8 B_p^2 q_1.$$

Формы а, б. Сила на якоре суммируется из сил в среднем и крайних зазорах системы. Формулу Максвелла приведем к среднему зазору

$$Q = 2,03 \cdot 10^8 [B_{p1}^2 q_1 + 2B_{p2}^2 q_2].$$

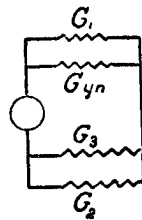
Учитывая, что

$$q_2 = \alpha q_1; B_{p2} = \frac{0,5B_{p1}}{\alpha},$$

получаем

$$Q = 2,03 \cdot 10^8 B_{p1}^2 q_1 \left(1 + \frac{0,5}{\alpha}\right). \quad (10-19)$$

Форма д. Рабочими являются зазоры 1 и 2, нерабочим — 3. В начальном положении якоря зазор 3 в несколько раз меньше зазора 2. Поэтому в этом положении силу считают только по зазору 1. В последующих положениях якоря необходимо учитывать и силу у зазора 2. При этом приходится пользоваться формулой, учитывающей силу в функции производной проводимости $\frac{dG}{db}$; здесь G — полная проводимость всей системы, которая определяется на основании схемы замещения магнитной системы (фиг. 10, 6).



Фиг. 10, 6.
Схема замещения.

В конечном положении якоря удобнее проводить расчет по формуле Максвелла, причем данные расчета получают относительно точными, поскольку значение краевого потока при малых зазорах между якорем и ярмом весьма невелико.

10-9. ДАННЫЕ ДЛЯ ПРОЕКТИРОВАНИЯ

При проектировании магнитной системы приходится задаваться некоторыми величинами. Рассмотрим основные величины B_p и σ .

Индукция B_p в рабочей части зазора. Ориентировочно значение B_p можно принять по величине конструктивного фактора. Соответствующие зависимости $B_p = f(\text{к. ф.})$, разработанные для систем постоянного тока, приведены на фигуре 5, 5. Для пользования этими данными при расчете систем переменного тока вводится понятие об эквивалентном конструктивном факторе. Учи-

тывая, что при одном и том же сечении площади рабочего зазора и одном и том же значении максимальной индукции сила в системе постоянного тока получается вдвое большей, чем в системе переменного тока, принимают эквивалентное значение к. ф. равным $\frac{\sqrt{2Q}}{\delta}$.

Коэффициент рассеяния σ . При весьма большом значении к. ф., характеризующемся малым значением хода δ , величина σ получается порядка 1,1. Ориентировочное значение σ при заданном к. ф. определяется следующим образом.

По кривым, изображенным на фигуре 5, 5, находят значение индукции B_{p1} для максимального значения к. ф. и B_{p2} — для заданного.

Тогда

$$\sigma = 1,1 \frac{B_{p1}}{B_{p2}}. \quad (10-20)$$

Пример. Определить значение коэффициента рассеяния для проектируемой системы.

Задано: $Q = 4,5 \text{ кГ}$; $\delta = 1 \text{ см}$.

Для к. ф. = 400 значение $B_{p1} = 1,38 \cdot 10^{-4}$. Эквивалентное значение к. ф. проектируемой системы

$$\frac{\sqrt{2 \cdot 4,5}}{1} = 3; B_{p2} = 0,5 \cdot 10^{-4}.$$

$$\sigma = 1,1 \frac{1,38}{0,5} \approx 3.$$

10-10. РАСЧЕТ КОРОТКОХОДОВОЙ СИСТЕМЫ ПЕРЕМЕННОГО ТОКА

Расчет системы состоит из двух этапов: предварительного и проверочного. Рассмотрим расчет системы, приведенной на фигуре 10, 5с.

Заданными являются следующие величины:

U_n — рабочее напряжение сети; начальная сила Q , начальный ход — δ — см, режим работы катушки.

а) Предварительный расчет. 1) Определяют значение B_p по эквивалентному к. ф. (см. кривые фиг. 5, 5).

2) Пользуясь формулой (10-20), устанавливают ориентировочное значение σ .

3) Из формулы $Q = 2,03 \cdot 10^9 B_p^2 q$ определяют сечение рабочего воздушного зазора.

4) Поток в системе $\Phi = B_p q \sigma$ вебер.

5) Число витков катушки определяется из условия получения требуемой силы Q при напряжении, равном $0,85U_n$:

$$\omega = \frac{\sqrt{2} \cdot 0,85U_n}{\omega B_p q \sigma}. \quad (10-21)$$

б) Предварительные размеры системы. Сечение якоря принимается квадратным. Значение нерабочего зазора 0,05 см, толщина наружной части ярма принимается равной половине ширины якоря.

Величина вхождения якоря в катушку, в начальном его положении, должна быть не меньше $\frac{1}{3}$ длины катушки. Таким образом,

$$l_k > 3x.$$

Толщина намотки принимается приблизительно равной ширине якоря.

б) Проверочный расчет. При проведении проверочного расчета уточняются все данные системы.

1) Проводится расчет катушки — по вычисленному в предварительном расчете определяют диаметр проволоки, электрическое сопротивление катушки.

2) Уточняются значения коэффициентов рассеяния.

3) Уточняются значения силы в начальном положении якоря.

4) Проводится расчет экрана, применяемого для устранения вибрации (см. главу 7). Выбор значения пульсации принимается из условия, чтобы минимальное значение электромагнитной силы в конечном положении якоря было больше максимальной величины противодействующей силы.

5) Вычисляется значение тока в катушке: I_1 — при отпущенном и I_2 — при притянutom якоря (см. главу 3). При вычислении индукции в стали учитывается коэффициент заполнения пакета жестей

$$q_c = k_c q.$$

6) Определяется температура превышения (см. главу 2).

7) Вычисляется степень использования системы по кажущейся мощности

$$k_u = \frac{I_1 - I_2}{I_1}.$$

Значение k_u в рационально спроектированных магнитных системах получается порядка 0,85—0,9.

Если полученные при расчете данные не удовлетворяют заданным условиям, меняют значение индукции B_p и проводят новый вариант расчета.

ГЛАВА 11

ВРЕМЯ СРАБАТЫВАНИЯ ЭЛЕКТРОМАГНИТНЫХ СИСТЕМ

11-1. КЛАССИФИКАЦИЯ ПО ВРЕМЕНИ СРАБАТЫВАНИЯ

Временем срабатывания реле считают продолжительность действия от момента подачи импульса на срабатывание до момента полного его срабатывания. Различают: 1) время срабатывания на включение, протекающее с момента включения катушки до момента замыкания нормально открытых контактов; 2) время сра-

батывания на отключение, протекающее с момента отключения катушки до момента замыкания нормально замкнутых контактов. Время t складывается из двух частей:

1) t_n — времени, протекающего с момента подачи импульса на катушку (на включение ее или отключение) до момента начала движения якоря;

2) t_d — времени движения якоря до полной его остановки. По времени срабатывания различают следующие основные группы реле:

1) быстродействующие $t = 0,003'' \div 0,050''$;

2) обыкновенные $t = 0,050'' \div 0,15''$;

3) замедленно действующие $t = 0,25''$ и более.

Приведенная классификация условна и принята здесь из следующих соображений: при проектировании электромагнита без учета времени t получают t порядка $0,05'' \div 0,15''$, т. е. обыкновенное реле; для создания быстродействующих или замедленно действующих реле приходится применять специальные конструкции или специальные схемы включения электромагнитов. Время срабатывания электромагнита зависит от кратности k приложенного к катушке напряжения

$$k = \frac{U}{U_{\text{ср. мин}}}$$

Заводские гарантийные данные по времени t относят к определенному значению k ($1,3 \div 3$).

Виды замедленно действующих реле. Такие реле могут быть исполнены с различными временными характеристиками в функции тока катушки $t = f(I)$ (фиг. 11, 13):

а) зависимая характеристика; б) независимая, в которой время срабатывания не зависит от силы тока в катушке; в) ограниченно зависимая, в которой начальная часть является зависимой и далее переходит в независимую.

Методы получения замедленного действия. Существует два основных метода.

Увеличивают составляющую t_n за счет увеличения постоянной времени контура или посредством специальных схем. Назовем эту группу реле с магнитным демпфированием.

Увеличивают составляющую t_d , для чего применяют различные виды торможения: масляные или воздушные демпферы, ветрянки и другие. Назовем эту группу реле с механическим демпфированием.

В обеих группах реле время t на срабатывание отличается от времени на отпуск; в реле с магнитным демпфированием, как правило, стремятся увеличить t на отпуск, а в реле с механическим демпфированием — на срабатывание.

11-2. ВЫЧИСЛЕНИЕ ВРЕМЕНИ t_n НА ВКЛЮЧЕНИЕ

Время t_n определяется, в основном, величиной постоянной времени T магнитной системы. Это положение справедливо для цепей постоянного тока. В цепях переменного тока закон нарастания тока при включении индуктивной катушки на напряжение $U = U_m \sin \omega t$ в момент, характеризуемый фазой ψ :

$$i = I_m \left[\sin(\omega t + \varphi + \psi) - \sin(\psi - \varphi) e^{-\frac{t}{T}} \right].$$

Таким образом, время нарастания тока определяется в данном случае в основном фазой ψ , и амплитудное значение I_m достигается максимум за время порядка четверти периода. Поэтому учет времени t_n производится только для систем постоянного тока с катушкой параллельного включения. В системах с катушкой последовательного включения закон нарастания тока определяется параметрами силовой цепи, а не реле. Ниже рассматриваются несколько схем включения катушек и показываются способы вычисления времени.

а) Обычное включение катушки. Закон нарастания тока

$$i = I \left(1 - e^{-\frac{t}{T}} \right). \quad (11-1)$$

При расчете реле устанавливают значения:

$I = \frac{U}{r}$; i — ток срабатывания реле; T — постоянная времени электромагнита; при наличии воздушных зазоров в магнитной системе индуктивность L определяется практически только магнитным сопротивлением воздушных зазоров. Тогда

$$T = \frac{L}{r} = \frac{\omega^2 G}{r} \text{ генри,}$$

где G — полная проводимость всех путей потока по воздуху.

Из (11-1) получаем

$$t_n = T \ln \frac{I}{I-i} \text{ сек.} \quad (11-2)$$

Уменьшение t_n достигается введением добавочного сопротивления r_1 последовательно с катушкой реле.

Тогда

$$T = \frac{L}{r + r_1}.$$

Пример. Требуется определить время t_n на включение.

Данные электромагнита: $I = 0,08$ а; $i = 0,06$ а; $r = 1000$ ом, $\omega = 10^4$; $G = 33 \cdot 10^{-8}$.

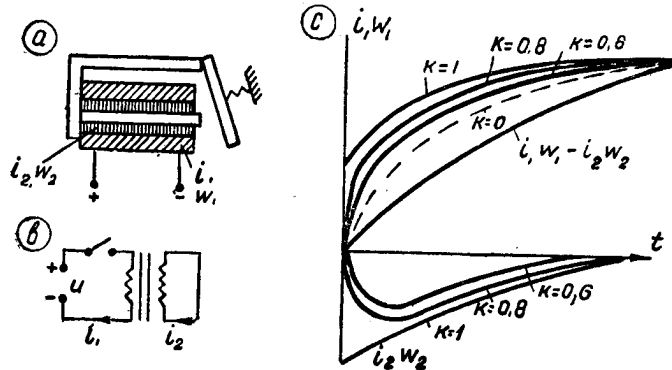
$$t_n = \frac{10^8 \cdot 33 \cdot 10^{-8}}{1000} \ln \frac{0,08}{0,08 - 0,06} = 45,5 \cdot 10^{-3} \text{ сек.}$$

б) Применение электромагнитного демпфера. Схема устройства и кривые приведены на фигуре 11, 1. Катушка 1 с числом витков ω_1 включается на напряжение сети, катушка 2 с числом витков ω_2 замкнута накоротко.

Мгновенные значения токов i_1 и i_2 определяются из следующих дифференциальных уравнений:

$$r_1 i_1 + L_1 \frac{di_1}{dt} + M \frac{di_2}{dt} = U; \quad (11-3)$$

$$r_2 i_2 + L_2 \frac{di_2}{dt} + M \frac{di_1}{dt} = 0. \quad (11-4)$$



Фиг. 11, 1. Система с магнитным демпфером:
а — эскиз; б — схема замещения; в — кривые $F = F(t)$.

Взаимоиндуктивность $M = k \sqrt{L_1 L_2}$, где k — коэффициент электромагнитной связи. Для крайних случаев:

отсутствие электромагнитной связи $k = 0$, $M = 0$;

полная связь — $k = 1$, $M = \sqrt{L_1 L_2}$.

При концентричной посадке обеих катушек — рабочей и демпферной $k = 0,95$, при раздельной посадке на общем сердечнике $k = 0,85 \div 0,9$. Полное значение н. с. системы

$$F = i_1 \omega_1 - i_2 \omega_2 = \omega_1 \left(i_1 - i_2 \frac{\omega_2}{\omega_1} \right).$$

Решение уравнений (11-3) и (11-4) приводится в учебниках теоретической электротехники и имеют следующий вид:

$$i_1 = I_{10} + I_{11} e^{p_1 t} + I_{12} e^{p_2 t}; \quad (11-5)$$

$$i_2 = I_{21} e^{p_1 t} + I_{22} e^{p_2 t}, \quad (11-6)$$

где p_1 и p_2 — корни характеристического уравнения.

Закон протекания токов i_1 и i_2 зависит от значения коэффициента электромагнитной связи k и постоянных времени T_1 и T_2 . При $k = 0$, $i_2 = 0$ — получается экспоненциальная кривая тока при отсутствии демпфера (см. пунктир).

Время $t_{н2}$ для получения требуемого значения тока срабатывания получается при наличии демпфера больше, чем $t_{н1}$ без демпфера.

В формулах (11-5) и (11-6) выражения для p_1 и p_2 имеют весьма громоздкий вид, вследствие чего теряется возможность анализа с целью установления влияния отдельных параметров на время t_n . Для практических расчетов можно пользоваться упрощенным методом, рассматривая кривую $i_1 - i_2 \frac{\omega_2}{\omega_1}$ как экспоненту с постоянной времени

$$T_n = T_1 + T_2 = \frac{\omega_1^2 G}{r_1} + k \frac{\omega_2^2 G}{r_2}. \quad (11-7)$$

Определив требуемое значение тока i_c — срабатывания, вычисляют

$$t_n = T_n \ln \frac{I}{I - i_c}. \quad (11-8)$$

Вихревые токи, индуцируемые в переходном процессе в стали магнитопровода, оказывают влияние на закон нарастания и убывания тока в катушке. Это влияние заметно сказывается при наличии массивных магнитопроводов, выполненных из стали с относительно малым удельным сопротивлением. Учет влияния стали производится в расчетах в виде постоянной времени T_c : сталь магнитопровода рассматривается как короткозамкнутая катушка с одним витком (см. ниже). Таким образом, в общем виде приведенная постоянная времени T_n магнитной системы состоит из трех составляющих:

$$T_n = T_1 + T_2 + T_c. \quad (11-9)$$

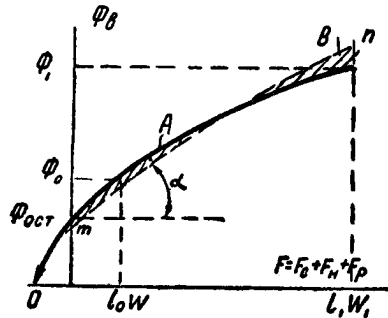
11-3. ВЫЧИСЛЕНИЕ ВРЕМЕНИ t_n НА ОТКЛЮЧЕНИЕ

Общие положения. При отключении катушки электромагнита ток исчезает практически мгновенно, так как длительность дуги на контактах отключающего устройства весьма мала. Поток в магнитопроводе также резко снижается до значения остаточного потока. В таких случаях время t_n на отключение практически не учитывают.

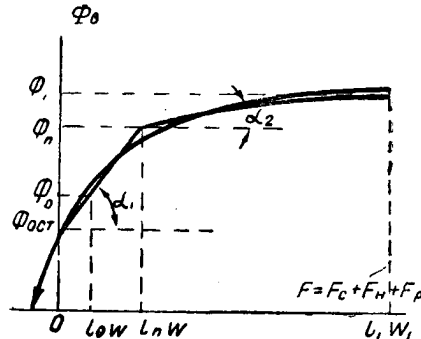
В том случае, если магнитопровод выполнен массивным, магнитная энергия системы при отключении катушки сохраняется в первый момент и затрачивается затем на джоулево тепло, определяемое вихревыми токами, индуцируемыми в стали. Закон убывания вихревых токов определяется постоянной времени T_c . Поток спадает не мгновенно, а меняется в соответствии с изменением $F = i_2 \omega_2$, определяемой вторичными токами. В рассматриваемой системе приходится учитывать величину t_n на отключение. Аналогичная картина получается при наличии в системе демпфирующей катушки.

Для увеличения t_n применяют также специальные схемы: шунтируют рабочую катушку электрическим сопротивлением, емкостью или выпрямительным вентилем. При применении таких схем магнитная энергия системы переходит после момента отключения катушки от сети в джоулево тепло, определяемое токами в катушке; намагничивающая сила определяется током i_1 : $F = i_1 \omega_1$.

Зависимость $\Phi = f(F)$ нелинейная. Поэтому задачу вычисления t_n решают графически или аналитически, с линеаризацией кривой размагничивания.



Фиг. 11, 2. Линеаризация кривой размагничивания.



Фиг. 11, 3. Дробная линеаризация кривой.

Данные для расчета. Время t_n на отключение определяется временем спадания потока от величины Φ_1 — потока в рабочем зазоре при притянутом якоре, до Φ_0 — потока отпуска, при котором электромагнитная сила Q_0 становится меньше противодействующей, механической силы Q_n .

При проведении расчета требуется строить кривую размагничивания системы (фиг. 11, 2)

$$\Phi = f(F_c + F_n + F_p).$$

Полному значению н. с., получающемуся при напряжении U_n , соответствует поток Φ_1 ; току отпуска i_0 — поток Φ_0 ; значение остаточного потока в системе — $\Phi_{ост}$.

Для аналитического решения задачи линеаризируют кривую размагничивания, для чего проводят прямую mn так, чтобы площадь A была равна площадке B .

Уравнение кривой размагничивания в этом случае:

$$\Phi = \Phi_{ост} + i\omega \operatorname{tg} \alpha. \quad (11-9)$$

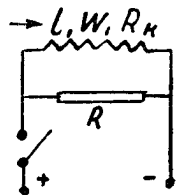
Для насыщенных систем применяют дробную линеаризацию (фиг. 11, 3).

Значение Φ_0 определяется из условия

$$Q_0 = Q_n = 4,06 \cdot 10^8 \frac{\Phi_0^2}{q} \quad (10-10); \quad \Phi_0 = \sqrt{\frac{qQ_n}{4,06 \cdot 10^8}}. \quad (11-11)$$

Величина Q_n может быть регулируема изменением натяжения пружины.

Время t_n при наличии шунтирующего сопротивления. При отключении катушки от сети ток катушки продолжает циркулировать в контуре: катушка — сопротивление R (фиг. 11, 4).



Фиг. 11, 4. Схема шунтирования рабочей катушки.

Исходное уравнение

$$i(R_k + R) + \omega \frac{d\Phi}{dt} = 0. \quad (11-12)$$

Интегрируя (11-12), получаем

$$t_n = \frac{\omega}{R_k + R} \int_{\Phi_0}^{\Phi_1} \frac{d\Phi}{i}. \quad (11-13)$$

Для удобства расчета уравнение (11-13) преобразуем так:

$$t_n = \frac{\omega^2}{R_k + R} \int_{\Phi_0}^{\Phi_1} \frac{d\Phi}{i\omega}. \quad (11-14)$$

Из (11-9)

$$d\Phi = \operatorname{tg} \alpha d(i\omega) \quad (11-15)$$

Из (11-14) и (11-15)

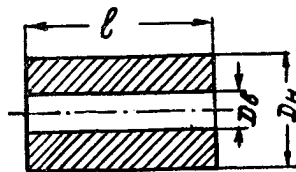
$$t_n = \frac{\omega^2}{R_k + R} \operatorname{tg} \alpha \int_{i_0 \omega}^{i_1 \omega} \frac{d(i\omega)}{i\omega},$$

или

$$t_n = \frac{\omega^2}{R_k + R} \operatorname{tg} \alpha \ln \frac{i_1}{i_0}. \quad (11-16)$$

Изменяя натяжение пружины, т. е., меняя значение Q_n , а вместе с тем Φ_0 и i_0 , получают различные значения t_n . Это время может быть изменено регулированием величины шунтирующего сопротивления.

Время t_n при наличии магнитного демпфера. Магнитный демпфер может быть выполнен в виде короткозамкнутой катушки. Практически демпфер выполняют в виде гильзы, насаженной на сердечник. Число витков демпфера $\omega_d = 1$ (фиг. 11, 5); электрическое сопротивление



Фиг. 11, 5. Магнитный демпфер.

$$r = \rho \pi \frac{D_n + D_d}{D_n - D_d} 10^{-8}. \quad (11-17)$$

В момент отключения рабочей катушки от сети в демпфере индуцируется ток, величина которого зависит от значения коэффициента электромагнитной связи k .

Если принять $\omega_\partial = 1$, то

$$i_\partial = ki\omega. \quad (11-18)$$

Время t_n может быть вычислено из (11-16) подстановкой данных демпфера:

$$t_n = \frac{1}{r} \operatorname{tg} \alpha \ln \frac{ki_1}{i_0}. \quad (11-19)$$

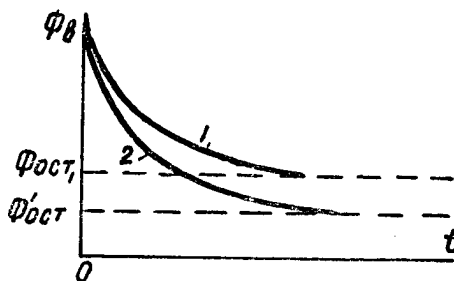
Строится кривая зависимости t_n от Φ_0 (фиг. 11,6). При $\Phi_0 = \Phi_{ост}$ время t_n становится равным бесконечности, что соответствует залипанию якоря. Учитывая различные отклонения от расчетных данных, получающиеся при изготовлении реле, принимают при расчетах коэффициент запаса порядка $1,3 \div 1,4$, т. е. максимальное расчетное время t_n принимают из условия $\Phi_0 = 1,3 \div 1,4 \Phi_{ост}$.

В формуле (11-19) не учитывается влияние вихревых токов, индуцируемых в момент отключения катушки в стали магнитопровода. Эти токи несколько увеличивают время t_n .

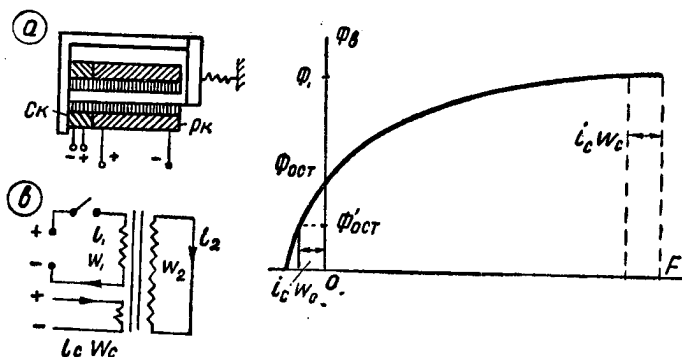
Как указывалось выше, для расчета сильно насыщенных систем рационально применить дробную линеаризацию. В данном случае

$$t_n = \frac{1}{r} \left(\operatorname{tg} \alpha_1 \ln \frac{i_n}{i_0} + \operatorname{tg} \alpha_2 \ln \frac{ki_1}{i_n} \right). \quad (11-20)$$

Система с магнитным демпфером и ссаживающей катушкой. Для увеличения t_n магнитную систему снабжают, помимо демп-



Фиг. 11, 6. Кривые $t = f(\Phi)$:
1 — без ссаживающей катушки; 2 — со ссаживающей катушкой.

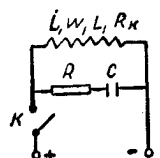


Фиг. 11, 7. Система с демпфером и ссаживающей катушкой.

фера, еще дополнительной катушкой, включаемой так, что н. с. этой катушки действует против рабочей (фиг. 11,7а).

Такие катушки называют ссаживающими. Действие ссаживающей катушки поясняется кривой, изображенной на фигуре 11, 7: в насыщенной системе при наличии отрицательной н. с., $i_c \omega_c$ — ссаживающей катушки, величина потока Φ_1 при токе i_1 уменьшается незначительно; значительно влияние $i_c \omega_c$ на величину остаточного потока, который снижается от $\Phi_{ост}$ до $\Phi'_{ост}$. Благодаря этому можно снижать величину Φ_0 до $1,3 \div 1,4 \Phi'_{ост}$, что позволяет увеличивать время t_n ослаблением пружины. Сравнительные характеристики: кривые $t = f(\Phi)$ приведены на фигуре 11, 6.

Шунтирование катушки емкостью. В практике встречаются схемы, где катушка электромагнита шунтируется емкостью (фиг. 11, 8). Последовательно с емкостью рекомендуется ставить электрическое сопротивление R , назначение которого:



Фиг. 11, 8.
Схема шунтирования катушки емкостью.

1) уменьшать величину зарядного тока при включении контакта K ; 2) устранять возможность возникновения колебательных процессов в момент отключения контакта K . Назначение шунтирующего контура — изменить время t_n на отпуск.

В зависимости от соотношения параметров цепи $L_k - (R + R_k) - C$ получается колебательный или аperiodический процесс затухания тока катушки. В системах с массивным магнитопроводом, где из-за вихревых токов, индуцируемых в стали, получается значительное время t_n , подбирают параметры контура так, чтобы получился колебательный процесс, что приводит к уменьшению времени t_n .

В других случаях, где требуется увеличение t_n на отпуск, подбирают параметры контура для получения аperiodического процесса.

Уравнение переходного процесса

$$i(R_k + R) + L \frac{di}{dt} + \frac{1}{c} \int i dt = 0. \quad (11-21)$$

Метод определения L см. в главе 3.

Дифференцируя (11-21) по t , имеем

$$L \frac{d^2 i}{dt^2} + (R_k + R) \frac{di}{dt} + \frac{i}{c} = 0. \quad (11-22)$$

Характеристическое уравнение

$$LP^2 + (R_k + R)P + \frac{1}{c} = 0;$$

$$P_{1,2} = -\frac{R_k + R}{2L} \pm \frac{2}{2L} \sqrt{(R_k + R)^2 - \frac{4L^2}{c}}. \quad (11-23)$$

Переходный процесс зависит от параметров цепи. Рассмотрим два случая:

$$1) (R_k + R)^2 - \frac{4L^2}{c} > 0.$$

Обозначим

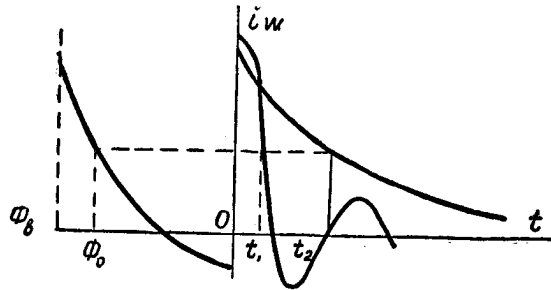
$$\frac{R_k + R}{2L} = \sigma; \quad \frac{1}{2L} \sqrt{(R_k + R)^2 - \frac{4L^2}{c}} = \rho,$$

тогда

$$P_1 = -\sigma + \rho; \quad P_2 = -\sigma - \rho. \\ i = e^{-\sigma t} \{A_1 e^{\rho t} + A_2 e^{-\rho t}\}. \quad (11-24)$$

Получается аperiодический процесс (фиг. 11, 9). Постоянные A_1 и A_2 определяют из начальных условий.

$$2) (R_k + R)^2 - \frac{4L^2}{c} < 0; \quad P_1 = -\sigma + j\rho; \quad P_2 = -\sigma - j\rho; \\ i = e^{-\sigma t} \{A_1 e^{j\rho t} + A_2 e^{-j\rho t}\}. \quad (11-24a)$$



Фиг. 11, 9. График для вычисления времени.

Получается колебательный процесс (фиг. 11, 9). При вычислении t_n необходимо построить кривую $i\omega = f(t)$. На этом же графике строится кривая размагничивания, соответствующая замкнутому состоянию магнитной системы (фиг. 11, 9). Вычисляется поток Φ_0 — отпуска якоря и графически определяется время t_n .

11-4. ВЫЧИСЛЕНИЕ ВРЕМЕНИ ДВИЖЕНИЯ

Основные уравнения.

1) Уравнение электрического равновесия:

$$U = ir + \frac{d\psi}{dt}. \quad (11-25)$$

В процессе движения якоря меняют свои значения как индуктивность системы L , так и ток катушки i . Поэтому

$$d\psi = Ldi + i dL. \quad (a)$$

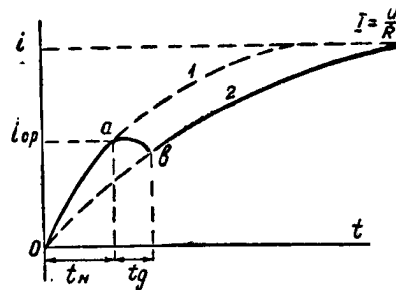
Из (11-25) и (a)

$$U = ir + L \frac{di}{dt} + i \frac{dL}{dx} \frac{dx}{dt} = ir + L \frac{di}{dt} + i \frac{dL}{dx} \frac{dx}{dt} \quad (11-26)$$

Из (11-26) видно, что в процессе движения якоря возникает добавочная противодействующая э. д. с., пропорциональная ско-

рости V — движения. Поэтому значение тока в переходном процессе не только не увеличивается, как это имело бы место при неподвижном якоре, но уменьшается.

На фигуре 11, 10 приведена осциллограмма изменения тока. До начала движения якоря значение тока изменяется по экспоненте 1, с постоянной времени T_1 , соответствующей разомкнутому состоянию магнитной системы.



Фиг. 11, 10. Осциллограмма тока при включении катушки.

По истечении времени t_n начинается движение якоря (точка a). При этом ток катушки, несколько повышаясь вначале, падает в дальнейшем. Точке b) кривой 2 соответствует полное время срабатывания. После остановки якоря значение тока изменяется по экспоненте 2, с постоянной времени T_2 , соответствующей замкнутому состоянию магнитной системы.

2) Уравнение механического равновесия.

$$Q_o = Q_n + Q_u + Q_d + Q_{тр}. \quad (11-27)$$

$$M_o = M_n + M_u + M_d + M_{тр}. \quad (11-28)$$

Здесь Q_n, M_n — значение силы и момента противодействующих устройств;

Q_u, M_u — значения инерционных сил и моментов;

Q_d, M_d — значения сил и моментов, созданных демпфирующими устройствами;

$Q_{тр}, M_{тр}$ — силы и моменты трения.

3) Уравнение силы $Q_s = 4,06 \cdot 10^8 \frac{\Phi^2}{q}$.

4) Значение потока $\Phi = \Phi(F)$. (11-29)

Динамическая тяговая характеристика. Тяговая характеристика, построенная без учета динамики движения якоря, называется статической. Такие характеристики рассматривались в предыдущих главах.

При учете инерционных сил движения, а также изменения силы тока в процессе движения получают динамическую характеристику. При больших скоростях движения якоря обе характеристики значительно отличаются по своим данным.

Методы расчета. Уравнения (11-26), (11-27) и (11-29) — нелинейные, вследствие чего аналитическое решение задачи не представляется возможным. Поэтому для вычисления времени t_0 пользуются методами графического или численного интегрирования, имеются также приближенные аналитические методы. При рассмотрении указанных методов можно отметить следующее.

1) Методы численного интегрирования громоздки и при решении задачи не выявляется влияние отдельных параметров на время t_{∂} .

2) Приближенные аналитические методы построены на ряде допущений. Так, в этих методах не учитывается изменение падения магнитного потенциала в стали по ходу якоря, которое фактически меняется в больших пределах — от $0,1 \div 0,2 F$ до $0,6 \div 0,8 F$. Далее, значение противодействующих сил принимается постоянным по ходу якоря или же меняющимся линейно. Между тем, в конструкциях реле они меняются ступенчато. Полученные в приближенных аналитических методах соотношения также громоздки и не показывают связи между временем t_{∂} и параметрами магнитной системы.

Пользуясь указанными методами, можно с некоторой точностью установить время t_{∂} для готовой модели электромагнита. В заводских условиях эта задача точнее решается экспериментальным путем.

При проектировании же системы необходимо определить время t_{∂} , если не точно, то хотя бы порядок этого времени. Однако при этом в распоряжении конструктора должны быть простые соотношения, показывающие, как надо изменять параметры системы для получения требуемого времени t_{∂} . Такому условию не удовлетворяют соответствующие методы графического интегрирования и упрощенные аналитические, поэтому практическое значение их невелико.

Упрощенный графо-аналитический метод. В практике пользуются упрощенным методом с использованием статической тяговой характеристики. При применении этого метода получается ошибка порядка 30—50%, так как не учитывается влияние всех факторов динамики процесса. Однако простые соотношения, получаемые при этом методе, позволяют установить влияние основных параметров магнитной системы на время t_{∂} , что важно знать при проектировании. Рассмотрим два случая.

1) *Магнитная система не снабжена механическим демпфером.* В этом случае пренебрегают значениями Q_{∂} , а также Q_{mp} . Тогда (11-27) будет

$$Q_{\partial} - Q_n = Q_u.$$

Учитывая, что $Q_u = m \frac{d^2 x}{dt^2}$, где m — масса подвижной системы, x — переменный ход якоря, получаем:

$$Q_{\partial} - Q_n = m \frac{d^2 x}{dt^2}. \quad (11-30)$$

Отсюда

$$t_{\partial} = \sqrt{\frac{2mx}{Q_{\partial} - Q_n}}. \quad (11-31)$$

По аналогии имеем

$$t_{\partial} = \sqrt{\frac{2J\alpha}{M_{\partial} - M_n}}. \quad (11-32)$$

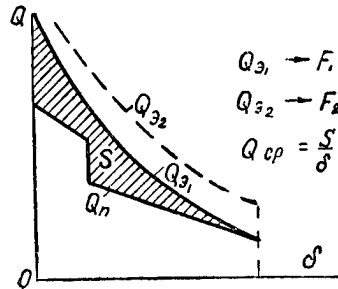
Значения $Q_0 - Q_n$, а также $M_0 - M_n$ принимаются как среднее по ходу якоря.

Построив механическую и электромагнитную характеристики (фиг. 11, 11), находят величину заштрихованной площадки S .

Тогда

$$Q_0 - Q_n = \frac{S}{\delta}.$$

Из (11-31) и графика, приведенного на фигуре 11, 11, видно, как надо менять параметры системы: массу, ход, значение н. с. F , характер тяговой характеристики, для получения требуемого времени t_0 .



Фиг. 11, 11. График.

$Q_{0_1} \rightarrow F_1$
 $Q_{0_2} \rightarrow F_2$
 $Q_{cp} = \frac{S}{\delta}$

2) Магнитная система снабжена механическим демпфером. Ускорения якоря получаются в таких системах незначительными. Поэтому принимают $Q_u = 0$. Пренебрегая также Q_{mp} , получаем

$$Q_0 - Q_n = Q_0.$$

Демпфирующая сила пропорциональна коэффициенту демпфирования k_0 и скорости движения:

$$Q_0 = k_0 \frac{dx}{dt}; \quad Q_0 - Q_n = k_0 \frac{dx}{dt}; \quad t_0 = k_0 \frac{x}{Q_0 - Q_n}. \quad (11-33)$$

По аналогии имеем

$$t_0 = k'_0 \frac{a}{M_0 - M_n}. \quad (11-34)$$

Формулы (11-33) и (11-34) дают относительно точное решение задачи, так как при малой скорости движения якоря статическая характеристика электромагнита незначительно отличается от динамической.

11-5. ВОПРОСЫ ПРОЕКТИРОВАНИЯ БЫСТРОДЕЙСТВУЮЩИХ СИСТЕМ

К быстродействующим относят реле, время срабатывания которых получается порядка 0,003 — 0,05 сек, причем это время относят к определенной кратности тока катушки $1,3 \div 3 I_n$. При проектировании таких реле необходимо принимать особые меры для снижения составляющих t_n и t_0 времени срабатывания. Полное время

$$t = t_n + t_0 = T \ln \frac{I}{I-i} + \sqrt{\frac{2mx}{Q_0 - Q_n}}. \quad (11-35)$$

Из (11-35) намечаются рекомендации по снижению времени t_0 .

1) Постоянная времени T должна быть минимальной. Для этого магнитопровод выполняют не сплошным, а шихтованным,

благодаря чему уменьшается влияние вихревых токов. Катушка выполняется с минимально возможными габаритами; последовательно с катушкой иногда ставят добавочное электрическое сопротивление.

В тех случаях, когда имеется контактная система большой разрывной мощности, магнитопровод делают сдвоенным, благодаря чему постоянная T значительно уменьшается.

2) Стремятся к уменьшению тока срабатывания i_c , для этого индукцию в стали принимают незначительной, порядка $(0,2 - 0,3) 10^{-4}$, выбирают материал магнитопровода с большим значением магнитной проницаемости при слабой напряженности поля, например пермаллой.

3) Стремятся к увеличению значения тока I , предельная величина которого определяется тепловым режимом катушки.

4) Якорь выполняют так, чтобы он имел возможно меньшую массу. Однако чрезмерно уменьшать сечение якоря не рекомендуется, так как при этом может получаться насыщение его.

5) Величина хода якоря x , или угла его поворота α , принимается минимальной. В некоторых исполнениях $x = 0,2 - 0,3$ мм.

6) Значение противодействующей силы Q_n должно быть возможно меньшим. Поэтому быстродействующие реле снабжаются маломощными контактами.

7) Значительное уменьшение времени срабатывания получается при применении постоянных магнитов (см. главу 13).

11-6. ВОПРОСЫ ПРОЕКТИРОВАНИЯ СИСТЕМ С МАГНИТНЫМ ДЕМПФИРОВАНИЕМ

Исполнения. Для получения замедленного срабатывания применяют в основном системы с демпфирующей катушкой. В таких системах получается выдержка времени как на срабатывание, так и на отпуск. Время t_n на срабатывание значительно меньше, чем на отпуск. Определяется это тем, что при разомкнутом якоре индуктивность катушки в несколько раз меньше, чем при притянтом. Время t_n на срабатывание непостоянно и колеблется при изменениях напряжения сети. Поэтому в схемах управления электроприводами, где реле с магнитным демпфером широко применяется, составляют схемы управления так, чтобы использовать выдержку времени только на отпуск якоря.

При проектировании схем телеуправления учитывают и время t_n на срабатывание, а иногда даже исполняют специальные схемы для увеличения его.

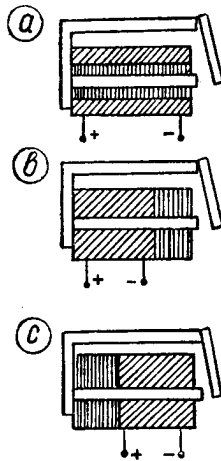
В практике нашли применение две серии реле.

1) Серия кодовых реле. t_n на срабатывание $0,015 - 0,32$ сек; на отпуск $- 0,013 - 1,2$ сек.

2) Серия первичных реле управления. Эта серия имеет более мощные контакты, чем серия кодовых реле. Время t_n на отпуск $- 0,3 - 3,5$ сек. Имеются исполнения с большими габари-

тами, в которых время t_n на отпуск может быть получено порядка 10—15 сек. Реле с магнитным демпфером может быть включено на цепь переменного тока через выпрямитель; рекомендуется применять двухполупериодное выпрямление.

Рабочая катушка. В реле обычного исполнения расчет катушки проводят так, чтобы при напряжении $U = 0,85 U_n$ получалось значение н. с., достаточное для притяжения якоря. В реле с магнитным демпфером расчет катушки проводят из условия получения необходимого значения н. с. при $U = 0,5 \div 0,6 U_n$; при этом получают повышенные значения потока при номинальном напряжении в замкнутом состоянии системы.



Фиг. 11, 12. Размещение катушки и демпфера.

Демпфер выполняется в виде гильзы, насаживаемой на сердечник. Имеются исполнения в виде набора шайб. Как правило, демпфер выполняется из меди. В некоторых случаях, когда требуется пониженное значение t_n , применяют алюминий. Такое исполнение не допускается в том случае, если реле работает в установке, где возможно появление щелочных паров и газов.

Размещение обмоток. Применяется два способа размещения обмоток.

1) Рабочая катушка и демпфер расположены на сердечнике отдельно (фиг. 11, 12), местоположение демпфера на сердечнике принимается различным: ближе к якорю — у рабочего зазора — исполнение *b*, или у основания сердечника — исполнение *c*. В случае исполнения *b* получается увеличенное значение t_n на срабатывание, в случае *c* — на отпуск. Раздельное размещение принято в кодовых реле.

2) Обе катушки расположены концентрично: непосредственно на сердечнике — демпфер, поверх него — рабочая катушка (фиг. 11, 12*a*). При таком размещении катушек, применяемом в конструкциях реле управления, получается большее значение коэффициента электромагнитной связи k , что приводит к увеличению времени t_n .

Магнитопровод должен выполняться из материала с малой коэрцитивной силой, чтобы уменьшить значение остаточного потока в магнитной системе. Как правило, применяется железо Армко марки Э. В кодовых реле ярмо выполняется из железа Армко, сердечник из кремнистой стали круглого профиля.

Конструкция магнитопровода должна быть такой, чтобы нерабочие зазоры магнитной системы были минимальными. Удачное решение дано инженером М. П. Книгиным, предложившим исполнение магнитопровода цельным (фиг. 8, 1*e*). Такое исполнение

принято заводами Советского Союза для всех реле с магнитным демпфером.

Регулирование времени t_n осуществляется, в основном, изменением натяжения пружины. Ступенчатое регулирование достигается изменением толщины немагнитной прокладки между торцом сердечника и якорем.

Точность по времени t_n в реле с магнитным демпфером относительно невелика и получается порядка 5—10%.

Колебания по времени действия определяются в основном изменениями напряжения сети, а также нагрева катушки. Уменьшение влияния этих факторов достигается выполнением магнитной системы насыщенной, когда колебания значения н. с. в пределах $\pm 20\%$ мало отражаются на величине потока в замкнутом магнитопроводе.

Время t_n на отпуск зависит еще и от времени заряда, т. е. времени, протекающего от момента включения катушки до момента, когда поток Φ_1 системы достигает максимального значения. Требуемое время заряда зависит от размеров магнитопровода и демпфера и бывает порядка 0,3—0,8 сек.

Ниже приводятся данные испытания реле серии РЭ500:

Время заряда, сек.	0,31	0,5	5
Время t_n на отпуск	1,87	2,18	2,2

При составлении схем автоматики учитывают влияние на время t_n времени заряда.

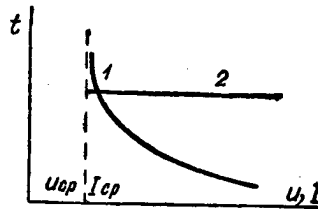
11-7. ВОПРОСЫ ПРОЕКТИРОВАНИЯ СИСТЕМ С МЕХАНИЧЕСКИМИ ДЕМПФЕРАМИ

а) Системы с механическими демпферами выполняются, как правило, с временем t_d на срабатывание. Возврат в исходное положение якоря и связанной с ним контактной системы должен осуществляться при отключении катушки мгновенно. Это необходимо для быстрой подготовки реле к следующему срабатыванию.

б) Для получения значительных выдержек времени t_d требуется большое значение хода якоря. Поэтому в таких системах применяются прямоходовые электромагниты.

в) Исполнения. Системы с механическими демпферами исполняются как реле тока, напряжения, времени.

Реле тока снабжаются токовой катушкой, они применяются для защиты двигателей, работающих в повторно-кратковременном режиме. Временная характеристика таких реле должна быть зависящей от силы тока, т. е. с увеличением силы тока время t_d должно уменьшаться (фиг. 11, 3, 1).

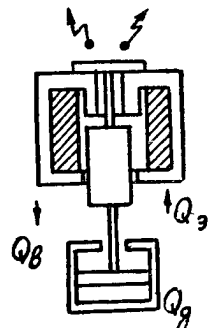


Фиг. 11, 13. Характеристика реле.

Реле времени называют такие устройства, в которых получается время срабатывания, не зависящее от приложенного напряжения. Временная характеристика таких реле получается независимой (фиг. 11, 13, 2).

г) **Демпфирующие механизмы** в практике применяются разного вида — масляные демпферы, воздушные крылатки, с постоянными магнитами, спусковые механизмы. Выбор механизма определяется требованиями точности. Реле с механическим демпфером могут быть установлены в помещениях, где температура среды колеблется в больших пределах: от -30° зимой до $+40^\circ$ летом. Наиболее часто применяются спусковые механизмы, работа которых не зависит от температуры. Работа других видов механизмов зависит от температуры: в масляных демпферах коэффициент демпфирования зависит от вязкости масла, меняющейся с изменением температуры; в воздушных — от плотности воздуха, в механизмах с постоянными магнитами — от удельного сопротивления материала диска или барабана.

д) **Кинематическая схема реле.** В общем случае реле состоит из следующих элементов: электромагнита, контактной системы, зубчатой передачи, храпового устройства (или втулки торпедо), демпфирующего механизма.



Фиг. 11, 14. Реле с масляным демпфером.

Связь между электромагнитом и демпфером может быть жесткой или гибкой.

Жесткая связь. Примерное устройство приведено на фигуре 11, 14. Временная характеристика при жесткой связи получается зависимой, так как сила на поршне демпфера определяется силой на якоре

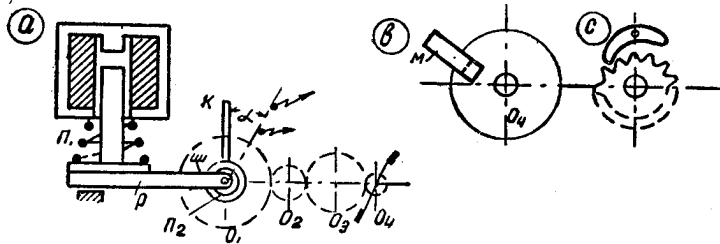
$$Q = kI^2.$$

Гибкая связь может быть осуществлена двояко.

1) Электромагнит при срабатывании заводит пружину P_2 (фиг. 11, 15), под действием которой происходит срабатывание

спускового механизма. В данном случае контактная система связана не с якорем, а с демпфирующим устройством. Временная характеристика получается независимой, так как движение спускового механизма определяется натяжением пружины P_2 , величина которой не зависит от тока в катушке.

2) В разомкнутом состоянии электромагнита под действием пружины P_1 заводится пружина P_2 . При срабатывании электромагнита освобождается рычаг P , который начинает поворачиваться под действием пружины P_2 . Такое исполнение создает условия для получения большой точности во времени; применяется в реле времени для защиты электрических систем, где требуется высокая точность — порядка 2—3%.



Фиг. 11, 15. Реле с механическим торможением:
а — с крылаткой; б — с диском и магнитом; в — со спусковым механизмом.

е) **Вычисление времени движения.** Для вычисления времени t_0 применяются различные формулы, в зависимости от вида демпфирующего устройства. В масляных демпферах поршень имеет поступательное движение, электромагнит выбирается таким, чтобы якорь был непосредственно связан с поршнем (фиг. 11, 14). Демпфирующие устройства, исполняемые в виде крылатки, диска или барабана с постоянным магнитом, центробежного регулятора с механическим трением, или в виде спускового механизма, имеют вращательное движение. Связь между электромагнитом и устройством нежесткая.

На фигуре 11, 15 приведено примерное исполнение реле с различными видами демпфирующих устройств: а — с крылаткой, б — с диском и постоянным магнитом, в — со спусковым механизмом.

Электромагнит имеет пружину P_1 , служащую для завода механизма. На оси O механизма насажены: рычаг P , контактный рычаг K , заводная пружина P_2 и ведущее колесо. Угол поворота контактного рычага до замыкания контактов α ; передаточное отношение зубчатой передачи i .

1) **Масляный демпфер** (фиг. 11, 14). Силы в системе: электромагнитная $Q_e = kU^2$ кг; демпфирующая $Q_d = k_d V$; $k_d = \frac{\text{кг сек}}{\text{см}}$; $V = \frac{\text{см}}{\text{сек}}$; вес якоря с поршнем Q_0 . Ход якоря до замыкания контактов — x см.

Из (11-33)

$$t_{\partial} = k_{\partial} \frac{x}{kU^2 - Q_{\partial}} \text{ сек.} \quad (11-36)$$

2) *Воздушная крылатка* (фиг. 11, 15 а). Среднее значение вращающего момента заводной пружины P_{∂} с учетом коэффициента полезного действия зубчатой передачи — $M_n \eta$; демпфирующий момент крылатки

$$M_{\partial} = k_{\partial} n^2 i^2,$$

где n — число оборотов ведущего колеса (в минуту).
Условие равновесия: $M_n \eta = M_{\partial} = k_{\partial} i^2 n^2$.

Отсюда

$$n = \sqrt{\frac{M_n \eta}{k_{\partial} i^2}}; \quad t_{\partial} = \frac{\alpha^{\circ}}{60n} = \frac{\alpha}{60} \sqrt{\frac{k_{\partial} i^2}{M_n \eta}} \text{ сек.} \quad (11-37)$$

3) *Диск с постоянным магнитом* (11, 15 б).

Демпфирующий момент $M_{\partial} = k_{\partial} n i$.Условие равновесия: $M_n \eta = M_{\partial} = k_{\partial} n i$.

$$t_{\partial} = \frac{\alpha}{60n} = \frac{\alpha k_{\partial} i}{60 M_n \eta} \text{ сек.} \quad (11-38)$$

4) *Спусковой механизм*. Период колебаний анкера

$$T = 2 \sqrt{\frac{J \varphi}{M_x}} \text{ сек.}$$

где J — момент инерции анкера, φ — угол поворота анкера. M_x — момент на оси ходового колеса.

За период T ходовое колесо поворачивается на один зуб. При числе зубьев ходового колеса z , передаточном отношении механизма i

$$t_{\partial} = \frac{\alpha z T i}{360} \text{ сек.}$$

Пример. Данные выполненного для реле часового механизма, состоящего из заводной пружины, зубчатой передачи, втулки торпедо и анкерного механизма:

$$M_x = 1 \text{ Г см}; \quad \eta = 0,912; \quad M_n = \frac{M_x i}{\eta} = \frac{1 \cdot 7,775}{0,912} = 8,53 \text{ Г см.}$$

$$T = 0,172 \text{ сек}; \quad z = 24; \quad \alpha = 45^{\circ}.$$

$$t_{\partial} = \frac{45 \cdot 24 \cdot 0,172 \cdot 7,775}{360} = 4 \text{ сек.}$$

Г Л А В А 12

ПОЛЯРИЗОВАННЫЕ РЕЛЕ

12-1. ВВЕДЕНИЕ

Принцип действия. Поляризованные реле представляют собой разновидность электромагнитных и отличаются тем, что в них действуют на якорь два независимых друг от друга потока:

- 1) постоянный поток, называемый поляризующим, не зависящий от рабочего состояния схемы устройства;
- 2) рабочий поток, определяемый н. с. рабочей катушки; величина и направление рабочего потока зависит от рабочего состояния схемы.

Особенности поляризованных реле. Поляризованные реле, по сравнению с обычными электромагнитными обладают рядом особенностей.

- 1) Направленностью действия: при изменении направления тока в рабочей катушке меняется знак результирующего момента на якоре, который перемещается вправо или влево от нейтрали воздушного зазора между полюсами.
- 2) Высокой чувствительностью, определяемой минимальной требуемой для срабатывания реле мощностью P_c . В поляризованных реле максимальная чувствительность определяется значениями порядка $P_c = 1 \cdot 10^{-6}$ *вт*;
- 3) Высокой кратностью термической устойчивости, представляющей собой отношение

$$k_T = \frac{\text{допустимый длительный ток включения}}{\text{минимальное значение тока срабатывания}}$$

В электромагнитных реле значение k_T бывает порядка $1 \div 1,5$, в поляризованных — $20 \div 50$.

- 4) Быстрым действием: в некоторых моделях реле получают минимальное время срабатывания порядка 2 *мсек*.

Применение. Широкое применение получили поляризованные реле в устройствах связи: телефон, телеграф, радио. Такие реле применяют также в устройствах электросиловой автоматики, релейной защиты, в следящих системах, системах телеуправления, железнодорожной сигнализации и блокировке (СЦБ).

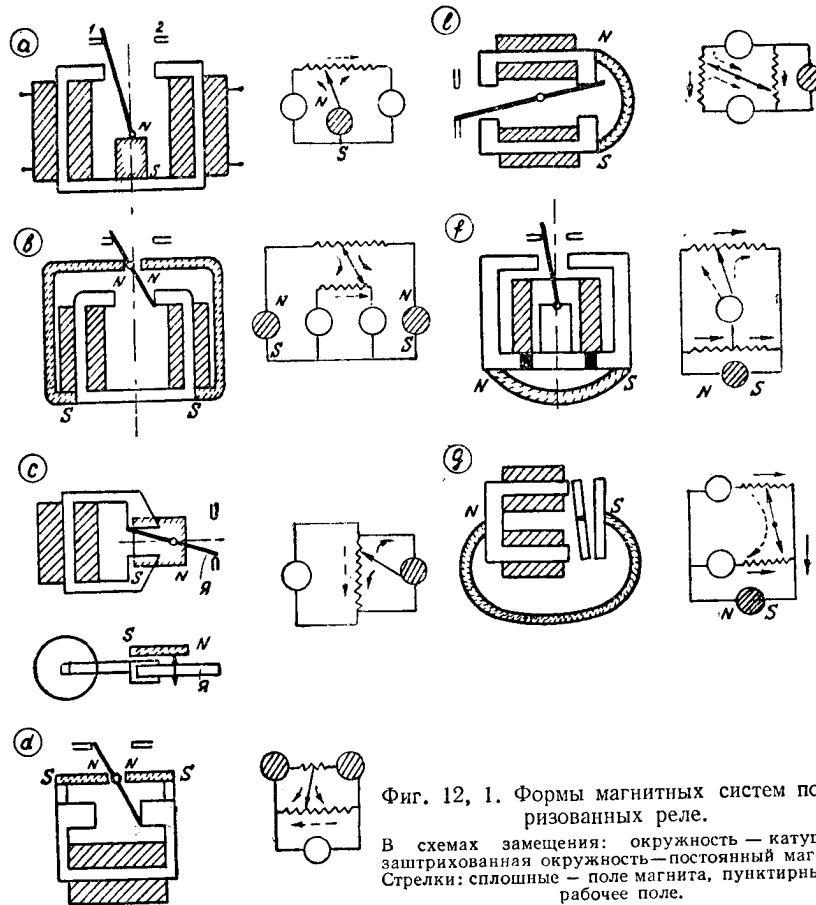
12-2. ФОРМЫ МАГНИТНЫХ СИСТЕМ

Для выполнения поляризованных реле в практике применяются различные формы магнитных систем. Основная классификация этих систем может быть проведена по двум признакам: по источнику н. с. поляризующего поля и по форме системы.

По источнику н. с. поляризованного поля различают:

- 1) системы с поляризующим полем от постоянного магнита (фиг. 12, 1);

2) системы с поляризующим полем от электромагнита (фиг. 12, 2); в этом случае катушка поляризующего поля приключается к постоянному источнику э. д. с. U_n , не зависящему от рабочего состояния схемы, где включена рабочая катушка



По форме различают системы дифференциальные и мостовые (по виду схемы замещения магнитной системы). Они отличаются друг от друга различным характером распределения в рабочем зазоре потоков — рабочего и поляризующего.

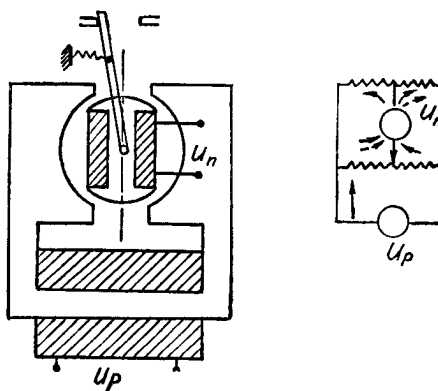
В дифференциальной системе (фиг. 12, 1 а, б, с, д) поляризующий поток в левой и правой частях рабочего зазора имеет разные направления, а рабочий поток одно — от одного полюса к другому.

В мостовой системе (фиг. 12, 1 е, г) картина обратная: поляризующее поле имеет один знак во всем рабочем зазоре, рабочее поле — разные знаки.

Наибольшее распространение получили в практике дифференциальные системы, более простые по конструкции. Однако мостовые системы можно осуществить с большей чувствительностью. Поэтому в последнее время наблюдается тенденция к разработке поляризованных реле с мостовой схемой магнитной системы.

Принципы построения и схемы замещения наиболее распространенных в практике форм магнитных систем поляризованных реле приведены на фигуре 12, 1: системы *b, c* — телеграфных реле завода «Красная звезда»; *f* — Вестингауза для устройств автоматики.

На фигуре 12, 2 приведена система Чебоксарского завода, с электромагнитной поляризацией от катушки, включенной на постоянный источник напряжения. Система предназначена для исполнения реле обратного тока, применяемого для защиты генераторов постоянного тока при параллельной их работе с другими генераторами или аккумуляторными батареями.



Фиг. 12, 2. Магнитная система с поляризующей катушкой.

12.3. НАСТРОЙКИ ПОЛЯРИЗОВАННЫХ РЕЛЕ

Поляризованные реле могут быть настроены на два режима работы: *о д н о с т о р о н н и й* — якорь реле срабатывает при одном направлении тока в рабочей катушке; *д в у х с т о р о н н и й* — срабатывание происходит при разных направлениях тока в рабочей катушке. Указанные режимы работы достигаются соответствующей настройкой реле, изменением положения контактных винтов, а также добавлением специальных пружин.

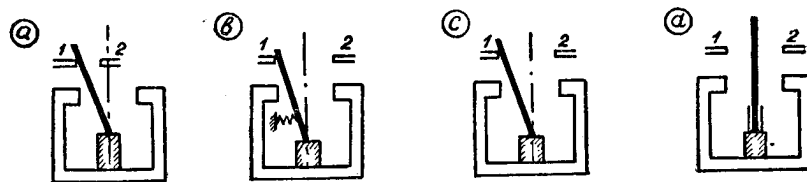
Однопозиционная настройка — магнитное преобладание (фиг. 12, 3а). Расположение контактных винтов таково, что якорь при срабатывании не переходит за нейтральную линию рабочего зазора. При обесточенной рабочей катушке поток Φ_1 магнита всегда больше потока Φ_2 — якорь всегда располагается на контакте 1. Срабатывание якоря — переход на контакт 2 — происходит при одном определенном направлении тока в рабочей катушке (см. направление поля Φ_p на фиг. 12, 5).

Однопозиционная настройка, комбинированное преобладание (фиг. 12, 3в). В обесточенном состоянии катушки на якорь действуют два момента:

1) момент, определяемый полем постоянного магнита и 2) момент пружины. При этом момент пружины всегда больше момента,

создаваемого полем магнита. Поэтому при обесточенном состоянии катушки якорь всегда возвращается в положение 1. Срабатывание якоря происходит при одном определенном направлении тока в рабочей катушке.

Двухпозиционная настройка (фиг. 12, 3c). В обесточенном состоянии катушки якорь располагается справа или слева от нейтрали (положения 1, 2) в зависимости от направления тока в предшествующем режиме. Срабатывание якоря происходит при разных направлениях тока в рабочей катушке.



Фиг. 12, 3. Настройка поляризованных реле.

a — однопозиционная, магнитное преобладание; b — однопозиционная, комбинированное преобладание; c — двухпозиционная, d — с нейтральным положением якоря.

Двухпозиционная настройка с нейтральным положением якоря. Якорь подвешивается на плоской пружине так, чтобы он занимал при обесточенной катушке нейтральное положение (фиг. 12, 3d). Срабатывание якоря — двухстороннее. В некоторых литературных источниках реле с такой настройкой называют трехпозиционными.

12-4. ЭКСПЛУАТАЦИОННЫЕ ПАРАМЕТРЫ РЕЛЕ

При проектировании поляризованных реле учитывают следующие эксплуатационные параметры, значения которых определяются условиями работы системы.

а) **Давление Q_n** — в нормально замкнутом контакте (н. з.) при обесточенной рабочей катушке. Это давление определяется разностью потоков поляризующего поля в правом и левом зазорах. Значение Q_n в высокочувствительных реле принимают порядка 1—2 Г; в реле, где допускается относительно большая мощность срабатывания, значение Q_n достигает значения 15—20 Г.

Ход контактов — x_k . В реле связи, где рабочее напряжение контактной цепи невелико (порядка 12—48 в) значение хода контактов принимают небольшим 0,05 ÷ 0,08 мм. В устройствах электросиловой автоматики, где рабочее напряжение контактной цепи достигает значения 220 вольт, принимают x_k порядка 0,4 ÷ 0,5 мм.

Минимальная мощность срабатывания P_c .
Кратность термической устойчивости $k_T = \frac{I_0}{I_c}$.

Коэффициент возврата, представляющий собой отношение тока отпуска якоря I_0 к току срабатывания I_c : $k_a = \frac{I_0}{I_c}$. В ряде

устройств не предъявляют особых требований к величине k_a ; необходимо лишь обеспечить условие надежного возврата якоря в исходное положение при обесточенной катушке. В данном случае значение $k_a = 0,1$ считается вполне достаточным.

В некоторых устройствах, учитывая возможные величины остающихся токов небаланса схемы, требуется, чтобы значение k_a было порядка 0,3—0,4, в специальных схемах требуется повышенное значение порядка 0,6. Значение k_a определяется настройкой реле. Повышение k_a приводит к ухудшению других параметров реле: времени срабатывания, чувствительности и конечного давления на контактах.

Время срабатывания. При токе I_c время t_c относительно велико 20—50 мсек и уменьшается с увеличением значения тока в рабочей катушке. Обычно учитывают время t_c при токе $I = 3 \div 4 I_c$.

Испытательное напряжение, зависящее от рабочего напряжения катушки и контактной системы.

12-5. РАСЧЕТ ПОЛЯРИЗОВАННЫХ РЕЛЕ РАЗНЫХ СИСТЕМ

Задачей расчета является установление для системы с определенными размерами соответствующих эксплуатационных параметров.

При расчете поляризованного реле необходимо принимать во внимание некоторые факторы, трудно учитываемые аналитически, например потоки рассеяния постоянного магнита, влияние толщины якоря на величину добавочной силы, вызываемой намагничиванием якоря полем электромагнита. Поэтому для расчета реле пользуются упрощенными соотношениями, которые служат только для ориентировки и установления взаимосвязи между отдельными параметрами реле.

При выводе соотношений, в целях их упрощения, не учитываются падения н. с. в стали магнитопровода, а также составляющие силы, определяемые краевыми потоками и намагничиванием якоря полем электромагнита. Ниже приводятся основные соотношения для расчета реле при разных видах настройки.

Общие положения, принятые при выводе соотношений:

1) Все значения сил, а также перемещения якоря приведены относительно линии mn , соединяющей середины полюсов (фиг. 12, 5). Пересчет этих сил и перемещений по отношению к контактам проводят затем с учетом расстояний от оси якоря.

2) Расчетный воздушный зазор a принят

$$a = A - \Delta, \quad (12-1)$$

где A — раствор между полюсами;

Δ — толщина якоря.

3) Рабочие проводимости приняты без учета проводимостей краевых потоков.

При определении индуктивности цепи и коэффициентов рассеяния учитываются дополнительно проводимости потоков краевых и утечки.

В разделе 12-8 приводятся соотношения и ход расчета дифференциальной и мостовой систем с разными видами настройки.

12-6. ВОПРОСЫ ПРОЕКТИРОВАНИЯ

Форма магнитной системы. При выборе формы магнитной системы учитывают, с одной стороны, эксплуатационные параметры, с другой — условия производства: конструкция должна быть простой в изготовлении и настройке. Удачное реле получено в модели дифференциальной системы, приведенной на фигуре 12, 1 с.

Настройка. Как правило, применяется настройка с магнитным преобладанием. Конструкция выполняется так, чтобы в ней можно было осуществить одно- и двухпозиционную настройку.

Комбинированное преобладание рационально применять в конструкциях, где требуется большой ход якоря — порядка 0,3 мм и выше. Сравнительная оценка магнитного и комбинированного преобладания, при однопозиционной настройке, может быть проведена при рассмотрении требуемой величины раствора между полюсами А, значение которого определяет собой чувствительность реле.

Настройка	Раствор	Коэффициент возврата
магнитная	$A = 2x + 2x_1 + 2\delta_0 + \Delta$	$\kappa_0 = \frac{x_1}{x + x_1}$
комбинированная . .	$A = x + 2\delta_0 + \Delta$	$\kappa_0 = \frac{j \cdot x - Q_{M0}(1 - a)}{Q_{M0}(1 + a)}$

Катушки. В обычных исполнениях реле ставится одна катушка (фиг. 12, 1 с) или две магнитносвязанные. При необходимости увеличения чувствительности реле увеличивают размеры катушки.

Рассмотрим связь между размерами катушки и потребляемой мощностью реле.

$$P = I^2 r (a).$$

В главе 2 (катушки) показано, что

$$r = \kappa \omega^2 (b),$$

где κ — конструктивный коэффициент, значение которого тем меньше, чем больше размеры катушки. Из (а) и (b):

$$P = \kappa (I \omega)^2.$$

Таким образом, при требующейся намагничивающей силе срабатывания якоря $F = I\omega$ увеличение размеров катушки приводит к уменьшению мощности срабатывания.

Однако увеличение размеров катушки приводит к увеличению постоянной времени системы, T , что влечет за собой увеличение времени срабатывания t . Поэтому при условии получения высокой чувствительности и малых времен срабатывания рациональное решение получается при выполнении сдвоенной системы с двумя катушками, не имеющими магнитной связи (фиг. 12, 4). В такой конструкции постоянная времени T системы определяется значением постоянной времени одной катушки (см. выражение 12-7).

Расчет катушки. В зависимости от схемы устройства, где включено реле, расчет катушки проводят по напряжению или по току.

Расчет по напряжению проводится в тех случаях, когда катушка реле приключается к определенному источнику тока или в диагонали моста; расчет по току — когда катушка включается в выходной цепи какого-либо усилителя — магнитного или электронного.

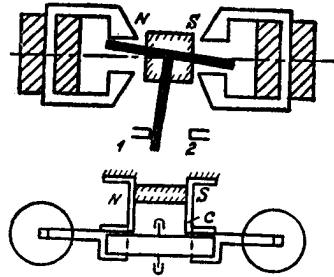
При включении катушки на выпрямленное напряжение переменного тока — двухполупериодное выпрямление — получается пульсирующее поле, приводящее к вибрации якоря. Устранение вибраций может быть осуществлено двумя способами:

1) сглаживанием пульсаций поля, что достигается насадкой на сердечнике магнитопровода коротко замкнутой гильзы из меди, толщиной порядка 0,3—0,5 мм;

2) сглаживанием пульсаций тока — включением емкости параллельно катушке.

Магнитопровод. Выбору материала для магнитопроводов поляризованных реле уделяется большое внимание. Для повышения точности работы реле исполняют магнитопровод из материалов, обладающих малой коэрцитивной силой H_c — из железа Армко, пермаллой. При этом на н. с. срабатывания реле меньше сказывается влияние остаточного поля, определяемого н. с. предшествующего режима.

Сечение магнитопровода принимается небольшим, но увеличивается у полюсов с целью увеличения площади сечения рабочего воздушного зазора. Большое значение имеет выбор толщины якоря. С повышением Δ увеличивается добавочная намагничивающая сила, определяемая полем рабочей катушки, т. е. увеличивается чувствительность реле. Однако увеличение Δ приводит к увеличению J , момента инерции, что, в свою очередь, приводит к по-



Фиг. 12, 4. Сдвоенная магнитная система:
с — латунная скоба.

вышению времени срабатывания. Толщина якоря принимается порядка 1,5—2 мм. Крепление якоря рационально проводить на плоской пружинной растяжке. При большом ходе якоря — порядка 0,5 мм и выше, приходится крепить якорь в оси. На этой же оси крепится спиральная моментная пружина, применяемая при выполнении комбинированной настройки.

12-7. ВРЕМЯ СРАБАТЫВАНИЯ РЕЛЕ

Полное время срабатывания реле t , отсчитываемое от момента включения рабочей катушки до момента замыкания нормально открытого контакта (н. о.), складывается из двух частей: t_n — времени нарастания тока до момента начала движения якоря, t_d — времени движения якоря.

а) Время t_n . Включение катушки к источнику постоянного тока:

$$i = I \left(1 - e^{-\frac{t}{T}} \right), \quad (a)$$

где i — ток срабатывания реле; I — значение установившегося тока; T — постоянная времени магнитной системы.

Из (a) имеем

$$t_n = T \ln \frac{I}{I-i}.$$

Таким образом, время t_n в основном определяется значением T . Постоянная времени T может быть вычислена на основании данных магнитной системы. При этом не учитывается магнитное сопротивление магнитопровода, которое во много раз меньше магнитного сопротивления рабочего воздушного зазора. Тогда

$$L = \omega^2 \sigma G_p,$$

где G_p — проводимость рабочего зазора, определенная без учета краевых проводимостей G_k , а

$$\sigma = \frac{G_p + G_k + G_{yn}}{G_p}; \quad G_p = \mu_0 \frac{S}{a}; \quad r = \kappa \omega^2.$$

Рассмотрим постоянную времени трех видов магнитных систем.

1) Система с одной катушкой (фиг. 12, 1 c):

$$T = \frac{L}{r} = \frac{\omega^2 \mu_0 S}{\kappa \omega^2} = \sigma \frac{\mu_0 S}{\kappa};$$

2) Система с двумя катушками, магнитно связанными (фиг. 12, 1 a). Считая κ — значение конструктивного коэффициента для одной катушки, получаем

$$T = 2\sigma \frac{\mu_0 S}{\kappa}.$$

3) Система с двумя катушками, магнитно не связанными (фиг. 12, 4 или 12, 1 e):

$$T = \sigma \frac{\mu_0 S}{ak}$$

Таким образом, лучшие результаты в отношении получения минимального значения T получаются при сдвоенной магнитной системе.

б) Время t_{∂} . При вычислении t_{∂} следует учитывать, что по ходу якоря меняется значение силы тока в рабочей катушке и значение противодействующих сил. Вычисление t_{∂} см. в главе 11.

12-8. ОСНОВНЫЕ РАСЧЕТНЫЕ СООТНОШЕНИЯ

Дифференциальная система однопозиционная. Настройка с магнитным преобладанием (фиг. 12, 5).

Обозначения.

Φ, Φ_1, Φ_2 — потоки в об магнита: полный и в левой, и в правой частях зазора;

Расчетный зазор: $a = A - \Delta$, см;

S — рабочая площадь полюса, см²;

x — ход контакта, см.

x_1 — смещение от нейтрали.

1) Распределение поляризующего поля.

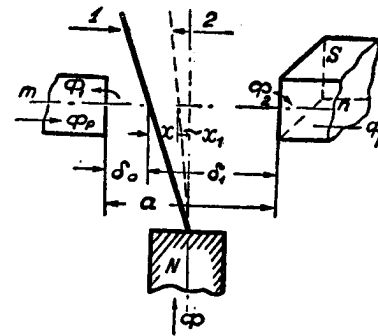
$$\Phi = \Phi_1 + \Phi_2 \quad (a); \quad \Phi_1 / \Phi_2 = G_1 / G_2 \quad (b);$$

$$G_1 = \mu_0 \frac{S}{\delta_0} \quad (c); \quad G_2 = \mu_0 \frac{S}{\delta_1} \quad (d).$$

Из (a, b, c, d)

$$\Phi_1 = \Phi \frac{\delta_1}{a}; \quad (12-2)$$

$$\Phi_2 = \Phi \frac{\delta_0}{a}. \quad (12-3)$$



Фиг. 12, 5. Магнитная система реле. Магнитная настройка.

2) Сила на якоре при обесточенной катушке (отнесенная к линии mn).

$$Q_{\text{мн}} = 4,06 \cdot 10^8 \frac{\Phi_1^2}{S} - 4,06 \cdot 10^8 \frac{\Phi_2^2}{S};$$

$$Q_{\text{мн}} = 4,06 \cdot 10^8 \Phi^2 \frac{\delta_1 - \delta_0}{aS} \quad \text{кГ}. \quad (12-4)$$

3) Силы на якоре при включенной катушке. При наличии рабочего потока Φ_p

$$Q = 4,06 \cdot 10^8 \frac{(\Phi_p + \Phi_2)^2}{S} - 4,06 \cdot 10^8 \frac{(\Phi_p - \Phi_1)^2}{S};$$

или

$$Q = 8,12 \cdot 10^8 \frac{\Phi_p \Phi}{S} - 4,06 \cdot 10^8 \Phi^2 \frac{\delta_1 - \delta_0}{aS} \quad \text{кГ}, \quad (12-5)$$

где

$8,12 \cdot 10^8 \frac{\Phi_p \Phi}{S} = Q_s$ — сила, определяемая рабочим потоком, она остается постоянной по ходу якоря;

$4,06 \cdot 10^8 \Phi^2 \frac{\delta_1 - \delta_0}{aS} = Q_m$ — сила, определяемая поляризующим потоком; эта сила меняется по ходу якоря. В конечном положении якоря

$$Q_{mk} = 4,06 \cdot 10^8 \Phi^2 \frac{2x_1}{aS}. \quad (12-6)$$

4) *Минимальное значение рабочего потока.* Значение Φ_{pc} срабатывания определяется из критического условия

$$Q = Q_s - Q_m = 0, \quad (12-7)$$

Из (12-7), учитывая (12-5),

$$\Phi_{pc} = \Phi \frac{\delta_1 - \delta_0}{2a} = \Phi \frac{x + x_1}{a}. \quad (12-8)$$

5) *Коэффициент возврата.* Рассматривая коэффициент возврата по току и учитывая, что рабочие потоки пропорциональны токам, получаем

$$\kappa_a = \frac{I_0}{I_c} = \frac{\Phi_{po}}{\Phi_{pc}}. \quad (12-9)$$

Значение Φ_{po} получаем из условия $Q_{эк} = Q_{mk}$.

$$8,12 \cdot 10^8 \frac{\Phi_{po} \Phi}{S} = 4,06 \cdot 10^8 \Phi^2 \frac{2x_1}{aS};$$

отсюда

$$\Phi_{po} = \Phi \frac{x_1}{a} \quad (12-10)$$

Учитывая (12-9), имеем

$$\kappa_a = \frac{x_1}{x + x_1}. \quad (12-11)$$

6) *Минимальная мощность срабатывания*

$$P = I^2 r \quad (a); \quad r = \kappa \omega^2, \quad (b)$$

где κ — конструктивный коэффициент, определяемый размерами катушки (см. главу 2).

$$\Phi = I_c \omega \mu_0 \frac{S}{a}. \quad (c)$$

Из (a), (b), (c) имеем

$$P_c = \kappa \Phi_{pc}^2 \frac{a^2}{\mu_0^2 S^2}. \quad (12-12)$$

В системе с двумя катушками, принимая κ для одной катушки, получаем

$$P_c = 0,5 \kappa \Phi_{pc}^2 \frac{a^2}{\mu_0^2 S^2}. \quad (12-13)$$

Учитывая (12-4), (12-8) и (12-13), можно найти значение P_c в функции начального давления на контакте Q_{mn} .

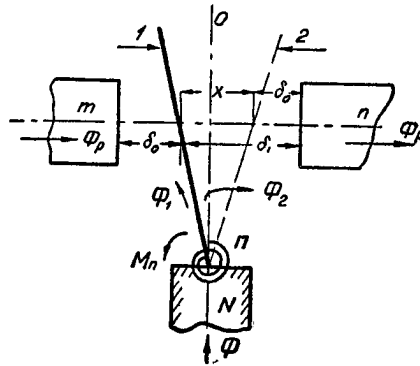
Дифференциальная система. Однопозиционная настройка. Комбинированное преобладание (фиг. 12, 6).

1. Пружина (n): Q_{nn} — начальная сила в положении якоря 1, отнесенная к линии mn ; j — жесткость пружины.

2. Сила на якоре при обесточенной катушке.

$$Q_n = Q_{mn} + Q_{nn}.$$

(Вычисление Q_{mn} и распределение потоков см. выше).



Фиг. 12, 6. Магнитная система реле. Комбинированная настройка.

3. Сила на якоре при включенной катушке. Положение 1:

$$Q_1 = 8,12 \cdot 10^8 \frac{\Phi_p \Phi}{S} - 4,06 \cdot 10^8 \Phi^2 \frac{x}{aS} - Q_{nn}. \quad (12-14)$$

Положение 2:

$$Q_2 = 8,12 \cdot 10^8 \frac{\Phi_p \Phi}{S} + 4,06 \cdot 10^8 \Phi^2 \frac{x}{aS} - Q_{nn} - jx. \quad (12-15)$$

Значение j — жесткости пружины, $\frac{\kappa \Gamma}{\text{см}}$, отнесенное к линии mn , принимается из условия получения требуемого значения κ_s . Обозначив $\alpha = \frac{Q_{nn}}{Q_{mn}}$, получаем из (12-14 и 12-15):

$$Q_1 = 8,12 \cdot 10^8 \frac{\Phi_p \Phi}{S} - 4,06 \cdot 10^8 \Phi^2 \frac{x}{aS} (1 + \alpha); \quad (12-16)$$

$$Q_2 = 8,12 \cdot 10^8 \frac{\Phi_p \Phi}{S} + 4,06 \cdot 10^8 \Phi^2 \frac{x}{aS} (1 - \alpha) - jx. \quad (12-17)$$

4. Коэффициент возврата $\kappa_s = \frac{I_o}{I_o} = \frac{\Phi_{po}}{\Phi_{pc}} \quad (12-9).$

Значение Φ_{pc} определим из (12-16), приравнявая электромагнитную силу противодействующей, а Φ_{po} — из (12-9).

Обозначив

$$4,06 \cdot 10^8 \Phi^2 \frac{x}{aS} = Q_{\text{мн}},$$

имеем

$$8,12 \cdot 10^8 \frac{\Phi_{\text{рс}} \Phi}{S} = Q_{\text{мн}} (1 + \alpha)$$

и

$$8,12 \cdot 10^8 \frac{\Phi_{\text{рс}} \Phi}{S} = jx - Q_{\text{мн}} (1 - \alpha),$$

тогда

$$\kappa_{\theta} = \frac{jx - Q_{\text{мн}} (1 - \alpha)}{Q_{\text{мн}} (1 + \alpha)}. \quad (12-18)$$

Дифференциальная система. Двухпозиционная настройка
(фиг. 12, 7)

1. Исходные соотношения (см. выше).

$$2. \Phi_1 = \Phi \frac{\delta_1 + x}{a}; \quad \Phi_2 = \Phi \frac{\delta_0}{a}.$$

3. Сила на якоре при обесточенной катушке

$$Q_{\text{м}} = 4,06 \cdot 10^8 \Phi^2 \frac{x}{aS}. \quad (12-19)$$

4. Сила на якоре при включенной катушке. Положение 1:

$$Q_1 = 8,12 \cdot 10^8 \frac{\Phi_{\text{рс}} \Phi}{S} - 4,06 \cdot 10^8 \Phi^2 \frac{x}{aS}. \quad (12-20)$$

После переброски якоря в положение 2:

$$Q_2 = 8,12 \cdot 10^8 \frac{\Phi_{\text{рс}} \Phi}{S} + 4,06 \Phi^2 \frac{x}{aS}. \quad (12-21)$$

5. Минимальное значение рабочего потока. Значение $\Phi_{\text{рс}}$ — потока срабатывания — определяется из условия $Q_1 = Q_2$:

$$\Phi_{\text{рс}} = \Phi \frac{x}{2a}. \quad (12-22)$$

6. Минимальная мощность срабатывания (см. однопозиционную настройку с магнитным преобладанием).

Мостовая система. Двухпозиционная настройка (фиг. 12, 8)

1. Распределение потоков поляризующего поля

$$\Phi_1 = \Phi \frac{G_1}{G_1 + G_2} \quad (12-23); \quad \Phi_2 = \Phi \frac{G_2}{G_1 + G_2}. \quad (12-24)$$

$$G_1 = \mu_0 \frac{S}{2\delta_0}; \quad G_2 = \mu_0 \frac{S}{2\delta_1}.$$

2. Сила на якоре при обесточенной катушке

$$Q_{\text{м}} = 2 \cdot 4,06 \cdot 10^8 \frac{\Phi_1^2 - \Phi_2^2}{S} = 8,12 \cdot 10^8 \Phi^2 \frac{\delta_1 - \delta_0}{S(\delta_1 + \delta_0)}. \quad (12-25)$$

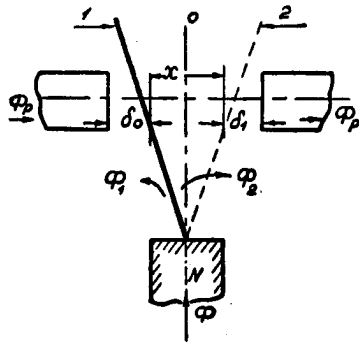
3. Сила на якорь при включенной катушке

$$Q_1 = \frac{2 \cdot 4,06 \cdot 10^8}{S} [(\Phi_p + \Phi_2)^2 - (\Phi_p - \Phi_1)^2];$$

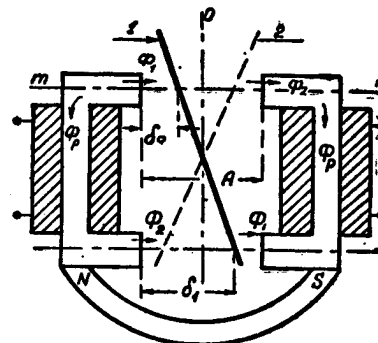
$$Q_1 = 4 \cdot 4,06 \cdot 10^8 \frac{\Phi_p \Phi}{S} - 8,12 \cdot 10^8 \Phi^2 \frac{\delta_1 - \delta_0}{S(\delta_1 + \delta_0)}, \quad (12-26)$$

или

$$Q_1 = Q_o - Q_m; \quad Q_2 = Q_o + Q_m.$$



Фиг. 12, 7. Магнитная система, двухпозиционная настройка.



Фиг. 12, 8. Мостовая система реле.

4. Минимальное значение рабочего потока. Исходя из условия $Q_o = Q_m$ из (12-26),

$$\Phi_{pc} = \Phi \frac{\delta_1 - \delta_0}{2(\delta_1 + \delta_0)}.$$

5. Минимальная мощность срабатывания (см. однопозиционную настройку с магнитным преобладанием).

ГЛАВА 13

ОСОБЫЕ ИСПОЛНЕНИЯ И СХЕМЫ

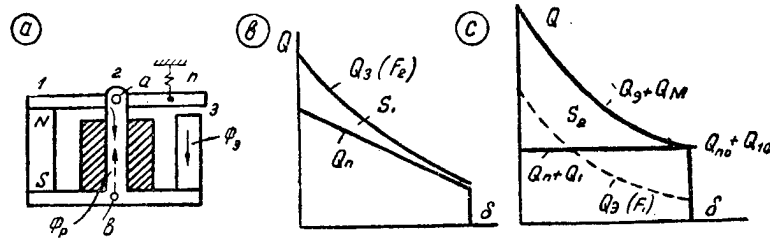
13-1. ЭЛЕКТРОМАГНИТНЫЕ СИСТЕМЫ С ПОСТОЯННЫМИ МАГНИТАМИ

Постоянные магниты применяются в электрических реле уже с давних пор, например при построении поляризованных и магнитоэлектрических реле.

Значительный интерес представляет применение постоянных магнитов в электромагнитных системах. Использование свободной магнитной энергии магнитов в таких системах открывает ряд возможностей построения конструкций реле, контакторов и механизмов уменьшенных размеров, с весьма малыми временами срабатывания, меньше 1 мсек (в обыкновенных электромагнитных системах минимальное время получается порядка 3÷10 мсек). Это определяется тем, что при наличии потока постоянного магнита

не требуется затраты времени на накопление магнитной энергии для срабатывания, или времени на израсходование накопленной магнитной энергии в массивных системах — для отпуска якоря. Конструкции электромагнита выполняются так, что для срабатывания или отпуска якоря нужно лишь перераспределить энергию постоянного магнита в различных частях магнитопровода, т. е. управлять этой энергией, для чего требуется весьма малое значение энергии, созданной током катушки. Для иллюстрации приведем несколько примеров исполнения электромагнитных систем с постоянными магнитами.

1) Быстродействующая электромагнитная система (фиг. 13, 1). Рассмотрим следующие основные положения работы системы.



Фиг. 13, 1. Система с постоянным магнитом:

a — эскиз; b — диаграмма системы без магнита; c — диаграмма системы с магнитом.

a) Катушка обесточена. Якорь притянут к магниту силой $Q_0 = Q_{m0} - Q_{n0} - Q_{m3}$, где сила $Q_{m0} = 4,06 \cdot 10^8 \frac{\Phi^2}{q_1}$; Q_{n0} — начальная сила пружины, $Q_{m3} = 4,06 \cdot 10^8 \frac{\Phi^2}{q_3}$, причем $Q_{m3} \ll Q_{m0}$.

Распределение потоков $\Phi_2 = \Phi \frac{G_2}{G_2 + G_3}$; $\Phi_3 = \Phi \frac{G_3}{G_2 + G_3}$;
 G_2 и G_3 — проводимости воздушных зазоров 2 и 3.

b) Включение катушки. После того, как н. с. катушки достигает величины $F_p > \frac{\Phi_3^2}{G_3}$, весь поток постоянного магнита и поток электромагнита проходят через зазор δ — якорь притягивается.

c) При размыкании цепи катушки — якорь возвращается в исходное положение.

На фигуре 13,1 приведены диаграммы $Q = f(\delta)$ для обычного электромагнита (b) и для системы с постоянным магнитом (c): требующаяся для срабатывания н. с. катушки $F_1 \ll F_2$, работа на ускорение якоря, характеризующаяся площадками S , в системе (b) значительно больше чем в системе (c): $S_1 \gg S_2$. Оба условия обеспечивают большую скорость срабатывания системы с постоянным магнитом.

2) Система реле (фиг. 13, 2). Система выполнена с магнитным шунтом (1), имеющим немагнитный зазор δ_1 . При обесточенных катушках поле магнита протекает в зазоре шунта δ_1 и якоря δ_2 , причем $\Phi_1 \gg \Phi_2$. При включении рабочей катушки (р. к.) поле

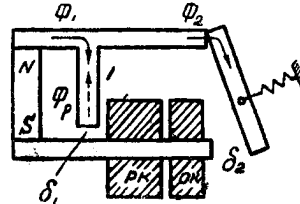
Φ_p вытесняет Φ_1 , и в якоре протекает весь поток магнита $\Phi = \Phi_1 + \Phi_2$. Якорь притягивается к сердечнику.

Возможны два исполнения реле: быстродействующее без залипания и с залипанием якоря.

В первом случае магнит выбирается таким образом, чтобы при замкнутом якоре сила, определяемая полем Φ_2 , $Q_{м2} = 4,06 \cdot 10^8 \Phi_2^2 / q$, была меньше силы Q_n — растянутой пружины.

Во втором случае магнит выбирается так, чтобы при обесточенной катушке и притянутом якоре $Q_{м2} > Q_n$.

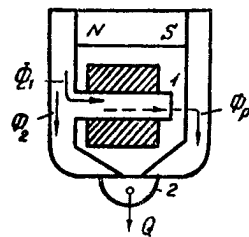
Описанная система, в случае исполнения её в качестве реле с залипающим якорем, снабжается добавочной отбойной катушкой (о. к.), поле которой действует против поля магнита. Исполненные согласно эскизу (фиг. 13, 2) конструкции отличаются весьма малыми габаритами, малой потребляемой мощностью и весьма малыми временами срабатывания.



Фиг. 13, 2. Система реле.

3) Реле с залипающим якорем завода «Трансвязь» выполнено в виде кодового реле (фиг. 1, 28), но сердечник катушки изготовляется из хромистой стали. При кратковременном включении катушки сердечник намагничивается и остаточное поле его достаточно для удерживания якоря при отключенной катушке. Рассматриваемая система

менее рациональна, чем система, изображенная на фигуре 13, 2, так как в данном случае требуется затрата энергии на создание поля в зазоре якоря, причем н. с. катушки должна быть высокой, учитывая относительно большое значение магнитного сопротивления хромистой стали.



Фиг. 13, 3. Система удерживающего магнита.

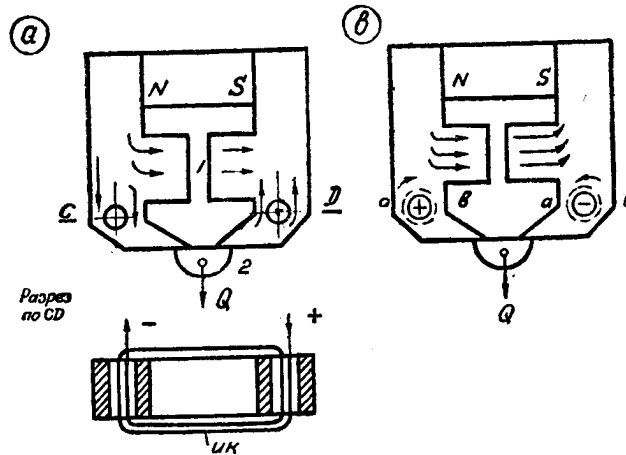
Исполнение описанного реле оправдывается тем, что в данном случае завод не должен разрабатывать специальной модели и использует существующую модель кодового реле.

4) **Удерживающие магниты.** В ряде устройств автоматики необходимы удерживающие механизмы с весьма быстрым отключением в требуемый момент времени. Применение удерживающих электромагнитов для таких целей нерационально, лучшее решение задачи получается при исполнении конструкции с постоянными магнитами: при этом достигается минимальная затрата магнитной энергии и значительно большие скорости срабатывания (система с подъемной силой 25 кг может быть выполнена с временем срабатывания меньше 1 мсек).

На фигуре 13, 3 приведен эскиз удерживающего магнита с электромагнитным отключением. При замыкании цепи катушки поток постоянного магнита отводится целиком по шунту (I) и

якорь (2) отпадает. В данной конструкции необходимо учитывать полярность включения катушки.

Другое исполнение удерживающего магнита приведено на фигуре 13, 4. Катушка наматывается в самом магнитопроводе (см. разрез *CD*). При обесточенной катушке поток магнита удерживает

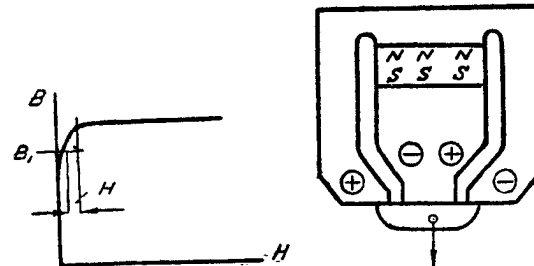


Фиг. 13, 4. Система удерживающего магнита.

якорь, оттягиваемый внешней силой Q (13, 4 *a*). При включении катушки поток последней вытесняет поле магнита из зоны якоря: в частях *a* (фиг. 13, 4 *b*) — вследствие наличия противодействующей н. с. катушки,

в частях *b* — вследствие насыщения участка. Весь поток магнита замыкается в этом случае через шунт.

Для выполнения описанной системы применяются магнитные материалы, обладающие крутой характеристикой (фиг. 13, 5), например никелевые стали с индукцией насыщения $B_s = 0,8 \cdot 10^{-4} \frac{\text{об}}{\text{см}^2}$ или $B_s = 1,5 \cdot 10^{-4}$.



Фиг. 13, 5. Кривая намагничивания.

Фиг. 13, 6. Сдвоенная система.

Индукция B_1 , определяемая полем постоянного магнита, должна быть в сечении q_1 близка к индукции насыщения B_s . Для создания условия насыщения участков *b* достаточно небольшого значения напряженности поля H катушки.

Вариант исполнения сдвоенной модели приведен на фигуре 13, 6.

Пример. Задано — лампа 6С5; реле: $F_c = 25$ ав; коэффициент возврата $\kappa_a = \frac{I_o}{I_c} = 0,4$; коэффициент катушки $\kappa = 8,97 \cdot 10^{-4}$. Напряжение сети 220 в. Найти электрическое сопротивление реле, число витков, напряжения смещения лампы — срабатывания и отпуска.
Решение: Из (13-2):

$$P = U_c I_c = \kappa F_c^2 = 8,97 \cdot 10^4 \cdot 625 = 0,56 \text{ вт.}$$

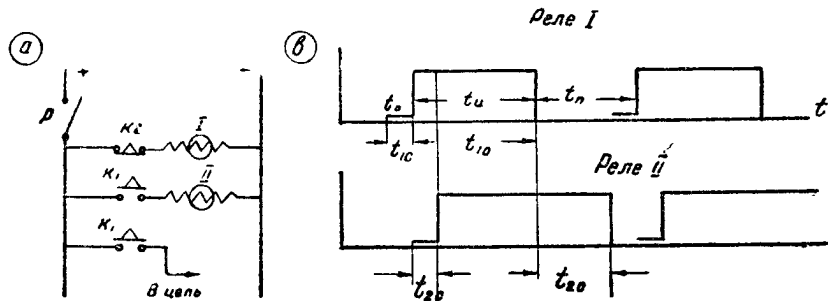
Строим кривую $I_c = f(U_c)$, задаваясь различными значениями U_c , откладывая значение U_c влево по оси абсцисс (от точки с). При проведении прямой cm к точке касания характеристик реле и лампы получаем:

ток срабатывания $I_c = 7,5$ ма; $U_p = 75$ в,
 $r = \frac{75 \cdot 10^3}{7,5} = 10000$ ом; число витков $w = \sqrt{\frac{r}{\kappa}} = 3400$;

ток отпуска $I_o = I_c \kappa_a = 7,5 \cdot 0,4 \cdot 10^{-3} = 3$ ма.
 Напряжения смещения: срабатывания $U_d = -2$ в; отпуска — $U_d = -7$ в.

б) Схема пульс-пара. В устройствах автоматики часто применяются схемы, называемые пульс-пара, служащие для создания импульсов с определенной длительностью и паузой импульса (генераторы импульсов). Задача решается посредством двух электромагнитных реле с магнитным демпфером (см. главу II). Время t_c на срабатывание в таких реле относительно невелико, порядка 0,15 секунды. Время t_o на отпуск может быть получено в пределах до нескольких секунд.

На фигуре 13, 8 приведена примерная схема пульс-пара. Контакты реле: I → 2 н. о. замыкающихся мгновенно, размыкающихся — при обрыве цепи катушки — с выдержкой времени; II → I н. з.



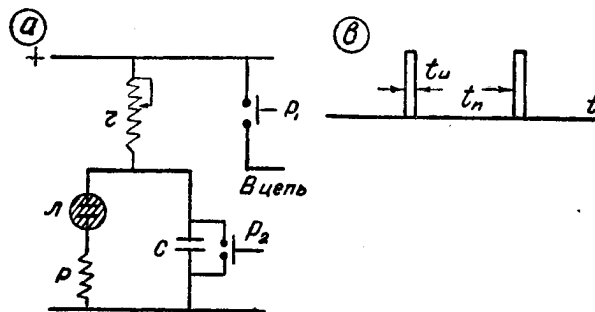
Фиг. 13, 8. Схема и временная диаграмма пульс-пары.

контакт, размыкающийся мгновенно, замыкающийся с выдержкой времени (на схеме стрелка на контакте показывает направление выдержки).

Работа схемы поясняется временной диаграммой. В момент времени t_0 замыкается рубильником P цепь катушки реле I, через время t_{1c} , включаются контакты: κ_1 — цепи катушки 2 и κ_1 — цепи куда подается импульс.

При срабатывании реле $II(t_{2c})$ размыкается цепь катушки (I), но якорь реле отпадает через время t_{10} . Это же время определяет собой длительность замыкания контактов κ_1 — **длительность** импульса t_u . Пауза t_n определяется временем t_{20} замедления на отпуск реле (II) плюс время t_{c1} .

В ряде случаев требуется получение кратковременного импульса с относительно длительной паузой. Пример выполнения соответствующей схемы приведен на фигуре 13, 9.



Фиг. 13, 9. Схема и временная диаграмма импульсного генератора.

В схеме применена лампа с тлеющим разрядом, контур rC и быстродействующее высокочувствительное реле, например в исполнении с постоянным магнитом (фиг. 13, 2). Конденсатор C контура при достижении определенного напряжения заряда U_c , соответствующего пробое лампы, разряжается через обмотку реле. Последнее своим контактом P_1 посылает импульс в цепь, контактом P_2 накоротко замыкает емкость с целью устранения остаточного заряда.

Длительность паузы: $t_n = rC$.

ЧАСТЬ ТРЕТЬЯ

ИНДУКЦИОННЫЕ СИСТЕМЫ И РЕЛЕ

ГЛАВА 14

МАГНИТНЫЕ СИСТЕМЫ

14-1. ЗАДАЧИ РАСЧЕТА

Индукционными называются такие электромеханические системы, в которых вращение подвижной части, ротора, получается от взаимодействия магнитных полей статора с токами, индуцируемыми этими полями в роторе. Индукционные системы применяются для построения реле различных исполнений: тока, напряжения, мощности, сопротивления, частоты, коэффициента мощности и других.

Наибольшее распространение получили индукционные реле в устройствах релейной защиты, некоторые исполнения применяются в устройствах автоматики.

Разработка индукционных реле проводится в два этапа.

1) Проектирование и расчет индукционной системы — разрабатывают систему определенной формы, определенных габаритов с учетом возможности применения этой формы для различных видов реле. Устанавливают определенные значения расчетных и конструктивных параметров разработанной системы.

2) Проектирование и расчет реле — по выбранной форме индукционной системы разрабатывают схему включения катушек и вычисляют параметры реле.

При выборе индукционной системы для данного вида реле учитывают ряд факторов: требуемую электромагнитную характеристику, т. е. зависимость рабочего момента на роторе и измеряемой электрической величины, сложность конструкции, временные характеристики.

Момент на роторе индукционной системы получается при наличии потоков, действующих на ротор, и сдвинутых в пространстве и во времени.

При расчете системы необходимо знать соотношения между электрическими параметрами сети, а именно, подведенным к системе током, напряжением, и магнитными параметрами — потоками, создаваемыми в магнитной системе. Эти соотношения определяются из расчета магнитной системы.

Каждый полюс магнитной системы снабжается или катушкой тока, включенной последовательно в цепь тока I , или катушкой напряжения, включенной на напряжение сети U . В первом слу-

чае элемент магнитной системы рассматривается как трансформатор тока, во втором — как трансформатор напряжения. Вторичной обмоткой в обоих случаях является ротор. Все соотношения характеризуются ниже векторными диаграммами. При построении этих диаграмм принимаются следующие допущения.

1) Явления рассматриваются с учетом работы магнитной системы в ненасыщенном состоянии; значение индукции в магнитопроводе системы не превышает $(0,4 \div 0,5) 10^{-4}$ вб/см², и в первом приближении не учитывается падение магнитного потенциала в стали.

2) Не учитывается индуктивное сопротивление ротора, которое во много раз меньше электрического: принимается, что вектор тока, индуцированного в роторе, совпадает по фазе с вектором э. д. с.

3) Не учитывается индуктивность рассеяния катушки (в воздухе).

Наиболее трудной при расчете системы является задача определения магнитных проводимостей путей потока по воздуху, так как картина поля полюсов в воздушных путях системы весьма сложна. Поэтому расчетные данные указанных проводимостей корректируют данными экспериментальной проверки, которую проводят для определенной модели индукционной системы.

14-2. МАГНИТНЫЕ ПРОВОДИМОСТИ ВОЗДУШНЫХ ПУТЕЙ МАГНИТОПРОВОДА

Необходимо вычислять значение проводимости в воздухе для магнитного поля отдельных полюсов. Вычисление проводят методом разбивки поля на простые фигуры или на основании эмпирических кривых Шмиделя.

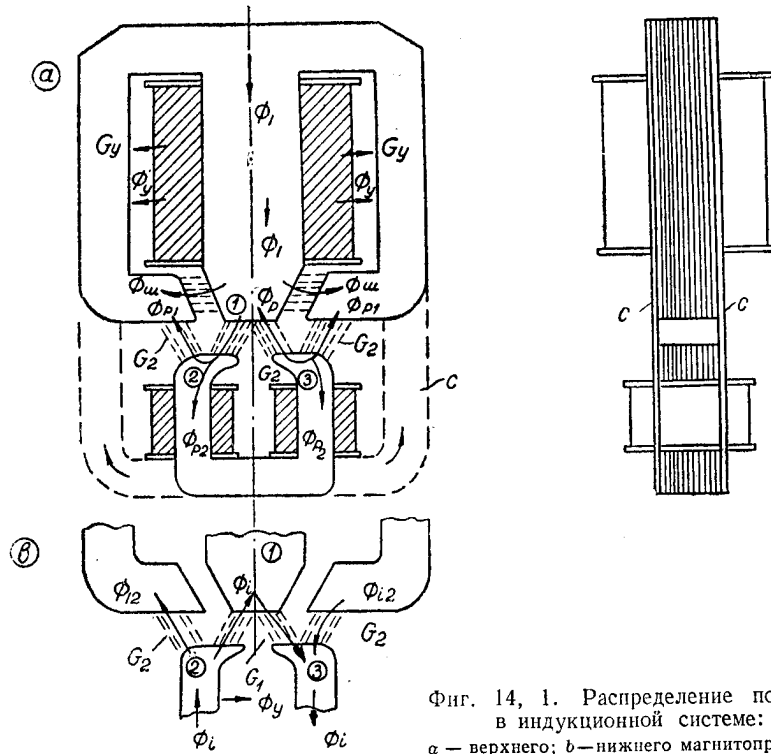
Метод Шмиделя удобен для определения проводимости зазора у полюсов. Поле в зазоре полюса неоднородно, и в расчетах принимают усредненное значение индукции B , которое получается как отношение потока к приведенной площади полюса S_n . Если площадь полюса $S = ab$, то по формулам Шмиделя получают $S_n = a_1 b_1$. По кривым находят переход от a и b к a_1 и b_1 .

Тангенциальная система. Отдельно рассматривается картина распределения поля, определяемого намагничивающей силой катушки верхнего магнитопровода, отдельно нижнего.

Поток Φ_1 (фиг. 14, 1 а) состоит из рабочего потока Φ_p , полюса 1, пронизывающего диск, и шунтирующих — Φ_w и Φ_y .

Поток Φ_p возвращается в магнитопровод по двум путям: через полюсы 2 и 3 по воздуху и через стальную скобу C , связывающую верхний и нижний магнитопровод.

Поток Φ_i , определяемый н. с. катушек нижнего магнитопровода, выходит из полюса 2 (фиг. 14, 1 в), разветвляется на потоки Φ_{i1} , Φ_{i2} и Φ_v .



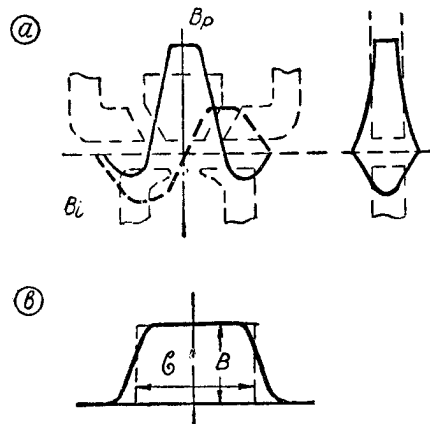
Фиг. 14, 1. Распределение потоков в индукционной системе:
а — верхнего; б — нижнего магнитопровода.

Некоторая неопределенность при расчете получается вследствие трудностей, связанных с учетом магнитного сопротивления пути потока Φ_{p2} по скобе (С). Это сопротивление зависит от качества сборки магнитной системы.

На фигуре 14, 2 показана примерная картина распределения индукции в зазорах у полюсов системы. Ниже приведены данные измерений, проведенных в индукционной системе реле мощности.

Потоки верхнего магнитопровода:
н. с. катушки = 275 ас;
напряжение $U = 100$ в;

$$\begin{aligned} \Phi_p &= 5,8 \cdot 10^{-5} \text{ вб}; \\ \Phi_{p2} &= 2,1 \cdot 10^{-5} \text{ вб}; \\ \Phi_1 &= 1,98 \cdot 10^{-5} \text{ вб}. \end{aligned}$$

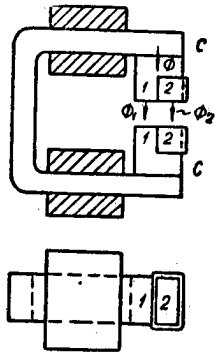


Фиг 14, 2. Распределение индукции в индукционной системе:
а — в рабочих зазорах тангенциальной;
б — у полюсной дуги системы с барабаном.

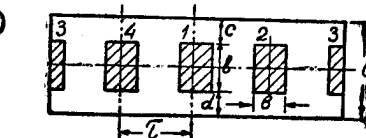
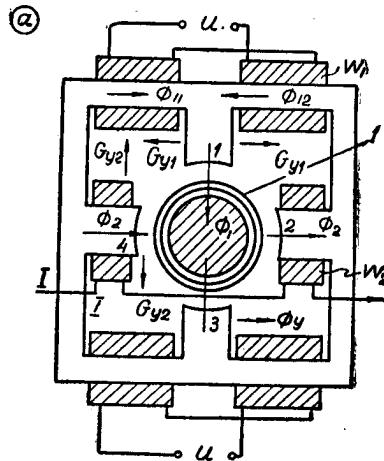
Потоки нижнего магнитопровода:

н. с. катушек = 170 ас; $I = 5$ а; $\Phi_i = 4,4 \cdot 10^{-5}$ вб.

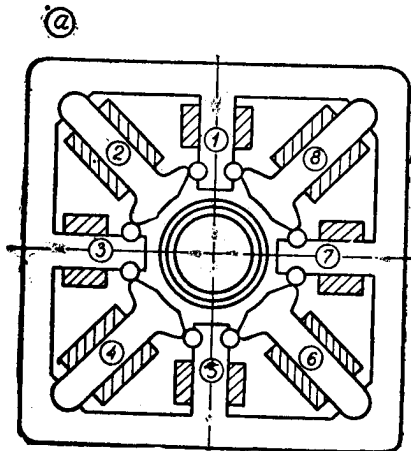
Экранированная система рассматривается состоящей из двух полюсов: первый — без экрана, второй — охватываемый экраном (фиг. 14, 3). При вычислении магнитных сопротивлений для каждого полюса учитывают сопротивление у зазора, для сборного магнитопровода учитывают дополнительно магнитное сопротивление стыков (С).



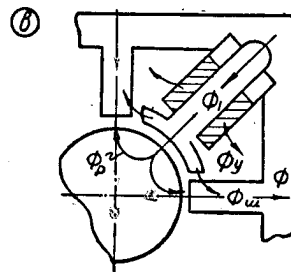
Фиг. 14, 3. Распределение потоков в экранированной системе:
1 — сердечник; 2 — экран.



Фиг. 14, 4. Магнитная система в 4-полюсном исполнении:
1 — ротор.



Фиг. 14, 5. Магнитная система в 8-полюсном исполнении.



Система с барабаном. Рассмотрим для примера магнитную систему четырехполюсного исполнения реле мощности (фиг. 14, 4). Катушки тока размещаются на сердечниках полюсов, катушки напряжения — на ярме магнитопровода, который выполняется сборным.

Поток Φ_1 пронизывает ротор два раза. Магнитное сопротивление воздушного промежутка учитывается как удвоенное сопротивление воздушного промежутка у одного полюса. Вычисление этого сопротивления или обратной ему величины магнитной проводимости удобней производить, пользуясь разверткой полюса и барабана.

На фигуре 14.5 показано исполнение восьмиполюсной индукционной системы и распределение потоков одного из полюсов.

14.3. КОЭФФИЦИЕНТ РАССЕЯНИЯ

Значения коэффициентов рассеяния полюсов σ — определяют для вычисления индуктивного сопротивления катушек. Коэффициент рассеяния определяют как отношение

$$\sigma = \frac{\psi_p + \psi_y}{\psi_p},$$

где ψ_p и ψ_y — соответственно числа потокосцеплений потоков рабочего полюса и утечки.

Для удобства расчетов принимают во внимание не числа потокосцеплений, а соответствующие им проводимости. В ряде случаев потоки утечки не сцепляются со всеми витками катушки. Поэтому принимают в таких случаях значение G_{yc} — приведенное по числу потокосцеплений значение проводимости.

Примеры. 1) Тангенциальная система, верхний магнитопровод (фиг. 14, 1):

$$G_{yc} = \frac{G_y}{3}; \quad \sigma = \frac{G_p + G_{u1} + G_{yc}}{G_p},$$

где G_{u1} и G_y — вычисленные по геометрическим размерам проводимости в обе стороны сердечника.

2) Система с барабаном — катушки на ярме (фиг. 14, 4 а)

$$\sigma = \frac{G_p + G_{y1}}{G_p},$$

где G_{y1} — значение проводимости утечки, вычисленное по геометрическим размерам.

3) Катушки на полюсах. Поток Φ_2 сцепляется со всеми витками катушек. Поток утечки каждого из полюсов сцепляется только с частью витков одной из катушек.

Число потокосцеплений рабочего потока $2\omega_2\Phi_2$;
Число потокосцеплений потока утечки $2\omega_2\Phi_y$;

$$G_{yc} = \frac{G_{y2}}{3}; \quad \sigma = \frac{G_p + G_{yc}}{G_p}.$$

14-4. УГОЛ ПОТЕРЬ

Углом потерь будем в дальнейшем считать угол между векторами тока катушки и потока системы. Для определения угла потерь необходимо вычислить следующие значения (см. фиг. 14, 6):

- н. с. для проведения потоков в воздухе F_a и стали F_c ;
- н. с., компенсирующую потери на гистерезис и вихревые токи F_n и потери в роторе F_p .

Для вычисления указанных значений н. с. задаются определенной величиной рабочего потока, такой, чтобы индукция в магнитопроводе была порядка $0,4 \frac{\text{вб}}{\text{см}^2}$. Это обычное значение индукции в индукционных системах.

Тогда

$$F_a = \frac{\Phi}{\sqrt{2} G}; F_c = av \cdot l,$$

где av — удельные ампервитки, определяемые по кривым намагничивания материала;

l — средняя длина силовой линии в стали магнитопровода.

$$F_n = \frac{P}{U} \omega = \frac{\sqrt{2} P}{\omega \sigma \Phi},$$

где P — потери в ваттах на гистерезис и вихревые токи;

σ — коэффициент рассеяния системы.

Значение I_p вычисляется по формулам, приведенным в параграфе 15-5. Число витков ротора принимается равным 1.

$$\operatorname{tg} \gamma = \frac{F_p + F_n}{F_a + F_c}. \quad (14-1)$$

Значение γ несколько меняется при изменении потока в системе. Практически, однако, это изменение незначительно и угол потерь считают постоянным. В некоторых случаях удобнее вычислять угол потерь по приведенным значениям токов.

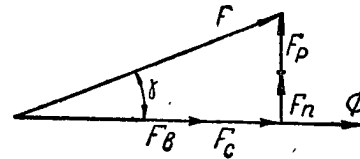
14-5. ВЕКТОРНЫЕ ДИАГРАММЫ

Векторные диаграммы строят для получения данных относительно составляющих тока катушки и углов между векторами.

Тангенциальная система. Верхний магнитопровод (фиг. 14, 7). Построение диаграммы начинают с потока Φ_p — рабочего потока, пронизывающего диск; учитывая, что система рассматривается как линейная, задаются любым значением Φ_p , но при условии, что значение индукции в системе получится небольшим, порядка $0,4 \cdot 10^{-4} \text{ вб/см}^2$.

При построении векторной диаграммы:

- составляющие тока $I_{\mu 0}$ определяют по потоку Φ_p и проводимостям G_p ;

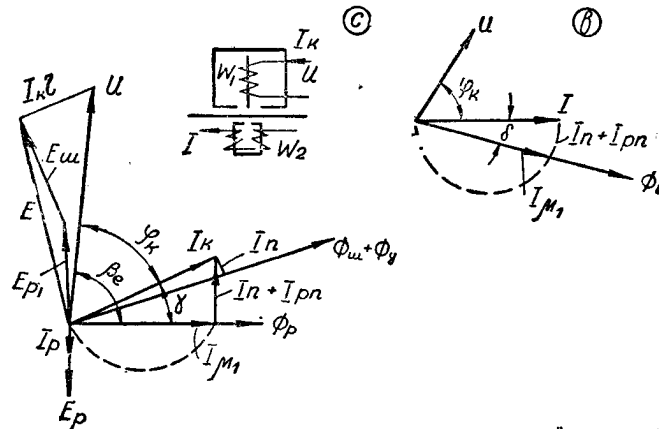


Фиг. 14, 6. Векторная диаграмма.

2) составляющие I_{pc} определяют по полному потоку $\Phi = \Phi_p + \Phi_{ш} + \Phi_y$, т. е. при нахождении значения H по кривой намагничивания учитывают значение B , а не B_p ;

3) составляющие тока потерь I_n также находят по потоку Φ . Находят полное значение потерь P в ваттах. Распределяют эти потери пропорционально потокам Φ_p и $\Phi_y + \Phi_{ш}$:

$$P_1 = P \frac{\Phi_p}{\Phi}; P_2 = P \frac{\Phi_y + \Phi_{ш}}{\Phi}.$$



Фиг. 14, 7. Векторные диаграммы тангенциальной системы.

Приведенные значения токов:

$$I_{pn} = \frac{I_p}{\omega}; I_{n1} = \frac{P_1}{U}; I_{n2} = \frac{P_2}{U}.$$

Из векторной диаграммы определяют значения углов:

- 1) φ_k — между U и I_k ; 2) потерь γ .

В данном исполнении системы оба угла остаются приблизительно постоянными при разных напряжениях (при условии полного заполнения окна магнитопровода). В расчетах систем часто применяют значение β_e — угла $U - \Phi_p$; $\beta_e = \varphi_k + \gamma$. В исполняемых индукционных системах значение β_e колеблется в пределах $85 - 105^\circ$. Меньший угол получается при большем зазоре шунта, больший — при минимальном зазоре. Угол γ — бывает порядка 12° .

Нижний магнитопровод обычно снабжается катушками тока, которые включаются последовательно (фиг. 14, 7 c). Построение диаграммы (фиг. 14, 7 b) начинают с вектора Φ_i , пронизывающего диск. Задаются определенным значением Φ_i . Данные векторной диаграммы:

$$E_p = 4,44 f \Phi_i \text{ вольт};$$

$$I_{pn} = \frac{I_p}{2\omega_2}; I_n = \frac{P}{U_c};$$

U_c — падение напряжения на катушках.

При вычислении потерь на гистерезис и вихревые токи учитывается также и поток рассеяния магнитопровода.

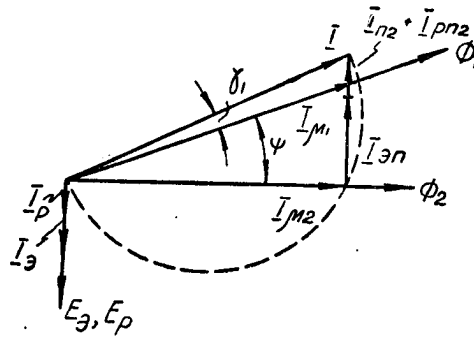
I_{μ} определяется из соотношения $\sqrt{2}I_{\mu}2\omega_2G_p = \Phi_i$. Из векторной диаграммы определяют угол потерь γ_n :

$$\operatorname{tg} \gamma_n = \frac{I_{pn} + I_n}{I_{\mu}}$$

В определенной магнитной системе угол потерь γ_n остается приблизительно постоянным при различных значениях тока I , значение его в исполненных системах — порядка 12° .

Угол φ_n — между вектором тока I и вектором падения напряжения на катушке U_c — порядка $60-62^\circ$.

Векторная диаграмма экранированной системы (фиг. 14, 8). Поток Φ магнитопровода разделяется у полюсов на два потока: Φ_1 — протекающего в неэкранированной части зазора и Φ_2 — в экранированной. Магнитная система рассматривается как система с двумя полюсами: 1 и 2 (фиг. 14, 3). Магнитные сопротивления по пути прохождения потоков Φ_1 и Φ_2 различны, поэтому оба потока не равны друг другу и сдвинуты по фазе на угол ψ . Векторные диаграммы для реле с катушкой тока и напряжения строятся аналогично. Построение диаграммы начинают с потока Φ_2 .



Фиг. 14, 8. Векторная диаграмма экранированной системы.

Данные векторной диаграммы:
 $I_{эп}$ — приведенное значение тока экранов:

$$I_{эп} = \frac{E_{э}}{r_{э}\omega} = \frac{\omega\Phi_2}{\sqrt{2}r_{э}\omega} \text{ ампер.}$$

При вычислении $r_{э}$ учитывают, что оба экрана — на нижнем и верхнем полюсах — включены параллельно. Тогда

$$r_{э} = \frac{r_a + r_n}{2}; \omega_{э} = 1.$$

Далее пренебрегают индуктивным сопротивлением экрана, весьма малым сравнительно с омическим. Вычисление остальных параметров диаграммы производят, как указано выше.

Угол ψ между потоками Φ_1 и Φ_2 определяется в основном значением тока $I_{эп}$ и с некоторым приближением вычисляется

по следующей формуле:

$$\operatorname{tg} \psi = \frac{I_{\partial n}}{I_{\mu 2}} = \frac{\omega \Phi_2 \sqrt{2} \omega_1 G_2}{r_s \cdot \omega \Phi_2} = \frac{\omega G_2}{r_s}. \quad (14-2)$$

Угол ψ бывает порядка 45—50°.

Векторные диаграммы систем с барабанным ротором строятся по одному из указанных методов.

14-6. ВЛИЯНИЕ РЕАКЦИИ РОТОРА

Во всех рассмотренных случаях влияние реакции ротора учтено при построении диаграммы введением приведенного тока I_{pn} , индуктированного в диске.

Влияние реакции ротора оказывается на уменьшении потоков полюсов. Поэтому экспериментальную проверку параметров магнитной системы необходимо производить над системами со вставленным ротором. При проведенных испытаниях четырехполюсной системы с барабанным ротором получены следующие данные, характеризующие влияние реакции ротора.

Полюс	н. с. ампер- витков	Поток, вебер •	
		Ротор вставлен	Ротор вынут
С катушкой тока . . .	156	$4,7 \cdot 10^{-5}$ 80%	$5,9 \cdot 10^{-5}$ 100%
С катушкой напряже- ния	205	$7 \cdot 10^{-5}$ 90%	$7,75 \cdot 10^{-5}$ 100%

14-7. СПОСОБЫ ИЗМЕНЕНИЯ УГЛОВ β_e и β_i

Важными для расчета параметрами индукционной системы являются углы:

β_e — между вектором напряжения U и рабочим потоком Φ_p полюса с катушкой напряжения. Значение β_e (фиг. 14,7) определяется как сумма углов φ_k и γ и зависит от величины шунтирующего потока магнитопровода, величины омического сопротивления катушки и других данных системы.

β_i — между вектором тока I катушки тока и потоком Φ полюса. Обычно угол β_i равен углу потерь γ (фиг. 14,7). Каждая форма индукционной системы характеризуется определенным значением углов φ_k и γ , а следовательно, β_e и β_i . Эти углы остаются приблизительно постоянными при различных режимах работы реле, так как индукция в магнитопроводе системы не достигает значения насыщения.

При разработке некоторых исполнений реле необходимо иметь значения β_e и β_i отличные от нормальных значений данной системы. Существует несколько способов изменения этих углов.

Изменение угла β_e . Для изменения угла в небольших пределах — порядка 20—30° — включают последовательно с катуш-

кой напряжения добавочное омическое сопротивление; при этом уменьшается угол φ_k , а следовательно и угол β_e .

Введением последовательно с катушкой напряжения емкостного сопротивления можно уменьшить значение β_e в больших пределах.

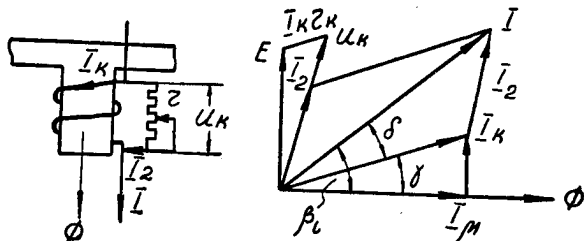
Изменение угла β_i . Для изменения этого угла применяется несколько способов.

1) Шунтирование катушки тока омическим сопротивлением (фиг. 14, 9). Из векторной диаграммы

$$I_2^2 = I_k^2 + I^2 - 2I_k I \cos \delta; \quad (a)$$

$$U_k = I_k Z_k = I_2 r_2. \quad (б)$$

В полученных соотношениях неизвестны I_k , I_2 и r_2 , значения Z_k и угла γ известны.



Фиг. 14, 9. Схема для изменения угла β_i .

Задавая значение I_k в виде определенной величины от тока I сети (например $I_k = 0,8I$), можно, пользуясь выражениями (а) и (б), найти значение r , необходимое для получения угла δ .

Тогда

$$\beta_i = \delta + \gamma.$$

2) На сердечнике насаживается коротко замкнутая гильза.

При наличии на сердечнике такой гильзы (фиг. 14, 10) угол β_i увеличивается. Приведенное к числу витков ω основной катушки значение тока в гильзе

$$I_{2n} = \frac{I_2}{\omega} = \frac{E_2}{r_2 \omega} = \frac{\omega \Phi}{\sqrt{2} r_2 \omega}; \quad (a)$$

далее

$$\Phi = \sqrt{2} I_p \omega G, \quad (б)$$

где r_2 — омическое сопротивление гильзы; индуктивным сопротивлением пренебрегают.

Подставляя (б) в (а), получим

$$I_{2n} = \frac{I_2}{\omega} = \frac{\omega I_p G}{r_2}. \quad (в)$$

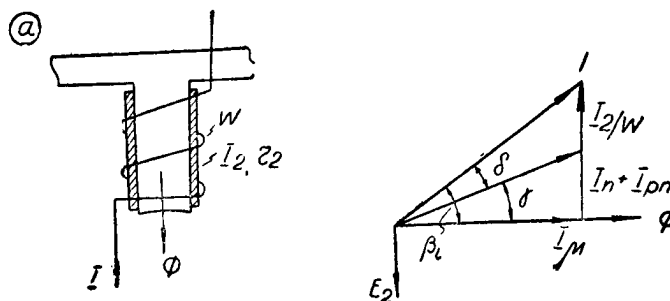
В данной системе известно значение угла γ , значением угла β_i задаются:

$$\operatorname{tg} \beta_i = \frac{I_{2n} + I_n + I_{2n}}{I_\mu} = \frac{I_\mu \operatorname{tg} \gamma + \omega I_\mu G / r_2}{I_\mu}, \quad (z)$$

или

$$\operatorname{tg} \beta_i = \operatorname{tg} \gamma + \frac{\omega G}{r_2}. \quad (d)$$

Из (d) определяют значение r_2 и геометрические размеры гильзы. Гильза выполняется обычно из меди.



Фиг. 14, 10. Исполнение полюса с гильзой:
I — гильза.

3) На сердечнике насаживается дополнительная катушка, включаемая на емкость. В дополнительной катушке (фиг. 14, 11) индуцируется ток I_2 ; фаза вектора I_2 определяется соотношением между индуктивным и емкостным сопротивлениями.

Обычно подбирают емкость к условиям резонанса. Приведенное значение тока I_{2n} вторичной обмотки определится из соотношения

$$I_{2n} = \frac{I_2 \omega_2}{\omega_1} = \frac{E_2 \omega_2}{Z_2 \omega_1} = \frac{\omega \Phi \omega_2^2}{\sqrt{2} Z_2 \omega_1} = \omega I_\mu \omega_2^2 \frac{G}{Z_2}.$$

При настройке вторичного контура в резонанс

$$I_{2n} = \omega I_\mu \omega_2^2 \frac{G}{r_2}. \quad (a)$$

В главе 2 показано, что при определенных заданных габаритах катушки

$$r_2 = k \omega_2^2.$$

Тогда

$$I_{2n} = \omega I_\mu \frac{G}{k}. \quad (b)$$

Из векторной диаграммы (фиг. 14, 11 б)

$$\operatorname{tg} \beta_i = \frac{I_{2n} + I_n + I_{2n}}{I_\mu} = \operatorname{tg} \gamma + \frac{\omega G}{k},$$

или

$$k = \frac{\omega G}{\operatorname{tg} \beta_i - \operatorname{tg} \gamma}$$

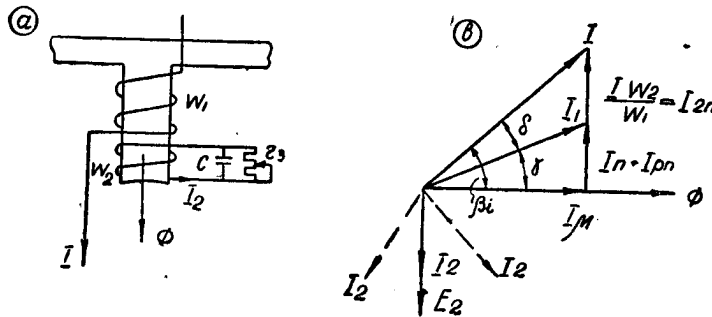
Для установления числа витков катушки задаются определенным значением емкости порядка $1 \div 2 \mu F$. Из условия

$$\omega \omega^2 G = \frac{1}{\omega C}$$

находим

$$\omega = \frac{1}{\omega \sqrt{GC}}$$

По известным, заданным в конструкции, габаритам катушки и по числу витков вычисляют значение диаметра проволоки. Точная



Фиг. 14, 11. Исполнение полюса с добавочной катушкой.

настройка для получения необходимого угла осуществляется регулируемым электрическим сопротивлением, шунтирующим емкость C .

Из рассмотренных трех способов изменения угла β_i наиболее часто применяется третий — введение дополнительной катушки, включенной на емкость, так как при этом способе получается меньшее значение мощности в $ва$, потребляемой катушкой.

ГЛАВА 15

ТОКИ В РОТОРЕ

15-1. ЭЛЕКТРОДВИЖУЩИЕ СИЛЫ, НАВЕДЕННЫЕ В РОТОРЕ

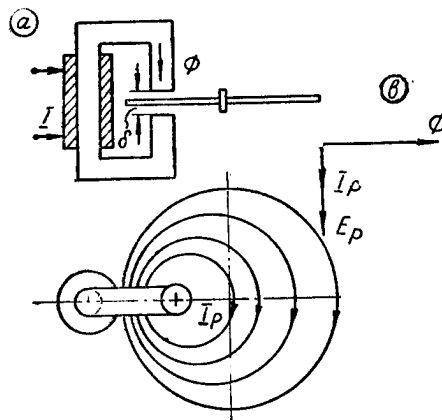
При наличии в магнитной системе потока, пересекающего ротор, в последнем возникают токи индукции, вследствие наведенных в нем электродвижущих сил. Различают два вида э. д. с., наведенных в роторе:

а) э. д. с. трансформации, получающаяся при пронизывании ротора магнитным потоком полюсов;

б) э. д. с. резания, получающаяся при пересечении вращающимся ротором поля полюсов.

При неподвижном роторе существует только э. д. с. трансформации, при вращении ротора добавляется еще э. д. с. резания.

а) Электродвижущая сила трансформации. На фигуре 15, 1 показан элемент системы с диском. Поток, созданный в магнитной системе обмоткой с током I , пересекает диск, расположенный в воздушном зазоре δ . В диске индуцируется э. д. с. E_p , отстающая по фазе от потока на 90° .



Фиг. 15, 1. Токи трансформации в диске.

Под действием э. д. с. в диске протекает ток I_p . При проведении расчетов принимают число витков диска равным единице. Далее пренебрегают индуктивным сопротивлением диска и полагают, что токи I_p и э. д. с. E_p совпадают по фазе.

Действующее значение э. д. с.

$$E_p = 4,44 f\Phi \text{ вольт,}$$

где f — частота в гц, Φ — в веберах; $\omega = 1$.

Обозначив электрическое сопротивление диска r_p , получаем

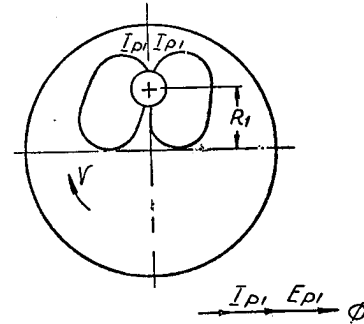
$$I_p = \frac{E_p}{r_p} \text{ ампер.}$$

Аналогичная картина получается и в системе с барабаном (фиг. 15, 3).

Поток Φ , определяемый н. с. катушек 1 и 3, проходит от полюса 1 через ротор и сердечник к полюсу 3, наводит в роторе э. д. с. E_p . Направление токов показано на разрезе барабана, а также на развертке барабана в плане (фиг. 15, 3).

б) Электродвижущая сила резания. При вращении ротор пересекает поток магнитной системы, и в нем (роторе) индуцируются э. д. с. резания E_{p1} , создающие токи резания I_{p1} , препятствующие вращению ротора (согласно закону Ленца).

Направление э. д. с. перпендикулярно вектору V , скорости вращения ротора. Вектор E_{p1} совпадает по фазе с вектором потока. Картина распределения токов резания показана на фигуре 15, 2.



Фиг. 15, 2. Токи резания в диске.

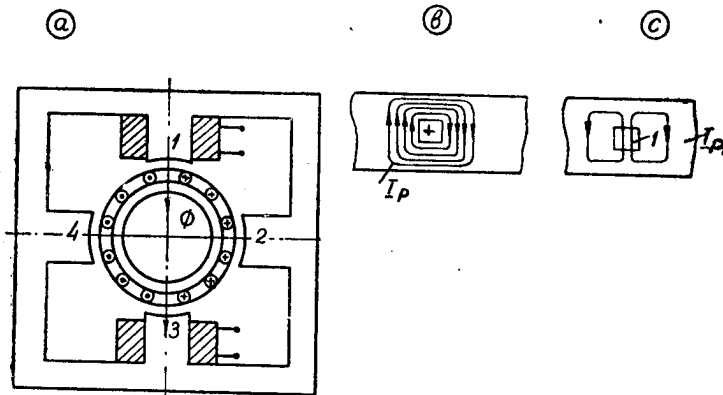
Действующее значение э. д. с. резания

$$E_{p1} = \frac{B}{\sqrt{2}} V l_n \text{ вольт,}$$

где B — индукция в $\frac{\text{об}}{\text{см}^2}$;

V — скорость вращения ротора под полюсом в см/сек ;

l_n — длина в см проводящей нити ротора под полюсом в направлении, перпендикулярном движению ротора.



Фиг. 15, 3. Разрез и развертки барабана:
а — разрез; б — токи трансформации; с — токи резания.

Обозначив электрическое сопротивление диска для трубок токов резания r_{p1} , получаем

$$I_{p1} = \frac{E_{p1}}{r_{p1}} \text{ ампер.}$$

На фигуре 15, 3 показана в развертке картина распределения токов резания в барабане.

15-2. ВЫЧИСЛЕНИЕ ЗНАЧЕНИЙ ТОКОВ ТРАНСФОРМАЦИИ

Токи трансформации в роторе рассматриваются в виде отдельных трубок. При расчете системы необходимо определить значение плотности тока, протекающего в определенном месте ротора, например под определенным магнитным полюсом.

Затруднения возникают при определении значения сопротивления определенной трубки тока.

Наиболее просто определяется значение r_p при симметричном расположении круглого полюса по отношению к диску. В этом случае токи трансформации I_p имеют вид концентрических трубок. Однако в практических случаях полюс всегда располагается несимметрично относительно диска: имеются исполнения дисков не круглой формы, сечение полюса имеет вид прямоугольника, квадрата.

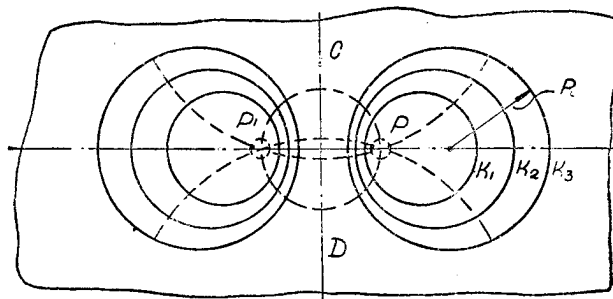
Во всех этих случаях отступления от полной симметрии токи в диске или барабане располагаются в виде трубок, граничные поверхности которых имеют весьма разнообразные и часто не поддающиеся анализу формы. В сложных случаях распределения токов в роторе подсчет сопротивлений трубок тока трансформации производится графо-аналитическим путем.

Наиболее часто исполняются системы с круглым диском с полюсами, расположенными несимметрично относительно центра (фиг. 15, 1). В этих случаях токи I_p рассматриваются в виде трубок с граничными уровнями, которые представляются в виде неконцентрических цилиндрических поверхностей. В системах с барабаном получается аналогичная картина распределения токов (фиг. 15, 3).

При расположении полюсов согласно фигурам 15, 1 и 15, 3 можно получить выражение для аналитического определения значения тока. Для этого пользуются методом отображений магнитного поля — по аналогии с методом отображений электрического поля.

15-3. МЕТОД ОТОБРАЖЕНИЙ ДЛЯ СИСТЕМ С ДИСКОМ

Рассмотрим круглый диск радиусом R с полюсом P , расположенным вне центра диска (фиг. 15, 4). Трубки токов имеют граничные окружности $k_1, k_2, k_3 \dots$. По методу отображений диск рассматривается как часть бесконечно протяженного листа, на



Фиг. 15, 4. Линия токов в диске.

который действует два поля: полюса P — действительного, и полюса P' — воображаемого — обратного полюсу P полярности.

Полюсы P и P' индуктируют в бесконечно большом листе токи, равные индуктированному значению э. д. с., деленному на омическое сопротивление трубки тока.

Эквипотенциальные линии э. д. с. представляют собой радиальные линии, выходящие из каждого полюса и образующие окружности с центрами, расположенными на линии CD . В каждом месте эквипотенциальные линии э. д. с. располагаются нормально к линиям токов — k_1 — k_2 — k_3 , которые также представляют собой окружности.

г) Плотности токов. При вычислении плотности токов j_{a1} и j_{a2} учитывают, что трубки токов располагаются концентрично с полюсами P и P' . Это вытекает из метода отображений (см. выше). Тогда, учитывая расстояния точки A от центров полюсов τ и τ' , пользуясь (15-2) и (15-3), получаем

$$j_{a1} = \frac{dl_{a1}}{\Delta d\tau} = \frac{E}{\Delta dr d\tau} = \frac{\omega \Phi \gamma \Delta d\tau}{\sqrt{2} 2\pi r \Delta d\tau} \text{ ампер.}$$

Или при $\gamma = 34 \cdot 10^{-4} \frac{1}{\text{ом} \cdot \text{см}}$ (уд. проводимость алюминия),
 $f = 50 \text{ гц}$ $j_{a1} = 1,2 \cdot 10^7 \frac{\Phi}{\tau}$. (15-4)

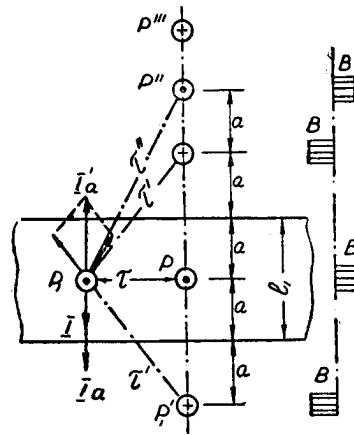
Аналогично

$$j_{a2} = 1,2 \cdot 10^7 \frac{\Phi}{\tau'}. \quad (15-5)$$

Из выражений (15-4) и (15-5) видно, что плотности токов в диске обратно пропорциональны расстоянию точки A от центров полюсов.

15.4. МЕТОД ОТОБРАЖЕНИЙ ДЛЯ СИСТЕМ С БАРАБАНОМ

Рассмотрим систему с барабаном. Для удобства вычислений будем пользоваться разверткой барабана в плоскости (фиг. 15,6).



Фиг. 15, 6. Диаграмма для определения токов в барабане.

В такой системе барабан в развертке рассматривается как часть бесконечно большого листа, на который действуют поля: действительного полюса P и воображаемых полюсов P' и P'' разной полярности.

Все эти полюсы индуцируют в бесконечно большом листе токи, равные индуцированному значению э. д. с., деленному на сопротивление по пути тока. Вокруг полюсов P и P', P'' располагаются трубки токов, ограниченные окружностями.

Если вырезать одну или несколько трубок тока, это не влияет на протекание процесса во всем листе бесконечно больших размеров. Поэтому в листе можно вырезать часть, по расположению и размерам соответствующую развертке барабана с полюсом P .

В данном случае получается следующее правило для вычисления значений токов трансформаций в определенной точке A развертки барабана: находят положение воображаемых полюсов и рассматривают значение тока I в точке A состоящим из суммы токов:

I_a — определяемым э. д. с. E от поля полюса P ;
 I'_a и I''_a — определяемым э. д. с. от поля воображаемых полюсов P', P'' и т. д.

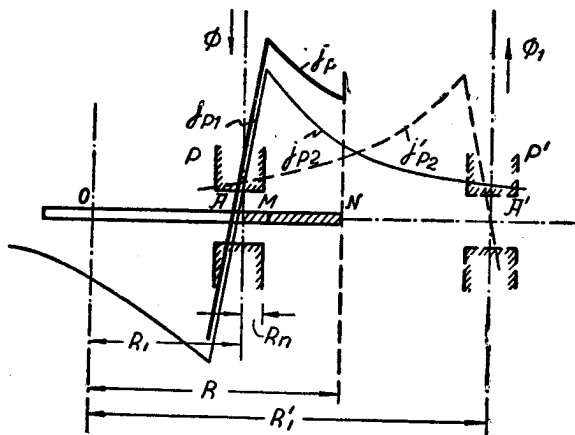
Число воображаемых полюсов равно бесконечности. Практически достаточно рассмотреть действие только двух ближайших воображаемых полюсов, так как величины токов I_a , индуцируемых полем отдаленных воображаемых полюсов, незначительны.

Месторасположение воображаемых полюсов находится в зависимости от расстояния a от центра действительного полюса до края развертки барабана. Обычно располагают полюс посередине длины барабана. Поэтому расстояние ближайших воображаемых полюсов до края развертки также равно a (фиг. 15, 6).

Значения плотностей токов вычисляют согласно формуле (15-4).

15-5. ПОЛНОЕ ЗНАЧЕНИЕ ТОКА ТРАНСФОРМАЦИИ

Для учета влияния реакции ротора необходимо определить величину тока трансформации I_p , индуцированного полем полюса P во всем роторе. Вычисление будем производить, пользуясь методом отображений и полученными ранее соотношениями.



Фиг. 15, 7. Распределение плотностей токов трансформации в диске.

а) *Вычисление токов трансформации в диске.* Все трубки токов, индуцируемых в диске, проходят через сечение AN . Задача заключается в вычислении тока I_p , протекающего по сечению AN (фиг. 15, 7).

Ток I_p состоит из следующих частей:

- i_{p1} — определяемого полем полюса P — зона AM ($R_k < R_n$);
- i_{p2} — » » » P — зона MN ($R_k > R_n$);
- i_{p3} — » » » P' — зона AN .

$$1) I_{p1}; E_{p1} = \frac{\omega \Phi}{\sqrt{2}} \frac{R_k^2}{R_n^2}; j_{p1} = 1,2 \cdot 10^7 \Phi \frac{R_k^2}{R_n^2} \frac{1}{R_k};$$

$$dI_{p1} = j_{p1} \Delta dR_k;$$

$$I_{p1} = 1,2 \cdot 10^7 \Phi \Delta \frac{1}{R_n^2} \int_{R_{n0}}^{R_n} R_k dR_k = \frac{1,2 \cdot 10^7 \Phi \Delta}{2} \quad (a)$$

$$2) I_{p2}; j_{p2} = 1,2 \cdot 10^7 \Phi \frac{1}{R_k};$$

$$dI_{p2} = j_{p2} \Delta dR_k = 1,2 \cdot 10^7 \Phi \Delta \frac{dR_k}{R_k}$$

$$I_{p2} = 1,2 \cdot 10^7 \Phi \Delta \int_{R_k = AM}^{R_k = AN} \frac{dR_k}{R_k} = 1,2 \cdot 10^7 \Phi \Delta \ln \frac{AN}{AM} \quad (b)$$

$$3) I'_{p2} = 1,2 \cdot 10^7 \Phi \Delta \ln \frac{A'A}{A'N} \quad (c)$$

$$4) I_p = I_{p1} + I_{p2} + I_{p3},$$

$$\text{или } I_p = 1,2 \cdot 10^7 \Phi \Delta \left(0,5 + \ln \frac{AN A'A}{AM A'N} \right).$$

Величины AN , $A'A$, AM и $A'N$ можно выразить в функции основных параметров системы: R , R_1 , R'_1 и R_n ; полное значение тока трансформации в диске при наличии поля Φ полюса P

$$I_p = 1,2 \cdot 10^7 \Phi \Delta \left[0,5 + \ln \frac{(R - R_1)(R'_1 - R_1)}{R_n(R'_1 - R)} \right], \quad (15-6)$$

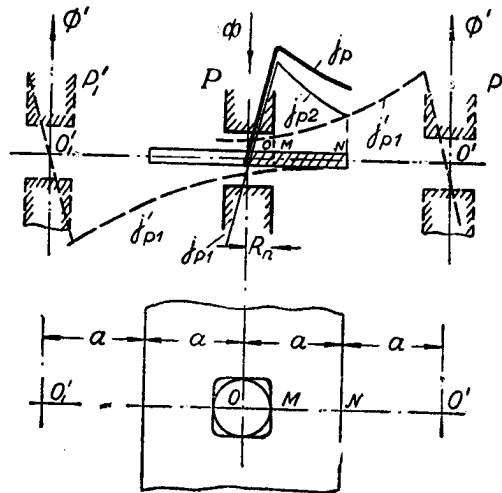
где Φ в веберах.

б) *Вычисление токов трансформации в барабане.* Принимают для упрощения задачи только два воображаемых полюса (фиг. 15, 8). Вычисление токов проводится по аналогии с методом, примененным для диска.

Полюс	Зона	Радиус	Составляющие токи
P	OM	$R_k < R_n$	$I_{p1} = \frac{1,2 \cdot \Phi \Delta \cdot 10^7}{2};$
P	MN	$R_k > R_n$	$I_{p2} = 1,2 \cdot 10^7 \Phi \Delta \ln \frac{ON}{OM};$
P'	$O'N$		$I'_{p2} = 1,2 \cdot 10^7 \Phi \Delta \ln \frac{O'O}{O'N}$
P'_1	O'_1N		$I'_{p1} = 1,2 \cdot 10^7 \Phi \Delta \ln \frac{O'_1N}{O'_1O}$

Учитывая, что $ON = a$; $OM = R_n$; $O'O = 2a$; $O'N = a$ и т. д., получаем полное значение тока трансформации в роторе при наличии поля Φ полюса P :

$$I_p = 1,2 \cdot 10^7 \cdot \Phi_{\Delta} \left[0,5 + \ln \frac{a}{R_n} + \ln 2 - \ln 1,5 \right]. \quad (15-7)$$



Фиг. 15, 8. Распределение плотностей токов трансформации в барабане.

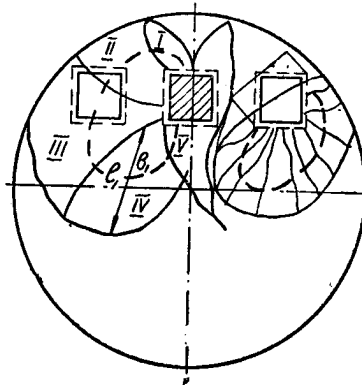
15-6. ТОКИ РЕЗАНИЯ

Трубки токов резания имеют в роторе весьма сложную конфигурацию, что затрудняет проведение аналитических расчетов. В таких случаях применяют графические методы расчета. На чертеже, изображающем диск или развертку барабана, вычерчивают эквипотенциальные линии э. д. с. резания, а по ним определяют граничные кривые трубок токов. Трубки разбивают на участки и вычисляют сопротивления этих участков.

На фигуре 15, 9 показано примерное построение трубок токов резания под полюсами тангенциальной системы. Трубка разбита на шесть участков.

Сопротивление каждого участ-

$$r_i = \frac{l_i}{\gamma b_i \Delta} \text{ ом};$$



Фиг. 15, 9. Графическое определение трубок токов резания.

Полное сопротивление трубки $r = \sum r_i$;
 Значение тока резания $I = \frac{E}{r}$.

Пример. Вычислить сопротивление r тока резания, определяемого полем среднего полюса (фиг. 15, 9). Ток резания рассматривается состоящим из двух частей — левой и правой трубки. Толщина диска равна 0,1 см. Вычислим сопротивление левой трубки. Вся трубка разбита на 6 частей. Данные расчета сведены в таблицу.

Часть	Средняя длина l_i	Средняя ширина b_i	r_i , ом
I	1,2	1,2	$0,0294 \cdot 10^{-8}$
II	1,1	1,7	$0,0182 \cdot 10^{-8}$
III	2,7	2,9	$0,0274 \cdot 10^{-8}$
IV	2,0	3,3	$0,0178 \cdot 10^{-8}$
V	0,9	1,15	$0,0230 \cdot 10^{-8}$
VI	1,7	1,15	$0,0435 \cdot 10^{-8}$

$$\sum r_i = 0,1593 \cdot 10^{-8}$$

$$r = \frac{0,1593 \cdot 10^{-8}}{2};$$

ГЛАВА 16

МОМЕНТЫ НА РОТОРЕ

Условием получения рабочего момента на роторе индукционной системы является наличие в системе двух или нескольких потоков, смещенных в пространстве и не совпадающих во времени.

Взаимодействием индуктированных в роторе токов трансформации с потоками полюсов создаются вращающие моменты на роторе. В основу расчетов для определения рабочих моментов принят метод отображений: учет влияния краевого эффекта производится с помощью воображаемых полюсов.

Б. К. Булем разработан ряд простых соотношений для определения рабочих моментов на диске. Расчет систем с барабаном детально разработан Л. А. Гельбухом, который дал ряд соотношений, позволяющих вычислять значения рабочих и тормозных моментов и анализировать влияние отдельных параметров для установления оптимальных значений.

Ниже приводится метод определения рабочих моментов на роторе, основанный на вычислении плотностей индуктированных токов. Этот метод является общим для расчета систем с диском и с барабаном и предусматривает применение простых геометрических фигур, что позволяет ориентироваться в выводе основных формул.

16-1. СРЕДНЕЕ ЗНАЧЕНИЕ СИЛЫ

Электромагнитные моменты на роторе вычисляются по закону Био и Саварра. Мгновенное значение силы Q_t при длине l_n проводника под полем

$$Q_t = iB_l l_n.$$

Пусть

$$B_t = B \sin \omega t; \quad i = I_m \sin(\omega t - \epsilon),$$

где ϵ — угол отставания тока по отношению к потоку, с которым он взаимодействует (фиг. 16,1).

Тогда

$$Q_t = I_m \cdot Bl_n \sin \omega t \sin(\omega t - \epsilon),$$

среднее значение за период

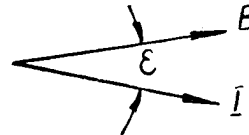
$$Q = \frac{1}{T} \int_0^T Q_t dt = \frac{I_m B \cdot l_n}{T} \int_0^T \sin \omega t \cdot \sin(\omega t - \epsilon) dt.$$

Значение

$$\int_0^T \sin \omega t \sin(\omega t - \epsilon) dt = \frac{T}{2} \cos \epsilon.$$

Тогда

$$Q = \frac{I_m B l_n}{2} \cos \epsilon.$$



Фиг. 16, 1. Векторная диаграмма.

Учитывая, что действующее значение тока $I = \frac{I_m}{\sqrt{2}}$, а также — $1 \frac{\partial \mathcal{J}}{\partial \text{см}} = 10,2 \text{ кГ}$, получим

$$Q = 7,2 \cdot I B l_n \cos \epsilon \text{ кГ}, \quad (16-1)$$

где $B = \frac{\partial \mathcal{B}}{\partial \text{см}^2}$; l_n — в см.

Соотношение (16-1) является основным для вычисления средних значений сил на роторе.

16-2. ОБЩЕЕ УРАВНЕНИЕ МОМЕНТОВ

Рассмотрим индукционную систему с полюсами P_1 с полем Φ_1 , и P_2 — с полем Φ_2 (фиг. 16,2). Значения токов трансформации в роторе — I_1 и I_2 ; ψ — угол между векторами потоков. Моменты на роторе определяются взаимодействием поля Φ_1 и тока I_2 , а также поля Φ_2 и тока I_1 .

Значение токов трансформации $I_1 = k_2 \Phi_1$; $I_2 = k_3 \Phi_2$. Тогда общий момент на роторе

$$M = M_1 + M_2 = k_p \Phi_1 \Phi_2 \sin \psi. \quad (16-2)$$

Необходимо обратить внимание на знаки моментов M_1 и M_2 ; мгновенные значения M_{1t} , M_{2t} могут быть разных знаков, но среднее за период значение M_1 и M_2 одного и того же знака. Поэтому при определении суммарного момента складывают значения составляющих моментов.

Фиг. 16, 2. Векторная диаграмма.

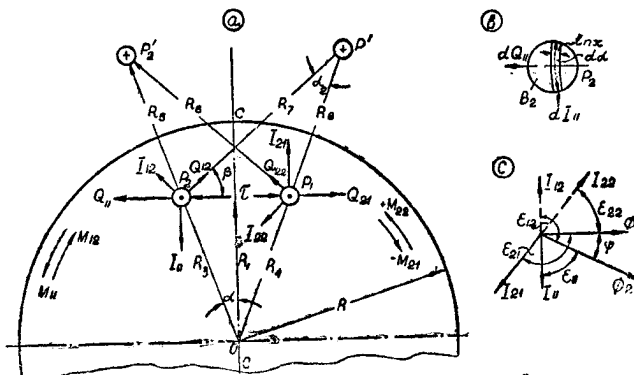
По формуле (16-2) нельзя судить о направлении вращения ротора, обычно его определяют по следующему мнемоническому правилу: ротор вращается в сторону от полюса с опережающим потоком к полюсу с отстающим.

В соответствии с диаграммой, приведенной на фигуре (16-2), ротор вращается от полюса P_1 (поле Φ_1) к полюсу P_2 (поле Φ_2).

Вычисление рабочих моментов на диске экранированной системы

1. Момент M_{11} . Элементарное значение силы, определяемое током I_{11} полюса P_1 и полем Φ_2 полюса P_2 (фиг. 16-3) согласно формуле (16-1):

$$dQ_{11} = 7,2 dI_{11} B_2 l_{nx} \cos \epsilon_{11} \quad (a).$$



Фиг. 16, 3. Моменты на диске экранированной системы.

Элементарное значение объема диска под полюсом $\Delta d\alpha$. Тогда $dI_{11} = j_{11} \Delta d\alpha \quad (b).$

Согласно (15-4)

$$j_{11} = 1,2 \cdot 10^7 \frac{\Phi_1}{\tau} \quad (c); \quad \cos \epsilon_{11} = \sin \psi \quad (d).$$

Из (a, b, c, d)

$$dQ_{11} = 8,65 \cdot 10^7 \frac{\Phi_1}{\tau} B_2 \Delta \sin \psi l_{nx} d\alpha;$$

$$Q_{11} = \int dQ_{11} = 8,65 \cdot 10^7 \frac{\Phi_1}{\tau} B_2 \Delta \sin \psi \int_{S_2} l_{nx} d\alpha.$$

Согласно фигуре 16,3 b:

$$\int_{S_2} l_{nx} d\alpha = S_2; \quad B_2 S_2 = \Phi_2.$$

Окончательно

$$Q_{11} = 8,65 \cdot 10^{10} \Phi_1 \Phi_2 \frac{\Delta}{\tau} \sin \psi \quad \Gamma. \quad (16-3)$$

$$M_{11} = Q_{11} R_1 \quad \Gamma \text{ см.}$$

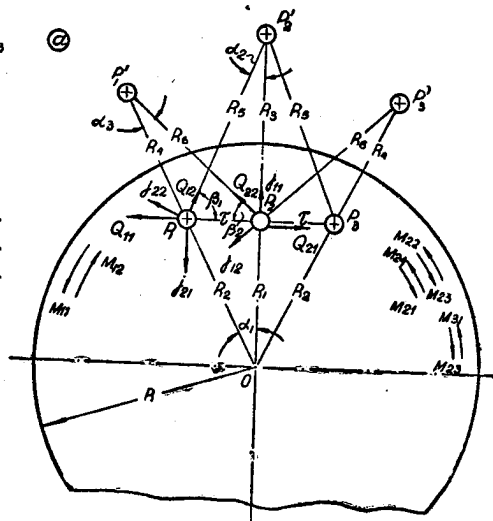
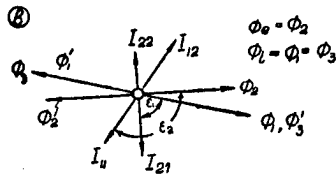
II. Момент M_{12} , определяемый током, индуцируемым полем воображаемого полюса P'_1 и полем Φ_2 полюса P_2 :

$$Q_{12} = 8,65 \cdot 10^{10} \Phi_1 \Phi_2 \frac{\Delta}{R_7} \sin \psi;$$

$$M_{12} = Q_{12} R_1 \cos \beta.$$

III. Момент M_{21} , определяемый током, индуцируемым полем полюса P_2 и полем Φ_1 полюса P_1 :

$$Q_{21} = 8,65 \cdot 10^{10} \Phi_1 \Phi_2 \frac{\Delta}{\tau} \sin \psi; \quad M_{21} = Q_{21} R_1.$$



IV. Момент M_{22} — определяемый током, индуцируемым полем воображаемого полюса P'_2 и полем полюса P_1 :

$$Q_{22} = 8,65 \cdot 10^{10} \cdot \Phi_1 \Phi_2 \frac{\Delta}{R_6} \sin \psi;$$

$$M_{22} = Q_{22} R_1 \cos \beta.$$

V. Суммарный момент M является суммой моментов, определяемых основными полюсами (M_{11} ; M_{21}), за вычетом моментов, определяемых воображаемыми полюсами (M_{12} и M_{22}):

$$M = 2 \cdot 8,65 \cdot 10^{10} \Phi_1 \Phi_2 R_1 \Delta \left(\frac{1}{\tau} - \frac{\cos \beta}{R_6} \right) \sin \psi. \quad (16-4)$$

VI. Значения отдельных величин. Заданными являются R_1, τ, R , остальные значения определяют из геометрических соотношений:

а) Расположение воображаемых полюсов P'_1 и P'_2

$$R_3 (R_3 + R_5) = R^2; \quad R_5 = R_8,$$

б) $R_6 = R_7$; R_6 определяем из треугольника $P_1 P'_2 O$; $\cos \beta$ — из треугольника $P_1 P_2 P'_1$.

VII. Значение k_p определяем из выражения (16-2)

$$M = k_p \Phi_1 \Phi_2 \sin \psi.$$

Учитывая (16-4), получаем

$$k_p = 17,3 \cdot 10^{10} \cdot \Delta R_1 \left(\frac{1}{\tau} - \frac{\cos \beta}{R_6} \right). \quad (16-5)$$

Пример. Вычислить значение k_p для экранированной дисковой системы.

Задано: $R = 4,8 \text{ см}$; $\Delta = 0,1 \text{ см}$; $\tau = 1,1 \text{ см}$; $R_1 = 3,14 \text{ см}$.

Данные вычислений:

$$R_3 = 3,2 \text{ см}; \quad R_5 = R_8 = 4 \text{ см};$$

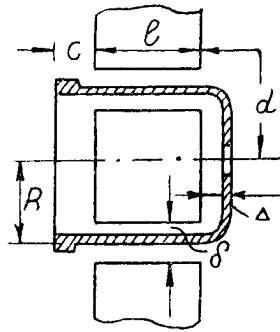
$$R_6 = R_7 = 4,3 \text{ см}; \quad \cos \beta = 0,393.$$

Подставляя полученные значения в (16-5), имеем $k_p = 4,44 \cdot 10^{10}$.

Если известны Φ_1 , Φ_2 , ψ , то момент на диске определяют по формуле (16-2);

Пример. $\Phi_1 = 2 \cdot 10^{-5} \text{ вб}$; $\Phi_2 = 1 \cdot 10^{-5} \text{ вб}$; $\psi = 45^\circ$;

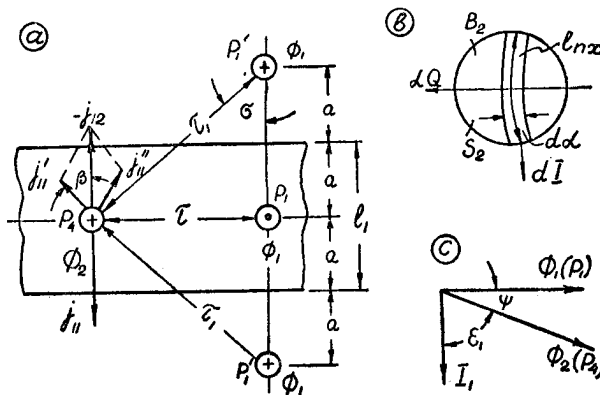
$$M = 4,44 \cdot 10^{10} \cdot 2 \cdot 10^{-5} \cdot 1 \cdot 10^{-5} \cdot 0,707 = 6,26 \text{ Гсм}.$$



Фиг. 16, 5. Разрез магнитной системы.

Вычисление рабочих моментов на диске тангенциальной системы

1. Моменты на диске. Вычисление сил производят по формуле (16-3). Силы, определяемые индуктированными токами основных полюсов (P_1 , P_2 и P_3), считаем положительными, а воображаемых — отрицательными (фиг. 16,4).



Фиг. 16, 6. Развертка системы и диаграммы. •

Полус	Величины, определяющие силу		Момент	Знак момента
	Поток	Ток, индуктир. потоком полюса		
P_1	Φ_1	P_2	$M_{11} = 8,65 \cdot 10^{10} \Delta \Phi_1 \Phi_2 R_1 \frac{1}{\tau} \sin \psi$	+
		P'_2	$M_{12} = 8,65 \cdot 10^{10} \Delta \Phi_1 \Phi_2 \frac{R_1 \cos \beta_1}{R_6} \sin \psi$	-
P_2	Φ_2	P_1	$M_{21} = M_{11}$	+
		P'_1	$M_{22} = 8,65 \cdot 10^{10} \Delta \Phi_1 \Phi_2 \frac{R_1 \cos \beta_2}{R_6} \sin \psi$	-
P_3	Φ_3	P_3	$M_{31} = M_{21} = M_{11}$	+
		P'_3	$M_{32} = M_{22}$	-
P_3	$\Phi_3 = \Phi_1$	P_2	$M_{33} = M_{11}$	+
		P'_1	$M_{32} = M_{12}$	-

Суммарный момент $M = 4M_{11} - 2M_{12} - 2M_{22}$.

$$M = 17,3 \cdot 10^{10} \Phi_1 \Phi_2 \Delta \cdot R_1 \left(\frac{2}{\tau} - \frac{\cos \beta_1}{R_5} - \frac{\cos \beta_2}{R_6} \right) \sin \psi \text{ Гсм.}$$

Здесь Φ_1 и Φ_2 — в веберах, линейные размеры в см.

$$k_p = 17,3 \cdot 10^{10} \Delta R_1 \left(\frac{2}{\tau} - \frac{\cos \beta_1}{R_5} - \frac{\cos \beta_2}{R_6} \right). \quad (16-6)$$

Определение значений R_5 , R_6 , $\cos \beta_1$, $\cos \beta_2$ производится аналогично показанному выше.

Вычисление рабочих моментов на роторе барабанной системы

I. Развернутая длина ротора

$$l_1 = c_1 + d_1 + l,$$

где c_1 , d_1 — приведенные к размеру Δ , размеры c и d (фиг. 16,5).

II. M_{12} — момент, определяемый взаимодействием поля Φ_2 полюса P_2 (фиг. 16,6) и тока I_1 , индуктированного в роторе полем Φ_1 , полюса P_1 .

а) Плотности токов:

$$j_{11} = 1,2 \cdot 10^7 \frac{\Phi_1}{\tau}; \quad j'_{11} = j_{11} = -1,2 \cdot 10^7 \frac{\Phi_1}{\tau_1};$$

$$j_{12} = j'_{11} + j''_{11} = -2 \cdot 1,2 \cdot 10^7 \frac{\Phi_1}{\tau_1} \cos \beta; \quad \cos \beta = \sin \sigma = \frac{\tau}{\tau_1};$$

$$j_1 = j_{11} - j_{12} = 1,2 \cdot 10^7 \Phi_1 \left(\frac{1}{\tau} - \frac{2\tau}{\tau_1^2} \right).$$

Обозначим

$$c_2 = \frac{1}{\tau} - \frac{2\tau}{\tau_1^2};$$

тогда

$$j_1 = 1,2 \cdot 10^7 c_2 \Phi_1.$$

б) Сила. Элементарное значение силы согласно (16-1, см. фиг. 16-6),

$$dQ = 7,2 dl_1 B_2 j_{nx} \cos \varepsilon;$$

элементарный объем ротора под полюсом Δda .

$$dl_1 = j_1 \Delta da = 1,2 \cdot 10^7 c_2 \Phi_1 \Delta da; \quad \cos \varepsilon = \sin \psi;$$

$$Q = 8,65 \cdot 10^7 c_2 \Delta \cdot \Phi_1 \sin \psi B_2 \int_{S_2} l_{nx} da \quad \text{кГ.}$$

$\int_{S_2} l_{nx} da = S_2$ — площади кругового полюса P_2 ; $B_2 S_2 = \Phi_2$, тогда

$$Q_{12} = 8,65 \cdot 10^{10} c_2 \Delta \cdot \Phi_1 \Phi_2 \cdot \sin \psi \quad \text{грамм}; \quad M_{12} = Q_{12} R_1.$$

III. Суммарный момент на роторе — M . Поле Φ_2 , взаимодействуя с током I_3 , индуцируемого полем Φ_1 полюса P_3 , создает момент $M_{13} = M_{12}$.

Таким образом, момент, создаваемый на каждом полюсе, равен $2M_x$, при числе полюсов n , общий момент $M = 2nM_x$. Поэтому общий момент четырехполюсного реле:

$$M = 8 \cdot 8,65 \cdot 10^{10} \Delta \cdot R c_2 \Phi_1 \Phi_2 \cdot \sin \psi.$$

Обозначая

$$k_p = 8 \cdot 8,65 \cdot 10^{10} \Delta \cdot R c_2 = 69,2 \cdot 10^{10} \Delta R \left(\frac{1}{\tau} - \frac{2\tau}{\tau^2 + l_1^2} \right),$$

имеем

$$M = k_p \Phi_1 \Phi_2 \sin \psi.$$

Пример. Определить значение k_p для четырёхполюсного реле. Данные реле: $\Delta = 0,05$ см, $R = 1,425$ см; $\tau = 2,36$ см.

$$c_1 = d_1 = 2,15 \text{ см}; \quad l = 3 \text{ см}; \quad f = 50 \text{ гц.}; \quad \gamma = 34 \cdot 10^4 \frac{1}{\text{ом см}}.$$

Условная длина ротора $l_1 = 3 + 2 \cdot 2,15 = 7,3$ см.

$$k_p = 1,65 \cdot 10^{10}.$$

Формула Б. К. Буля для вычисления рабочих моментов на диске

1. Общее выражение для момента

$$M = 2 \frac{10,2 \Delta f 10^8}{\rho} k \Phi_1 \Phi_2 \sin \psi \quad \text{Гсм.}$$

где ρ — удельное сопротивление материала диска, *ом. см.*
 f — частота в *гц*; Φ_1 и Φ_2 — максимальное значение потоков в *вб*;
 k — геометрическая постоянная.
 Ниже, в таблице, приводятся данные для определения геометрической постоянной.

Коэффициенты к формуле Буля

Таблица 24

$k_1 = \frac{h_0}{2a_0} \left[1 - \frac{4a_0^2}{(1 - a_0^2 - h_0^2)^2 + 4a_0^2} \right];$ $a_0 = \frac{a}{R};$ $b_0 = \frac{b}{R};$ $h_0 = \frac{h}{R};$	
$k_2 = \frac{2h_0}{a_0} \left[1 - \frac{a_0^2}{(1 - a_0^2 - h_0^2)(1 - h_0^2) + a_0^2} \right]$	
$k_3 = 2k_4 + k_{14} + k_{23};$ $k_4 \text{ и } k_{23} \text{ как } k_1;$ $k_4 = \frac{h_0}{a_0 - b_0} \left[1 - \frac{(a_0 - b_0)^2}{(1 - a_0^2 - h_0^2)(1 - b_0^2 - h_0^2) + (a_0 + b_0)^2} \right]$	

Фиг. 16, 7. К формуле Б. К. Буля.

Формулы Л. А. Гельбуха для вычисления рабочих моментов на роторе барабанной системы

Л. А. Гельбухом предложена следующая формула для вычисления рабочего момента одной пары полюсов:

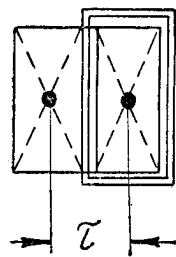
$$M = \frac{4}{981} \frac{\omega \gamma \Delta l_1^3 R B_1 B_2}{\pi^4} \sin \phi \sum \frac{1}{k^4} \operatorname{sh}^2 \frac{k\pi b}{2l_1} e^{-\frac{k\pi \tau}{l_1}} \times$$

$$\times \left[\cos \frac{k\pi d_1}{l_1} (-1)^{k+1} + \cos \frac{k\pi c_1}{l_1} \right]^2 \cdot 10^{-5} \text{ Гсм.}$$

Здесь B_1, B_2 — индукция в гауссах, k — коэффициент, характеризующий порядок гармоник, b — размер полюсной дуги в *см.*
 l_1 — условная длина диска ротора, $l_1 = l + c_1 + d_1$; l — длина полюса; $\gamma = 34 \frac{\text{м}}{\text{ом м.м.}^2}$.

Общий момент системы: при наличии в системе n полюсов значение момента увеличивается в n раз.

Формула для вычисления тормозного момента одного полюса:



Фиг. 16, 8. Определение полюсного деления.

$$M_{\partial} = 2,04 \cdot \frac{\gamma \Delta \omega l_1^2}{\pi} R^2 B^2 \sum_{k=1}^{\infty} \frac{1}{k^3} \operatorname{sh} \frac{-k\pi b}{l} e^{-\frac{k\pi b}{2l}} \times$$

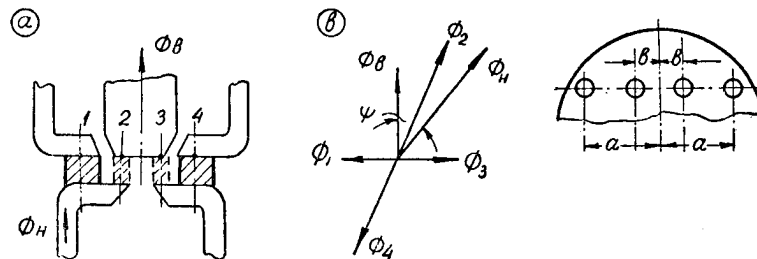
$$\times \left[\cos \frac{k\pi d_1}{l} (-1)^{k+1} + \cos \frac{k\pi c}{l} \right] 10^{-8} \text{ грамм} \cdot \text{см.}$$

Общий тормозной момент равен сумме тормозных моментов отдельных полюсов.

О форме полюсов

Как правило, в имеющихся конструкциях индукционных систем полюсы бывают прямоугольной формы. Однако при выводе формул для расчета рассматриваются точечные или круглые полюсы, благодаря чему значительно упрощается вывод этих формул.

Имеются кривые поправочных коэффициентов, по которым можно учитывать переход от точечных полюсов к прямоугольным.

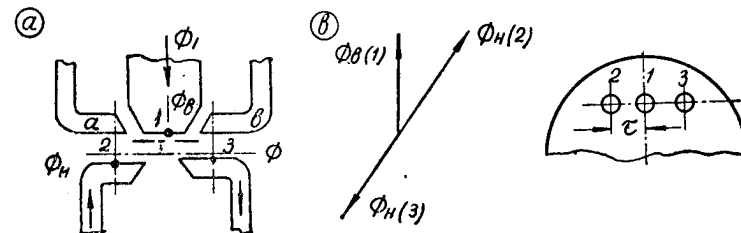


Фиг. 16, 9. Рассмотрение тангенциальной системы в 4-полюсном исполнении.

Соотношение размеров прямоугольных полюсов обычно бывает таким, что значение поправочного коэффициента близко единице.

Полюсное деление

Под полюсным делением τ понимают расстояние между центрами полюсов. Точное значение τ вычисляется в системах с ба-



Фиг. 16, 10. Рассмотрение тангенциальной системы в 3-полюсном исполнении

рабаном, где величина полюсного деления определяется расстоянием по окружности между центрами полюсов.

Несколько сложнее решается задача в экранированных системах. Условно считают τ — как расстояние между центрами тяжести фигур, представляющих собой площади полюсов (фиг. 16,8). Еще более сложной является задача установления τ в тангенциальных системах. Сложная картина распределения поля в воздушных зазорах между верхним и нижним магнитопроводами приводит к некоторой неопределенности в установлении центров полюсов.

Б. К. Буль рассматривает такую систему состоящей из четырех условных полюсов (фиг. 16,9). В каждом из них учитывается действие двух потоков, являющихся частью потоков Φ_s и Φ_n , причем в одних полюсах эти потоки складываются, в других вычитаются.

В дальнейшем принят другой вариант решения задачи: система принимается состоящей из трех полюсов (фиг. 16,10):

- 1 — выходит поток Φ_s ;
- 2 — » » Φ_n ;
- 3 — входит » Φ_n .

При трехполюсном варианте не учитывается то, что в местах (а) и (б) потоки Φ_s и Φ_n суммируются и вычитаются. Это упрощение принято потому, что моменты, определяемые потоками $\Phi_s + \Phi_n$ и $\Phi_s - \Phi_n$, в указанных местах имеют разные знаки и на величину общего момента не влияют. Трехполюсный вариант удобен для анализа и для определения направления вращения ротора при повороте одного из векторов поля на определенный угол.

Сложная картина распределения поля не позволяет точно учитывать значения потоков, протекающих в полюсах системы с четырьмя полюсами. Так, часть потока Φ_s замыкается не по воздуху, а по скобе С (фиг. 14,1). Поэтому приходится проводить экспериментальную проверку расчетов. Такую проверку удобно производить, измеряя отдельно потоки Φ_s у полюса 1 и Φ_n у полюсов 2 и 3 и, таким образом, получают непосредственные данные для полей указанных трех полюсов.

Влияние краевого эффекта

Краевой эффект на роторе выражается в том, что в роторе конечных размеров векторы индуктированных токов под соответствующим полюсом отличаются от соответствующих векторов в случае бесконечных размеров ротора.

Как указывалось выше, учет краевого эффекта проводится с использованием воображаемых полюсов, действие которых аналогично действию основных. При этом моменты, определяемые воображаемыми полюсами, вычитают из моментов, определяемых основными.

Влияние удаленных полюсов

Выше показана методика вычисления момента на роторе с учетом взаимодействия полей и токов двух соседних полюсов.

В многополюсных индукционных системах с барабаном следует учитывать дополнительное влияние удаленных, не смежных полюсов. Так, в восьмиполюсной системе (фиг. 14, 5) момент на роторе определяется не только взаимодействием полей полюсов 1—2 и 1—8, но и несмежных полюсов 1—3, 1—4 и т. д. Вычисление моментов, определяемых несмежными полюсами, производят по приведенным выше соотношениям. По мере удаления несмежных полюсов соответствующее значение момента уменьшается.

Ниже приводятся экспериментальные данные, полученные для восьмиполюсной системы:

α°	45	90	135	180
$M\%$	100	32	5	0

Здесь α — пространственный угол между полюсами. Влияние элементов несмежных полюсов искажает электромагнитную характеристику реле. Поэтому приходится располагать катушки на полюсах системы так, чтобы моменты двух полюсов, смещенных на 90° , взаимно уравновешивались. Например, моменты полюсов 1—5 и 1—7 должны быть равны, но направлены в разные стороны. Значениями моментов полюсов, смещенных на 135° , например 1—4, пренебрегают.

16-3. ОБЩЕЕ ВЫРАЖЕНИЕ ДЛЯ ВЫЧИСЛЕНИЯ ТОРМОЗНОГО МОМЕНТА

Тормозные моменты возникают на роторе при его вращении в поле полюсов. Поток Φ полюса индуцирует в диске э. д. с. резания:

$$E_p = \frac{B}{\sqrt{2}} l_n V \text{ вольт,}$$

где

$$B = \frac{\Phi}{S} \text{ — индукция;}$$

$$l_n \text{ — длина под полюсом трубки тока резания;}$$

$$V \text{ — скорость движения ротора в месте, расположенном под полюсом } V = \omega R_1 \text{ см/сек.}$$

Ток резания

$$I_p = \frac{E_p}{r} = \frac{1}{\sqrt{2}r} B l_n \omega R_1 \text{ ампер,} \quad (a)$$

где r — сопротивление расположенной под полюсом трубки тока резания.

Среднее значение тормозной силы для переменного поля

$$Q_\theta = \frac{1}{\sqrt{2}} B I_p l_n \cos \varepsilon \frac{d\varepsilon}{d\omega}, \quad (b)$$

ε — угол между векторами потока и тока резания. Подставляя

из (а) в (б) значение I_p , учитывая значение плеча силы R_2 и переход от $\frac{\partial \mathcal{M}}{\partial \epsilon}$ в $\kappa \Gamma \cdot \text{см}$, получаем

$$M_{\partial} = \frac{1,02 \cdot 10^4 \Phi_n^2 l_n^2}{2 S^2 r} R_1 R_2 \omega \cos \epsilon. \quad (16-7)$$

При взаимодействии поля полюса с индуктированным им током резания угол $\epsilon = 0$.

Обозначая

$$k_{\partial} = \frac{1,02 \cdot 10^4 l_n^2 R_1 R_2}{2 S^2 r},$$

получаем

$$M_{\partial} = k_{\partial} \Phi^2 \omega. \quad (16-8)$$

Выражение (16-8) является основным для определения момента торможения, которое формулируется так: момент торможения пропорционален квадрату потока и окружной скорости ротора.

В индукционных системах возникают тормозные моменты, определяемые взаимодействием поля Φ_1 одного полюса с током, индуктируемым полем Φ_2 другого полюса.

Если $\angle \Phi_1 \Phi_2 = \psi$, то выражение для M_{∂} :

$$M_{\partial} = k_{\partial} \Phi_1 \Phi_2 \cos \psi. \quad (16-9)$$

Пример.

$$\Phi_1 = 1,8 \cdot 10^{-5} \text{ вб}; l_n = 1,56 \text{ см}; R_1 = R_2 = 2,65 \text{ см}.$$

$$\omega = 3,62; S = 2,84 \text{ см}^2.$$

$$r = 8,7 \cdot 10^{-5} \text{ (вычислено графическим методом)}$$

$$M_{\partial} = 0,146 \text{ Гсм}.$$

16-4. ТОРМОЗНЫЕ МОМЕНТЫ РАЗНЫХ СИСТЕМ

Экранированная система. Такая система состоит из двух полюсов с потоками Φ_1 и Φ_2 , сдвинутыми по фазе на угол ψ :

$$M_1 = k_{\partial 1} \Phi_1^2 \omega; M_2 = k_{\partial 2} \Phi_2^2 \omega; M_{12} = k_{\partial 12} \Phi_1 \Phi_2 \omega \cos \psi;$$

$$M_{21} = k_{\partial 21} \Phi_1 \Phi_2 \cdot \omega \sin \psi; M_{\partial} = M_1 + M_2 + M_{12} + M_{21}.$$

Тангенциальная система. Такая система рассматривается состоящей из трех полюсов: среднего с полем Φ_1 и крайних с полем Φ_2 , причем $\angle \Phi_1 \Phi_2 = \psi$.

Суммарное значение тормозных моментов, определяемых взаимодействием поля Φ_1 с токами резания I_{p2} и поля Φ_2 с токами резания I_{p1} , равно нулю. Таким образом, необходимо учитывать только тормозные моменты, определяемые полями полюсов с индуктированными ими токами резания.

$$M_{\partial} = (k_{\partial 1} \Phi_1^2 + 2k_{\partial 2} \Phi_2^2) \omega \text{ грамм см}.$$

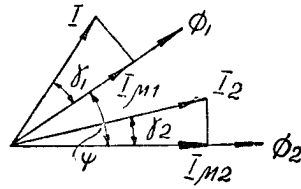
Система с барабаном. Вычисление тормозных моментов проводится по выведенным выше соотношениям (см. параграф 16-3).

Л. А. Гельбух предложил выражение для вычисления значения тормозного момента одного полюса системы с барабаном (см. выше).

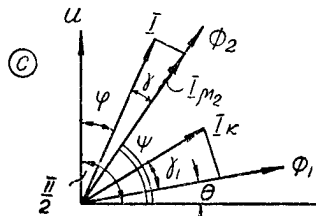
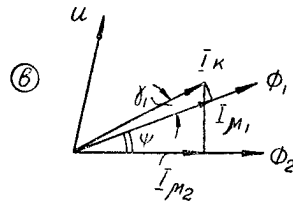
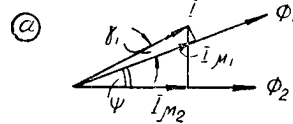
Общий тормозной момент многополюсного реле равен сумме тормозных моментов отдельных полюсов.

16-5. РАЗЛИЧНЫЕ ВИДЫ УРАВНЕНИЙ МОМЕНТОВ ИНДУКЦИОННЫХ СИСТЕМ

Полученные выше уравнения для определения рабочих моментов не всегда удобны для проведения расчетов индукционных систем. Поэтому применяют и другие виды уравнений. Ниже приводятся различные виды уравнений, выражающих значение рабочего момента в функции параметров систем.



Фиг. 16, 11. Векторная диаграмма.



Фиг. 16,12. Векторные диаграммы.

а) Исходное уравнение

$$M = k_r \Phi_1 \Phi_2 \sin \psi \quad (16-2).$$

б) Уравнение рабочего момента в функции н. с. системы. Если известны проводимости G_1 и G_2 для путей потоков Φ_1 и Φ_2 и углы потерь γ_1 и γ_2 , то можно составить следующие соотношения (фиг. 16,11):

$$\Phi_1 = \sqrt{2} \cdot I_{\mu 1} \omega_1 G_1; \quad (a)$$

$$\Phi_2 = \sqrt{2} \cdot I_{\mu 2} \omega_2 G_2; \quad (b)$$

$$I_{\mu 1} = I_1 \cos \gamma_1 \quad (c); \quad I_{\mu 2} = I_2 \cos \gamma_2. \quad (d)$$

Подставляя (a, b, c, d) в (16-2) получаем

$$M = 2k_r G_1 G_2 I_1 \omega_1 I_2 \omega_2 \cos \gamma_1 \cos \gamma_2 \sin \psi. \quad (16-10)$$

Обозначая

$$k_a = 2k_r G_1 G_2 \cos \gamma_1 \cos \gamma_2,$$

имеем

$$M = k_a I_1 \omega_1 I_2 \omega_2 \sin \psi. \quad (16-11)$$

Отметим здесь, что реакция ротора учтена при выводе формулы введением коэффициентов $\cos \gamma_1$ и $\cos \gamma_2$.

в) Уравнение рабочих моментов в функции параметров сети¹.

¹ Выводы значений моментов реле тока и напряжения приведены с учетом экранированной системы.

Реле тока. В таких реле оба потока Φ_1 и Φ_2 определяются в функции н. с. катушки реле, проводимости G_1 и G_2 и различны для обоих потоков, значение угла ψ остается неизменным при различных режимах работы реле.

Пользуясь диаграммой, изображенной на фигуре 16,12 а, имеем

$$\Phi_1 = \sqrt{2} I \omega G_1 \cos \gamma_1 (a); \quad \Phi_2 = \sqrt{2} I \omega G_2 \cos (\gamma_1 + \psi). \quad (b)$$

Подставляя (а) и (b) в (16-10), получаем

$$M = 2k_p G_1 G_2 \cdot \cos \gamma_1 \cos (\gamma_1 + \psi) (I \omega)^2 \sin \psi.$$

Учитывая, что $\sin \psi$ для экранированной системы = *const.* и обозначая $k_m = 2k_p G_1 G_2 \cos \gamma_1 \cos (\gamma_1 + \psi) \sin \psi \cdot \omega^2$, имеем

$$M_m = k_m I^2. \quad (16-12)$$

Реле напряжения. Построение реле напряжения аналогично реле тока. Основные соотношения:

$$\begin{aligned} \Phi_1 &= \sqrt{2} I_n \omega \cos \gamma_1 G_1; \\ \Phi_2 &= \sqrt{2} I_n \omega \cos (\gamma_1 + \psi) G_2; \\ I_n &= \frac{U}{Z_n}. \end{aligned}$$

Обозначая

$$k_n = 2k_p \frac{\omega^2}{Z_n^2} G_1 G_2 \cos \gamma_1 \cdot \cos (\gamma_1 + \psi) \sin \psi,$$

имеем

$$M_n = k_n U^2. \quad (16-13)$$

Реле мощности. В таких исполнениях реле имеется катушка напряжения, включаемая на напряжение U сети, катушка тока, через которую протекает ток I сети. Следует учитывать угол φ между векторами U и I . Основные соотношения (фиг. 16, 12 с):

$$\begin{aligned} \Phi_1 &= \sqrt{2} I_n \omega_1 G_1 \cos \gamma_1; \\ \Phi_2 &= \sqrt{2} I \omega_2 G_2 \cos \gamma_2; \\ I_n &= \frac{U}{Z_n}; \end{aligned}$$

$$\sin \psi = \cos (\varphi + \gamma + \theta) = \cos (\varphi + \alpha),$$

Здесь $\alpha = \gamma + \theta$.

Обозначая

$$k_m = 2k_p \frac{\omega_1 \omega_2}{Z_n} G_1 G_2 \cos \gamma_1 \cos \gamma,$$

имеем

$$M_m = k_m U I \cos (\varphi + \alpha) \quad (16-14)$$

Формулы (16-12), (16-13) и (16-14) удобны для пользования при расчете различных исполнений реле. Коэффициенты k_m , k_n и k_p определяют при проектировании индукционной системы.

В тех случаях, когда при расчете какого-либо исполнения реле получают значения потоков в комплексной форме и непо-

средственное вычисление значения угла ψ неудобно, применяют другое уравнение момента.

Так, если $\dot{\Phi}_1 = M_1 + jN_1$; $\dot{\Phi}_2 = M_2 + jN_2$, то момент на роторе пропорционален мнимой части произведения комплекса одного потока на сопряженный комплекс другого.

Мнимая часть произведения равна $N_1M_2 - M_1N_2$.

Тогда

$$M = k_p (N_1M_2 - M_1N_2).$$

При расчете сложных исполнений реле, например некоторых видов дистанционных, получаются комплексные выражения для токов \dot{I}_1 и \dot{I}_2 — двух катушек реле в функции тока I и напряжения U сети.

Например,

$$\dot{I} = M_1I + M_2U + j(N_1I + N_2U); \quad (a)$$

$$\dot{I}_2 = M_3I + M_4U + j(N_3I + N_4U). \quad (b)$$

Значение момента на роторе пропорционально мнимой части произведения комплекса (a) на сопряженный комплекс (b).

Значения потоков выражены в форме

$$\dot{\Phi}_{t1} = \Phi_1 e^{j(\omega t + \alpha)}; \quad \dot{\Phi}_{t2} = \Phi_2 e^{j(\omega t + \beta)}.$$

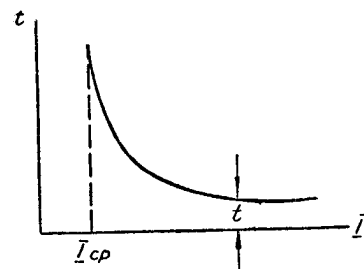
$$M = k_p \Phi_1 \Phi_2 \sin(\beta - \alpha).$$

ГЛАВА 17

ВРЕМЯ СРАБАТЫВАНИЯ

17-1. ОСНОВНЫЕ ТРЕБОВАНИЯ

Индукционные реле применяются в основном в схемах релейной защиты электрических установок и должны удовлетворять различным требованиям этих схем в отношении времени срабатывания. Характеристика $t = f(I)$, где I — параметр реле, представляет собой кривую (фиг. 17, 1). Обычно, время t реле относят к независимой части характеристики, т. е. к той части, в которой с увеличением параметра I значение t не изменяется. По условиям работы схем защиты выполняют две группы реле.



Фиг. 17, 1. Временная характеристика реле тока.

а) Быстродействующие, т. е. такие, у которых время срабатывания должно быть минимальным. Практически в различных исполнениях реле минимальное время — в независимой части характеристики — получается порядка 0,02—0,06 секунд.

б) Замедленно действующие, т. е. такие, у которых время срабатывания должно быть относительно большим. В исполненных реле время t — в независимой части характеристики — бывает порядка 0,5—10 секунд. Дисковые индукционные системы исполняются как быстродействующие, так и замедленно действующие, барабанные — только быстродействующие.

17-2. ОБЩЕЕ УРАВНЕНИЕ ДВИЖЕНИЯ РОТОРА

Рассмотрим для примера индукционное реле (фиг. 17,2) с магнитной системой 1, постоянным магнитом 2, пружиной 3 и контактным рычагом 4. Полный ход контактного рычага до замыкания контактов определяется углом α .

При вращении ротор находится под воздействием следующих моментов:

рабочего $M_p = k_p \Phi_1 \Phi_2 \sin \psi$;

тормозного переменных токов

$$M_\theta = (k_{\theta 1} \Phi_1^2 + 2k_{\theta 2} \Phi_2^2) \omega = k_{\theta n} \omega;$$

тормозного постоянных магнитов

$$M_{\theta, м} = k_{\theta 3} B^2 \omega = k_{\theta, м} \omega;$$

момента инерции ротора

$$M_u = J \frac{d^2 \alpha}{dt^2};$$

момента противодействующей пружины M_n .

В дальнейшем принимаются следующие допущения:

не учитывается момент трения, величина которого относительно мала;

противодействующий момент M_n компенсирован так, что при вращении ротора разность $M_p - M_n$ остается постоянной.

По принципу Даламбера

$$M_p = M_u + M_\theta + M_n,$$

или

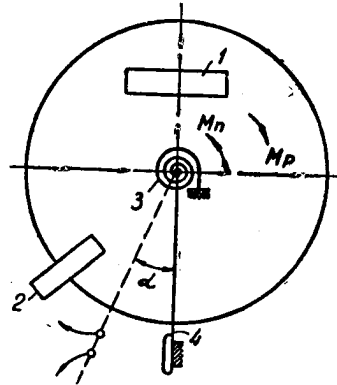
$$M_p - M_n = J \frac{d^2 \alpha}{dt^2} + k_\theta \frac{d\alpha}{dt}. \quad (17-1)$$

Здесь принято $k_{\theta n} + k_{\theta, м} = k_\theta$; $\omega = \frac{d\alpha}{dt}$

Решение уравнения (17-1) с учетом начальных условий:

$$\alpha = \frac{M_p - M_n}{k_\theta} \left[t - \frac{J}{k_\theta} \left(1 - e^{-\frac{k_\theta t}{J}} \right) \right]; \quad (17-2)$$

$$\omega = \frac{M_p - M_n}{k_\theta} \left(1 - e^{-\frac{k_\theta t}{J}} \right). \quad (17-3)$$



Фиг. 17, 2. Индукционная система.

Полученные выражения неудобны для анализа влияния на время t параметров системы. Поэтому для практических расчетов принимают следующие упрощения.

а) Значение $k_0 \frac{d\alpha}{dt} = 0$; такое упрощение допустимо для быстродействующих реле, так как время t в таких реле настолько мало, что ротор не успевает развивать значительных окружных скоростей. Кроме того, в таких реле отсутствует постоянный магнит.

б) В замедленно действующих реле время разгона ротора до установившейся скорости весьма мало сравнительно с общим временем t . Поэтому принимают

$$J \frac{d^2\alpha}{dt^2} = 0.$$

17.3. БЫСТРОДЕЙСТВУЮЩИЕ РЕЛЕ

Уравнение (17-1) с учетом $k_0 \frac{d\alpha}{dt} = 0$ принимает вид

$$M_p - M_n = J \frac{d^2\alpha}{dt^2}.$$

Решение этого уравнения

$$t = \sqrt{\frac{2J\alpha}{M_p - M_n}}. \quad (17-4)$$

Из (17-4) следует: для уменьшения времени срабатывания значения угла α и момента инерции J вращающихся частей должны быть минимальными. Для уменьшения момента инерции ротора последний выполняют из алюминия.

Время t уменьшается с увеличением рабочего момента. Минимальное значение рабочего момента $M_{p \text{ мин}} = M_n$. Обозначив

$$k = \frac{M_p}{M_{p \text{ мин}}},$$

получаем (17-4) в виде:

$$t = \frac{1}{\sqrt{k-1}} \sqrt{\frac{2J\alpha}{M_n}}.$$

Данные существующего исполнения четырехполосной системы с барабаном: раствор контактов — 1 мм, радиус траверзы — 30 мм; $\alpha = 0,035 \text{ рад}$.

$$J = 88 \cdot 10^{-3} \text{ Гсм. сек}^2; M_n = 0,35 \text{ Гсм},$$

$$\text{при } k = 3,4, t = 0,084 \text{ сек};$$

$$k = 5,1, t = 0,064 \text{ сек};$$

$$k = 30, t = 0,024 \text{ сек}.$$

17-4. ЗАМЕДЛЕННО ДЕЙСТВУЮЩИЕ РЕЛЕ

Из выражения (17-1) с учетом $J \frac{d^2\alpha}{dt^2} = 0$ имеем

$$M_p - M_n = k_\theta \frac{d\alpha}{dt}.$$

Решение этого уравнения

$$t = \frac{k_\theta \alpha}{M_p - M_n}. \quad (17-5)$$

Из (17-5) следует, что для увеличения t значения k_θ и α должны быть максимальными. С этой целью в системе ставят постоянные магниты с большой индукцией в зеве магнита (порядка $B = 2 \cdot 10^{-5}$).

Г Л А В А 18

ПАРАМЕТРЫ ИНДУКЦИОННОЙ СИСТЕМЫ

Параметрами устройства называют данные, характеризующие определенное исполнение этого устройства. При рассмотрении индукционных систем и реле учитывают следующие основные группы параметров.

Конструктивные, характеризующие конструктивные размеры частей системы: ротора, полюсов и другие. Выбор размеров отдельных частей и соотношений между этими размерами производится из условия получения оптимального решения.

Расчетные, характеризующие расчетные данные рассматриваемого исполнения системы. Расчетные параметры необходимы для расчета определенного вида реле. В таблицах 26 и 27 приведены примерные расчетные данные индукционных систем.

Эксплуатационные, характеризующие работу определенного вида реле: электрические характеристики реле, время срабатывания, потребляемую мощность, коммутационную способность контактной системы и другие.

Ниже рассматриваются отдельные виды параметров.

18-1. КОНСТРУКТИВНЫЕ ПАРАМЕТРЫ

При проектировании индукционной системы выбирают конструктивные размеры такими, чтобы получились наилучшие условия использования активных материалов системы. Исходной величиной является значение k_p — конструктивного коэффициента системы, определяющего собой величину вращающего момента на роторе в функции значений токов катушек.

Основные параметры системы с диском

а) Значение воздушного зазора δ должно быть минимальным, так как при этом условии получаются наибольшие значения потоков в системе при заданном значении намагничивающей силы. Минимальное значение δ принимается из конструктивных соображений так, чтобы при возможных отклонениях в размерах магнитной системы диск, вращаясь, не касался магнитопровода.

Обычные параметры систем с диском: толщина материала диска — 1 мм, расстояние от диска до магнитопровода в обе стороны по 0,5 мм, $\delta_{\min} = 2$ мм.

б) оптимальное значение угла ψ для экранированных систем получается при $\psi = 45^\circ$. Так, выражая значение $M_p = k_p \Phi_1 \Phi_2 \sin \psi$ в функции н. с. катушки, получаем

$$M_p = k_p 2mI^2 \omega^2 G_1 G_2 \cos \psi \sin \psi.$$

Здесь m — постоянная, определяемая реакцией ротора и потерями в стали.

Максимальное значение произведения $\cos \psi \sin \psi$ получается при угле $\psi = 45^\circ$.

в) Размеры полюсов и их расположение характеризуются следующими параметрами:

$$\alpha = \frac{R_n}{R},$$

R_n — радиус полюса; R — радиус диска;

$$\beta = \frac{\tau}{R},$$

τ — расстояние между центрами полюсов (полюсное деление);

$$\eta = \frac{R_1}{R},$$

R_1 — расстояние от линии центров полюсов до центра диска. На основании ряда проведенных расчетов получены следующие данные для оптимальных значений:

$$\alpha = 0,1; \beta = 0,15; \eta = 0,75.$$

Основные параметры системы с барабаном

а) Значение δ воздушного зазора между полюсом и сердечником также принимается из конструктивных соображений.

Пример исполнения системы с барабаном (фиг. 18, 1):
 Δ барабана равна 0,5 мм.

$$\delta = \frac{38,5 - 36,5}{2} = 1 \text{ мм.}$$

Размеры D_n и D_c исполняются по третьему классу точности, посадка C_8 .

Верхний предел отклонения в размере зазора получается равным 0,05 мм.

Следует отметить, что выполнение системы с зазором $\delta = 1$ мм связано с технологическими трудностями. Поэтому часто исполняют системы, в которых $\delta = 2$ мм.

б) Размеры полюсов и их расположение характеризуются следующими параметрами:

$$\alpha = \frac{b}{l_1} = \frac{\text{длина полюсной дуги}}{\text{длина барабана}};$$

$$\beta = \frac{l}{l_1} = \frac{\text{длина полюса}}{\text{длина барабана}};$$

$$\eta = \frac{b}{\tau} = \frac{\text{длина полюсной дуги}}{\text{полюсное деление}}.$$

Для определения оптимальных значений α , β , η значение k_p приводят к виду

$$k_p = \alpha(\alpha); k_p = \beta(\beta); k_p = \eta(\eta).$$

Принимая различные значения α , β и η , строят кривые и находят оптимальное решение.

По данным ряда расчетов оптимальное значение находится в пределах $\alpha = 0,3 \div 0,4$.

Значение β в исполненных системах:

$$4\text{-полюсная } \beta = 0,625; 8\text{-полюсная } \beta = 0,316.$$

18-2. РАСЧЕТНЫЕ ПАРАМЕТРЫ

Расчетные параметры вычисляются при проектировании индукционной системы и уточняются экспериментальной проверкой. Основные расчетные параметры:

Магнитные проводимости G для потоков системы. Вычисляют значения проводимостей для потоков рабочего и рассеяния.

Определяют значение коэффициента рассеяния полюсов

$$\sigma = \frac{G_p + G_{yc}}{G_p},$$

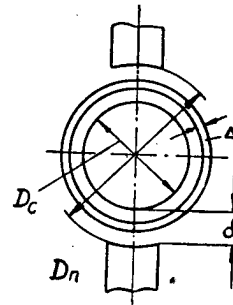
где G_{yc} — приведенное значение коэффициента рассеяния.

Углы потерь. Вычисляют значение γ для каждой группы полюсов. В исполненных системах, как правило, значение γ одинаково для всех групп полюсов. Так, в тангенциальной системе значение γ_0 — верхнего магнитопровода равно γ_n — нижнего.

Коэффициенты сопротивлений. Катушки каждого из полюсов системы характеризуются значениями коэффициентов сопротивлений: k_r — электрического, k_x — индуктивного, k_z — полного. Пользуясь этими значениями, можно определить все параметры катушки в функции числа ее витков.

Сопротивления: $r = k_r \omega^2$; $x = k_x \omega^2$; $z = k_z \omega^2$.

Потребляемая мощность $P = k_z (I\omega)^2$ вольт-ампер.



Фиг. 18, 1. Размеры частей магнитной системы.

Н. с. катушки $F = \frac{U}{k_r \omega}$; ампер-витков.

Вычисление k_r . Исходные соотношения:

$$r = \rho \frac{l_c \cdot 10^{-3}}{q} \omega \quad (a); \quad q = f_k \frac{l_k \cdot h_k}{\omega} \quad (b),$$

или

$$r = \rho \frac{l_c \cdot 10^{-3}}{f_k \cdot l_k h_k} \omega^2 = k_r \omega^2.$$

Здесь ρ — для нагретой меди; l_c, l_k, h_k — в мм;

f_k — в предварительных расчетах принимают $f_k = 0,5$.

Вычисление k_x : $x = \omega L = \omega \omega^2 G = \omega \omega^2 \sigma G_p = k_x \omega^2$.

Здесь G_p — проводимость для рабочего потока;

σ — коэффициент рассеяния полюса.

Коэффициенты моментов. Вычисляют значения k_p, k_n, k_m, k_m (см. параграф 16—5), которыми пользуются при расчете отдельных исполнений реле.

Термическая устойчивость. Для отдельных катушек полюсов вычисляют предельное значение допустимой из условий нагрева намагничивающей силы.

Исходное соотношение

$$P = \mu S \theta, \quad (a)$$

где P — суммарные потери, определяющие собой нагрев катушки и состоящие из электрических потерь в катушке и потерь на гистерезис и вихревые токи (P_n). Обычно значение P учитывают с повышенным значением коэффициента k_r . Принимают

$$P = k_r' (I\omega)^2; \quad (b)$$

μ — коэффициент теплоотдачи катушки;

θ — предельное значение превышения температуры катушки.

Для индукционных реле принимают $\theta = 60^\circ \div 65^\circ$.

Из (a) и (b)

$$I^2 k_r' \omega^2 = \mu S \theta,$$

или

$$(I\omega)_{np} = \sqrt{\frac{\mu S \theta}{k_r'}}. \quad (18-1)$$

При выборе сечения провода для катушки тока следует учитывать значение кратности тока короткого замыкания. Так, при кратности, равной 50, сечение провода принимают таким, чтобы при $I_{кз} = 50 I_n$ и времени протекания тока 1 сек, предельный нагрев меди обмотки не превосходил температуры 200° .

Таким образом, необходимо вычислять два значения термической устойчивости токовых катушек: номинальное, с учетом превышения температуры не более 65° , и односекундное, с предельным нагревом обмотки меди 200° .

Динамическая устойчивость. Предел динамической устойчивости определяется допустимыми вибрациями ротора. Некоторая

незначительная вибрация всегда наблюдается в системе, но при больших токах и напряжениях, подведенных к катушкам реле, вибрация усиливается настолько, что может привести к порче подшипников. Кроме того, сильные вибрации нарушают правильность работы реле и создают неприятное для слуха дребезжание.

Динамическая устойчивость определенной системы устанавливается экспериментально и характеризуется предельным значением ампер-витков катушек реле, допустимых для этой системы.

В обычных исполнениях индукционных систем предельное значение н. с. F_T катушек напряжения, допустимое из условий термической устойчивости, меньше значения н. с. F_θ — допустимого из условий динамической устойчивости. Так, при проведенных испытаниях реле с четырехполюсной барабанной системой получены следующие данные: катушка напряжения: $F_T = 450$; $F_\theta = 500$; катушка тока: F_T — длительные — 450; $F_\theta = 500$.

Таким образом, в данной системе $F_T < F_\theta$, и предельное значение н. с. выбирают, исходя из условий термической устойчивости. При этом следует учесть, что в полюсах с катушкой тока возникают при режимах короткого замыкания значительные потоки, приводящие к вибрации ротора. Поэтому нужно стремиться к тому, чтобы сталь полюсов с катушкой тока насыщалась уже при 500 ав. Достигают этого путем вырезов в жестях полюсов.

На фигуре 18, 2 приведен пример такого исполнения. Ширина выреза выбрана такой, чтобы в узкой части полюса индукция была порядка $1,5 \cdot 10^{-5}$, при н. с. катушки, равной 500.

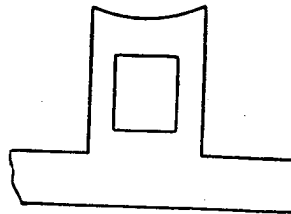
Данные системы: сечение стали полюса $q = 4,5 \text{ см}^2$; поток полюса, при $F = 500$, $\Phi = 16 \cdot 10^{-5} \text{ вб}$; сечение полюса в узкой части $q_1 = 1 \text{ см}^2$.

Добротность. Основным конструктивным параметром, характеризующим степень использования данного вида реле, считается коэффициент добротности $k_{дб}$ или, как говорят, «добротность» системы. При оценке системы учитывают две основные величины:

- моменты на роторе, определяющие работу контактной системы, а также чувствительность системы;
- потребляемую мощность катушек, определяющую нагрузку измерительных трансформаторов.

Поэтому для оценки качества индукционной системы принята та же формула, что и для асинхронных двигателей с коротко замкнутым ротором:

$$k_{дб} = \frac{\text{пусковой момент}}{\text{потребляемая мощность при пуске}}$$



Фиг. 18, 2. Полюс с вырезом.

Для индукционной системы значение $k_{\partial\delta}$ характеризует момент при 1 ва потребляемой мощности катушек:

$$k_{\partial\delta} = \frac{M}{\omega a} \frac{\Gamma_{cm}}{\omega a}.$$

Для определения значения $k_{\partial\delta}$ надо выразить значение $M_p = k_p \Phi_1 \Phi_2 \sin \psi$ в функции потребляемой мощности катушек. Найдем зависимость Φ_1 и Φ_2 от потребляемой мощности системы.

а) Полус с катушкой тока

$$\Phi = \sqrt{2} \cdot I \omega G \cos \gamma. \quad (a)$$

Падение напряжения на катушке тока

$$U_k = \frac{\omega \Phi \omega}{\sqrt{2}}. \quad (b)$$

Потребляемая мощность катушки тока из (a), (b)

$$UI = \frac{\omega}{2G \cos \gamma} \Phi^2,$$

или

$$\Phi = \sqrt{\frac{2G}{\omega}} \sqrt{UI}. \quad (c)$$

б) Полус с катушкой напряжения

$$U = \frac{\omega \Phi \omega}{\sqrt{2}}; \quad I_k = \frac{U}{\omega \omega^2 G} = \frac{\Phi}{\sqrt{2} \omega G},$$

или

$$\Phi = \sqrt{\frac{2G}{\omega}} \sqrt{UI}. \quad (d)$$

Пользуясь (c) и (d), можно найти значение $k_{\partial\delta}$.

а) Экранированная система с катушкой тока: проводимости полюсов G_1 и G_2 ; $\sin \psi = \text{const.} = k_1$.

$$k_{\partial\delta} = k_p \cdot k_1 \frac{2 \cos \gamma}{\omega} \sqrt{G_1 G_2};$$

$$M = k_{\partial\delta} UI.$$

Такое же значение $k_{\partial\delta}$ получается и для экранированной системы с катушкой напряжения.

б) Реле с двумя катушками питания, например реле мощности. Поток Φ_1 определяется током I сети, поток Φ_2 — напряжением U сети; угол ψ переменный.

$$k_{\partial\delta} = k_p \sqrt{\frac{2G_1 \cos \gamma_1}{\omega}} \sqrt{\frac{2G_2}{\omega}};$$

$$M = k_{\partial\delta} \sqrt{U_k I} \sqrt{UI_k} \sin \psi.$$

Из выражения для $k_{\partial\delta}$ видно, что сущность создания реле с высокой добротностью заключается в получении большего вращающего момента при том же потреблении катушек в вольт-ам-

перах. Высокие значения $k_{об}$ получают при правильном подборе всех соотношений магнитной системы.

В тангенциальных системах большую роль играют конфигурации верхнего и нижнего магнитопровода, расстояние между полюсами, расстояние полюсов от центра; в экранированных — геометрические размеры экрана; в барабанных — размеры полюсных делений, размеры барабана; во всех видах систем — соотношение между толщиной материала ротора и величиной зазора. Степень использования различных видов индукционных систем и исполнений этих видов характеризуется значением $k_{об}$. Тангенциальные системы имеют значение $k_{об}$ порядка $5 \frac{Гсм}{ва}$;

Чебоксарским аппаратным заводом были проведены исследования для определения коэффициента добротности индукционных систем барабанного типа. Измерения проводились в четырехполюсном реле разных видов. Полученные данные приведены в таблице.

Таблица 25
Значения коэффициентов добротности

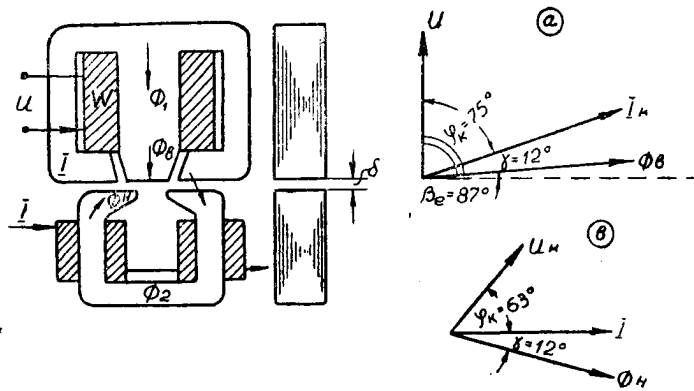
Зазор δ мм	Толщина ротора Δ мм	Материал магнитопровода	$k_{об} \frac{Гсм}{ва}$
2	0,5	Э4	7,9
1	0,5	Э4	12,4
1	0,4	ХВП	14,3

По расчетным данным индукционное восьмиполюсное реле имеет значение $k_{об} = 6,6 \frac{Гсм}{ва}$.

Таблица 26
Примерные значения параметров тангенциальной системы¹

Наименование параметра	Магнитопровод	
	верхний	нижний
Магнитная проводимость $аб/ав$	потока Φ_1	$G_1 = 5,1 \cdot 10^{-7}$
	потока Φ_2	$G_2 = 1,54 \cdot 10^{-7}$
Коэффициенты катушек	потоков $\Phi_в, \Phi_н$	$G_в = 1,5 \cdot 10^{-7}$
	омический, с учетом потерь в стали	$k_r = 43,5 \cdot 10^{-6}$
Углы	индуктивный	$k_x = 162 \cdot 10^{-6}$
	полного сопротивления	$k_z = 168 \cdot 10^{-6}$
Допустимое значение н. с.	полного сопротивления	$\varphi_k = 75^\circ;$
	потерь	$\gamma = 12^\circ;$
Допустимое значение н. с.	по тепловому режиму ($\theta = 60^\circ$)	$F = 610 ав;$
	по динамической устойчивости	$F = 500 ав;$

¹ См. фигуру 18, 3.



Фиг. 18, 3. Эскиз магнитопровода и векторные диаграммы.

Таблица 27

Примерные значения параметров барабанной системы¹

Наименование параметра	Значение параметра
Магнитные проводимости с учетом реакции ротора $\frac{\omega\delta}{av}$	для потоков ярма ($\Phi_{11}; \Phi_{12}$) $G_{11} = 3,14 \cdot 10^{-7}$ для потоков полюсов ($\Phi_1; \Phi_2$) $G_1 = 2,16 \cdot 10^{-7}$
Коэффициенты рассеяния	$\sigma_1 = \frac{\Phi_{11} + \Phi_{12}}{\Phi_1} = 1,45; \sigma_2 = \frac{\Phi_{21}}{\Phi_2} = 1,3$
Коэффициенты катушек	для двух последовательно включенных катушек ярма ($2\omega_1$) $k_z = 147 \cdot 10^{-8}$ для двух последовательно включенных катушек полюсов $k_z = 108 \cdot 10^{-8}$.
Угол полного сопротивления	$\varphi_k = 54 \div 55^\circ$
Термическая устойчивость	для $2\omega_1$ и $2\omega_2$ $F = 450 \div 500$ ав
Динамическая устойчивость	$F = 490$ ав

¹ См. фигуру 14,4.

ГЛАВА 19

ПРИМЕРЫ ИСПОЛНЕНИЯ НЕКОТОРЫХ ВИДОВ РЕЛЕ

19-1. ДИФФЕРЕНЦИАЛЬНО-ТОКОВЫЕ РЕЛЕ

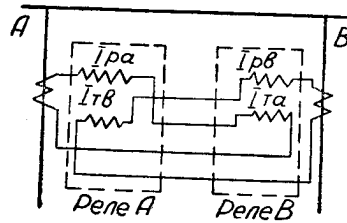
Дифференциальные токовые реле действуют в функции разности двух токов

$$M_T = k_T (I_1 - I_2).$$

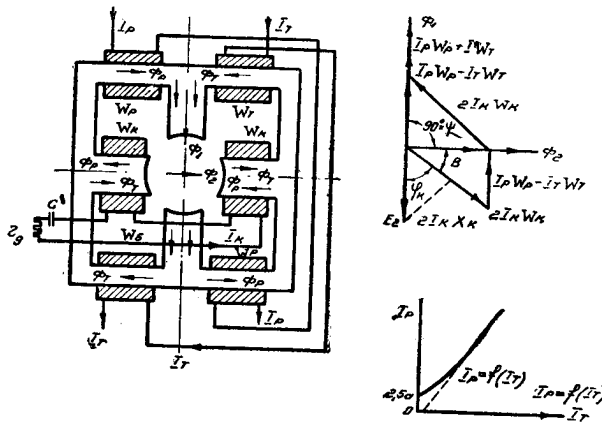
В СССР выпускают ряд таких реле для дифференциальной защиты линий и трансформаторов.

Дифференциальные реле имеют две группы катушек: первая питается током I_p , который называется рабочим током, так как величина его определяет собой рабочий момент на роторе, вторая питается током I_T , который называется тормозным, так как величина его определяет собой тормозной момент.

Для пояснения сказанного рассмотрим схему, изображенную на фигуре 19,1. От общих шин отходят две питающие линии. В каждой линии установлено дифференциальное реле, действие которого должно быть в функции разности токов $(I_p - I_T)$, где I_p — ток защищаемой линии, а I_T — второй.



Фиг. 19,1. Схема включения дифференциального токового реле.



Фиг. 19, 2. Магнитная система и векторная диаграмма.

Электромагнитная характеристика реле, называемая тормозной, оказана на фигуре 19, 2: при $I_T = 0$, ток I_{p0} должен быть определенным — порядка $2 \div 4 a$, при наличии тормозного тока действие реле огрубляется. Степень огрубления реле выражается в

процентах. Так, при 30-процентной характеристике отношение $I_p/I_T = 1,3$.

Ниже приводятся общие соотношения для расчета дифференциального токового реле для защиты линий электропередач. Схема включения катушек, векторная диаграмма и характеристика реле показаны на фигуре 19, 2.

Определить параметры реле при заданных величинах: $M_{но} = 0,35$ Гсм; ток срабатывания — $I_{по} = 2,5$ а при $I_T = 0$; $I_p/I_T = 1,3$.

1. Принятые упрощения: полагаем угол $I_p I_T = \alpha = 0$. Для получения максимального значения M_p принимаем $\psi = 90^\circ$; векторы Φ_1 и Φ_2 принимаем совпадающими с н. с. катушек, так как углы потерь γ_1 и γ_2 приближенно равны. Тогда угол $\Phi_1 \Phi_2$ равен углу $I_1 I_2$;

II. Принята 4-полюсная система. Известны параметры

$$k_p, \sigma, G, k_z, k_x, \gamma.$$

III. Общее уравнение момента

$$\Phi_1 = \sqrt{2}(I_p \omega_p + I_T \omega_T) G \cos \gamma; \quad (a)$$

$$\Phi_2 = \sqrt{2}(I_p \omega_p - I_T \omega_T + 2I_k \omega_k) G \cdot \cos \gamma. \quad (b)$$

Из фигуры 19,2

$$I_p \omega_p - I_T \omega_T + 2I_k \omega_k = \frac{I_p \omega_p - I_T \omega_T}{\operatorname{tg} \theta}. \quad (c)$$

$$\sin \phi = 1, \quad (d)$$

тогда

$$M_p = 2k_p G^2 \cos^2 \gamma (I_p \omega_p + I_T \omega_T) (I_p \omega_p - I_T \omega_T) \frac{1}{\operatorname{tg} \theta}.$$

Учитывая, что угол θ остается постоянным при всех таких катушках, имеем

$$M_p = k_a (I_p^2 \omega_p^2 - I_T^2 \omega_T^2) \text{ Гсм},$$

где

$$k_a = 2k_p G^2 \cos^2 \gamma \frac{1}{\operatorname{tg} \theta}.$$

IV. Установление угла θ .

Для увеличения M_p стремятся к уменьшению θ . Ограничивающие условия: 1) значение емкости c , определяющей собой $\varphi_k = 90 - \theta$; принимают $c = (4 \div 6) \mu F$ с таким расчетом, чтобы конденсатор можно было встроить в корпус реле; 2) электрическое сопротивление катушек контура θ принимают $45 \div 55$.

V. Число витков катушек контура — ω_k . r_1 — сопротивление одной катушки; r_∂ — добавочное, для настройки контура; $r_\partial \cong 0,5r_1$.

Тогда

$$\begin{aligned} r_k &= 2r_1 + r_0 = 2,5k, \omega_k^2; \\ L_1 &= \omega_k^2 G_{yc}; \quad x_k = \frac{1}{\omega c} - 4\omega \omega_k^2 G_{yc}. \\ Z_k &= \sqrt{r_k^2 + x_k^2}; \quad I_k = \frac{E_k}{Z_k} = \frac{\omega 2\omega_k \Phi_2}{\sqrt{2} Z_k}. \end{aligned} \quad (a)$$

$$\Phi_2 = \sqrt{2} I_k 2\omega_k G \cos \theta. \quad (b)$$

Подставляя (b) в (a), решая относительно Z_k и сравнивая с уравнением $Z_k = \sqrt{r_k^2 + x_k^2}$, имеем

$$2\omega G \cdot \cos \theta \omega_k^2 = \sqrt{(2,5k, \omega_k^2)^2 + \left(\frac{1}{\omega c} + 4\omega \omega_k^2 G_{yc}\right)^2}. \quad (c)$$

Из (c) определяем ω_k . Число витков ω_k бывает в пределах 1400—2000.

VI. Числа витков ω_p , ω_T . Учитывая условия $I_{p0} = 2,5 a$ при $I_T = 0$, имеем

$$\omega_p = \sqrt{\frac{M_{p0}}{k_a I_{p0}^2}}.$$

Из условия

$$I_p \omega_p = I_T \omega_T$$

имеем

$$\omega_T = 1,3\omega_p.$$

19-2. РЕЛЕ МОЩНОСТИ

В устройствах релейной защиты реле мощности применяют в качестве элементов, фиксирующих направление мощности. Поэтому в данном случае интересуются не точным замером значения мощности в линии, а только направлением энергии.

Необходимо обеспечить надежное срабатывание реле в случае короткого замыкания, когда напряжение сети резко снижается. Стремятся подобрать параметры реле такими, чтобы при коротком замыкании, когда угол φ сети изменяется, значение угла ψ было близко к 90° . Как правило, реле мощности включаются на линейные напряжения, например U_{ab} , U_{cb} , и линейные токи — I_a , I_b . Однако заводская характеристика реле относится к случаю включения катушек реле на напряжение и ток одной фазы, например U_a , I_a . Такую характеристику назовем основной.

Вывод основной характеристики можно получить из векторной диаграммы реле (фиг. 19, 3):

$$\begin{aligned} \Phi_1 &= k_1 U; \quad \Phi_2 = k_2 I; \quad \sin \psi = \sin [90 - (\varphi + \gamma + \theta)] = \cos (\varphi + \alpha); \\ M_m &= k_m U I \cos (\varphi + \alpha). \end{aligned}$$

Угол α называется дополнительным углом реле и выбирается зависимости от схемы включения реле. Из той же диаграммы

видно, что угол α зависит от угла катушки φ_k . Таким образом, для изменения α необходимо менять значение φ_k введением в цепь катушки емкостей или электрических сопротивлений.

Выбор угла α производится при построении векторной диаграммы действительного включения реле. Так, если включение производится по 90-градусной схеме (U_{bc}, I_a), то зная значение угла φ сети при коротком замыкании, можно найти α из условия получения в этом случае значения угла $\psi = 90^\circ$.

Из этого условия (фиг. 19, 4) следует:

$$\begin{aligned} \varphi_k &= 90 - \gamma - \delta; & (a) \\ \delta &= 90 - \gamma - \varphi. & (b) \end{aligned}$$

Из (a) и (b) получаем $\varphi_k = \varphi$. Тогда

$$\alpha = 90 - \varphi_k = 90 - \varphi = 90 - 75 = 15^\circ.$$

Для данного случая, т. е. включения реле по 90-градусной схеме и угла короткого замыкания сети $\varphi = 75^\circ$, нужно применять реле мощности с основной характеристикой

$$M_M = k_M UI \cos(\varphi + 15).$$

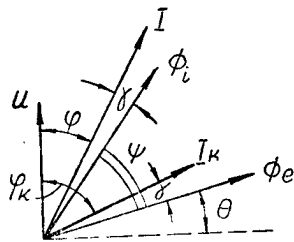
Исполнения реле направления мощности

Существуют реле с тангенциальной системой и с барабанной системой. Первые применяются в тех случаях, когда не предъявляют особых требований к быстрому действию реле.

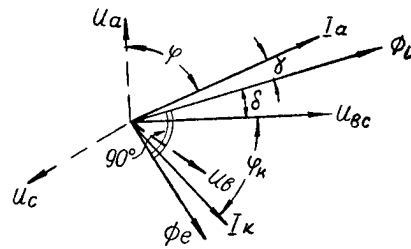
Ниже приводится примерный краткий расчет направления мощности. Схему и векторную диаграмму см. на фигуре 19, 5. Задано: $U = 100$ в, $I = 5$ а, $f = 50$ гц; $\varphi = 60^\circ$, включение по 90-градусной схеме; начальный момент возвратной пружины $M_{no} = 0,5$ Гсм. Требуется определить данные катушек и минимальную мощность срабатывания.

1. Магнитная система. Принимаем для проектирования тангенциальную систему, представленную на фигуре 14, 1, параметры согласно таблице 26.

2. Вычисление угла φ_k . Строится векторная диаграмма реле, напряжение линейное U_{bc} , ток фазы I_a . Значение φ_k подбирается таким, чтобы при коротком замыкании ($\varphi = 60^\circ$) значение $\psi = 90^\circ$.



Фиг. 19, 3. Векторная диаграмма.



Фиг. 19, 4. Векторная диаграмма.

Из диаграммы $\varphi_k = 60^\circ$. Нормально φ_k реле равен 75° , для уменьшения его значения вводится дополнительно добавочное сопротивление r_∂ .

3. Определение ω_s ; задаются значением потребляемой мощности катушки, обычно в пределах $20 \text{ вa} - 30 \text{ вa}$.

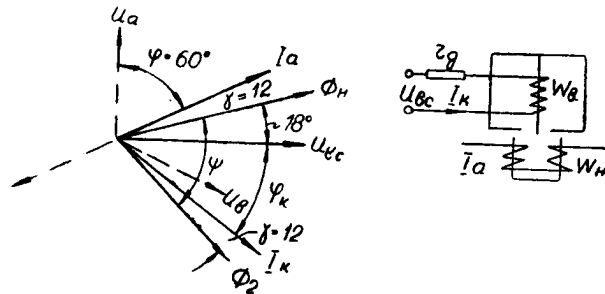
Примем $P_s = 30 \text{ вa}$. Расчет проводится вначале без учета r_∂ . Тогда

$$Z = \frac{U^2}{P_s} = \frac{100^2}{30} = 334 \text{ ом.}$$

Из соотношения

$$k_2 \omega_s^2 = Z; \omega_s = \sqrt{\frac{334}{168 \cdot 10^{-8}}} = 1410$$

принимая $\omega_s = 1400$.



Фиг. 19, 5. Схема включения и векторная диаграмма реле.

4. Определение ω_n . Значение термической устойчивости принимается равным 50. По данным, приведенным в главе 2, допускается плотность тока $j = 155 \frac{\text{а}}{\text{мм}^2}$; тогда

$$q = \frac{250}{155} = 1,61 \text{ мм}^2; d = 1,43 \text{ мм.}$$

Принимаем $d = 1,45 \text{ мм}$. По габаритам нижнего магнитопровода в нем можно поместить две катушки по 17 витков: $\omega_n = 34$.

5. Вычисление r_∂ .

$$\frac{x}{r_s + r_\partial} = \text{tg } \varphi_k = \text{tg } 60^\circ = 1,63; x = 162 \cdot 10^{-6} \omega_s^2 = 316 \text{ ом;}$$

$$r_s = 43,6 \cdot 10^{-6} \omega_s^2 = 87 \text{ ом; } r_\partial = 107 \text{ ом.}$$

6. Полное сопротивление.

$$Z = \sqrt{x^2 + r^2} = 370 \text{ ом.}$$

7. Проверка ω_s по тепловому режиму. F при $U = 100 \text{ в}$:

$$F = \frac{U}{Z} \omega_s = \frac{100}{370} \cdot 1400 = 380,$$

допустимое значение $F = 610$.

8. Значение k_m .

$$\sin \psi = 1; k_p = 7 \cdot 10^{-6}; \gamma = 12^\circ; G_s = 15; G_n = 8.$$

Тогда $k_m = 0,2$.

$$M_m = k_m UI = 0,2 UI \text{ Гсм.}$$

9. Минимальная мощность срабатывания реле

$P_{мин}$.
Учитывая $M_{по} = 0,5 \text{ Гсм,}$

$$P_{мин} = \frac{0,5}{0,2} = 2,5 \text{ вa.}$$

19.3. РЕЛЕ СОПРОТИВЛЕНИЯ

Реле сопротивления применяются для выполнения дистанционных защит, действие которых определяется дистанцией — расстоянием от места установки защиты до места повреждения. Линия электропередачи характеризуется сопротивлением одного километра линии $Z \text{ км.}$

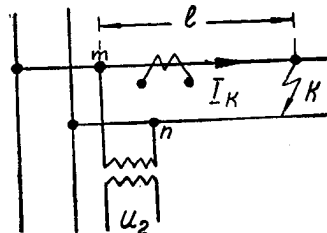
При коротком замыкании на участке линии (K) (фиг. 19, 6) на зажимах mn получается напряжение

$$U_{mn} = 2I_K Z_{км.}$$

Полное сопротивление петли $ткп$

$$Z_1 = \frac{U_{mn}}{I_K} = 2IZ_{км.}$$

При расчете реле принимают во внимание вторичные значения сопротивлений, т. е. величины сопротивлений с учетом напряжений и токов во вторичных цепях трансформаторов тока и напряжения:



Фиг. 19, 6. Схема участка линии электропередачи.

$$U_2 = \frac{U_{mn}}{k_1}; I_2 = \frac{I_K}{k_2}; z_2 = Z_1 \frac{k_2}{k_1}.$$

Ниже рассматриваются только вторичные значения напряжений, токов и сопротивлений.

В практике, как правило, применяются реле сопротивления минимального действия, т. е. срабатывающие в том случае, если сопротивление петли короткого замыкания $ткп$ меньше установленного в реле.

При построении реле сопротивления учитываются два режима работы линии электропередачи: 1) нормальный, характеризующийся номинальными значениями электрических величин; 2) аварийный.

Для нормального режима принимают

$$U = 100 \text{ в; } I = 5 \text{ a; } z = \frac{100}{5} = 20 \text{ ом.}$$

В аварийном режиме линии резко снижается ее сопротивление, напряжение на зажимах mn резко падает, а ток короткого замыкания увеличивается.

Одно и то же значение z может иметь место при различных вторичных значениях напряжений и токов.

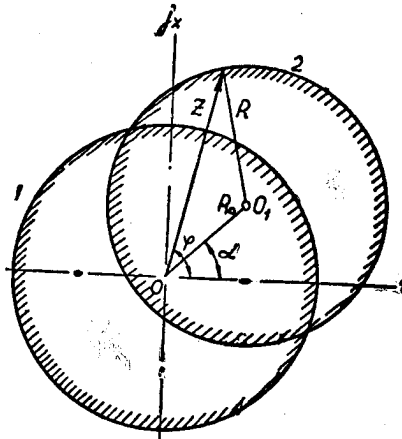
Например, $z = \frac{50}{20} = \frac{20}{8} = 2,5 \text{ ома}$.

Отсюда возникает требование к реле: его срабатывание должно определяться не величинами U и I , а их отношением.

Отметим, что в длинных, сильно загруженных линиях электропередачи значения z при нормальном и аварийном состояниях линии мало отличаются; в данном случае для срабатывания реле пользуются тем, что углы полного сопротивления при этих режимах резко отличны.

Характеристики реле сопротивления строят обычно в координатах r — омического сопротивления и x — индуктивного линии.

Эти характеристики имеют вид окружностей (фиг. 19, 7) с различным расположением центров: 1) реле полного сопротивления с центром в начале координат; штриховкой показана граница срабатывания реле — оно срабатывает при значениях Z меньших радиуса окружности; 2) направленное реле сопротивления с центром вне начала координат.



Фиг. 19, 7. Характеристика реле сопротивления.

Принципы построения реле

Общее уравнение окружности в полярных координатах $r-x$ (фиг. 19, 7).

$$z^2 + R_0^2 - 2zR_0 \cos(\varphi - \alpha) = R^2. \quad (19-1)$$

Здесь R_0 — полярный радиус-вектор центра окружности, α — угол между R_0 и осью абсцисс.

Частные случаи:

- 1) $R = R_0$ — окружность проходит через начало координат;
- 2) $R_0 = 0$ — центр окружности в начале координат.

Уравнение реле сопротивления должно соответствовать (19-1).

Для решения задачи выполняют реле так, чтобы на общую ось действовали моменты трех магнитных систем с характеристиками

$$M_m = k_m UI \cos(\varphi - \alpha); \quad M_T = k_T I^2; \quad M_N = k_N U^2.$$

Знаки моментов должны быть такими, чтобы моменту M_n противодействовали M_T и M_m .

В нормальном режиме

$$U = U_n; M_n > M_m + M_T;$$

в аварийном

$$U \ll U_n; M_n < M_m + M_T.$$

Расчет реле проводится по критическому положению

$$M_n = M_m + M_T;$$

$$k_n U^2 - k_m U I \cos(\varphi - \alpha) = k_T I^2.$$

Разделив все члены уравнения на $k_n I^2$, считая $\frac{U}{I} = Z$, и добавив в каждой части $\left| \frac{k_m}{2k_n} \right|^2$, получаем

$$z^2 + \left| \frac{k_m}{2k_n} \right|^2 - \frac{k_m}{k_n} z \cos(\varphi - \alpha) = \left| \sqrt{\frac{k_m^2 + 4k_T k_n}{4k_n^2}} \right|^2. \quad (19-2)$$

Полученное выражение соответствует уравнению окружности в полярных координатах. Радиус окружности

$$R = \sqrt{\frac{k_m^2 + 4k_T k_n}{4k_n^2}}.$$

Положение центра характеризуется полярным радиусом-вектором

$$R_0 = \frac{k_m}{2k_n}.$$

Выражение (19-2) является исходным для построения разных видов дистанционных реле. При наличии всех трех элементов — тока, напряжения и мощности, получается характеристика, соответствующая окружности 2 (фиг. 19, 7).

В частном случае, при отсутствии элемента мощности, т. е. когда $k_m = 0$, получают реле полного сопротивления (окружность 1).

$$R = \sqrt{\frac{k_T}{k_n}} = z; R_0 = \frac{k_m}{2k_n} = 0.$$

Ранее выполнялись многосистемные реле, т. е. такие, в которых на общую ось действовали механические моменты трёх систем. Более рациональной является конструкция односистемного реле, в котором сравнение моментов M_n , M_m и M_T осуществляется соответствующим распределением магнитных полей в индукционной системе.

Советские инженеры и ученые провели ряд работ по усовершенствованию реле сопротивления и схемы дистанционных защит.

Индукционная система имеет две группы катушек:

1) Четыре катушки ярма; ток в цепи этих катушек определяется напряжением U и э. д. с. E_1 ; емкость c подбирается так, чтобы поток в системе опережал по фазе приложенные напряжения.

2) Две катушки, расположенные на полюсах 2 и 4, включаются на напряжение U и э. д. с. E_2 .

Для пояснения работы реле рассмотрим отдельно токи в контурах и потоки в системе, определяемые напряжением U и э. д. с. E (см. векторную диаграмму фиг. 19,9).

Потоки полюсов 1, 3: $\Phi_{a1} = c_1 U$; $\Phi_{a2} = c_2 I$.

Потоки полюсов 2, 4: $\Phi_{n1} = c_3 U$; $\Phi_{n2} = c_4 I$.

Результирующий момент на роторе имеет следующие составляющие:

$$M_m = k_p [\Phi_{a1} \Phi_{n2} \sin \psi_1 + \Phi_{a2} \Phi_{n1} \sin \psi_2] = k_m U I \cos (\varphi - \alpha);$$

$$M_n = k_p \Phi_{a1} \Phi_{n1} \sin \psi_3 = k_n U^2;$$

$$M_T = k_p \Phi_{a2} \Phi_{n2} \sin \psi_4 = k_T I^2.$$

ЧАСТЬ ЧЕТВЕРТАЯ

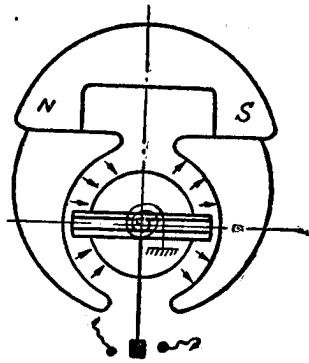
РАЗНЫЕ РЕЛЕ, КОНСТРУКТИВНЫЕ ЭЛЕМЕНТЫ

ГЛАВА 20

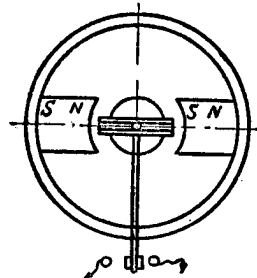
МАГНИТОЭЛЕКТРИЧЕСКИЕ И ЭЛЕКТРОДИНАМИЧЕСКИЕ РЕЛЕ

20-1. МАГНИТОЭЛЕКТРИЧЕСКИЕ РЕЛЕ

В магнитоэлектрических реле вращение подвижной части, рамки, определяется взаимодействием поля постоянного магнита с током рамки. Такие реле применяются только для цепей постоянного тока. Включение их в цепь переменного тока возможно через выпрямители.



Фиг. 20, 1. Система магнито-электрического реле.



Фиг. 20, 2. Система магнито-электрического реле.

Особенностью магнитоэлектрических реле является их высокая чувствительность. Имеются исполнения таких реле с мощностью срабатывания $P = i^2 r$ порядка $1 \cdot 10^{-10} \div 1 \cdot 10^{-11}$ ватт. Время срабатывания их относительно велико — порядка $0,4 \div 1$ секунды, давление в контактах незначительное, порядка $0,5 \text{ Г}$. Поэтому область применения магнитоэлектрических реле весьма ограничена, они применяются в схемах, где требуется элемент весьма высокой чувствительности.

Формы магнитных систем. Различают две основные формы: с внешним магнитом и с внутримрамочным магнитом. Вторая форма применяется в тех случаях, когда требуется выполнение конструкции весьма малых габаритов. Такие условия предъявляются к конструкциям ряда типов реле.

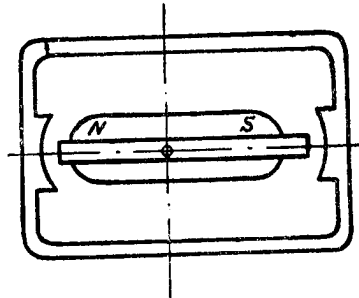
На фигуре 20,1 приведена форма системы реле обычного исполнения, заимствованная из измерительной техники.

На фиг. 20,2 показано исполнение системы, дающее лучшее решение в отношении компактности устройства.

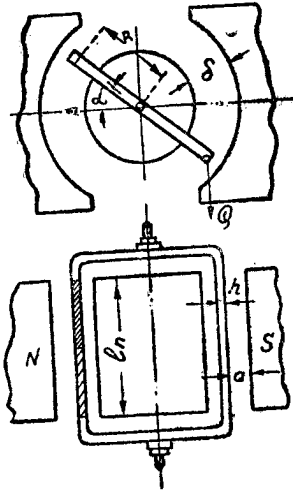
На фиг. 20,3 приведено исполнение реле с внутрирамочным магнитом.

Представляет интерес применяемая в настоящее время в измерительной технике конструкция системы с подвижным магнитом весьма малого веса (порядка 1 Г). Катушка системы неподвижна и может иметь увеличенные габариты. Добавочный постоянный магнит, расположенный под катушкой, заменяет спиральные моментные пружинки.

Основные соотношения для расчета. Рабочий угол рамки α принимается небольшим, порядка



Фиг. 20, 3. Система с внутрирамочным магнитом.



Фиг. 20, 4. Подвижная система реле.

5—10° от нейтрального положения. Полусы магнита выполняются так, чтобы поле магнита было равномерным в пределах рабочих углов поворота рамки. Сила на рамке

$$Q = 1,02 \cdot 10^4 B i l_n \omega \Gamma, \quad (20-1)$$

где B — $\frac{60}{\text{см}^2}$ — индукция в рабочем зазоре;

i — ток в рамке, ω — число витков рамки,

l_n — рабочая длина проволоки рамки в см (фиг. 20, 4).

Момент

$$M = 2 \cdot 1,02 \cdot 10^4 B i l_n \omega R, \quad (20-2)$$

или

$$M = k i, \quad (20-3)$$

т. е. момент пропорционален току в рамке.

При питании катушки выпрямленным током расчет проводится по среднему значению тока. Система реле снабжается обычно двумя спиральными моментными пружинами, действующими друг против друга и устанавливающими рамку в нейтральное положение.

$$M = 2j\alpha,$$

где J — жесткость пружин в $\frac{\Gamma \text{ см}}{\text{рад}}$; α — рабочий угол поворота рамки в радианах.

Рабочие моменты в магнитоэлектрических реле относительно невелики. Поэтому при вычислении значения тока срабатывания реле учитывается момент трения $M_{тр}$ (см. главу 24).

Ток срабатывания

$$i = \frac{2J\alpha + M_{тр}}{K}. \quad (20-4)$$

Время срабатывания реле

Полное время, t , срабатывания реле определяется в основном временем движения рамки t_{θ} . Временем t_n пренебрегают, учитывая малое значение постоянной времени рамки. При составлении дифференциального уравнения движения необходимо учитывать дополнительные демпфирующие моменты, возникающие на рамке в процессе ее движения.

1) При движении рамки в ее обмотке индуцируется э. д. с.

$$e_{\theta} = -\frac{d\Phi_p}{dt} \omega; \quad i_{\theta} = \frac{e_{\theta}}{r},$$

где Φ_p — поток, пронизывающий рамку, r — электрическое сопротивление рамки. Согласно фигуре 20,4

$$\Phi_p = 2BRl_p \sin \alpha \text{ вебер.}$$

Здесь R и l_p в сантиметрах.

Учитывая, что при малых углах $\sin \alpha \approx \alpha$, получаем величину демпфирующего тока рамки:

$$i_{\theta} = -\frac{2BRl_p \omega \alpha'}{r} = k_1 \alpha'. \quad (20-5)$$

Демпфирующий момент, согласно выражению (20-3):

$$M_{\theta} = k k_1 \alpha'. \quad (20-6)$$

2) В алюминиевом каркасе рамки индуцируются э. д. с.: первая, определяемая изменением тока в обмотке рамки

$$e_{\kappa 1} = -M \frac{di}{dt};$$

вторая, определяемая движением каркаса в магнитном поле

$$e_{\kappa 2} = -\frac{d\Phi_p}{dt}.$$

Значение $e_{\kappa 1}$ относительно невелико и им пренебрегают. Демпфирующий момент каркаса определяется в основном э. д. с. $e_{\kappa 2}$.

$$i_{\kappa} = \frac{e_{\kappa 2}}{r_{\kappa}} = \frac{-2BRl_p \alpha'}{r_{\kappa}} = k_2 \alpha', \quad (20-7)$$

$$M_{\theta \kappa} = \frac{k}{\omega} k_2 \alpha'. \quad (20-8)$$

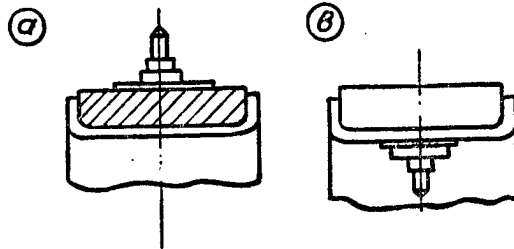
Дифференциальное уравнение движения рамки

$$J\alpha'' + \left(kk_1 + \frac{k}{\omega} k_2\right)\alpha' + 2Ja = ki - M_{тр}. \quad (20-9)$$

Конструктивные вопросы

1) *Постоянный магнит* выполняется из высококоэрцитивного сплава — магнито, альнико. Индукция B в рабочем зазоре принимается порядка $(0,2 - 0,3) \cdot 10^{-4} \frac{вб}{см^2}$.

2) *Рамки*. Применяется медная проволока с эмалевой изоляцией, диаметр провода — $0,02 - 0,03$ мм. Каркас рамки — алюминиевый толщиной $0,2 - 0,3$ мм.



Фиг. 20, 5. Крепление ядра:
а — снаружи рамки; б — внутри рамки.

При выборе толщины намотки h (фиг. 20, 4) необходимо учитывать следующее: при увеличении h можно повысить число витков рамки. При этом, однако, увеличивается вес рамки, а следовательно, момент трения, увеличивается

зазор δ , что в свою очередь приводит к уменьшению индукции B . Необходимо проводить несколько вариантов расчета с целью получения оптимального решения. Вес рамки обычно бывает порядка 1 г, размер a принимается из конструктивных соображений так, чтобы не имело место затирание рамки.

3) *Опоры* выполняются в виде ядер. Последние крепятся в буксе, которая приклеивается к рамке. Обычно ядра крепятся снаружи рамки (фиг. 20, 5 а), но в некоторых случаях для уменьшения размеров конструкции применяют крепление ядра внутри рамки (фиг. 20, 5 б).

4) *Контакты* реле выполняются из платины или платины-иридия.

20-2. ЭЛЕКТРОДИНАМИЧЕСКИЕ РЕЛЕ

В электродинамических реле момент на подвижной рамке определяется взаимодействием поля неподвижного электромагнита с током в рамке.

Электродинамические системы применялись ранее для построения реле направления мощности, дистанционных и других. Однако ввиду недостаточной надежности подвижной рамки и относительно большого времени срабатывания эти реле исполняются сейчас с использованием других принципов действия.

В настоящее время электродинамические системы имеют весьма ограниченное применение: известны только исполнения систем для построения автоматических регуляторов постоянного тока.

Форма магнитных систем

В устаревших конструкциях электродинамических реле применялась форма, заимствованная из измерительной техники — с вращающейся рамкой. Более совершенными являются конструкции с поступательным движением рамки (фиг. 20, 6). На фигуре 20, 7 приведено исполнение системы для автоматических регуляторов мощности.

Рамка перемещается в поле Φ катушки, расположенной на неподвижном сердечнике. В зазоре между кольцом неподвижного магнитопровода и сердечником имеет место равномерное поле с индукцией B , причем

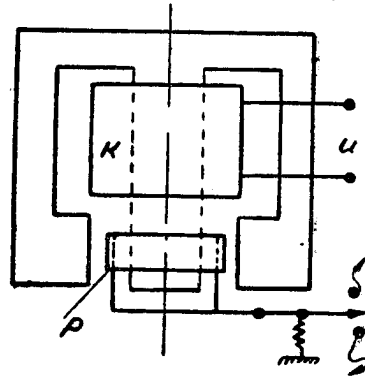
$$B = cU. \quad (20-10)$$

Рабочий момент на рамке

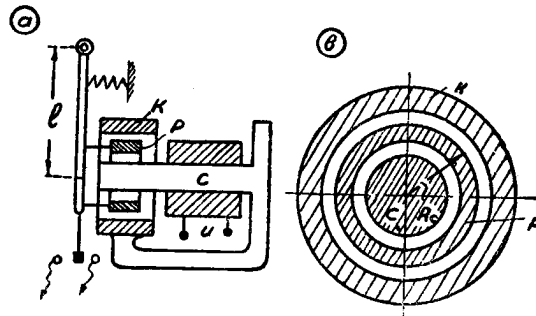
$$M = 1,02 \cdot 10^4 Bi \, 2\pi R_{\text{ср}} \, \text{Гсм} \quad (20-11)$$

При равенстве этого момента моменту пружины M_n рычаг с рамкой устанавливается в среднем положении между контактами. Заменяя в (20-11) значение B , из (20-10) получаем

$$M = \kappa U i. \quad (20-12)$$



Фиг. 20, 6. Электродинамическая система.



Фиг. 20, 7. Система регулятора мощности:
а — эскиз; б — разрез по кольцу.

Таким образом, момент на рамке пропорционален мощности регулируемой системы.

Г Л А В А 21

ТЕРМИЧЕСКИЕ РЕЛЕ

21-1. КЛАССИФИКАЦИЯ

Термическими называются реле, срабатывающие при определенной температуре нагрева измерительного элемента реле. Нагрев этого элемента может быть осуществлен двояко: 1) теплом, развиваемым протекающим током (в самом элементе или в окружающем его нагревательном устройстве), 2) теплом окружающей среды. Различают поэтому две группы термических реле.

1) **Тепловые**, работающие в функции тока: токовые реле для защиты при перегрузках и токах короткого замыкания, пусковые реле в функции тока, реле времени.

Различают непосредственно тепловые реле и терморасцепители. Тепловые реле снабжаются контактами небольшой разрывной мощности. Терморасцепители являются частью воздушных автоматических выключателей (установочных автоматов); во включенном положении эти выключатели удерживаются рычажной системой, запертой терморасцепителем; при соответствующей деформации биметалла рычажная система освобождается от зацепления и выключатель отключается.

2) **Термореле** или **термостаты**, работающие в функции температуры окружающей среды. Термореле исполняются как реле максимального действия, срабатывающие при определенной максимальной температуре, или как регуляторы, контролирующие температуру определенного устройства — электрической печи, обмотки электродвигателя, и других. Каждый термостат снабжается регулировочным устройством для настройки на определенную температуру срабатывания.

В качестве основного чувствительного к температуре элемента применяют: 1) легкоплавкие сплавы; 2) тела с изменяющейся магнитной проницаемостью; 3) испаряющиеся или расширяющиеся жидкости; 4) тела с изменяющимся электрическим сопротивлением; 5) твердые тела с различными коэффициентами линейного расширения; 6) термобиметаллы. Наибольшее распространение получили термореле с применением элементов в виде твердых тел с различными коэффициентами линейного расширения и с применением биметаллов. Ниже рассматриваются вопросы проектирования и расчета таких реле.

21-2. ТЕРМОБИМЕТАЛЛ И ЕГО ПАРАМЕТРЫ

Биметаллами называют ленты или полосы, состоящие из отдельных сваренных или спаянных вместе полос металлов. В машиностроении применяют два вида биметалла:

1) **антикоррозийный**, состоящий из стальных полос с наложенными по обе стороны полосами из антикоррозийного металла: меди, латуни и других;

2) *термобиметалл*, состоящий из двух металлов-компонентов с разными коэффициентами линейного расширения.

В термореле применяется только термобиметалл, который в дальнейшем для сокращения называем биметаллом. Подробные данные о составе компонентов и технологии изготовления биметалла изложены в литературе. Ниже приводятся только основные, необходимые для расчета данные.

Марки биметалла. Наша промышленность изготавливает биметаллические полосы и ленты разных сортов, марок, отличающихся металлом или сплавом активного и пассивного компонента. Некоторые технические данные разных марок приведены в ГОСТ 5198—50 (см. табл.28) и в технических условиях ЦМТУ (см. табл. 29). Указанные в ГОСТ данные о коэффициенте чувствительности M учитывают разность коэффициентов линейного расширения компонентов $M = \alpha_1 - \alpha_2$.

Таблица 28
Данные термобиметалла (из ГОСТ 5198—50)

Марка	Коэффициент чувствительности M не менее $^{\circ}\text{C}^{-1} \cdot 10^6$	Удельное сопротивление ρ , ом $\frac{\text{мм}^2}{\text{м}}$	Температурный интервал постоянства коэффициента чувствительности, $^{\circ}\text{C}$	Температура нагрева биметалла соответствен. верхнему пределу упругих деформаций, $^{\circ}\text{C}$	
				при нагреве без нагрузки	при нагреве с нагрузкой
ТБ1	21	0,58—0,68	—60+(190÷210)	530÷550	280÷310
ТБ2	18	0,56—0,64	—50+(320÷340)	430÷450	240÷260
ТБ3	20	0,75—0,83	—50+(190÷210)	450÷470	210÷230
ТБ4	12	0,14—0,20	—60+(190÷210)	410÷430	280÷300
ТБ5	19	0,08—0,14	—50+(180÷210)	180÷200	140÷160
ТБ6	19,5	0,14—0,20	—50+(190÷210)	230÷250	120÷140
ТБ7	18	0,14—0,18	—50+(210÷230)	290÷310	160÷170

Таблица 29
Размеры термобиметаллических полос и лент
(по техническим условиям ЦМТУ 1900—44)

	Размеры в мм					
	Толщина	Допуск отклонений по толщине	Ширина	Допуск отклонений по ширине	Длина	Допуск отклонений по длине
Полосы	0,6—1,0	$\pm 0,04$	10÷120	$\pm 0,5$	300—1000	± 10
Ленты	1,0—1,5	$\pm 0,05$	10÷120	$\pm 0,5$	500—2000	± 10
	0,2—0,5	0,03				

Выбор соответствующей марки определяется предельным значением рабочей температуры, требуемой чувствительностью, влияющей на величину деформации, электрическим сопротивлением и устойчивостью против коррозии.

Параметры. Помимо указанных в таблице 28, необходимо иметь данные о некоторых других параметрах.

1) *Модуль упругости E* колеблется в пределах $(12,5 \div 16) \times 10^6 \frac{\Gamma}{\text{мм}^2}$. В среднем, для наиболее ходких биметаллов принимают $E = 15 \cdot 10^6 \frac{\Gamma}{\text{мм}^2}$.

2) *Температурный коэффициент по сопротивлению:* от 0,03 до 0,7% на 1°.

3) *Внутренние напряжения.* Реле должно быть исполнено так, чтобы внутренние напряжения, возникающие в биметалле, не превосходили пределов упругости компонентов; при этом должен быть введен коэффициент безопасности. Значение Q — равномерной аксиальной силы в каждом компоненте: сжатия в первом, растяжения во втором, определяется по упрощенной формуле

$$Q = \frac{E(\alpha_1 - \alpha_2)}{h} \vartheta, \quad (21-1)$$

где h — толщина биметалла, в мм.

4) *Термическая устойчивость.* В тепловых реле с непосредственным нагревом биметалла могут возникнуть при коротком замыкании в защищаемой линии значительные температуры, приводящие к расплавлению биметалла или частичному его выгоранию в местах наибольшей плотности тока. Количество выделяемого в биметалле тепла зависит от величины тока к. з. i_k , длительности t_k и сопротивления r биметалла.

Среднее значение $(i_k)_{cp}$ определяется из формулы

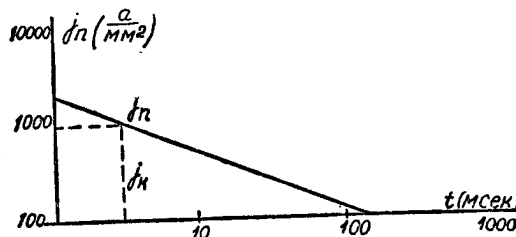
$$(i_k)_{cp} = \sqrt{\frac{1}{t_k} \int_0^{t_k} i_k^2 dt}. \quad (21-2)$$

Для расчетов пользуются данными допустимой плотности тока для каждой марки биметалла

$$j_k = \frac{(i_k)_{cp}}{b_1 h} \frac{a}{\text{мм}^2} \quad (21-3)$$

При этом j_k должно быть меньше j_n плотность тока, при которой начинается плавление металла:

$$\frac{j_n}{j_k} = 1,5 \div 2.$$



Фиг. 21, 1. Зависимость $j_n = f(t)$.

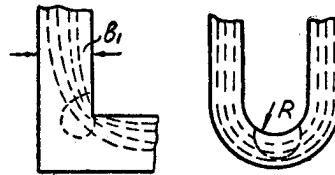
На фигуре 21,1 приведена примерная зависимость $j_n = f(t_k)$, которую следует строить для разных марок.

Необходимо учитывать местные увеличения плотности тока, определяемые неравномерным распределением линий тока в изогну-

той пластинке (фиг. 21, 2). Уменьшение величины местных перегревов достигается увеличением радиуса R . Ниже приводятся опытные данные, показывающие влияние значения R на кратность n увеличения температуры в месте перехода

R	1,1	2	5
n	2	1,5	1,2

Таким образом, при нагреве пластинки током короткого замыкания до $\vartheta_{к}$, например до 800° , температура в месте перехода получается при $R = 2 \text{ мм}$, $\vartheta_{к1} = 1200^\circ$.

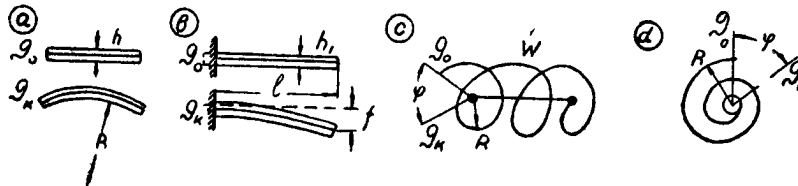


Фиг. 21, 2. Распределение линий токов в изогнутом проводнике.

21-3. ДЕФОРМАЦИИ БИМЕТАЛЛА

Вывод основных уравнений для определения деформации термобиметалла дан профессором Тимошенко. Ниже приводятся основные соотношения для расчета.

Свободная пластинка. При нагреве прямой пластинки (с температурой ϑ_0) до температуры $\vartheta_{к}$ пластинка изгибается (фиг. 21, 3 а).



Фиг. 21, 3. Исполнения биметаллических элементов.

Если учитывать определенное приведенное значение модуля упругости E , то кривизна пластинки

$$c = \frac{1}{R} = \frac{6M \cdot h_1 h_2}{(h_1 + h_2)^3} \vartheta; \quad (21-4)$$

$$\vartheta = \vartheta_{к} - \vartheta_0, \quad (21-5)$$

где M — коэффициент чувствительности биметалла; h_1 и h_2 — толщина компонентов. Обычно выполняют биметалл: $h_1 = h_2 = 0,5h$; тогда

$$c = \frac{1}{R} = 1,5M \frac{\vartheta}{h}. \quad (21-6)$$

Пластинка изгибается не только по длине, но и по ширине. При $\gg b$ изгиб по ширине не учитывают.

Свободная шайба. При нагреве деформируется вся шайба с кривизной c (см. выражение 21-6).

Плоская пластинка закреплена в одном конце. Деформация конца пластинки (фиг. 21, 3 б).

$$f = 0,75 M \frac{l^2}{h} \vartheta \quad (21-7)$$

Геликоидальная пружина (винтовая цилиндрическая, фиг. 21, 3 с)

$$\varphi = 1,5 M \frac{2\pi R w}{h} \vartheta \text{ радиан,} \quad (21-8)$$

где w — число витков пружины.

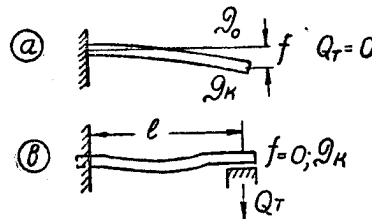
Спираль (фиг. 21, 3 d)

$$\varphi = 1,5 \cdot M \frac{l}{h} \vartheta \text{ радиан,} \quad (21-9)$$

где l — длина спирали.

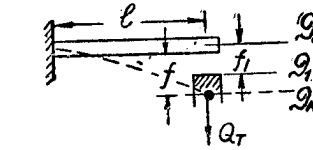
21-4. МЕХАНИЧЕСКИЕ СИЛЫ, РАЗВИВАЕМЫЕ БИМЕТАЛЛОМ

Свободный конец биметаллической пружины — плоской или спиральной, деформируется при нагреве вследствие возникающих в пружине внутренних напряжений. Величина деформации изменяется в том случае, если конец пружины не может перемещаться свободно, а должен преодолевать действие внешних механических



Фиг. 21, 4. Деформации биметаллической пластинки:

a — свободной; b — подпертой.



Фиг. 21, 5. К определению сил на пластинке:

a — изображение в плане сплошной пластинки; b — с вырезом.

сил. Для определения сил, развиваемых биметаллом, система рассматривается так, как в случае приложения сил, изгибающих пружину на величину деформации (f, φ), соответствующей температуре $\vartheta = \vartheta_k - \vartheta_0$. Рассмотрим отдельные виды пружин.

Плоская пружина, закрепленная в одном конце (фиг. 21,4). Свободная деформация

$$f = 0,75 \cdot M \frac{l^2}{h} \vartheta \quad (21-7)$$

сила в конце $Q_T = 0$.

Если пружина подперта, то она давит на опору силой Q_T . С некоторым приближением значение Q_T определяют из условия изгиба консольной балки на величину f :

$$Q_T = \frac{3EJ}{l^3} f. \quad (21-10)$$

Из (21-7) и (21-10)

$$Q_T = 2,25 \frac{EJM}{lh} \vartheta.$$

В случае сплошной пластинки $J = \frac{bh^3}{12}$ (фиг. 21,5 a),

тогда

$$Q_T = \frac{3}{16} ME \frac{bh^3}{l} \vartheta. \quad (21-11)$$

В случае U-образной пластинки $J = \frac{b_1 h^3}{6}$ (12-12)

$$Q_T = \frac{3}{8} ME \frac{b_1 h^3}{l} \vartheta. \quad (21-13)$$

Введем понятие — жесткость пластинки по температуре:

$$j_T = \frac{Q_T}{\vartheta}. \quad (21-14)$$

Для сплошной пластинки

$$j_T = \frac{3}{16} ME \frac{bh^3}{l} \left(\frac{\Gamma}{\circ C} \right). \quad (21-15)$$

Для V-образной

$$j_T = \frac{3}{8} ME \frac{b_1 h^3}{l} \left(\frac{\Gamma}{\circ C} \right). \quad (21-16)$$

Спираль. Уравнение механического момента, развиваемого спиральной пружиной при угле закручивания φ :

$$M_T = \frac{EJ}{l} \varphi = \frac{Ebh^3}{12l} \varphi \quad \text{Гмм.} \quad (21-17)$$

Деформация спирали при нагреве

$$\varphi = 1,5 M \frac{l}{h} \vartheta. \quad (21-18)$$

Из (21-17) и (21-18)

$$M_T = \frac{1}{8} bh^2 ME \vartheta. \quad (21-19)$$

$$j_T = \frac{M_T}{\vartheta} = \frac{1}{8} bh^2 ME. \quad (21-20)$$

Примеры. Полученными соотношениями пользуются в практически встречающихся случаях, когда биметаллическая пружина при своем перемещении должна преодолевать механические силы или создавать их в контактной системе.

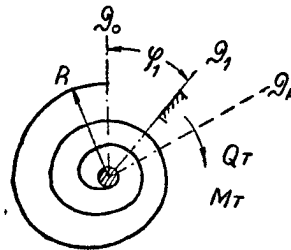
1) Определить давление на контакте (фиг. 21,5). Биметалл нагревается до температуры ϑ_k , соответственно этому свободная деформация должна быть f . Контакт расположен на расстоянии f_1 от начального положения пружины при ϑ_0 . Для деформации f_1 необходим нагрев ϑ_1 :

$$\vartheta_1 - \vartheta_0 = f_1 \frac{h}{0,75 M l^2};$$

$$Q_T = j_T (\vartheta_k - \vartheta_1).$$

2) Спиральная пружина (фиг. 21, 6). Начальные условия: ϑ_0 ; $\varphi_0 = 0$. Нагрев до температуры ϑ_k , свободный ход — φ_1 .

$$M_T = j_T (\vartheta_k - \vartheta_1); \quad \vartheta_1 = \varphi_1 \frac{h}{1,5 M l}; \quad Q_T = \frac{M_T k}{R}.$$



Фиг. 21, 6. Биметаллическая спираль.

21-5. ВОПРОСЫ ПРОЕКТИРОВАНИЯ ТЕПЛОВЫХ РЕЛЕ

Технические условия. Срабатывание теплового реле должно происходить при определенных значениях токов и времени срабатывания. При проектировании реле различают значения токов I_n — номинальное, I — ток перегрузки.

Введем понятие о коэффициенте перегрузки, представляющем собой кратность тока перегрузки к номинальному.

$\rho_0 = \frac{I}{I_n}$ — минимальное значение кратности тока, при котором должно произойти срабатывание реле. Обычно принимают $\rho_0 = 1,2 \div 1,25$. При выборе меньших значений ρ_0 требуется тщательная подгонка всех частей реле и наличие температурной компенсации. Время t_c при кратности ρ_0 обычно не лимитируется. Важным является установление времени t_c при кратности $\rho > \rho_0$, причем выбор значения ρ определяется условиями работы реле. Так, для защиты электродвигателей малой и средней мощности, характеризующихся тяжелыми условиями пуска — значительная кратность пускового тока, продолжительный пусковой режим, требуется время $t_c = 1 \div 5$ секунд при $\rho = 8$.

Для защиты силовых цепей установок, характеризующихся незначительной кратностью пускового тока, кратковременным пусковым режимом, контрольной точкой характеристики принимают $\rho = 2,5$; $t_c = (3 \div 20)$ секунд.

При проектировании реле для подвижных установок необходимо учитывать условия вибро- и ударостойкости. При вибрации и тряске установок появляются колебания небольшой амплитуды определенной частоты, например амплитуды 1 мм, $f = 45$ гц. При наличии вибраций и ударов в установке могут иметь случаи непроизвольного срабатывания реле.

Конструктивные вопросы. В практике применяется три способа нагрева биметалла: прямой и косвенный, в редких исполнениях — комбинированный. Выбор способа нагрева определяется, в основном, условиями создания шкалы токов реле.

Необходимо всегда выполнять реле с большой градацией сил токов, чтобы номинальный ток реле мало отличался от номинального тока двигателя. Так, в реле ТРА в пределах токов 7 — 52 а применяется 15 исполнений.

При прямом нагреве применяются U-образные биметаллические пластины. Большое число исполнений по току срабатывания достигается изменением размеров биметалла по ширине b_1 или толщине h ; в некоторых случаях применяют пакеты из нескольких пластин. При больших значениях номинального тока ($I_n > 50a$) шунтируют биметалл электрическим сопротивлением.

Способ прямого подогрева имеет следующие недостатки: 1) необходимо иметь в производстве большое число штампов для изготовления пластин разных размеров; 2) при замене в эксплуатации исполнения реле для какого-либо значения I_n на другое не-

обходимо заменить биметаллический элемент, что связано с дополнительной настройкой.

При косвенном подогреве биметаллическая пластинка подогревается от отдельного нагревательного элемента, расположенного непосредственно над пластинкой. Одно и то же исполнение реле, с одним размером биметалла, пригодно для большого числа исполнений по току, меняется только нагревательный элемент. Так, реле исполнения на 25 а пригодно при установке соответствующих нагревательных элементов для сил токов от 0,5 до 25 а. При замене нагревательных элементов не требуется переградуировка реле.

Контактные системы. Исполнение конструкций с непосредственным воздействием биметаллической пластинки на контактную цепь нерационально, так как контактная цепь при нагреве или охлаждении биметалла замыкается или размыкается весьма медленно, что приводит к подгоранию контактов.

Моментное действие контактов достигается двумя способами:

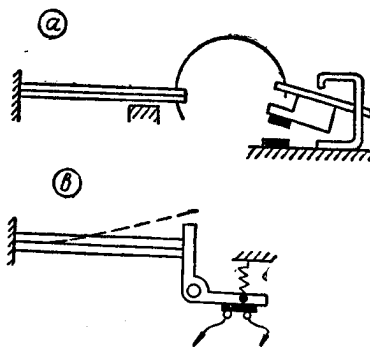
- 1) конец биметаллической пластинки связан с прыгающей пружиной (фиг. 21, 7 а);
- 2) конец биметаллической пластинки защемляет контактный рычаг реле, освобождая последний при нагреве (фиг. 21, 7 б);

В терморасцепителях для автоматических выключателей при нагреве биметаллической пластинки освобождается рычаг зацепления. Возврат контактной системы в исходное положение осуществляется от руки, что принимается из условий эксплуатации установки: не допускается автоматическое повторное включение реле или выключателя, так как размыкание цепи происходит в случаях аварийных режимов.

Температуры срабатывания биметалла тепловых реле. В биметаллической пластинке непосредственного нагрева температура распределяется неравномерно — наибольшая получается в средней части пластинки (фиг. 21, 8). Расчетные температуры: ϑ_n при I_n , ϑ_c при I_c принимают для точек, расположенных на расстоянии $1/4$ длины l_c от середины. Выбор значения ϑ_c связан с требуемой точностью работы реле при колебаниях температуры окружающей среды.

Установим связь между температурой биметалла и кратностью перегрузки. В установившемся режиме температура нагрева определяется электрическими потерями $I^2 r$ и теплоотдачей μS .

Обозначим: ϑ_0 , ϑ_n , ϑ_c — температуры окружающей среды, при номинальном токе I_n , при перегрузке $I = \rho_0 I_n$.



Фиг. 21, 7. Исполнения контактных систем.

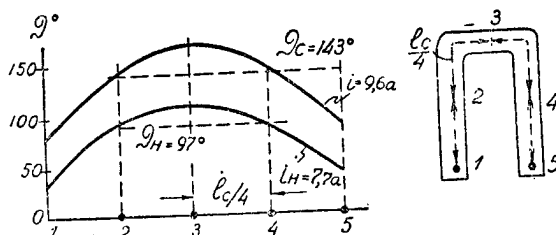
Тогда

$$\vartheta_c = \vartheta_0 + \frac{\rho_0 I_H^2 r}{\mu S} \quad (a); \quad \vartheta_H = \vartheta_0 + \frac{I_H^2 r}{\mu S} \quad (b).$$

Из (a) и (b)

$$\frac{\vartheta_c - \vartheta_0}{\vartheta_H - \vartheta_0} = \rho_0^2. \quad (21-21)$$

Таким образом, значение ρ_0 — минимальной кратности тока, при которой происходит срабатывание реле, зависит от темпера-



Фиг. 21, 8. Распределение температуры в пластинке при непосредственном подогреве.

туры окружающей среды. По заводским данным для реле типа ТРТ значения ρ_0 получаются:

$$\begin{aligned} \text{при } \vartheta_0 = 40^\circ & \quad \rho_0 = 1,2; \\ \vartheta_0 = 20^\circ & \quad \rho_0 = 1,5. \end{aligned}$$

Для уменьшения влияния колебаний температуры внешней среды на величину ρ_0 надо повышать значение ϑ_c . Рассмотрим для примера два реле: в первом $\vartheta_c = 90^\circ$, во втором $\vartheta_c = 150^\circ$.

1) 90° ; $\rho_0 = 1,2$ при $\vartheta_0 = 40^\circ$; определить ρ_{01} при $\vartheta_0 = 20^\circ$.

$$\frac{\vartheta_c - 20}{\vartheta_c - 40} = \frac{\rho_{01}^2}{\rho_0^2}; \quad \frac{90 - 20}{90 - 40} = \frac{\rho_{01}^2}{1,2^2}; \quad \rho_{01} = 1,42.$$

2) 150° ; $\rho_0 = 1,2$ при $\vartheta_0 = 40^\circ$; определить ρ_{01} при $\vartheta_0 = 20^\circ$.

$$\frac{150 - 20}{150 - 40} = \frac{\rho_{01}^2}{1,2^2}; \quad \rho_{01} = 1,26.$$

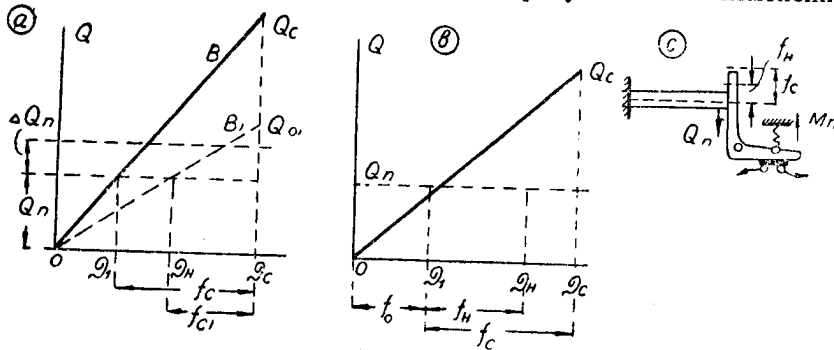
В настоящее время ряд заводов выполняет тепловые реле с учетом $\vartheta_c = 90^\circ$; однако более рационально принять ϑ_c порядка $150 \div 160^\circ$.

Коэффициент надежности. При расчете тепловых реле учитывают необходимость получения требуемой деформации биметаллической пластинки, f_c , при заданных ϑ_c и ρ_0 . Задача может быть решена при исполнении биметалла с разной термической жесткостью $f_T = \frac{Q_T}{\vartheta}$. Увеличение значения f_T приводит к увеличению

потребляемой мощности реле, уменьшение I_T снижает устойчивость и точность работы реле. Для решения задачи выбора требуемого значения I_T введем понятие о коэффициенте надежности m , величина которого определяется из условий работы реле.

1) *Точность работы реле.* Биметаллическая пластинка при своем перемещении должна преодолевать противодействующие механические силы Q_n , значение которых определяется силами трения или действием пружин терморасцепителя.

Работа механизма иллюстрируется диаграммой, приведенной на фигуре 21, 9а. Линии B и B_1 характеризуют закон изменения



Фиг. 21, 9. Диаграмма работы биметалла.

силы, создаваемой биметаллом, если пластинка не может свободно деформироваться. Углы наклона линий B и B_1 характеризуют термическую жесткость I_T и I_{T_1} пружины. При наличии противодействующей силы Q_n деформация пластинки начинается с температуры ϑ_1 или ϑ_{11} ; полная деформация пластинки при температуре срабатывания ϑ_c получается соответственно f_c и f_{c1} .

При изготовлении реле могут получиться значительные отклонения ΔQ_n от расчетного значения Q_n . При этом меняются значения ϑ_1 и ϑ_{11} , а следовательно, f_c и f_{c1} , что приводит к изменению тока срабатывания.

Из диаграммы (фиг. 21, 9а)

$$\Delta \vartheta = \frac{\Delta Q_n}{I_T}; \Delta \vartheta_1 = \frac{\Delta Q_n}{I_{T_1}}; \Delta f = 0,75M \frac{l^2}{h} \Delta \vartheta = 0,75M \frac{l^2}{h} \frac{\Delta Q_n}{I_T}. \quad (21-22)$$

Таким образом, при увеличении жесткости I_T уменьшается разброс точек срабатывания реле.

Коэффициент надежности

$$m = \frac{Q_c}{Q_n}, \quad (21-23)$$

где Q_c — условное значение силы на конце биметалла, при нагреве его до температуры ϑ_c и защемлении его конца.

$$Q_c = I_T \vartheta_c. \quad (21-24)$$

Из (21-23) и (21-24)

$$j_T = m \frac{Q_n}{\vartheta_c}. \quad (21-25)$$

Из условий точности работы реле принимают $m = 4 \div 5$.

2) *Вибростойкость*. При вибрациях и тряске возможны случаи отхода контактного рычага от конца биметаллической пластинки (фиг. 21, 9с). При этом Q_n становится равным нулю и пластинка может дополнительно переместиться на величину f_0 , определяемую из условия

$$f_0 = 0,75M \frac{l^2}{h} \vartheta_1. \quad (21-26)$$

При нагреве пластинки номинальным током температура повышается соответственно ϑ_n , а деформация f_n (фиг. 21, 9в).

Условием вибростойкости реле является

$$f_0 < f_c - f_n.$$

Рассмотрим критический случай

$$f_0 = f_c - f_n. \quad (21-27)$$

Так как деформации пропорциональны температурам, то из (21-27),

$$\vartheta_1 = \vartheta_c - \vartheta_n$$

или

$$\frac{Q_n}{j_T} = \vartheta_c - \vartheta_n.$$

Учитывая, что

$$\frac{\vartheta_c}{\vartheta_n} = \rho_0^2,$$

получаем

$$j_T = \frac{\rho_0^2}{\rho_0^2 - 1} \frac{Q_n}{\vartheta_c}.$$

Следовательно, в данном случае

$$m = \frac{\rho_0^2}{\rho_0^2 - 1} \quad (21-28)$$

и

$$j_T = m \frac{Q_n}{\vartheta_c}. \quad (21-29)$$

3) *Ударостойкость*. Наличие ускорений при ударе приводит к дополнительной деформации f_y , определяемой действием сил ускорения Q_y . Значение Q_y отнесем к величине Q_n : $Q_y = kQ_n$.

В данном случае

$$m = k \frac{\rho_0^2}{\rho_0^2 - 1}. \quad (21-30)$$

21-6. ВРЕМЯ СРАБАТЫВАНИЯ ТЕПЛОВЫХ РЕЛЕ

Время срабатывания тепловых реле t_c зависит от кратности перегрузки ρ . При заданной минимальной кратности ρ_0 время t_c не нормируется. В теоретических расчетах рассматривается критический случай: при ρ_0 время $t_c = \infty$.

Из теории нагрева однородного тела известно соотношение:

$$\vartheta = \vartheta_m (1 - e^{-\frac{t}{T}}), \quad (21-31)$$

где ϑ — переходное значение температуры нагрева тела;
 ϑ_m — максимальное значение при $t = \infty$;
 T — постоянная времени нагрева в секундах.
 Из (21-31)

$$t = T \ln \frac{\vartheta_m}{\vartheta_m - \vartheta}. \quad (21-32)$$

Определим время t_c срабатывания реле, которое зависит от температуры срабатывания ϑ_c .

Зависимость $\vartheta = \vartheta(I)$ без учета ϑ_0 — окружающей температуры (см. выражение 21-21)

$$\frac{\vartheta_m}{\vartheta} = \rho^2.$$

Согласно принятым выше определениям

$$\frac{\vartheta_m}{\vartheta_n} = \rho^2; \quad \frac{\vartheta_c}{\vartheta_n} = \rho_0^2.$$

Следовательно,

$$\vartheta_m = \vartheta_c \frac{\rho^2}{\rho_0^2}. \quad (21-33)$$

Подставляя в (21-32) значение ϑ_m из (21-33), получаем

$$t_c = T \ln \frac{\rho^2}{\rho^2 - \rho_0^2}. \quad (21-34)$$

Значение t_c — время срабатывания реле с холодного состояния биметалла.

Если биметалл предварительно нагрет током I_n до температуры ϑ_n , то время t_c уменьшается. Уравнение кривой нагрева в данном случае имеет вид

$$\vartheta_c = \vartheta_m (1 - e^{-\frac{t}{T}}) + \vartheta_n e^{-\frac{t}{T}}. \quad (21-35)$$

Разделив это выражение на ϑ_n , получаем

$$\frac{\vartheta_c}{\vartheta_n} = \frac{\vartheta_m}{\vartheta_n} (1 - e^{-\frac{t}{T}}) + e^{-\frac{t}{T}}.$$

или

$$p_0^2 = p^2 \left(1 - e^{-\frac{t}{T}}\right) + e^{-\frac{t}{T}}.$$

После проведения преобразований

$$t_c = T \ln \frac{p^2 - 1}{p^2 - p_0^2}. \quad (21-36)$$

Если биметаллическая пластинка должна преодолевать при своем перемещении механические противодействующие силы, то время с холодного состояния (21-34) несколько увеличивается. Время охлаждения t_0 от температуры ϑ_c до ϑ_n определяется из уравнения

$$\vartheta_n = \vartheta_0 e^{-\frac{t}{T}};$$

отсюда

$$t_0 = T \ln \frac{\vartheta_c}{\vartheta_n} = T \ln p_0^2. \quad (21-37)$$

Постоянная времени нагрева тепловых реле. Время срабатывания t_c определяется в основном постоянной времени нагрева биметалла T , значение которого зависит от различных факторов:

- 1) от режима работы — значение T меняется при разных значениях кратности перегрузки p ;
- 2) от способа нагрева биметалла — прямого или косвенного.
- 3) от условий теплоотдачи: значение μ коэффициента теплоотдачи меняется с изменением расположения биметалла и с изменением температуры нагрева.

Аналитическое определение значения T связано со значительными трудностями, необходимо провести ряд экспериментальных работ для установления в данном реле зависимости T от указанных факторов. Приблизительно значение T можно определить как отношение теплоемкости тела к количеству тепла, отдаваемого во внешнее пространство.

Биметаллическая пластинка — непосредственный нагрев.

$$T = \frac{cG}{\mu S}, \quad (21-38)$$

Здесь: c — теплоемкость тела, $\frac{вт \cdot сек}{г \cdot ^\circ C}$;

μ — коэффициент теплоотдачи $\frac{вт}{см^2 \cdot ^\circ C}$;

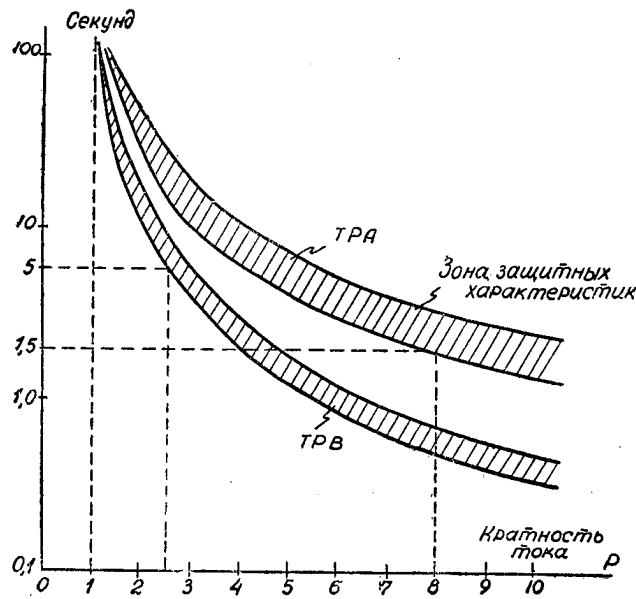
G , грамм; S — поверхность охлаждения, $см^2$.

Биметаллическая пластинка — косвенный нагрев. При нагреве пластинки от нагревательного элемента значение T увеличивается. Учет изменения T производится с помощью коэффициента β , учитывающего теплоотдачу от нагревательного элемента к биметаллической пластинке. В данном случае

$$T = \frac{cG}{\beta \mu S} \text{ секунд.} \quad (21-39)$$

Значение $\beta < 1$ и увеличивается с приближением нагревательного элемента к биметаллу.

Для каждого исполнения теплового реле или расцепителя строится зависимость $t = t(p)$, которую называют ампер-секундной характеристикой.



Фиг. 21, 10. Ампер-секундные характеристики реле.

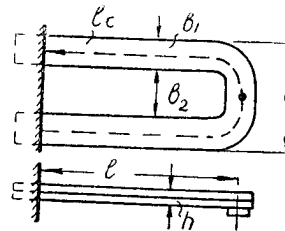
На фигуре 21, 10 приведены ампер-секундные характеристики реле типов ТРА и ТРВ с непосредственным нагревом биметалла. Реле ТРА предназначено для защиты электродвигателей; контрольной точкой для этого реле является $p = 8$. Реле ТРВ предназначено для защиты генераторов и силовых цепей; контрольная точка для него $p = 2,5$.

**21-7. ОСНОВНЫЕ СООТНОШЕНИЯ
ДЛЯ РАСЧЕТА ТЕПЛОВЫХ РЕЛЕ
И РАСЦЕПИТЕЛЕЙ**

Ниже приводятся основные соотношения для биметаллических реле с прямым подогревом (фиг. 21, 11).

Исходные соотношения. Деформация свободного конца биметаллической пластинки

$$f_c = 0,75M \frac{l^2}{h} (\vartheta_c - \vartheta_1), \quad (21-7)$$



Фиг. 21, 11. Эскиз биметаллической пластинки.

где ϑ_c — температура срабатывания реле;
 ϑ_1 — температура нагрева для создания на биметалле силы Q_n ,
 преодолевающей действие внешних механических сил.

$$\vartheta_1 = \frac{Q_n}{I_T}. \quad (21-14 a)$$

Жесткость пластинки по температуре

$$I_T = \frac{3}{8} ME \frac{b_1 h^2}{l}. \quad (21-16)$$

Учитывая m — значение коэффициента надежности, имеем

$$I_T = m \frac{Q_n}{\vartheta_c}. \quad (21-25)$$

Электрическое сопротивление биметалла

$$r = \rho \frac{l_c}{b_1 h} \text{ ом}; \quad (\rho - \text{ом} \cdot \text{мм}) \quad (21-40)$$

$$l_c = 2l + b_1 + h; \quad (21-41)$$

Поверхность охлаждения

$$S = 2(b_1 + h)l_c 10^{-2} \text{ кв. см.} \quad (21-42)$$

Условия теплового равновесия при токе срабатывания.

$$\rho_0^2 I_n^2 r = \mu S \vartheta_c; \quad (21-43)$$

с учетом (21-40) и (21-42):

$$\rho_0^2 I_n^2 \rho \frac{l_c}{b_1 h} = 2\mu l_c (b_1 + h) \vartheta_c \cdot 10^{-2}; \quad \left(\mu - \frac{\text{см}}{\text{см}^2 \cdot \text{°C}} \right).$$

Принимая

$$k_1 = \frac{\rho_0^2 I_n^2 \rho \cdot 10^2}{2\mu \vartheta_c}, \quad (21-44)$$

получаем

$$b_1 h (b_1 + h) = k_1. \quad (21-45)$$

Учитывая

$$t = T \ln \frac{\rho^2}{\rho^2 - \rho_0^2} \quad (21-34); \quad T = \frac{cG}{\mu S}, \quad (21-38)$$

получаем

$$t = \frac{c\gamma l_c b_1 h 10^{-8}}{2\mu (b_1 + h) l_c 10^{-2}} \ln \frac{\rho^2}{\rho^2 - \rho_0^2}.$$

Принимая

$$k_2 = \frac{20\mu t}{c\gamma \ln \frac{\rho^2}{\rho^2 - \rho_0^2}}, \quad (21-46)$$

имеем

$$\frac{b_1 h}{b_1 + h} = k_2. \quad (21-47)$$

Из (21-16) и (21-25)

$$\frac{3}{8} ME \frac{b_1 h^2}{l} = m \frac{Q_n}{\vartheta_c}.$$

Принимая

$$k_3 = \frac{8mQ_n}{3ME\vartheta_c}, \quad (21-48)$$

имеем

$$\frac{b_1 h^2}{l} = k_3. \quad (21-49)$$

Из (21-7) с учетом (21-14 а) и (21-25) имеем

$$f_c = 0,75M \frac{l^2}{h} \left(\vartheta_c - \frac{\vartheta_c}{m} \right).$$

Принимая

$$k_4 = 0,75M\vartheta_c \left(\frac{m-1}{m} \right), \quad (21-50)$$

имеем

$$f_c = k_4 \frac{l^2}{h}. \quad (21-51)$$

Пример расчета. (Исполнение реле см. фигуры 21, 9 и 21, 11).Заданы: $I_n = 10 \text{ а}$; $\rho_0 = 1,3$; $\rho = 3$; $t = 5 \text{ сек}$.По конструктивным условиям $m = 5$; $Q_n = 300 \text{ Г}$.Принимаем $\vartheta_c = 150^\circ$; биметалл марки ТБЗ;

$$M = 20 \cdot 10^{-6}; \quad E = 15 \cdot 10^6 \frac{\Gamma}{\text{мм}^2}; \quad c = 0,42 \frac{\text{ат. сек}}{\Gamma^\circ \text{С}};$$

$$\rho_{20} = 0,8 \cdot 10^{-3} \text{ ом} \cdot \text{мм}; \quad \rho_{150} = 0,88 \cdot 10^{-3}.$$

По опытным данным для тепловых реле малых габаритов

$$\mu = 5 \cdot 10^{-3} \frac{\text{ат}}{\text{см}^2 \text{ } ^\circ \text{С}}.$$

Вычисление постоянных. Из (21-44)

$$k_1 = \frac{1,69 \cdot 100 \cdot 0,88 \cdot 10^{-3} \cdot 10^2}{2 \cdot 5 \cdot 10^{-3} \cdot 150} = 9,9.$$

Из (21-46)

$$k_2 = \frac{20 \cdot 5 \cdot 10^{-6} \cdot 5}{0,42 \cdot 81 \ln \frac{9}{9-1,69}} = 0,712.$$

Из (21-48)

$$k_3 = \frac{8 \cdot 5 \cdot 300}{3 \cdot 20 \cdot 10^{-6} \cdot 15 \cdot 10^6 \cdot 150} = 8,9 \cdot 10^{-2}.$$

Из (21-50)

$$k_4 = 0,75 \cdot 20 \cdot 10^{-6} \cdot 150 \cdot \frac{4}{5} = 1,8 \cdot 10^{-3}.$$

Вычисление b_1 и h (из (21-45) и (21-47)).

$$b_1 h (b_1 + h) = 9,9; \quad \frac{b_1 h}{b_1 + h} = 0,712.$$

Из этих выражений после решения квадратного уравнения получаем $h = 1,86 \pm 0,89$; принимая меньшее значение h и округляя его, получаем $h = 1$ мм, тогда $b_1 = 2,66$; $b = 2b_1 + b_2 = 8$ мм.

Вычисление l . Из (21-49)

$$\frac{b_1 h^2}{l} = 8,9 \cdot 10^{-2}; \quad l \cong 30 \text{ мм.}$$

Вычисление f_c . Из (21-51)

$$f_c = 1,8 \cdot 10^{-3} \cdot 900 = 1,62 \text{ мм.}$$

Потери в биметалле при номинальном токе.

$$P = I_{nr}^2 r = 100 \frac{0,85 \cdot 10^{-3} \cdot 63,66}{2,66} = 2,05 \text{ вт,}$$

где $\rho = 0,85 \cdot 10^{-3}$ принято при ϑ_n , соответствующей I_n .

21-8. ТЕРМОРЕЛЕ И РЕГУЛЯТОРЫ

Термореле и терморегуляторы работают в функции температуры окружающей среды. Такие устройства называют также термостатами. Термореле применяются для защиты различных устройств при превышении температуры в этих устройствах сверх допустимой. Терморегуляторы предназначаются для поддержания постоянным значения температуры регулируемого объекта: электрической печи, прессформы для пластмасс и других.

Основные технические требования. Принципы построения термореле и терморегуляторов одинаковы. Однако при конструировании этих устройств учитывают особые требования, предъявляемые к реле и регуляторам.

1) Коэффициент возврата — отношение температуры отпуска к температуре срабатывания реле. Терморегуляторы должны обладать высоким коэффициентом возврата — порядка 0,97—0,98, так как это условие определяет точность регулирования. К термореле такие требования не предъявляются.

2) Влияние скорости изменения температуры окружающей среды. Конструкция должна быть выполнена так, чтобы точность работы устройства незначительно уменьшалась при больших скоростях изменения температуры окружающей среды. Гарантийные данные точности для регуляторов приводятся к скорости порядка $0,5 — 1 \frac{\text{град}}{\text{мин}}$. И эта скорость незначительна, в ряде случаев наблюдается скорость порядка $50 \div 100 \frac{\text{град}}{\text{мин}}$.

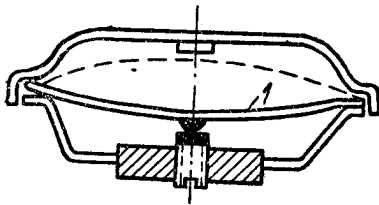
3) Контактная система должна выполняться так, чтобы осуществлялся мгновенный разрыв цепи. Это требование удовлетворяется в ряде конструкций термореле применением так называемых опрокидывающихся биметаллических шайб и пластин

При этом, однако, уменьшается значение коэффициента возврата, вследствие чего терморегуляторы часто выполняют с медленно расходящимися контактами. Такие контакты обладают незначительной коммутационной способностью; при применении металло-керамических контактов допускается значительная нагрузка. Так, в терморегуляторах для утюгов с медленно расходящимися металло-керамическими контактами осуществляется надежное замыкание цепи 220 в переменного тока с электрической нагрузкой 600 *вт*.

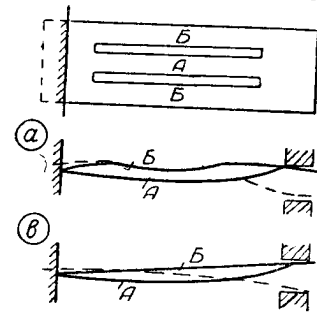
4) Регулирование уставки. Терморегуляторы должны снабжаться устройством для изменения уставки температуры срабатывания в значительных пределах. Так, в регуляторе ТР200 осуществляется регулирование в пределах от 20 до 200°. Такое требование не предъявляется к конструкции термореле. Обычно такие реле выполняются для определенной температуры срабатывания, должна быть только предусмотрена возможность корректировки при градуировке устройства.

Биметаллические термостаты применяются для контроля температуры до 250°.

На фигуре 21, 12 приведен эскиз термореле с шайбой. Шайба выполнена так, чтобы она была изогнута в холодном состоянии. При нагреве и достижении критической температуры шайба резко опрокидывается и размыкает контакты (см. пункт); возврат в исходное положение получается при охлаждении шайбы. Регули-



Фиг. 21, 12. Термореле с опрокидывающейся шайбой (1).

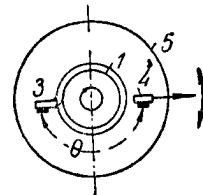
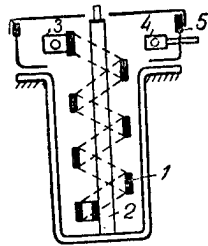


Фиг. 21, 13. Биметаллические пластинки:
а — с гофрированными краями,
б — с деформированной средней частью.

рование температуры срабатывания осуществляется контактным винтом.

В качестве измерительных элементов применяют также опрокидывающиеся биметаллические пластинки (фиг. 21, 13). В такой пластинке с прорезями внешние части гофрируют, вследствие чего внутренняя часть А изгибается. При нагреве эта часть стремится изогнуться в сторону уменьшения радиуса кривизны. Изгибу препятствуют сжимающие крайние части. При определенной критической температуре средняя часть скачком перегибается в обратную сторону (см. пункт).

Другое исполнение пластинки, примененное в термореле ХПИ, приведено на фигуре 21, 13*b*. Средняя часть пластинки растягивается под прессом, пластинка обладает свойствами гофрированной, но технология изготовления ее значительно проще. При конструировании термореле обращают особое внимание на то, чтобы перепад температуры между корпусом и биметаллической пластинкой был минимальным. Это условие весьма важно при применении реле в качестве устройства защиты обмоток электродвигателей. Реле укрепляется на головке статора асинхронного двигателя.



Фиг. 21, 14. Термо-биметаллический регулятор:

1 — спираль; 2 — стержень; 3 — контакт спирали; 4 — контакт.

При нормальном режиме работы двигателя перепад температуры между обмоткой и биметаллическим элементом невелик — порядка $10 \div 15^\circ$. При резких перегрузках или коротких замыканиях, когда скорость изменения температуры обмотки достигает величин порядка $50 \frac{^\circ\text{C}}{\text{мин}}$, перепад температуры резко увеличивается.

При выполнении обмотки двигателя с изоляцией класса А предельная допустимая температура равна 105° ; практически допускают кратковременное превышение температуры до 160° . Таким образом, перепад температуры между обмоткой двигателя и измерительной частью реле при токах короткого замыкания не должен превосходить $55 \div 70^\circ$. Это условие выполняется при установке реле в тихоходных двигателях (до 1000 об/мин), в двигателях более быстроходных имеют место весьма большие скорости изменения температуры (при включении на сеть заторможенного двигателя) — перепад температуры между обмоткой и реле достигает $100\text{—}120^\circ$.

Исполнение терморегулятора для электрических печей и прессформ приведено на фигуре 21, 14. Биметаллический элемент в виде спирали закреплен внизу на оси. В верхней части спирали имеется контакт. Угол перемещения свободного конца спирали θ регулируется перемещением в корпусе (5) регулировочного винта (4). В имеющихся исполнениях таких регуляторов пределы регулирования $25 \div 250^\circ$, точность $\pm 2 \div 3^\circ$.

Дилатометрические термостаты. Работа этих устройств основана на разности линейных перемещений при нагреве двух разных металлов. В качестве термоинертных материалов применяют: при температурах до 200° — инвар (при более высоких температурах коэффициент линейного расширения инвара, α , резко возрастает — см. таблицу 30); при температурах больше 200° — до 1000° — кварц, фарфор, обладающие весьма низким коэффициентом $\alpha = 5 \cdot 10^{-7}$. В качестве термоактивных материалов применяют:

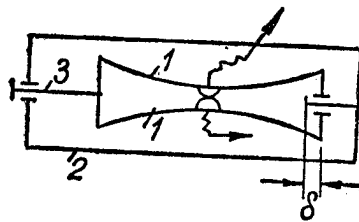
в интервале температур $0 \div 400^\circ$ — латунь, $\alpha = 21 \cdot 10^{-6}$; в интервале температур $0 \div 1000^\circ$ — никель, $\alpha = 16 \cdot 10^{-6}$.

Коэффициенты линейного расширения инвара Таблица 30

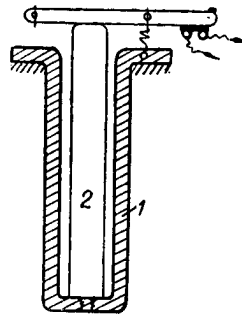
Пределы температур	-100—0	0—100	100—200	250—350	300—400	400—500
Коэффиц. на 1°C	$1,06 \cdot 10^{-6}$	$1,5 \cdot 10^{-6}$	$2,8 \cdot 10^{-6}$	$9,6 \cdot 10^{-6}$	$13,7 \cdot 10^{-6}$	$16,1 \cdot 10^{-6}$

Дилатометрические термостаты исполняют в качестве реле и регуляторов.

Принцип устройства регулятора ТР-200 иллюстрируется фигурой 21, 15. Две пластинки из инвара (1) изогнуты так, что в холодном состоянии они смыкаются, создавая давление на контактах. При нагреве латунной трубки последняя удлиняется больше пластинок инвара,



Фиг. 21, 15. Принцип устройства терморегулятора ТР-200:
1 — пластинка из инвара; 2 — латунная трубка; 3 — регулировочный винт.



Фиг. 21, 16. Принцип устройства дилатометрического регулятора:
1 — трубка; 2 — стержень.

которые растягиваются и разрывают контактную цепь. Регулировка уставки от $25 \div 200^\circ$ осуществляется винтом 3. Погрешность $\pm 5^\circ$ при скорости изменения температуры окружающей среды не более $5 \frac{\text{град}}{\text{мин}}$.

Другое исполнение дилатометрического регулятора приведено на фигуре 21, 16.

ГЛАВА 22

ПРУЖИНЫ

22-1. КЛАССИФИКАЦИЯ

Пружина, как правило, является составной частью каждого электромеханического реле. Пружина может быть использована в качестве элемента контактной системы, в качестве противодействующего устройства — для возврата якоря или диска, в качестве промежуточного звена для гибкой связи двух деталей, в качестве аккумулятора энергии — заводные пружины. Различают следующие

щие основные виды пружин, отличающиеся между собой или исполнением, или характером деформации.

Прямые и изогнутые пружины (фиг. 22, 1; 22, 2; 22, 5) имеют вид балки, закрепленной в одном конце или посередине, и исполняются из материала с прямоугольным или круглым сечением. При работе пружины в материале возникают напряжения изгиба — пружина работает на изгиб. Прямые и изогнутые пружины применяются в реле в основном в качестве элементов контактной системы. Иногда такие пружины ставят для передачи силы при незначительном почти прямолинейном ходе, например, в запорных собачках или тормозах.

Винтовые цилиндрические пружины (фиг. 22, 9) получают при навивке проволоки вокруг цилиндрического стержня по винтовой линии. При работе пружины в материале возникают деформации кручения. В зависимости от способа крепления пружины и воздействия на нее внешних сил различают: 1) пружины, работающие на растяжение; 2) пружины, работающие на сжатие.

Винтовые конические пружины (фиг. 22, 19) получают при навивке проволоки вокруг конического стержня. В плане такая пружина имеет вид спирали. В материале пружины возникают напряжения кручения и среза. Конические пружины обладают по сравнению с цилиндрическими большей устойчивостью от деформаций на боковой изгиб.

Винтовые цилиндрические пружины, работающие на изгиб (фиг. 22, 20). В винтовой цилиндрической пружине, закрепленной так, что при действии пружины оба конца ее поворачиваются вокруг оси стержня, происходят деформации изгиба — пружина работает на изгиб.

Спиральные пружины (фиг. 22, 22) получают при изгибе плоского ленточного материала в виде спирали. При работе пружины в материале ее возникают деформации изгиба. Спиральные пружины служат для сообщения какой-либо вращающейся оси крутящего момента: одним своим концом пружина прикрепляется к оси, другим — к какой-либо неподвижной точке.

Спиральные заводные пружины (фиг. 22, 24) играют роль аккумуляторов энергии и являются приводными пружинами в разных устройствах. Метод расчета спиральных заводных пружин несколько отличается от метода расчета обычных спиральных пружин.

22-2. МАТЕРИАЛЫ ДЛЯ ИЗГОТОВЛЕНИЯ ПРУЖИН

Виды материалов. В качестве материалов для изготовления пружин применяются или специальные сорта стали — ленточная пружинная сталь, рояльная проволока, или специальные сплавы меди — фосфористая бронза, латунь, бериллиевая бронза, нейзильбер, куниталь и другие. Выбор материала для пружины определяется механическими и электрическими свойствами этих материалов.

Электро- и теплопроводность. При работе пружины в качестве

элемента контактной системы необходимо выбрать материал с относительно большой электро- и теплопроводностью, например фосфористую бронзу, латунь.

Модуль упругости. При наличии относительно больших сил в системе и требующемся небольшом прогибе пружины применяется сталь (модуль упругости стали $E = 20 \times 10^3 \text{ кг/мм}^2$). Наоборот, если нужно получить относительно большой прогиб пружины при незначительных по величине силах, применяются материалы с меньшим значением модуля упругости, например фосфористая бронза (модуль упругости бронзы $E = 11 \cdot 10^3 \text{ кг/мм}^2$).

Значение допускаемых напряжений. В том случае, если при работе пружины возможны относительно большие деформации ее — прогиб, растяжение, закручивание, — следует выбирать материалы с относительно большими значениями допускаемых напряжений.

В рестроении применяются следующие материалы.

1) Пружинная стальная углеродистая проволока по ГОСТ 5047—49, холоднотянутая диаметром от 0,2 до 8 мм, применяемая для изготовления пружин, навиваемых в холодном состоянии и не подвергаемых закалке. В зависимости от предела прочности при растяжении эта проволока изготавливается трех классов: Н — нормальной прочности, П — повышенной, В — высокой.

2) Проволока стальная углеродистая. ГОСТ В 1546—42 марки ОВС — особо высокого сопротивления и ВС — высокого сопротивления.

3) Ленточные стали 60С2А, 65Г и У8А по ГОСТ 2283—43 для изготовления плоских пружин.

4) Ленточная сталь 70С2 × А по ГОСТ 2283—43 для изготовления спиральных заводных пружин. Для этих пружин применяется сталь У8А, У10А.

5) Фосфористая бронза — ГОСТ 1761—50 для плоских и спиральных пружин.

6) Ленты из нейзильбера ГОСТ 5187—49.

7) Бериллиевая бронза — ленты ГОСТ 1789—50, проволока по ЦМ/ТУ 673—41, обладающая высокими механическими свойствами.

8) Бронза КМ₅ 3-1 — ленты ГОСТ 4748—59, проволока ГОСТ 5222—50, дешевле фосфористой и бериллиевой бронзы; широко применяется для изготовления плоских пружин.

Выбор допускаемых напряжений. При выборе рассматриваются три категории нагрузки пружин: I — постоянная, II — повторная, пульсирующая, одного знака, III — знакопеременная. Среднее значение допускаемых напряжений:

$$R_{I \text{ узг}} = \frac{\sigma_b}{2}; R_{II \text{ узг}} = \frac{\sigma_b}{3}; R_{I \text{ кр}} = \frac{3}{4} R_{I \text{ узг}}; R_{II \text{ кр}} = \frac{4}{5} R_{II \text{ узг}},$$

где $\sigma_b = \frac{\text{кг}}{\text{мм}^2}$ — предел прочности при растяжении. Для ответственных пружин принимается больший запас прочности:

$$R_{I \text{ узг}} = \frac{\sigma_b}{2,5}; R_{II \text{ узг}} = \frac{\sigma_b}{4}; R_{I \text{ кр}} = \frac{3}{4} R_{I \text{ узг}}; R_{II \text{ кр}} = \frac{4}{5} R_{II \text{ узг}}.$$

При исполнении спиральных моментных пружин принимают весьма высокие запасы прочности — от 5 до 12.

Данные электро- и теплопроводности и допускаемых напряжений приведены в таблицах 31 и 32.

Таблица 31

Электропроводность и теплопроводность материалов

Материал	Электропроводность в % к проводимости чистой меди	Теплопроводность $\frac{\text{кал}}{\text{см}^2 \cdot \text{°C} / \text{см}}$	
Медь	98	0,9—0,94	ρ — чистой меди равно 0,018 $\frac{\text{ом} \cdot \text{мм}^2}{\text{м}}$ (при 20°С)
Фосфористая бронза . .	13	0,15—0,20	
Бериллиевая бронза . .	23—25	0,25	
Латунь	25	0,31	
Сталь	9	0,15	

Таблица 32

Допускаемые напряжения при расчете пружин

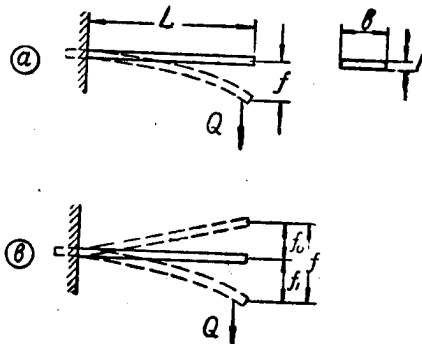
Материал	Марка	Предел прочности при растяжении σ_b	Допускаемые напряжения $\frac{\text{кг}}{\text{мм}^2}$				Модули	
			на кручение		на изгиб		E	G
			$R_{II \text{ кр}}$	$R_{II \text{ кр}}$	$R_{II \text{ изг}}$	$R_{II \text{ изг}}$	кг/мм ²	кг/мм ²
Стальная проволока ГОСТ 5047—49	В II Ø 0,2—0,6	265	80	50	95	65	20·10 ³	8·10 ³
	В II Ø 0,7—1,5	220 260	75	45	90	60		
	П—II Ø 1—3,0	160— 195	55	36	70	45		
Сталь 65Г, У8А, 60С2А			53	35	65	45	20·10 ³	8·10 ³
Бронза фосфористая	ОФ 6,5—0,15		21	15	26	18	10·10 ³	4·0—10 ³
Бронза бериллиевая	БрБ2		25	18	31	22	12,5·10 ³	5·10 ³
Нейзильбер	65—20 НМЦ		18	13				5·1—10 ³
Бронза	КМЦ 3—1		21	15	26	18	12·10 ³	4,8—10 ³

Сортамент проволоки по ГОСТ 5047—49:

Ø 0,2, 0,22, 0,25, 0,28, 0,3, 0,35, 0,4, 0,45, 0,5, 0,55, 0,6, 0,8, 0,9, 1,0, 1,1, 1,2, 1,3, 1,4, 1,5, 1,6, 1,8, 2,0.

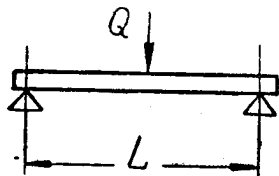
22-3. КОНСТРУКТИВНЫЕ ВОПРОСЫ

а) **Прямые и изогнутые пружины.** Прямые пружины чаще всего исполняются в виде консольной балки (фиг. 22, 1). В ряде случаев плоская пружина устанавливается при отсутствии внешней нагрузки не свободно, а имеет предварительный изгиб в обратную сторону на величину f_0 (фиг. 22, 1 в). Этот предварительный изгиб может быть осуществлен либо каким-нибудь упором, либо подгибом пружины, т. е. приданием ей предварительного натяжения f_0 . Общий прогиб f пружины равен предварительному прогибу f_0 плюс полезный прогиб f_1 . Полезный прогиб считаем от горизонтального положения вниз. Сила Q , требующаяся для получения полезного прогиба f_1 , определяется из условия прогиба на величину $f = f_0 + f_1$.

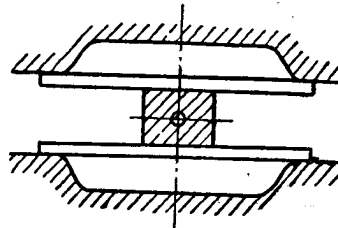


Фиг. 22, 1. Прямая пружина:
а — без предварительного изгиба;
б — с предварительным изгибом.

Прямая пружина, нагруженная посередине (фиг. 22, 2). Такие пружины применяются в качестве элементов контактных систем или в качестве защелки, когда требуется фиксация положения оси какого-либо переключающего устройства (фиг. 22, 3). В этом случае ось переключателя исполняется в виде квадрата. При расчете таких пружин исходят из



Фиг. 22, 2. Прямая пружина, нагруженная посередине.

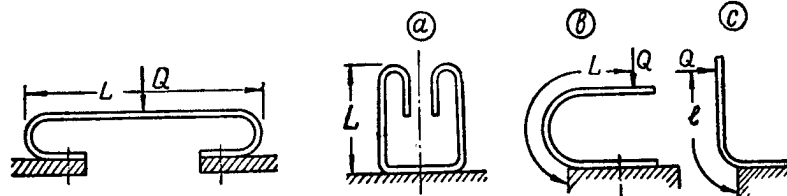


Фиг. 22, 3. Прямая пружина в качестве защелки.

предположения, что пружина зажата в средней части и на каждом конце нагружена силой $\frac{Q}{2}$.

Натянутые пружины (фиг. 22, 4). Исполняются в виде листовых пружин, закрепленных на обоих концах; сила приложена посередине. В этом случае применяются формулы изгиба для балок, укрепленных на двух концах.

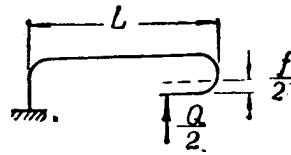
Изогнутые пружины (фиг. 22, 5). Расчет таких пружин производится аналогично прямым — их рассматривают как консольные балки, но допустимое напряжение изгиба $R_{изг}$ принимается с большим коэффициентом запаса.



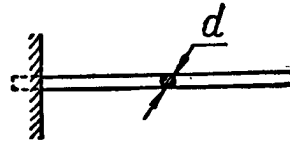
Фиг. 22, 4. Натянутая пружина.

Фиг. 22, 5. Изогнутые пружины.

Крепление прямых пружин обычно осуществляется двумя винтами или заклепками. При креплении одним винтом необходимо принимать меры против проворачивания пружины. Отверстия для

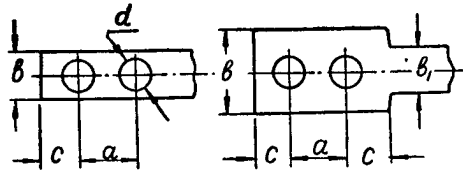


Фиг. 22, 6. Изогнутая пружина.

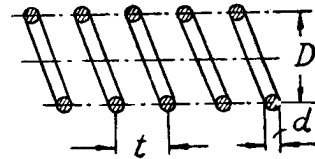


Фиг. 22, 7. Прямая пружина круглого сечения.

крепления в пружине производится штамповкой или сверлением. При этом в некоторых случаях требуется производить отжиг или отпуск конца пружины, соответствующего месту крепления.



Фиг. 22, 8. Крепление пружин.



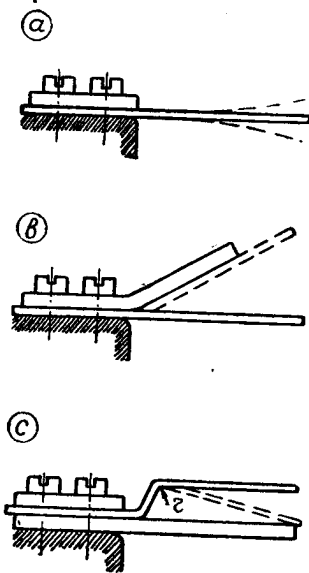
Фиг. 22, 9. Пружина с шаговой навивкой.

Таблица 33

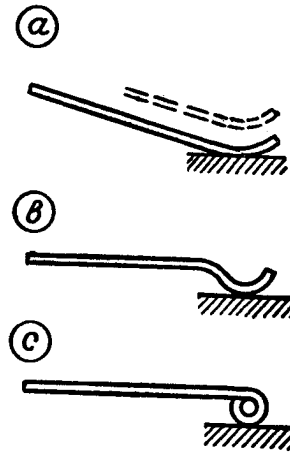
	Крепление	
	винтами	заклепками
d	0,5 b	0,3 b
a	От 3 до 4d	
c	0,5 b	

В некоторых случаях пружина в месте крепления выполняется более широкой (фиг. 22, 8), однако следует по возможности избегать резких изменений сечения. Рекомендуемые данные расположения отверстий и их размеров приводятся в таблице 33 в соответствии с обозначениями фигуры 22, 8 (d , a , c).

Рекомендуется крепить пружины посредством прижимных пластинок (фиг. 22, 10). Края пластинок, подкладок и других деталей, к которым прижимается пружина, следует закруглять. Свободные концы пружин, перемещающиеся по плоскостям (фиг. 22, 11), должны быть загнуты или иметь закругления, что предохраняет пружину от заклинивания и поломки.

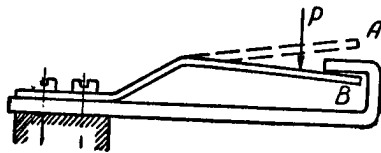


Фиг. 22, 10. Крепление пружин прижимными пластинками.

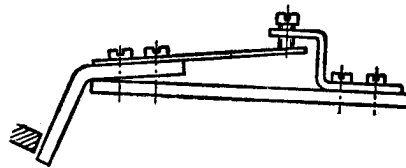


Фиг. 22, 11. Исполнение свободных концов пружин.

Предварительное натяжение пружины осуществляется отгибом ее в сторону, противоположную действию силы. В некоторых случаях для уменьшения хода пружины свободный конец ее удерживается упорной пластинкой. Такое исполнение показано на фигуре 22, 12.



Фиг. 22, 12. Пружина с предварительным изгибом:
А — свободное состояние пружины.



Фиг. 22, 13. Устройство для регулирования нажатия пружины.

Регулирование нажатия пружины осуществляется винтами (фиг. 22, 13).

Изогнутые пружины следует изготовлять с определенным минимальным радиусом r закругления в месте изгиба.

Рекомендуемые минимальные значения r зависят от рода материала, толщины листа и направления штамповки: перпендикулярно или параллельно направлению вальцовки листа, из которого изготавливается пружина.

- 1) При штамповке перпендикулярно направлению вальцовки
 - $r_{\text{макс}} = 4h$ — для стальных пружин;
 - $r_{\text{мин}} = h$ — для пружин из латуни или фосфористой бронзы.
- 2) При штамповке параллельно направлению вальцовки
 - $r_{\text{мин}} = 8h$ — для стали;
 - $r_{\text{макс}} = 3h \div 5h$ — для фосфористой бронзы или латуни.

Термическая обработка. Плоские пружины, в которых необходимо вырубить отверстия, а также изогнутые пружины штампуют, как правило, в отожженном состоянии. После штамповки пружина подвергается закалке. Закалка пружин из ленточной стали производится в соляных ваннах при температуре 780—800° в течение 1—3 минут, охлаждение пружин — в масле при комнатной температуре. После закалки пружины дополнительно подвергаются отпуску в серитренной или масляной ванне, в зависимости от требуемой твердости. Обычно получается твердость в пределах $40 \div 55R_c$; соответственно этому температуры отпуска порядка 400—275°.

Для предохранения стальных пружин от коррозии их покрывают защитным слоем в гальванических ваннах. В этом случае возможно наводороживание металла, поэтому пружина должна пройти процесс обезводороживания в масляной ванне при температуре $210 \pm 10^\circ$ с выдержкой в течение 5 часов.

Штамповка и гнутье пружин из бериллиевой бронзы осуществляются в закаленном (мягком) их состоянии. Режим закалки: нагрев до температуры 750—780°, выдержка при этой температуре в течение 40—50 минут и охлаждение в холодной воде. После штамповки пружина подвергается облагораживанию (старению). Режим облагораживания: нагрев до 300° с выдержкой при этой температуре в течение 8—9 часов, охлаждение в спокойном воздухе.

Аналогичные процессы проводятся и с другими пружинами из цветных сплавов, например из кундуала.

б) Винтовые цилиндрические пружины. Такие пружины чаще всего изготавливаются из проволоки круглого сечения. В реле-строении применяются пружины с небольшим шагом витка, поэтому при определении напряжений в поперечном сечении витка пружины считают, что каждый отдельный виток расположен в плоскости, перпендикулярной к оси витка.

Соотношение основных геометрических размеров пружин. Следующие геометрические размеры являются основными для винтовых цилиндрических пружин: средний диаметр витка D , длина пружины L и диаметр проволоки d .

Характерными параметрами пружины принято считать следующие:

$$\frac{D}{d} = c_1; \quad \frac{L}{D} = c_2.$$

В рационально исполненных пружинах величины c_1 и c_2 находятся в определенных пределах. В обычном исполнении пружин

$$c_1 = 5 \div 10;$$

$$c_2 \leq 5.$$

Важно выдерживать второе условие для значения c_2 в пружинах, работающих на сжатие. Когда $c_2 > 5$ при работе пружины может получаться боковое выпучивание. При этом создается трение в витках, искажающее характеристику пружины. При проектировании пружин рекомендуется выбирать следующие размеры D и d : $D \rightarrow 5; 6; 8; 10; 12,5; 16; 20; 25; 32; 40; 50; 62$ мм; $d \rightarrow 0,3; 0,4; 0,5; 0,6; 0,8; 1; 1,2; 1,6; 2; 2,5; 3; 4,5; 6$ мм.

Винтовая пружина, работающая на растяжение. Навивка цилиндрических винтовых пружин растяжения может быть осуществлена двумя способами.

1. *Закр*ытая навивка выполняется так, чтобы витки пружины соприкасались друг с другом с некоторым натяжением. Последнее условие достигается тем, что при закручивании вокруг стержня навивки проволоку вращают одновременно и вокруг своей оси. Сила натяжения пропорциональна углу поворота α проволоки вокруг оси. Пружины с закрытой навивкой работают с предварительным натяжением, которое они получают в процессе изготовления. Для удаления витков друг от друга надо приложить силу Q_0 , компенсирующую предварительное натяжение. Применяются такие пружины в тех случаях, когда требуется получить большую силу растяжения при малой длине пружины.

2. *Шаговая навивка* выполняется так, что в готовой навивке пружине имеется расстояние между витками — шаг t (фиг. 22, 9).

В определенной конструкции аппарата работа пружины характеризуется обычно двумя положениями: начальным и конечным. Начальному положению пружины может соответствовать или свободное состояние пружины, или предварительное растяжение. При свободном состоянии пружины начальная сила Q_0 будет соответственно равна (фиг. 22, 14):

$Q_n = Q_0$ — для пружины с предварительным натяжением (закрытая навивка);

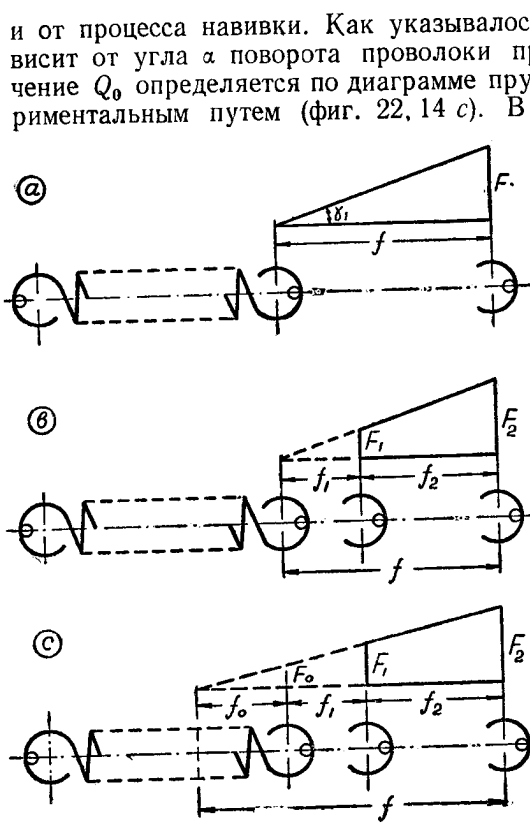
$Q_n = 0$ — для пружины с шаговой навивкой.

Для пружин с предварительным растяжением начальная сила определяется величиной начального растяжения f_1 , а в пружинах с предварительным натяжением еще дополнительно величиной f_0 , соответствующей силе Q_0 :

$Q_n = kf_1$ — для пружин с шаговой навивкой;

$Q_n = k(f_0 + f_1)$ — для пружин с закрытой навивкой.

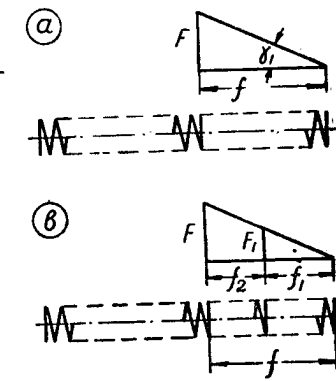
Значение силы Q_0 не может быть найдено аналитически, так как ее величина зависит не только от размеров пружины, но



Фиг. 22, 14. Диаграммы пружин растяжения: а - свободной пружины; б - с предварительным растяжением - шаговая навивка; с - с закрытой навивкой.

и от процесса навивки. Как указывалось выше, значение Q_0 зависит от угла α поворота проволоки при навивке. Обычно значение Q_0 определяется по диаграмме пружины, полученной экспериментальным путем (фиг. 22, 14 с). В начальной точке, когда пружина находится в спокойном состоянии и витки соприкасаются друг с другом, в ней имеется предварительное натяжение Q_0 . При растяжении пружины на величины f_1 и f_2 получаем силы Q_1 и Q_2 .

Поскольку работа пружины происходит



Фиг. 22, 14А. Диаграммы для винтовых пружин сжатия: а - свободная пружина; б - с предварительным сжатием.

в пределах упругих деформаций, все значения Q_0 , Q_1 и Q_2 пропорциональны величинам f_0 , f_1 и f_2 (см. диаграммы фиг. 22, 14):

$$\frac{f_0}{f_0 + f_1 + f_2} = \frac{Q_0}{Q_2}; f = f_0 + f_1 + f_2; Q_0 = \frac{f - (f_1 + f_2)}{f} Q_2.$$

Обычно пружины с предварительным натяжением (с закрытой навивкой) исполняют так, что сила Q_0 создает в материале пружины напряжение τ порядка 15 кг/мм^2 (для стали).

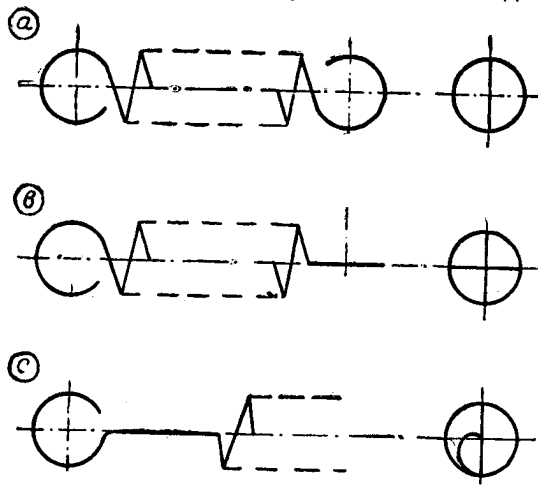
В предварительных расчетах пружин принимают ориентировочно:

$$Q_0 = \frac{Q_{\text{макс}}}{3} \text{ — для пружин с диаметром проволоки } d < 5 \text{ мм.}$$

$$Q_0 = \frac{Q_{\text{макс}}}{4} \text{ при } d > 5 \text{ мм.}$$

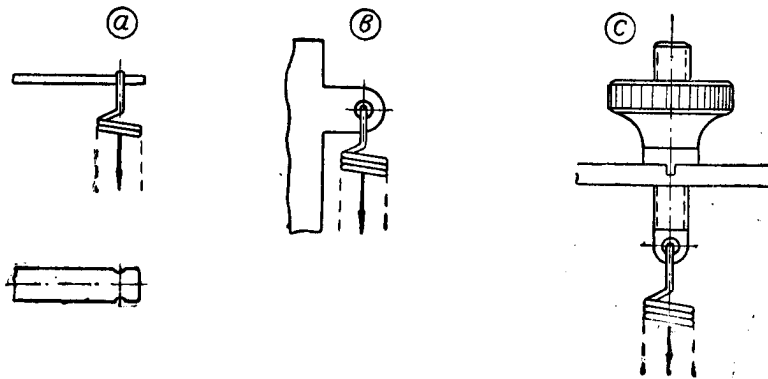
Шаг пружины растяжения (при шаговой навивке) принимается обычно равным диаметру проволоки.

Концы пружины исполняются в виде отдельных витков (фиг. 22, 15), которые должны быть расположены по оси пружины. Оба концевых витка могут быть расположены в одной плоскости



Фиг. 22, 15. Заделка концов пружины растяжения.

или в разных: один виток повернут в этом случае на 90° . По конструктивным соображениям приходится иногда удлинить конец пружины (фиг. 22, 16 c).



Фиг. 22, 16. Формы крепления пружин растяжения.

Концы пружины укрепляют на деталях аппарата, изображенных на фигуре 22, 16.

Концевые витки также принимают участие в работе пружины. При расчете пружины оба концевых витка считают за один виток.

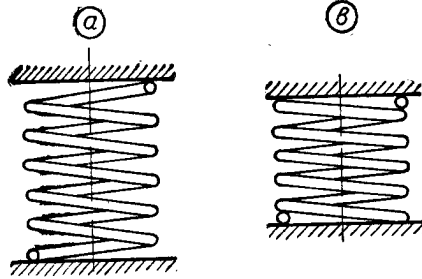
Для пружин с предварительным растяжением величины растяжений (фиг. 22, 14)

$$f_1 = f_2 \frac{Q_1}{Q_2 - Q_1}; f = f_2 \frac{Q_2}{Q_2 - Q_1}.$$

Для пружин с предварительным натяжением и предварительным растяжением (фиг. 22, 14 c).

$$f = f_0 + f_1 + f_2;$$

$$Q_0 = Q_2 \frac{f - (f_1 + f_2)}{f}.$$



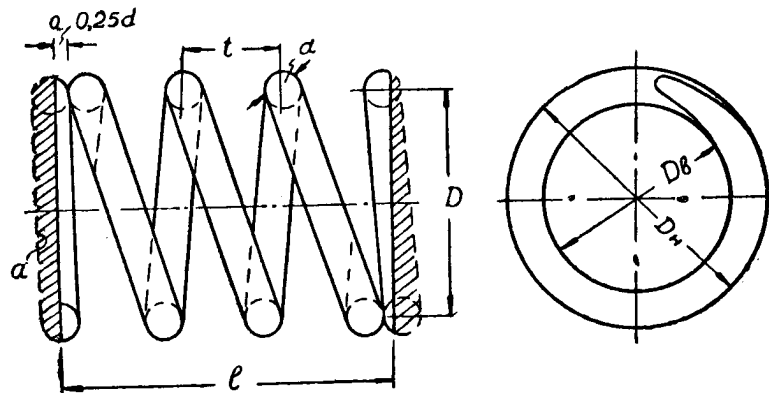
Фиг. 22, 17. Заделка концов пружины сжатия.

Винтовые пружины, работающие на сжатие. Длина пружины L должна быть не более $5D$. Величина t шага не нормируется. Принятые в практике значения $t = \frac{D}{2} \div \frac{D}{3}$.

Заделка концов. Возможны три исполнения:

1. Свободное (фиг. 22, 17a). Применяется редко.
2. Концевые витки частично навиты без уклона, поджаты (фиг. 22, 17 b). При таком исполнении устраняется опасность искривления пружины моментом, определяемым силой Q и радиусом пружины. Применяется в пружинах из проволоки $d < 0,5$ мм.

$$w = 4,5; w_p = 2,5$$



Фиг. 22, 18. Заделка концов пружины сжатия.

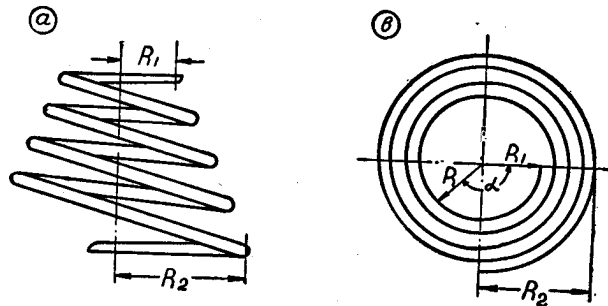
3. Концевые витки поджаты и подшлифованы. Крайние, нерабочие витки поджимаются к рабочим и сошлифовываются до размера $0,25d$ (фиг. 22, 18). Применяется в пружинах из проволоки $d > 0,5$ мм.

Пружина удерживается в своем положении соответствующими частями аппарата. В ряде случаев для удержания пружины приходится ставить специальные шайбы или приспособлять детали аппарата для этой цели, делая в них специальные выемки или выступ.

Защита от коррозии. Пружины из проволоки диаметром от 0,2 до 0,5 мм покрываются черным лаком методом окунания с последующей сушкой в печи, где температура поднимается равномерно от 50 до 220°; выдержка в печи длится 30 минут. Получается твердая лаковая пленка черного цвета. Пружины из проволоки диаметром выше 0,5 мм цинкуют, толщина покрытия порядка 10—15 микрон.

Термическая обработка. Готовые пружины подвергаются термической обработке для снятия внутренних напряжений, получающихся после процесса навивки, а также для обезводороживания пружин, прошедших гальванические ванны. Термическая обработка производится в масляной ванне при температуре порядка $210 \pm 10^\circ$ с выдержкой в течение 5 часов. Пружины, подлежащие окраске, должны быть подвергнуты термической обработке до процесса покраски.

в) Винтовые конические пружины по сравнению с цилиндрическими пружинами сжатия обладают большей устойчивостью

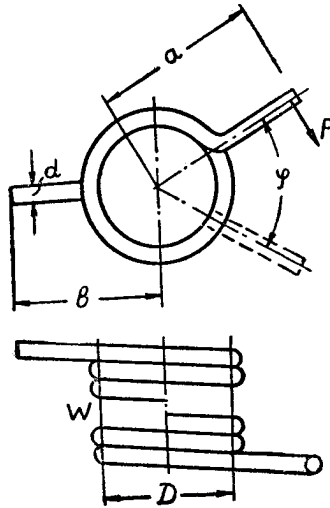


Фиг. 22, 19. Коническая винтовая пружина:
а — вид спереди; б — план.

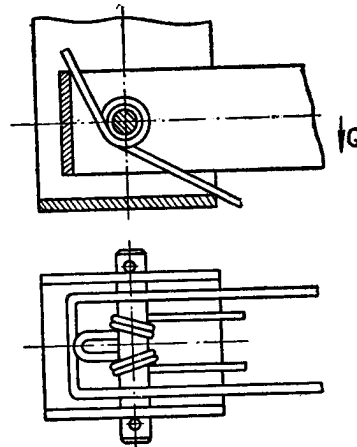
против бокового изгиба. Винтовая коническая пружина состоит из винтовых витков, как бы навитых на круговой направляющий конус (фиг. 22, 19). Концы пружины замкнуты в круг и для того, чтобы обеспечить прилегание их по всей периферии к опорным плоскостям, сошлифовываются. Конические пружины могут иметь два исполнения — с постоянным шагом и с постоянным углом наклона. Горизонтальная проекция пружины с постоянным шагом представляет собой архимедову спираль с постоянным углом наклона — логарифмическую.

В практике релестроения приняты пружины с постоянным шагом.

г) **Винтовые пружины, работающие на изгиб.** В винтовой пружине, исполненной так, что оба ее конца поворачиваются вокруг оси (фиг. 22, 20), происходит деформация изгиба — пружина работает на изгиб. С целью уменьшения передачи сил на ось следят за тем, чтобы оба конца пружины при монтаже ее были параллельны, витки должны свободно поворачиваться вокруг стержня. Обычно бывает достаточно, чтобы пружина своими свободными концами только прилегала к штифту неподвижной части



Фиг. 22, 20. Винтовая пружина, работающая на изгиб.



Фиг. 22, 21. Винтовая двойная пружина.

или к краю неподвижной части механизма. Для более равномерного распределения сил, действующих на ось, пружину исполняют двойной (фиг. 22, 21).

д) **Спиральные моментные пружины.** Спиральные пружины применяются для сообщения какой-либо оси крутящего момента. Исполняются такие пружины в виде завитой в форме спирали плоской ленты. Один конец пружины — внутренний — закрепляется на оси (фиг. 22, 22). Для перемещения наружного конца A на угол φ вокруг оси требуется сила Q на плече R_n . На ось тогда действует крутящий момент

$$M = QR_n.$$

При этом безразлично, происходит ли вращение по часовой стрелке или против нее, т. е. свертывается ли пружина — заводится, или разворачивается — отдает работу. В материале пружины получаются напряжения изгиба.

Ход расчета. Обычно спиральная пружина характеризуется моментом, требующимся для поворота конца A пружины

на 90° . При этом безразлично, закручивается или разворачивается пружина. Угол закручивания в действительных условиях работы пружины бывает значительно больше 90° . Значение предельного допустимого угла закручивания ограничивается двумя условиями:

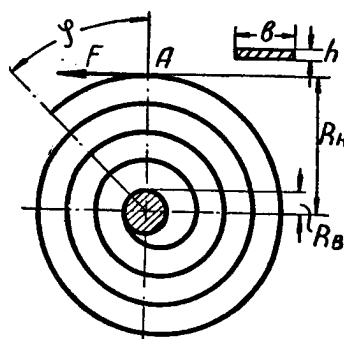
1. При увеличении закручивания повышается напряжение материала. Требуется, чтобы остаточная деформация материала не превышала при этом определенной величины (порядка $0,1\%$).

2. При значительном угле закручивания пружины (порядка 360° и выше) уменьшается расстояние a между витками. Необходимо следить за тем, чтобы витки пружины не соприкасались, так как при этом вследствие возникающего добавочного момента трения меняется характеристика пружины. При расчете задаются моментом пружины, определяемым данными конструкции аппарата, и материалом пружины. Для пружин с относительно большим моментом применяется ленточная сталь.

В электромеханических реле — электромагнитных, индукционных, электродинамических — применяются материалы с меньшим модулем упругости — фосфористая бронза, нейзильбер, куньяль и другие сплавы. В тех случаях, когда пружина является токоведущей, применяются материалы с меньшим удельным сопротивлением, например кадмиевая медь, серебро (для пружин с малым моментом). Выбор материала определяет собой модуль упругости и величину допускаемого напряжения на изгиб.

Обычно при расчете спиральных пружин принимается большое значение запаса прочности, так как в рабочем состоянии пружина закручивается на угол φ в несколько раз больший 90° . Так, для фосфористой бронзы принимается $R_{изг}$ порядка $8-10 \text{ кг/мм}^2$, а для пружин особенно точных приборов — порядка $4-5 \text{ кг/мм}^2$. Конструкция пружины характеризуется отношением $\frac{L}{h}$, которое колеблется в пределах от 1200 до 3000. Меньшие значения $\frac{L}{h}$ принимаются для пружин со значительным моментом, большие — для пружин с малым моментом. Так, для пружин с моментом порядка $10-20 \text{ Гмм}$ значение $\frac{L}{h}$ бывает порядка 3000.

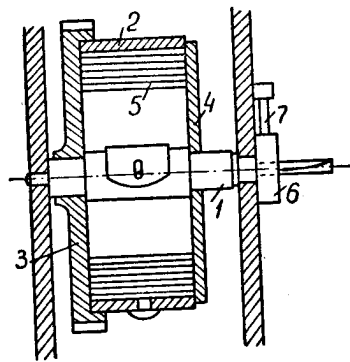
При расчете пружины задаются значением L и определяют величины h и b . Если значение b получается большим, увеличивают значение L .



Фиг. 22, 22. Спиральная пружина.

Конструктивные вопросы. Внутренний конец пружины загибается и закладывается в прорез стержня или оси пружины. Имеется также другая форма крепления конца пружины — посредством штифтов или винтов. Наружный конец пружин крепится к конструкции аппарата посредством хомутиков или пайкой.

е) **Заводные спиральные пружины.** Заводные спиральные пружины играют роль аккумуляторов энергии и являются приводными механизмами в разных устройствах. Такие пружины закладываются обычно в заводной барабан. Неподвижный наружный конец пружины закрепляется в корпусе барабана, а внутренний конец — на оси, которой и придается вращающий момент.



Фиг. 22, 23. Барабан с заводной пружиной.

Заводные пружины обычно бывают большой длины, но помещаются в корпусе с относительно малыми габаритами и поэтому должны быть завиты без промежутков между витками. На фигуре 22, 23 показано примерное устройство барабана. Ось имеет две цапфы, расположенные на стенках корпуса. На оси свободно вращается барабан 2. Дно барабана 3 исполнено с зубчатым венцом и входит в зацепление с механизмом (механизм на фигуре не показан). В верхней части барабана укрепляется крышка 4. Внутри барабана находится пружина 5, закрепленная внутренним концом к оси, а наружным к барабану. На квадратной части оси насажен храповик 6, удерживаемый собачкой 7. Завод пружины осуществляется поворотом оси. Для этого на квадратной части оси насаживается рукоятка (при ручном заводе) или ось связывается с приводным вспомогательным электродвигателем. Обратному вращению оси мешают храповик и собачка. Под действием пружины проворачивается барабан, а с ним и зубчатый венец, приводящий во вращение механизм устройства.

Расчет. В заведенном состоянии пружина не занимает всего объема барабана (фиг. 22, 24 а). При этом число витков пружины ω_2 , внутренний радиус r и наружный r_2 .

При спуске пружина изменяет число витков от ω_2 до ω_1 , внутренний и наружный радиусы пружины меняются от r до r_1 и от r_2 до R (фиг. 22, 24 б). Число оборотов ω заводного барабана равно разности числа витков заведенной и спущенной пружины: $\omega = \omega_2 - \omega_1$.

Число витков пружины в обоих случаях можно определить в функции ее геометрических размеров, считая, что в заведенном и спущенном состоянии пружины все витки ее плотно при

легают друг к другу. Считая толщину материала пружины h , имеем

$$\omega_2 = \frac{r_2 - r}{h} (a); \quad \omega_1 = \frac{R - r_1}{h} (b); \quad \omega = \omega_2 - \omega_1 = \frac{r_2 - r - R + r_1}{h} (c).$$

При конструировании заводных пружин стремятся получить максимально возможное число оборотов барабана. Это условие достигается при соблюдении определенного соотношения между размерами барабана. Объем материала заводной пружины до и после ее завода не меняется. Поэтому имеет место равенство

$$b(\pi R^2 - \pi r_1^2) = b(\pi r_2^2 - \pi r^2).$$

Следовательно,

$$r_2 = \sqrt{R^2 - r_1^2 + r^2}.$$

Подставляя значение r_2 в уравнение (c), получаем

$$\omega h = \sqrt{R^2 - r_1^2 + r^2} - r - R - r_1 \quad (d).$$

Заданными являются геометрические размеры барабана и оси — R , r , а также толщина материала пружины h . При этих заданных величинах число оборотов ω является функцией переменного значения r_1 . Определим условие получения $\omega_{\text{макс}}$. Для этого берем производную функцию по r_1 и приравняем ее нулю, учитывая, что $r = \text{const}$; $R = \text{const}$:

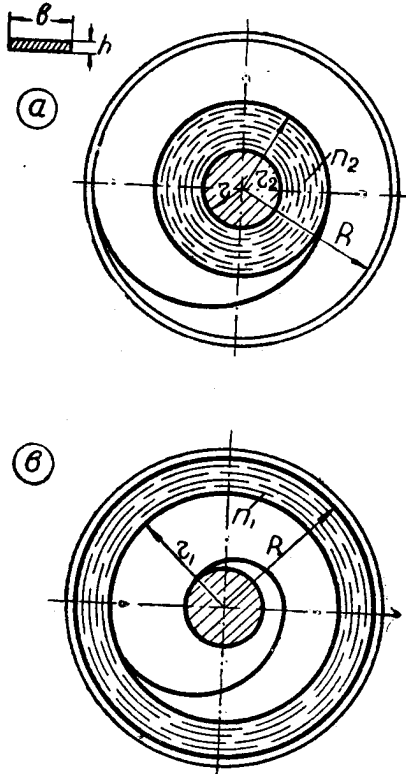
$$\frac{d\omega}{dr_1} = \frac{2r_1}{2\sqrt{R^2 - r_1^2 + r^2}} + 1 = 0.$$

Отсюда

$$r_1 = \sqrt{\frac{R^2 + r^2}{2}} \quad (e).$$

Подставляя значение r_1 в уравнение (d), имеем

$$r_2 = \sqrt{R^2 + r^2 - \frac{R^2 + r^2}{2}} = \sqrt{\frac{R^2 + r^2}{2}} \quad (g).$$



Фиг. 22, 24. Спиральная заводная пружина: а — пружина заведена; б — пружина спущена.

Таким образом, условием получения максимального числа оборотов пружины согласно (e) и (g) является равенство:

$$r_2 = r_1,$$

т. е. внешний радиус заведенной пружины должен быть равен внутреннему радиусу спущенной. Пружина, удовлетворяющая этому условию, называется нормальной заводной пружиной.

Расчетные данные 1. Материал пружины — обычно заводные пружины исполняются для получения больших значений моментов и поэтому, как правило, изготавливаются из ленточной стали или из стали для часовых пружин. Допускаемое напряжение на изгиб принимается с учетом коэффициента запаса $m = 1,5$, т. е. $R_{изг} = 70 \div 120 \text{ кг/мм}^2$.

2. Толщина материала пружины принимается в зависимости от внешнего радиуса пружины и допускаемого напряжения на изгиб. По опытным данным рекомендуется принимать $h = \frac{2R}{R_{изг}}$.

$$\text{Если } R_{изг} = 70 \div 120, \text{ то } h = \frac{2R}{70} \div \frac{2R}{120}.$$

Обычно избегают применения толстых стальных лент, так как при этом пружина работает неплавно и вероятность ее поломки большая, чем при более тонких пружинах. В рестроении применяются пружины с размером в пределах $0,3 \div 0,6 \text{ мм}$.

3. Значение радиуса оси определяется из соотношения

$$R/r = k.$$

Обычно принимается $k = 3$.

22-4. РАСЧЕТНЫЕ СООТНОШЕНИЯ И ПРИМЕРЫ

Прямая пружина (фиг. 22, 25 а).

$$f = \frac{1}{3} \frac{QL^3}{EJ}; \quad (22-1)$$

$$\text{момент } QL = WR_u. \quad (22-2)$$

$$J = \frac{bh^3}{12}; \quad W = \frac{bh^2}{6}.$$

Из (22-1) и (22-2)

$$f = \frac{2}{3} \frac{L^2 R_u}{h E}; \quad (22-3)$$

$$Q = \frac{bh^2 R_u}{6 L}. \quad (22-4)$$

Известны L, Q, f ; задавая R_u , определяют

$$h = \frac{2L^2 R_u}{3E f}; \quad (22-5)$$

$$b = \frac{6QL}{h^2 R_u}. \quad (22-6)$$

Пример.

$L = 40; f = 6; Q = 0,2 \text{ кг}.$

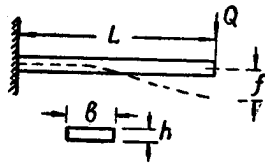
Материал: фосфористая бронза $E = 11 \cdot 10^3 \frac{\text{кг}}{\text{мм}^2}.$

$$R_u = 30 \frac{\text{кг}}{\text{мм}^2},$$

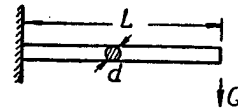
$$h = \frac{2 \cdot 40^2 \cdot 30}{3 \cdot 11 \cdot 10^3 \cdot 6} = 0,485, \text{ принимаем } h = 0,5 \text{ мм},$$

$$b = \frac{6 \cdot 0,2 \cdot 40}{0,5 \cdot 30} = 6,4 \text{ мм}.$$

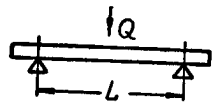
Фиг. 22, 25 (от а до к). Эскизы к расчету пружин:



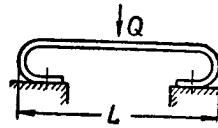
а — прямая;



б — круглого сечения;



с — прямая пружина, нагруженная посредине;



д — натянутая;

Прямая пружина круглого сечения (фиг. 22, 25 б)

$$f = \frac{2}{3} \frac{L^2 R_u}{dE}; \quad (22-7)$$

$$Q = \frac{\pi d^3 R_u}{32 L} \approx 0,1 d^3 \frac{R_u}{L}. \quad (22-8)$$

Прямая пружина, нагруженная посредине (фиг. 22, 25 с)

$$f = \frac{1}{6} \frac{L^2 R_u}{h E}; \quad (22-9)$$

$$Q = \frac{2}{3} \frac{bh^2 R_u}{L}. \quad (22-10)$$

Натянутая пружина (фиг. 22, 25 д)

$$f = \frac{1}{12} \frac{L^2 R_u}{h E}; \quad (22-11)$$

$$Q = \frac{4}{3} \frac{bh^2 R_u}{L}. \quad (22-12)$$

Контактная пружина изогнутая (фиг. 22, 25 e)

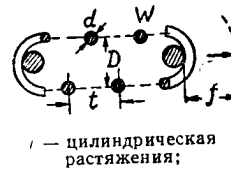
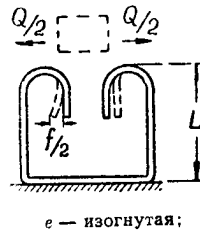
Пример. $f = 1$ мм; $Q = 0,4$ кг; $L = 16$ мм, фосфористая бронза.
Расчет для половины пружины. $0,5Q$; $0,5f$.

$$m_0 = 4; R_u = 12 \frac{\text{кг}}{\text{мм}^2},$$

согласно (22-5) и (22-6)

$$h = \frac{2 \cdot 16^2 \cdot 12}{3 \cdot 11 \cdot 10^8 \cdot 0,5} = 0,372 \text{ мм, принимаем } h = 0,4 \text{ мм}$$

$$b = \frac{6 \cdot 0,2 \cdot 16}{0,4^2 \cdot 12} = 10 \text{ мм.}$$



Винтовая цилиндрическая пружина растяжения (фиг. 22, 25 f)

$$c_1 = \frac{D}{d} = 5 \div 10; \quad f = \frac{8D^3 Q}{d^4 G} \omega. \quad (22-13)$$

$$c_2 = \frac{L}{D} \leq 5; \quad Q = \frac{\pi d^3}{8D} R_h. \quad (22-14)$$

Тогда

$$f = \frac{\pi \omega D^2 R_h}{Gd}. \quad (22-15)$$

Рекомендуемые данные:

D — 5; 6; 8; 10; 12,5; 16; 20; 25; 32; 40 мм;

d — 0,3; 0,4; 0,5; 0,6; 0,8; 1; 1,2; 1,6; 2 мм.

Из (22-14), задаваясь c_1 , получаем

$$d = \sqrt[3]{\frac{8c_1 Q}{\pi R_h}}. \quad (22-16)$$

Из (22-13) вычисляется ω , оба концевых витка считают за один рабочий.
Свободная длина пружины

$$\text{с шаговой навивкой} \rightarrow l_0 = t\omega_p + 2D \quad (22-17)$$

$$\text{с открытой навивкой} \rightarrow l_0 = d\omega_p + 2D \quad (22-18)$$

Длина проволоки с учетом концевых витков

$$L = \pi D (\omega + 1). \quad (22-19)$$

Жесткость пружины

$$i = \frac{Q}{f} = \frac{Gd^4}{8\omega D^3}. \quad (22-20)$$

Пример.

$$Q_1 = 0,55 \text{ кг}; Q_2 = 0,8 \text{ кг}; j = 0,05 \text{ кг/мм.}$$

Материал: рояльная проволока

$$G = 8 \cdot 10^3 \text{ кг/мм}^2; R_k = 80 \text{ кг/мм}^2.$$

В а р и а н т 1. Шаговая навивка $t = 1,5 d$; $c_1 = 10$.

$$d = \sqrt{\frac{8 \cdot 10 \cdot 0,8}{\pi 80}} = 0,55; \text{ принимаем } d = 0,5 \text{ мм.}$$

$$D = 10 \text{ мм}; t = \frac{Q_2}{j} = \frac{0,8}{0,55} = 16 \text{ мм}; f_1 = \frac{Q_1}{j} = 10 \text{ мм};$$

$$f_2 = 16 - 10 = 6 \text{ мм}; \omega_p = \frac{16 \cdot 0,5^4 \cdot 8000}{8 \cdot 5^3 \cdot 0,8} = 10;$$

$$l_0 = 1,5 \cdot 0,5 \cdot 10 + 10 = 17,5 \text{ мм.}$$

Полная длина пружины в растянутом состоянии

$$l = l_0 + f = 17,5 + 16 = 33,5 \text{ мм.}$$

В а р и а н т 2. Требуется, чтобы $l \leq 30 \text{ мм}$; принимаем закрытую навивку.

При $Q_2 = 0,8 \text{ кг}$; $f = l - l_0 = 30 - 17,5 = 12,5 \text{ мм}$;

$Q_2 - Q_0 = 12,5j = 12,5 \cdot 0,05 = 0,625 \text{ кг}$.

Предварительное натяжение

$$Q_0 = Q_2 - 0,625 = 0,175 \text{ кг.}$$

$$\frac{Q_0}{Q_2} = \frac{0,175}{0,8} = 0,218 \text{ — допустимое.}$$

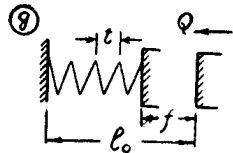
Цилиндрические пружины сжатия (фиг. 22, 25g)

Расчетные соотношения см. выше (пружины растяжения).

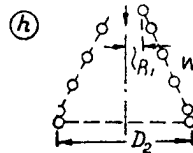
$$t = \frac{D}{2} \div \frac{D}{3}; \omega_p = \omega - 2;$$

$l_0 = t\omega_p + 1,5 d$ — свободное состояние,

$l = d\omega_p + 1,5 d$ — полностью сжатая.



g — цилиндрическая сжатия;



h — коническая;

Рекомендуемые данные:

$$\omega_p \rightarrow 2,5; 3,5; 4,5; 6,5; 8,5; 11,5; 15,5; 18,5.$$

Коническая пружина (фиг. 22, 25 h)

$$f = \frac{\pi \omega D_2^2 R_k}{4Gd}; \quad (22-21)$$

$$c_1 = \frac{D_2}{d}; \quad d = \sqrt{\frac{8Qc_1}{\pi R_k}}; \quad (22-22)$$

$$\omega = \frac{4dfG}{\pi d_2^2 R_k}. \quad (22-23)$$

Длина проволоки

$$L = \pi w \left(\frac{D_1 + D_2}{2} \right). \quad (22-24)$$

Винтовая пружина, работающая на изгиб (фиг. 22,20)

$$Q = W_z \frac{R_u}{a} = \frac{\pi d^3}{32a} R_u. \quad (22-25)$$

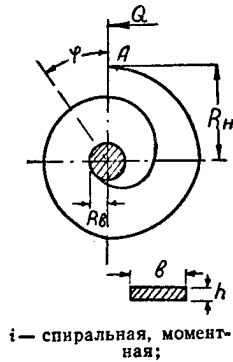
$$M = Qa = \frac{J_z E \varphi}{L}. \quad (22-26)$$

Прогиб свободного конца $f = a\varphi$.

$$f = 2 \frac{aLR_u}{Ed}; \quad (22-27)$$

$$L = \frac{J_z E \varphi}{Qa} = \frac{\pi d^4 E \varphi}{64 Qa}. \quad (22-28)$$

$$w = \frac{L - a - b + D}{\pi D}. \quad (22-29)$$



i - спиральная, моментная;

Пример. Задано: $Q_0 = 0,155 \text{ кг}$; $Q_1 = 0,2 \text{ кг}$;
 рабочий угол $\varphi_1 = 10^\circ = 0,174 \text{ рад}$;
 $D = 6 \text{ мм}$; $a = 9 \text{ мм}$; $b = 12 \text{ мм}$;
 Найти: $w_1 - d_1 - \varphi_0 - L$.

Материал: рояльная проволока $E = 20 \cdot 10^8 \frac{\text{кг}}{\text{мм}^2}$;

$$R_u = 30 \frac{\text{кг}}{\text{мм}^2}; \quad \frac{\varphi}{\varphi_0} = \frac{Q_1}{Q_1 - Q_0}; \quad \varphi = \varphi_0 \frac{Q_1}{Q_1 - Q_0} = 0,78 \text{ рад} = 45^\circ;$$

$$\varphi_0 = 35^\circ; \quad d = \sqrt[3]{\frac{32 \cdot 0,2 \cdot 9}{\pi 30}} = 0,85 \text{ мм};$$

$$L = \frac{\pi \cdot 0,85^4 \cdot 20 \cdot 10^8 \cdot 0,78}{64 \cdot 0,2 \cdot 9} = 223 \text{ мм}; \quad w = \frac{223 - 9 - 12 + 6}{\pi 6} = 11.$$

Спиральная пружина (фиг. 22, 25 i)

$$M = WR_u = \frac{bh^3}{6} R_u; \quad (22-30)$$

$$M = \frac{EJ_z}{L} \varphi; \quad (22-31)$$

$$J_z = \frac{bh^3}{12}.$$

L — развернутая длина ленты, мм.

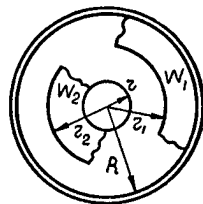
$$\varphi = 2 \frac{R_u L}{hE}. \quad (22-32)$$

$$h = \frac{2LR_u}{\varphi E}; \quad (22-33)$$

$$b = \frac{6M}{R_u h^2}. \quad (22-34)$$

$$L = \frac{\varphi E h}{2R_u}; \quad (22-35)$$

$$w = \frac{L}{\pi(R_e + R_u)}. \quad (22-36)$$



k - спиральная заводная.

Расстояние между витками

$$a = \frac{\pi(R_n^2 - R_o^2)}{L}. \quad (22-37)$$

Пример. Момент $M = 12,2$ Гмм при угле $\frac{\pi}{2}$. Материал — фосфористая бронза $E = 11 \cdot 10^8 \frac{\kappa\Gamma}{\text{мм}^2}$.

$$R_n = 3 \frac{\kappa\Gamma}{\text{мм}^2}; \text{ задаемся } L = 400 \text{ мм};$$

$$h = \frac{2 \cdot 400 \cdot 3}{0,5 \cdot \pi \cdot 11 \cdot 10^8} = 0,139; \text{ принимаем } h = 0,15 \text{ мм};$$

$$b = \frac{6 \cdot 0,012}{0,15^2 \cdot 3} \approx 1 \text{ мм}; R_n = 11,5 \text{ мм}; R_o = 2,5 \text{ мм};$$

$$\omega = \frac{L}{\pi(R_n + R_o)} = 9,1; a = 1 \text{ мм}.$$

Спиральная заводная пружина (фиг. 22,25 б)

$$h = \frac{2R}{\sigma}; \frac{R}{r} = 3;$$

$$\text{развернутая длина } L = \pi\omega_1(R + r_1) + 2\pi r. \quad (22-38)$$

$$M = \frac{EJ_2}{L} \varphi; \quad (22-39)$$

$$J_2 = \frac{\sigma h^3}{12}; \varphi = 2\pi\omega_s.$$

$$\omega_{s \text{ макс}} = \omega_2 - \omega_{св}; \omega_{s \text{ мин}} = \omega_1 - \omega_{св};$$

$\omega_{св}$ на 2—3 витка меньше ω_1 .

$$M_{\text{макс}} = \frac{EJ_2}{L} 2\pi\omega_{s \text{ макс}}; M_{\text{мин}} = \frac{EJ_2}{L} 2\pi\omega_{s \text{ мин}}.$$

Учитывая трение между витками, уменьшают значение M на 20—30%.

ГЛАВА 23

КОНТАКТНЫЕ СИСТЕМЫ

Контактной системой называется элемент реле, осуществляющий операции замыкания или размыкания контактов. Контактная система состоит из подвижных частей — контактных рычагов, пружин, и неподвижных — контактных стоек, пружин, а также самих контактов, скрепленных с этими частями.

23-1. ОСНОВНЫЕ ПАРАМЕТРЫ КОНТАКТНЫХ СИСТЕМ

Различают эксплуатационные параметры — разрывную способность P , предельно допустимые значения напряжения U_{np} тока I_{np} , частоту включений, и конструктивные — ход кон-

1*

тактов x , давление Q . Надежная работа контактной системы при заданных эксплуатационных параметрах зависит от принятых конструктивных параметров, формы и материала контактов.

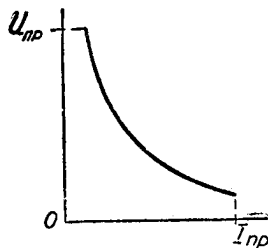
При проектировании контактной системы следует учитывать не только процесс размыкания, но и процесс включения. При неудачно выполненной конструкции в процессе включения наблюдаются явления отскакивания контактов, вибрации пружин и приваривания контактов.

Разрывная способность контактов определяется условием надежного гашения дуги, возникающей в процессе размыкания контактов. Разрывную способность характеризуют условно

$$P = UI,$$

где U — напряжение сети, I — ток в цепи до момента размыкания.

Для контактной системы, обладающей определенной разрывной способностью P , имеет место гиперболическая зависимость $U = f(I)$ (фиг. 23, 1). Гипербола эта имеет ограничения, определяемые значениями: U_{np} — предельно допустимого для данной контактной системы напряжения сети,



Фиг. 23, 1. Характеристика контактной системы.

I_{np} — предельно допустимого из условий нагрева контактов значения тока. Значение P в основном зависит от раствора контактов x : гашение дуги происходит в момент, когда дуга растягивается до размеров, при которых приложенное к контактам напряжение становится недостаточным для ее поддержания. Значение x определяется различными факторами: родом тока, индуктивностью нагрузки, напряжением сети.

При разрыве цепи постоянного тока необходимо растянуть дугу до величины, при которой напряжение сети будет меньше напряжения погасания дуги U_n . При переменном токе, вследствие погасания дуги при переходе тока через нуль, напряжение сети должно быть меньше напряжения зажигания дуги U_z при определенной ее длине.

Напряжение U_z значительно выше U_n . Поэтому для данной контактной системы значение P цепи переменного тока принимается в несколько раз большим, чем для постоянного (см. ниже). Большое значение при разрыве дуги имеет индуктивность нагрузки.

Для цепей постоянного тока индуктивность нагрузки условно оценивают значением $\cos \varphi$, которую экспериментально определяют включением катушки на переменный ток 50 гц. По опытным данным, магнитные системы контакторов и реле управления постоянного тока обладают индуктивностью, характеризующейся условно $\cos \varphi = 0,1 \div 0,15$.

В литературе приводятся эмпирические зависимости $U_d = f(I, l_d)$,

где l_0 — длина дуги. Однако в этих зависимостях не учтен ряд факторов, влияющих на разрывную способность контактов. К настоящему времени не проведены также систематические экспериментальные работы, на основании которых можно было бы установить данные для расчета.

Ниже приводятся для иллюстрации опытные данные по отдельным конструкциям контактных систем.

1) Раствор контактов $x = 0,5$ мм; размыкание цепи постоянного тока 220 в; 15 вт; $\cos \varphi = 0,1$.

2) Контакты кодовых реле $x = 0,6 \div 1,2$ мм;
материал — серебро;

предельное напряжение $U_{np} = 220$ в;

предельный ток разрыва $I_{np} = 3$ а;

разрывная способность — постоянный ток — 80 вт;
переменный ток — 400 вт.

Указанные данные относятся к случаям работы реле с малой частотой. При работе реле с большой частотой рекомендуется уменьшить значение P до 50—70% приведенных данных. Для улучшения работы контактов применяется искрогасительный контур: $C = 0,5$ мкф, $r = 5$ ом. Значительное улучшение работы контактной системы получается при замене материала контактов: вместо серебра ставят металлокерамические.

3) Контакты реле управления, исполняемые в виде контактных мостов с двумя последовательно включаемыми разрывами. Ход контактов — 3—4 мм. Давление в контактах — порядка 100—125 Г.

Данные разрывной способности при индуктивной нагрузке.

Постоянный ток: 110 в — 2,5 а;

220 в — 1 а;

500 в — 0,5 а.

Переменный ток: 220 в — 20 а;

380 в — 20 а.

Минимальное значение раствора контактов. При коммутации контактами реле небольших значений токов — порядка десятых долей ампер — можно принимать небольшие значения раствора контактов. При проектировании быстродействующих и высокочувствительных реле доводят величину x до минимума. Так, в телефонных поляризованных реле значение x принимают порядка (0,05 — 0,1) мм. При установлении значения x следует учитывать возможные коммутационные перенапряжения в цепи постоянного тока. При отключении мощных электромагнитов возникающее при этом напряжение $E = -L \frac{di}{dt}$ достигает высоких значений. В таких цепях наблюдаются случаи пробоя контактного промежутка и связанного с этим неправильного срабатывания.

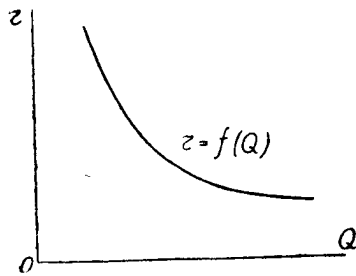
Прочность изоляции каждой электрической установки проверяется испытательным напряжением, в несколько раз большим,

чем рабочее. Ниже приводятся опытные данные для установления $x_{мин}$ в зависимости от величины испытательного напряжения.

$U_1, кВ$ 0,8 1,2 1,9 2

$x_{мин}, мм$ 0,125 0,25 0,5 0,6

Давление в контактах. Нагрев контактов зависит от величины джоулевых потерь в переходном сопротивлении r . Допускаемая по ГОСТ предельная температура превышения 70° . Переходное сопротивление r зависит от материала контактов, состояния их



Фиг. 23, 2. Зависимость переходного сопротивления от давления.

поверхности и от давления в контактах. На фигуре 23,2 приведена кривая, показывающая зависимость $r = f(Q)$ для контактов реле завода «Красная заря».

В среднем, для контактов, коммутирующих токи порядка до $2 \div 3 a$, принимается давление $25 \div 30 Г$. В высококочувствительных реле это давление снижают до $1 \div 2 Г$, в реле управления с контактами, коммутирующими токи порядка $10 a$ — повышают до $100 \div 150 Г$.

При коммутировании токов порядка $1 a$ и выше необходимо выполнять контактную систему так, чтобы в момент соприкосновения контактов давление получалось определенной величины, а не нарастало от нуля. Определяется это условие необходимостью устранения приваривания контактов, которое может получиться при $Q_n \cong 0$ и большом значении r .

В кодовых реле это условие достигается установкой пружин неподвижных контактов с предварительным напряжением (фиг. 23, 7), в устройствах мостикового типа, пружина n создает на мостике предварительную силу нажатия Q_0 (фиг. 23, 19). Обычно Q_0 принимают равным приблизительно половине конечного давления в контактах.

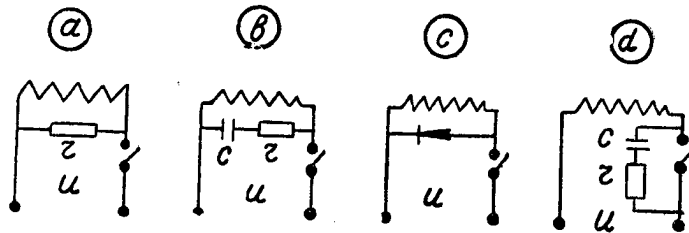
23-2. СПОСОБЫ УМЕНЬШЕНИЯ ИСКРООБРАЗОВАНИЯ

Дуга, образующаяся на контактах, приводит к их порче, поэтому, в случаях, когда контактам приходится разрывать большие мощности и при большой частоте срабатываний применяют искусственные средства для уменьшения размеров дуги. Применяется два основных метода: шунтирование индуктивности разрываемой цепи или шунтирование контактов. В обоих случаях магнитная энергия индуктивной цепи расходуется в основном не в дуге, а в шунтирующем устройстве.

Шунтирование индуктивности цепи (фиг. 23, 3а, б, с). Такие схемы применяются относительно редко, так как шунтирование

катушки электрическим сопротивлением, емкостью или выпрямителем изменяет время t_n отпуска якоря.

Шунтирование контактов (фиг. 23, 3 *d*). Наиболее часто применяется схема с емкостью, последовательно с которой ставится сопротивление r ; значение C принимают порядка $0,5 \div 2$ мкф.



Фиг. 23, 3. Схемы.

Для определения минимального значения r пользуются формулой Крюгера

$$r = \frac{U^2}{a},$$

где U — напряжение на конденсаторе, a — постоянная, зависящая от материала контакта. Для серебра $a = 140$; платины — $a = 180 \div 200$. Для шунтирования контактов кодовых реле применяют обычно: $C = 0,5$ мкф.

23-3. ВИБРАЦИИ КОНТАКТОВ

При оценке работы контактных систем большое внимание уделяется также процессу замыкания контактов. В процессе замыкания имеет место явление удара, определяемое наличием кинетической энергии подвижных контактных частей. При ударе может получиться отскакивание подвижного контакта с разрывом цепи или вибрация пружин подвижного контакта.

Исследование явлений, происходящих в момент удара в контактной системе, производится осциллографом (см. схему на фиг. 23, 4). Kontakтами реле замыкают цепь постоянного тока низкого напряжения — порядка 4 в. При включении цепи катушки срабатывает реле.

На фигуре 23, 5 показаны примерные осциллограммы тока, протекающего в цепи контактов в момент их включения:

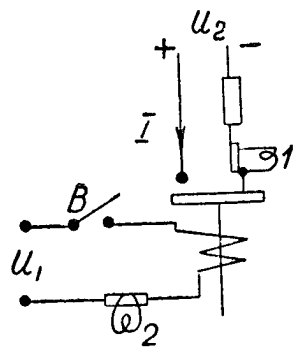
a, b — процесс происходит без отскакиваний и ударов;

c — наблюдается явление отскакивания;

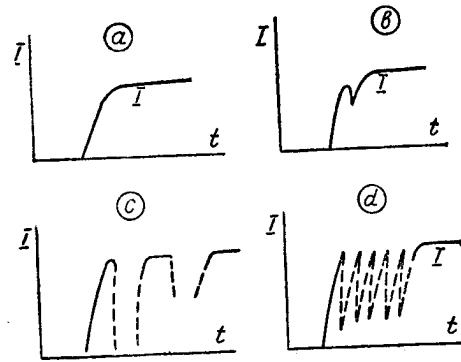
d — наблюдается явление вибраций контактной пружины.

Явления отскакивания контактов и вибраций пружин весьма нежелательны, так как приводят к удлинению процесса замыкания цепи, а в ряде случаев к значительному подгоранию контактов.

В литературе приводятся аналитические соотношения для определения условий устранения указанных явлений. Практического применения эти соотношения не получили, так как при выводе этих соотношений не учтены условия работы контактной системы, определяющие в основном возникновение явления отскакивания и вибраций. Поэтому при проектировании контактных



Фиг. 23, 4. Схема для исследования работы контактной системы: 1, 2 — шлейфы; В — выключатель.

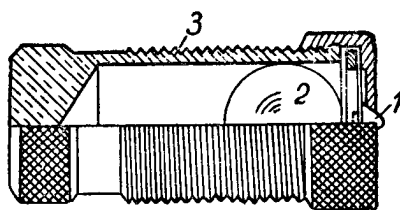


Фиг. 23, 5. Осциллограммы работы контактной системы в момент включения.

систем пользуются общими практическими соображениями, которые приводятся ниже.

Способы устранения явлений отскакивания и вибраций.

1) Применение успокоителей, в виде демпфирующих устройств, связанных с подвижной частью контактной системы. Такой способ используется в индукционных дисковых реле: в системе устанавливают постоянные магниты, поле которых индуцирует в диске при его вращении вихревые токи.



Фиг. 23, 6. Устройство неподвижного контакта.

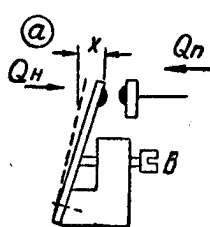
Успокоители не должны применяться в быстродействующих реле, так как при наличии успокоителей уменьшается скорость срабатывания.

2) Контактная система исполняется так, что кинетическая энергия подвижной части поглощается неподвижной.

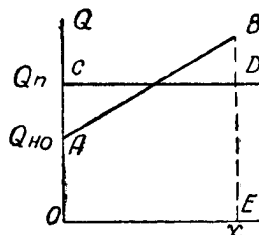
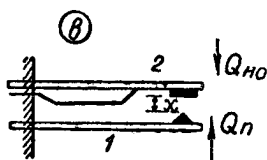
На фигуре 23, 6 приведен пример такого устройства. Неподвижный контакт исполняется в виде полой трубки (3), установленной с небольшим уклоном так, чтобы шарик (2) располагался всегда у мембраны. При ударе подвижного контакта вся энергия движения передается через гибкие мембраны шарик, который отскакивает с определенной скоростью, а затем медленно возвращается в исходное положение. Размеры шарика подбираются

экспериментально. В выполненных быстродействующих индукционных реле лучшее решение получилось при $\varnothing = 9 \text{ мм}$.

3) Неподвижный контакт выполняется в виде пружины с предварительным натяжением (фиг. 23, 7). Диаграмма работы контактной системы приведена на фигуре 23, 8: подвижный контакт подходит с силой Q_n , совместное перемещение обоих контактов x ; неподвижный контакт имеет силу $Q_{но}$ предварительного натяжения,

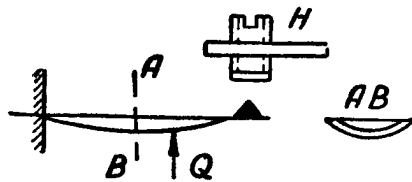


Фиг. 23, 7. Исполнения контактных систем.

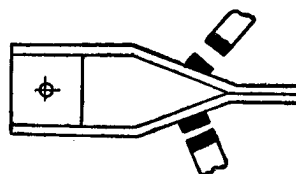


Фиг. 23, 8. Диаграмма работы.

жесткость пружины характеризуется линией AB . Площадка $ocDE$ соответствует работе $Q_n x$, которая эквивалентна кинетической энергии подвижного контакта. Для поглощения этой энергии на пути x необходимо, чтобы площадка $OABE = ocDE$.



Фиг. 23, 9. Исполнение контактной системы.



Фиг. 23, 10. Исполнение контактной системы поляризованных реле.

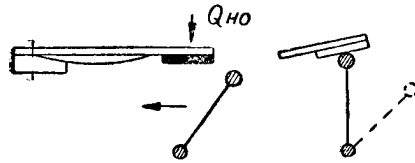
Таким образом, при определенной жесткости пружины неподвижного контакта, требуется наличие предварительного натяжения этой пружины $Q_{но}$. Выполняют контакты так, чтобы можно было регулировать значение $Q_{но}$ винтом (фиг. 23, 7a) или подгибом пружины (фиг. 23, 7b).

4) Неподвижный контакт жесткий, подвижный укреплен на пружине малой жесткости (фиг. 23, 9). Это исполнение встречается в высокочувствительных быстродействующих реле. В таких реле подвижная контактная пружина выполняется толщиной порядка $0,15 \text{ мм}$, и в процессе замыкания контактов такая пружина вибрирует. Для устранения вибраций необходимо ожесточить пружину. Рациональное решение получено при выполнении пружины с выдавкой в средней ее части, как это показано на рисунке.

5) Исполнение контактов двухстороннего действия (фиг. 23, 10). Такая конструкция выполняется в поляризованных реле. Две

пружины подвижного контакта заканчиваются в виде плоских площадок, прижимающихся друг к другу. В процессе включения одного из контактов кинетическая энергия системы поглощается работой сил трения в площадках при перемещении одной из них по отношению к другой.

6) Подвижный контакт имеет вращательное движение (фиг. 23, 11). В таком исполнении подвижный контакт с некоторым трением заходит в зону неподвижного. Возникающие в момент замыкания силы удара стремятся отбросить подвижный контакт, но разрыва цепи не происходит, так как напряженный неподвижный контакт перемещается совместно с подвижным.



Фиг. 23, 11. Исполнение контактной системы.

23-4. КОНТАКТЫ

При проектировании контактной системы учитывается материал, форма и размеры контактов.

Материал контактов. Срок службы и надежность работы контактной системы в значительной степени зависят от материала, из которого изготовлен контакт. Материал выбирается в зависимости от принятого давления в контактах, а также от условий работы по частоте включений и разрывной мощности. Большое внимание уделяется износоустойчивости контактов. Разрушение контактов определяется следующими факторами:

Механический износ получается вследствие ударов подвижного контакта о неподвижный; при частой работе контактов и при большом давлении применяют материалы большой твердости и механической прочности — вольфрам, платину, иридий.

Химический износ — коррозия — получается вследствие окисления материала при высоких температурах и отпадания пленки в момент удара. Для уменьшения коррозии контактов применяют платину и ее сплавы.

Электрический износ — эрозия — выражается в переносе металла с одного контакта на другой с образованием бугров, кратеров и мостов между контактами. Минимальный электрический износ получается при материалах с высокой температурой плавления — вольфрам, металлокерамические контакты.

Материалы для контактов должны быть твердыми и механически прочными, устойчивыми против эрозии и коррозии, должны обладать высокой тепло- и электропроводностью, легко обрабатываться. В современной практике применяется большое число материалов, в той или иной степени удовлетворяющих перечисленным требованиям. Наиболее ходкими из них являются: серебро, платина и ее сплавы, вольфрамы и композиции из двух или нескольких металлов. Ниже приводится краткая характеристика перечисленных материалов.

Серебро — наиболее дешевый из применяемых материалов, легко обрабатывается, обладает хорошими свойствами в отношении электро- и теплопроводности. К недостаткам серебра относятся малая твердость, склонность к образованию дуги и, как следствие, недостаточная стойкость в отношении эррозии; серебряные контакты портятся в атмосфере сероводорода, при котором образуется сульфидная пленка, резко повышающая переходное сопротивление контактов. Серебро применяется для контактов с легкими условиями работы и при давлениях порядка 5 Г и выше. Не рекомендуется применять серебро в конструкциях, имеющих эбонит, вулканизированную резину, провода с резиновой изоляцией.

Платина — в воздухе не корродирует, обладает повышенной по сравнению с серебром сопротивляемостью эррозии. Из-за недостаточной твердости редко применяется в чистом виде. Область применения — легкие условия работы, при давлениях порядка 1 — 5 Г.

Платина-иридий. Этот сплав также не окисляется. Хорошо противостоит дугообразованию, обладает высокой твердостью и поддается, при содержании иридия до 30%, механической обработке. Сплав платина-иридий применяется для маломощных контактов, а также для мощных при большой частоте срабатываний.

Вольфрам весьма тверд, не сваривается и не изнашивается механически, но имеет окисную пленку, вследствие чего требует повышенных давлений в контактах. Обычно чистый вольфрам (99,5 — 99,8%) применяется в конструкциях с повышенными давлениями в контактах (не менее 100 Г) при тяжелых условиях работы — большой мощности дуги и частоты срабатываний.

Композиции. В последние годы широкое применение получили композиции, чаще всего двухфазные, например, серебро — молибден, серебро — графит и другие. Для контактов реле чаще всего применяют композицию серебро — окись кадмия (ГОСТ 3884—47), в которой роль токоведущей составляющей играет основная масса — серебро, а роль тугоплавкой — окись кадмия.

Вследствие несплаваемости составляющих композиция изготавливается методами металлокерамики: порошок, смешанный в заданном соотношении, прессуется в стальных прессформах в виде контактов требуемой формы, полученная деталь проходит высокотемпературный нагрев для спекания составляющих, затем калибруется дополнительным обжимом в прессформах и отжигается для снятия наклепа. Область применения композиции серебро — окись кадмия — контактные системы с повышенной разрывной мощностью. Минимальное требуемое давление на контактах — порядка 35—40 Г.

В таблице 33 приводятся основные данные наиболее ходких материалов для контактов реле.

Таблица 33

Физические свойства некоторых материалов для контактов реле

Материал	Содержание	Твердость по Бриггелю	Температура плавления, °С	Теплопроводность, $\frac{вт}{см \cdot °С}$	Удельное сопротивление, $ом \cdot см \cdot 10^8$	Удельный вес
Серебро	—	25	961	4,16	1,58	10,5
Платина	—	50	1773	0,7	10,5	21,4
Сплав платина — иридий	Ir 10%	150	1780	0,31	24,5	21,6
	20%	230	1815	0,175	30	21,7
	25%	270	1840		33,5	—
	30%	310	1890		35	21,8
Вольфрам	—	350	3400	1,68	5,32	19,1
Композиция серебро — окись кадмия	OK—12(CdO—12%)	45			2,5	9,6
	OK—15(CdO—15%)	55			3	9,4

В практике приняты следующие положения для выбора материала контактов:

- 1) малые давления — 1—3 Г — платина;
- 2) давления 5—100 Г, малая частота срабатываний, — серебро;
- 3) давление 30—100 Г, большая частота срабатываний, — металлокерамические контакты.



Фиг. 23, 12. Формы контактов.

4) давление больше 100 Г, большая частота срабатывания, — вольфрам.

Форма контактов. При давлениях до 30—40 Г лучшее решение получается при применении плоско-острых пар контактов (фиг. 23, 12): один контакт имеет коническую форму, другой плоскую. При таком исполнении получают повышенные удельные давления в контактах, а также облегчаются условия сборки, так как устраняется необходимость точной установки контактов по общей осевой линии.

В системах с давлениями больше 30—40 Г применяют формы контактов: полушар — плоскость, плоскость — плоскость.

Размеры контактов зависят от величины предельно допустимого тока I_{np} (см. таблицу 34).

Таблица 34

Размеры серебряных контактов

I_{np} , ампер	Диаметр, мм	Толщина, мм
< 1	до 3	1,0
1—5	» 6	1,5
5—10	» 8	2,0
10—20	» 12	3,0

**Размеры металлокерамических контактов
(ОК12 и ОК15) (фиг. 23, 17)**

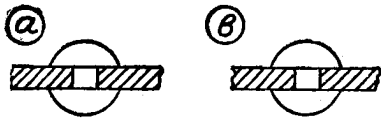
Диаметр, мм	5	6	8	10
Длительно-допустимый ток, ампер	10—15	15—25	25—50	50—75

**Размеры контактов из сплава: платина — иридий
или из вольфрама, $I_{np} = 5 а$**

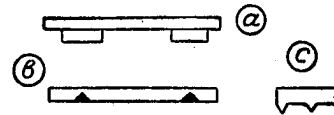
Частота включения	До 8000 в минуту	>8000 в минуту
Диаметр, мм .	4	5
Толщина, мм .	1,5—2	2

Вопросы технологии. Контакты прикрепляются к другим деталям контактной системы расклепкой, плотной посадкой и приваркой.

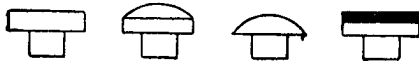
а) При креплении серебряных контактов на контактных пружинах толщиной 0,15 — 0,4 мм применяется расклепка (фиг. 23, 13).



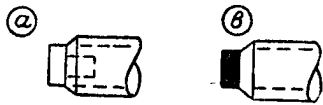
Фиг. 23, 13. Крепление контактов на пружинах.



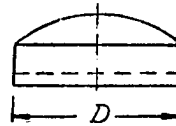
Фиг. 23, 16. Контактный мостик:
а — с приваренным контактом;
б, с — до приварки.



Фиг. 23, 14. Формы заклепок.



Фиг. 23, 15. Крепление контактов на винтах:
а — вдавливанием; б — приваркой.



Фиг. 23, 17. Исполнение металлокерамического контакта ОК12 и ОК15.

б) В ряде случаев контакт выполняется в виде заклепки (фиг. 23, 14) или целиком из материала для контакта, например из серебра, или с основой из латуни с приваркой серебра, платины и других материалов. Заклепка закрепляется в контактном винте или посадкой с натягом, или приваркой в сварочном агрегате (фиг. 23, 15).

в) Прикрепление контактов к подвижным контактным пластинкам (фиг. 23, 16) осуществляется сваркой.

г) Металлокерамические контакты исполняются чаще всего в виде пластинок (фиг. 23, 17) с основанием из латуни. Такая пластинка приваривается серебряным припоем ПСр 45, флюс-бура, нагрев в сварочном агрегате.

23-5. ИСПОЛНЕНИЯ КОНТАКТНЫХ СИСТЕМ

Требования к контактным системам. Для обеспечения надежной работы контактной системы необходимо выполнять следующие основные условия.

1) Должно быть обеспечено определенное давление Q_k . Переходное сопротивление в момент соприкосновения контактов должно быть незначительным. Поэтому контактные системы необходимо выполнять так, чтобы в момент соприкосновения контактов имело место начальное давление Q_n определенного значения. Обычно принимают $Q_n = 0,4 \div 0,5 Q_k$.

2) Должен быть обеспечен достаточный ход контактов. Хорошее решение получается при исполнении контактов с двойным разрывом (фиг. 23, 19).

3) Замыкание и размыкание контактов должно происходить с большой скоростью. При замыкании контактов не должно иметь место явление отскакивания или вибраций.

4) Температура превышения частей контактной системы не должна быть больше допустимой (70°).

Возврат контактной системы в исходное положение. Различают:

Системы с автовозвратом, когда контакты автоматически возвращаются в исходное положение при возврате подвижной части реле. Такое исполнение принято для основных видов реле.

Системы с ручным возвратом — возврат контактной системы производится от руки. Такое исполнение принято для тепловых реле, для специальных конструкций реле автоматики.

Системы с дистанционным возвратом — возврат осуществляется посредством добавочного электромагнита, включаемого на расстоянии — дистанции.

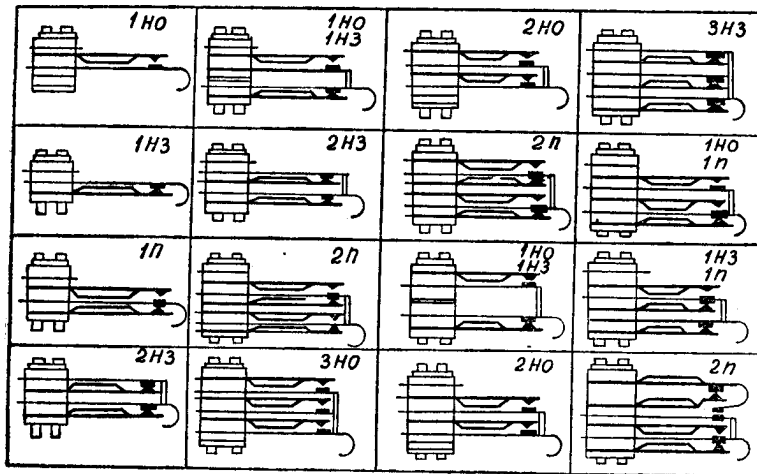
Формы исполнения контактных систем весьма разнообразны и определяются как конструкцией реле, так и условиями монтажа и эксплуатации.

1) *Контактные системы кодовых реле* являются наиболее распространенной формой маломощных электромагнитных реле. Исполняются эти системы в виде отдельных контактных групп из н. о., н. з. и п-контактов (фиг. 23, 18). Каждое реле может быть снабжено несколькими контактными группами (см. фигуру 1, 28, где установлено три контактных группы).

2) *Контактные системы реле управления* исполняются, как правило, с мостиковыми контактами (фиг. 23, 19): рычаг P , связанный с подвижной частью реле, имеет гибкую связь с мостиком — посредством пружины n . Последняя давит с определенной силой Q_n — «встроенное сжатие» — на пластинку при разомкнутом контакте, чем обеспечивается начальное давление в контактах в момент их соприкосновения.

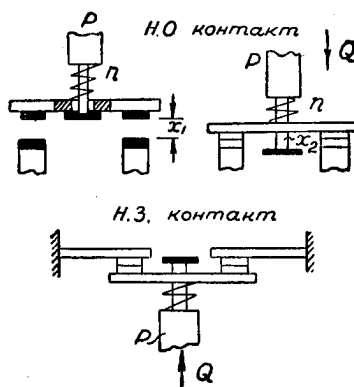
Ход контактного мостика — x_1 , ход рычага P должен быть равен $x_1 + x_2$. Последняя величина — x_2 — называется провалом. Величина провала принимается в зависимости от жесткости f

пружины n . Если эта пружина встроена с предварительной деформацией x_3 , то конечное давление на контакте $Q_k = j(x_2 + x_3)$.



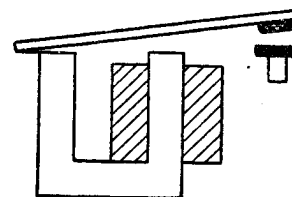
Фиг. 23, 18. Контактные группы кодовых реле.

Средние данные: $Q_n = 50 + 75 \text{ Г}$, $Q_k = 100 - 150 \text{ Г}$. Наши заводы выпускают комплекты контактные системы — контактные блоки — состоящие из ряда н. о. и н. з.-контактов.



Фиг. 23, 19. Мостиковые контактные системы.

3) Контактные системы быстродействующих реле исполняют в ряде случаев без добавочных контактных пружин: подвижный контакт укрепляется непосредственно на якоре (фиг. 23, 20).



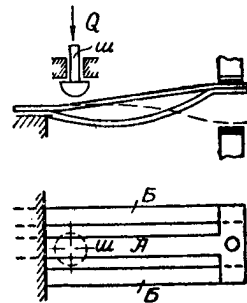
Фиг. 23, 20. Исполнение контактной системы.

4) Контактные системы индукционных реле весьма разнообразны. Примерные исполнения приведены на фигурах 23, 7а, 23, 11. Примерное исполнение контактной системы поляризованного реле риведено на фигуре 23, 10.

5) *Прыгающие контакты* — особая группа контактов, исполненных так, что при определенном небольшом ходе нажимающего на контакт штифта получается опрокидывание контакта, перемещающегося на большую величину. Такие контактные системы применяются в тех случаях, когда ход невелик $0,1 \div 0,2$ мм и требуется большая величина раствора контактов $3 \div 5$ мм, а также тогда, когда подвижная часть реле перемещается с небольшой скоростью. Следует учитывать, что

требуемая сила нажатия на штифт относительно велика — порядка $150 \div 200$ Г.

На фигуре 23, 21 приведен эскиз исполнения прыгающего контакта. Средняя часть контактной пружины выполнена так, что она длиннее крайних. Благодаря этому вся пружина изгибается и контакт прижимается кверху. При нажатии на пружину в средней ее части штифтом *ш*, пружина вначале деформируется незначительно, а затем, при увеличении хода штифта, резко опрокидывается (см. пунктир). Имеются исполнения прыгающих контактных систем, встроенных в герметический изоляционный корпус весьма небольших размеров. Такое исполнение называется микроключателем и выпускается нашими заводами под маркой МП. Применяется в ряде конструкций реле и устройств автоматики.

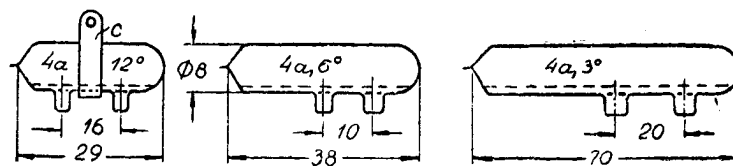


Фиг. 23, 21. Эскиз прыгающего контакта.

6) *Ртутные контакты*, в которых замыкание или размыкание цепи осуществляется ртутью, залитой в стеклянную эвакуированную колбу. Колба укрепляется на металлической скобе, связанной

с подвижной частью реле. Требуемый для срабатывания контактов угол наклона колбы принимается порядка $3 \div 12^\circ$ и зависит от размеров колбы (фиг. 23, 22). Ртутные контакты исполняются для сил токов 4, 6, 10, 15, 20, 30 а, до 220 вольт, срок службы, определяемый числом включений, достигает 1 — 2 миллионов.

Ртутные контакты обычно применяются в реле и устройствах автоматики, устанавливаемых в химических производствах, где имеется опасность возникновения взрыва от искры или дуги контактов открытого исполнения.



Фиг. 23, 22. Эскизы ртутных контактов, 4 а, 220 в.

Ртутные контакты обычно применяются в реле и устройствах автоматики, устанавливаемых в химических производствах, где имеется опасность возникновения взрыва от искры или дуги контактов открытого исполнения.

23-6. ПОСТРОЕНИЕ МЕХАНИЧЕСКИХ ХАРАКТЕРИСТИК

При проектировании и расчете реле необходимо строить механические характеристики контактных систем: $Q = f(\delta)$ или $M = \varphi(\theta)$. В обычных исполнениях реле точка приложения сил реакции пружин не совпадает с точкой приложения электромагнитных сил. Поэтому при проведении расчетов данные сил реакции приводят к точке приложения на якоре электромагнитной силы. Ниже приводятся примеры построения характеристик контактных систем с нормально открытым (н. о.), нормально закрытым (н. з.) и переключающимся (п) контактами.

В таблице 35 приведены формулы для расчетов плоских пружин.

Таблица 35

Вычисление сил реакции и деформаций пружины при разных положениях внешней приложенной силы

Внешняя сила приложена в конце пружины	Внешняя сила приложена в средней части пружины
$f_1 = \frac{Q_1 l_1^3}{3EJ}$	$f_1 = \frac{Q_2 l_2^2}{EJ} \left(\frac{l_1}{2} - \frac{l_2}{6} \right)$
$f_2 = \frac{Q_1 l_2^2}{EJ} \left(\frac{l_1}{2} - \frac{l_2}{6} \right)$	$f_2 = \frac{Q_2 l_2^3}{3EJ}$
$K_1 = \frac{f_2}{f_1} = \frac{l_2^2}{l_1^2} \left(1,5 - 0,5 \frac{l_2}{l_1} \right)$	$K_2 = \frac{f_1}{f_2} = 1,5 \frac{l_1}{l_2} - 0,5$
$Q_1 = \frac{3EJ f_1}{l_1^3}$	$Q_1 = Q_2 / K_2$
$Q_2 = K_2 Q_1$	$Q_2 = \frac{3EJ f_2}{l_2^3}$

1) Н. о. контакт (фиг. 23, 23).

Заданные величины: пружины — из фосфористой бронзы,

$$\begin{aligned} l_1 &= 40 \text{ мм}; & l_2 &= 34 \text{ мм}; & b_2 &= 6 \text{ мм}; \\ b_1 &= 6 \text{ мм}; & h_1 &= 0,35 \text{ мм}; & h_2 &= 0,4 \text{ мм}. \end{aligned}$$

Ход контактов $f_{21} = 1 \text{ мм}$; начальное давление на якорь у изоляционной планки 3: $Q_{10} = 10 \text{ Г}$. Конечное давление в контактах — $Q_{21} = 25 \text{ Г}$, начальное нажатие пружины 2, отнесенное к l_2 — $Q_{20} = 12 \text{ Г}$, конечное значение рабочего зазора у якоря $\delta_3 = 0,1 \text{ мм}$, отношение плеч l_3 к l_4 равно $k_3 = 1,68$.

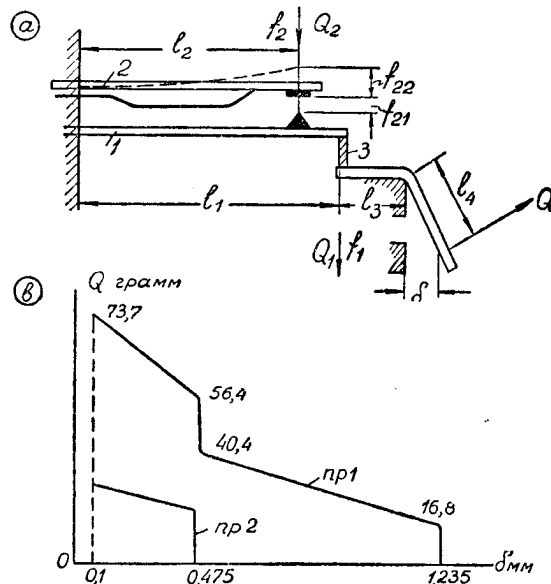
Расчет.

Пружины: $J_1 = \frac{b_1 h_1^3}{12} = 2,15 \cdot 10^{-2} \text{ мм}^4$; $J_2 = 3,2 \cdot 10^{-2} \text{ мм}^4$.

$E = 11 \cdot 10^6 \frac{\Gamma}{\text{мм}^2}$; $f_1 = \frac{3EJ_1}{l_1^3} = 11,1 \frac{\Gamma}{\text{мм}}$; $f_2 = \frac{3EJ_2}{l_2^3} = 26,4 \frac{\Gamma}{\text{мм}}$.

Коэффициенты (см. таблицу 35):

$$k_1 = \frac{f_2}{f_1} = 0,785; \quad k_2 = \frac{f_1}{f_2} = 1,26; \quad k_3 = 1,68 \text{ (задано).}$$



Фиг. 23. Нормально открытый контакт кодового реле.

Перемещения: пружины (1) до замыкания контактов $f_{21} = 1 \text{ мм}$;

$$\text{якоря } \delta_1 = \frac{f_{11}}{k_3} = \frac{f_{21}}{k_1 \cdot k_3} = 0,76 \text{ мм.}$$

Дополнительное перемещение пружины 1 до получения давления в контактах, равного 25 Г;

$$f_{22} = \frac{Q_{21} - Q_{20}}{f_2} = \frac{25 - 12}{26,4} = 0,493.$$

$$\delta_2 = \frac{f_{22}}{k_1 k_3} = 0,375 \text{ мм.}$$

Силы на якоре: Начальное значение $Q_0 = Q_{10} \cdot k_3 = 10 \cdot 1,68 = 16,8 \text{ Г}$, при перемещении пружины 1 до момента замыкания контактов — f_{11} (при ходе контактов f_{21})

$$Q_1 = Q_0 + Q_{11} k_3 = Q_0 + f_{11} j_1 k_3 = 16,8 + \frac{f_{21} j_1 k_3}{k_1} = 40,4 \text{ Г.}$$

При замыкании контактов добавляется сила предварительного нажатия пружины 2

$$Q_1' = Q_1 + \frac{Q_{20}}{k_2} k_3 = 40,4 + \frac{12 \cdot 1,68}{1,26} = 56,4 \text{ Г.}$$

Сила на якоре в конечном — притяннутом — положении

$$Q_2 = Q_1' + \frac{f_{22} l_2}{k_3} k_3 = 73,7 \text{ Г.}$$

Диаграмма сил, отнесенных к якору, приведена на фигуре 23, 23 б.

2) Н. з. контакт (фиг. 23, 24).

Заданные величины: пружины — фосфористая бронза;

- $l_1 = 40 \text{ мм};$
- $l_2 = 34 \text{ мм};$
- $b_1 = 6 \text{ мм};$
- $b_2 = 6 \text{ мм};$
- $h_1 = 0,35 \text{ мм};$
- $h_2 = 0,4 \text{ мм}.$

Ход контактов — разомкнутом их состоянии — $f_2 = 1 \text{ мм}$; начальное давление на якоре — у изоляционной планки 3 — 10 Г ; нажатие в контактах $Q_{21} = 25 \text{ Г}$, начальное нажатие пружины (2) $Q_{20} = 12 \text{ Г}$; конечное значение рабочего зазора $\delta_3 = 0,1 \text{ мм}$.

Расчет.

Пружины: $J_1 = 2,15 \cdot 10^{-2} \text{ мм}^4$; $J_2 = 4,55 \cdot 10^{-2} \text{ мм}^4$; $E = 11 \cdot 10^8 \frac{\text{Г}}{\text{мм}^2}$;

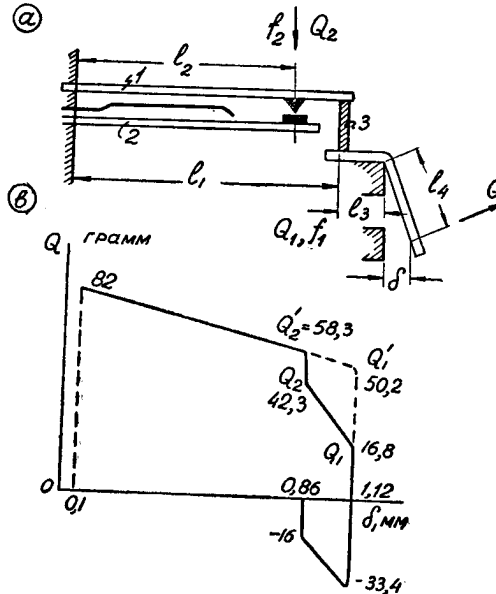
$$f_1 = 11,1 \frac{\text{Г}}{\text{мм}}; \quad f_2 = \frac{3EJ_2}{l_2^3} = 38,3 \frac{\text{Г}}{\text{мм}}.$$

Коэффициенты (см. таблицу):

$$k_1 = \frac{f_2}{f_1} = 0,785; \quad k_2 = \frac{f_1}{f_2} = 1,26; \quad k_3 = 1,68 \text{ (задано)}.$$

Перемещения: пружины 1 до момента размыкания контактов определяются деформацией пружины 2 под действием силы Q_{21} ; при этом надо учесть начальное нажатие на упор пружины 2 — Q_{20} :

$$f_{21} = \frac{Q_{21} - Q_{20}}{f_2} = \frac{25 - 12}{38,3} = 0,34.$$



Фиг. 23, 24. Нормально закрытый контакт кодового реле.

Изоляционной планки 3: $f_{11} = \frac{f_{21}}{k_1} = 0,434$; якоря $\delta_1 = \frac{f_{11}}{k_3} = 0,26$.
 Ход контактов: $f_2 = 1$ мм (задано); $\delta_2 = \frac{f_2}{k_1 \cdot k_3} = 0,76$; полный ход якоря $\delta = \delta_0 + \delta_1 + \delta_2 = 0,1 + 0,26 + 0,76 = 1,12$.

Силы. Пружина 1 изогнута так, чтобы создавались силы: на якоре у планки 3 $Q_{11} = 10$ Г; на контактах $Q_{21} = 25$ Г.

Полное значение этой силы, отнесенное к якору у рабочего зазора:

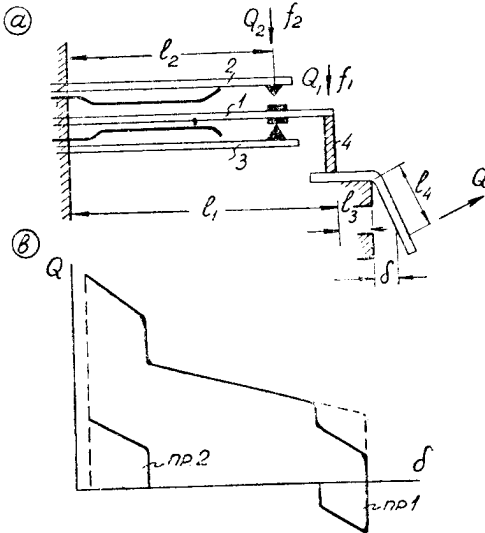
$$Q'_1 = \left(\frac{Q_{21}}{k_2} + Q_{11} \right) k_3 = 50,2 \text{ Г.}$$

Учитывая реакцию пружины 2, получаем результирующую силу на якоре

$$Q_1 = Q'_1 - \frac{Q_{21}}{k_2} k_3 = 50,2 - \frac{25}{1,26} 1,68 = 16,8 \text{ Г.}$$

При перемещении контактных пружин значение силы реакции пружины 2 падает и полностью снимается в момент расхождения контактных пружин (перемещение f_{21}).

На диаграмме, приведенной на фигуре 23, 24,



Фиг. 23, 25. Переключающийся контакт кодового реле.

значения сил реакции пружины 2 показаны как обратные силам пружины 1. Сила реакции пружины 2 при перемещении f_{21} , отнесенная к якору, равна

$$\frac{Q_{20} k_3}{k_2} = 16 \text{ Г.}$$

Сила на якоре при перемещении f_{21} , чему соответствует перемещение δ_2 , без учета реакции пружины (2)

$$Q'_2 = Q'_1 + j_1 \frac{f_{21}}{k_1} k_3 = 50,2 + 11,1 \frac{0,34}{0,785} 1,68 = 58,3 \text{ Г;}$$

$$Q_2 = Q'_2 - 16 = 42,3 \text{ Г.}$$

Конечное значение силы на якоре

$$Q_3 = Q'_1 + j \frac{f_2 + f_{21}}{k_1} k_3 = 50,2 + 11,1 \frac{1,27 + 0,34}{0,785} 1,68 = 82 \text{ Г.}$$

Диаграмма сил, отнесенных к якору в средней точке против рабочего полюса, приведена на фигуре 23, 23.

3) П-контакт (фиг. 23, 25). Расчеты для построения диаграммы $Q = f(\delta)$ проводятся по аналогии с расчетами для других видов контактов.

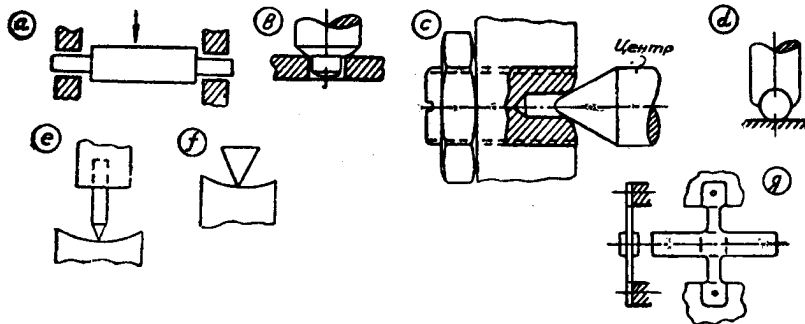
Построение контактных систем, приведенное на фигурах 23, 23, 23, 24 и 23, 25, характерно и для других исполнений.

ГЛАВА 24

НАПРАВЛЯЮЩИЕ ДЛЯ ВРАЩАТЕЛЬНОГО ДВИЖЕНИЯ

24-1. КЛАССИФИКАЦИЯ

Подвижные части реле — якорь, диск, рамка — имеют вращательное движение с ограниченным углом поворота. В некоторых



Фиг. 24, 1. Виды опор реле:
 а, б — цилиндрическая; с — на центрах; d — шаровая; е — на кернах;
 f — на призме; g — растяжка.

видах реле и регуляторов имеются детали, совершающие непрерывное вращательное движение, например зубчатые колеса механизмов, диски демпфирующих элементов и другие. Направляющие для вращательного движения — опоры — деталей реле имеют различные формы исполнений.

Выбор той или иной конструкции опоры определяется условиями конструкции: расположением оси, величиной нагрузки на опору, требуемой точностью направления, моментом трения, стойкостью против износа рабочих поверхностей, стойкостью при вибрациях и ударах, стоимостью изготовления и сборки.

В зависимости от вида трения между основными деталями опоры различают:

- 1) опоры с трением скольжения — цилиндрическая, на центрах, шаровая, на кернах (фиг. 24,1);
- 2) опоры с трением качения — на призме, шарикоподшипниковая;
- 3) опоры с трением упругости — на растяжках.

Общие положения по выбору вида опор

Цилиндрические опоры имеют по сравнению с другими большой момент трения, невысокую точность направления и центрирования, износ опор в процессе работы относительно невелик. Цилиндрические опоры применяются в основном при больших нагрузках на опору и значительных моментах вращения.

Опоры на центрах обладают относительно малым моментом трения, высокой точностью центрирования, допускают регулировку оси в пространстве за счет перемещения одного из подшипников. Малая поверхность соприкосновения деталей обуславливает применение таких опор только для небольших нагрузок и очень малых скоростей вращения. Точность центрирования при особом тщательном исполнении достигает 1—2 м.

Шаровые опоры обладают относительно малым моментом трения. Широко применяются в ряде конструкций реле с вертикальной осью вращения.

Опоры на шпиле имеют весьма малый момент трения. Точность центрирования невелика. Широко применяются в ряде конструкций реле с малой нагрузкой на опору.

Опоры на призме обладают весьма малым моментом трения, износостойкостью при большом числе срабатываний. Широко применяются в электромагнитных реле постоянного тока, где требуются опоры, выдерживающие десятки миллионов срабатываний реле.

Шарикоподшипниковые опоры. До последнего времени шариковые подшипники не применялись в релестроении, что объясняется относительно большими габаритами стандартных подшипников. В настоящее время имеются малогабаритные шариковые подшипники с внутренним диаметром 0,8 мм и менее. Такие подшипники обладают очень малым моментом трения, способностью восприятия больших нагрузок, высокой точностью направления и центрирования, устойчивостью при вибрационных нагрузках. Следует поэтому ожидать внедрения этих подшипников в релестроении. Относительно высокая стоимость шарикоподшипников окупится улучшением качества реле и увеличением срока его службы.

Опоры на растяжках представляют собой металлическую ленту, на которой закреплен якорь реле. Применяются при малых углах поворота якоря, например в поляризованных реле.

Смазка опор. Как правило, в опорах конструкций реле не предусматривается смазка.

24-2. МАТЕРИАЛЫ ДЛЯ ОПОР

Цапфы опор и керны исполняют из сталей ОВС, У8А, У10а закаленных до твердости $R_c = 50 \div 55$, из сталей СТ35 и других. Буксы делают чаще всего из латуни Л59, в отдельных случаях из оловянистой бронзы.

В ряде случаев для уменьшения момента трения выполняя конструкции опор с большими значениями удельных давлени

В таких конструкциях подшипники изготовляют из камней — синтетического агата, сапфира и корунда. Конструкции подшипников из агата и корунда приведены на фигуре 24, 11.

Данные для расчетов опор приведены в таблице 36.

Таблица 36

Коэффициенты трения для разных материалов

Материалы	Сталь конструкционная	Сталь инструмент. закаленная
Латунь Л59	0,19	0,14
Бронза оловянистая	0,16	0,15
Дюралюминий	0,19	0,17
Агат	—	0,13
Рубин	—	0,14
Сапфир	—	0,15

Таблица 37

Физические свойства материалов

Материалы	Твердость по МООСУ	Модуль упругости
Бронза	3	$1 \cdot 10^6$
Сталь инструментальн.	6—7	$2 \cdot 10^6$
Латунь	—	$0,95 \cdot 10^6$
Агат	6,5	$1 \cdot 10^6$
Корунд	9	—
Сапфир	9	$5 \cdot 10^6$
Рубин	9	—

Таблица 38

Допустимые значения удельных давлений, кг/см²

Сталь по стали	$p = 130 \div 200$
Каленая сталь по бронзе	$p = 90 \div 150$
Сталь по бронзе	$p = 60 \div 90$
Сталь каленая по сапфиру	$p = 3000$

24.3. РАСЧЕТНЫЕ ФОРМУЛЫ

Размеры деталей опор определяются, как правило, по конструктивным соображениям, расчет на прочность производится только в тех конструкциях, где действуют значительные силы и моменты. Особое внимание уделяют при конструировании высокочувствительных и быстродействующих реле вопросу уменьшения моментов трения.

Цилиндрическая опора. Расчет цапф на механическую прочность производится по формулам, принятым для расчета деталей общего машиностроения. При диаметре цапфы меньше 1 мм расчет на механическую прочность не выполняется.

Удельное давление на цапфу при равномерном прилегании цапфы к вкладышу (фиг. 24, 2)

$$p = \frac{Q}{d \cdot l}. \quad (24-1)$$

Учитывая неравномерное прилегание цапфы, считают

$$p = \frac{4}{\pi} \frac{Q}{d \cdot l}. \quad (24-2)$$

Уравнение прочности цапфы на изгиб

$$\frac{Ql}{2} = \frac{\pi d^3}{32} R_{изг}. \quad (24-3)$$

При расчете цапфы вводится коэффициент $k = \frac{l}{d}$, называемый конструктивным параметром.

Из (24-2) и (24-3)

$$k = 0,5 \sqrt{\frac{R_{изг}}{p}}. \quad (24-4)$$

Фиг. 24, 2. Цилиндрическая опора.

Из (24-4) и (24-3)

$$d = 2,27 \sqrt{\frac{Qk}{R_{изг}}}. \quad (24-5)$$

Ход расчета. Задаваясь значениями p и $R_{изг}$ вычисляют k , далее определяют d , а затем l . В высокочувствительных реле принимают

$$p = 15 \div 25 \text{ кг/см}^2.$$

Момент трения цилиндрической опоры

$$M = \frac{Qd}{2} \mu. \quad (24-6)$$

Плоская пята (фиг. 24, 3). Конечная цапфа и опорная поверхность выполняются в виде плоскостей. Удельное давление в пяте

$$p = \frac{Q}{\pi R^2}; \quad (24-7)$$

$$d(M_{mp}) = p \, dq r \mu = 2\pi p \mu r^2 dr.$$

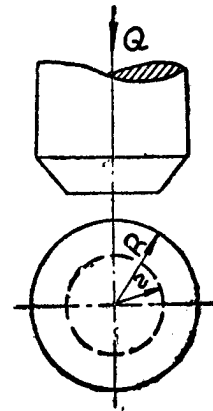
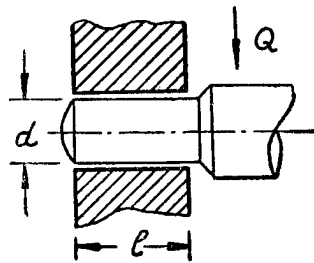
$$M_{mp} = 2\pi p \mu \int_0^R r^2 dr = \frac{2}{3} \pi p \mu R^3. \quad (24-8)$$

Из (24-6) и (24-8)

$$M_{mp} = 0,376 \frac{Q^{3/2}}{\sqrt{p}} \mu. \quad (24-9)$$

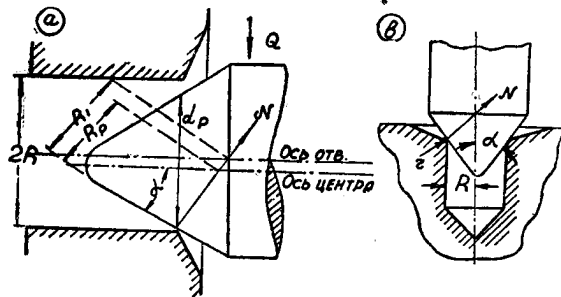
Опоры на центрах.

1) Горизонтальная ось (фиг. 24, 4 а). Место соприкосновения центра с подшипником рассматривается как сопряжение



Фиг. 24, 3. Плоская пята.

двух тел, сжимаемых силой N . Форма этих тел определяется главной кривизной в двух взаимно перпендикулярных плоскостях: для центра $\frac{1}{R_p}$ и 0 ; для подшипника $-\frac{1}{R_1}$ и $\frac{1}{r}$; максималь-



Фиг. 24, 4. Эскизы к расчету опоры на центрах.

ное напряжение в центре площадки соприкосновения определяется по формуле Герца:

$$\sigma_{\max} = n_0 \sqrt[3]{\frac{QE^2}{\rho^2}}, \quad (24-10)$$

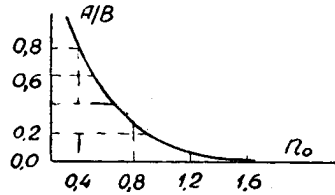
где $E = \frac{2E_1E_2}{E_1 + E_2}$; E_1 и E_2 — модули упругости подшипника и центра; $\frac{1}{\rho}$ — приведенная кривизна; n_0 — коэффициент, определяемый из фиг. 24, 5.

Принимают

$$\frac{1}{\rho} = \frac{1}{r}; \quad \frac{A}{B} = \frac{1/r}{1/R_p - 1/R_1}.$$

Момент трения

$$M_{тр} = \mu \frac{Q}{\cos \alpha} d_p. \quad (24-11)$$



2) Вертикальная ось (фиг. 24, 4 б). Напряжение на площадке соприкосновения

Фиг. 24, 5. Кривая для определения n_0 .

$$\sigma_{\max} = 0,59 \sqrt{\frac{Q}{2\pi R \sin \alpha \left(\frac{1}{E_1} + \frac{1}{E_2}\right) r}}. \quad (24-12)$$

Момент трения

$$M_{тр} = \mu \frac{Q}{\sin \alpha} R. \quad (24-13)$$

3) Опоры на кернях.

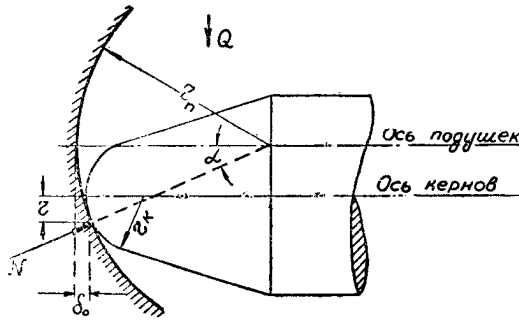
Горизонтальная ось (фиг. 24,6). Обычно длина керня несколько меньше расстояния по оси между подушками. Разность между этим расстоянием и длиной керня — осевой зазор — принимается в среднем

$$\Delta l = 0,6(r_n - r_k). \quad (24-14)$$

Величина максимального напряжения на площадке соприкосновения керна и подушки определяется по формуле Герца:

$$\sigma_{\max} = 0,617 \sqrt[3]{\frac{Q \left(\frac{1}{r_k} - \frac{1}{r_n}\right)^2}{\left(\frac{1}{E_n} + \frac{1}{E_k}\right)^2}}. \quad (24-15)$$

Момент трения определяется из следующих условий: в процессе вращения керна вскатывается по подушке с трением, радиус которого r . Момент трения $M_{тр} = Nr\mu$. Учитывая



получаем

$$\frac{r}{r_k} = \sin \alpha; \quad N = \frac{Q}{\sin \alpha},$$

получаем

$$M_{тр} = Qr_n\mu. \quad (24-16)$$

Таким образом, момент силы трения определяется только радиусом керна и не зависит от радиуса подушки и

Фиг. 24, 6. Эскиз к расчету опоры на кернах.

величины осевого зазора. Отношение $\frac{r_k}{r_n}$ принимают порядка $\frac{1}{4} - \frac{1}{10}$, в среднем $\frac{1}{7}$.

Вертикальная ось. При вертикальном расположении оси соприкосновение керна и подушки происходит не в одной точке, а по некоторой плоскости радиусом r , что объясняется деформацией керна под действием осевой составляющей силы — веса подвижной системы. По теории Герца значение r зависит от веса подвижной системы, радиусов r_k и r_n , а также от модулей упругости материалов керна и подушки:

$$r = \sqrt[3]{0,682 \frac{\frac{1}{E_k} + \frac{1}{E_n}}{\frac{1}{r_k} + \frac{1}{r_n}} Q}. \quad (24-17)$$

Удельное давление получается равным

$$p = \sqrt[3]{0,235 \frac{\left(\frac{1}{r_k} - \frac{1}{r_n}\right)^2}{\left(\frac{1}{E_k} + \frac{1}{E_n}\right)^2} Q} \text{ кг/см}^2. \quad (24-18)$$

Момент трения определяется по формуле (24-9), выведенной для плоской пяты. Уменьшение момента трения керна по сравне-

нию с таковым для плоской пяты достигается за счет повышения удельного давления. Поэтому в опорах на кернах последний изготавливается из специальных сортов стали и закаляется, а подушка — из твердых пород камней — сапфира, рубина или корунда. Конус керна исполняют с углом у вершины от 60 до 90°, значение радиуса керна принимают в функции нагрузки.

Q, грамм	r _к , микрон
1	5—10
1—2	15—30
2—5	30—40
3—6	40—60
6—200	60—200

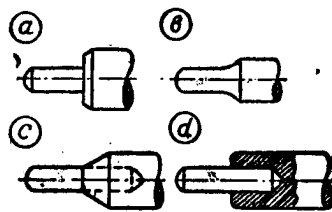
Опоры на призме. Напряжения в месте соприкосновения ножа и подушки определяются по формуле Герца:

$$\sigma_{\text{макс}} = 0,59 \sqrt{\frac{Q}{l} \frac{1}{r} \frac{1}{1/E_n + 1/E_k}}, \quad (24-19)$$

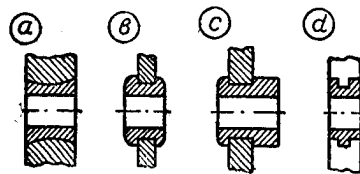
где Q — нагрузка на опору, l — длина площадки соприкосновения, r — радиус закругления ножа.

24.4. КОНСТРУКТИВНЫЕ ВОПРОСЫ

Цилиндрические опоры. В простейшем случае ось выполняется из сплошного материала без специальной обточки цапф. Цапфы диаметром больше 0,5 мм обычно представляют собой одно целое



Фиг. 24, 7. Конструкция цапф.



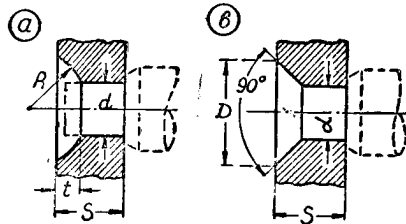
Фиг. 24, 8. Крепление боек:
a — в толстой плате; b, c — в тонких;
d — в фарфоре.

с осью (фиг. 24, 7). При малом диаметре цапфы ось выполняется составной: вставную цапфу из рояльной проволоки, а при больших диаметрах, больше 1 мм, из стали серебрянки, впрессовывают в стальную или алюминиевую ось. Способы крепления боек и конструкции подшипников приведены на фигурах 24, 8 и 24, 9.

Опоры на центрах. Основные рекомендуемые размеры приведены на фигуре 24, 10.

Опоры на кернах. Исполнение подшипниковых винтов простейшей конструкции приведено на фигуре 24, 11: a — без камня, b — с камнем. В более сложных конструкциях подшипников камни укрепляются на пружинном основании. Формы исполнения камней

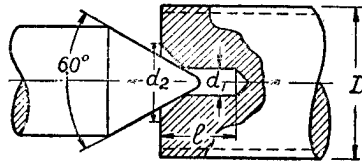
приведены на фигуре 24, 12. Ввиду недостаточной центровки опор на кернах ответственные конструкции выполняют при вертикальной оси: верхнюю опору на цапфе, нижнюю — на керне.



Фиг. 24, 9. Конструкция подшипников:

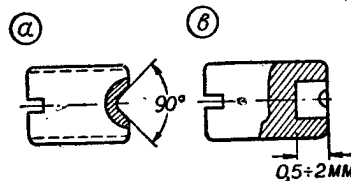
$$a - d < 5 \text{ мм}; t = \frac{S}{3}; R \cong d;$$

$$b - d > 5 \text{ мм}; D = 0,8 (S + d).$$

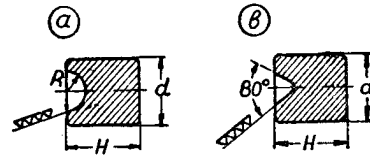


Фиг. 24, 10. Опора на центрах:

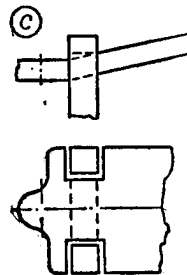
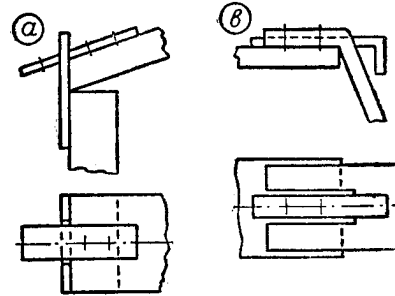
D	До 2,5	2,5—5	5—10	10—2a
d ₁	0,5	0,75	1,0	1,5
d ₂	1,2	2	2,5	4
t	1,5	2,5	3	4,5



Фиг. 24, 11. Исполнение подшипниковых винтов.



Фиг. 24, 12. Исполнение камней
H = 1 ÷ 4,5; d = 2,3 ÷ 4,5.



Фиг. 24, 13. Формы исполнения якорей реле — опора на призме.

Опоры на призме. На фигуре 24, 13 приведены примерные исполнения опор для электромагнитных реле постоянного тока.

Обработка. Цапфы, центры, керны и буксы должны быть тщательно отполированы. Чистота поверхности при обработке должна быть обеспечена для опор в центрах, кернах и на призме в пределах 9—10 класса.

ЛИТЕРАТУРА

К части первой «Магнитные системы»

1. Бабилов М. А. Основы электроаппаратостроения, М., Госэнергоиздат, 1950.
2. Бессонов Л. А. Электрические цепи со сталью, М., Госэнергоиздат, 1948.
3. Буйлов А. Я. Основы электроаппаратостроения, М., Госэнергоиздат, 1946.
4. Буль Б. К. Методы учета магнитных цепей с учетом магнитного сопротивления стали, ж. «Электричество», 1952 г., № 11.
5. Буль Б. К. К расчету магнитных проводимостей поля, ж. «Электричество», 1952 г., № 7.
6. Буль Б. К. Метод расчета катушки со сталью на переменном токе, ж. «Электричество», 1954 г., № 8.
7. Гальперин Н. А. Специальные характеристики холоднокатанной стали, ж. «Вестник электропромышленности», 1951 г., № 2.
8. Довгалецкий Я. М. Сплавы для постоянных магнитов, М., Металлургиздат, 1954.
9. Займовский Я. С. и Усов В. В. Металлы и сплавы в электротехнике, М., Госэнергоиздат, 1949.
10. Кобленц М. Г. Определение перегрева катушек контакторов постоянного тока, ж. «Вестник электропромышленности», 1947 г., № 7.
11. Сенкевич А. М. Постоянные магниты, изд-во ВВА им. Жуковского, 1949.
12. Сотсков Б. С. Элементы автоматической и телемеханической аппаратуры, М., Госэнергоиздат, 1950.
13. Справочник по изоляции, М., Госэнергоиздат, 1946.
14. Шрамков А. Я. Прибор для определения короткозамкнутых витков. Научные записки ЛПИ, вып. VIII, 1950.

К части второй «Электромагнитные системы и реле»

1. Витенберг М. Н. Расчет и конструкция телефонных и кодовых реле, М., Госэнергоиздат, 1947.
2. Котов П. А. К теории поляризованных реле, Научно-технический сборник по электросвязи, вып. 4/20, 5/21, 1937.
3. Лысов Н. Е. Расчет электромагнитных механизмов, М., Оборонгиз, 1949.
4. Москвитин А. И. К расчету динамических характеристик, «Известия АН СССР», Отдел технических наук, 1948 г., № 4.
5. Остроумов Г. А. К вопросу о линейной теории поляризованных реле, ж. «Автоматика и телемеханика», 1948 г., № 5.
6. Ротерс. Электромагнитные механизмы, М., Госэнергоиздат, 1949.
7. Сахаров П. В. и Пеккер П. И. Сравнение магнитных систем контакторов переменного тока, ж. «Вестник электропромышленности», 1949 г., № 4.
8. Сахаров П. В. и Пеккер П. И. Упрощенный метод расчета пря-

моходовых систем переменного тока, ж. «Вестник электропромышленности», 1949 г., № 12.

9. Сотсков Б. С., Декабрун И. Е. и Путилова Н. А. Зависимость времени движения подвижной системы от параметров реле, ж. «Автоматика и телемеханика», 1951 г., № 4.

10. Ступель Ф. А. Реле защиты и автоматики, М., Госэнергоиздат, 1949.

11. Ступель Ф. А. Расчет и конструкция электромагнитных реле, М., Госэнергоиздат, 1950.

12. Ступель Ф. А. Вопросы проектирования поляризованных реле для электросиловой автоматики, ж. «Электричество», 1954 г., № 7.

К части третьей «Индукционные системы и реле»

1. Атабеков Г. И. Релейная защита электрических систем, М., Госэнергоиздат, 1952.

2. Бресслер А. М. Омметры дистанционной защиты, Технический бюллетень завода МЭП, Чебоксары, 1950, октябрь.

3. Буль Б. К. Расчет вращающего момента индукционного реле, ж. «Электричество», 1946 г., № 4.

4. Буль Б. К. Теория индукционной системы с экраном, ж. «Электричество», 1947 г., № 7.

5. Гельбух Л. А. Расчет электромеханических моментов индукционного реле с цилиндрическим ротором, ж. «Автоматика и телемеханика», 1949 г., № 1.

6. Ступель Ф. А. R — X характеристики дистанционных реле, Сборник ХЭТИ, М., Госэнергоиздат, 1946.

7. Ступель Ф. А. Расчет индукционных систем и реле. Учебное пособие, ХПИ, 1953.

К части четвертой «Разные реле, элементы конструкций»

1. Алексеевский В. А. Применение термобиметалла в электроаппаратостроении, АН Армянской ССР, 1953.

2. Алексеевский В. В. Новая серия тепловых реле, ж. «Вестник электропромышленности», 1949, № 11.

3. Воробьев А. Н. и Кацнельсон О. Г. Термобиметалл и его применение, М., Госэнергоиздат, 1951.

4. Дроздов Ф. В. Детали приборов, М. Оборонгиз, 1948.

5. Миллер А. Е. и Габриелян Д. И. Термобиметаллы, ж. «Цветные металлы», 1947, № 1 и 2.

6. Сивоконенко И. М. Опоры подвижных систем приборов. Гос. изд-во судостроительной литературы, 1952 г.

СОДЕРЖАНИЕ

Предисловие	3
Перечень обозначений	4
Введение	6

ЧАСТЬ ПЕРВАЯ

МАГНИТНЫЕ СИСТЕМЫ

Глава 1. Магнитопровод

1-1. Основные величины, характеризующие свойства материалов для магнитопроводов	12
1-2. Классификация магнитных материалов	12
1-3. Материалы для магнитопроводов	13
1-4. Вопросы технологии	16
1-5. Область применения магнитных материалов	18
1-6. Исполнение деталей магнитопроводов	18

Глава 2. Катушки

2-1. Основные требования и классификация	28
2-2. Материалы для намотки катушек	29
2-3. Вспомогательные изоляционные материалы	32
2-4. Намоточные станки	35
2-5. Каркасные катушки	36
2-6. Бескаркасные катушки	40
2-7. Выводы катушек	43
2-8. Отделка катушек	45
2-9. Коэффициент заполнения	48
2-10. Нагрев катушек постоянного тока	52
2-11. Теплоотдача катушек постоянного тока	54
2-12. Нагрев и теплоотдача катушек переменного тока	58
2-13. Сопротивление намотки	59
2-14. Расчет катушки при разных режимах работы	61
2-15. Основные соотношения для расчета	67
2-16. Допуски и отклонения	67
2-17. Испытание и проверка катушек	68

Глава 3. Расчет магнитной цепи со сталью

3-1. Основные законы магнитных цепей	72
3-2. Схема замещения магнитной цепи	73
3-3. Магнитная проводимость воздушных зазоров	75
3-4. Коэффициент рассеяния	82
3-5. Местоположение максимального значения потока в магнитной системе	85
3-6. Падение магнитного потенциала в стали и в зазорах	87
3-7. Кривые проводимостей	90

3-8. Потери в стали магнитопровода	91
3-9. Токи и потребляемая мощность катушек	92
3-10. Кривые намагничивания и размагничивания	95
3-11. Кривые полного цикла работы электромагнита	101
3-12. Распределение потоков в магнитопроводе	101
3-13. Метод последовательных приближений	103

Глава 4. Расчет цепей с постоянными магнитами

4-1. Материалы для постоянных магнитов	104
4-2. Магнитные потоки отдельного магнита	108
4-3. Свободная магнитная энергия магнита	112
4-4. Кривые возврата	113
4-5. Старение постоянных магнитов и их стабилизация	113
4-6. Вопросы расчета магнитных цепей с постоянными магнитами	115
4-7. Коэффициент рассеяния системы с постоянным магнитом	117
4-8. Расчет магнитной цепи с постоянными магнитами	119
4-9. Проектирование постоянных магнитов	120

ЧАСТЬ ВТОРАЯ

ЭЛЕКТРОМАГНИТНЫЕ СИСТЕМЫ И РЕЛЕ

Глава 5. Вопросы проектирования электромагнитов

5-1. Задачи расчета	125
5-2. Общая классификация электромагнитных систем	126
5-3. Факторы, влияющие на выбор формы электромагнита	126
5-4. Условная работоспособность электромагнита	128
5-5. Удельный расход материала	130
5-6. Данные для проектирования электромагнита	131
5-7. Установление оптимального значения принятых данных	133
5-8. Коэффициент возврата	136

Глава 6. Вычисление сил и моментов

6-1. Исходные положения	138
6-2. Влияние краевого потока	138
6-3. Формула Максвелла	139
6-4. Выражение силы в функции производной проводимости	140
6-5. Трехфазные электромагниты	142
6-6. Соленоидная сила	142
6-7. Влияние расположения катушки в системе	143

Глава 7. Расчет экрана для устранения вибраций якоря

7-1. Вибрации якоря	145
7-2. Способы устранения вибраций	146
7-3. Пульсации силы	146
7-4. Расчет системы с экраном	148
7-5. Проектирование системы с экраном	150
7-6. Пример расчета экрана	152

Глава 8. Магнитные системы с поворотным якорем

8-1. Вопросы проектирования систем постоянного тока	154
8-2. Исходные данные и соотношения для систем постоянного тока	155
8-3. Предварительный расчет системы постоянного тока	156
8-4. Проверочный расчет системы постоянного тока	158
8-5. Пример расчета системы постоянного тока	159
8-6. Вопросы проектирования систем переменного тока	163

8-7.	Основные соотношения для предварительного расчета	165
8-8.	Предварительный расчет системы переменного тока	166
8-9.	Поверочный расчет системы переменного тока	167
Глава 9. Магнитные системы с поперечным движением якоря		
9-1.	Формы магнитных систем	168
9-2.	Основные соотношения и данные	169
9-3.	Системы со сплошным якорем	169
9-4.	Системы с выступающим якорем	171
9-5.	Системы со втяжным якорем	172
Глава 10. Прямоходовые магнитные системы		
10-1.	Формы магнитных систем постоянного тока	173
10-2.	Силы на якоре	174
10-3.	Выбор формы и длины стопа	175
10-4.	Предварительный расчет короткоходовой системы постоянного тока	176
10-5.	Проверочный расчет короткоходовой системы постоянного тока	177
10-6.	Расчет длинноходовой системы постоянного тока	179
10-7.	Формы магнитных систем переменного тока	181
10-8.	Силы на якоре	181
10-9.	Данные для проектирования	182
10-10.	Расчет короткоходовой системы переменного тока	183
Глава 11. Время срабатывания электромагнитных систем		
11-1.	Классификация по времени срабатывания	184
11-2.	Вычисление времени t_n на включение	186
11-3.	Вычисление времени t_n на отключение	188
11-4.	Вычисление времени движения	193
11-5.	Вопросы проектирования быстродействующих систем	196
11-6.	Вопросы проектирования систем с магнитным демпфированием	197
11-7.	Вопросы проектирования систем с механическими демпферами	199
Глава 12. Поляризованные реле		
12-1.	Введение	203
12-2.	Формы магнитных систем	203
12-3.	Настройки поляризованных реле	205
12-4.	Эксплуатационные параметры реле	206
12-5.	Расчет поляризованных реле разных систем	207
12-6.	Вопросы проектирования	208
12-7.	Время срабатывания реле	210
12-8.	Основные расчетные соотношения	211
Глава 13. Особые исполнения и схемы		
13-1.	Электромагнитные системы с постоянными магнитами	215
13-2.	Схемы	219
ЧАСТЬ ТРЕТЬЯ		
ИНДУКЦИОННЫЕ СИСТЕМЫ И РЕЛЕ		
Глава 14. Магнитные системы		
14-1.	Задачи расчета	222
14-2.	Магнитные проводимости воздушных путей магнитопровода	223
14-3.	Коэффициент рассеяния	226

14-4.	Угол потерь	227
14-5.	Векторные диаграммы	227
14-6.	Влияние реакции ротора	230
14-7.	Способы изменения углов β_e и β_i	230

Глава 15. Токи в роторе

15-1.	Электродвижущие силы, наведенные в роторе	233
15-2.	Вычисление значений токов трансформации	235
15-3.	Метод отображений для систем с диском	236
15-4.	Метод отображений для систем с барабаном	238
15-5.	Полное значение тока трансформации	239
15-6.	Токи резания	241

Глава 16. Моменты на роторе

16-1.	Среднее значение силы	242
16-2.	Общее уравнение моментов	243
16-3.	Общее выражение для вычисления тормозного момента	252
16-4.	Тормозные моменты разных систем	253
16-5.	Различные виды уравнений моментов индукционных систем	254

Глава 17. Время срабатывания

17-1.	Основные требования	256
17-2.	Общее уравнение движения ротора	257
17-3.	Быстродействующие реле	258
17-4.	Замедленно действующие реле	259

Глава 18. Параметры индукционной системы

18-1.	Конструктивные параметры	259
18-2.	Расчетные параметры	261

Глава 19. Примеры исполнения некоторых видов реле

19-1.	Дифференциально-токовые реле	267
19-2.	Реле мощности	269
19-3.	Реле сопротивления	272

ЧАСТЬ ЧЕТВЕРТАЯ

РАЗНЫЕ РЕЛЕ, КОНСТРУКТИВНЫЕ ЭЛЕМЕНТЫ

Глава 20. Магнитоэлектрические и электродинамические реле

20-1.	Магнитоэлектрические реле	277
20-2.	Электродинамические реле	280

Глава 21. Термические реле

21-1.	Классификация	282
21-2.	Термобиметалл и его параметры	282
21-3.	Деформации биметалла	285
21-4.	Механические силы, развиваемые биметаллом	286
21-5.	Вопросы проектирования тепловых реле	288
21-6.	Время срабатывания тепловых реле	293
21-7.	Основные соотношения для расчета тепловых реле и распепителей	295
21-8.	Термореле и регуляторы	298

Глава 22. Пружины

22-1. Классификация	301
22-2. Материалы для изготовления пружин	302
22-3. Конструктивные вопросы	305
22-4. Расчетные соотношения и примеры	318

Глава 23. Контактные системы

23-1. Основные параметры контактных систем	323
23-2. Способы уменьшения искрообразования	326
23-3. Вибрации контактов	327
23-4. Контакты	330
23-5. Исполнения контактных систем	334
23-6. Построение механических характеристик	337

Глава 24. Направляющие для вращательного движения

24-1. Классификация	341
24-2. Материалы для опор	342
24-3. Расчетные формулы	343
24-4. Конструктивные вопросы	347

Литература	349
----------------------	-----

Файвель Аронович Ступель
Электромеханические реле
Основы теории, проектирования и расчета
Учебное пособие

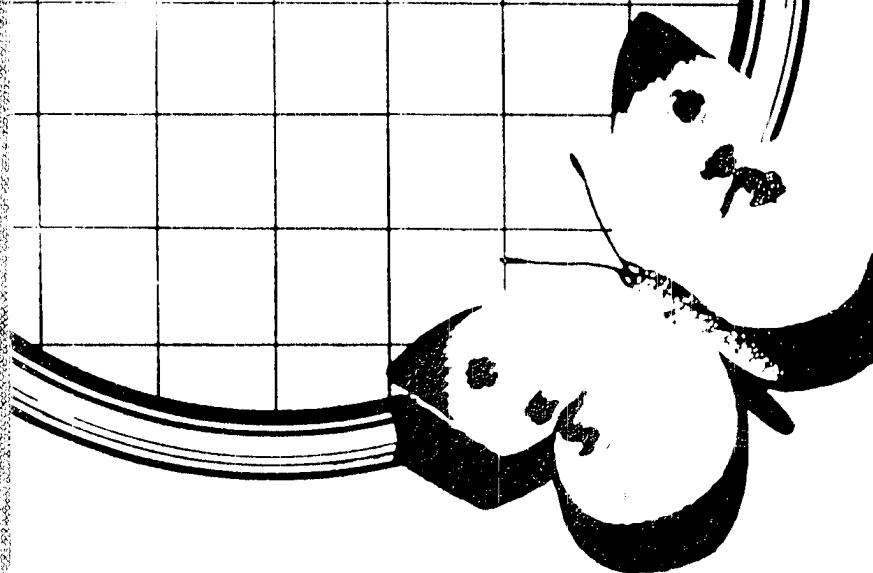
Редактор *Д. А. Вайнберг*
Технический редактор *Я. Т. Чернышенко*
Корректоры *В. М. Потапов, Н. Н. Евсеева*

Подписано к печати 2/VIII—1956 г. БЦ 06783. Тираж 20.000. Формат 60×92¹/₁₆.
Объем 11,12 б. л.=22¹/₄ п. л. 23,5 уч. изд. л.=В 1 п. л. 42600 зн.
Цена в переплете 9 руб. 75 коп.

Напечатано с матриц, изготовленных на Харьковской книжной фабрике
им. Фрунзе, в типографии «Коммунист» Главиздата Министерства
культуры УССР, Харьков, Пушкинская, 29.

Цена 9 руб. 75 коп.

У.А. Торешаев



СИГНАЛ

Sanitized Copy Approved for Release 2010/06/16 : CIA-RDP80T00246A045100330001-4

И. А. Полежаев

СИГНАЛ



ИЗДАТЕЛЬСТВО „СОВЕТСКОЕ РАДИО“
МОСКВА — 1958

Sanitized Copy Approved for Release 2010/06/16 : CIA-RDP80T00246A045100330001-4

И. А. ПОЛЕТАЕВ

СИГНАЛ

(О некоторых понятиях кибернетики)

Книга представляет собой попытку изложить некоторые из основных черт природы информационных или «кибернетических» устройств и естественных информационных механизмов в живых существах. В книге разобраны понятия сигнала, информации, количества информации, обсуждается роль случайных событий при искажениях сигналов, связь информационной и физической энтропии. Кратко рассматриваются процессы передачи сигнала в технических каналах связи и в нервной системе, работа систем с обратной связью, устройство и работа счетных машин при решении математических и логических задач. Обсуждаются вопросы работы простых и сложных автоматов — роботов, а также черты работы нервной системы, позволяющие сделать заключение о сигнальном характере процессов в нервной системе. Отдельно разобраны вопросы, относящиеся к поведению человека и автоматов в ситуации соревнования или игры, а также возможности самоорганизации систем, построенных без заранее установленного плана. В заключение обсуждаются принципиальные отличия автоматов от живых существ и перспективы развития техники сложных автоматов.

Книга рассчитана на широкий круг читателей. Разделы, требующие специальных знаний математики, выделены и могут быть опущены без большого ущерба для понимания остального материала.

Sanitized Copy Approved for Release 2010/06/16 : CIA-RDP80T00246A045100330001-4

*Узнать, понять и охватить гармонию
научного здания с его недостроенными
частями — значит получить такое наслаж-
дение, какое даёт только высшая красота
и правда.*

Д. И. МЕНДЕЛЕЕВ.

Sanitized Copy Approved for Release 2010/06/16 : CIA-RDP80T00246A045100330001-4



ПРЕДИСЛОВИЕ

Кибернетика* есть наука о процессах управления и передачи сигналов в машинах и живых организмах, использующая математические методы.

Сигнал, т. е. физический процесс, несущий информацию, является центральным понятием кибернетики, отсюда — название этой книги.

Возникновение, передача и использование сигнала для управления представляет собой явление весьма общее для самых, на первый взгляд, различных объектов. Закономерности использования сигнала в системах управления и связи обширны, многообразны и отличны от законов преобразования энергии. Эта специфика нуждается в отдельном изучении и отдельной трактовке, достаточно общей и строгой.

Задача данной книги не в том, чтобы дать исчерпывающие, точные и окончательные определения понятий

* От греческого корня *κυβερνήτης*, означающего «искусство рулевого».

кибернетики и изложить решения всех проблем. Такая задача сегодня вряд ли выполнима. Цель этой книги заключается лишь в том, чтобы дать читателю предварительное, по возможности понятное изложение общих идей, на которых строится учение об информации и управлении. Большой интерес к вопросам кибернетики, по-видимому, оправдывает такую попытку. Общее знакомство со всем кругом идей кибернетики в целом обычно вызывает большой интерес и помогает углубленному изучению какого-либо одного ее раздела.

Широкое поле применений кибернетики—от теории связи до рефлексологии—создает большие трудности для тех, кто пытается охватить материал в целом. Однако большая часть ценности понятий кибернетики в том и состоит, что они позволяют усматривать общее в самых разнообразных явлениях и взаимно обогащать самые, казалось бы, отдаленные друг от друга области знаний. Поэтому ограничиваться рассмотрением каких-либо отдельных применений кибернетики означает наносить вред изложению. Признав нежелательным такое обеднение содержания и будучи убежденным в том, что опаснее совершить мелкие ошибки в деталях является меньшим злом, чем выбрасывание из рассмотрения целых областей применения кибернетики, автор был вынужден, естественно, входить в те области знаний, которые далеки от его привычной узкой специальности. Разумеется, добиться исчерпывающей полноты изложения было затруднительно по многим причинам. К содержанию книги можно и, быть может, нужно многое добавить.

Изложение понятий кибернетики в этой книге, в основном, преследует цель возбудить интерес к конкретным проблемам и привлечь к ним внимание. Нам представляется, что биолог и врач по-новому увидят тот живой материал, с которым они работают, получив общее представление об информации, связи и управлении в их технических применениях, а инженер сумеет увидеть новые перспективы, сравнив системы, им создаваемые, с естественными системами аналогичного назначения.

Изложение понятий кибернетики в общедоступной форме и без больших упрощений представляется делом нелегким. Берясь за его выполнение, часто приходится жертвовать развлекательной стороной изложения, надеясь, что содержание самих понятий кибернетики не позволит читателю скучать.

Кибернетике еще не исполнилось десяти лет со дня ее

если не рождения, то «крещения»*. Количество проблем как принципиально-познавательного, так и утилитарно-прикладного характера в области кибернетики поистине огромно, и большинство из них ожидает скорейшего разрешения. Поэтому задача направить силы ученых и техников на решение этих насущных проблем, создать единое понимание общих законов управляющих систем и систем связи, установить общую систему понятий и терминов для различных отраслей знаний представляется задачей чрезвычайно актуальной. Если настоящая книга хоть в незначительной степени послужит привлечению внимания различных специалистов к их общим интересам, задачу ее можно считать выполненной.

Мы являемся свидетелями первых шагов кибернетики. Возможности, которые она сулит в перспективе, настолько велики, что самая смелая фантазия может оказаться бессильной их представить.

Изложение в настоящей книге не копирует какой-либо источник или серию источников, хотя общие идеи почерпнуты из многих работ.

Кибернетика сложилась впервые в США, Франции и Англии. Первые шаги кибернетики в СССР были омрачены возникшим по недоразумению предвзятым отношением к ней. К счастью, недоразумение сейчас рассеяно, и кибернетика заслуженно завоевывает все большее внимание советской интеллигенции. Сегодня уже невозможно огульное охаивание кибернетики; остались только споры принципиального характера, закономерные и необходимые на всех этапах развития науки.

Спорных и нерешенных вопросов в кибернетике много. Это делает ее особенно интересной. В настоящей книге намеренно не обходятся спорные вопросы. Относительно будущих, не созданных еще устройств, нерешенных еще вопросов невозможно делать бесспорные утверждения. Широкое и свободное обсуждение проблем в научных спорах полезно для развития науки и поучительно для участников этих споров. Поэтому автор заранее благодарен всем, кто возьмет на себя труд добросовестной и аргументированной критики. Речь идет, разумеется, о науч-

* Слово «кибернетика» вошло в обиход с 1948 года после выхода в свет книги Н. Винера под этим названием. Впервые слово «кибернетика» употребил французский физик Андре Мари Ампер для обозначения науки о государственном управлении.

ном обсуждении вопроса, а не о декларациях с предвзятых позиций, которыми изобиловали первые сообщения о кибернетике у нас.

Из числа отрицательных суждений о кибернетике хочется отметить одно. Иногда работники узких прикладных специальностей, примыкающих к кибернетике, задают вопрос: «Зачем нужно называть «кибернетикой» то, что существует много лет без этого названия и чем мы с успехом занимаемся ежедневно? Нам это ни в чем не поможет!» Не хотелось бы резко возражать против подобных суждений. Можно много лет работать кочегаром и не иметь понятия об энергетике в целом. Можно всю жизнь говорить прозой и не подозревать об этом, без малейшего вреда для себя. Совершенно так же можно изо дня в день занимаясь, например, следящими системами, не задумываться над связями понятий техники и биологии. И никто не будет вправе сказать, что такая деятельность в пределах одной узкой специальности не полезна. Но если бы мы занимались только деятельностью в пределах «узких специальностей», не выходя за их рамки, то мы не имели бы ни реактивных самолетов, ни меченых атомов, ни, быть может, атомной энергии, словом, ничего, что создано смелой мыслью, ломающей рамки «узких специальностей». Порой смелое объединение разнородных явлений общими понятиями приносит неизмеримо больше общественной пользы, чем движение по проторенным путям.

Автор считает своим приятным долгом выразить глубокую благодарность всем, оказавшим ему помощь в написании этой книги советами, указаниями и обсуждением трудных вопросов, и в первую очередь академику А. И. Бергу, без инициативы которого эта книга не увидела бы света, а также т.т. А. А. Ляпунову, А. И. Китову, Л. В. Крушинскому, М. О. Герцбергу и многим другим.

Москва, 1956 г.



ГЛАВА 1

ЭНЕРГЕТИКА И КИБЕРНЕТИКА

Нельзя представить себе современной культуры без использования огромных количеств энергии из естественных источников. Никакая отрасль современной индустрии не обходится без энергетических устройств значительной мощности. Огромные количества энергии расходуют мы при плавке чугуна в доменных или стали в мартеновских печах, обжимая многотонные болванки на прокатных станах, снимая стружку с миллионов металлических деталей на металлообрабатывающих станках, поднимая и перенося стройматериалы на стройплощадках, перемещая миллионы тонн грунта на строительстве гидросооружений, перевоза грузы и пассажиров из одного конца страны в другой по земле, воде и по воздуху, выполняя сельскохозяйственные работы. Энергия, дробясь на мелкие порции, проникает в наши жилища, освещает, согревает их, позволяет без затраты физического труда выполнять мелкие домашние работы. Энергия, преобразуемая машинами из одного вида в другой, послушно служит не только нашим материальным потребностям, но и помогает удовлетворять потребности

духовные. Современная полиграфическая промышленность, телефон, телеграф, радио, телевидение, кино не могли бы существовать без умелого использования больших количеств энергии. Попробуйте представить себе, как изменилась бы жизнь современного города, если бы были выключены все источники энергии и нам пришлось бы обходиться только силой собственных мышц, и вам станет ясно, как тесно и прочно связана жизнь современного общества с использованием энергии.

Разнообразные и многочисленные применения энергии на пользу человека суть результат многолетнего развития различных отраслей науки и техники, которые объединяются общим названием энергетика. Электротехника, гидро- и аэродинамика, термодинамика, ядерная физика и многие другие отрасли науки и техники являются составными частями энергетики.

Основные знания в области энергетики были заложены уже в девятнадцатом веке. Открытие основных законов преобразования энергии — первого и второго законов термодинамики — позволило подходить строго, с числом и мерой, к конструированию энергетических машин.

Первый закон, закон сохранения энергии гласит: невозможно возникновение или уничтожение энергии. Этот закон устанавливает эквивалентность различных видов энергии при ее преобразованиях. Один киловаттчас электроэнергии может превратиться в 367 100 килограммометров механической работы, не больше и не меньше. Однако обычно не удается целиком использовать все количество энергии, участвующей в преобразовании. При работе машины часть энергии расходуется попусту, но она не исчезает, а переходит в такие виды энергии (чаще всего в тепло), которые уже нельзя полностью «собрать» и использовать. Использовать удается лишь часть энергии, участвующей в преобразовании. Эта часть (обычно выражаемая в процентах), называется коэффициентом полезного действия машины, преобразующей энергию. Одним из следствий закона сохранения энергии является следующее утверждение: «Коэффициент полезного действия (к. п. д.) машины не может быть больше 100%». Практически он всегда меньше 100%.

Конкретные отрасли энергетики, занимающиеся преобразованием энергии, в качестве одной из основных задач имеют повышение к. п. д. В этом отношении остается еще многое сделать. Так, например, к. п. д. паровоза обычно не превышает нескольких процентов, и из всей энергии,

полученной от сжигания угля в топке, лишь менее одной десятой расходуется на перемещение поезда, а остальные девять десятых бесполезно тратятся на нагревание окружающего воздуха.

Энергию, израсходованную на нагревание воздуха, обычно уже невозможно использовать снова или же, если и возможно, то лишь частично. Если два резервуара тепла имеют одинаковую температуру, то, хотя в них запасено большое количество энергии, использовать ее невозможно, не привлекая на помощь резервуар с более высокой или более низкой температурой. Это утверждение является одним из следствий второго закона термодинамики. Второй закон может быть сформулирован так: невозможен процесс, единственным результатом которого явилось бы превращение тепла в работу.

Первый и второй законы дают возможность ставить разумные требования к энергетическим системам и показывают, что от них можно требовать и чего нельзя.

Правда, и до наших дней еще встречаются горе-изобретатели вечных двигателей, пытающиеся обойти либо первый, либо второй закон и получить энергию из таких источников, из которых ее невозможно извлечь. Попытки «изобретения» вечных двигателей не приносят пользы.

Энергетика еще не сказала своего последнего слова и продолжает неуклонно и быстро развиваться. Достаточно напомнить, что сейчас начата новая глава энергетики—использование атомной энергии. По этому пути сделаны только первые шаги, и в настоящее время еще трудно предсказать, насколько сильно использование энергии атома изменит облик энергетики. Ясно только, что эти изменения будут значительными и что в результате их «мышцы» человечества станут еще более могучими, чем в наши дни.

Управление потоками энергии

Всякое применение или использование энергии требует управления её потоком. Любую энергетическую машину нужно пускать в ход и останавливать, регулировать ее работу. Любой технологический процесс требует изменения количества подводимой энергии во времени.

Самым существенным в процессе управления потоком энергии является то, что для осуществления управления необходимо всегда меньшее количество энергии, чем то, которым управляют. Если бы это было не так, то управ-

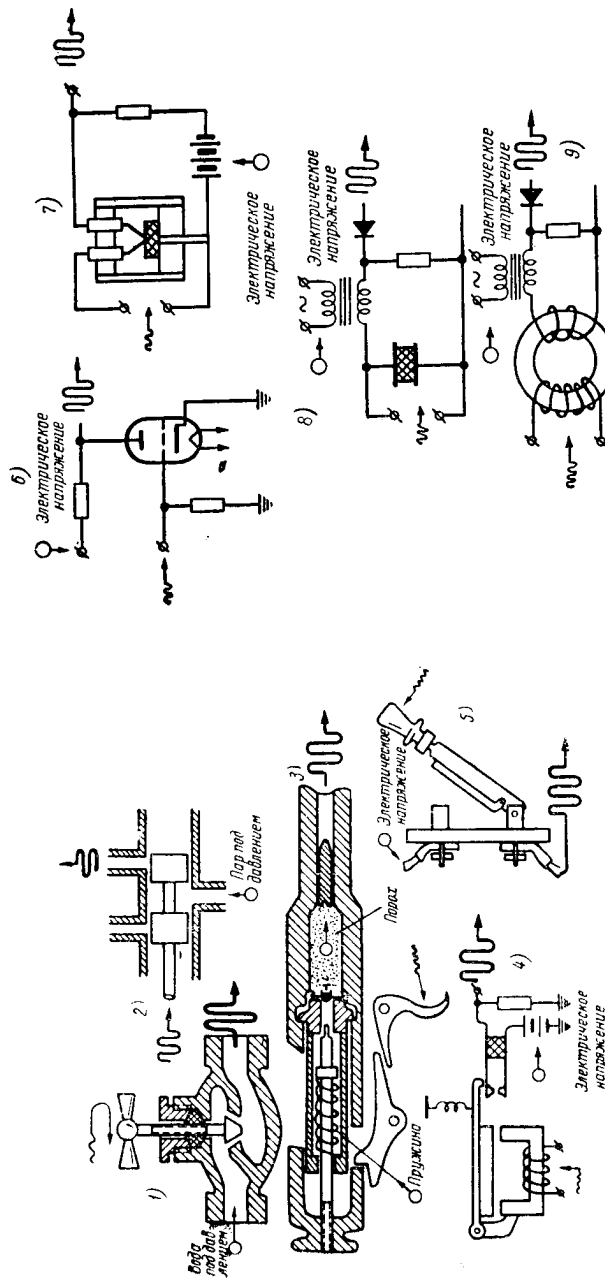


Рис. 1.1. Примеры устройств, позволяющих управлять энергией внешнего источника (обозначен кружком со стрелкой) посредством слабого воздействия (обозначено тонкой стрелкой) и получить выходной эффект значительной мощности (обозначен жирной стрелкой).

1 — кран, 2 — золотник, 3 — курковый механизм огнестрельного оружия, 4 — реле, 5 — усилитель на вакуумной электронной лампе, 6 — усилитель на полупроводниковом триоде, 7 — реле, 8 — дзлектрический усилитель, 9 — магнитный усилитель.

ление было бы невозможно. Любое управляющее устройство имеет в своем составе «клапан», который открывает или закрывает путь большому количеству энергии и который требует для своего «срабатывания» сравнительно малого усилия. Такого рода «клапанами» являются электронные лампы, реле, контакторы, выключатели, дроссельные заслонки двигателей внутреннего сгорания, золотники паровой машины, водопроводные краны и т. д. Все устройства подобного рода можно рассматривать как усилители, которые получают на «вход» слабое воздействие и дают на «выходе» соответствующее действие большой силы, для которого энергия черпается из постороннего источника. Усилие, которое прикладывается к рукояткам управления паровой машины локомотива, соизмеримо с силой мышц человека, и человек, стоящий у машины, движением своих рук управляет энергией, во много раз превышающей его собственные энергетические ресурсы.

Первоначально, на заре энергетики, только человек-машинист стоял у рукояток управления машиной. В «атмосферической машине» Ньюкомена даже управление золотником для впуска пара в цилиндр при каждом ходе поршня осуществлялось машинистом, и скорость хода машины зависела, таким образом, от его расторопности. По преданию, мальчик-машинист Гемфри Поттер заставил машину самоё дергать ручку золотника в нужный момент, привязав ручку веревочкой к шатуну машины. Это было одним из первых применений обратной связи в машине, которое сохранилось в принципе до наших дней и которое будет, вероятно, существовать еще долго. Изобретательный и ленивый мальчик передал свои обязанности по регулированию впуска пара самой машине, причем он разгрузил этим не только свои руки, но—что принципиально гораздо важнее— и свое внимание.

Регулятор Джемса Уатта, часто приводимый как классический пример автомата регулирования, относится к тому же классу устройств, заменяющих деятельность человека при выполнении им одной операции. В данном случае эта операция заключается в поддержании скорости хода машины постоянной посредством регулировки подачи пара при изменениях нагрузки. Регулятор отводит небольшую часть энергии машины и направляет ее на приведение в действие регулирующего клапана—заслонки. Этого небольшого количества энергии достаточно для управления машиной потому, что заслонка, подобно другим устрой-

ствам «клапанного» типа, обладает свойством «усиления». Энергия, отводимая для регулирования, несет «сведения» об изменениях хода машины. Эти сведения регулятор использует для того, чтобы воздействием на заслонку регулировать ход. В сущности то же делал и человек, передавший свои функции регулятору.

И сейчас еще человек продолжает управлять машиной непосредственно, находясь в будке крановщика или экскаваторщика, сидя за рулем автомашины или трактора, держа в руках штурвал корабля или ручку контроллера электропоезда. Однако со времен Уатта и Ньюкомена машины настолько «поумнели», что некоторые из них могут

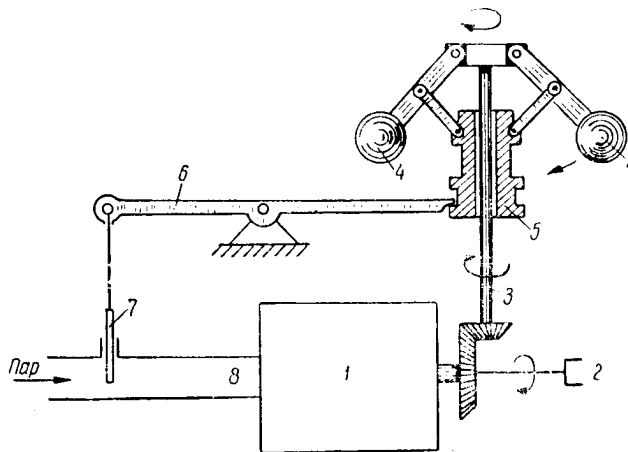


Рис. 1.2. Центробежный регулятор.

Паровая машина 1 работает на нагрузку 2, одновременно вращая регулятор. Вместе с валом 3 вращаются грузы 4 и обойма 5. Если нагрузка на валу машины уменьшится, вал начнет вращаться быстрее, грузы под действием центробежной силы разойдутся в стороны и поднимут обойму. Скользящий в пазу обоймы конец рычага 6 также поднимется. Другой конец рычага опустится и опустит заслонку 7, перекрывая трубу 8 и уменьшая доступ пара в машину. В результате работы регулятора машина лишь немного увеличит обороты, несмотря на уменьшение нагрузки.

обходиться вовсе без помощи человека. Человек же в тех случаях, когда он присутствует, выполняет лишь операции, которые почему-либо еще не автоматизированы. Машина «сама» принимает меры предупреждения против вредных последствий невнимательности и ошибок человека.

Решающую роль в развитии техники автоматического управления энергетическими устройствами сыграло появление

ние электроники—техники применения приборов, использующих свойства электронных и ионных токов в вакууме и полупроводниках, а также достижения техники связи. Создание быстродействующих, надежных в эксплуатации усилителей и реле создало огромные возможности для автоматического управления и управления на расстоянии.

Когда мы набираем с помощью диска номер телефона, мы получаем соединение с нужным абонентом—одним из нескольких десятков или сотен тысяч, причем ни один человек не участвует непосредственно в осуществлении соединения. Вращение наборного диска под нашими пальцами быстро и надежно управляет работающими телефонными искателями на АТС, расположенной за несколько километров от нашей квартиры.

Стоящий у вас в комнате холодильник сам следит за постоянством температуры своего шкафа, не допуская ни нагрева, ни переохлаждения его содержимого. Он управляется автоматически.

Сложные сообщения, приказы, распоряжения передаются по телефону, телеграфу, радио. Эти приказы управляют действиями людей—исполнителей. Рисунки, чертежи, фотографии передаются по фототелеграфу. Еще более сложные сообщения передаются по системе телевидения, доводя до зрителя движущееся изображение, серое или цветное, одновременно со звуковым сопровождением.

Передача сообщений на расстояние, управление на расстоянии, автоматическое управление—все эти отрасли техники выросли в недрах энергетики, на основе ее технических достижений и для обслуживания ее нужд. Однако совершенно ясно, что управление не сводится к одним только преобразованиям энергии, у него есть специфические особенности, которые заслуживают отдельного изучения.

Нужда в автоматах и управлении на расстоянии особенно возросла и стала насущной потребностью, когда родилась новая отрасль энергетики—использование ядерной энергии. Известно, что непосредственное соприкосновение человека с радиоактивными препаратами вызывает ожоги, лучевую болезнь и смерть. Только очень небольшие порции лучистой энергии переносятся живыми организмами без вреда. Поэтому управление ядерными реакциями, процессами изготовления и исследования радиоактивных препаратов не может производиться руками человека и под его непосредственным наблюдением.

К счастью, к моменту появления ядерной энергетики руки у человека были уже достаточно «длинными» и глаза достаточно «дальнозоркими», чтобы справиться с новыми задачами. Техника автоматического и дистанционного управления появилась раньше ядерных реакторов. Большие и малые электросиловые станции, управляемые полностью автоматически, работали уже более двух десятков лет тому назад. На таких станциях все операции по регулированию напряжения в сети при изменениях нагрузки, аварийная защита, переключения агрегатов, регулирование их режима работы, словом, все текущие операции обслуживания производятся автоматически. Только в случае глубокого повреждения дежурный инженер вызывается автоматически действующей аварийной сигнализацией. Такая же станция, но с ручным управлением, требует для своего обслуживания непрерывного внимания и усилий более десятка работников.

Автоматические метеорадиостанции сбрасываются с самолета на парашюте, сами (с помощью автомата) устанавливаются в рабочее вертикальное положение, включаются в работу и в течение долгого времени регулярно передают по радио результаты метеонаблюдений.

Очевидно, что некоторый и притом не слишком узкий круг обязанностей человека, связанных с применением немускульной силы, а внимания и сообразительности, полноценно выполняется без его присутствия устройствами автоматического управления. На наших глазах машины «отнимают» у человека все больше обязанностей по управлению энергией. Собственно, само понятие «машина» раздвоилось. Машиной мы называем и двигатель внутреннего сгорания и электронную счетную машину. Мы, во избежание путаницы, будем называть «энергетическими машинами» машины типа двигателя внутреннего сгорания или паровой, преобразующие один вид энергии в другой. Что же касается счетных машин и систем управления, то мы дадим им несколько позже другое название.

Появление быстродействующих электронных цифровых машин, или — как их еще называют — машин дискретного счета, резко усилило «наступление» машин на область интеллектуальной человеческой деятельности или, проще говоря, позволило резко усложнить поручаемые машинам операции управления. Достаточно сказать, что автоматические устройства такого типа доказали на деле свою способность не только заменить труд многих десятков и

сотен техников—вычислителей, но даже выполнять удивительным образом такие, казалось бы, неотъемлемо человеческие обязанности, как перевод текста с одного языка на другой, например, с английского на русский. Несколько лет назад предположение о возможности поручить перевод машине многим показалось бы плодом большой фантазии.

Если машина и человек станут соревноваться в выполнении одной какой-либо операции, представляющей собой более или менее сложную реакцию на изменение внешней обстановки, то первое место, безусловно, останется за машиной. Машина работает во много раз быстрее и надежнее, чем человек, она не подвержена усталости, не отвлекается, ей не свойственны и другие человеческие слабости.

Человек, призванный выполнять одну единственную операцию, например, вести по ровной, свободной от препятствий дороге автомашину, работает как автомат. Его работа тем лучше, чем точнее, своевременнее и автоматичнее его реакции. Всякое отвлечение внимания, всякий «выход из режима автомата», чреват для него и пассажиров неприятными последствиями. Его работа может быть точно описана в тех же терминах, что и работа автомата.

Но слабость человека, призванного все внимание отдавать выполнению роли автомата, слабость, заключающаяся в том, что он может переключить свое внимание на нечто другое—отвлечься, она же является его главной силой, которой никакая машина не может похвастаться. Человек может выполнять функции любого автомата, перестраиваясь с одной задачи на другую и обучаясь делать каждое дело наилучшим образом: вести любой экипаж в любых условиях, следить за ходом производственного процесса, вычислять, изобретать и т. д. Кроме того, он может выбрать, какую из задач ему нужно решать в настоящий момент, сформулировать эту задачу и найти методы ее решения. И хотя уже существует много автоматов, каждый из которых решает одну из задач, обычно решаемых человеком, и число типов таких автоматов непрерывно растет, до настоящего времени еще нет такого, который решал бы их все. Впрочем, мы не можем привести убедительных доводов, которые опровергали бы возможность существования и такого автомата.

В теле животных, в том числе и человека, непрерывно протекает большое число процессов, которые вполне по-

подобны работе технических автоматов, хотя они и протекают в совершенно иной физической среде и включают в себя работу живых тканей и клеток, которые по своей природе совершенно отличны от реле, радиолампы, рычага, потока пара. И тем не менее регулировка энергетического процесса нагрева и охлаждения и поддержание постоянства температуры тела теплокровных животных в принципе вполне подобны регулировке температуры в термостате. Это сходство можно проследить даже в деталях.

Уместно, однако, спросить, не является ли грубой ошибкой проведение сравнений между живыми организмами и механизмами? Не впадаем ли мы при этом в «вульгарный механицизм»? По-видимому, подобные сравнения проводить можно, тем более что они проводились и проводятся уже давно. Мы давно изучаем химические превращения веществ в тканях живых организмов и многие реакции повторяем в пробирке. Мы изучаем обмен энергии в организме, подсчитываем теплотворную способность пищи, поглощаемой организмом, изучаем механизмы рычагов, образуемых костями скелета и скелетными мышцами, мало того, мы вмешиваемся и в механику и в химию организма, исправляя его недостатки на основе знаний механики, физики и химии, предпринимаем хирургическое и медикаментозное лечение и при этом не задаемся никакими вопросами о «вульгарном механицизме» или «вульгарном химизме». Если и другие законы природы, законы управления посредством сигналов, оказываются применимыми равно как для живой, так и для мертвой материи, то вряд ли что-либо, кроме, разве, предрассудков, может запретить рассматривать и использовать и эти законы. Было время, когда из-за тех же предрассудков строжайше запрещалось вскрытие трупов. Правда, это было очень давно. Сегодня мы можем и должны рассматривать факты без предубеждений, четко устанавливая как сходство, так и различия в явлениях природы.

Автоматизмы наблюдаются не только в деятельности внутренних органов, но и в поведении животных. Это так называемые «отраженные движения», или рефлексy, которые появляются с «машинной» регулярностью и всегда одинаково при одинаковых раздражениях. По их изменениям судят об отклонениях от нормы в работе организма. Значит, не только деятельность внутренних органов тела животного, но и его поведение, можно поставить, конечно,

с должной осторожностью, в один ряд с работой технических автоматических устройств, рассматривать их, хотя бы отчасти, с единой точки зрения, применяя общую методику, одинаковые критерии оценки и т. д.

Несомненно, что обобщение сведений и методов как технических, так и биологических наук на основе строгого математического подхода приведет к взаимному обогащению естественных наук. Поэтому объединение точек зрения следует считать весьма желательным.

Но возможно ли это в действительности? Не является ли все же живая природа глубоко качественно отличной от неживой? Не стоит ли меж ними непроходимая стена, которая не позволит установить единые взгляды и методы? И можно ли свести все поведение животных, их приспособление к условиям среды к автоматизмам, пусть даже очень сложным?

И. М. Сеченов отвечал на этот вопрос утвердительно. Он писал в своей работе «Рефлексы головного мозга»: «... вопрос о полнейшей зависимости произвольнейших из произвольных поступков от внешних условий человека решен утвердительно. Отсюда же роковым образом следует, что при одних и тех же внутренних и внешних условиях человека деятельность его должна быть одна и та же. Выбор между многими концами одного и того же психического рефлекса, следовательно, положительно невозможен, а кажущаяся возможность есть лишь обман сознания...»

И. П. Павлов разделял это мнение. Он писал: «... Человек есть, конечно, система (грубее говоря—машина), как и всякая другая в природе, подчиняющаяся неизбежным и единым для всей природы законам, но система, в горизонте нашего современного научного видения, единственная по высочайшему саморегулированию. Разнообразно саморегулирующиеся машины мы уже достаточно знаем между изделиями человеческих рук. С этой точки зрения метод изучения системы—человека тот же, как и всякой другой системы: разложение на части, изучение значения каждой части, изучение связи частей, изучение соотношений с окружающей средой и в конце концов понимание, на основании всего этого, ее общей работы и управление ею, если это в средствах человека...»

Что касается различия и сходства живой и неживой природы вообще и возможности «объяснения» в сех явле-

ний с единой точки зрения, то по этому поводу можно высказать некоторые общие соображения, хотя многие стороны процессов, протекающих в живых организмах, нам еще не совсем ясны.

Всякое физическое тело больших размеров (или, как говорят, макроскопическое тело) состоит из молекул (или микроскопических частиц), безразлично, принадлежит ли это тело живой или мертвой природе. Все свойства макроскопических тел определяются в конечном итоге свойствами микрочастиц и характером их взаимодействия. Однако для живых организмов эта зависимость проявляется, по-видимому, иначе, чем для тел мертвой природы.

Всякая молекула представляет собой устойчивое образование. Она не изменяет своей структуры или состояния до тех пор, пока на нее не будет оказано воздействие извне, обладающее достаточно большой энергией (тепловое соударение с другой молекулой, соударение с элементарной частицей, воздействие поля).

Молекулы, собранные в большом количестве, образуют физическое тело, обладающее новыми свойствами, которых не было у единичной молекулы. Молекулы, входящие в состав тела, постоянно обмениваются энергией, обмениваются количеством движения, взаимно перемещаются. В теле непрерывно происходят изменения, даже при полном отсутствии внешних воздействий: выравнивается температура его частей, выравнивается концентрация различных веществ в объеме тела и т. д. В результате тело приходит в некоторое состояние равновесия. Физики говорят: «Трение, диффузия, теплопроводность приводят микроскопическую систему посредством термодинамически необратимых процессов в состояние с наибольшей энтропией в соответствии со вторым законом термодинамики». В этом состоянии равновесия микрочастицы—молекулы, из которых состоит тело, движутся, обмениваются местами, переносят энергию, но при этом в любом направлении происходят в среднем одинаковые перемещения. Поэтому макроскопически, т. е. в масштабе всего тела, не происходит ни переноса энергии, ни изменения концентраций. В этом и заключается сущность равновесного состояния или состояния с «наибольшей энтропией».

Живые организмы, в отличие от тел неживой природы, будучи предоставлены самим себе, не приходят в состояние равновесия. В них непрерывно продолжаются многочисленные и разнообразные физико-химические процессы.

Процесс «повышения энтропии», т. е. выравнивания температуры с окружающими телами и распада устойчивых структур начинается только после смерти организма. Это не означает, впрочем, что второй закон термодинамики недействителен для живых организмов. Они поддерживают постоянство своей структуры за счет поглощения и разложения пищи и за счет поглощения энергии извне. Если рассматривать, в соответствии со вторым законом термодинамики, «замкнутую систему», т. е. систему, полностью разобщенную с остальным миром, состоящую из организма, пищи и отходов, то в этой системе будет наблюдаться повышение энтропии. Однако энтропия самого живого тела остается приблизительно неизменной до момента смерти.

Различие в свойствах живых и неживых тел, отмеченное нами выше, определяется тем, что в живом организме в целом и в каждой живой клетке протекают физико-химические процессы, направление и ход которых в конечном итоге определяются молекулярными структурами ядра клетки. В живой клетке непрерывно происходит управление макроскопическими процессами со стороны микроскопических объектов. При этом устойчивость, неизменность, характерная для молекул как объектов микромира, проявляется в виде неизменности структур и постоянства процессов организма в макроскопических масштабах. Посредником являются процессы управления в живой клетке.

Постепенно выясняется все больше деталей этих процессов управления. Возможно, что ближайшие годы принесут нам знание новых интересных подробностей, а может быть и разгадку «тайны жизни», разгадку, полученную путем изучения процессов управления и передачи информации в живой клетке. Ведь именно эти процессы и делают клетку живой, устойчивой, устойчиво и однообразно функционирующей в окружающей среде до момента смерти.

Независимо от решения вопроса о том, существует ли непроходимая стена между живым и неживым миром и можно ли свести все поведение живых существ к сложным автоматизмам, мы можем утверждать, что сегодня нам известно большое число примеров, когда машина успешно выполняет сложные операции интеллектуального характера, заменяя человека в операциях управления энергетическими устройствами. Это обстоятельство породило общую систему взглядов на процессы связи и управления, которая включает в себя понятие и н ф о р м а ц и и.

Информация

Для всех систем, в которых протекают процессы управления, будь то устройства регулирования или живые организмы, характерна одна весьма общая черта: отдельные части этих систем связаны друг с другом таким образом, что они передают друг другу некоторые сообщения о процессах, в них происходящих, с помощью сигналов. Именно по этому признаку можно проследить глубокое сходство и единство процессов управления. Энергетические процессы, сопровождающие сигнализацию, играют второстепенную и непринципиальную роль. Важна не энергия, а сигнал. Чтобы показать на примере справедливость последнего замечания, зададим себе вопрос: чему равен коэффициент полезного действия телевизора или радиолокатора? Ответить на этот вопрос невозможно (как, впрочем, и на всякий неправильно поставленный вопрос) уже потому, что на выходе телевизора, радиолокатора и им подобных систем энергия как таковая не представляет никакого интереса. Назначение радиолокатора состоит не в отдаче энергии в той или иной форме, как это свойственно энергетической машине, а в решении совершенно иной задачи. И радиолокатор и телевизор потребляют энергию, и даже в значительных количествах, но отдают они не энергию, а информацию, сведения в виде сигналов.

Понятие информации очень широкое. Информацию переносят телеграф, телефон и радио. Информация записана на грампластинках, магнитных лентах, на фотоснимках и литографских оттисках. Информация передается с помощью человеческого языка устно или письменно, ее пересылают по почте, издаются в виде книг, газет и журналов, хранят в библиотеках. Информация заключена в отсчете измерительного прибора, в результатах контроля продукции, в числовых подсчетах, в математических формулах и таблицах. Наше зрение, слух, осязание приносят нам информацию о внешних событиях, внутренние органы обмениваются информацией, координируя свою совместную работу. Ничтожные количества химических веществ доставляют нам посредством обоняния и вкуса информацию о качестве пищи. Изменения физических величин (электрического напряжения и тока, электромагнитного поля, давления), механические перемещения вводят информацию в автоматические устройства и позволяют получить из них новую информацию.

Информация — это то, что несет на себе след какого-то факта или события, события, которое уже произошло или должно произойти, все то, что доставляет нам об этом факте сведения или сообщения. Создание, передача, хранение, использование и главным образом преобразование информации происходит и в машинах и в живых организмах по определенным строгим законам. Правила, по которым происходит преобразование информации, называют алгоритмами преобразования*. Законы существования и преобразования информации объективны и доступны изучению. Они интенсивно изучаются. Собственно, определение этих законов, их точное описание, использование алгоритмов преобразования информации, в особенности алгоритмов управления, и составляет содержание кибернетики.

Здесь уместно заметить, что точное определение содержания и границ таких наук, как кибернетика, затруднительно, и поэтому вокруг них до сих пор не затихает дискуссия. Точное определение границ любой науки обычно можно дать только после того, как эта наука вполне оформится. Этого пока нельзя сказать о совсем еще молодой кибернетике.

Кибернетика выросла на основе изучения конкретных процессов передачи сигналов, процессов управления и обобщения законов, по которым протекают эти процессы. По мере накопления и обобщения фактов, естественно, расширяется область применения уже изученных законов. Обилие приложений кибернетики иногда заставляет задавать вопрос: а что же не относится к кибернетике? Подобные вопросы вызваны, конечно, только новизной ситуации, так как аналогичный вопрос в отношении, например, математики никому не пришло бы в голову задать, хотя математика имеет не меньше областей применения, чем кибернетика.

Специалисты отдельных прикладных наук иногда отождествляют кибернетику со своей специальностью. Так, например, часто приходится слышать, что кибернетика — это теория автоматического регулирования (разумеется, сильно расширенная). Некоторые, увлекаясь наиболее эффективными перспективами, утверждают, что кибернетика — это наука о моделировании функций человеческого мозга. Подобные определения оказываются очень ограниченными.

Вряд ли попытки дать точное и строгое определение ки-

* Примером алгоритма может служить любая математическая формула.

бернетика, которое оказалось бы верным раз и навсегда, могут сейчас оказаться плодотворными. Однако водораздел между кибернетикой и «не кибернетикой» всегда нетрудно провести, если помнить, что интересы кибернетики лежат в области общих законов передачи информации, ее преобразования и использования для управления.

Можно сказать, что одной из основных задач кибернетики являются поиски строго формализованных алгоритмов преобразования информации и реализация этих алгоритмов.

Системы или устройства, имеющие дело с сигналами, воспринимающие, преобразующие, передающие, принимающие, хранящие, обрабатывающие или использующие информацию и работающие в соответствии с определенным алгоритмом, мы будем называть кибернетическими системами или устройствами.

Таким образом, электронная счетная машина является кибернетической машиной в отличие от паровой машины—машины энергетической.

Энергетические и кибернетические системы чаще всего существуют и работают совместно. Автоматическая энергоподстанция, беспилотный самолет, авторегулируемый производственный процесс могут служить тому примерами. В живом организме также сочетаются энергетическая и кибернетическая системы.

Энергетика и кибернетика идут рука об руку. И как энергия не может использоваться без управления ею, так и управление не может осуществляться помимо материального, физического процесса, без энергии, пусть даже и самых малых количествах.

Однако специфика и закономерности этих двух областей явлений природы различны, и это различие нужно очень четко видеть. Понятие информации сложилось позже, чем понятие энергии. И законы работы кибернетических систем познаны еще далеко недостаточно. В наши дни только закладываются основы их понимания.

Широкое поле уже существующих применений и настоящие необъятные перспективы развития кибернетики требуют быстрого движения вперед в познании законов кибернетики и их использовании.

Одним из основных понятий кибернетики является понятие сигнала. К разбору этого понятия мы и перейдем.



ГЛАВА 2

СИГНАЛ

Мы можем знать о событиях, в которых не принимали участия, свидетелями которых не были, отделенных от нас большими пространствами и промежутками времени. Мы можем знать об этих событиях полно и подробно, подчас подробнее, чем участники или свидетели этих событий.

Подобные знания мы получаем ежедневно, когда читаем газеты и журналы, изучаем историю, слушаем радиопередачи или смотрим телевизионную программу. Иной раз телевизионный экран позволяет нам заметить такие подробности футбольного матча, о которых не сумеют рассказать зрители, присутствовавшие на стадионе.

Попытаемся вдуматься в этот простой и привычный факт.

Если наши знания о каком-нибудь событии приходят к нам уже после того, как само событие закончилось, скажем, исход футбольного матча становится нам известным на другой день после матча, то это значит, что какие-то следы события продолжают существовать и тогда, когда само событие закончилось. Сохраняются не только ре-

зультаты и следствия данного события, но и незначительные подробности—забавный эпизод у ворот одной из команд, ошибка судьи, реакция публики и т. д. Все это может сохраниться в точном соответствии с тем, как оно происходило.

Событие может сохраняться в памяти очевидцев, в описаниях, в фотоснимках, в зарисовках, в кинофильмах, в звукозаписи самого события или репортажа о нем или в какой-либо иной форме. В большинстве случаев описание события соответствует самому событию более или менее полно. Смысл слова «соответствует» нам интуитивно ясен. Попросту говоря, описание события правдиво: на фотоснимке изображены фигуры игроков в тех позах, в каких они находились в тот момент, когда мяч летел в ворота; в рассказе очевидца сохранен порядок и характер действий игроков, который может быть лишь немного окрашен эмоциями рассказчика, и т. д.

Однако постараемся понять точнее, что значит «соответствовать», а главное, что чему соответствует. Прежде всего, фотография момента игры у ворот есть прямоугольный кусок плотной белой бумаги размером 9×12 см, покрытый слоем желатина, в который вкраплены мельчайшие зерна серебра. Все это никакого отношения к футболу не имеет. Не имеет к нему также отношения и то, что частицы серебра поглощают падающий на них свет, а белая бумага этот свет отражает. О футболе на фотоснимке говорит только распределение зерен серебра по поверхности бумаги. Оно определяет, в свою очередь, распределение потоков света, идущего от фотоснимка к глазу наблюдателя, и это распределение оказывается близким к тому, которое существовало в момент, когда был сделан снимок, в том месте, где стоял фотокорреспондент. При этом соответствие существует лишь приблизительное: не передан цвет, четко передано распределение интенсивности, отсутствует стереоэффект (снимок получился плоский, а не объемный) и т. д.

Отметим следующее: фотоснимок (заметка в газете, лента звукозаписи, кинолента, устный рассказ и т. д.) как физическое явление не имеет ничего общего с изображаемым событием. Он имеет самостоятельную физическую природу и самостоятельное существование в пространстве и времени. Упомянутое выше «соответствие» фотоснимка изображенному на нем событию заключается в том, что снимок при осуществлении некоторых условий,

производит на наблюдателя приблизительно такое же воздействие, как и само событие.

В сказанном заключается нечто общее для характеристики всех видов сигнала. Сигнал порождается некоторым событием, фактом, действием. Сигнал имеет самостоятельную физическую природу и самостоятельное существование в пределах некоторой организованной системы; он всегда воплощен в некотором материальном объекте или процессе. В этом виде сигнал может быть зафиксирован и существовать долгое время. Сигнал может быть передан на большие расстояния. Сигнал в конце своего существования может превратиться в действие или в событие. Он может воздействовать на живого наблюдателя или на мертвую физическую систему и породить некоторую реакцию этого наблюдателя или системы. Иногда, например, сигнал может вызвать ту же реакцию, какую произвело бы породившее этот сигнал событие при непосредственном воздействии. Так, узнав об удаче любимой футбольной команды из газет, мы обрадуемся почти так же, как обрадовались бы на стадионе при окончании матча.

Существование сигнала после окончания события позволяет ему не только пассивно «ждать» того момента, когда он вызовет действие, но и вести «полноценную жизнь», в том смысле, что, комбинируясь с другими сигналами, он может порождать новые сигналы, обладающие новыми свойствами, новой формой и несущие новое действие, новый смысл. Например, в результате вычислений из исходных данных задачи получается ее решение, возникает новый сигнал, соответствующий новому факту.

Разумеется, такое «самостоятельное» существование сигналы ведут отнюдь не самостоятельно, а только в пределах некоторой организованной системы, по законам, свойственным этой системе. Самостоятельность заключается лишь в независимости от породившего сигнал события, после того, как сигнал возник.

Если мы имеем событие и его подробное, правдивое описание, то по описанию мы можем восстановить событие. Это происходит, например, при изготовлении машины по чертежу, постройке дома по проекту, выполнении военной операции по плану и т. д.

Итак, между событием и сигналом при некоторых условиях существует однозначная связь, взаимное соответствие, в пределах, задаваемых степенью подробности описания.

Изоморфизм

Взаимное соответствие физически разнородных явлений часто встречается и используется.

Возьмем в качестве примера геометрические объекты—линии, поверхности, тела—в евклидовом пространстве, в котором мы сами существуем. Можно построить модели этих объектов, изобразить их на чертеже, представить себе их мысленно. Мысленно или реально эти объекты можно перемещать, видоизменять, сопоставлять и т. д., т. е. производить с ними некоторые операции. Мы можем, например, нарисовать на плоскости окружность определенного размера и пересечь ее прямой, определенным образом расположенной. Затем мы можем найти точки пересечения окружности и прямой.

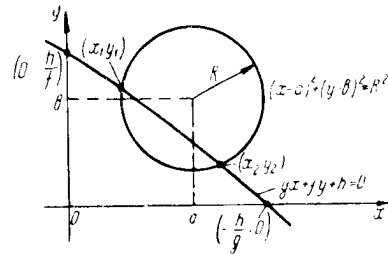


Рис. 2.1. Геометрический образ на плоскости (окружность и пересекающая ее прямая) и изоморфная аналитическая запись в виде уравнений в прямоугольных координатах.

Теперь сопоставим каждую точку плоскости с парой чисел x и y —координатами точки, например, в прямоугольной декартовой системе координат. Каждое число представляет собой расстояние этой точки от одной из осей.

Числа, удовлетворяющие уравнению

$$(x - a)^2 + (y - b)^2 = R^2, \quad (2.1)$$

соответствуют точкам окружности, имеющей радиус R и центр в точке $x_0 = a$; $y_0 = b$.

Числа, удовлетворяющие уравнению

$$gx + fy + h = 0, \quad (2.2)$$

соответствуют точкам, лежащим на прямой, проходящей через точки:

$$x' = 0; \quad y' = -\frac{h}{f} \quad \text{и} \quad x'' = -\frac{h}{g}; \quad y'' = 0.$$

Если уравнения (2.1) и (2.2) имеют общие решения (их может быть в действительных числах либо два, либо одно), то эти решения соответствуют точкам, лежащим одновре-

менно и на окружности, и на прямой, т. е. точкам пересечения прямой с окружностью. Если мы решим оба уравнения совместно, то точки пересечения будут найдены нами аналитически.

Производя вычисления с числами x и y , мы могли бы вовсе не думать о геометрических образах (и наоборот). Однако ясно, что любой геометрический образ и любая операция с ним имеют соответствующие образы в виде пар чисел и операции над ними. Мы ставим некоторую геометрическую задачу, затем переводим ее на аналитический «язык» и решаем эту задачу в числах, не вспоминая о геометрии. Дойдя до результата и снова переведя его на язык геометрии, мы получим тот же ответ, который получили бы, решая задачу геометрически—построением фигур. Мы получили правильный результат.

В математике существует понятие **изоморфизма** двух множеств X и Y элементов $x \in X$ и $y \in Y$.

Множества X и Y называются **изоморфными**, если выполняются следующие условия:

1) каждый элемент $x \in X$ может быть взаимно однозначно сопоставлен с элементом $y \in Y$, т. е. $x \rightarrow y$ и $y \rightarrow x$;

2) каждая операция f (из некоторого класса операций), преобразующая элемент $x_1 \in X$ в $x_2 \in X$ во множестве X , $f(x_1) = x_2$, может быть взаимно однозначно сопоставлена с операцией F , преобразующей элемент $y_1 \in Y$ в $y_2 \in Y$, $F(y_1) = y_2$, т. е. $f \rightarrow F$, $F \rightarrow f$;

3) если $x_1 \in X$ соответствует $y_1 \in Y$ и $x_2 \in X$ соответствует $y_2 \in Y$, если $f(x_1) = x_2$ и $f \rightarrow F$, то для всех x, y, f $F(y_1) = y_2$.

Два множества называются **изоморфными**, если их элементы попарно взаимно однозначно соответствуют друг другу и преобразования элементов одного множества соответствуют преобразованиям соответствующих элементов второго множества.

Если считать, что геометрические образы получаются непосредственно из практики как результаты измерений и операций производства, то аналитические операции—вычисления с числами на бумаге—являются **изоморфным** изображением реальных явлений. Каждое событие—операция с реальными объектами (из определенного класса объектов и операций)—может быть сопоставлено с записью и вычислением. И, наоборот, каждая разумная запись может быть сопоставлена с событием, которое можно воспроизвести. Длинный ряд вычислений, проведенных по определенным правилам, воспроизводящим реальные операции,

приводит к результату, который точно соответствует результату изоморфной этим вычислениям практической операции. Мы можем сопоставить оба результата и проверить правильность вычислений.

Можно назвать записи инженера, производившего расчет, сигналами реальных событий—операций производства. Система этих записей и вычислений изоморфна реальным объектам и операциям с ними. Можно говорить об изоморфизме теории и практики вообще, не только для геометрических построений, но и для других отраслей знаний.

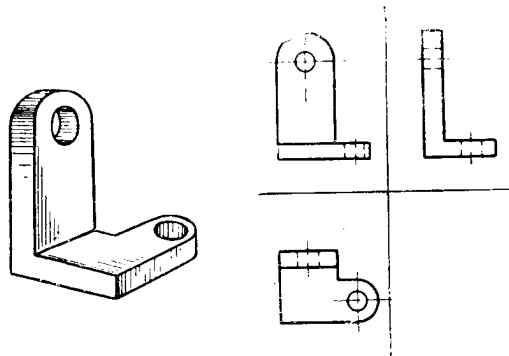


Рис. 2.2. Деталь и ее изображение в виде чертежа, находящиеся в изоморфном соответствии друг с другом.

Совокупность сигналов в кибернетической системе является изоморфным отображением некоторых сторон реального события. Мы говорим «некоторых сторон» потому, что никакое самое подробное описание события с помощью совокупности сигналов не исчерпает всех характеристик реального события. Описание с помощью сигналов отражает лишь некоторые стороны события, существенно важные для кибернетической системы, в пределах которой существует сигнал.

Характерным примером изоморфизма систем может служить часто упоминаемая аналогия различных видов колебаний. Физически колебания могут протекать в виде механического движения тел (механические колебания), изменений плотности и давления газообразной, жидкой или твердой среды (акустические колебания), изменений токов и напряжений в электрических цепях (электромагнитные колебания), изменений напряженности электромагнит-

ного поля (радиоволны, оптические процессы), изменений физиологического состояния живых тканей и т. д. Во всех этих физически конкретных ситуациях ход колебательного процесса имеет общие черты, которые могут быть прослежены очень далеко. Именно поэтому возможно существование общей теории колебаний, в формальном построении которой можно отвлекаться от размерностей величин и выводы которой в значительной части одинаково справедливы для всех конкретных случаев. Все виды колебаний объективно изоморфны друг другу, а теория колебаний с ее формальным математическим аппаратом изоморфна им всем.

Изоморфизм различных физических явлений позволяет осуществлять модели промышленных и технических устройств и протекающих в них процессов. Речь идет не о моделях судов, самолетов и т. п., т. е. о моделях с масштабным уменьшением размеров. Понятие моделирования несравненно шире и включает воспроизведение реального явления на изоморфной модели, которая отлична от воспроизводимого явления по своей физической природе. Само изучаемое явление может быть недоступно для изучения либо вследствие своих энергетических масштабов или масштабов времени (просачивание грунтовых вод, полет управляемого снаряда), либо вследствие недоступности для непосредственного наблюдения (работа сердца в живом организме). Изоморфная модель—механическая, электронная, гидравлическая—воспроизводит законы протекания реального процесса и делает доступной оценку его течения в широких пределах изменения условий.

Существенным в природе сигнала является не сам физический объект, в котором сигнал воплощен, а лишь тот факт, что сигнал представляет собой множество различных состояний физического объекта или процесса. Множества состояний для различных физических объектов или их подмножества могут быть изоморфно сопоставлены друг с другом и реально переведены одно в другое соответствующим физическим преобразованием. Так, из одной физической формы сигнала может быть получена другая форма. Именно таким образом множество значений звукового давления, упорядоченное во времени, переводится с помощью микрофона во множество значений электрического тока, а затем—во множество положений иглы звукозаписывающего аппарата и множество точек звуковой дорожки, упорядоченные в пространстве. Одно-

значное соответствие позволяет применить обратное преобразование и вернуться к прежнему множеству в прежнем, звуковом, воплощении сигнала.

При записи звуков с помощью нотных знаков или речи с помощью букв сопоставляются не мгновенные значения давления и символы записи, а комбинации значений давления (тон, фонема) и символы. При этом некоторые характеристики звука в записи теряются (тембр, интонация), отсутствует полное соответствие звуков нотам и тексту. Воспроизведение нотной записи и текста оказывается, та-

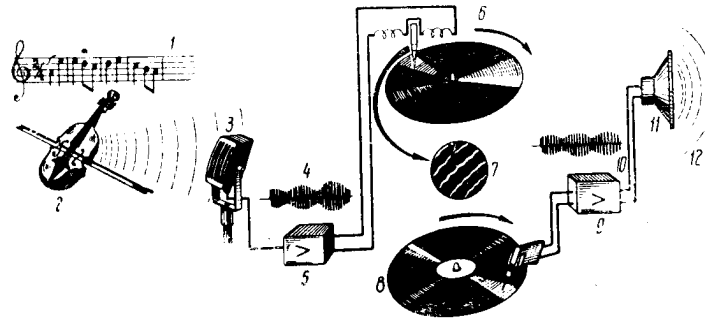


Рис. 2.3. Электрические преобразования при записи звука на диск. 1 — запись звука в виде нот, 2 — воспроизведение нот в виде звука на музыкальном инструменте, 3 — микрофон преобразует звук в электрические колебания (4), 5 — усилитель, 6 — записывающая головка преобразует электрические колебания в механические и наносит «звуковую дорожку» на вращающийся диск, 7 — звуковая дорожка на диске, 8 — воспроизводящая головка (игла) скользит по дорожке и преобразует ее рисунок в механические колебания, пьезоэлемент преобразует механические колебания в электрические, 9 — усилитель, 10 — усиленные электрические колебания, 11 — репродуктор преобразует электрические колебания в звуковые (12).

ким образом, оставленным в некоторой степени на произвол исполнителя. Различные национальные языки могут быть поставлены в соответствие друг другу так, как это делается при переводе. Каждому языку соответствует примерно одинаковое множество понятий, отражающих события индивидуального и общественного бытия. Все, что можно выразить на французском языке, можно сказать по-русски и наоборот.

Итак, сигнал есть изоморфное отображение некоторых сторон физического факта или события. Сигнал всегда существует в некотором физическом воплощении. Сигнал существует в пределах организованной кибернетической системы. Сигнал порождается в пределах этой системы под воздействием внешнего события или других сигналов. Сигнал существует самостоятельно, независимо от поро-

дившего его события, передается, фиксируется (запоминается), воспроизводится и в конечном итоге воздействует на кибернетическую систему, вызывая некоторое действие—реакцию. Ни физическая форма, ни количество энергии не имеют решающего значения. Значение, смысл сигнала заключается в том, что имеется соответствие сигнала и события или реакции.

Таким образом, смысловое или семантическое содержание понятия информации, по-видимому, исчерпывается изоморфным соответствием внешнего события и сигнала.

Имея, с одной стороны, некоторое событие или явление и множество возможных «состояний» этого явления и, с другой стороны, сигнал и множество его состояний или символов и установив соответствие элементов этих двух множеств, мы тем самым припишем каждому сигнальному символу его «смысл», т. е. соответствующее ему конкретное состояние описываемого события.

Посылка определенного сигнала представляет собой выбор одного символа или элемента из всего множества возможных символов. Этот выбор производится тем внешним событием, с которым сигнал находится в соответствии.

Особенно нужно отметить тот факт, что сигнал как физический процесс, несущий на себе информацию, т. е. сохраняющий соответствие с тем или иным событием, всегда существует только в пределах некоторой организованной системы. Эта система, в силу своих свойств, позволяет реализовать все свойства сигнала, из которых главным является то, что сигнал используется для управления, т. е. для создания реакции или ответа на изменение внешней обстановки. Вне организованной системы сигнал может существовать и сохранять свои физические свойства, но теряет все свойства сигнала, утрачивается информация, связь с событием и реакцией, утрачивается возможность «понять» сигнал, использовать его для выработки реакции.

Кибернетическая система, в пределах которой существует данный сигнал и которая, таким образом, обеспечивает условия его существования, должна быть заранее организована, т. е. построена по некоторому плану и иметь определенную степень сложности. Иначе говоря, структура системы уже включает в себе некоторую информацию, которую мы назовем «предварительной» по отношению к «текущей» информации, которую несет на себе сигнал, существующий в данной системе.

Возникновение сигнала и реакция

Внутренность шкафа термостата стала нагреваться вследствие теплопроводности стенок. Поток тепла грозит нарушить постоянство температуры, которое должно тщательно поддерживаться. Небольшая часть этого тепла нагрела чувствительный элемент термореле. Реле замкнуло контакты, в цепи возник ток — сигнал опасного перегрева. Этот сигнал вызовет включение охлаждающего устройства, и равновесие будет восстановлено.

При возникновении пожара ничтожная часть выделяемого тепла воздействует на термочувствительный прибор

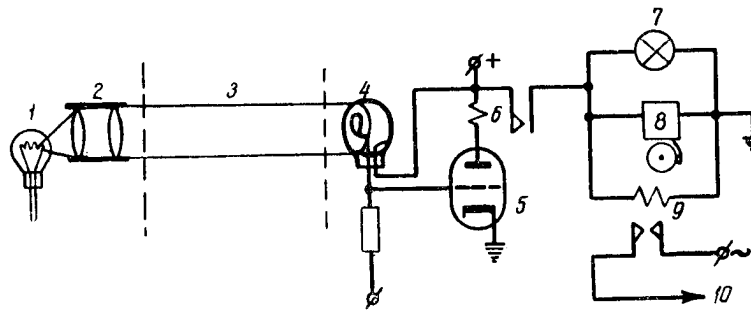


Рис. 2.4. Сторожевое фогореле.

1 — осветитель, 2 — линза, 3 — луч света, пересекающий охраняемое пространство, 4 — фотозлемент, 5 — ламповый усилитель, 6 — реле, 7 — световой сигнал, 8 — звуковой сигнал (звонок), 9 — контактор мощного исполнительного устройства, 10 — провод к исполнительному устройству.

и преобразуется в нем в сигнал тревоги или в сигнал, включающий систему спринклеров, разбрызгивающих воду на очаг пожара.

Бывают случаи, когда сигнал по своей энергии не только не меньше, а гораздо больше того события, появление которого он отмечает. Так, счетчик Гейгера — Мюллера отмечает импульсом электрического тока пролет через его объем частицы космического или радиоактивного излучения. В этом случае сигнал по своей энергии во много раз превышает само событие.

Свет, падающий на сетчатку глаза животного, вызывает появление возбуждения нервных окончаний, которое соответствует характеристикам падающего света и является сигналом внешнего события для организма. Нервное возбуждение передается в мозг, где в результате вырабатывается реакция на внешнее раздражение.

Любая рефлекторная деятельность организма начинается с ощущения, т. е. с появления сигнала нервного возбуждения, возникающего под действием внешнего или внутреннего события. Все впечатления от внешнего мира суть сигналы совершающихся вокруг событий.

Внутренние процессы в живом организме также сопровождаются появлением сигналов, даже в тех случаях, когда ощущения отсутствуют. Работа мышц вызывает расширение кровеносных сосудов, вследствие чего падает кровяное давление. Падение кровяного давления отмечается раздражением чувствительных нервных окончаний, расположенных в большом числе в стенках периферических сосудов, и в виде сигнала нервного возбуждения передается по так называемым прессорным нервам к нервным центрам продолговатого мозга. Эта сигнализация не сопровождается субъективным ощущением.

Можно привести еще много примеров, показывающих, как в некотором элементе организованной системы (реле, фотоэлемент, нервное окончание) внешнее событие возбуждает сигнал.

Не менее богата примерами и область явлений, в которых сигнал превращается в действие и вызывает течение энергетических процессов большой мощности.

Простейшим примером из этой области является пуск, остановка и регулирование хода мощных моторов, обслуживающих электровоз, подъемный кран или прокатный стан, с помощью вспомогательных маломощных цепей управления. Нажатие кнопки или поворот рукоятки управления является сигналом, который по своей мощности не способен привести в движение мощный мотор, но который достаточен для того, чтобы включить реле или контактор, и замыкание контактов силовой цепи произведет желаемое действие. Мотор будет вращаться за счет энергии источника питания. Сигнал же способен только направить энергию источника в желаемое русло. В данном случае основным элементом системы, позволяющим сигналу превратиться в действие, является контактор, который можно рассматривать как нелинейный усилитель мощности сигнала.

Любое дистанционное управление, осуществляемое передачей сигналов по радио или по проводам, относится к тому же кругу явлений. В московском Политехническом музее одно время демонстрировалось включение моторного привода от фотоэлемента. Исходный сигнал включения создавался посетителем, пересекавшим луч света, падаю-

ний на фотоэлемент. В этом случае сигнал представлял собой не поступление порции энергии, а прекращение подачи ее на фотоэлемент. Этот пример убеждает нас в том, что для сигнала существенным является не наличие энергии, а изменение ее поступления, т. е. различие состояний.

Аналогичные явления наблюдаются и в работе нервной системы. Нервное возбуждение может возникать не только при увеличении энергии воздействия (появление света, звука, давления), но и при ее уменьшении или исчезновении.

Все движения живого организма, произвольные или непроизвольные, являются результатом сокращения мышц или групп мышц, а эти последние происходят только под воздействием очень слабых сигналов нервного возбуждения, приходящих по нервным волокнам. Сами сокращения происходят за счет энергии, накопленной в мышечных волокнах.

Сигналы порождают сигналы

Появление первоначального сигнала и реакция системы суть начальный и конечный этапы существования сигнала. Между своим «рождением» и «смертью» сигнал может вести долгую жизнь, полную событий. Далеко не всегда сигнал сейчас же вслед за появлением вызывает реакцию. Часто он существует долго вместе с другими сигналами, с которыми он может взаимодействовать по определенным законам, порождая новые сигналы. Бесчисленные примеры такого рода можно привести из работы решающих устройств. Еще более богата подобными примерами работа нервной системы живого организма. Нервные процессы сопровождаются таким тесным переплетением воздействий многих сигналов, что поистине приходится потрудиться, чтобы выделить отдельный акт возбуждения.

Рассмотрим самый простой пример. Тяговый мотор электропоезда при разгоне, на малых оборотах, потребляет повышенный ток. Чтобы его ограничить последовательно включают пусковой реостат. По мере разгона растет противоэлектродвижущая сила, ток падает, реостат выводят, замыкая накоротко его секции с помощью контакторов. В системе управления имеется так называемое реле ускорения, которое регулирует режим пуска. Оно получает один сигнал от контроллера машиниста, а другой — от мотора. Первый сигнал несет команду: «увеличить ско-

рость» или «замкнуть очередную секцию реостата». Второй сигнал несет информацию: «ток мотора достиг нормы и может быть увеличен». Отсутствие сигналов имеет значение, противоположное указанным. Под воздействием комбинации двух сигналов — «нужно» и «можно» — реле срабатывает и включает контактор, замыкающий секцию реостата. Если хотя бы один из сигналов отсутствует (либо «не нужно», либо «нельзя», либо и то и другое), реле не срабатывает и поезд не увеличивает скорости.

В коммутаторе АТС при наборе занятого номера два сигнала—набор номера и снятая трубка—порождают новый сигнал—короткие гудки «занято».

Если неожиданно уколоть человека в плечо, он произвольным движением отдернет руку. Нервное раздражение прошло по нервным волокнам в спинной мозг и обратно к мышцам, замкнув рефлекторную дугу, и вызвало сокращение мышц. Теперь предположим, что человеку назначена врачом подкожная инъекция. Ясно, что в этом случае при уколе пациент не отдернет (или почти не отдернет) руки. «Акт воли», выраженный в нервном сигнале, пришедшем от головного мозга, в комбинации с болью укола — другим нервным сигналом, вызвал новый сигнал и новое действие — торможение рефлекторного движения.

Подобные примеры можно без труда умножить и усложнить.

Полнота описания

Итак, сигнал есть физический процесс, существующий в кибернетической системе, имеющий множество состояний, которые он принимает в соответствии с внешними событиями, воздействующими на систему. Основным свойством сигнала является то, что он несет информацию о со-

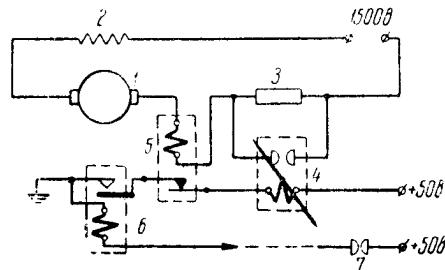


Рис. 2.5. Схема реле ускорения электропоезда.

1 — тяговый мотор, 2 — обмотка возбуждения мотора, 3 — секция пускового реостата, 4 — контактор, замыкающий реостат накоротко при разгоне поезда, 5 — обмотка реле, соединенная последовательно с мотором, и контакты, работающие на размыкание, 6 — обмотка реле, возбуждаемая от контроллера машиниста, и контакт, работающий на замыкание, 7 — контакт контроллера машиниста.

бытия, т. е. находится в соответствии с ним. Одна и та же информация может заключаться в различных формах и, наоборот, один и тот же сигнал может нести на себе различную информацию.

Очень часто в теории информации различают понятия «сообщения» и «сигнала». Нам представляется весьма существенным тот факт, что сообщение никогда не существует иначе, как в виде сигнала события, т. е. в виде материального объекта.

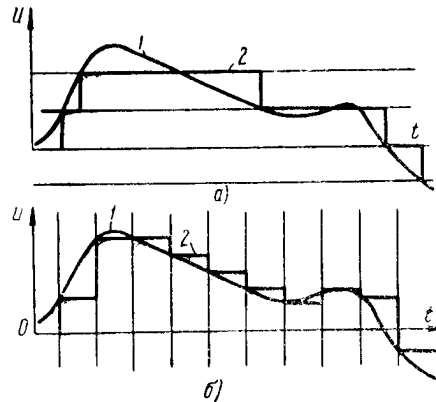


Рис. 2.6. Квантование непрерывного сигнала.

а) квантование по уровню, б) квантование по времени: 1 — непрерывный сигнал, 2 — квантованный сигнал.

более обширное множество сообщений (около ста), сообщив счет. Можно наращивать подробности в любом количестве. Таким образом, многообразие реальных событий ставится в соответствие или отображается на меньшее по объему многообразие сообщений в виде сигналов. Одиночное описание неполно, и, только продолжая описание как угодно долго, мы можем приблизиться к полному соответствию описания реальному событию.

Поскольку сигнал существует в кибернетической системе, свойства которой полностью определяют и возможные свойства сигнала, постольку полнота описания события с помощью сигнала зависит от свойств системы. Система воспринимает только те стороны события, которые она способна прямо или косвенно «ощутить». Простейшие системы с ограниченной сложностью структуры способны только к примитивному и неполному восприятию и воспроизведению («описанию») события. Так, например, работа термостата связана только с одной единственной величи-

ной—температурой; другие физические величины—давление, масса, ускорение и т. п.— не имеют смысла для этой системы в том отношении, что сигнал их не отображает (хотя они косвенно или прямо влияют на работу термостата). Другие, более сложные системы, обладают большими возможностями в смысле описания событий.

Кроме качественной стороны—способности систем воспринимать различные физические свойства и стороны явлений,— полноту описания события с помощью сигналов можно рассматривать также и с количественной стороны, оценивая числом «количество информации» о данном событии, заключенное в сигнале. Этот вопрос составляет содержание теории информации, которой мы коснемся в гл. 4.

Иногда говорят, что сигнал, описывающий внешнее воздействие, несет «осведомительную информацию», а сигнал определяющий реакцию системы—информацию «исполнительную». Однако этими двумя типами не исчерпываются все возможные сигналы. Существуют сигналы, порождаемые осведомительной информацией и порождающие в свою очередь сигналы исполнительные.

Бывают случаи, когда вообще затруднительно сопоставить некоторые сигналы с определенными событиями или реакциями. Например, инструментальная музыка, которая заполняет эфир в радиопередачах и которая, несомненно, имеет все признаки сигнала, не позволяет сразу и просто указать, какому именно событию или реакции она соответствует. Затруднение это, однако, не принципиальное, а скорее «техническое». Музыка, рассматриваемая как сигнал, несомненно, несет информацию и ей соответствуют эстетические переживания слушателей. Действие музыки можно наблюдать, рассматривая нервные механизмы, управляющие эмоциями человека. Эти механизмы работают закономерно, но мы сравнительно мало знаем об этих закономерностях. Тем не менее мы можем, хотя бы качественно, проследить связь между эмоциями и физиологическим состоянием нервной системы. Слезы восторга являются лишь крайним проявлением объективного действия, которое оказывает музыка как сигнал на человека как кибернетическую систему. Мы не можем проследить точно детали и оттенки этого действия, так как еще не все знаем о работе нервной системы. Современная эстетика весьма далека от того, чтобы изучать произведения искусства как звено цепи «событие—сигнал—реакция», и подобное предположение большинству искусствоведов показалось бы, с непривычки, совершенно неприемлемым. Однако,

для по этому пути, можно постепенно поставить эстетику в ряд точных наук, оперирующих количественной мерой и точными критериями. В настоящее время мы весьма далеки от этого, чему не приходится удивляться, так как трудности на этом пути необычайно велики.

Итак, сигнальное сообщение о событии всегда беднее, чем само событие. Сигнал описывает только некоторые стороны события, существенные для кибернетической системы. При возникновении сигнала происходит примерно то же, что случается с толстым романом, когда его излагают на иностранном языке для легкого чтения.

Дискретные и непрерывные сигналы

Событие или одна из его сторон, описываемая сигналом, может иметь характер либо непрерывно изменяющейся величины, либо состоять из резко различающихся ступеней или градаций, т. е. иметь дискретный характер. Число посетителей выставки может быть только целым числом. Количество жидкости может принимать любые, не только целочисленные значения. Различие в характере событий отражается и на структуре сигналов. Сигналы могут быть дискретными или непрерывными. В первом случае они состоят из отдельных, резко различных символов, во втором — два сигнала могут отличаться друг от друга сколь угодно мало. Речь, записанная в виде текста, или число, записанное цифрами, представляют собой примеры дискретного сигнала. Отсчет по шкале амперметра или число, отложенное на шкале счетной линейки, являются сигналами непрерывной структуры.

Дискретный сигнал может быть без труда выражен с помощью сигнала непрерывного, так как среди возможных значений непрерывных величин содержатся и все дискретные. Обратный переход может быть совершен только приближенно, с округлением до ближайшего дискретного значения. Во всех реальных случаях, однако, можно с достаточной точностью сопоставить непрерывные и дискретные сигналы, если использовать достаточное число дискретных символов.

Непрерывный процесс, например телефонный разговор, может быть представлен при определенных условиях в виде набора дискретных величин. Это утверждение впервые было использовано В. А. Котельниковым для анализа передачи сигналов по каналам связи.

Пусть сигнал представляет собой некоторую величину $y(t)$, непрерывно меняющуюся во времени от момента $t=0$ до $t=T$. Его можно представить в виде разложения в ряд Фурье на интервале от 0 до T

$$y(t) = a_0 + \sum_{k=1}^{\infty} \left\{ a_k \cos 2\pi k \frac{t}{T} + b_k \sin 2\pi k \frac{t}{T} \right\}, \quad 0 < t < T, \quad (2.3)$$

где a_0, a_k, b_k — коэффициенты Фурье, определяемые выражениями:

$$a_0 = \frac{1}{T} \int_0^T y(t) dt,$$

$$a_k = \frac{2}{T} \int_0^T y(t) \cos 2\pi k \frac{t}{T} dt,$$

$$b_k = \frac{2}{T} \int_0^T y(t) \sin 2\pi k \frac{t}{T} dt.$$

Формулу (2.3) можно переписать в виде

$$y(t) = c_0 + \sum_{k=1}^{\infty} c_k \cos \left(2\pi k \frac{t}{T} - \theta_k \right), \quad (2.4)$$

где $c_0 = a_0$, $c_k = \sqrt{a_k^2 + b_k^2}$, $\theta_k = \arctg \frac{b_k}{a_k}$.

Множество величин $c_0, c_1, c_2, c_3, \dots, c_k$ называется амплитудным спектром функции $y(t)$. Каждая из величин c_k представляет собой амплитуду колебаний с частотой $f_k = \frac{k}{T}$. Сложение колебаний дает первоначальную периодическую функцию $y(t)$ с периодом T .

Если функция в виде физической величины передается через реальный канал связи, в составе которого всегда имеются резонансные элементы, то амплитуды некоторых составляющих колебаний изменятся. Колебания очень высоких частот полностью подавятся, и амплитуды их окажутся равными нулю. Обычно об этом факте в технике связи говорят, как об „ограниченной частотной полосе пропускания“ канала связи. Ширина полосы пропускаемых частот Δf определяется как разность между самой высокой и самой низкой из пропускаемых частот. В простейшем случае

$$\Delta f = f_{\max} - f_{\min}.$$

Если $f_{\min} = 0$, то $\Delta f = f_{\max}$.

Реальный сигнал всегда имеет ограниченный спектр, т. е. содержит в спектре лишь те частоты, которые лежат в полосе пропу-

сказания канала. Таким образом, сигнал с ограниченным спектром Δf на ограниченном интервале времени T может быть полностью описан перечислением величин всех амплитуд c_k и фаз θ_k или же всех коэффициентов a_k и b_k , где $k=0, 1, 2, \dots, n$, а n конечное целое число. Поскольку самая низкая частота (кроме постоянной составляющей $f_0 = 0$) равна $f_1 = \frac{1}{T}$, а самая высокая

$$f_n = n f_1 = f_{\text{макс}} = \frac{n}{T},$$

то число гармоник, необходимое для характеристики данной функции, равно

$$N = 2n = 2f_{\text{макс}} T = 2\Delta f T,$$

где Δf — ширина спектра функции (полоса пропускания канала).

Можно описать непрерывный процесс с помощью дискретных величин иначе, указав мгновенные значения функции $y(t)$ в определенные моменты времени. Можно показать, что для полного определения функции достаточно задать ее значения в моменты, отстоящие друг от друга по времени на величину $\tau = \frac{1}{2\Delta f}$. Число таких „вы-

борочных значений“ равно, очевидно, $N = \frac{T}{\tau} = 2\Delta f T$. Если, таким образом, заданы дискретные значения функции $y(t)$ с ограниченным спектром и на ограниченном интервале времени T :

$$y(0), y(\tau), y(2\tau), \dots, y(k\tau), \dots, y[(N-1)\tau],$$

то любое значение функции, в том числе и промежуточное, может быть получено по формуле

$$y(t) = \sum_{k=0}^{N-1} \frac{\tau}{T} y(k\tau) \frac{\sin \frac{\pi}{\tau} (t - k\tau)}{\sin N \frac{\pi}{\tau} (t - k\tau)}. \quad (2.5)$$

И в этом случае набор $N = 2\Delta f T$ дискретных величин $y(k\tau)$ полностью описывает непрерывный процесс $y(t)$.

Представление непрерывного сигнала в виде дискретного часто называют «квантованием» сигнала. Практически это производится таким образом, что непрерывная физическая величина заменяется конечным множеством ее значений («квантовых уровней»). Различают два случая: «квантование по уровню» и «квантование по времени». При квантовании по уровню дискретный сигнал остается неизменным до момента, когда значение непрерывного сигнала достигнет некоторой заданной величины; это может произойти в любой момент времени. В случае квантования по времени фиксируются те величины непрерыв-

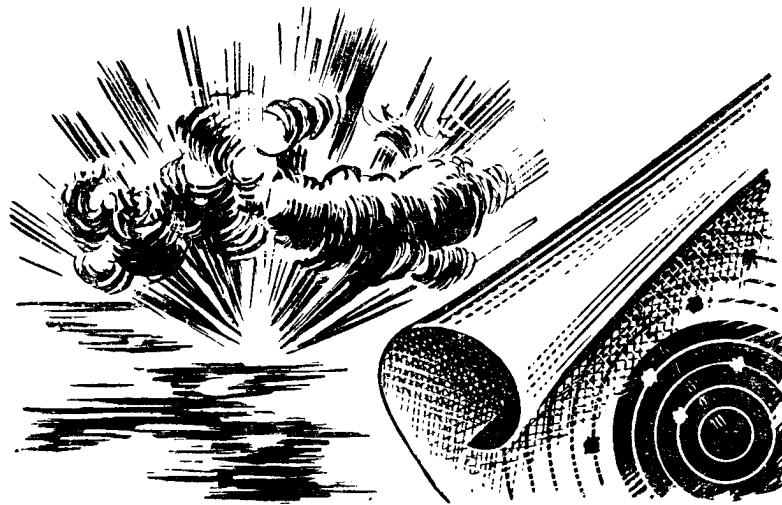
ного сигнала, которые существуют в заданные мгновения времени. При квантовании по времени можно «округлять» дискретную величину до ближайшего квантового уровня.

Итак, в большинстве случаев можно подобрать достаточно близкое выражение непрерывного сигнала посредством сигнала дискретного. Это верно в той же степени, в какой любое число, например иррациональное, может быть выражено с достаточно малой ошибкой десятичной дробью с конечным числом разрядов.

Допустимые искажения сигнала при квантовании определяются, с одной стороны, допустимой неточностью воздействия на систему управления, а с другой стороны, — еще одним очень серьезным обстоятельством — наличием шума. Шумом называются случайные воздействия и искажения, которые испытывает сигнал при передаче, возникновении или хранении. В присутствии шума тонкие различия сигналов стираются и сигналы становятся неразличимыми. Уровень шума кладёт, таким образом, предел различимости отдельных символов сигнала.

Поскольку шум всегда сопутствует сигналу, можно сказать, что дискретный сигнал практически всегда может быть сделан равноценным заменителем непрерывного.

О природе и характеристиках шума мы скажем несколько слов в следующей главе.



ГЛАВА 3

СЛУЧАЙ

Случайные события

В тексте бывают опечатки. В речи встречаются оговорки, неточности, замички. На фотографиях, на кинокадрах заметны пятнышки, которых не было на объекте съемки. В наушниках радиоприемника, в телефонной трубке слышен шорох. На экране телевизора мелькают посторонние полосы... Все перечисленное представляет собой искажение сигнала случайными воздействиями—шумами или помехами.

При наложении помехи на сигнал мы получаем вместо информации о событии нечто такое, что не имеет отношения к событию. Информация разрушается. Выражаясь вольно, помеха есть враг сигнала. Но одновременно она связана с сигналом и «родством». И помеху, и сигнал можно рассматривать как случайный процесс. Познакомимся с понятием «случая».

Если много раз повторять какое-либо измерение, наблюдение, испытание, всякий раз пытаюсь точно воспроизвести одни и те же условия, то, вообще говоря, вместо того, что-

бы получать одни и тот же результат, при каждом отдельном измерении будет получаться результат, отличный от прочих. На результат каждого испытания оказывают воздействие не только те условия, которые мы воспроизвели, но и те, которые мы воспроизвести не в состоянии. Они лежат либо вне нашего внимания, либо находятся не в нашей власти. Воспроизводя условия испытания, мы не можем вернуть весь мир в первоначальное состояние. Воспроизведению поддаются лишь некоторые из большого числа причин, влияющих на результат.

Событие, подверженное такому разбросу, мы называем случайным событием. Случайный характер явления или события порождается причинами, лежащими вне наблюдаемого события или явления. Чаще всего наличие большого числа посторонних причин приводит к тому, что явление приобретает характер случайности. Так, при измерении малых токов весьма чувствительным гальванометром отклонение легкой измерительной катушки будет зависеть не только от величины тока, ради которой и предпринято измерение, но и от броуновского движения катушки, которое вызвано ударами о катушку большого числа молекул окружающего газа.

Иногда явление приобретает случайный характер от того, что результат сильно зависит от начальных условий, и достаточно лишь очень немного, случайным образом, изменить эти условия, чтобы результат приобрел большой разброс. Это имеет место при запуске рулетки или при бросании игральной кости. Бросание кости, монеты, сдача игральных карт из перетасованной колоды, пуск рулетки — типичные случайные события, которые используются в играх.

Если случайное событие представляет собой появление одной из нескольких возможных величин (выпадение числа очков игральной кости, падение снаряда на некотором расстоянии от цели), то говорят о случайной величине или о величине, принимающей случайные значения.

Имея дело со случайными событиями или величинами, никогда нельзя быть уверенным в получении определенного результата в данном испытании. Однако в серии одинаковых опытов выявляются специфические закономерности, и, следовательно, случайные явления не представляют собой полного сумбура и хаоса, о котором невозможно сказать что-либо убедительно и достоверно. Каждому случайному явлению можно приписать определенные чис-

ловые характеристики, которые позволяют отличать случайные явления друг от друга, предсказывать характеристики сложных событий по характеристикам простых и решать большое количество задач, полезных для практики. Изучением закономерностей случайных событий занимаются теория вероятностей и математическая статистика — разделы математики, тесно связанные с естествознанием.

Теория вероятностей формулирует законы массовых явлений. Эти законы тем более точны, чем к большему числу однородных явлений они относятся. О случайных величинах, взятых в большом числе, возможно точное знание

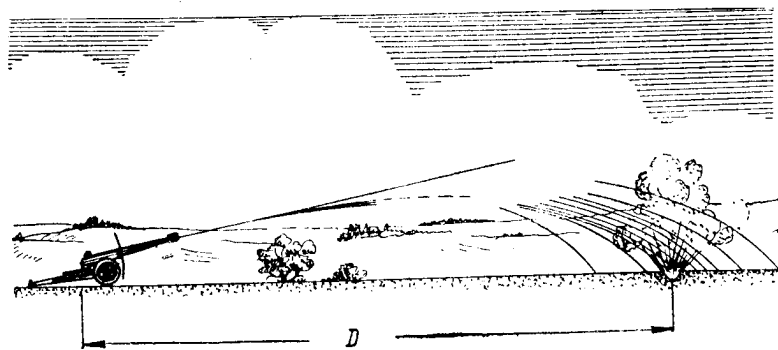


Рис. 3.1. Пример случайной величины — дальность D точки падения артиллерийского снаряда.

Так, законы газового состояния (например, закон Бойля — Мариотта) являются законами статистическими и в то же время точными, если речь идет о достаточно больших числах молекул. Если же переходить ко все более малым объемам, содержащим все более малые числа молекул, то в конце концов закон Бойля—Мариотта для мгновенных значений давления газа перестанет выполняться. С манометром, позволяющим измерить давление, производимое газом на площадку очень малых размеров, произойдет примерно то же, что и с гальванометром: он будет испытывать случайные воздействия небольших групп молекул — флуктуации давления, которые будут выражены тем сильнее, чем меньше будут размеры чувствительного элемента манометра.

Сигнал всегда испытывает более или менее сильные случайные искажения, вызываемые помехами и шумами. Искаженный помехой сигнал на приемном конце канала

связи перестанет быть однозначно связанным с переданным сигналом и с событием, о котором он несет информацию. Это является одной из причин того, что теория информации строится, в основном, на базе статистики и теории вероятностей. Вторая причина заключается в том, что для системы, воспринимающей сигнал, появление того или другого сигнала является также случайным. К последнему утверждению мы вернемся в гл. 4.

Вероятность

Основным понятием теории вероятностей является понятие вероятности случайного события. При реализации некоторых условий случайное событие может либо произойти, либо не произойти. Так, например, если мы, с некоторой экспозицией τ , сделаем фотографический снимок объема камеры Вильсона*, то, проявив пластинку, мы либо обнаружим фотографию следа ионизирующей частицы, пролетевшей через объем камеры (событие наступило), либо не обнаружим ее (событие не наступило).

Повторяя опыт много раз в одинаковых условиях, мы будем получать случайные результаты. Подсчитаем, какая часть опытов дала нам фотографии следов. Для этого разделим число полученных фотографий следов на общее число истраченных пластинок, т. е. возьмем отношение числа наступлений события n к полному числу испытаний N . Полученное отношение $\frac{n}{N}$ называется частотой появления события.

Наблюдая многократные повторные опыты (или серии опытов, проводимые в большом числе одновременно), иначе говоря, наблюдая массовые явления, нередко приходится отмечать, что частота наступления события от одной серии опытов к другой остается приблизительно постоянной, мало отличаясь от некоторого среднего значения. Отклонения от среднего оказываются обычно тем меньше, чем больше число испытаний. Очевидно, что для таких явлений объективно существует некоторая средняя частота, которая может быть с большей или меньшей точностью определена в серии опытов.

* Камера Вильсона представляет собой прибор для наблюдения траекторий ионизирующих частиц (космического излучения, продуктов атомного распада). В камере создается атмосфера пересыщенных паров, которые конденсируются на ионах, созданных пролетевшей частицей вдоль ее пути. Следы частиц мы наблюдаем в виде белых нитей тумана, которые могут быть сфотографированы.

На основе этого опытного факта в теории вероятностей строится математическая модель явления в виде системы аксиом. В математике многие системы аксиом строятся подобным же образом. Так аксиомы геометрии Эвклида являются не чем иным, как математической моделью реальных геометрических построений, моделью, которая ценна только до тех пор, пока она соответствует практике.

Итак, теория вероятностей принимает в качестве аксиомы, что каждому случайному событию A может быть поставлено в соответствие неотрицательное число $P(A)$, на-

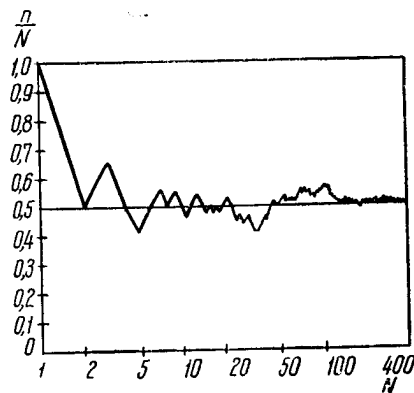


Рис. 3.2. Частота выпадения герба в последовательности случайных бросаний монеты. n/N — отношение числа выпадений герба n к общему числу бросаний N . Шкала N по оси абсцисс отложена в логарифмическом масштабе. Вероятность выпадения герба при одном бросании равна 0,5.

зываемое его вероятностью. При большом числе испытаний частота наступления события A стремится по величине к вероятности $P(A)$.

Если некоторое событие U наступает в заданных условиях всегда, то его частота, а следовательно, и вероятность равны единице, $P(U) = 1$. Такое событие U называется достоверным событием. Если событие V никогда не наступает в заданных условиях, то его частота и его вероятность равны нулю, $P(V) = 0$. Событие V называется невозможным событием.

Частота наступления и вероятность любого случайного события лежат между нулем и единицей, $0 \leq P(A) \leq 1$.

В примере с наблюдением следов частиц в камере Вильсона достоверным событием будет появление на фотографии изображения координатной сетки, нанесенной на стекло камеры, и изображения самой камеры (мы считаем, что в каждом опыте фотоаппарат исправен, пластинки доброкачественны и освещение достаточно). Невозможным

событием явилось бы непоявление изображения камеры. Вероятность появления следа частицы мы приблизительно определим, сделав много снимков и подсчитав частоту появления изображения следа на снимке. При этом мы предполагаем, что источник частиц остается постоянным и находится на неизменном расстоянии от камеры.

Теория вероятностей делает возможным, зная вероятности простых или элементарных случайных событий, определять вероятность сложных событий. Что такое сложное событие?

Пусть, например, мы имеем не одно, а два устройства, фиксирующих пролет ионизирующих частиц. Это могут быть не камеры Вильсона, а счетчики Гейгера—Мюллера, которые дают при пролете частицы через их объемы импульсы тока, фиксируемый автоматическим регистратором. Назовем событием A_1 срабатывание первого, а событием A_2 —срабатывание второго счетчика в течение времени t . Расположим счетчики так, чтобы одна и та же частица не могла последовательно пролететь оба счетчика. Кроме того, введем несколько искусственное предположение, что одновременный вылет двух частиц не случается, и за короткое время наблюдения t две различные частицы наблюдаться не могут. Таким образом, всегда дело идет о регистрации одной частицы первым A_1 или вторым A_2 счетчиком.

Зададим теперь вопрос: какова вероятность того, что в течение времени наблюдения t сработает любой из двух счетчиков? Иначе говоря, какова вероятность сложного события, состоящего в пролете частицы через первый или второй счетчик?

Два события A_1 и A_2 такие, что появление одного из них исключает возможность появления другого, называются несовместимыми. В нашем случае несовместимыми являются пролеты частицы через оба счетчика одновременно. Сложное событие, состоящее в наступлении хотя бы одного из событий A_1 и A_2 , называется суммой событий A_1 и A_2 (обозначается $A_1 + A_2$).

Вероятность суммы несовместимых событий равна сумме вероятностей каждого

$$P = (A_1 + A_2) = P(A_1) + P(A_2) \quad (3.1)$$

(аксиома сложения вероятностей).

Это соотношение понятно без пояснений, если вместо вероятностей подставить частоты событий.

Событие, состоящее в ненаступлении события A (обозначается \bar{A}), называется дополнительным к событию A и, очевидно, является с ним несовместимым. Поскольку в каждом испытании событие A либо наступает, либо не наступает, т. е. наступает событие \bar{A} , то на основании (3.1) будем иметь

$$P(A + \bar{A}) = P(A) + P(\bar{A}) = 1 \quad (3.2)$$

как вероятность достоверного события. Отсюда, вероятность дополнительного события (непоявления следа на фотографии) равна

$$P(\bar{A}) = 1 - P(A). \quad (3.3)$$

В систему аксиом теории вероятностей входят следующие:

Аксиома 1. Каждому случайному событию A поставлено в соответствие неотрицательное число $P(A)$, называемое его вероятностью.

Аксиома 2. Вероятность достоверного события равна единице

$$P(U) = 1.$$

Аксиома 3 (аксиома сложения). Если события A_1 и A_2 несовместимы, то

$$P(A_1 + A_2) = P(A_1) + P(A_2).$$

Пусть теперь два счетчика находятся на большом расстоянии друг от друга и каждый фиксирует пролет частиц от отдельного источника. Пролет частицы через счетчики № 1 и 2 в течение интервала времени τ снова имеет некоторые вероятности $P(A_1)$ и $P(A_2)$. На этот раз пролет частицы через первый счетчик не исключает возможности одновременного пролета другой частицы через второй счетчик и наоборот. Более того, пролет частицы через один из счетчиков не изменяет вероятности пролета через другой.

Если наступление или ненаступление одного из событий не меняет вероятности другого, то такие события называются независимыми. Сложное событие, состоящее в одновременном наступлении событий A_1 и A_2 , называется произведением событий A_1 и A_2 (обозначается A_1A_2). Вероятность произведения независимых событий равна произведению вероятностей каждого из них.

$$P(A_1A_2) = P(A_1)P(A_2). \quad (3.4)$$

Этим соотношением определяется вероятность одновременного пролета частиц в течение τ секунд через первый и второй счетчики.

Можно обобщить эти результаты на любое число элементарных событий. Если события A_1, A_2, \dots, A_k попарно несовместимы, т. е. никакие два из них не могут появиться одновременно, то

$$P(A_1 + A_2 + \dots + A_k) = P(A_1) + P(A_2) + \dots + P(A_k). \quad (3.5)$$

Если события A_1, A_2, \dots, A_k независимы, т. е. наступление или ненаступление любой системы этих событий не меняет вероятностей наступления остальных, то

$$P(A_1 A_2 \dots A_k) = P(A_1) P(A_2) \dots P(A_k). \quad (3.6)$$

В задачах теории вероятностей приходится иметь дело с системами событий, которые наступают или не наступают при реализации определенных условий. Обычно принимаются следующие допущения относительно системы событий:

а) если система содержит события A и B , то она содержит также и события $AB, A + B, A - B$;

б) система содержит достоверное и невозможное события.

Система событий, удовлетворяющая этим допущениям, называется полем событий.

Мы все время пользовались в качестве примера случайного события пролетом ионизирующей частицы через заданный объем. Те же рассуждения и заключения справедливы и для других случайных событий: для бросания игральной кости или монеты, вытаскивания карты из колоды или тиража лотереи, появления бракованного изделия в продукции завода или появления ценного вещества в образце породы и т. д.

Вернемся снова к подсчету ионизирующих частиц, чтобы ввести еще одно понятие—условную вероятность. Будем с помощью счетчиков исследовать образцы минералов, содержащих радиоактивное вещество. Пусть в нашем распоряжении имеются одинаковые образцы двух сортов с различным содержанием радиоактивного вещества, причем мы не знаем, какие именно из образцов принадлежат первому и какие—второму сорту. Если мы наугад берем один из образцов, то существует вероятность $P(B_1)$, что он окажется первого сорта (с низкой радиоактивностью), и вероятность $P(B_2)$ —второго (с высокой радиоактивностью). Вероятность зафиксировать пролет частицы $P(A)$ зависит от того, к какому сорту принадлежит взятый нами наугад образец, который подвергается исследованию. Иначе говоря, мы наблюдаем случайное событие A —пролет частицы—при условии, что наступило другое случайное собы-

ние B —случайно выбран образец первого (или второго) сорта. Вероятность такого события называется условной вероятностью и обозначается $P(A|B)$.

Вероятность сложного события, состоящего в совместном появлении событий A и B_1 —выбор образца первого сорта и пролет частицы за время τ ,— будет равна произведению вероятности первого события на условную вероятность второго при условии, что первое событие наступило

$$P(AB_1) = P(B_1)P(A|B_1). \quad (3.7)$$

Условная вероятность, так же как и безусловная, имеет величину, лежащую в пределах от нуля до единицы. Для нее остаются справедливыми правила сложения и умножения вероятностей, о которых мы говорили выше. Это естественно, так как, в сущности, любую безусловную вероятность можно рассматривать как условную, ибо она существует только при условии осуществления определенной обстановки, в которой наблюдается данное случайное событие. Говоря об условной вероятности, мы подчеркиваем тот факт, что к однажды оговоренным условиям испытания мы добавляем некоторые дополнительные условия (выбор определенного сорта образца в нашем примере), которые осуществляются не при всяком испытании, являются случайным событием и имеют определенную вероятность.

Вероятность совместного наступления двух событий—выбор образца первого сорта и появление следа на фотоснимке или срабатывание счетчика в заданном интервале времени— можно записать иначе, чем мы это сделали в (3.7), а именно:

$$P(AB_1) = P(A)P(B_1|A). \quad (3.8)$$

Здесь $P(B_1|A)$ есть условная вероятность того, что мы выбрали образец первого сорта при условии, что после опыта мы обнаружили на фотографии след частицы.

Взяв образец наугад, мы не знаем, к какому именно сорту он принадлежит, и можем по этому поводу делать лишь предположения или высказывать гипотезы. Проведя ряд экспериментов и получив их результаты (обнаружив следы частиц), мы можем на основании этих результатов сделать заключение о вероятности того, что наша гипотеза справедлива.

Вопрос о статистической проверке гипотез является обширным и важным разделом теории вероятностей.

В практике, разумеется, задачи несравненно сложнее, чем в нашем примере. Вопрос о проверке гипотез особенно важен в технике передачи сигнала в помехах. Сигнал, посланный передатчиком, оказывается на приемном конце искаженным помехой. Приняв искаженный сигнал, мы можем определить условные вероятности передачи того или иного сигнала и на основании этих условных вероятностей вынести суждение об истинном сообщении, которое было послано.

Если условная вероятность $P(B|A)$ равна безусловной вероятности $P(B)$, то события A и B являются независимыми. Если $P(B)$ не равна $P(B|A)$, то события A и B оказываются статистически связанными, между ними существует корреляция. Наступление одного из них меняет вероятность другого. В пределе наступление одного события делает наступление другого достоверным или невозможным. В этом случае мы имеем обычную функциональную связь двух событий.

Распределение вероятностей. Математическое ожидание и дисперсия

Если случайным событием является выбор определенного числа или величины из некоторого множества чисел или величин, то мы имеем дело со случайной величиной. Число пассажиров трамвая, число зафиксированных ионизирующих частиц, ошибка измерения, отклонение размера обработанной детали от заданного, вес початка кукурузы, число выигрышей, выпавших на ваши лотерейные билеты, число козырей, доставшееся игроку при сдаче карт, число родившихся сегодня в Москве мальчиков, возраст умершего—все это суть случайные величины. Они могут принимать как дискретные, так и непрерывные множества значений.

Задавшись одним определенным значением случайной величины или областью ее значений, мы можем подсчитать вероятность события, состоящего в том, что случайная величина в очередном испытании примет это значение или окажется в заданной области значений. При массовых испытаниях эта вероятность будет равна частоте появления данного значения случайной величины. Так можно найти, например, вероятность того, что все ваши лотерейные билеты выиграют или что не выиграет ни один. Можно определить вероятность того, что размер детали выйдет из допусков, и т. д.

Определив вероятность каждого значения случайной величины, мы можем построить диаграмму зависимости вероятности от значения случайной величины. Обычно строится кривая распределения вероятностей, на которой по горизонтальной оси откладывается значение величины ξ , а по вертикальной оси — вероятность того, что эта величина ξ окажется меньше некоторого заданного значения x , $P(\xi < x) = F(x)$. По мере роста x кривая распределения монотонно возрастает от нуля до единицы. Если случайная величина ξ может принимать только отдельные дискретные значения (число выигравших билетов, например, может быть только целым), то кривая распределения оказывается ступенчатой. Если же величина ξ изменяется непрерывно (отклонение размера детали), то кривая распределения оказывается плавной. С помощью кривой распределения можно ответить на вопрос, чему равна вероятность того, что случайная величина окажется лежащей в пределах от x_1 до x_2 , $P(x_1 < \xi \leq x_2)$. Эта вероятность равна разности ординат, соответствующих абсциссам x_1 и x_2 .

Иногда пользуются кривой плотности вероятности, на которой откладывается производная (скорость роста) распределения вероятности по x . Для дискретных величин эта кривая имеет вид отдельных точек, для которых абсциссы равны дискретным значениям ξ , а ординаты — их вероятностям $P(\xi)$. Для непрерывных распределений кривая плотности вероятности непрерывна. Для того чтобы по кривой плотности определить вероятность попадания ξ в интервал $x_1 \leq \xi \leq x_2$, необходимо определить площадь, заключенную между кривой и осью абсцисс на участке x_1, x_2 , т. е. интеграл плотности вероятности. Вся площадь под кривой плотности равна, очевидно, вероятности достоверного события, т. е. единице.

В качестве характеристик кривой распределения вероятности пользуются величинами математического ожидания и дисперсии. Математическое ожидание случайной величины [обозначается $M(\xi)$] представляет собой сумму произведений значений случайной величины на соответствующие вероятности

$$\begin{aligned} M(\xi) &= \xi_1 P(\xi_1) + \xi_2 P(\xi_2) + \xi_3 P(\xi_3) + \dots + \xi_n P(\xi_n) = \\ &= \sum_{i=1}^n \xi_i P(\xi_i). \end{aligned} \quad (3.9)$$

В случае непрерывной величины ξ сумма заменяется интегралом

$$M(\xi) = \int_{-\infty}^{+\infty} \xi p(\xi) d\xi, \quad (3.10)$$

где $p(\xi)$ — плотность вероятности.

Дисперсией [обозначается $D(\xi)$] называется математическое ожидание квадрата отклонения случайной величины ξ от $M(\xi)$

$$D(\xi) = M[\xi - M(\xi)]^2 = \sum_{i=1}^n [\xi_i - M(\xi_i)]^2 P(\xi_i),$$

где $P(\xi_i)$ — вероятность появления ξ_i . Для непрерывных величин

$$D(\xi) = \int_{-\infty}^{+\infty} [\xi - M(\xi)]^2 p(\xi) d\xi,$$

где $p(\xi)$ — плотность вероятности.

Чем больше дисперсия, тем чаще в среднем встречаются большие отклонения от среднего значения. Величина дисперсии ошибок измерения характеризует качество измерения, с точки зрения точности. Чем меньше дисперсия, тем точнее измерение.

Приведем некоторые типичные законы распределения вероятностей.

Пусть существует N лотерейных билетов, каждый из которых имеет вероятность выиграть независимо от остальных, равную p , одинаковую для всех билетов (и, следовательно, вероятность не выиграть, равную, как вероятность дополнительного события, $q = 1 - p$). Какова вероятность того, что в тираже из N билетов выиграют ровно n (и, следовательно, не выиграют $m = N - n$) билетов? Если мы заранее укажем определенные билеты, которые должны выиграть, то вероятность того, что именно указанные билеты выиграют, а остальные — нет, равна, как вероятность произведения независимых событий, $p^n q^m$. Однако мы могли бы выбрать не только те билеты, которые указали, но и любые другие из числа N , имеющихся в нашем распоряжении. Из общего числа N можно выбрать n билетов $C_N^n = \frac{N!}{n!(N-n)!}$ способами. Искомая вероятность того, что любые n билетов выиграют (для несовместимых событий — отдельных выборов) равна

$$P_N(n) = \frac{N!}{n!(N-n)!} p^n q^{N-n}. \quad (3.11)$$

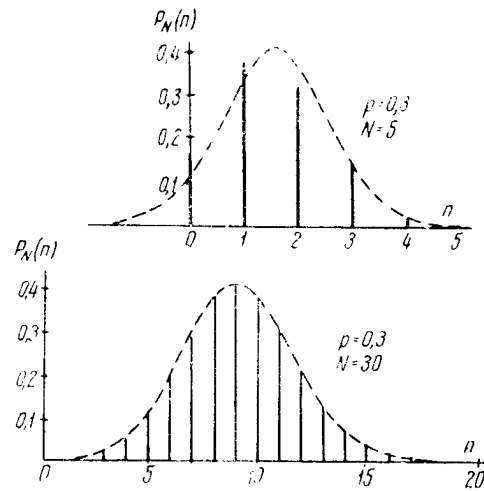


Рис. 3.3. Биномальное распределение. Вероятность получить данный результат в серии испытаний [см. формулу (3.11)].

Нетрудно заметить, что $P_N(n)$ в (3.11) равно n -му члену разложения бинома $(p + q)^N$. По этому признаку распределение вероятностей, соответствующее формуле (3.11), называется биномальным распределением. Функция распределения вероятностей $F_N(n)$ равна вероятности того, что случайная переменная ξ меньше величины n или равна ей.

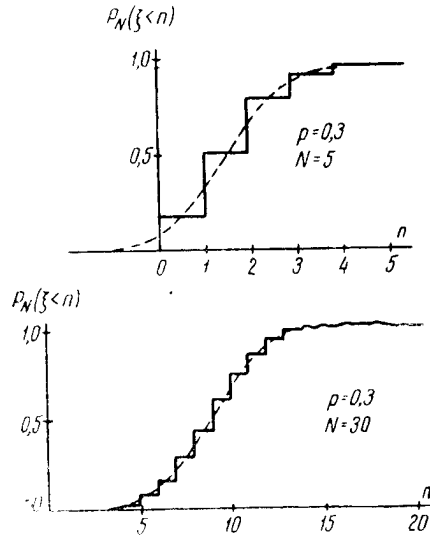


Рис. 3.4. Биномальное распределение. Вероятность получить результат не меньше данного в серии испытаний [см. формулу (3.12)].

$$P(\xi \leq n) = F_N(n) = \sum_{k=1}^n \frac{N!}{k!(N-k)!} p^k q^{N-k}. \quad (3.12)$$

Величины $P_N(n)$ и $F_N(n)$ заранее вычислены и для них составлены специальные таблицы. Для биномиального распределения математическое ожидание равно $M(n) = Np$, дисперсия $D(n) = Npq$.

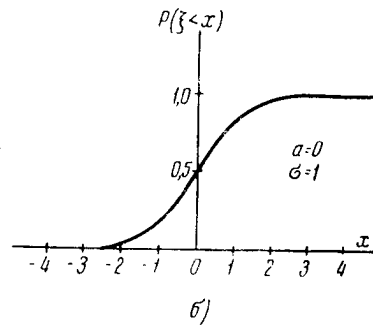
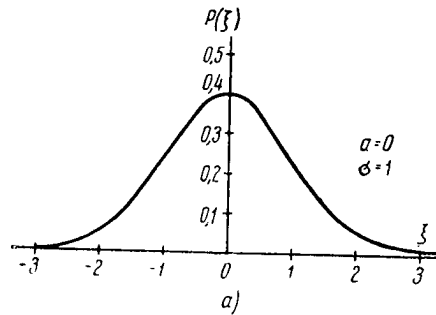


Рис. 3.5. Нормальное распределение.
а) плотность вероятности, б) распределение вероятности [см. формулы (3.13) и (3.14)].

Если мы будем в формуле (3.11) увеличивать N и n , оставив постоянными p и q , то в пределе придем к непрерывному распределению

$$P(\xi \leq x) = F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{(x-a)^2}{2\sigma^2}} dx \quad (3.13)$$

с плотностью вероятности

$$p(\xi) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(\xi-a)^2}{2\sigma^2}}, \quad (3.14)$$

называемому нормальным распределением или законом Муавра—Лапласа—Гаусса.

Нормальное распределение встречается в практических задачах весьма часто. Этому закону следуют случайные величины, являющиеся суммой очень большого числа взаимно-независимых случайных величин, таких, что каждая из них мала по сравнению со всей суммой. Математическое ожидание для нормального распределения равно $M(\xi) = a$, дисперсия $D(\xi) = \sigma^2$.

Если теперь мы будем увеличивать N в формуле (3.11), положив величину вероятности p очень малой, $p \ll 1$, то в пределе получим так называемый закон распределения Пуассона

$$P_N(\xi \leq n) = F_N(n) = \sum_{i=0}^n \frac{(Np)^i}{i!} e^{-Np}. \quad (3.15)$$

В этом случае вероятность для данной величины n равна

$$P_N(n) = \frac{(Np)^n}{n!} e^{-Np} = \frac{\lambda^n}{n!} e^{-\lambda}; \quad (3.16)$$

$$M(n) = Np = \lambda, \quad D(n) = Np = \lambda.$$

Закону Пуассона следует, например, распределение вероятностей числа вызовов телефонной станции, число излученных позитронизирующих частиц, число посетителей выставки за данный отрезок времени и другие дискретные случайные величины. Для нормального закона и закона Пуассона также имеются таблицы.

В практике встречаются задачи, когда экспериментально изучается случайная величина, вероятность которой распределена по известному закону. Из результатов измерения бывает необходимо определить только параметры закона: $M(\xi)$ и $D(\xi)$. Другие задачи состоят в определении типа закона распределения по результатам наблюдений или в определении влияния на этот закон условий опыта.

Шум

На выходе усилителя приемника сигналов при большом усилении всегда имеется напряжение шумов, которое представляет собой случайную величину и принимает в различные моменты времени случайные значения. Эти шумы происходят от флуктуаций тока во входных цепях усилителя, обусловленных тепловому хаотическому движению электронов. Если измерить значения шумового напряжения в различные моменты времени и подсчитать частоту появления тех или иных значений, то закон распределения окажется близким к нормальному. Тепловые шумы называются поэтому нормальными или гауссовыми.

Если на вход усилителя поступает сигнал и притом настолько слабый, что он оказывается мал по сравнению с шумами, то распределение вероятностей напряжения на

выходе изменится. Наблюдая напряжение на выходе однократно или в течение короткого промежутка времени, мы не сможем с достоверностью установить присутствие или отсутствие сигнала и, только продолжая наблюдения достаточно долго, по изменению вида распределения сможем надежно решить этот вопрос.

Подобная ситуация представляется довольно общей для современного естествознания. В любом эксперименте результаты наблюдений подвержены случайным отклонениям, и единичное наблюдение может дать результат, вооб-

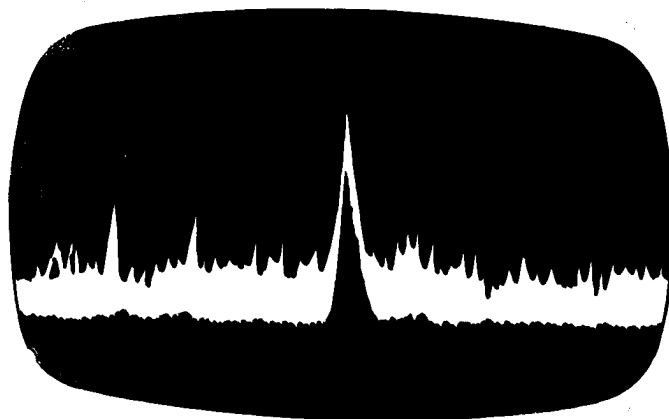


Рис. 3.6. Импульсный сигнал и шумы на выходе радиоприемника.

ще говоря, сколь угодно сильно отличающийся от среднего. Поэтому результаты эксперимента должны быть после обработки представлены в виде кривых распределения вероятностей этих величин. Избежать ошибок можно только посредством корректной статистической обработки результатов.

Когда на выходе приемника одновременно присутствуют и сигнал, и шум, мы имеем дело с вероятностью сложного события—напряжения, зависящего от двух случайных причин. Для подобных ситуаций следует пользоваться понятиями вероятности сложного события и условной вероятности.

Получив на приемном конце канала связи некоторый сигнал, искаженный шумами, мы оказываемся не в состоянии точно указать, какой именно сигнал был передан. Мы

можем только строить предположения или гипотезы на этот счет, каждая из которых после получения искаженного сигнала приобретает некоторую апостериорную или «послеопытную» вероятность. По-видимому, наблюдая однократно переданный сигнал, большего мы и не в состоянии определить.

Проблема приема сигнала в шумах или помехах является одной из центральных проблем не только теории и техники связи, но и техники автоматического регулирования и других областей применения кибернетики. Величина шумов или помех по сравнению с сигналом определяет скорость, с которой возможно передавать сигналы без ошибок. Статистические характеристики помех заставляют применять особые методы передачи и приема.

Минимальный по мощности сигнал, различимый при заданном числе наблюдений на фоне существующих шумов, носит название порогового сигнала. Чем ниже его мощность, тем выше чувствительность приемника. Чувствительность приемника является одним из самых важных его качеств. Интересно отметить, что, как правило, технические устройства для приема сигналов далеко уступают в чувствительности естественным приемникам, которые существуют у живых организмов в виде так называемых рецепторов или органов чувств. С этой точки зрения изучение процессов восприятия сигналов у живых организмов может принести огромную пользу технике.

Не только шумы, накладывающиеся на сигнал в канале связи, являются примером случайных воздействий на кибернетические системы. Иногда случайные воздействия поступают другими путями. При работе автопилота или автоматического рулевого система в целом испытывает случайные воздействия в виде порывов ветра, ударов волн, течений, которые хаотически изменяют курс корабля. Возникает задача противодействия этим случайным влияниям и их «отработки», т. е. внесения поправок в управление с таким расчетом, чтобы курс корабля не отклонялся случайным образом от желаемого. Иначе говоря, оказывается необходимым стабилизировать курс, сделать его независимым от случайных воздействий.

Подобные же задачи стабилизации решают многие искусственные и естественные кибернетические системы, например орган: тела животного, обеспечивающие постоянство температуры тела, давления и состава крови при случайных изменениях внешних и внутренних условий (температуры среды, питания, нагрузки).

Случайные воздействия, неизвестные и непредсказуемые изменения обстановки—вот нормальные условия существования как живых организмов, так и созданных человеком кибернетических систем.

Имея дело с шумами, помехами или приемом заранее неизвестного сигнала, приходится рассматривать последовательность случайных событий во времени. В этом случае мы имеем дело уже не со случайной величиной, а со случайной функцией времени. Если множество случайных функций может быть охарактеризовано с помощью распределения вероятностей, то процесс, включающий в себя это множество, есть случайный процесс.

Для характеристики случайного процесса обычно задаются вероятности сложных событий, состоящих в совместном появлении определенных величин в различные моменты времени:

1) вероятность того, что величина x в момент t_1 окажется в интервале $x_1; x_1 + dx_1$,

$$p_1(x_1, t_1) dx_1, \quad (3.17)$$

2) вероятность того, что величина x в момент t_1 окажется в интервале $x_1; x_1 + dx_1$ и затем в момент t_2 — в интервале $x_2; x_2 + dx_2$,

$$p_2(x_1, t_1; x_2, t_2) dx_1 dx_2, \quad (3.18)$$

и т. д.

Если для некоторого случайного процесса замена всех значений $t_1, t_2 \dots$ в (3.17), (3.18) и т. д. на $t_1 + \tau; t_2 + \tau \dots$ не изменяет величин вероятностей p_1, p_2, \dots , то такой случайный процесс называется стационарным. Смещение по времени τ может быть выбрано произвольным. Характеристики стационарного процесса не зависят от времени; их можно изучить, наблюдая процесс достаточно долго. В случае нестационарного процесса мы должны были бы для той же цели наблюдать большое число одновременно протекающих процессов или воспроизводить процесс с начала до конца много раз. Шумы приемника являются стационарным процессом.

Случайный процесс может давать как непрерывное изменение случайной величины, так и дискретные значения через определенные интервалы времени. В общем случае вероятность получить определенную величину в данный момент зависит от всей предыдущей истории случайной величины. В простейшем случае все значения величины не

зависят друг от друга. В том случае, когда значение случайной величины в данный момент времени t_0 полностью определяет распределение вероятностей величины x для следующего момента времени и если это распределение не зависит от значений x в предыдущие моменты времени $t < t_0$, случайный процесс является процессом без последовательности и называется процессом типа Маркова.

Вероятность в физике. Энтропия

Мы упоминали выше о термодинамике. Эта область физики дает нам настолько интересную иллюстрацию плодотворного применения математического понятия вероятности, что уместно сказать об этом подробнее.

Как известно, энергия существует в различных видах: механическая энергия, электрическая, магнитная, химическая, тепловая и т. д. Многочисленные физические процессы сопровождаются превращениями энергии из одного вида в другой, причем, согласно первому началу термодинамики, энергия не возникает и не исчезает. Однако, хотя количество энергии всегда сохраняется, в некоторых процессах происходит, как говорят, «обесценение» энергии. Это значит, что уменьшаются возможности дальнейшего преобразования энергии в другие виды. В этом отношении особую роль играет тепловая энергия.

Тепловая энергия представляет собой кинетическую энергию движения молекул вещества или же элементарных частиц. Собственно, о тепловой энергии можно говорить только в тех случаях, когда мы имеем дело с макроскопическим телом, состоящим из очень большого числа движущихся молекул, скорости которых распределены по некоторому закону. Для малого числа молекул, и тем более для одной молекулы, следует говорить не о тепловой, а о кинетической энергии.

Особая роль тепловой энергии в процессах преобразования энергии из одного вида в другой заключается в том, что тепловая энергия тела не может быть целиком превращена в механическую работу или в другой вид энергии без того, чтобы этому процессу не сопутствовали какие-то другие процессы, приводящие в конечном итоге к превращению в тепло некоторых количеств других видов энергии. Второе начало термодинамики, сформулированное на основании практики и результатов эксперимента, утверждает именно это: «Невозможен процесс, единственным результатом которого являлось бы превраще-

ше тепла в работу». Однако возможны процессы, единственным результатом которых является переход в тепло других видов энергии (трение, нагревание провода электрическим током и т. д.).

Преобразование тепла в работу возможно, например, при использовании изменений состояний газа. Нагревая газ, заключенный в цилиндре машины, мы заставляем его расширяться, перемещать поршень и поднимать груз. Тепло при этом отбирается у некоторого нагретого тела (котла). Чтобы продолжить этот процесс, необходимо вернуть поршень в исходное положение, т. е. охладить газ, а это значит, что часть тепловой энергии газа должна быть передана холодильнику. Таким образом, завершив цикл, мы отняли у горячего источника тепла некоторое количество тепловой энергии Q_1 , часть ее превратили в работу A , а другую часть Q_2 отдали холодильнику в виде тепловой энергии. При этом $Q_1 = A + Q_2$ (все величины измерены в калориях). Коэффициент полезного действия машины η оказывается меньше единицы, так как часть энергии не перешла в работу,

$$\eta = \frac{A}{Q_1} = \frac{Q_1 - Q_2}{Q_1}. \quad (3.19)$$

Последнее, однако, не означает, что в данном случае тепло Q_2 безвозвратно потеряно. Можно осуществить обратный процесс, сообщив тепловой машине механическую энергию, т. е. совершив за счет посторонних источников работу A и заставив машину работать как холодильник. При этом не только работа A полностью превратится в тепло и будет передана горячему котлу, но, кроме этого, количество тепла Q_2 будет отнято у холодильника, который при этом еще сильнее охладится, и передано также котлу, который еще более нагреется. Цикл работы холодильной машины полностью восстанавливает исходное состояние, с которого начала свой цикл тепловая машина. Таким образом, превращение тепла в работу тепловой машиной является процессом обратимым. Это значит, что с помощью процесса, обратного данному, можно вернуть систему в первоначальное состояние, не внося при этом никаких изменений в окружающие тела.

Однако обратимый процесс является лишь теоретической идеализацией реальных процессов. Для того чтобы приблизиться к осуществлению обратимого процесса, нужно исключить трение, теплопроводность и производить все изменения в газе чрезвычайно медленно, не нарушая его

равновесного состояния. В действительности при превращении тепла в работу, кроме тепла Q_2 , бесполезно рассеивается еще некоторое количество тепла за счет трения и теплопроводности. Это рассеяние не сопровождается превращением в работу соответствующей части тепла, и, следовательно, для того чтобы вернуть систему в исходное состояние, необходимо почерпнуть дополнительную работу откуда-то извне. Процесс с участием трения оказывается необратимым.

Принципиально важно то, что коэффициент полезного действия (к. п. д.) идеальной тепловой машины, осуществляющей обратимый процесс превращения тепла в работу, не зависит от свойств газа («рабочего тела») и определяется только температурами котла T_1 и холодильника T_2

$$\eta_i = \frac{T_1 - T_2}{T_1} \quad (3.20)$$

Если котел и холодильник имеют одинаковую температуру, то к. п. д. равен нулю, т. е. невозможно никакую часть тепла превратить в работу. Всякое выравнивание температур котла и холодильника ведет к понижению к. п. д. машины, т. е. к уменьшению той части тепла котла, которая может быть превращена в работу. Отсюда ясен смысл утверждения, что теплопроводность «обесценивает» тепловую энергию.

Приведенные выше утверждения справедливы не только для тепловой машины с рабочим телом в виде газа. То же самое справедливо, например, и для процессов непосредственного превращения тепла в энергию электрического тока. Можно практически осуществить приближение к идеальному обратимому процессу, используя термоэлектрический эффект (явление Зеебека) и явление Пельтье.

Явление Зеебека заключается в появлении электрического тока в замкнутой цепи, состоящей из двух последовательно включенных проводов различного материала, когда один спай нагревается, а другой охлаждается извне. При этом часть тепла, получаемая нагретым спаем, превращается в электрическую энергию и часть передается холодильнику.

Явление Пельтье заключается в самопроизвольном нагреве одного из спаев в цепи, подобной только что описанной, и в охлаждении другого спаю при пропускании через цепь электрического тока от постороннего источника. Энергия электрического тока превращается при этом в

тепло, выделяющееся в нагретом спаяе, и, кроме того, некоторое количество тепла переносится от холодного спая к горячему. Последний факт часто вызывает утверждение, будто бы термоэлектрический холодильник «имеет коэффициент полезного действия больше единицы». Такое же утверждение можно было бы сделать, как мы видели, и относительно газовой холодильной машины. Однако, если мы попытаемся осуществить замкнутый обратимый процесс: превращения электрической энергии в тепло (явление Пельтье), а затем обратное превращение полученного тепла в электрическую энергию (явление Зеебека), то даже в идеальном случае «выигрыш» оказался бы равным нулю, а к. п. д. — равным единице.

В реальном же эксперименте нагревание проводов за счет джоулева тепла вызовет понижение к. п. д. В этом случае, как и в любом другом, реализовать вполне обратимый процесс не удастся.

Внутреннюю энергию тела можно разделить на две части: ту, которая может быть превращена в работу (свободная энергия), и ту, которая при этом неизбежно будет передана более холодному телу в виде тепла (связанная энергия). Физическая энтропия тела характеризует долю этой «бесполезной» энергии, которая не может быть превращена в работу. Энтропия является реальной величиной, которая может быть определена расчетом или измерением. При физических изменениях тела энтропия его меняется. Энтропия аддитивна: энтропия суммы нескольких тел равна сумме их энтропий. Размерность энтропии — *калории/градус*.

Приращение энтропии тела при изменениях его физического состояния определяется суммой «приведенных теплот» dQ/T :

$$S_1 - S_2 = \int_1^2 \frac{dQ}{T}. \quad (3.21)$$

Самым существенным свойством энтропии является то, что в замкнутой системе, для которой невозможен обмен энергией с внешним миром, энтропия никогда не уменьшается. Она остается постоянной при обратимых процессах и повышается при необратимых.

Связанная энергия тела пропорциональна его энтропии. Таким образом, трение, теплопроводность, увеличивая долю связанной энергии, повышают энтропию системы.

Изменение энтропии является мерой «необратимости»

данного процесса или мерой «неполноценности» внутренней энергии данного тела. Второе начало термодинамики в формулировке Клаузиуса звучит так: «С течением времени энтропия всякой изолированной системы стремится к максимуму». Достигнув максимума энтропии, изолированная система пребывает в состоянии теплового равновесия.

Как мы уже упоминали, некоторыми особенностями в этом отношении обладают тела живой природы. Они самопроизвольно поддерживают свою энтропию неизменной за счет увеличения энтропии перевариваемых питательных веществ или за счет поглощения энергии извне (главным образом, солнечной радиации). Однако живые тела сами по себе не представляют собой изолированной системы, так как не могут существовать без взаимодействия со средой. Поэтому утверждение, будто бы «живые организмы не подчиняются второму началу термодинамики», является неверным. Поддержание энтропии некоторого тела постоянной за счет соответствующего увеличения энтропии окружающих тел вовсе не представляет собой чего-то исключительного и может быть воспроизведено. Особенностью живых тел является собственно лишь тот способ, которым они достигают этого непрерывно и «автоматически».

Свойство изолированных систем изменять свою энтропию только в одном направлении—в сторону увеличения—было блестяще объяснено Больцманом на основе теории вероятностей. Сущность этого объяснения вкратце сводится к следующему: «Физические тела переходят от состояний менее вероятных к более вероятным».

Энтропия, так же как и температура, свойственна только телам макроскопическим, т. е. состоящим из очень большого числа молекул. Для одной молекулы энтропии не существует. Расположение и движение молекул газа в объеме могут быть самыми разнообразными. Большое число соударений с другими молекулами делает координаты и скорости молекул случайными величинами, которые имеют некоторые распределения вероятностей. Число молекул чрезвычайно велико: в кубике с ребром 4 микрона при атмосферном давлении содержится больше молекул, чем имеется людей на земном шаре. Понятие «массовое явление» как нельзя более подходит к газовому состоянию.

Мы можем представить себе различные распределения молекул газа в объеме. Каждое такое распределение является сложным событием и будет иметь определенную

вероятность, которую можно подсчитать по известным правилам теории вероятностей.

Пусть в сосуде с объемом V находится N молекул газа. Разобьем объем сосуда на большое число ν одинаковых элементарных объемов или ячеек. Будем считать, что все молекулы размещаются независимо друг от друга и что вероятность того, что молекула находится в определенной ячейке, равна отношению объема ячейки к полному объему:

$$p = \frac{\Delta V}{V} = \frac{\Delta V}{\nu \Delta V} = \frac{1}{\nu}. \quad (3.22)$$

Теперь подсчитаем вероятность сложного события, заключающегося в том, что в первой ячейке окажется n_1 молекул, во второй n_2 , ..., в последней n_ν . При этом, разумеется,

$$\sum_{i=1}^{\nu} n_i = N,$$

где N — общее число молекул в сосуде. Эта задача аналогична подсчету вероятности выигрыша n билетов из общего числа N . Обобщая формулу (3.11), напомним:

$$P_N(n_1, n_2, \dots, n_\nu) = \frac{N!}{n_1! n_2! \dots n_\nu!} \left(\frac{1}{\nu}\right)^N. \quad (3.23)$$

Меняя числа n_i , мы будем, очевидно, менять и величину вероятности P_N . По-видимому, одно из возможных распределений будет иметь наибольшую вероятность. Определив максимум величины P_N , убедимся, что наибольшей вероятностью обладает равномерное распределение молекул в объеме сосуда. При этом

$$n_i = \frac{N}{\nu} = N p_i. \quad (3.24)$$

Именно к такому распределению плотности приходит газ при отсутствии внешних возмущений в любом реальном эксперименте. В частности, явление диффузии приводит именно к такому наиболее вероятному распределению плотностей в смеси газов.

Аналогичным образом можно определить наиболее вероятное состояние газа с учетом влияния поля тяготения и с учетом распределения скоростей. В этом случае получается закон распределения Максвелла — Больцмана, который наблюдается и на опыте. При этом физические вели-

чины, свойственные всей массе газа, оказываются параметрами распределения вероятностей. Так, например, температура газа пропорциональна математическому ожиданию кинетической энергии молекул.

Состояние газа, отличное от равновесного, обладает меньшей вероятностью, причем вероятность падает тем быстрее, чем больше число молекул. М. Смолуховский подсчитал, что вероятность самопроизвольного отклонения мгновенного значения давления газа от среднего на 1% такова, что в объеме газа, соответствующем кубу с ребром 0,2 микрона, такое отклонение происходит в среднем каждые 10^{-9} секунды; в объеме, равном одному кубическому сантиметру, то же самое случается лишь каждые 10^{14} секунд, т. е. практически никогда не случается.

Можно доказать методами теории вероятностей, что система молекул газа, находящаяся в состоянии, отличном от наиболее вероятного, устойчивого, вскоре самостоятельно придет к устойчивому состоянию и останется в нем. Если мы, например, приведем в соприкосновение через теплопроводную стенку две порции газа с различными средними энергиями молекул (т. е. с различными температурами), то обе порции газа вскоре окажутся при одинаковой температуре. Если мы уберем перегородку, то газ будет самопроизвольно перемещаться и распределится по объему равномерно. Таким образом, теплопроводность и диффузия являются «макроскопическим» выражением статистической закономерности перехода от маловероятного состояния к более вероятному.

Можно вычислить энтропию S равновесного состояния газа, выразив ее через функцию распределения молекул по скоростям, или вероятность состояния. Это выражение имеет вид

$$S = k_1 \ln P + C, * \quad (3.25)$$

где P — вероятность состояния.

C — некоторая постоянная.

Энтропия равновесного состояния газа пропорциональна логарифму вероятности состояния газа. Самопроизвольное увеличение энтропии получает, таким образом, наглядное толкование как увеличение вероятности состояния.

* Энтропия, выраженная через числа молекул в ячейках объема и вероятности пребывания в них, принимает вид:

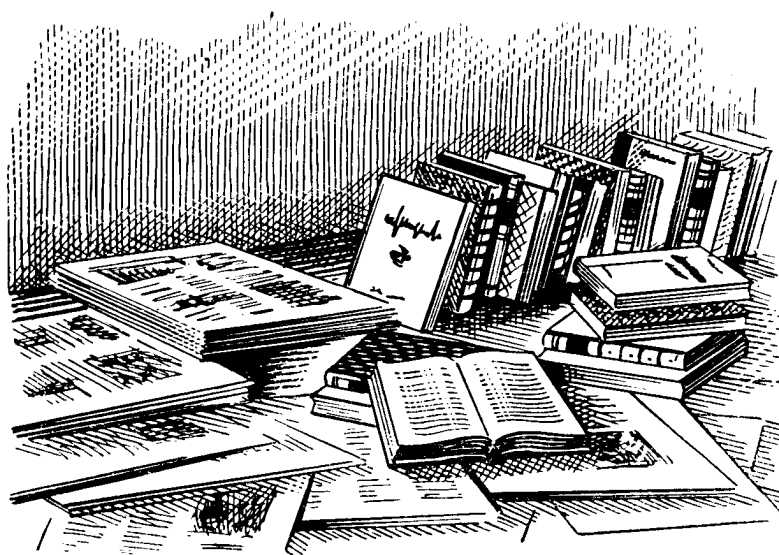
$$S = k \sum_{i=1}^v n_i \log n_i = -k \sum_{i=1}^v p_i \log p_i + C. \quad (3.26)$$

Sanitized Copy Approved for Release 2010/06/16 : CIA-RDP80T00246A045100330001-4

Второе начало термодинамики является законом статистического типа, справедливым только для массовых явлений. Для малых количеств молекул энтропия может иногда испытывать случайные самопроизвольные уменьшения и второе начало будет справедливым только в среднем по времени. Законы газового состояния являются законами статистического типа.

Другим примером плодотворного применения статистических методов, на этот раз к вопросам передачи сигнала, является теория информации, некоторые положения которой мы рассмотрим в следующей главе.

Sanitized Copy Approved for Release 2010/06/16 : CIA-RDP80T00246A045100330001-4



ГЛАВА 4

КОЛИЧЕСТВО ИНФОРМАЦИИ

При проектировании и создании систем, работающих с сигналами, в частности каналов связи, недостаточно подходить к решению задачи только с точки зрения физики или энергетики. Правда, физические процессы, происходящие в канале, весьма существенны и их характеристиками нельзя пренебрегать, так как сигнал есть физический процесс. Однако ограничиваться только рассмотрением физических процессов как таковых — значит упускать из виду самое главное, ради чего создается канал связи. Этим главным является количество информации, передаваемой без искажений в единицу времени. Таким образом, мы приходим к необходимости иметь численную меру количества информации.

Информационная емкость

С установления оценки количества информации посредством числа начинается новая отрасль теории связи — теория информации, которая рассматривает вопро-

сы о пропускной способности каналов связи, о наилучшем кодировании информации и выводит предельные соотношения для количества информации, которое может быть передано по каналу в единицу времени. Эти соотношения в теории связи по их значению можно сравнить с законом сохранения энергии в энергетике. Подобно тому как в энергетике, количество энергии, отдаваемое энергетической машиной, не может быть больше количества потребляемой ею энергии, в теории информации количество информации, передаваемое без ошибок в единицу времени, не может быть больше пропускной способности канала.

Как же оценить количество информации? Как это ни странно с первого взгляда, но для численной оценки количества информации, т. е. количества содержания, смысла сигнала, необходимо прежде всего отвлечься от этого содержания, от смысла, от семантики. Впрочем странность этого только кажущаяся. При любых количественных расчетах все единицы теряют свою индивидуальность и становятся качественно одинаковыми. Чтобы воспользоваться, например, арифметикой или алгеброй для какого-либо конкретного расчета, мы должны прежде всего отбросить эту конкретность и перейти к числу вообще. То же справедливо относительно формальной логики и других теоретических дисциплин.

Если мы перестанем интересоваться семантикой сообщений, т. е. отвлечемся от конкретного содержания того события, о котором передаются или записываются сообщения, то у нас останется только множество отличных друг от друга состояний. Какое бы реальное событие мы ни описывали: строительство дома или производство автомобилей, ход спортивного соревнования или военной операции, состояние физической системы или ход химической реакции, судьбу отдельного человека или историю народа—любое событие в каждый момент времени может находиться в одном из большого числа возможных состояний. Описание события есть не что иное, как указание на то, в каком из возможных состояний находится описываемый объект.

Протекание события во времени есть смена этих состояний, выбранных из числа возможных. Различные конкретные события отличаются друг от друга числом возможных состояний и особенностями выбора. Множество возможных состояний может быть дискретным или непрерывным, представлять собой одну величину или совокупность многих величин.

Таким образом, информация в отвлеченном виде, без конкретного смысла сообщения, есть просто выбор некоторого определенного элемента из множества возможных. Каждому состоянию события может быть приведено в соответствие определенное состояние физического процесса, представляющего собой сигнал. Посылка сообщения о событии есть, таким образом, выбор определенного символа сигнала из множества возможных символов. Элементы множества сигнальных символов, подобно состояниям описываемого события, могут быть либо дискретными символами (буквы печатного текста, точки и тире телеграфной азбуки, цифры), либо непрерывными (высота звукового тона, амплитуда колебания).

Будем называть множество возможных состояний или символов сигнала алфавитом сигнала, а состояние физического объекта или процесса, реализующее этот алфавит, физическим алфавитом. Отдельный символ можно называть буквой алфавита, а серию или комбинацию символов—словом в алфавите. Простейшим примером дискретного алфавита является алфавит, состоящий из двух символов: 0 и 1, или + и —, или *да* и *нет* и так далее.

Чтобы подойти к определению числовой меры количества информации, необходимо прежде всего составить представление об объеме множества возможных сообщений, из которого производится выбор, а затем уточнить характер самого выбора.

Для дискретных алфавитов подсчет числа возможных сообщений, состоящих из заданного числа символов, не составляет труда. Если в алфавите N букв, то число возможных слов, состоящих из n букв каждое, равно N^n (одну букву можно выбрать N способами, две буквы — N^2 способами, ... n букв — N^n способами).

Для непрерывных алфавитов число возможных сообщений бесконечно велико. Однако непрерывная величина, являющаяся сигналом, всегда известна с некоторой ошибкой, обязанной шуму. Поэтому с точностью до допустимой ошибки эта величина может быть представлена в виде дискретных значений (или, как говорят, «квантована»). Наличие шума не позволяет реализовать и использовать бесконечно большое множество сообщений в непрерывном алфавите. Мы уже видели, каким образом можно непрерывную функцию времени с ограниченным спектром и с ограниченной амплитудой в ограниченном интервале времени представить в виде конечной последовательности конеч-

ных дискретных величин. Квантованием этих величин можно представить любую такую функцию в виде конечного множества чисел с конечным числом разрядов. Таким образом, можно свести непрерывный алфавит к дискретному и пересчитать все возможные непрерывные сообщения и в этом случае.

Пересчитав подобным образом все возможные трехминутные телефонные разговоры при полосе частот 4000 герц и разбиении амплитуды квантованием на 32 значения, мы получим число $10^{2000000}$ (единицу с двумя миллионами нулей). Разумеется, среди пересчитанных таким образом разговоров подавляющая часть окажется волюющей бессмыслицей или просто набором звуков.

Для телеграммы из 10 слов по 5 букв каждое количество всех возможных (в том числе и бессмысленных) телеграмм равно $N^n = 32^{50} \approx 10^{75}$. Для передачи одного сообщения, состоящего в выборе одного из двух символов, число возможных сообщений равно 2.

Итак, зная алфавит и длительность передачи, можно подсчитать число возможных сообщений, из которых производится выбор при каждой передаче. Мы подошли к определению информационной емкости системы, передающей или запаасающей информацию. В 1928 г. Хартли предложил оценивать информационную емкость системы логарифмом числа возможных состояний системы. Так, если система может находиться в N возможных состояниях и все ее состояния взаимно независимы, то ее емкость равна

$$C_1 = \log N. \quad (4.1)$$

Две такие системы вместе будут, очевидно, иметь N^2 состояний (каждому состоянию первой соответствует N возможных состояний второй). Следовательно, емкость двух систем составит

$$C_2 = \log N^2 = 2 \log N = 2C_1, \quad (4.2)$$

т. е. она будет вдвое больше, чем емкость одной системы. Очевидно, что емкость трех систем в таком определении будет втрое больше, а емкость k систем — в k раз больше; другими словами, емкость аддитивна.

С помощью меры, предложенной Хартли, работа системы, состоящая в прохождении тактов, может быть описана таким образом, что емкость будет возрастать линейно

с увеличением числа тактов, т. е. пропорционально времени работы. Эти линейные соотношения делают удобными расчеты и оправдывают применение логарифмической меры емкости. Поскольку в качестве систем, запаасающих информацию, применяются реле или другие устройства, имеющие только два устойчивых состояния, основание логарифма удобно взять равным двум. Тогда присоединение к схеме дополнительного реле увеличит емкость схемы на единицу.

Чтобы подойти к определению понятия количества информации, необходимо представить себе, как заполняется информационная емкость канала и какая часть ее действительно используется при передаче. Мы уже видели, что часть пересчитанных нами разговоров и телеграмм бессмысленны, т. е. никогда не встретятся в реальной передаче. Стало быть, при использовании канала некоторые состояния или сообщения никогда не встречаются, другие встречаются редко, третьи—часто. По-видимому, для передачи только «разумных», т. е. реально встречающихся сообщений, в принципе можно было бы обойтись каналом с меньшей емкостью.

Представим себя в роли адресата, находящегося на приемном конце канала связи. Мы получаем различные сообщения, которые исполняем или запоминаем. Интуитивно ясно, что не все серии сообщений встречаются с нашей стороны одинаковое субъективное отношение. Случается, что сообщения оказываются «интересными», «новыми», «неожиданными». Случается и так, что мы не получаем «ничего нового», «все без перемен» «все и так было ясно заранее». и т. д. Если все получаемые сообщения одинаковы или заранее известны или если из предыдущих можно однозначно вывести последующие, то, естественно, мы не узнаем ничего нового, т. е. будем получать количество информации, равное нулю.

Помните, как один из героев Жюль Верна—журналист Геден Спиллет—передавал по телефону главу из библии только для того, чтобы не дать возможности конкурирующим репортерам воспользоваться линией. Вся острота ситуации заключается в том, что Спиллет использует линию нелепым образом, т. е. передает общезвестный текст.

Сообщения «интересны», мы узнаем из них новости и получаем информацию только тогда, когда они «неожиданны» или, лучше сказать, случайны. Этот термин объективен и точен и, кроме того, позволяет перейти от ин-

туитивного и психологического определения ситуации «новизны» или «неожиданности» к объективному статистическому описанию явлений, из которого полностью исключается какая бы то ни была произвольность оценок.

Разумеется, случайность, которая входит, таким образом, в выбор сообщения из множества возможных, имеет источником случайный характер описываемого события. В поисках меры количества информации мы должны обратиться к методам теории вероятностей.

Мера количества информации

Всякое сообщение, передаваемое по каналу или фиксируемое в запоминающем устройстве, соответствует некоторому определенному состоянию внешнего явления, о котором передаются сообщения. Внешнее явление или событие представляет собой источник сообщения или источник информации. Те или иные состояния внешнего явления наступают часто или редко в зависимости от характера тех законов природы, по которым это явление протекает. Поскольку течение описываемого явления в будущем неизвестно (иначе о нем не было бы смысла передавать сообщения), можно говорить только о частотах или о вероятностях наступления тех или иных состояний. В связи с этим различные сообщения, каждое из которых соответствует определенному состоянию описываемого явления, также приобретают определенные вероятности.

Расположим все возможные сообщения A_i в некотором порядке и припишем каждому из них вероятность P_i , равную вероятности появления этого сообщения в реальной передаче. Мы будем иметь конечную схему, состоящую из полной системы попарно несовместимых событий (сообщений) A_i , заданных вместе с их вероятностями.

Конечная схема может быть записана так:

$$\begin{pmatrix} A_1 & A_2 & A_3 & \dots & A_i & \dots & A_n \\ P_1 & P_2 & P_3 & \dots & P_i & \dots & P_n \end{pmatrix}. \quad (4.3)$$

При каждой передаче производится выбор одного из элементов A_i схемы в соответствии с распределением вероятностей. Для каждой передачи какое-то сообщение будет

обязательно выбрано, поэтому сумма всех вероятностей должна быть равна единице

$$\sum_{i=1}^n P_i = 1. \quad (4.4)$$

Некоторые из P_i могут быть равны нулю.

Всякая конечная схема представляет неко орую неопределенность выбора ее элементов, которая снимается после того, как выбор произведен. Степень этой неопределенности различна в различных схемах. Неопределенность, по-видимому, возрастает, если возрастает число элементов схемы при равномерном распределении вероятностей. При одинаковом числе элементов схемы неопределенность возрастает по мере приближения распределения вероятностей к равномерному. Так, по-видимому, схема $\begin{pmatrix} A_1 & A_2 \\ 1/2 & 1/2 \end{pmatrix}$ имеет большую неопределенность, чем схема $\begin{pmatrix} A_1 & A_2 \\ 0,01 & 0,99 \end{pmatrix}$. Неопределенность исчезает вовсе, если одна из вероятностей становится равной единице.

Мера неопределенности выбора в данной конечной схеме и является мерой количества информации на один выбор, так как выбор снимает неопределенность. Мера неопределенности выбора или количества информации на один выбор должна быть функцией всех вероятностей P_i , изменяться непрерывно при непрерывном изменении P_i и, кроме того, принимать максимальное значение в том случае, когда все вероятности равны друг другу.

К. Шэннон, которому принадлежит заслуга постановки этого вопроса, показал, что этим и некоторым дополнительным требованиям удовлетворяет функция

$$H = -k \sum_{i=1}^n P_i \log P_i, \quad (4.5)$$

где k — постоянная, определяющая единицу измерения.

Вид этой функции совпадает с выражением для энтропии в статистической физике, причем это совпадение существенно, а не только формально. Поэтому величина H носит название энтропии совокупности вероятностей P_i .

В тех случаях, когда будет возникать опасность путаницы, мы будем называть соответствующие величины физической и информационной энтропиями.

В качестве единицы измерения количества информации или энтропии принимается единичный выбор из двух равновероятных возможностей. Для этого случая $H=1$; логарифм, как и в случае меры Хартли, берется при основании 2. Этим определяется постоянная $k=1$.

$$H\left(\frac{1}{2}; \frac{1}{2}\right) = -\left(\frac{1}{2} \log \frac{1}{2} + \frac{1}{2} \log \frac{1}{2}\right) = 1. \quad (4.6)$$

Аналогично мере Хартли энтропия возрастает пропорционально числу последовательных выборов. Для однократного выбора из двух возможностей с неравными вероятностями p и q ($p+q=1$) энтропия принимает различные значения в пределах между нулем и единицей (рис. 4.1). Когда p стремится к нулю или единице, т. е. когда выбор оказывается заранее predetermined, энтропия равна нулю в полном соответствии с нашими интуитивными представлениями. Максимум энтропии соответствует наибольшей неопределенности выбора, т. е. равным вероятностям. В этом случае энтропия равна информационной емкости множества возможных сообщений. Для любого множества выравнивание вероятностей ведет к увеличению энтропии. Энтропия одиночного выбора из множества N сообщений или символов, приведенная в формуле (4.5), представляет собой энтропию на символ передаваемого сообщения. Иногда эту величину называют также «содержательностью одиночного сообщения».

Интересно отметить, что Хартли, предложив меру емкости системы для запасаения информации при равновероятном использовании элементов, остановился перед дальнейшими трудностями. Он считал, что характер реального вы-

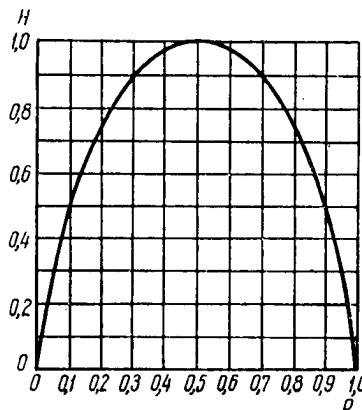


Рис. 4.1. Зависимость энтропии от вероятности в случае выбора из двух возможностей с вероятностями p и $1-p$.

бора сообщения из множества представляет собой проблему психологическую, а не математическую и не инженерную. Заслугой Шэннона является то, что он вместо психологических использовал статистические данные. В результате статистической постановки задачи мера Хартли

$$C = \log N = -\log \frac{1}{N} = -\log P \quad (4.7)$$

заменена энтропией

$$H = \sum_{i=1}^n P_i (-\log P_i),$$

которая является не чем иным, как математическим ожиданием (или средним по вероятности) отрицательного логарифма вероятности, т. е. меры Хартли.

Следует отметить, что величина энтропии относится не к единичному сообщению, которое фактически выбрано, а к единичному выбору из данного множества с заданным распределением вероятностей. Следовательно, энтропия характеризует не то, что в данный момент передано, а то, что могло бы быть передано. Иногда называют энтропию «мерой свободы выбора» или «мерой априорного незнания».

Выбор сообщения при передаче производится на входе канала связи источником сообщений, т. е. случайным процессом, обладающим соответствующими статистическими характеристиками. Следует, таким образом, говорить об «энтропии источника сообщений». Реальным источником сообщений для канала связи может быть и репортер, передающий сообщения в газету, и клиенты, подающие телеграммы, и измерительные приборы автоматической метеостанции: барометр, гигрометр, термометр и другие, выдающие результаты измерения для передачи по радио, и т. д. Важно, что в любом случае существует множество возможных сообщений и определенные частоты или вероятности отдельных сообщений. Делом нашего производства будет выбор того, что считать «сообщением»: целую телеграмму, слово, букву или, наконец, точку, тире и паузу телеграфной азбуки Морзе.

Энтропию источника сообщений можно измерять или на символ, т. е. на один выбор, или в единицу времени. Энтропия на символ, умноженная на число символов, выдаваемых в единицу времени, равна энтропии в единицу времени.

Емкость русского алфавита, включая знаки препинания и пропуски, при 32 возможных символах равна

$$-\sum_{i=1}^{32} \frac{1}{32} \log \frac{1}{32} = \log 32 = 5 \text{ единиц/символ.} \quad (4.8)$$

Если источник выдает текст на русском языке, то его энтропия на букву в два—три раза меньше емкости алфавита, т. е. составляет около 2 единиц на букву, поскольку в реальном «осмысленном» тексте символы не равновероятны (например, мягкий знак встречается реже, чем буква «а»).

Энтропия телеграммы из 10 слов (50 знаков) равна приблизительно 100 единицам. Страница книги содержит около 5000 единиц, книга в 200 страниц — около 1 000 000 единиц, библиотека в 100 000 томов — 10^{11} единиц.

Реальные события, о которых передаются сообщения, не всегда могут описываться простой конечной схемой. Различные стороны события могут потребовать введения нескольких конечных схем («погода» описывается температурой, давлением, влажностью, облачностью и т. д.). Таким образом, встает задача об определении количества информации сложного события.

Пусть мы имеем два взаимно-независимых события A и B и для каждого из них задана конечная схема, т. е. множества возможных исходов или состояний A_i и B_k и их вероятности. Совместное появление состояний A_i и B_k представляет собой сложное событие и будет иметь определенную вероятность

$$P(A_i B_k) = P(A_i) P(B_k).$$

Совокупность таких сложных событий образует новую конечную схему, более обширную, чем первоначальные. Выбору из этой новой схемы соответствует новое количество информации сложного события $H(A, B)$.

Для независимых событий энтропия схемы сложного события равна сумме энтропий простых событий

$$H(A, B) = H(A) + H(B). \quad (4.9)$$

Для любого числа независимых событий A, B, C, \dots

$$H(A, B, C, \dots) = H(A) + H(B) + H(C) + \dots \quad (4.10)$$

Для независимых событий энтропия аддитивна.

Код

На пути от источника информации к адресату сигнал может проходить несколько различных участков канала, каждый в отличном от других физическом алфавите. Так, при передаче по радио сигнал представляет собой то звуковые колебания, то колебания электрического тока, то колебания электромагнитного поля и т. д. На границах участков сигнал испытывает преобразование из одного физического алфавита в другой, причем содержание (информация), т. е. соответствие первоначальной форме сообщения или описываемому событию, сохраняется.

Преобразуя сигнал из одного алфавита в другой, мы сопоставляем символы первого алфавита с символами второго. Такое сопоставление алфавитов называется кодированием сигнала, а правило, по которому оно производится, — кодом*. В частности, всякое измерение представляет собой кодирование измеряемой величины в «выходном алфавите» измерительного прибора. Выходным алфавитом чаще всего служит механическое смещение стрелки — индикатора, поскольку оно лучше всего воспринимается человеком.

При кодировании дело, однако, не сводится только к тому, чтобы заменить одну физическую величину другой. При кодировании может измениться множество символов и распределение вероятностей. При кодировании мы можем сопоставить один символ первого алфавита с комбинацией или группой символов второго, и наоборот, или же сопоставить группу с группой. Так, при кодировании буквенного текста азбукой Морзе буква русского алфавита сопоставляется с двумя—семью (считая пропуски) символами телеграфного алфавита. Таким образом могут быть сопоставлены различные по информационной емкости алфавиты. Поскольку при кодировании меняется множество символов и распределение вероятностей, постольку меняется, вообще говоря, и энтропия на символ сообщения. Отсюда следует, что, выбирая один из возможных кодов, мы можем найти такой код, который имеет наибольшую энтропию на символ, или, иначе, наименьшее число символов на данное количество информации. Этот код будет «наибо-

* Часто называют «кодовой группой» или просто «кодом» группу символов — слово в некотором алфавите. Говорят, например, что «устройство памяти хранит 1024 кода». Мы будем применять слово «код» только в смысле правила кодирования, но не в смысле слова в алфавите.

«... более экономным» для передачи сообщения, так как потребует наименьшего числа символов и наименьшего времени передачи.

Рассмотрим пример, приведенный Шэнноном. Пусть источник выдает четыре взаимно-независимых символа A, B, C, D с вероятностями, соответственно, $1/2, 1/4, 1/8, 1/8$. Иначе говоря, источник изображается конечной схемой

$$\begin{pmatrix} A & B & C & D \\ 1/2 & 1/4 & 1/8 & 1/8 \end{pmatrix}. \quad (4.11)$$

Энтропия источника, очевидно, равна

$$H_0 = - \left[\frac{1}{2} \log \frac{1}{2} + \frac{1}{4} \log \frac{1}{4} + \frac{2}{8} \log \frac{1}{8} \right] = \frac{7}{8}. \quad (4.12)$$

Рассмотрим два кода, которые переводят эти символы в двоичный алфавит — (0,1).

Код № 1:

$$A \rightarrow 00, \quad B \rightarrow 01, \quad C \rightarrow 10, \quad D \rightarrow 11.$$

Код № 2:

$$A \rightarrow 0, \quad B \rightarrow 10, \quad C \rightarrow 110, \quad D \rightarrow 111.$$

Подсчитаем вероятности появления символов нового алфавита для обоих кодов. Для этого выпишем все символы исходного алфавита в соответствии с их вероятностями или частотой их появления и соответствующие комбинации символов нового алфавита для обоих кодов, а затем подсчитаем частоты появления последних.

Исходный алфавит:

$$A \quad A \quad A \quad A \quad B \quad B \quad C \quad D$$

Код № 1:

$$00 \quad 00 \quad 00 \quad 00 \quad 01 \quad 01 \quad 10 \quad 11$$

Код № 2:

$$0 \quad 0 \quad 0 \quad 0 \quad 10 \quad 10 \quad 110 \quad 111$$

Подсчитаем частоты появления (вероятности) символов.

Исходный алфавит:

$$P(A) = 4/8 = 1/2, \quad P(B) = 2/8 = 1/4, \quad P(C) = 1/8, \quad P(D) = 1/8.$$

Код № 1:

$$P_1(0) = 11/16, \quad P_1(1) = 5/16.$$

Код № 2:

$$P_2(0) = 7/14 = 1/2, \quad P_2(1) = 7/14 = 1/2.$$

В результате кодирования мы получили две новые конечные схемы.

Код № 1:

$$\begin{pmatrix} 0 & 1 \\ 11/16 & 5/16 \end{pmatrix}.$$

Код № 2:

$$\begin{pmatrix} 0 & 1 \\ 1/2 & 1/2 \end{pmatrix}. \quad (4.13)$$

Подсчитаем теперь энтропии на символ.

Для исходного алфавита [по формуле (4.12)]:

$$H_0 = 7/4 \text{ единиц/символ.}$$

Для кода № 1:

$$H_1 = -(11/16 \log 11/16 + 5/16 \log 5/16) = 0,875.$$

Для кода № 2:

$$H_2 = -(1/2 \log 1/2 + 1/2 \log 1/2) = 1.$$

Очевидно, что код № 2 является более экономичным, чем код № 1, так как каждый символ кода № 2 несет большее количество информации.

Закодировав сообщения кодом № 2 в алфавит (0,1), можно снова вернуться к исходному алфавиту (A, B, C, D), воспользовавшись кодом № 1 в обратном порядке: 00 — A, 01 — B, 10 — C, 11 — D. После такого двойного перекодирования мы получим в исходном алфавите новую конечную схему с новыми вероятностями. Чтобы подсчитать эти вероятности, снова прибегнем к тому же приему и в результате получим:

$$P'(A) = 1/4, \quad P'(B) = 1/4, \quad P'(C) = 1/4, \quad P'(D) = 1/4.$$

Таким образом, новая конечная схема, полученная после двойного перекодирования, будет

$$\begin{pmatrix} A & B & C & D \\ 1/4 & 1/4 & 1/4 & 1/4 \end{pmatrix}. \quad (4.14)$$

Как нетрудно видеть, энтропия на символ в новой схеме H_3 равна 2, т. е. информационной емкости применяемого алфавита.

Максимальная энтропия данного алфавита равна его информационной емкости и достигается при равенстве вероятностей всех символов. Всякий конкретный код с определенной энтропией можно оценить с точки зрения его экономичности, т. е. близости его к оптимальному коду, обладающему максимальной энтропией. Для этого достаточно взять отношение значения энтропии данного кода к максимальному ее значению в том же алфавите. Это отношение носит название относительной энтропии h

$$h = \frac{H}{H_{\max}}. \quad (4.15)$$

Наибольшее значение относительной энтропии равно единице.

В вышеприведенном примере относительная энтропия сообщений в исходном алфавите по схеме (4.11) равна $7/4 : 2 = 7/8$; после двойного перекодирования относительная энтропия стала равна 1. Таким образом, двойное перекодирование дало «сжатие» сообщений в $7/8$ раза, т. е. для передачи той же информации по новому коду потребуется число символов на $1/8$ меньше, чем по исходному.

Вычитая величину относительной энтропии из ее максимального значения, мы получим число, называемое избыточностью кода r и служащее мерой числа излишних символов в передаче

$$r = 1 - h = \frac{H_{\max} - H}{H_{\max}}. \quad (4.16)$$

В приведенном выше примере: для исходных сообщений $r_0 = 1 - 7/8 = 1/8$, для кода № 1 $r_1 = 1 - 0,875 = 0,125$, для кода № 2 $r_2 = 0$, для окончательно полученной схемы (4.14) $r_3 = 0$.

Избыточность равна нулю только при оптимальном коде, когда появление всех символов независимо и равновероятно, как в схеме (4.14). Неравномерность распределения вероятностей, как в схеме (4.11), или же наличие внутренних вероятностных связей (корреляции между появлениями букв) уменьшает относительную энтропию и увеличивает избыточность.

При составлении текста мы делаем по нашему произволу гораздо меньше выборов, чем число букв текста. В среднем лишь $1/4 - 1/3$ общего числа букв появляется случайно вследствие выбора нами смысла фразы, подбора выражений и т. д., а остальные приходят вынужденно, благодаря

наличию грамматических законов словообразования и составления фраз.

Избыточность русского и европейского языков превышает половину и составляет, по некоторым подсчетам, от 0,68 до 0,75. Мы могли бы выражаться в 3—4 раза лаконичнее, умей мы пользоваться в качестве языка оптимальным кодом. В этом случае любая комбинация букв составляла бы слово, имеющее смысл, любая таблица, состоящая из букв, представляла бы собой кроссворд, все буквы и все их комбинации применялись бы одинаково часто.

В действительности, конечно, в реальном языке вероятности появления букв и их сочетаний далеко не одинаковы. Если подсчитать частоты появления слов, например в телеграммах, то окажется, что такие слова, как *задерживаюсь, целую, деньги* будут иметь большие частоты, а такие, как *пахта* или *верей* очень малые. Примерно то же самое имеет место и в разговорной речи, и в газетно-журнальных текстах, и в научно-технической литературе. Все реальные сообщения, закодированные в символы национальных языков, имеют значительную избыточность.

Этот факт, конечно, не случаен и не является недостатком принятого способа выражать человеческие мысли. Избыточность полезна, и польза ее состоит в том, что «излишние» символы позволяют без труда исправлять ошибки, опечатки, которые всегда появляются при передаче или воспроизведении информации.

Конечно, никакой аптекарь не рискнул бы выполнить рецепт, написанный типичным врачебным почерком, если бы ошибка в одной букве меняла слово *аспирин* на слово *стрихин*.

Скорость передачи

Всякое устройство, осуществляющее передачу сигналов на расстояние, называется каналом связи. Примером могут служить телеграф, телефон, телеметрическая система, нерв животного и т. д. Канал связи на входном передающем конце получает информацию в виде сигналов от некоторого источника. Далее эти сигналы проходят по каналу до другого его конца — выходного или приемного, где выдаются адресату. Алфавит выходного сигнала не обязательно совпадает с алфавитом сигнала на входе; иначе говоря, в состав канала связи могут входить кодирующие устройства. Так, канал телеизмерения может получать на вход от датчика-измерителя электриче-

ское напряжение, а на приемном конце выдавать кривую, записанную на бумажную ленту.

При передаче сигнала по каналу связи на сигнал тем или иным путем воздействуют шумы или помехи. В результате на приемном конце канала адресат получает нечто отличное от посланного сообщения — сигнал, искаженный помехой. Если рассматривать выход канала как источник сообщений и подсчитать его энтропию, то окажется, что энтропия на выходе больше, чем энтропия на входе. Количество информации возросло, однако ее содержание отнюдь не обогатилось, так как содержанием дополнительной информации является помеха, которая не соответствует описываемому событию.

Если помехи велики, то на выходе канала они преобладают над полезным сигналом и большая часть выходных сигналов канала становится не относящейся к исходному сообщению. Возникает вопрос: какое же количество полезной информации передается при этом от входа к выходу канала?

Сигнал на выходе канала при наличии помех перестает быть однозначно функционально связанным с сигналом на входе, между ними сохраняется только статистическая связь.

Рассмотрим два случайных события: появление на входе канала сигнала X и появление на выходе сигнала Y . Каждое из этих событий имеет определенную вероятность $P(X)$ и $P(Y)$. X и Y не являются независимыми, поэтому вероятность сложного события — одновременного и совместного появления обоих сигналов, $P(X, Y)$, можно выразить с помощью условной вероятности:

$$P(X, Y) = P(X)P(Y|X) = P(Y)P(X|Y), \quad (4.17)$$

$P(X|Y)$ — вероятность появления X , если появился Y ,
 $P(Y|X)$ — вероятность появления Y , если появился X .

Для всех возможных сигналов X и Y , образующих две взаимно-зависимые конечные схемы, можно подсчитать энтропию на символ или в единицу времени. То же можно сделать и для объединения схем X и Y в одну общую конечную схему сложного события (X, Y) .

$$\begin{aligned} H(X, Y) &= \sum_i \sum_k P(X_i, Y_k) \log P(X_i, Y_k) = \\ &= \sum_i \sum_k P(Y_k) P(X_i|Y_k) [\log P(Y_k) + \log P(X_i|Y_k)] = \end{aligned}$$

$$= \sum_i P(X_i|Y_k) \sum_k P(Y_k) \log P(Y_k) + \\ + \sum_k P(Y_k) \sum_i P(X_i|Y_k) \log P(X_i|Y_k), \quad (4.18)$$

а так как $\sum_i P(X_i|Y_k) = 1$ для любого Y_k , как сумма вероятностей полной группы событий, то

$$H(X, Y) = H(Y) + \sum_k P(Y_k) \left[\sum_i P(X_i|Y_k) \log P(X_i|Y_k) \right]. \quad (4.19)$$

Величина $\sum_i P(X_i|Y_k) \log P(X_i|Y_k)$ есть энтропия посланного сигнала (схемы X) при заданном значении принятого сигнала Y_k . Поскольку Y_k — случайная величина, эта энтропия также является случайной величиной. В формулу (4.19) входит ее математическое ожидание по всем возможным сигналам на выходе Y . Эта величина

$$H(X|Y) = \sum_k P(Y_k) \sum_i P(X_i|Y_k) \log P(X_i|Y_k)$$

носит название условной энтропии. Она дает меру количества информации на один выбор (или меру неопределенности до выбора) из множества возможных X , если известен Y , в среднем по всем Y . Итак,

$$H(X, Y) = H(Y) + H(X|Y); \quad (4.20)$$

аналогично получаем

$$H(X, Y) = H(X) + H(Y|X). \quad (4.21)$$

Условная энтропия $H(X|Y)$ не может быть больше безусловной $H(X)$, т. е. всегда $0 \leq H(X|Y) \leq H(X)$. Иначе говоря, знание выходного сигнала Y никогда не увеличивает энтропию X . Из (4.20) и (4.21) следует

$$H(X, Y) \leq H(X) + H(Y), \quad (4.22)$$

причем равенство имеет место только в том случае, если X и Y независимы. В случае взаимозависимости X и Y энтропия перестает быть аддитивной.

Итак, мы имеем перед собой две ситуации. Во-первых, до передачи сигнала нам известна конечная схема X , т. е. возможные сообщения X_i и их вероятности, а следовательно, и энтропия источника $H(X)$, которая характеризует степень неопределенности выбора X из множества возможных. Во-вторых, после получения искаженного сигнала Y мы все еще остаемся перед некоторой, хотя и меньшей, неопределенностью относительно того, какое же из возможных сообщений X_i было в действительности послано. Эта «остаточная» неопределенность характеризуется условной энтропией $H(X|Y)$, которую называют также не-надежностью.

Поскольку $H(X|Y)$ не может быть больше $H(X)$, неопределенность выбора X никогда не может возрасти, а может только уменьшиться в результате получения искаженного сигнала Y . Уменьшение неопределенности происходит потому, что, несмотря на помехи, некоторое количество информации все же было передано по каналу. Это количество информации при исчислении энтропии в единицу времени называется действительной скоростью передачи R . Она равна разности безусловной и условной энтропий

$$R = H(X) - H(X|Y) = H(X) + H(Y) - H(X, Y). \quad (4.23)$$

Иначе говоря, R есть уменьшение неопределенности выбора X после приема сигнала. Это уменьшение неопределенности за счет приема сигнала относительно неопределенности исходной схемы называют иногда отрицательной или «негативной» энтропией или же сокращенно негэнтропией.

Если шум в канале отсутствует, то, как мы видели, $H(X|Y)$ равно нулю и скорость передачи равна энтропии источника; иначе говоря, все количество информации, выдаваемое источником в единицу времени, доходит до адресата. Если шум настолько велик, что условная энтропия равна энтропии источника, то по каналу информация не проходит вовсе и $R=0$. Это, между прочим, не означает, что все сигналы обязательно искажаются. Так, например, если при передаче сигнала в алфавите (0,1) только половина знаков проходит без искажений, а другая половина меняется на противоположные случайным образом, то скорость передачи оказывается равной нулю. Это очевидно хотя бы потому, что в этом случае можно было бы не пользоваться каналом связи, а просто восстанавливать каждый символ с помощью жребия; результат был бы такой же,

как и при использовании канала связи. В самом деле, пусть $P(0) = P(1) = 1/2$. Тогда $H(X) = 1$ *единиц/символ*. После принятия любого символа условные вероятности равны:

$$P(0|1) = P(0|0) = P(1|0) = P(1|1) = \frac{1}{2}.$$

Следовательно,

$$\begin{aligned} H(X|Y) &= P(0)[P(0|0)\log P(0|0) + P(1|0)\log P(1|0)] + \\ &+ P(1)[P(0|1)\log P(0|1) + P(1|1)\log P(1|1)] = \\ &= \frac{1}{2} \left[\frac{1}{2} \log \frac{1}{2} + \frac{1}{2} \log \frac{1}{2} \right] + \frac{1}{2} \left[\frac{1}{2} \log \frac{1}{2} + \frac{1}{2} \log \frac{1}{2} \right] = 1, \end{aligned}$$

т. е.

$$R = H(X) - H(X|Y) = 1 - 1 = 0.$$

Скорость передачи зависит от распределения вероятностей как сообщений, так и шумов. При данных шумах в канале и различных множествах сообщений с различными распределениями вероятностей скорости передачи будут также различны. Если мы будем поочередно испытывать различные источники на входе канала, то для одного из них скорость передачи окажется максимальной и будет равна пропускной способности канала C .

Пропускная способность канала, следовательно, равна

$$C = \max [H(X) - H_y(X)], \quad (4.24)$$

где максимум отыскивается по всем возможным источникам.

Величина C является характеристикой канала связи. Она имеет тот смысл, что при наличии любых шумов, но при правильном кодировании, т. е. при статистическом согласовании источника и канала, по каналу можно передавать в единицу времени количество информации, равное C *ед/сек*, со сколь угодно малой вероятностью ошибки. Если величина C будет превзойдена, т. е. мы захотим передать по каналу количество информации, превышающее C (например, используя более экономичные коды), то вероятность ошибок должна остаться конечной и не сможет быть сведена к нулю. В результате при приеме сигнала появится ненадежность (см. рис. 4.2).

Таким образом, проблема передачи информации по каналу связи приобретает вполне законченную, по крайней мере

ре в принципе, количественную трактовку. Информацию можно рассматривать как физическую величину, которая может быть точно измерена, и вопрос о ее передаче становится аналогичным, скажем, перевозке грузов. Шеннон сравнивает передачу информации с перевозкой леса на конвейере. Если вес и кубатура леса не превышает грузоподъемности и емкости конвейера, лес может быть транспортирован полностью. Нужно только, чтобы форма бревен соответствовала конструкции конвейера. Для правильной упаковки леса и полной загрузки конвейера придется, быть может, некоторые бревна распилить на бруски или даже измельчить до опилок. В случае канала связи это изменение формы соответствует перекодировке сообщений для статистического согласования источника и канала*.

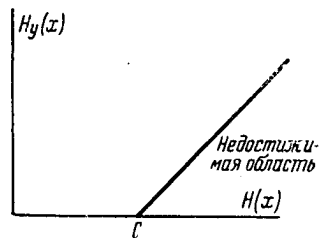


Рис. 4.2. Пенадежность принятого сигнала в зависимости от энтропии на входе канала.

Самокорректирующиеся коды

Понижение вероятности ошибок при передаче достигается разумным использованием избыточности при кодировании сообщений. Чем выше уровень шумов в канале, тем больше вероятность ошибки при передаче и тем больше лишних символов нужно применить для того, чтобы устранить ошибки. Именно этим ценна высокая избыточность человеческого языка. Отдельные опечатки в тексте делают слова бессмысленными, т. е. такими, которые ничему не соответствуют и не применяются. Это позволяет без труда обнаружить их и исправить.

То же свойство используется при составлении помехоустойчивых или, как их еще называют, самокорректирующихся кодов. Простейшим способом защиты от помех является многократное повторение передачи. При достаточном числе повторений правильный вариант сигнала будет приниматься несколько чаще, чем любой из искаженных. Есть,

* Мы опускаем рассмотрение вопроса о количестве информации, содержащейся в сообщении, которое закодировано в непрерывном алфавите. Исчерпывающее изложение вопроса читатель найдет у А. Н. Колмогорова [5].

однако, и более тонкие методы составления самокорректирующихся кодов. Рассмотрим примеры.

Пусть мы передаем сигнал в алфавите (0,1), причем символы следуют группами, по N символов в группе. Вероятность появления в группе более одной ошибки достаточно мала и ею можно пренебречь. Иначе говоря, в группе может появиться только одна ошибка (0 вместо 1 или 1 вместо 0), но почти никогда не появится две или больше таких ошибок.

Припишем к каждой группе из N символов еще один, который не несет информации, а служит контрольным. Будем выбирать его для каждой группы так, чтобы вся группа из $N+1$ символов при суммировании нулей и единиц, в нее входящих, давала бы, скажем, четное число. На приемном конце канала все группы проверяются на четность суммы и та группа, в которой ошибочно изменен один символ, немедленно будет опознана как ошибочная. Мы получили код, обнаруживающий одиночную ошибку. Его недостаточно для того, чтобы указать место ошибки в группе и, следовательно, исправить ее.

Можно составить код и для исправления ошибки, только для этого придется увеличить избыточность и приписать дополнительные символы. Пусть в предыдущем примере $N=4$; припишем к каждой группе из 4 символов еще 3 контрольных. Пусть символы, стоящие на 3, 5, 6 и 7-м местах, используются для передачи информации, а символы, стоящие на 1, 2 и 4-м местах — для контроля. Их мы выбираем по следующим правилам: составляем три вспомогательные суммы, в каждую из которых входит один контрольный символ, выбранный так, чтобы сделать ее четной:

$$\begin{aligned} \text{выбор } X_1: & X_3 + X_5 + X_7 + X_1 = \gamma \text{ четное;} \\ \text{выбор } X_2: & X_3 + X_6 + X_7 + X_2 = \beta \text{ четное;} \\ \text{выбор } X_3: & X_5 + X_6 + X_7 + X_4 = \alpha \text{ четное.} \end{aligned} \quad (4.25)$$

Группы из 7 символов передаются по каналу с искажением. На приемном конце канала вычисляются суммы α , β и γ . Четные обозначаются нулем, нечетные — единицей. Двоичное число $\alpha\beta\gamma$ дает номер неверно переданного символа в данной группе.

Пусть, например, $X_3=0$, $X_5=1$, $X_6=1$, $X_7=0$. Тогда $X_1=1$, $X_2=1$, $X_4=0$. Передается группа 1100110. Пусть в третьем символе произошла ошибка и группа принята в следующем виде: 1110110. Вычислив по принятой груп-

не величины контрольных сумм $\alpha=0$, $\beta=1$, $\gamma=1$, получим число 011*, которое в двоичном начертании равно трем. Это индекс ошибочно принятого символа. Ошибка в любом другом символе была бы исправлена подобным же образом.

В самокорректирующемся коде не все слова «имеют смысл», т. е. являются допустимыми к употреблению. Проверка принятого сигнала на наличие ошибок сводится к опознаванию недопущенных слов и к определению, из какого допущенного слова вероятнее всего получилось в результате искажения данное недопущенное. Это, по-видимому, тем легче сделать, чем большим числом символов отличаются друг от друга допущенные слова.

Пример кода, который мы приводили выше, носит название кода типа Голей—Хемминга. Коды этого типа применимы, когда вероятность ошибки невелика. При больших вероятностях ошибки, когда в слове из N символов достаточно часто появляется более одной ошибки, код Хемминга оставляет без изменений слова, в которых нет ошибок, корректирует одиночные ошибки, не исправляет слова с четным числом ошибок и может внести одну дополнительную ошибку, если число ошибок больше одной и нечетное.

В таких случаях можно применять более сложные коды, например с повторным применением кодирования. Так, например, после исправления одиночных ошибок избыточные контрольные символы отбрасываются и из оставшихся символов составляются новые группы, состоящие из информационных и контрольных символов, и т. д. В результате достигается возможность передать некоторую информацию почти без ошибок, несмотря на значительную вероятность ошибок в отдельном символе, правда ценой значительной задержки во времени. Задержка при кодировании—общее явление для всех сложных кодов.

Передача телеграмм с последовательным исправлением ошибок повторным кодированием дает результат только спустя некоторое время. Вот как примерно выглядит подобная передача по мнению одного из изобретателей кодов.

Телеграмма, принятая с возможными искажениями, вручается адресату с припиской: «Вероятность ошибки на букву равна 0,1; исправления—завтра». На завтра, после того как приняты следующие телеграммы с дополнитель-

* О двоичных числах см. гл. 7.

ными контрольными символами и найдены поправки, адресату вручается извещение: «Вместо *гетры* читайте *метры*, вместо *мыло* читайте *мило*; вероятность ошибки на букву теперь равна 0,01; дополнительные исправления—на следующей неделе». Через неделю адресату вручается новое извещение: «Вместо *кости* читайте *гости*, вероятность ошибки на букву—0,001; ждите следующего месяца». Это продолжается до тех пор, пока адресат не попросит больше не беспокоить его извещениями.

Теорема Шэннона

Самокорректирующиеся коды позволяют в ряде случаев с успехом пользоваться каналом связи даже тогда, когда уровень шумов в нем относительно высок. Вопрос о построении помехоустойчивых кодов для всех возможных случаев и в любом алфавите еще не решен до конца. Использование самокорректирующихся или других рациональных кодов требует, разумеется, разработки соответствующих автоматических кодирующих и декодирующих устройств.

Каково же то предельное количество информации, которое может быть передано без ошибок по каналу с шумами? Иначе говоря, чему равна пропускная способность канала? Ответ на этот вопрос, по крайней мере для одного весьма часто встречающегося случая, дает теорема Шэннона для канала с шумами. Пропускная способность канала вычислена точно только для случая гауссовых шумов, мощность которых равномерно распределена по частоте в пределах полосы пропускания канала при ограниченной средней мощности шумов и сигнала и их статистической независимости. Для этого случая пропускная способность канала равна

$$C = W \log_2 \frac{P+N}{N} \text{ ед/сек}, \quad (4.26)$$

где P — средняя мощность сигнала,
 N — средняя мощность шумов,
 W — полоса пропускания канала.

Не заботясь о строгости доказательства, можно вкратце пояснить вывод этой формулы следующим образом. Любая случайная функция с ограниченной полосой частот W может быть полностью описана $2WT$ ее значениями за время T . Значения амплитуды сигнала, смешанного с шумом,

имеют в среднем величину $\sqrt{P+N}$ (складываются мощности, которые пропорциональны квадратам амплитуд). Из-за наличия шумов различение амплитуд выходной функции происходит с точностью до величины средней амплитуды шума — \sqrt{N} . Стало быть число различаемых значений амплитуды на выходе в среднем равно $\sqrt{P+N}/\sqrt{N}$. А так как в единицу времени проходит $2W$ дискретных значений функции, то пропускная способность равна

$$C = 2W \log_2 \sqrt{\frac{P+N}{N}} = W \log_2 \frac{P+N}{N}. \quad (4.27)$$

Для очень малых отношений сигнал/шум можно написать приближенно

$$C = W \log_2 \left(1 + \frac{P}{N}\right) \approx W \frac{P}{N} \log_2 e = 1,443W \frac{P}{N}. \quad (4.28)$$

Действительная скорость передачи может как угодно мало отличаться от пропускной способности канала, если применить правильный код. Передавать информацию без ошибок с большей скоростью невозможно.

В формулу Шэннона входит как полоса пропускания W , так и мощность сигнала P . Можно сохранить пропускную способность канала, уменьшив мощность передатчика и соответственно увеличив полосу пропускания или наоборот. При этом, разумеется, кодирование должно также измениться. Такая возможность действительно используется в системах связи. Например, различные типы так называемой импульсно-кодовой модуляции позволяют это осуществить.

Исследования показывают, что существующие системы связи далеки от достижимого предела в смысле использования их пропускной способности. Для подтверждения этого мы можем привести результат простого расчета, приводимого Шэнноном. Если телефонный канал будет работать оптимальным кодом для передачи нормальной речи со скоростью 100 слов в минуту при наличии шумов, меньших в 10 раз по амплитуде, чем сигнал, то при передаче с точностью до смысла (т.е. не передавая интонаций) потребуется полоса пропускания всего 2,3 герца, вместо обычно используемых при тех же условиях 1,5—5 тысяч герц. Очевидно, разработка кодов, кодирующих устройств и их использование могут принести немалую экономию.

Информационная и физическая энтропия. Организация

Понятие количества информации в теории информации тесно связано с понятием физической энтропии. Этому не приходится удивляться, так как сигнал есть отражение явлений физического мира и информация в семантическом значении этого слова есть соответствие сигналов и событий. Следовательно, случайный характер явления, распределение вероятностей и физическая энтропия просто переносятся на сигнал, когда физическое явление представляет собой источник сообщений, которые перекодируются в алфавит сигнала. Если бы мы задались целью описать такое явление в газе, которое переводит его в новое состояние, связанное с изменением физической энтропии на величину $H_{\text{ф}}$, то минимальное количество информации, содержащееся в таком сообщении, или его информационная энтропия $H_{\text{и}}$, были бы численно равны изменению физической энтропии $H_{\text{ф}}$.

Физическая энтропия определяется обычно только для физических явлений, связанных с молекулярными процессами (тепловые явления, диффузия, химические реакции). Но не только такие явления могут служить источниками информации. Можно, по-видимому, воспользовавшись понятием информационной энтропии, обобщить понятие физической энтропии на более широкий круг явлений—всех явлений, могущих быть источником сообщений, или же ввести некоторое новое понятие, служащее для той же цели.

Именно таким путем было введено появившееся недавно понятие организации, которое применимо к широкому кругу объектов или систем и которое дает числовую меру статистических связей отдельных элементов, входящих в систему.

Интуитивно мы отождествляем отсутствие организации с хаосом. Пусть мы имеем некоторую систему, состоящую из отдельных частей или элементов, не обязательно однородных. Такой системой может быть не только совокупность молекул физического тела, но и собрание клеток в ткани живого организма, механизм, состоящий из связанных между собою твердых тел, производственное предприятие, состоящее из отдельных цехов, коллектив людей, воинское соединение, состоящее из подразделений, и т. д.

Каждый из элементов в процессе существования или деятельности системы может принимать некоторое множество состояний, над которым задано распределение вероятностей. Если все элементы системы разобщены и состояние одного никак не влияет на состояния других, словом, если элементы независимы, то очевидно, что система является предельно неорганизованной. Другим предельным случаем является такой, когда состояние одного элемента однозначно определяет состояния всех остальных.

Подсчитаем обычным образом энтропию множества состояний каждого элемента системы X_i .

$$H_i(X_i) = - \sum_{k=1}^n P_{ik} \log P_{ik},$$

где $k = 1, 2, \dots, n$ — число возможных состояний элемента X_i и P_{ik} — вероятность элементу X_i оказаться в состоянии k .

Подсчитаем теперь энтропию системы в целом. В случае независимости элементов в предельно неорганизованной системе общая энтропия будет максимальной и равна сумме энтропий отдельных элементов. Так, для системы, состоящей из двух элементов

$$H_0(X_1, X_2) = H(X_1) + H(X_2).$$

Если в системе устанавливаются связи элементов и взаимные влияния, т. е. если система организуется, то общая энтропия становится меньше суммы энтропий элементов. В соответствии с (4.20), (4.21)

$$H_1(X_1, X_2) = H(X_1) + H(X_2 | X_1) = H(X_2) + H(X_1 | X_2) \leq H(X_1) + H(X_2).$$

В результате организации системы ее энтропия H_1 понизилась по сравнению с максимальной энтропией H_0 . Это уменьшение энтропии может служить числовой мерой степени организованности системы Δ .

$$\Delta = H_0 - H_1 = H(X_1) + H(X_2) - H_1(X_1, X_2) \quad (4.29)$$

или, согласно (4.23),

$$\begin{aligned} \Delta &= H(X_1) - H(X_1 | X_2) = H(X_2) - H(X_2 | X_1) = \\ &= H_1(X_1, X_2) - H(X_2 | X_1) - H(X_1 | X_2). \end{aligned} \quad (4.30)$$

Мы снова пришли к негэнтропии, как и в случае передачи информации по каналу связи. Передача информации является, таким образом, частным случаем организации.

Рассмотрим простой пример. Пусть мы имеем систему, состоящую из двух элементов. Множество возможных состояний каждого есть отрезок прямой, над которым задано распределение вероятностей (или плотность вероятности). Распределение вероятностей состояний системы будет задано над прямоугольником, стороны которого суть отрезки, соответствующие множествам состояний элементов.

Для неорганизованной системы распределение плотности вероятности над прямоугольником есть просто произведение плотностей вероятности на отрезках. При максимальной организации совместное распределение стягивается в линию. Промежуточный случай соответствует преимущественному группированию вероятности в области вокруг линии (рис. 4.3).

Максимальная организация системы соответствует, таким образом, соотношениям функционального типа между состояниями элементов и системы в целом. Уменьшение организации ведет к тому, что выбор определенного состояния одного элемента еще не определяет однозначно состояний остальных. Каждый элемент имеет еще некоторую «свободу», т. е. состояние его остается случайным и лишь статистически связанным с состояниями остальных. При полной дезорганизации все элементы становятся независимыми.

Для физических систем уменьшение организации соответствует повышению физической энтропии (уменьшению негэнтропии).

Механизм с разболтанными шарнирами обладает меньшей организацией по сравнению с механизмом, у которого шарниры пригнаны точно. На заводе, обладающем высокой организацией, цехи работают строго согласованно по времени, номенклатуре и качеству изделий. Разброс качества изделий определяется только разбросом качества сырья. Цехи завода, обладающего низкой организацией, работают вразброд (то недодавая другим цехам полуфабрикаты, то забывая ими промежуточные склады), несогласованно во времени, с большим количеством брака. Из хорошего сырья такой завод производит много некондиционных и бракованных изделий.

Подразделения высокоорганизованной воинской части ведут бой взаимосогласованно, строго по плану операции и не совершают никаких произвольных маневров. По мере снижения организации часть постепенно превращается

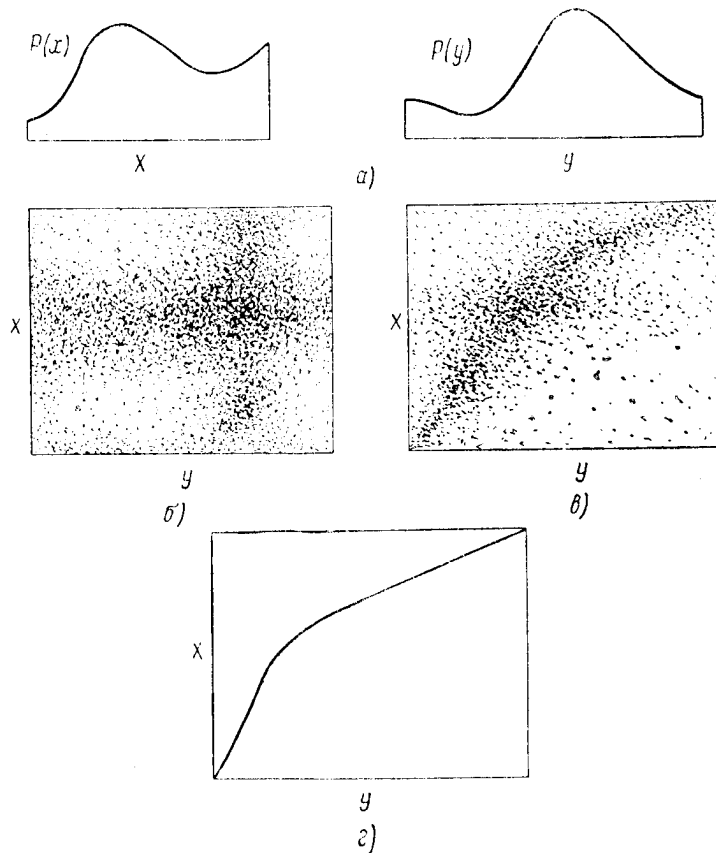


Рис. 4.3. Пример систем, обладающих различной степенью организации.

а) распределение плотности вероятности над отрезками X и Y ; б) распределение плотности вероятности над прямоугольником со сторонами X и Y (плотность вероятности пропорциональна плотности точек на рисунке), случай организации, равной нулю; в) то же, организация отлична от нуля; г) то же, организация максимальная — между X и Y установлена функциональная зависимость.

в независимо действующие подразделения, а затем в толпу вооруженных людей.

Живой организм, пораженный раковой опухолью; нме-

ет меньшую организацию по сравнению со здоровым. Функции больного организма, его реакции на внешние воздействия также дезорганизируются, т. е. приобретают характер неоднозначности, случайности.

В любом конкретном случае статистическое описание системы и ее деятельности позволяет определить числовую меру степени ее организации. Разумеется, конкретные задачи такого типа могут оказаться очень сложными и потребовать для своего решения затраты большого труда.

Повышение организации системы является следствием усиления связей между ее элементами, в частности, введения управления с помощью сигналов. Именно поэтому кибернетические системы являются системами организованными. Введение числовой меры организации придает этому интуитивному понятию строгий смысл.

Можно было бы связать понятие организации и количества информации теснее, поставив, например, вопрос о количестве информации, необходимом для организации определенной системы. Такая постановка вопроса законна в тех случаях, когда приходящая извне информация не только запоминается системой, но и влияет на ее организацию. Так дело обстоит в так называемых самоорганизующихся системах (см. гл. II). Можно было бы рассматривать любую организованную систему как систему, накопившую «предварительную» информацию, которая и позволяет ей выполнять свои функции.

Эти построения затрудняются при наличии нескольких «этажей» управления, когда одна часть системы непосредственно осуществляет обработку информации по определенному алгоритму, другая определяет этот алгоритм по какому-либо иному правилу, третья изменяет это последнее правило и т. д. Возможно, что для таких случаев потребовались бы своеобразные разделения оценок организации по «этажам».

Пока что количественная оценка организации еще не нашла широкого применения на практике.

Подведем некоторые итоги. События, внешние по отношению к кибернетической системе, воздействуют на систему таким образом, что заставляют ее изменять свое состояние. При этом одни состояния системы встречаются чаще, другие — реже. Каждое состояние соответствует определенному внешнему событию. Это соответствие представляет собой информацию о событии. Тот физический объект или процесс, который в пределах кибернетической системы испытывает измене-

ния состояния и, следовательно, является носителем информации, есть сигнал. Чем больше разнообразие состояний системы, чем обширнее множество и чем длительнее отрезок времени, тем подробнее может быть описано внешнее событие. Количество информации, передаваемое или запаасаемое системой в виде сигнала, зависит не только от множества состояний, но и от распределения вероятностей этих состояний. Содержание сообщений (или конкретная информация), физический алфавит сигнала и энтропия (или количество информации) являются независимыми явлениями и каждое может изменяться независимо от других.

Сигнал как физический процесс связан с внешним событием энергетически. Хотя сигнал существует в системе за счет энергии, которая черпается из источников, принадлежащих системе, а не за счет энергии, связанной с событием, породившим сигнал, тем не менее возникновение сигнала, первичное изменение состояния системы под воздействием изменений во внешней среде происходят с поглощением энергии. Иногда энергия, идущая на создание сигнала, ничтожно мала по сравнению со всей энергией, которая связана с событием. Иногда же, когда само событие включает в себе ничтожно мало энергии, создание сигнала требует затраты значительной доли или даже всей этой энергии. Это имеет место, например, при измерениях параметров элементарных частиц — электронов, протонов и т. д. По-видимому, с этим связаны трудности полного описания микросостояний вещества, выраженные в физическом принципе неопределенности.

Если крупное событие большого энергетического масштаба, например производственную операцию по обработке крупной детали на станке, спортивное соревнование и т. п., мы можем наблюдать со многих точек зрения, пользоваться измерительными и записывающими приборами, «снимать» с этого события многие сигнальные процессы, увеличивая объем информации и подробность описания события, и при этом никак не влияя на ход самого события, то наблюдение положения движущейся частицы мы сможем произвести только в том случае, если эта частица оставит след или, иначе говоря, если произойдет взаимодействие ее с другими частицами, что потребует затраты заметной доли энергии частицы и изменит ее состояние.

То обстоятельство, что информация в виде сигнала не может существовать вне физических объектов, позволяет, между прочим, подойти к решению очень интересного па-

*

радокса о так называемом энтропийном «демо» Максвелла — воображаемом существе, которое, находясь в равновесной физической системе, достигшей состояния с максимальной физической энтропией, «сортирует» молекулы газа, пропуская их через отверстие с заслонкой в отдельный сосуд, если они обладают высокой скоростью, и не пропуская, если скорость мала. Такая сортировка привела бы в конце концов к понижению энтропии в изолированной системе, вопреки второму закону термодинамики, так как в одном сосуде самопроизвольно сосредоточились бы молекулы с высокой энергией, а в другом — с низкой. В результате, можно было бы использовать образовавшуюся разность температур для получения полезной работы без затраты энергии извне. Если мыслить максвелловского демона не как фантастическое существо, способное творить чудеса, а как реальный прибор, то вопрос о его работе приобретает большую принципиальную и практическую остроту. Второй закон термодинамики издавна подвергается произвольным толкованиям, и в наши дни то и дело появляются варианты «вечных двигателей второго рода», в которых делаются попытки более или менее замысловато обойти этот закон природы.

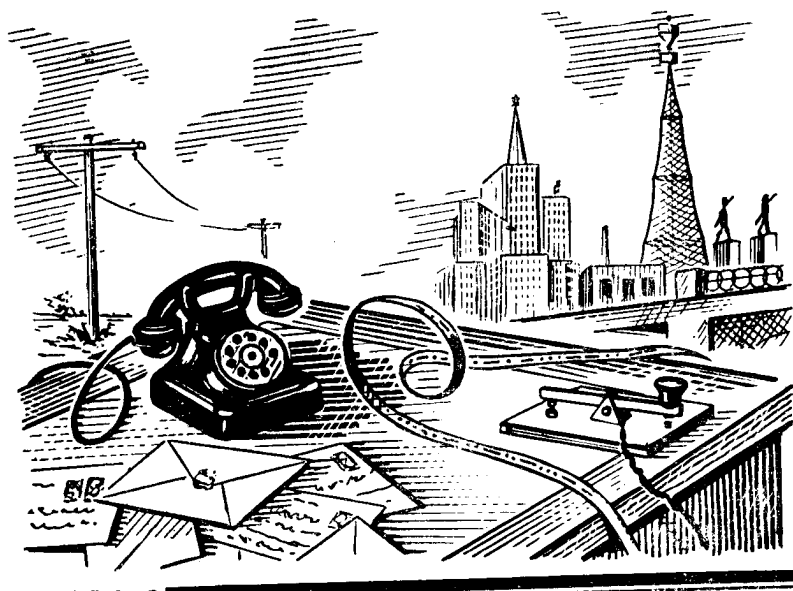
Демон Максвелла, находясь внутри равновесной системы, не может различить молекул друг от друга, так как на него со всех сторон падает одинаковое количество излучения одинакового спектрального состава. Это обстоятельство есть одно из свойств равновесного состояния. Для того чтобы иметь возможность сортировать молекулы по скоростям, демон должен получать информацию о положениях и скоростях молекул. Однако такая информация может быть получена только с помощью сигнала, т. е. процесса, связанного с энергией.

Внутри равновесной системы демон может получить информацию только при условии, если он будет применять некоторое дополнительное информационное устройство, сигнализирующее ему о состояниях молекул. Так, например, он может освещать каждую молекулу светом от отдельного источника с достаточной яркостью для того, чтобы молекула была надежно видна на фоне равновесного излучения, или же со спектральным составом, отличным от равновесного излучения (светом более короткой волны). Если провести подсчет энергии, необходимой для создания такой информации даже исходя из минимальных заниженных требований, предположив, например, что достаточно одного кванта света на молекулу, то окажется, что выиг-

Sanitized Copy Approved for Release 2010/06/16 : CIA-RDP80T00246A045100330001-4

рыш свободной энергии, даваемый демоном, не окупит затрат энергии на получение информации. Замкнутая система «газ—демон—информационное устройство» будет постепенно повышать свою физическую энтропию, делая это лишь несколько медленнее, чем делал бы неуравновешенный газ без демона. Максвеллов демон лишь немного замедляет естественный процесс повышения энтропии, но не приостанавливает его и уж во всяком случае не заставляет энтропию понижаться.

Sanitized Copy Approved for Release 2010/06/16 : CIA-RDP80T00246A045100330001-4



ГЛАВА 5

ПЕРЕДАЧА СИГНАЛА

Сигнал может быть передан от одного кибернетического устройства к другому. Это одно из самых существенных свойств сигнала. Именно это свойство позволяет пространственно разделенным частям системы или различным системам действовать согласованно. Иногда передача сигнала производится на очень большие расстояния (в системах телеуправления, в сетях связи и так далее), иногда же — на малые расстояния в пределах одного устройства (передача числа из блока памяти в арифметическое устройство в электронной вычислительной машине, передача возбуждения по перву).

Техника передачи сигнала имеет дело с физическими процессами, которые переносят сигнал, с анализом протекания этих процессов во времени и пространстве, с физическими свойствами систем, в которых эти процессы протекают. Естественно, что для передачи сигналов используются те физические процессы, свойством которых является распространение, перемещение в пространстве. К их числу относятся волновые процессы различного рода (звуковые

волны, радиоволны, свет) или механические перемещения предметов (почта).

Поскольку всякий сигнал для того, чтобы переносить информацию, должен иметь возможность принимать различные значения или состояния, то же требование предъявляется и к физическому процессу, используемому для передачи сигнала. В упрощенном виде дело можно было бы представить так, как будто бы от передатчика к приемнику непрерывно движется нечто, на чем передатчик может делать пометки в виде символов некоторого алфавита. Эти пометки, распространяясь вместе с их носителем, доходят до приемника, где пометки снимаются с носителя

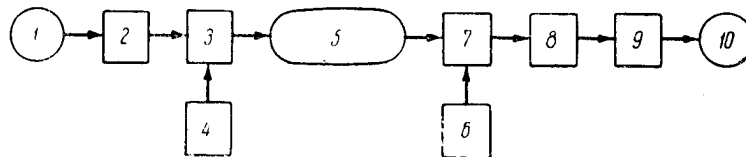


Рис. 5.1. Схема канала связи.

1 — источник сообщений — внешнее событие, 2 — кодирующее устройство — измеритель-датчик, корреспондент, 3 — модулятор-передатчик, 4 — генератор несущих колебаний, 5 — среда передачи, 6 — источник шумов, 7 — приемник, 8 — демодулятор, 9 — декодирующее устройство, 10 — исполнительное устройство — адресат

и переносятся или преобразуются в иную форму. Так, например, в случае радиотелефона «нечто», которое перемещается от передатчика к приемнику, представляет собой электромагнитное излучение, а «пометки» — модуляцию.

Модуляция

Модуляцией называется изменение во времени одного из параметров (характеристик или свойств) физического процесса, несущего информацию. Чаще всего используется изменение мощности колебаний. Автомобильный гудок, звонок, сирена либо молчат, либо включаются на полную громкость. Их включение и выключение есть модуляция звуковых колебаний. Мигание светового сигнала — пример модуляции света. Вибрация звука при пении или игре на скрипке — пример модуляции звуковых колебаний по частоте.

Чаще всего понятие модуляции применяется к волновому процессу, т. е. к процессу распространения колебаний. Волновой процесс описывается зависимостью от времени и расстояния некоторой физической величины: давления

газа в случае звука, напряженности электрического или магнитного поля в случае радиоволны. В простейшем случае эта зависимость представляет собой синусоидальную функцию

$$u = A \sin(\omega t + \varphi_0). \quad (5.1)$$

Эта функция является решением линейного дифференциального уравнения, которое описывает подавляющее большинство реально встречающихся процессов колебаний.

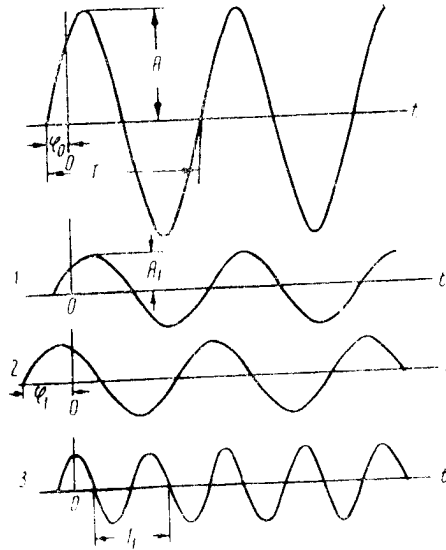


Рис. 5.2. Временная диаграмма синусоидального колебания (A — амплитуда колебания, T — период колебания, φ_0 — начальная фаза колебания, $f = 1/T$ — частота колебаний).

1 — колебание с измененной амплитудой, 2 — колебание с измененной начальной фазой, 3 — колебание с измененным периодом (частотой).

Точнее, решением является не синусоидальная, а экспоненциальная функция вида:

$$u = C e^{z t},$$

где C и z — комплексные числа.

Синусоидальная функция представляет собой частный случай экспоненциальной для чисто мнимого аргумента z .

Всякую более сложную периодическую функцию можно представить как сумму синусоидальных.

Параметр A в (5.1) (наибольшее значение u) называется амплитудой синусоиды. Значения синусоидальной функции повторяются через равные промежутки времени T . Если заметить какое-либо значение, например $u = 0$, на восходящем участке кривой, то то же самое значение u тоже на восходящем участке снова появится через промежутки T , $2T$

$3T$ и т. д. T называется периодом колебания, число периодов в секунду $1/T = f$ — частотой колебания, а частота, умноженная на 2π ($2\pi f = \omega$) — круговой частотой. Величина $\omega t + \varphi_0 = 2\pi f t + \varphi_0$ называется фазой или текущей фазой колебания, а φ_0 — начальной фазой.

Для волнового процесса аргументом синусоидальной функции является величина $\Phi = \left(t - \frac{x}{v}\right)$, которая дает зависи-

мость и от расстояния, и от времени. Для некоторой точки пространства, положив x равным постоянной величине (координате точки наблюдения), мы получим синусоидальную зависимость от времени (синусоидальное колебание); для некоторого фиксированного момента времени, положив t равным постоянной величине, получим синусоидальную зависимость от расстояния. Зафиксировав текущую фазу вол-

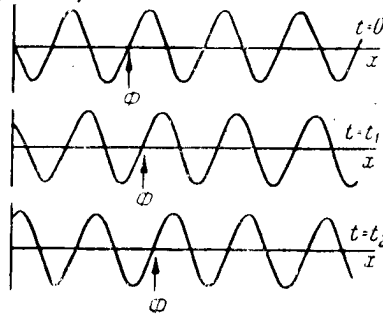


Рис. 5.3. Бегущая волна в пространстве в последовательные моменты времени (Φ — положение отмеченной фазы колебания).

ны $\psi = \omega \left(t - \frac{x}{v}\right) + \varphi_0$, мы заметим, что координата этой фазы меняется со временем; она перемещается со скоростью v . Волна „бежит“, сохраняя по пути свою форму. Функция $u = \sin \omega \left(t - \frac{x}{v}\right) + \varphi_0$ описывает немодулированную волну, распространяющуюся вдоль оси x со скоростью v . В качестве модуляции можно использовать изменение любой из величин, параметров, в выражении для u (5.1) A , ω или φ_0 . Можно, например, изменять по синусоидальному или другому закону амплитуду волны. Мы получим амплитудную модуляцию

$$u = [A_0 + A_0 m \cos \Omega \Phi] \sin (\omega \Phi + \varphi_0). \quad (5.2)$$

Изменяя синусоидально начальную фазу, будем иметь фазовую модуляцию

$$u = A_0 \sin (\omega \Phi + \varphi_0 \cos \Omega \Phi). \quad (5.3)$$

Изменяя частоту, получим частотную модуляцию

$$u = A_0 \sin [(\omega_0 + \omega_1 \cos \Omega \Phi) \Phi + \varphi_0]. \quad (5.4)$$

Иногда частотную и фазовую модуляции ввиду их сходства объединяют под общим названием угловой модуляции (так как при обоих видах модуляции происходит изменение фазы или фазового угла, но по разным законам).

Функция, описывающая закономерность изменения параметра волны, называется модулирующей функцией.

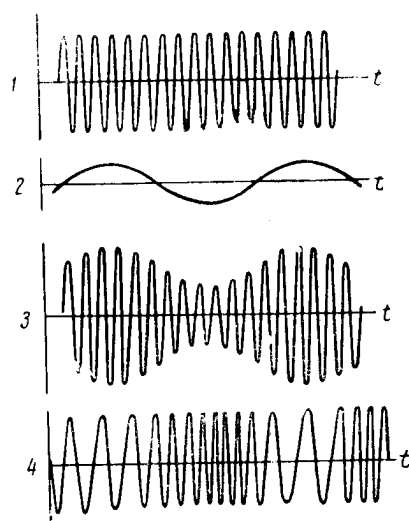


Рис. 5.4. Модуляция.

1 — несущее колебание, 2 — модулирующая функция, 3 — амплитудно-модулированное колебание, 4 — частотно-модулированное колебание.

Только модулирующая функция несет информацию. Это может быть колебание, повторяющее речь или музыку, или же условные знаки — символы телеграфной передачи. Основное колебание с частотой $f = \frac{\omega}{2\pi}$ представляет собой лишь средство перенесения информации. Оно называется несущим колебанием. Его роль в передаче сообщения подобна роли бумаги, на которой написано письмо.

На приемном конце канала связи выделяется модулирующая функция, которая в том или ином виде доводится до адресата — человека или системы управления.

Процесс выделения модулирующей функции называется демодуляцией или детектированием. Способ детектирования зависит от вида модуляции.

Рассмотрим несколько примеров. Свет представляет собой, как известно, электромагнитные колебания весьма высоких частот (порядка 10^{15} герц). При передаче сообщений с помощью световых волн источником колебаний несущей частоты является либо сам источник информации, либо посторонний источник света. Приемником является либо глаз животного, либо фотоэлемент. При сигнализации гелиографом используется амплитудная модуляция. Модулирующей функцией чаще всего являются точки и тире

азбуки Морзе. Модуляция в этом случае имеет вид «прямоугольных импульсов»: интенсивность света очень быстро возрастает от нуля до максимума и через некоторый промежуток времени снова быстро падает до нуля.

Сигнализация цветными фонарями типа светофора или железнодорожная сигнализация использует частотную модуляцию, так как свет горит непрерывно с постоянной интенсивностью, но меняет цвет, т. е. частоту несущих колебаний.

Зрение представляет собой связь многоканальную. Глаз воспринимает одновременно очень много отдельных сигналов, которые, будучи связаны друг с другом, дают в комбинации сложный сигнал — образ.

Звук также является переносчиком информации. Сигнализация звуком использует амплитудную модуляцию (включение гудка). Устройство типа эхолота, измеряющие глубину водоемов по времени прихода отраженного звукового сигнала, используют амплитудную импульсную модуляцию или же частотную.

Общение людей посредством разговора использует как амплитудную, так и частотную звуковую модуляцию. То же относится и к музыке, но в музыке модуляция заметна отчетливее, так как звуки разной высоты и силы в музыке выявлены рельефнее, чем в речи. Особенностью речи (как, впрочем, и оркестровой музыки) является то, что в звучании присутствуют одновременно колебания различных частот с различными амплитудами. Этот набор колебаний меняется во времени. Отдельные звуки речи — фонемы, звучания отдельных инструментов оркестра, отличаются друг от друга своими спектрами, т. е. различным содержанием колебаний с различными частотами. Тембры голосов позволяющие узнавать знакомых при разговоре по телефону, также определяются содержанием различных частотных составляющих, т. е. спектром. К фазовым соотношениям звуковых колебаний ухо человека нечувствительно.

Процессом, служащим для передачи сигнала, не обязательно должны быть колебания. Передача сигнала без использования колебаний может осуществляться, например, с помощью электростатического или магнитостатического поля. Электрическое поле распространяется вдоль провода и вызывает в проводе изменение электрического тока. Это используется при телеграфировании постоянным током или при телефонной передаче. Модуляция при этом наложена на постоянный ток в проводе.

Иногда используется и постоянное магнитное поле. Не-

давно в печати сообщалось об использовании мощного электромагнита для передачи команд режиссера актерам в киностудии при съемках звуковых фильмов. Посторонние звуки в студии недопустимы, световые сигналы неудобны и недостаточны. Передача с помощью магнитного поля оказалась очень удобной в этом случае. Режиссер наблюдает за съемкой из звукоизолированной кабины и дает указания актерам в микрофон. Токи звуковой частоты модулируют магнитостатическое поле, которое пронизывает студию. Приемник, превращающий изменения магнитного поля в звук, удалось сделать настолько малым, что он

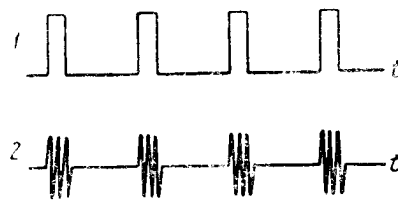


Рис. 5.5. Импульсы.

1 — импульсы прямого («постоянного») тока (видеоимпульсы). 2 — импульсы колебаний (радиоимпульсы).

может быть незаметно размещен возле уха актера и замаскирован прической и костюмом. Телефонная трубка размещается в ушной раковине.

Носителями информации могут быть и такие процессы, как теплопроводность, диффузия, течение жидкости или газа и т. д. Так, сигнал, управ-

ляющий термостатом, достигает чувствительного элемента посредством теплопроводности. Запах дыма, предупреждающий пожарную охрану об опасности, передается диффузией или током воздуха. Возбуждение некоторых нервных центров происходит в том случае, когда током крови к ним приносится измененное по сравнению с нормальным количество химического вещества — гормона. Так происходит, например, возбуждение дыхательного центра головного мозга при увеличении концентрации углекислого газа в крови, омывающей мозг. Вкус — сигнал, несущий информацию о качестве пищи, также передается с помощью химических воздействий при изменениях концентрации вещества.

Из всех процессов, несущих информацию, наибольшее значение в технике связи в настоящее время имеет, пожалуй, распространение электромагнитных колебаний — радиоволны. В радиосвязи применяется и амплитудная, и частотная модуляции. При этом частоты несущих колебаний могут лежать в очень широких пределах (10^4 — 10^{10} герц).

В последнее время в радиотехнике получила распространение так называемая импульсная модуляция, которая

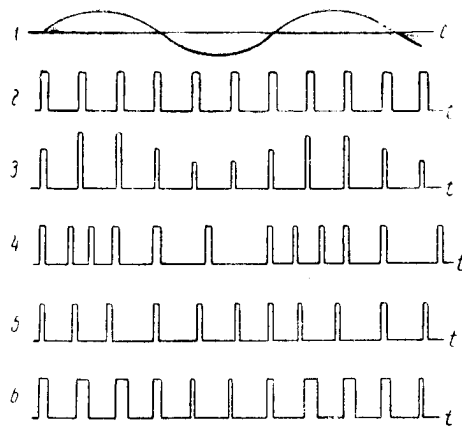


Рис. 5.6. Импульсно-кодовая модуляция.
 1 — модулирующая функция, 2 — несущие импульсы, 3 — модуляция по амплитуде, 4 — модуляция по частоте, 5 — модуляция по положению, 6 — модуляция по ширине.

заслуживает особого упоминания. Импульсы представляют собой протекание тока большой силы в течение очень коротких промежутков времени, разделенных продолжительными паузами. Если рассматривать серию импульсов как процесс, протекающий во времени, то становится ясно, что импульсы можно промодулировать. Модуляцию при этом

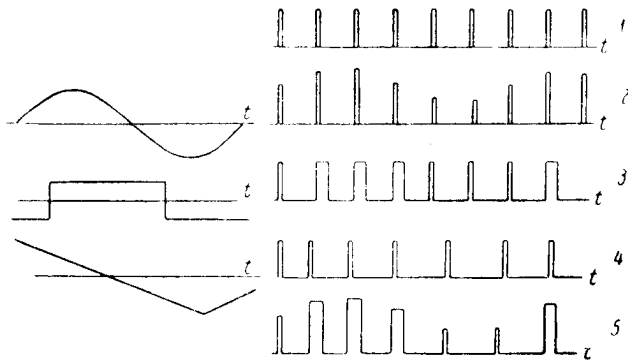


Рис. 5.7. Одновременное наложение модуляций различных типов на один процесс.
 1 — несущие импульсы, 2 — модулирующая функция и амплитудная модуляция, 3 — вторая модулирующая функция и модуляция по ширине, 4 — третья модулирующая функция и частотная модуляция, 5 — совмещение амплитудной, частотной и широтной модуляций с различными модулирующими функциями в одном процессе.

можно осуществить различными способами, например, по амплитуде импульсов, по частоте их повторения, по положению каждого импульса в пределах периода повторения, по длительности. Можно при этом одновременно применять два или три вида модуляции к одной и той же последовательности импульсов независимо друг от друга и для передачи двух или трех различных сообщений одновременно. Число одновременно передаваемых сообщений можно еще увеличить, если в промежутки между импульсами поместить новые серии импульсов, модулированные независимо от первых.

Одновременная передача двух независимых сообщений на одной общей несущей может осуществляться не только при импульсно-кодовой модуляции, но и при использовании непрерывных колебаний. В этом случае используются одновременно как амплитудная, так и фазовая модуляция. Передаваемое колебание складывается из двух независимых колебаний

$$\begin{aligned} u_1 &= A_0 [1 + m_1 f_1(t)] \sin \omega t, \\ u_2 &= B_0 [1 + m_2 f_2(t)] \cos \omega t. \end{aligned} \quad (5.5)$$

Суммарное колебание $u = u_1 + u_2$ оказывается модулированным и по амплитуде и по фазе, причем на приемном конце из него оказывается возможным выделить первоначальные составляющие и их модулирующие функции $f_1(t)$ и $f_2(t)$.

Искажения

По пути от входа передатчика до выхода приемника сигнал претерпевает двоякого рода искажения: во-первых, на него воздействуют шумы, т. е. случайные изменения в среде распространения и в аппаратуре, и, во-вторых, он подвергается искажениям закономерного характера, которые вносятся самой аппаратурой и определяются ее характеристиками. С шумами мы уже знакомы, рассмотрим здесь вкратце закономерные искажения.

Если мы имеем какую-либо систему, обладающую инерцией, и воздействуем на нее некоторой внешней силой, то смещение системы из положения равновесия произойдет не мгновенно, а лишь постепенно.

Так, если мы попытаемся двигать массивное тело по шероховатой поверхности с помощью эластичного шнура или пружины, то, перемещая свободный конец шнура равномерно со скоростью v_0 , мы заставим тело двигаться

с ускорением, и оно достигнет скорости v_0 лишь спустя некоторый (теоретически — бесконечно большой) промежуток времени.

Такое же явление произойдет, если к электрической цепи, состоящей из емкости, индуктивности и сопротивления, приложить электродвижущую силу. Ток в цепи в этом

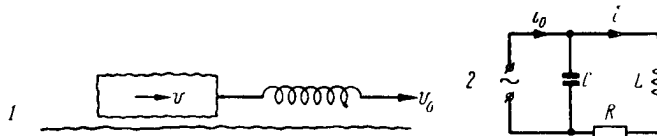


Рис. 5.8. Системы, обладающие инерцией.
1 — тяжелое тело, приводимое в движение по шероховатой поверхности с помощью эластичного шнура, 2 — электрическая цепь, состоящая из сопротивления, индуктивности и емкости.

случае будет нарастать постепенно и приближаться к предельному значению асимптотически. В зависимости от соотношения параметров m , ρ , ϵ (массы, сопротивления трения, эластичности) или L , R , C (индуктивности, сопротивления, емкости) могут возникать различные переходные процессы. Колебательный переходный процесс происходит при малом трении (сопротивлении) по сравнению с инерцией и упругостью (индуктивностью и емкостью). Аперiodический процесс (без колебаний) происходит при большом сопротивлении.

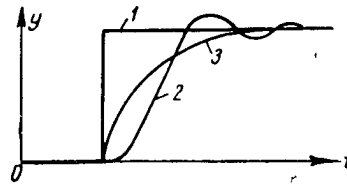


Рис. 5.9. Движение системы, обладающей инерцией, под действием внешней силы.
1 — внешнее воздействие; 2, 3 — различные типы движения.

Запаздывание реакции физической системы на приложенное извне воздействие представляет собой явление весьма общего характера. Оно имеет место и для механических, и для электрических, и для акустических систем. Реакции всех этих систем протекают по сходным законам и могут быть строго сопоставлены друг с другом. При формальном описании различных систем уравнения, описывающие изменения их состояния, совпадают, и только коэффициенты в уравнениях отличаются размерностью. Так, массе тела в механической системе (или моменту инерции в случае вращательного движения) соответствует индуктивность электрической системы и инертность акустической. Эластичности

механической системы (т.е. свойству деформироваться под действием внешней силы) соответствует емкость электрической системы и акустическая емкость. Силе трения соответствует электрическое и акустическое сопротивление.

Аналогия между механическими, электрическими и акустическими явлениями может быть прослежена очень далеко. Многие сложные механические системы можно повторить или «моделировать» электрически и наоборот. Явле-

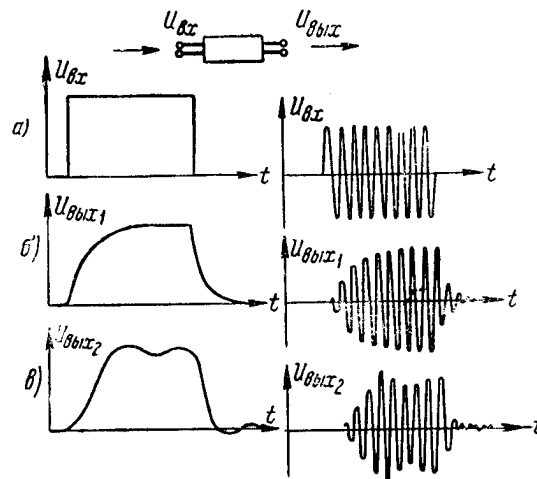


Рис. 5.10. Линейные искажения сигнала.
а) входной сигнал; б, в) искаженный сигнал на выходе

ния инерционности проявляются при передаче сигналов принципиально одинаково в любой системе связи, независимо от ее природы, и поэтому могут рассматриваться с единой точки зрения.

Наличие инерционности ведет к тому, что сигнал, приложенный ко входу системы, оказывается искаженным на ее выходе. Прямоугольный импульс растягивается и изменяет форму. При этом характер искажений однозначно и закономерно определяется устройством и характеристиками системы. Один и тот же сигнал всегда искажается одинаково. Эти искажения в корне отличны от тех, которые вносятся шумами.

Таким образом, если знать устройство системы или ее характеристики, то в принципе возможно по искаженному сигналу восстановить точный вид сигнала, поступившего

на вход. Практически же для такого «восстановления» нужно измерять сигнал с очень большой точностью, что бывает невозможно из-за присутствия шумов.

Спектр и полоса пропускания

Можно подойти к изучению свойств физических систем с иной точки зрения, рассматривая не реакцию системы на изменения входного сигнала, а установившееся

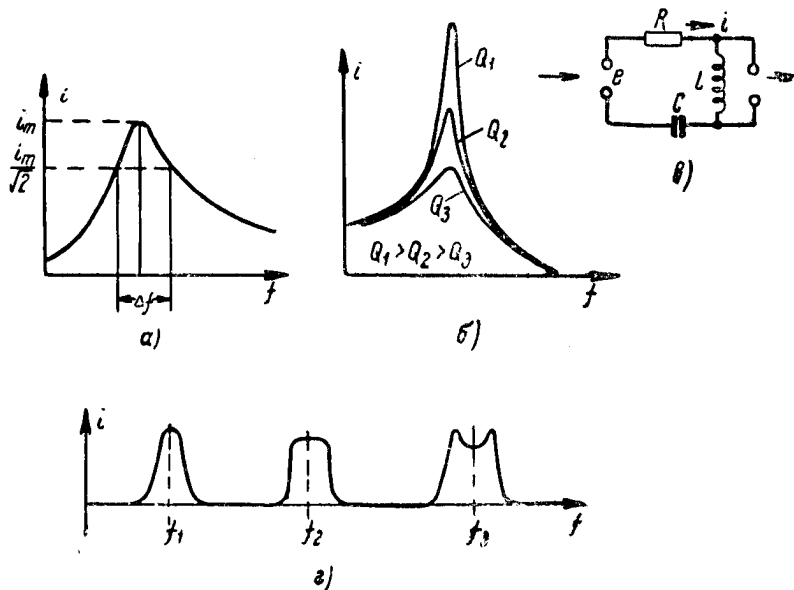


Рис. 5.11. Резонансные кривые.

а) резонансная кривая (отмечена полоса пропускания). б) резонансные кривые при различных добротностях контура, в) последовательный контур, г) резонансные кривые сложных систем.

состояние — величину сигнала на выходе по прошествии достаточно долгого времени после изменения сигнала на входе.

Подавая на вход электрической цепи, изображенной на рис. 5.11,в, синусоидальное напряжение, мы получим на выходе, спустя некоторое время, достаточное для затухания переходных процессов, также синусоидальное напряжение некоторой амплитуды $A_{\text{вых}}$. Изменяя частоту входного напряжения и сохраняя постоянной его амплитуду,

мы заметим, что амплитуда установившегося выходного напряжения изменяется в зависимости от частоты. В некоторых случаях напряжение на выходе может превышать напряжение на входе, в иных же — будет меньше него. Зависимость амплитуды выходного напряжения от частоты, изображенная графически, представляет собой кривую с максимумом и называется «резонансной кривой» или частотной характеристикой цепи. Максимум резонансной кривой может быть выше или ниже в зависимости от соотношения индуктивности, емкости и сопротивления контура (добротности контура Q^*). Добротность контура, а с нею и величина максимума на резонансной кривой, тем выше, чем меньше потери в контуре, т. е. чем меньше сопротивление, включенное последовательно с индуктивностью и емкостью (или проводимость, включенная параллельно им).

Если на резонансной кривой отметить точки, лежащие справа и слева от максимума, для которых амплитуда выходного напряжения меньше максимальной в $1/\sqrt{2}$ раза (т. е. мощность колебаний вдвое меньше, чем в максимуме кривой, так как мощность пропорциональна квадрату амплитуды), и отметить соответствующие этим точкам частоты f_1 и f_2 , то разность этих частот $f_2 - f_1 = \Delta f$, называемая полосой пропускания цепи, окажется также зависящей от добротности: чем выше добротность, тем меньше полоса пропускания. Система с высокой добротностью резко различает частоты, она дает сильную реакцию на колебания с частотой, близкой к резонансу, и слабую — на другие частоты. Такую систему называют «узкополосной». В противоположность этому широкополосные системы имеют размытую резонансную кривую, характерную для низкой добротности, и реагируют на колебания с заметно отличающимися частотами почти одинаково.

Реакция системы на изменение входного напряжения, т. е. переходный процесс в системе, полностью определяется формой резонансной кривой и полосой пропускания. Узкополосные системы обладают большой инерционностью и реагируют медленно, широкополосные — обладают малой инерционностью и реагируют быстро. Если подать на вход цепи напряжение, которое изменяется скачком, то выходное напряжение будет возрастать лишь постепенно.

* Добротностью контура $Q = \omega L/R$ называется величина, обратная затуханию α : $Q = 1/\alpha$. Чем меньше добротность, тем большая доля энергии, запасенной в контуре, рассеивается за один период колебания.

Время, в течение которого выходное напряжение достигает 63% от установившегося значения ($0,63 \approx 1 - 1/e$), называется постоянной времени цепи. Величина постоянной времени обратно пропорциональна ширине полосы пропускания.

Увеличивая полосу пропускания цепи, мы можем понизить инерционность системы и ускорить ее реакцию; однако этого можно достичь лишь ценой понижения амплитуды выходного напряжения. Создание весьма узкополосных и весьма широкополосных цепей представляет значительные трудности, как это часто бывает в технике при решении задач, связанных с приближением к пределу возможности. В реальных системах всегда приходится считаться с наличием ограниченной полосы пропускания и связанной с этим инерционностью системы.

Искажения формы сигнала при прохождении им реальных физических систем можно рассматривать с двух различных точек зрения: во времени и, во-вторых, в области частот. Имея сигнал в виде функции времени $u(t)$, мы можем пользоваться для описания вольт системы понятием „отклика на единичный импульс“ $w(t)$ и получать значение выходного сигнала либо с помощью интегральной вёртки (интеграла Дюамеля)

$$u_{\text{вых}}(\tau) = \int_{-\infty}^{\tau} u_{\text{вх}}(t) w(\tau - t) dt, \quad (5.6)$$

либо с помощью спектральной функции сигнала, применив к сигналу преобразование Фурье

$$F_{\text{вх}}(\omega) = \int_{-\infty}^{+\infty} e^{j\omega t} u(t) dt. \quad (5.7)$$

Умножая в последнем случае спектральную функцию $F(\omega)$ на частотную характеристику системы $Y(\omega)$, получим спектральную функцию выходного сигнала

$$F_{\text{вых}}(\omega) = F_{\text{вх}}(\omega) Y(\omega). \quad (5.8)$$

Производя затем обратное преобразование Фурье, получим выходной сигнал как функцию времени

$$u_{\text{вых}}(t) = \int_{-\infty}^{+\infty} F_{\text{вых}}(\omega) e^{-j\omega t} d\omega. \quad (5.9)$$

Интегрирование в последнем случае производится в комплексной плоскости. По результатам оба способа вычислений вполне равноценны.

Преобразование Фурье интегральной свертки двух функций времени равно произведению преобразований Фурье обеих этих функций. Преобразование Фурье отклика системы на единичный импульс равно частотной характеристике системы.

Операции в области спектров имеют некоторое преимущество благодаря своей наглядности. В сигнале содержатся колебания с различными частотами. При прохождении через систему с конечной полосой пропускания часть колебаний, частоты которых лежат вне полосы пропускания, сильно подавляется, и в спектре выходного сигнала остаются главным образом колебания с частотами, лежащими в полосе пропускания. Искажения сигнала соответствуют исчезновению из спектра некоторых частотных составляющих. Обычно исчезают высокочастотные составляющие, что соответствует сглаживанию формы кривой.

Искажения сигнала, обусловленные конечной шириной полосы пропускания системы, кладут предел возможности передачи большого количества информации через систему. Так, например, при импульсной модуляции посылка все более коротких и частых импульсов приводит к тому, что амплитуда импульсов на выходе канала уменьшается, так как за время очень короткого импульса его величина не успевает достаточно нарастать, импульсы становятся малозаметными от пауз, и увеличивается вероятность появления ошибок из-за наличия шумов, т. е. падает надежность передачи.

Полоса пропускания частот, время передачи и мощность сигнала по отношению к шумам — вот те три величины, которые определяют пропускную способность канала.

Селекция

Если нам нужно передать две телеграммы или провести два телефонных разговора, мы можем сделать это либо по очереди (за счет увеличения времени), либо одновременно по двум линиям связи, либо, наконец, одновременно по одной и той же линии, но на двух различных несущих частотах. В обоих последних случаях необходимо увеличение суммарной полосы частот, используемой при передаче. Каждый разговор в последнем случае используется как модулирующая функция для одного из двух колебаний с различными несущими частотами. Затем оба модулированных колебания передаются по общему проводу (или по радио) одновременно. На приемном конце установлены фильтры — резонансные системы, которые про-

пускают разные полосы частот. На выходе фильтра первой несущей вторая несущая отсутствует, так как она лежит вне полосы пропускания фильтра. В результате из выходного напряжения может быть получено напряжение, соответствующее модулирующей функции, т. е. принят только первый разговор. То же происходит во втором фильтре со второй несущей и со вторым разговором. На входном конце канала два разговора сводятся в общий канал, но не смешиваются в нем, на выходном — они разделяются с ис-

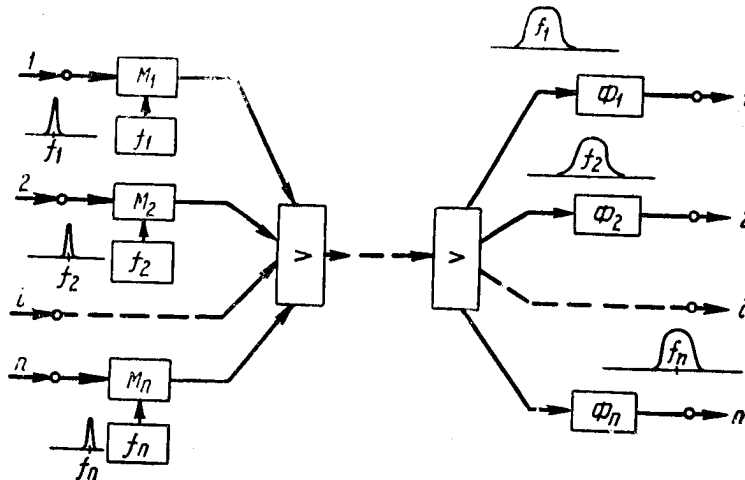


Рис. 5.12. Схема многоканальной связи с частотной селекцией.
 M_1, \dots, M_n — модуляторы, Φ_1, \dots, Φ_n — фильтры.

пользованием разницы несущих частот. Такое разделение сигналов по частоте называется частотной селекцией.

Таким способом можно по одному каналу одновременно передавать и более двух разговоров. Если на передачу каждого разговора требуется полоса частот 3—5 килогерц, то для общего канала эта полоса будет равна сумме всех полос или несколько больше.

При импульсной передаче можно «вписывать» одну последовательность коротких импульсов в паузы другой последовательности и таким образом совмещать две и более передачи одновременно в общем проводе или радиоканале. Затем на приемном конце пропускать каждый первый импульс в первый выходной канал, второй — во второй и т. д., каждый раз начиная с некоторого «нулевого»

синхронизирующего импульса, который отличается от всех других каким-либо признаком, например амплитудой. Такой способ разделения сигналов, идущих по общему каналу, называется селекцией по времени или временной селекцией.

Системы связи, позволяющие передавать по одному проводу несколько различных сообщений одновременно, называются многоканальными. Говоря о многоканальных системах, нельзя не упомянуть и о системах, которые яв-

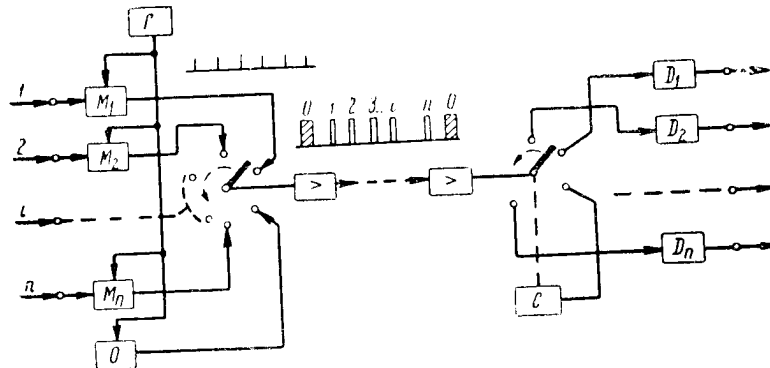


Рис. 5.13. Схема многоканальной связи с временной селекцией. Г — генератор, M_1, \dots, M_n — модуляторы, О — опорный сигнал, С — селектор, D_1, \dots, D_n — детекторы.

ляются в буквальном смысле многоканальными, т. е. состоящими из многих отдельных проводов — каналов связи. С такими системами мы встречаемся, например, в электронных счетных машинах параллельного типа, где каждый разряд числа передается по отдельной шине. С помощью одного из видов селекции можно осуществить переход от многопроводной системы связи к однопроводной многоканальной и обратно.

Во всяком реальном канале связи происходит рассеяние энергии сигнала и поэтому энергия (амплитуда) сигнала на выходе канала должна быть меньше, чем на входе. Для того чтобы компенсировать потери энергии в канале, часто используют усиление сигнала по пути его следования с помощью усилителей, которые черпают энергию из собственных источников. Расход этой энергии никак не окупается энергетически, т. е. эта энергия почти целиком теряется в цепях канала, однако в результате на выходе

канала сигнал оказывается достаточно большим по величине по сравнению с шумами. Иногда общее усиление в канале достигает очень больших величин.

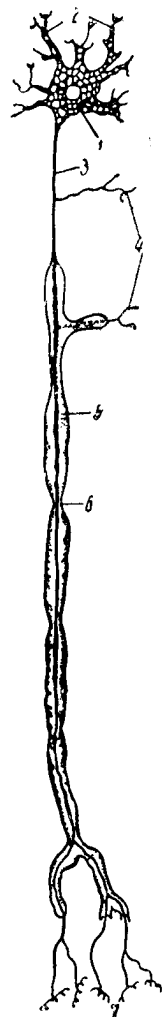
Пример канала связи

Рассмотрим подробно работу одного вида связи, чрезвычайно важного для существования живых организмов — передачу возбуждения или раздражения по нерву. Нерв, проводящий раздражение, представляет собой типичный канал связи и может сравниваться по своим характеристикам с техническими каналами. Эта далеко идущая аналогия так же несомненна, как несомненна, например, аналогия костей и суставов скелета позвоночных животных с рычагами механизмов.

Нервная система животного (в том числе и человека) состоит из очень большого числа живых клеток особого вида (нейронов), связанных в определенном порядке друг с другом и с органами тела животного. В нервной системе существует много нейронов различных типов, но всем им присущи общие черты, несмотря на различия в деталях строения. Каждый нейрон состоит из клеточного тела размером около 0,1 мм и отходящих от тела отростков. Большинство этих отростков (дендритов) ветвится, образуя как бы дерево длиной несколько миллиметров. Кроме того, от клеточного тела отходит один осевой отросток (аксон) большой длины (от нескольких сантиметров до метров). От аксона отходит небольшое число боковых волокон — коллатералей; на конце аксон делится на несколько ветвей.

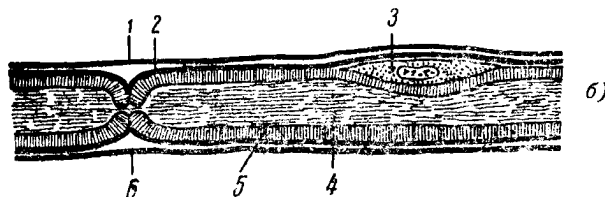
Таким образом, нервная клетка представляет собой как бы длинный провод, оканчивающийся на одном конце телом клетки. Эти провода — нервные волокна — обычно объединены в пучки, называемые нервами. Нерв, таким образом, больше напоминает многожильный телефонный кабель, чем одиночный провод. Как и в многожильном кабеле, в нерве каждое волокно живет своей собственной жизнью и выполняет свое собственное назначение.

Основным назначением нервного волокна и нервной клетки в теле животного является передача сообщения (возбуждения, раздражения) от одного органа к другому. По каналам связи, для которых нервные волокна служат проводами, передаются сообщения, которые оцениваются человеком как субъективные ощущения: сообщения о механическом давлении (осязание), об изменениях температуры (ощущение тепла, холода), о биологически опасных



a)

Рис. 5.14. Схема строения нейрона.
а) общее строение: 1 — тело клетки, 2 — дендриты, 3 — аксон, 4 — коллатерали, 5 — оболочка аксона, 6 — перехват Ранвье, 7 — концевое разветвление; б) разрез аксона: 1 — перехват Ранвье, 2 — нервилемма, 3 — ядро шванновской клетки, 4 — нейроплазма, 5 — миелиновая оболочка, 6 — внешняя оболочка.



б)

воздействиях (боль), об акустических, световых, химических явлениях, происходящих на границе живого организма и внешней среды. По тем же каналам передается команда о включении в работу мышечных волокон или клеток желёз. В случае ощущения сигнал идет от рецепторов («органов чувств»), непосредственно воспринимающих внешние воздействия, к нервным узлам центральной нервной системы. Эти нервные связи называются восходящими, центростремительными или афферентными. В случае действия сигнал идет от нервных центров к периферии и соответствующие нервы называются нисходящими, центробежными или эфферентными.

К афферентным относятся также и нервы так называемой внутренней чувствительности, передающие информацию о положении членов тела, т. е. о состоянии сокращения или расслабления отдельных мышц и натяжении сухожилий (так называемые кинэстетические ощущения). Кроме того, большое число нервов обслуживает связи между внутренними органами, необходимые для их совместной целесообразной деятельности. Работа этой группы нервов (так называемой вегетативной или автономной нервной системы) не дает субъективных ощущений непосредственно, и мы «не замечаем» ее работы, во всяком случае до тех пор, пока она работает нормально.

Итак, нервная система представляет собой сложную, разветвленную систему связи, которая обеспечивает обмен сигналами между органами тела животного и согласованную их работу.

Как и в технических каналах связи, в нервной системе все сообщения передаются с помощью одного и того же процесса. Передача раздражения по одному нерву качественно не отличается от передачи по любому другому. Особенностью нервной системы как системы связи является то, что по каждому нервному волокну передаются сообщения, как правило, лишь в одном направлении и лишь об одном типе внешних или внутренних событий. Так, например, нервное волокно, соединяющее чувствительное болевое окончание, расположенное в коже ладони, со спинным мозгом, будет работать только при сильных механических давлениях на этот участок кожи (укол) и не будет реагировать на слабые раздражения ни теплом, ни холодом, ни прикосновением. В этом отношении нервные каналы связи напоминают скорее устройства аварийного предупреждения типа пожарной тревоги, чем телефонную или телеграфную связь. Правда, нерв позволяет передавать сигналы о раз-

ных степенях воздействия и этим отличается от пожарной сигнализации. Таким образом, каждый нерв работает только в одной «службе связи» со строгой специализацией: в «болевой», «слуховой», «вкусовой», «моторной» и т. д.

Провод первого канала связи — нервное волокно или аксон — представляет собой ниточку клеточного вещества — протоплазмы, покрытую оболочками. Толщина аксона для разных нервных клеток различна. Самые толстые из них имеют диаметр от 0,02 до 0,001 миллиметра. Клеточное вещество аксона, как и клеточного тела, покрыто оболочкой или мембраной, отделяющей его от окружающей среды. Поверх мембраны аксон одет, как чехлом, оболочкой особого жирового вещества — миэлина. Толщина миэлиновой оболочки неодинакова для различных волокон. У волокон вегетативной нервной системы миэлиновая оболочка очень тонка; иногда ее даже не удается обнаружить. Самой толстой оболочкой обладают нервные волокна, проводящие раздражение от чувствительных окончаний к мозгу и от мозга к мышцам. Поверх миэлиновой оболочки аксон одет еще одной оболочкой — неврилеммой. Миэлинизированные волокна отличаются серебристо-белой окраской от клеточных тел, имеющих серый цвет.

Со времен Гальвани известно, что деятельность нерва связана с электрическими явлениями. Эта связь выражается в том, что, во-первых, приложение к нерву электрического потенциала вызывает его раздражение и, во-вторых, в том, что раздражение нерва любым способом вызывает в нем появление изменений электрического потенциала. Однако нервное раздражение совсем не то же самое, что электрический ток в проводе.

С помощью современных электроизмерительных приборов, позволяющих отмечать и записывать очень слабые и быстро протекающие изменения электрических напряжений и токов, снимая потенциалы с поверхности нерва, можно установить, что изменения этих потенциалов представляют собой короткие толчки или импульсы, длящиеся около 0,001 секунды. Эти импульсы распространяются вдоль нерва и, достигая его конца, вызывают, например, сокращение мышечных волокон. При этом передача раздражения мышечным волокнам и распространение электрических импульсов по нерву всегда присутствуют совместно. Более того, скорость распространения электрических импульсов и скорость передачи физиологического раздражения совпадают, что позволяет считать электрические импульсы не побочным явлением, а составной частью процесса передачи

раздражения. Скорость передачи раздражения по нервному волокну тем больше, чем больше диаметр волокна и чем толще его миелиновая оболочка. Эта скорость вообще невелика; для разных волокон она колеблется в пределах от 2 до 160 метров в секунду, что несравненно меньше скорости распространения электрического поля (и тока) вдоль провода (10^8 метров в секунду).

Передача раздражения по нервному волокну есть электрохимический процесс, протекающий за счет энергии, накопленной в самом волокне. Начальное возбуждение, наносимое на окончание нервного волокна, вызывает только начало, «поджиг», процесса распространения раздражения, который затем уже продолжается самостоятельно, подобно искре, бегущей по бикфордову шнуру. Каждая такая «искра», импульс, пробегает самостоятельно.

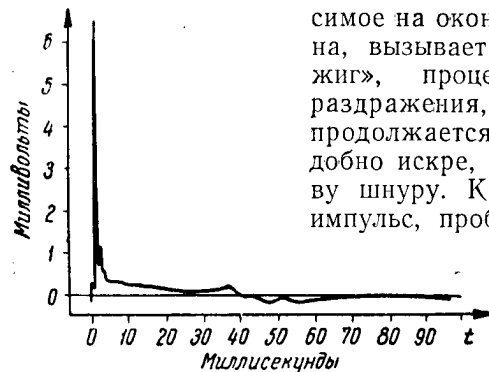


Рис. 5.15. Одиночный электрический импульс нервного возбуждения.

и независимо от величины энергии приложенного внешнего воздействия, если только это воздействие превышает минимальное, пороговое, значение. Энергия, израсходованная нервом на проведение импульса, возмещается впоследствии процессами питания нерва.

Эта ситуация типична для сигнальных устройств или систем связи, которые почти всегда питаются энергией от внутренних источников, не связанных с поступлением на вход сигнала.

После каждого прохождения импульса по нервному волокну последнее в течение некоторого времени должно «отдохнуть» и подготовиться к проведению следующего импульса, который пройдет или не пройдет в зависимости от наличия сигнала на входе. Эта пауза, или рефрактерный период, длится от 0,02 до 0,1 секунды. Таким образом, нервное волокно может находиться лишь в одном из двух возможных состояний: состоянии полного покоя или полного возбуждения. Все сообщения, передаваемые по нерву, ко-

дируются с помощью двоичного алфавита, состоящего из двух символов: «покой», «возбуждение».

Однако, несмотря на то, что нервное волокно работает в двоичном алфавите, по нерву могут передаваться сообщения о различной степени возбуждения, причем передаваемая величина может изменяться непрерывно. Различные степени возбуждения, передаваемые по нервному волокну, изменяют частоту следования импульсов. Таким образом, в нерве мы имеем дело с частотно-импульсной модуляцией, получившей в последнее время широкое применение в технике связи. С увеличением степени раздражения увеличивается и число волокон нерва, передающих раздражение.

Работа нервного волокна

Распространение импульса возбуждения по нервному волокну протекает примерно следующим образом. Внутренняя часть волокна, как мы уже говорили, отделена оболочкой от наружной. Эта оболочка в нормальном состоянии покоя является электрическим изолятором и находится в состоянии электрической поляризации. Она разделяет области высокой концентрации ионов отрицательного знака (внутри оболочки) и области положительных ионов (снаружи оболочки). Таким образом, нервное волокно в состоянии покоя представляет собой длинный и тонкий заряженный цилиндрический конденсатор. Емкость перегородки равна приблизительно одной микрофараде на квадратный сантиметр ее поверхности.

Если в любой точке оболочки каким-либо путем нарушить ее изоляционные свойства, электрически «повредить» оболочку, возникнет местный электрический ионный ток («пробой» конденсатора). Ток замыкается через соседние участки оболочки и тем самым нарушает их равновесное состояние, изоляционные свойства и поляризацию. В результате участок проводимости оболочки перемещается вдоль волокна в обе стороны от места первоначального нарушения. Вместе с тем перемещается и место протекания ионного тока и вся картина распределения потенциалов, с ним связанная. Это перемещение распределения потенциала и воспринимается измерительным прибором в точке, где наложен электрод, как проходящий импульс напряжения. Амплитуда импульса напряжения равна разности потенциалов между внутренней частью (аксоплазмой) и внешней оболочкой нервного волокна и составляет

30—60 милливольт. Для данного волокна величина импульса остается всегда постоянной, если только нерв нормально питается и не поврежден.

Процесс восстановления активности нервного волокна еще не полностью изучен. Чтобы избежать излишних подробностей, мы не будем обсуждать здесь этот вопрос.

Итак, одиночное нервное волокно представляет собой канал связи для передачи сообщений об одной единственной величине. На входном конце канала находится чувствительный элемент (рецептор), который воспринимает внешнее событие и кодирует сообщение в алфавите канала в виде периодического импульсного возбуждения конца нервного волокна, примыкающего к рецептору. Увеличение раздражения соответствует возрастанию частоты периодических возбуждений. Каждый отдельный акт возбуждения — нервный импульс — распространяется по волокну и достигает выходного конца, в случае афферентных нервов — тела нервной клетки. Импульс может пройти сквозь клетку дальше, на другие ее отростки, и перейти через соединения (синапсы), существующие между телами клеток, дендритами и аксонами клеток, на одно из волокон следующей по цепи клетки или нескольких клеток сразу. Распространяясь по волокнам и переходя с одной нервной клетки на другую, сигнал достигает места своего окончательного назначения — стоящего на выходе канала исполнительного элемента (мышечного волокна или клетки железы), который под воздействием пришедшего сигнала включает в работу. Происходит декодирование сигнала и его исполнение.

Передача сигнала раздражения с рецептора на волокно, с одной клетки на другую и с волокна на исполнительный элемент (эффектор) происходит обычно с участием промежуточных химических веществ (нейрогуморов), которые выделяются в синапсах и нервных окончаниях с приходом нервного импульса и воздействуют на последующее звено канала. Характер воздействия нейрогуморов может быть различным. Так, например, возбуждение некоторых нервов ведет к торможению работы органа (остановке сокращений изолированного сердца при раздражении блуждающего нерва), возбуждение других нервов — к усилению их работы (ускорение биений сердца при раздражении симпатических нервов).

Прохождение импульсов нервного возбуждения через тела клеток определяется состоянием этих клеток. Клетки могут пропускать или не пропускать импульсы в зависи-

мости от того состояния, в котором они находятся. Это состояние определяется питанием клетки, наличием химических стимуляторов в окружающей клетку среде и приходящими стимулирующими или тормозящими нервными раздражениями со стороны других клеток. Здесь мы подходим к сложному вопросу о коммутации нервных раздражений — их переключению с одной линии связи на другую. Сейчас мы не будем вдаваться в подробности и вернемся к этому вопросу позже.

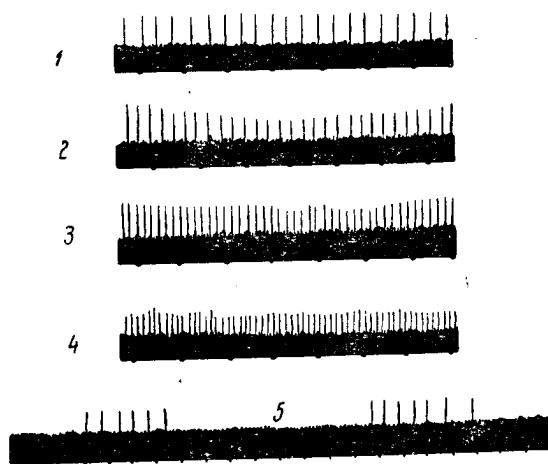


Рис. 5.16. Импульсы возбуждения чувствительного нервного волокна нервного узла сердца (каротидного синуса) при различных уровнях давления крови.

Давление повышается от 1 к 4; 5—ритмические группы импульсов в одиночном волокне двигательного нерва наружных межреберных мышц кошки (группы следуют в ритме дыхания). Отметки времени—0,2 сек (по Бронку).

На всем протяжении нервных связей распространение возбуждения происходит за счет энергии нервной системы, а не за счет энергии поступающего извне сигнала раздражения. Нервная система имеет, таким образом, внутреннее усиление сигнала, обеспечивающее его передачу без затухания. Нервный канал связи обладает ограниченной полосой частот. Изменения входного сигнала более быстрые, чем, скажем, самый короткий рефрактерный период нервного волокна не могут быть переданы по нерву. «Высоко-частотный» сигнал при передаче подвергается искажению.

Можно подсчитать пропускную способность нервного волокна тем же методом, который применяется для техни-

ческих каналов связи. При этом, разумеется, необходимо учитывать тот факт, что передача по нерву является связью многопроводной. Так, например, зрительное ощущение передается от сетчатки глаза к зрительным центрам мозга по зрительному нерву, содержащему несколько десятков тысяч волокон.

Можно также оценить пропускную способность не отдельных нервов, а всей их совокупности, обслуживающей данное «внешнее чувство» — слух, зрение, осязание, вкус и т. д. Затем можно сравнить отдельные каналы по их пропускной способности. В результате сравнения оказывается, что на первом месте стоит зрение, имеющее самую большую информационную пропускную способность, далее идет осязание и еще далее — слух. Именно поэтому протезирование (в информационном смысле) утраченного зрения легче осуществить с помощью осязания, чем с помощью слуха. Этот вывод может оказаться полезным для конструкторов машин, предназначенных для «чтения» печатного шрифта слепыми.

Все сказанное выше о нервных связях позволяет рассматривать передачу возбуждения по нерву, как типичный случай передачи сообщений по каналу связи. И этот канал связи может полноценно и полноправно рассматриваться в одном ряду с техническими каналами, тоже передающими сигнал от корреспондента адресату. Особенностью нервного канала связи является только своеобразие физиологических процессов, являющихся носителем информации, которое в сущности одно только и отличает этот канал связи от многих других.



ГЛАВА 6

ОБРАТНАЯ СВЯЗЬ. РЕГУЛИРОВАНИЕ

Предположим, что нам необходимо во время нашего отсутствия поддерживать постоянной температуру в комнате. Для этой цели мы хотим изготовить автомат, который регулировал бы комнатную температуру. Допустим, что для поддержания постоянной температуры воздух в комнате необходимо все время подогревать и мы можем регулировать нагрев, изменяя подачу топлива в печь или силу тока в электронагревателе.

Самое простое решение поставленной задачи — это установить такой регулятор (например, часовой механизм), который с течением времени менял бы нагрев в соответствии с ожидаемой потребностью в тепле. Обычно к утру становится холоднее, и подачу топлива в утренние часы следует увеличить. Такое «регулирование по программе» предполагает знание статистических сведений об изменении внешней температуры в течение суток. Регулятор, работающий по программе, не способен справиться со своей задачей, если внешние условия сильно отклонятся от своего среднего значения, например, если к утру произойдет неожиданное похолодание или потепление.

Для того чтобы улучшить работу регулятора, необходимо вводить в него информацию о внешней температуре и заставить его реагировать на ее изменения. Тогда случайное понижение внешней температуры вызовет не только охлаждение комнаты, но и через воздействие на регулятор — увеличение нагрева. Такая система называется системой «прямого регулирования». Ценой больших усилий при налаживании автомата можно подобрать его характе-

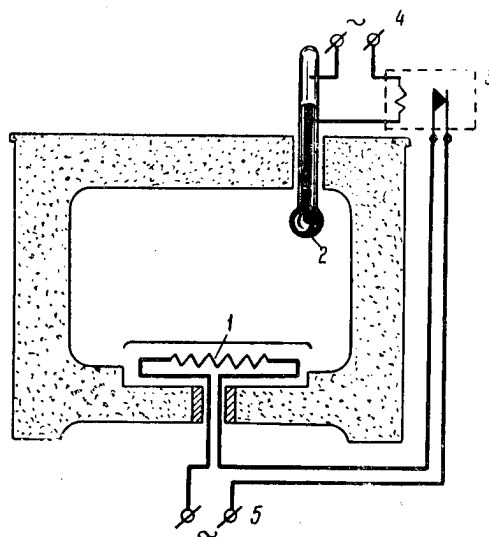


Рис. 6.1. Термостат.

Повышение температуры внутри шкафа вызывает реакцию чувствительного элемента — подъем ртути в трубке термометра 2 и замыкание электрического контакта, который включает питание 4, подаваемое на обмотку реле 3. Реле, сработав, замыкает контакты и включает питание 5 нагревателя 1, тем самым прекращая повышение температуры.

ристики так, чтобы степень нагрева соответствовала степени охлаждения воздуха при изменении внешней температуры. Однако открытая дверь или форточка сделают и эту систему неэффективной, ибо охлаждение комнаты через форточку увеличится, а нагрев, зависящий только от внешней температуры, — нет, и равновесие будет нарушено.

Радикальное решение задачи достигается системой с обратной связью, или «системой регулирования замкнутого цикла». Такая система отличается от предыдущих тем, что подача топлива определяется не заранее установленной программой и не внеш-

ними условиями, а разностью между действительным значением температуры в самой комнате и желаемым ее значением. Если температура в комнате равна заданной величине, топливо подается с некоторой нормальной скоростью. Если температура становится ниже (выше) заданного уровня, подача топлива увеличивается (уменьшается) пропорционально образовавшейся разности. Таким образом, регулятор не только воздействует на температуру комнаты, но и «ощущает» результат этого воздействия и автоматически изменяет воздействие так, чтобы результат как можно меньше отличался от желаемого.

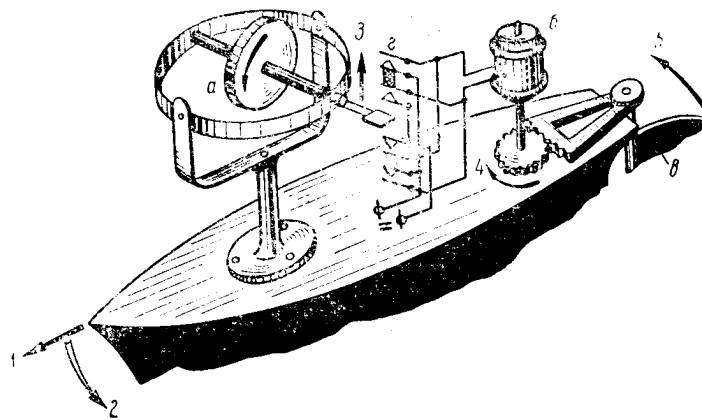


Рис. 6.2. Автоматический рулевой.

Изменение курса судна 1 в сторону 2 вызывает прецессию волчка а в направлении 3 и замыкание контактов 4, которое заставляет мотор 6 работать в направлении 5 и поворачивать руль 8 в направлении 5, выравнивая судно на прежнем курсе.

Теперь можно открывать форточки и двери и не бояться капризов погоды. Наш автомат «привязан» непосредственно к температуре воздуха в комнате и будет ее поддерживать автоматически, независимо от любых внешних влияний.

Другим типичным примером применения обратной связи является автоматический рулевой, или автопилот. Выходное устройство воздействует на руль корабля или руль самолета. Входное устройство работает совместно с компасом, отмечая отклонения фактического курса от заданного. При отклонениях от заданного курса входное устройство выдает сигнал, который после усиления вызывает поворот рулей и выравнивание курса

Системы с обратной связью чрезвычайно широко распространены. Можно перечислить многие и многие примеры их применения в технике. Более того, регулирование с обратной связью характерно для многих биологических процессов. Можно без преувеличения сказать, что обратная связь является некоторым универсальным общим принципом кибернетических систем.

Простая система регулирования замкнутого цикла состоит из нескольких узлов, соединенных в общую, харак-

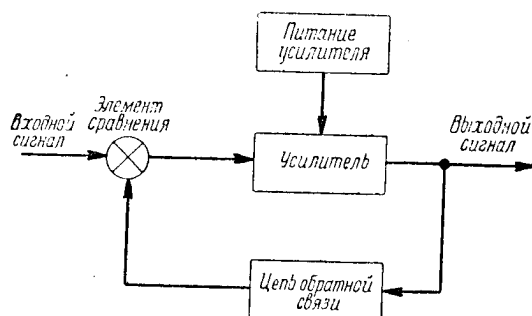


Рис. 6.3. Схема системы с обратной связью.

терную для всех таких систем схему. Эта схема приведена на рис. 6.3.

Система регулирования замкнутого цикла, как правило, включает в себя некоторое энергетическое устройство, приводимое в действие от внутреннего источника энергии и управляемое с выхода усилителя сигнала. Это выходное исполнительное устройство может, например, производить нагрев или охлаждение, перемещать тяжелые тела, производя механическую работу, открывать и закрывать клапаны, заслонки, краны, вращать рули и т. д. Результатом работы выходного устройства всегда является действие, подчас заключающее в себе очень большие количества энергии и направленное на некоторый внешний, по отношению к системе регулирования, объект.

Как мы уже сказали, управление выходным устройством, скажем печью в нашем примере, осуществляется сигналом с выхода усилителя. Сигнал только управляет выходным устройством, но не снабжает его энергией. Мощность входного сигнала, подаваемого на усилитель, может быть очень мала.

Существенной частью системы регулирования с обратной связью является сама цепь обратной связи. Действие, производимое выходным устройством, улавливается некоторым чувствительным элементом, в котором оно преобразуется в сигнал, и этот сигнал подается с выхода системы обратно на ее вход (отсюда термин «обратная связь»). На входе системы стоит элемент схемы, который служит для сравнения двух сигналов: сигнала обратной связи, несущего информацию о результате работы выходного устройства, и входного сигнала системы, поступающего извне и несущего информацию о желаемом результате действия выходного устройства. В результате сравнения получается разность обоих сигналов, которая и подается на вход усилителя. Если желаемый и действительный результаты совпадают, сигнал на входе усилителя равен нулю и выходное устройство бездействует.

Если же, вследствие любых воздействий, желаемый и действительный результаты не совпадают, разностный сигнал на входе усилителя не равен нулю, что заставляет работать выходное устройство до тех пор, пока действительный результат не станет равен желаемому.

Входной узел системы представляет собой вычитающее устройство. Из выходного сигнала системы, дающего сведения о желаемом результате, вычитается величина, определяемая сигналом обратной связи. Поэтому такой вид обратной связи называется «отрицательной» обратной связью. Существуют и системы с «положительной» обратной связью, которые мы рассмотрим несколько позже.

Система с обратной связью испытывает, очевидно, воздействия двоякого рода. Во-первых, влияние случайных изменений внешних условий; в случае термостата — это изменения условий теплообмена (изменения внешней температуры и теплопроводности стенок), в случае автопилота — это влияние ветра и т. д. Во-вторых, воздействие внешнего сигнала управления, который дает, так сказать, опорный уровень желаемого результата, исходя из которого система «отрабатывает» внешние воздействия. Пользуясь сигналом управления, мы можем установить терморегулятор на новое значение температуры, и автомат будет поддерживать это новое значение.

Входной сигнал можно непрерывно менять. В этом случае выходное устройство будет работать так, что на выходе его величина, характеризующая результат его действия, например, положение тяжелого груза, будет послушно следовать за входным сигналом, который может задавать-

ся, например, поворотом небольшой ручки или легкого диска или ничтожным световым потоком, падающим на фотозлемент.

Как бы при этом ни был слаб входной сигнал и как бы ни была велика мощность выходных устройств, выходное устройство будет не только послушно следовать за изменениями входного сигнала, но и стабилизировать выходную величину при случайном изменении внешних условий.

Естественно назвать систему с отрицательной обратной связью устройством, исполняющим сигнал, т. е. превращающим сигнал в действие значительного энергетического масштаба.

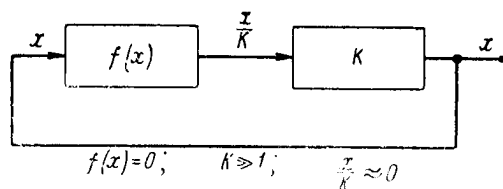


Рис. 6.4. Решение трансцендентного уравнения с помощью системы с обратной связью (K —коэффициент усиления).

Впрочем, не следует думать, что система с обратной связью непременно обладает большой мощностью выходного устройства. Существует много применений таких систем, где этого не требуется, хотя и в этих применениях имеется усилитель сигнала. К числу таких применений относится использование обратной связи в решающих устройствах непрерывного действия. Для решения сложного уравнения (например, трансцендентного) используются устройства, функционально преобразующие сигнал, изображающий независимую переменную. Решаемое уравнение приводится к виду $f(x)=0$; сначала в качестве аргумента x вводится произвольная величина. На выходе функциональных преобразователей получается некоторая величина, отличная от нуля $f(x)=y \neq 0$. Эта величина подается в качестве входного сигнала на цепь обратной связи, с выхода которой снимается величина аргумента x , вводимого в функциональное решающее устройство. Система изменяет входную и выходную величины до тех пор, пока не будет удовлетворено уравнение и на выходе функционального устройства не появится нулевое значение (точнее, значение, близкое к нулю). Решающее устройство с обратной связью как

бы ищет подходящее значение корня, решая задачу методом последовательных приближений.

Разнообразные приборы такого типа, и подчас очень сложного устройства, широко применяются в технике.

Системы с обратной связью по своим свойствам резко отличаются от систем прямого регулирования. Отличия эти можно проследить с различных точек зрения. Они характерны и имеют большое принципиальное значение.

В системе прямого регулирования, как мы уже видели выше, некоторый сигнал, подаваемый на вход системы, проходит усиление и преобразование в системе. В результате, на выходе системы имеет место некоторое энергетическое действие, определенно определяемое входным сигналом, например выделение тепла или поворот руля. В этом отношении система прямого регулирования, казалось бы, подобна системе с обратной связью. Однако это справедливо только при том условии, что все внешние влияния и внутреннее состояние системы остаются строго неизменными и не оказывают воздействия на результат работы системы. В практике такие условия чаще всего не выполняются, поэтому любое изменение внешних условий или изменение характеристик узлов самой системы влечет за собой изменение результата работы системы регулирования разомкнутого цикла (по сравнению с желаемым результатом). Так, например, порыв ветра сбивает корабль с курса или же уменьшение коэффициента усиления усилителя сигнала влечет за собой уменьшение выходного эффекта при неизменном сигнале на входе.

Замыкание петли обратной связи создает качественно новую систему, обладающую новыми свойствами, которых не имели ее составные части. Характеристики системы с обратной связью в широких пределах не зависят от характеристик ее узлов или эта зависимость оказывается очень малой. Так, изменение коэффициента усиления усилителя оказывается практически незаметным, если только коэффициент усиления остается еще достаточно большим. Кроме того, система с обратной связью работает, учитывая внешние влияния на результат работы и реагируя на них таким образом, чтобы устранить эти влияния. Система с обратной связью как бы «приспосабливается» к внешним условиям. Автоматический рудерный восстановит пужный курс, но какой бы причине ни произошло отклонение.

Различие систем прямого регулирования от систем регулирования с обратной связью хорошо знают любители

строить парусные и моторные модели судов. Пуская модель в открытом бассейне, они всегда сталкиваются с трудностями, когда пытаются обеспечить ход модели по заданному курсу.

Если это удастся сделать при равномерном ветре и отсутствии течения, то при порывистом ветре или волне, т. е. при наличии случайных влияний извне, модель сбивается с курса и далеко отклоняется от желаемой траектории. Никакая предварительная установка руля и снастей не обеспечивает стабилизации курса.

В этом случае задача решается двояким путем: либо вводится автомат стабилизации, либо управление по радио. Простой автомат передает воздействие при изменении силы ветра с паруса через натяжение шкота на румпель руля и доворачивает руль на некоторый угол, что компенсирует изменение курса при порыве ветра. Такое устройство, строго говоря, хотя и отрабатывает изменения ветра, все же не является системой с обратной связью, так как оно не ощущает и не отрабатывает изменений курса судна.

Более сложный автомат для механических моделей конструируется так же, как автоматический рулевой для настоящих судов. Изменение курса (по каким бы причинам оно ни происходило) относительно курса, заданного по компасу (или маяку), т. е. ошибка в курсе, вызывает в конечном итоге поворот руля, доворачивающего судно до нужного направления. Это — типичная система с обратной связью.

При использовании управления моделями по радио петля обратной связи замыкается через оператора, наблюдающего с берега за отклонениями модели от курса и воздействующего на рули с помощью управления по радио. Для полноты следует упомянуть, что на парусном судне (не модели) рулевой, наблюдающий отклонения от курса и исправляющий их поворотом руля, является живым звеном петли обратной связи, стабилизирующей судно на курсе.

При конструировании станков-автоматов различие в программном регулировании и регулировании с обратной связью приобретает уже серьезное технико-экономическое значение. Станок-автомат выполняет операцию по обработке сырья, превращая его в изделие. Операция может представлять собой изменение формы и размеров, обточку, шлифовку, окраску (печатание обоев, рисунка на ткани), химическую обработку, тепловую обработку и т. д. В каждом случае сырье, орудие, характер обработки и

результат различны. Общим для всех случаев является то, что производится некоторое воздействие на сырье, которое подвергает сырье обработке и превращает его в продукт.

Работа станка-автомата может регулироваться по программе. В этом случае подача суппорта и резца токарного станка, например, определяется формой шаблона или установкой предельного упора, задающего нужный размер детали, который должен быть получен в результате обработки; температура печи и время пребывания в печи

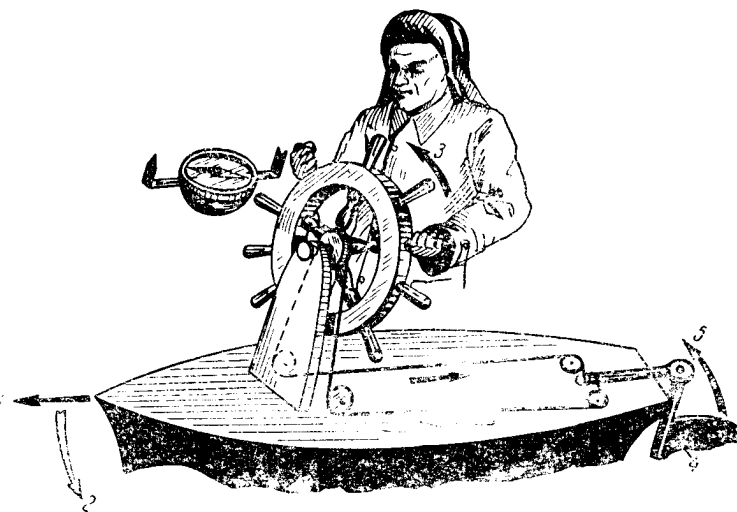


Рис. 6.5. Рулевой, управляющий курсом судна.

Изменение курса судна *1* в сторону *2* воспринимается зрением рулевого (по береговому ориентиром, звездам или компасу). Рулевой переключает штурвал в сторону *3*, что вызывает поворот руля *4* в сторону *5* и выравнивание судна на прежнем курсе.

изделий хлебопекарни устанавливается заранее и однажды; совпадение рисунков различных цветов определяется положением валиков печатного станка и точностью протяжки материала и т. д.

Точность результата зависит во всех этих случаях от постоянства структуры и режима работы станка-автомата. Достаточно измениться одному размеру (например, в случае износа) или сбиться установке — и результат обработки выйдет из допусков; станок начнет делать брак. За таким «автоматом» необходим неослабный надзор, проверка и регулировка.

Надзор за автоматом резко сокращается, если автомат

работает по системе регулировки с обратной связью. Так, например, при печатании цветных обзев на выходе напечатанной ленты можно поставить устройство с фотоэлементами и светофильтрами, которое будет фиксировать момент прохождения рисунка каждого цвета. В случае расхождения рисунков (брак) несовпадение будет отмечаться в виде появления тока «сигнала ошибки» в цепях фотоэлементов, который используется после усиления для регулировки относительного расположения валиков, печатающих рисунки разных цветов, с целью устранения расхождения.

Случайные проскальзывания или растягивание бумажной ленты, износ механизмов и другие случайные и систематические влияния будут автоматически компенсироваться следящим устройством. Станок сам контролирует качество своей работы и исправляет ошибки.

Подобные же системы с наблюдением результата работы и соответственной автоматической регулировкой установки могут быть применены и применяются во многих производствах. И в этих случаях обратная связь, осуществляемая техническими устройствами, заменяет труд человека, наблюдающего за работой станка-автомата и регулирующего его работу.

Работа систем регулирования

Работа следящих систем или систем с обратной связью поддается точному описанию и расчету. Математической формулировкой таких описаний занимается теория автоматического регулирования. Подробная теория сформулирована, в частности, для широкого класса так называемых «линейных» систем, т. е. систем, описываемых линейными дифференциальными уравнениями. Математический аппарат теории линейных следящих систем является общим для теории автоматического регулирования и теории линейных цепей в электро- и радиотехнике, а также теории линейных систем в механике упругих тел и акустике.

Мы не будем уделять много внимания математическому описанию и исследованию свойств систем. Нам больше будет интересовать качественная сторона явлений. Поэтому мы рассмотрим характерные типы реакций системы на некоторые типичные входные сигналы — единичный скачок и единичный импульс. Оба эти сигнала являются математической идеализацией, и в практике мы можем создавать их только приближенно, но они удобны для описания систем и ими поэтому широко пользуются.

На рис. 6.6 представлены типичные реакции на единичный скачок (функцию Хевисайда):

$$u_{вх} = 0 \text{ для } t < 0; \quad u_{вх} = 1 \text{ для } t \geq 0;$$

и на единичный импульс (функцию Дирака):

$$u_{вх} = 0 \text{ для } t \neq 0; \quad u_{вх} = \infty \text{ для } t = 0;$$

$$\int_{-\infty}^{+\infty} u_{вх}(t) dt = 1.$$

Реакция следует не мгновенно, а с некоторым запазданием и достигает установившегося значения, лишь спустя значительное (теоретически бесконечно большое) время.

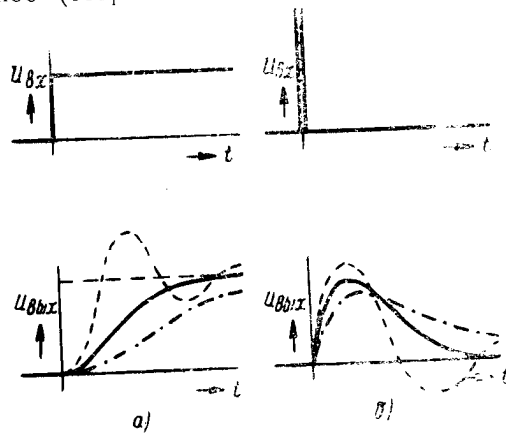


Рис. 6.6. Стандартные входные сигналы и их обработка системами с обратной связью, обладающими различными характеристиками.
а) единичный скачок (функция Хевисайда), *б)* единичный импульс (функция Дирака).

Отрезок времени, за который система достигает некоторой доли установившегося значения (63%), называется постоянной времени системы. Чем меньше постоянная времени, тем быстрее система реагирует. Система может достигать установившегося состояния, либо подходить к установившемуся значению с одной стороны (апериодическая реакция), либо колеблется вокруг установившегося значения. Колебания должны затухать со временем. Наличие колебательного процесса в реакции системы и скорость

затухания этих колебаний определяются структурой системы.

Реакции следящей системы на входные сигналы подобны реакциям электрических (механических, акустических) фильтров с той лишь разницей, что следящая система обладает внутренним источником энергии, и потому, в принципе, может расходовать эту энергию на поддержание самостоятельных колебаний. Вопрос о возможности возникновения самостоятельных колебаний или вопрос устойчивости системы, не возникавший для пассивных

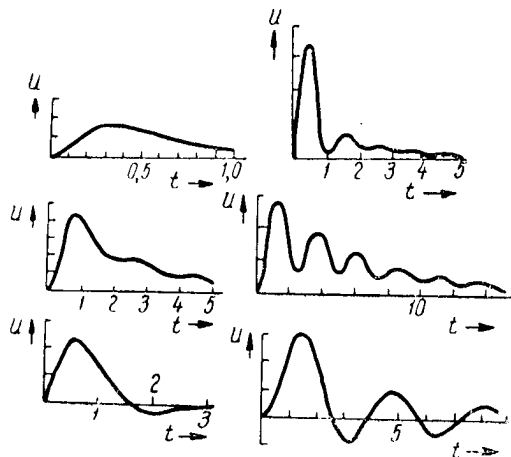


Рис. 6.7. Примеры реакций технических систем с обратной связью на единичный импульс (t — время в секундах).

фильтров, становится одним из главных вопросов при изучении активных следящих систем.

Можно непрерывным изменением параметров системы изменять реакцию системы от медленной апериодической до колебательной, которая в пределе может привести к неустойчивости системы.

Для определения степени устойчивости системы разработаны, по крайней мере для достаточно простых случаев, точные критерии.

Явления, приводящие к неустойчивости и возникновению колебаний от внешнего толчка, в простейшем описании могут быть изложены следующим образом.

Сигнал, приложенный на вход системы, проходит летью обратной связи за некоторый отрезок времени (с некоторой

задержкой) и с некоторым искажением формы. Если амплитуда сигнала обратной связи не мала, то входной сигнал усилителя может изменить за счет сигнала обратной связи свою величину и знак. Сигнал обратной связи, приложенный ко входу системы, снова усиливается и задерживается и снова поступает на вход системы с обратным знаком. На входе усилителя сигнал возрастает по абсолютной величине и периодически меняет знак. Система «раскачивается», т. е. возникают колебания с нарастающей амплитудой.

Если в цепи обратной связи высокочастотные компоненты сигнала имеют коэффициент усиления больше единицы и сдвиг фазы больше 90° , то обратная связь из отрицательной превращается в положительную, так как на входе системы происходит вычитание противофазных компонент, т. е. сложение синфазных, что и дает постепенное нарастание амплитуды.

Рассмотрим на примере автоматического управления рулем судна, что происходит в инерционных системах, т. е. системах, обладающих большой задержкой реакции. Случайное отклонение от курса вызывает реакцию системы и поворот руля. Это ведет к постепенному довороту судна до направления заданного курса; входной сигнал исчезает, однако вследствие инерции системы сигнал на выходе, т. е. поворот руля, еще некоторое время сохраняется. Судно продолжает поворачивать и отклоняться от курса в противоположную сторону. Оно успеет отклониться тем больше, чем дольше не сработает (запоздаст) система управления. Затем отклонение вызовет реакцию, руль будет переложено в противоположном направлении и все повторится в том же порядке, но в противоположную сторону. Судно будет «рыскать» на курсе и может совсем потерять управление. Дело можно исправить различными способами, например введя регулирование с учетом производной входного сигнала, т. е. так, чтобы поворот руля вызывался не только величиной отклонения от курса, но и скоростью изменения этого отклонения, а главное, зависел бы от знака этой скорости. Тогда нарастание угла ошибки вызовет более сильный поворот руля, а уменьшение ее — уменьшит поворот руля или даже заставит руль слегка перейти за нулевое положение, чтобы притормозить вращение судна по инерции. Вводя это изменение, мы меняем частотную характеристику системы так, чтобы фаза сигнала обратной связи для всех усиливаемых частот обеспечивала отрицательную обратную связь.

Системы с положительной обратной связью могут использоваться и используются как генераторы незатухающих колебаний. Эти системы более сложны в описании, чем системы с отрицательной обратной связью, так как они нелинейны. Положительная обратная связь в линейной системе вызывает неограниченное нарастание амплитуды колебаний, а так как всякая реальная система линейна лишь в некоторых пределах, то эти пределы и задают наибольшую, предельную амплитуду нарастающих колебаний. Математический аппарат линейных систем способен лишь указать на наличие положительной обратной связи и на возможность возникновения колебаний. Для описания подробностей развития и установления процесса требуется другой, более сильный математический аппарат.

Технические следящие системы встречаются в большом разнообразии типов и воплощений. По применению их можно разделить на следующие классы: усилители мощности, системы дистанционного управления, преобразователи, измерительные устройства, счетно-решающие устройства. По характеру физических процессов, которые используются в сигнальной и в силовой частях системы, они делятся на электромашинные (сервосистемы) постоянного и переменного тока, электронные (вакуумные и тиратронные), механические, гидравлические, пневматические и т. д. По структуре систем их можно разделить на системы непрерывного, прерывистого (импульсного) и релейного действия. В системах непрерывного действия выходной эффект пропорционален «ошибке», выделенной во входном устройстве, ее производным и интегралам. В системах импульсного действия выходной эффект постоянен в течение некоторого периода времени, и его изменение происходит только в момент короткого импульса регулирования. В релейных системах выходной эффект достигает максимального значения всякий раз, как «ошибка», выделенная во входном устройстве, превышает некоторую минимальную, пороговую величину. Все перечисленные типы следящих систем широко применяются на практике.

Обратная связь в живых организмах

Существование систем регулирования с обратной связью далеко не ограничивается техникой. Гораздо более широко, чем в технике, такие системы представлены в биологии, в живых организмах.

Любой, даже самый простой живой организм (живая клетка), представляя собой нечто единое и цельное, всегда заключает в себе различные составные части, отличающиеся структурно и функционально. Эти составные части взаимодействуют друг с другом и с внешней средой сложным образом, в результате чего некоторые жизненно важные для организма величины поддерживаются строго постоянными, несмотря на изменения, происходящие во внешней среде. Процесс биологической приспособляемости организмов к изменениям внешней среды весьма многообразен и сложен. Мы сумеем упомянуть лишь некоторые процессы, участвующие в этом взаимодействии.

Переходя от технических устройств к живым организмам, нельзя не отметить целого ряда особенностей предмета изучения, которые должны наложить отпечаток на подход к вопросу и на методы исследования.

Наблюдая поведение, реакции живых организмов, в особенности высших животных, на внешние раздражения и сравнивая эти реакции с работой машин, даже самых сложных, мы невольно отмечаем огромную разницу, заметную уже при самом поверхностном наблюдении. Разница эта настолько велика, что часто мы вовсе не решаемся проводить никаких аналогий или же, если и делаем это, то сопровождаем многочисленными оговорками и «извинениями».

Реакция простой машины на любое внешнее воздействие всегда однозначна. Зная устройство машины, зная внешнее воздействие, мы всегда заранее можем предсказать ее реакцию. Даже если машина среагирует неожиданно, то и тогда мы уверены, что это произошло из-за ее поломки, которая, собственно, и оказалась неожиданной и однозначно определила неожиданную реакцию.

Иное дело — живые существа. С ними никогда нельзя, казалось бы, быть уверенным, как именно они поступят в следующий момент. Любое неучтенное воздействие со стороны меняет их реакцию, их поведение. Чтобы добиться постоянства условий эксперимента, биологи отгораживают лаборатории звуконепроходимыми стенками, возводят «башни молчания», но даже стены не ограждают от случайного воспоминания, уловить влияние которого на поведение практически невозможно. И чем выше стоит животное на лестнице развития, тем сложнее и многообразнее, неожиданнее и, казалось бы, «беспричиннее» с виду его реакция. Даже два близнеца, выросших вместе, по разному ответят на одинаковые события.

Как же можно и можно ли подходить с одной меркой к животным и машинам? Не пытаемся ли мы сравнивать несравнимое и не является ли мир живых существ качественно новым и несравнимым с миром неживым?

Несомненно, живые существа составляют особый мир. Несомненно, что их существование качественно отлично от неживой природы. И тем не менее эти качественные отличия не могут заслонить несомненных черт сходства между работой машины и поведением живых существ.

Если мы будем помнить о различиях физической природы процессов и если мы всегда будем иметь в виду, что мы говорим о сходстве общего характера, касающемся передачи информации с помощью сигналов в пределах организованной системы, то мы, во-первых, избежим грубой ошибки, заключающейся в ложном утверждении, что «животное подобно техническому автомату», и, во-вторых, сумеем внести много нового в наше понимание процессов жизни.

Не все реакции живых организмов «неожиданны» или «произвольны» даже с первого взгляда. Произвольных в строгом смысле слова, т. е. беспричинных реакций, вообще не существует, есть только очень сложные реакции, вызываемые незаметными и неизвестными нам причинами, в частности прошлой историей существования живого организма, т. е. реакции, имеющие для нас характер случайности. Но многие жизненные проявления имеют вполне устойчивый, закономерный характер, напоминающий строго закономерные реакции машин-автоматов и получившие за это названия «автоматизмов». Вот об этих автоматизмах мы, в первую очередь, и поведем разговор.

Для нормального существования живого организма необходимо строгое поддержание постоянства некоторых физико-химических величин. Регулирование этих величин производится автоматически и протекает помимо ощущений, сознания и воли. Человек обычно не может прямым волевым актом повысить или понизить температуру своего тела, кровяное давление, содержание сахара в крови и т. д. Сознание и воля могут оказывать влияние на эти величины, но не так непосредственно, как, скажем, на работу скелетных мышц.

Функции организма, не подчиненные непосредственно центральной нервной системе и обеспечивающие постоянство внутренней среды организма, так называемые вегетативные функции, протекают под управлением сложного механизма автономной (висцеральной, или вегетативной) нервной системы.

Интересным и очень важным свойством этих процессов является их постоянство и закономерность. Для данного нормального индивидуума любой из вегетативных процессов при повторных экспериментах будет давать одну и ту же зависимость от времени. Те же эксперименты для другого индивидуума дадут другую, но тоже постоянную зависимость. Отклонение зависимости от обычной для данного индивидуума всегда свидетельствует об отклонениях функций от нормы и может быть использовано для диагностики заболевания. Именно поэтому врачи часто больше внимания уделяют изучению объективных лабораторных анализов, чем жалоб пациента.

Автоматические реакции по поддержанию постоянства внутренней среды организма — гомеостазис — осуществляются в связи с энергетическими процессами усвоения питательных веществ и их использования в организме. Получение питательных веществ, воды, кислорода, химическое расщепление и синтез новых веществ, доставка нужных продуктов к внутренним органам, использование запасов энергии и физической энтропии, заключенных в этих продуктах и получаемых организмом извне для нужд организма, удаление продуктов распада — вот основные энергетические процессы, протекающие в различных органах и тканях.

Все эти процессы протекают одновременно в различных органах и должны быть точно согласованы друг с другом во времени и пространстве. В любую минуту включение в работу той или иной группы мышц или других «потребителей» энергетической системы должно быть удовлетворено соответствующим «обслуживанием» во всех отношениях. Эта внутренняя «интеграция функций» различных органов не может иметь места без обмена информацией между сотрудничающими органами, без каналов связи и автоматов регулирования. Можно предположить, что если бы все вегетативные регулировки человеческого тела зависели от внимания человека, то человеку не только некогда было бы подумать ни о чем другом, кроме этих внутренних регулировок, но, попросту, не хватило бы внимания и времени обо всех о них подумать во-время.

Итак, кроме взаимодействий энергетических, связанных со снабжением тканей питательными веществами и кислородом, необходима сигнальная связь между органами и взаимное согласование их работы. Такая связь, как мы уже знаем, осуществляется, в первую очередь, по нервным волокнам вегетативной нервной системы. Кроме того,

информация о химических реакциях переносится химическими же веществами — гормонами, которые разносятся током крови и оказывают воздействие на соответствующие органы, даже в ничтожно малой концентрации. Оба типа каналов связи, нервный и химический, тесно взаимодействуют: под действием нервных раздражений возникают химические вещества (нейро-гуморы), воздействие гормонов на нервные узлы вызывает передачу раздражений по нервам. Управление функциями с помощью этих двух типов связи называется нейро-гуморальной регуляцией.

Работа мышечных волокон сопровождается усиленным потреблением кислорода и выделением углекислого газа. Снабжение клеток происходит через посредство крови. Если бы при работе мышцы ток крови не усиливался, мышца не могла бы долго усиленно работать. Однако это возможно, так как при включении в работу мышцы одновременно увеличивается приток крови к работающей мышце. Как это происходит? Выделяемый при работе мышечной клетки в кровь углекислый газ является переносчиком информации о работе мышцы. Повышение концентрации CO_2 в крови служит сигналом, на который сосуды прилежащие к работающей мышце, реагируют расширением. В результате, ток крови усиливается и обмен веществ обеспечивает возросшую потребность. Местный регулирующий механизм представляет собой автомат с обратной связью, причем внешним сигналом является иннервация мышцы, а его исполнением — увеличение притока крови. Передача сигнала происходит с помощью физико-химического процесса изменения содержания CO_2 в крови.

Если в работу включаются большие группы мышц, регуляция такого типа окажется недостаточной в масштабе всего организма, вследствие понижения кровяного давления. Должен быть увеличен доступ кислорода извне, т. е. усилены дыхание и кровообращение. Это действительно происходит под действием регулирующих механизмов более крупного масштаба. Увеличение концентрации CO_2 в крови воздействует, во-первых, на хеморецепторы, т. е. чувствительные к изменениям химизма нервные окончания вегетативной нервной системы, расположенные в аорте и сонных артериях («каротидные синусы»), и, во-вторых, на группу нервных клеток, расположенных в продолговатом мозге. Первое воздействие ведет через нервные раздражения, передаваемые сердцу, к учащению его сокращений и усилению тока крови, второе — к усилению дыхания (и

появлению субъективного ощущения «удушья»). И здесь, уже в больших масштабах, возникшая вследствие включения в работу мышц потребность в кислороде вызывает передачу сигналов по замкнутой петле обратной связи, что приводит к восстановлению равновесия между потребностью и ее удовлетворением.

Содержание сахара (глюкозы) в крови регулируется многими механизмами. Сахар обратимо усваивается печенью, где он превращается в гликоген, сахар усваивается тканями и т. д. Взаимодействие всех процессов этого рода организуется с помощью автоматических регулировок типа регуляторов с обратной связью. Так, увеличение содержания сахара в крови (гипергликемия) воздействует на гипоталамическую область мозга и далее в виде нервного возбуждения передается поджелудочной железе, где под действием нервных сигналов в островковой ткани выделяется в кровь гормон — инсулин, что приводит к увеличению усвоения сахара тканями тела, частичному превращению сахара в гликоген и, в результате, к устранению избытка сахара. Увеличение концентрации глюкозы воздействует и непосредственно на поджелудочную железу, стимулируя секрецию инсулина, но в меньшей степени. Гипогликемия, т. е. понижение концентрации глюкозы, вызывает реакцию другой группы нервных узлов, также в гипоталамической области мозга; нервное возбуждение передается печени и надпочечникам. Надпочечники выделяют гормон — адреналин, который воздействует на печень вместе с непосредственным нервным возбуждением и вызывает усиленное расщепление гликогена (гликогенолиз) и превращение его в глюкозу, которая выводится в кровь, восполняя недостаток.

Можно перечислить очень много примеров подобных автоматических регулировок. Содержание в организме воды, кислорода, солей металлов, температура тела, давление крови и т. д. — все эти величины поддерживаются постоянными с помощью аналогичных автоматических механизмов, использующих передачу информации посредством химических раздражителей и нервных возбуждений, а также регулировки замкнутого цикла. Во всех этих процессах мы ясно видим характерные черты управления с помощью сигналов, когда энергетические процессы организма вызываются очень слабыми нервными или химическими воздействиями, являющимися только первым толчком к развитию процесса, энергия которого во много раз превышает энергию управляющего сигнала.

В каждом из этих процессов можно проследить возникновение сигнала регулирования, его передачу, усиление, превращение его в энергетическое (механическое, химическое) действие, которое снова вызывает появление уже нового сигнала регулирования. Можно вызывать искусственно (в эксперименте) повреждение механизма регулирования и наблюдать отклонение от нормы в его работе. Именно такими методами (экстрипация желез, перерезание нервных связей, инъекции гормонов) и изучаются механизмы регулирования в живых организмах.

В эксперименте выявляются также характерные черты сходства работы технических и биологических систем регулирования. Так, при искусственном нарушении состояния покоя какой-либо системы регулирования, например при увеличении содержания сахара путем инъекции глюкозы или приема внутрь некоторого количества сахара, можно наблюдать ответную реакцию системы регулирования, снимая зависимость от времени регулируемой величины (концентрация сахара в крови — «сахарная кривая»). Ход кривой напоминает реакцию технической системы регулирования на единичный импульс. Протекание реакции может быть как аperiodическим, так и периодическим с затухающими колебаниями.

Автоматические регулировки можно наблюдать и у растений и низших животных (тропизмы). Примером может служить цветок подсолнечника, поворачивающийся к свету, или бактерия, перемещающаяся в область более благоприятных условий среды.

Сравнивая регулирование при помощи обратной связи в живых организмах и в технических автоматах, можно заметить ряд интересных различий. Так, например, в живых организмах характеристики отдельных звеньев цепи регулирования не обладают линейными свойствами. Величина нервного возбуждения чаще всего бывает пропорциональна не силе раздражения (как в линейных цепях), а ее логарифму. При этом изменения внешних воздействий в весьма широких пределах (в тысячи и десятки тысяч раз) дают изменения величины сигналов, существующих внутри организмов, в более узких пределах (всего на несколько единиц). Эти свойства цепей с логарифмическими амплитудными характеристиками в последнее время все чаще находят применение и в технике.

В живых организмах регулировки чаще всего осуществляются по двум противоположно действующим, как бы «борющимся», каналам. Так, например, к сердцу подходят

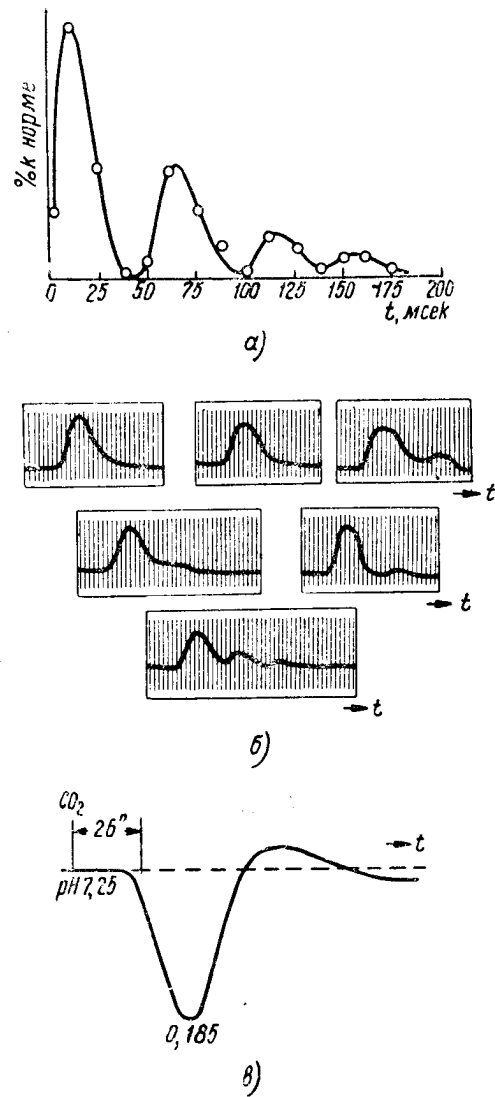


Рис. 6.8. Примеры реакций органических систем с обратной связью на внешнее возбуждение типа единичного импульса.

a) возбудимость сетчатки глаза в зависимости от времени после однократного раздражения электрическим током (по Мотовалов), *б)* изменение скорости роста растения после переноса его из темноты на свет (по Блаув), *в)* изменение водородного показателя (pH) крови собаки после выдыхания углекислоты (по Гольвиггеру). Момент выдыхания отмечен черточками на оси времени.

два регулирующих нервных ствола — блуждающий нерв и ветвь симпатического нерва. Периодические сокращения сердца могут протекать независимо от воздействия обоих этих нервов. Однако оба нерва оказывают воздействие на работу сердца. Возбуждение ветви симпатического нерва вызывает ускорение и усиление сердцебиения, возбуждение блуждающего — замедляет сердцебиение и даже прекращает его вовсе. Воздействия обоих нервов на сердце передаются с помощью нейтротуморов — ацетилхолина и

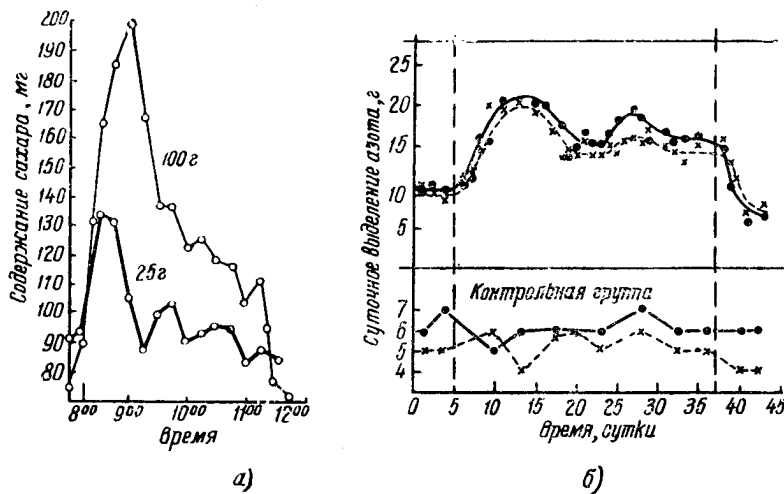


Рис. 6.9. Примеры реакций органических систем с обратной связью на внешнее возбуждение типа единичного импульса и единичного скачка.

а) изменение содержания сахара в крови после приема внутрь 25 и 100 граммов глюкозы (по Францу), б) выделение азота с мочой, после резкого увеличения содержания белков в пище (по Йошимура).

симпатина, соответственно. Ритм работы сердца определяется, таким образом, уровнем возбуждения обоих нервов, действующих в противоположных направлениях: уменьшение возбуждения одного действует примерно так же, как усиление возбуждения другого.

Не только сердце регулируется подобным образом. Вся автономная или симпатическая нервная система делится на два раздела: ортосимпатическую и парасимпатическую системы. Клетки первой располагаются главным образом вдоль позвоночного столба (вне спинного мозга), клетки второй образуют узлы (ганглии)

в различных участках тела, располагаясь рядом с внутренними органами. Функционально каждый внутренний орган испытывает противоположные воздействия от обеих систем; если одна тормозит его функции, другая — возбуждает их. Чувствительность обеих систем к воздействию химических веществ совершенно различна.

Пользуясь техническими терминами, можно сказать, что мы имеем некоторое подобие балансной системы, когда положительные воздействия передаются отдельно от отрицательных, по отдельным цепям. Такая система раздельной передачи положительных и отрицательных воздействий по разным каналам в живом организме неизбежна, так как по нервам могут передаваться возбуждения лишь одного знака, от состояния покоя и до максимума.

Рефлексы

Не только вегетативные функции внутренних органов регулируются механизмами с обратной связью. Такие же механизмы в ряде случаев действуют и в управлении скелетной мускулатурой. К их числу относятся антигравитационные механизмы, обеспечивающие устойчивость тела в поле тяжести, регулировки, создающие наилучшие условия работы рецепторов, а также многие другие безусловные рефлексы.

Скелетные мышцы, обеспечивающие устойчивое положение нашего тела в поле тяжести, например мышцы ног и спины, всегда несколько напряжены. Состояние напряжения мышц вызывает раздражение особых чувствительных окончаний афферентных нервов, расположенных в мышцах и суставах. (Благодаря этим проприоцептивным или интрорецептивным нервам внутренней чувствительности мы знаем о положении наших членов тела, не глядя на них). Если мышца ноги окажется дополнительно растянутой или сокращенной за счет внешнего воздействия (как это случается, например, когда мы спускаемся на лыжах с горы по неровному склону), вследствие автоматической реакции произойдет удлинение или сокращение мышцы таким образом, что нога не потеряет контакта с почвой и не передаст нашему телу опрокидывающего толчка; мышца автоматически приспособится к изменению внешней нагрузки. Регулирующий механизм безусловного рефлекса состоит из афферентных нервов, поднимающихся от чувствительных окончаний к спинному мозгу, промежуточных нервных связей в тканях спинного мозга, эфферентных

волокон, спускающихся от спинного мозга к мышцам, нервных окончаний в мышцах и самих мышц. Реакцию этого нервно-мышечного автомата проверяют невропатологи, постукивая молоточком по сухожилию коленного сустава и наблюдая сокращение мышц (искусственное, механическое растяжение мышцы при ударе по сухожилию вызывает ее физиологическое сокращение).

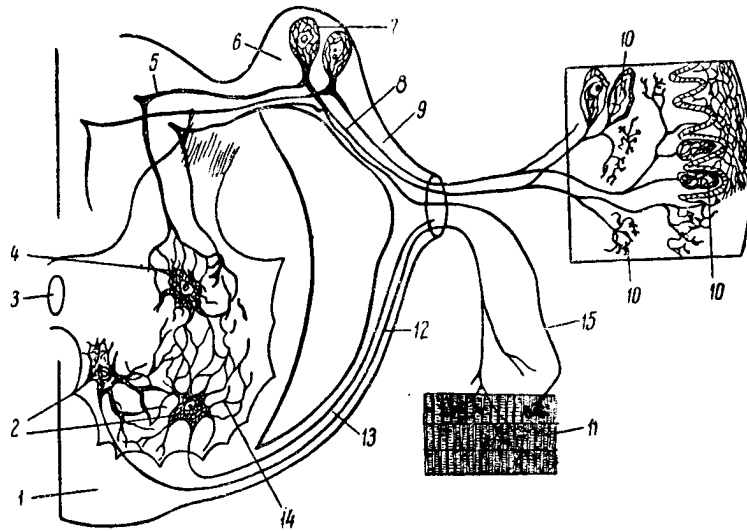


Рис. 6.10. Схема рефлекторной дуги.

1—белое вещество спинного мозга; 2—двигательные нейроны; 3—спинномозговой канал; 4—вставочный нейрон (включенный между чувствительными и двигательными нейронами); 5—аксон чувствительного нейрона; 6—нервный узел; 7—чувствительный нейрон; 8—отросток чувствительного нейрона; 9—задний корешок спинного мозга; 10—чувствительные окончания в коже; 11—двигательные окончания в мышечных волокнах; 12—передний корешок спинного мозга; 13—аксон двигательного нейрона; 14—дендриты двигательного нейрона; 15—волокно проприоцептивного нейрона.

Антигравитационные автоматизмы позволяют нам «не думать» о своем равновесии.

Сказанное выше есть лишь крайне упрощённое описание рефлекторной дуги. В действительности в осуществлении даже элементарных рефлекторных актов участвует более сложная нервная структура. К изложенной выше простейшей схеме следует добавить, в первую очередь, обратную связь, идущую от исполнительных органов (мышц) к спинному мозгу. Эта связь осуществляется нервными окончаниями и клетками-рецепторами внутренней чувствительности (интро- или проприоцепторами). Чувст-

вительные проприоцептивные окончания расположены между мышечными волокнами, в связках и сухожилиях суставов. Раздражения, идущие от проприоцепторов, сигнализируют о положении костей, суставов и состоянии мускулатуры. Таким образом, даже простые движения происходят не как однократный акт «включения в работу» той или иной группы мышц, а скорее как процесс работы системы с обратной связью, в котором сравнивается «желаемый» и «достигнутый» результаты.

К этому нужно еще добавить, что чувствительность проприоцепторов, по крайней мере некоторых из них, может дополнительно регулироваться (натяжением промежуточных мышечных волокон, в толще которых непосредственно расположен рецептор) с помощью отдельных особого типа нервных волокон. Вероятно, нервная структура, осуществляющая простейший рефлекс, не исчерпывается полностью приведённым описанием и имеет дополнительные осложнения.

Безусловные рефлексы в большинстве случаев имеют сложную структуру, причем нервная дуга замыкается через несколько промежуточных нейронов спинного мозга и реакция включает работу не одной мышцы, а целой группы мышц. Если обезглавленной лягушке нанести на кожу живота каплю кислоты, лягушка стирает каплю ближайшей к месту повреждения лапкой. Если эту лапку придержать или же отсечь, то лягушка стирает кислоту другой лапкой или даже лапками другой стороны тела.

Безусловные рефлексы не протекают изолированно. Нервные возбуждения ответвляются от рефлекторной дуги и по нервным связям спинного мозга передаются в кору головного мозга, который в ряде случаев посылает соответствующие эфферентные импульсы и меняет, таким образом, реакцию — усиливает или тормозит ее. Пациент может затормозить коленный рефлекс, напрягая мышцы ноги. Врачу приходится отвлекать внимание больного, заставляя напрягать произвольно другие группы мышц (растягивать руки), чтобы растормозить коленный рефлекс для исследования. Безусловные рефлексы могут быть связаны в цепь и вызывать последовательность стереотипных движений, каждое из которых выполняется с автоматическим приспособлением к требованиям обстановки. Такие цепи произвольных движений — рефлексов — представляют собой ходьба и бег, лазание по лестнице и т. д. Реакции на интенсивные раздражения — кашель, чихание, глотание — представляют собой цепь движений, которые вызывают

друг друга последовательно и образуют вместе целесообразное автоматическое действие, как бы выполнение одной общей программы.

Работа рецепторов (органов внешних чувств), например глаз, протекает по той же схеме автоматических регулировок. При взгляде на светлое пятно глаза совершают целый ряд движений. Во-первых, оси обоих глаз поворачиваются вместе в направлении пятна; во-вторых, оси глаз взаимно смещаются на такой угол, чтобы точка их пересечения пришлась на пятно (конвергенция); хрусталики глаз изменяют форму* и фокусируют изображение на сетчатке (аккомодация); зрачки сокращаются или расширяются в соответствии с яркостью пятна (адаптация) и, наконец, кровеносные сосуды, снабжающие сетчатку, меняют диаметр в соответствии с потребностями в кровоснабжении. Все эти регулировки протекают с помощью нервных возбуждений, идущих в мозг и возвращающихся к мышцам глаз. И здесь имеются противоположно направленные воздействия, только они осуществляются не двумя системами нервных регулировок, а двумя группами мышц, работающих в противоположных направлениях и возбуждаемых, конечно, отдельно. И в этом случае отдельные петли обратной связи работают совместно в едином комплексе и взаимосвязанно. И работа каждой из них зависит от управления со стороны высших отделов центральной нервной системы.

При выполнении сложных движений, требующих точности, обратная связь осуществляется с помощью зрения. Вдевание нитки в иголку протекает как типичный процесс сведения к нулю разности между достигнутым и желаемым результатами, причем эта разность координат ушка иголки и конца нитки воспринимается зрением. С закрытыми глазами вдеть нитку в иголку можно только случайно. Кинестетического чувства, т. е. обратной связи через проприоцепторы, или рецепторы внутренней чувствительности, оказывается недостаточно. Даже овести концы указательных пальцев с закрытыми глазами удастся далеко не всегда. При некоторых расстройствах проприоцептивной чувствительности оказывается затруднительным даже коснуться пальцем носа с закрытыми глазами.

* У некоторых животных, например у рыб, фокусировка производится не деформацией неподвижного хрусталика, а перемещением хрусталика, имеющего неизменную форму, относительно сетчатки вперед и назад.

Письмо и рисование протекают под неослабным контролем зрения. Существует забавный способ убедиться в этом наглядно. Предлагается с закрытыми глазами нарисовать на листе бумаги, не отрывая карандаша, очертание поросенка или собаки. В завершение рисунка следует пририсовать глаз. При этом запрещается прикасаться к бумаге свободной рукой. Обычно на рисунке глаз оказывается на совершенно неподходящем месте, часто даже вне очертаний животного.

Слепые пишут обычно, пользуясь трафаретом, который позволяет им избежать ошибок.

В живом организме существует большое число примеров систем с положительной обратной связью. К их числу, в первую очередь, следует отнести замкнутые цепи пейсеров в нервной системе, которые могут постоянно поддерживать свою активность за счёт самовозбуждения. В настоящее время имеются гистологические данные, доказывающие существование таких цепей. Первоначальный импульс, подаваемый на такую систему, может вызвать очень длительное возбуждение системы, которое облегчает прохождение всех последующих импульсов. Активность в процессе циркуляции импульсов возбуждения по замкнутым цепям поддерживается за счёт метаболизма (обмена веществ, связанного с витанием) нейронов, составляющих цепь.

Самовозбуждающиеся нейронные цепи с положительной обратной связью могут играть определенную роль в процессе образования и закрепления навыков животного, т. е. образовывать материальный субстрат памяти живого существа. Это утверждение подтверждается некоторыми экспериментами. Так, перерезание специфических замкнутых нервных путей приводит к исчезновению процессов образования навыков (опыт производился на головоногих моллюсках).

Различного рода циклические процессы, протекающие в живом организме, являются следствием работы систем с положительной обратной связью. К числу нормальных циклических процессов относится, в первую очередь, работа сердца, которое периодически сокращается, даже если все подходящие к нему нервы перерезаны. Начало сокращения сердца — сокращение предсердий — наступает при раздражении нервных узлов стенки предсердия, которое по коротким выростам клеток проводящей системы распространяется к мышечным волокнам предсердия. Мышечные волокна желудочков возбуждаются другим чувствительным узлом (атриовентрикулярный узел), расположенным в сердечной стенке, который возбуждается от узлов

предсердия, но с некоторой задержкой. Таким образом, приток крови в предсердия вызывает сокращение сердца (систола), а последнее прогоняет кровь сквозь сердце в сосудистую систему и, в свою очередь, вызывает новый приток крови, замыкающий цикл.

Интересно отметить, что в нормальном состоянии члены тела не находятся в покое, а совершают хаотические колебания малой амплитуды с частотами порядка десятка герц. Такие колебания регистрируются, например, при изучении антигравитационных рефлексов. Положение членов тела определяется тонусом (степенью напряжения) мышц, который регулируется по схеме рефлекса. Однако прохождение сигнала нервного возбуждения по рефлекторной дуге происходит с задержкой. Основная доля задержки происходит в соединениях нейронов—синапсах. Время задержки и определяет средний период колебаний. Дело обстоит здесь совершенно так же, как в технических системах с обратной связью, при наличии задержек, о чём мы говорили выше. Механическая инерция членов тела препятствует появлению колебаний большой амплитуды при частоте порядка 10—15 герц. При физическом или умственном утомлении или при заболевании синаптическая задержка прохождения нервного возбуждения возрастает. Это ведёт к понижению частоты собственных колебаний, что, в свою очередь, влечёт за собой увеличение амплитуды. Становится заметным дрожание пальцев, рук, головы.

К циклическим процессам относятся такие патологические явления, как дрожь от переохлаждения тела, заикание, икота и т. п.

Смена сна и бодрствования представляет собой нормальный цикл.

Весьма поучительными, с точки зрения демонстрации действия обратной связи в организме, являются опыты венгерских физиологов Секей и Сентагота. Производилась пересадка ткани образующегося глаза от одного зародыша тритона другому. Пересадка производилась антисимметрично; правый глаз донора пересаживался на место левого глаза реципиента и наоборот, причем нижний край глаза оставался внизу. Таким образом, бывший задний сектор глаза донора оказывался на новом месте, расположенном спереди. После пересадки ткань приживалась и глаз развивался нормально. В процессе развития зародыша нервные связи сетчатки глаза и мозга устанавливались в соответствии с первоначальным расположением глаза донора (задний сектор сетчатки донора, расположенный

после пересадки спереди, соединялся нервным пучком с передним разделом зрительного анализатора, с которым нормально соединяется задний сектор сетчатки). Развившееся животное имело извращенные безусловные рефлексы. Видя червяка, расположенного сзади, передним сектором глаза (расположенным сзади) животное воспринимало его расположение ложным образом, спереди, и реагировало прыжком вперед.

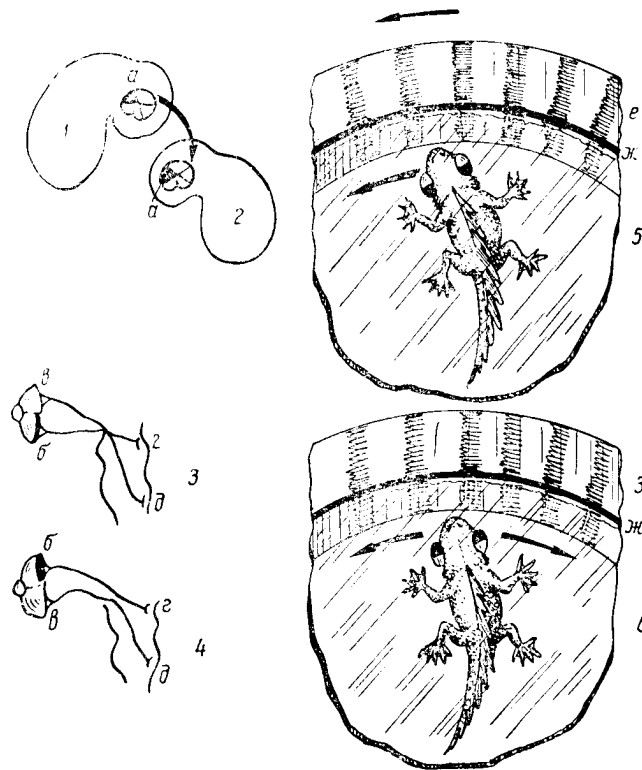


Рис. 6.11. Схема опытов по пересадке глаза у тритона.

1—эмбрион-донор; 2—эмбрион-реципиент; а—передний сектор сетчатки донора, который становится задним сектором у реципиента; 3—прорастание нервных пучков при развитии нормального зрительного анализатора (пучок от переднего сектора сетчатки а прорастает в задний участок зрительного отдела мозга d, пучок от заднего сектора с прорастает в передний участок e); 4—прорастание нервных пучков при развитии зрительного анализатора с пересаженным глазом (пучок от бывшего заднего сектора a прорастает в задний участок d), пучок от бывшего переднего сектора прорастает в задний участок e); 5—нормальное животное совершает движение в направлении движения бумаги e, помещенной снаружи аквариума ж; б—животное, подвергнутое двусторонней пересадке глаз с переменой мест переднего и заднего секторов при неподвижной бумаге e совершает колебательные движения вправо и влево.

Особенно поучительным является сравнение поведения нормального и оперированного животного в следующих условиях: снаружи стенки аквариума помещался лист белой бумаги с нарисованными на нем вертикальными полосами. Если двигать лист вправо и влево, то нормальное животное в ответ начинает перемещаться вслед за движением бумаги. Оно стремится сохранить свое положение относительно окружающих предметов, как в том случае, когда его в естественной обстановке сносит течением. При неподвижной бумаге тритон сохраняет неподвижность. В тех же условиях, при неподвижной бумаге, тритон, у которого оба глаза пересажены так, что передний сектор оказывается сзади и наоборот, не остается неподвижным. Он совершает периодические движения, как маятник, вправо и влево, которые изредка прерываются судорожными хаотическими перемещениями.

И у нормального, и у оперированного животных в описанных случаях имеют место проявления врожденного безусловного рефлекса, определяемого свойствами нервной структуры. У нормального животного рефлекс протекает по схеме с отрицательной обратной связью, у оперированного — с положительной обратной связью, превращающей животное в «генератор колебаний». У нормального животного перемещение бумажной ленты вызывает зрительное раздражение (перемещение образа на сетчатке) и рефлекторное движение «здогонку» бумажной ленты. В результате перемещения тела животного движение образа на сетчатке прекращается, происходит, по терминологии физиологов, «уравновешивание воздействия среды», типичное для отрицательной обратной связи. У оперированного животного случайное движение или поворот его вызывают движение образа на сетчатке, но в обратную сторону (вследствие извращенного положения секторов глаза). Рефлекторная реакция животного (попытка двигаться за кажущимся движением ленты) также происходит в неверном направлении, животное усиливает первоначальное относительное движение, что ускоряет кажущееся движение ленты и ещё более усиливает реакцию (положительная обратная связь).

Поскольку скорость движения животного не может нарастать неограниченно (нелинейность системы), это движение в какой-то момент времени должно замедлиться. Замедление может произойти, например, из-за тормозящего воздействия проприоцептивных раздражений. Замедление движения животного влечёт за собой замедление

перемещения образа на сетчатке, которое стимулирует дальнейшее замедление движения, и так далее. В результате ложные восприятия направления движения стимулируют ускорение начатого движения и заставляют тритона совершать ритмические движения, как это случается с любой другой системой с положительной обратной связью. Неподвижное положение тритона становится неустойчивым.

Преобразование отрицательной, стабилизирующей, обратной связи в положительную, вызывающую появление колебаний, происходит в описанном эксперименте из-за изменения знака возбуждения в одном звене петли обратной связи—перемены мест переднего и заднего секторов сетчатки обоих глаз. Все остальные звенья природной нервной структуры безусловного рефлекса остаются без изменений и не поддаются исправлению путём «переучивания».

Некоторые болезненные состояния вызывают типичные явления нарушения обратной связи. Так, заболевание, называемое локомоторной атаксией и представляющее собой нарушение проприоцептивной чувствительности конечностей, лишает больного возможности чувствовать, например, положение своих ног. Этим разрушаются многие автоматизмы ходьбы. Больной может двигаться, только наблюдая за положением собственных ног. Походка его неуверенна и неловка. В темноте или с завязанными глазами он не может ходить и теряет равновесие.

Рост и развитие организма, размножение клеток, входящих в состав его тканей, также протекает под воздействием многих перекрещивающихся и взаимодействующих регуляторов. Жизнедеятельность одних тканей тормозит или стимулирует рост других в различные периоды развития организма, и только те организмы оказываются жизнеспособными и выживают, у которых процессы взаимных влияний тканей друг на друга приводят к образованию биологически целесообразных структур.

Если, например, некоторая ткань *а* стимулирует развитие ткани *б*, а ткань *б* задерживает развитие ткани *а*, то взаимодействие между ними, протекающее по схеме замкнутой петли отрицательной обратной связи, приведет к некоторому равновесному состоянию, которое может быть нарушено внешними влияниями. Так, например, взаимодействуют друг с другом в теле петуха половые железы и ткань петушиного гребня. Рост гребня тормозит развитие тканей желез, функция желез стимулирует рост гребня

Если хирургически удалить гребень — усилится рост желез; если удалить железы — уменьшится гребень.

Примеров подобного «противоречивого взаимодействия» между органами тела животного можно привести очень много.

В условиях патологических внешних воздействий, в особенности длительных, регуляровки подобного типа могут быть перегружены или даже вовсе нарушены. За нарушением регуляровок специфического синтеза веществ в организме следует патологический рост органов или тканей или всего организма. Так, чрезмерная активность одной из долей гипофиза вызывает гигантизм — ненормальное увеличение размеров тела. Недоразвитие щитовидной железы приводит к недоразвитию тела и кретинизму. Многократное повторное местное разрушение тканей ведет к усиленному течению восстановительных процессов, что в некоторых случаях стимулирует рак — заболевание, связанное с патологическим, неумеренным размножением клеток. Размножение клетки определяется как внутренними процессами в самой клетке, так и внешними условиями. Нарушение нормальных регуляровок роста и размножения клеток в организме, разрыв петли обратной связи, приводит к вредным, а иногда и губительным для организма последствиям.

Системы с обратной связью и в технике, и в живой природе существуют не только в том простом виде одиночной замкнутой петли регулирования, о котором мы вели речь. Взаимная попарная связь различных органов, осуществляемая как через посредство нервной системы, так и с помощью гуморальных воздействий, позволяет проследить длинные цепи регуляровок функций в организме, образующие сложные системы регулирования. Наличие таких систем подтверждается на обширном экспериментальном материале.

Схемы подобных систем и их функционирование весьма сложны. Однако в принципе возможно точное описание их поведения с помощью того же формального аппарата, что и для простых систем. Примером многопетлевой системы могут служить, во-первых, сложные решающие устройства технического назначения и, во-вторых, взаимосвязанные гомеостатические регуляровки в живом организме. Можно сказать, что принцип «взаимно противоречивого» взаимодействия органов, или принцип отрицательной обратной связи, в процессе развития организма связывает части организма в единое целое.

По-видимому, дальнейшее изучение взаимодействия органов, проводимое не только качественно, но и количественно, с привлечением аппарата теории автоматического регулирования, может принести большую пользу не только для изучения функций организма, но и для улучшения методов распознавания причин заболеваний и лечения последних.

Изучение организации систем с обратной связью в живых организмах даёт много полезных конструктивных идей

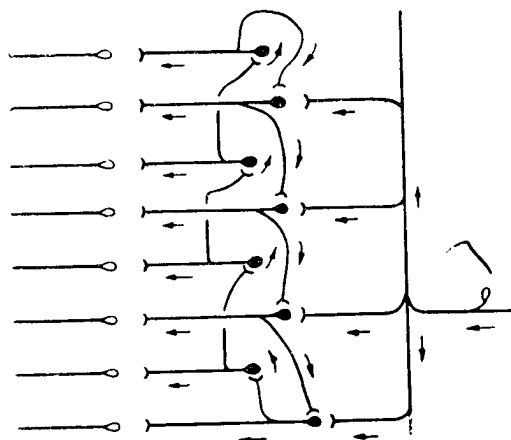


Рис. 6.12. Предполагаемая схема замкнутой цепи нейронов (по Рэнсону).

для создания технических систем. В живых организмах простейшие системы с обратными связями типа безусловных рефлексов, как мы видели, управляются и «настраиваются» более высоко расположенными нервными структурами. В числе технических систем регулирования в последнее время всё чаще встречаются так называемые «самонастраивающиеся» системы.

Обычная система с обратной связью, собранная из звеньев, обладающих некоторыми фиксированными характеристиками, может выполнять свою задачу регулирования наилучшим образом лишь в определённых внешних условиях. При изменении внешних условий качество работы может ухудшиться и для его улучшения могут потребоваться новые регулировки (изменение характеристик системы и её звеньев). Так, автопилот при изменении высоты полёта и плотности воздуха может потребовать таких ре-

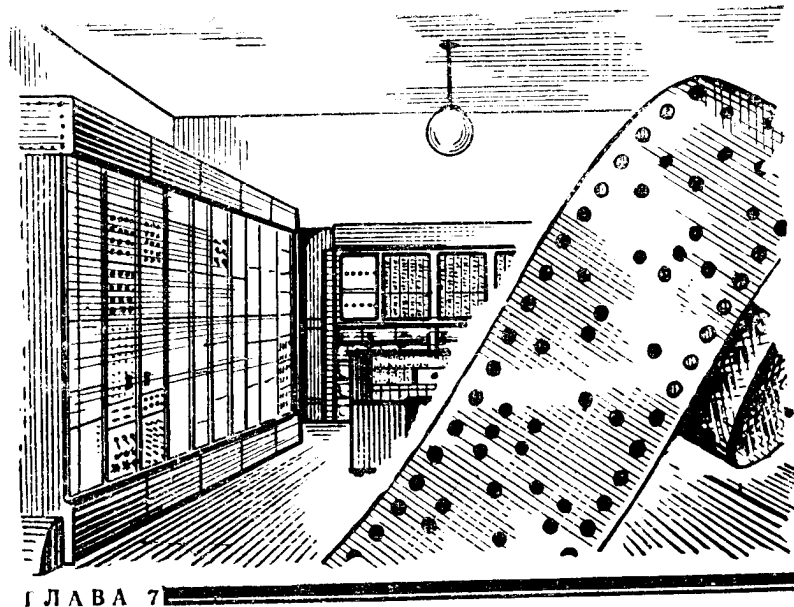
гулировок. Регулировку или подстройку обычно осуществляет оператор, наблюдающий за работой системы и оценивающий по тем или иным признакам качество её работы. Функция текущей подстройки системы может быть передана автоматическому устройству. Для этого над основной системой настраивается ещё одна система регулирования, для которой входной информацией является качество работы основной системы, оцениваемое по объективным признакам, а выходной реакцией — регулировка основной системы для оптимальной работы в соответствии с заданным критерием качества. Все устройство, включающее в себя и основную систему и надстройку, называют обычно самонастраивающейся системой.

Естественно, что «второй этаж» управления может сам нуждаться в подстройке при изменении внешних условий в более широких пределах. Для его регулировки может быть надстроен «третий этаж» и так далее. Каждый следующий «этаж» управления использует всё более общие критерии качества, всё более сложную информацию и должен требовать для своей работы всё больше времени, так как определение качества работы нижерасположенной системы предусматривает проведение «проб» или испытаний работы этой системы в различных режимах.

Любая самонастраивающаяся система требует предварительной принципиальной организации, т. е. воплощения некоторого принципа, некоторого общего критерия качества работы.

В принципе можно, при достаточном усложнении системы, сделать единственным критерием качества естественный критерий: способность системы устойчиво существовать не разрушаясь в данных условиях. При наличии некоторых исходных качеств система может сама, без участия конструктора, перестроиться и дойти до устойчивого состояния посредством изменения своей внутренней структуры. К системам такого типа удобно применить название «самоорганизующиеся». (Часто в просторечии гермины «самонастраивающаяся» и «самоорганизующаяся» система употребляют без различия.) Самоорганизующиеся системы в отличие от самонастраивающихся не несут в себе никакого конструктивного критерия качества работы, кроме естественного критерия — способности «выжить» в данных условиях.

К вопросам реализации самоорганизующихся систем мы вернёмся в гл. 11.



ГЛАВА 7

СИГНАЛ В МАШИНЕ

В системах с отрицательной обратной связью, рассмотренных нами вкратце в предыдущей главе, на входе устройства автоматического регулирования взаимодействовали два сигнала: сигнал, соответствующий «желаемому» или заданному действию автомата, и сигнал, соответствующий уже достигнутому в результате работы автомата действию. Оба сигнала выражаются в простейших случаях в одном и том же физическом алфавите (обычно в виде непрерывно изменяющейся физической величины: электрического напряжения, поворота или скорости вращения вала, концентрации или скорости изменения концентрации химического вещества и т. д.) и в одинаковых масштабах. Взаимодействие двух сигналов сводилось к сравнению двух физических величин с выделением нового сигнала (их разности или некоторой эквивалентной функции), который управлял работой исполнительного устройства. Из двух сравниваемых сигналов внешним был только один.

Можно представить себе ситуацию, когда на вход или входы системы подходят одновременно или последователь-

но во времени несколько сигналов, причем система должна автоматически реагировать не на каждый сигнал, а только на их комбинации.

Примером простейших устройств такого рода является дешифратор, реагирующий лишь на одну единственную комбинацию символов. На кораблях союзников во время Второй мировой войны были установлены автоматические дешифраторы сигналов бедствия. Спасательные лодки и плоты снабжались полуавтоматическими радиопередатчиками, которые после завода часового механизма автоматически в течение нескольких минут передавали на специально отведенной для этого волне сигналы бедствия — серию тире и пауз стандартной длительности. Сторожевые приемники на судах не выключались круглые сутки. На выходе каждого приемника стоял дешифратор. Однажды принятое тире включало программный механизм, который работал только в том случае, когда принимаемые сигналы совпадали с его программой. В конце программы (по окончании последнего тире) включался сигнал, дававший знать вахтенному о приеме сигналов бедствия. Любое отклонение от программы (отсутствие нужного тире или наличие лишних, не выдерживание пауз и т. д.) прекращало работу дешифратора и возвращало его в исходное положение.

Аналогично, но более сложно работает селектор автоматической телефонной станции. На любую (из числа возможных) серию электрических импульсов, получаемых им от наборного диска абонента, он отвечает включением одной из линий — той, которая соответствует набираемому номеру. Селектор реагирует уже не на единственную комбинацию, а на любую. Сопоставление всех элементарных символов сигнала (импульсов) ставит в соответствие определенной комбинации символов одну единственную реакцию.

Закон сопоставления заложен в структуре селектора. Для того чтобы его изменить, необходимо вводить изменения в его схему.

Всякое устройство подобного рода, которое, получив серию или комбинацию входных сигналов, вырабатывает новый сигнал на выходе, в основе своей структуры и работы имеет некоторый закон обработки, сопоставления, взаимодействия входных сигналов, или, как говорят, алгоритм преобразования.

При конструировании автоматических устройств для преобразования сигналов ставится задача реализации заданного алгоритма, т. е. создание схемы машины, ее струк-

туры, которая обеспечивала бы необходимое преобразование. С этим кругом вопросов встречаются конструкторы вычислительных устройств, релейных коммутаторов и, в особенности, автоматических цифровых вычислительных машин. Обычным назначением таких устройств является решение математических, или для цифровых машин, лучше сказать, арифметических, и логических задач. Структурные схемы решающих устройств связаны со свойствами переработки сигналов алгоритмами их преобразований.

В настоящее время существует и широко применяется много типов решающих устройств, отличающихся друг от друга характером и сложностью решаемых задач, использованием тех или иных физических процессов и, главное, характером сигналов, с которыми они работают. По этому последнему признаку вычислительные устройства делятся на два обширных класса — устройства непрерывного действия и устройства дискретного счета. Первые работают с сигналами, имеющими непрерывную структуру, вторые — с дискретными сигналами. Из этого принципиального различия вытекает много различий технического характера и в устройстве, и в областях применения. Различия эти настолько существенны, что описание устройств непрерывного действия и устройств дискретного счета должно проводиться отдельно.

Вычислительные устройства непрерывного действия

Вычислительные устройства непрерывного действия применяются исключительно для решения задач, связанных с физическими величинами, поддающимися измерению, т. е. математических задач, решаемых средствами арифметики, алгебры, интегрального и дифференциального исчисления. Логических задач такие устройства не решают.

Физическая задача, вернее, ее математическая идеализация, изображается в виде дифференциального уравнения или системы дифференциальных уравнений. Затем строится некоторая физическая модель, работа которой описывается тем же дифференциальным уравнением или той же системой уравнений. Эта модель и служит вычислительным устройством. Такие устройства часто поэтому называют моделирующими. Обычно физические процессы, протекающие в модели, совершенно иные, чем в решаемой задаче. Чаще всего строятся электрические модели; однако широко

применяются также механические, гидравлические и другие модели.

Каждая физическая величина, фигурирующая в решаемой задаче (например, дальность, скорость, высота, угол), изображается другой, соответствующей ей величиной в модели (например, углом поворота вала, числом оборотов, на-

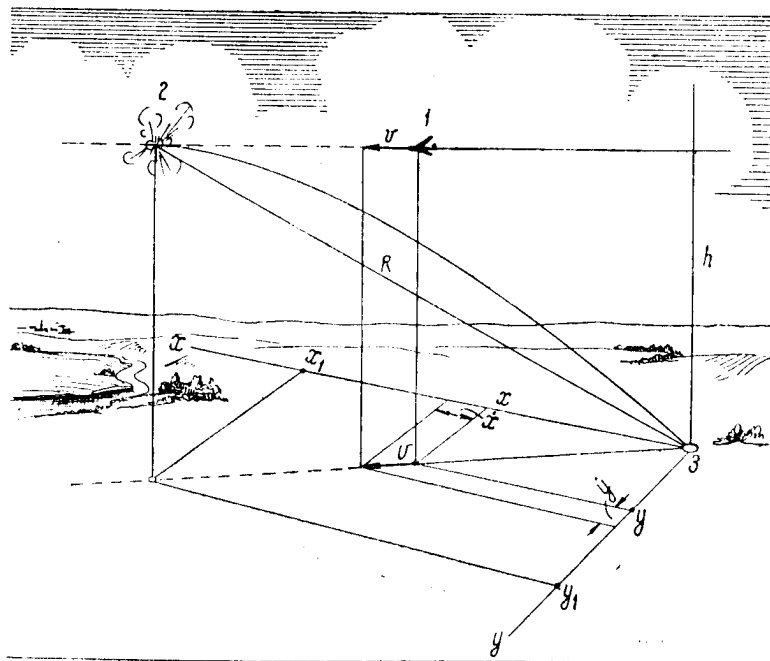


Рис. 7.1. К решению „задачи встречи“ снаряда и самолета.
1—настоящее положение самолета; 2—будущее положение самолета и положение точки разрыва снаряда; 3—точка расположения батареи.

пряжением). Пределы изменения и масштабы величин подбираются, исходя из условий задачи. То, что происходит при работе модели, представляет собой математическую аналогию исходной физической задачи, поэтому решающие устройства непрерывного действия часто называют машинами-аналогами. Каждая величина может принимать любое значение в заданных пределах и непрерывно меняться в этих пределах; соответственно этому непрерывно изменяются и величины в модели, являющиеся сигналами

величины задачи. Процесс решения задачи также протекает во времени непрерывно.

Одной из наиболее популярных задач, решаемых на машинах-аналогах, является «задача встречи», решение которой необходимо для наведения орудия на движущуюся цель и наведения истребителя на самолет противника. Допустим, что нас интересует только величина полетного времени снаряда T , которое сложным образом зависит от наклонной дальности R до точки встречи снаряда и цели. Если цель находится в точке с координатами x, y, h

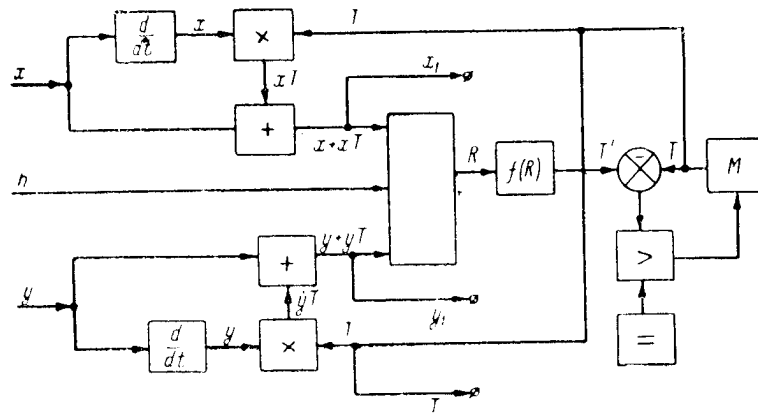


Рис. 7.2. Упрощенная схема вычислительного устройства прибора управления артиллерийским огнем.

(в декартовой системе координат) и имеет составляющие скорости \dot{x}, \dot{y} , то, спустя время T , она будет находиться в точке

$$x_1 = x + T\dot{x}; \quad y_1 = y + T\dot{y}; \quad h_1 = h.$$

Для определения координат точки встречи x_1, y_1 и времени T необходимо решить систему уравнений

$$R = \sqrt{(x + \dot{x}T)^2 + (y + \dot{y}T)^2 + h^2},$$

$$T = F(R),$$

где $F(R)$ задается баллистическими таблицами.

Возможная схема вычислительного устройства прибора управления артотнем приведена на рис. 7.2. Текущие координаты цели x, y, h (h остается постоянной) вводятся

непрерывно в виде условных сигнальных величин от измерительных устройств (радиолокатора, дальномера). Координаты x и y дифференцируются и умножаются на T , которое задается следящей системой. Далее к произведениям xT , yT прибавляются текущие координаты x и y , и затем три полученные координаты «упрежденной» точки геометрически складываются (определяется диагональ параллелепипеда R); по величине R определяется T' , которое подается на вход следящей системы; величина T с выхода следящей системы, как мы уже говорили, подается в вычислительное устройство. Поскольку следящая система непрерывно поддерживает равенство входного и выходного сигналов $T=T'$ (по окончании начального переходного процесса), выходные величины x_1 , y_1 , T , являющиеся результатом решения задачи, также непрерывно выдаются, если только на вход подаются текущие координаты цели.

Для реализации приведенной в этом грубом описании схемы нам нужны устройства, способные производить указанные в квадратах схемы преобразования: дифференцирование, умножение, алгебраическое и геометрическое суммирование, функциональное преобразование, интегрирование и т. д. Весь арсенал средств, служащих для такого рода преобразований, в настоящее время хорошо разработан для вычислительных устройств различных типов: механических, электрических и др.

Каждое вычислительное устройство типа модели-аналога приспособлено для решения только одной задачи. Изменение задачи влечет за собой изменение конструкции модели, что практически выполнить нелегко, особенно в механических моделях. Точность получаемых в результате решения величин обычно невелика, ошибка составляет не менее 0,1%, в лучших случаях. Все узлы и детали машины должны выполняться с высокой точностью и потому они дороги и трудны в исполнении.

Удобством в работе с аналогами является то, что входные данные на них можно подавать прямо с измерительных устройств, согласовав масштабы, без сложных промежуточных преобразований, а выходные — непосредственно использовать для управления объектами (например, орудиями) в качестве входного сигнала силовой передачи. Поэтому чаще всего вычислительные устройства непрерывного действия применяются в системах автоматического управления, где приходится решать на данной модели только одну задачу.

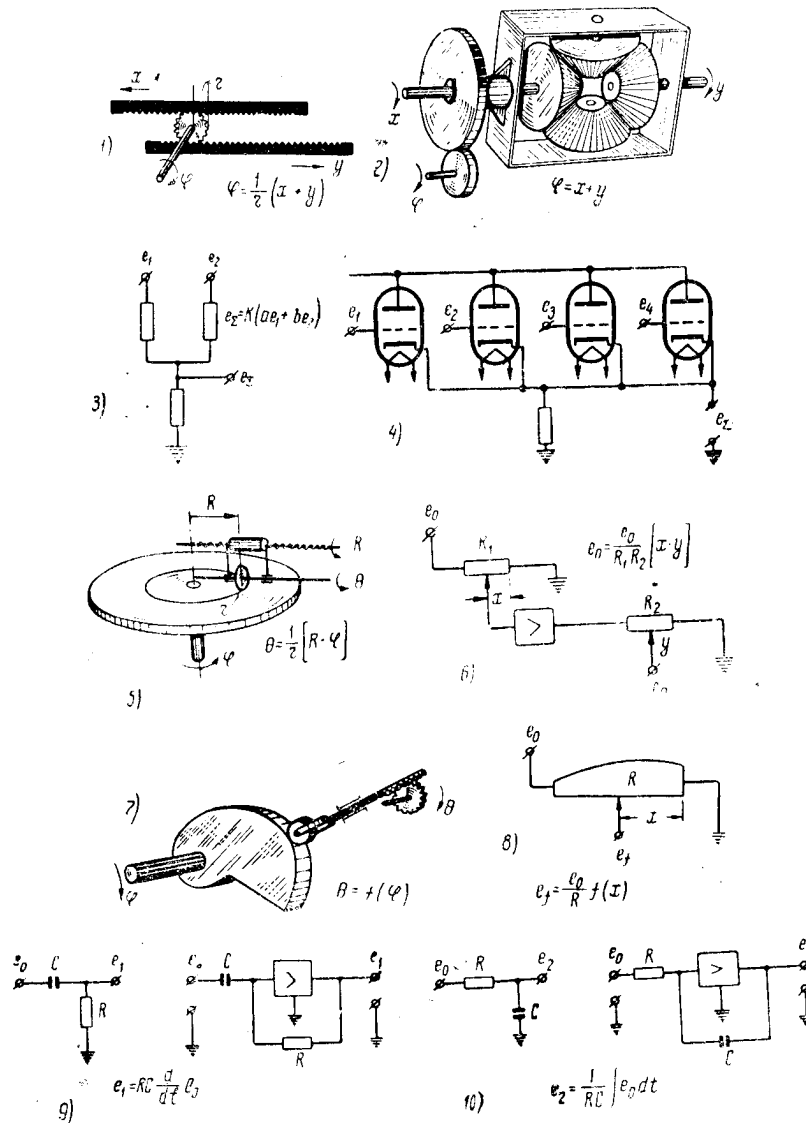


Рис. 7.3. Примеры функциональных узлов счетно-решающих устройств.
 1 - суммирование механических перемещений; 2 - дифференциал (суммирование углов поворота); 3 - суммирование электрических напряжений на сопротивлениях; 4 - суммирование электрических напряжений в ламповой схеме; 5 - множительный механизм; 6 - умножение с помощью потенциометров; 7 - эксцентриковый механизм (задание функции от одного аргумента); 8 - функциональный потенциометр; 9 - дифференцирующая ячейка; 10 - интегрирующая ячейка.

Функциональная схема модели не определяется однозначно системой уравнений задачи. Одну и ту же задачу можно решать различными путями на конструктивно различных моделях, построенных по различным функциональным схемам. Машины с различными структурами могут, таким образом, выполнять одинаковую функцию. Обратное для машин-аналогов неверно: определенная схема может решать только одну математическую задачу, т. е. выполнять только одну функцию.

Во время решения задачи машина-модель непрерывно отображает текущее состояние процесса, относительно которого сформулирована задача. Внутри модели как бы воспроизводятся в виде сигналов все существенные для решения задачи стороны процесса, вся полезная для решения информация. При этом физическая структура и состояние модели могут не иметь ничего общего с отображаемым процессом (кроме изоморфной связи события и сигнала). Если модель сконструирована и запущена правильно, то она даст правильный ответ в том смысле, что снаряд, выпущенный в соответствии с рассчитанными величинами установок орудия, точно или почти точно попадает в цель. На этом мы убеждаемся в правильности утверждения об изоморфизме сигналов и событий в данном случае.

Термины «правильность» и «точность» не равнозначны. Точность (или обратная ей величина—ошибка) характеризует допуск в определении искомой величины (вместе с вероятностью этого допуска). Правильность состоит в том, что истинная (реальная) величина и величина, полученная с преобразованиями сигнала, совпадут в пределах допуска.

Дискретный счет

Физическая величина может быть сопоставлена не только с другой физической величиной, но и выражена числом, записанным в определенной системе счисления в виде ряда условных символов. Для этого измеряемая величина сравнивается с некоторой единицей измерения и в результате сравнения получается число, умноженное на единицу, имеющую размерность физической величины. Число записывается в виде ряда символов—цифр, расположенных на определенных местах—позициях. Каждая позиция соответствует разряду и придает стоящей в этом разряде цифре определенный вес. В повсеместно приме-

няемой десятичной системе счисления числа, выражаемые цифрами, стоящими в соседних разрядах, отличаются по величине в 10 раз.

Так число 325 равно $3 \cdot 10^2 + 2 \cdot 10^1 + 5 \cdot 10^0$. Число десять является основанием десятичной системы счисления. В этой системе в каждом разряде может стоять один из десяти возможных символов 1, 2, 3, 4, 5, 6, 7, 8, 9 и 0. При

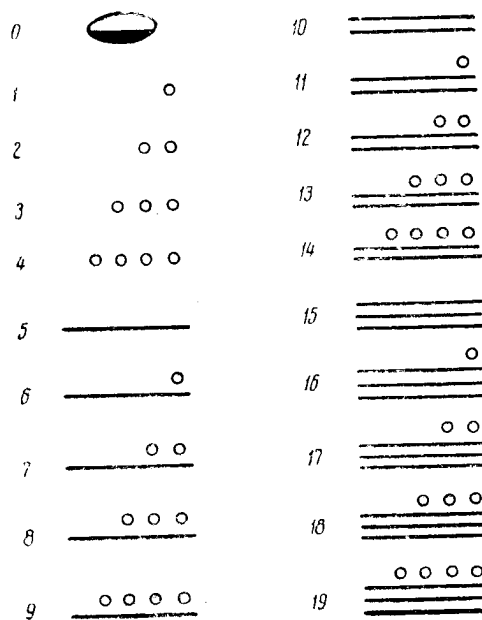


Рис. 7.4. Изображение чисел у племени Майя.

перечислении последовательно возрастающих чисел символы меняются в указанном порядке, и по завершении цикла, при переходе от 9 к 0, меняется символ в соседнем слева разряде.

Эти общеизвестные правила мы напомним для того, чтобы воспользоваться ими для других систем счисления, отличающихся тем, что в качестве основания в них используется не 10, а другое число: 2, 3, 12, 20 и т. д. Некоторые из таких систем счисления получили за последнее время большое распространение. Все эти системы одинаково пригодны для практического использования, и применение

той или иной из них определяется, во-первых, привычкой и, во-вторых, некоторыми техническими удобствами. Так, например, общепринятая десятичная система имеет или, вернее, имела в далеком прошлом то преимущество, что пальцы двух рук представляли собой удобную «счетную машину», всегда находящуюся в распоряжении вычислителя, который мог данные любой простой задачи «по пальцам пересчитать».

Отчасти сохранились до настоящего времени и другие системы счисления: счет дюжинами и grossами (дуодецимальная система, основание—12), счет шестью десятками — минуты, секунды при исчислении времени, углы в градусной мере (гексагезимальная система).

У древнего южноамериканского племени индейцев Майя, не знавших обуви, счет велся по основанию 20 по очевидным причинам. Символы, употреблявшиеся для записи разряда, прямо указывают на их происхождение (например, число 7 — рука и два пальца). Мудрые лошади «Гуньяньямы», в государстве которых гостил герой Дж. Свифта Гулливер, наверное, должны были бы пользоваться двоичным счетом, так как вместо человеческих десяти пальцев в их распоряжении было только два копыта.

В любой системе счисления позиционная запись числа имеет вид

$$a_3 a_2 a_1 a_0,$$

что означает

$$a_3 c^3 + a_2 c^2 + a_1 c^1 + a_0 c^0,$$

где c —основание счисления, $a_0, a_1, a_2, a_3, \dots$ —один из возможных в разряде символов $0 \leq a_i < c$, означающий целое число единиц.

В табл. 7.1 приводятся для сравнения записи одних и тех же чисел в различных системах счисления.

Вводя разряды справа от запятой, представляющие доли единицы, мы можем записать любую величину с любой заданной погрешностью в виде ряда дискретных символов, т. е. в виде разрядного представления числа. Таким образом, непрерывная физическая величина может быть изображена в виде сигнала дискретной природы. Реальные преобразования физических величин получают коморфное отображение в действиях над символами записи числа, производимых по определенным правилам.

Таблица 7.1

Таблица сравнения записи чисел в различных системах счисления

Десятичная или десятичная (основание 10)	Восьмиричная или октавальная (основание 8)	Трочная или тринерная (основание 3)	Двоичная или бинарная (основание 2)	Двудесятичная (разряды десятичные, числа в разрядах двоичные)
0	0	0	0	0000
1	1	1	1	0001
2	2	2	10	0010
3	3	10	11	0011
4	4	11	100	0100
5	5	12	101	0101
6	6	20	110	0110
7	7	21	111	0111
8	10	22	1000	1000
9	11	100	1001	1001
10	12	101	1010	0001 0000
11	13	102	1011	0001 0001
12	14	110	1100	0001 0010
13	15	111	1101	0001 0011
14	16	112	1110	0001 0100
15	17	120	1111	0001 0101
16	20	121	10000	0001 0110
17	21	122	10001	0001 0111
18	22	200	10010	0001 1000
19	23	201	10011	0001 1001
20	24	202	10100	0010 0000
21	25	210	10101	0010 0001
30	36	1010	11110	0011 0000
50	62	1212	110010	0101 0000
100	144	10201	1100100	0001 0000 0000
1000	1750	1101001	111101000	0001 0000 0000 0000

Вычислительные устройства дискретного счета

Символы, применяемые для записи числа, не обязательно должны иметь вид знаков, написанных на бумаге. Каждому из знаков может быть поставлено в соответствие некоторое состояние физической системы с соответствующим числом степеней свободы. Так, каждому символу-цифре может соответствовать число отложенных в конторских счетах косточек, нажатие одного из клавиш счетной машины, поворот валика с фиксатором на несколько ступеней — «щелчков» и т. д. Счетные устройства имеют при этом структуру, соответствующую применяемой системе

ме счисления; они имеют разряды и в каждом разряде число состояний, равное основанию счисления. Все это совершенно ясно видно на конструкции конторских счет. Тот же принцип лежит в основе механических счетных машин.

Вводя число в арифмометр или откладывая его на счетах, мы делаем, по существу, то же, что и при записи

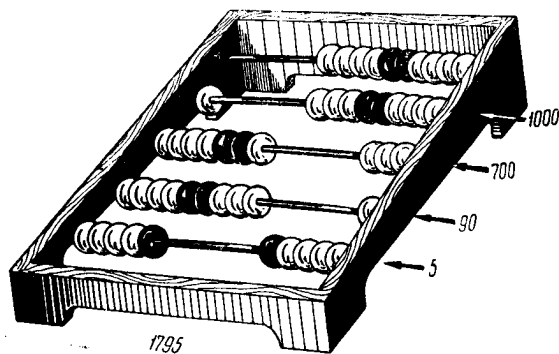


Рис. 7.5. Конторские счета.

числа цифрами на бумаге. Только те же действия мы выражаем другими символами в другом физическом алфавите.

Вычисляя на бумаге, производя, например, сложение, мы по двум записанным столбиком цифрам, пользуясь «таблицей сложения», выученной наизусть в школе, определяем третью и записываем ее в столбик под слагаемыми. Если нужно, мы переносим единицу в разряд слева. Буквально то же самое делается и в арифмометре с той лишь разницей, что «таблица сложения» уже содержится в конструкции колес арифмометра вместе с правилом переноса единицы в старший разряд. Алгоритм сложения мы не должны держать в голове и применять его по очереди к каждому разряду; он автоматически выполняется арифмометром при повороте ручки, так как он заложен в конструкцию арифмометра. Это возможно потому, что арифмометр имеет дискретный ряд состояний, изменения которых связаны друг с другом так же, как связаны символы в разрядах складываемых чисел.

Хотя принципиально здесь все представляется предельно простым, изобретение механического сумматора явилось в свое время выдающимся достижением человеческой мыс-

ли и сослужило неоценимую службу развитию техники и науки.

Инженерные задачи создания счетных машин сводятся в настоящее время не к поискам принципов осуществления алгоритмов арифметических действий, а к наилучшему воплощению этих принципов и созданию наилучших (простых, удобных, дешевых) конструкций.

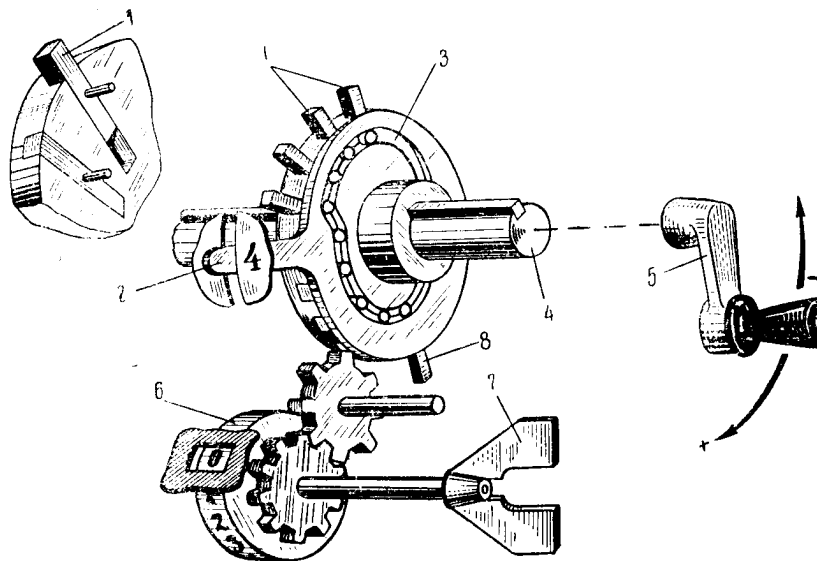


Рис. 7.6. Суммирующий механизм арифмометра („колесо Соднера“).
1—выдвижные зубцы; 2—спица для установки числа; 3—прорезь, выводящая зубцы 1 вверх при повороте спицы 2; 4—вал; 5—ручка; 6—счетчик (при повороте ручки 5 счетчик 6 поворачивается на такое число зубцов, которое соответствует числу выведенных вверх зубцов 1 колеса; при многократном повороте в счетчике откладывается сумма); 7—рукоятка сброса (установка счетчика на нуль); 8—дополнительный зубец для передачи единицы в соседний слева разряд при переполнении счетчика.

В настоящее время механические счетные машины автоматически суммируют и вычитают, множат, делят и печатают на ленте результаты вычислений. Вычислитель, работающий на машине, не обязан больше помнить таблицы сложения и умножения; их однажды заложил разработчик в конструкцию машины. В обязанности вычислителя входит теперь только вводить в машину исходные данные, включать нужные операции, фиксировать результаты вычислений и вводить их снова, соблюдая общую программу решения задачи.

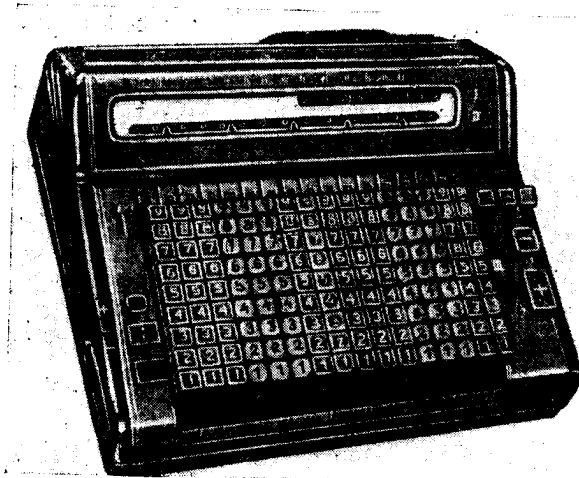
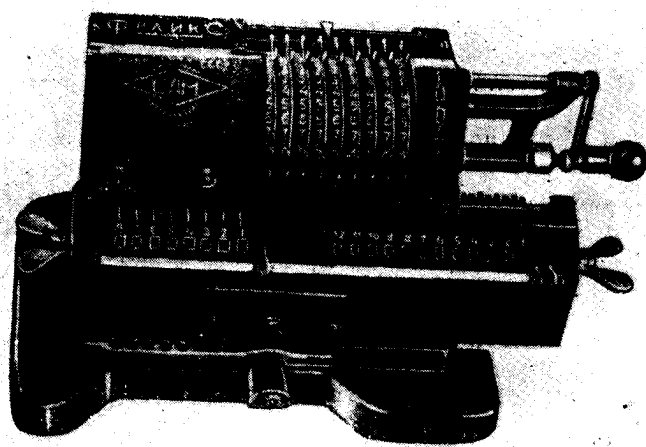


Рис. 7.7. Клавишные вычислительные машины.
а) арифмометр; б) автоматическая вычислительная машина.

Счетно-аналитические машины

Ввод данных в машину и передача их из одной машины в другую могут быть также автоматизированы. Чаще всего это делается с помощью перфорированных карточек или лент. Перфокарта представляет собой кусок тонкого картона стандартной формы, на котором пробиты одинаковые, но различно расположенные отверстия. Отверстия пробиваются только в определенных точках карточки, расположенных в виде правильной прямоугольной сетки по строкам и столбцам. Наличие или отсутствие пробивки на данном месте карты может условно обозначать ту или иную цифру, букву или другой символ. Пробивки на карточках представляют собой, таким образом, сигнал, несущий ту или иную информацию в соответствии с принятым кодом, а сама перфокарта со всеми возможными пробивками представляет собой дискретный алфавит.

Вычислительные устройства, работающие с перфокартами, носят общее название «счетно-аналитических машин». В число их входят машины различного назначения:

«перфоратор» — для нанесения вручную перфораций на карты,

«сортировка» (сортировочная машина) — раскладывающая группу карт по различным карманам в зависимости от наличия пробивок в тех или иных местах карты,

«контрольщик» — для проверки пробивок,

«интерпретер» — расшифровщик, печатающий текст с перфокарты,

«табулятор» — осуществляющий подсчет карт, сложение и вычитание данных, пробитых на картах, и пробивку результатов,

множительный и вычислительный перфораторы — выполняющие сложение, умножение, вычитание и деление данных, пробитых на картах, и пробивающие результаты и т. д.

Существует много типов таких машин с различными особенностями конструкции, различной скоростью и для различных операций. Во многих машинах изменением внутренних соединений — переключением на коммутационной доске — изменяется их схема и функции, которые они выполняют.

Работа машины основана на том, что пробивки на введенной в машину карте оказывают на механизмы машины такое же действие, как нажатие клавиша или включение контакта. При прохождении карточки в машине сквозь

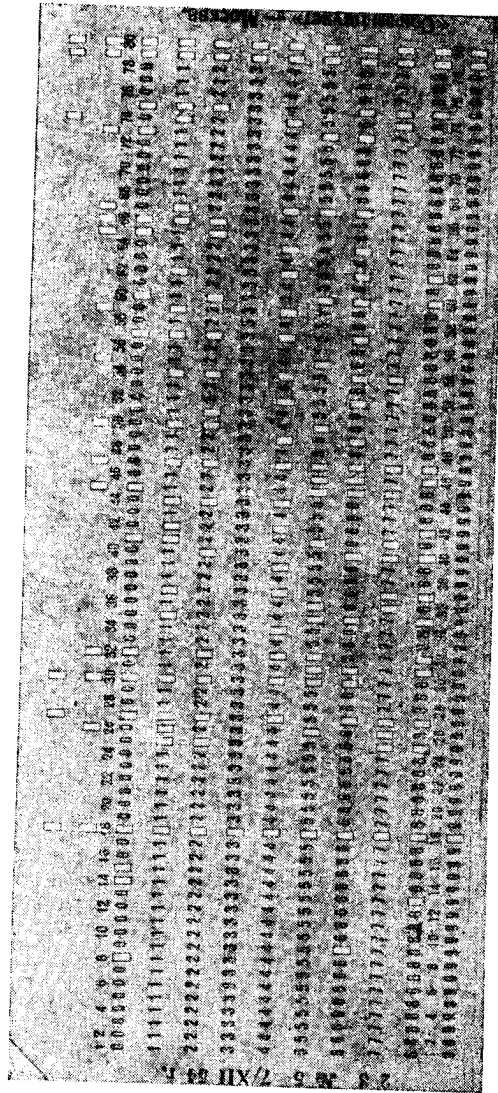
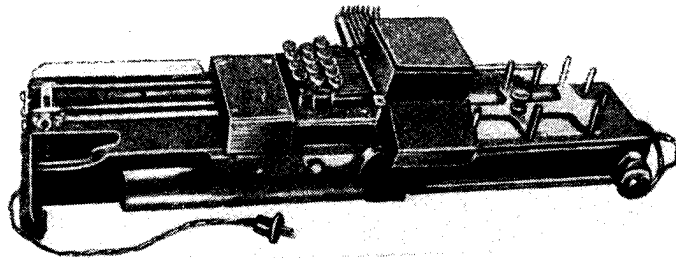
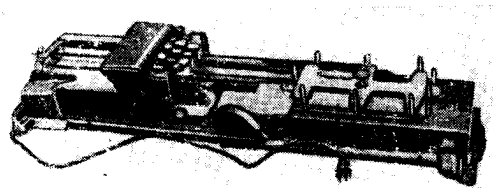


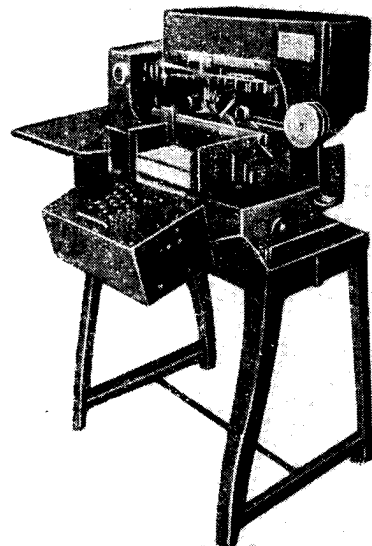
Рис. 7.8. Перфокарта с нанесением пробивок для записи буквенного текста.



a)



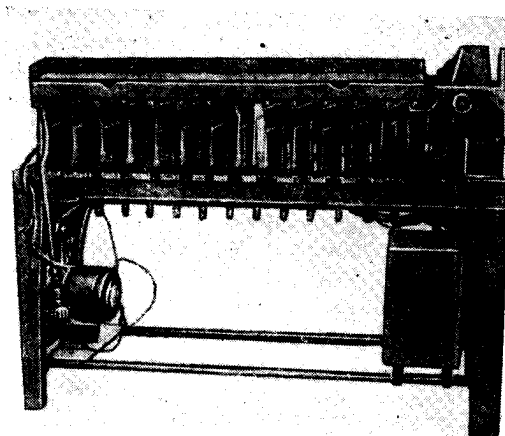
b)



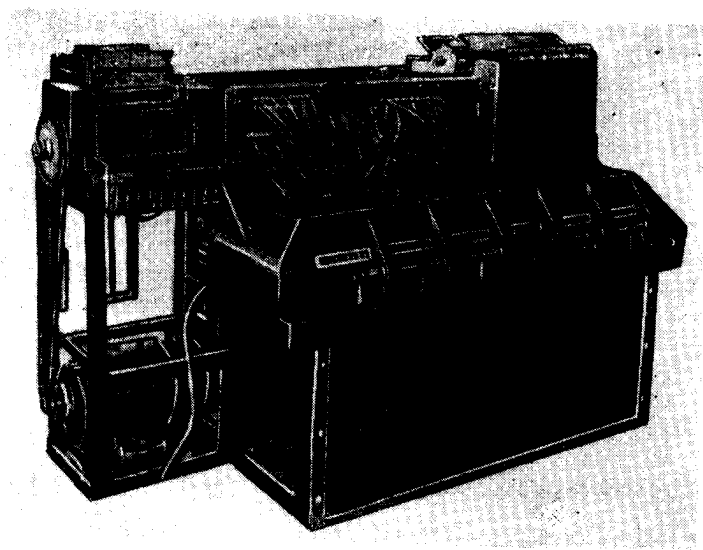
в)

Рис. 7.9. Счетно-аналитические машины.

a) перфоратор САМ; б) контрольный САМ; в) перфоратор ПД-45-1.



а)



б)

Рис. 7.10. Счетно-аналитические машины.
а) сортировка САМ; б) табулятор Т-2.

*

пробивку либо проваливается шуп, либо замыкается электрический контакт. Таким образом, автоматическая подача карты равносильна вводу данных на клавиатуру.

Счетно-аналитические машины освобождают оператора еще от нескольких обязанностей. Они способны автоматически регистрировать информацию и сами ее считывать, хранить информацию, вести подсчет, сортировать карты по данному признаку или ряду признаков и т. д. Передавая карты из одной машины в другую, можно, не вмешиваясь в ход решения, получить полное решение сложной задачи с большим числом исходных данных.

Возможности счетно-аналитических машин не ограничиваются, таким образом, только арифметическими действиями. Сравнение, выбор по признакам, сортировка являются уже не арифметическими, а логическими операциями.

Особенно широко используются счетно-аналитические машины в операциях бухгалтерского учета. Однако они находят применение и в решении научно-технических задач.

Счетно-аналитические машины могут производить вычисления со скоростью порядка тысячи умножений шестизначных чисел в час; сортировальные машины пропускают около 500 карт в минуту. Все операции, производимые с помощью перфокарт, можно вести и с перфолентой, что в ряде случаев удобнее, так как ленту проще подавать в машину.

Перфоленты и перфокарты служат для обработки и хранения не только чисел, но и других видов информации, например текста, условных обозначений и т. п. Это сильно расширяет область их применения.

В конструкции карточки отражается позиционная система записи. Число пробивается на ней поразрядно. В каждом разряде позиция пробивки соответствует цифре. Перфокарта, изображенная на рис. 7.8, использует десятичную систему счисления. Можно пользоваться и другими системами, например двоичной.

Быстродействующие электронные вычислительные машины

Совершенно новый этап в обработке дискретных сигналов на машине наступил с появлением электронных вычислительных машин. В этих машинах сигнал имеет форму электронных импульсов весьма малой продолжительности (порядка 1 микросекунды, т. е. 10^{-6} секунды). Такого же

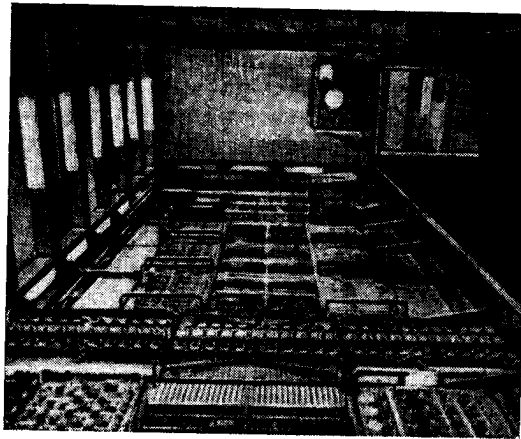


Рис. 7.12. Цифровая электронная вычислительная машина "Вихрь" (Whirlwind).
Внешний вид.

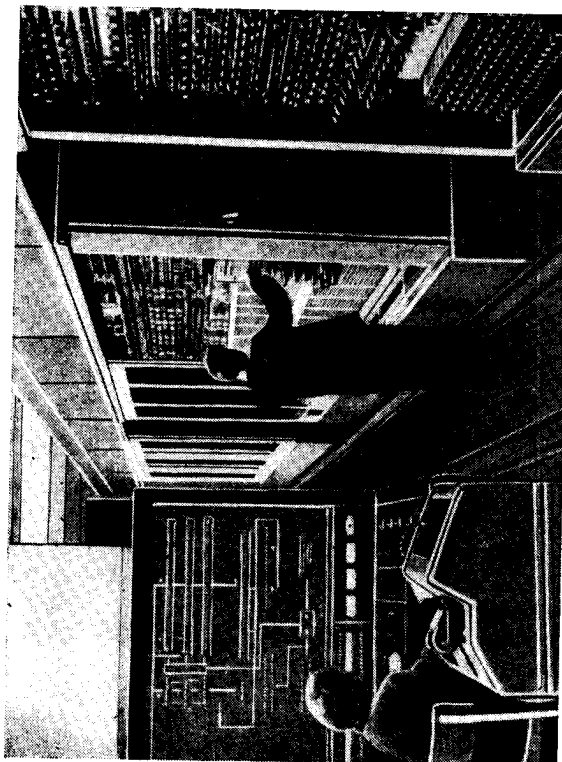


Рис. 7.11. Цифровая электронная вычислительная машина БЭСМ.
Внешний вид.

и время реакции схемы. Это в десятки и сотни тысяч раз короче, чем время работы механических рычагов или электромеханических реле. Соответственно выше и скорость вычислений. Современные электронные машины способны производить вычисления со средней скоростью 5—7 тысяч, а некоторые и до 20 тысяч операций в секунду с девятиразрядными десятичными числами. Использование чрезвычайно гибкой и богатой электронной техники открыло новые пути и настолько широкие перспективы применения вычислительных устройств дискретного счета, что об этом типе машин следует сказать несколько подробнее.

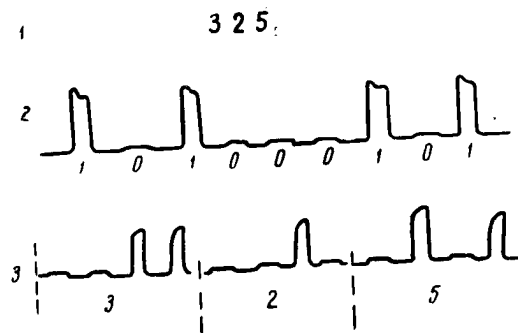


Рис. 7.13. Запись чисел с помощью электрических импульсов.
1—обычная запись числа в десятичном счислении; 2—запись в виде электрических импульсов в двоичной системе счисления (наличие импульса означает „1“, отсутствие его—„0“); 3—запись числа в виде электрических импульсов в двоично-десятичной системе счисления (четыре двоичных разряда образуют один десятичный разряд, в котором десятичное число записано в двоичной системе).

Основным носителем сигнала в электронной машине является электрическое напряжение, основным элементом, фиксирующим символ алфавита, — электронный триггер.

Схема триггера представляет собой соединение двух электронных усилительных ламп. Обе лампы должны быть включены, как усилители напряжения с положительной обратной связью: входы усилителей, сетки ламп, соединены с выходами, анодами, крест-накрест, т. е. сетка первой — с анодом второй и наоборот. Схема имеет два состояния устойчивого равновесия:

- 1) первая лампа проводит, вторая заперта отрицательным потенциалом на сетке,
- 2) симметричное относительно первого состояния.

Состояние схемы, при котором через обе лампы текут одинаковые токи, неустойчиво. Достаточно току через одну

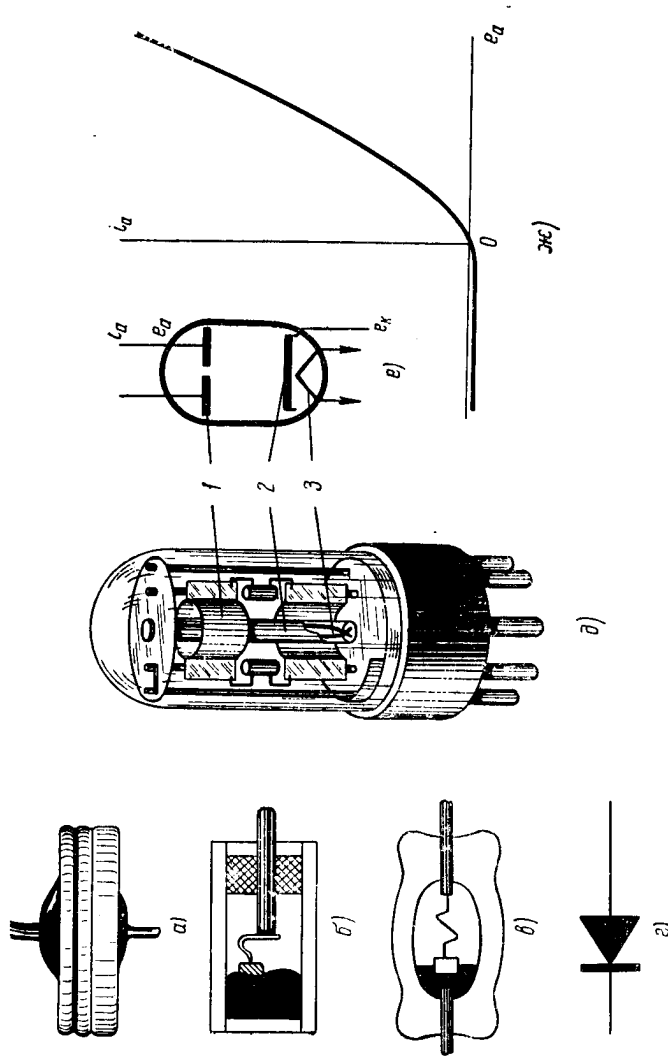


Рис. 7.14. Элементы схемы машины. Диод—двуэлектродный прибор, проводящий ток только в одном направлении.

а) селеновый диод; б) кремниевый диод в металлической капсуле; в) германиевый диод в стеклянной оболочке; г) обозначение кристаллического диода в схемах; д) вакуумный диод; е) изображение вакуумного диода в схемах (1—анод, 2—катод, 3—нить накала); ж) характеристика диода—зависимость тока через диод от напряжения на его аноде.

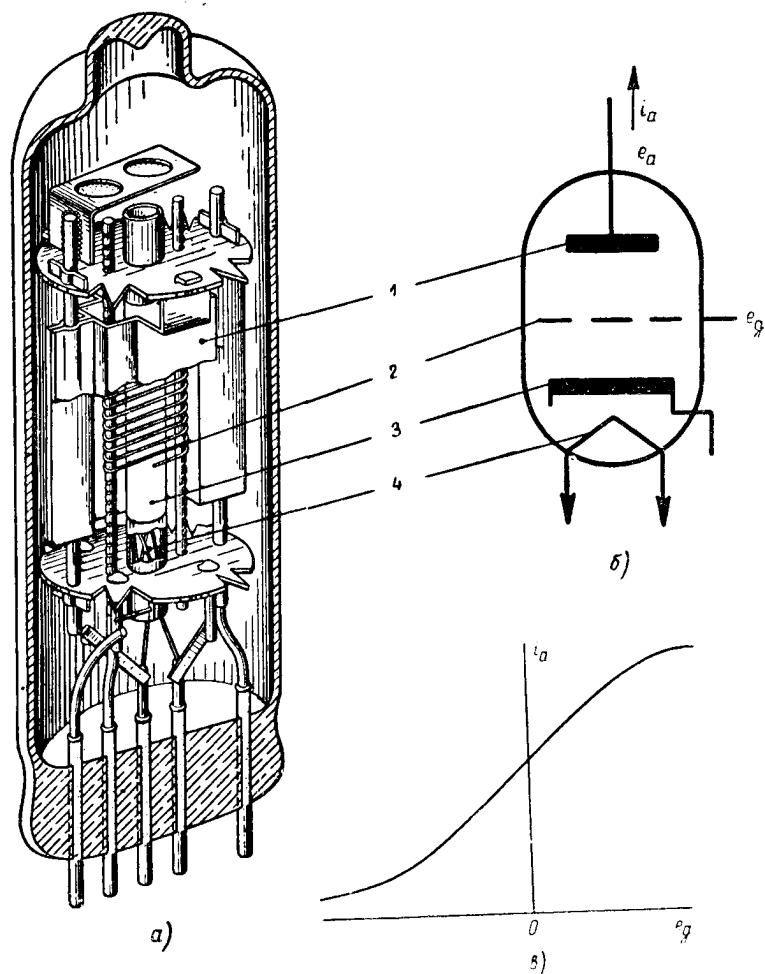


Рис. 7.15. Элементы схемы машины. Вакуумный триод—трехэлектродный прибор, осуществляющий усиление сигнала за счет энергии внешнего источника.

а) конструкция триода; б) обозначение триода в схемах (1—анод, 2—сетка, 3—катод, 4—нить накала); в—характеристика триода—зависимость тока анода от напряжения на сетке.

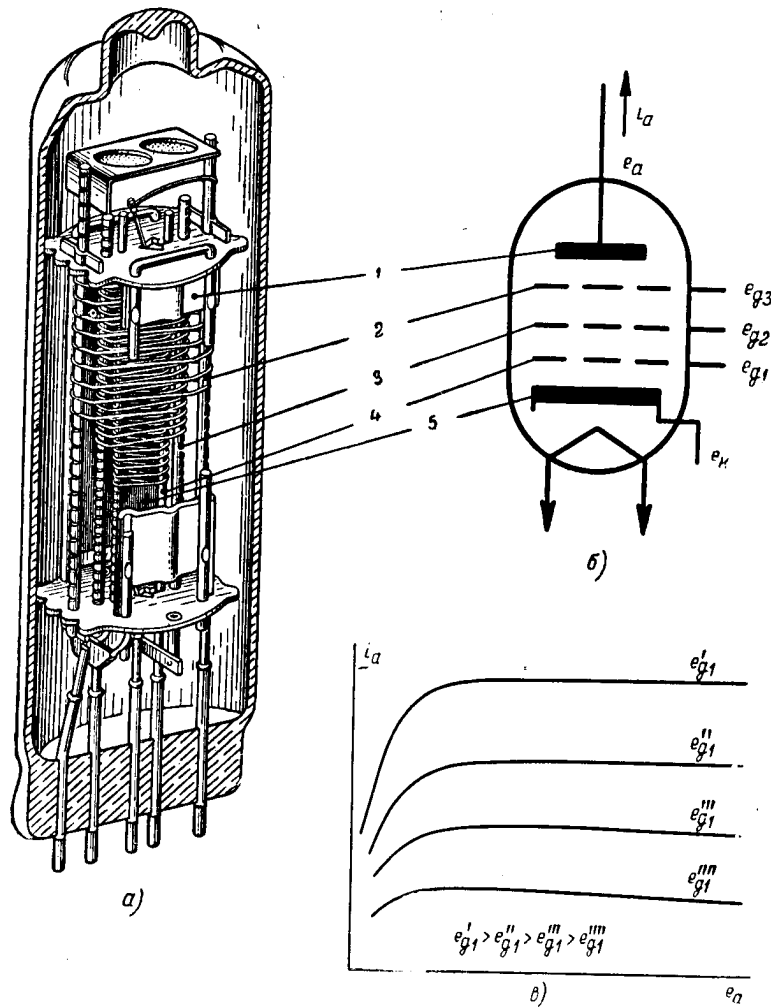


Рис. 7.16. Элементы схемы машины. Вакуумный пентод — электронная лампа с тремя сетками. Лампа используется как усилитель сигнала и как прибор, позволяющий осуществлять управление от двух сигналов одновременно.

а) конструкция пентода; б) обозначение пентода в схемах [1 — анод, 2 — третья сетка (используется для управления или для защиты от вредного действия вторичной эмиссии электронов с анода), 3 — вторая сетка (используется для экранировки первой сетки от анода), 4 — первая управляющая сетка, 5 — катод]; в) характеристика пентода — зависимость анодного тока от напряжения на аноде при различных напряжениях на управляющей сетке.

из ламп случайно измениться, например уменьшиться, как сейчас же возникнет процесс, который переведет схему в устойчивое состояние: уменьшение тока через лампу 1 (см. рис. 7.18) вызовет увеличение напряжения на ее аноде и соответственно на сетке лампы 2; это вызовет увеличение тока через вторую лампу, понижение напряжения на ее аноде и соответственно на сетке лампы 1; последнее в свою оче-

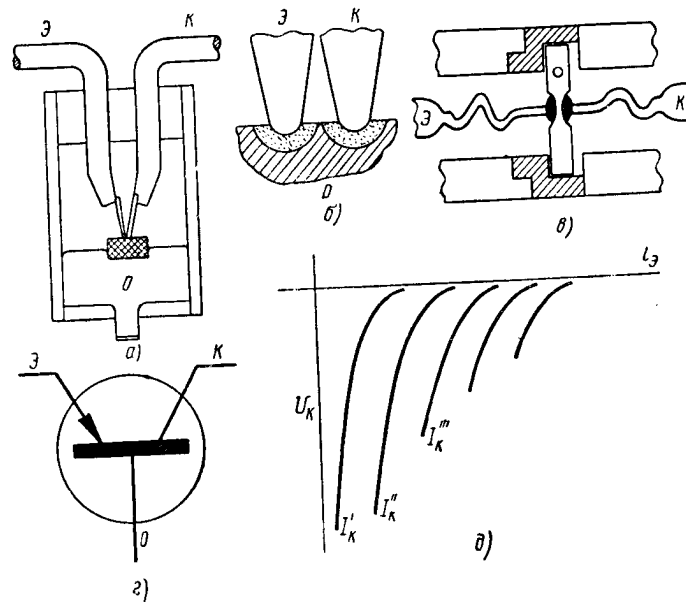


Рис. 7.17. Элементы схемы машины. Полупроводниковый триод (транзистор) — трехэлектродный прибор, осуществляющий усиление сигнала за счет энергии внешнего источника.

а) точечный триод; б) расположение электродов (протекание тока между металлическим эмиттером ε и полупроводниковым основанием o увеличивает концентрацию носителей заряда в объеме полупроводника и уменьшает сопротивление между коллектором $к$ и основанием o); в) плоскостной триод; г) обозначение транзистора в схемах; д) характеристики триода — зависимость напряжения на коллекторе от тока эмиттера.

редь, вызовет новое уменьшение тока в лампе 1, т. е. поддержит первоначальное изменение (положительная обратная связь). Ток в лампе 1 быстро упадет до нуля, а в лампе 2 возрастет до максимума.

При подаче короткого импульса напряжения (отрицательного) на сетки обеих ламп запертая лампа не изменит своего состояния (ток через нее не изменится, он был и останется равным нулю), отпертая лампа с приходом на

сетку отрицательного импульса среагирует уменьшением тока, что вызовет повышение напряжения на ее аноде и далее — всю последовательность описанных выше явлений в цепи положительной обратной связи, которая приведет к переходу триггера в новое состояние равновесия. Лампы поменяются ролями. Та, которая проводила ток, окажется

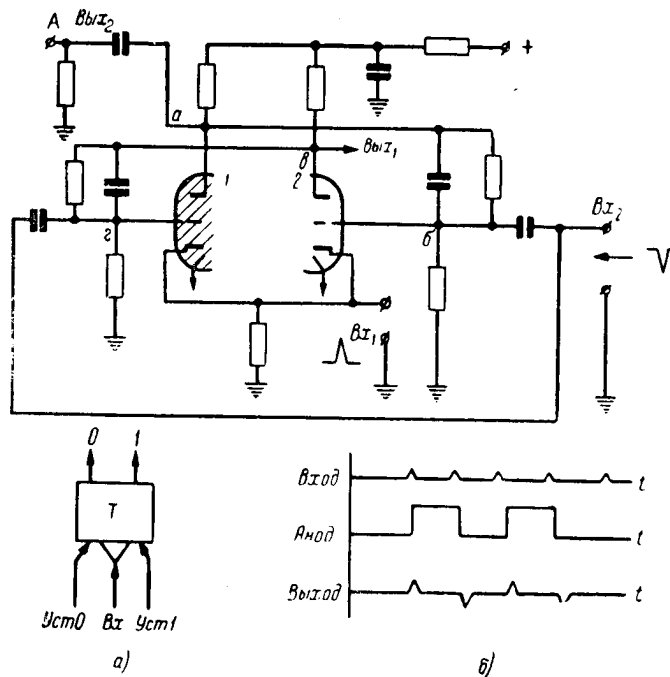


Рис. 7.18. Схема триггерной ячейки.

а) обозначение триггерной ячейки на схемах; б) временная диаграмма напряжений на входе триггера, на его аноде *a* и на выходе (после прохождения дифференцирующей ячейки).

Два усилителя на триодах 1 и 2 работают совместно. Выход первого *a* соединен со входом второго *b* и, наоборот, выход второго *b* соединен со входом первого *z*.

запертой, а запертая будет проводить ток. Следующий импульс снова возвратит схему в прежнее состояние и т. д.

Два устойчивых состояния триггера дают возможность сопоставить их с двумя символами двоичного алфавита, которые могут находиться в разряде числа, записанного в двоичной системе счисления, 0 и 1. Состоянию «0» соответствует отсутствие тока в первой лампе, «1» — отсутствие тока во второй лампе.

Если подавать на сетки ламп триггера серию импуль-

сов напряжения, то каждый нечетный импульс будет переводить триггер в состояние «1», а каждый четный — в состояние «0». Снимая напряжение с анода лампы *I* через дифференцирующую цепочку, мы получим на выходе отрицательный импульс всякий раз, когда триггер будет переходить из положения «1» в положение «0». Этот импульс можно использовать как единицу для переноса в соседний слева (старший) двоичный разряд и подать его на вход (сетки лампы) еще одного триггера.

Объединив в одну схему несколько триггеров, мы можем на этой схеме «записать» число, представленное символами двоичного счисления. Каждый триггер будет соответствовать разряду, состояние этого триггера «0» или «1» будет соответствовать 0 или 1 в этом разряде «записанного» числа. Цепочка триггеров, на которой таким способом фиксируется число, называется регистром. Цепочка триггеров, способная пересчитывать поступающие на ее вход импульсы, называется счетчиком. В сущности, счетчик и регистр работают так же, как конторские счета или, вернее, арифмометр с передачей единиц в старшие разряды, только с несравненно большей скоростью и в двоичной, а не в десятичной системе счисления.

Другим важным элементом электронной машины является «схема совпадения». Это устройство пропускает на выход импульс в том и только в том случае, если на оба его входа одновременно поступают импульсы. Один из этих импульсов может быть символом числа, а второй — управляющим сигналом, отпирающим схему совпадений и пропускающим импульс в нужную цепь. Схема совпадений выполняется либо на многосеточной электронной лампе, либо на диодах.

Схема совпадений может управляться выходным напряжением триггера и пропускать импульсы в линию, если триггер находится в состоянии «1», или не пропускать, если триггер находится в состоянии «0».

В нашу задачу не входит разбирать детали устройства и работы узлов машины. Однако мы должны сказать немного о некоторых из них, чтобы сделать хотя бы отчасти понятной структуру и работу машины в целом.

Если на регистре уже записано число и на него подается еще одно, то в результате в нем будет отложена сумма этих двух чисел. Так можно осуществлять сложение в устройстве, именуемом сумматором. Числа в сумматоре подаются поразрядно, т. е. в каждый разряд приходит только 0 (нет импульса) или 1 (есть импульс). Таб-

лица сложения для двоичного счисления до предела проста:

$$\begin{aligned} 0 + 0 &= 0; & 0 + 1 &= 1; \\ 1 + 0 &= 1; & 1 + 1 &= 10. \end{aligned}$$

Если в триггере регистра был 0, то приходящее число не опрокинет триггера, если оно 0, и опрокинет, если оно

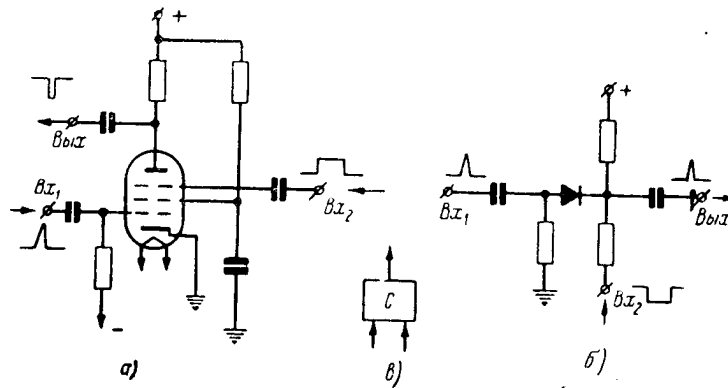


Рис. 7.19. Схема совпадений или вентильный каскад.

а) схема на пентоде (входные положительные импульсы поступают на первую и третью сетки лампы, выходной отрицательный импульс снимается с анода лампы); б) схема с использованием диода; в) обозначение каскада совпадений в схемах. Сигнал на выходе появляется только тогда, когда на обоих входах одновременно присутствуют входные сигналы.

1, и т. д. Самый сложный случай, когда складываются 1 + 1. В этом случае в триггере оказывается 0, а 1 с выхода пере-

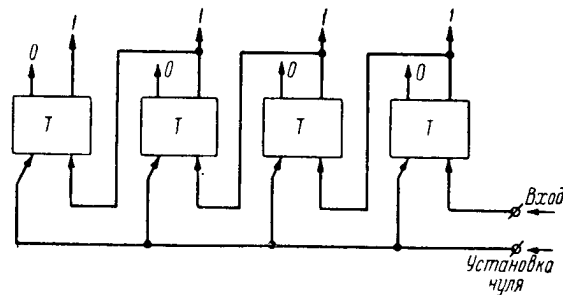


Рис. 7.20. Счетчик импульсов, состоящий из нескольких триггеров, включенных последовательно.

Выход каждого предыдущего триггера соединен со входом следующего. Наличие напряжения на выходах „0“ или „1“ каждого триггера дает в двоичной записи число поступивших на вход импульсов. Подачей импульса на шину УСТАНОВКА НУЛЯ все триггеры переводятся в нулевое положение.

дается в старший разряд (где она в свою очередь может сложиться с 1 и дать перенос в следующий разряд).

Умножение осуществляется последовательным сложением и сдвигом. Умножать при этом приходится только на 0 и 1. Двоичная таблица умножения такова:

$$\begin{array}{ll} 0 \times 0 = 0; & 1 \times 0 = 0; \\ 0 \times 1 = 0; & 1 \times 1 = 1. \end{array}$$

Умножение многоразрядных чисел, например, выглядит так:

$$\begin{array}{r} \times 25 \\ \hline 00 \\ 25 \\ \hline 250 \end{array} \quad [11001 = 1 \cdot 16 + 1 \cdot 8 + 0 \cdot 4 + 0 \cdot 2 + 1 \cdot 1 = 16 + 8 + 1 = 25; \\ 1010 = 1 \cdot 8 + 0 \cdot 4 + 1 \cdot 2 + 0 \cdot 1 = 8 + 2 = 10];$$

$$\begin{array}{r} \times 11001 \\ \hline 00000 \\ + 11001 \\ 00000 \\ 11001 \\ \hline 1111010 \end{array}$$

Множимое откладывается или не откладывается в регистр в зависимости от того, стоит в разряде единиц множителя 1 или 0; затем множимое сдвигается на разряд вправо и тем же путем умножается на следующий разряд множителя и т. д. В регистре-сумматоре происходит при этом сложение как обычно.

Вычитание производится в электронных машинах чаще всего как сложение с числом, дополнительным до самого большого, причем единица самого старшего разряда отбрасывается.

В десятичной системе

$$\begin{array}{r} 45 - 22 = 23 \\ - 45 \\ \hline 22 \\ - 22 \\ \hline 23 \end{array}$$

или

$$45 + (100 - 22) - 100 = 45 + 78 - 100 = 23;$$

$$\begin{array}{r} + 45 \\ + 78 \\ \hline 123 \\ - 100 \\ \hline 23 \end{array}$$

В двоичной системе:

$$0101101 - 0010110 = 0010111; \quad [45 - 22 = 23];$$

$$\begin{array}{r} 0101101 \\ - 0010110 \\ \hline 0010111 \end{array}$$

или

$$0101101 + (1000000 - 0010110) - 1000000 =$$

$$= 0101101 + 0101010 - 1000000 = 0010111;$$

$$[45 + (64 - 22) - 64 = 45 + 42 - 64 = 23];$$

$$\begin{array}{r} + 0101101 \\ + 0101010 \\ \hline 1010111 \end{array} \quad \begin{array}{r} - 1010111 \\ - 1000000 \\ \hline 0010111 \end{array}$$

Деление может производиться как повторное вычитание со сдвигом разрядов.

Могут быть применены и другие приемы вычислений, но в них не будем сейчас вникать.

Числа, передаваемые в виде электрических импульсов по проводам от одного узла машины к другому, могут быть представлены различным образом. Во-первых, они могут посылаться последовательно во времени разряд за разрядом по одному проводу; так делается в так называемых машинах последовательного действия. Во-вторых, они могут посылаться одновременно всеми разрядами, но каждый разряд по своему собственному проводу; так делается в машинах параллельного действия. Первые содержат меньше шин, но требуют больше времени для передачи чисел и потому медленнее работают. Вторые работают быстрее, но имеют более сложную схему коммутации.

Все импульсы, идущие по разным шинам и приходящие в различные узлы машины, должны точно совпадать по времени. Работа машины протекает в строгом ритме, задаваемом генератором тактовых или синхронизирующих импульсов. Из гребенки этих импульсов и «нарезается» число пропусканьем или непропусканьем очередного импульса на соответствующую шину.

Следует упомянуть о том, что числа в машине могут быть представлены различными способами. В регистре, имеющем, например, 40 двоичных разрядов, можно каждый разряд регистра закрепить за определенным разрядом числа. В этом случае, скажем, разряд единиц будет всегда

фиксироваться на одном и том же триггере. Такой способ записи называется записью с фиксированной запятой. Его преимуществом является простота выполнения элементарных операций сложения и вычитания чисел, которые производятся так, как описано выше. Недостатком этого способа записи является то, что он ограничивает число значащих разрядов и вынуждает прибегать к округлениям. Другим способом записи является запись с пла-

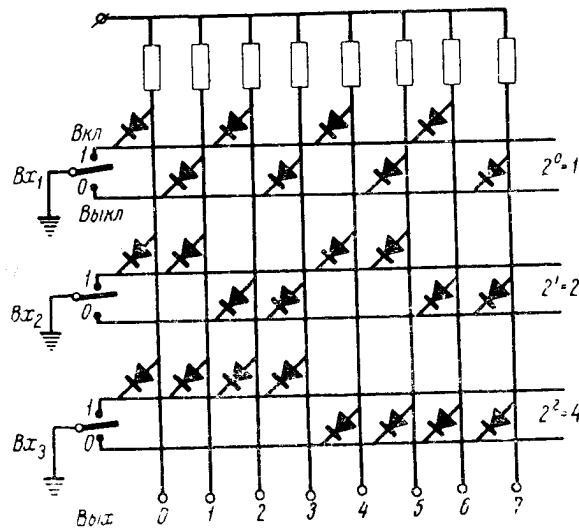


Рис. 7.21. Схема преобразователя.

Преобразователь выдает напряжение на выходной шине с номером, равным числу, поданному на входы в виде двоичной поразрядной записи. Например, перевод всех входных переключателей в верхнее положение (положение 1) вызывает появление напряжения только на выходной шине 7. Схема может служить для преобразования двоичных чисел в восьмеричные.

в а ю щ е й з а п я т о й. При этом записывается максимальное число разрядов независимо от порядка числа, причем первая значащая цифра ставится сразу после запятой, а порядок числа записывается отдельно. В этом случае перед сложением двух чисел требуется дополнительная операция «выравнивания порядков», что является недостатком этого способа. Его достоинство — в большой точности записи.

Следующая таблица дает представление об этих двух формах записи различных чисел. Для простоты число разрядов взято равным 6 и числа записаны в десятичной системе счисления.

Число	Запись с фиксированной запятой	Запись с плавающей запятой
3571,25	3571,25	$0,357125 \cdot 10^4$
357,125	0357,18	$0,357125 \cdot 10^3$
35,7125	0035,71	$0,357125 \cdot 10^2$
3,57125	0003,57	$0,357125 \cdot 10^1$
0,357125	0000,36	$0,357125 \cdot 10^0$
0,00357125	0000,00	$0,357125 \cdot 10^{-2}$

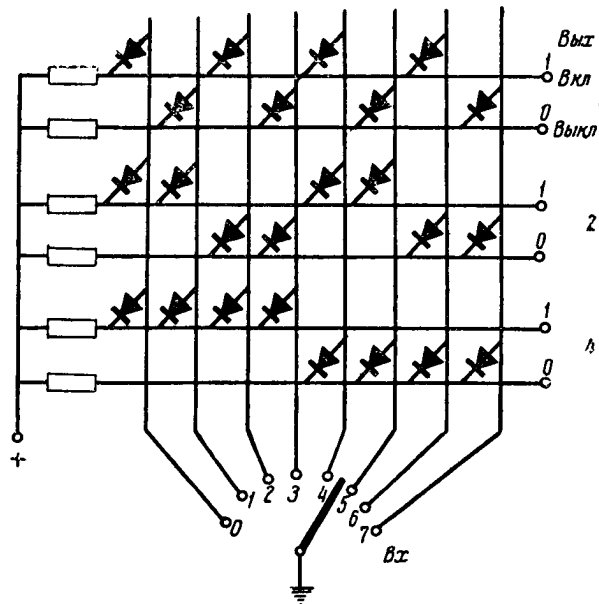


Рис. 7.22. Схема преобразователя, обратного приведенному на рис. 7.21.

Установка входного переключателя в одно из восьми положений вызывает появление на выходных шинах напряжений, изображающих в двоичной записи номер положения переключателя.

Управление и память

Электронные машины позволяют вести расчеты с огромной быстротой не только потому, что скорость вычислений в них высока, но еще и потому, что в них все промежуточные операции, необходимые при решении задачи, производятся автоматически с большой скоростью. К числу таких операций относятся: запись промежуточных результатов.

татов вычислений, подача исходных данных и промежуточных результатов для последующих вычислений, установка порядка операций, выбор необходимых данных и т. д. Все решение задачи от начала до конца идет самостоятельно, без участия человека, и с большой скоростью.

Такая полностью автоматическая работа достигается в результате того, что схема машины с каждым тактом ее работы изменяется, в ней производятся переключения с помощью устройств типа схемы совпадений, которые заставляют машину переключаться с одного действия на другое. Переключения производятся специальным управляющим устройством, встроеным в машину и работающим по заданной ему оператором программе. Управляющее устройство командует всеми узлами рабочей части машины, заставляя их посредством переключений в очень короткое время перестраиваться для выполнения новой операции.

Наличие управляющего устройства, которое воздействует на вентили в различных узлах машины, приводит к тому, что машина оказывается способной выполнять различные операции. Для каждой операции управляющее устройство меняет схему машины, отпирая одни и запирая другие шины. Оно, таким образом, воздействует на структуру машины, изменяя ее для каждой операции. При этом установленная данной командой управляющего устройства схема операции однозначно соответствует этой операции и только ей. Команда управляющего устройства представляет собой такие же импульсы напряжения, как и те, которые представляют числовые данные; поэтому команды могут быть также записаны условно в виде числа в двоичной системе счисления. Команды передаются по особым шинам, но хранятся в записанном виде в таких же устройствах памяти, в каких хранятся и числовые данные задачи.

Кроме арифметического устройства и устройства управления, в машине обязательно имеются устройства «памяти» или запоминающие устройства, а также устройства ввода и вывода данных.

Запоминающее устройство обычно бывает двух видов: оперативное (или внутреннее) и внешнее. Первое обладает небольшой емкостью, но является быстродействующим; записать число и снять число можно за короткое время. Второе обладает большой емкостью, но требует большего времени для считывания. Требования быстродействия и большой емкости технически противоречивы, и потому приходится идти на такое разделение устройств памяти.

Для того чтобы записать число в двоичной системе, т. е. серию разрядов, в каждом из которых либо присутствует, либо отсутствует импульс напряжения, можно использовать различные физические процессы. Это можно сделать на ламповых схемах-регистрах, состоящих из цепи триггеров (по одному на разряд числа). Ламповые регистры можно заменить катушками и сердечниками из магнитного

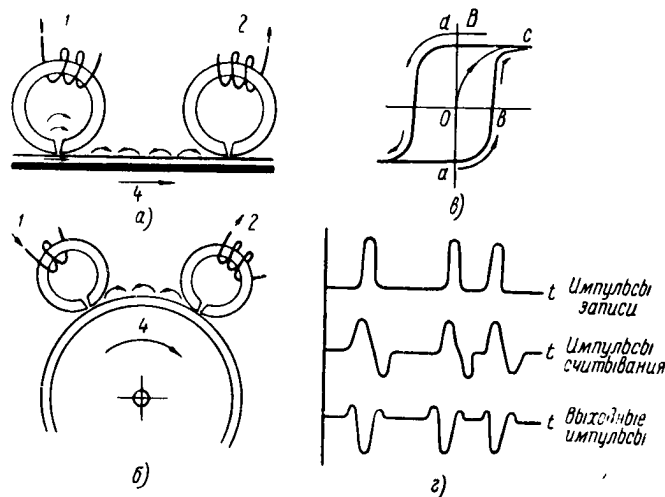


Рис 7.23. Устройство памяти машины — запись двоичных чисел на магнитную ленту.

а) схема записи на ленту, б) схема записи на барабан, в) характеристика намагничивания, г) импульсы записи, считывания и выходные импульсы после прохождения формирующего каскада.

Головка записи 1 — электромагнит с узкой щелью, перед ней движется магнитная лента или барабан 4; изменения тока в обмотке головки переманичивают участки ленты. Намагничивание участков ленты сохраняется и после прекращения воздействия магнитного поля головки вследствие большой коэрцитивной силы материала ленты (прямоугольная характеристика намагничивания). Головка считывания 2 подобна по устройству головке записи; при прохождении магнитной ленты с намагниченными участками мимо щели в обмотке возникает ток, служащий выходным сигналом считывания.

материала; после пропускания через катушку тока сердечник остается намагниченным и фиксирует импульс. Намагничение можно использовать в форме записи на ленту типа магнитофонной или на барабан, покрытый намагничивающейся пленкой. Наконец, импульсы можно фиксировать в виде зарядов на диэлектрике, на который в вакууме падает узкий электронный пучок. Последнее можно делать в обычной электронно-лучевой трубке, подобной кинескопу, стоящему в телевизорах, записывая точки вместо единиц на экране, где они сохраняются достаточно дол-

го. Можно, наконец, регистры выполнять на электромеханических реле.

Для запоминания сигналов применяются ультразвуковые устройства — линии задержки. Это — трубки, наполненные жидкостью (например, ртутью), или металлические стержни, на концах которых в контакте с ними расположены пластинки пьезоэлектрического материала (например, кварца). Прикладывая к пьезоэлектрическому материалу электрическое напряжение — импульс, мы получим

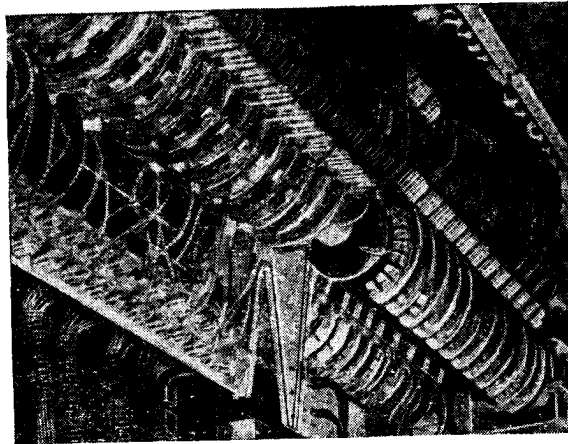


Рис. 7.24. Общий вид устройства памяти электронной вычислительной машины — магнитного барабана. (Видны ряды магнитных головок записи и воспроизведения сигнала.)

механический толчок, который передастся жидкости или стержню и будет распространяться в нем в виде звуковой волны. (Обычно используется не импульс постоянного тока, а импульс колебаний высокой — ультразвуковой — частоты). Поскольку звук имеет относительно небольшую скорость, волна доходит до конца линии задержки через заметный промежуток времени (доли миллисекунды). На конце линии волна падает на приемный пьезоэлектрический преобразователь, где снова превращается в электрический импульс. Выходной импульс попадает снова на вход той же линии, предварительно пройдя узел, восстанавливающий его форму. В результате импульс, вернее серия импульсов, будет как угодно долго циркулировать по линии задержки, периодически появляясь на ее выходе,

откуда число может быть снято и послано в регистр арифметического устройства.

Каждая запись в запоминающем устройстве должна быть сделана в определенном месте так, чтобы она не мешала предыдущим и последующим записям и чтобы записанное

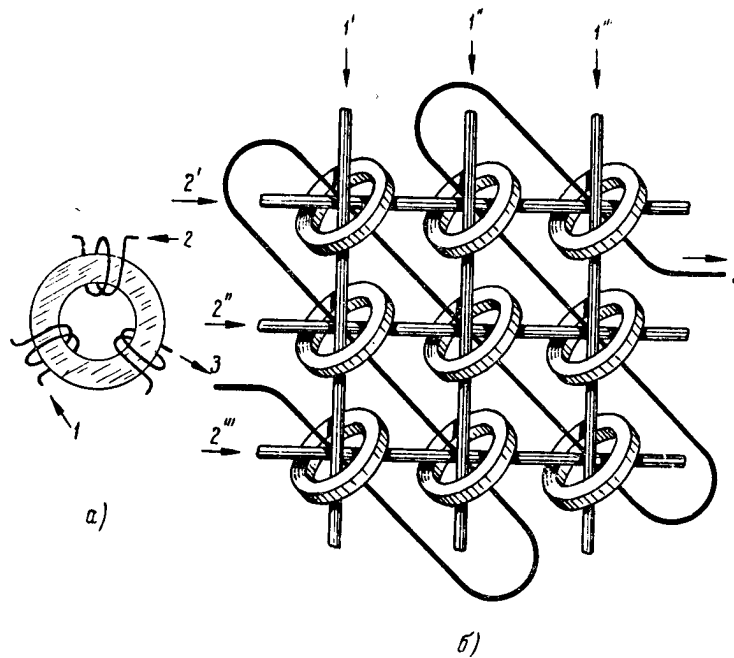


Рис. 7.25. Устройство памяти машины на магнитных сердечниках (ферритовых торах).

а) ячейка для запоминания одного двоичного разряда; б) матрица для запоминания чисел (1, 2—входные обмотки, 3—выходная обмотка). При подаче напряжений на одну пару входных проводов 1 и 2 и протекании в них тока тор, стоящий на пересечении этих проводов, намагничивается. При вторичном опросе этого элемента импульс тока в выходном проводе 3 появляется только в том случае, если соответствующий тор был предварительно намагничен.

можно было отыскать. Говорят, что записанное число имеет адрес, условный номер того места, где оно записано.

Адрес—это номер строки и столбца на квадратной таблице точек, записанных на экране электронно-лучевой трубки, а также номер трубки, если их несколько, номер лампового, магнитного или релейного регистра, номер отрезка магнитной ленты и номер дорожки при записи с помощью магнитофона и магнитного барабана, номер линии

задержки и номер серии импульсов («слова») из числа циркулирующих в этой линии и т. д.

Очевидно, что необходимы коммутирующие устройства, которые, получая на вход команду (адрес числа), под действием этой команды производили бы нужные переключения для выдачи числа из указанного командой адреса в арифметическое устройство. Такие коммутаторы устроены не очень сложно. В машине параллельного действия,

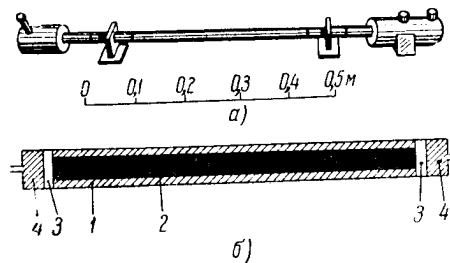


Рис. 7.26. Устройство памяти — ртутная ультразвуковая линия задержки.

а) общий вид, б) схема (1—металлическая труба, 2—ртуть, 3—кварцевые пьезоэлектрические преобразователи, 4— входные электроды). Электрические импульсы ультразвуковой частоты прикладываются ко входному электроду. Кварцевая пластина преобразует электрические колебания в механические, которые передаются ртути и проходят по линии в виде ультразвуковых волн. Дойдя до приемного кварца, акустические волны снова преобразуются приемным кварцем в электрические колебания, которые снимаются с выходного электрода на усилитель. Время распространения звуковых колебаний в ртути является временем задержки.

где каждый разряд числа передается по отдельной шине, можно свести командные шины на диодный переключатель. При подаче на входные шины «слова», изображающего число, на выходной шине, имеющей номер, обозначенный этим числом, появляется напряжение. Выходная шина соединяется с регистром соответствующего номера и открывает его выходные вентили на кодовые шины арифметического устройства. Посылая по одним и тем же входным шинам разные числа, мы будем получать один и тот же выходной сигнал на разных выходных шинах, открывающих разные регистры.

Можно использовать этот диодный переключатель «наоборот», если выходные шины сделать входными. Подавая напряжение на одну из входных шин, мы получаем на выходных шинах комбинацию сигналов, дающую номер возбужденной шины.

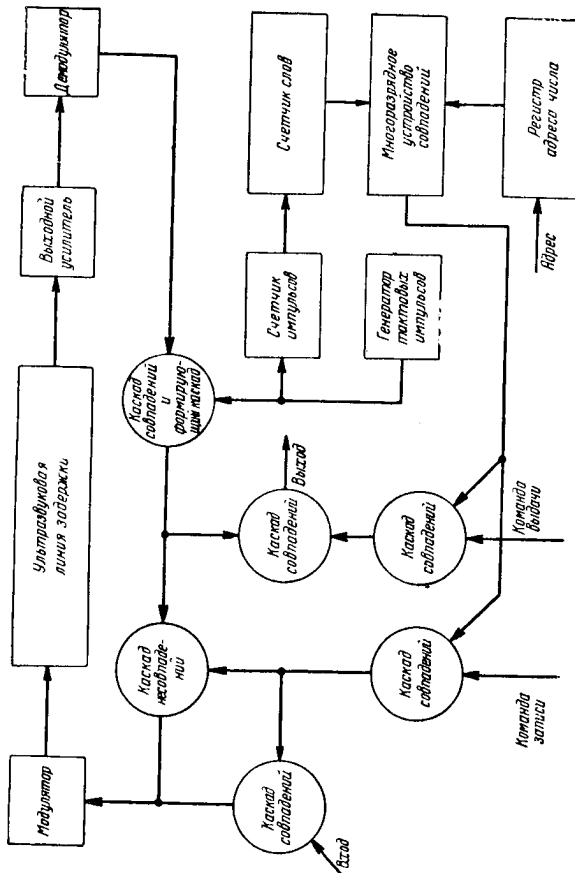


Рис. 7.27. Схема циркуляционной памяти машины с ультразвуковой линией задержки. Числа или «слова», записанные в памяти, непрерывно циркулируют в замкнутом кольце; появившись на выходе линии, импульс претерпевает регенерацию (исправление формы, которая несколько ухудшается при прохождении линии) и снова подается на вход линии. Прохождение каждого импульса и каждого слова, состоящего из определенного количества разрядов, отмечается и фиксируется счетчиком импульсов. Для эписи или считывания слова в регистр подается адрес (номер по порядку эписи) того места, с которого должно производиться считывание или на которое должна быть произведена запись. Кроме того, подается команда записи или считывания—импульс на соответствующий вход. При совпадении чисел в счетчике и в регистре адреса устройство совпадений выдает на выход импульс, который пропускает команду через входные каскады команд. При записи числа каскад совпадений не пропускает на вход линии число, записанное на том месте, на котором должно быть записано новое число, и таким образом старое число «забывается» машиной.

Подобные же устройства могут использоваться для декодирования чисел из двоичной записи, например, в восьмеричную и обратно.

В тех случаях, когда в устройствах памяти слова изображаются последовательным расположением символов во времени и сами слова расположены во временной последовательности, как в линии задержки или на ленте магнитофона, прихо-

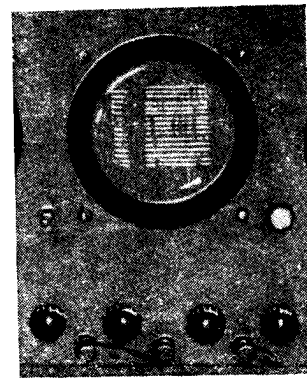
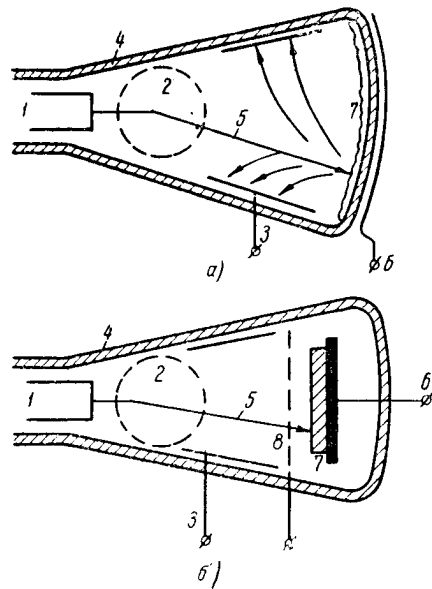


Рис. 7.28. Устройство памяти машины на электронно-лучевых трубках.

а) запись импульсов на экране кинескопа. б) запись с помощью потенциалоскопа. в) вид экрана контрольной трубки с воспроизведением записанных импульсов. 1—электронная пушка, создающая узкий пучок электронов, 2—отклоняющая система, направляющая пучок электронов в определенную точку экрана трубки, 3—коллектор—электрод, собирающий рассеянные электроны, 4—стеклянная колба, 5—электронный луч, 6—входной электрод, 7—диэлектрик, на котором сохраняются электрические заряды, нанесенные при эписи. Электронный луч, падающий на диэлектрик, выбивает из него вторичные электроны, которые уходят на коллектор или возвращаются на поверхность диэлектрика в зависимости от потенциала диэлектрика. Подавая различные потенциалы на входную пластину 6, можно получить различную величину заряда и потенциала в точке падения пучка. Заряд в этой точке сохраняется продолжительное время. При свитывании пучок направляется в ту же точку; на входную пластину подается постоянный потенциал; ток коллектора меняется в зависимости от заряда, который сохранился в данной точке диэлектрика. Запись можно периодически останавливать, чтобы воспрепятствовать ее стиранию с течением времени.

дится ожидать, когда подойдет нужное слово, чтобы вывести его на шины из запоминающего устройства. В этих случаях на магнитной ленте имеется отдельная дорожка с записью синхронизирующих импульсов, которые пересчитываются счетчиком. Адрес искомого числа по-

дается в специальный регистр адреса и сравнивается на каскадах совпадения с очередным номером в счетчике. При совпадении число выдается, так как по сигналу совпадения отпираются вентили и сигнал выдается на шины.

Время ожидания подхода нужного числа не очень велико на линиях задержки, несколько больше — на магнитном барабане и довольно велико — при записи на ленту. Самое короткое время выборки — у запоминающего устрой-

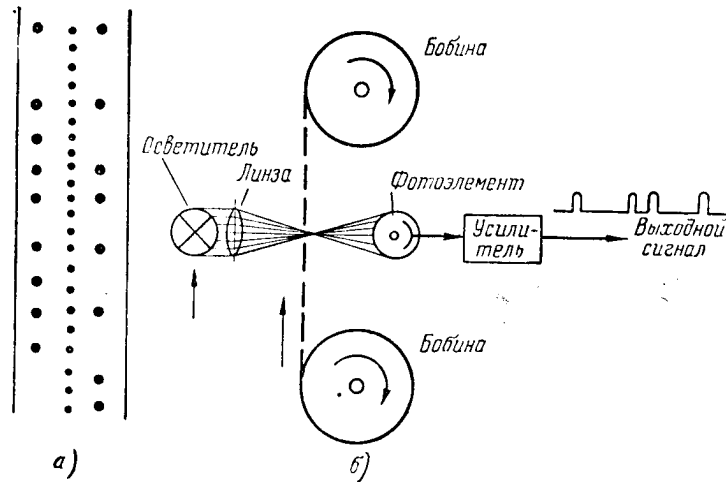


Рис. 7.29. Ввод данных в машину с помощью перфорированной ленты.

а) лента с записью в двоичной системе входных данных (чисел или команд); б) ввод данных в машину.

ства на электронно-лучевых трубках; оно равно времени установки электронного луча на нужную точку экрана, что достигается подачей напряжений определенной величины на отклоняющие электроды трубки. В машинах параллельного действия на каждый разряд имеется своя трубка. Поэтому адрес числа одинаков для всех трубок и выдается на них одновременно. Для быстродействующей оперативной памяти часто используют этот способ запоминания.

Оперативное и внешнее запоминающие устройства обмениваются зафиксированными в них данными. Готовые результаты вычислений передаются на магнитную ленту с трубок, а с ленты на трубки переписываются новые входные данные.

Ряд мест записи, или, как говорят, ячеек памяти, используется для записи команд управления машиной. Команды расположены в порядке номеров; очередная команда передается в регистр команд, отсюда в управляющее устройство для исполнения.

В различных типах машин используются различные виды команд. Стандартная команда определяет одно действие машины и вызывает все необходимые для этого дей-

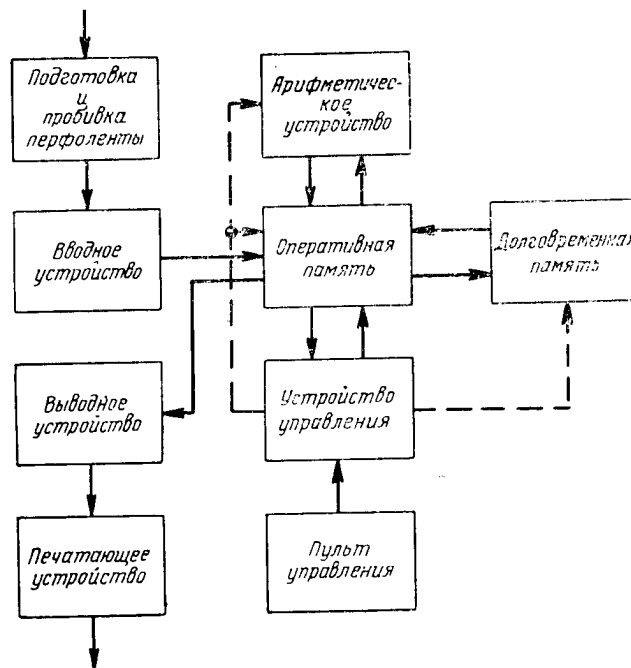


Рис. 7.30. Общая схема электронной вычислительной машины.

ствия переключения (отпирание и запираание коммутирующих узлов — вентилях). Под одним действием можно подразумевать более или менее сложный процесс. Так, если считать одним действием, например, передачу числа из запоминающего устройства в арифметическое, то команда для выполнения этого действия должна содержать указание (или номер) операции и адрес числа. Такие команды, и машины, работающие с ними, называются одноадресными. Для того чтобы сложить два числа и записать результат, нужно три команды, но сами команды короткие.

Если под одним действием понимать весь процесс сложения и записи, то команда должна содержать номер операции и три адреса: первые два — откуда взяты слагаемые и третий — куда записать сумму. Эта система применяется в трехадресных машинах.

Все команды в последовательности их выполнения составляют программу работы машины над данной задачей. Программа составляется математиком-программистом и записывается в виде слов в алфавите машины, т. е. в виде чисел двоичного счисления на перфоленте или перфокартах. Затем программа вводится во внешнее запоминающее устройство машины (уже автоматически) и выполняется в порядке следования команд после запуска машины. Последней командой машины будет «остановка».

Таким же путем вводятся и исходные данные для вычислений. Результаты вычислений выводятся из машины печатающей автоматической машиной (телетайпом) или же с помощью быстродействующего фотопечатающего устройства.

Особенности электронных машин

Электронная вычислительная машина выполняет более широкий круг операций по сравнению с арифмометром и даже со счетно-аналитическими машинами. Она работает как расчетное бюро, выполняя не только все вычисления, но и все операции записи, сортировки и считывания промежуточных результатов, которые на клавишных машинах еще делает оператор.

Подобно тому, как оператору-расчетчику дается программа обработки исходных данных задачи и таблица для записи результатов, электронной машине также задается программа действий, которую она строго выполняет. Эта программа составляется несколько детальнее, чем для вычислителей. Машина выполняет программу несравненно быстрее. За несколько часов работы она сделает больше, чем вычислитель за всю свою жизнь. Машина, пока она исправна, не допускает ошибок. Ошибки появляются только при поломках машины или при наличии ошибок в программе. Для снижения частоты поломок принимается ряд эффективных мер: предварительный отбор деталей, работа в ослабленных режимах, периодические проверки точности с помощью специальных «испытательных» задач — тестов и т. д.

Машины сверхвысокой производительности не просто заменяют труд многих вычислителей. Несомненно, что

использование машин дает качественно новые возможности, которые заключаются в том, что с помощью машин могут ставиться и решаться задачи такого объема и сложности, решение которых при ручном счете было бы практически невозможно, хотя бы из-за трудностей кооперирования большого числа вычислителей, обмена информации между ними и согласования их действий при совместной работе над одной задачей.

Решающее значение в расширении функций электронных машин сыграло введение в их состав быстродействующей памяти и операций автоматической записи и отыскания слов в ней. В процессе обработки информации машина, использующая запоминающие устройства, т. е. обладающая памятью, реагирует не только на приходящие в данный момент сигналы, но и на те, которые приходили в прошлом. Такая машина способна накапливать сведения и использовать их по мере надобности. Эта способность не только чрезвычайно сильно расширяет возможности машин, но придает машинам совершенно новое качество по сравнению с автоматами, реагирующими стереотипно всякий раз на существующее в данный момент воздействие. Наличие памяти в принципе дает возможность менять реакцию по мере накопления информации. Такой способностью простейшие автоматы обладают лишь в начаточном состоянии в том случае, если управление производится с интегрированным сигналом.

Введение в машины долговременной памяти, даже с той примитивной ее организацией, какая используется в машинах сейчас, позволяет имитировать на машинах, например процесс обучения, бывший характерным до сего времени только для живых организмов.

Одним из наиболее важных свойств электронных машин является также их способность до некоторой степени самостоятельно изменять собственную программу действий. Дело в том, что расписывать подробно ход за ходом длинную программу решения сложной задачи представляется очень долгой и трудоемкой работой. Поэтому программисты стремились сократить этот труд и добились в этом отношении заметных успехов. Для большинства математических задач ход вычислений разбивается на повторяющиеся серии одинаковых операций — циклы. Очевидно, нет смысла выписывать цикл за циклом одни и те же команды. Можно поручить эти обязанности той же машине, дав ей для этого соответствующую программу. Поскольку команды записываются так же, как числа, с ними можно произ-

водить преобразования как с числами, например после выполнения команды преобразовать ее, прибавляя по единице к адресам для того, чтобы по очереди перемножить столбики чисел, размещенных в ячейках памяти в порядке номеров.

Особенно интересный и принципиально очень важный результат в этом отношении дает применение команды условного перехода. Эта команда не однозначна, она имеет два значения: *A* — *Перейти к выполнению команды № A* и *B* — *Перейти к выполнению команды № B*. Выбор варианта *A* или *B* зависит от исхода специальной логической операции сравнения двух чисел x и y . Если $x > y$, выбирается вариант *A*, если $x \leq y$ — вариант *B*. Сравнение проводится машиной в конце цикла вычислений, и результат сравнения уже однозначно определяет выбор *A* или *B*.

Так, если при вычислении очередного члена ряда в итоге цикла вычислений получается величина больше заданной минимальной, машина возвращается к началу цикла (команда *A*) и вычисляет следующий член. Если очередной член ряда оказывается меньше заданной минимальной величины, машина, не вычисляя следующего члена, переходит к новой операции, определяемой командой *B*. Условный переход представляет собой отклонение от принципа программного управления и переход к управлению с обратной связью. В самом деле, при условном переходе действия машины (следующая операция) зависят не только от внешнего влияния — программы, но и от результатов работы машины. Правда, способ введения обратной связи в этом случае сложнее, чем в случае следящих систем, и отличается от линейной связи в последних. Однако в принципе — это обратная связь (зависимость действия от результата предыдущего действия).

Используя в программе условный переход, программист может заранее не знать и не интересоваться тем, сколько повторных циклов (итераций) проделает машина до получения результата нужной точности. Машина без участия оператора и без предварительного плана сама решит по заложенному в нее правилу и исходя из обстановки, когда ей перейти к новому циклу операций. Не исключен и такой случай, когда сама же машина может составить себе это правило перехода на основании других, более общих указаний. Таким образом, с использованием в программе условных переходов машина приобретает некоторую самостоятельность в выполнении простых

операций, подчиняясь указаниям программы лишь более общего характера.

Введение операции выбора на основе сравнения представляет собой, как мы уже сказали, не арифметическую, а логическую операцию. Выбор — не единственная операция логики, которую выполняют машины. Выполнение логических операций наряду с математическими — это еще одно очень интересное и важное свойство счетных машин. Операцию выбора продельвает и счетно-аналитическая сортировальная машина.

Преобразование команд и условные переходы являются очень важными принципиальными особенностями электронных цифровых машин, которые открывают много совершенно новых возможностей их использования.

Какие же операции логики и как выполняют машины? Для ответа на этот вопрос следует сказать несколько слов о теоретической логике и алгебре релейных схем.

Теоретическая логика и алгебра релейных схем

Теоретическая логика представляет собой ветвь математических наук, объектом изучения которой уже являются числа. Теоретическая логика разбивается на несколько разделов: исчисление высказываний, исчисление предикатов, исчисление классов и т. д. Предметом изучения теоретической логики являются средства, с помощью которых из заданных логических предложений, выражающих связь между объектами, можно получить новые предложения. Теоретическая логика строится как математическая дисциплина, т. е. с точной формулировкой основных положений и правил и введением формально-математических выражений для их записи.

Один из простейших разделов теоретической логики — исчисление высказываний — находит весьма полную аналогию в соотношениях, существующих для схем, состоящих из электрических контактов и соединяющих их проводов (релейно-контактных схем). Эта аналогия используется двояко: во-первых, математические формализмы теоретической логики используются для исследования свойств релейных схем, как рабочий аппарат для анализа и синтеза схем для технических применений; во-вторых, релейные схемы, построенные по логическим формулам, позволяют автоматически решать логические задачи.

Несомненно, что область аналогий между положениями теоретической логики и решающими схемами не ограничит-

ся исчислением высказываний. Усложнение вычислительных схем приводит и приведет в дальнейшем к расширению соответствующего логического аппарата, подлежащего использованию в машинной логике.

В исчислении высказываний изучаются простейшие утверждения или высказывания, каждое из которых может быть либо ложным, либо истинным. Например, такими высказываниями могут быть:

X = Дверь в лифт закрыта,

Y = Пассажир находится в лифте,

Z = Мотор вращается,

V = На фотоэлемент падает свет,

W = Кнопка электрического контакта нажата и т. д.

Из этих простых высказываний посредством сопоставления их друг с другом в различных связях могут быть построены новые высказывания, которые также могут быть либо истинными, либо ложными.

Принимаются следующие связи между высказываниями:

1) Конъюнкция $X \& Y$ (читается: *икс и игрек*). Например, *Дверь в лифт закрыта, и пассажир находится в лифте*. Сложное высказывание $X \& Y$ истинно в том и только в том случае, если истинны X и Y .

2) Дизъюнкция $X \vee Y$ (читается: *икс или игрек*). Например, *Кнопка контакта нажата или на фотоэлемент падает свет*. Высказывание $X \vee Y$ истинно тогда, когда истинно по крайней мере одно из высказываний X , Y (или оба вместе). Слово *или*, применяемое для обозначения дизъюнкции, не следует понимать в смысле исключения — *либо-либо* («либо пан, либо пропал»), а только в смысле *или X, или Y, или оба вместе*. Для исключаяющего *или* в смысле *либо-либо* можно составить особое выражение, более сложного вида.

3) Отрицание \bar{X} (читается: *не икс*). Не икс истинно, когда икс ложно, и наоборот. Например, X = *Мотор вращается* — ложно; тогда \bar{X} = *Мотор не вращается* — истинно.

4) Импликация $X \rightarrow Y$ (читается: *если икс, то игрек*). Например: *Если кнопка нажата, то мотор вращается*. Однако это соотношение не следует понимать в смысле причины и следствия, т. е. в смысле: *Если кнопка нажата, то тогда и только тогда мотор вращается*. Высказывание $X \rightarrow Y$ истинно всегда, когда X ложно или когда Y истинно. Оно ложно только тогда, когда X истинно, а Y ложно.

Можно ввести еще и другие элементарные связи между высказываниями.

Применяя основные связи несколько раз, можно образовать из данных высказываний все более сложные, например:

$$[(V \vee W) \& X \& Y] \rightarrow Z \quad \text{или} \quad \overline{(X \& Y \vee Z)} \quad \text{и т. д.} \quad (7.1)$$

Каждое такое сложное высказывание — логическая функция — может быть либо истинным, либо ложным в зависимости от комбинаций истинности и ложности входящих в него логических аргументов — независимых переменных X, Y, Z, V, W . Можно перечислить все комбинации значений аргументов и для каждого вычислить значение функции (в приведенном примере их всего $2^5 = 32$).

Значения сложного логического высказывания $\varphi(X, Y)$ — логической функции от двух аргументов — можно наглядно представить в виде квадратной таблицы. Два столбца этой таблицы соответствуют значениям независимого переменного X , а две строки — значениям Y . На пересечениях строк и столбцов помещено соответствующее значение функции $\varphi(X, Y)$.

Обычно истинное высказывание обозначается единицей, а ложное нулем в двоичном алфавите, поскольку значения каждого высказывания укладываются в этот алфавит.

Для конъюнкции, дизъюнкции и импликации таблицы значений выглядят следующим образом:

		$X \& Y$	
$Y \backslash X$		0	1
0		0	0
1		0	1

		$X \vee Y$	
$Y \backslash X$		0	1
0		0	1
1		1	1

		$X \rightarrow Y$	
$Y \backslash X$		0	1
0		1	0
1		1	1

Любую четырехзначную таблицу такого вида можно представить с помощью некоторого логического высказывания. Так, например, таблица сложения в двоичной системе счисления, которая записывается

$Y \backslash X$	0	1
0	0	1
1	1	0

совпадает с таблицей для логического выражения $X \vee Y \& \overline{(X \& Y)}$.

Двоичная таблица умножения совпадает с таблицей конъюнкций.

Вообще, любую функцию от k аргументов $W(X_1, X_2, \dots, X_k)$, если сама функция и каждый из её аргументов может принимать только два значения, обозначаемые 0 и 1 соответственно, можно задать в виде таблицы:

X_1	X_2	\dots	X_k	$W(X_1, X_2, \dots, X_k)$
0	0		0	$W(0, 0, \dots, 0)$
0	0		1	$W(0, 0, \dots, 1)$
\vdots	\vdots		\vdots	$W(\vdots, \vdots, \dots, 1)$.

Число значений данной функции от k аргументов (число строк в таблице) равно 2^k . Общее число всех возможных функций от k аргументов равно 2^{2^k} .

Любую функцию, заданную таблицей, можно выразить с помощью формулы, используя элементарные логические функции: конъюнкцию, дизъюнкцию и отрицание и, наоборот, любую формулу можно записать в виде таблицы.

Таблица, состоящая только из единиц или только из нулей задается так называемыми всегда истинными и всегда ложными высказываниями, примерами которых могут служить

$$X \& \bar{X} \equiv 0; \quad \overline{X \& \bar{X}} \vee \overline{Y \& \bar{Y}} \equiv 1 \text{ и т. д.} \quad (7.2)$$

Можно показать, что для составления всех сложных высказываний, которые получаются с применением отрицания, конъюнкции, дизъюнкции и импликации, достаточно только двух операций — отрицания и конъюнкции. Тогда, например, импликация выразится так:

$$X \rightarrow Y \equiv \bar{X} \vee Y, \quad (7.3)$$

а дизъюнкция

$$X \vee Y \equiv \overline{\bar{X} \& \bar{Y}}, \quad (7.4)$$

то нетрудно проверить по таблицам. Здесь знак равенства означает эквивалентность обоих выражений в смысле ложности и истинности. Обычно, однако, применяют дизъюнкцию для упрощения записи.

В других разделах логики — исчислении предикатов и исчислении классов — применяются формально те же отношения между логическими переменными, но в них вкладывается иной смысл. Интерпретация логических

связей в исчислении классов в ряде случаев помогает освоиться с формализмами теоретической логики.

В исчислении классов под переменной X понимают уже не высказывание, а класс или множество элементов или предметов, для которого истинно данное утверждение (предикат). При этом в рассмотрение вводятся также пустой класс, не содержащий элементов, и класс всех предметов, включающий все возможные элементы. Можно изобразить класс всех предметов в виде площади прямоугольника и класс X как часть этой площади, ограниченную, например, окружностью. Тогда выражение $X \& Y$ будет пересечением классов X и Y , т. е. площадью, принадлежащей обоим классам одновременно. Подобные толкования получают и другие логические выражения.

Иногда для обозначения конъюнкции принимают алгебраический знак умножения (\cdot), для дизъюнкции—сложения ($+$), и вводят алгебраический знак равенства ($=$)

$$X \& Y = X \cdot Y; \quad X \vee Y = X + Y. \quad (7.5)$$

Применяя эти обозначения, мы приходим к привычным и потому легко запоминающимся соотношениям:

$$\begin{array}{lll} X \& Y & X \vee Y & 0 + X = X \\ 0 \cdot 0 = 0 & 0 + 0 = 0 & 0 \cdot X = 0 \\ 0 \cdot 1 = 0 & 0 + 1 = 1 & 1 \cdot X = X \\ 1 \cdot 0 = 0 & 1 + 0 = 1 & X + Y = Y + X \\ 1 \cdot 1 = 1 & 1 + 1 = 1 & X \cdot Y = Y \cdot X \end{array} \quad (7.6)$$

$$\begin{array}{l} X + (Y + Z) = (X + Y) + Z \\ X(YZ) = (XY)Z \\ X(Y + Z) = XY + XZ \end{array}$$

Однако соотношение $1 + 1 = 1$ уже выходит за пределы привычных. Кроме того, в этих обозначениях верны также и следующие соотношения:

$$\begin{array}{l} X + X + X + X = X; \quad X \cdot X \cdot X \cdot X = X; \quad X + 1 = 1; \\ X + YZ = (X + Y) \cdot (X + Z) \text{ и т. д.} \end{array} \quad (7.7)$$

Эти соотношения не верны для обычной алгебры. Поэтому, во избежание недоразумений, логические связи лучше записывать без помощи алгебраических знаков, оставив только знак равенства.

Любое сложное логическое высказывание можно привести к так называемой «нормальной форме», предста

ляющей собой конъюнкцию членов, содержащих только дизъюнкции переменных и их отрицаний или же дизъюнкции конъюнкций членов. Приведение к нормальной форме производится последовательным применением эквивалентных преобразований вида:

$$X \& Y = Y \& X; X \vee Y = Y \vee X \text{ (коммутативность); (7.8)}$$

$$\begin{aligned} X \& (Y \vee Z) &= X \& Y \vee X \& Z; X \vee (Y \& Z) = \\ &= X \vee Y \& X \vee Z \text{ (дистрибутивность); (7.9)} \\ (X \& Y) \& Z &= X \& (Y \& Z); \end{aligned}$$

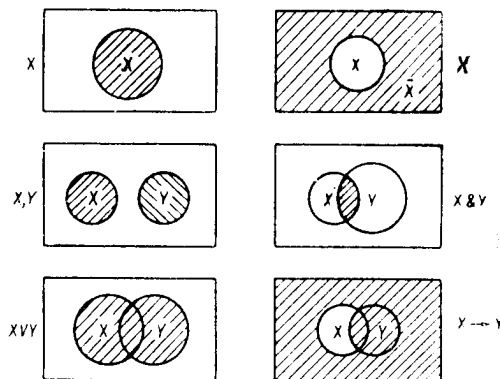


Рис. 7.31. Символическое изображение логических операций в исчислении классов.

$$X \vee (Y \vee Z) = (X \vee Y) \vee Z \text{ (ассоциативность); (7.10)}$$

$$\overline{\overline{X}} = X; \overline{\overline{\overline{X}}} = \overline{X}; \overline{X \& Y} = \overline{X} \vee \overline{Y}; \overline{X \vee Y} = \overline{X} \& \overline{Y}. \text{ (7.11)}$$

Обозначая сложное выражение W (в виде функции f) от аргументов $X_1, X_2, X_3, \dots, X_n$,

$$W = f(X_1, X_2, X_3, \dots, X_n),$$

можно [показать, что функция $f(X_1, X_2, X_3, \dots, X_n)$ равна ее разложению по одному или более переменным:

$$\begin{aligned} f(X_1, X_2, X_3, \dots, X_n) &= [X_1 \& f(1, X_2, X_3, \dots, X_n)] \vee \\ &\vee [\overline{X_1} \& f(0, X_2, X_3, \dots, X_n)], \end{aligned} \text{ (7.12)}$$

$$\begin{aligned}
f(X_1, X_2, X_3, \dots, X_n) &= [X_1 \& X_2 \& f(1, 1, X_3, \dots, X_n)] \vee \\
&\vee [X_1 \& \bar{X}_2 \& f(1, 0, X_3, \dots, X_n)] \vee \\
&\vee [\bar{X}_1 \& X_2 \& f(0, 1, X_3, \dots, X_n)] \vee \\
&\vee [\bar{X}_1 \& \bar{X}_2 \& f(0, 0, X_3, \dots, X_n)] \quad (7.13)
\end{aligned}$$

и т. д. или же

$$\begin{aligned}
f(X_1, X_2, X_3, \dots, X_n) &= [X_1 \vee f(0, X_2, X_3, \dots, X_n)] \& \\
&\& [\bar{X}_1 \vee f(1, X_2, X_3, \dots, X_n)] = \\
&= [X_1 \vee X_2 \vee f(0, 0, X_3, \dots, X_n)] \& \\
&\& [X_1 \vee \bar{X}_2 \vee f(0, 1, X_3, \dots, X_n)] \& \\
&\& [\bar{X}_1 \vee X_2 \vee f(1, 0, X_3, \dots, X_n)] \& \\
&\& [\bar{X}_1 \vee \bar{X}_2 \vee f(1, 1, X_3, \dots, X_n)] \quad (7.14)
\end{aligned}$$

и т. д.

Продолжая разложение и проводя его по всем переменным, получим выражение W в одной из нормальных форм.

Отметим еще логическую операцию, обозначаемую так называемым знаком Шеффера. Эта операция интересна тем, что все логические связи можно представить с помощью одной этой операции. Связь двух переменных X и Y с помощью знака Шеффера обозначается X/Y и записывается словами: X и Y несовместны. Эта связь эквивалентна отрицанию конъюнкции или дизъюнкции отрицаний

$$X/Y = \bar{X} \& \bar{Y} = \bar{X} \vee \bar{Y}. \quad (7.15)$$

В этой записи

$$X/X = \bar{X}; \quad (X/X)/(Y/Y) = \bar{X}/\bar{Y} = X \vee Y \text{ и т. д.}$$

Очевидно, если дизъюнкция и отрицание могут быть выражены с помощью только знака Шеффера, то то же можно сделать и со всеми остальными связями.

Релейные схемы

Реле с двумя контактирующими электродами может иметь лишь два состояния: включено и выключено, подобно высказыванию логики, которое может быть либо истинным, либо ложным. Логические связи между высказыва-

ниями могут быть успешно сопоставлены со схемами взаимных соединений между реле.

Сопоставление состояний реле с истинностью высказываний может быть двояким:

либо

1 = истина = включено (контакт)
0 = ложь = выключено (разрыв),

либо

1 = истина = выключено (разрыв)
0 = ложь = включено (контакт).

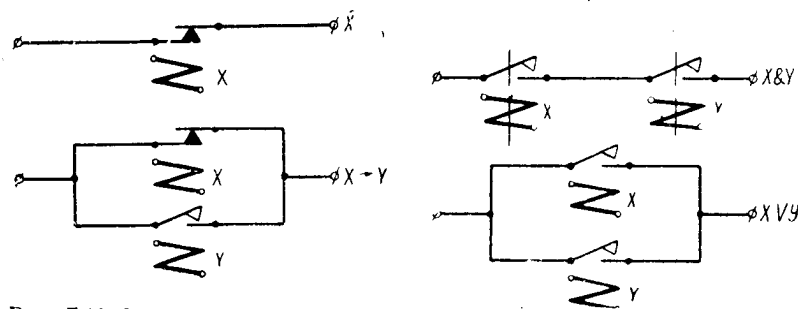


Рис. 7.32. Релейные схемы, изображающие логические операции.

Оба типа сопоставлений равноценны и равноправны. Они образуют так называемые взаимные схемы соединений, о чем мы скажем несколько слов позже.

Воспользовавшись, например, сопоставлением первого типа, мы приписываем каждому реле одно определенное высказывание ($X, Y, Z...$) и, соединяя реле друг с другом, получим схемы, эквивалентные сложным высказываниям (двухполюсники). Вся схема в целом от входа до выхода будет давать соединение или разрыв в соответствии с ложностью или истинностью сложного высказывания, которое она изображает, и в зависимости от включения (истинности) или выключения (ложности) реле — «аргументов».

В такой системе сопоставления конъюнкция изображается последовательным включением двух реле (см. рис. 7.32). Конъюнкция истинна тогда и только тогда, когда истинны оба высказывания. Последовательная цепь проводит тогда и только тогда, когда включены оба реле. Дизъюнкция изобразится параллельным включением кон-

тактов. Отрицание изображается контактом, работающим при включении реле (истина) не на замыкание, а на размыкание (на схемах изображается черным контактом).

Сформулировав и записав логические высказывания, можно перейти от них к релейной схеме.

Примем, например, следующие обозначения логических высказываний (и соответствующих им контактов):

D_1, D_2 — двери лифта на первом и втором этажах закрыты,

$D_н$ — дверь кабины лифта закрыта,

Π — пассажир находится в кабине,

K_1 — кнопка спуска в лифте нажата,

K_2 — кнопка вызова на первом этаже нажата.

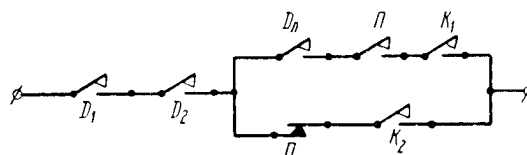


Рис. 7.33. Схема управления лифтом.

Сформулируем сложное высказывание, эквивалентное условию включения мотора лифта — *Вниз*:

(дверь на первом этаже закрыта и дверь на втором этаже закрыта и дверь кабины закрыта и пассажир в кабине и нажата кнопка спуска) или [дверь на первом этаже закрыта и дверь на втором этаже закрыта и пассажира нет в кабине и кнопка вызова нажата и (дверь в лифте закрыта или не закрыта)]:

$$M = (D_1 \& D_2 \& D_н \& \Pi \& K_1) \vee \vee [D_1 \& D_2 \& \bar{\Pi} \& K_2 \& (D_н \vee \bar{D}_н)]. \quad (7.16)$$

Преобразуем

$$M = (D_1 \& D_2) \& [(D_н \& \Pi \& K_1) \vee (\bar{\Pi} \& K_2)].$$

Схема, реализующая эту логическую формулу, приведена на рис. 7.33.

Другой пример. Необходимо тремя контакторами K_1 , K_2 , K_3 включать три исполнительных механизма M_1 , M_2 , M_3 в комбинациях, задаваемых таблицей

K_1	K_2	K_3	M_1	M_2	M_3
0	0	0	0	1	0
0	0	1	0	0	1
0	1	0	0	1	0
0	1	1	0	1	1
1	0	0	1	0	0
1	0	1	1	0	1
1	1	0	1	1	0
1	1	1	0	1	0

Для этих механизмов составляем по таблице выражения таким способом, чтобы для каждого состояния, когда механизм включен, получить конъюнкцию выражений коммутаторов, и затем выражения этого вида объединим знаками дизъюнкции.

$$M_1 = (K_1 \& \bar{K}_2 \& \bar{K}_3) \vee (K_1 \& \bar{K}_2 \& K_3) \vee (K_1 \& K_2 \& \bar{K}_3); \quad (7.18)$$

$$M_2 = (\bar{K}_1 \& \bar{K}_2 \& \bar{K}_3) \vee (\bar{K}_1 \& K_2 \& \bar{K}_3) \vee (\bar{K}_1 \& K_2 \& K_3) \vee (K_1 \& K_2 \& \bar{K}_3) \vee (K_1 \& K_2 \& K_3); \quad (7.19)$$

$$M_3 = (\bar{K}_1 \& \bar{K}_2 \& K_3) \vee (\bar{K}_1 \& K_2 \& K_3) \vee (K_1 \& \bar{K}_2 \& K_3). \quad (7.20)$$

Упрощая, получим:

$$\begin{aligned} M_1 &= K_1 \& (\bar{K}_2 \vee \bar{K}_3); \\ M_2 &= K_2 \vee (\bar{K}_1 \& \bar{K}_3); \\ M_3 &= K_3 \& (\bar{K}_1 \vee \bar{K}_2). \end{aligned} \quad (7.21)$$

По этим формулам составляем схему, приведенную на рис. 7.34.

Если учесть, что число возможных сложных выражений с ростом числа логических аргументов n возрастает как 2^{2^n} и уже для $n=6$ составляет $2,5 \cdot 10^{19}$, то станет ясно, что наличие строгого формального метода для проектирования сложных схем является весьма полезным.

Правда, аппарат исчисления высказываний не дает возможности проектировать на и лучшие схемы, например, в смысле наименьшего числа элементов, и для реше-

ния этой важной для техники задачи приходится прибегать к дополнительным приемам. В этой очень трудной области сделано уже много, но в ней еще имеется много нерешенных проблем.

Ясно, что в качестве коммутирующих элементов могут применяться электронные лампы. Приведем некоторые схемы с их применением в качестве примера. Уже знакомый нам клапан или каскад совпадений реализует логическую операцию конъюнкции. Собирательный каскад выполняет многократную дизъюнкцию. Реле ускорения, упоми-

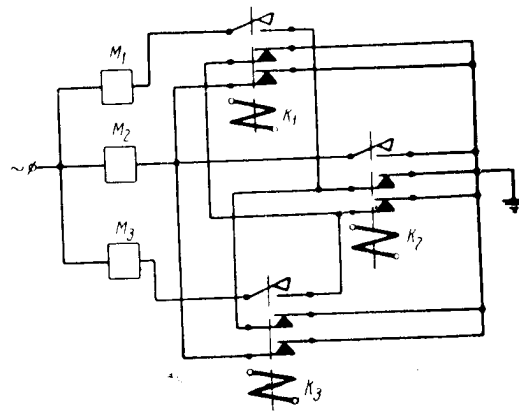


Рис. 7.34. Схема управления тремя механизмами M_1 , M_2 , M_3 с помощью трех контакторов K_1 , K_2 , K_3 .

навшему нами в гл. 2, выполняет операцию конъюнкции сигнала (*нужно*) от контроллера и отрицания сигнала (*нельзя*) от мотора.

Как мы уже говорили выше, можно вместо применявшегося нами до сих пор обозначения (1) для истинности применять не соединение, а разрыв контактов, для ложности (0) — не разрыв, а соединение. В этом случае конъюнкция изобразится параллельным, а дизъюнкция последовательным соединением контактов.

Две схемы, приведенные на рис. 7.36, эквивалентны. На левой схеме истина изображается замыканием контактов, на правой — размыканием

$$X \& \bar{Y} \vee Y \bar{U} = \xi \& \bar{\nu} \vee \eta \bar{\omega}. \quad (7.22)$$

Узловые точки схемы *I* становятся контурами схемы *II* и наоборот (включая внешний контур, показанный пунктиром). Замкнутые контакты схемы *I* (истина) превращаются в разомкнутые контакты схемы *II* (истина) и наоборот. Такие схемы называются «взаимно-обратными» или «взаимными».

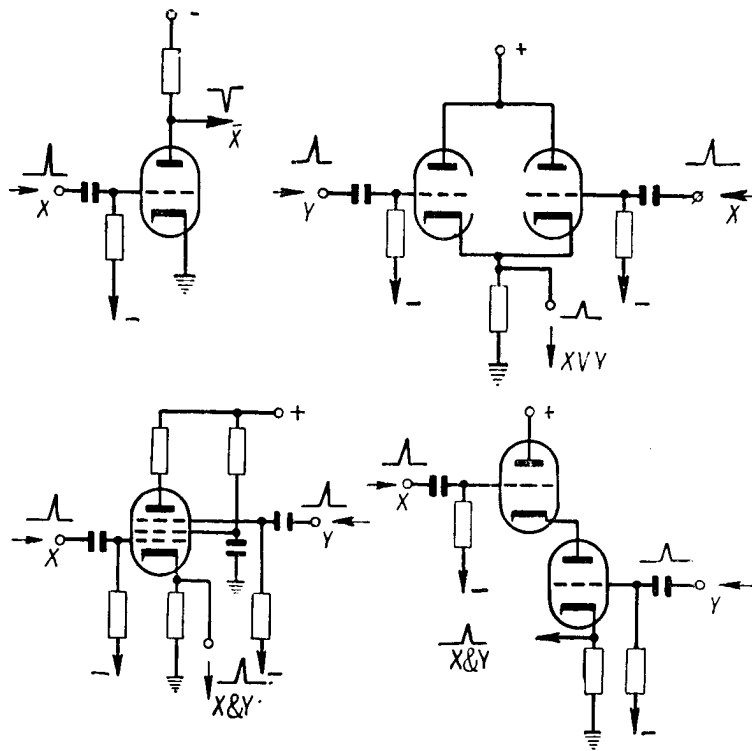


Рис. 7.35. Электронные схемы, изображающие логические операции.

Логические выражения исчисления высказываний позволяют описать в алгебраической форме любые релейно-контактные схемы. Одновременно на релейных схемах можно построить аналог любого выражения исчисления высказываний.

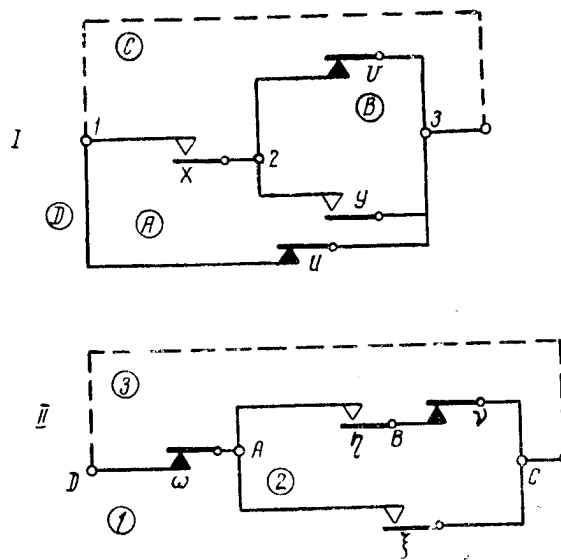


Рис. 7.36. Взаимно-обратные схемы.

Машины и мышление

Современные электронные вычислительные машины уже начали новую эпоху в кибернетике несмотря на то, что им от роду всего 12 лет. Быстрота обработки информации, гибкость, с какой они могут переключаться с одной операции и с одной задачи на другую, разнообразие выполняемых операций и, наконец, наличие собственной памяти и некоторая самостоятельность машин, работающих продолжительное время без вмешательства оператора, — все эти черты создают качественно новое явление. Машины получают все новые и все более удивительные применения, они вырывают у человека одну за другой его, казалось бы, неотъемлемые человеческие обязанности, заменяют его труд во все большем количестве операций, которые всегда считались областью умственной работы.

В самом деле, если бы мы услышали заявление о том, что можно не думая переводить текст с английского языка на русский, выбирать нужные книги или библиографические карточки из каталога, вести бухгалтерский учет,

продавать железнодорожные билеты, давать справки о расписании поездов, управлять движением самолетов над аэродромом и выводить машины в нужном порядке на посадку и все это делать без ошибок и с невероятной скоростью, мы не согласились бы с такой возможностью и запротестовали. Однако современные электронные вычислительные машины делают все это и многое другое, и, конечно, вслед за этими операциями будут появляться все новые. Значит ли это, что машины думают, выполняя человеческую работу, думают так, как думает человек, делающий то же самое?

Это — вопрос, на который нелегко ответить, даже стоя перед лицом фактов. Ясно, что результаты работы живого и машинного переводчиков с английского языка на русский совпадают. Ясно, что физические процессы в работе человека и машины не имеют ничего или почти ничего общего. И в том и в другом случае речь идет об обработке информации, которая заключена в сигналах, воплощенных в определенном физическом процессе, закодированных некоторым кодом в определенном алфавите и преобразуемых в соответствии с некоторым алгоритмом. Все эти черты кибернетического процесса перевода различны для живого и машинного переводчиков, кроме, быть может, алгоритма преобразования, который, по-видимому, имеет много общего для обоих переводчиков. Оба переводчика перекодировывают информацию с кода, называемого английским языком, в код, именуемый русским языком.

Если же мы согласимся считать, что эквивалентность результатов влечет за собой эквивалентность терминов, то, будучи последовательными, мы обязаны будем признать, что простейшие сигнальные устройства, например пожарная сигнализация или железнодорожная автоблокировка, также обладают способностью думать. Это явно противоречит интуитивным представлениям о процессе мышления и потому с этим нельзя согласиться. По-видимому, «самое понятие «думать» требует четкого и однозначного определения. Однако такое определение дать нелегко.

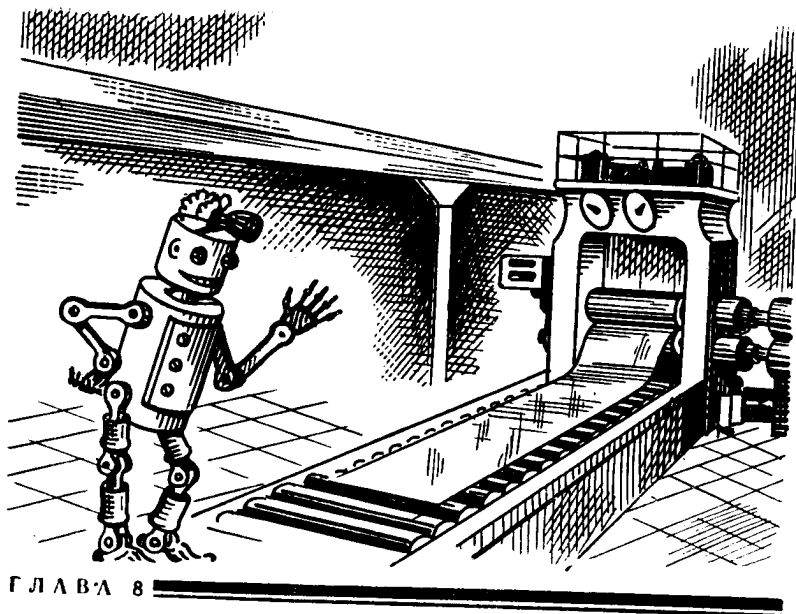
То, что делает человек, происходит посредством физиологического процесса мышления. Мышление более обширно, чем те функции, которые оно выполняет при переводе с английского на русский. То, что делает машина, она производит с помощью электронных процессов. И физиологический процесс мышления, и электронные процессы вычислений в случае перевода имеют одинаковое информаци-

ное содержание. По своим возможностям мышление шире, чем возможности современной вычислительной машины. По своей физической природе мышление не имеет ничего общего с работой машины.

Однако поскольку оба процесса ведут от одинаковых исходных данных к одинаковым результатам, и поскольку оба процесса являются кибернетическими, т. е. связанными с информацией, нельзя не признать, что они родственны.

Кибернетическое, информационное содержание мышления и содержание электронных вычислений пересекаются в узких областях, но не совпадают полностью. Физические процессы в обоих случаях различны.

Может ли машина, хотя бы в принципе, охватить все области человеческого интеллекта и можно ли расширить ее возможности за пределы человеческого мышления? Этот новый вопрос пока не имеет решения. Сегодня у нас нет достаточных оснований ответить на него твердо да или нет, но его можно и нужно обсуждать. Для того же, чтобы подойти к обсуждению этого вопроса, необходимо знать, что представляют собой процессы, протекающие в человеческом мозге. К работе мозга мы подойдем в гл. 9.



ГЛАВА 8

РОБОТ

Слово «робот» принадлежит чешскому писателю-сатирику и публицисту Карелу Чапеку*. Это слово прочно вошло в языки многих народов. У Чапека роботы (изобретенные инженером Россумом живые автоматы) идеально приспособлены для работы на станках вместо рабочих. Они не обладают человеческими чувствами и лишены человеческих запросов. По мере усовершенствования роботов они из автоматов перерастают в нечто иное и в финале драмы уничтожают своих хозяев, но вынуждены погибнуть и сами, так как только люди владеют секретом изготовления роботов. Последний из оставшихся людей вкладывает в роботов свойство желать и размножаться.

Художники-карикатуристы любят изображать роботов в виде нелепых фигур, скрепленных огромными шарнирами и опутанных свитыми в спиральки проводами.

Немало досужих изобретателей потрудились над изготовлением игрушек в рост человека, нарочито причудливо-

* Karel Čapek. «R. U. R.» Hry. Československý Spisovatel. Praha.

го вида, умеющих вставать, садиться, поворачивать голову, произносить фразы и т. д.

Все эти выдумки, рассчитанные на то, чтобы поразить воображение, а подчас и на самом деле производящие сильное впечатление, совсем не так удивительны, как действительно полезные роботы, которые не имеют вида ка-

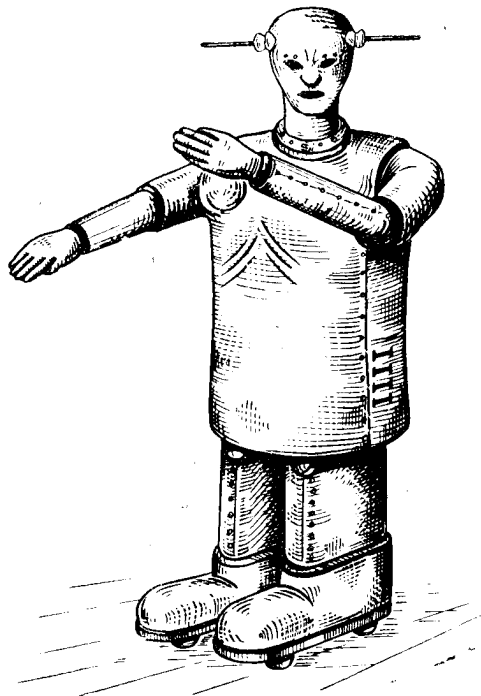


Рис. 8.1. Робот „Сабор IV“ изобретателя Августа Губера (1938).

рикатуры на человека, не подражают внешне манерам и поведению человека, но зато точно и безошибочно, неутомимо и быстро выполняют обязанности, которые всего несколько лет назад мог выполнять только человек и никто другой. То и дело появляются сообщения о новых достижениях информационной кибернетической техники, и поток этих новостей нарастает. Мы ведь делаем только первые робкие шаги по пути создания и эксплуатации роботов, и трудно еще предвидеть, сколько открытий, удивительных изобретений лежит на этом пути, какие новые факты при несет нам завтрашний день.

Мы будем называть «роботами», пользуясь терминологией Чапека, автоматические устройства, в которых реализуются следующие возможности:

1) автоматические измерения, непрерывные или дискретные, некоторого множества физических величин и

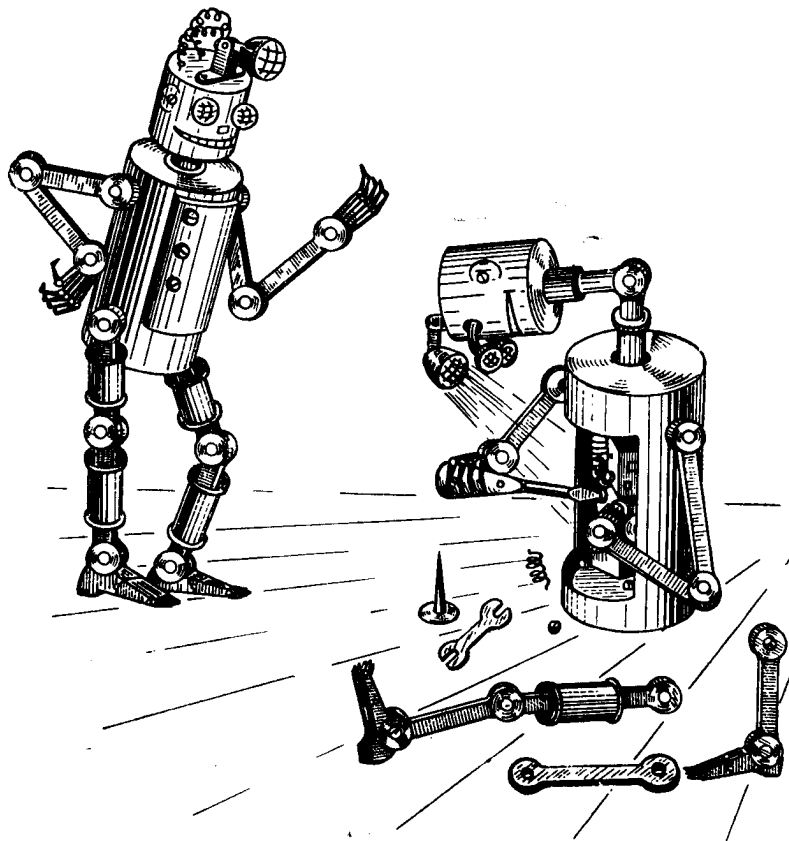


Рис. 8.2. Роботы (в изображении карикатуриста).

автоматический непрерывный ввод, по мере получения, результатов измерения (сигналов) в решающее устройство в схеме робота;

2) передача входных сигналов внутри схемы робота и запоминание информации;

3) обработка входной информации в соответствии с некоторым алгоритмом. Алгоритм может быть не единствен-

ным, он может быть выбран из множества возможных алгоритмов по определенному правилу в соответствии с поступающей извне или накопленной ранее информацией. В результате обработки входной информации образуется выходная, исполнительная информация — решение;

4) исполнение решения, реализуемое в виде действий орудий с использованием, быть может, значительных количеств энергии;

5) автоматическое измерение результатов действия (контроль исполнения) и введение полученной информации во входные устройства в виде сигнала (обратная связь).

Сложный машинный автомат-робот способен автоматически и самостоятельно выполнять сложные операции, приспособляясь к меняющимся внешним условиям и производя на всякое изменение внешних воздействий некоторое ответное действие — реакцию, рациональную в смысле решаемой им задачи.

Робот воздействует на некоторый объект, наблюдает за состоянием объекта и результатом его обработки, вычисляет на основании наблюдений, какие действия необходимо предпринять для дальнейшей обработки наилучшим (в смысле некоторого определенного критерия) образом. Результаты вычислений выдаются управляющему устройству (состояние которого также «известно» вычислительной машине) и используются для дальнейшего воздействия на объект.

По такой схеме работает автоматический прокатный стан, автоматический фрезерный станок, артиллерийский комплекс противовоздушной обороны. По такой схеме может работать огромное количество автоматов-роботов, выполняющих каждый только одну задачу, — «одноцелевых роботов».

В общем случае в схеме робота может использоваться многократная самонастройка с большим числом «этажей» иерархии регулирования.

Робот этого типа должен воспринимать информацию, понимать ее, запоминать, обрабатывать и производить действия на основании результатов вычислений. Все эти операции по отдельности и вместе машина способна выполнить.

Машины могут воспринимать информацию о внешних событиях и переводить ее на свой язык, т. е. кодировать рабочим кодом вычислительного устройства.

Огромное количество разнообразных по размерностям физических величин может автоматически, непрерывно или периодически, измеряться и превращаться в сигнал. Только перечисление этих величин займет много места, а описанию методов измерений пришлось бы посвятить не одну книгу. Устройства, превращающие отсчет физической величины в сигнал, называются в телемеханике и телеметрии датчиками и величинами. Каждый датчик характеризует-

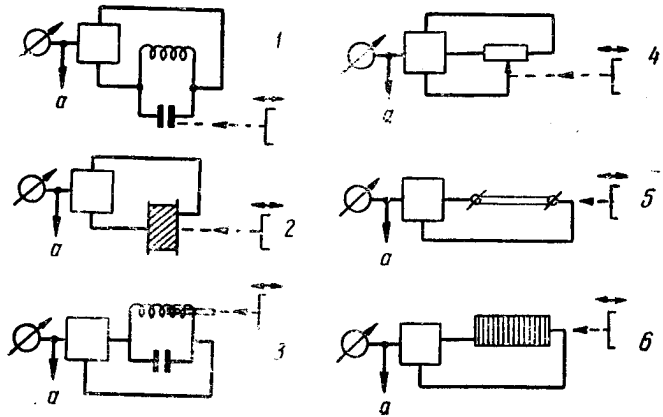


Рис. 8.3. Примеры датчиков — преобразователей механического перемещения в электрическое напряжение или ток.

1 — емкостный датчик (смещение пластины электрического конденсатора вызывает изменение его емкости и изменение резонансной частоты электрического контура); 2 — пьезоэлектрический датчик (деформация пьезокристалла вызывает появление электрического потенциала на его поверхности); 3 — индукционный датчик (смещение сердечника вызывает изменение индуктивности катушки и изменение резонансной частоты электрического контура); 4 — потенциометрический датчик (смещение движка потенциометра вызывает изменение тока, текущего через внутренние механические напряжения и электрическое сопротивление проволоки, а также и текущий через проволоку ток); 5 — тензометрический датчик (деформация угольного столбика меняет его электрическое сопротивление и текущий через него ток).

ся, во-первых, той входной физической величиной, которую он измеряет и, во-вторых, тем типом сигнала, который он выдает на выходе — физическим алфавитом и кодом.

Входной величиной может быть: геометрический размер (длина), скорость — первая производная длины по времени, ускорение — вторая производная, угол (отклонение от заданного направления), угловая скорость, сила, давление жидкости или газа, концентрация водородных ионов, сила света различных длин волн, звуковое давление и высота тона, электрический ток и напряжение, напряженность электрического или магнитного поля, число ионизирующих

частиц, количество (уровень) жидкости, масса, вес, поток (газа, жидкости), протекающий через поперечное сечение канала, частота, амплитуда, фаза колебаний (механических, электрических, акустических), твердость, упругость, прозрачность, электропроводность, диэлектрическая и магнитная проницаемости, вязкость, концентрация ионов и т. д. и т. п. При этом каждая величина может измеряться не в одной только точке, а во многих и, следовательно, измеряться будет не местное значение, а распределение и с величины по контуру, поверхности или объему и изменение этого распределения во времени. Измерения каждой из величин могут, в случае необходимости, производ-

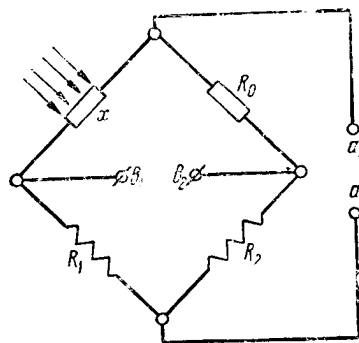


Рис. 8.4. Мостовая схема измерений.

$a_1 - a_2$ — место подключения питания.
 $b_1 - b_2$ — место подключения измерительного прибора, x — чувствительное сопротивление.

даться в различных пределах, с различной точностью, в различных масштабах, непрерывно или периодически, с различным темпом. Измерения могут быть как абсолютные, так и относительные или разностные.

Выходным сигналом датчика чаще всего является сигнал, представляющий собой непрерывную электрическую величину, например величину прямого тока, с последующей перекодировкой в любой алфавит.

Техника автоматических измерений весьма богата. Одну и ту же физическую величину часто представляется возможным измерить различными методами, используя различные эффекты. Так, например, перемещение или изменение геометрического размера можно превратить в величину электрического напряжения, связав перемещаемый элемент с движком потенциометра (металлического, полупроводникового, жидкостного), или с перемещающейся пластинной переменной конденсатора, или с подвижной катушкой переменной индуктивности (вариометра), или

с пьезоэлектрической пластиной и т. д. В зависимости от требований точности, пределов измерения, постоянной времени выбирают тот или иной метод измерения.

В измерениях большой популярностью пользуются так называемые компенсационные методы, сущность которых сводится к тому, что измеряемая величина непосредственно в приборе сравнивается с некоторой стандартной величиной той же размерности. Чаще всего такое сравнение производится в электрической схеме моста. Компенсационные схемы особенно полезны в тех случаях, когда чувствительный элемент реагирует не только на изменение измеряемой («полезной») величины, но и на другие внешние условия. Так, например, при измерении деформаций и усилий с помощью проволочных электрических тензометров проволока, меняющая электрическое сопротивление под воздействием натяжения, меняет сопротивление также и под воздействием изменений температуры; или при измерении влажности чувствительный элемент, кроме влажности, оказывается также чувствительным к температуре, давлению и т. д. В этих случаях в схему моста полезно включать два одинаковых чувствительных сопротивления, причем одно из них используется для измерений (тензометр, гигрометр), а другое — для компенсации паразитных влияний (нагрев и пр.), которые оказываются одинаковыми для обоих сопротивлений. При этом, разумеется, компенсирующее сопротивление не подвергается воздействию измеряемой величины. Меняя компенсирующее сопротивление, можно заранее «настроить» мост на желаемое значение измеряемой величины (сместить нулевую точку).

В схемах измерения часто применяют следящие системы. При этом узлом сравнения может являться мостовая схема. Выходной сигнал моста усиливается и приводит в действие исполнительное устройство, которое воздействует на компенсирующий элемент (например, мотор перемещает движок реостата). При этом измеряемая величина «отрабатывается» исполнительным устройством значительной мощности, и, таким образом, при измерении очень малых энергий на выходе активного измерительного устройства такого рода в качестве сигнала получается довольно значительная энергия. Кроме того, указанный метод измерения позволяет получать сигнал в виде угла поворота вала, что иногда оказывается удобным.

Полученный в результате измерения сигнал, отображающий измеренную величину, распределение величины

или распределение нескольких величин, должен быть закодирован в алфавите вычислительного устройства, которое использует этот сигнал. В результате кодирования сигнал может быть представлен либо в виде непрерывно меняющейся величины, либо в дискретной форме. При кодировании могут меняться и физический алфавит сигнала и масштабы, а также производиться функциональные преобразования шкал.

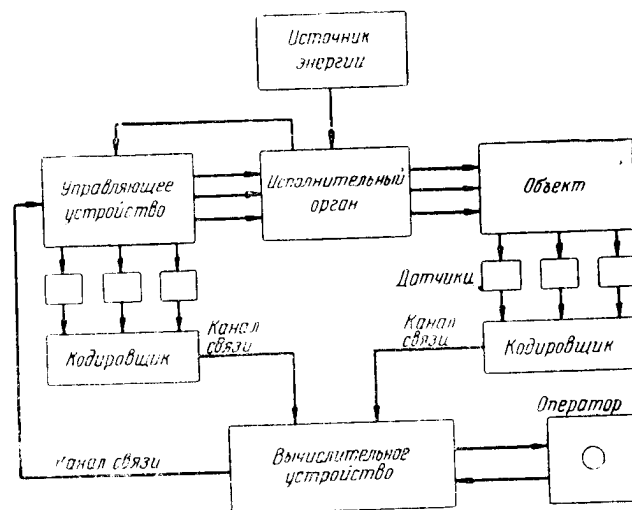


Рис. 8.5. Общая схема автомата-робота, воздействующего на некоторый объект.

Так, например, измерение размера, взятое с выхода датчика в виде электрического напряжения, может быть преобразовано в поворот вала и затем переведено в алфавит дискретной бинарной записи и пробито на перфокарте. Величина электромагнитного поля и звуковое давление, изменяющиеся в таких широких пределах, что максимальный отсчет отличается от минимального в тысячи раз, могут быть изображены сигналами, пропорциональными логарифмам этих величин, и представлены на шкале длиной в несколько единиц.

Распределение освещенности по участку экрана, измеренное с помощью фотоэмиссии электронов или с помощью фотосопротивлений и превращенное, таким образом, в распределение тока или сопротивления, перекодируется во

временную последовательность величин электрического напряжения посредством последовательного просмотра точек экрана электронным пучком и затем превращается в амплитудную модуляцию метровых радиоволн для передачи изображения на расстояние по радио. Одновременно передавая три распределения освещенности различных цветов, мы кодируем цветное плоское изображение. В принципе можно закодировать и передать шесть распределений, соответствующих цветному стереокадру.

Распределение черных и белых пятен в кадре, например рисунок буквы печатного текста, может быть превращено фотоэлементом с разверткой телевизионного типа в последовательность во времени электрических импульсов (последовательный двоичный код) или в сигнал наличия и отсутствия напряжения в ряде шин — по числу точек измерения (параллельный двоичный код). Каждой букве данного шрифта будет соответствовать одно двоичное число, букве, записанной любым шрифтом, — некоторый набор двоичных чисел. Таким способом в машину можно вводить печатный текст, предназначенный для обычного чтения.

Можно сказать, что любое множество значений измеряемой физической величины или множество распределений этой величины может быть сопоставлено посредством измерения и кодирования с любым, равным по объему, множеством значений сигнала. Разнообразие измеряемых величин, технических приемов измерений и сигналов составляет богатую и пеструю картину техники автоматических измерений.

Сигнал, полученный в результате кодирования измеренной величины, может быть передан по каналу связи (телеметрический канал) и на приемном конце записан с помощью регистрирующего устройства. Часто уже такая фиксация результата дистанционного измерения оказывает огромную услугу. Эта техника дает возможность пользоваться результатами измерений (например, космического излучения), проведенных в стратосфере измерительными приборами — датчиками, установленными на борту ракеты, или изучать режим работы двигателя беспилотного самолета в полете.

Передача сигнала от измерительных устройств к вычислительному устройству может осуществляться и не на многие километры, а всего лишь на расстояние нескольких метров, по кабелю. В вычислительном устройстве сигнал,

несущий информацию об измеренных величинах, подвергается обработке для определения наилучшего действия — реакции в данной ситуации.

Прежде всего, каждая из измеренных величин, являясь объектом обработки, дает материал для определения производных по времени различных порядков и интегралов, спектрального состава, функций корреляции и т. д. Затем различные величины могут быть сопоставлены друг с другом и вычислены функции от нескольких переменных. Так, по трем текущим координатам могут быть вычислены векторы скорости и ускорения и определена экстраполяцией будущая траектория движущейся точки. По данной температуре можно определить скорость нагрева или остывания и будущую температуру и т. д.

Интересным примером обработки измеренных величин является определение текущих спектров звуковых колебаний. Звуки, издаваемые человеческим речевым аппаратом, сложны по своему спектральному составу. Каждый звук, или фонема, представляет собой сложный процесс во времени, который может быть представлен суммой гармонических составляющих. Эти составляющие, спектр фонемы, зависят не только от произносимой буквы, но и от места в слове, характера произношения этой буквы в слове (например, буква «о» в словах *слон* и *отец*), от тембра голоса, интонации, ударения и т. д. Однако для одной и той же фонемы, произносимой различными голосами и с различной интонацией, сохраняются некоторые особенности, которые и позволяют на слух различать одну фонему от другой. Эти особенности заключаются в различном относительном расположении отдельных спектральных составляющих.

Подвергая электрическое напряжение, снятое в качестве сигнала с выхода акустического датчика — микрофона, спектральному анализу и обрабатывая полученный спектр с целью опознавания характерных особенностей фонем, можно в некотором алфавите получить сигналы, соответствующие определенным фонемам, и заставить исполнительное устройство — автоматическую пишущую машинку типа телетайпа печатать в виде текста речь, произносимую перед микрофоном. Так может быть устроен простейший автоматический стенограф. Правда, этот автомат не будет знать грамматики и будет записывать слова так, как он их слышит. Вместо слова *отец* он запишет, наверное, *атец*, а слово *слон*, весьма возможно, запишет как *илон*, если

говорящий шепелявит или если в канале записи имеются искажения.

Можно заставить автоматического стенографа учитывать и грамматику; это задача более сложная, но она разрешима. К ней мы еще вернемся.

По мере поступления сигналов в вычислительное устройство и по мере их обработки может производиться и автоматическая регистрация, или запоминание, как самих сигналов, так и их функций. Это может оказаться необходимым либо для того, чтобы контролировать впоследствии работу робота, либо для самого решения задачи, если в процессе решения приходится прибегать к использованию прошлых значений сигнала.

Далее производится решение задачи о выборе реакции, необходимой для достижения нужного результата в данных условиях. Вычисления, связанные с нахождением этого решения, могут быть простыми или сложными в зависимости от характера задачи. Они могут производиться с помощью устройств дискретного счета или непрерывного действия и определяться программой решения или структурой вычислительного устройства, предусмотренными заранее для данной задачи. Вычисления могут включать в себя логические операции сравнения и выбора, обращение к памяти и сравнение с прошлыми входными или промежуточными данными или результатами.

Существенным моментом в решении задачи является то, что для определения оптимального или наилучшего решения некоторая переменная величина должна в результате принимать экстремальное (наибольшее или наименьшее) значение. Такой величиной может быть, например, время, потребное для всего цикла обработки, которое нужно свести к минимуму, или же количество выданной продукции из данного количества сырья, которое должно быть наибольшим, и т. п. Важно то, что такая величина — критерий качества работы автомата — всегда существует, будучи заложенной конструктором в структуру машины или в программу ее работы, и на основании этого критерия машина может самостоятельно выбирать именно то действие из всех возможных, которое соответствует наилучшему, или оптимальному, значению критической величины.

Полученное вычислительным устройством решение подается в виде сигналов к управляющим элементам робота, которые воздействуют на исполнительные механизмы. Передача сигнала к управляющим устройствам произво-

дится по каналу связи и может осуществляться как на ближние, так и на дальние расстояния.

Вычислительное устройство контролирует исполнение команды, получая информацию от управляющего устройства о положении исполнительных элементов (обратная связь).

Само управляющее устройство может представлять собой систему с обратными связями, исполняющую сигнал — команду, которая приходит от вычислительной машины. Команда превращается в действие — реакцию робота на обстановку. Действие исполнительных органов робота вызывает появление изменений в объекте обработки. Эти изменения регистрируются измерительными приборами — датчиками и снова вводятся в автомат в качестве новой информации. Процесс обработки продолжается без перерыва до его окончания.

Итак, наш робот не только ощущает и принимает решения, но и действует. Он действует решительно, эффективно, правильно, но, кроме управления своим объектом, он ни на что не отвлекается. Он работает, как очень сосредоточенный, пожалуй даже, как фанатически настроенный труженик: он не способен ни размышлять, ни отвлекаться, ни заниматься чем-либо другим, кроме своей задачи. Он работает, как чапековский робот.

Система автоматизации прокатного стана требует устройства именно такого типа. Прокатываемая болванка проходит между валами с определенной скоростью, причем скорость выхода после валов зависит как от скорости вращения валов, так и от степени обжимки. Если несколько пар валов расположены последовательно друг за другом, то режим работы последующих (зазор, число оборотов, вращающий момент) зависит от скорости подачи прокатываемого материала, т. е. от режима работы предыдущей пары валов. Кроме того, режим зависит еще и от температуры, механических свойств материала и формы его сечения. В связи с тем, что все эти величины меняются и зависят от режима прокатки, ясно, во-первых, что задача о последовательном пропускании болванки через несколько пар валов весьма сложна, и, во-вторых, поскольку это так, что точное ее решение позволит значительно ускорить производство и удешевить его, так как стан будет использоваться на полную мощность.

Нужно сказать, что роботы, подобные тому, который мы выше описали в общих чертах, являются чрезвычайно

полезными устройствами. Они чувствительнее, безынерционнее, неутомимее, чем любой работник. Для управления сложной производственной операцией, например плавкой или прокаткой стали, необходим большой производственный опыт. В чем он заключается? В том, что старый мастер знает наперед, как пойдет процесс, если произойдут какие-либо отклонения от нормы. По первым признакам отклонения он, долго не раздумывая, примет нужные меры, не дожидаясь, пока станет очевидным появление брака.

Вот такого опытного мастера и может заменить робот. Он тоже может предвидеть ход процесса и принять вовремя нужные меры не только для предотвращения брака, но и для предотвращения замедления процесса. Опыт показывает, что введение автоматического управления производственным процессом всегда ведет к значительному ускорению производства и повышению качества продукции. Роботу не нужно учиться и приобретать опыт. Он всегда работает очень хорошо.

Но хотя для производства и хорошо, что робот не требует обучения, тем не менее в принципе невозможность обучения и самосовершенствования есть несомненно неспособность робота, т. е. его несовершенство.

Работающие сегодня роботы-мастера, роботы-производители не учатся. Но их экспериментальные братья-роботы способны обучаться. Для того чтобы придать роботу эту способность, достаточно снабдить его дополнительными «органами» — устройствами, которые меняли бы программу вычислений в зависимости от качества выполнения задачи роботом.

Достаточно простые устройства такого рода осуществлялись с демонстрационной целью. Они были способны менять свою программу действий по мере производства последовательных проб. Это изменение поведения в некоторой мере имитирует процесс обучения на основе опыта (методом проб и ошибок). В результате накопления опыта закрепляется некоторое «наилучшее» в определенном смысле поведение. Смысл понятия «наилучшее» определяется критерием оценки качества работы, заложенным конструктором.

Кибернетические игрушки

Автоматы-роботы, имеющие практическое значение, по необходимости весьма сложны и не только конструктивно, но и принципиально. Задачи автоматизации произ-

водства в основном еще не решены до конца. Решаются только отдельные части, частные задачи большой проблемы полной автоматизации производственных операций и производства в целом. Пока еще только создается теория больших автоматов, намечаются принципы их создания и разрабатываются необходимые частные вопросы, весьма поучительным оказывается конструирование простых кибернетических роботов, которые демонстрируют возможности, превышающие обычные способности автоматов.

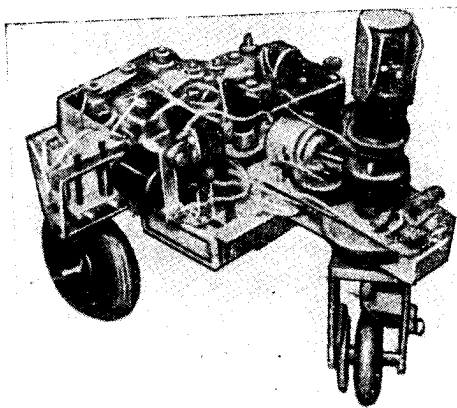


Рис. 8.6. Черепашка Эльси Грея Уолтера со снятым внешним щитом.

Можно создать робот-игрушку, состоящую всего из нескольких деталей: чувствительных, исполнительных и коммутирующих устройств. Такие игрушки создавались и создаются в значительном числе как кибернетиками-любителями для развлечения, так и учеными для решения серьезных задач. Многие из них описаны в литературе и имеют большое принципиальное значение.

Мы не станем перечислять многочисленные варианты автоматов с часовыми механизмами, подчас довольно замысловатых и сложных, ведущих свою историю еще с XVII века. Гораздо интереснее появившиеся недавно роботы с обратной связью.

Одними из первых описанных в литературе игрушек-роботов были «черепашки» англичанина Грея Уолтера. Они представляют собой миниатюрные тележки на колесах, снабженные двумя электромоторчиками. Один —

вращает колеса, другой — поворачивает руль. На тележке, кроме источников питания, установлены фотоэлемент, усик, замыкающий электрический контакт при соприкосновении с препятствием, несколько реле и электрическая лампочка. Несмотря на такое предельно примитивное устройство, черепашки демонстрируют свойства не только забавные,

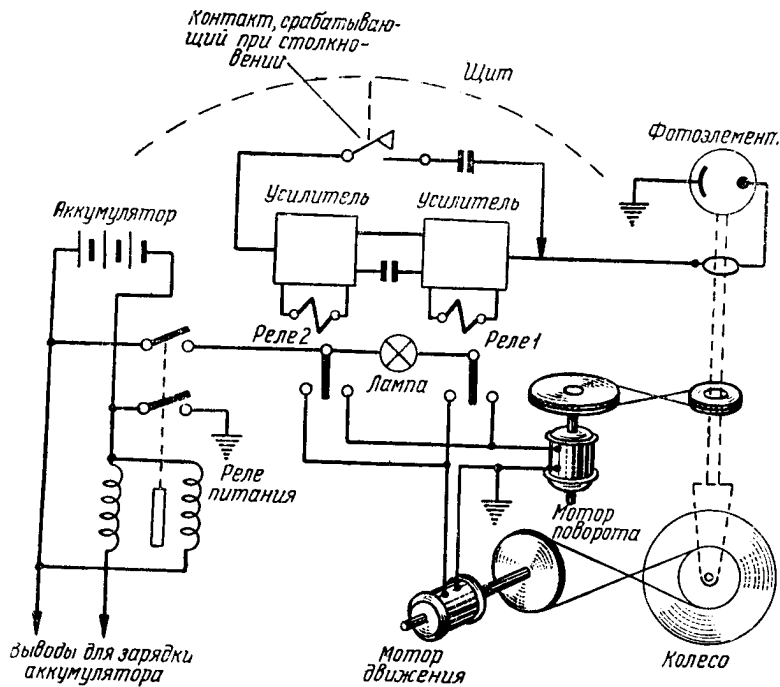


Рис. 8.7. Схема соединений элементов черепашки Эльси.

но и поучительные. В темноте или при слабом свете они беспорядочно ползают, как будто чего-то ищут. Наткнувшись на препятствия, они сворачивают и пытаются их обойти. Если имеется достаточно сильный источник света, они его скоро замечают и решительно направляются в его сторону, обнаруживая «положительный тропизм». Однако подойдя к свету слишком близко, они от него отворачиваются («отрицательный тропизм»), двигаются вокруг источника света, находя для себя «оптимальные условия» и непрерывно поддерживая их («гомеостазис»). Между двумя источниками света черепашка совершает путешествие от

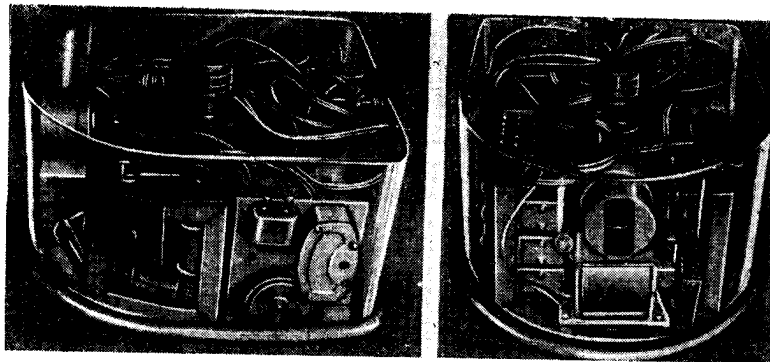


Рис. 8.9. Внешний вид игрушечного робота — черепашки Эйхлера, обладающей „условными рефлексам“.

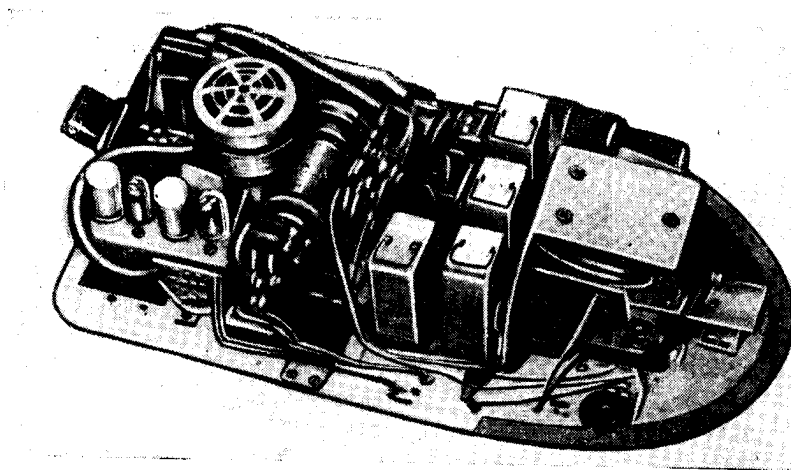


Рис. 8.10. Советская черепашка с „условными рефлексам“ (без щитка).

зирована» на определенную работу — собирание с пола разбросанных в беспорядке белых шариков (мячей для игры в гольф) и складывание их в определенном месте. Авторы назвали своего игрушечного робота «белка», так как живая белка тоже собирает орехи на зиму и складывает их в свое гнездо. «Белка» Беркли имеет два фотоэлемента, две лапки, которые могут раздвигаться или сдвигаться на уровне пола, образуя чашечку, язычок внутри ашечки и металлический хвостик, волочащийся по полу.

Кроме того, в «организме» белки, т. е. в ее схеме, имеется, помимо коммутирующих реле, еще фильтр, позволяющий различать постоянный и переменный токи.

Обстановка, в которой работает белка, такова. Большая пустая комната освещена лампами накаливания. По полу в беспорядке разбросаны белые шарики. В одном из углов комнаты положен металлический лист, освещенный газосветной люминесцентной лампой (лампой дневного света) и обозначающий «гнездо» белки. Белка совершает наугад поиск, пока в поле зрения фотоэлементов, служащих ей «глазами», не попадет белый шарик. Тогда белка направляется прямо к нему, раздвигает лапки, останавливается, сдвигает лапки, захватывая в них шарик. Язычком, помещенным между лапками, белка чувствует, что шарик у нее «в лапках». Затем белка поворачивается и ищет гнездо. Поскольку гнездо освещено мигающим светом газосветной лампы, в отличие от непрерывного света ламп накаливания, электрический фильтр позволяет ей распознать направление на гнездо. Белка направляется к гнезду, заползает на металлический лист и останавливается на нем, так как замыкание контакта между листом и ее металлическим хвостиком дает ей знать, что она «дома». Белка раздвигает лапки, выбрасывая шарик, и снова направляется на поиски следующего шарика.

Схема коммутации между чувствительными и исполнительными элементами может быть очень быстро и без труда составлена с помощью релейной логической алгебры, как только сформулированы все операции, которые должна выполнять белка.

Еще более интересной и поучительной, а кстати, принципиально и гораздо более серьезной является машина Клода Шэннона, имитирующая поведение мыши в лабиринте. «Мышь» Шэннона, которую автор назвал именем древнегреческого героя Тезея, представляет собой всего лишь намагниченный кусочек стали размером в несколько сантиметров, на колесиках, с декоративными глазками ушками и хвостиком. Мышь Тезей ползает по квадратной доске, разбитой на 25 одинаковых квадратов, под действием магнита, помещенного под доской и направляемого ведущим механизмом, моторы которого управляются релейным устройством, состоящим из 50 телефонных реле.

По границам квадратов устанавливаются перегородки в произвольном порядке, которые образуют лабиринт. В одном из квадратов помещают «сыр» — электрод, замы

кание которого представляет собой цель путешествия мыши. Сыр — мышиное лакомство — может быть заменен выходом из лабиринта или, если хотите, другой целью хождения по лабиринту, например Минотавром — чудовищем с головой быка и телом человека, которого древнегреческий легендарный герой Тезей — тезка мыши Клода Шэннона — отыскивал в лабиринте Миноса, царя Крита. Независимо от литературного оформления цели путешествия такая цель существует, находится в определенном квадра-

Рис. 8.11. Внешний вид искусственной обучающейся мыши Тезея Клода Шэннона. Видны колесики, магнитная пластинка, усик, ощущающий препятствие, и нижний контакт, ощущающий приход к цели.



те, и мышь отправляется на поиски ее, начиная с того квадрата, куда ее пустили.

Механизм ведет мышь наугад, переводя ее в соседний квадрат по «горизонтали» или по «вертикали». Решение о направлении движения дается некоторым механизмом, «бросающим жребий», т. е. имитирующим случайный процесс. Если мышь встречает на своем пути из квадрата в квадрат перегородку и не может пройти, специальное устройство сигнализирует об этом в решающее устройство, где замыкается соответствующее этой границе квадрата реле и, таким образом, «запоминается» частичка плана лабиринта. Мышь поворачивается под прямым углом вправо или влево (случайно) и продолжает свой путь, испытывая успехи и неудачи, наугад продвигаясь все дальше и запоминая по пути все больше сведений о расположении перегородок лабиринта. Наконец, совершенно случайно мышь добирается до цели своего странствования (Тезей убивает Минотавра или мышь съедает сыр, — как Вам

удовно). При этом все замкнутые реле остаются замкнутыми и исследованная мышью часть лабиринта — зафиксированной в памяти решающего устройства.

Если мышь снова вернуть в исходное положение, она будет вести себя совершенно иначе, чем в первый раз. Она уже не движется наугад, не ищет, не ошибается. Она движется к цели кратчайшим путем по той части лабиринта, которая ею уже исследована.

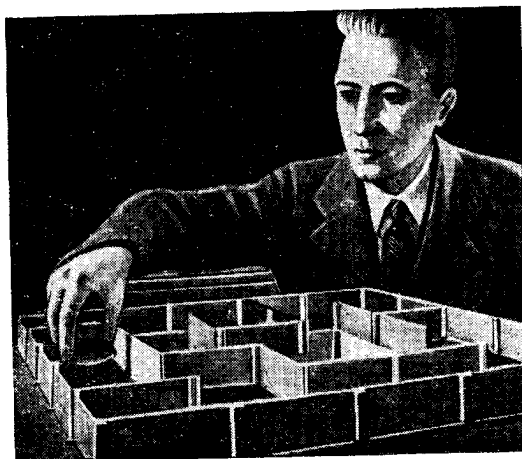


Рис. 8.12. Клод Шэннон пускает искусственную мышь Тезея в лабиринт.

Если теперь мышь поместить в один из квадратов, в котором она еще ни разу не побывала и который она, следовательно, «не знает», мышь снова начнет бродить наугад, запоминая путь, как и в первый раз, пока случайно не попадет в один из знакомых квадратов, откуда она сразу находит кратчайший путь к цели.

«Мышь» Клода Шэннона демонстрирует нам, так сказать, «машинный вариант» способности к запоминанию предыдущего опыта и к рациональному его использованию в дальнейшем. Это не что иное, как «обучение», понимаемое в самом широком смысле. Недаром внешнее оформление машины Шэннона нарочито пародирует «лабиринтовые эксперименты» Лешли и других зоопсихологов и физиологов, проводившиеся в очень больших количествах и заключающиеся в наблюдениях главным образом над

крысами и определении времени, в течение которого животное находит выход из лабиринта при последовательно проводимых попытках. Это время обычно закономерно уменьшается, образуя так называемую «кривую научения». «Мышь» Шэннона, в отличие от живых мышей или крыс, идеально «обучается» за один эксперимент, без повторных попыток.

Правда, искусственная мышь ничему, кроме выхода из лабиринта, не может больше научиться, но это и не суще-

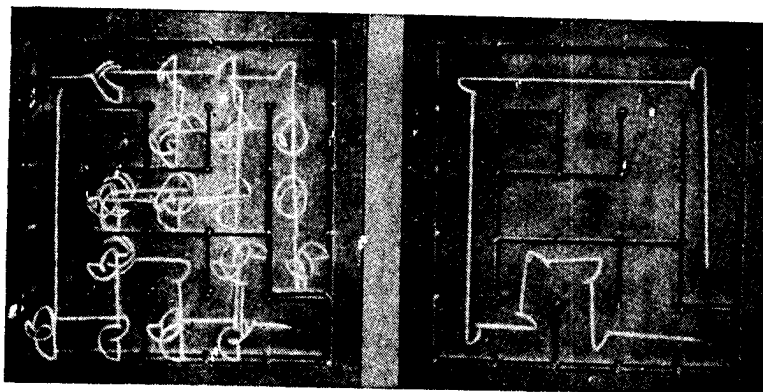


Рис. 8.13. Путь мыши Тезея в лабиринте.
а) при первом запуске (мышь бродит наугад). б) при втором запуске (мышь сразу находит кратчайший путь к цели).

ственно. Дело в том, что эксперимент Шэннона преследовал не столько цель имитировать некоторую отдельно взятую черту поведения животного, сколько совсем другую и гораздо более существенную для технических приложений цель, а именно: создать такую систему коммутации телефонных линий, которая автоматически избирала бы для соединений (например, междугородных) кратчайшие пути с наименьшим числом обходных петель в зависимости от наличной загрузки кабелей.

В последнее время экспериментальные модели типа черепахи и мыши построены в Советском Союзе в ходе экспериментальных работ по кибернетике.

Кроме мыши Шэннона, было сделано много других более или менее простых роботов, имитирующих отдельные черты поведения животных. Так, например, неоднократно демонстрировался условный рефлекс как на машине

дискретного счета, так и на простых электрических игрушках. В этих случаях, как и всегда при сравнениях поведения животных и роботов, следует помнить, что робот обычно имеет в своем распоряжении лишь ничтожное количество возможных «раздражений» или входных сигналов, в то время как животное может воспринимать необозримое разнообразие внешних воздействий.

Условным рефлексом называется выработанная в результате тренировки реакция животного на безразличный

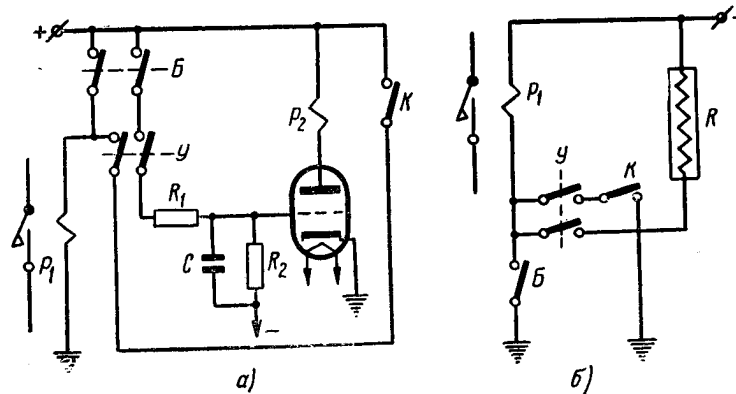


Рис. 8.14. Возможные схемы для имитации условных рефлексов в игрушечных роботах.

Б — контакт, замыкаемый при появлении безусловного раздражителя, *У* — контакт, замыкаемый при появлении условного раздражителя, *Р₁* — исполнительное реле, включающее реакцию на безусловное раздражение. После многократного одновременного включения контактов *Б* и *У* в схеме *а*) заряжается емкость *С*, в схеме *б*) нагревается сопротивление *Р*, в результате замыкается контакт *К* и после этого реле *Р₁* возбуждается только замыканием контакта *У*.

(условный) раздражитель. Основным условием образования условного рефлекса является многократное совпадение во времени действия безразличного раздражителя и агента, вызывающего безусловный рефлекс, который имеет непосредственное биологическое значение. Обычная схема опыта такова: у животного (например, у собаки) многократно вызывают биологически целесообразную реакцию на внешний раздражитель (например, выделение слюны при приёме пищи, оборонительные движения при уколе в лапу и т. д.), причём действию внешнего раздражителя (пища укол), вызывающего автоматическую или безусловную реакцию, каждый раз предшествует действие безразличного раздражителя (например, звонка, зажигания лампы и т. п.). После некоторой тренировки у животного выраба-

тывается автоматический ответ выделения слюны только на включение звонка, без подкрепления пищей, или отдёргивания лапы только на зажигание лампы, без укола. Если много раз повторять демонстрацию условного рефлекса, не «подкрепляя» его, т. е. не проводя время от времени совместного применения безразличного и безусловного раздражителей, то условный рефлекс затухает и в конце концов исчезает вовсе.

Можно выработать условный рефлекс на сложное восприятие, например аккорд, форму светового пятна и т. д.; можно выработать условный рефлекс, совмещая безразличный раздражитель не с безусловным, а с другим условным, на который предварительно уже был выработан условный рефлекс.

Если отвлечься от всего разнообразия возможных сочетаний простых и сложных раздражений и ограничиться схемой с одним условным и одним безусловным раздражителями, то для процесса выработки и затухания условного рефлекса нетрудно осуществить машинную модель. В любой такой схеме необходим узел, который бы позволял запоминать совпадение двух раздражителей — безусловного и условного.

На схеме рис. 8.14,а этим элементом является емкость C , которая через сопротивление R и одновременно включенные контакты B и $У$ заряжается до положительного потенциала и заставляет радиолампу L проводить ток. В результате срабатывает реле P_2 , замыкая контакт K , и после этого реле P_1 может быть включено только контактом $У$ без замыкания контакта B . Пока конденсатор C не разрядится через сопротивление R (на что потребуется некоторое время, тем большее, чем больше величины емкости C , сопротивления R и заряда на емкости C), исполнительное реле P_1 может включаться либо контактами $У$, либо контактами B , либо обоими вместе. После разряда емкости C лампа L перестает проводить, контакт K размыкается и исполнительное реле P_1 может включаться только контактом B . «Условный рефлекс» оказывается затухшим.

На схеме рис. 8.14,б в качестве запоминающего элемента использовано нагревание биметаллического контакта K электронагревателем R . Это нагревание происходит только в момент одновременного замыкания контактов B и $У$. После нагрева контакт K замыкается и до остывания позволяет включать исполнительное реле P_1 замыканием контактов $У$ при незамкнутых контактах B .

Используя в качестве модели «черепашку», подобную описанным выше, можно включить реле P_1 в одной из приведенных схем как исполнительное реле, вызывающее остановку черепашки (выключение ведущего мотора) или ее поворот (включение управляющего магнита или мотора). В качестве контакта B в схеме можно использовать контакт усика, замыкающийся при соприкосновении с препятствием, а в качестве контакта $У$ — контакт реле, срабатывающего от микрофона на свисток. В результате мы получим «черепашку с условным рефлексом». Наш робот будет всегда останавливаться или поворачивать, столкнувшись с препятствием («безусловный рефлекс»). При этом он никак не будет реагировать на свисток (безразличный раздражитель). После достаточно продолжительного или повторного совпадения условного раздражителя с безусловным, т. е. если мы несколько раз будем давать свисток в тот момент, когда черепашка натывается на препятствие, робот начнет останавливаться по свистку, без соприкосновения с препятствием. Если долго не «подкреплять» рефлекс, т. е. не давать свистка одновременно с остановкой на препятствии, то робот «забудет» тренировку и вернется к прежнему состоянию. Такие черепашки с условным рефлексом изготавливались и испытывались.

Нет нужды, по-видимому, напоминать здесь, что физические процессы в животном и роботе не только различны, но могут не иметь вообще ничего общего. Но с точки зрения кибернетической, т. е. с точки зрения использования информации о внешней обстановке для выработки реакции, условный рефлекс животного и «условный рефлекс» робота несомненно имеют больше сходства, чем различий. Схема внешних проявлений в обоих случаях в общих чертах, попросту, одинакова. Различия заключаются лишь в том, что у животного имеется способность к выработке одновременно или последовательно многих различных рефлексов, а наши черепашки способны лишь к немногим сочетаниям. Никто, впрочем, не может запретить, пожертвовав простотой устройства игрушки, сделать ее способной к выработке достаточно большого числа различных рефлексов. Говоря о различиях, необходимо сказать еще и о том, что в силу многообразия восприятий животного и обилия образов, запасаемых в памяти, выработка условного рефлекса всегда протекает под перекрестным огнем посторонних внешних и внутренних влияний и потому включает в себя элементы случайности. Случайность отсутствует в простей-

шем роботе. Ее, впрочем, можно искусственно внести в модель. По-видимому, при достаточно развитой схеме и обилии восприятий и реакций у модели эта случайность неизбежно появится сама, и робот, как и подопытная собака, будет «отвлекаться» посторонними шумами, вибрациями, светом и даже «воспоминаниями по ассоциации», и для опытов с ним придется, как и для животных, строить «башню молчания».

Теперь мы оставим поучительные игрушки, имитирующие животных, и обратимся к роботам, имитирующим деятельность человека.

Роботы и язык

Рассматривая самые простые модели, мы видели, что их возможности лимитируются ограниченностью доступных им восприятий и реакций. Для того чтобы оказаться на высоте положения в тех случаях, когда робот перестает быть игрушкой и претендует на оказание серьезной помощи человеку в его труде, он, робот, должен быть способен к некоторым сложным операциям обработки информации, которые уже не укладываются в простые схемы. Помогая человеку, робот имеет дело с тем материалом, над которым работает человек. Если оставить в стороне производственные операции, то останется еще широкое поле деятельности, связанное с человеческим языком, записанной и напечатанной человеком и для человека информацией.

В истории человеческих знаний было два выдающихся события, которые, каждый по-своему, открыли новые пути и перспективы сохранения полезных сведений. Это — изобретение письменности, позволившее передавать знания потомкам в виде документов, свободных от искажающих воздействий устной передачи и превратностей человеческой памяти, и изобретение книгопечатания, сделавшее возможным распространение знаний в широкие массы человеческого общества.

Однако в последнее время ощущается настоятельная необходимость еще в одном, новом шаге в вопросе хранения и распространения информации. Дело в том, что количество полезных сведений, накопленных за годы существования точных наук и изложенных в печатных трудах, стало настолько большим, что, хотя любые из этих сведений в принципе и доступны для изучения, вся их масса вряд ли может быть обозрима одним человеком в разумные

сроки. Книги по всем специальностям лежат в библиотеках, и для того, чтобы отыскать нужные сведения, приходится тратить недели и месяцы нелегкого труда. Для исправления положения составляются справочники, каталоги, обзорные труды, на что уходит также немало времени, средств и усилий, но что, однако, далеко не исправляет печального положения вещей.

Обилие накопленных знаний начинает вступать в конфликт с возможностями человека, с его «пропускной способностью» и требует для дальнейшего развития наук применения новых, более сильных и гибких методов обращения с информацией с тем, чтобы в минимальное время сделать возможным доступ к любой информации, не только по ее адресу, т. е. названию труда и имени автора, но и по содержанию. Иначе говоря, необходимы средства для быстрого просмотра большого массива книг и журналов и выбора из них нужных сведений. Эта задача является развитием и обобщением работы по составлению обзоров, реферативных справочников, аннотаций и каталогов. В решении этой задачи машины могут и должны оказать человечеству не меньшую услугу, чем печатный станок. Именно быстродействующие машины с достаточно большой памятью способны и обязаны хранить, сортировать, пополнять запас информации в библиотеках и совершать все необходимые информационные поиски, выдавая по требованию читателей любые сколь угодно подробные справки о всех уже накопленных сведениях.

По-видимому, мы находимся на грани нового переворота в этой области. Этот переворот еще не свершен, так как не решены основные проблемы передачи машине человеческих знаний и хранения информации. Многие в этом направлении уже сделано, еще больше делается и, несомненно, уже в ближайшие годы мы будем освобождены от обязанности просматривать массу книг и журналов в поисках необходимых нам сведений.

Основной проблемой механизации хранения и поисков информации является, пожалуй, создание не самих информационных машин, а удобного алфавита и кода, которыми экономно, полно и точно можно было бы записать в «понятной» для машины форме все необходимые для передачи содержания подлежащих розыску предметов признаки и их комбинации. Дело идет о создании информационного машинного языка, такого, что, задав на этом языке вопрос машине-библиотекарю, мы получили бы в ответ либо спи-

сок литературы по заданному вопросу (в простейшем случае), либо микроотпечатки всей этой литературы или ее аннотации.

Чтобы представить себе всю сложность этой задачи, следует вспомнить, что при поисках литературы нам приходится призывать на помощь нашу эрудицию и знания предмета, помнить большое число синонимов и равнозначных выражений. Очевидно, что недостаточно снабдить каждую отрасль науки и техники четким однозначным индексом, чтобы решить эту задачу. Необходимо еще предусмотреть большое количество разнообразных связей между различными отраслями знания. Как правило, любая техническая проблема охватывает сразу несколько областей. Если, например, нас интересует вопрос о допустимой скорости резания при обработке металла резцом из твердого сплава, то полезные сведения по этому вопросу нам удастся обнаружить в разделах: технология холодной обработки металлов, станкостроение, химия сплавов, термодинамика, кристаллография. Больше или меньшее знание предмета определит успех поисков.

Если машина будет в состоянии запомнить все связи между отраслями знаний в достаточно компактной форме, то она будет в состоянии по одной характеристике или совокупности характеристик предмета или книги отыскивать и указывать все другие характеристики или совокупности характеристик этого предмета или печатного труда.

По-видимому, машинный информационный язык, удовлетворяющий самым широким требованиям, будет обладать некоторым грамматическим строем, подобным грамматическому строю человеческого языка, но более строгим, рациональным и компактным. Он будет позволять записывать сложные высказывания при достаточно малом количестве символов.

Потребуется осуществлять перевод с языка машин на национальные языки и дублирование аннотаций на печатные труды или самих трудов, т. е. параллельную запись на национальном и машинном языках. В наилучшем случае машинный язык должен быть интернациональным. По-видимому, машинный язык следует строить в коде, близком к оптимальному, с разумным использованием избыточности кода для снижения вероятности искажений.

Современные машинные системы индексов, которые применяются в уже существующих информационных машинах, далеки от совершенства, и, таким образом, проб-

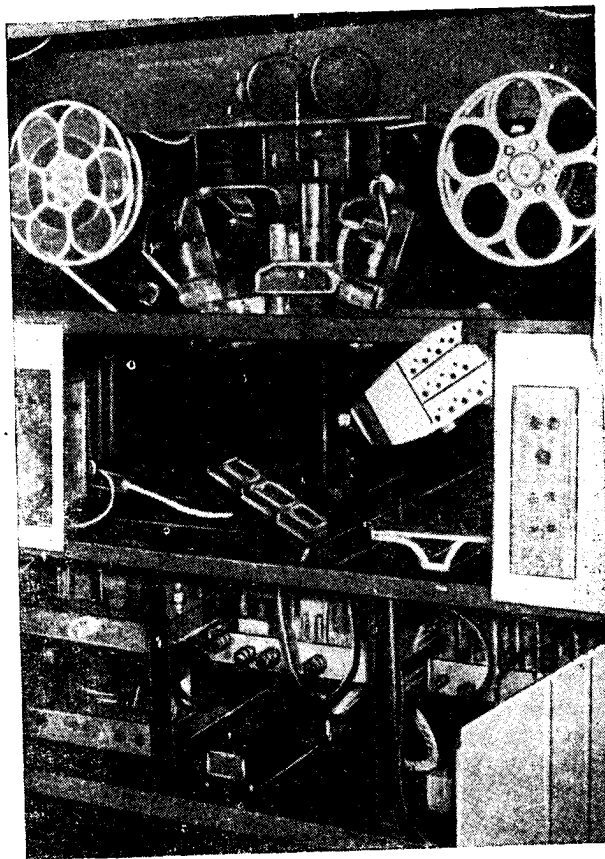


Рис. 8.15. Автоматический библиограф — машина, выбирающая из каталога аннотации книг и статей на заданную тему.

лема разработки машинного языка стоит во весь рост. Конструирование самих информационных машин представляет собой проблему более легкую, которая при наличии алфавита и кода может быть решена без принципиальных затруднений. Подобные машины находят применение пока только в узких специализированных областях с уже разработанной системой индексов, в которой обычно почти отсутствуют внутренние связи и, таким образом, весь язык описывает только перечень некоторого множества элементов.

Одна из нескольких действующих информационных машин представляет собой автоматический каталог — справочник. На киноленте двумя рядами отпечатаны: с одной стороны краткие аннотации статей в виде обычного текста, на другой — указатель содержания аннотированной статьи,

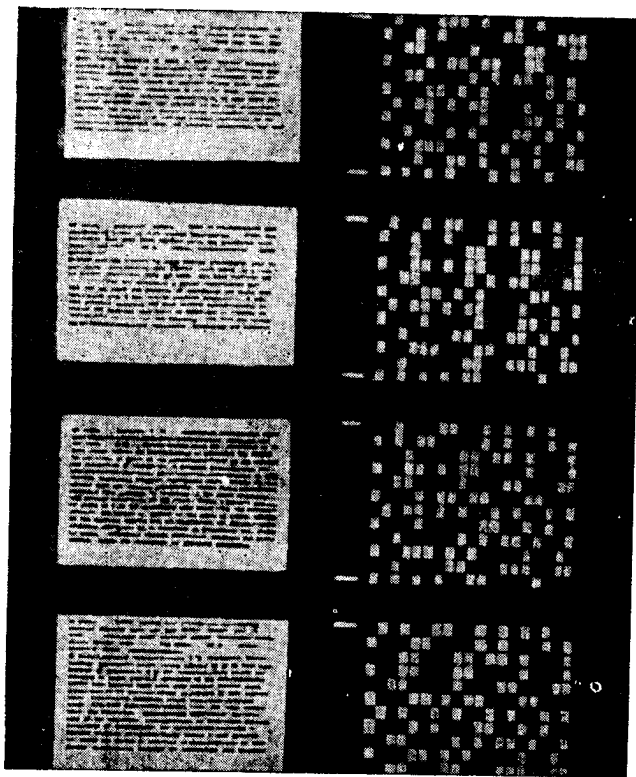


Рис. 8.16. Каталог на киноленте автоматического библиографа. Слева аннотации, записанные обычным текстом, справа — библиографический индекс, записанный двоичным алфавитом в виде белых и черных пятен и читаемый фотоэлементом.

записанный условным кодом в двоичном алфавите в виде белых и черных точек.

Кинолента, содержащая около 70 000 библиотечных карточек такого вида, просматривается автоматическим устройством с фотоэлементами за 6—7 минут. Перед просмотром вводятся шифры, соответствующие содержанию

отбираемых статей. При просмотре шифров фотоэлементы улавливают нужные комбинации в указателе, лента автоматически притормаживается и производится фотографическая печать соответствующих аннотаций, записанных обычным текстом. Таким образом, через 6—7 минут автоматический библиотекарь выдает библиографические справки по любому вопросу из имеющихся в его распоряжении материалов.

Целый ряд проблем возникает и в связи с тем способом фиксировать и запоминать информацию, которым мы пользуемся,—человеческим национальным языком, устным или письменным (рукописным и печатным). Машина способна работать с человеческим языком и выполнять полезные операции. Наиболее важной и сложной из них является автоматический перевод с одного национального языка на другой. Более простыми и, быть может, менее нужными являются машины, читающие печатный текст, стенографирующие устную речь и т. д.

Машина для чтения печатного текста была создана у нас и за границей во многих вариантах. Первая машина такого назначения — «Оптофон» Ф. д'Альба — была разработана еще в 1904 году. В настоящее время такие машины могут применяться не только для чтения текста слепыми, но и для ввода буквенного или цифрового текста в счетную машину.

Для чтения буквы печатного текста участок поверхности листа, на котором изображена буква, разбивают на элементы — прямоугольники. Каждый из элементов может быть белым или черным. Просматривая поочередно или одновременно все элементы изображения буквы, мы получаем двоичное число с числом разрядов, равным числу элементов, и с двумя возможными символами в каждом разряде: 0 (белое) и 1 (черное). Различные начертания (шрифты) одной и той же буквы дадут, вообще говоря, различные числа, но так как число вариантов начертания букв ограничено и ни один вариант начертания одной из букв не совпадает ни с одним из вариантов начертания никакой другой буквы, то каждая буква может быть однозначно опознана.

Элементы изображения буквы просматриваются с помощью миниатюрного устройства, имеющего в своем составе один или несколько фотоэлементов. Черное соответствует отсутствию тока в фотоэлементе, белое — наличию тока. Порядок просмотра элементов изображения может

быть различным: либо последовательно строка за строкой, как в телевизионной развертке, одним фотоэлементом, либо параллельно несколькими фотоэлементами, движущимися горизонтально по строкам или сверху вниз по столбцам. Выбирая тот или иной порядок просмотра, мы всякий раз получаем некоторый код, в котором кодируется выходной сигнал.

Для просмотра можно использовать либо движение всей головки вручную вдоль строки текста (при равномер-

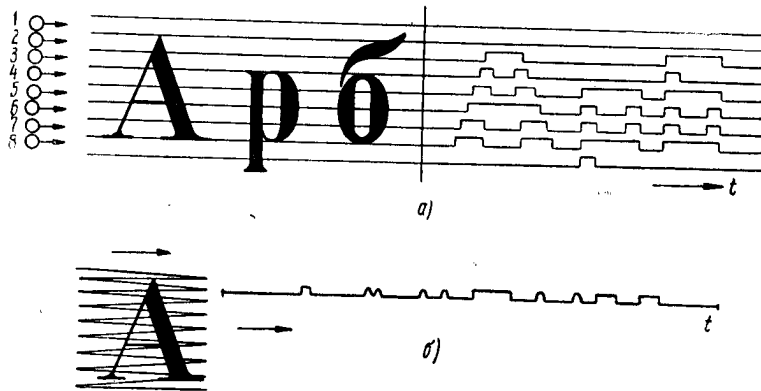


рис. 8.17. Считывание буквы типографского шрифта в автоматическом устройстве для чтения.

а) считывание одновременно восемью фотоэлементами по строкам (справа — сигналы в цепи каждого фотоэлемента); б) считывание одним фотоэлементом с помощью телевизионной развертки (справа — сигнал в цепи фотоэлемента).

ном освещении страницы), либо периодическое движение ветового пятна (по неосвещенной странице). Световое пятно, движущееся по листу бумаги, создается источником света, отраженным в качающемся зеркале, которое управляется электромагнитом, питаемым переменным током, или ветовым пятном, движущимся по экрану миниатюрной электронно-лучевой трубки, изображение которого проектируется оптически на поверхность страницы.

Сигналы «черное» и «белое» (точки и тире), снимаемые фотоэлементов головки по мере движения ее по строке г буквы к букве, дают в двоичном алфавите и в определенном коде изображение прочтенных букв. Эти сигналы — числа в двоичной записи — можно использовать по-разному. В простейших случаях их просто используют для непосредственного восприятия на слух. В одной из пер-

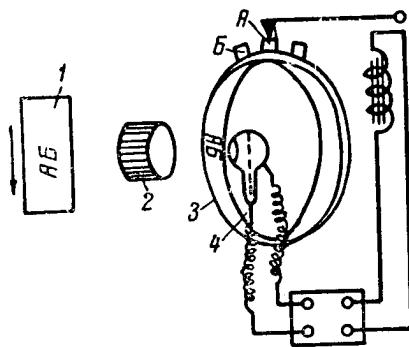
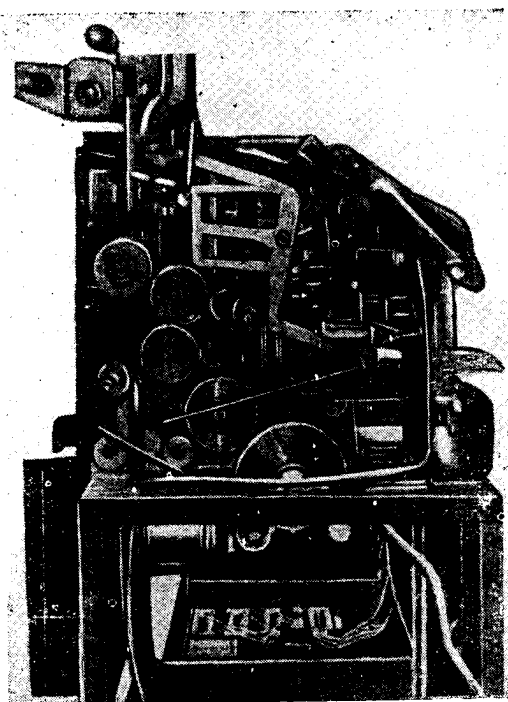


Рис. 8.18. Одна из первых машин для чтения (машина Г. Таучека) и схема ее работы.

Буква текста 1 проектируется на фотоэлемент 4 с помощью линзы 2 сквозь маску 3 на которой вырезаны отверстия в форме стандартных букв. Маска вращается пробегая полный оборот за время экспозиции одной литеры. При совпадении формы выреза в маске и проектируемого изображения буквы фотоэлемент полностью за темняется, что дает сигнал совпадения с полученным положением маски.

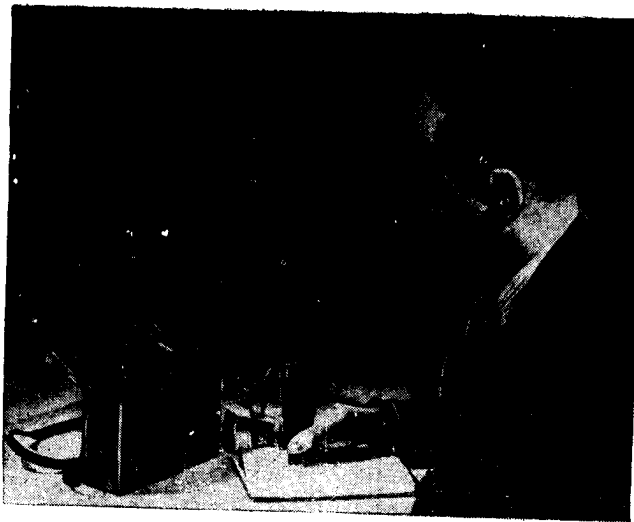


Рис. 8.19. Слепой за чтением текста с помощью современной, машины для чтения.



Рис. 8.20. Внешний вид одного из типов воспринимающей головки для чтения текста.

вых читающих машин буква просматривалась периодическим движением сверху вниз светового пятна от вибрирующего зеркальца при одновременном перемещении головки вправо по строке от руки. Синхронно с вертикальным движением пятна изменялся звуковой тон в телефоне, причем

звук появлялся только тогда, когда пятно проходило черный элемент. Каждая буква при чтении получала своеобразное звучание, состоящее из звуков разной высоты и длительности, комбинации которых нетрудно было различать на слух. Средняя высота тона позволяла скорректировать положение головки по вертикали.

Можно использовать точки и тире несколько более сложным способом и в результате вторичной перекодировки каждой букве поставить в соответствие замыкание оп-

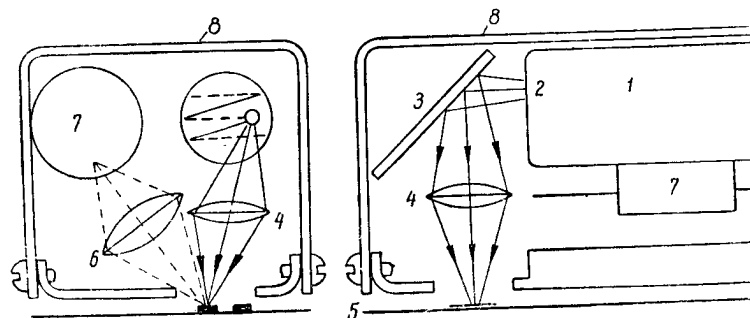


Рис. 8.21. Схема устройства воспринимающей головки для чтения текста.

1 — миниатюрная электронно-лучевая трубка, 2 — световое пятно на экране трубки движущееся по экрану как в системах телевидения, 3 — зеркало, расположенно наклонно, 4 — линза, фокусирующая изображение светового пятна на поверхность листа бумаги 5 с читаемым текстом, 6 — линза, отбрасывающая изображение освещенной части литеры на катод фотоумножителя, 7 — фотоумножитель (фотоэлемент с усилителем), 8 — кожух.

ределенного быстродействующего реле, которое включает считывание с магнитной ленты или барабана заранее записанного звука, соответствующего прочитанной букве. Такой вариант читающей машины тоже был осуществлен.

Можно, наконец, превращать сигналы букв не в звуки, а в конфигурации выпуклостей, применяемые при печати книг для слепых по системе Брайля.

Для осуществления двух последних способов приходится преодолевать некоторые трудности. Произношение слов или начертание их в системе Брайля неоднозначно определяется начертанием слова в печатном тексте. При вторичной перекодировке желательно читать и воспроизводить не буквы, а их сочетания — по две или три буквы — или даже целые слова, что не может не усложнить машину.

При проектировании машин для чтения приходится сталкиваться с тем фактом, что пропускная способность

различных каналов внешних восприятий человека (органов чувств) неодинакова. Наибольшей пропускной способностью обладает зрение, далее следует осязание, затем — слух. Перекодировка текста в алфавит Брайля для восприятия осязанием, оказывается, дает лучшие результаты в смысле скорости чтения, малой утомляемости, чем перекодировка в звуки, воспринимаемые на слух.

Как уже говорилось, можно использовать сигналы букв для ввода текста в машину, для чего, разумеется, необходима предварительная кодировка букв кодом машины.

Значительную трудность при чтении машиной представляет проблема различия размеров шрифта и точной установки головки на букву. Эти трудности обычно преодолеваются путем однократной регулировки машины под шрифт данной книги, например с помощью переключателя или посредством изменений размеров головки и с помощью вспомогательного трафарета для вождения головки по строкам. Это — решение простое и достаточно хорошее.

Можно было бы для регулирования положения головки в соответствии с расположением и размерами буквы применить автоматическое устройство с обратной связью, которое при развертывании световым пятном участка страницы реагировало бы на наличие белого пространства по краям буквы и автоматически совмещало головку с буквой так, чтобы белые поля с каждой стороны были одинаковы, а размер этих полей был приведен (регулировкой размера раstra) к некоторой стандартной величине. Однако такое устройство было бы сложным и дорогим.

Еще сложнее техническое решение проблемы чтения рукописи. В принципе она решается так же, как и чтение печатного текста, однако разнообразие начертаний букв, даже в пределах одного почерка, делает решение трудным.

Различение машиной рисунка буквы является частным случаем более общей проблемы различения формы вообще — на плоскости или в пространстве. Как машина может отличить, например, круг от квадрата? Речь идет о различении не двух конкретных фигур, а двух классов этих фигур. Как круги, так и квадраты могут быть любого размера, как угодно расположены и даже слегка деформированы. Машина должна рассортировать эти образы на два класса — «круги» и «квадраты» — так же четко и безошибочно, как это делает глаз человека. Для того чтобы достичь этого, геометрические образы должны быть приведены в совпадение с некоторым стандартом, для чего каж-

дый конкретный образ, или его запись в виде сигнала, необходимо подвергнуть ряду преобразований, состоящих в перемещениях, поворотах и растяжениях или сжатиях. Если набор этих преобразований будет достаточно полным, то при одной из их комбинаций любой круг или эллипс будет превращен в стандартный круг, т. е. круг, центрально расположенный, со стандартным размером; когда он совпадет со стандартным изображением и даст сигнал совпадения, он будет, таким образом, опознан как круг.

Проблема различения формы, кроме чтения текста, имеет и другие практические приложения. Одно из них состоит в автоматизации дешифровки аэрофотоснимков при составлении карт местности, для ускорения и удешевления трудоемкой операции определения рельефа местности по ее стереоскопическому фотоснимку. Для этого отдельные элементы обоих изображений стереоскопического снимка должны быть сопоставлены с точки зрения расхождения формы и размеров отдельных пятен. Это расхождение, поскольку оно однозначно связано с рельефом, используется для вычисления высоты каждой точки изображения. Каждая точка местности будет изображена на обоих снимках, но с некоторым смещением, из-за параллакса. Это смещение будет тем больше, чем ближе к фотоаппарату расположена данная точка местности. Если ось объектива при съемке была вертикальна, то смещение зависит от высоты точки. Таким образом, высота точки может быть определена по относительному смещению двух ее изображений на стереофотоснимке.

Для механизации этой операции оба снимка считываются по строкам точка за точкой с помощью фотоэлемента подобно тому, как читающая текст машина считывает букву, и записываются на магнитных лентах. При многократном воспроизведении записей с лент они могут сдвигаться друг относительно друга и, таким образом, приводить к совпадению отдельные элементы записи, т. е. пятна, изображающие отдельные участки местности. По степени сдвига автоматически вычисляется высота этого элемента. На выходе машины вычерчивается или вырезается из целлулоида рельеф, соответствующий данной строчке снимка. Складывая строка за строкой все вырезанные рельефы строк, получают рельефный макет сфотографированной местности.

Другим примером проблемы различения является различение звуков разговорного языка — фонем. Решени

этой проблемы, как и предыдущей, имеет непосредственное прикладное значение при создании протезов — аппаратов для понимания устной речи (например, принимаемой по радио или телефону) глухими. В дальнейшем автоматы, регистрирующие и различающие слова разговорной речи, вероятно, найдут много применений в качестве входных устройств роботов, которые, таким образом, окажутся способными выполнять устные команды.

Простейший протез для глухих представляет собой устройство, которое непосредственно записывает осциллограмму звуковых колебаний на флюоресцирующем экране. По виду осциллограммы не очень трудно научиться на глаз различать фонемы и целые слова. В этом случае восприятие формы кривой, записанной на экране, с помощью зрения заменяет работу слухового аппарата по различению звуков и слов.

Можно переложить эту обязанность на автомат другого устройства. Для этого звуки, воспринимаемые микрофоном, превращенные в электрические колебания и усиленные, поступают на ряд параллельно включенных узкополосных фильтров, настроенных на разные частоты. Каждая фонема речи имеет характерный для неё спектр и сопровождается появлением на выходных шинах узкополосных фильтров некоторой комбинации напряжений: на выходе одних фильтров (тех, которые настроены в резонанс со спектральными составляющими фонемы) напряжение присутствует, на выходе других — отсутствует. Эта комбинация возбуждаемых и невозбуждаемых шин соответствует фонеме.

Если меняется высота голоса, говорящего перед микрофоном, то та же комбинация появится выше или ниже на шинах фильтров подобно тому, как мажорное трезвучие может быть взято на фортепиано в разных тональностях. Решающее устройство реагирует на комбинацию выходов, а не на среднюю высоту тона. Поэтому одна и та же фонема, произнесенная различными голосами, будет одинаково зарегистрирована и превращена на выходе решающего устройства в сигнал данной фонемы, который может быть использован для приведения в действие исполнительного печатающего аппарата или же записан в устройстве памяти для дальнейшего различения слов и фраз. В практически осуществленном варианте такого устройства использовалось 20 фильтров с полосой пропускания около $\frac{1}{3}$ октавы каждый в диапазоне от 125 герц до 8 килогерц. Это

устройство хорошо различало звуки «т, к, м, н, с, ш, и, а, у» и не могло различать звуки «п, ф и р».

Однако такой метод различения фонем обладает по крайней мере двумя недостатками: во-первых, фонемы не всегда соответствуют правописанию, а во-вторых, некоторые фонемы в живой речи могут оказаться неразборчивыми и не будут зафиксированы. К исправлению этих недостатков есть пути. Грамматически правильное написание слова может быть восстановлено сравнением с образцами или выполнением некоторого набора стандартных правил. Для записи неразборчиво произнесенных фонем в одной из конструкций автомата этого назначения было предусмотрено специальное устройство. Оно состояло из набора потенциометров, которые заранее устанавливались в такие положения, которые соответствуют условным вероятностям появления различных фонем после данной. Таким образом, напряжение, соответствующее сигналу следующей фонемы, взвешивается по вероятности, и, если неразборчивая фонема плохо выявлена, оценка по вероятности в большинстве случаев позволяет избежать ошибки.

Проблемы грамматической коррекции записываемого текста принадлежат к кругу тех вопросов, которые решаются при автоматическом переводе текста с языка на язык. Успехи в этой области, достигнутые в последнее время, имеют большое принципиальное и прикладное значение. Правда, проблема литературного перевода любого текста еще не решена, но достигнутые успехи позволяют надеяться на разрешимость и этой проблемы.

Автоматический перевод

Перевод делается на электронной быстродействующей цифровой машине, управляемой специальной программой. Составление программы перевода производится один раз, и любая фраза переводится по одной и той же программе. Собственно, вся проблема автоматического перевода и сводится к составлению этой программы и словаря.

Прежде всего, текст должен быть введен в машину. Для этого он кодируется буква за буквой в двоичном алфавите, причем каждая буква изображается двухзначным десятичным числом (табл. 1), закодированным в двоичном алфавите.

Буквы, изображенные в двоичном алфавите, и слова изображаемые этими буквами, пробиваются в виде перфо

Таблица 1

Кодирование букв латинского алфавита с помощью чисел
(числа записаны в десятичной системе)

<i>a</i>	16	<i>i</i>	12	<i>q</i>	23	<i>y</i>	04
<i>b</i>	06	<i>j</i>	18	<i>r</i>	07	<i>z</i>	25
<i>c</i>	22	<i>k</i>	19	<i>s</i>	05	<i>.</i>	31
<i>d</i>	30	<i>l</i>	27	<i>t</i>	21	<i>,</i>	03
<i>e</i>	08	<i>m</i>	11	<i>u</i>	20	пробел	00
<i>f</i>	14	<i>n</i>	15	<i>v</i>	29		
<i>g</i>	10	<i>o</i>	28	<i>w</i>	13		
<i>h</i>	26	<i>p</i>	24	<i>x</i>	09		

раций на ленте, которую «читает» машина. Разумеется, пробивка производится специальным входным устройством с клавиатурой, на которой, как на пишущей машинке, оператор, могущий не знать языка, печатает, например, английскую фразу. Можно, по-видимому, использовать для этой цели и машину, читающую печатный текст.

Чтобы перевести фразу, прежде всего необходимо перевести слова. Для этой цели в устройство памяти электронной цифровой машины предварительно вводится специальный словарь. Словарь размещается в запоминающем устройстве машины и разбивается на английскую и русскую части. В ячейках памяти английской части словаря для каждого слова записываются следующие сведения (разумеется, кодом в двоичном алфавите): само слово (иногда только неизменяемая часть слова без флексии); грамматические сведения о слове, необходимые при переводе; адрес соответствующего русского слова. В русской части словаря в ячейках памяти, адрес которых указан в английской части словаря, записывается русский перевод

Таблица 2

Кодирование букв русского алфавита с помощью чисел
(числа записаны в десятичной системе)

<i>а</i>	16	<i>к</i>	19	<i>у</i>	20	<i>ю</i>	01
<i>б</i>	06	<i>л</i>	27	<i>ф</i>	14	<i>я</i>	03
<i>в</i>	13	<i>м</i>	11	<i>х</i>	26	<i>й</i>	18
<i>г</i>	10	<i>н</i>	15	<i>ц</i>	22	<i>ъ</i>	00
<i>д</i>	30	<i>о</i>	28	<i>ч</i>	23	<i>ь</i>	09
<i>е</i>	08	<i>п</i>	24	<i>ш</i>	09		
<i>ж</i>	29	<i>р</i>	07	<i>щ</i>	2		
<i>з</i>	25	<i>с</i>	05	<i>э</i>	17		
<i>и</i>	12	<i>т</i>	21	<i>ы</i>	04		

слова тоже кодом в двоичном алфавите, но для букв русского алфавита (табл. 2).

Слово, введенное в машину, после того как машина «прочла» это слово с перфоленты и записала его в регистре, сравнивается поочередно со всеми словами словаря. Быстрота вычислений электронной машины позволяет проделать это за короткий срок. Слово текста, записанное кодом в двоичном алфавите, попросту вычитается поочередно из слов, записанных в словаре. Машиной управляет при этом специальная часть программы — пользование словом. Как только разность равна нулю, т. е. слова совпали, из словаря считывается адрес русского слова, которое и может использоваться для перевода.

Однако описанная схема работы реально осуществляется лишь в редких случаях. Чаще дело обстоит гораздо сложнее. Многие английские слова не находятся в словаре сразу, и их нужно отыскивать по корням, отбрасывая окончания. Многие слова имеют не один, а несколько переводов, в зависимости от контекста, т. е. от места в предложении и от окружающих слов. Для таких слов в словаре вместо адреса русского слова стоит адрес дополнения к словарю, которое содержит необходимые данные для решения вопроса в этих случаях.

Решение возможно потому, что язык и его грамматика построены закономерно, по определенным правилам. Ведь язык представляет собой систему сигналов, отражающую в определенном алфавите и с помощью определенного кода внешние события, наши действия, идеи и понятия.

Сигнальное отображение (кодирование) возможно потому, что оно закономерно. И хотя эти закономерности многочисленны, разнообразны и сложны, тем не менее они существуют и потому могут быть выявлены и четко сформулированы.

Число грамматических правил, необходимых для перевода фраз, определяется характером текста. Чем обширнее круг понятий, заключенных в данном тексте, тем больше требуется грамматических правил. Технический текст обладает более строгой, «сухой» структурой, и для его перевода требуется меньше грамматических правил, чем для перевода художественных литературных произведений. Кстати, и словарь технической литературы меньше по объему. Из нескольких сотен тысяч слов языка в технических специализированных текстах используется обычно немного более тысячи «общих» слов и примерно столько же техниче-

ских терминов. Именно по этим причинам перевод технических текстов значительно проще. Первые переводы на машине выполнялись со словарем, состоявшим всего из 250 слов, и с шестью грамматическими правилами.

Если машина не находит в словаре искомого слова, то по окончании поиска (с помощью команды условного перехода) включается код программы отбрасывания окончаний. Слово проверяется на возможные окончания (*ing, ed, er, est, s, 's* и т. д.). После каждой проверки получается ответ *да* или *нет*, в зависимости от чего машина либо переходит к следующей проверке, либо производит отбрасывание найденного окончания и начинает новый поиск в словаре. Если и он не увенчается успехом и надежды найти слово в словаре не остается, машина просто печатывает в перевод непереведенное слово латинскими буквами.

Если перевод слова неоднозначен, т.е. слово в переводе имеет несколько значений, машина выбирает нужное значение, соответственно тексту. Для этой цели по указанию словаря дается команда перехода к подпрограмме проверки слов, окружающих данное слово. Делается перевод уже не одного слова, а кусочка фразы, например слова и предлога; переводится не слово *instance* (*инстанция, требование, пример*), а *for instance* (*например*). И в этих случаях проверки окружающих слов построены таким образом, что они производятся по очереди и каждый раз в результате получается ответ: *да* или *нет*.

Когда слово полностью определено, оно запоминается со всеми грамматическими характеристиками, необходимыми для перевода. После перевода всех слов английской фразы производится синтез русской фразы. Каждому русскому слову системой последовательных проверок грамматических признаков придается необходимая грамматическая форма. Специальная часть программы определяет порядок слов, другая — расстановку знаков препинания.

В опытах автоматического перевода машина порой совершает грамматические ошибки. Они происходят по вине программистов, забывших предусмотреть какую-либо грамматическую деталь в программе. Так, один из первых переводов с английского на русский, выполненный с помощью машины, содержал, по сообщению Д. Ю. Панова, следующую фразу: «.....*Это верно безусловно для обширной категории задач связанных с силой и движением, так то хотим ли мы знать будущий путь Юпитера в небе-*

сах или путь электрона.....» и т. д. Другая машина, переводя с французского, напечатала: «...Мы можем сделать...» Ошибки очень напоминают человеческие, каковы-ми они, в сущности, и являются.

В конце концов английское предложение, «впечатанное» в машину с помощью клавиатуры, отпечатывается машиной на ленте русским шрифтом в грамматически правильном русском переводе.

Особенно удобно и быстро печатный текст можно выводить из машины, используя специальные электронно-лучевые трубки: тайпотроны или характроны. В этих трубках, в общем подобных телевизионным кинескопам, электронный луч испытывает последовательно два отклонения, независимых друг от друга. Первое отклонение выводит луч на участок металлического экрана, в котором пробиты миниатюрные отверстия в форме знаков, цифр или букв. При этом луч как бы «продавливается» сквозь трафарет, и его поперечное сечение принимает форму отверстия. Второе отклонение выводит луч на желаемое место экрана. Луч дает светящееся изображение знака, пробитого на трафарете. Если последовательно проводить луч по строкам экрана и каждый раз пропускать его через нужный участок трафарета, мы получим на экране текст. Печать на экране тайпотрона производится практически мгновенно, причем изображение текста можно повторять и задерживать на экране произвольно долго, столько, сколько нужно, чтобы его прочесть или сфотографировать. Первое и второе отклонения независимы, и поэтому на любом месте экрана может быть изображен любой знак трафарета.

Задача машинного перевода с языка на язык сегодня еще не решена до конца. Грамматика человеческого языка слишком сложна для того, чтобы можно было ее формализовать для машины небольшим числом приемов. Однако сегодня уже доказана принципиальная возможность вложить в машину необходимые правила. Дальнейшая работа по усовершенствованию машинного перевода продолжается. Несомненно, что «правила машинной грамматики» могут быть и будут сформулированы для все языков. Возможно, некоторые языки окажутся особенно «удобными», с точки зрения формализации грамматики и они приобретут особое значение. Возможно также, что будет выработан некоторый международный «машинный язык», т. е. код, кодировка с которого на националь-

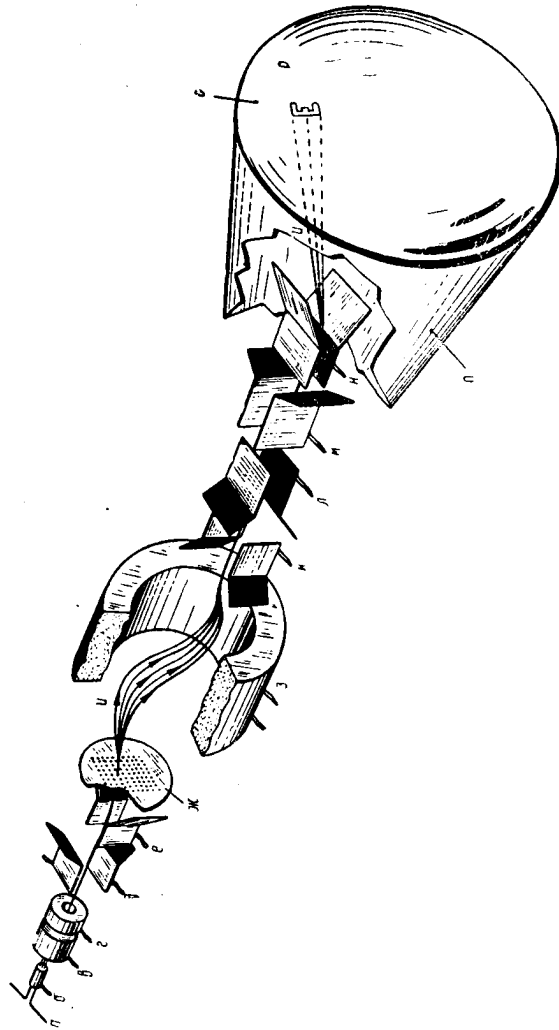


Рис. 8.22. Схема характронна.
а) подогреватель катода; б) катод; в, г) ускоряющие электроды электронной пушки; д, е) первая система отклоняющих пластин для выбора знака; ж) матрица с проемами отверстиями в виде знаков алфавита; з) катушка магнитной фокусировки; и) траектория электронов луча; к, л) вторая система пластин, компенсирующая отклонение луча, вызванное первой системой; м, н) третья система отклоняющих пластин для выбора места экрана, на котором появляется изображение выбранного знака; о) светящийся экран; п) колба; р) изображение знака на экране.

языки будет максимально простой, и, таким образом, международный обмен информацией будет значительно облегчен.

Весьма возможно, что проблема машинного перевода находится в настоящее время приблизительно в таком же состоянии, в каком находилась проблема механического воспроизведения речи до изобретения фонографа. Подобно тому, как тогда применялись громоздкие и малоэффективные артикуляционные машины, которые стали ненужными после замены принципа воспроизведения, так и в проблеме перевода, возможно, наступит переломный момент, когда окажется возможным непосредственно фиксировать смысл предложения универсальным лаконичным кодом в удобном алфавите, переходя от него к национальным языкам по стандартным правилам.

Если машина может переводить, то не может ли она сочинять? Это — законный и совсем не праздный вопрос. Пытаясь на него ответить, мы можем уяснить ряд интересных и не простых понятий.

Машина, работающая над переводом по той схеме, которая описана выше, конечно, не «понимает» ничего из того текста, который она переводит. Иначе говоря, в машину не заложено ничего, что давало бы ей возможность сопоставлять информацию, содержащуюся в тексте, с информацией, полученной ранее из каких-либо других источников. И переводя с французского фразу «*va t'en!*», машина напечатает «*уходи*», но никак не свяжет то, что она напечатала, ни с движением, ни с повелением, хотя в машине глагол записан, если не черным по белому, то «намагниченным по ненамагниченному» в повелительной форме, ни, наконец, с переживаниями действующих лиц, одно из которых хочет, чтобы другое удалилось, а второе вынуждено уйти. Словом, машина не «читает» текст так, как читает его разумный читатель, и она не ассоциирует слово ни с каким действием, предметом, событием. Она не могла бы поэтому, скажем, сыграть на сцене переводимую ею пьесу. Машина-переводчик не имеет необходимых для этого возможностей.

Однако в некотором смысле такая машина может, если не «сочинять», то «излагать» человеческим языком нечто, что ей задается в качестве «смысла». Если подойти к тексту, как к простому набору букв и слов, то, анализируя его статистически, можно установить наличие внутренних связей, которые вытекают, во-первых, из грамматики и, во-вторых,

из того, насколько часто используются в речи применяемые обороты, т.е. комбинации слов. Можно перекодировать любой текст так, что в результате он будет лишен всех или почти всех внутренних связей. Текст укоротится, энтропия его повысится, он будет содержать только эссенцию смысла. С точки зрения статистики, текст по своим свойствам приблизится к «белому шуму». Если теперь взять любую случайную последовательность символов, обладающих теми же статистическими свойствами, и произвести с ней перекодировку, обратную той, которую мы производили с исходным текстом, внося в код избыточность в соответствии со свойствами человеческого языка, то исходный «белый шум» приобретет структуру и качества фраз. Однако в качестве смысла этих фраз окажется нечто случайное, то, что мы взяли за исходный процесс. Такие фразы машина может «сочинять» или лучше сказать «излагать». Примерно так же работала машина Лапутянской академии у Свифта. В этом, в сущности, нет ничего обескураживающего. Мы вложили в машину случайную последовательность и получили текст со случайным смыслом. Если бы на вход машины было подано что-либо иное, представляющее собой соответствующим образом закодированную рациональную реакцию на внешнее воздействие, смысл полученной фразы был бы в соответствии с обстановкой и фраза не была бы бессмысленной. Машина могла бы реагировать на обстановку грамотными русскими высказываниями.

Сейчас уже не представляют редкости демонстрации сочиненной подобным образом музыки, например, в стиле Баха и Вивальди. Утверждается, что мелодии ковбойских песен хорошо имитируются с помощью случайного процесса марковского типа, т.е. процесса, у которого выбор следующей ноты зависит по вероятности только от непосредственно предшествующей.

И, наконец, последнее, о чем нам хотелось бы упомянуть, говоря о машинах, работающих с текстом,— это об автоматических шифровальщиках и дешифровальщиках секретных передач. Задача автоматической кодировки любым сложным (однократным или многократным), но известным кодом и раскодирования шифрограмм не содержит принципиальных трудностей, и сложность ее определяется лишь сложностью заданного кода.

Гораздо интереснее задача расшифровки секретных сообщений, закодированных неизвестным кодом. Задача

эта имеет интересную особенность по сравнению с другими видами обработки информации. Обычно любая система сигналов бывает заранее установлена или предварительно организована. В случае искусственных кибернетических систем эта организация (алфавит, код) предусматривается конструкцией системы. Естественные системы организуются в процессе существования (язык вырабатывается в связи с действиями, трудовыми процессами и т. д.). Отсутствие всякой предварительной организации, по-видимому, исключает возможность существования сигнальной связи двух систем. Мы могли бы понять сигналы, посланные с Марса, только в той мере, в какой жизнь на Марсе имела бы общие черты с нашей.

Расшифровка секретных сообщений с неизвестными кодами является задачей с минимальными предварительными сведениями о множестве возможных сообщений, алфавите и коде. При сравнительно простых кодах эту задачу решают обычно, опираясь на статистические свойства текста (частоту появления символов и их комбинаций). Расшифровка в этом случае не слишком трудна. В сложных кодах для того, чтобы изменить статистическую структуру, применяют кодировку по таблицам случайного набора символов. В этом случае задача расшифровки усложняется, однако не становится полностью неразрешимой. Расшифровка ведется, грубо говоря, пробами на разных кодах. По мере хода проб отсеиваются неверно выбранные коды и из всех возможных вариантов неизвестного текста постепенно оказываются более вероятными лишь некоторые. Успех расшифровки зависит от размеров текста (чем длиннее текст, тем легче его расшифровать) и, конечно, от количества проб. Расшифровка с неизвестным кодом напоминает поиски неизвестно где спрятанного предмета. Спрятавший может его отыскать сразу, но тот, от кого спрятан предмет, должен обыскать все возможные тайники и только тогда найдет искомое.

Работы по поиску неизвестного кода чрезвычайно трудоемки. На этом, собственно, и основывается «секретность» передачи.

Ясно, что механизация поисков неизвестного кода с помощью быстродействующей машины повышает производительность в десятки тысяч раз. Соответственно снижается и секретность кодов.

Интересен тот факт, что техника дешифровки шифрограмм имеет много общего с «расшифровыванием» надпи-

сей на древних, ныне забытых языках. И здесь один из основных методов состоит в выявлении частот комбинаций символов, их связей и статистической структуры. И в этом случае большая часть долгой и утомительной работы может быть поручена аккуратным, неутомимым быстродействующим машинам.

Несколько лет назад была сделана попытка поручить машине набор и размещение текста на газетной полосе, т. е. работу метранпажа. Отредактированный текст вводился в машину, куда подавалась также информация о размерах газетной полосы, отведенной под текст. Машина «сама» определяла тип шрифта, размеры и расположение заголовков и делала все необходимые для этого расчеты. Результаты работы машины автоматически реализовывались специальным станком, и газетный текст оказывался не только набранным, но и сверстанным в готовом для печати виде.

Другие работы

Оставим машины, работающие с текстом, и обратимся к другим, которые также успешно выполняют некоторые «человеческие» обязанности. Сейчас появляются все новые работы; возможности их создания весьма широки. Мы сможем упомянуть только некоторые из них.

Не так давно сообщалось о машине, заменяющей кассира железной дороги. Автоматический кассир не только получает деньги и выдает билеты и сдачу, но и исполняет все, что требуется от кассира: дает справки о расписании поездов и необходимых в пути пересадках, об опозданиях поездов, если такие случаются, дает справки о наличии мест в разных типах вагонов и, выдав билеты, «вычеркивает» номера проданных мест из списка, дабы не продать их дважды.

Человек в работе этой машины не принимает участия. Тем не менее все обязанности живого кассира машина выполняет и притом предельно четко, безошибочно, быстро, без промедлений и лишних разговоров.

В некоторых городах США на перекрестках больших улиц установлены автоматические регулировщики уличного движения. На подходах к перекрестку располагаются маломощные радиолокационные установки, которые отмечают каждую автомашину, направляющуюся к перекрестку, и одновременно регистрируют ее скорость (по эффек-

ту Допплера). Информация, собранная локационными установками, направляется по каналам связи в решающее устройство, которое по определенным правилам (алгоритму) определяет необходимый сигнал светофора в данной ситуации, исходя из того, чтобы свести к минимуму задержку машин на перекрестке. Применение автоматического регулировщика увеличило пропускную способность перекрестков примерно на 30% и освободило от необходимости держать на перекрестке полицейский пост. Очень важным, по-видимому, является тот факт, что автоматический регулировщик получает сведения о прибывающих на перекресток машинах до того, как машины скопятся у перекрестка в ожидании сигнала. Важно, конечно, и то, что в авторегулировщике заложен хороший алгоритм управления, полученный на основе научного решения задачи о сведении к минимуму времени ожидания машин.

Более серьезной обязанностью является диспетчерская служба, сопряженная с большой ответственностью и требующая постоянного напряжения внимания и подчас немалой находчивости и опыта. Посмотрим, как работает автоматический диспетчер аэропорта.

Приближающиеся к аэропорту самолеты обнаруживаются аэродромными радиолокаторами. Каждый новый самолет, появившийся в районе аэродрома, отмечается и «берется на учет». Координаты самолета, определяемые радиолокаторами, автоматически вводятся в счетно-решающую машину, которая, собственно, и является диспетчером. Машина с помощью радиолокаторов успевает следить за всеми самолетами от момента их появления в зоне наблюдения до посадки. Каждый самолет снабжается номером и записывается «в очередь на посадку». Поскольку для совершения посадки самолет должен оказаться в нужное время в нужной точке, подход его обеспечивается строгими указаниями. Исходя из его настоящего местоположения, высоты, скорости, для него вычисляется траектория подхода, которая ему и передается в виде команды с проверкой исполнения и корректировкой. Траектория подхода должна обеспечивать не только заход на посадку в соответствии с расчетом, но и безопасность подхода, поэтому она выбирается так, чтобы не было опасности излишнего сближения с другими самолетами на подходе или в зоне ожидания. Наконец, оставляются резервные «места» в очереди на случай, если один из прибывающих

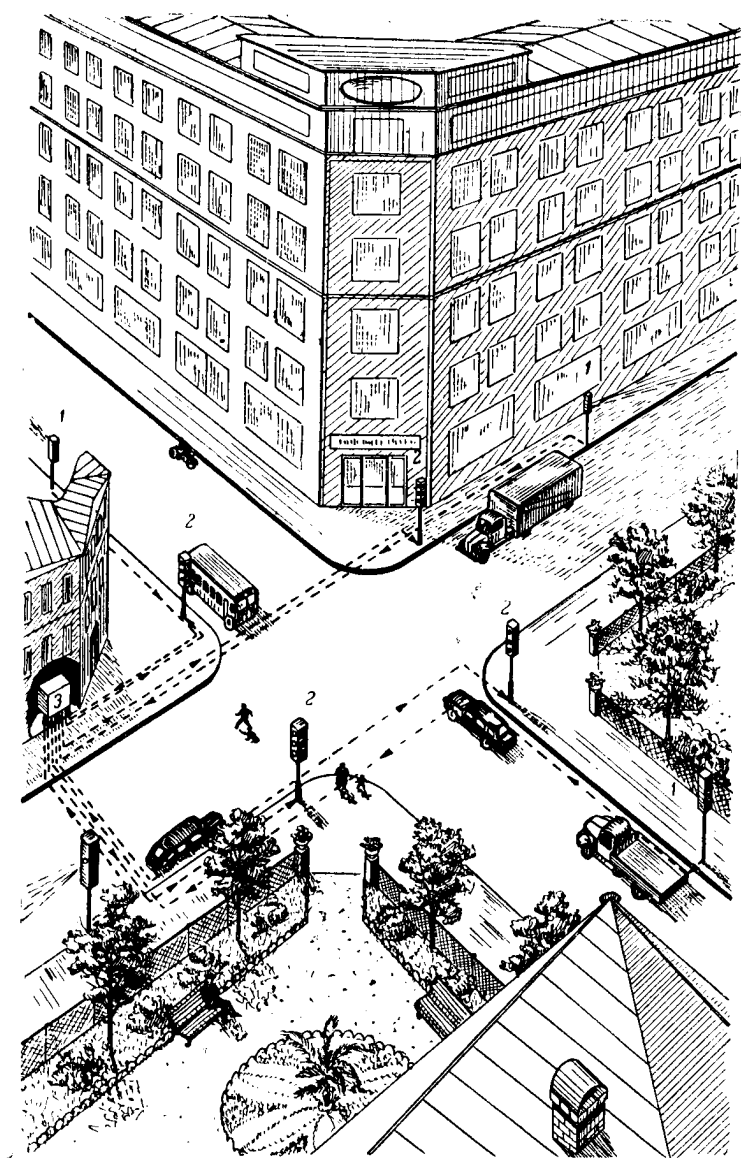


Рис. 8.23. Перекресток, обслуживаемый автоматическим регулировщиком движения.
1 — радиолокационные установки, 2 — светофоры, 3 — решающее устройство.

самолетов потребует срочной посадки из-за недостатка горючего или из-за аварии. Если «внеочередных» посадок не оказывается, в резервный момент может быть посажен один из ожидающих самолетов, для чего соответственно меняется его маршрут.

Ведение «текущей работы» по выводу на посадку многих десятков самолетов одновременно легко укладывается в производительность электронной машины, которая заменяет при этом работу многих диспетчеров аэропорта, ускоряет оборот порта и повышает надежность работы. Связь с самолетами поддерживается по радио либо командами пилоту и штурману, либо непосредственным управлением автопилотом. Посадка также может осуществляться автоматами вслепую, без участия пилота.

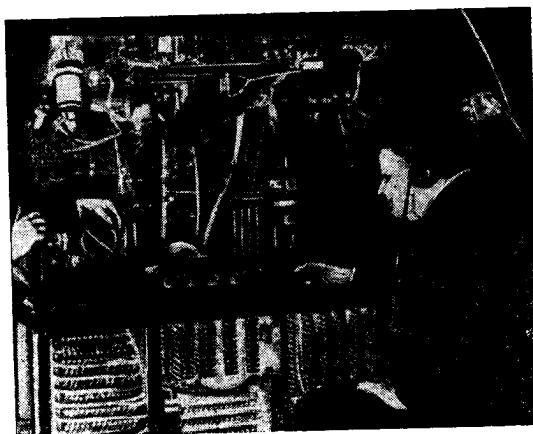


Рис. 8.24. Аппаратура, обеспечивавшая автоматическое пилотирование самолета от взлета до посадки (в процессе наладки).

Не только посадка самолетов в аэропорту, но и полет по трассе может быть полностью автоматизирован. Задача автоматического выдерживания заданного курса самолета с помощью автопилота решена давно. Однако это не решает полностью всех задач полета, так как автопилот сам требует управления (введения курса, который определяется штурманом по данным пеленга радиомаяков, наблюдению ориентиров, наблюдениям небесных светил и с помощью радионавигационных систем, дающих возмож-

ность сразу определять координаты самолета). Необходимые расчеты по данным измерений производит штурман. Расчеты эти сложны, но они производятся по строгим правилам и потому могут быть поручены машине.

Принципиально задача полностью автоматического пилотирования самолета от взлета до посадки не представляет особых трудностей; трудности сосредоточены в технической стороне дела: в создании устройств, сопрягающих выход навигационных приборов со входом вычислительной машины, в создании достаточно легкой бортовой машины, программировании ее работы и обеспечении высокой надежности работы всей системы. Первый полностью автоматический полет пассажирского самолета на дальнее расстояние был осуществлен в 1947 г. в США.

Транспортный самолет с пассажирами совершил десятичасовой перелет через Атлантический океан из Нью-фаундленда в окрестности Лондона. Полет начался единственным действием человека — нажатием стартовой кнопки, после этого, вплоть до приземления самолета и вывода его к месту выгрузки, «человеческое» управление полностью отсутствовало. В течение всего полета самолет управлялся роботом. На отдельных этапах полета менялась программа управления, менялись маяки и ориентиры, и эта смена управления, так же как и посадка, производилась автоматически.

С тех пор проблема вождения самолета без человека-пилота разрабатывалась и продолжает весьма активно разрабатываться, и в этом направлении достигнуты новые успехи.

Пилотирование самолетов не является, конечно, единственной областью, где необходимо производить замену человека роботом. Но это та область, где очень ярко выступает необходимость такой замены, где человеческие слабости явно тормозят использование возможностей техники. Человек на борту скоростного высотного самолета слишком медленно реагирует, слишком быстро утомляется, требует слишком много забот о себе (аппаратуры для обеспечения его работы в условиях пониженного давления, недостатка кислорода, чрезмерных нагрузок при ускорениях, низкой температуры и т. д.). Человек «не согласован по техническим характеристикам» с современным самолетом, и для «согласования» требуется ряд дополнительных тяжелых, сложных и дорогих устройств. И даже эти устройства (скафандры, герметизированные

кабины, отопление и т.п.) не могут решить задачу до конца. Для грузовых перевозок, для боевых самолетов пора заменить человека на борту более выносливым, менее медлительным, неутомимым роботом. Это и происходит на наших глазах.

Не менее насущные потребности того же рода имеются и в промышленности радиоактивных веществ и в ядерной энергетике. Здесь также роботы, нечувствительные к биологически вредным излучениям, должны стать на место человека.

Говоря о роботах «высшей квалификации», выполняющих очень сложную и ответственную работу по координации действий большого коллектива людей и механизмов, таких, например, как диспетчер аэропорта, нельзя не упомянуть о некоторых социальных аспектах этой проблемы. Диспетчер аэропорта или промышленного предприятия или офицер противовоздушной обороны (его обязанности по своему характеру мало отличаются от обязанностей диспетчера аэропорта, который обеспечивает «обслуживание» приходящих самолетов) является лицом ответственным, в юридическом смысле этого слова, и эта особенность не только входит органически в техническую сторону организации его работы, но и составляет социальную сторону вопроса. На месте диспетчера аэропорта немислим работник, не несущий ответственности за свои действия. Это исключило бы возможность пользоваться аэропортом.

И если роботы по мере их «роста», по мере развития теории автоматов, электроники и других отраслей знаний будут все более и более полноценно заменять человека на самых сложных операциях, то это не значит, что человек сможет полностью устраниваться и оставить робота в одиночестве выполнять свои функции. Вопрос об «ответственности» робота за качество выполняемой им работы лежит вне сферы технических наук, а в области социального использования техники, и даже сама постановка этого вопроса, по-видимому, представляет нелегкую задачу.

Именно как ответственное лицо человек является необходимым звеном в работе робота, и это место, по-видимому, не может быть занято никем другим. Робот может очень быстро, точно и безошибочно думать за человека, подсказывать ему самые сложные, точные и безошибочные решения, но принимать ответственные решения или, иначе, одобрять или отвергать решения ро-

бота, может только человек — юридически ответственное лицо, командир производства или офицер армии*.

Это обстоятельство первостепенной важности накладывает некоторые требования на деятельность, а следовательно, и на устройство робота. Прежде чем действовать, робот обязан предъявить свое решение человеку-контролеру для одобрения и критики. В некоторых случаях это будет не одно решение, а несколько возможных решений, из которых человек выберет одно, исходя из соображений, учет которых по тем или иным причинам не заложен в машину.

Следовательно, робот обязан предъявлять человеку некоторую информацию о своей работе. Естественно, что эта информация не может быть полной, ибо информационная пропускная способность, или, говоря проще, скорость работы робота, в тысячи раз превышает соответствующую способность человека. Человек просто не успеет ничего «прочитать» или увидеть и понять. Информация, которую робот обязан предъявить, должна быть очень краткой и представленной максимально наглядно и понятно для человека. Обратное воздействие человека на работу робота также должно быть максимально простым, например, сведено к нажатию сигнальной кнопки.

Необходимость общения человека и робота в ходе работы ставит целый ряд проблем по отбору нужной информации, ее перекодированию в «человеческий» алфавит

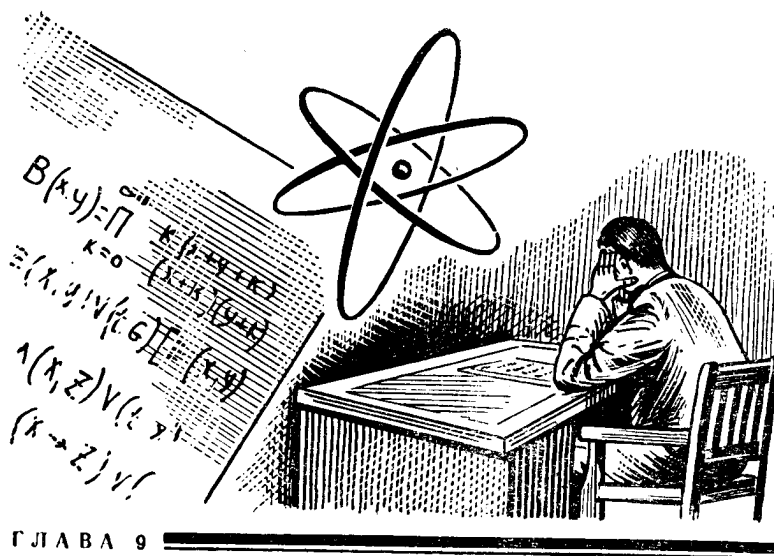
* В настоящее время, когда кибернетическая техника еще только выходит из своей «колыбели», контроль за работой машин особенно необходим. Это иллюстрируется забавным, хотя и не очень приятным, происшествием на открытии Всемирной выставки в Брюсселе 18 апреля 1958 года (Паэзэ Сэра, 19 — 20/4 1958).

Приезжие обращались за направлением в гостиницу в посреднические бюро, расположенные в большом числе возле вокзалов и аэропорта. Бюро сносилось по телефону с центральной конторой, которая обслуживалась электронной счетной машиной. Машина, получив запрос, отыскивала в памяти адрес ближайшей к месту, откуда поступил запрос, гостиницы, где имелись свободные номера требуемого типа (с ванной или без ванны, одиночные, парные и т. п.) и выдавала адрес. Адреса гостиниц и загрузки хранились в памяти машины и непрерывно корректировались по мере прибытия новых постояльцев. Все было хорошо кроме... надежности машины. Случилось так, что машина из-за неисправности, которую нельзя было сразу заметить, стала выдавать адреса з а н я т ы х номеров, вместо свободных. К вечеру город оказался в смятении из-за скопления около 50 000 вновь прибывших, которые не могли найти себе пристанища по вине безответственной машины. Неисправность была устранена только к утру следующего дня. (Примечание при корректуре).

из алфавита машины, очищению информации от шума, а вместе с этим и вопросы о скорости восприятия человеком различных типов представления информации. С новой остротой возникают старые вопросы о работе оператора с индикаторами и о реакции оператора на различные внешние раздражения. Только теперь оператор работает не с элементарными сведениями, получаемыми, например, от радиолокатора или по телеграфу, телефону и т. д., а с сообщениями о результатах решения очень сложных, запутанных задач, которые «подсказывает» ему робот и на проверку которых у человека ушло бы, быть может, несколько суток драгоценного времени.

Мы снова сталкиваемся с проблемой «перевода» с языка машины на «язык» восприятия человека, и на этот раз проблема по своему значению оказывается выросшей на две головы выше. В данном случае речь идет не только о переводе информации, т. е. перекодировании ее в широком смысле слова, но и о селекции только строго необходимых сведений, о согласовании пропускной способности человека и робота, об устранении всякой избыточности сведений.

Это — проблемы сегодняшнего дня. Пока наши роботы делают первые шаги и мы только учим их разговаривать с нами толково, лаконично и корректно, а вместе с ними многому учимся и сами.



ГЛАВА 9

МЫСЛЬ

В работе кибернетических систем управления и связи почти всегда так или иначе принимает участие человек. До появления автоматов и технических средств связи человек был единственным агентом управления, своими руками направлявшим и силу животных, и тепло, идущее от огня, и воду рек на поля. С тех пор уже многие операции управления переданы человеком в руки автоматов, и человек теперь направляет свою деятельность не только непосредственно на управление используемой им энергией, но и на создание управляющих систем и их регулировку.

Если посмотреть на деятельность человека с самой общей точки зрения, то можно увидеть, что по существу всякая производственная деятельность человека представляет собой управление силами природы в его собственных биологических и социальных интересах. Обработка почвы, разведение скота, строительство плотин, электрических станций, добыча полезных ископаемых, обработка сырья и т. п. — все это есть не что иное, как направление естественных природных процессов по некоторому желательному

руслу, то есть, в конечном итоге, управление ими с целью извлечения пользы.

Мы уже говорили, что характерным свойством всякого процесса управления является то, что энергия управляющего сигнала может быть ничтожно мала по сравнению с энергией, подчиняющейся управлению. Именно это свойство управляющих систем влечет за собой тот факт, что производственная деятельность человека всегда дает «энергетический выигрыш», приносит энергии больше, чем затрачивает, или, если не ограничиваться только рассмотрением энергии, извлекает из природы больше «материальных благ», необходимых для существования человеческого рода, чем затрачивается тех же «благ» человеком, производящим работу. Именно об этом свойстве человека по существу говорил К. Маркс, когда он писал о рабочей силе, которая в условиях капитализма является товаром: «...решающее значение имела специфическая потребительная стоимость этого товара, его свойство быть источником стоимости, притом большей стоимости, чем имеет он сам». (К. Маркс. «Капитал», т. 1).

С этой точки зрения введение автоматизации и механизации производственных процессов не только в смысле замены мускульной силы людей при транспортировке грузов, на строительстве и в других физических работах, но и в смысле замены индивидуальной человеческой сообразительности, опыта и т. п. обобщенным научным опытом, вложенным в машину, замена человека роботом в процессах управления, естественно, приводит к тому, что ту же работу может сделать меньшее число людей или те же люди в то же время сделают гораздо больше, т. е. повысится производительность труда человека.

С этой общей точки зрения научно-исследовательская деятельность человека представляет собой не непосредственно производственную деятельность, а процесс собирания и обработки информации, которая затем (может быть очень не скоро, но с тем большим эффектом) будет использована в управлении природой для достижения большей пользы при меньших затратах. Конечной целью науки является познание, т. е. описание наблюдаемых фактов, их систематизация, выявление их связей и путей к полезному промышленному использованию накопленных сведений.

Итак, люди, энергетические машины и кибернетические машины-роботы работают совместно на пользу человеку

Человек в этом коллективе играет особую, руководящую роль, он может делать многое такое, чего не могут (если угодно, «еще не могут») делать роботы. В частности, человек должен исполнять те обязанности, которые нельзя поручать роботам и за которые он сам обязан нести ответственность.

Огромное количество операций в равной степени может выполнять и робот, и человек, и по характеру выполнения этих операций, по качеству результатов работу человека-оператора и робота можно сравнивать. Иногда мы бываем вынуждены проводить такие сравнения. Как же выглядит работа человека рядом с работой автомата?

Человек у машины

Пусть, например, человек сидит на вышке деррик-крана, который поднимает грузы со строительной площадки на стены строящегося здания. Под его руками три рукоятки; каждая из них включает один из трех моторов и имеет три положения: среднее — «выключено», правое — «вперед» и левое — «назад». Один из моторов вращает лебедку, поднимающую груз вверх или опускающую его вниз, второй вращает стрелу крана вправо или влево, третий движет весь кран по рельсам вдоль строящейся стены вперед или назад. Крановщик получает от своего помощника, сигнализирующего взмахами флажка, команду, показывающую, откуда и куда нужно перенести груз, и выполняет этот сигнал-команду. Он нажимает рукоятки, делая это подчас почти не задумываясь. Кран движется, крюк, на который цепляют груз, подходит к нужному месту площадки и останавливается. Помощник закрепляет груз и дает новую команду. Снова кран приходит в движение и перемещает груз.

Крановщик не может работать с закрытыми глазами, не глядя на груз. По мере движения груза, наблюдая за его положением, крановщик включает и выключает моторы в зависимости от положения и скорости груза так, чтобы в конце концов действительное положение груза совпало с желаемым. Обратная связь осуществляется с помощью зрения крановщика и его рук, воздействующих на ручки пуска моторов.

Мы уже знаем, что буквально то же самое может делать автомат с обратной связью. Неловкий, неопытный или невнимательный крановщик, остановив или пустив не вовремя мотор может вызвать раскачивание груза, не

безопасное для окружающих. Чтобы избежать этого, он должен научиться гасить возникающие колебания, рассчитывая моменты включения и выключения моторов, исходя из скорости груза, его ускорения, длины троса, величины груза и т. д. Этому нужно научиться, и научиться этому не очень легко.

Правильно сконструированная сервосистема с достаточным числом обратных связей будет делать то же самое без ошибок. Для этого ее структура должна быть подобрана соответствующим образом, исходя из заданной допустимой величины колебаний груза. Если мы создадим такую следящую систему, которая перемещает груз по программе, определяемой сигналом-командой, и обеспечивает отсутствие опасной раскачки груза, и будем рассматривать ее отдельно от крана и груза, то мы сможем измерить ее характеристики и, если нужно, изменить их и отрегулировать систему.

Теперь обратимся к крановщику. Он должен делать и делает то же, что наш автомат. Он получает внешние воздействия — информацию о судьбе груза — и сам воздействует на кран. Мы можем описать действия крановщика в терминах и символах теории автоматического регулирования и, подвергнув его работу объективным измерениям, определить его «функцию передачи» совершенно так же, как если бы на его месте был автомат. Это не абстрактная возможность, это технический прием, которым широко пользуются инженеры при расчете полуавтоматических систем регулирования, в работе которых участвует человек. Ведь для того, чтобы разумно и грамотно построить систему, необходимо знать характеристики всех звеньев, в нее входящих, в том числе и характеристики человека.

Какими же особенностями обладают функции передачи «человеческого звена» следящей системы по сравнению с характеристиками мертвых звеньев, выполняющих ту же работу?

Во-первых, характеристики человека и автомата оказываются близкими друг другу. Иного, собственно говоря и ожидать нельзя, поскольку практически совпадают выполняемые функции. Во-вторых, характеристики человека не являются строго постоянными. От раза к разу человек несколько изменяет свою реакцию, сам того не замечая. Таким образом, функция передачи человека подвержена случайным изменениям и потому о ней приходится делать

утверждения лишь в статистическом смысле, как о средней величине с некоторым разбросом.

Кроме случайных отклонений от среднего, функция передачи человека подвержена также и систематическим изменениям. Она меняется по мере обучения, сначала медленно, затем быстрее, после этого снова замедляет изменение и приближается к некоторому стационарному значению. Об этом установившемся значении мы и говорили выше. Процесс же изменения функции передачи, скорость изменения и пределы изменения часто используются как характеристики скорости обучения или способности субъекта к обучению. Для автомата эта кривая бывает «пройдена» еще до его рождения; мышшь Шэннона проходит обучение за один урок; человеку нужен для этого длительный «производственный стаж».

Далее, функция передачи меняется от субъекта к субъекту. Она меняется также для одного и того же оператора в зависимости от настроения, состояния здоровья, утомления и других внешних и внутренних факторов. Однако эти изменения не меняют вида функции передачи, а дают лишь количественные отклонения от среднего.

Самым интересным свойством «человеческого звена» являются не эти сравнительно мелкие особенности, а другое свойство, имеющее весьма важное значение. Человек-оператор может посредством обучения достичь того, что его функция передачи может приблизиться к любому (разумеется, в некоторых пределах) наперед заданному виду. Иначе говоря, человек не обладает одной определенной функцией передачи, он может «настроиться» или научиться работать в соответствии с любой функцией из некоторого класса. Таким образом, человек, хотя и требует времени на обучение, но зато оказывается «универсальным» звеном.

В зависимости от поставленной перед ним задачи человек может работать как звено с линейной или нелинейной характеристикой, как дифференцирующее или интегрирующее звено и т. д. Разумеется, он при этом не может выйти за некоторые естественные пределы, например, повысить скорость реакции сверх свойственной ему максимальной скорости или давать меньшую ошибку считывания, чем та, которая соответствует остроте его зрения и т. д. Оставаясь в количественном отношении ограниченным, человек обладает чрезвычайно гибкими характеристиками и потому легко приспосабливается к любой машине

и любой операции. Правда, различные операции требуют от него различного напряжения сил: физических и интеллектуальных (внимание), и потому вызывают большее или меньшее утомление. Поэтому, несмотря на наличие ценных универсальных возможностей оператора к приспособлению, при проектировании систем необходимо «согласовывать» характеристики машины с параметрами человеческого интеллекта так же, как согласуют геометрические размеры частей машины и необходимые для управления ею усилия с размерами человеческого тела и его энергетическими возможностями.

Если задача оператора, например, состоит в том, чтобы подобно крановщику приводить к совпадению некоторый наблюдаемый им предмет (не обязательно груз, висящий на крюке, это может быть подвижная метка на индикаторе или перекрестье прицела орудия) с желаемым его положением (местом опускания груза, стрелкой индикатора или с целью) и при этом оператор воздействует на штурвал, то будет справляться со своей задачей как в случае, когда положение штурвала будет определять положение управляемого им объекта, так и в случае, когда положение штурвала соответствует скорости или ускорению или комбинации положения, скорости и ускорения объекта. В каждом случае оператор должен «научиться», «привыкнуть», «приобрести опыт», т. е. довести свое умение до такой степени, чтобы ему не нужно было размышлять над тем, какое действие предпринять, чтобы реакция возникла «автоматически», без размышлений, раньше, чем сам он успеет «сообразить», что он делает.

Задача шофера, сидящего за рулем — вести машину по отведенной ему части дороги. Случайные отклонения от желаемого пути из-за поворотов дороги, ее неровностей, ошибок самого шофера он ликвидирует поворотом баранки в нужном направлении. При этом угол поворота пропорционален кривизне траектории машины, т. е. приблизительно второй производной функции, определяющей положение машины на полосе дороги. Таким образом, шофер воздействует непосредственно на вторую производную регулируемой величины. Поэтому для того, чтобы, например, сохранив направление движения, «принять» на метр влево, шофер должен повернуть руль четыре раза (второе и третье движение обычно сливаются, так как производятся в одном направлении). Стрелок же для

того, чтобы переместить точку прицеливания должен сделать всего одно движение, если его штурвал воздействует на положение ствола, или два движения, если штурвал определяет скорость поворота.

Стрелок, крановщик, водитель транспорта и любой другой оператор, будучи «живым звеном» в системе с обратной связью, выполняет свою задачу тем лучше, чем меньше он проявляет свои разнообразные человеческие способности, чем больше его работа напоминает работу автомата, чем меньше он рассуждает и отвлекается. Ошибки в работе оператора порождаются как внешними, случайными, так и внутренними причинами. Шофер может совершить ошибку не только потому, что он вынужден был отвлечься, чтобы посмотреть на что-то, происходящее на краю дороги, или потому, что его отвлекли вопросом, или в глаз ему залетела мошка, но и просто потому, что он поддался посторонним мыслям или воспоминаниям, переключил внимание на заботы, не связанные с выполняемыми им обязанностями.

Действительно, хорошо работает в качестве шофера или вообще оператора у машины только тот человек, который работает, как автомат. Есть очень эффективный, хотя и рискованный, способ продемонстрировать автоматизм реакций при управлении велосипедом. Попробуйте, если вы хорошо ездите на велосипеде, на ходу переместить правую руку на левую ручку руля, а левую — на правую ручку и, перекрестив таким образом руки, управлять рулем велосипеда. Обычно вначале дело идет хорошо две или три секунды, пока велосипед может двигаться без управления. Но как только произойдет первое вынужденное движение рулем, оно окажется направленным в сторону, противоположную той, в которую следовало его сделать. Автоматические реакции «перепутанных» рук осуществляют не отрицательную, стабилизирующую, а положительную обратную связь, и велосипед неизбежно опрокидывается. Разумеется такой эксперимент может предпринять только закаленный спортсмен, способный перенести небольшую аварию. Прodelывать такой опыт всякому не рекомендуется.

Итак, в полном смысле слова сознательное, рассудительное мышление действует только в период обучения. При выполнении привычных операций действуют автоматические, быстрые реакции, за которыми детальный анализ рассудка попросту не успевает. При этом автомати-

ческие реакции совсем не обязательно являются элементарно простыми: шофер, ведя машину в потоке транспорта, не только учитывает положение соседних машин в данный момент, но и экстраполирует их положение на последующие моменты времени, осуществляя предсказание будущей ситуации и предпринимая те или иные действия (тормозит, ускоряет, позорачивает) в зависимости от того, что «должно» случиться спустя несколько секунд. Такое же предсказание будущего осуществляет подсознательно пешеход, переходящий улицу в потоке транспорта (в нарушение правил движения). То же делает охотник, стреляющий в летящую птицу и выносящий точку прицеливания вперед, «упреждая» цель. Ту же задачу решает прибор управления артиллерийским огнем.

Все эти и подобные им естественные или, чаще, приобретенные обучением и опытом автоматизмы, т.е. определенные реакции на определенные раздражения, которые регулярно повторяются от опыта к опыту и потому имеют характер устойчивости, осуществляются только потому, что в организме протекают определенные, закономерные и стабильные процессы, начинающиеся восприятием и кончающиеся реакцией. Функционирование, работа организма человека как хорошо построенного автомата протекает на базе материальной структуры, в которой происходят материальные физико-химические, физиологические процессы. Функции человека, его реакции связаны со структурой его нервной системы. Со строением нервной системы связаны, конечно, не только автоматические реакции, но и процессы обучения и все психические явления, как бы сложны они ни были. Вряд ли это нужно доказывать, так как всякому известно, что нормальная сознательная активность человека или животного возможна только при нормальном здоровом состоянии его нервной системы. Любое повреждение, заболевание, отравление системы нервных клеток обязательно влечет за собой расстройство функций нервной системы с соответствующими внешними проявлениями. Точно так же ведет себя машина: она не может нормально функционировать, если ее структура нарушена.

Мы уже говорили в гл. 5 о работе нервных волокон и клеток при передаче сигналов — нервных возбуждений. В гл. 6 мы упоминали о нервных структурах, ответственных за простейшие врожденные автоматизмы — безусловные рефлексы. Теперь мы вкратце остановимся на некоторых чертах строения нервной системы в целом.

Строение нервной системы

Мы уже говорили, что нервная система, как и любой другой орган живого тела, состоит из отдельных клеток. Нервная клетка — нейрон, как и всякая другая клетка, имеет ядро, окруженное протоплазмой. Размеры тела клет-

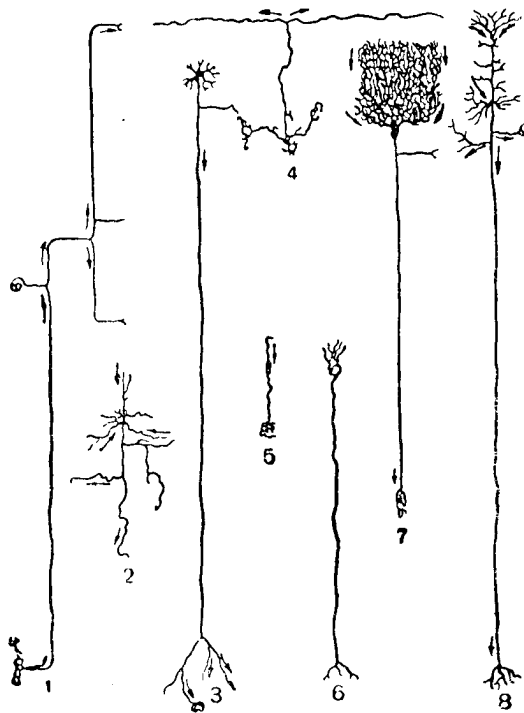


Рис. 9.1. Нейроны различных отделов нервной системы позвоночных.
1 — нейрон спинального ганглия, 2 — нейрон с коротким аксоном, 3 — моторный нейрон передних рогов спинного мозга, 4 — нейрон мозжечка, 5 — периферический афферентный нейрон обоняния, 6 — нейрон ганглия автономной нервной системы, 7 — клетка Пуркинье мозжечка, 8 — пирамидный нейрон коры полушарий головного мозга.

ки порядка 0,1 мм. От тела клетки отходят один или несколько отростков, имеющих диаметр порядка 0,01 мм и длину от долей миллиметра до метра и больше. Обычно только один отросток (аксон) имеет большую длину и осуществляет связь данного нейрона с удаленными клетками. Аксон имеет короткие ответвления (коллатерали). Остальные отростки данной клетки (кроме аксона), на-

зываются древовидными отростками (дендритами), имеют небольшую длину и сильно ветвятся. Они обеспечивают контакты с аксонами других нейронов.

Нервные волокна — аксоны, собираясь в пучки, образуют более или менее толстые нити жемчужно-белого цвета.

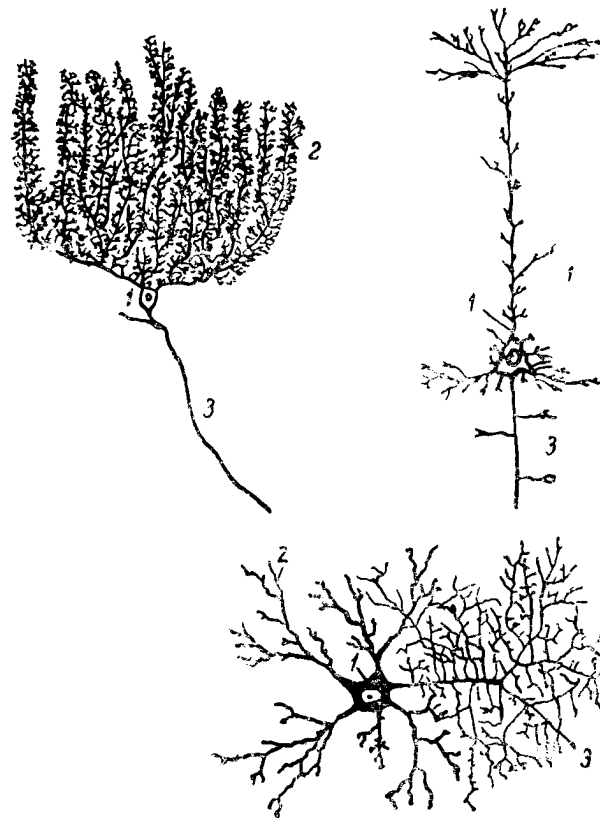


Рис. 9.2. Нейроны различных отделов нервной системы позвоночных.
1 — тело клетки, 2 — дендриты, 3 — аксон.

та — нервы. Нервы пронизывают ткани тела животного и соединяют между собой скопления нервных клеток, а также нервные клетки с чувствительными окончаниями (рецепторами) разного рода и с исполнительными клетками (эффекторами). Таким образом, составными частями или простейшими элементами нервной системы являются: чувствительные окончания — рецепторы (осязательные

или тактильные, болевые, химические, световые и т. д.), расположенные как на поверхности тела, так и вблизи нее, нервные клетки и волокна — нейроны, места соединений нервных волокон различных клеток между собой (синапсы) и нервные окончания в клетках мышц и желез.

Число типов нейронов довольно велико. В разных частях нервной системы преобладает тот или иной тип нейронов, имеющих, очевидно, некоторую специализацию.

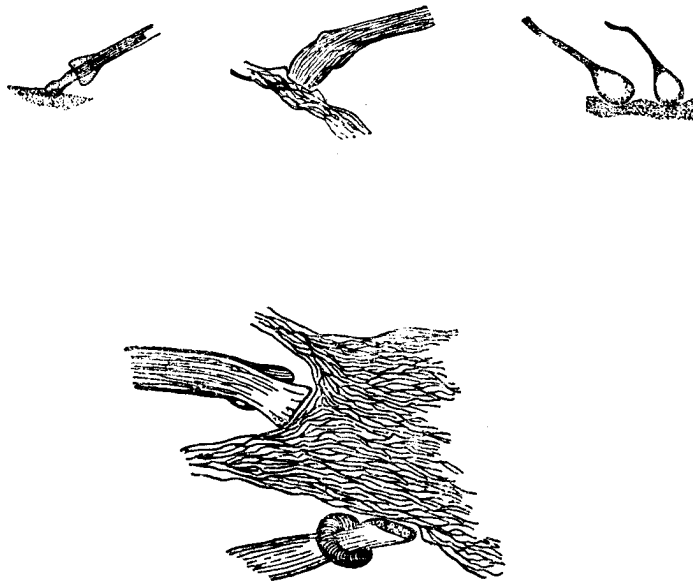


Рис. 9.3. Различные виды синапсов в нервной системе рыбы.

Клетки нервной ткани и их волокна образуют некоторые структуры, различимые как по их обособленному расположению, так и по совместному функционированию. Эти структурные образования чрезвычайно многообразны и весьма различны у различных животных. Чем сложнее организм, тем более сложную структуру представляет собой его нервная система и тем разнообразнее реакции организма на внешние воздействия.

У гидроидных полипов, например, нервная система состоит из двух типов клеток и представляет собой единую рассеянную сеть без отдельных, выделяющихся структур. У кольчатых червей нервные клетки образуют скопления

в каждом сегменте тела, называемые ганглиями, связанными в единую цепочку. От каждого ганглия отходят волокна к мышцам своего сегмента тела. В головной части имеется более крупное скопление клеток — надглоточный ганглий, являющийся органом, координирующим в какой-то мере работу всей системы в целом.

Все позвоночные имеют примерно такую же общую схему строения нервной системы. Она распадается на сегменты, связанные в единое целое по продольной оси тела животного и имеющие на головном конце наиболее обширное скопление клеток — головной мозг. Начиная от костистых рыб и кончая человеком, у всех позвоночных нервная система имеет одни и те же составные части. Правда, эти части у разных животных выполняют не совсем одинаковые функции.

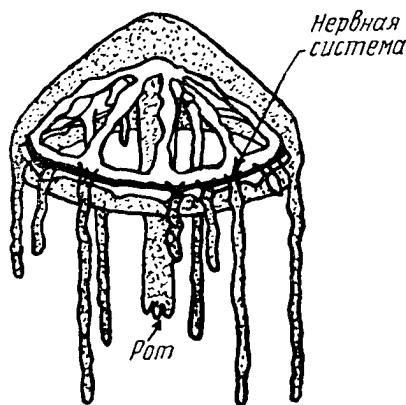


Рис. 9.4. Нервная система медузы.

Среди органов тела животного нервная система занимает особое место. Только нервная система в своей центральной части целиком защищена сплошным костным чехлом (череп, позвоночник). При голодании организма нервные ткани в последнюю очередь начинают терять в весе. Дело обстоит так, будто нервная система больше других органов оберегается от вредных влияний. Несомненно, что повреждение нервной системы наносит больший вред организму в целом, чем повреждение многих других органов.

Нервная система млекопитающих делится на два раздела: низший и высший. Иногда первый называют автономной или вегетативной системой, подчеркивая тот факт, что этот раздел нервной системы обеспечивает согласованную работу внутренних органов животного, работающих независимо от сознания. Однако эта независимость весьма относительна, так как низший раздел тесно связан и структурно и функционально с высшим разделом — центральной нервной системой. Автономная нервная система состоит из скоплений нервных клеток — узлов или ганглиев,

расположенных в различных частях тела, рядом с другими органами, и сети волокон, пронизывающих внутренние органы и проникающих, в частности, в спинной мозг. Центральная нервная система состоит из головного мозга, спинного мозга и сети волокон, идущих от рецепторов к мозгу и от него к эффекторам, так называемая периферическая нервная система.

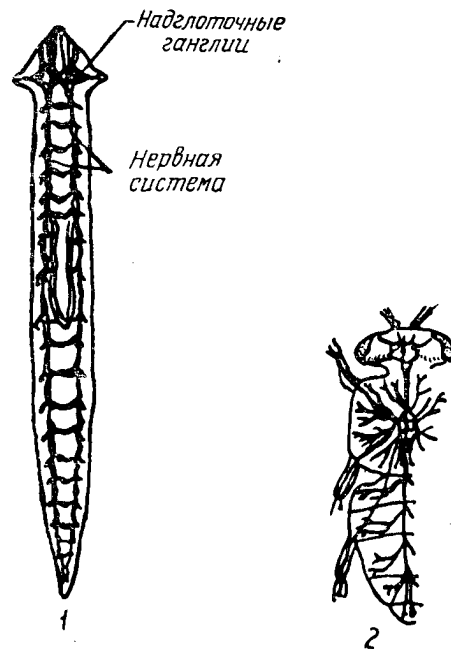


Рис. 9.5. Нервная система червя (1) и насекомого (2).

Спинной мозг представляет собой круглый тяж, лежащий внутри позвоночника. В разрезе ясно видно серое вещество, представляющее собой скопление нейронов и лежащее в средней части мозга в виде буквы Н. Внешняя часть состоит из белого вещества — нервных волокон, идущих отчасти в радиальном направлении, но главным образом в продольном, в виде так называемых столбов. Спинной мозг имеет периодическую структуру. На уровне каждого позвонка в него входят нервы, всего 31 пара. Нервы подходят симметрично справа и слева, причем сзади входят задние корешки нервов периферической чувствительности (в задние рога — отростки серого вещества), а спереди — передние корешки (отростки серого вещества).

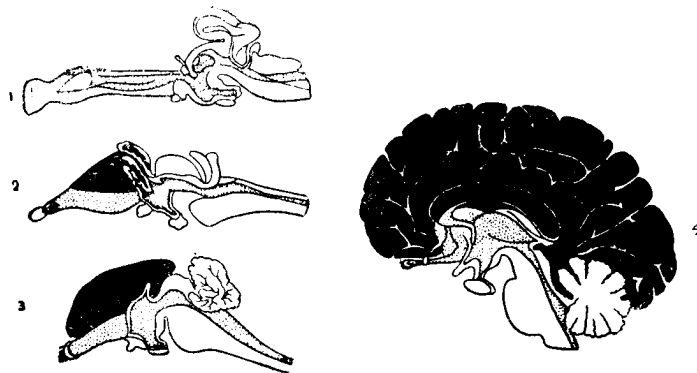


Рис. 9.6. Развитие мозга рыбы (1), ящерицы (2), кролика (3) и человека (4).

реди выходят передние корешки — двигательные (из передних рогов). Передние и задние корешки сливаются в общие узлы, от которых идут к периферии нервы, имеющие в своем составе и чувствительные и двигательные волокна.

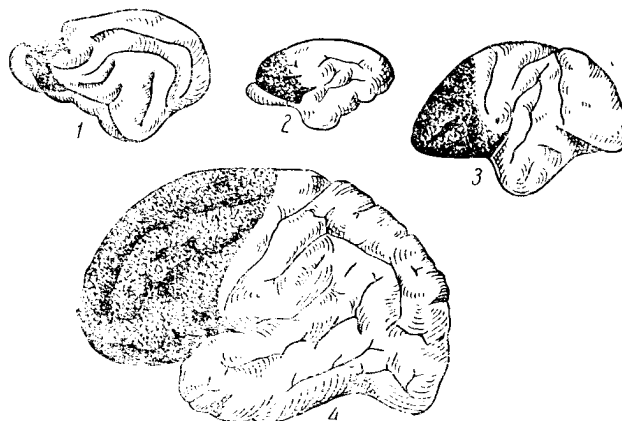


Рис. 9.7. Лобные доли головного мозга кошки (1), лемура (2), макаки (3) и человека (4).

В передней части спинной мозг переходит в продолговатый мозг, представляющий собой часть головного мозга.

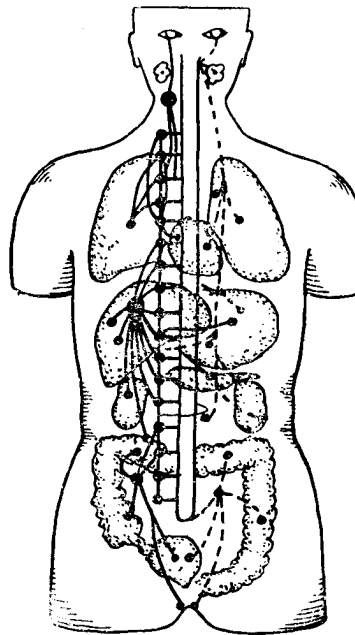
Головной мозг состоит из продолговатого мозга, являющегося непосредственным продолжением спинного, сред-

него мозга со зрительным бугром, промежуточного мозга и переднего мозга. Головной мозг в целом представляет очень сложное образование, изобилующее скоплениями клеток, вкрапленных в виде ядер в белое вещество, и пучков волокон, проходящих в толще белого вещества в различных направлениях. В нашу задачу не входит подробно описывать структуру мозга.

Наибольшим сходством строения и функций у различных видов животных обладают более древние, появившие-

Рис. 9.8. Схема автономной нервной системы.

С правой стороны тела показаны парасимпатические нервы, с левой — ортосимпатические. Ортосимпатические нервы образуют сплетения (ганглии) вблизи позвоночного столба, парасимпатические — возле тканей тех органов, на которые они воздействуют. Орто- и парасимпатические нервы обычно оказывают противоположное воздействие на внутренние органы: если одни стимулируют, то другие тормозят функцию (или наоборот). Окончания ортосимпатических волокон выделяют химическое вещество (нейрогумор) — симпатин, окончания парасимпатических волокон — ацетилхолин.



ся ранее в процессе эволюции части мозга, которые ближе примыкают к спинному мозгу: продолговатый, задний, средний и промежуточный. Все они сходны у различных видов животных, и функции их более или менее одинаковы.

Эволюционно наиболее молодая часть мозга — передний или конечный мозг — и в особенности часть его, называемая плащом, по размерам и строению резко различается у различных животных. Эта часть мозга образует так называемые большие полушария. У человека полушария занимают большую часть черепа и содержат около

половины всех нервных клеток (из общего числа около 10^{10}). Полушария имеют характерное строение, отличное от нижележащих частей. Тела нервных клеток полушарий сосредоточены в тонком поверхностном слое мозгового вещества серого цвета, называемом корой (то же имеет место в мозжечке). Поверхность коры больших полушарий человека достигает $\frac{1}{4}$ квадратного метра вследствие того, что она расположена в виде складок, образующих извилины (выступы) и борозды (впадины). Под корой находится белое вещество, состоящее из нервных волокон. Размеры и строение больших полушарий, в особенности лобных долей, лежат в основе многообразия и сложности высших нервных реакций животного.

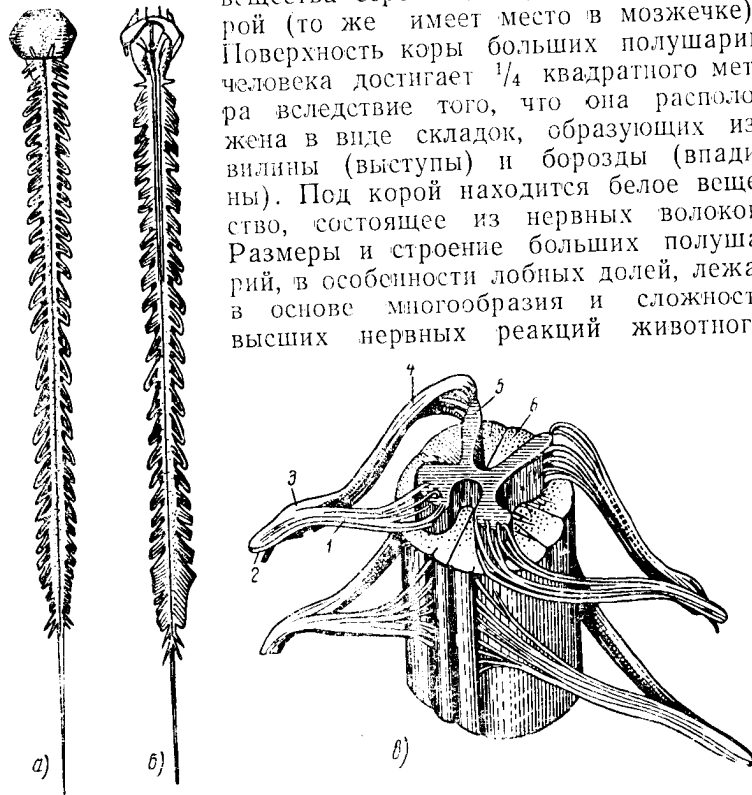


Рис. 9.9. Спинальный мозг человека.

a) вид спереди; *б)* вид сзади; *в)* разрез (1 — передний, двигательный корешок, 2 — спинномозговой нерв, 3 — межпозвоночный узел, 4 — задний, чувствительный корешок, 5 — серое вещество, 6 — белое вещество).

Поверхность лобных долей составляет у кошки 3% общей поверхности мозга, у собаки 7%, у шимпанзе 16%, у человека 29%.

Нервные волокна, из которых состоит белое вещество, соединяют между собой различные участки коры, а также кору с подкорковыми ядрами. Особенно обильны связи с другими участками у лобных долей. Обычно различают следующие типы связей: ассоциативные волокна, соеди-

няющие различные области одного и того же полушария; комиссуральные волокна, связывающие между собой участки двух различных полушарий; проекционные волокна, связывающие участки коры с подкорковыми ядрами и нижележащими отделами нервной системы.

Кора полушарий не представляет собой однородного образования. Микроскопические срезы коры обнаруживают характерное слоистое строение. В различных слоях

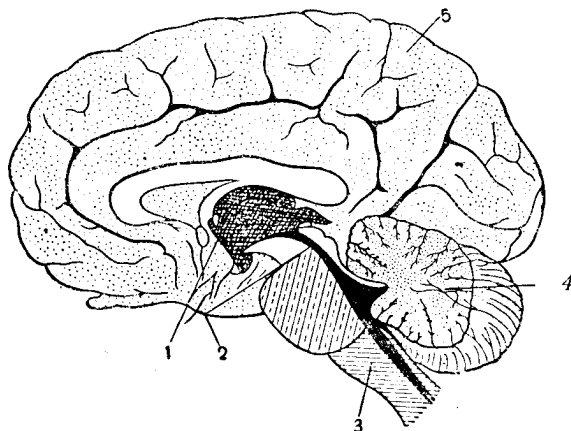


Рис. 9.10. Отделы головного мозга человека.
1 — промежуточный мозг, 2 — средний мозг, 3 — продолговатый мозг,
4 — мозжечок, 5 — передний мозг.

обнаруживаются различные типы клеток и волокон. Особенно интересно то, что в различных участках коры микроскопическое строение оказывается различным. Из восьми характерных слоев некоторые могут отсутствовать, быть слабо или сильно развитыми. В соответствии с особенностями клеточного строения всю кору делят на 11 цитоархитектонических областей и 52 поля. По мере изучения коры число известных полей постепенно увеличивается, так как внутри, казалось бы, однородных структур при детальном исследовании обнаруживаются новые.

Цитоархитектоническая и миелоархитектоническая структуры мозга, т. е. структуры клеточных скоплений и расположения пучков волокон, изучены довольно подробно. Изучение нервных структур производится параллельно и в связи с изучением функций. И хотя изучение мозга еще чрезвычайно далеко от завершения, уже в настоя-

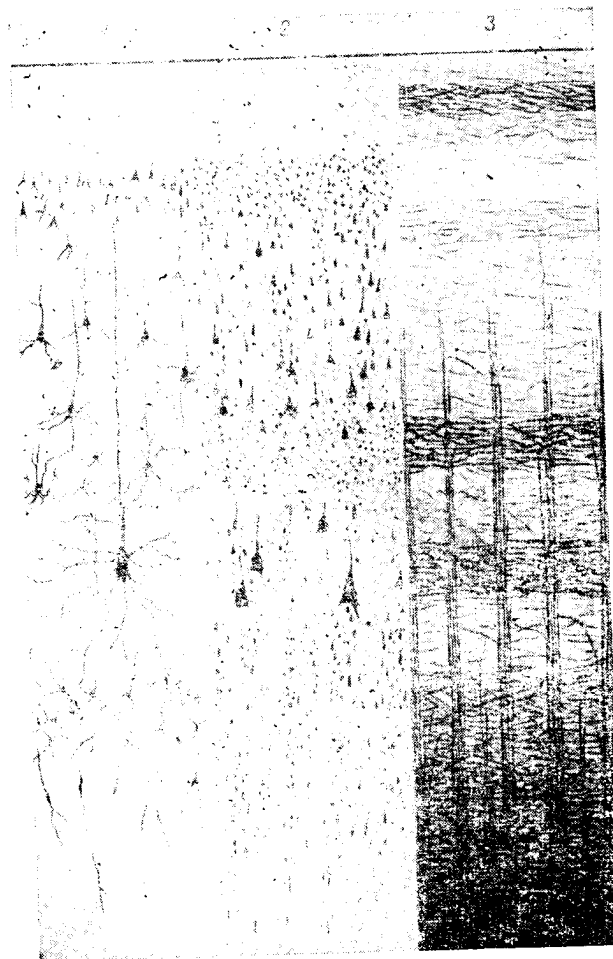


Рис. 9.11. Послойное строение коры головного мозга человека.
1 — нервные клетки, окрашенные по методу Гольджи, 2 — нервные клетки, окрашенные по Нисслю, 3 — миелиновые волокна по Кульчицкому.

щее время известны интересные сведения, позволяющие утверждать, что между работой мозга и его структурой имеется определенная связь. Физиологи называют сопоставление работы мозга и его структуры «локализацией функций». Вопрос о «локализации функций», т. е. зависимость функции от работы определенных участков мозга имеет длинную и богатую историю. Для нас он интересен прежде всего тем, что подход к работе мозга с точки зрения связи структуры и функции, подход более чем знакомый конструкторам вычислительных устройств, позво-

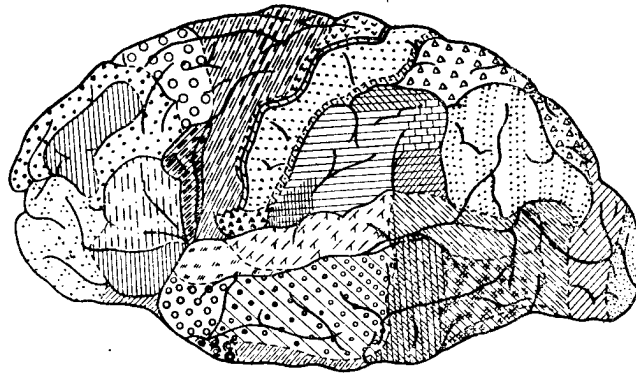


Рис. 9.12. Цитоархитектонические поля коры больших полушарий головного мозга человека (участки коры, отличающиеся друг от друга клеточным строением).

ляет четко сопоставить особенности живой и неживой машин, провести между ними параллели, а главное, определить, где кончается сходство, где начинается различие между ними и в чем же состоит это различие. В частности, весьма интересен вопрос о том, существуют ли такие функции живого организма, для которых не имеется соответствующих физико-химических, физиологических процессов, протекающих в пределах определенных органических структур. Весь опыт изучения деятельности нервной системы говорит однозначно о том, что ответ на последний вопрос отрицателен.

Работа мозга

Прежде чем ставить вопрос о характере работы мозга, следует посмотреть, какими методами изучаются структура и функции мозга. Таких методов несколько. Прежде

всего это микроскопия. Ткань мертвого мозга разрезается на очень тонкие слои (срезы), которые подвергаются окраске различными красителями, избирательно окрашивающими либо клеточные тела, либо волокна, либо промежуточную ткань и т. д. Цитоархитектоника мозга установлена именно таким способом. Однако на срезе можно проследить нервное волокно лишь на очень коротком участке; установить же топографию взаимосвязей в белом веществе мозга не представляется возможным, потому что пучки волокон тянутся подчас на несколько сантиметров. Для этой цели пользуются методом перерезки волокон (на живых животных).

Небольшой участок белого вещества рассекается на операционном столе и затем оперированное животное сохраняют живым в течение нескольких дней. Нервные волокна, отделенные ножом хирурга от нейронов, перерождаются и гибнут (подвергаются некрозу). Если после этого рассматривать под микроскопом окружающие ткани при соответствующей окраске, то перерезанные, переродившиеся волокна оказываются ясно видимыми. Длинной серией таких опытов можно детально проследить нервные связи в белом веществе.

Трудоемкость и трудность этой работы очевидны. Если бы сложное техническое устройство, например телевизионный центр, уменьшить до размеров спичечной коробки, залить в парафин и затем изучать под микроскопом на тонких срезах, вряд ли бы мы могли рассчитывать на скорый успех в изучении схемы и работы неизвестного нам устройства. Мы наблюдали бы в различных срезах то или иное количество деталей, провода, идущие в различных направлениях, но от этих сведений до схемы в том понимании, которое известно радиотехникам, было бы еще очень далеко. Примерно так же обстоит дело и с изучением мозга, сложность строения которого несравненно выше, чем сложность любой существующей сегодня радиосхемы.

К счастью, в распоряжении физиологов имеются и другие методы изучения мозга, еще в живом нормальном функционирующем состоянии.

Если на операционном столе под наркозом вскрыть череп животного и снять мозговые оболочки, то можно непосредственно воздействовать на поверхность мозга, раздражая ее механическим, химическим или электрическим способом. Можно получить доступ и к внутренним

частям мозга, вводя тонкие изолированные электроды в толщу мозга и раздражая ткань мозга электрически. Мозговая ткань не обладает чувствительностью ни тактильной, ни болевой, так как в ткани мозга не имеется соответствующих рецепторов. Подопытное животное или

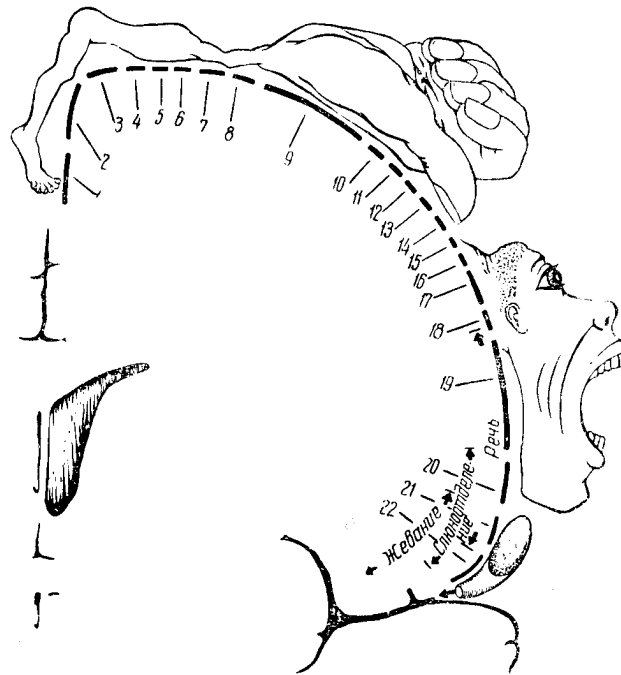


Рис. 9.13. Диаграмма Пенфильда — представление двигательных функций тела в передней центральной извилине коры головного мозга человека. (Функции символически для наглядности изображены в виде частей тела человека.)

1 — пальцы ноги, 2 — лодыжка, 3 — колено, 4 — бедро, 5 — туловище, 6 — плечо, 7 — локоть, 8 — запястье, 9 — кисть, 10 — мизинец, 11 — безымянный палец, 12 — средний палец, 13 — указательный палец, 14 — большой палец, 15 — шея, 16 — бровь, 17 — веко и глазное яблоко, 18 — лицо, 19 — губы, 20 — челюсть, 21 — язык, 22 — глотка.

человек, подвергающиеся такому воздействию, могут быть в полном сознании, без воздействия наркоза и не испытывать страданий при умеренном раздражении мозговых тканей. Таким образом, производя раздражения и наблюдая реакцию или узнавая субъективные ощущения пациента, можно многое узнать о работе нервной системы.

Что же известно из результатов подобного рода исследований?

Если раздражать электрическим током мозговую извилину, лежащую впереди от так называемой роландовой борозды, то всегда при раздражении одной и той же точки будет происходить одно и то же произвольное автоматическое сокращение определенной группы мышц на противоположной стороне тела. Переходя от одной точки поверхности мозга к другой, мы будем, таким образом, поочередно «включать» мышцы ног, спины, руки, лица и т. д. Можно построить отображение мышц тела на поверхность извилины, сохраняя наблюдающееся соответствие. Тогда мы получим проекцию мышц скелетной мускулатуры на переднюю извилину — «моторную зону» коры. От нервных клеток моторной зоны (так называемых «больших пирамид» или клеток Бетца) отходят нервные волокна, образующие большой пучок (пирамидный путь). Волокна пирамидного пути проходят через различные отделы мозга, частично перекрещиваются на границе продолговатого и спинного мозга (переходят с правой стороны тела на левую и наоборот) и входят в число проводящих путей спинного мозга. В пределах серого вещества спинного мозга волокна пирамидального пути образуют соединения (синапсы) с клетками серого вещества, соединяясь через них с эффекторными окончаниями на скелетных мышцах.

Если раздражать током заднюю извилину, лежащую позади роландовой борозды, пациент будет ощущать прикосновение (фактически отсутствующее) в различных участках кожи. И снова каждой точке поверхности мозга однозначно соответствует участок кожи, от которого можно проследить проекционные чувствительные связи в коре. В задней извилине мы имеем чувствительную или «сензорную» зону кожной чувствительности.

Подобным образом можно хотя бы отчасти «распутать» схему нервных взаимодействий в теле животного. Искусственно раздражая участок мозга, мы получаем не только видимую реакцию, но и субъективное ощущение. При этом мы как бы обходим нормальные пути прохождения сигнала нервного возбуждения и подаем его где-то посредине пути. Сигнальный характер процессов в мозге виден при этом весьма рельефно.

Опыты с раздражением коры позволяют обнаруживать устойчивые структуры, несущие некоторые определенные функции. Подобно тому, как в сегменте спинного мозга можно проследить дугу простейшего безусловного рефлекса от кожного рецептора к задним корешкам и рогам

спинного мозга, через серое вещество к передним рогам и корешкам и далее до мышечных эффекторов, можно установить также и более сложные и разветвленные связи.

Существуют связи между различными сегментами спинного мозга; благодаря этим связям в реакцию включаются большие группы различных мышц. Существуют побочные

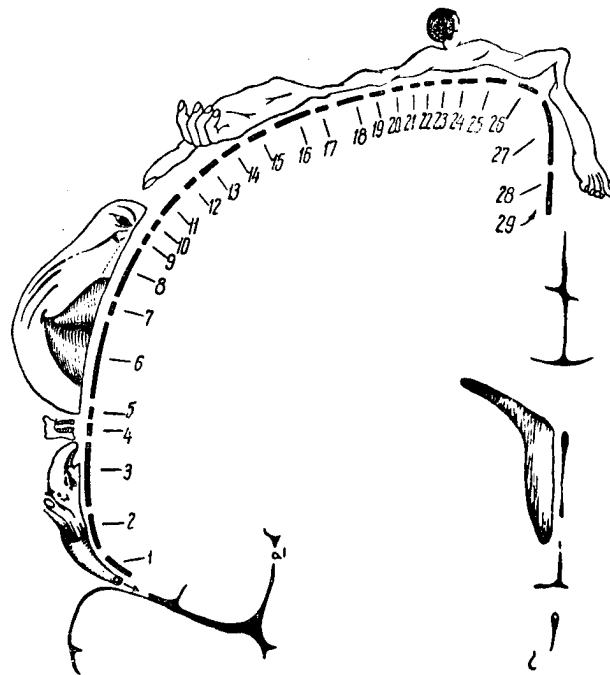


Рис. 9.14. Диаграмма Пенфильда — представление чувствительных функций тела в задней центральной извилине коры головного мозга человека.

1 — внутренние органы, 2 — глотка, 3 — язык, 4 и 5 — челюсти, 6 — нижняя губа, 7 — верхняя губа, 8 — щека, 9 — нос, 10 — глаз, 11 — большой палец, 12 — указательный палец, 13 — средний палец, 14 — безымянный палец, 15 — мизинец, 16 — кисть, 17 — запястье, 18 — предплечье, 19 — локоть, 20 — плечо, 21 — лопатка, 22 — голова, 23 — шея, 24 — тазовице, 25 — бедро, 26 — голень, 27 — ступня, 28 — пальцы, 29 — половые органы.

связи с ганглиями симпатической нервной системы, расположенными вне позвоночного столба, вдоль него по сторонам. Эти связи оказывают воздействие на регулировки внутренних органов, включая их в работу при интенсивных реакциях. Существуют, наконец, многочисленные связи с головным мозгом по восходящим и нисходящим

нервным путям спинного мозга. Эти связи, во-первых, дают возможность осуществлять, кроме чисто автоматических рефлекторных движений, движения сознательные, субъективно определяемые волей и объективно исходящие из коры головного мозга и идущие по пирамидальным путям через спинной мозг к эффекторам. Эти связи, во-вторых, регулируют протекание простейших рефлексов, задерживая их или включая. Именно поэтому акт ходьбы проходит часто незаметно для сознания, автоматически,

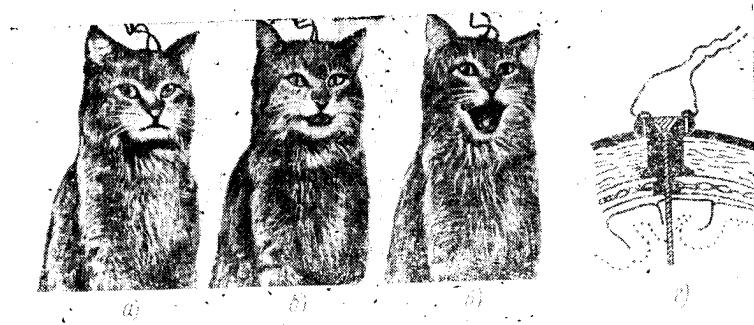


Рис. 9.15. Кошка, подвергающаяся непосредственному воздействию электрического тока на нервные клетки головного мозга через введенные в мозг электроды.

а) до включения тока, б и в) после включения тока, г) способ крепления электродов на черепе. Электроды введены в область подбугорья (по А. Б. Королу).

или рефлекторно. Сознательно ходьба может регулироваться лишь «в общем», а в деталях протекает автоматически.

Разбирая элементарный акт передачи сигнала по нервному волокну, мы уже говорили о том, что нервное возбуждение, представляющее собой физиологический процесс, имеющий электрохимическую природу, протекает в нервном волокне периодически, распространяясь по нерву со скоростью от 2 до 160 метров в секунду и повторяясь с большей или меньшей частотой в зависимости от силы возбуждения рецептора.

Прохождение возбуждения через нервную клетку подчиняется определенным законам и зависит от состояния клетки, которое, в свою очередь, определяется питанием, предыдущей историей клетки и падающими на нее раздражениями. Клетка под влиянием этих условий может

проводить или не проводить нервный импульс. Ряд приходящих со стороны раздражений может повысить ее порог возбудимости и «запереть» ее, как запирается для прохождения сигнала вентиль в схеме электронной цифровой машины. Недаром нейрон так любят последнее

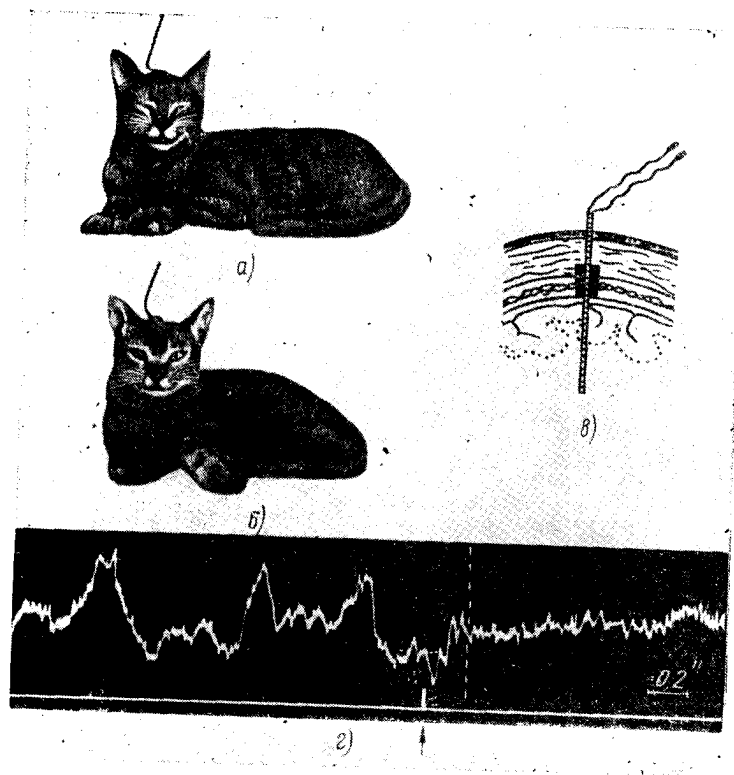


Рис. 9.16. Запись электрических потенциалов внутренних частей мозга кошки с помощью введенных в мозг электродов.
а) подошвная кошка в состоянии покоя, б) кошка, возбужденная зовом «кис—кис»,
в) способ крепления электродов, г) осциллографическая запись потенциала подбурья. Отметка соответствует моменту возбуждения зовом «кис—кис»
(по А. Б. Когану).

время сравнивать с радиолампой. К тому же результату могут привести химические (гуморальные) влияния на клетку через омывающую ее жидкость.

Пока организм живет, на вход нервной системы непрерывно поступают различные раздражения, которые распро-

страняются по нервным структурам, вызывают реакции, которые, в свою очередь, дают начало новым сигналам и т. д. В результате, нервная сеть постоянно находится в активном состоянии и проводит возбуждение. Это состояние активности не исчезает даже во время сна или коматозного состояния. Таким образом, каждое внешнее воздействие падает не на ожидающие его прихода безжизненные реле, а на непрерывно движущуюся и самостоятельно работающую систему.

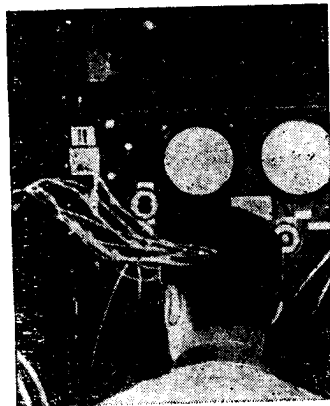


Рис. 9.17. Крепление электродов на коже головы человека при записи электрических потенциалов (энцефалограмм).

Внешние воздействия вносят изменения в уже существующее динамическое функционирование нервной системы. Часто эту динамику называют «спонтанной», т. е. самостоятельной деятельностью нервной системы. Термин этот неудачен, ибо эта деятельность является в конечном итоге отзвуком внешних воздействий, и, таким образом, она в некотором смысле рефлекторна.

Эта пестрая картина непрерывной работы мозга проявляется в разнообразии картин электрической активности. Электрохимические процессы проведения нервных импульсов вызывают появление изменяющихся электрических потенциалов в окружающих тканях и могут быть наблюдаемы и записаны даже с поверхности кожи головы при должном усилении. Такого рода записи — энцефалограммы — похожи на записи случайных процессов, однако в них можно выделить систематические составляющие, так называемые ритмы мозга. Систематические колебания потенциала отражают некоторые периодические процессы, протекающие в мозге. Можно классифицировать характерные колебательные процессы по их частоте. Удастся уловить колебания с частотами порядка 8—10 герц («альфа-ритм»), 18—32 герца («бета-ритм»), 35—55 герц («гамма-ритм»), 1—6 герц («дельта-ритм»).

Период и амплитуда систематических составляющих и характер непериодических процессов сильно меняются

В зависимости от того, с какой части черепа снимается потенциал и в каком состоянии находится подопытный. При засыпании пациента амплитуда медленных дельта-колебаний увеличивается, при открывании глаз — уменьшается или исчезает. Состояния нормального и гипнотического сна дают отличные друг от друга записи. Некоторые душевные болезни, например эпилепсия, вызывают

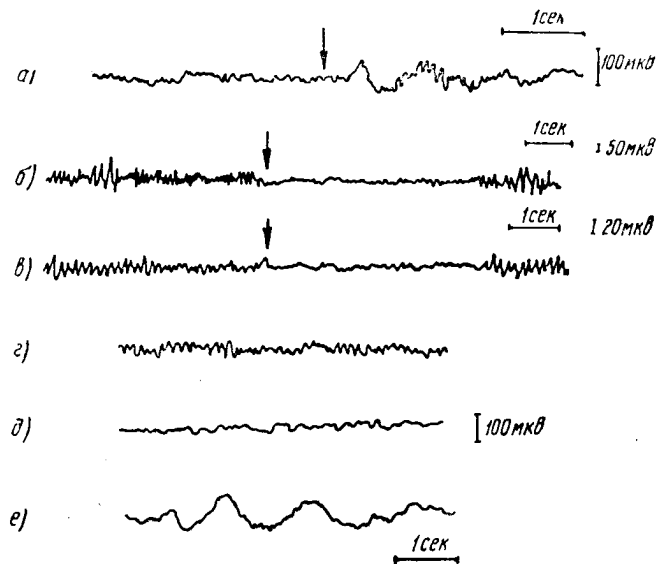


Рис. 9.18. Энцефалограммы — осциллографическая запись электрических потенциалов с кожи головы.
 а) человек; слуховое раздражение во сне (отмечено стрелкой), б) кошка, открывание глаз (отмечено стрелкой), в) человек, открывание глаз (отмечено стрелкой), г) человек в бодром состоянии, д) человек в состоянии легкой дремоты, е) человек в состоянии глубокого сна.

характерные изменения вида энцефалограммы: в ней появляются «высоковольтные» пики.

В последнее время появился новый метод наблюдения электрических потенциалов, снимаемых с кожи головы. На голове пациента в шахматном порядке укрепляют несколько одинаковых электродов. Потенциал каждого электрода подается на усилитель и далее через коммутатор на вход электронного осциллографа. Световое пятно, образуемое лучом осциллографа, перемещается по экрану, воспроизводя геометрическое расположение электродов, укрепленных на голове пациента. Яркость пятен на экране

изменяется при этом в соответствии с потенциалами, снимаемыми с соответствующих точек кожи. На экране осциллографа получается светящаяся картина распределения электрических потенциалов на коже головы, которая изменяется во времени в зависимости от состояния пациента и его психической деятельности.

Еще один метод исследования деятельности мозга дается клиническими наблюдениями над нервными больными, а также экспериментами над животными с удалением частей мозга. Животное и человек могут существовать при

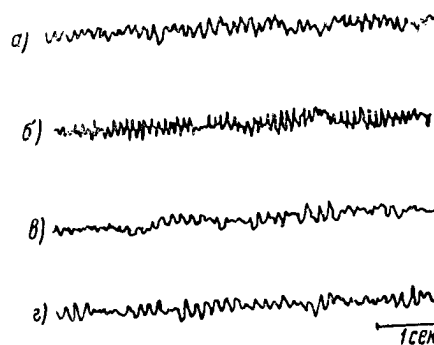


Рис. 9.19. Энцефалограммы различных животных.

а) морская свинка, б) кошка, в) обезьяна, г) человек.

значительных повреждениях отдельных частей мозга. При этом их поведение заметно отклоняется от нормы. Изучение таких отклонений и позволяет выяснить роль отдельных частей мозга, поврежденных болезнью или операцией.

Если перерезать спинной мозг, полностью нарушается чувствительность и произвольные движения в той части тела, которая управляется (иннервируется) частью спин-

ного мозга, отделенной от головного. Однако в отделенной части остаются простейшие рефлекторные движения, так как нервные пути рефлекторных дуг не нарушены. Более того, простейшие рефлексы оказываются повышенными. То же самое наблюдается не только в том случае, если полностью перерезать спинной мозг, но и при заболеваниях, нарушающих в большей или меньшей степени проводимость пирамидного пути, ведущего, как мы уже знаем, от коры больших полушарий к эффекторам и передающего сигналы произвольных движений. Усиление рефлекторных движений и появление нормально невыявленных рефлексов служит в медицинской практике признаком (симптомом) заболевания пирамидного пути (сдавливание опухолью и т.п.). Одним из таких симптомов является симптом Бабинского. Если провести тупым

предметом по подошве ноги от пятки к пальцам, то в результате происходит произвольное тыльное разгибание большого пальца и расхождение веером остальных так, как это происходит при ходьбе. Это явление происходит только при нарушении управления со стороны коры. Рефлексы проявляются «самостоятельно», без команды вышележащих центров, и поведение становится автоматическим. Подобных явлений можно наблюдать много.

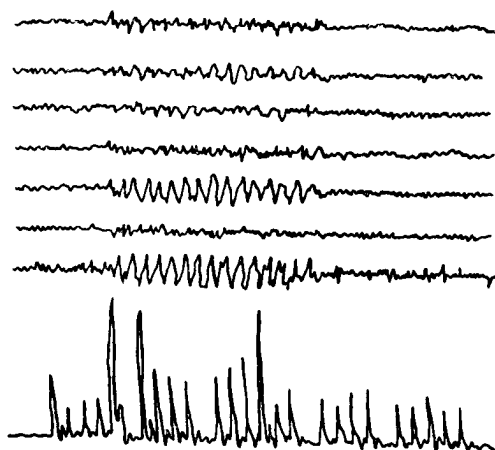


Рис. 9.20. Энцефалограмма с различных участков черепа больного эпилепсией, снятая во время припадка, длившегося три секунды.

Разрушение продолговатого мозга ведет к смерти животного. Исследования раздражением показывают, что в продолговатом мозге находятся участки, возбуждение которых дает сложные автоматические движения глотания, рвоты, слюноотделения, дыхания, сердцебиения, расширения и сужения сосудов и т. д.

Если разрезать мозг выше промежуточного, то появятся судорожные напряжения мышц (так называемая децеребрационная ригидность) за счет усиления рефлексов, обеспечивающих противодействие силе тяжести (антигравитационных рефлексов).

Удаление коры головного мозга превращает животное в своего рода автомат. Оно ходит, совершает координированные движения, но неспособно ничему научиться, не узнает даже знакомых предметов, неспособно добывать пищу и, будучи голодным, пробегает мимо кормушки с пи-

шей, которую тем не менее исправно ест, если пищу вкладывают ему в рот.

Удаление или поражение болезнью отдельных участков коры головного мозга дает очень пеструю картину нервных расстройств, которая в общем согласуется с теми сведениями о функциональных особенностях различных структур, которые получены из опытов с раздражением электрическим током.

Участки мозговой ткани, которые тесно связаны с данной функцией, например, с сокращением определенной группы мышц, обычно называют «центром» данной функции

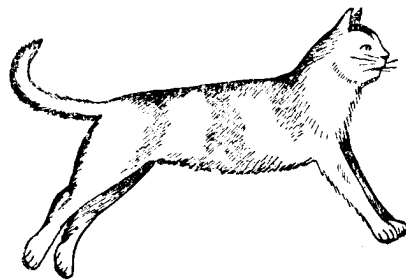


Рис. 9.21. Децеребрационная ригидность кошки.

Перерезание ствола мозга вызывает напряжение мышц-разгибателей за счет снятия торможения антигравитационных рефлексов со стороны головного мозга.

и говорят о «локализации данной функции» в этом центре. Постепенно было открыто много таких «центров». Особенно интересен тот факт, что границы многих функциональных «центров» совпадают с границами цитоархитектонических полей. Очевидно, что структурные особенности соответствуют функциональному назначению участков. Имеются рецепторные, или сенсорные поля, имеются

эффektorные, или моторные поля, но, кроме тех и других, существуют также участки коры, не связанные непосредственно с простыми ощущениями и движениями. Некоторые физиологи называют их «ассоциативными полями», отмечая роль этих полей в связывании ощущений и реакций. Многие из таких полей имеются только у определенных видов животных, например, у человека. Разрушение или повреждение нервных структур или функциональных систем коры влечет за собой расстройство интеллектуальных способностей. Симптомы таких заболеваний чрезвычайно разнообразны. Так, например, повреждение задней части третьей лобной извилины слева вызывает особое расстройство речи — так называемую «афазию Брока»; пациент не способен выражать свои мысли словами. Он понимает то, что ему говорят, но не в состоянии сформулировать свои мысли в виде фразы. Поражение некоторых частей нижней теменной доли сопровождается выпадением

дением из памяти отдельных слов (амнестическая афазия). Больной способен произносить эти слова, если их подскажут и он понимает их смысл.

Повреждение отдельных участков коры вызывает потерю способности писать (аграфия), читать (алексия), считать (акалькулия), распознавать ритм и мелодию музыки (амузия), невозможность совершать произвольные движения при отсутствии параличей (апраксия) и т. д. Описаны случаи, когда специалист лингвист с повреждением некоторого участка мозга оказывается в состоянии грамматически правильно выполнять перевод текста с одного языка на другой и при этом не понимает ни слова из переведенного им текста.

Все эти и многие другие подобные им факты, казалось, позволяют достоверно утверждать о существовании геометрически ограниченных «центров» речи, письма, чтения и т. д. Однако это утверждение может быть подвергнуто критике. Оказывается, одни и те же симптомы можно получить с различных участков коры. Иногда вместо изолированного движения получают координированные движения различных мышц, например, поворот глаз и головы в одну и ту же сторону и т. п. Можно было бы попытаться назвать новую точку «центром координированного движения глаз и головы», но это не дало бы ничего для понимания работы мозга, как, собственно, ничего не объясняет и введение понятия «центра».

Некоторые сложные рефлексы при повреждении коры вскоре восстанавливаются. Очевидно, что работу поврежденной части стала выполнять какая-то другая — неповрежденная. При этом, чем сложнее нарушенная функция и чем она «моложе» филогенетически, т. е. чем позже она появилась в процессе эволюции, тем чаще и быстрее она восстанавливается. Филогенетически «старые» функции, относящиеся к подкорковым структурам, восстанавливаются хуже или вовсе не восстанавливаются.

По-видимому, различные функции различным образом связаны с нервными структурами. Некоторые функции — типа простых рефлексов — связаны со структурой однозначно, другие — сложные, вырабатывающиеся по мере роста организма в виде приобретенных навыков, — могут протекать различными путями и использовать как одни, так и другие структуры. Следует говорить не о «локализации функций» в мозге, а о локализации нервных структур, одни из которых выполняют лишь небольшое разнообра-

зие функций, другие же, подобно счетной машине, выполняющей на одной и той же конструкции различные действия, способны реагировать очень сложно и разнообразно на разнообразные внешние воздействия и, в частности, выполнять некоторую функцию различными путями, используя различные каналы связи или обходные пути. К структурам последнего типа относится вся кора головного мозга. Удаление ее целиком лишает животное всех приобретенных навыков и возможности их приобретения, т. е. всех интеллектуальных функций.

Попытки «локализации функций», т. е. разыскания нервных структур, работающих при реализации тех или иных психических функций и их взаимодействий, очень напоминает задачу, которая стоит перед инженером, когда он пытается локализовать причину неправильного функционирования сложной системы автоматического управления. Правда, есть и разница: инженер всегда имеет в своем распоряжении полную схему устройства, в случае же анализа нервных структур такой схемы нет, а если бы она и была, то оказалась бы в огромное число раз сложнее схемы самых сложных технических устройств. Но в существенных чертах обе задачи близки друг другу. В системе управления одно и то же нарушение или изменение функции может быть вызвано различными изменениями в структуре, вносимыми в различных точках цепи прохождения сигнала. Бывает и так, что некоторое нарушение функции однозначно связано с неисправностью одного определенного узла системы. При анализе схемы машин в связи с ее функционированием становится ясным, что любая функция реализуется посредством сложной схемы, включающей в себя большое число узлов машины и их связей. И если неисправность в определенном узле всегда вызывает нарушение определенной функции, то это совсем не дает права утверждать, что данная функция «локализована» в данном «центре».

Одна и та же задача может решаться животным по-разному в зависимости от наличных ресурсов нервной системы. Об этом говорят, например, результаты опытов с обучением крыс в лабиринте (те самые, которым подражал Шэннон своей искусственной мышью). Понижение способности крыс ориентироваться в лабиринте зависит не столько от того, какие части мозга у них удалены или повреждены, сколько от того, какая часть общего количества мозгового вещества выключена. Это говорит прежде всего

о том, что зрение, слух, обоняние, осязание могут дополнять и заменять друг друга при ориентировке в лабиринте. Эти наблюдения оказали влияние на установившиеся в свое время взгляды относительно существования строго фиксированных «центров» любой функции. В настоящее время установлено, что, помимо «основных» центров, существуют разбросанные по всей коре участки, которые в некоторой степени могут заменить собой основной центр.

Отдельные индивидуумы могут отличаться друг от друга особенностями строения и развития отдельных частей мозга. Эти различия могут сказываться на способностях индивидуума, характере его восприятий, памяти, склонностях и т. д. Поскольку общее число клеток мозга более или менее постоянно, чрезмерное развитие какой-либо одной способности часто предполагает некоторое снижение других.

Эмоции

Эмоции связаны с работой определенных нервных структур мозга, т. е. клеточных образований, возбуждение которых вызывает реакцию вегетативной нервной системы и внутренних органов, дающую внешнюю картину проявления эмоции и одновременно субъективное ощущение перемены настроения. Некоторая часть нервных структур, ответственных за возникновение эмоций, расположена в отделе мозга, называемом «гипоталамус», по соседству с некоторыми центрами телесных гомеостатических регуляторов.

Эмоция и ее проявления имеют определенную биологическую целесообразность, определяя общий характер реакции организма на обстановку в целом. Так, появление чувства страха объективно сопровождается возбуждением автономной нервной системы, выделением активирующего вещества, симпатина, в эффекторах ортосимпатической системы, выделением адреналина надпочечными железами, усилением расщепления гликогена в глюкозу в печени и усиленным поступлением глюкозы в кровь, сокращением селезенки, выбрасывающей добавочную резервную кровь в кровеносное русло, усилением работы сердца (и задержкой работы пищеварительного тракта), сужением кровеносных сосудов, усилением дыхания, расширением зрачков, уменьшением охлаждения тела (ослабление притока крови к коже и поднятие волос для сохранения тепла), общим усилением обмена веществ и т. д. Все телесные проявле-

ния страха способствуют мобилизации организма на то, чтобы обеспечить максимальную отдачу усилий ради самосохранения. Работоспособность организма в состоянии страха (или гнева), как известно, заметно повышается. Телесные проявления эмоции охватывают весь организм.

Деятельность «центров» эмоций может быть подавлена воздействием коры головного мозга по соответствующим нервным связям. Заболевание таламической области или связанной с ней лобной доли коры мозга ведет к стойким нарушениям настроения. Воздействие алкоголя на кору ведет к ослаблению воздействия коры на таламус и к его «расторжению», чем и объясняются неразумные и разнуданные поступки людей, подвергающих себя такого рода «алкогольной децеребрации».

Связь эмоциональных «центров» с корой явствует также и из того, какими путями возбуждаются эмоции. Кроме прямых воздействий (испуг от слишком громкого, неожиданного шума и т. п.), тот же страх может возникнуть в результате восприятия смысла какого-либо устрашающего рассказа, новости и т. п.

Существует особый вид общения людей, призванный апеллировать непосредственно к эмоциям. Это — музыка, поэзия, живопись, театр и другие виды искусства. Некоторые произведения искусства (инструментальная, симфоническая музыка) очень мало связаны с разумным восприятием (если оставить в стороне профессиональную, ремесленную сторону музыки). Другие же (жанровая живопись, эпическая поэзия, проза) пользуются средствами рассудка для достижения той же цели — возбуждения эмоций; чувство радости или гнева возникает у читателя лишь тогда, когда он подробно узнает историю героев или рассматривает во всех подробностях полотно жанровой живописи.

Некоторые химические вещества избирательно воздействуют на отдельные области мозга, вызывая возбуждение или торможение их деятельности и давая характерную картину изменения поведения человека и его реакций на изменения внешней обстановки. Так, например, одно из таких веществ («Обливион») вызывает подавление эмоций страха, не вызывая при этом снижения способности правильно оценивать обстановку и принимать решения. Другое вещество (мескалин) вызывает временное состояние, весьма напоминающее душевное заболевание — шизофрению.

Сигнальные процессы в мозге

Если подытожить те не слишком обширные и далеко не исчерпывающие сведения, которые имеются у нас о работе нервной системы в целом и мозга в частности, то в результате складывается убеждение, что нервная система, рассматриваемая как физический объект, представляет собой типичную кибернетическую систему, работающую с сигналами. Непосредственное и детальное сравнение мозга и счетной машины вряд ли имеет смысл, и название «электронный мозг», которое часто в последнее время применяют к машине, имеет, разумеется, только иносказательный, пожалуй даже иронический, смысл. Но в самом общем плане такое сравнение вполне законно, ибо и мозг, и машина работают с информацией, используют сигналы, выполняют преобразования сигналов и управляют исполнительными органами.

При таком сравнении нельзя игнорировать тот факт, что самая сложная машина имеет всего несколько тысяч элементов, что в миллионы раз меньше числа нервных клеток мозга. Правда, машина совершает элементарные такты работы в сотни тысяч и миллионы раз быстрее мозга. Кроме того, деятельность машины и деятельность мозга специализированы на решение различных классов задач с различными алгоритмами и в различных алфавитах.

Все, что мы знаем о мозге, не противоречит утверждению, что известные нам мозговые процессы суть процессы сигнальные. И у нас нет никаких оснований утверждать, что сигнальные процессы не исчерпывают всей деятельности мозга.

Внешние события воздействуют на рецепторы, и возбуждение передается в соответствующие отделы мозга, так называемые анализаторы, где происходит предварительная обработка информации. О механизмах этой обработки, об алгоритме ее и о коде нам известно очень мало главным образом потому, что наблюдать работу соответствующих структур в живом мозге очень трудно. Если раздражать токком различные участки, например, зрительного анализатора, то при раздражении частей, расположенных близко к рецептору (сетчатке глаза), пациент сообщает об ощущении зрительного образа, ясно различимого, как бы реального и зафиксированного в пространстве; при раздражении более высокорасположенных разделов пациент

сообщает о том, что видимый им фиктивный образ имеет характер не реальности, а скорее яркого представления, не связанного с пространством и окружающей обстановкой.

По мере продвижения от рецептора к коре число нервных волокон не увеличивается, а уменьшается. Очевидно, объем доводимой до сознания информации меньше, чем объем информации, получаемой на входе. Это сжатие информации, отсеив ненужной и сохранение лишь необходимой представляет чрезвычайно большой интерес с принципиальной стороны.

Представим себе, что не слишком крепко спящий человек испытывает внешнее воздействие: он слышит сильный звук или ощущает прикосновение. Обычно в этой ситуации воздействие не осознается спящим непосредственно, а вызывает сновидение, которое он по пробуждении хорошо помнит и может воспроизвести рассказом во всех деталях. Можно зафиксировать объективно время от начала воздействия до пробуждения, т. е. время, пока длилось сновидение. Обычно оно невелико. Воспользовавшись аналогией с телевизионной передачей изображения, можно оценить, сколько кадров и сколько элементов в кадре потребовалось бы, чтобы воспроизвести сновидение с теми деталями, которые может рассказать проснувшийся, а затем подсчитать, исходя из времени сна, какую полосу частот должна была занимать подобная «телепередача». Подобные приблизительные подсчеты дают астрономическую цифру порядка 10^{20} — 10^{23} герц, т. е. частоты жесткого гамма-излучения. Поскольку физиологические процессы в мозге имеют верхнюю границу полосы порядка не выше 100 герц, а число возможных параллельно работающих каналов не превышает 10^9 — 10^{10} , остается заключить, что способ кодирования информации в мозге в огромное число раз экономнее, чем в современном телевидении. Не приходится говорить, какое значение для технических отраслей кибернетики имела бы разгадка этой тайны физиологии.

Аналогичные примеры можно было бы привести и относительно «синтеза» реакций на основе обработки информации.

К сожалению, «синтез и анализ» раздражений — понятие, введенное в физиологию еще Павловым, — с тех пор не получило никакого развития и конкретизации, и мы сегодня так же мало знаем о деталях обработки входной осведомительной информации и переработки ее в выход-

ную исполнительную, как и во времена Павлова. «Синтез и анализ» применяется только вообще.

Мы уже обсуждали вкратце проблему различения формы, когда говорили о машинах для чтения. Эта проблема решена природой в мозговых структурах, но мы о ней не знаем почти ничего и должны отнести ее к очередным задачам кибернетики.

Еще более интересен и актуален вопрос о механизме памяти мозга. «Анализ» раздражений и «синтез» реакций тесно связаны с памятью. Объем памяти мозга чрезвычайно велик. Механизм, ответственный за сохранение сигналов нервных раздражений, до сих пор не установлен и не локализован. По-видимому, каждый анализатор имеет свою собственную память (может быть даже несколько видов ее), тесно переплетающуюся с механизмом обработки информации. Существует память зрительная, слуховая, моторная и т. д. При этом «вызов» информации из памяти происходит не по ее «адресу», как в счетной машине, а по ассоциации данного сигнала с другими, поступившими одновременно с данным, на основе ранее выработанных условных рефлексов. Однако детали этого процесса, его схема, порядок протекания его в нервных структурах мозга нам еще неизвестны.

Каким образом, различные раздражения, поступившие одновременно в различные отделы мозга, фиксируются памятью совместно и впоследствии появление одного из них вызывает представление о другом и притом достаточно сильное для того, чтобы вызвать реакцию? Ответ на этот вопрос затруднителен потому, что количество сочетаний различных элементарных раздражений необозримо велико. И на любое сочетание с безусловным раздражением может быть выработан условный рефлекс. Кроме того, условный рефлекс, т. е. совместная фиксация в активной памяти комбинации раздражителей, непрерывно изменяется. С течением времени он укрепляется, ослабевает, сочетается с новыми раздражителями и рождает новые рефлексы, тормозится и т. д. Время реакции уменьшается при обучении, т. е. выработке рефлекса, и увеличивается при его торможении. По-видимому, все это сопровождается динамической перестройкой функционирования нервных структур, изменением путей прохождения сигналов в мозге. При огромном количестве активных элементов (клеток в мозге) картина таких динамических изменений должна

... чрезвычайно сложной. Энцефалограмма дает о ней лишь очень слабое представление.

Существуют гипотезы относительно элементарных процессов, которые обеспечивают запоминание нервных сигналов. Согласно одной из них память осуществляется циркулирующей нервной возбуждения по замкнутому пути, состоящему из нервных волокон и клеток, подобно тому, как «слова», закодированные в бинарном алфавите, циркулируют в устройстве памяти с линией задержки вычислительной машины (см. гл. 7). Некоторыми исследователями были обнаружены замкнутые петлеобразные нейронные структуры в нервных тканях рецепторов, однако решительного подтверждения этой гипотезы памяти у нас нет. Другая гипотеза приписывает память свойствам протеиновых молекул, находящихся внутри клеток. Изомерные перестройки этих весьма сложных агрегатов (т. е. перестройки, не изменяющие числа атомов, а лишь порядок расположения атомов в молекуле) дают набор дискретных состояний, отличающихся, вообще говоря, химическими свойствами и могущих проявиться в физиологических функциях клетки. Эта гипотеза, также впрочем еще не проверенная основательно, весьма заманчива, так как она объясняет многие факты, необъяснимые гипотезой циклических нейронных структур, и, в частности, наличие памяти у простейших одноклеточных организмов, которое отмечалось многими естествоиспытателями*.

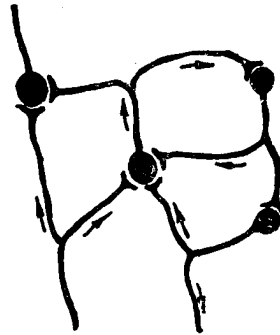
Если гипотеза о внутриклеточной молекулярной памяти верна, то необходимо допустить, что молекулярные квантово-механические процессы в клетке оказывают влияние на функционирование целых органических структур и всего макроскопического организма. Происходит обмен информацией между квантово-механическими, молекуляр-

* Парамеция, посаженная в узкую, закрытую с одного конца капиллярную трубку, доплыв до тупика, совершает 30—50 неудачных попыток повернуться, затем находит способ это сделать. Во второй раз в той же ситуации парамеция делает только 10—15 неудачных попыток перед правильным движением. При последующих повторениях опыта она все быстрее находит выход из затруднения, явно приобретая опыт, т. е. запоминая предыдущие действия и их результаты.

Инфузория-трубач (стенатор) при раздражении ее струей воды вначале пытается уклониться от раздражения, совершая движения в стороны. Затем прячется в слизистую оболочку, вылезая из нее по временам, и прячется снова, если раздражение продолжается. Наконец, она покидает свое убежище и уплывает прочь, ища более благоприятной среды. Поведение изменяется по мере накопления опыта, что не может происходить без запоминания предыдущих ситуаций.

ными структурами и «классическими» макроскопическими структурами, а также взаимное управление (запоминание и вспоминание). Следует заметить, что эта ситуация — управление макроскопической системой со стороны квантово-механических процессов — не является единственной. То же имеет место при взаимодействии молекулярных структур хромосомы и плазмы клетки. Процесс обмена веществ в плазме, расщепление пищевых веществ и синтез клеточного вещества, синтез энзимов — органических катализаторов, определяющих направление и характер химических реакций, — определяются в конечном итоге актив-

Рис. 9.22. Замкнутые нервные цепи в ретикулярном веществе глазодвигательного нерва (предполагаемая схема по Лоренте де Но).



ным воздействием сложнейших молекул, которые входят в состав хромосом.

Таким образом, жизнедеятельность клетки складывается из направляющего воздействия молекулярных структур хромосомы и воздействий внешней среды (поступление пищи того или другого состава и создание условий хода реакций, температуры, давления и т. п.).

Обширная и гибкая память по ассоциациям есть то свойство мозга, которое создает громадное качественное отличие поведения животного от поведения механических автоматов.

Выработка условных рефлексов создает тот тип рационального поведения, которым характеризуются животные. Условный рефлекс всегда создается только на основе биологически целесообразной безусловной реакции, например реакции на пищу. Вид и запах пищи не являются безусловными раздражителями. Щенки, воспитанные на вегетарианском питании, не реагируют на вид и запах мяса. Лишь после того, как они узнают вкус мяса, вид и запах его будут прочно ассоциироваться с поглощением пищи и вызы-

вать обильное выделение слюны; вырабатывается «натуральный» условный рефлекс, который будет поддерживаться и существовать в течение всей жизни собаки. Впоследствии, после первой поимки дичи, подобные же натуральные рефлексы вырабатываются на вид дичи, ее поведение (убегание) и т. д.

Затем будут вырабатываться условные рефлексы на условные рефлексы (продолженные рефлексы), и если сочетания условных раздражителей во внешней обстановке устойчивы, эти рефлексы также будут непрерывно поддерживаться, создавая в своей совокупности общий облик поведения животного. Условные рефлексы вырабатываются не только в виде реакции активного действия, но и в виде торможения действия. Рефлексы могут вырабатываться с задержкой реакции по времени (до получаса), т. е. реакция будет происходить не на раздражение, а на след раздражения, зафиксированный в памяти. Рефлекторной реакцией может быть действие внутренних органов, не воспринимаемое сознанием, в частности изменение эмоций. Следует заметить, что «сознательная» часть работы мозга составляет лишь малую долю общего объема его работы.

Все эти хорошо изученные факты позволяют строить гипотезы о внутренних процессах, определяющих поведение животного или человека.

Информация о внешних событиях в обработанном виде, с отсевом несущественной части, зафиксированная в памяти в виде сигналов — физиологических состояний и процессов нервной системы, либо может превратиться в действие, реакцию немедленно или с задержкой, либо реакция может быть заторможена. Процесс сопоставления сигналов различных событий, вызов из памяти ассоциированных сигналов других событий (например, причин и следствий), выбор на основании эмоциональных оценок наиболее целесообразной реакции и представляет собой физиологическую основу того процесса, который в субъективной терминологии называется «размышлением». Мысль представляется, таким образом, объективно как сигнальный процесс обработки информации, запасенной в памяти за все время жизни человека, на основе свойственных индивидууму природных инстинктов, эмоциональных оценок, темперамента. Мысль, по Сеченову и Бехтереву, есть задержанный исполнением условный рефлекс «... очевидно, что мысль и вообще субъективные переживания должны быть понимаемы как задержанные рефлексы, которые рано или поздно, освободившись от торможения, перейдут в объективный мир в форме

ли пересказа или в форме действий и других реакций». (В. Бехтерев. «Общие основы рефлексологии», гл. II). При эмоциональном возбуждении торможение ослабевает и реакция прорывается наружу. Интенсивно и возбужденно мыслящий человек шепчет, жестикулирует, гримасничает.

Объективная психология, или рефлексология, имеющая дело не с анализом субъективных внутренних ощущений и переживаний, а с наблюдаемыми реакциями, рефlekсами и подобными им явлениями, может быть в значительном объеме, а скорее всего целиком, построена на сигнале как основном понятии. Субъективные ощущения представляют собой информацию, несомую сигналом — физиологическими процессами нервной деятельности. В этом факте отражается общее свойство сигнала—соответствие, изоморфизм сигнала и внешних событий или реакций.

«Объяснить» субъективное сознание индивидуума, сознание как ощущение личности, собственного «я», как ощущение различия между внутренними и внешними явлениями в жизни индивидуума вряд ли возможно, ибо для каждого индивидуума оно единственно и ни с чем не сравнимо. Сознание же других людей, наблюдаемое не как явление внутри моего «я», а в окружающем меня коллективе людей (индивидуальное сознание каждого члена этого коллектива), может быть понято по сопоставлению с богатым миром кибернетических систем и процессов, существующих помимо живых существ. Правда, сегодня ни одна самая сложная машина не подошла еще к той грани, за которой начинается сознание, и мы не знаем, перейдет ли она когда-нибудь эту грань. Но ничто, кроме предубеждения и предрассудков, не позволяет сегодня уверенно отрицать подобную возможность. Здесь уместно вспомнить слова академика И. П. Павлова о человеке как машине, единственной по высочайшему саморегулированию, которые мы цитировали в первой главе.

В последующих главах мы еще вернемся к этому вопросу.

Чем не обладает машина?

Когда обсуждается вопрос о сравнении потенциальных возможностей счетных машин со способностями человеческого интеллекта, обычно в качестве недоступных машине качеств называют чисто человеческие свойства: фантазию, способность к творчеству, «свободу воли». Сюда можно добавить еще свойства, которые не всегда упоминаются: эгоизм, эмоции и, наконец, способность размножаться.

Если говорить только о существующих и эксплуатируемых сегодня машинах, то этот перечень свойств значительно расширится. Но мы будем обсуждать не только свойства современных утилитарных устройств, но и более широкого класса возможных конструкций машин, игнорируя их утилитарную ценность и экономическую целесообразность их эксплуатации.

Попытаемся оценить, в какой степени фантазия и способность к творчеству являются для машины «запретной

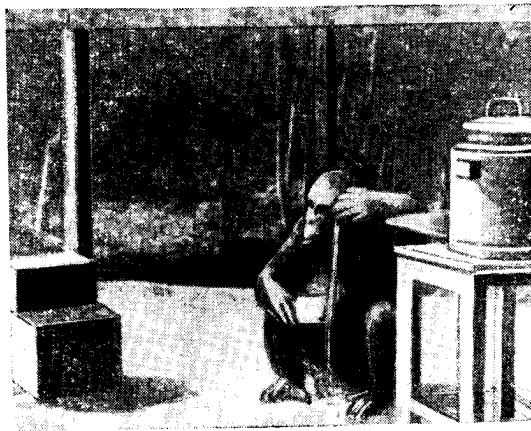


Рис. 9.23. Обезьяна наливает в кружку воду из бака.

областью». Для этого прежде всего следует дать определение этих понятий.

Если человек фантазирует, он создает нечто (рассказ, рисунок и т. п.), не соответствующее непосредственно наблюдаемой реальности. Плод фантазии — это, например, чудовище с телом быка, крыльями орла и головой человека, стоящее на страже ворот ассирийского храма; это — существо с человеческим телом и козлиными рогами, играющее в роще на свирели, или существо с двойным телом коня и человека; это — история войны с титанами, история божеств и их взаимоотношений со смертными и т. д.

Ничего подобного ни один фантазер никогда не наблюдал, и у нас нет оснований утверждать, что подобные вещи могли быть почерпнуты им из окружающей реальности, заимствованы, скопированы. Несомненно, он их

создал сам, но создал из материала, полученного все тем же путем — путем наблюдения реальности.

Способность любого организма реагировать одинаково на различные раздражения определенного класса лежит в основе различения не конкретных предметов, а классов предметов, т. е. абстракций. Одним из первых абстрактных образов, существовавших еще на заре жизни, был образ, или зародыш понятия, реакция — «пища». Однородные



Рис. 9.24. Обезьяна заливает водой огонь, горящий вокруг апельсина.

признаки различных предметов (признаки, имеющие одинаковое биологическое значение) в процессе выработки условных рефлексов непроизвольно группируются и закрепляются в виде однородной группы. Впоследствии с появлением речи эти группы признаков ассоциируются со словом. Возможно, первыми двумя различными абстрактными классами предметов еще задолго до того, как на земле было произнесено первое слово, были классы «съедобные» и «несъедобные» или подобные им категории. Сама необходимость различать предметы и отбирать из большого числа конкретных объектов полезные и вредные рождает способность к классификации. Животные не могли бы существовать, если бы не обладали этой способностью в большей или меньшей степени.

Таким путем высшие животные могут различать огромное количество не только конкретных предметов, но и классов предметов, и понятия «голова» и «нога» любого животного различаются теперь уже не по степени их съедобности, а в результате длинной цепи натуральных и продолженных условных рефлексов, прочно закрепленных в памяти и связанных с эмоциональными оценками. То же относится, разумеется, не только к предметам, но и к действиям. В результате рождаются оторванные от конкретной обстановки образы: «разделить», «соединить», а первоначально, вероятно, «съесть», «поймать» и т. п.

Если сигналы классов предметов и действий существуют в нервной системе, точнее, в памяти мозга животного, то, естественно, они могут случайно сочетаться друг с другом в фантастических образах и при этом «нравиться» или «не нравиться», привлекать или пугать вследствие эмоциональной окраски.

Посмотрим внимательнее на произведения человеческой фантазии и попытаемся найти в них то, чего мы никогда не знали. К продуктам фантазии относятся только сочетания элементарных образов или понятий, но не сами образы. Нельзя придумать ничего такого, что не было бы ни в чем (ни в деталях, ни в общем) похоже на действительность. Лучшее, что известно в этом жанре, это китайский «белый дракон», изображаемый пустым фоном без рисунка.

Фантазия есть просто активная переработка информации о реальных предметах, применяющая случайный выбор сочетаний при синтезе образов, понятий, сюжета и т. п., или же намеренное извращение привычных сочетаний. Как известно, основными признаками «смешного» являются неожиданность, нелепость (несоответствие привычному и полезному) при отсутствии угрозы.

Необходимость опираться на реальные образы или хотя бы на их «осколки», даже при самой смелой фантазии, часто приводит к серьезным затруднениям при формировании понятий. Известно, например, с каким трудом студенты и учащиеся привыкают к понятиям корпускулярно-волновой природы элементарных частиц. Свойства частицы каждый знает хорошо на основании опыта. Свойства волны несколько хуже, но тоже знает. Но никто никогда не наблюдал непосредственно ничего, что обладало бы одновременно и теми и другими свойствами, как это имеет место для электрона, который как частица остав-

ляет индивидуальный след на фотоэмульсии и образует на ней же интерференционный рисунок, да ко всему остальному еще и не обладает индивидуальностью! *

Свойства элементарных частиц мы не можем «понять» в обычном смысле слова, т. е. сопоставить известными операциями с уже известными явлениями. Мы можем только привыкнуть к новым понятиям и научиться ими правильно пользоваться, несмотря на их «неочевидность», просто потому, что они являются для нас элементарными, изначальными понятиями, информацией о физической реальности. И при ответе на вопрос: «Как Вы себе представляете электрон или квант?» — мы вынуждены воздержаться от описания формы, цвета, шероховатости и ограничиться только разъяснением свойств уравнений Шредингера — де Бройля как наиболее точного описания событий внешнего мира.

Итак, фантазия сводится, очевидно, к особым методам обработки информации, поступающей из внешнего мира. Зная признаки этих методов, можно вложить их в машину и заставить ее фантазировать. На этом пути, вероятно, могут получиться забавные результаты. Музыка, сочиняемая с помощью случайных процессов с заданными распределениями вероятностей, представляет собой один из первых примеров творчества без человека в области искусства.

Человек сам иногда страдает отсутствием достаточно богатой фантазии и оказывается слишком сильно привязанным к информации, получаемой им непосредственно из внешнего мира.

Обратимся к процессу творчества и постараемся представить, как протекает этот процесс у человека. Творчество есть создание чего-то (предмета или идеи), что является новым, т. е. не существовавшим до сих пор. Это может быть энергетическая машина или метод управления ею, использование огня или силы животного, статуя или мелодия, орудие труда или геория и т. д.

В основе творчества, конечно, лежит фантазия. Творчество, помимо фантазии, включает в себе, во-первых, отбор по некоторым признакам «нужных» или «хороших»

* Как известно из опыта, распределение энергий электронов подчиняется статистике Бозе — Эйнштейна, которая в отличие от статистики Больцмана, применимой для молекул и атомов, не учитывает перестановки индивидуальных электронов в ячейках фазового пространства.

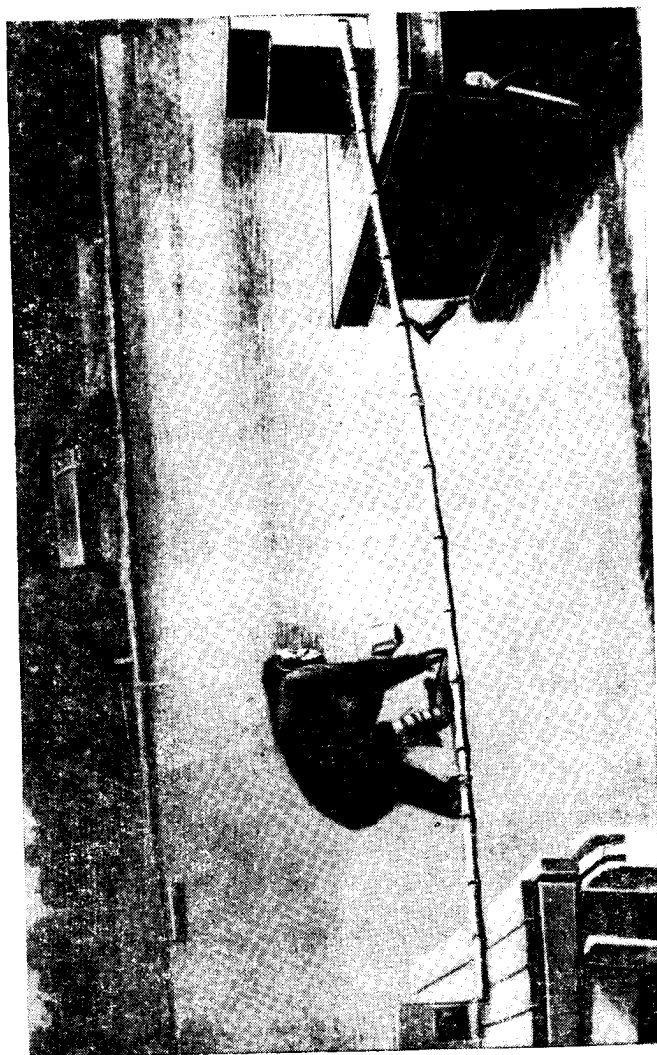


Рис 9 25 Обезьяна переправляется с плота на плот.

продуктов фантазии, а во-вторых, создание реальных предметов в соответствии с выбранной «фантастической» идеей.

Академик И. П. Павлов ставил много опытов с человекообразными обезьянами для проверки их способности к решению творческих задач. Эти опыты весьма поучительны для выяснения самого понятия творчества.

Обезьяна получала на тарелке лакомство — апельсин, брала его с тарелки и с удовольствием съедала. После



Рис. 9.26. Фантастический образ — страж Ассирийского дворца.



Рис. 9.27. Фантастический образ — Кентавр.

этого апельсин окружали кольцевым сосудом со спиртом, спирт горел, и обезьяна при попытке достать апельсин обжигала лапу. Она искала способа избежать ожога. Обезьяна научилась заливать огонь водой; для этого ей пришлось научиться последовательно, в порядке постепенного усложнения задачи, пользоваться кружкой и бачком, из которого она черпала воду, ходить к бачку за водой, пользоваться краном бачка, когда крышка его была закрыта, добираться до бачка, стоящего на плоту, по досочке, перекинутой через пруд*. Наконец, был поставлен

* Фотографии любезно предоставлены Н. Н. Ладыгиной-Котс.

последний опыт. На пруду плавали два плота, соединенные узким мостиком. На одном плоту находилась обезьяна, апельсин, окруженный горящим спиртом, и кружка в руках у обезьяны. На другом — бачок без воды. Все привычные способы добычи воды не давали успеха, несмотря на затрату усилий и преодоление мостика (обезьяны — не любители купаться). Кругом был целый пруд воды, но это была не та вода, которую знала обезьяна, видевшая ее в кружке. Вода в пруду не была связана с кружкой, огнем и апельсином, была неизвестна и потому бесполезна. Отсутствовала ассоциация — реальная связь в мозге обезьяны. Налицо была ситуация для творческого открытия воды в пруду. Долго обезьяна повторяла привычные безуспешные попытки, затем оставила их, по-видимому, вместе с надеждой получить апельсин. И стала развлекаться, чем могла. И вот тогда, хотя далеко не сразу, обезьяна совершила «научное открытие». Оно, как все гениальные открытия, было до смешного простым и пришло, как это часто бывает, случайно. Играя, обезьяна случайно зачерпнула в кружку воды из пруда и узнала ее. Ассоциация вызвала из памяти стереотипную последовательность действий... Апельсин был съеден.

Что же мешало обезьяне отождествить воду в пруду и воду в кружке? Отсутствие прямого опыта и вытекающих из него ассоциаций. Случай восполнил этот пробел, а пужда в воде (в кружке), т.е. условно-рефлекторная связь образа воды (в кружке) с цепью реакций, имеющих результатом апельсин и удовольствие от этого лакомства, заставила обезьяну обратить внимание на воду (в кружке), когда она зачерпнула ее (из пруда). Будь обезьяна сыта, она могла бы не заметить своего открытия и «обезьянья наука» на этот раз не совершила бы прогресса.

Человеческие изобретения и открытия совершаются двумя путями: путем подражания и путем находок. Один из первых автомобилей имел четыре ноги и переставлял их, как четвероногое животное. Один из первых паровозов имел привод на кулисный механизм, упирившийся одним концом в землю, как веслом в воду или шестом в дно. Это — изобретения по подражанию, напоминающие действия обезьяны, бегавшей через пруд за водой из бака, по привычке. Известно много изобретений — находок, когда ошибка в задуманном процессе приводила к неожиданным хорошим результатам. Однако изобретения и открытия делаются не столько путем экспериментальных по-

исков, сколько «выдумыванием» при наличии настоятельной необходимости.

Если у человека болит зуб и не дает ему покоя, большой мечется от боли, он совершает массу движений, не имеющих ни плана, ни смысла. Если животное хочет вырваться из клетки, оно совершает такой же беспорядочный и настойчивый поиск выхода. Одно из хаотических движений, или их комбинация, может привести к желаемому результату.

Если изобретатель или ученый горят желанием решить некоторую задачу, то это неотвязное желание часто не

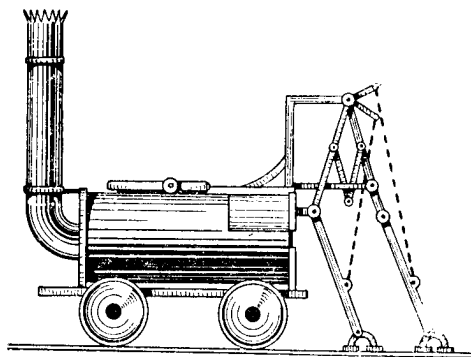


Рис. 9.23. Паровоз системы Брумтона использует в качестве двигателя кулисный механизм, упирающийся концами в почву наподобие ног животного.

дает покоя не меньше, чем зубная боль. Человек, размышляя, перебирает в уме без определенных методов и плана всю информацию относительно интересующего его предмета. Рано или поздно он натывается на решение, отбросив предварительно много вариантов. Эмоциональная установка — желание найти решение — заставляет его отметить благоприятную находку, обратить на нее внимание и остановиться на ней. Впоследствии путь к решению аналогичной задачи уже не будет столь извилистым. Память позволит восстановить кратчайший путь к цели.

Этот метод, называемый методом проб и ошибок, в общем, одинаков и для животного, пытающегося выйти из клетки, и для ученого, отыскивающего, например, новый математический аппарат для решения абстрактной задачи. Разным является только материал.

Творчество есть решение новой задачи, поставленной

внешней обстановкой, методом поисков и находок. При этом важную роль играет внимание. Когда экспериментирование с реальными объектами заменяется в процессе творческого размышления упорядоченным или беспорядочным перебиранием запасенных в памяти образов, понятий — точнее их изоформных изображений в виде сигналов нервных возбуждений — внимание автоматически и бессознательно останавливается на тех сигнальных описаниях событий, на тех идеях, которые дают решение задачи или ее части. Эта невольная фиксация внимания является непременным условием успеха.

Однако тот же самый факт — невольное фиксирование внимания на самом важном — может порой сослужить дурную службу. Достаточно сказать, что техника циркового искусства фокусников (иллюзионистов) основана именно на этом явлении — невольном для зрителя сосредоточении внимания на том, что представляется самым важным в происходящем на арене и что в действительности является приемом артиста, направленным на отвлечение внимания от действительно важных для результата фокуса действий.

Такие же «фокусы» проделывает инстинктивное сосредоточение внимания на «самом важном» и без помощи фокусника. Многие предрассудки и суеверия имеют именно такое происхождение. Совпадения «примет» или снов с важными событиями, особенно с несчастьями, привлекают к себе внимание (часто болезненно возбужденное) и надолго остаются в памяти. Несовпадение как факт, не имеющий значения, не привлекает внимания и забывается или просто не замечается. Объективная регистрация примет и событий позволяет обойти фокусы внимания и строго устанавливает отсутствие взаимосвязи между приметами, предчувствиями и событиями, ими «предсказываемыми». Естественные испытатели должны осторожно применять объективные статистические методы, дабы избежать серьезных ошибок и заблуждений, обязанных объективному несовершенству человеческой психики, которое, однако, в других случаях является необходимым.

Несомненно, что поиск и нахождение наилучшего решения, метод проб и ошибок, может быть заложен в машину, и в этом смысле акт творчества в принципе машине не чужд.

Способность к размножению или самовоспроизведению резко отличает существующие сегодня машины от живых

организмов. Однако в принципе возможно создание машины, которая сама создавала бы себе подобную. Для этого, прежде всего, необходимо предоставить машине сырье или полуфабрикаты, которые она должна будет использовать для построения новой машины, подобно тому, как животные используют сложные органические вещества, входящие в состав пищи, для синтеза тканей будущего организма. При условии, если машине будет поставаться в нужном количестве «пища» в виде радиоламп, проводов, моторов, фотоэлементов и так далее, и при условии, что сама машина будет обладать достаточной сложностью структуры и богатством функций, машина сможет самостоятельно построить новую машину, которая, в свою очередь, если она не «умрет от голода», построит следующую и так далее.

Строгое доказательство этой возможности было проведено математиком Джоном фон Нейманом, который построил математическую модель устройства саморазмножающейся машины. В описании фон Неймана каждая деталь машины занимает элементарную ячейку пространства, что позволяет свести задачу к геометрии. Используются три вида элементов машины: логические или комбинаторные элементы, проводящие элементы и исполнительные элементы. Первые имитируют нейроны животных, вторые — нервные проводящие пути, последние — мышцы и железы.

Логический элемент модели имеет возможность пребывать в различных состояниях, которые могут меняться во времени. Время, как и пространство, квантовано, т.е. разбито на дискретные участки — «циклы». Состояние элемента в данном цикле зависит от состояний его самого и состояний соседних элементов в предыдущие циклы. Состояние логических элементов может передаваться по цепям проводящих элементов к исполнительным элементам. Исполнительные элементы воздействуют на окружающие детали (сырье), выбирая нужные и соединяя их в схемы в соответствии с получаемыми сигналами.

Исходную информацию машина получает от программы, записанной на серии элементов машины и занимающей часть общего объема машины. В программе могут быть записаны любые возможные операции над деталями, окружающими машину (над сырьем). Согласно программе машина может собрать любую возможную схему. Для того чтобы машина воспроизводила сама себя, достаточно

записать в программе сборку такой же машины, как исходная, копирование программы, передачу копии программы новой машине и запуск ее в работу по программе. В результате, процесс самовоспроизведения будет повторяться и продолжаться до тех пор, пока хватит сырья или пока размножившиеся машины не войдут в конфликт друг с другом, начав мешать друг другу синтезировать новые машины.

По подсчетам фон Неймана минимальное число элементов, необходимое для построения самовоспроизводящейся машины, составляет около 200 000 (около 50 000 элементов самой машины и около 150 000 элементов для записи программы. Модель самовоспроизводящейся машины фон Неймана по своей схеме во многом сходна с воспроизведением живой клетки при ее делении. Программа машины, которая копируется и в копии передается новой машине, напоминает набор хромосом клетки, определяющих, в общем, строение и развитие самой клетки.

Итак, свойство самовоспроизведения или размножения не чуждо машине, хотя бы в принципе. До сего времени самовоспроизводящиеся машины не были реализованы. В этом не было необходимости, так как практического применения такие машины сейчас, по-видимому, не нашли бы, а принципиальная сторона вопроса достаточно ясна и в математической формулировке, без реализации. Однако некоторые практические работы прикладного характера по автоматизации производства непосредственно примыкают к кругу вопросов, связанных с самовоспроизводящимися системами. К числу таких работ относятся, например, создание полностью автоматизированного цеха для производства радиоаппаратуры.

Цех-автомат снабжен автоматическими станками, автоматическими измерительными устройствами, осуществляющими операции контроля и браковки, и транспортными устройствами. Станки и контрольные устройства связаны друг с другом цепями управления и обратной связи. Это позволяет автоматически поддерживать режим работы автоматов, соответствующий минимальному браку изделий.

Каждый станок-автомат может управляться программой, записанной на ленте, выдавая, в зависимости от содержания программы, тот или иной вид продукции (из числа возможных). Часть станков производит полуфабрикаты, другая часть осуществляет автоматическую сборку

узлов и блоков. В процесс сборки входят также операции промежуточного контроля и управления с помощью обратной связи.

Работа всего цеха координируется общим вычислительным устройством, работающим по некоторой генеральной программе. Цех выдает продукцию типа радиоприемников, телевизоров или других электронных устройств без всякого участия в его работе человека при условии питания энергией и доброкачественным сырьем.

Преимуществом такой системы автоматизации является ее большая гибкость. Можно за короткий срок перестроить цех на выпуск нового изделия, изменив программу управляющего устройства. Одно время в печати публиковались результаты экспериментальных работ по эксплуатации подобного цеха. Эти работы дают некоторое представление о возможности реализовать модель фон Неймана, если бы по каким-либо причинам это оказалось желательным.

К сказанному выше о размножении машин можно добавить еще несколько слов. Можно заметить различия между моделью самовоспроизводящейся машины и живыми организмами. Эти различия идут по двум направлениям. Во-первых, организм не является фиксированной структурой. Его строение и функции непрерывно меняются по мере роста, развития организма и накопления им опыта. Машина же не заменяет своих составных частей, не увеличивает их числа, в ней отсутствует «обмен веществ», хотя структура ее в процессе работы несомненно претерпевает изменения (коммутация элементов управляющим устройством). Во-вторых, машина, в отличие от живых существ, сама себе не добывает «пищи», т.е. сырья, и, таким образом, не может существовать самостоятельно, без помощи человека.

Оба упомянутых различия имеют относительный характер и могут быть, по крайней мере в принципе, сведены к нулю. Вспомнив автоматический цех радиодеталей и блоков, легко представить себе, как такой цех производит запасные узлы и блоки для собственного управляющего устройства и заменяет ими вышедшие из строя или подключает, в случае необходимости, дополнительные блоки (например, увеличивая объем памяти). По-видимому, машина с помощью автоматического цеха, становится способной к регенерации и росту.

Что касается самостоятельного добывания машиной сырья, то, даже не обсуждая этой возможности (которая,

однако, не представляется заранее исключенной), следует сразу же заметить, что не только машины, но и животные не способны самостоятельно обеспечивать себе все необходимые для жизни вещества. Обязательным условием жизни животных является наличие растений, которые служат им пищей. Только растения, как известно, способны синтезировать органические вещества из неорганических, используя энергию солнечного света (фотосинтез). Животные не способны этого сделать и потому должны получать готовые органические вещества, поедая растения или других животных. Можно, таким образом, утверждать, что животные могут существовать только за счет жизнедеятельности растений подобно тому, как машины могут получать сырье для развития и «размножения» только при условии существования и деятельности человека.

По-видимому, отмеченные выше отличия машин от животных не являются коренными.

Обладает ли человек «свободой воли»? Иначе говоря, свойственно ли человеку хотя бы иногда, в отличие от автомата, выбирать такой образ действий, который не определяется однозначно и фатально сложившейся внешней ситуацией? Несомненно, человек способен действовать, и не слишком редко, вопреки узко понимаемым эгоистическим интересам. Голодный он способен делиться с товарищем, способен погибнуть ради спасения других и т. д.

Является ли, однако, такое поведение действительно «свободным», т. е. независимым от цепи причин и следствий? Если мы дополним наш вопрос и спросим, может ли человек действовать независимо от сложившейся внешней и внутренней обстановки, понимая под внутренней обстановкой весь опыт, накопленный в его памяти, и природные качества, то ответ, несомненно, будет отрицательным. Природные качества индивидуума причинно обусловлены. Приобретенные им понятия, знания, склонности являются конечным звеном причин и следствий в истории его существования. И если он может принимать решения о своих поступках в соответствии со своими желаниями, в желаниях своих он не властен, они с немумолимостью закона причинности диктуются суммой внутренних и внешних причин. Если же читатель захочет доказать противное, нарочно совершить «свободный», беспричинный поступок, то он это сделает потому, что ему это навеяно желанием возразить.

«Идея детерминизма, устанавливая необходимость че-

ловеческих поступков, отвергая вздорную побасенку о свободе воли, нисколько ни уничтожает ни разума, ни совести человека, ни оценки его действий» *.

Так называемая «свобода воли» есть иллюзия, порожденная, с одной стороны, необозримым разнообразием причин, настоящих и прошлых, которые определяют поведение человека, а с другой, — ограниченностью объема внимания, не позволяющего учесть все эти причины одновременно. Подсознательная деятельность мозга несравненно богаче, чем сознательная **. Чтобы понять это, достаточно представить себе, как много мы могли бы вспомнить фактов, о которых редко вспоминаем и которые постоянно хранятся в памяти вне сферы внимания.

Машина не испытывает эмоций, и действия машины не связаны с ее собственной пользой. Впечатления и реакции человека всегда эмоционально окрашены и имеют в виду собственные интересы, эгоистические в широком смысле этого слова. (Если человек правильно воспитан и морально здоров, его собственные интересы окажутся социально обусловленными и его «эгоизм», в нашем понимании, состоит в соблюдении интересов коллектива.)

Эмоциональная окраска тесно связана с «эгоистическими» интересами индивидуума и является непосредственным критерием (очень часто неосознанным) для выбора реакции. Эмоции, представляющие, таким образом, некоторый руководящий принцип поведения, являются тем не менее продуктом деятельности филогенетически древних мозговых центров, т. е. образований, имеющих у всех животных. Подавление эмоций, по крайней мере исключение влияния их на поступки, является чертой чисто человеческой, обязанной коре головного мозга. У человека

* В. И. Ленин. Соч. Т. 1, стр. 142.

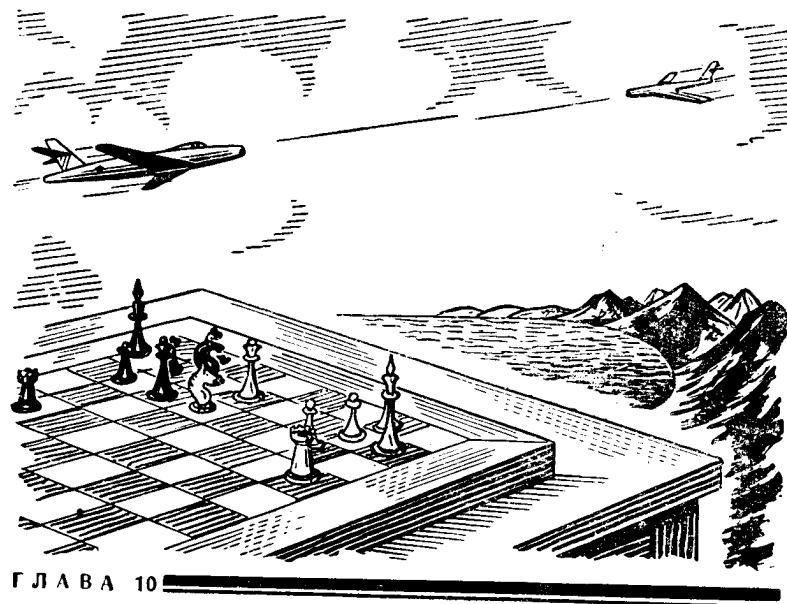
** Подтверждением тому служат случаи, когда различные люди в сходной обстановке, действуя «свободно», поступают одинаково. Люди сходных темпераментов, долго живущие или работающие вместе, иногда, не стовариваясь, одновременно начинают одну и ту же фразу. Большой класс людей одинаково решает следующую задачу. Испытуемому для отвлечения внимания предлагают вести счет, называя только четные числа, в порядке их возрастания, или числа, кратные трем. Одновременно, не прекращая счета, ему предлагают быстро назвать домашнюю птицу, часть лица, фрукт, русского поэта. В качестве ответа в подавляющем большинстве случаев оказываются следующие: курица, нос, яблоко, Пушкин. Испытуемый бывает неизменно искренне удивлен, когда ему показывают бумажку, на которой заранее написаны его ответы (бумажку эту ради эффекта кладут под его ладонь до начала опыта).

сильные эмоции проявляются иногда в виде произвольных действий, не имеющих очевидного характера биологической целесообразности (смех, плач). Однако в целом направление поступков, руководимых эмоцией, хотя и нуждается в серьезных поправках со стороны разума, все же имеет тенденцию к обеспечению благополучия индивидуума. И если машину, снабженную исполнительными механизмами, можно запрограммировать так, чтобы она сама себя разрушала, то человек с нормальной психикой охраняется от подобного поведения отрицательными эмоциями страдания. Случаи самобичевания и самоистязания на почве религиозного фанатизма относятся к числу патологических, но даже в подобных случаях человек испытывает добровольно причиняемое страдание только ради ложно понятого «высшего блага».

Несомненно, было бы очень интересно наблюдать работу машины, которой свойственны внешние проявления эмоций, машины, которая работала бы то с лихорадочным подъемом, то с меланхолической медлительностью в зависимости от успеха собственной деятельности и предлагаемых задач. Вряд ли, впрочем, это улучшило бы эксплуатационные качества машины как полезного орудия. Пути реализации такого эксперимента еще не вполне ясны. С утилитарной точки зрения в ряде задач от машины требуются некоторые свойства, подобные животному эгоизму, однако более широкие. В ряде задач (некоторые из них описаны в следующей главе) машине необходимо выбрать решение, являющееся оптимальным в смысле выполнения некоторого критерия. Этот критерий не обязательно «эгоистичен», т. е. устанавливается не исходя из «интересов» самой машины. Установление критерия может быть поручено самой машине, но опять-таки принцип выбора должен быть в нее заложен извне. Может ли такой принцип появиться у машины самостоятельно?

Живые организмы, от простейших до высших, обладают таким принципом установления критериев целесообразности. Этот принцип можно, по-видимому, сформулировать различным образом, но общий смысл его сводится к сохранению биологического вида.

Наличие такого рода «видового эгоизма» у животных и отсутствие его у машин представляется нам естественным различием между ними. К возможности появления «эгоизма» у технических систем мы еще вернемся в гл. 11.



ГЛАВА 10

ИГРА

В процессе работы машины могут решать и решают разнообразные по характеру математические и логические задачи. Человеческие существа в процессе существования, индивидуального или коллективного, рационально реагируют на внешнюю обстановку, приспосабливаясь к ее изменению и извлекая из внешнего окружения необходимые для существования вещества, энергию и информацию. Реакция на внешнюю обстановку также требует решения задач различных типов. Часть этих задач, как мы уже видели, может быть переложена на плечи машины.

Из всего разнообразия решаемых человеком задач мы остановимся подробнее на одном типе, отличном от прочих, требующем принципиально отличного подхода и к тому же чрезвычайно распространенном и важном.

Мы хорошо знаем задачи, в которых результат решения однозначно вытекает из исходных данных. Задачи этого типа характерны тем, что в их постановке учтены все условия, влияющие на результат. Ответ функционально связан с исходными данными.

Задачи другого типа возникают в тех случаях, когда на результат оказывают влияние некоторые случайные факторы, не могущие, по тем или иным причинам, быть точно учтенными в условиях задачи. При заданных точно условиях задачи результат получается неоднозначным. В сущности результатом решения задач такого рода является не число и не функция, а распределение вероятностей, заданное над возможными значениями результата. Естественно, что в исходных данных задачи неучитываемые точно факторы случайного характера также должны задаваться распределением вероятностей.

В таких задачах стохастического типа наличие неконтролируемых факторов не препятствует получению практически ценных результатов, только формулировка этих результатов оказывается отличной от формулировки результатов задач функционального типа: результат формулируется не для одного частного случая, а для большого числа аналогичных случаев. В каждом отдельном случае или при отдельном испытании результат может больше или меньше отклоняться от среднего значения в ту или другую сторону.

Существует также ситуация иного характера, когда не находящиеся под контролем факторы почти всегда дают наименее благоприятный результат. Это происходит в тех случаях, когда неконтролируемая часть условий задачи управляется не случаем, а разумным существом — оппонентом, противником, интересы которого противоположны нашим. Эта ситуация соревнования, борьбы, весьма характерна для существования отдельных биологических индивидуумов и коллективов. К этому необходимо еще добавить, что от успеха или неуспеха в этой ситуации часто зависит дальнейшее существование индивидуума.

Учитывая жизненную важность ситуации соревнования, следует попытаться ответить на вопрос: является ли возможность выбора правильного поведения в ситуации соревнования присущей только живому существу или и в этом случае, как и во многих других, на помощь человеку могут прийти машины и решать за него задачи лучше, чем он сам. Вопрос этот возникает потому, что обычные логико-математические и статистические методы подхода к решению задачи, которыми могут оперировать машины, как правило, оказываются непригодными для выбора поведения в ситуации соревнования с разумным оппонентом. Поставив задачу обычным методом, приняв некото-

рые предположения о действиях противника и сделав из них выводы, мы решим задачу о нашем поведении, но результат будет зависеть не только от наших действий, но и от действий противника, который, поступая разумно и вопреки нашим предположениям, может достичь того, что мы получим результат, противоположный тому, на который рассчитывали.

Очевидно, что для ситуации соревнования должен быть выработан особый подход и методика решения задачи. Без этого невозможно ни научное доказательное знание, ни использование машин. Такой формально-математический аппарат в основах своих заложен довольно давно (в 1928 г.) работами Вальда, Неймана и др. в виде математической теории игр, имеющей большое принципиальное значение и ряд ценных практических применений в экономике, военном деле и других областях. Формализация соотношений, имеющих место в соревнованиях или «играх», позволяет надеяться на то, что мощная поддержка человеческому интеллекту со стороны машин может быть оказана и в этой области.

Понятия теории игр

Рассмотрим вкратце основные понятия, вводимые в математической теории игр. Игрой в смысле этой теории именуется не только игра в обычном житейском смысле, т. е. шахматы, шашки, домино и карты, но и многие другие виды соревнований, гораздо менее безобидные. Именно поэтому теория игр и имеет практическое значение, так как вряд ли стоило бы прилагать большие усилия ради создания теории только развлекательных игр.

В игре участвуют два или более игрока или команды. Каждая игра протекает согласно правилам. Каждый игрок (или команда) в соответствии с правилами по очереди или одновременно совершает некоторые действия или ходы, выбирая каждый раз один какой-нибудь ход из некоторого множества различных возможных ходов, допущенных правилами. После каждого хода игрокам становятся известны некоторые сведения о ходе игры. Сумма этих сведений представляет собой очередную ситуацию, в которой оказывается игрок и на основании знания которой игрок должен выбирать очередной ход. Правила игры определяют, когда игра кончается и каков результат данной партии этой игры, т. е. кто и что выиграл.

Выбор хода игроком на основании знания ситуации является его реакцией на внешнюю обстановку, в состав которой входит и реакция его противника. Эта последняя зависит, в свою очередь, от действий первого игрока.

Оставим пока в стороне усложнения этих определений, вносимые случаем. В развлекательных играх, например в шашках или шахматах, понятия правил игры, определяющих очередность, порядок ходов и выигрыш, не требуют пояснений, так как являются общеизвестными и строго формализованными. Для игр, представляющих собой реальную борьбу, «правилами» игры являются условия физической реализуемости тех или иных действий, а проигрышем — физическая невозможность продолжать борьбу. В ходе борьбы наличные силы могут использоваться тем или иным способом, причем в каждый данный момент всегда существует выбор из нескольких возможностей для использования этих сил. Таким образом, каждая из борющихся сторон (игроков) выбирает свое действие (ход) из числа физически возможных и в результате в конце игры одерживает победу или терпит поражение.

Стратегия

Представим себе, что игрок до начала игры составляет исчерпывающий план игры, т. е., во-первых, перечисляет все возможные ситуации, в которых он должен делать ход, и, во-вторых, каждой такой ситуации приписывает определенное решение, т. е. заранее решает, какой ход он сделал бы из числа возможных в данной ситуации. Вместо полного перечисления всех ситуаций он может дать правило или систему правил, которые однозначно позволяют получить тот же результат и провести игру. Будем называть такое перечисление ходов или правило, определяющее игру нашего игрока, обобщенным ходом или стратегией игрока. Определение стратегии — задача несравненно более трудоемкая, чем проведение реальной партии данной игры. В реальной партии игрок должен принимать решения только для тех ситуаций, которые фактически встречаются, в то время как для определения стратегии, кроме упомянутых, он должен принять решения еще и для всех прочих возможных ситуаций, которые в реальной партии могут и не встретиться. Так, например, играя в шахматы черными фигурами и делая первый ход, игрок в реальной партии имеет дело с одной

единственной ситуацией (после первого хода белых) и должен принять одно решение. Составляя же стратегию игры, тот же игрок для первого хода должен иметь дело с двадцатью ситуациями (после каждого из двадцати возможных первых ходов белых).

Составление стратегии, например, для шахмат,— дело невыполнимое по своей трудоемкости, даже если этот труд механизировать в доступной сегодня степени. Нам, однако, это обстоятельство, имеющее место для многих игр, не будет смущать, и мы удовлетворимся тем, что стратегии принципиально могут быть определены; они существуют, и этого для наших целей достаточно.

Различных стратегий в данной игре может быть очень много, и перечисление их или подсчет их числа представляет собой задачу не менее трудоемкую, чем определение одной стратегии. В некоторых играх стратегии могут попарно отличаться друг от друга бесконечно мало, образуя континуум — пространство непрерывно меняющихся величин.

Однако для любой игры можно указать, опять-таки хотя бы принципиально, все стратегии данного игрока. Одновременно таким же образом можно перечислить и все стратегии другого игрока.

Если до начала партии оба игрока (или все игроки, если их больше двух) совершили «обобщенные ходы», т. е. выбрали каждый из всех возможных стратегий одну, то дальше их участие в игре оказывается необязательным и партия может быть проведена в их отсутствие посредником (судьей), так как все ходы, которые делали бы игроки, заранее определены стратегией и судьей известны. Таким образом, выбор стратегии всеми игроками однозначно определяет ход игры и ее результат. Заметим попутно, что один и тот же результат игры может быть достигнут применением различных стратегий. Однако, вообще говоря, меняя стратегию, игрок может достигнуть изменения результата.

Результат игры не всегда представляет собой выигрыш или проигрыш, иногда результатом игры может явиться некоторая ситуация, например, обладание некоторыми предметами, территорией, уничтожение некоторых объектов и т. д. В таких случаях вопрос о выигрыше и проигрыше в игре решается не просто, даже при знании результата игры.

Функция пользы

Для того чтобы решить вопрос о том, выиграл данный игрок или проиграл, необходимо знать, чего добивался он в игре, т. е. узнать предпочтительный для него результат. Лучше всего достигнутый результат оценить некоторой числовой функцией, такой, что игрок предпочитает всегда тот результат, который дает большую по величине функцию. Такую функцию часто называют функцией пользы данного игрока. Для того чтобы в реальных случаях получить такую функцию, приходится расположить все результаты на шкале, отражающей степень предпочтения.

Существует специальный прием для количественной оценки степени предпочтения различных объектов. Расположить в порядке предпочтения несколько объектов обычно не представляет труда. Для этого достаточно сравнить их попарно и в каждом случае решить, какой из них предпочтительнее. Если такого решения для какой-либо пары объектов вынести нельзя, оба объекта имеют одинаковую степень предпочтительности, т. е. выбор между ними безразличен.

Допустим, из упорядоченной по предпочтению последовательности объектов мы выделим три объекта: A , B и C , причем

$$u(A) > u(B) > u(C),$$

где $u(x)$ — функция пользы.

Последнее неравенство говорит о том, что объект A предпочтительнее объекта B и B предпочтительнее C .

Теперь попытаемся решить, на каком „расстоянии“ по предпочтительности расположен объект B от объектов A и C . Считая интервал $u(A) - u(C)$ равным единице, составим выбор из следующей пары объектов: объекта B и комбинации объектов A и C , взятых каждый с некоторой вероятностью.

Можно этот выбор сформулировать иначе. Пусть в течение 100 промежутков времени мы будем получать либо 100 раз объект B , либо 35 раз объект A и 65 раз объект C (или в общем виде: 100 p раз объект A и 100 q раз объект C ; $p + q = 1$). Что мы предпочтем? Меняя p и q , мы найдем такое p_0 , для которого нельзя вынести решения о предпочтении, т. е. выбор будет безразличен. Тогда объект B разместится по шкале предпочтения на расстоянии $p[u(A) - u(C)]$ от объекта C и на расстоянии $q[u(A) - u(C)]$ от объекта A , т. е. будет иметь координату

$$u(C) + (1 - p)[u(A) - u(C)] = u(A) - p[u(A) - u(C)].$$

Итак, построив функцию пользы для всех результатов данной игры, мы получим суждение о том, к какому именно результату стремится в игре данный игрок. Очевидно,

что в игре игрок будет стремиться применить ту стратегию, которая обеспечит ему наибольший выигрыш в установленном выше смысле, т. е. даст максимальное значение функции пользы.

Особенный интерес представляют так называемые парные игры с нулевой суммой, в которых принимают участие две стороны (два игрока или две команды) и где сумма выигрышей обоих игроков в результате игры равна нулю. Иначе говоря, один выигрывает только за счет второго и сколько выиграл первый, столько проиграл второй. В играх этого типа нам достаточно рассматривать только выигрыш одной стороны, так как выигрыш другой имеет ту же величину, но обратный знак (мы считаем проигрыш отрицательным выигрышем). При этом, разумеется, мы предполагаем, что оба игрока совершенно одинаково выражают свое предпочтение относительно исходов игры, и поэтому мы просто говорим об одной и той же функции пользы для обоих как о «выигрыше», который выражается числом.

В некоторых парных играх множество возможных стратегий является конечным и стратегии каждого игрока можно просто перенумеровать. Выбор игроком A некоторой стратегии номер i и игроком B некоторой стратегии номер j полностью определяет исход партии и величину выигрыша a_{ij} (т. е. функции пользы), скажем, игрока A . Перебирая по очереди все пары стратегий (i, j) , мы перечислим все возможные варианты партий и их исходов. Далее можно построить таблицу — матрицу, каждая строка которой будет соответствовать стратегии i игрока A , каждый столбец — стратегии j игрока B , а на пересечении i -той строки и j -того столбца будет записан a_{ij} — выигрыш игрока A (т. е. проигрыш игрока B). Матрица конечной игры, полученная нами, перечисляет, таким образом, все возможные варианты партий данной игры.

Рассмотрим стратегию с номером i игрока A . Его выигрыш будет меняться в зависимости от действий противника, т. е. от стратегии игрока B . Применяя стратегию номер i , игрок A не может быть уверенным, что он получит выигрыш больше, чем наименьшее значение по строке номер i , ибо его противник всегда может выбором своей стратегии добиться этого результата. Минимум по j от величины a_{ij} является гарантированной величиной, хуже которой результат для игрока A не может быть, если

он применит стратегию номер i . Находя этот минимум для всех строк матрицы, т. е. для всех стратегий игрока A , мы можем найти его максимальную величину по i . Обозначим ее через a_A .

Стратегия номер $i=i_1$ игрока A , для которой получен этот максимум (по i) от минимума (по j), или „максимум“, дает, очевидно, наилучший результат из гарантированных наилучших результатов.

Теперь проведем те же рассуждения для игрока B , учитывая, что для него желательным является минимальный выигрыш игрока A , который соответствует максимальному выигрышу игрока B . Для каждого столбца, соответствующего стратегии игрока B с номером j , найдем максимум по i величины a_{ij} , т. е. наилучший результат для игрока B при его стратегии номер j . Проведем это для всех столбцов и найдем минимум по j этой величины. Обозначим найденный минимум (по j) от максимума (по i), или „минимум макс“, через a_B для соответствующей стратегии $j=j_2$.

Теперь сравним величины a_A и a_B и обе пары стратегий (i_1, j_1) и (i_2, j_2) . Могут быть два случая:

1) обе величины равны друг другу и обе пары стратегий совпадают

$$a_B = a_A = a, \quad i_1 = i_2; \quad j_1 = j_2;$$

2) те же величины не совпадают

$$a_B > a_A; \quad i_1 \neq i_2; \quad j_1 \neq j_2.$$

Игры первого типа мы будем называть определенными играми или играми с оценкой, игры второго типа — неопределенными или играми без оценки. Величины a_A и a_B называются соответственно нижней и верхней оценками игры; величина a — собственной оценкой.

В первом случае, т. е. для определенных игр, матрица игры имеет седловую точку, т. е. a максимально на протяжении столбца и минимально на протяжении строки. Во втором случае, т. е. для неопределенных игр, такая точка отсутствует. В седловой точке, в первом случае, пересекаются стратегии, которые называются наилучшими или оптимальными. Игрок, придерживающийся оптимальной стратегии, не рискует получить результат хуже a . Если один из игроков, например A , отклонится от

оптимальной стратегии, он только ухудшит свой результат, если только при этом B играет разумно, так как a представляет собой максимум по стратегиям игрока A . То же произойдет с игроком B , если он отойдет от своей оптимальной стратегии, так как a минимально по стратегиям игрока B , если игрок A играет разумно. (Напомним, что для игрока B выгодны меньшие величины a_{ij} .)

Наличие седловой точки матрицы игры и оптимальных стратегий имеет большое принципиальное значение. Игрок, знающий оптимальные стратегии, имеет преимущество перед своим противником, не знающим их. Построение стратегий и матрицы игры (так называемое решение игры), определение оптимальных стратегий является задачей игрока, желающего вести игру действительно разумно.

Решение игры определенного типа в общем случае может быть получено поочередным вычеркиванием строк и столбцов в матрице игры и исключением соответствующих стратегий обоих игроков, которые являются «не лучшими» по сравнению с другими стратегиями. Вычеркивается та стратегия, у которой некоторые элементы матрицы такие же, а другие — хуже, чем у другой стратегии (с точки зрения данного игрока). Последовательное вычеркивание «не лучших» стратегий в матрице определенной игры по очереди у каждого из двух игроков приводит к тому, что в результате у каждого игрока остается всего по одной стратегии, которые и являются оптимальными.

К числу игр определенного типа, т. е. игр, обладающих «собственной оценкой», относятся все так называемые игры с полной информацией, т. е. игры, в которых игрок перед очередным ходом знает все предыдущие ходы, свои и своего противника.

Рассмотрим для иллюстрации предельно упрощенный пример парной конечной игры с нулевой суммой и полной информацией.

На доске нарисовано в ряд шесть клеток, пронумерованных по порядку. Правила игры формулируются следующим образом. В исходном положении в первой клетке расположена белая шашка (игрока A), в шестой клетке — черная (игрока B). Игрок A двигает белую шашку от первой клетки к шестой, игрок B двигает черную шашку от шестой клетки к первой. Игроки делают ходы по очереди, начиная с игрока A , передвигая каждый свою шашку на одну клетку, в своем направлении. Если следующая (по пути шашки) клетка занята шашкой другого цвета, то оче-

редной ход совершается, минуя занятую клетку (перепрыгивая через шашку противника). Первый ход каждый игрок может сделать по своему выбору, передвигая шашку либо на одну, либо сразу на две клетки. Выигрывает в партии тот игрок, шашка которого первой

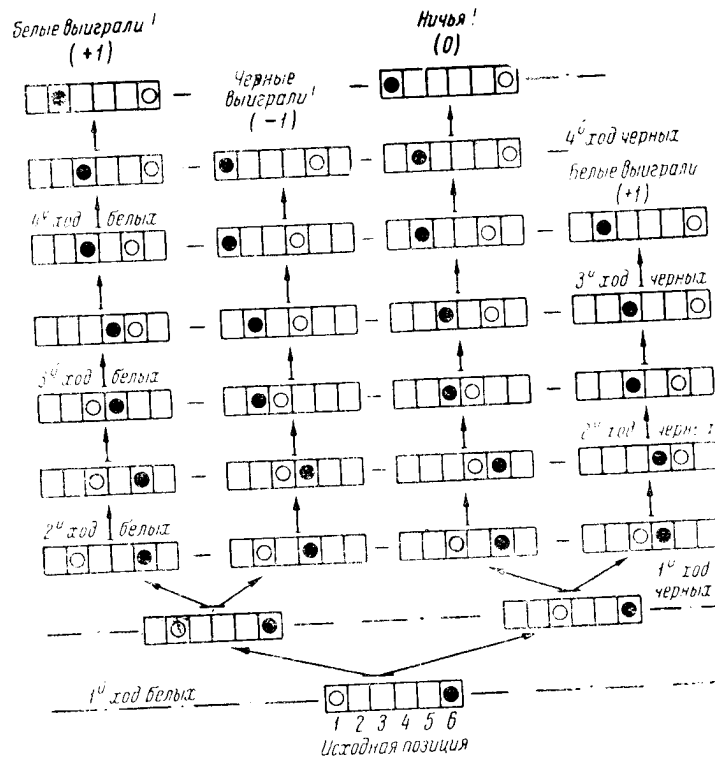


Рис. 10.1. „Дерево игры“.

достигнет конечной клетки, если только шашка противника не достигнет того же следующим ходом, то есть если шашка противника отстанет больше, чем на один ход. В противном случае, т. е. если шашка противника достигнет последней клетки следующим же ходом, имеет место ничья (см. рис. 10.1).

Легко показать, что исход игры определяется выбором обоими противниками первого хода (одна или две клетки). После первого хода белых и первого хода черных игра

протекает однозначно, т. е. каждый игрок может делать только один ход, без выбора.

У белых, таким образом, есть две стратегии, совпадающие с их первым ходом. Стратегии белых:

1А) первый ход на одну клетку (1);

2А) первый ход на две клетки (2).

На каждый ход белых у черных имеется два ответа, т. е. черные имеют четыре стратегии. Стратегии черных:

1В) если белая шашка первым ходом продвинулась на одну клетку, продвинуть черную шашку на одну клетку; если белая шашка продвинулась на две, продвинуть черную на одну (1—1; 2—1);

2В) если белая шашка продвинулась первым ходом на одну клетку, продвинуть черную на одну; если белая шашка продвинулась на две, то продвинуть черную на две клетки (1—1; 2—2);

3В) если белая шашка продвинулась первым ходом на одну клетку, продвинуть черную на две; если белая шашка продвинулась на две, продвинуть черную на одну (1—2; 2—1);

4В) если белая шашка продвинулась первым ходом на одну клетку, продвинуть черную на две; если белая шашка продвинулась на две, продвинуть черную на две (1—2; 2—2).

Составим матрицу игры. Обозначим выигрыш белых через +1, их проигрыш — 1, ничью через 0. Строки матрицы, как мы условились, будут соответствовать стратегиям игрока А (белым), столбцы — стратегиям игрока В (черным).

	1В	2В	3В	4В
1 А	$a_{11} = +1$	$a_{12} = +1$	$a_{13} = -1$	$a_{14} = -1$
2 А	$a_{21} = 0$	$a_{22} = +1$	$a_{23} = 0$	$a_{24} = +1$

Для стратегии 1А игрока А (первая строка матрицы) наихудшими для игрока А по стратегиям игрока В (т. е. наименьшими) являются результаты a_{13} и a_{14} ; для стратегии 2А (вторая строка матрицы) наихудшими для А будут a_{21} и a_{23} . Из числа этих наихудших по стратегиям В результатов наилучшими для игрока А (т. е. наибольшими) по стратегиям игрока А будут результаты a_{21} и a_{23} .

Они представляют собой верхнюю оценку игры, или максимин.

Для стратегии 1В игрока B (первый столбец матрицы) наилучшим для игрока B (т. е. наибольшим) по стратегиям игрока A является результат a_{11} ; соответственно, для стратегии 2В — a_{12} и a_{22} , для 3В — a_{23} , для 4В — a_{24} . Из числа этих наилучших по стратегиям A результатов наилучшим для игрока B (т. е. наименьшим) по стратегиям B является результат a_{23} . Он представляет собой нижнюю оценку игры, или минимакс.

Результат a_{23} является одновременно верхней и нижней оценкой игры. Иначе говоря, нижняя и верхняя оценки игры совпадают и являются собственной оценкой игры. Этого следовало ожидать, так как разбираемая игра принадлежит к играм с полной информацией, т. е. к играм определенного типа.

Стратегии 2А белых и 3В черных являются оптимальными стратегиями, а собственной оценкой игры будет ничья. Результат a_{23} представляет собой седловую точку матрицы игры. Для любого игрока отход от оптимальной стратегии может только ухудшить его результат по сравнению с собственной оценкой игры (ничьей).

Другими примерами игр с полной информацией могут служить шашки, шахматы, крестики и нолики, игра в Ним и т. п. Для простейших из этих игр известно, что в них выигрывает всегда один игрок, если оба играют правильно; либо тот, кто делает первый ход, либо его противник. Этот результат и представляет собой седловую точку матрицы игры, результаты которой имеют только два значения *выиграл* (+1) и *проиграл* (—1). Шашки и шахматы тоже имеют предрешенный исход игры при правильной игре обоих противников, однако огромное число стратегий не позволяет провести до конца решение игры и определить оптимальные стратегии и собственное значение. Таким образом, в шахматы и шашки мы играем только потому, что «не умеем» играть в эти игры правильно, так как неизвестны решения этих игр. И хотя такое утверждение звучит парадоксально, оно тем не менее вполне обоснованно.

Смешанное расширение игры

Кроме игр определенного типа, имеющих седловую точку матрицы, есть еще, как мы знаем, игры неопределенного типа, лишенные этого свойства. Игроки не имеют опти-

мальных стратегий в одной партии. Любая стратегия игрока может быть «битой» противником путем выбора им соответствующей стратегии.

Примером такой игры может быть следующая: два игрока одновременно выбирают целое число и записывают каждый на бумаге, не сообщая другому. Затем выбор обоих объявляется, выбранные числа складываются. Если результат четный — первый игрок выигрывает очко, если нечетный — очко выигрывает второй*.

Стратегией игрока в каждой партии этой игры может быть либо выбор четного числа, либо выбор нечетного. Для любой из этих стратегий найдется стратегия противника, которая приводит к проигрышу. Матрица игры имеет вид

$$M_{ij} = \begin{vmatrix} +1 & -1 \\ -1 & +1 \end{vmatrix}. \quad (10.1)$$

Седловой точки в ней нет. Верхняя и нижняя оценки не совпадают. У игроков нет оптимальных стратегий в одной партии игры, и, ведя игру, ни один из игроков не может быть уверен в том, что его противник выберет какую-то определенную оптимальную, т. е. наилучшую, стратегию.

Аналогичную ситуацию (эквивалентную игру) мы имеем в том случае, если два противника, например на поле боя или на спортивной арене, имеют каждый по два приема или по два вида оружия: первый — 1А и 2А, второй — 1В и 2В. При этом прием 1А дает преимущество игроку А, если противник В применяет прием 1В, или ведет к проигрышу, если противник применяет прием 2В; прием 2А дает выигрыш или преимущество против приема 2В и проигрыш против приема 1В. Примеров такой ситуации очень много.

Еще чаще встречаются примеры, когда приемов, родов оружия, или стратегий существует не две, а больше. Простой пример — игра «Камень, бумага и ножницы»**. Играть

* Ту же игру можно вести с монетами. Игроки кладут на стол каждый свою монету, накрыв ладонью, затем показывают монеты и сравнивают. Если обе монеты лежат вверх гербом или решеткой, выиграл первый, если одна гербом, другая решеткой, выиграл второй. Аналогичную игру можно вести, показывая один или два пальца («Морра»). Игры, одинаковые по способу их ведения, называются эквивалентными.

** Напомним игру. Каждый из двух игроков имеет три карты: «камень», «бумага», «ножницы». Оба одновременно кладут по одной карте на стол. При одинаковых картах — ничья. При разных картах «ножницы» выигрывают у «бумаги», «бумага» — у «камня» и «камень» — у «ножниц».

в такую игру было бы нетрудно, если бы мы имели основание предположить, что наш противник выберет какой-либо определенный образ действий, т. е. приписать его поведенно некоторую априорную вероятность. Однако величина этой вероятности зависит от решения противника, а это решение нам неизвестно.

Неопределенный тип игры приводится к результатам, практически совпадающим с результатами для определенного типа, введением понятия смешанного расширения игры.

Будем рассматривать не одну партию данной игры, а длинный ряд партий и оценим выигрыш и проигрыш в среднем за большое число партий. Такой подход не удаляет нас от реальности, так как реальные игры, независимо от того, развлекательные они или нет, почти всегда имеют характер повторяемости. Если в последовательности партий игрок применяет все время одну и ту же стратегию (например, «герб» или «четное» в неопределенной игре с двойным выбором, описанной выше), то будем называть такую стратегию чистой стратегией. Очевидно, что упрямо придерживаясь чистой стратегии в неопределенной игре, игрок как бы раскрывает свои карты, т. е. делает для противника ясным ход своих действий на будущее. Поэтому противник уже не строит предположений о его действиях в будущем, а получает уверенность в них, просто экстраполируя его поведение. При этом противник выбирает свою стратегию, чтобы все время выигрывать. Чистая стратегия оказывается, таким образом, непригодной. Следует затруднить или сделать невозможным для противника предугадывание действий игрока. Для этого, очевидно, необходимо, во-первых, от партии к партии менять стратегию, а во-вторых, добиться того, чтобы эти изменения не носили никаких следов закономерности, т. е. были бы абсолютно случайными. Всякая закономерность изменений стратегии (например, цикличность) и даже корреляция (в тех случаях, когда изменения стратегии случайны) дает возможность противнику хотя бы иногда правильно «угадывать» действия игрока и «ловить» его на этом. Ведение игры со случайной сменой чистых стратегий от партии к партии называют смешанной стратегией в игре, состоящей из длинной серии партий и называемой смешанным расширением данной игры.

Комбинирование чистых стратегий в смешанной стратегии можно, очевидно, производить по-разному. Одна из

чистых стратегий, чередуясь случайно с другими, может повторяться чаще или реже других, имея ту или иную частоту или вероятность появления. Таким образом, беспорядочное чередование, скажем, двух чистых стратегий с вероятностями p и q ($p+q=1$) дает некоторую смешанную стратегию. Изменяя p и q , мы получим бесконечное множество смешанных стратегий.

Применение смешанных стратегий в смешанном расширении игры сводит игру неопределенного типа к игре определенной, в которой имеются оптимальные стратегии у обоих игроков и оценка игры. Задача правильного ведения смешанного расширения данной неопределенной игры сводится, таким образом, к определению оптимальных вероятностей, с которыми чистые стратегии входят в смешанную. Кроме того, необходимо каким-то образом имитировать случайный процесс с заданными вероятностями чистых стратегий. В случае игры в «чет и нечет», описанной выше, в которой оптимальные вероятности «чета» и «нечета» (или герба и решетки) равны 0,5, случайный процесс может быть имитирован бросанием монеты или игральной кости.

Попытаемся определить оптимальные стратегии смешанного продолжения конечной неопределенной игры, матрица которой имеет вид

$$M_{ij} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = \begin{vmatrix} 9 & -6 \\ -21 & 4 \end{vmatrix}. \quad (10.2)$$

Игрок A , чистые стратегии которого соответствуют строкам матрицы, имеет в своем распоряжении две чистые стратегии 1А и 2А, которые он применяет с вероятностями p_1 (для первой строки матрицы) и q_1 (для второй); $p_1+q_1=1$.

Игрок B , чьи чистые стратегии соответствуют столбцам матрицы, также имеет две стратегии 1В и 2В, которые он применяет с вероятностями p_2 и q_2 , соответственно; $p_2+q_2=1$.

Если игрок A применяет чистую стратегию (верхняя строка), а игрок B — смешанную, результат равен

$$a_{10} = a_{11}p_2 + a_{12}q_2. \quad (10.3)$$

То же для стратегии 2А игрока A (нижняя строка)

$$a_{20} = a_{21}p_2 + a_{22}q_2. \quad (10.4)$$

Если оба игрока применяют смешанные стратегии

$$\begin{aligned} a_{00} &= a_{10}p_1 + a_{20}q_1 = \\ &= a_{11}p_1p_2 + a_{12}p_1q_2 + a_{21}p_2q_1 + a_{22}q_1q_2. \end{aligned} \quad (10.5)$$

Для седловой точки максимум по p_1 совпадает с минимумом по p_2 ; это соответствует

$$\frac{da_{10}}{dp_1} = \frac{da_{10}}{dp_2} = 0. \quad (10.6)$$

Учитывая, что

$$\frac{dq}{dp} = -1,$$

получим

$$p_{10} = \frac{a_{22} - a_{21}}{a_{11} + a_{22} - a_{12} - a_{21}}, \quad (10.7)$$

$$p_{20} = \frac{a_{22} - a_{12}}{a_{11} + a_{22} - a_{12} - a_{21}}. \quad (10.8)$$

Подставляя p_{10} в (10.7), мы обнаруживаем, что для этого (оптимального) значения p_1 в смешанной стратегии значение выигрыша $a_{00}(p_1)$ не зависит от p_2 . Иначе говоря, что бы ни делал игрок B , результат не изменится, если игрок A применяет оптимальную стратегию. То же справедливо для второго игрока.

Для данной матрицы выигрыш для оптимальных смешанных стратегий равен нулю (ничья). Оптимальные стратегии: $p_1 = 0,70$ $p_2 = 0,4$. Иначе говоря, в 100 партиях игрок A применяет чистую стратегию 1А (верхняя строка) 70 раз, стратегию 2А (нижняя строка) 30 раз; игрок B применяет стратегию 1В (левый столбец) 40 раз, стратегию 2В (правый столбец) 60 раз (в случайном чередовании).

Если один из игроков в игре «чет и нечет», описанной выше, случайным образом (на основании бросания монеты) ходит «с чета» и «нечета» с вероятностью 0,5, то игра в среднем идет ничью, независимо от того, что делает противник (если только действия противника не имеют корреляции с ходами игрока, т. е. противник не «шпионит»). Безразлично в этом случае, ходит ли противник все время «с чета», или все время «с нечета», или же случайно меняет «чет» и «нечет».

Если игрок изменит свою смешанную стратегию, отойдя от оптимальной, и станет применять «чет» (или «нечет») чаще, чем в половине случаев, то противник может на этом выиграть, перейдя на соответствующую чистую стратегию и добившись, таким образом, некоторого выигрыша.

В общем случае решение игры неопределенного типа можно получить путем совместного решения системы уравнений;

$$\sum_i p(x_i) M(x_i, y_j) = v, \quad (10.9,a)$$

$$\sum_i p(x_i) = 1, \quad (10.9,b)$$

$$\sum_j q(y_j) M(x_i, y_j) = v, \quad (10.9,c)$$

$$\sum_j q(y_j) = 1, \quad (10.9,d)$$

где $p(x_i)$ и $q(y_j)$ — вероятности появления чистых стратегий x_i и y_j , соответственно, в оптимальной смешанной стратегии.

Уравнения (10.9,a) и (10.9,c) записываются для всех y_j и всех x_i , входящих в оптимальную стратегию противника; v — собственное значение игры.

Для простых игр выводы представляются тривиальными. Однако для более сложных игр, где одного здравого смысла недостаточно, математический аппарат оказывается ценным. Правда, задачи о решении конкретных игр, имеющие практическое значение, чаще всего бывают чрезвычайно сложны. Многие весьма серьезные и важные практические задачи в настоящее время ждут своего решения.

Случайные ходы

До сих пор мы говорили об играх, в которых встречались только личные ходы, т. е. ходы, сделанные игроком по собственному произволу. В общем случае, кроме личных ходов, в игре встречаются, согласно правилам игры, также ходы случайные, примером которых может служить сдача карт в карточной игре.

Когда правила игры требуют совершения случайного хода, включается некоторый случайный процесс, который определяет исход случайного хода в соответствии с некоторым заданным распределением вероятностей. В этом случае результат игры не определяется однозначно выборами личных ходов на протяжении партии, т. е. личной стратегией игрока. Одна и та же пара стратегий, применяемая обоими игроками, будет давать различные результаты игры в зависимости от исходов серии случайных хо-

дов в данной партии («стратегии случая»). Зная распределение вероятностей при отдельных случайных ходах, можно найти распределение вероятностей всех комбинаций исходов всех случайных ходов партии. Это распределение при известных личных стратегиях игроков дает возможность определить распределение вероятностей результатов игры. Таким образом, в игре со случайными ходами применение игроками определенных личных стратегий дает не определенный результат, а некоторое определенное распределение вероятностей π над возможными результатами x . Выбирая личную стратегию, игрок должен отдать предпочтение одному из этих распределений вероятностей результатов, а не самим результатам игры. При числовой оценке результатов с помощью функции пользы $u(x)$ можно воспользоваться математическим ожиданием этой функции, вычисленным для распределения вероятностей результатов π .

Полученное число

$$u(\pi) = \sum_{x=0}^n \pi(x) u(x) \quad (10.10)$$

представляет собой функционал, т. е. число, зависящее от некоторой функции, в данном случае от функции $\pi(x)$ распределения вероятностей результатов игры. $u(\pi)$ есть функция пользы для игры со случайными ходами. Таким образом, решение игр со случайными ходами производится аналогично играм без случайных ходов, с привлечением теории вероятностей. Некоторые игры (например, лото, рулетка) целиком сводятся только к случайным ходам; личные ходы и стратегии в них отсутствуют.

В ряде конкретных игр случайные ходы совершаются одновременно с личными. Так, например, посылка мяча в спортивных играх (теннис, баскетбол) совершается игроком в определенную точку площадки, и выбор этой точки есть результат личного хода. Однако мяч всегда попадает в выбранную точку с некоторой ошибкой, являющейся результатом неточностей движений игрока и имеющей случайный характер. Разброс мячей не зависит от воли игрока (хотя зависит, конечно, от уровня его спортивной техники) и является случайным ходом в игре, который совершается одновременно с личным ходом. Уменьшение ошибок отражает рост спортивной техники игрока, а правильный выбор направлений ударов — его спортивную тактику, или «стратегию» в смысле теории игр.

Почти все спортивные игры (а также игры типа военных действий) предъявляют к игрокам смешанные требования: с одной стороны, требования высокого качества спортивной техники, точности удара, силы, выносливости, а с другой, — умения применять эту технику правильно, т. е. выбирать стратегии, близкие к оптимальным, или же, если

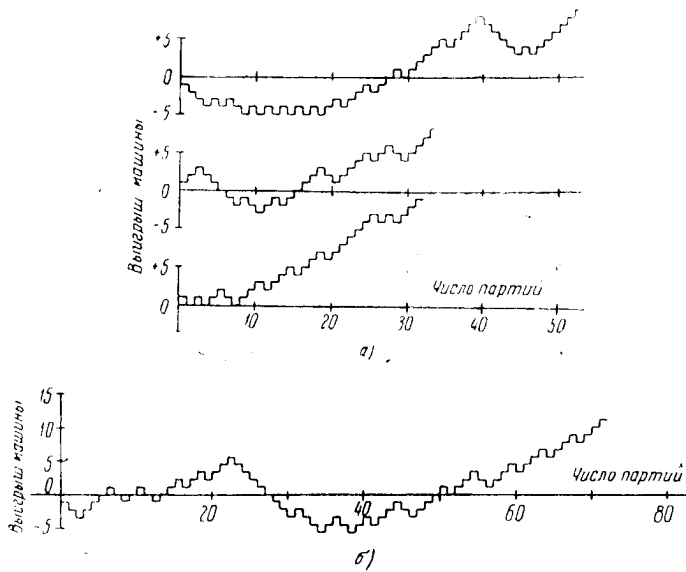


Рис. 10.2. Ход игры машины с человеком в «чет и нечет». а) последовательные процессы выработки стратегии при применении противником машины одной определенной комбинации ходов, б) ход игры при изменении комбинаций, применяемых противником машины.

стратегия, выбранная противником, уклоняется от оптимальной, то стратегию, дающую преимущество.

Существуют спортивные соревнования такие, как стрельба, городки, соревнования в прыжках, метаниях и т. д., в которых «стратегическая часть» игры отсутствует и игра сводится к соревнованию в точности, силе или выносливости. Эти соревнования по сути дела не являются «стратегическими играми». Однако такое, казалось бы, далекое от стратегии соревнование, как перетягивание каната, позволяет улучшить результат за счет маневра величиной усилия, т. е. допускает применение стратегии.

Бокс, фехтование, борьба являются, несомненно, стратегическими играми. Соотношение сил и владение приемами

ми ведения боя создают те исходные предпосылки, на которых должна строиться смешанная стратегия применения приемов и контрприемов на протяжении боя. И нередко применение «с умом» ограниченного арсенала не очень сильных приемов оказывается равноценным большой силе, быстрой и точности движений.

Мы не можем входить здесь в подробности спортивной или военной техники и тактики. Несомненно, что ведение любой игры на основе знания ее решения всегда дает преимущество игроку над его противником, не знающим этого решения.

Машины, играющие в стратегические игры

Мы уже говорили, что решение большинства конкретных игр, представляющих практический интерес, требует очень большого труда. Однако практический интерес к нахождению решения некоторых игр, как нетрудно понять, весьма велик. Отсюда, естественно, следует стремление использовать для получения решения игр и для ведения сложных игр, решение которых почему-либо не найдено, современные счетные машины и автоматы.

В настоящее время работа по созданию специализированных машин, играющих в игры, или программ игры для универсальных машин развернута довольно широко. О результатах этой работы мы знаем по большому числу сообщений о машинах, играющих в Ним, «крестики и нолики» и т. п. Эти, казалось бы, ничтожные по практическому значению результаты не должны, однако, казаться нам не имеющими никакого значения. Напротив, следует отдавать себе отчет в том, что серьезные применения машинной техники в этом направлении невозможны без предварительного решения элементарных и даже тривиальных задач. Без этого первого шага невозможны последующие. И сегодня сделан уже не только первый шаг на этом пути.

Машины, «умеющие» играть в игры, делятся на три типа. Первый — это машины, способные вести конечную игру с полной информацией с оптимальной стратегией (типа крестиков и ноликов). Второй — машины, играющие в «нерешенные» игры типа шашек или шахмат, т. е. игры, для которых оптимальная стратегия не определена. Третий — машины, рассчитанные на ведение игр неопределенного типа, для которых оптимальная стратегия существует только в смешанном расширении игры.

Ним — игра, состоящая в том, что два игрока берут по очереди со стола некоторое число фишек. Фишки в Ниме разбиваются на несколько групп. Каждый игрок берет за один ход любое число фишек, но только из одной группы, и не менее одной фишки (можно взять сразу всю группу). Игрок, вынужденный взять последнюю фишку, проиграл.

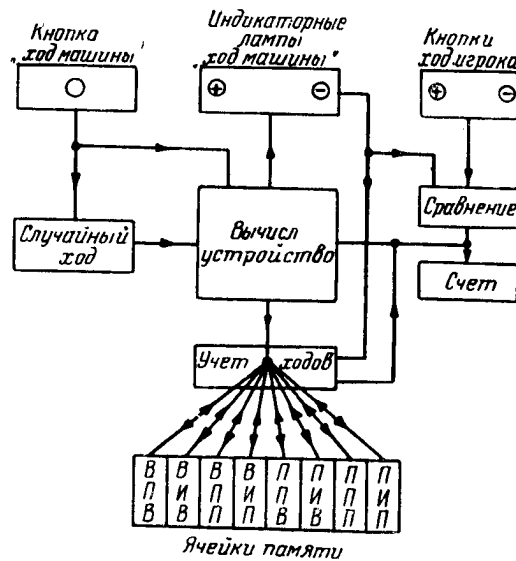


Рис. 10.3. Упрощенная схема машины, играющей в „чет и нечет“.

В ячейках памяти записаны последовательности из двух ходов. Например: „выиграла“ (В), „сделала прежний ход“ (И), „выиграла“ (В) или „проиграла“ (И) „изменила ход“ (И) и „снова проиграла“ (И).

Ним — конечная парная игра с нулевой суммой и полной информацией. Она принадлежит к играм определенного типа с собственной оценкой и чистыми оптимальными стратегиями. В этой игре при заданных начальных условиях всегда выигрывает один из игроков, если игра ведется без просчетов*.

* Рассмотрим вариант Нима с группами из трех, четырех и пяти фишек. Выигрывает всегда тот, кто делает первый ход, если этот ход правильный (взять две фишки из группы, состоящей из трех). Положение (1, 4, 5) после первого хода не дает противнику возможности выиграть. После любого его хода первый игрок может привести игру к одному из следующих положений: (0, 4, 4), (0, 0, 1), (1, 1, 1), (1, 2, 3), из которых проигрыш второго игрока очевиден. Любой первый ход, кроме указанного, ведет к проигрышу игрока, его сделавшего.

Игры такого типа могут быть решены, и механизм для осуществления ходов оптимальной стратегии реализован в виде электронной схемы. Составление такого рода схем не менее интересно, чем решение шахматных этюдов, и в настоящее время существует много любителей, посвящающих часы досуга этому занятию. Некоторые схемы оказываются предельно простыми*.

Обычно машины такого типа имеют индикаторные лампы, зажигание или гашение которых производится кнопками и системой реле. После очередного хода живого игрока, погасившего те лампы, которые соответствуют взятым фишкам, нажимается кнопка ХОД МАШИНЫ. На это воздействие релейная схема машины реагирует очередным ходом в пределах оптимальной стратегии.

Аналогично ведется игра в крестики и нолики и остальные игры с полной информацией. Можно сделать универсальную схему, которая коммутируется для ведения одной из нескольких игр такого типа. При составлении схем этих машин-игроков удобно пользоваться релейной алгеброй логики.

Подобным же образом в принципе можно было бы построить машину и для игры в шашки и шахматы. Однако большое число ходов в этих играх и огромное число стратегий делают решение этих игр чрезмерно трудоемким. Для решения шахматной игры (определения оптимальных стратегий и собственной оценки), даже с помощью современных быстродействующих электронных вычислительных машин, потребовалось бы много миллионов лет вычислительной работы. Однако задача ведения игры в шахматы или шашки машинной вызывает большой интерес, так как многие практически интересные игры еще более сложны, чем шахматы или шашки. В качестве подхода к решению задачи о ведении машинной сложных игр задача о ведении

* Для примера рассмотрим Ним с одной группой из девяти, тринадцати, семнадцати и, вообще $4n+1$ фишек ($n=1, 2, 3 \dots$), причем брать за один ход можно не меньше одной и не больше трех фишек; вынужденный взять последнюю фишку — проиграл. Проигрывает в игре без просчетов всегда игрок, делающий первый ход. На любой ход первого игрока второй отвечает так, чтобы общее число фишек взятых за оба первых (оба вторых) и т. д. хода, было равно четырем. Машина для ведения этой игры при очередном ходе просто выключает очередную группу из четырех индикаторных ламп, в том числе и те, которые уже потушены ходом первого игрока. Несмотря на предельную простоту игры, для новичка исход ее не очевиден, и машина, обыгрывающая его подряд много раз и с неизменным успехом, производит большое впечатление.

игр в шашки и шахматы весьма заманчива, поскольку правила этих игр четко установлены и игры носят наглядный обозримый характер.

Ведение игры в шахматы без живого игрока решается на универсальной электронной вычислительной машине с помощью специальной программы. Обстановка на доске вводится в машину с помощью цифрового кода, в котором машина работает. Допустимые правилами ходы для каждой фигуры превращаются в операции над числами, изображающими позиции фигур на доске. Машина может поочередно перебирать все возможные в данной позиции ходы, как свои, так и противника. Каждая новая позиция получает числовую оценку тем выше, чем лучше эта позиция с точки зрения тактики шахмат. В системе этих оценок в сущности заключаются все тактические принципы ведения игры, которые сообщает машине составитель программы. Перебирая возможные позиции на два-три хода вперед, машина вычисляет и запоминает все оценки этих позиций и в результате выбирает тот ход, который ведет к позиции с наилучшей оценкой. Вычисленный и выбранный машиной ход декодируется в обычные обозначения и печатается автоматически на бумажной ленте. Оценивается не только взятие фигур и пешек в соответствии с их достоинством (король имеет оценку, значительно превосходящую все другие фигуры), но и такие особенности расположения фигур и пешек, как сдвоенные пешки, проходная пешка, продвижение пешки вперед и т. д. Каждая из этих особенностей имеет свою «цену», отрицательную или положительную, отдельно для своих фигур и отдельно для фигур противника. Общая оценка позиции равна сумме частных оценок.

Оценки можно, разумеется, делать различными способами, и эти способы неравноценны.

Машина, играющая по такой программе, выполняет некоторые операции из числа тех, которые выполняет и человек-шахматист; при этом она фантазирует при выдумывании ходов в меру отведенных ей на то творческих возможностей. Увеличение этих возможностей требует расширения памяти машины с тем, чтобы анализ позиций мог распространяться на большее число ходов вперед. Потребность в емкости памяти с ростом числа анализируемых ходов возрастает чрезвычайно быстро.

Программы этого типа для игры в шахматы и шашки страдают отсутствием тактического замысла, общего плана

ведения игры, рассчитанного на большое число ходов. Этот недостаток не является неустранимым, и, возможно, существуют другие типы программ, не обладающие этим недостатком.

Машина, работающая по описанной программе, играет в шахматы вполне корректно, соблюдая правила игры, не делая грубых промахов и используя ошибки противника. Однако сильный шахматист может выиграть партию у такой машины, поскольку машина играет в силу спортсмена-шахматиста четвертой-третьей категории.

Особенно заметны недостатки программ такого типа в партиях «позиционных», в которых отсутствует возможность заметного изменения обстановки на протяжении нескольких ходов. Это иллюстрируется забавным эпизодом с одним из первых составителей программы для игры в шашки, о котором рассказывает К. Шэннон. После составления программы ее проверку было решено провести посредством игры машины с машиной. Программа игры и исходная позиция были введены в машину, и была нажата пусковая кнопка для первого хода. Машина работала в течение нескольких секунд и затем бодро напечатала: «Я сдаюсь!».

Шахматные задачи двух- и трехходовки очень хорошо решаются машиной, работающей по программе с оценками.

В области ведения машинами «нерешенных» игр, несомненно, еще многое будет сделано в ближайшее время.

Весьма сложные задачи выбора стратегии в военных операциях, в вопросах оптимального планирования промышленной, хозяйственной и административной деятельности в различных масштабах являются как раз тем типом задач, которые практически неразрешимы точно при решении их «вручную», без помощи машин. Применение машинного решения конкретных игр в этих областях позволит заменить приближенные «интуитивные» решения (которые иногда могут быть попросту неверными) точными, оптимальными решениями, что, учитывая масштабы, должно принести гораздо больше пользы, чем автоматизация любых отдельно взятых процессов производства.

Еще раз о неопределенных играх

Несколько иной характер имеет задача ведения машинной неопределенных игр типа описанной выше игры в «чет и нечет», т. е. игр, в которых отсутствуют оптимальные

чистые стратегии. Оказывается, в играх этого типа машина имеет более высокие способности, чем человек. Дело в том, что человек в силу врожденной и воспитанной опытом привычки действовать рационально, т. е. в соответствии с некоторым намеченным планом, почти не способен вести игру наугад, имитируя своим поведением чисто случайный процесс. А именно такой чисто случайный выбор ходов в игре «чет и нечет» и составляет оптимальную смешанную стратегию.

Одна из первых машин такого рода играла именно в «чет и нечет». Эта машина была рассчитана на игру против противника, применяющего определенные закономерные стратегии — последовательности ходов, и не применяла оптимальной смешанной стратегии. В состав машины входило около сотни реле. Человек, игравший против машины, делал выбор из двух возможностей («чет» или «нечет», «да» или «нет», «красное» или «зеленое») и сообщал вслух о своем выборе, но не вводил его в машину. Далее нажималась пусковая кнопка машины, машина самостоятельно (наугад) делала выбор из тех же двух возможностей и объявляла о своем выборе зажиганием одной из двух сигнальных ламп. Если выбор машины совпадал с выбором человека, машина выигрывала, в противном случае — проигрывала. Результат «партии» вводился в машину нажатием одной из двух предназначенных для этого кнопок. Далее игра продолжалась тем же порядком. Машина запоминала свои ходы и результаты каждой партии, а затем обнаруживала в серии запомненных данных закономерности, единственным источником которых была тенденция человека играть «разумно» и перехитрить машину. На основе обнаруженных закономерностей машина производила экстраполяцию и предугадывала ходы человека. Разумеется, это удавалось ей не всегда, но в некотором числе случаев удавалось.

Сначала игра протекала с переменным успехом, но затем машина постепенно начинала выигрывать все чаще и чаще, а проигрыш человека систематически возрастал.

Из общего числа 9795 партий, сыгранных машиной со многими партнерами, машина выиграла 5218 и проиграла 4577 партий. Правда, соотношение это не очень показательное, поскольку игры были большей частью короткими и продолжались лишь до тех пор, пока машина «училась» играть с данным партнером, прекращаясь, когда машина начинала выигрывать. Если для оценки предположить, что

полученное соотношение выигранных и проигранных партий явилось результатом случая, а не было обязано качествам машины, и подсчитать вероятность такого результата в предположении, что выигрыш и проигрыш машины в каждой партии равновероятны, то вероятность окажется ничтожно малой — порядка 10^{-10} . Таким образом, машина на практике доказала свою способность использовать отклонения от оптимальной смешанной стратегии, свойственные человеку. Здесь снова уместно напомнить, что незинное оформление этого эксперимента в виде детской игры «чет и нечет» не должно никого обманывать, ибо в принципе применение смешанных стратегий свойственно большому числу игр, ничего общего не имеющих с детскими развлечениями.

Интересно, что вокруг машины для игры в «чет и нечет», о которой мы говорили выше, возник спор о наилучших методах экстраполяции случайных процессов и об оптимальном числе партий, по которому должна производиться экстраполяция. В результате спора была построена вторая машина, работавшая по меньшему числу партий. Для их сравнения было сконструировано устройство, связывавшее обе машины и вводившее ходы одной в другую и наоборот. Обе машины были запущены в игру друг против друга. Необычный матч машин вызвал бурный интерес сотрудников лаборатории, где производился эксперимент, и заключение большого числа пари.

В результате соревнования выявилось заметное преимущество машины, работавшей по малому числу партий и таким образом, имевшей более быструю реакцию (ее автором был все тот же Клод Шэннон).

Описанные машины не применяют оптимальной стратегии; ее применение всегда приводило бы игру к ничьей. Поэтому существует стратегия противника, играющего против машины, которая обеспечивает ему выигрыш у машины. Однако эта стратегия не очевидна, что и приводит к проигрышу неподготовленного оппонента.

Итак, человек в соревновании с машиной в случае неопределенных игр оказывается более слабым. Это, однако, не является непоправимым. Мы уже говорили, что человек имеет свойство менять свои реакции в результате накопления опыта. В данном случае машина «от рождения» имеет способность вести неопределенные игры наилучшим образом, и человек, вступивший с ней в соревнование без предварительной тренировки, оказывается попросту непод-

готовленным. Он может улучшить свои результаты в игре с машиной, если сумеет с помощью тренировки развить в себе соответствующие способности. Правда, в данном случае задача довольно сложна, поскольку она состоит в выработке чисто случайного поведения, что противоречит устойчивым привычкам. Целый ряд фактов убеждает нас в том, что такое обучение, как и всякое другое, возможно. Об этом говорит опыт игры в покер — типичную игру неопределенного типа. Кроме того, некоторые эксперименты с обучением дают возможность подтвердить этот вывод.

Представим себе, что человеку (испытуемому) много раз повторно предлагают испытание одного и того же типа, например, ему предлагается на фоне хаотически расположенных пятен обнаружить рисунок или узор, который может присутствовать в предъявляемой картинке или же отсутствовать и быть либо ясно, либо неявно выраженным. Просматривая серию картинок, в которой четкость рисунка на скрывающем его фоне распределена в соответствии с некоторым вероятностным законом, испытуемый по своему усмотрению сортирует картинки на две группы: 1) группа, в которой есть рисунок, 2) группа, в которой нет рисунка. При этом он может совершать ошибки двух типов: а) относить картинку с рисунком во вторую группу и б) относить картинку без рисунка в первую группу. Итак, все его реакции (сортировка картинок) состоят из четырех действий: двух правильных и двух ошибочных. Обозначим их следующим образом:

I (+) и II (—) — правильные действия; III (+—) и IV (—+) — неправильные (первый знак в скобках соответствует наличию рисунка, второй — характеризует отнесение к соответствующей группе).

Представим себе, что действия испытуемого являются ходами в некоторой игре против неизвестного противника. Каждый правильный ход I или II приносит ему некоторый выигрыш a_1, a_2 , а ошибочный III и IV — проигрыш a_3, a_4 ($a_1, a_2 > 0$; $a_3, a_4 < 0$). Установим величины a_1, a_2, a_3, a_4 , сообщим их испытуемому и будем его поощрять и штрафовать в соответствии с установленными оценками. Это можно делать, либо выплачивая ему денежные суммы и делая вычеты из выплат в качестве штрафов, либо применяя иные меры воздействия типа армейских взысканий и поощрений и т. п.

Ясно, что величины a_1, a_2, a_3, a_4 должны быть подобраны в соответствии с распределением вероятностей появ-

ления рисунков в картинках так, чтобы испытуемый при хорошей работе получал от нее некоторый выигрыш, т. е. работал с положительной заинтересованностью. Не будем входить в детали подбора величин штрафов и поощрений, ибо это отвлекло бы нас от главного.

Теперь дадим испытуемому поработать некоторое время, чтобы он прошел период «обучения», а затем подсчитаем и зафиксируем число его правильных и ошибочных действий.

Затем повторим тот же эксперимент с тем же испытуемым (или лучше с той же группой испытуемых), но с другими оценками, и снова подсчитаем число правильных и ошибочных реакций.

Как показывают проведенные эксперименты такого типа, реакции испытуемого меняются при изменении оценок. Эксперимент производился с операторами, обнаруживавшими сигнал определенной формы, но произвольной величины и расположения, на фоне шумов на экране электронно-лучевой трубки индикатора.

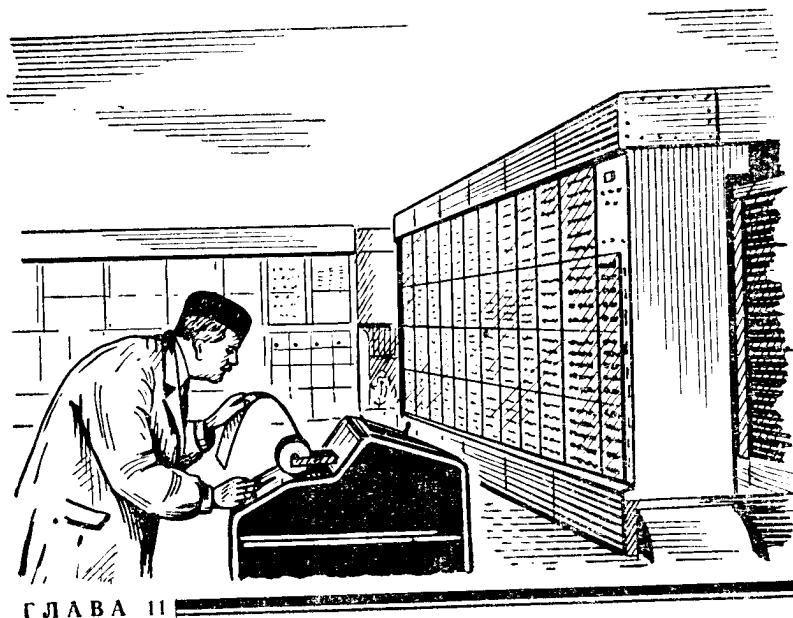
Увеличение штрафа a_1 за «ложную тревогу», т. е. за отнесение случая *нет* по ошибке в категорию *есть* (-+), ведет за собой уменьшение ошибок этого типа и увеличение ошибок типа (+-), т. е. отнесение случая *есть* ошибочно в категорию *нет*. Это означает понижение порога чувствительности оператора к сигналу. Понижение платы a_1 за правильное обнаружение сигнала (++) ведет к тому же результату.

Можно сформулировать результат эксперимента следующим образом. Оператор в результате обучения начинает работать так, чтобы обеспечить себе максимальный выигрыш в игре, вырабатывая для этого те реакции, которые ведут к этому результату, исходя из действующих оценок.

Самое интересное в этой серии экспериментов, пожалуй, то, что, вообще говоря, возможно решение обратной задачи. Статистически изучая реакции оператора, мы можем установить те оценки возможных результатов его работы, в соответствии с которыми он работает. В ряде случаев решение такой задачи представляет интерес. Мотивы, которые руководят деятельностью человека, решающего ту или иную задачу, иначе говоря, ведущего ту или иную игру, часто бывают многообразны и скрыты от непосредственного наблюдения. Изучая реакции человека в ходе решения задачи, т. е. в процессе ведения игры, можно сле-

дать определенные выводы о том, каковы же истинные цели, которые он преследует, и каковы те оценки, которые он придает тем или иным результатам.

Таким образом, статистика реакций дает ключ к скрытым психологическим мотивам человеческой деятельности. Такого рода анализ может быть произведен количественно и вполне строго на основании математической теории игр. Подобный анализ может количественно установить изменения критериев оценок при эмоциональных сдвигах в психике. Этими методами можно вести строго количественное изучение особенностей психологии человека, в частности, подсознательных мотивов, ускользающих от самонаблюдения испытуемого.



ГЛАВА II

РОБОТ, КОТОРЫЙ МОГ БЫ БЫТЬ УМНЕЕ СВОЕГО КОНСТРУКТОРА

В предыдущих главах мы обсуждали явления и понятия, связанные с управлением энергетическими системами посредством сигналов. Эти явления и эти понятия общи как для управляющих кибернетических систем, созданных человеком искусственно ради тех или иных применений в технике, так и для кибернетических систем естественного происхождения — живых организмов. Поскольку такая общность явлений для живой и неживой природы существует, необходимо проследить, насколько далеко эта общность простирается. Мы уже делали такую попытку в главе о работе мозга и сейчас продолжим ее, сделав несколько замечаний относительно возможностей изменять свою структуру в соответствии с внешними условиями и воздействиями внешней среды, которые имеют живые организмы, с одной стороны, и неживые механизмы, с другой.

Здесь еще раз уместно повторить следующее: мы далеки от того, чтобы непосредственно сравнивать ма-

шины и организмы и делать какие бы то ни было утверждения, хотя бы отдаленно приближающиеся к утверждению о том, что «мозг человека — это счетная машина» и т. п. Живой организм и любая современная машина представляют собой в целом совершенно различные вещи. И по составу, и по структуре, и по функциям, и по характеру протекающих в них физических процессов, и по условиям существования, и, наконец, по истории их возникновения и развития они совершенно несходны между собой. И тем не менее в принципе они представляют собой явления одного и того же порядка, так как и те и другие суть кибернетические системы. И именно поэтому допустимо и нужно сравнивать их свойства и возможности.

Машина всегда, идейно и материально, создается ее конструктором — человеком. Человек — это животное, умеющее создавать орудия и инструменты, в том числе и такие сложные, как кибернетические машины, помогающие ему выполнять не только физическую, но и интеллектуальную работу. Создавая машину, конструктор всегда имеет перед собой задачу, решение которой диктуется необходимостью, вызывающей внутренний эмоциональный стимул, именуемый желанием.

Решение стоящей перед конструктором задачи может быть достигнуто, как и решение всякой другой задачи, разными неравноценными путями и методами. Конструктор из известных ему методов выбирает тот, который он считает наилучшим. Этот выбор оптимального варианта, будь он правилен или ошибочен, предопределяет структуру машины и связанные с ней особенности функционирования. Другие, отвергнутые конструктором варианты со всеми их частичными достоинствами остаются в стороне, в замыслах, в голове конструктора. Реализуемый вариант воплощает в себе лишь часть знаний конструктора, он всегда беднее по своему содержанию, чем то, что мог бы создать конструктор в идеале, без учета материальных ограничений. В этом смысле духовное дитя конструктора — машина, вопреки известному правилу, не идет дальше своих родителей.

Создание устройства, выполняющего данную функцию, есть задача всегда более обширная, чем само выполнение этой функции.

Имея в виду именно такой подход к созданию машин-роботов, имеющих утилитарную ценность, нетрудно убедиться, что все операции, которые способны выполнять

роботы, заложены в них конструктором. Даже в тех случаях, когда робот по роду своей деятельности делает выборы и самостоятельно принимает решения, изменяет реакцию и производит самонастройку, эти решения принимаются им на основании критериев и правил, сообщенных ему конструктором и реализованных либо в виде структурных особенностей машины, либо в виде сигналов или команд, зафиксированных в памяти робота.

Конструктор всегда знает больше, чем робот. И хотя конструктор не успевает следить за работой робота во всех деталях, хотя некоторые элементарные или сложные операции робот выполняет, не сообщая об этом человеку, и «услужливые» роботы выполняют все больше чисто человеческих операций, беря все более и более сложные задачи из рук человека в свои,— все же человек выдумал робота и создал его, а не робот выдумал и создал человека.

Может быть, в завершение следовало бы добавить, как это делают многие: «...и никогда не создаст!» Но мы воздержимся от такого утверждения, несмотря на казалось бы крайнюю фантастичность противоположной возможности. Воздержимся для того, чтобы обсудить, действительно ли человек, конструктор кибернетической системы, является единственно возможным источником и причиной эволюции роботов и нет ли путей к созданию роботов, которые сами, без вложенного в них «идейного содержания», были бы способны под воздействием внешних влияний менять свою структуру и функции, развиваясь и эволюционируя «по пути к совершенству», как это свойственно живому миру. Факт саморазвития, эволюция без вмешательства творца, под действием законов природы не представляет собой для естествоиспытателя чего-то нового и невиданного.

Возможность саморазвития, «самоорганизации» кибернетических систем реально существует; в наши дни ведутся экспериментальные работы, в которых исследуется эта возможность. Идеи самоорганизующихся систем связаны с именем англичанина В. Р. Ашби, который одним из первых их высказал.

Самоорганизующиеся системы

Мы будем называть самоорганизующимися кибернетические системы, обладающие некоторыми описанными ниже свойствами. Их первоначальная структура представ-

ляет собой случайный набор элементов, связанных друг с другом случайным образом. Предоставленная самой себе система пребывает в первоначальном состоянии, и ее организация, т. е. связь между элементами, может меняться лишь случайным образом под воздействием внешних или внутренних случайных событий (шумов). Если на «вход» системы подаются некоторые сигналы, система реагирует, т. е. на ее выходе появляются новые сигналы, представляющие собой преобразование входных (в том числе и с использованием памяти); сигналы с выхода передаются во внешнюю среду и воздействуют на нее. Если в процессе такого взаимодействия с внешней средой, и только при наличии такого взаимодействия, структура кибернетической системы и ее функции меняются, приближаясь к некоторому устойчивому состоянию, то такую систему мы будем называть самоорганизующейся.

Из определения непосредственно следует, что возможность самоорганизации зависит не только от свойств самой системы, но и от характера воздействий внешней среды и вообще от характера взаимодействия системы со средой.

Представим себе, что некоторая кибернетическая система, робот, некоторым образом произвольно организованная, находится в постоянном взаимодействии с внешним миром. Пусть во внешнем мире произошло что-то, что изменило условия существования системы и оказало на нее воздействие, скажем, уменьшило поступление энергии, питающей систему, или внесло некоторые другие изменения (например, ослабление или усиление связей между всеми или некоторыми элементами системы), повлиявшие как на функционирование системы, так и на ее структуру. Система испытывает это влияние двояким образом: во-первых, уменьшение поступления энергии заставляет ее работать более вяло, если только она не обладает собственным запасом, накопленным в каком-либо аккумулялирующем устройстве, во-вторых, это событие, воздействуя на чувствительные элементы системы, превращается в сигнал и, пройдя после преобразования в системе на выход, вызывает некоторое ответное действие — реакцию. Эта реакция окажет, в свою очередь, воздействие на внешнюю среду, которое может быть произвольным, так как внутренняя организация кибернетической системы произвольна или случайна. В некоторых случаях реакция может быть «нейтральной» в том смысле, что она ничего не

изменит в дальнейшем ходе событий. В других случаях реакция системы может повлечь за собой изменение воздействия внешней среды на систему, и первоначальное воздействие либо усилится, либо ослабится, либо, наконец, появится некоторое новое воздействие на систему.

Если реакция системы такова, что она ослабляет первоначальное внешнее воздействие или полностью его снижает, то можно сказать, выражаясь языком физиологов, что эта система «уравновешивает воздействие среды» или «целесообразно реагирует», по крайней мере, в отношении данного воздействия. Во взаимодействии системы и среды мы имеем отрицательную обратную связь, которая стабилизирует существующее состояние. Повторные воздействия того же типа снова будут обрабатываться системой, которая, таким образом, будет «уходить» от воздействий внешней среды. Слово «уходить» можно понимать и буквально, т. е. в смысле пространственного перемещения в более благоприятную область.

Другой вид реакции (типа положительной обратной связи) окажет на внешнюю среду такое воздействие, которое усилит первоначальное воздействие со стороны среды на систему. В примере с уменьшением поступления энергии реакция системы приведет к еще большему снижению питания, а затем, по мере «самовозбуждения» этого процесса, — к полному выключению системы в энергетическом смысле. Реакция системы на внешнее воздействие в этом случае не стабилизирует положение, а «подхлестывает», ускоряет его изменение и ведет к непрерывному развитию процесса взаимодействия в одну сторону. Если первоначальное воздействие наносит повреждение системе, то в дальнейшем может наступить полное ее разрушение. Повторное воздействие застанет систему уже не в том виде, в каком оно заставляло ее в первый раз. После первого воздействия система «выключилась», или разрушилась, или приобрела совсем иную, отличную от первоначальной структуру.

Какую именно реакцию имеет система — типа отрицательной обратной связи, положительной обратной связи или нейтральную — зависит в конечном итоге от первоначальной организации системы, т. е. от случая, поскольку исходная организация системы по предположению случайна.

Системы, случайно оказавшиеся обладателями реакции типа отрицательной обратной связи, могут, приспособив-

ваясь к воздействиям среды, долгое время существовать без изменений. Системы с реакцией типа положительной обратной связи «выходят из игры» при первом же вредном воздействии. Таким образом, при прочих равных условиях время существования тех и других неодинаково.

Представим себе, что некоторый естественный случайный процесс создает одну за другой различные случайно организованные системы, т. е. некоторые собрания взаимосвязанных активных элементов. Допустим, что известны вероятности p_1 и p_2 появления систем первого и второго типов с реакциями типа отрицательной и положительной обратной связи, соответственно. Иначе говоря, в единицу времени создается всего N систем, из них в среднем $n_1 = p_1 N$ первого и $n_2 = p_2 N$ второго типа. Каждая из систем данного типа существует в среднем в течение времени T_1 и T_2 , соответственно. Начиная с некоторого момента будем наблюдать только те системы, которые появляются после начала наблюдения, игнорируя остальные. Их число будет нарастать со временем для каждого типа до тех пор, пока по истечении времени T_1 и T_2 не начнется их исчезновение. После этого количество их будет постоянным и равным $n_1 T_1$ и $n_2 T_2$. Поскольку срок существования систем первого типа гораздо больше того же срока для систем второго типа ($T_1 \gg T_2$), количество одновременно наблюдаемых систем с отрицательной обратной связью намного превысит количество систем с положительной связью, даже если вероятность появления систем второго типа (p_2) несколько больше, чем p_1 . И даже если вероятности p_1 и p_2 равны, т. е. системы обоих типов рождаются приблизительно в одинаковых количествах, все же долго существовать и потому преимущественно наблюдаться будут только устойчивые системы.

Это простое рассуждение приходит на помощь тогда, когда у естествоиспытателя появляется чувство удивления, вызванное «рациональным устройством» живых организмов. Действительно, все наблюдаемые живые существа обладают многочисленными и совершенными реакциями типа отрицательной обратной связи на вредные воздействия (и положительной на полезные). Для различных проявлений таких реакций существуют названия: тропизмы, рефлексy и т. д., о чем мы говорили ранее. Реакции эти часто бывают совершеннее реакций, придуманных человеком и реализованных им в роботах. Недаром много

веков существовала идея о «творце» всего живого, который создал все «разумно». Однако этим «творцом» является тот факт, о котором мы только что упоминали: на свет появляются и «рационально», и «нерационально» устроенные существа; выживают же и существуют только те, которые устроены действительно рационально; остальные очень скоро, порой почти мгновенно, исчезают.

Но вернемся от живой природы, от уже сформировавшихся и отобранных биологических видов, закрепленных наследственностью и лишь иногда подвергающихся случайным «экспериментальным» изменениям, к искусственным, случайно организованным системам со случайными реакциями. Как мы установили, одни из них, «ощутлив» опасное воздействие, от него «уходят», а другие, «почувствовав ожог, лезут в огонь» и в нем сгорают. Вот эти последние нас и должны заинтересовать.

Несомненно, что не все «нерационально» реагирующие системы должны обязательно подвергнуться полному физическому разрушению и уничтожению. По мере «разгития» процесса взаимодействия системы и среды в одном направлении и усиления внешнего воздействия это воздействие, поскольку оно усиливается, может вызвать частичную перестройку структуры системы и тем самым изменить ее реакцию. В некоторых случаях реакция может измениться в сторону «рационального» поведения, т. е. в сторону реакции типа отрицательной обратной связи. Разумеется, перестройка, или, лучше сказать, ломка системы под нарастающим внешним воздействием, происходит не обязательно в этом направлении, она происходит случайно. Однако, подобно песчинке в сите, которая рано или поздно проваливается через сетку, если только этому не препятствуют ее размеры, система, подвергаясь хаотическим реконструкциям, в конце концов может обрести устойчивое состояние, т. е. приобрести реакцию, выводящую ее из-под опасного, травмирующего воздействия. Таким образом, система может приобрести «рациональную» реакцию без какого-либо руководящего влияния разумного существа, под воздействием чисто случайных причин. Может быть, очень немногие из случайно организованных систем дойдут до конечного устойчивого состояния, но те из них, которым это удастся, останутся в устойчивом состоянии надолго.

Можно сказать, что система, прошедшая путь случайных преобразований под внешними воздействиями и по-

лучившая структуру, обеспечивающую реакцию «самозащиты», является самоорганизующейся системой.

Путь, по которому идет преобразование системы, имеет много общего, с одной стороны, с методом «проб и ошибок», т. е. с приобретением опыта живыми организмами без их намеренного обучения или дрессировки со стороны, а с другой стороны, несомненно, представляет собой пример обратной связи, в очень широком понимании этого явления, направленной на реорганизацию структуры системы.

Поскольку процесс протекает без вмешательства разумного существа и, быть может, без всякой связи с его существованием, самоорганизующаяся система не заимствует структурной организации и информации ниоткуда, кроме внешней среды. Структурные изменения системы под воздействием внешней среды, несомненно, представляют собой некоторый род памяти. Использование информации, запасенной в структурной памяти, требует определенного уровня сложности системы, т. е. достаточного количества активных чувствительных элементов и их связей. Перешагнув некоторый порог сложности, система приобретает способность самостоятельно накапливать поступающую в нее информацию и так или иначе ее использовать.

Заметим попутно, что из числа случайно организованных кибернетических систем, подвергающихся хаотическим изменениям и естественному отбору, выживают только те, которые случайно научились употреблять внешние влияния в свою пользу, т. е. приобрели положительные тропизмы к воздействиям, поддерживающим их существование, и отрицательные — к воздействиям, которые угрожают им разрушением. Таким образом, естественно, зарождается то свойство, которое мы в гл. 9 называли «эгоизмом» в широком понимании и которое были вынуждены признать несвойственным современным роботам.

Если допустить, что рациональная организация, сложившаяся в процессе самоорганизации, закрепляется передачей ее по наследству, то мы получим грубую модель биологических видов.

Осуществление самоорганизующихся систем

Следует сказать несколько слов о первых попытках экспериментальной проверки идей, связанных с самоорганизующимися системами.

Самым первым экспериментом было создание д-ром Ашби устройства, названного им гомеостатом. Гомеостат состоял из четырех электромагнитов, сердечники которых, перемещаясь, передвигали ползунки реостатов. Питание электромагнитов включалось через эти реостаты. Отклонение сердечника в крайнее положение вызывало в схеме

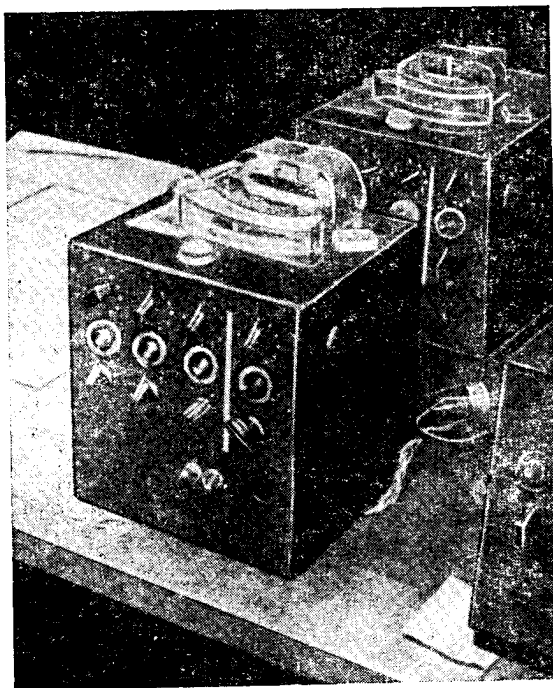


Рис. 11.1. Внешний вид одного из четырех блоков гомеостата Ашби.

случайное переключение соединений реостатов и обмоток электромагнитов. Положения всех четырех сердечников были, таким образом, взаимосвязанными, так как ток в каждом соленоиде зависел от положений всех четырех реостатов, а положение каждого реостата зависело от тока в соответствующем соленоиде.

Включение питания на схему гомеостата Ашби вызвало движение всех сердечников и перемещение ползунков всех реостатов. При этом могло быть два случая: либо

после некоторого переходного процесса все сердечники останавливались в некотором промежуточном устойчивом состоянии и движение прекращалось, либо система не находила устойчивого состояния и один из сердечников выходил за пределы нормального отклонения, доходя до упора. В последнем случае в схеме происходило случайное переключение, и поиски равновесия возобновлялись.



Рис. 11.2. Д-р В. Р. Ашби со своим гомеостатом.

В конце концов, после нескольких переключений, осуществляющих случайные изменения схемы соединений гомеостата, система самостоятельно доходила до такого состояния, при котором оказывалось возможным равновесие, и затем находила это равновесие.

Различные манипуляции с гомеостатом: перестановка упоров, изменение связей, даже частичные поломки — не нарушали этого свойства гомеостата.

Еще более поучительным является другой эксперимент. Исходным материалом в этом эксперименте являлся набор из некоторого количества (от 8 до 64) нелинейных элементов. Каждый элемент мог приходиться в состояние

возбуждения или быть в покое. Находясь в состоянии возбуждения, элемент мог передавать это состояние или его долю другим элементам, с ним связанным. Каждый элемент мог также воспринимать возбуждение от одного или нескольких элементов. Состояние возбуждения каждого элемента затухало по экспоненциальному закону, возник-

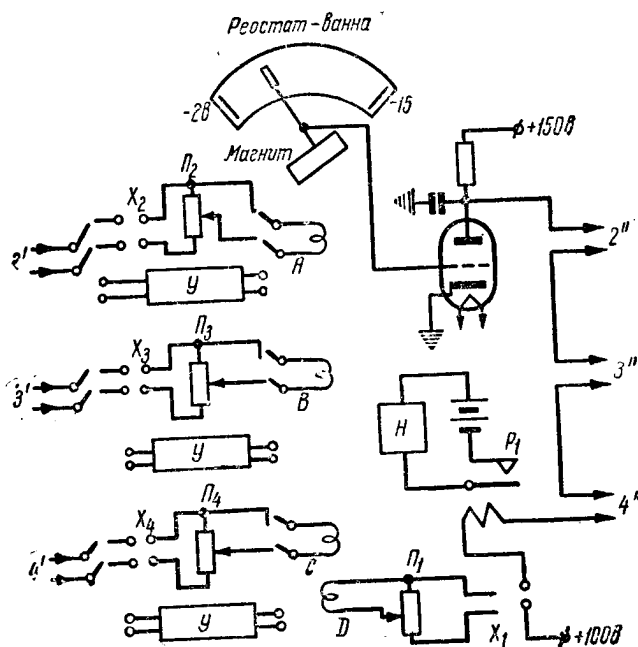


Рис. 11.3. Схема одного из четырех блоков гомеостата Ашби.
 А, В, С, D — обмотки соленоида, каждая из которых питается от одного из четырех блоков гомеостата. $\Pi_1, \Pi_2, \Pi_3, \Pi_4$; X_1, X_2, X_3, X_4 — потенциометры и переключатели, устанавливаемые вручную. У — автоматический переключатель, дающий случайное изменение параметров схемы при каждом переключении и срабатывающий при включении узла Н от реле P_1 в случае увеличения тока в цепи соленоидов выше порога. $2'', 3'', 4''$; $2''', 3''', 4'''$ — провода, соединяющие блок с остальными тремя блоками гомеостата. Остальные блоки имеют такие же схемы.

нув одновременно с импульсом, подаваемым извне. Для того чтобы элемент, находящийся в покое, мог быть возбужден, уровень подаваемого на него извне или от соседей возбуждения должен был превышать некоторый порог. Величина порога весьма велика в момент возбуждения (только что возбужденный элемент не может быть еще раз возбужден) и спадает с течением времени, стремясь

к некоторому, вообще говоря, переменному или регулируемому минимальному значению.

Все элементы связаны друг с другом. Эти связи в начале эксперимента устанавливались по жребию с помощью таблицы случайных чисел, имитировавшей случайный процесс, который определял первоначальную органи-

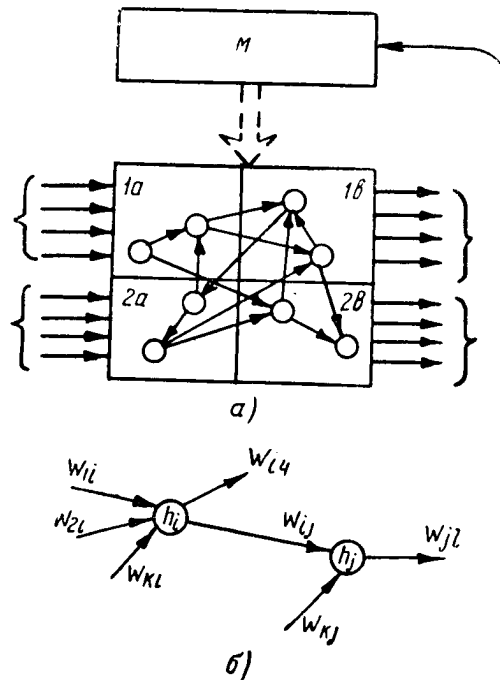


Рис. 11.4. Упрощенная схема модели самоорганизующейся системы (по Фарли и Кларку).

a) общая схема: *1a, 2a* — две входные группы элементов, *1b, 2b* — две выходные группы элементов. *M* — модификатор — устройство, вносящее изменения в структуру системы; *b)* два элемента системы и их связи.

зацию системы. Во-первых, устанавливалось наличие связи между каждой парой элементов. Для этого бросался жребий — бралось наугад число, расположенное между нулем и единицей; если оно превышало некоторую величину $k < 1$, связь для данной пары не устанавливалась, если не превышало, то устанавливалась (величина k называлась постоянной связи системы; чем больше k , тем больше связей устанавливалось). Для каждой установ-

ленной связи между элементами с номерами i и j выбирался также по жребию «вес» этой связи, имевший 16 различных значений и определявший, какая доля уровня возбуждения по этой связи передается.

Все связанные таким образом по жребию элементы разбивались на четыре группы: две «входные» ($1a$ и $2a$) и две «выходные» ($1b$ и $2b$). На элементы входных групп в определенном порядке подавалось периодическое возбуждение. Со всех выходных элементов снималось их возбуждение на выход системы. Возбуждение извне на входные элементы подавалось периодически, причем за каждую серию циклов возбуждение подавалось на все элементы только одной входной группы, а на элементы другой не подавалось. Затем две входные группы менялись ролями. Таким образом, хаотически организованная система подвергалась систематическому воздействию извне.

„Задачей“ системы было проводить возбуждение от первой входной группы только к первой выходной ($1a \rightarrow 1b$) и от второй входной только ко второй выходной ($2a \rightarrow 2b$). Качество выполнения этой задачи оценивалось числом N_t . После каждого такта работы системы подсчитывалось число возбужденных элементов выходных групп $1b$ и $2b$ (n_1 и n_2) и их разность (т. е. превышение числа „правильно“ возбужденных над числом „неправильно“ возбужденных элементов в группе, куда сигналы не должны проходить). Эта величина, накапливаясь от такта к такту, образует число N_t . Это число в начале решения задачи может быть установлено произвольным по желанию экспериментатора. Когда число N_t в процессе работы системы меняет знак, пара входных групп меняется местами, т. е. возбуждение начинает подаваться не на группу $1a$, а на группу $2a$.

Если бы система работала идеально, то при подаче возбуждения на группу $1a$ все элементы группы $1b$ были бы возбуждены, а все элементы группы $2b$ не возбуждены. Разность $n_1 - n_2$ имела бы максимальное значение, число N_t быстро изменяясь, достигло бы нулевого значения и изменило знак. Входные группы попарно поменялись бы местами, число N_t снова изменило знак и т. д. Затем число N_t поддерживалось бы непрерывно вблизи нуля. Если бы при этом искусственно изменить число N_t от нуля в любую сторону, то оно с максимальной скоростью снова вернулось бы к нулю.

Система первоначально весьма далека от идеала и, работая описанным способом, она будет медленно менять число N_t , вообще говоря, в произвольную сторону.

Теперь в систему вводится дополнительное усложнение. Специальное устройство вносит изменение во внутреннюю организацию системы. Его влияние на систему определяется направлением изменения числа N_t . Если N_t меняется в правильном направлении, т. е. уменьшается по абсолютной величине, то в системе подкрепляются все фактически действующие связи. Если N_t меняется в неправильном направлении, действующие связи ослабляются. Усиление и ослаб-

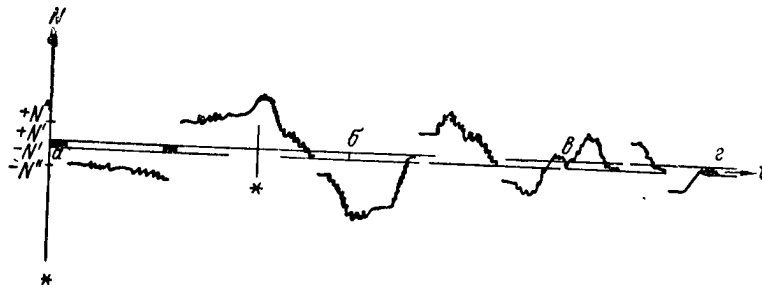


Рис. 11.5. График работы самоорганизующейся системы при последовательных циклах работы (по Фарли и Кларку). Число N оценивает качество работы в зависимости от времени. В начале каждой пробы N искусственно отводится от нулевого значения. Скорость, с которой N стремится к нулю, возрастает по мере хода опыта и самоорганизации системы. Звездочкой отмечен момент включения модификатора M .

ление связей производится изменением «веса» этих связей. Кроме того, в систему вводится некоторый «шум» — случайные хаотические изменения структуры, независимые от результата работы системы, а также регулировка уровня максимального значения порога возбуждения.

Введение в действие механизма, изменяющего структуру, придает системе новые свойства. По мере хода работы в процессе повторного решения задач показатель качества работы изменяется в лучшую сторону. Число N_t на первых порах слегка удаляется от нуля, но затем почти неизменно стремится к нулю. Скорость возвращения к нулю возрастает по мере организации системы под действием «тренировки».

На рис. 11.5 приведены графики изменения N_t во времени. После сведения N_t к нулю в ходе работы системы оно

искусственно скачком изменяется в ту или другую сторону. Улучшение качества функционирования системы, выражающееся в нарастании скорости, с которой N_t стремится к нулю, достигается за счет изменения связей в системе. Таблица весов связей между элементами от начала эксперимента к концу существенно меняется.

Повторное проведение эксперимента с различными случайно выбранными начальными связями внутри системы давало одинаковые результаты: каждый раз по мере использования поступающей извне информации система самостоятельно изменяла свою внутреннюю структуру так, что качество ее функционирования улучшалось. Из 64 опытов лишь 3 или 4 не привели к «самоорганизации» вследствие того, что в исходной системе случайно оказались отсутствующими важные связи.

Эксперимент проводился посредством моделирования описанной системы на быстродействующей электронной вычислительной машине со специальной программой. Выходная величина N_t автоматически записывалась с помощью осциллографа.

Интересно также то, что по прошествии периода «самоорганизации» система оказывалась довольно устойчивой по отношению к повреждениям. Искусственное удаление до 10% всех связей не наносило еще заметного ущерба функционированию системы.

Описанный эксперимент с «самоорганизующейся» системой весьма элементарен по характеру решаемой задачи. Результаты эксперимента, несомненно положительные, не позволяют еще усмотреть непосредственно какие-либо полезные технические применения продемонстрированных принципов*. Тем не менее уже эти простые эксперименты позволяют сделать важные принципиальные обобщения.

Продолжение исследований в этой области, несомненно, должно дать не только результаты большой принципиальной важности, но и ценные утилитарные возможности. Эти работы ведутся, и результаты их, возможно, явятся основой целого раздела кибернетики — изучение свойств «самоорганизующихся» систем, — быть может, самого интересного ее раздела.

* Однако, несомненно, что эта работа имела в виду практические применения. Об этом говорит хотя бы то, что финансирование работы производилось военным министерством США.

Интересным является тот факт, что явление «самоорганизации» воспроизводится при малом числе элементов (восемь — в описанном выше эксперименте). По-видимому, уже такое малое число элементов позволяет создать достаточную избыточность структуры. Избыточные элементы являются принципиально необходимыми в самоорганизующихся системах, которые обязательно отклоняются от оптимальных, наиболее экономно построенных структур, выполняющих ту же функцию.

Трудно предсказать, в каком именно воплощении могут быть и будут практически использованы возможности, представляемые самоорганизующимися системами. Скорее всего, это будут наиболее гибкие, электронные элементы. Однако не исключена возможность, что, кроме средств электроники, могут оказаться применимыми и другие, физико-химические, а быть может, и биологические средства.

Понятие самоорганизации случайно сложившейся структуры представляет собой не только принцип, обещающий ценные применения, но и средство понять многие процессы, происходящие в живой природе. Модель самоорганизующейся системы, о которой мы говорили выше, создавалась в явном расчете на имитацию некоторых черт механизма нервной деятельности (образование нервных структурных связей, ответственных за возникающую в процессе обучения новую функцию).

Возможно, что структуры, содержащиеся в органических пищевых веществах и поглощаемые в процессе питания живыми тканями, являются источником информации для процессов самоорганизации вещества живых тканей в какой-то мере подобно тому, как информация, подаваемая на вход самоорганизующейся системы, служит исходным звеном в процессе организации структуры. Известно, что живые существа при питании нуждаются не только в энергии, скрытой в пище, но и в физической энтропии (точнее, в отрицательной энтропии, т. е. в структурной организации поглощаемых питательных веществ)*.

Нам представляется чрезвычайно плодотворным применение понятий информации и самоорганизации к элементарным жизненным актам живой материи.

Даже столь краткое упоминание о самоорганизующихся системах, которое мы сделали в этой главе, дает возмож-

* См. Э. Шредингер. «Что такое жизнь с точки зрения физики?». Изд. иностранной литературы, 1946.

ность ясно себе представить, какое огромное поле для приложения совместных усилий ученых различных специальностей представляет изучение процессов управления в живых организмах.

Жизнедеятельность живых организмов от простейших до высших и процессы управления, непрерывно протекающие в ходе этой жизнедеятельности и составляющие сущность «уравновешивания со средой», включают в себя не только непосредственное использование информации, поступающей извне, и ее автоматическую обработку, но и гораздо более сложные, не однозначные и не прямые пути использования информации, одним из которых является самоорганизация типа «проб и ошибок».

Избыточность организации

При анализе живых структур, выполняющих ту или иную функцию, всегда обращает на себя внимание избыточность состава их по сравнению с соответствующими оптимальными, наиболее экономичными системами технического характера.

Такая избыточность живых структур возникла, по-видимому, не случайно. Во-первых, системы, возникающие в порядке индивидуального или видового развития, если только они возникают путем «самоорганизации» за счет поступающей извне информации в описанном выше смысле, всегда по необходимости избыточны. Во-вторых, избыточность конструкции — это то свойство, которое в конечном итоге обеспечивает долговечность системы при бесперебойном функционировании.

Здесь уместно было бы вспомнить об известном факте «двойного обеспечения» функций в живых организмах. Передача информации в живом организме почти всегда «дублируется» различными каналами. Именно поэтому нарушение связи и управления в каком-либо звене чаще всего ведет только к временному нарушению функции, которая затем восстанавливается посредством включения в работу обходных путей и «запасных» регулировок. Примером тому могут служить результаты опытов Лешли с ориентировкой крыс в лабиринте, о которых мы упоминали выше. Удаление у крыс значительной части мозгового вещества не приводило к полному нарушению способности животного к ориентировке.

В современной технике весьма остро стоит вопрос о долговечности кибернетических систем. Сложные систе-

мы, построенные из малонадежных элементов, имеющих небольшой срок службы, оказываются все менее и менее надежными по мере того, как с их усложнением возрастает число составляющих их элементов. Это происходит потому, что системы технической кибернетики создаются с максимальной экономией деталей и выход из строя хотя бы одной детали выводит обычно из строя всю систему. Таким образом, срок службы до ремонта системы равен сроку службы наименее надежного элемента системы. Повышение средней надежности деталей еще не решает вопроса, так как разброс по сроку службы может свести на нет достижения. Проблема обеспечения высокой надежности сложной системы, построенной из большого числа ненадежных элементов, стоит весьма остро и представляет большой прикладной интерес.

Ответ на этот вопрос дается нам живыми кибернетическими системами. И этот ответ — избыточность конструкции. В живых организмах иногда удаление некоторой части живого вещества, даже в условиях отсутствия регенерации, — как это имеет место, например, для нервных структур — дает лишь временное ухудшение функции, которая затем восстанавливается, вырабатываясь заново в остатках тканей. Конечно, дело обстоит не так, если удалена полностью некоторая структура, без которой данная функция не может быть осуществлена.

Увеличение избыточности конструкции в некотором, не слишком изящном варианте уже используется в технике. В особенно ответственных случаях, например, при работе автоматического оборудования без оператора и без возможности ремонта (автоматические радиометеостанции), приходится, попросту, дублировать целиком весь радиопередатчик. Один из двух передатчиков непрерывно эксплуатируется, другой — находится наготове. Выход из строя первого автоматически включает в работу второй.

Более выгодным явилось бы дублирование отдельных блоков и узлов передатчика. В случае дублирования передатчика целиком выход из строя второго передатчика (после того, как первый уже вышел, т. е. случай безнадёжного отказа системы) может произойти вследствие поломки совсем другой детали в другом блоке, чем при аварии первого. Чаще всего так и бывает, если выход из строя деталей случаен. Но в этом случае в системе с подменой не всего передатчика, а блоков или узлов, работа не была бы нарушена, так как оба поврежденных узла

были бы заменены полноценными. Таким образом, чем «мельче» блоки и узлы, для которых имеются запасные, тем надежнее работает система, состоящая из данного числа комплектов аппаратуры. Правда, в этом случае аппаратура дополнительно усложняется за счет возрастания числа коммутирующих аварийных устройств.

Вероятно, подробное знакомство с путями использования избыточности в живой природе могло бы дать много поучительных и полезных идей для конструкторов систем технической кибернетики.

Еще одним интересным вопросом, касающимся «самостоятельного» существования роботов без идейной помощи со стороны конструктора, является вопрос о приспособленности данной структуры, достаточно сложной и избыточной, к выполнению не одной, а нескольких различных функций. Всякая структура с избыточностью имеет возможность использовать избыточные элементы для формирования своего поведения. Поэтому часто искусственная кибернетическая система, построенная для выполнения одной функции, оказывается способной производить и другие (например, логические операции наряду с арифметическими и т. д.). Некоторые конструкторы высказывают «жалобы» на то, что не успеет новая конструкция вступить в строй для выполнения своей задачи, как уже оказывается, что она может и должна использоваться для совершенно иных применений, на которые она не рассчитывалась.

Проблема приспособляемости кибернетических систем к новым условиям в известной мере аналогична проблемам самоорганизации и выживания. Несомненно, что система, способная в известных внешних условиях к самоорганизации, способна также и к реорганизации. Проводя последовательно приспособление системы к различным условиям, можно усложнять и «наращивать» функции системы в пределах, определяемых количеством элементов. Идя по этому пути, мы могли бы получить некоторую модель эволюции организованных систем, их усложнения как структурного, так и функционального.

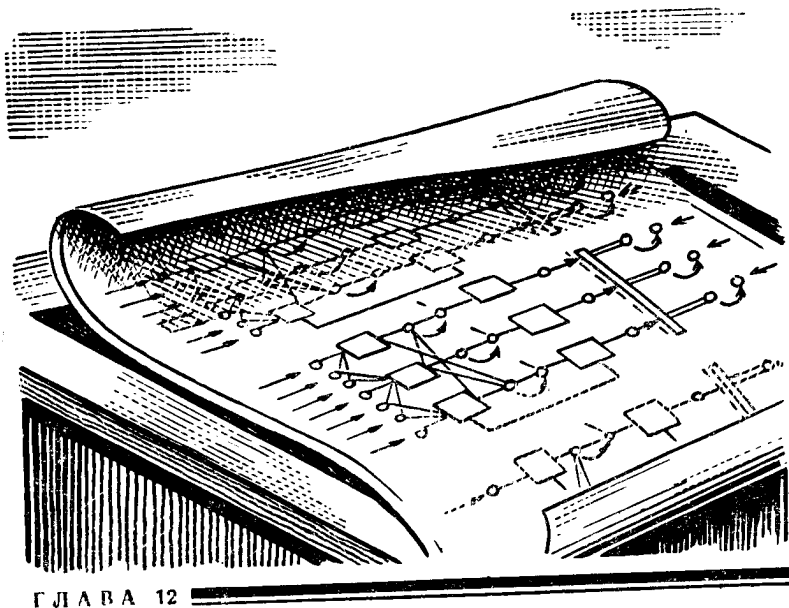
Сейчас очень трудно предсказать конкретные перспективы таких исследований, однако несомненно, что последовательное повторение процесса самоорганизации в отношении все более и более сложных систем может привести к системе, как угодно сложной с как угодно слож-

ными функциями, если только не ставить ограничений со стороны количества элементов, входящих в систему.

Учитывая, что процесс самоорганизации происходит без участия разумного творца, приходится признать, что на вопрос: «Может ли робот стать умнее своего конструктора?» — мы не можем с полным основанием ответить отрицательно. Такая возможность, хотя бы в принципе, существует, но пути ее реализации еще ждут дальнейших исследований*.

* Пока печаталась эта книга, появились работы, продолжающие исследования самоорганизующихся систем. Интересные результаты получены в Гарвардском университете. На машине ИБМ-704 моделировался творческий процесс поисков методов решения новой задачи. В машину вводилось много различных программ, в том числе и бессмысленных, и ставилась задача. Машина решала задачу наугад, чаще всего неправильно. Результат решения оценивался и на основе оценки изменялся метод решения. Вначале правильные результаты получались лишь случайно, приблизительно один раз из 400 попыток. После нескольких сотен тысяч попыток у машины «вдруг» появилось самостоятельное суждение о правильном методе решения, и далее она придерживалась всегда этого метода, несколько изменяя его самостоятельно, если изменялась задача.

Общая схема этого эксперимента напоминает описанные выше опыты Фарли и Кларка, но в данном случае условия опыта ближе к практически интересным явлениям. Задача, ставившаяся машине, состояла в обработке 14-значного числа посредством 63 математических операций. Авторы работы считают, что легче построить машину, способную выработать собственное суждение по сложному вопросу, чем точно формализовать алгоритм решения этого вопроса. (Нью-Йорк Таймс, 4/1, 1958). (Примечание при корректуре).



ГЛАВА 12

БОЛЬШОЙ РОБОТ

Что же дальше?

Человек приобрел отличный от прочих животных облик и образ жизни с тех пор, как стал способен применять орудия и инструменты. По мере накопления человеком опыта использования различных орудий производства, передвижения и т. д., эти орудия совершенствовались, усложнялись, приносили все больше пользы и давали, при правильном их использовании, все большую производительность труда. В наше время человек, которому предстоит, скажем, проложить дорогу в какой-либо местности, уже не приступает к этой работе сразу, хватаясь за топор и лопату. Сначала он изготавливает машины-орудия и машины-двигатели, с помощью которых затем выполняет работу по устройству дороги с несравненно меньшими затратами времени и усилий, чем если бы он прокладывал дорогу голыми руками или с помощью простейших орудий. В наш век механизации трудоемких работ иное и представить себе невозможно. Все больше машин приходит в промышленность, сельское хозяйство, транспорт — всюду, где работает человек.

Очень редко человеческая рука непосредственно прикасается к сырью для его обработки. Руки мастера всегда либо держат инструмент, либо, чаще всего, лежат на рукоятках управления машиной. Однако работает в процессе производства не инструмент и не машина, а человек с помощью инструмента и машины. Машина «удлиняет» его руки, делает возможными миллионы различных операций обработки, которые никогда нельзя было бы совершить ни пальцами, ни ногтями, ни зубами, делает руки человека настолько сильными, что ему не приходится прилагать собственной физической силы. Те производственные операции, которые сегодня еще требуют физической силы рабочего, мы называем «отсталыми», немеханизированными, имея в виду их скорое исчезновение.

Работа человека в процессе производства все больше и больше становится работой управления армией орудий, и руки человека становятся из непосредственных исполнителей лишь проводниками сигналов-команд, передаваемых от человека к исполнительным механизмам.

Физический труд человека стал непроизводительным, «отсталым», рядом с механическими машинами огромной мощности и сверхтвердыми «когтями» и «зубами» машин, рядом с тончайшей технологией сложных производств, немислимой при выполнении голыми руками человека. Человек в производстве ценится сегодня не как силач-грузчик или молотобоец, способный работать «за двоих», а как умный мастер, вдумчивый администратор, имеющий в запасе много знаний и опыта и способный вести армию машин к победе кратчайшим путем и с наименьшими потерями. Труд управления производством велик и сложен. И труд этот становится все сложнее и обширнее, а главное, ответственнее, по мере роста общественного производства, роста тех мощностей и разнообразия средств, которые принимают участие в процессе производства. Встает вопрос о механизации управления производством.

И как в те времена, когда человек впервые привязал камень к палке, изготовив первый инструмент, чтобы выполнять физическую работу, в наши дни человек тоже изготовил и применил орудия, облегчающие его труд управления производством. На наших глазах все чаще в области интеллектуального труда происходит то же самое, что имеет место для труда физического. Вместо того, чтобы своими руками (лучше сказать теперь «своей головой») решать задачу, переписывая множество цифр и про-

изводя в уме тысячи вычислений, человек изготавливает машину, которая быстрее, надежнее и лучше его выполняет эту работу. Все больше так называемых чисто человеческих задач решают информационные кибернетические машины, и решаемые ими задачи становятся все объемистее и сложнее. Недаром Н. Винер говорит о «второй промышленной революции», совершающейся в наши дни и состоящей в механизации интеллектуальной работы наравне с работой физической.

Если бы пещерного человека, изготавливавшего каменные топоры, спросили, какие топоры человек будет применять в будущем и как он будет ими пользоваться, вряд ли бы даже самый искусный мастер топорных дел того времени сумел рассказать о паровом молоте, гидравлическом прессе и других машинах-орудиях нашего времени. И это не его вина. В его распоряжении было недостаточно информации для того, чтобы сделать эти столь далеко идущие выводы.

Мы в наши дни тоже можем спросить себя: «Что дадут нам и нашим потомкам кибернетические машины?» И надо признаться, что ответить на этот вопрос мы сможем вряд ли лучше, чем пещерный мастер. Мы еще слишком мало знаем возможности машин, мы еще мало их применяли. По мере накопления опыта и знаний придут догадки и находки науки, и подобно тому, как человек шел от камня и бронзы к стали и сплавам, от огня к электричеству и атомной энергии, он пройдет новый путь от современных громоздких и дорогих кибернетических машин к чему-то новому, на фоне чего сегодняшние машины будут заслуженно казаться пещерной техникой.

Какие элементы придут на смену современным радиолампам, транзисторам, магнитным и диэлектрическим ячейкам, проводам и контактам? Об этом можно только мечтать и фантазировать.

Быть может, это будут волноводные устройства, работающие с волнами длиной в сантиметр, миллиметр, доли миллиметра или со световыми волнами. Быть может, это будут живые клетки, которые биологи-кибернетики научатся выращивать в питательной среде по заданной схеме со специализированными функциями. Может быть, это будут замороженные в криостате почти до абсолютного нуля квантово-вырожденные сложные кристаллы с непериодической структурой (сложные и сложнейшие молекулы), способные претерпевать изомерные переходы в соот-

ветствии с поступающими сигналами. Может быть, это будет что-то еще, чему мы не знаем названия, подобно тому, как пещерный человек не знал названий марок легированных сталей...

У нас есть только один путь поскорее узнать все это, и этот путь — неустанно и упорно трудиться над решением сегодняшних проблем кибернетики для того, чтобы быстрее передать машинам сначала свои будничные и простые задачи, а затем и более сложные, чтобы полнее использовать помощь машин для сбора и обработки информации, чтобы поскорее научить машины человеческим делам и научиться у машин всему, что необходимо человеку для дальнейшего движения вперед.

Человек, несомненно, многому учится у машин и не только потому, что он использует результаты расчетов, напечатанные на выходной ленте автоматического решающего устройства. Оператор-программист, формулирующий задачу для машины, не только обладает специальными знаниями. Выражаясь образно, машина воспитывает его мыслить иначе, строже, не допускать произвольных толкований и необоснованных решений. Передавая машине ту или иную задачу, человек обязан строго сформулировать, что должно быть сделано, как сделано, как сделать это наилучшим образом, а стало быть, каким критерием надо пользоваться для определения «наилучшего» и почему именно этим критерием, а не другим. Невозможно записать в программу машины действия, ход которых непонятен программисту или которые нечетко сформулированы. Говоря коротко, работа с машиной воспитывает в человеке точность и строгость мышления и умение критически оценивать свои собственные решения и действия. Против такого воспитания вряд ли можно что-либо возразить.

Отдав себе отчет в своих действиях с достаточной полнотой и четкостью, поняв, как именно человек производит ту или иную работу, сформулировав правила, по которым эта работа производится, можно эту работу запрограммировать, т. е. записать кодом команд машины или реализовать в решающем устройстве непрерывного действия и передать полностью машине. Это особенно выгодно в тех случаях, когда аналогичная задача решается повторно, много раз.

Вполне разумно ставить задачи механизации с помощью кибернетических машин информационных поисков (автоматический библиотекарь и библиограф); диспет-

черизации; административно-хозяйственного управления промышленными предприятиями и планирования производства в разных масштабах; постановки диагноза заболевания по данным объективного обследования, анамнеза и жалоб больного; инженерно-технического руководства производством; ведения сложных технологических процессов; оптимальной дозировки лекарственных веществ и распределения доз во времени при медикаментозном лечении; применения исправительных мер в отношении преступников на основании данных следствия и объективных характеристик подсудимого, а также ведения следствия по уголовным делам.

Все эти задачи требуют знания критериев оптимизации, правил, по которым вырабатывается решение, и объективного взвешивания по вероятности различных гипотез в случае неполной информации. Польза от попыток механизации решения этих задач заключается не только в разгрузке государственного и технического аппарата и повышении производительности его труда, но еще и в том, что попытка сформулировать точные правила и критерии качества работы инженера, врача, судьи, администратора принесет огромную пользу соответствующим отраслям деятельности, позволит четко поставить вопросы, устранить произвольные решения и выявить существующие неясности. Машина не терпит даже тени неясности и произвола.

Большое число задач связано с военным делом. Их решение должно механизироваться, ибо скорость реакции в играх типа военных действий почти всегда имеет решающее значение, а объем задачи, как правило, не позволяет человеку в ограниченное время решить ее строго, без произвола. Задачи о решении игр и выборе оптимальной стратегии характерны не только для военного дела, но и для ряда мирных применений.

Многие из задач, упомянутых выше, не только не решены, но еще и не поставлены достаточно точно и строго. Производится лишь разведка путей подхода к их постановке и решению.

Для решения этих задач, несомненно, потребуется расширение математического аппарата. Машины вызовут к жизни новые разделы математической науки не только потому, что они позволяют производить вычисления, недоступные ранее по трудоемкости, и исследовать вопросы, недоступные для исследования сегодня из-за огромного

количества необходимой вычислительной работы, но еще и потому, что в ходе применения машин выявится много неожиданных возможностей, для использования которых будет необходим новый формальный аппарат.

Возможности кибернетических систем

Пытаясь заглянуть в будущее, мы должны, хотя бы на основе наших ограниченных знаний, оценить возможности произвольно сложной машины — большого робота, отвлекаясь от его конструктивных и экономических характеристик. Мы не будем учитывать экономические характеристики не потому, что этот вопрос не важен, а лишь на том основании, что возможности реализации могут нас интересовать только в том случае, когда мы знаем, что в принципе конструируемая машина будет в состоянии выполнять желаемую функцию. Естественно, что при этом большая часть высказываний явится плодом фантазии.

Человек, занимаясь судопроизводством или администрацией, планированием хозяйственных или военных операций, разработкой конструкций или технологии, или же научным исследованием, кроме практической или экспериментальной деятельности, непременно должен мыслить или, лучше сказать, размышлять. Для того чтобы решать те же задачи, робот, по-видимому, должен быть способен производить некоторые операции, равноценные размышлению, если не по характеру, то по результату. Можно ли допустить существование такой возможности? Чего не хватает машине для того, чтобы полноценно заменить размышляющего человека?

Представим себе большого робота — машину, имеющую произвольно большое число чувствительных и исполнительных элементов, достаточный запас энергии и как угодно много запоминающих и комбинирующих элементов. Пусть мы можем каким-то образом собрать все эти элементы восдино и обеспечить их нормальное функционирование в течение некоторого промежутка времени. Пусть, кроме чувствительных элементов (рецепторов) и исполнительных (эффекторов), имеются элементы, возбуждение которых соответствует возбуждению комбинаций рецепторов и эффекторов. Поскольку число всех возможных сочетаний элементов гораздо больше, чем число самих элементов, естественно, что число таких «центральных» элементов будет значительно превышать число рецептов и эффекторов.

Допустим для простоты, что каждый элемент может находиться только в одном из двух возможных состояний: «возбуждение» и «покой». Возбуждение может передаваться от элемента к элементу непосредственно, при наличии соединения, прямого или через усилитель. Каждый элемент может получать возбуждение от любого числа других элементов и передавать его также любому числу элементов.

Любое внешнее событие, доступное восприятию робота, окажет воздействие на его чувствительные элементы и превратится, таким образом, в сигнал возбуждения в его системе управления. Если чувствительных элементов много и они разнообразны, то различные события будут возбуждать различные комбинации чувствительных элементов, т. е. будут различаться роботом.

Каждая комбинация возбужденных чувствительных элементов будет давать возбуждение одного внутреннего центрального элемента, соответствующего этой комбинации рецепторов и представляющего, таким образом, соответствующий чувствительный «центр». Для того чтобы данный «центр» возбуждался тогда и только тогда, когда возбуждена только данная комбинация рецепторов (ни одного лишнего и ни одного недостающего), достаточно включить необходимые чувствительные элементы по схеме, изображающей логическое высказывание

$$a_1 \& a_2 \& a_3 \dots \& a_k \& \bar{b}_1 \& \bar{b}_2 \& \bar{b}_3 \dots \& \bar{b}_l,$$

где a_k — элементы избранной комбинации,
 b_l — прочие элементы.

Далее, исполнительные элементы могут быть также включены группами, возбуждаемыми каждая от своего моторного «центрального» элемента. В результате мы получаем робота, способного различать сложные воздействия и совершать сложные реакции.

Остается только соединить чувствительные и моторные «центры», сопоставив внешние раздражения и реакции робота с некоторой стратегией, и мы получим робота с мгновенной реакцией (т. е. без памяти) с заданным поведением в данной обстановке. Если при этом исполнительные элементы — эффекторы робота — соединены с энергетическими выходными устройствами, реакции робота будут воздействовать на внешнюю среду. Можно, однако, оставить эти реакции в виде сигналов возбуждения эф-

факторных элементов робота, не доводя их до энергетических устройств.

Создадим много различных роботов подобного типа с различными стратегиями (правилами поведения) и предусмотрим возможность «включать» их в работу отдельно, по очереди. Мы можем конструктивно совместить этих роботов таким образом, чтобы один и тот же элемент или группа элементов поочередно работали в различных схемах. Поручим управление комбинированной группой роботов новому роботу «старшего» типа и снабдим его некоторыми дополнительными возможностями, например возможностью запоминать внешние воздействия и реакции «младших» роботов, а также включать или выключать любого младшего робота в соответствии с некоторой собственной внутренней программой, ну скажем, случайным образом или в некотором порядке, по очереди. Пусть также старший робот обладает некоторым особым родом чувствительности к воздействиям внешней среды, состоящим в том, что он не только регистрирует эти воздействия в виде сигналов, но и различает их таким образом, что различные классы внешних влияний оказывают на него «стимулирующие» воздействия, изменяя тем или иным способом программу его работы. Для достижения последней цели, возможно, придется создать отдельную схему или самостоятельного робота, связанного со старшим роботом и управляющего его действиями. Это будет уже «старший» робот второго класса.

Пусть теперь под перекрестным огнем внешних влияний наши совместно работающие роботы оказываются включенными в работу, пока без робота второго класса. Старший робот, работая по программе, включает одного из младших роботов, который автоматически выполняет свойственные ему реакции на внешнюю обстановку. Старший робот запоминает и воздействия, и реакции. Затем включается поочередно второй, третий и остальные младшие роботы. Когда методом проб использованы все роботы, включается старший робот второго класса, который обследует память первого старшего робота и оценивает или классифицирует работу младших роботов. В результате его работы некоторые из младших роботов, получившие высокие оценки по различным критериям, отбираются в особую группу пригодных для дальнейшей работы, а остальные, получившие «неуд» по всем предметам, либо попросту уничтожаются и никогда больше не включа-

ются в работу, либо структурно изменяются и испытываются повторно.

Отобранные младшие роботы остаются в работе и включаются старшим роботом в соответствии с измененной программой. Дальнейшая работа приносит новый опыт применения отдельных младших роботов, уточнение оценок, дополнительный отбор. Все меньше становится случайных неудачных реакций, все рациональнее — в смысле удовлетворения заложенным критериям — будет поведение системы, все больше знаний накапливается в памяти старшего робота.

Мы могли бы еще и еще усложнять систему, наращивая верхние этажи управления, например, с целью испытания и подбора критериев оценки работы младших роботов, введения оценок старших роботов и т. д. Можно также воспользоваться принципами самоорганизации, т. е. сделать общим и всеобъемлющим критерием работы всей системы естественное условие длительного устойчивого существования системы в целом в окружающей среде. Естественно, что для процесса организации или обучения системы потребовался бы некоторый период времени.

По прошествии времени обучения система реагирует на внешние воздействия оптимальным или почти оптимальным образом, пока она находится в «привычных» условиях.

Любое сложное воздействие может быть сначала зафиксировано в памяти, затем память обследована в поисках аналогичных предыдущих воздействий и по этим воспоминаниям восстановлены все возможные реакции, их результаты и оценки результатов. Из воспоминаний о результатах может быть выбран один с наивысшей оценкой в соответствии с действующим критерием оценки («настроением» или «расположением духа» системы, если позволить себевольные сравнения) и осуществлена та реакция на действующее раздражение, которая согласно прошлому опыту давала этот результат.

Но вот, допустим, во внешней обстановке произошло некоторое отступление от привычных условий, например, появился какой-то неизвестный объект, оказывающий на систему непривычные возбуждения. Эти возбуждения обычным образом воспринимаются машиной, запоминаются и сравниваются с памятью. Но непосредственного совпадения не получается. Новое воздействие не может быть опознано, и оказывается невозможным выбрать реакцию

непосредственно на основе предыдущего опыта. Человек в подобных условиях вынужден бывает задуматься и поразмыслить. Он должен что-то придумать.

В нашем сложном роботе возникает состояние неустойчивости («неудовлетворенности», «волнения»), вызываемое отсутствием сигнала совпадения после полного обследования памяти. Это состояние можно назвать состоянием «наличия вопроса» или «необходимости решить проблему».

Это состояние системы, поскольку оно выражено в конкретном сигнале, может включить программу действий, направленную на решение задачи о наилучшей реакции на известное воздействие. Для этого производится анализ воздействия, оно расчленяется на элементы или группы элементов, которые сопоставляются с памятью, выделяется то, что уже знакомо, и то, что представляет новизну, и таким образом задача формулируется. Далее, совместно группируется все известное и делаются попытки дополнить недостающие сведения сведениями, взятыми наугад или по ассоциациям из памяти, до получения некоторого комплекса сведений, который совпадает с одним из имеющихся в памяти. Иначе говоря, строятся гипотезы или предположения. Каждая гипотеза проверяется, т. е. с ней сопоставляется реакция, вероятный результат реакции, и этот результат сопоставляется с комплексом уже имеющихся сведений о новом объекте. При наличии противоречий испытывается новая гипотеза, которая «выдвигается» тем же путем дополнения неполных сведений. Все это происходит без исполнения реакций, на основе анализа содержимого памяти, где зафиксирован обширный материал о воздействиях, реакциях, их результатах. Обсуждение гипотез без принятия решения и без реакции представляет собой характерную особенность размышления. Система уже не пробует реакции наугад с риском ошибиться и прийти к плохим последствиям, а перебирает в памяти все возможные последствия реакций и сопоставляет их со сведениями о новом объекте. Если одна из гипотез приведет к согласию с предыдущим опытом, состояние «вопроса» снимается, решение найдено, оно исполняется.

Если же гипотезы иссякают, вопросительное состояние приводит либо к энергичной хаотической реакции, действиям наугад (панической реакции), в случае если воздействие приносит большой вред, либо к осторожному экс-

перименту — выжидательной реакции на получение дополнительных сведений для анализа.

То обстоятельство, что испытание гипотез может производиться без действительных реакций, на основе запасов памяти в сигнальном алфавите и коде системы, описывающем внешние события и их взаимосвязь с реакциями, имеет большое значение. Это не пробы наугад, не эмпиризм, а некоторое подобие творческих теоретических построений — размышление.

Решенная задача запоминается и входит в актив памяти для дальнейшего использования, как используется однажды доказанная математическая теорема, т.е. сразу от формулировки к результату без промежуточных рассуждений.

То, что мы сказали о гипотетической системе сложного робота, способного размышлять, представляет собой просто фантазию, быть может, слабо обоснованную. Если бы мы вздумали приступить к осуществлению подобной системы на современных машинах, мы бы встретились с целым рядом неудобств, трудностей и нерешенных вопросов. Мы были бы вынуждены совмещать в одной системе свойства различных существующих систем, а кроме того, проделать большую работу по конкретному программированию этапов работы системы. Возможно, что такая задача оказалась бы сложнее, чем, например, задача машинного перевода текста или создания информационного справочника.

Для того чтобы система выполняла описанные функции, она должна работать непрерывно, в отличие от существующих счетных машин, запускаемых старт-сигналом после ввода данных. Кроме того, система должна обладать рецепторами — чувствительными элементами, которые выдают в машину сигналы в алфавите машины в соответствии с внешними событиями. Большой робот должен работать, в общем, так, как работает автомат, непрерывно ощущающий обстановку и непрерывно реагирующий на нее, только, разумеется, гораздо более замысловато.

Обследование памяти, выбор в соответствии с критерием, группировка по признакам и другие подобные операции не новы, так как они широко применяются в действующих сегодня машинах. Однако следует отметить, что в современных машинах «вспоминание» сводится к перебиранию всего содержимого памяти (как поиски слова

в машинном словаре). «Вспоминание» по ассоциации современной машине чуждо.

«Вопросительное состояние», о котором мы говорили выше, как об исходном моменте размышления, насколько нам известно, до сего времени в машинах не реализовано*.

Принципиально было бы очень интересно подробнее исследовать это состояние, порождающее поиски, изменения программы поведения и т. д. Ощущение наличия проблемы является очень важным моментом мыслительной работы, который до сих пор чужд современному «машинному разуму». Возможно, что введение в машину способности «ощущать вопрос» приведет к появлению у нее некоторого подобия различных «эмоциональных» или «аффективных» состояний или потребует их введения наряду с соответствующими центральными структурами, регулирующими состояния «нормальной реакции», «вопросительного возбуждения» (или «неудовлетворенности»), «ускоренного поиска реакции» (или «паники») и т. д.

Сложный робот должен быть непосредственно связан с эффекторами — исполнительными органами, причем их включение для осуществления реакции должно производиться отдельно от «принятия решения», т. е. достижения окончательного состояния выходных сигнальных элементов робота. Таким образом, и ощущения, и реакции должны быть зафиксированы в памяти и существовать в виде сигналов, соответствующих реальным воздействиям и реакциям. Несомненно, что содержание памяти при этом должно оказаться богаче, чем реальное поведение, так как на выход системы подаются только некоторые реакции, отобранные из большого числа возможных.

Сложные роботы, подобные описанному выше, будут неизбежно обладать собственными «темпераментами». Смещая критерии оценок и изменяя программы поиска при решении задач и выборе реакции, можно было бы при необходимости построить робота «задумчивого» — испытывающего много гипотез до выбора реакции, робота «решительного» — быстро осуществляющего первую же найденную более или менее подходящую реакцию, робота «пугливого» — склонного к резким «необдуманным» реакциям, робота «изобретательного» — способного к выполнению

* Ближайшим подобием этого состояния является переход к подпрограмме отбрасывания окончаний при отсутствии искомого слова в словаре в случае машинного перевода.

большого числа разнообразных операций с имеющейся в его распоряжении информацией и к образованию разнообразных гипотез и предположений, и т. д.

Машина, способная не только решать, но и ставить задачи, ощущая необходимость этого, явится новым шагом в освобождении человека от повседневной умственной работы. Машина не станет уклоняться от трудных вопросов, как это часто склонен делать человек, или заменять собственные выводы, полученные на основе анализа фактов и опыта, предубеждениями, предрассудками или подражаниями; во всяком случае, ей можно запретить это делать.

На основе изучения таких машин, несомненно, был бы приобретен ценнейший материал о мыслительной работе человека и о протекающих при этом сигнальных процессах. Интересен вопрос о взаимоотношениях двух подобных роботов, созданных и «воспитанных» независимо друг от друга. Будут ли такие роботы понимать друг друга и в какой степени? По-видимому, если при их создании не предусмотрен общий алфавит и общий код для связи (общий язык), как это всегда предусматривается при создании каналов связи и других единых кибернетических систем, то понимание будет очень ограниченным или будет полностью отсутствовать, по крайней мере, в начале «знакомства». Согласованность действий обоих «незнакомых» или «чужих» роботов и обмен информацией между ними могут возникнуть самостоятельно, но только на основе их действий, на основе активного общения. Характер этого общения будет зависеть от устройства обоих роботов.

«Коллектив», состоящий из многих роботов, представляет собой еще более сложный и интересный объект для изучения. Каким образом характеристики членов коллектива повлияют на их взаимодействие и на самоорганизацию коллектива? Какие сведения, полезные для общества людей, можно извлечь из изучения коллектива роботов?

Исследования подобного рода моделей творческих личностей и коллективов представляются чрезвычайно интересными и плодотворными. Следует проводить такие исследования, не пугаясь трудностей и не считаясь с предрассудками.

Хочется сказать здесь, что в этой, еще очень мало исследованной области должны быть приложены совместные усилия ученых, инженеров и изобретателей самых различных специальностей. Даже сравнительно простые модели

с небольшим числом элементов могут оказаться очень поучительными и обладающими удивительным разнообразием поведения. Ведь число комбинаций работающих элементов растет несравненно быстрее, чем само число элементов. Может быть, именно этим в значительной степени и объясняется сложность поведения животных и человека, кажущаяся нам порой «бесконечной» и «неисчерпаемой». Поле деятельности для изучения свойств кибернетических систем и их полезного использования поистине огромно и необозримо. И это поле уже давно должно было стать ареной организованной, планомерной деятельности и специалистов, действующих по единому государственному плану, и энтузиастов-любителей, организованных в добровольные общества и кружки.

Быть может, скоро мы станем свидетелями спортивных соревнований роботов-моделей, созданных кибернетиками-любителями и выполняющих некоторую стандартную интеллектуальную задачу данного спортивного класса, например, играющих в какую-либо игру друг против друга.

Трудно перечислить все задачи и все области кибернетики, требующие активной исследовательской деятельности. Попытаемся назвать лишь некоторые из них.

Во-первых,— это изучение способов кодирования информации, т. е. выражения необходимых сведений в различных физических алфавитах и перевода из одного алфавита в другой. Сюда относятся вопросы построения наиболее экономных и сжатых оптимальных кодов, а также кодов с использованием избыточности и обладающих возможностями заметить и поправить ошибки, возникающие от шумов (самокорректирующиеся коды). Сюда же относятся вопросы оптимизации каналов связи и полного использования их пропускной способности. К тому же кругу проблем принадлежат вопросы создания чувствительных устройств, регистрирующих события и описывающих их в некотором алфавите, т. е. по существу выполняющих задачу кодирования информации.

Во-вторых,— это множество проблем, связанных с переработкой информации: классификация информации по признакам, выявление связей между элементами информации, опознавание классов или комплексов этих элементов; различение образов, сочетаний, образование операций преобразования или алгоритмов переработки информации, оценка результатов переработки в соответствии с заданными критериями, выработка частных критериев из общих,

образование оптимальных реакций, выработка оптимальной стратегии поведения, запоминание информации и обследование памяти по заданным признакам (ассоциациям).

Еще один круг проблем — анализ структуры с целью выявления ее функции, исходя из свойств элементов структуры и их взаимосвязей.

Далее следует более сложный комплекс проблем — создание структур, выполняющих заданную функцию, проблема моделирования кибернетических систем, создание экономичных систем из малого числа элементов, создание надежно работающих систем из малонадежных элементов, физическая реализация алгоритмов из заданных структурных элементов.

И, наконец, — изучение собственного поведения систем, снабженных обратными связями того или иного вида, обучение, самоорганизация, самостоятельные структурные изменения системы по мере накопления внешней информации.

Несомненно, что изучение упомянутых здесь проблем, равно как и многих других, должно и будет протекать в процессе создания полезных устройств.

Развитие техники синтеза кибернетических систем на основе мобилизации сведений всех наук под общим организующим влиянием математики даст человечеству мощный и точный разум, свободный от человеческих слабостей, даст понимание свойств личности и коллектива, позволит выявить и использовать резервы человеческих способностей и в итоге поднимет человека еще на одну ступень культуры.



ЗАКЛЮЧЕНИЕ

Одним из наиболее интересных и привлекательных аспектов кибернетики является, по нашему мнению, возможность приблизиться к объяснению сущности человеческого сознания и мышления. До сих пор было очень трудно подойти к этому объяснению, и вокруг сущности сознания не смолкали бурные и обширные споры. Нам нет нужды, да у нас и нет возможности напоминать здесь многовековую историю борьбы философских воззрений по поводу природы сознания. Сама длительность этих споров говорит о том, что в этом вопросе существуют реальные трудности. В чем же они заключаются?

Всякое объяснение есть в конечном итоге сведение неизвестного к известному путем известных операций, например, путем сравнения. Однако до настоящего времени не существовало ничего, подобного сознанию, ничего такого, что позволило бы проводить сравнения, что давало бы необходимые понятия, термины, образы для установления сходства и различия, словом, для определения и, следовательно, для объяснения сознания. Поэтому сознание

являлось единственным в своем роде фактом, не поддающимся сравнению и объяснению. Немудрено, что единственность этого явления приводила многих к ошибочной мысли о его первичности и несводимости к другим, аналогичным явлениям.

Последнее время ситуация постепенно, но решительно меняется. Сделаны первые шаги по созданию систем, которые моделируют отдельные процессы восприятия и обработки информации, выработки реакций. Выработаны основы системы понятий, позволяющей поставить в один ряд некоторые проявления человеческого интеллекта и работу искусственно созданных систем. Моделирование «чисто человеческой» деятельности пока что идет успешно. Человеческое сознание, возможно, потеряет в будущем свое исключительное положение и станет одним из многих явлений, свойственных различным видам естественно или искусственно организованной материи. Появится большое число разнообразных сигнальных или кибернетических процессов в различных физических реализациях, результаты которых, а быть может, и самый ход, будут эквивалентны результатам мышления, а возможно, и превзойдут их.

Ведь и процессы в сложной машине, и процессы в мозге суть лишь различные физические, сигнальные процессы, несущие информацию либо осведомительного, либо управляющего характера. И подобно тому, как физиологические процессы в мозге не исчерпывают сущности мышления, электрический ток в машине не объясняет процесса перевода машиной иностранного текста на русский язык. Оба процесса являются сигнальными, они несут информацию и способны производить действия. Именно в этом сущность и мысли, и работы управляющей системы, в этом их единство, в этом их общность.

Мы идем к освобождению от фетишизации человеческого «духа», «души», идем к тому, чтобы полностью или пока хотя бы отчасти понять сущность тех материальных кибернетических явлений, которые собирательно обозначены этими терминами. И когда человек своими руками построит машину, к творческим идеям которой он сам должен будет относиться с уважением, он не сможет больше пользоваться ни ложным понятием, ни термином «душа» или «дух», как чем-то, что существует вне и помимо физического тела. От них придется отказаться всем и окончательно. Так будут, возможно, прекращены споры о «душе»

Часто говорят по этому поводу, что кибернетика не только «развенчивает» человеческий разум, но и «принижает» его достоинство. Такое мнение противоречиво, и с ним мы не можем согласиться. Человеческий разум представляет собой огромную силу, одержавшую немало побед в познании явлений и законов природы. И самой замечательной его победой, несомненно, явится познание собственной его сущности и силы. В результате, не только мышцы человечества будут иметь силу атомного взрыва, но и его разум приобретет мудрость такого же масштаба. Вряд ли можно называть это «принижением достоинства» разума. Ведь экскаватор не принижает достоинства человека; он ему помогает, заменяет его и позволяет ему заниматься другим, более важным, увлекательным, трудным человеческим делом, чем копанье земли.

Однако объяснение сущности человеческого сознания пока является делом будущего, будущими плодами тех трудов, которые мы должны совершить на поле научной деятельности. Сегодня мы говорили об этом, как о перспективе. И многие сегодня с нами не согласны. Почему? Когда ведешь беседу с людьми разных складов ума, разных темпераментов, различного воспитания и специальностей по поводу перспектив и возможностей кибернетики, то приходится выслушивать не только много положительных, но и некоторое количество отрицательных суждений. Подавляющая часть возражений, которые иногда приходится слышать, сводится к одному утверждению, которое всегда высказывается с крайней убежденностью и решительностью и без каких-либо попыток аргументации. Это утверждение сводится к тому, что «машина не сможет заменить человека потому, что она машина» или же «потому, что человек — не машина!» Подобно отставному Войска Донского поручику Семи-Булатову из села Блины-Съедены (А. П. Чехов «Письмо ученому соседу»), некоторые готовы утверждать, что «этого не может быть потому, что этого не может быть никогда», а иногда и действительно произносят нечто подобное и при том с неменьшей убежденностью, чем чеховский герой. Эта убежденность, по-видимому, не случайна и имеет свои причины, которые важно знать, так как порой они приводят к плачевным последствиям. Достаточно вспомнить содержание критических статей, которые в 1951—1952 гг. писались против кибернетики. Все они разделяют мнение и пользуются «аргументацией» Семи-Булатова. Но в ре-

зультате такой «критики» мы начали заниматься кибернетикой с опозданием и несем от этого сегодня заметный урон.

Чем же объясняется (хоть и не оправдывается) глубокое убеждение многих в том, что «машина не может делать всего, что делает человек»? Почему некоторым людям одинаково знакома сама идея объективности и закономерности явлений, связанных с человеческим мышлением, сознанием, волей?

Прежде всего, этому способствует то, что подсознательная деятельность мозга несравненно обширнее сознательной. Человек, если можно так сказать, знает о себе только ту часть, которую сознает, а это очень мало. В его памяти, вне сознания, хранится гораздо больше. Поэтому многие мыслительные акты представляются ему суверенными, изначальноными, ничем не обусловленными. Отсюда иллюзия свободы воли и фетишизация этой свободы.

Кроме того, признанию «машинности», причиной обусловленности мысли мешают инстинкты и, в первую очередь, один из самых сильных — инстинкт самосохранения, которому мы обязаны в значительной мере выживанием биологического вида. Подсознательный страх смерти, присутствующий в живом организме, иногда не позволяет некоторым людям охотно признать факты, которые приводят к заключению о неизбежности гибели любого индивидуума наряду с разрушением всякой существующей машины. Признание хотя бы отдаленного родства между «живым» и «мертвым» разрушает подсознательную веру в исключительность человеческого существования, надежду на то, что человек обладает особыми свойствами и особыми правами на бессмертие в той или иной фантастической форме. Вера в бессмертную душу, коренящаяся в инстинкте самосохранения, иногда бывает заперта глубоко в подсознании, но и тогда она является судьей и законодателем суждений, заставляя не принимать суждений, неприятных с этой точки зрения. Иногда эта вера проявляется в осознанных религиозных воззрениях, вступая в конфликт с фактами и существую, несмотря на явные противоречия с ними. Чем меньше человек знает научных фактов и чем сильнее в нем инстинкты, тем более подвержен он подсознательным инстинктивным предрассудкам.

Признание фактов, принятие к руководству основанных на фактах научных сведений, требует от человека некоторого внутреннего самоотречения, победы над собственными

ми древними животными инстинктами, замены подсознательных критериев оценки суждений разумными — чисто человеческими. Не всякому это удастся в одинаковой степени.

Когда слышишь утверждения без научной, т. е. основанной на фактах и логике аргументации, убеждения без их разумного подтверждения, невольно приходишь к выводу о том, что ведешь беседу с догматиком, руководствующимся в своих суждениях предубеждениями и предпочитающим интуитивную уверенность разумным, научным доказательствам.

Нечто подобное имело место в те времена, когда Чарльз Дарвин показал несомненное биологическое родство человека и низших животных. Идея родства человека с приматами вошла тогда у многих современников Дарвина в конфликт с той же инстинктивной верой в исключительность человека как вида и как индивидуума. И хотя сегодня уже никто не может усомниться в сходстве различных биологических видов, Дарвину в свое время пришлось выслушать немало суждений того же характера, какие можно услышать порой в наши дни по поводу идеи об аналогиях между мыслительной сигнальной деятельностью живого мозга и сигнальными процессами в машине.

Действительную исключительность человек сумеет сохранить за собой всегда, даже и в те времена, когда машины будут не только работать физически, но и думать за человека. Эта исключительность выразится в том, что человек сумеет остаться на вершине руководства всеми управляющими системами, как природными, которые он изучит и которыми овладеет, так и искусственно созданными им самим. Единственным общим критерием качества работы системы умных роботов будет правильно понятая польза, приносимая человеческому обществу. И роботы придут нам на помощь тем скорее, чем раньше мы сумеем освободиться от господства предрассудков, чем скорее мы признаем узы нашего «родства» с роботами.

Не исключена возможность, что это родство может оказаться более близким, чем мы сегодня в состоянии признать. Изучение самоорганизующихся систем, едва только начатое сегодня, может привести к результатам совершенно неожиданным, предвидеть которые попросту невозможно.

КРАТКАЯ БИБЛИОГРАФИЯ

- 1) Н. Винер. Кибернетика. Перевод с англ. Изд. «Советское радио», 1958.
- 2) С. Л. Соболев, А. И. Китов, А. А. Ляпунов. Основные черты кибернетики. «Вопр. философии», № 4, 1955.
- 3) Э. Кольман. Что такое кибернетика. Там же.
- 4) А. И. Китов. Техническая кибернетика. «Радио», № 11, 1955.
- 5) Колмогоров А. Н. Теория передачи информации. Сессия АН СССР по научным проблемам автоматизации производства, октябрь, 1956, пленарное заседание. Изд. АН СССР, 1957.
- 6) Б. В. Гнеденко и А. Я. Хинчин. Элементарное введение в теорию вероятностей. Гос. издательство технико-теоретической литературы, 1952.
- 7) А. Я. Хинчин. Об основных теоремах теории информации. «Успехи матем. наук», т. 11, вып. 1(67), январь 1956.
- 8) Теория передачи электрических сигналов при наличии помех. Сборник статей. Изд. иностранной литературы, 1953.
- 9) И. А. Кошечев. Основы теории электрической связи. Изд. литературы по вопросам связи и радио, 1954.
- 10) Л. Маккол. Основы теории сервомеханизмов. Изд. иностранной литературы, 1947.
- 11) Х. Джеймс, Н. Никольс, Р. Филипс. Теория следящих систем. Изд. иностранной литературы, 1951.
- 12) А. И. Китов. Электронные цифровые машины. Изд. «Советское радио», 1956.
- 13) Д. Гильберт и В. Аккерман. Основы теоретической логики. Изд. иностранной литературы, 1947.
- 14) Автоматы. Сборник статей под ред. Шэннона и Маккарти. Перевод с англ. под ред. А. А. Ляпунова. Изд. иностранной литературы, 1956.
- 15) Д. Ю. Панов. Автоматический перевод. Изд. АН СССР, 1956.
- 16) Цзянь Сюэ Сэнь. Техническая кибернетика. Изд. иностранной литературы, 1956.
- 17) М. Брейзье. Электрическая активность нервной системы. Изд. иностранной литературы, 1955.
- 18) Р. Джерард. Функции человеческого тела. Изд. иностранной литературы, 1947.
- 19) М. М. Завадовский. Противоречивые взаимодействия в живом организме. Изд. МГУ, 1941.
- 20) Блэквелл и Гиршик. Теория игр и статистических решений. Изд. иностранной литературы, 1958.
- 21) Л. Бриллюэн. Демон Максвелла не может работать. J. Appl. Phys., 1951, v. 22, № 3.

- 22) В. Г. Уолтер. Имитация жизни. *Scient. Amer.*, 1950, v. 182, № 5.
- 23) Е. С. Беркли. Зрячий электронный зверек. *Radio-Electronics*, 1951, v. 23, № 3.
- 24) Е. Эйхлер. Искусственная черепаха. *Radiotechnik (Radio-amateur)*, 1955, v. 31, № 5/6.
- 25) К. Шеннон. Программирование машин для игры в шахматы. *Bell Labs. Record*, 1948, v. 8.
- 26) К. Шеннон. Машины, играющие в игры. *J. Franklin Inst.*, 1955, v. 260, № 6.
- 27) Н. Винер. Машина умнее своего создателя. *Electronics*, 1953, v. 26, № 6.
- 28) Д. В. Хагельбергер. Робот с последовательной экстраполяцией. *Trans. I.R.E., EC-5*, 1956, № 1.
- 29) Б. Е. Фарли и В. А. Кларк. Воспроизведение самоорганизующихся систем на машине дискретного счета. *Trans. I.R.E. PGIT-4*, Sept. 1954.
- 30) Г. Гамов. Передача информации в живой клетке. *Scient. Amer.*, 1955, № 4.
- 31) Е. Л. Грюнеберг. Размышление в машине. *Computers and Automat.*, 1951, v. 3, № 2.

О Г Л А В Л Е Н И Е

Предисловие	9
Глава 1. Энергетика и кибернетика	9
Управление потоками энергии	11
Информация	22
Глава 2. Сигнал	25
Изоморфизм	28
Возникновение сигнала и реакция	34
Сигналы порождают сигналы	36
Полного описания	37
Дискретные и непрерывные сигналы	40
Глава 3. Случай	44
Случайные события	44
Вероятность	47
Распределение вероятностей. Математическое ожидание и дисперсия	53
Шум	58
Вероятность в физике. Энтропия	62
Глава 4. Количество информации	70
Информационная емкость	70
Мера количества информации	75
Код	80
Скорость передачи	84
Самокорректирующиеся коды	89
Теорема Шэннона	92
Информационная и физическая энтропия. Организация	94
Глава 5. Передача сигнала	102
Модуляция	103
Искажения	110
Спектр и полоса пропускания	113
Селекция	116
Пример канала связи	119
Работа нервного волокна	124
Глава 6. Обратная связь. Регулирование	128
Работа систем регулирования	137
Обратная связь в живых организмах	141
Рефлексы	150
Глава 7. Сигнал в машине	162
Вычислительные устройства непрерывного действия	164
Дискретный счет	169
Вычислительные устройства дискретного счета	172
Счетно-аналитические машины	176
	403

Быстродействующие электронные вычислительные машины	180
Управление и память	193
Особенности электронных машин	203
Теоретическая логика и алгебра релейных схем	206
Релейные схемы	212
Машины и мышление	218
Глава 8. Робот	221
Кибернетические игрушки	233
Роботы и язык	245
Автологический перевод	253
Другие роботы	267
Глава 9. Мысль	275
Человек у машины	277
Строение нервной системы	283
Работа мозга	293
Эмоции	307
Сигнальные процессы в мозге	309
Чем не обладает машина?	315
Глава 10. Игра	331
Понятия теории игр	333
Стратегия	334
Функция пользы	336
Смешанное расширение игры	342
Случайные ходы	347
Машины, играющие в стратегические игры	350
Еще раз о неопределенных играх	354
Глава 11. Робот, который мог бы быть умнее своего конструктора	360
Самоорганизующиеся системы	362
Осуществление самоорганизующихся систем	367
Избыточность организации	376
Глава 12. Большой робот	380
Что же дальше?	380
Возможности кибернетических систем	385
Заключение	395
Краткая библиография	401

Игорь Андреевич Полетаев

СИГНАЛ

(О некоторых понятиях кибернетики)

Редактор *В. И. Грознова*

Техн. редактор *Н. Н. Корзуев*

Художественная редакция и оформление
художника *В. В. Волкова*

Сдано в набор 11/II 1958 г. Подписано к печати 2/VII 1958 г.
Формат 84×108^{1/2} Объем 20,71 п. л. Бум. л. 6,31 Уч. изд. л. 22,65
Г.41916 Цена в переплете № 5—7 р. 80 к., в переплете № 7—8 р. 80 к. Заказ 1131.

Типография Госэнергоиздата. Москва, Шлюзовая наб., 10.

7. p. 80. k.

« СОВЕТСКОЕ РАДИО »

Я. И. ЛУКОМСКИЙ

ТЕОРИЯ
КОРРЕЛЯЦИИ
И ЕЕ ПРИМЕНЕНИЕ
К АНАЛИЗУ
ПРОИЗВОДСТВА

Я. И. ЛУКОМСКИЙ

ТЕОРИЯ КОРРЕЛЯЦИИ
И ЕЕ ПРИМЕНЕНИЕ
К АНАЛИЗУ ПРОИЗВОДСТВА

ГОСУДАРСТВЕННОЕ СТАТИСТИЧЕСКОЕ ИЗДАТЕЛЬСТВО
МОСКВА 1958

В книге излагаются основные положения теории статистического изучения взаимосвязей между величинами, используемые для научной обработки опытных и отчетных данных предприятий. Эти положения иллюстрируются примерами и задачами из практики машиностроительного, металлургического и других производств.

Изложенные в книге методы имеют целью помочь изучению закономерностей производственных процессов и в ряде случаев могут способствовать сокращению затрат на дорогостоящие специальные исследования.

Книга рассчитана на инженерно-технических работников, экономистов, преподавателей вузов и может быть использована в качестве учебного пособия по математической статистике.

ПРЕДИСЛОВИЕ

Теория корреляции является одним из центральных разделов математической статистики, в котором рассматриваются методы статистического изучения взаимосвязей между величинами. Применение методов математической статистики, в частности методов корреляции, при решении различных инженерно-технических и экономических вопросов производства приобретает важное значение в связи с развитием новой техники, автоматизацией производственных процессов и повышением требований к качеству инженерных расчетов. Однако широкому распространению методов корреляции препятствует недостаточное знакомство с ними наших инженерно-технических и экономических кадров, что в значительной степени связано с отсутствием учебной литературы, где изложение этих методов было бы достаточно элементарным и связанным с практикой их применения.

В настоящей книге излагаются теория и практика применения методов корреляции к анализу производства. Излагаемые в книге методы могут быть использованы для научной обработки опытных и отчетных данных, получаемых в процессе производства, что позволит полнее изучать закономерности производственных процессов и в ряде случаев сократить затраты на проведение специальных дорогостоящих экспериментальных исследований.

Автор стремился сделать книгу максимально доступной для инженеров и экономистов промышленности. Этим назначением книги определяется ее построение и характер изложения.

Математический аппарат, используемый в книге, является вполне элементарным и опирается на основные сведения из вузовского курса высшей математики. Теоретический материал иллюстрируется большим количеством примеров и задач, составленных по материалам заводской практики и исследовательских работ.

Книга состоит из четырех разделов.

Первый раздел содержит подготовительный материал из математической статистики, необходимый для понимания центрального раздела книги — теории корреляции. Он написан наиболее

элементарно и не предполагает у читателя никаких предварительных сведений по этому предмету.

Второй раздел посвящен основам теории корреляции. Материал в нем изложен в объеме, необходимом для решения практических задач производства.

Третий раздел, посвященный ошибкам статистических показателей, имеет в данной книге вспомогательное значение. В нем приведены без математических выводов формулы для оценки погрешностей основных показателей корреляционного анализа. Однако автор считал необходимым более подробно остановиться на принципиальных основаниях таких оценок и в связи с этим на некоторых понятиях теории вероятностей, что важно как для сознательного использования указанных формул и учета реальных условий, при которых их применение оправдано, так и для более точного уяснения основных положений теории корреляции.

В четвертом разделе даны некоторые направления использования методов теории корреляции на производстве. Приведенные в разделе примеры имеют характер комплексных исследований в отличие от ряда примеров, приводимых в других разделах книги и предназначенных для иллюстрации отдельных вопросов теории. Разумеется, представленные направления ни в какой мере не претендуют на какую-либо полноту охвата; их надо рассматривать лишь как первые шаги на пути к широкому внедрению методов теории корреляции в производственные исследования.

В приложении дан ряд таблиц, содержащих первоначальные опытные данные, по которым читатель сможет самостоятельно воспроизвести расчеты, приведенные в книге. Данные этих таблиц могут быть также использованы для проведения практических и лабораторных работ при преподавании курса.

В основу книги легли лекции, читанные автором на протяжении ряда лет в высших технических и экономических учебных заведениях, и работы, выполненные при его участии для промышленности. Были использованы также учебные пособия автора, выпущенные ранее ограниченным тиражом на правах рукописи: «Теория корреляции», МАИ, 1948 и «Введение в теорию статистики. Ряды распределения», МАИ, 1947.

Большой труд по подготовке рукописи к изданию, проведению и проверке всех расчетных и графических работ взяла на себя инж. Р. В. Каткова, которой автор приносит глубокую благодарность.

РАЗДЕЛ I
РЯДЫ РАСПРЕДЕЛЕНИЯ

В математической статистике рассматриваются множества единиц одного и того же вида, называемые *статистическими совокупностями*. Это могут быть совокупности деталей одного и того же наименования, подвергающиеся одному и тому же процессу обработки, совокупности станков одного и того же вида, совокупности рабочих одной и той же специальности и т. д.

Отдельные единицы, входящие в состав совокупности, называются ее *элементами*. Число всех элементов называется *объемом совокупности*. Каждый элемент характеризуется рядом свойств или *признаков*, по которым элементы совокупности обнаруживают дисперсию — варьируют.

Дисперсией, вариацией, изменчивостью или *колеблемостью* называется способность признаков видоизменяться под влиянием большого количества различных причин или условий. Размер, вес, прочность, чистота поверхности обрабатываемых деталей, точность работы станков, выработка, квалификация, стаж рабочих служат примерами варьирующих признаков.

Отдельные значения или видоизменения признака называются его *вариантами*. По характеру своих вариантов признаки делятся на *количественные*, выраженные числами, и *атрибутивные*, которые числами не выражаются. Так, размер, вес, прочность являются количественными признаками, внешний вид, отделка, наличие или отсутствие дефектов — атрибутивными.

Элементы статистической совокупности *классифицируются* или *группируются* по вариантам признаков, и для каждой классификационной группы подсчитывается число элементов, или *частота*. При помощи частот рассчитываются различные *обобщающие* показатели статистического материала — относительные, средние и другие величины (доля брака, средний размер деталей и т. п.), которые выражают определенные соотношения или пропорции, складывающиеся в изменчивости варьирующих признаков и характеризующие изучаемую совокупность.

При увеличении числа наблюдений обобщающие показатели становятся более точными: они позволяют обнаружить те или иные правильности, закономерности, свойственные рассматрива-

емым явлениям, которые не заметны при рассмотрении отдельных случаев или при малом числе наблюдений. В этом проявляется действие *закона больших чисел*. Если, например, при текущем контроле качества продукции произведено 2—3 замера изготавливаемых на станке изделий и полученная средняя близка к требуемому размеру, то это еще не означает, что станок правильно налажен, так же как и отклонение такой средней от требуемого размера не говорило бы еще о противоположном. Но если средняя из нескольких десятков или сот измерений мало отклоняется от требуемого размера, то это уже является показателем правильности настройки. Если выносливость стали марки А выше, чем марки В, и сопротивление на разрыв у стали марки А также выше, то это еще не свидетельствует о закономерной связи между тем и другим показателем. Но если при сопоставлении нескольких десятков марок почти во всех случаях обнаруживается такое же соотношение, то это уже будет показывать наличие определенной закономерности.

По определению акад. А. Н. Колмогорова, законом больших чисел следует называть общий принцип, в силу которого совокупное действие большого числа случайных факторов приводит при некоторых весьма общих условиях к результату, почти не зависящему от случая¹.

Теория закона больших чисел служит основой для оценок надежности выводов и заключений, получаемых при помощи методов математической статистики и, в частности, методов корреляции.

¹ Большая Советская Энциклопедия, 2-е изд., т. 5, «Больших чисел закон».

ГЛАВА I

ПОСТРОЕНИЕ И ГРАФИЧЕСКОЕ ИЗОБРАЖЕНИЕ РЯДОВ РАСПРЕДЕЛЕНИЯ

§ 1. ПОСТРОЕНИЕ РЯДОВ РАСПРЕДЕЛЕНИЯ

В настоящем разделе излагаются приемы изучения изменчивости элементов статистической совокупности по отдельно взятому варьирующему признаку.

Изложение мы начнем с примеров статистической оценки качества продукции.

Вопросы улучшения качества продукции стоят в центре внимания нашей промышленности. Предприятие, выпуская продукцию того или иного вида, обязано строго соблюдать установленные стандарты ее качества. Для этой цели отделами технического контроля и в заводских лабораториях производятся испытания важнейших механических, физических и химических свойств (признаков) продукции. Эти признаки являются, как правило, признаками варьирующими, и, следовательно, их значения могут быть неодинаковыми для различных образцов продукции одного вида. Требуется большое количество таких испытаний для того, чтобы составить себе представление о качестве данного вида продукции.

Пример 1. При обработке болтов на многшпиндельном автомате требовалось получить размер диаметра $11,5^{-0,24}$. Были проведены замеры 190 деталей. Результаты замеров приводятся в табл. 1.

Чтобы разобраться в полученном материале и сделать необходимые выводы, данные табл. 1 должны быть прежде всего приведены в порядок, систематизированы. Этой цели служит табл. 2, в которой перечислены по порядку значения диаметра и указаны числа повторений каждого значения (столбцы 1 и 2 табл. 2).

Подобная таблица называется рядом распределения 190 деталей по размеру диаметра.

Таблица 1
Данные о размерах диаметра 190 деталей
(в мм)

11,34	11,42	11,40	11,38	11,42	11,38	11,44	11,44	11,38
38	40	40	42	36	44	40	40	44
38	40	38	44	38	46	34	38	40
38	38	38	44	42	44	38	34	38
34	40	36	44	48	40	40	40	36
32	42	34	46	38	46	38	42	42
34	36	38	38	36	38	40	32	36
38	42	42	38	32	36	30	38	40
36	40	44	42	36	40	34	38	40
34	34	36	36	40	42	38	36	42
42	36	42	40	42	40	34	34	40
36	42	40	44	40	40	40	32	34
36	42	32	38	46	40	46	32	42
42	44	40	38	36	42	38	38	38
36	32	44	36	38	40	42	36	
40	34	40	40	40	40	42	36	
38	42	34	36	42	42	32	38	
44	44	36	40	40	38	40	38	
42	40	42	40	40	36	42	38	
36	42	34	42	38	40	40	36	
40	40	44	42	44	40	40	36	
34	42	42	44	46	38	40	34	

Таблица 2
Распределение 190 деталей болта по размеру диаметра

Варианты (в мм)	Абсолютные частоты (число деталей)	Относительные частоты (в процентах)
11,30	1	0,53
11,32	8	4,21
11,34	17	8,95
11,36	26	13,68
11,38	36	18,95
11,40	44	23,16
11,42	34	17,89
11,44	17	8,95
11,46	6	3,15
11,48	1	0,53
Итого .	190	100,00

Итак, рядом распределения называется таблица, в которой перечислены по порядку варианты признака и указаны числа повторений каждого варианта.

Число повторений варианта называется его абсолютной частотой. Сумма всех частот равна, очевидно, объему совокупности.

Отношение абсолютной частоты к объему совокупности называется *относительной частотой*. Относительные частоты могут выражаться в процентах (столбец 3 табл. 2).

При анализе таблицы замечаем следующее:

1. Все полученные замеры находятся внутри поля допуска (11,26—11,50), следовательно, точность изготовления деталей удовлетворяет техническим условиям.

2. В изменчивости признака замечается следующая закономерность. Один из вариантов признака, находящийся приблизительно в центре ряда (11,40), встречается с наибольшей частотой. По мере удаления от этого варианта в сторону больших или меньших размеров диаметра частоты вариантов постепенно уменьшаются. Характер нарастания и убывания частот, в общем правильный, выражает закономерность, которой определяется связь между величиной вариантов и частотой, с которой они встречаются; эта закономерность отчетливо обнаружилась уже при данном количестве наблюдений.

Подобные закономерности, как мы дальше увидим, являются характерными для большинства рядов распределения, встречающихся при оценке качества продукции, а также и во многих других вопросах производства.

Заметим, что признак в табл. 2 представлен небольшим числом отдельных, различающихся между собой вариантов, каждый из которых встречается с некоторой частотой. Такие признаки и соответствующие им ряды называются признаками и рядами, заданными в *дискретной форме*. Легко видеть, однако, что каждый вариант таблицы получается в результате округления

близких к нему значений признака в процессе измерения с определенной точностью. Поэтому каждый из таких вариантов можно рассматривать как центр интервала, в котором заключается действительное значение измеряемого признака (диаметра). Следовательно, частоты ряда распределения будут относиться не к отдельным значениям признака, а к интервалам, которые называются *интервалами округления*. Соответствующая форма записи ряда распределения называется *интервальной*.

Ряд примера 1 в интервальной форме представлен в табл. 3.

Выражая признак в интервальной форме, мы обычно предполагаем, что признак, по существу, может принимать любое про-

Таблица 3

Интервалы (в мм)	Абсолютные частоты
11,29—11,31	1
11,31—11,33	8
11,33—11,35	17
11,35—11,37	26
11,37—11,39	36
11,39—11,41	44
11,41—11,43	34
11,43—11,45	17
11,45—11,47	6
11,47—11,49	1
Итого . .	190

межуточное значение внутри интервала. Такие признаки называются непрерывными.

Возможны ряды распределения принципиально дискретные, например: распределение станков по числу разладок за отчетный период, распределение числа зерен в поле зрения абразивного шлифа, распределение контрольных проб продукции по числу дефектных единиц в каждой пробе и т. д.

Представление таких рядов в интервальной форме в некоторых случаях также целесообразно (см. § 2 и 3), хотя и имеет условный характер.

Пример 2. Одним из основных признаков качества поршневого кольца является его упругость. Ниже приводим данные испытаний упругости (на Толедо) 200 поршневых колец (по материалам А. Н. Журавлева).

Построение ряда распределения имеет здесь свои особенности по сравнению с предыдущим примером. Упругость измеряется с большой степенью точности, и поэтому в табл. 4 мы видим большое количество различных значений признака. Необходимо провести укрупненную группировку, объединяя близкие между

Таблица 4

Данные испытаний 200 колец по упругости (в г/мм²)

1 450	1 460	1 520	1 500	1 410	1 510	1 580
1 450	1 520	1 420	1 480	1 360	1 400	1 400
1 460	1 420	1 380	1 340	1 550	1 440	1 480
1 660	1 440	1 380	1 380	1 480	1 460	1 500
1 575	1 420	1 380	1 340	1 340	1 400	1 430
1 500	1 480	1 360	1 510	1 350	1 290	1 430
1 460	1 450	1 490	1 320	1 340	1 360	1 370
1 340	1 350	1 300	1 340	1 330	1 500	1 390
1 450	1 480	1 200	1 430	1 340	1 400	1 480
1 440	1 580	1 390	1 230	1 450	1 430	1 120
1 510	1 410	1 500	1 460	1 490	1 430	1 370
1 420	1 250	1 510	1 260	1 330	1 240	1 390
1 410	1 340	1 440	1 420	1 340	1 470	1 430
1 460	1 390	1 430	1 360	1 430	1 450	1 460
1 210	1 410	1 400	1 420	1 310	1 470	1 260
1 360	1 400	1 300	1 430	1 350	1 370	1 270
1 390	1 490	1 350	1 410	1 370	1 440	1 320
1 430	1 390	1 380	1 440	1 450	1 340	1 380
1 490	1 310	1 450	1 280	1 530	1 400	1 320
1 340	1 440	1 350	1 470	1 450	1 370	1 320
1 270	1 390	1 300	1 415	1 620	1 510	1 390
1 320	1 470	1 360	1 580	1 230	1 390	1 570
1 440	1 500	1 515	1 510	1 460	1 380	1 420
1 450	1 330	1 400	1 425	1 410	1 330	1 500
1 390	1 450	1 500	1 485	1 420	1 380	1 410
1 590	1 550	1 350	1 220	1 340		
1 450	1 490	1 300	1 300	1 510		
1 400	1 480	1 500	1 540	1 340		
1 470	1 340	1 380	1 390	1 290		
1 400	1 440	1 280	1 480	1 400		

собой значения признака. Техника укрупненной группировки такова: отмечаем в табл. 4 наименьшее и наибольшее из встретившихся значений — числа 1120 и 1660. Этими двумя числами определяется промежуток вариации признака. Последний подразделяется на равные интервалы, число которых зависит от объема совокупности. Число интервалов не должно быть чрезмерно большим, так как в каждом интервале окажется слишком мало наблюдений для того, чтобы закономерность отчетливо проявилась; оно не должно быть и чрезмерно малым, дабы не утратить существенных подробностей распределения. При числе наблюдений от 100 до 500 рекомендуется делить промежуток вариации на 8—16 интервалов. В нашем случае, при числе наблюдений 200, мы разбили промежуток вариации на 11 интервалов; так как длина промежутка $1660 - 1120 = 540$, то длина каждого интервала должна приблизительно составить 50 единиц. При определении границ интервалов следует начинать ряд, отступая примерно на $\frac{1}{2}$ интервала влево от наименьшего встретившегося значения (1120), и заканчивать ряд, отступая примерно на $\frac{1}{2}$ интервала вправо от наибольшего встретившегося значения (1660). Таким образом, в итоге мы получаем одним интервалом больше против ранее намеченных 11.

При этом необходимо иметь в виду, что величина и границы интервалов должны быть согласованы с характером округления вариантов в процессе измерения признака. Наиболее простое правило согласования состоит в следующем: *границы интервалов группировки должны совмещаться с границами интервалов округления; следовательно, в интервале группировки должно заключаться целое число интервалов округления: $i = kj$* , где i — величина интервала группировки; j — величина интервала округления; k — целое число.

При соблюдении этого правила каждый интервал округления целиком относится к определенному интервалу группировки, и мы избегаем, таким образом, необходимости распределять его частоту между двумя смежными интервалами группировки (рис. 1,а).

Если не придерживаться этого правила, то частоты некоторых интервалов округления придется искусственным приемом разбивать на две части, распределяя их между смежными интервалами группировки (рис. 1,б).

В нашем примере величина интервалов округления равна 10 единицам (в виде исключения несколько вариантов таблицы округлены до 5 единиц, так как в процессе измерения было трудно решить, к какому десятку их следует отнести). Числа с конечной цифрой 5 являются границами интервалов округления, следовательно, границы интервалов группировки также должны выражаться числами, оканчивающимися цифрой 5.

Следует также обратить внимание на то, чтобы вблизи границ интервалов не стояли «круглые» числа (например, в нашем

Таблица 5

Распределение поршневых колец по упругости

Интервалы (в г/мм ²)		Частоты
1085—1135		1
1135—1185		0
1185—1235	□	4
1235—1285	□ □	9
1285—1335	□ □ □ □	18
1335—1385	□ □ □ □ □ □ □ □	40
1385—1435	□ □ □ □ □ □ □ □ □ □	49
1435—1485	□ □ □ □ □ □ □ □ □ □ □	43
1485—1535	□ □ □ □ □ □ □ □	26
1535—1585	□ □ □ □	8
1585—1635		1
1635—1685		1

200

черточек; в конечном подсчете половина таких кружков относится к верхнему, другая половина — к нижнему интервалу.

В нашем случае встретился один кружок (вариант 1485), который мы отнесли к интервалу 1435—1485. Столбец 2 табл. 5 содержит результаты разнесения черточек, в столбце 3 подведены итоги, т. е. даны абсолютные частоты интервалов.

В полученном ряде распределения мы также наблюдаем закономерность, до известной степени сходную с закономерностью предыдущего ряда: по мере увеличения вариантов ряда частота сначала возрастает, достигает в некотором интервале максимума, а затем постепенно уменьшается.

Интервальный ряд табл. 5 может быть представлен в дискретной форме (табл. 6); для этой цели варианты, входящие в любой интервал, заменяются, в порядке округления, средним значением (центром) интервала.

Пример 3. В машиностроении при контроле геометрической формы изделий мы встречаемся с параметрами, представляющими собой отклонения от правильной формы, принятой за нуль на шкале отсчета, выраженные положительными числами. Примером таких параметров, которые носят название существенно положительных величин, являются биения, эксцентриситеты, непараллельности, перпендикулярности и т. д.

Приводим ряд распределения 175 шестерен по одному из таких параметров — биению торца относительно вала после операции шлифования.

Заметим, что нарастание частот в этом ряду происходит на меньшем протяжении и более быстро, чем их убывание после максимальной частоты.

Мы говорим, что ряд обладает правосторонней (положительной) асимметрией (см. рис. 11).

Таблица 6 Распределение 200 колец по упругости (дискретная форма ряда)		Таблица 7 Распределение 175 шестерен по биению торца	
Варианты упру- ги (в $г/мм^2$)	Частоты	Биение (в $мм$)	Число деталей
1110	1		
1160	0	0,01	35
1210	4	0,02	57
1260	9		
1310	18	0,03	45
1360	40	0,04	22
1410	49		
1460	43	0,05	8
1510	26	0,06	4
1560	8	0,07	4
1610	1		
1660	1		
Итого . . .	200	Итого . . .	175

Этой особенности ряда можно дать следующее объяснение. Очевидно, что производственный процесс должен быть направлен на получение правильной формы, которой соответствует нулевое значение параметра. Поэтому наиболее часто встречаются значения параметра 0,02, сравнительно мало отклоняющиеся от нуля.

Так как вариация погрешностей ограничена слева близким нулем, то естественно, что вправо от наиболее частого значения погрешность варьирует более значительно. Отсюда явление правосторонней асимметрии, характерное для рядов распределения существенно положительных величин.

Явление асимметрии — лево-или правосторонней — в более или менее сильной степени наблюдается во многих рядах распределения. При очень сильной асимметрии целесообразна группировка ряда в неравных интервалах, так как ряд нуждается в более подробной дифференциации одной его части и менее подробной — другой.

§ 2. ГРАФИЧЕСКОЕ ИЗОБРАЖЕНИЕ РЯДОВ РАСПРЕДЕЛЕНИЯ

Наряду с табличной формой выражения числовых данных в математической статистике широко используются и другие формы — графическая и аналитическая. С аналитическим выражением рядов распределения, т. е. выражением их в виде математических формул, мы познакомимся в разделе III книги. В настоящем параграфе мы рассматриваем способы графического изображения рядов распределения.

Построение графиков для дискретных и интервальных рядов распределения имеет свои особенности. Рассмотрим их раздельно.

Графическое изображение дискретных рядов

В качестве примера возьмем распределение 190 деталей по размеру диаметра (табл. 2). Строим оси координат (рис. 2).

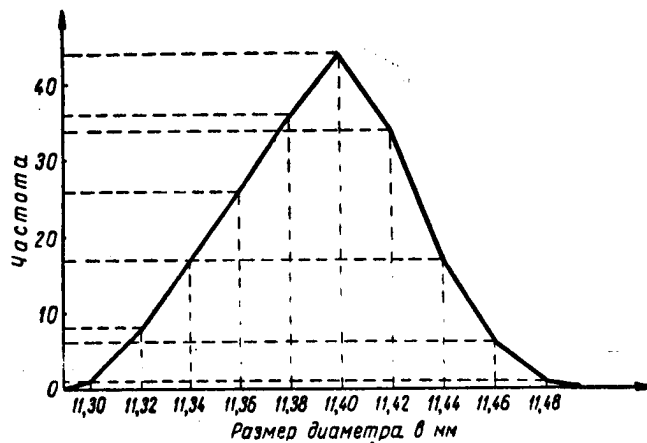


Рис. 2. Полигон распределения 190 деталей по размеру диаметра

По оси абсцисс откладываем значения признака (значения диаметра), по оси ординат — частоты. Из каждой точки на оси абсцисс, изображающей вариант признака, восстанавливаем перпендикуляр, длина которого выражает соответствующую частоту. Вершины перпендикуляров соединяем прямыми линиями. Полученная фигура называется *полигоном (многоугольником) распределения*.

Итак, статистический смысл элементов полигона: точки на оси абсцисс изображают варианты признака, ординаты — частоты.

Графическое изображение интервальных рядов

В качестве примера возьмем ряд распределения 200 экземпляров поршневых колец по величине упругости (табл. 6).

Строим оси координат. По оси абсцисс откладываем интервалы значений признака (упругости). Частоты изображаются в виде площадей прямоугольников, построенных на каждом интер-

вале (рис. 3). Для этой цели на оси ординат откладываем масштаб частоты; так как основания прямоугольников равны между собой, то их площади пропорциональны высотам. Полученная фигура называется *гистограммой распределения*.

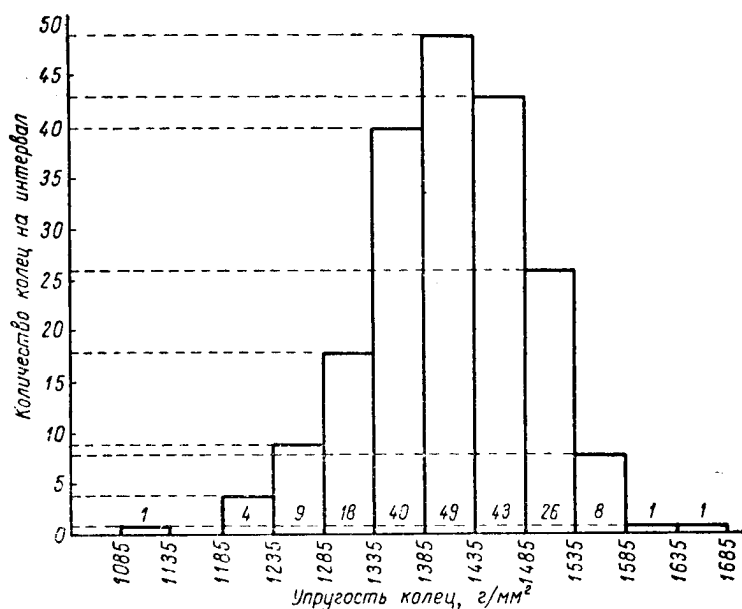


Рис. 3. Гистограмма распределения 200 поршневых колец по величине упругости

Итак, статистический смысл элементов гистограммы: точки на оси абсцисс изображают варианты признака, площади прямоугольников — частоты; ордината представляет собой отношение площади к длине основания — это частота, приходящаяся на единицу измерения признака. Она носит название *плотности частоты*.

Если за единицу измерения признака принять длину интервала, то в случае равенства интервалов плотности частоты количественно совпадают со значениями частот, как это имеет место на рис. 3. В случае неравных интервалов плотности частот количественно не совпадают с частотами.

Способ построения ряда в виде гистограммы основан на предположении, что плотность частоты остается постоянной внутри каждого интервала и меняется скачками на краях интервалов. Это условное допущение, очевидно, в известной мере искажает закономерность, лежащую в основе изучаемого статистического материала, искажает тем сильнее, чем больше длина интервала. Более близким к действительности является предполо-

жение о равномерном изменении плотности частоты от интервала к интервалу. Это допущение приводит нас к изображению интервальных рядов в виде полигонов распределения, которые построим следующим образом (пример 2). Из середины каждого интервала восстанавливаем ординату, пропорциональную *плот-*

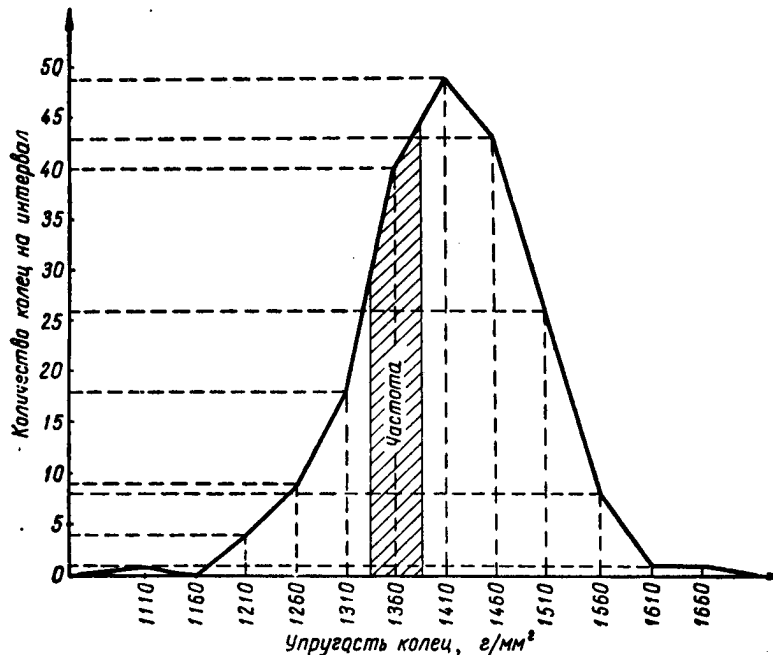


Рис. 4. Полигон распределения для интервального ряда

ности частоты. Вершины перпендикуляров соединяем прямыми линиями. Получается фигура, по внешнему виду сходная с полигоном, изображающим дискретный ряд распределения.

Однако статистический смысл элементов этого полигона иной. Ордината здесь означает не частоту, а плотность частоты. Частота же интервала изображается в виде площади криволинейной трапеции, ограниченной сверху отрезком полигона, снизу — осью абсцисс и по сторонам — двумя ординатами (рис. 4).

Полигоны, изображающие дискретные ряды распределения, называются *полигонами с нагруженными ординатами*.

Еще более точное представление о закономерности, лежащей в основе изучаемого статистического материала, дает нам третий способ изображения интервальных рядов в виде плавной кривой, передающей ход изменения плотности распределения (см. рис. 5). Построение таких кривых требует специальных математических расчетов, с которыми мы познакомимся в разделе III книги.

Лишь первый из трех способов изображения интервальных рядов (способ гистограмм) сохраняет частоты интервалов такими, какими они даны в таблице. В способе полигонов и в способе плавных кривых частоты отдельных интервалов, вычисляемые как площади криволинейных трапеций, будут несколько

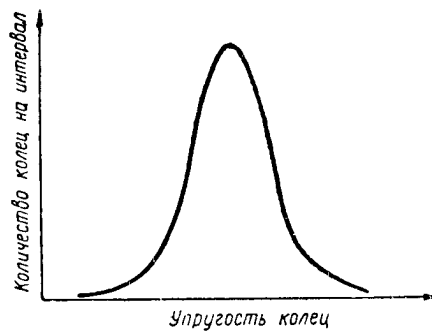


Рис. 5. Кривая распределения

отличаться от частот, заданных в таблице, «исправляя» последние. Общая же сумма всех частот остается при всех трех способах неизменной, равной объему совокупности.

ГЛАВА II

АНАЛИЗ РЯДОВ РАСПРЕДЕЛЕНИЯ

§ 1. ПРОБЛЕМА ИЗМЕРЕНИЯ МАССОВОГО УРОВНЯ

Анализ рядов распределения мы свяжем в дальнейшем с чтением графиков распределения. Уметь читать график — это значит уметь выделить в нем наиболее существенные особенности, последовательно переходя от более важных свойств к менее важным и отвлекаясь от второстепенных деталей.

В ряде проблем, которые решаются при помощи рядов распределения, основное место занимает проблема измерения массового уровня признака. С ней мы встречаемся при любом сравнении двух или нескольких совокупностей в статике или динамике. Сопоставляя, например, качество продукции в двух различных партиях, в первую очередь выясняют, в какой из двух партий уровень качества выше. Следя за изменением производительности труда рабочих какой-либо профессиональной группы на предприятии, устанавливают, повысился ли общий уровень производительности труда этих рабочих. Трудность вопроса состоит в том, что отдельные образцы продукции первой партии лучше, другие хуже, чем отдельные образцы продукции второй партии, что отдельные рабочие профессиональной группы неодинаково повысили производительность труда, и состав этой группы также мог измениться. Между тем ответ на поставленный вопрос требуется дать в форме, обобщающей данные всех элементов совокупности: речь идет не об измерении уровня признака для того или иного частного элемента совокупности, а об измерении уровня, характеризующего всю массу элементов в целом.

Выразим это графически. На рис. 6 сопоставляются кривые распределения по производственному стажу двух совокупностей рабочих. В какой совокупности уровень производственного стажа выше? Очевидно, стаж выше во второй совокупности, так как ее кривая распределения расположена правее и соответствует

более высоким значениям стажа, чем кривая распределения первой совокупности.

Итак, положение графика распределения на оси абсцисс является показателем высоты массового уровня признака. Чем правее расположен график на оси абсцисс, тем выше массовый уровень признака.

Мало, однако, установить, что уровень стажа в одной из совокупностей выше, чем в другой. Вопрос обычно ставится в количественной форме: на сколько лет стаж во второй совокупности выше, на сколько единиц (лет) смещена вторая кривая по сравнению с первой?

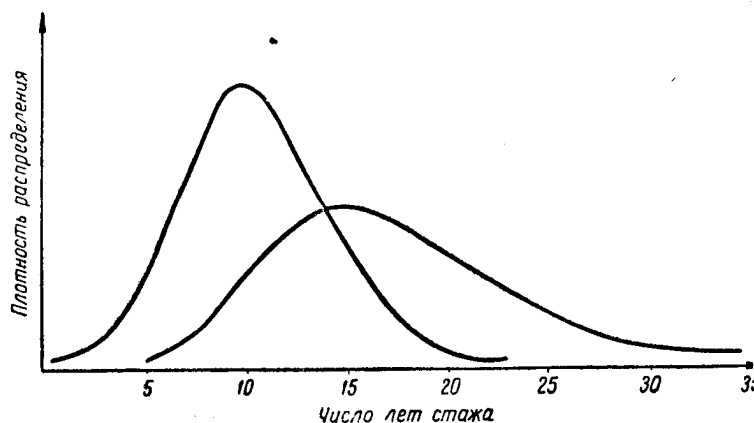


Рис. 6. Сравнение двух групп рабочих по производственному стажу

О смещении графика мы можем судить по смещению каких-либо точек, характеризующих положение графика на оси абсцисс. На первый взгляд это могут быть точки начала или конца графика. Так, например, если судить по положению начальных точек, то второй график смещен относительно первого на 5 единиц (лет). Если же судить по положению конечных точек, то размер смещения оценим в 12 лет.

Однако ни первый, ни второй способ не дает удовлетворительных результатов. Начало и конец графика — случайные точки, положение которых не характеризует положения всей кривой распределения. О последнем можно судить лишь по точкам, вокруг которых концентрируется основная масса наблюдений. Такие точки обычно расположены в центральной части графика, поэтому они называются средними точками, или *средними величинами*. Средние величины характеризуют положение графика распределения на оси абсцисс.

§ 2. СРЕДНЯЯ АРИФМЕТИЧЕСКАЯ

Наиболее известной в ряду средних величин является *средняя арифметическая*, которой называется сумма значений признака, разделенная на число этих значений

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n}$$

или, короче,

$$\bar{x} = \frac{\Sigma x}{n}, \quad (2.2.1)$$

где \bar{x} — обозначение средней арифметической.

Пусть, например, заработок четырех рабочих составляет 812, 980, 828, 892 руб. Средний заработок по формуле средней арифметической

$$\bar{x} = \frac{812 + 980 + 828 + 892}{4} = 878.$$

В случае, когда отдельные варианты признака повторяются по несколько раз, частота каждого варианта должна быть учтена при вычислении средней арифметической, и расчет производится по формуле

$$\bar{x} = \frac{\Sigma xt}{\Sigma t}, \quad (2.2.2)$$

где t — частоты вариантов x .

Например, для дискретного ряда распределения рабочих по зарплате:

Таблица 8

Зарплата рабочих x	Число рабочих t	xt
500	10	5 000
600	20	12 000
700	40	28 000
800	30	24 000
Итого . . .	100	69 000

Каждый вариант заработка умножается на свою частоту (действие производим в столбце 3 таблицы), и сумма произведений делится на число рабочих

$$\bar{x} = \frac{69\,000}{100} = 690 \text{ руб.}$$

Формула (2.2.1) дает простую среднюю арифметическую, формула (2.2.2) — взвешенную среднюю арифметическую.

Числа t в формуле (2.2.2.) называются весами вариантов. Чем больше вес, с которым входит какой-либо вариант в формулу средней арифметической, тем сильнее влияние этого варианта на окончательный результат.

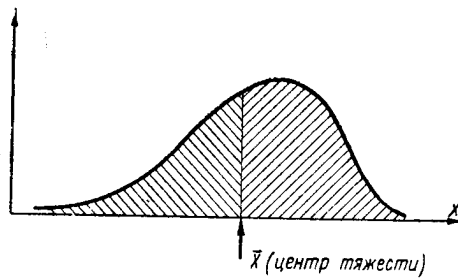


Рис. 7. Средняя арифметическая

Для интервальных рядов распределения расчет средней арифметической производится по взвешенной формуле, при этом интервалы заменяются своими центрами (т. е. ряд превращается в дискретный). Пример расчета среднего заработка для такого ряда:

Таблица 9

Центр интервала x	Зарплата (в руб.)	Число рабочих t	xt
650	600—700	200	130 000
750	700—800	500	375 000
850	800—900	300	255 000
		1 000	760 000

$$\bar{x} = \frac{760\,000}{1000} = 760 \text{ руб.}$$

В табл. 9 столбцы 2 и 3 содержат исходные данные, столбцы 1 и 4 составлены в процессе вычисления средней. Графическое истолкование средней арифметической уясняется из ее взвешенной формулы

$$\bar{x} = \frac{\sum xt}{\sum t}$$

Этой формулой определяется в механике центр тяжести масс. Вообразим себе ось абсцисс в виде балки, несущей нагрузку в форме графика распределения (рис. 7).

Точка на оси x , в которой давлением снизу уравновешивается вся нагрузка, является средней арифметической. Таким образом, средняя арифметическая представляет собой абсциссу центра тяжести графика распределения.

§ 3. МЕДИАНА

Другой средней величиной, характеризующей положение графика на оси абсцисс, является медиана. Медианой называется величина изучаемого признака, которая обладает следующим свойством: число элементов совокупности со значениями данного признака, большими этой величины, равно числу элементов со значениями признака, меньшими ее.

Графически медиану можно определить как точку на оси абсцисс, в которой ордината делит площадь графика распределения на две равные части. Это определение медианы пригодно

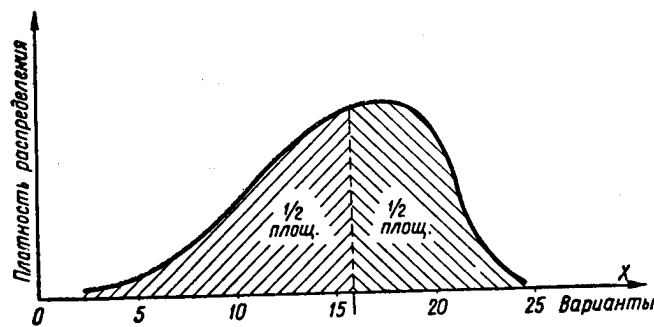


Рис. 8. Медиана

в тех случаях, когда площадь графика изображает частоту (в гистограммах, полигонах второго типа, плавных кривых).

При расчете медианы различаем два случая:

- 1) для несгруппированных данных,
- 2) для рядов распределения.

Расчет медианы для несгруппированных данных

Пример. Приводим результаты замеров диаметра девяти деталей (отклонения от номинала в микронах): 22, 20, 17, 21, 16, 23, 20, 25. Найти медиану.

Прежде всего располагаем эти данные в порядке возрастания: 16, 17, 20, 20, 21, 22, 23, 24, 25. Диаметр детали, занимающей в этом ряду центральное, 5-е место, оказался равным 21 мк, следовательно, медиана — 21.

При четном количестве данных на центральных местах оказывается два элемента; в качестве медианы в этом случае принимают полусумму значений признака у этих элементов.

Например, при десяти замерах диаметра — 16, 17, 20, 20, 21, 22, 23, 24, 25, 27 — медиана: $\frac{21 + 22}{2} = 21,5$.

Вычисление медианы по этому способу целесообразно производить при небольшом количестве данных. В случае большого

количества данных следует применить группировку и произвести расчет медианы для ряда распределения.

Расчет медианы для интервальных рядов распределения

Возьмем ряд распределения примера 2 и поясним способ расчета медианы на графике распределения (гистограмме).

Прежде всего определяем интервал, в котором находится медиана.

Для этой цели находим $\frac{n}{2} = 107$ и накапливаем частоты от начала графика, пока не дойдем до $\frac{n}{2}$ или ближайшего к нему меньшего итога частот. Получаем $1+0+4+9+18+40=72$.

Этим итогом определяется медианный интервал (1385—1435).

Следовательно, медиана равна $1385+z$, где z — неизвестная нам часть медианного интервала. Нам известна площадь прямоугольника, лежащего на всем медианном интервале, — 49 и площадь прямоугольника, лежащего на отрезке z (заштрихована на рис. 9): $100 - 72 = 28$.

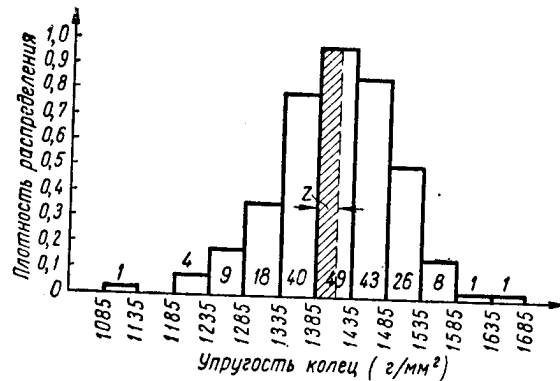


Рис. 9. Расчет медианы

z определяем из пропорции

$$\frac{z}{50} = \frac{28}{49}; \text{ откуда } z = 28,6$$

и медиана равна $1385 + 28,6 = 1413,6$.

Процесс вычисления медианы может быть записан готовой формулой

$$M_e = x_s + i \frac{\frac{n}{2} - \Sigma}{m_e}, \quad (2.3.1)$$

где M_e — обозначение медианы;
 x_e — начало медианного интервала;
 i — длина интервала;
 n — объем совокупности;
 Σ — сумма частот всех интервалов, предшествующих медианному;
 m_e — частота медианного интервала.
 В нашем примере

$$x_e = 1385, i = 50, n = 200, \Sigma = 72, m_e = 49.$$

$$M_e = 1385 + 50 \frac{100 - 72}{49} = 1413,6.$$

Порядок расчета медианы по формуле (2.3.1): сначала определяем $\frac{n}{2}$, затем Σ и, наконец, все остальные величины — x_e, m_e .

Формула (2.3.1) пригодна и в случае рядов распределения с неравными интервалами, i в ней тогда обозначает длину *медианного* интервала.

При вычислении медианы для дискретных рядов распределения их необходимо сперва представить в интервальной форме и затем использовать изложенный выше метод. Так, для примера 1 мы рассчитаем медиану по формуле (2.3.1.), если предварительно представим ряд в интервальной форме (см. табл. 3)

$$M_e = 39 + 2 \frac{95 - 88}{44} = 39 + 0,32 = 39,32.$$

§ 4. МОДА

Модой называется абсцисса точки максимума кривой распределения. Это графическое определение моды. Ее статистический смысл формулируется несколько иначе для дискретных и интервальных рядов распределения.

Для дискретных рядов распределения модой является то значение признака, которое чаще всего встречается. Для интервальных рядов модой является то значение признака, на которое приходится наибольшая плотность распределения (т. е. вокруг которого концентрируется наибольшая частота).

Для дискретных рядов мода находится без вычислений, на глаз, непосредственно по таблице или по графику распределения. Так, в примере 1 модой служит диаметр 40, так как он встречается чаще всего.

Для интервальных рядов нахождение моды требует расчета. Рассмотрим интервальный ряд примера 2. На глаз видно, что мода заключена в интервале (1385—1435), так как на этот интервал приходится наибольшая плотность распределения. Но ка-

кую точку этого интервала следует считать модой? Для ответа на поставленный вопрос следовало бы изобразить ряд распределения в виде плавной кривой и найти точку максимума этой кривой. Подобный путь, однако, требует длительных расчетов и потому непрактичен.

Приведем без доказательства формулу, по которой можно приближенно рассчитать моду, т. е. абсциссу точки максимума кривой распределения, не строя самой кривой

$$M_0 = x_0 + i \frac{\Delta_1}{\Delta_1 + \Delta_2}, \quad (2.4.1)^1$$

где M_0 — обозначение моды;

x_0 — начало модального интервала;

i — длина интервала;

Δ_1 — разность частот модального и домодального интервалов;

Δ_2 — разность частот модального и послемодального интервалов.

В рассматриваемом примере $x_0 = 1385$, $i = 50$, $\Delta_1 = 49 - 40 = 9$, $\Delta_2 = 49 - 43 = 6$.

$$M_0 = 1385 + 50 \frac{9}{9 + 6} = 1415.$$

Формула (2.4.1) пригодна для рядов распределения с равными интервалами. Заметим, что при вычислении моды для дискретных рядов последние часто приводят к интервальной форме и пользуются формулой (2.4.1). Этим путем достигается большая точность в определении моды.

Так, превратив дискретный ряд примера 1 в интервальный (см. табл. 3), мы вычисляем моду по формуле (2.4.1)

$$M_0 = 39 + 2 \frac{44 - 36}{(44 - 36) + (44 - 34)} = 39 + 0,89 = 39,89.$$

На рис. 6 моды двух графиков отстоят друг от друга на 6 лет. Эта величина приближенно характеризует различие в среднем уровне производственного стажа у двух групп рабочих.

§ 5. СОПОСТАВЛЕНИЕ СРЕДНЕЙ АРИФМЕТИЧЕСКОЙ, МЕДИАНЫ И МОДЫ

В настоящем параграфе рассмотрим вопросы количественного сопоставления трех средних и их практического применения.

Прежде всего выясним, в каких случаях мода, медиана и средняя арифметическая дают близкие значения и от чего за-

¹ Вывод формулы см., например [5] или [10] списка литературы в конце книги.

висит расхождение между этими средними. Рассмотрим график распределения симметричной формы.

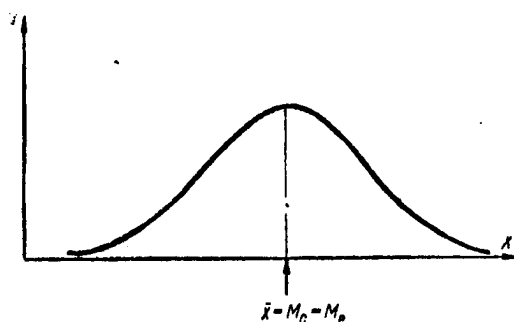


Рис. 10. Симметричная кривая

Очевидно, в таком графике точка максимума, линия, равноделящая площадь, и центр тяжести лежат на оси симметрии. Следовательно, для симметричного графика мода, медиана и средняя арифметическая совпадают.

Пусть теперь график имеет правостороннюю асимметрию («хвост» вправо). Покажем, что в таком случае мода расположена левее, а средняя арифметическая — правее медианы¹.

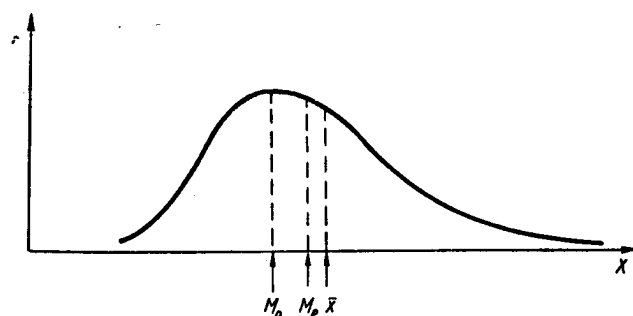


Рис. 11. Правосторонняя асимметрия

В самом деле, слева и справа от медианы расположены одинаковые площади (массы), но левая площадь поставлена на коротком, а правая на длинном основании. Поэтому левая часть кривой выше правой, и моду мы найдем слева от медианы. По этой же причине медиана не может служить точкой равновесия, ибо площадь, расположенная на длинном плече, перетянет площадь, расположенную на коротком; следовательно, средняя арифметическая расположена правее медианы.

Итак, при правосторонней асимметрии левее расположена

¹ Встречаются распределения «неправильной» формы, для которых наше рассуждение неприменимо.

мода, затем медиана и правее — средняя арифметическая. Обратное расположение имеет место при левосторонней асимметрии графика. Таким образом, расхождение между тремя средними зависит от асимметрии графика: чем более асимметричен график, тем больше расстояние между его средними точками.

Переходим к вопросам практического применения этих средних. Наиболее известной, имеющей наибольшее приложение к практике является средняя арифметическая.

Популярность средней арифметической складывается под влиянием различных причин. Средняя арифметическая имеет простой смысл отношения двух суммарных величин. Эти суммарные величины в ряде случаев имеют существенное значение для характеристики рассматриваемых явлений и совместно со средней арифметической образуют систему показателей, взаимно-увязанных друг с другом. Так, среднюю зарплату мы вычисляем по формуле: средней арифметической из заработков отдельных рабочих. В числителе средней — фонд зарплаты, в знаменателе — число рабочих. Эти величины увязаны в формуле средней арифметической простым соотношением, из которого, зная два показателя, можно определить третий. Так, в практике планирования по числу рабочих и средней зарплате определяют фонд зарплаты как произведение первых двух. Ни мода, ни медиана не в состоянии обеспечить подобной увязки между показателями. Поэтому в данном вопросе следует предпочесть среднюю арифметическую.

Аналогичными соображениями руководствуются при вычислении средней производительности труда, средней себестоимости, а также средней скорости и ряда других экономических и физических средних показателей. Эти соображения, однако, несостоятельны, когда относимые друг к другу суммы сами по себе не представляют интереса. Так, вычисляя средний рост по средней арифметической, мы в числителе получаем сумму всех значений роста, в знаменателе — число людей. Но первая сумма сама по себе никакого интереса не представляет. Поэтому вычисление среднего роста по формуле средней арифметической подобными соображениями обосновать нельзя, и если оно все же часто применяется, то по иным основаниям.

Другой причиной широкого использования средней арифметической является ее роль в теории точных измерений. В случае необходимости точного определения какой-либо физической величины (длины, веса, угла и т. д.) обычно не ограничиваются однократным измерением, а повторяют измерения несколько раз для того, чтобы исключить неизбежно сопутствующие им случайные погрешности. Теория точных измерений считает, что, вычисляя среднюю арифметическую из результатов многих измерений, каждый из которых осложнен случайной ошибкой, мы получаем наиболее правильное представление об измеряемой величине.

Далее следует указать, что математическая статистика при более глубоком анализе материала использует систему показате-

лей, объединенных общим методом — *методом наименьших квадратов*; первым в ряду этих показателей является средняя арифметическая. Эта увязка средней арифметической с другими важнейшими статистическими показателями (мы познакомимся с ними в дальнейшем) также является одной из причин частого использования средней арифметической. Мы указали некоторые основные причины, которые выдвигают среднюю арифметическую на первое место в ряду средних показателей, употребляющихся в статистике.

Благодаря такому положению средней арифметической и в тех случаях, когда по существу вопроса безразлично, какую среднюю применить, когда нет определенных указаний на целесообразность применения какой-либо иной средней, употребляют среднюю арифметическую.

Значительно меньшее, но достаточно существенное применение имеет медиана. Особенность медианы состоит в том, что для ее вычисления не требуется знать величину признака у всех элементов совокупности, а достаточно только уметь расположить элементы в порядке возрастания вариантов. Эта особенность делает медиану полезной в тех случаях, когда точное измерение величины признака для всех элементов затруднительно или невозможно. Так, пользуясь медианой, можно быстро определить средний рост большой группы людей, расположив их по росту на глаз и измеряя рост только того, кто оказался в центре шеренги.

Широко используется медиана в практике статистического контроля качества продукции. При оперативном статистическом контроле периодически через равные промежутки времени берутся пробы, состоящие из нескольких экземпляров продукции. Результаты замеров отмечаются точками на контрольной карте, где по оси абсцисс дано время, по оси ординат — шкала измерения параметра. Так как результаты замеров оказываются на шкале упорядоченными по величине параметра, то нахождение медианы для каждой пробы производится на глаз, без всяких расчетов. Отмечая значения медиан на контрольной карте условными значками и следя за изменением положения медиан на карте, рабочий получает возможность вовремя обнаружить начинающуюся разладку процесса и предотвратить появление брака.

В экономической литературе часто встречаются ряды распределения, в которых крайние интервалы не ограничены. Например, табл. 10 содержит распределение числа рабочих фасонно-литейного цеха за март 1955 г. по проценту выполнения норм.

В таблице последний интервал не ограничен (впрочем и для первого интервала граница 0 является заниженной), поэтому вычисление среднего процента выполнения норм по формуле средней арифметической здесь невозможно. Для такого подсчета пришлось бы либо искусственно приравнять последний интервал предыдущему (т. е. считать 200—250), что иногда применяется,

но ведет к значительным искажениям, если в последнем интервале сконцентрирована большая частота (это как раз имеет место в нашем примере), либо пришлось бы применить другие, более сложные приемы определения длины последнего интер-

Таблица 10

Процент выполнения норм	Число рабочих
До 100	3
100—120	209
120—150	156
150—200	180
Свыше 200	48
Итого . . .	596

вала, основанные на экстраполяции (искусственном продолжении) кривой распределения. Между тем наглядно видно, что медиана содержится в интервале (120—150) и расчет ее по формуле (2.3.1) не представляет затруднений.

Итак, в рядах распределения с неопределенными границами вычисление медианы в большинстве случаев решает вопрос о среднем уровне.

Нередки также случаи, когда распределение содержит отдельные варианты, далеко отстоящие от основной массы случаев. Такое явление наблюдается, например, при определении норм расхода времени на ту или иную производственную операцию и во многих других вопросах. Вычисление средней арифметической может тогда привести к ненадежным результатам, так как выдающиеся варианты оказывают сильное и в значительной степени случайное влияние на результат. Исключение таких «сомнительных» вариантов из подсчета, которое иногда применяется на практике, приводит к искажениям. Медиана же, очевидно, не зависит от того, как далеко отстоят от основной массы случаев исключительные варианты. Поэтому вычисление медианы заслуживает в таких случаях предпочтения.

Далеко не все количественные признаки измеряются твердой и общепризнанной единицей измерения, т. е. являются величинами в строгом смысле слова. Балл, которым оценивается величина зерна микроструктуры металла, или балл неметаллических включений в стали, разряд рабочего в тарифной сетке и т. д. служат примерами таких признаков.

Задача таких признаков скорее в установлении порядка, но не в точном измерении. Мы не можем, например, утверждать, что при оценке величины зерна балл 5 так же отличается от 4, как 3 от 2. Мы говорим только, что $5 > 4 > 3 > 2$. Единица измерения признака здесь не эквивалентна самой себе на разных участках шкалы. Для таких признаков медиана в некоторых отношениях имеет более определенный смысл, чем средняя арифметическая. Так, если несколько изменить систему эталонов, которыми определяется балл таким образом, что балл 4 приблизится по своему реальному значению к баллу 5 и отдалится от балла 3 и т. д., то образец, занимавший место медианы в старой

системе эталонов, сохранит это место в измененной системе; в то же время образец, который занимал место средней арифметической ряда, окажется теперь выше или ниже среднего.

Резюмируем: медиана — средняя, основанная не на величине, а на порядке вариантов. Поэтому в тех случаях, где мы не уверены в величине, но гарантируем порядок вариантов, следует прибегать к медиане.

Мода, как наиболее часто встречающаяся величина, всегда представляет самостоятельный интерес при исследовании материала независимо от вычисления других средних. Вычисление моды особенно существенно в несимметричных рядах распределения, когда она может сильно отличаться от других средних.

§ 6. ПРОБЛЕМА ИЗМЕРЕНИЯ КОЛЕБЛЕМОСТИ. СРЕДНЕЕ АБСОЛЮТНОЕ ОТКЛОНЕНИЕ, СРЕДНЕЕ КВАДРАТИЧЕСКОЕ ОТКЛОНЕНИЕ, РАЗМАХ И КОЭФФИЦИЕНТ ВАРИАЦИИ

Второй основной проблемой, которая решается с помощью рядов распределения, является проблема измерения колеблемости признака. Поясним ее содержание на примере. Показатели выполнения плана для двух групп предприятий выражаются следующими числами: в первой группе — 80, 90, 100, 100, 110 и 120%; во второй группе — 95, 97, 100, 100, 102 и 106%. Мы видим, что эти показатели в каждой группе колеблются вокруг одного и того же уровня 100%.

Однако в то время как во второй группе отдельные показатели весьма близко примыкают к среднему уровню, показатели первой группы отличаются сильной изменчивостью или колеблемостью. Это явление — большей или меньшей колеблемости вариантов вокруг среднего уровня — имеет основное значение во всех приложениях статистического метода и в его теории.

Приведем примеры:

1. Перед нашими промышленными предприятиями стоит важная задача достижения равномерного, ритмичного хода выполнения плана по выпуску продукции. Контроль над ее осуществлением требует расчета показателей, способных измерять степень нарушения на тех или иных участках производства хода равномерного выполнения плана, — показателей колеблемости. Построение таких показателей является задачей промышленной статистики.

2. Выпуская тот или иной вид продукции, предприятие обязано строго соблюдать установленные стандарты качества продукции. Соблюдать стандарты — это значит не только поддерживать средний уровень качества продукции на соответствующей высоте, но и следить за тем, чтобы колеблемость отдельных образцов продукции по тем или иным важным признакам не превосходила установленных пределов. Контроль над выполнением этой задачи требует расчета специальных показателей колеблемости качества продукции.

3. Следя за процессами роста производительности труда на предприятии, мы интересуемся не только движением среднего уровня выработки рабочих, но и показателями, характеризующими колеблемость рабочих по выработке. Достижения отдельных передовиков производства повышают средний уровень выработки и одновременно увеличивают колеблемость. Подтягивание основной массы к передовым рекордам увеличивает средний уровень и одновременно сокращает колеблемость.

Различия, существующие между рабочими по производительности труда, связаны с их колеблемостью по квалификации, производственному стажу и другим факторам производительности труда. Построение показателей колеблемости для варьирующих признаков является важнейшей задачей статистики.

На рис. 6 приведены кривые распределения по производственному стажу двух групп рабочих.

В какой группе колеблемость рабочих по стажу больше? Кривая рабочих второй группы имеет большее протяжение на оси абсцисс, она захватывает большее количество *различных* вариантов, чем кривая рабочих первой группы. Если размах рассеяния первой кривой от 0—23 лет, то размах рассеяния второй кривой от 5—35 лет. Отсюда ясно, что колеблемость данного признака больше во второй группе рабочих, чем в первой, так как ее кривая распределения более рассеяна по оси абсцисс.

Итак, *рассеяние кривой распределения по оси абсцисс является показателем колеблемости признака: чем более рассеяна кривая, тем больше колеблемость признака.*

Мы видели ранее, что первым и основным свойством графика распределения является его *положение* на оси абсцисс. Для точной оценки этого положения рассчитываются особые показатели — средние величины, например: средняя арифметическая, мода, медиана.

Вторым основным свойством графика распределения является его *рассеяние* по оси абсцисс. Покажем, как рассчитываются показатели рассеяния. Пусть месячный заработок шести рабочих выразился в сумме (столбец 2 табл. 11).

Таблица 11

Рабочий	Зарботок (в руб.) x	$x - \bar{x}$	$ x - \bar{x} $
1	800	0	0
2	800	0	0
3	760	-40	40
4	840	40	40
5	780	-20	20
6	820	+20	20
Итого . . .	4,800	0	120

Средний заработок равен $\bar{x} = \frac{4800}{6} = 800$ руб.

Зарплаты отдельных рабочих различаются между собой — колеблются. О величине этой колеблемости можно судить по отклонениям отдельных заработков от среднего. Соответствующие отклонения вычислены в столбце 3 табл. 11. Чтобы получить показатель колеблемости, естественно взять среднюю арифметическую из *абсолютных* величин этих отклонений (очевидно, нас интересует здесь абсолютная, а не алгебраическая величина отклонения; алгебраическая сумма отклонений от средней арифметической, как мы докажем далее, всегда равна нулю).

В столбце 4 таблицы даны абсолютные размеры отклонений, из которых и рассчитывается средняя

$$\Theta = \frac{120}{6} = 20.$$

Полученный показатель колеблемости называется *средним абсолютным отклонением* и обозначается буквой Θ (тэта). Его формула

$$\Theta = \frac{\sum |x - \bar{x}|}{n}. \quad (2.6.1)$$

Показатель Θ имеет размерность ту же, что и варианты признака. Так, в нашем примере $\Theta = 20$ руб. При вычислении столбец 3 в табл. 11 можно, разумеется, опустить и писать сразу абсолютные отклонения. Операция отбрасывания знаков при отклонениях имеет свои неудобства. Благодаря искусственному характеру этой операции математические свойства показателя Θ являются весьма сложными, что делает его использование неудобным при более детальном изучении материала. Поэтому в математической статистике отдают предпочтение другому способу устранения знака, а именно возведению каждого отклонения в квадрат. Беря затем среднюю арифметическую из этих квадратов и извлекая из нее квадратный корень, мы приходим к так называемому *среднему квадратическому отклонению*.

Формула среднего квадратического отклонения

$$\sigma = \sqrt{\frac{\sum (x - \bar{x})^2}{n}}, \quad (2.6.2)$$

где σ (сигма) — знак среднего квадратического отклонения.

Схема вычисления σ дана в табл. 12.

Таблица 12

x	$x - \bar{x}$	$(x - \bar{x})^2$
800	0	0
800	0	0
760	-40	1 600
840	40	1 600
780	-20	400
820	20	400
Итого 4 800	0	4 000

$$\bar{x} = \frac{4800}{6} = 800,$$

$$\sigma = \sqrt{\frac{4000}{6}} = \sqrt{667} = 25,8.$$

σ , как и Θ , имеет ту же размерность, что и варианты признака (в нашем примере $\sigma = 25,8$ руб.).

Преимущества σ перед Θ при математических выкладках состоят в том, что при построении первого используются лишь алгебраические операции: сложение, вычитание, деление, умножение, возведение в степень и извлечение корня. Поэтому математические свойства этого показателя значительно проще¹, что сказывается в ряде вопросов (например, при объединении совокупностей, см. § 9 главы 2). Кроме того, σ , так же как и x , принадлежит к системе показателей, увязанных общим методом — методом наименьших квадратов (§ 8 главы 2).

По этим причинам основным показателем колеблемости в статистике принято считать среднее квадратическое отклонение — σ .

Результаты вычислений в последнем примере могут быть записаны в виде

$$x = \bar{x} \pm \sigma = 800 \pm 25,8 \text{ руб.}$$

Эта запись в сжатом виде передает основные свойства ряда распределения: средний уровень признака и границы средней колеблемости вокруг этого уровня.

По формуле (2.6.2) мы вычисляем σ для *несгруппированных* данных. Она называется *простой*.

При вычислении σ для рядов распределения используется формула *взвешенной* σ

$$\sigma = \sqrt{\frac{\sum (x - \bar{x})^2 m}{\sum m}}, \quad (2.6.3)$$

где m — веса (частоты).

¹ Это, конечно, не значит, что расчет σ проще, чем расчет Θ .

Пример вычисления взвешенной σ для дискретных рядов распределения:

Таблица 13

Зарплата рабочего x	Число рабочих m	mx	$(x-\bar{x})$	$(x-\bar{x})^2$	$(x-\bar{x})^2m$
500	10	5 000	-190	36 100	361 000
600	20	12 000	-90	8 100	162 000
700	40	28 000	10	100	4 000
800	30	24 000	110	12 100	363 000
Итого..	100	69 000			890 000

$$\bar{x} = \frac{69\,000}{100} = 690,$$

$$\sigma = \sqrt{\frac{890\,000}{100}} = \sqrt{8900} = 94,3,$$

$$\bar{x} = 690 \pm 94,3.$$

В случае интервальной формы рядов распределения в вычислениях участвуют центры интервалов (т. е. ряд превращается в дискретный), например:

Таблица 14

Зарплата	Число рабочих m	x	mx	$(x-\bar{x})$	$(x-\bar{x})^2$	$(x-\bar{x})^2m$
600—700	5	650	3 250	-190	36 100	180 500
700—800	10	750	7 500	-90	8 100	81 000
800—900	20	850	17 000	10	100	2 000
900—1000	15	950	14 250	110	12 100	181 500
Итого	50		42 000			445 000

$$\bar{x} = \frac{42\,000}{50} = 840 \text{ руб.},$$

$$\sigma = \sqrt{\frac{445\,000}{50}} = 94,3,$$

$$\bar{x} = 840 \pm 94,3.$$

Графически σ изображается в виде отрезка на оси абсцисс, который откладывается в обе стороны от точки \bar{x} . Этот отрезок можно отложить несколько раз в обе стороны от \bar{x} , и тогда ось абсцисс будет измерена σ как новой единицей измерения.

Заметим, что в теоретических формулах σ часто выступает возведенной в квадрат. По этой причине целесообразно присвоить σ^2 особое наименование — *дисперсия*. Дисперсия, очевидно, также является мерой колеблемости признака. Тот же термин «дисперсия» употребляется, как мы видели ранее, и для обозначения самого свойства колеблемости.

Для грубо ориентировочной оценки величины рассеяния вычисляется *размах*. Размахом называется разность между наибольшим и наименьшим значениями признака в ряду наблюдений. Так, для примера (см. табл. 12) наибольшее значение признака оказалось 840, наименьшее 760, отсюда размах

$$R = 840 - 760 = 80.$$

Расчет размахов благодаря своей простоте получил широкое распространение при предупредительном контроле качества продукции малыми выборками.

Среднее квадратическое и среднее абсолютное отклонения, а также размах являются размерными показателями колеблемости признака. Они выражаются в тех же единицах, что и варианты признака. В ряде случаев полезен расчет безразмерного показателя колеблемости — *коэффициента вариации*. Коэффициентом вариации называется отношение σ к \bar{x} , обычно выражаемое в процентах

$$v = \frac{\sigma}{\bar{x}}, \text{ или } v = 100 \frac{\sigma}{\bar{x}} \%. \quad (2.6.4)$$

Для примера в табл. 14

$$v = \frac{94,3}{840} = 0,112, \quad v = 11,2\%.$$

§ 7. УПРОЩЕННЫЙ СПОСОБ ВЫЧИСЛЕНИЯ \bar{x} И σ

Вычисление \bar{x} и σ можно значительно упростить, если воспользоваться некоторыми элементарными свойствами этих показателей.

СВОЙСТВО I

Если все частоты (веса) умножить на одно и то же число, то \bar{x} и σ не изменятся.

Доказательство очевидно из рассмотрения формул для взвешенных показателей:

$$\bar{x} = \frac{\sum xm}{\sum m}, \quad \sigma = \sqrt{\frac{\sum (x - \bar{x})^2 m}{\sum m}}$$

Если все веса m умножить на постоянное число k , то это число выйдет за знак суммы в числителе и знаменателе формул и сократится. Из свойства I следует, что при вычислении \bar{x} и σ можно пользоваться относительными частотами вместо абсолютных. При этом результаты не изменятся.

СВОЙСТВО II

Если все варианты признака умножить на одно и то же число, то \bar{x} умножится на это число, а σ умножится на абсолютную величину этого числа.

Доказательство

Пусть k — число (положительное или отрицательное), на которое умножаются все варианты.

Новые варианты $z = kx$, новая средняя арифметическая

$$\bar{z} = \frac{\sum z}{n} = \frac{\sum kx}{n} = k \frac{\sum x}{n} = k\bar{x}$$

и новое среднее квадратическое отклонение

$$\begin{aligned} \sigma_z &= \sqrt{\frac{\sum (z - \bar{z})^2}{n}} = \sqrt{\frac{\sum (kx - k\bar{x})^2}{n}} = \sqrt{\frac{k^2 \sum (x - \bar{x})^2}{n}} \\ &= |k| \sqrt{\frac{\sum (x - \bar{x})^2}{n}} = |k| \sigma. \end{aligned}$$

СВОЙСТВО III

Если ко всем вариантам прибавить одно и то же число a , то \bar{x} увеличится на a , а σ останется без изменения.

Доказательство

Пусть $z = (x + a)$, тогда

$$\text{а) } \bar{z} = \frac{\sum (x + a)}{n} = \frac{\sum x}{n} + \frac{\sum a}{n},$$

но

$$\frac{\sum x}{n} = \bar{x}, \quad \frac{\sum a}{n} = \frac{\overbrace{a + a + \dots + a}^{n \text{ раз}}}{n} = \frac{na}{n} = a,$$

следовательно, $\bar{z} = \bar{x} + a$;

$$\begin{aligned} \text{б) } \sigma_z &= \sqrt{\frac{\Sigma(z - \bar{z})^2}{n}} = \sqrt{\frac{\Sigma(x + a - \bar{x} - a)^2}{n}} = \\ &= \sqrt{\frac{\Sigma(x - \bar{x})^2}{n}} = \sigma. \end{aligned}$$

Свойства II и III доказаны для невзвешенных \bar{x} и σ . Читателю рекомендуется провести доказательство для взвешенных \bar{x} и σ .

СВОЙСТВО IV

Дисперсия равна среднему квадрату минус квадрат средней

$$\sigma^2 = \frac{\Sigma x^2}{n} - (\bar{x})^2 \quad (2.7.1)$$

и для взвешенных показателей

$$\sigma^2 = \frac{\Sigma x^2 m}{\Sigma m} - (\bar{x})^2. \quad (2.7.2)$$

Доказательство

Докажем формулу (2.7.1); формула (2.7.2) доказывается аналогично

$$\sigma^2 = \frac{\Sigma(x - \bar{x})^2}{n} = \frac{\Sigma x^2}{n} - \frac{\Sigma 2x\bar{x}}{n} + \frac{\Sigma(\bar{x})^2}{n}.$$

Во втором члене выносим постоянное $2\bar{x}$ за знак суммы

$$\Sigma 2x\bar{x} = 2\bar{x}\Sigma x.$$

В третьем члене под знаком суммы постоянное $(\bar{x})^2$ повторяется n раз. Следовательно, $\Sigma(\bar{x})^2 = n(\bar{x})^2$.

Отсюда

$$\begin{aligned} \sigma^2 &= \frac{\Sigma x^2}{n} - 2\bar{x} \frac{\Sigma x}{n} + \frac{n(\bar{x})^2}{n} = \frac{\Sigma x^2}{n} - 2(\bar{x})^2 + (\bar{x})^2 = \\ &= \frac{\Sigma x^2}{n} - (\bar{x})^2, \end{aligned}$$

что и требуется доказать.

Последнее свойство позволяет заменить расчет квадратов отклонений от средней вычислением квадратов самих вариантов, что во многих случаях значительно проще.

Приведем пример.

Пусть $x=2, 5, 1, 3$.

Найти σ двумя способами: по основной формуле (2.6.2) и по формуле (2.7.1).

Первый способ

Таблица 15

x	$x - \bar{x}$	$(x - \bar{x})^2$
2	-0,75	0,5625
5	2,25	5,0625
1	-1,75	3,0625
3	0,25	0,0625
11	0	8,7500

Второй способ

Таблица 16

x	x^2
2	4
5	25
1	1
3	9
11	39

Изложим теперь упрощенный способ вычисления \bar{x} и σ , основанный на свойствах I—IV настоящего параграфа и непосредственно применимый к рядам распределения с равными интервалами.

Рассмотрим пример. Пусть дано распределение рабочих по зарплате (первые два столбца табл. 17).

Таблица 17

Зарплата рабочего	Число рабочих	x	t	x'	$t \cdot x'$	$t \cdot x'^2$
600—700	5	650	1	-2	-2	4
700—800	10	750	2	-1	-2	2
800—900	20	850	4	0	0	0
900—1000	15	950	3	1	3	3
Итого . . .	50		10		-1	9

Центры интервалов записываем в столбце 3 таблицы.

Прежде всего сокращаем частоты в 5 раз; согласно свойству I это не отразится на x и σ . Записываем веса t в столбце 4. Заменяем переменное x новым x' , которое связано с прежним соотношением.

$$x' = \frac{x - C}{i}, \quad (2.7.3)$$

где i — длина интервала (в нашем примере $i = 100$);

C — новое начало отсчета, которое обычно выбираем вблизи центра ряда (в нашем примере примем $C = 850$).

Значения x' записываем в столбце 5. (Практически мы записываем числа x' , не вычисляя их по формуле (2.7.3), а просто ставим «0» против какого-то значения $x=C$ и далее в сторону возрастания вариантов пишем 1, 2, 3., в сторону убывания вариантов: $-1, -2, -3...$)

Далее производим вычисление средней арифметической и среднего квадратического отклонения для новых вариантов x' . Для этой цели в столбце 6 записываем произведения mx' , в столбце 7 — произведения mx'^2 . Вычисляем

$$\begin{aligned}\bar{x}' &= -\frac{1}{10} = -0,1, \quad \sigma' = \sqrt{\frac{9}{10} - (-0,1)^2} = \\ &= \sqrt{0,9 - 0,01} = 0,943\end{aligned}$$

Для того чтобы перейти к средней арифметической и сигме прежних вариантов, пользуемся формулами перевода:

$$\bar{x} = \bar{x}' \cdot i + C, \quad (2.7.4)$$

$$\sigma = \sigma' \cdot i. \quad (2.7.5)$$

В нашем примере

$$\bar{x} = -0,1 \cdot 100 + 850 = 840,$$

$$\sigma = 0,943 \cdot 100 = 94,3.$$

Мы пришли к тем же значениям, которые были ранее получены прямым способом.

Приводим доказательство формул (2.7.4) и (2.7.5).

Из соотношения (2.7.3) имеем:

$$x = x' \cdot i + C.$$

На основании свойств II и III отсюда следует, что

$$\bar{x} = \bar{x}' \cdot i + C \quad \text{и} \quad \sigma = \sigma' \cdot i.$$

В ряде случаев упрощения, основанные на свойствах II—IV, могут быть применены к совокупности негруппированных данных. Так, σ обычно предпочтительнее вычислять по формуле (2.7.1), нежели по основной формуле (2.6.2). Если многозначные варианты отличаются между собой лишь в последних знаках, то целесообразно при вычислении \bar{x} и σ первые цифры отбросить, а потом результат исправить. Например, величина диаметра цапфы оси высотомера оказалась равной у нескольких образцов этой детали 2,974; 2,980; 2,966; 2,973; 2,975; 2,984. При

вычислении \bar{x} и σ можно использовать лишь последние две цифры этих вариантов: 74, 80, 66, 73, 75, 84. При этом к средней арифметической нужно будет прибавить число 2,900, а σ останется без изменения.

§ 8. НУЛЕВОЕ И МИНИМАЛЬНОЕ СВОЙСТВА СРЕДНЕЙ АРИФМЕТИЧЕСКОЙ

В настоящем параграфе мы докажем два основных свойства средней арифметической.

I. СВОЙСТВО (НУЛЕВОЕ)

Сумма отклонений от средней арифметической равна нулю.

Пример. В табл. 18 в столбце 1 указаны значения x , в столбце 2 — их отклонения от средней арифметической; легко проверить, что сумма последних равна нулю.

Таблица 18

x	$x - \bar{x}$	$(x - \bar{x})^2$	$x - 4$	$(x - 4)^2$	$x - 6$	$(x - 6)^2$
8	3	9	4	16	2	4
5	0	0	1	1	-1	1
4	-1	1	0	0	-2	4
3	-2	4	-1	1	-3	9
4	-1	1	0	0	-2	4
6	1	1	2	4	0	0
$\Sigma = 30$	0	16		22		22

$$\bar{x} = \frac{30}{6} = 5.$$

Доказательство свойства I

$$\Sigma(x - \bar{x}) = \Sigma x - \Sigma \bar{x}, \text{ но } \Sigma x = n\bar{x} \left(\text{так как } \bar{x} = \frac{\Sigma x}{n} \right),$$

$$\Sigma \bar{x} = n\bar{x}$$

(так как под знаком суммы **постоянное** \bar{x} повторяется n раз).

$$\text{Следовательно, } \Sigma(x - \bar{x}) = n\bar{x} - n\bar{x} = 0.$$

II. СВОЙСТВО (МИНИМАЛЬНОЕ)

Сумма квадратов отклонений от средней арифметической меньше, чем сумма квадратов отклонений от любого другого числа.

Проверим на примере. В столбце 3 табл. 18 даны квадраты отклонений от $\bar{x}=5$. Их сумма равна 16.

В столбце 5 даны квадраты отклонений от $a=4$, сумма равна 22.

В столбце 7 даны квадраты отклонений от $b=6$, сумма равна 22.

Сумма квадратов отклонений от \bar{x} оказалась **меньшей, чем** сумма квадратов отклонений от числа $a < \bar{x}$ и от числа $b > \bar{x}$.

Доказательство свойства II

Обозначим через f сумму квадратов отклонений вариантов от числа c и будем искать значение c , которое обращает функцию f в минимум

$$f = \Sigma (x - c)^2 = \min. \quad (2.8.1)$$

Решением является корень уравнения

$$\frac{df}{dc} = 0. \quad (2.8.2)$$

Дифференцируя, получаем

$$-2\Sigma(x - c) = 0 \quad \text{или} \quad \Sigma(x - c) = 0, \\ \Sigma x = \Sigma c, \quad \Sigma x = nc,$$

откуда

$$c = \frac{\Sigma x}{n} = \bar{x},$$

c оказалось средней арифметической, что и доказывает II свойство¹.

Условие (2.8.1) называется *требованием наименьших квадратов*. Смысл его заключается в следующем: нас интересует число c , которое ближе всего подходит к заданной системе значений x . Но близость c к x измеряется абсолютной величиной отклонения $|x - c|$, или, как это более принято, квадратом отклонения $(x - c)^2$.

¹ Уравнение (2.8.2) является, как известно, лишь необходимым условием минимума. Достаточным условием является при этом положительный знак второй производной: $\frac{d^2f}{dc^2} > 0$.

Легко видеть, что последнее условие имеет место. В самом деле

$$\frac{d^2f}{dc^2} = \frac{d}{dc} [-2\Sigma(x - c)] = 2\Sigma 1 = 2n > 0.$$

Но так как мы имеем несколько значений x , то естественно потребовать, чтобы сумма квадратов отклонений была наименьшей.

Мы видим, что число, которое ближе всего (в смысле требования наименьших квадратов) подходит к заданной системе значений x , является средней арифметической из этих значений.

Во II свойстве обнаруживается связь между средней арифметической и средним квадратическим отклонением; из этого свойства следует, что среднее квадратическое отклонение, вычисленное от средней арифметической, меньше, чем от какой-либо другой величины (например, медианы или моды)

$$\sqrt{\frac{\sum (x - \bar{x})^2}{n}} < \sqrt{\frac{\sum (x - c)^2}{n}},$$

если $c \neq \bar{x}$.

§ 9. СРЕДНЯЯ АРИФМЕТИЧЕСКАЯ И ДИСПЕРСИЯ ДЛЯ ОБЪЕДИНЕННОЙ СОВОКУПНОСТИ

В статистике большое значение имеет не только группировка данных, т. е. расчленение совокупности на части, но и обратная операция — объединение частных совокупностей в одну общую. При этом важно по показателям частных совокупностей уметь рассчитать соответствующие показатели объединенной совокупности. В настоящем параграфе мы установим формулы, связывающие \bar{x} и σ^2 объединенной совокупности с \bar{x} и σ^2 частных совокупностей.

Пример. Данные о производственном стаже для двух совокупностей рабочих приведены в табл. 19.

Таблица 19

Совокупность	Число рабочих	Средний стаж	Дисперсия по стажу
1-я	$n_1=100$	$\bar{x}_1=8$	$\sigma_1^2=4$
2-я	$n_2=50$	$\bar{x}_2=12$	$\sigma_2^2=6,25$
Итого . . .	$n=150$	$\bar{x}=?$	$\sigma^2=?$

Обе группы объединены в одну. Каковы \bar{x} и σ^2 объединенной совокупности? Показатели отдельных групп назовем частными, искомые показатели объединенной совокупности — общими.

Выведем сначала формулу для \bar{x}

$$\bar{x} = \frac{\sum x}{n}.$$

Σx разбивается на две суммы соответственно двум совокупностям, участвующим в ее образовании

$$\Sigma x = \Sigma_1 x + \Sigma_2 x.$$

Следовательно,

$$\bar{x} = \frac{\Sigma_1 x + \Sigma_2 x}{n_1 + n_2} = \frac{n_1 \frac{\Sigma_1 x}{n_1} + n_2 \frac{\Sigma_2 x}{n_2}}{n_1 + n_2} = \frac{n_1 \bar{x}_1 + n_2 \bar{x}_2}{n_1 + n_2}.$$

Или, обобщая,

$$\bar{x} = \frac{n_1 \bar{x}_1 + n_2 \bar{x}_2 + \dots}{n_1 + n_2 + \dots}. \quad (2.9.1)$$

Формулируем: *средняя арифметическая объединенной совокупности является средней арифметической из частных средних, взвешенных по объемам частей.*

Для нашего примера

$$\bar{x} = \frac{100 \cdot 8 + 50 \cdot 12}{150} = \frac{1400}{150} = 9,33.$$

Несколько сложнее дело обстоит с дисперсией. Было бы ошибкой предполагать, что общая дисперсия является некоторой средней из частных дисперсий. Так, например, если бы даже в каждой частной группе дисперсия была незначительной, но средний уровень признака в этих группах весьма различен (например, в одной группе все рабочие с малым стажем, в другой — с большим стажем), то при объединении групп дисперсия выросла бы за счет разницы *между* группами.

Естественно поэтому предположить, что общая дисперсия складывается из дисперсии, имеющей место *внутри* частных групп, и дисперсии, обнаруживающейся *между* средними данными этих групп.

Это очевидно для каждого индивидуального отклонения $x - \bar{x}$, которое складывается из двух частей: отклонения x от средней той частной совокупности, к которой относится данный x , и отклонения частной средней от общей средней \bar{x}

$$x - \bar{x} = (x - \bar{x}_i) + (\bar{x}_i - \bar{x}),$$

где i — номер части.

Докажем, что подобное равенство *осуществляется в среднем для квадратов отклонений.*

$$\text{Общая дисперсия } \sigma^2 = \frac{\Sigma (x - \bar{x})^2}{n}.$$

Разобьем сумму квадратов на частные суммы, относящиеся к каждой частной совокупности, из которых складывается общая совокупность

$$\Sigma (x - \bar{x})^2 = \Sigma_1 (x - \bar{x})^2 + \Sigma_2 (x - \bar{x})^2 + \dots$$

Первую из этих сумм $\Sigma_1 (x - \bar{x})^2$ преобразуем

$$\begin{aligned} \Sigma_1 (x - \bar{x})^2 &= \Sigma_1 \{(x - \bar{x}_1) + (\bar{x}_1 - \bar{x})\}^2 = \\ &= \Sigma_1 (x - \bar{x}_1)^2 + 2\Sigma_1 (x - \bar{x}_1)(\bar{x}_1 - \bar{x}) + \Sigma_1 (\bar{x}_1 - \bar{x})^2, \end{aligned}$$

$$\Sigma_1 (x - \bar{x}_1)^2 = n_1 \sigma_1^2 \left(\text{так как } \sigma_1^2 = \frac{\Sigma_1 (x - \bar{x}_1)^2}{n_1} \right).$$

Во втором члене постоянный множитель $(\bar{x}_1 - \bar{x})$ выносим за знак суммы

$$2\Sigma_1 (x - \bar{x}_1)(\bar{x}_1 - \bar{x}) = 2(\bar{x}_1 - \bar{x})\Sigma_1 (x - \bar{x}_1) = 0,$$

так как $\Sigma_1 (x - \bar{x}_1)$ равна нулю согласно нулевому свойству средней арифметической.

В третьем члене $\Sigma_1 (\bar{x}_1 - \bar{x})^2$ под знаком суммы постоянное; следовательно,

$$\Sigma_1 (\bar{x}_1 - \bar{x})^2 = n_1 (\bar{x}_1 - \bar{x})^2$$

или, обозначая отклонение частной средней от общей \bar{x} буквой δ_1 (дельта), запишем

$$\Sigma_1 (\bar{x}_1 - \bar{x})^2 = n_1 \delta_1^2.$$

Отсюда

$$\Sigma_1 (x - \bar{x})^2 = n_1 \sigma_1^2 + n_1 \delta_1^2.$$

То же, очевидно, справедливо и для остальных сумм

$$\Sigma_2 (x - \bar{x})^2, \quad \Sigma_3 (x - \bar{x})^2 \dots$$

Поэтому

$$\begin{aligned} \sigma^2 &= \frac{n_1 \sigma_1^2 + n_1 \delta_1^2 + n_2 \sigma_2^2 + n_2 \delta_2^2 + \dots}{n_1 + n_2 + \dots} = \\ &= \frac{n_1 \sigma_1^2 + n_2 \sigma_2^2 + \dots}{n_1 + n_2 + \dots} + \frac{n_1 \delta_1^2 + n_2 \delta_2^2 + \dots}{n_1 + n_2 + \dots}. \end{aligned}$$

Первый член представляет собой *среднюю из частных дисперсий* (взвешенную); ее можно обозначить через $\bar{\sigma}_i^2$

$$\bar{\sigma}_i^2 = \frac{n_1\sigma_1^2 + n_2\sigma_2^2 + \dots}{n_1 + n_2 + \dots} \quad (2.9.2)$$

Второй член является *дисперсией частных средних*; его обозначим через $\bar{\delta}_i^2$

$$\bar{\delta}_i^2 = \frac{n_1\delta_1^2 + n_2\delta_2^2 + \dots}{n_1 + n_2 + \dots} \quad (2.9.3)$$

Итак,

$$\sigma^2 = \bar{\sigma}_i^2 + \bar{\delta}_i^2 \quad (2.9.4)$$

или *общая дисперсия равна средней из частных дисперсий плюс дисперсия частных средних.*

Иначе, общая колеблемость складывается из колеблемости внутри частей и колеблемости между частями.

Рассчитаем σ^2 для нашего примера

$$\delta_1^2 = (\bar{x}_1 - \bar{x})^2 = (8 - 9,33)^2 = 1,77,$$

$$\delta_2^2 = (\bar{x}_2 - \bar{x})^2 = (12 - 9,33)^2 = 7,11,$$

$$\bar{\delta}_i^2 = \frac{100 \cdot 1,77 + 50 \cdot 7,11}{150} = \frac{532,5}{150} = 3,55,$$

$$\bar{\sigma}_i^2 = \frac{100 \cdot 4 + 50 \cdot 6,25}{150} = \frac{712,5}{150} = 4,75,$$

$$\sigma^2 = 4,75 + 3,55 = 8,3.$$

Формула (2.9.4) имеет весьма важное значение для теории корреляции.

§ 10. ПРИМЕРЫ ИСПОЛЬЗОВАНИЯ РЯДОВ РАСПРЕДЕЛЕНИЯ НА ПРОИЗВОДСТВЕ

Пример 1. Ряды распределения размеров деталей, обрабатываемых на станке, дают наглядное представление о точности работы станка и эффективности тех или иных мероприятий, направленных на повышение его точности.

В табл. 20 приведены ряды распределения замеров диаметра большого венца шестерни второй коробки передач, полученных после токарной чистовой обточки на многорезцовом полуавтомате, до и после текущего ремонта станка. Ремонт заключался в смене центров, очистке частей суппорта и изменении геометрии заточки резца.

Таблица 20

До ремонта			После ремонта		
x (в мм)	m	$Y\%$	x (в мм)	m	$Y\%$
1	2	3	4	5	6
15,5—19,5	14	3,5	7,5—10,5	1	0,6
19,5—23,5	20	5,0	10,5—13,5	4	2,2
23,5—27,5	26	6,5	13,5—16,5	8	4,4
27,5—31,5	25	6,25	16,5—19,5	11	6,1
31,5—35,5	4	1,0	19,5—22,5	20	11,2
35,5—39,5	7	1,75	22,5—25,5	10	5,6
39,5—43,5	3	0,75	25,5—28,5	6	3,3
43,5—47,5	1	0,25			
	100			60	

Замеры производились микрометром с точностью до 0,01 мм. Допуск на размер для данной операции установлен $61,10 \div \div 61,30$ мм. Для упрощения записи размеры выражаются в условных единицах. За начало отсчета принимается 61 мм, за единицу измерения — 0,01 мм. В этих единицах записаны значения признака в столбцах 1 и 4 таблицы.

При построении графиков распределения следует иметь в виду, что оба ряда различаются между собой по количеству замеров ($n_1=100$, $n_2=60$) и по величине интервала группировки ($i_1=4$, $i_2=3$). Поэтому для сравнения графиков необходимо по оси ординат откладывать относительные частоты, приходящиеся на единицу измерения признака, одинаковую для обоих рядов. Эти величины, называемые *относительными плотностями частоты*, рассчитываются по формуле

$$Y\% = mk,$$

где $Y\%$ — относительная плотность частоты, выраженная в процентах;

m — абсолютная частота интервала.

$$k = \frac{100}{i \cdot n},$$

где i — величина интервала группировки;
 n — общее число наблюдений.

В нашем случае для первого ряда

$$k_1 = \frac{100}{100 \cdot 4} = 0,25,$$

для второго ряда

$$k_2 = \frac{100}{60 \cdot 3} = 0,556.$$

Значения $Y\%$ записаны в столбцах 3 и 6 таблицы.
 Графики строим в виде гистограмм, представленных параллельно на рис. 12.
 На оси абсцисс наносим поле допуска. Следует особо остановиться на его определении. Номинально поле допуска задано в

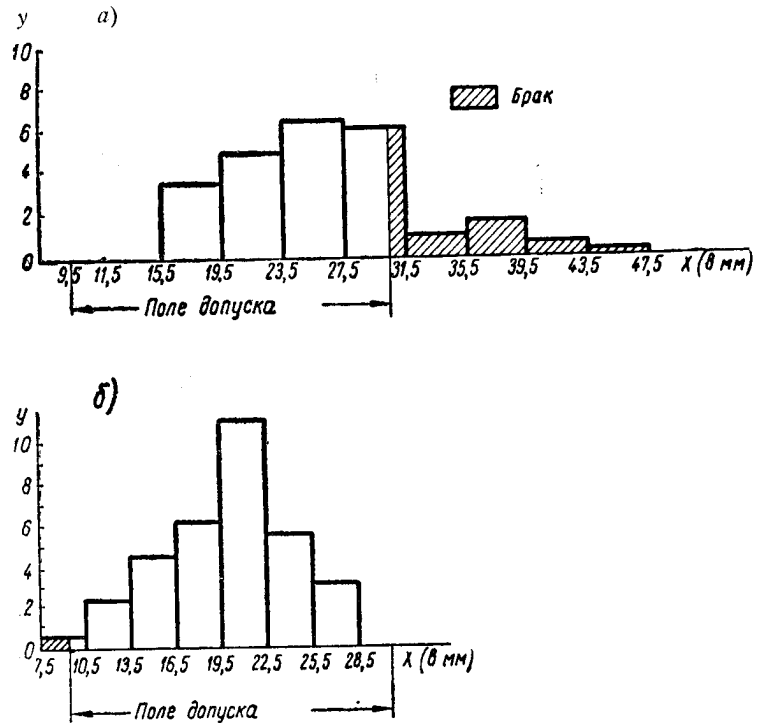


Рис. 12. Графики распределения деталей по размеру диаметра:
 а) до ремонта станка, б) после текущего ремонта станка

условных единицах от 10 до 30. Однако поскольку числа 10 и 30 включаются в поле допуска, а интервал округления значений в процессе измерения составляет одну единицу (0,01 мм), то действительными границами поля допуска являются числа a и b , полученные по формулам:

$$a = a' - \frac{j}{2} = 10 - \frac{1}{2} = 9,5,$$

$$b = b' + \frac{j}{2} = 30 + \frac{1}{2} = 30,5,$$

где a' и b' — номинальные границы;
 j — величина интервала округления.

Полученные границы отмечаются на графиках.

По данным табл. 20 рассчитываем показатели \bar{x} , σ , процент брака для каждого ряда. Результаты вычислений записываем в табл. 21.

Таблица 21

	\bar{x}	σ	Процент брака		
			слева	справа	общий
До ремонта	26,42	6,30	0	21,25	21,25
После ремонта	19,95	4,27	1,12	0	1,12

Приводим расчет процента брака.

До ремонта частота дефектных деталей с диаметром больше допуская составляла:

$$1 + 3 + 7 + 4 + 25 \frac{31,5 - 30,5}{31,5 - 27,5} = 21,25.$$

В последнем интервале частота брака рассчитана по правилу пропорциональности. Следовательно, процент брака справа равен

$$\frac{100 \cdot 21,25}{100} = 21,25.$$

Аналогично в других случаях.

На основе представленных графиков и рассчитанных показателей можно сделать следующие выводы:

1. График распределения размеров деталей до ремонта станка был сильно сдвинут к правой границе поля допуска. Точка средней арифметической \bar{x} оказалась правее середины поля допуска на 6,42 единицы, что составляет 30,6% от величины поля допуска. Столь резкое смещение центра тяжести кривой свидетельствует о плохой настроенности станка на требуемые размеры. График распределения обладает большим рассеянием и резкой асимметрией в правую сторону. Все это обусловило весьма значительный процент правостороннего брака.

2. График распределения размеров деталей после ремонта станка в основном укладывается в поле допуска. Точка \bar{x} близка к середине поля допуска, рассеяние графика заметно сократилось. Все это свидетельствует о значительно лучшей настроенности и точности процесса. В результате процент брака уменьшился до 1,12%.

Пример 2. Ряд распределения плавок по их продолжительности является важной характеристикой качества работы и производительности плавильного агрегата.

Приводим графики распределения продолжительности плавок для трех марок стали по двум печам, имеющим некоторые конструктивные особенности (по материалам О. И. Яцунской, завод «Серп и молот», рис. 13).

Во всех трех случаях графики распределения плавок первой печи оказались левее расположенными и менее рассеянными, чем графики распределения плавок второй печи. По этим данным можно сделать заключение, что работа первого агрегата более производительна и более ритмична, чем работа второго агрегата. Средняя продолжительность плавок первой печи по каждой марке стали меньше средней продолжительности плавок второй печи примерно на 20 мин., что достаточно ощутимо в производственных условиях.

Пример 3. Процесс изготовления деталей предусматривает ряд операций, направленных на достижение определенных размеров, формы, чистоты поверхности и других параметров с требуемой точностью. В хорошо спроектированном технологическом процессе каждая последующая операция должна приближать те или иные параметры изделия к требуемым значениям и повышать точность обработки, т. е. сокращать размеры погрешностей.

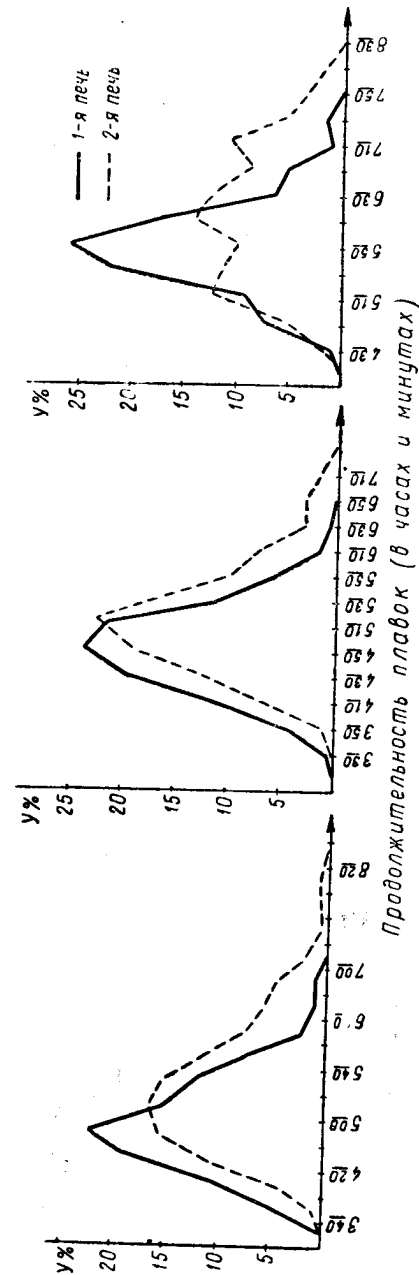


Рис. 13. Графики распределения продолжительности плавок для двух печей трех марок стали

Сопоставляя кривые распределения погрешностей после ряда последовательных операций, мы получаем представление о качестве технологического процесса.

В табл. 22 приведены ряды распределения детали валик-генератора по параметру — биение третьей шейки после ряда следующих друг за другом операций: центровки, токарной обточка, первой (предварительной) шлифовки и второй (окончательной) шлифовки.

Таблица 22

Центровка		Токарная обточка		Первая шлифовка		Вторая шлифовка	
x (биение в 0,01 мм)	m	x (в 0,01 мм)	m	x (в 0,01 мм)	m	x (в 0,01 мм)	m
0—2,5	24	0,5—3,5	20	0,125	2	0,125	13
2,5—5,5	93	3,5—6,5	90	0,5	19	0,5	128
5,5—8,5	76	6,5—9,5	97	1,0	82	1,0	128
8,5—11,5	37	9,5—12,5	58	1,5	80	1,5	14
11,5—14,5	23	12,5—15,5	20	2,0	85	2,0	9
14,5—17,5	13	15,5—18,5	10	2,5	11	2,5	2
17,5—20,5	4	18,5—21,5	1	3,0	10	3,0	3
20,5—23,5	10	21,5—24,5	1	3,5	3		
23,5—26,5	3			4,0	4		
26,5—29,5	4			4,5	0		
29,5—32,5	2			5,0	1		
32,5—35,5	3						
35,5—38,5	2						
38,5—41,5	1						
41,5—44,5	1						
44,5—47,5	1						
Итого . . .	297		297		297		297

На рис. 14 даны соответствующие графики, а в табл. 23 приведены основные показатели рядов распределения.

Таблица 23

Показатели После операции	Показатели	
	\bar{x}	σ
Центровка	8,98	7,61
Токарная обточка . . .	8,07	3,64
Первая шлифовка . . .	1,58	0,699
Вторая шлифовка . . .	0,83	0,458

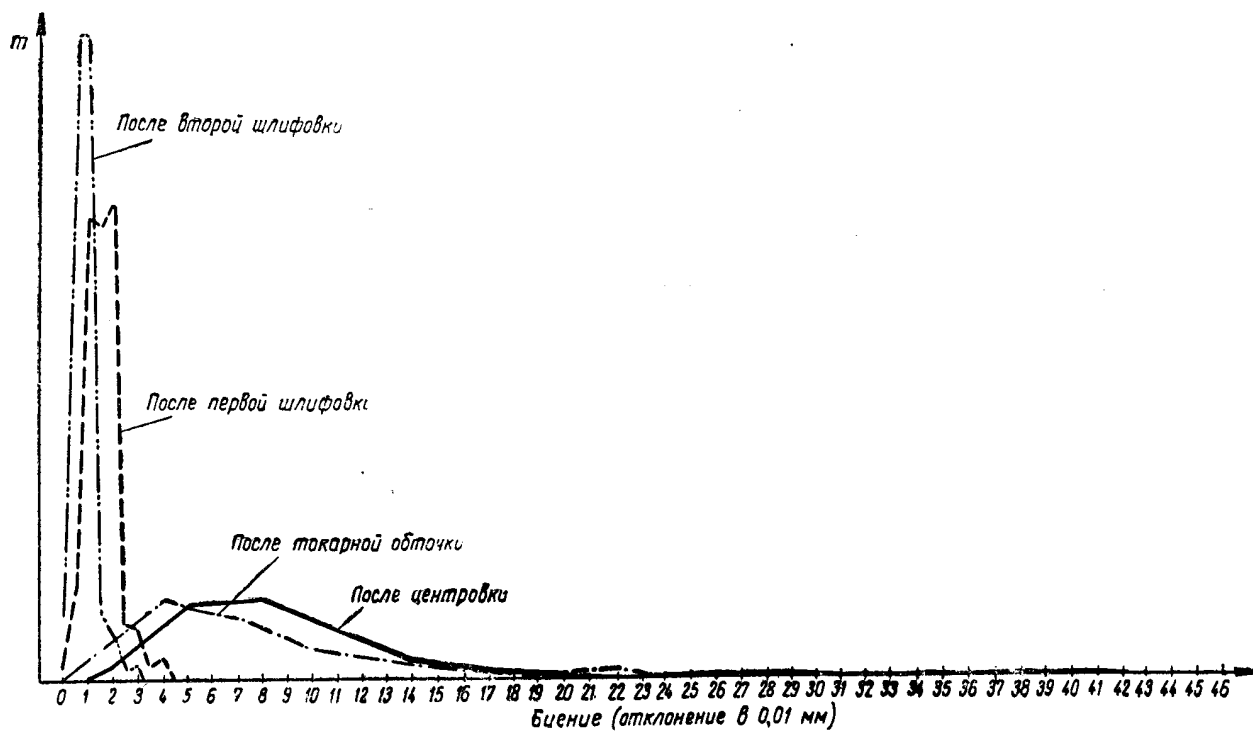


Рис. 14. Графики распределения валков генератора по параметру биения 3-й шейки

При анализе данных следует иметь в виду, что чем меньше величина биения, тем лучше качество продукции. Мы видим, что графики распределения биений имеют асимметричную форму, характерную для распределений существенно положительных величин. Значительная асимметрия была после операции центровки,

но уже следующая операция — токарная обточка — устранила особо резкие погрешности, образовавшие «хвост» графика. После каждой операции графики все более стягиваются к началу координат, уменьшаются средние величины, сокращается рассеяние. Особо резкое сокращение погрешностей, качественно изменяющее график распределения, наступает после операции первого шлифования.

В табл. 24 и на рис. 15 даны ряды и соответствующие графики распределения тех же деталей по другому параметру — размеру диаметра — после операций токарной обточки и первого шлифования.

Мы видим, что после первой шлифовки график распределения оказался значительно смещенным влево и менее рассеянным. Величина смещения, измеряемая разностью $x_1 - x_2$ (см. табл. 25), харак-



Рис. 15. Графики распределения валков генератора по размеру диаметра

теризует среднюю величину припуска на обработку, снятого на операции шлифования.

Сокращение рассеяния говорит об увеличении точности обработки, достигнутом на операции шлифования. Количественная оценка этого сокращения дается сопоставлением средних квадратических отклонений (см. табл. 25).

Таблица 24

После токарной обточки		После первой шлифовки	
x (отклонения от номинала в сотых мм)	m	x (отклонения от номинала в сотых мм)	m
43,5—47,5	1	14,5—15,5	1
47,5—51,5	13	15,5—16,5	1
51,5—55,5	39	16,5—17,5	2
55,5—59,5	66	17,5—18,5	16
59,5—63,5	55	18,5—19,5	26
63,5—67,5	51	19,5—20,5	38
67,5—71,5	51	20,5—21,5	47
71,5—75,5	16	21,5—22,5	73
75,5—79,5	5	22,5—23,5	55
79,5—83,5	3	23,5—24,5	31
		24,5—25,5	8
		25,5—26,5	1
		26,5—27,5	1
	300		300

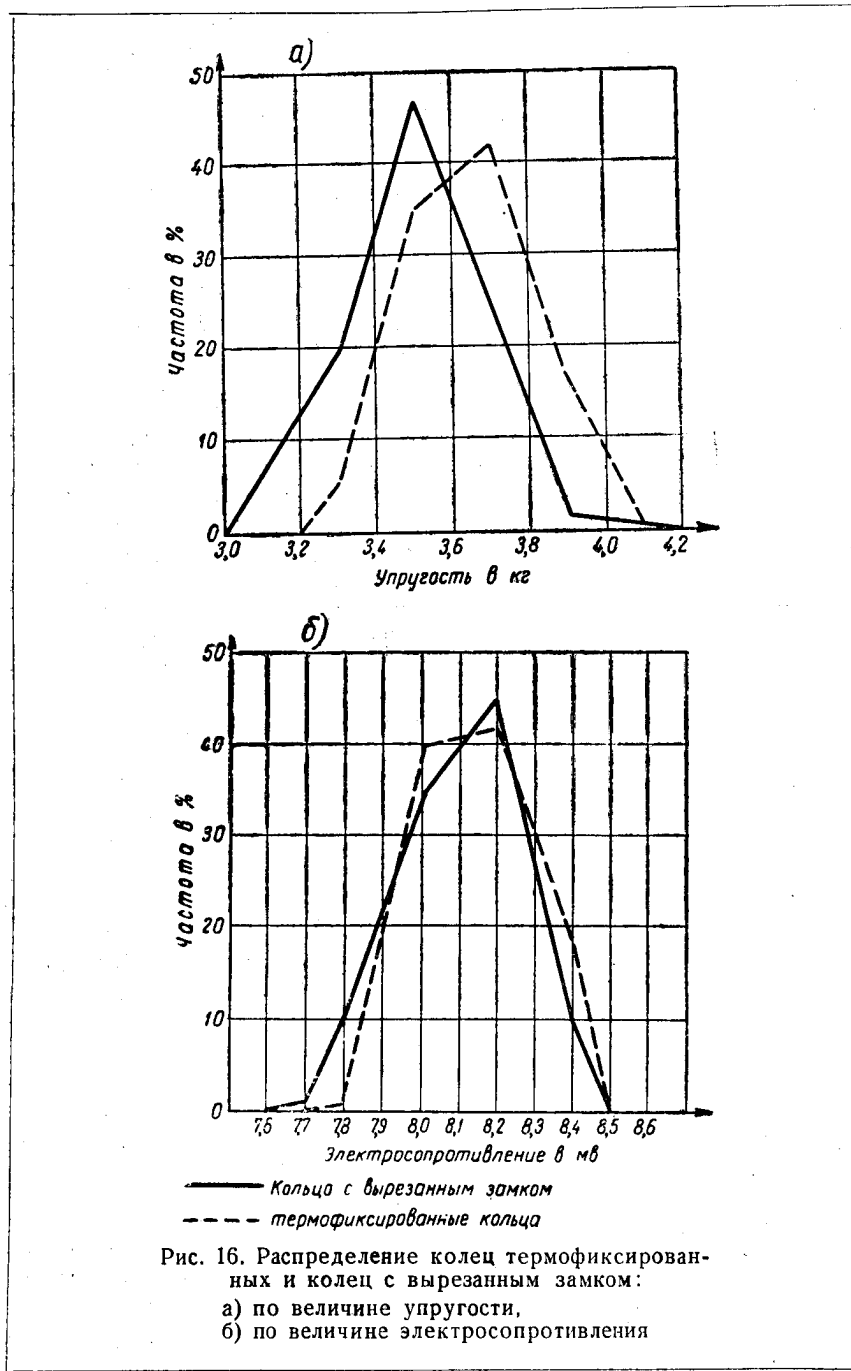
Таблица 25

Показатели		
	\bar{x}	σ
После операции		
Токарная обточка . . .	62,15	6,92
Первая шлифовка . . .	21,54	1,87

Пример 4. Ряды распределения используются для оценки эффективности различных вариантов технологического процесса.

Проверялись два способа образования раствора замка поршневых колец: изготовление отливок эллиптической формы с последующим вырезанием замка и изготовление отливок круглой формы с последующим надеванием на оправку и термообработкой, фиксирующей форму кольца и раствор замка, полученные при надевании (способ «термофиксации»). На рис. 16,а приведены графики распределения термофиксированных колец и колец с вырезанным замком по величине упругости, а на рис. 16,б — графики распределения тех же колец по величине электросопротивления (по данным И. Н. Богачева и В. Ф. Сенкевича)¹.

¹ И. Н. Богачев и В. Ф. Сенкевич, Структура и свойства поршневых колец, Машгиз, 1949, стр. 93.



Из графиков видно, что при вырезании замка достигается большая величина упругости колец, чем при термофиксации. Оба способа изготовления оказались приблизительно равноценными по показателям электросопротивления.

§ 11. ПРОБЛЕМА ФОРМЫ КРИВОЙ РАСПРЕДЕЛЕНИЯ

Положение кривой распределения на оси абсцисс и ее рассеяние являются двумя наиболее существенными свойствами кривой; однако наряду с ними существует и ряд других важных свойств кривой распределения (степень ее асимметрии, высокоили низковершинность), которые в совокупности характеризуют форму или тип кривой распределения.

Выяснение формы кривой распределения является задачей весьма важной как в практическом, так и в теоретическом отношении. Следует руководствоваться тем основным положением, что каждый статистический материал в обычных для него условиях дает по определенному признаку характерную, типичную для него кривую распределения. Всякое искажение этой характерной формы кривой означает нарушение или изменение нормальных условий возникновения материала. Так, следя за качеством продукции в заводской лаборатории, мы в случае изменения характерных для данного материала кривых можем сделать вывод и сигнализировать о нарушении нормального хода технологического процесса или изменении качества материала.

Появление двухвершинной или асимметричной кривой обычно говорит о разнотипном составе совокупности. Во всех подобных случаях важно прежде всего установить тип кривой распределения, обычный для данного материала в нормальных условиях его возникновения.

Однако форма кривой распределения обычно искажается случайными зигзагами — колебаниями, возникающими благодаря недостаточному числу наблюдений. Необходимо прежде всего освободиться от этих случайных зигзагов кривой и выявить закономерность, которая лежит в основе материала.

Прямым методом освобождения кривой от случайных зигзагов является, очевидно, увеличение числа наблюдений при тех же нормальных для данного материала условиях. На рис. 17 мы показываем, как уменьшается зигзагообразность кривой и выявляется ее форма при увеличении числа образцов для распределения медной втулки по величине угла конуса (после обточки на станке-автомате). Однако этот прямой путь чаще всего недоступен исследователю. Исследователь обычно имеет дело с материалом ограниченной численности и его задача — вскрыть закономерность, которая лежит в основе материала и которая проявилась бы при весьма большом количестве наблюдений.

Для решения этой задачи исследователь должен использовать косвенные методы. Такими методами являются:

- а) укрупнение интервалов,
- б) математическое выравнивание кривой.

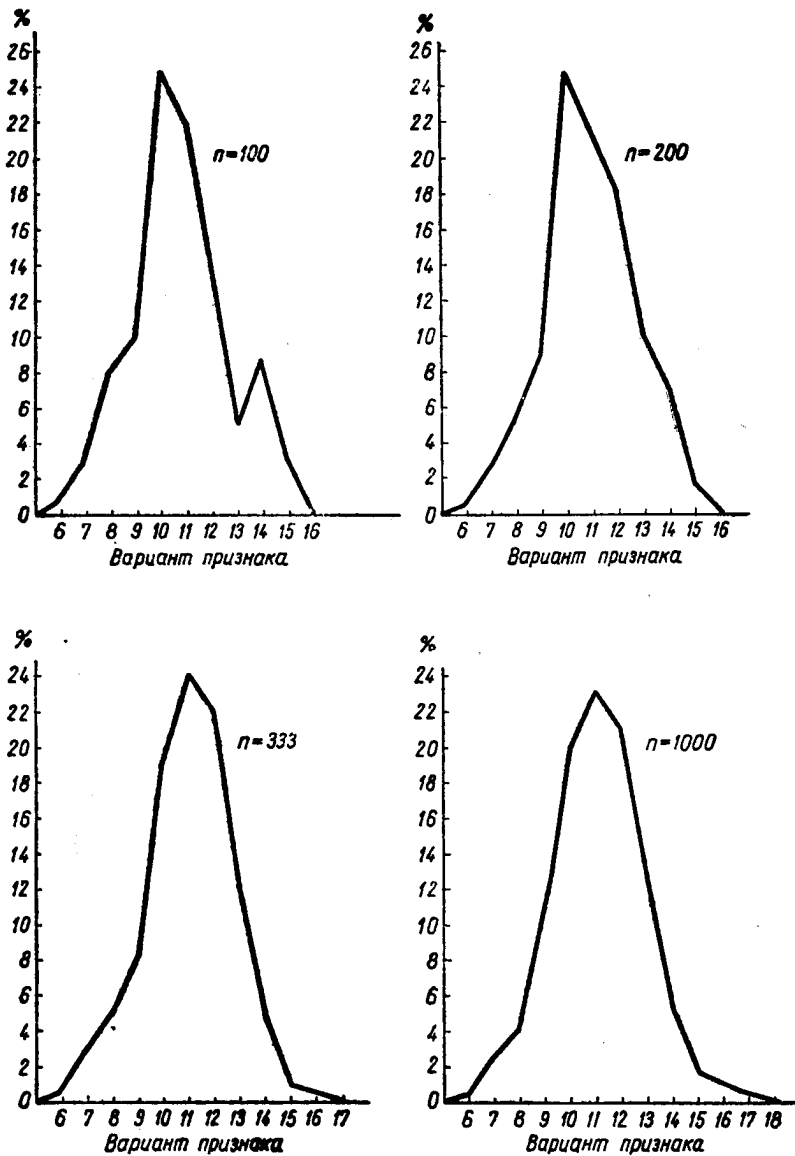


Рис. 17. Распределение угла конуса медной втулки при различном числе образцов в партии

При укрупнении интервалов число наблюдений в интервалах увеличивается и частоты изменяются более правильно.

На рис. 18 показаны графики распределения упругости поршневых колец при том же числе наблюдений (200), но при разном числе интервалов.

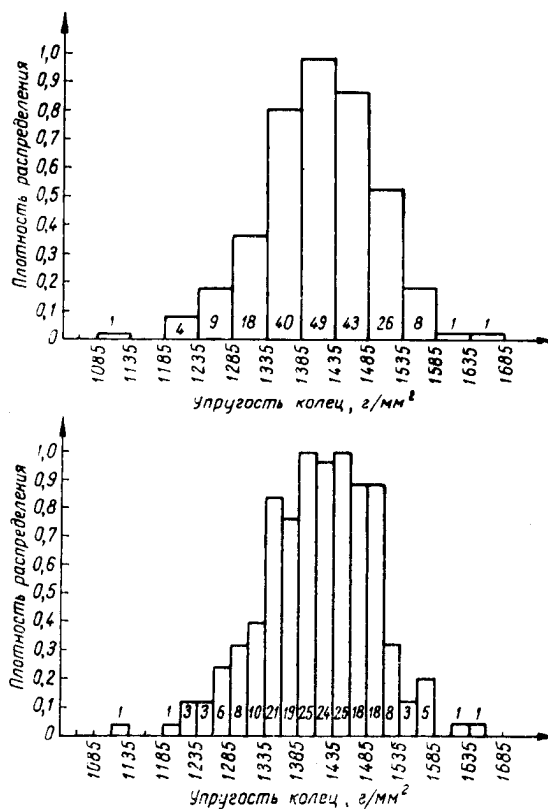


Рис. 18. Гистограмма распределения упругости 200 поршневых колец при различной величине интервалов группировки

Эффективность этого метода имеет свои границы. Нельзя слишком уменьшать число интервалов, так как при этом будут упущены существенные особенности кривой распределения. С другой стороны, при недостаточном числе наблюдений никакое укрупнение интервалов не в состоянии избавить кривую от случайных зигзагов. Более совершенным является метод математического выравнивания кривой, с которым мы познакомимся в разделе III книги.

РАЗДЕЛ II
ТЕОРИЯ КОРРЕЛЯЦИИ

ГЛАВА III

§ 1. ЗАДАЧИ СТАТИСТИЧЕСКОГО ИЗМЕРЕНИЯ СВЯЗИ

В настоящем разделе излагаются методы исследования статистического материала, характеризующего те или иные явления или процессы по нескольким варьирующим признакам. Эти признаки рассматриваются не изолированно, а во взаимной связи друг с другом, причем зависимость между признаками является основным предметом изучения.

Изучением зависимостей мы занимаемся в любой научной дисциплине. Диалектический подход к изучению природы и общества требует рассмотрения явлений в их взаимосвязи и непрерывном изменении. Современная наука, изучая качественную сторону изменения явлений, стремится выразить эти изменения и в количественной форме — в виде определенных соотношений или пропорций между величинами. Приемы количественного измерения связи между величинами — функциональной зависимости — рассматриваются в курсах математического анализа. В чем же особенности и каковы задачи статистического измерения связи?

Вспомним, как определяется функциональная зависимость. В основе этого исторически сложившегося понятия лежит идея однозначного соответствия между величинами, изменяющимися в каком-либо процессе так, что каждому значению одной величины, называемой аргументом, соответствует строго определенное значение другой величины — функции.

Однако каждому экспериментатору известно, что это понятие является лишь абстракцией и что как бы точно ни производился эксперимент, как бы строго ни закреплялись условия опыта и побочные факторы, неизбежен известный разброс результатов эксперимента, не позволяющий прийти к вполне однозначным выводам об интересующей нас зависимости.

Рассмотрим классический пример функциональной зависимости, которым нередко открываются курсы математического анализа, — закон Бойля-Мариотта, устанавливающий при изотер-

мическом процессе функциональную зависимость между объемом газа V и давлением P , под которым находится газ. Придавая в формуле, выражающей эту зависимость: $V = \frac{15}{P}$

(где 15 — условная постоянная), давлению P — различные значения, получаем строго определенные соответственные значения V , например:

P	V
1	15
2	7,5
3	5
...	...
...	...

Так обстоит дело согласно формуле, но не так получается в жизни. Осуществляя фактически опыт с газами, мы убеждаемся в том, что значения V не однозначно соответствуют значениям P . При повторении опыта получаем, как правило, для одного и того же давления несколько различающиеся объемы газа, т. е. наблюдаем колеблемость функции V ; это видно из примера:

P	V			
1	14,92	14,66	15,06	15,61
2	7,43	7,55	7,50	7,62
3	5,08	4,90	4,88	5,11
...
...

Причины колеблемости заключаются в том, что наблюдаемые показатели объема газа в действительности являются функцией не от одного лишь давления. Они зависят также и от ряда других аргументов: качества газа, атмосферных влияний, точности приборов и других условий, характеризующих обстановку опыта. Если бы все эти аргументы были закреплены, если бы они не изменялись от опыта к опыту, то объем газа оказался бы функцией одного лишь давления, т. е. мы имели бы дело со строгой функциональной зависимостью. Однако практически закрепление прочих аргументов не может быть идеально осуществлено; часть из них вообще ускользает из нашего поля зрения, другая часть с трудом поддается контролю. Прочие условия изменяются от опыта к опыту, колеблются по своим законам, случайным по отношению к интересующей нас зависимости, и тем самым создают случайную колеблемость функции V .

Все же, если эта колеблемость невелика, если она сказывается лишь в том или ином «знаке за запятой», можно с известным приближением считать зависимость строго функциональной и об-

ратиться к средствам математического анализа для дальнейшего ее исследования. Так именно обстоит дело в ряде вопросов точного естествознания — механики, физики. Здесь в эксперименте часто удается изолировать и закрепить влияние посторонних факторов с достаточной точностью.

Но во многих вопросах естествознания и техники, а также в общественных науках невозможно устранить влияние посторонних факторов при исследовании зависимости отчасти потому, что в значительной мере эти факторы неизвестны, отчасти потому, что самый эксперимент нередко затруднен или недоступен, и мы вынуждены ограничиться наблюдением явления в его естественных условиях. При этом возникает важная для науки задача — наблюдая за интересующей нас зависимостью при сложном взаимодействии посторонних влияний, уметь ответить на вопрос: какова была бы зависимость между функцией и аргументом, если бы прочие факторы не изменялись и этим своим изменением не осложняли и не затушевывали основную зависимость?

Рассмотрим пример. Известно, что с увеличением содержания марганца (в известных пределах) улучшается прокаливаемость стали (глубина слоя мартенсита при закале). Однако если исследовать эту зависимость по пробным данным, взятым из однотипных плавок, то натолкнуться бы на такие трудности. Могло бы случиться, что в какой-либо плавке с более высоким содержанием марганца в стали степень прокаливаемости оказалась ниже, чем в плавке с меньшим его содержанием. Прокаливаемость стали зависит не только от содержания марганца, но и от ряда других факторов, не остающихся постоянными от плавки к плавке: полноты и характера раскисления стали при плавке и удаления продуктов раскисления, величины зерен аустенита, колебаний температуры закалки, наличия небольших количеств никеля и других легирующих элементов и т. д.

Могло случиться, что эти факторы сложились более благоприятно во второй плавке по сравнению с первой и обусловили там более высокую прокаливаемость. Таким образом, данные двух или нескольких плавок недостаточны для выяснения зависимости. Но при большом числе таких наблюдений возможно обнаружить закономерность, которая позволяет установить, как изменялась бы прокаливаемость стали в зависимости от содержания марганца, если бы прочие аргументы прокаливаемости не изменялись.

Итак, первая основная задача измерения связи заключается в том, чтобы выяснить на основе наблюдения над большим количеством данных, как изменялась бы функция в связи с изменением одного из своих аргументов, если бы ряд других ее аргументов не изменялся. Причем задача решается на таком материале, где эти прочие аргументы на самом деле изменяются и своей изменчивостью затушевывают и искажают интересующую нас зависимость.

Одновременно возникает и другая задача. Искажающее влияние посторонних факторов может обнаруживаться с большей или меньшей силой. Оно может в большей или меньшей степени «размывать», «затушевывать» интересующую нас зависимость. Высвобождая последнюю от искажающего влияния посторонних факторов, мы всегда интересуемся силой этого влияния, или силой, с которой закономерность пробивает себе путь сквозь посторонние влияния. Если сила влияния посторонних факторов мала, то, зная аргумент, можно в каждом отдельном случае точно оценивать значение функции и, воздействуя на аргумент, если он является причиной изменения, управлять функцией. Если сила влияния посторонних факторов велика, то действие интересующего нас аргумента в каждом отдельном случае перекрывается действием прочих факторов, и мы далеки от точной оценки или управления функцией.

Итак, вторая основная задача измерения связи заключается в определении степени искажающего влияния прочих факторов на интересующую нас зависимость или в определении силы, с которой данная зависимость проявляется среди многообразных нарушающих ее воздействий.

В дальнейших параграфах излагаются методы решения этих двух основных задач статистического измерения связи.

§ 2. КОРРЕЛЯЦИОННАЯ ЗАВИСИМОСТЬ

Материал, подлежащий разработке, представляет собой данные наблюдений над элементами статистической совокупности по нескольким варьирующим признакам. Следует иметь в виду, что задачи корреляции решаются всегда *при заданном числе учитываемых* признаков, т. е. признаков, значения которых нам известны для каждого элемента совокупности. Число учитываемых признаков определяет не только средства, которыми располагает исследователь для решения задачи, но и постановку самой задачи, которая видоизменяется и углубляется по мере увеличения числа этих признаков.

Остальные признаки, значения которых для каждого элемента совокупности нам неизвестны, называются *неучитываемыми*. О действии последних исследователь в процессе анализа материала часто строит те или иные предположения для уяснения реального смысла получаемых закономерностей.

В ближайших главах мы рассматриваем вопросы корреляции между двумя, а в конце раздела между тремя и более учитываемыми признаками.

Пример 1. Одним из основных показателей качества поршневого кольца является его упругость. Выяснение причин, вызывающих различия колец по величине упругости, необходимо для усовершенствования производственного процесса и улучшения качества выпускаемой продукции. Особый интерес при контроле

качества продукции представляют такие факторы упругости кольца, измерение которых проще и дешевле, чем измерение самой упругости. Размер замка в свободном состоянии является одним из таких факторов. Поэтому установление зависимости между упругостью кольца и размером его замка представляет

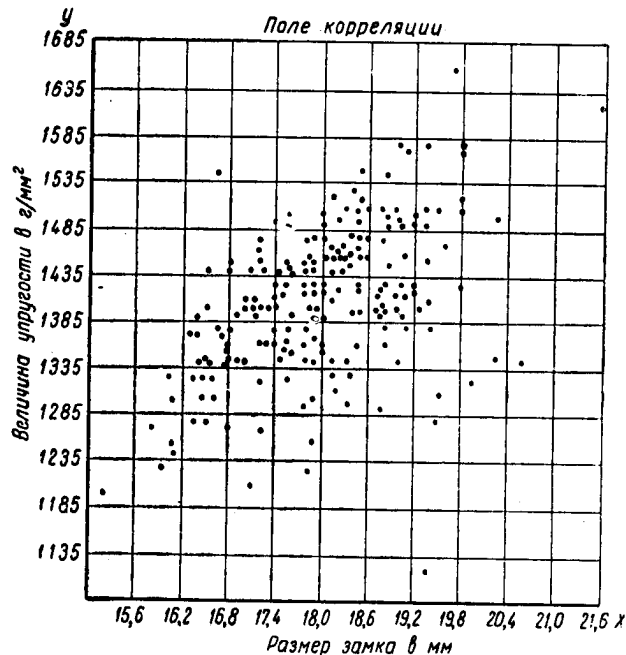


Рис. 19. Зависимость между упругостью и размером замка для 200 поршневых колец

задачу, имеющую практическое значение. Для исследования этой зависимости воспользуемся данными испытаний 200 поршневых колец; в каждом испытании учитывались два показателя: величина упругости (по Толедо) и размер замка в свободном состоянии (см. столбцы 2 и 3 табл. 4 приложения I)¹.

Прежде всего данные таблицы необходимо привести в порядок — систематизировать.

Для этой цели строим оси координат, отмечая на оси x значения аргумента — размер замка, а на оси y — значения функции — величину упругости (рис. 19).

Результаты каждого испытания отмечаем точкой в системе координат (x, y) . Нанеся на чертеж 200 точек, получим так называемое поле корреляции y, x .

¹ Данные заимствованы из работы А. Н. Журавлева, Исследование производственных методов и контроля качества поршневых колец в авиамоторостроении, диссертация.

Поле корреляции является графической формой систематизации материала. Чаще статистический материал систематизируется в табличной форме. Для этой цели намечаем систему интервалов группировки по признакам x и y . По размеру замка (x) начало группировки 15,0, величина интервала 0,6; по упругости (y) начало группировки 1085, величина интервала 50.

На поле корреляции накладывается координатная сетка, соответствующая принятой системе интервалов, и подсчитывается число точек (частота) в каждой клеточке координатной сетки. Результаты подсчета, а также горизонтальные и вертикальные итоги частот записываются в так называемой *корреляционной таблице*.

Таблица 26

		Размер замка (в мм)										Итого		
		15,0—15,6	15,6—16,2	16,2—16,8	16,8—17,4	17,4—18,0	18,0—18,6	18,6—19,2	19,2—19,8	19,8—20,4	20,4—21,0		21,0—21,6	
Упругость (в г. мм ²)	$x \backslash y$													
	1635—1685								1					1
	1585—1635													1
	1535—1585				1									8
	1485—1535							3	1					26
	1435—1485			2	6	13	17	3	2	2				43
	1385—1435			2	10	13	7	13	3	1				49
	1335—1385			10	7	12	4	4	1	1	1			40
	1285—1335			5	1	3	4	1	1	1		1		18
	1235—1285		2	3	2	2	1		1					9
	1185—1235	1	1		1	1								4
	1135—1185													0
1085—1135								1					1	
Итого		1	6	22	27	46	41	32	17	6	1	1	200	

Корреляционная таблица систематизирует результаты наблюдений над элементами статистической совокупности по двум варьирующим признакам. Прежде чем перейти к анализу этой таблицы, рассмотрим еще один пример.

Пример 2. При анализе хозяйственных итогов работы промышленности важно установить причины, в силу которых одни предприятия перевыполнили план, другие же оказались в числе отстающих.

Одним из основных факторов выполнения плана по объему продукции является обеспеченность предприятия рабочей силой. Для освещения вопроса о влиянии этого фактора приводим корреляционную таблицу, содержащую данные 214 предприятий Куйбышевской области о степени выполнения ими плана по валовой продукции в сочетании с процентом выполнения плана по

числу отработанных человеко-дней (таблица заимствована из работы К. Фавстова, помещенной в Ученых записках Планового института, т. 2, Куйбышев, 1940).

В верхних и боковых заголовках таблицы указаны интервалы группировки по x (процент выполнения плана по числу человеко-дней) и по y (процент выполнения плана по валовой продукции). В скобки заключены условные границы: «155—(165)» вместо «свыше 155», «(45)—55» вместо «до 55».

Внутренние клетки таблицы содержат частоты — числа предприятий, оказавшихся в соответствующих интервалах по x и y .

Подведены горизонтальные и вертикальные итоги частоты и указан общий итог 214.

Покажем теперь на примере табл. 27, как осуществляется «чтение» или предварительный анализ корреляционной таблицы. Прежде всего заметим, что итоговый столбец образует совместно с заглавным столбцом ряд распределения 214 предприятий по валовой продукции (аналогично итоговая строка совместно с заглавной образует ряд распределения 214 предприятий по числу отработанных человеко-дней). Любой «внутренний» столбец таблицы также образует (совместно с заглавным столбцом) ряд распределения предприятий по валовой продукции. Но каждый из этих рядов содержит предприятия с одинаковым значением аргумента x , т. е. с одинаковым процентом выполнения плана по рабочей силе. Так, первый ряд включает 6 предприятий со значением x в интервале 45—55% (округленно 50%), столбец 2 включает 7 предприятий, для которых x находится в интервале 55—65% (округленно 60%), и т. д.

Предприятия первого ряда обладают наименьшим значением аргумента (x). Мы видим, что и по валовой продукции (y) эти предприятия также находятся на низком уровне. Следующий по порядку ряд распределения, для которого значение аргумента несколько выше, оказывается и на несколько более высоком уровне значений функции. Переходя далее слева направо в сторону больших значений x , мы видим, что соответствующие ряды распределения y закономерно смещаются снизу вверх — в сторону больших значений функции. Обнаруживается зависимость между валовой продукцией (y) и числом человеко-дней (x).

Однако это не точная функциональная зависимость. Здесь каждому значению аргумента соответствует не одно строго определенное значение функции, а ряд распределения величины y ¹. Такая зависимость называется *корреляционной*.

Говорят, что y находится в корреляционной зависимости от x , если: 1) каждому значению аргумента x соответствует ряд распределения функции y и 2) с изменением x эти ряды закономерно изменяют свое положение.

¹ Зависимую переменную y мы будем условно называть функцией, не связывая этот термин с наличием строгой функциональной зависимости.

Корреляционная таблица
 для 214 предприятий промышленности Куйбышевской области за 1935 г.
 Признаки: валовая продукция—число отработанных человеко-дней (в процентах к плану)

		Процент выполнения плана по числу человеко-дней											Итого	
		(45)– 55	55–65	65–75	75–85	85–95	95– 105	105– 115	115– 125	125– 135	135– 145	145– 155		155– (165)
Процент выполнения плана по валовой продукции	155–(165)					1	1	2	1			1	4	10
	145–155						1					2	3	6
	135–145								4	1		1	2	8
	125–135					4	4	3	5	4				20
	115–125				1		8	5	6	2	1	1	2	26
	105–115		1		3	9	15	6	4	1	1		2	42
	95–105			1	6	6	20	8	2		1			44
	85–95			1	6	5	9	5	2		1		1	30
	75–85			1	1	5	3	2						12
	65–75	3	3	1	1									8
	55–65	1	2		1	1								5
	(45)–55	2	1											3
	Итого	6	7	4	19	31	61	31	24	8	4	5	14	214

Если же с изменением x эти ряды своего положения не изменяют или изменяют его лишь случайно, то y корреляционно не зависит от x ¹.

Табл. 27 обнаруживает корреляционную зависимость между процентом выполнения плана по валовой продукции и процентом выполнения плана по числу человеко-дней.

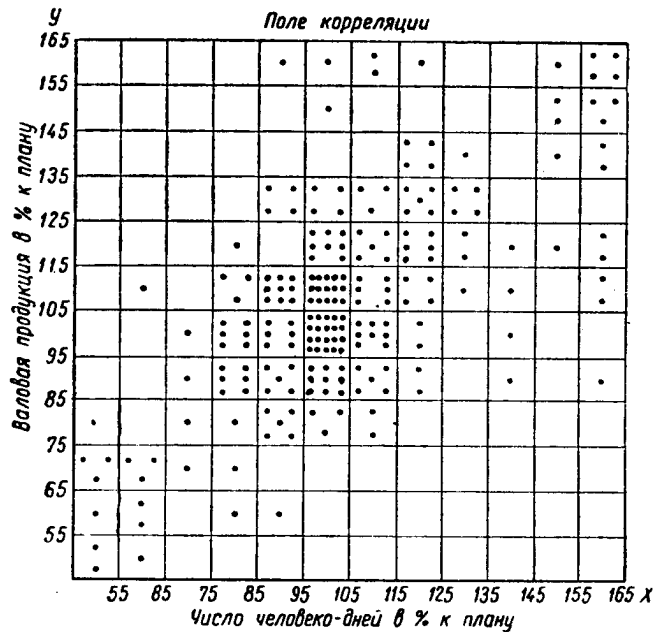


Рис. 20. Зависимость между валовой продукцией и числом отработанных человеко-дней (в % к плану) для 214 предприятий

Читателю рекомендуется, проведя аналогичный анализ табл. 26, убедиться в том, что между упругостью и размером замка поршневых колец имеет место корреляционная зависимость.

Те же выводы могут быть сделаны на основе анализа поля корреляции (см. рис. 19). Наличие у точек корреляционного поля тенденции смещаться снизу вверх по мере нашего передвижения слева направо свидетельствует о наличии корреляционной зависимости между y и x .

Корреляционная таблица примера 1 была составлена нами

¹ Приведенная формулировка содержит пока не вполне определенные понятия «случайного» и «закономерного» изменения «положения» рядов распределения. Более точное определение понятия корреляционной зависимости будет дано ниже.

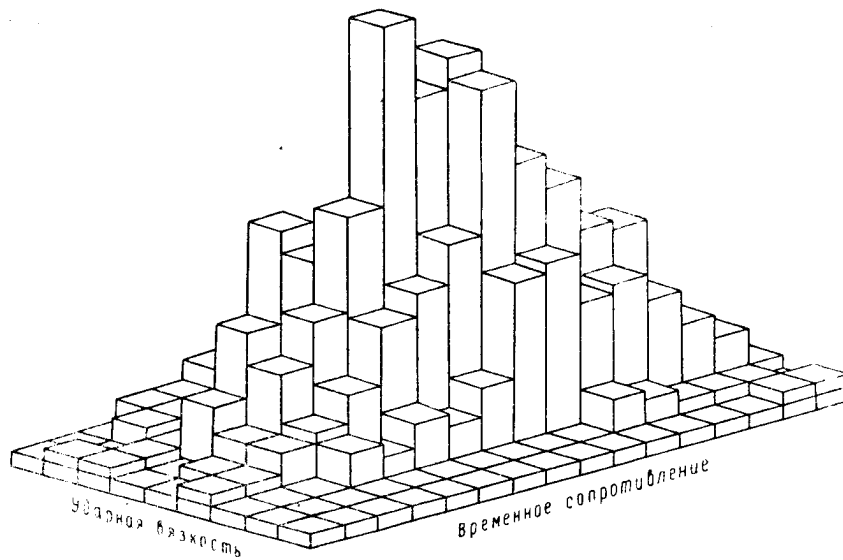


Рис. 21. Стереогрaмма распределения

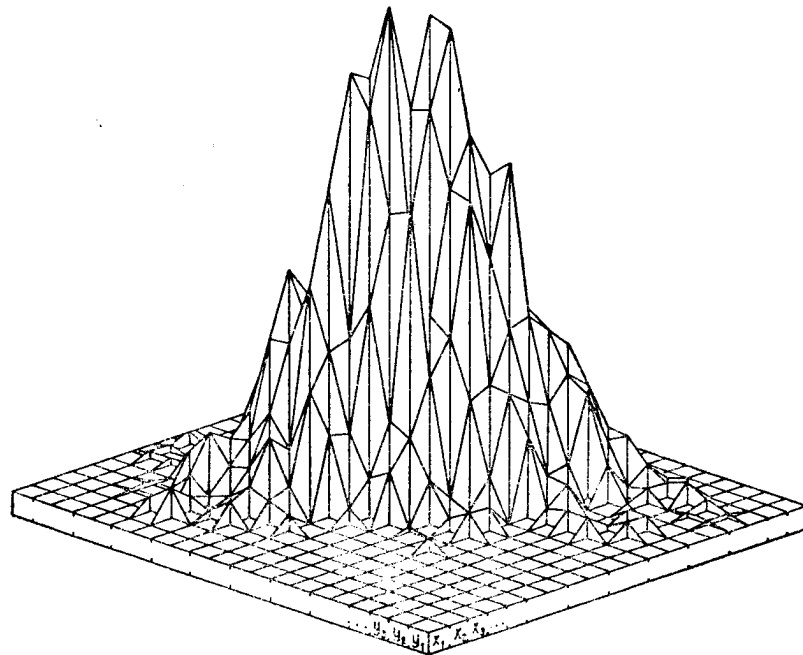


Рис. 22. Многогранник распределения

на основе поля корреляции. Для примера 2 построим поле корреляции на основе корреляционной таблицы.

Для этой цели на координатную плоскость (xy) накладываем сетку применительно к интервалам корреляционной таблицы и частоту каждой клетки изображаем в виде соответствующего числа точек, равномерно распределенных внутри клетки (рис. 20).

Подобное поле корреляции, составленное на основе корреляционной таблицы, называется *вторичным*.

Геометрическую форму систематизации опытного материала представляют также трехмерные модели — стереограммы и многогранники распределения. Частоты или плотности частот изображаются здесь столбиками или перпендикулярами, восстановленными к плоскости (xy) из соответствующих клеточек (см. рис. 21 и 22).

Этот способ изображения данных имеет принципиальное значение в теоретических вопросах, но на практике используется редко.

ГЛАВА IV

ИССЛЕДОВАНИЕ ФОРМЫ СВЯЗИ

§ 1. ЭМПИРИЧЕСКАЯ ЛИНИЯ РЕГРЕССИИ

Корреляционная зависимость выражается в закономерном смещении рядов распределения функции в связи с изменением аргумента. Это смещение может быть более или менее быстрым, оно может происходить в сторону увеличения или в сторону уменьшения функции. Изучая корреляционную зависимость, мы обязаны определить, в каком направлении и как быстро смещаются ряды распределения функции на тех или иных участках изменения аргумента. Необходимо поэтому иметь точную оценку положения рядов распределения функции на оси y . Из теории рядов распределения известно, что этой цели служит вычисление средних показателей.

Рассчитаем по табл. 27 средние величины для всех рядов распределения y , соответствующих заданным значениям x . Вычисления производим по формуле средней арифметической, как это обычно принято при исследовании корреляционной зависимости.

Для первого ряда распределения (при $x=50$):

y	m	my
70	3	210
60	1	60
50	2	100
Итого	6	370

$$\bar{y}_1 = \frac{\sum my}{\sum m} = \frac{370}{6} = 61,67.$$

Для второго ряда (при $x=60$):

y	m	my
110	1	110
100	0	0
90	0	0
80	0	0
70	3	210
60	2	120
50	1	50
Итого	7	490

$$\bar{y}_2 = \frac{\sum my}{\sum m} = \frac{490}{7} = 70,0.$$

Расчеты средних \bar{y}_i удобнее располагать в общей таблице, производя их параллельно для всех рядов и применяя упрощенный способ, как это показано в табл. 28.

Поясняем ход вычислений. Прежде всего варианты y заменяем упрощенными вариантами y' с помощью преобразования

$$y' = \frac{y - C_y}{i_y},$$

где C_y — новое начало отсчета;

i_y — интервал группировки по y .

В нашем случае $C_y = 100$, $i_y = 10$. Упрощенные варианты y' пишем в левом столбце табл. 28.

Далее умножаем частоту каждой клетки корреляционной таблицы на соответствующий вариант y' и произведение записываем в правом верхнем углу клетки. Подводим итоги этих произведений по столбцам и записываем в нижней строке таблицы (заголовков $\Sigma_i y'm$). Деля каждый итог на число наблюдений в столбце h_i , получаем средние \bar{y}_i в упрощенном масштабе (записываем строкой ниже).

Наконец, по формуле

$$\bar{y}_i = i_y \cdot \bar{y}'_i + C_y = 10\bar{y}'_i + 100$$

переводим упрощенные средние \bar{y}'_i в искомые средние \bar{y}_i (записываем строкой ниже).

Результаты подсчета изображаем графически на поле корреляции. Для этой цели из середины интервалов аргумента мысленно восстанавливаем ординаты, соответствующие значениям \bar{y}_i . Вершины ординат последовательно соединяем прямыми линиями. Полученная ломаная линия называется *эмпирической линией регрессии y по x* .

Линия регрессии показывает, как смещаются ряды распределения y с увеличением x или как в среднем изменяется y с увеличением x .

Рис. 23 обнаруживает отчетливую тенденцию к росту среднего процента выполнения плана по валовой продукции в связи с увеличением процента выполнения плана по числу человеко-дней. Поступательный ход линии регрессии несколько нарушается зигзагами, которые имеют случайный характер: они сильнее на тех участках линии, где наблюдений мало, и уменьшаются в тех частях графика, которые проходят среди значительного числа точек поля. Так, например, сильное западение линии регрессии наблюдается при $x=140$, где средняя вычислена всего лишь по четырем наблюдениям.

Эмпирическую линию регрессии нужно рассматривать в поле корреляции на фоне индивидуальных точек, которые она

Таблица 28

Расчет эмпирической линии регрессии

y'	x	50	60	70	80	90	100	110	120	130	140	150	160	Итого
6	160					1 6	1 6	2 12	1 6			1 6	4 24	10
5	150						1 5					2 10	3 15	6
4	140								4 16	1 4		1 4	2 8	8
3	130					4 12	4 12	3 9	5 15	4 12				20
2	120				1 2		8 16	5 10	6 12	2 4	1 2	1 2	2 4	26
1	110		1 1		3 3	9 9	15 15	6 6	4 4	1 1	1 1		2 2	42
0	100			1	6	6	20	8	2		1			44
-1	90			1 -1	6 -6	5 -5	9 -9	5 -5	2 -2		1 -1		1 -1	30
-2	80			1 -2	1 -2	5 -10	3 -6	2 -4						12
-3	70	3 -9	3 -9	1 -3	1 -3									8
-4	60	1 -4	2 -8		1 -4	1 -4								5
-5	50	2 -10	1 -5											3
Итого	h_i	6	7	4	19	31	61	31	24	8	4	5	14	214
	$\Sigma_i y'_m$	-23	-21	-6	-10	8	39	28	51	21	2	22	52	
	\bar{y}	-3,83	-3,00	-1,50	-0,53	0,26	0,64	0,90	2,13	2,63	0,50	4,40	3,71	
	\bar{y}_i	61,7	70,0	85,0	94,7	102,6	106,4	109,0	121,3	126,3	105,0	144,0	137,1	

осередняет. Большим «доверием» пользуются те части линии, которые проходят в «густых» местах корреляционного поля, и менее надежны те части линии, которые находятся в «разреженных» местах поля.

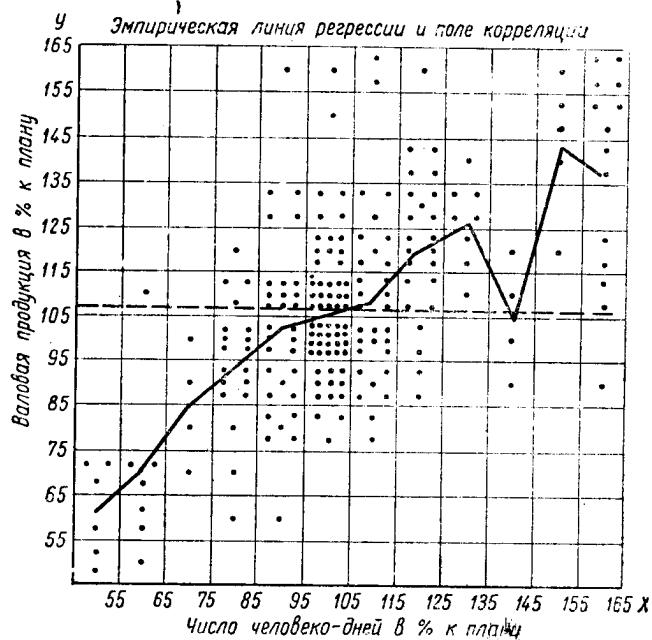


Рис. 23. Зависимость между валовой продукцией и числом отработанных человеко-дней (в % к плану) для 214 предприятий

В силу закона больших чисел можно утверждать, что по мере увеличения числа наблюдений случайные зигзаги линии будут сглаживаться и последняя будет принимать все более правильный, закономерный характер.

Предельное положение эмпирической линии регрессии, к которому она стремится при неограниченном увеличении числа наблюдений (и одновременно, но не столь быстром уменьшении величины интервалов аргумента), будем называть предельной теоретической линией регрессии¹. Ее нахождение составляет одну из основных проблем теории корреляции — проблему формы связи.

Мы можем теперь следующим образом уточнить определение корреляционной зависимости.

Переменная y находится в корреляционной зависимости от x , если: 1) каждому значению аргумента x соответствует ряд рас-

¹ В каком смысле здесь и в других аналогичных случаях понимать выражение «стремится к пределу при неограниченном увеличении числа наблюдений», будет уточнено в разделе III.

предела функции y и 2) предельная теоретическая линия регрессии y по x не параллельна оси абсцисс.

Если же предельная теоретическая линия регрессии y по x параллельна оси абсцисс, то y корреляционно не зависит от x .

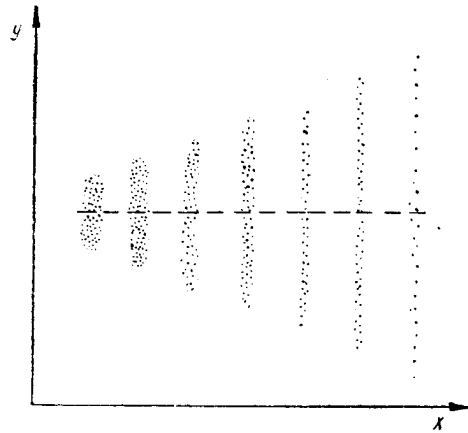


Рис. 24. Линия регрессии y по x параллельна оси абсцисс. С увеличением x рассеяние y увеличивается. y статистически зависит и корреляционно не зависит от x

Понятие корреляционной зависимости является частным случаем более общего понятия — зависимости статистической. Переменная y находится в статистической зависимости от x , если: 1) каждому значению x соответствует ряд распределения y и 2) с изменением x эти ряды закономерно изменяются.

Последний термин может означать не только изменение положения рядов, но и изменение рассеяния, степени асимметрии и других особенностей, характеризующих форму кривой распределения.

Если же с изменением x ряды распределения y не изменяются или изменяются случайно, то y статистически не зависит от x . Очевидно, что всякая корреляционная зависимость является одновременно статистической зависимостью. Однако не всякая статистическая зависимость является корреляционной зависимостью. На рис. 24 показан случай статистической зависимости y от x , не являющейся в то же время зависимостью корреляционной.

В тех случаях, когда характеристика положения ряда при помощи средней арифметической оказывается недостаточной и для этой цели прибегают к расчету других средних (например, медианы, моды), линии регрессии также рассчитываются по этим средним. Вопросы, требующие использования так называемых функциональных средних, рассматриваются в главе VI. Во всех таких случаях термин линия регрессии следует сопровождать указанием на форму принятых для ее расчета средних. Если нет дополнительного указания, то следует считать, что при построении линии регрессии используются средние арифметические.

§ 2. ИСТОЛКОВАНИЕ ЛИНИИ РЕГРЕССИИ

В § 1 главы III были указаны две основные задачи измерения связи в статистике. Покажем, что линия регрессии при определенных условиях дает решение первой из этих задач.

Задача была сформулирована следующим образом: установить, какова была бы зависимость между функцией и одним из ее аргументов, если бы прочие аргументы этой функции не изменялись (причем выяснить это на материале, где прочие аргументы на самом деле изменяются — варьируют). Выражение «прочие аргументы» не является достаточно определенным; его необходимо уточнить для того, чтобы поставленная задача имела смысл и могла получить определенное решение. В самом деле, в любом вопросе можно по-разному определять «прочие аргументы» и получать при этом существенно различные выражения искомой зависимости.

Пример. При массовом изготовлении цилиндрической детали (валика) линейные размеры и другие элементы этой детали обнаруживают вариацию. Между весом валика и другими варьирующими признаками имеет место следующая функциональная зависимость:

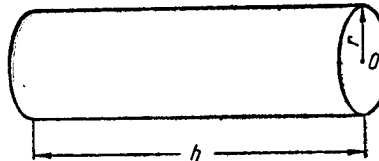


Рис. 25

$$g = \pi r^2 h \gamma (1 + \epsilon), \quad (4.2.1)$$

где g — вес валика;

r — радиус поперечного сечения;

h — длина валика;

γ — удельный вес материала;

$(1 + \epsilon)$ — коэффициент искажения цилиндрической формы (включающий и ошибки измерения);

π — постоянное число.

Из формулы (4.2.1) следует, что вес валика g изменялся бы пропорционально *квадрату* r , если бы «прочие аргументы» веса — h , γ , $(1 + \epsilon)$ — не варьировали.

Однако вместо этой формулы можно использовать зависимость

$$g = \frac{1}{2} S r \gamma (1 + \epsilon), \quad (4.2.2)$$

где S — величина боковой поверхности валика ($S = 2\pi r \cdot h$).

Из формулы (4.2.2) следует, что вес g изменялся бы пропорционально *первой степени* r , если бы прочие аргументы веса — S , γ , $(1 + \epsilon)$ — не варьировали.

Мы получили два существенно различных ответа на поставленный вопрос. Причина этому очевидна: не учтена зависимость, которая существует между «прочими» и основным аргументом; в данном случае между S и r . Между тем ясно, что, закрепляя аргументы, зависящие от основного, мы должны получить иное решение задачи по сравнению с тем, когда закрепляются аргу-

менты, от основного не зависящие. Отсюда следует, что в формулировке основной задачи мы обязаны так или иначе оговорить характер зависимости аргументов, которые мы называем «прочими», от аргумента, влияние которого на функцию мы собираемся проследить. Покажем, каким путем это осуществляется. Введем следующие допущения:

а) Будем предполагать, что влияние прочих аргументов на функцию сочетается с влиянием x по законам сложения и умножения. Иными словами, положим

$$y = u + v\varphi(x) + w\psi(x) + \dots, \quad (4.2.3)$$

где $\varphi(x)$, $\psi(x)$... — точные функции x (например, x , x^2 , x^3 , $\sin x$, $\ln x$, ... и т. д.);

u , v , w ... — прочие аргументы y , каждый из которых может оказаться результатом действия многих факторов.

Частым случаем соотношения (4.2.3) являются формулы (4.2.1) и (4.2.2) или, например, формула

$$y = u + vx^3 + w \frac{\sin x}{1+x} \text{ и др.}$$

б) «Прочие аргументы» положим *корреляционно не зависящими* от основного аргумента x .

При наличии условий (а) и (б) будем говорить, что прочие аргументы действуют *аддитивно* (сложением), *мультипликативно* (умножением) и *независимо от аргумента* x , или, короче, действуют по схеме (А — М — Н).

Основную задачу сформулируем теперь следующим образом:

Выяснить, какова была бы зависимость между функцией y и одним из ее аргументов x , если бы прочие аргументы функции, действующие по схеме (А — М — Н), не изменялись — находились на одном и том же среднем уровне. Причем выяснить это на материале, где прочие аргументы на самом деле изменяются, а их значения нам неизвестны (прочие аргументы являются неучитываемыми признаками).

Можно показать, что решение сформулированной выше основной задачи дается предельной теоретической линией регрессии y по x . Эта важная интерпретация линии регрессии была впервые указана проф. А. Я. Боярским [5].

Приводим доказательство.

$$\text{Пусть } y = u + v\varphi(x) + w\psi(x) + \dots,$$

где u , v , w ... — варьирующие аргументы функции, корреляционно не зависящие от аргумента x (схема А — М — Н).

В процессе вариации признак x может принимать различные значения: x_1, x_2, \dots, x_s , причем каждому значению x соответствует ряд распределения признака y (это легко себе представить, рассматривая последовательные столбцы частот в корреляционной таблице, связывающей y с x). Аналогично любому значению x соответствуют ряды распределения по каждому из признаков u, v, w, \dots

Дадим x значение x_1 и рассчитаем среднее значение \bar{y}_1 , соответствующее x_1 . Очевидно, что

$$\bar{y}_1 = \frac{\sum_1 y}{h_1},$$

где сумма $\sum_1 y$ берется для тех элементов совокупности, которые обладают значением x_1 ; h_1 — число таких элементов.

Но $\sum_1 y = \sum_1 u + \varphi(x_1) \sum_1 v + \psi(x_1) \sum_1 w + \dots$,

где $\sum_1 u, \sum_1 v, \sum_1 w, \dots$ берутся для тех элементов совокупности, которые обладают значением x_1 . Отсюда

$$\bar{y}_1 = \bar{u}_1 + \bar{v}_1 \varphi(x_1) + \bar{w}_1 \psi(x_1) + \dots, \quad (4.2.4)$$

где $\bar{u}_1, \bar{v}_1, \bar{w}_1, \dots$ представляют собой средние значения прочих аргументов, соответствующие заданному значению x_1 .

Аналогично

$$\left. \begin{aligned} \bar{y}_2 &= \bar{u}_2 + \bar{v}_2 \varphi(x_2) + \bar{w}_2 \psi(x_2) + \dots \\ \bar{y}_s &= \bar{u}_s + \bar{v}_s \varphi(x_s) + \bar{w}_s \psi(x_s) + \dots \end{aligned} \right\} \quad (4.2.4')$$

Следовательно, для эмпирической линии регрессии y по x имеет место уравнение

$$\bar{y}_i = \bar{u}_i + \bar{v}_i \varphi(x_i) + \bar{w}_i \psi(x_i) + \dots \quad (4.2.4'')$$

Но так как аргументы u, v, w, \dots корреляционно не зависят от x , то их средние значения в пределе при неограниченно большом числе наблюдений не должны изменяться с изменением x .

Следовательно, имеют место приближенные равенства:

$$\left. \begin{aligned} \bar{u}_1 &\approx \bar{u}_2 \approx \dots \approx \bar{u}_s \approx \bar{u}, \\ \bar{v}_1 &\approx \bar{v}_2 \approx \dots \approx \bar{v}_s \approx \bar{v}, \\ \bar{w}_1 &\approx \bar{w}_2 \approx \dots \approx \bar{w}_s \approx \bar{w}, \end{aligned} \right\} \quad (4.2.5)$$

которые с увеличением h_1, h_2, \dots, h_s становятся все более точными.

Отсюда для предельной теоретической линии регрессии y по x получаем уравнение

$$\bar{Y}(x) = \bar{u} + \bar{v}\varphi(x) + \bar{w}\psi(x) + \dots \quad (4.2.6)$$

Сопоставляя формулу (4.2.6) с формулой (4.2.3), мы убеждаемся в том, что предельная теоретическая линия регрессии дает решение поставленной задачи. Она показывает, как изменялась бы функция y с изменением аргумента x , если бы прочие ее аргументы u, v, w, \dots , действующие по схеме (А — М — Н), находились на одном и том же *среднем* уровне. Если числа наблюдений h_1, h_2, \dots, h_s невелики и равенства (4.2.5) осуществляются лишь приближенно, то колеблемость чисел $\bar{u}, \bar{v}, \bar{w}, \dots$ в уравнении (4.2.4'') создает зигзагообразный характер *эмпирической* линии регрессии.

Вспомним, что аргументы u, v, w, \dots являются *неучитываемыми* признаками. Нам, может быть, даже неизвестна материальная природа аргументов u, v, w, \dots , хотя понимание материальной природы этих аргументов существенно для уяснения реального смысла получаемых закономерностей.

Важно отметить, что указанный метод — метод расчета регрессии y по x — *позволяет освободиться от влияния всех аргументов функции y , действующих дополнительно к аргументу x по схеме (А — М — Н), каковы бы ни были конкретные значения этих аргументов и какова бы ни была материальная природа последних.*

В этом методе используются лишь данные о распределении статистической совокупности по двум признакам x и y .

Если аргументы u, v, w, \dots находятся в корреляции с x или, вообще говоря, не действуют по схеме (А — М — Н), то метод регрессии не освобождает зависимость между y и x от влияния изменчивости этих аргументов.

Для достижения поставленной цели необходимо применить более сильные статистические методы — методы множественной корреляции, которые, однако, опираются на *дополнительный учет всех находящихся в корреляции с x признаков*, влияние которых мы собираемся элиминировать, т. е. на более широкую статистическую информацию. В ряде случаев вопрос может быть решен и в рамках корреляции двух переменных при помощи метода функциональной регрессии, который мы рассматриваем в главе VI.

§ 3. ТЕОРЕТИЧЕСКАЯ ЛИНИЯ РЕГРЕССИИ

Из определения предельной теоретической линии регрессии, данного на стр. 72, вытекает и прямой путь ее нахождения: необходимо увеличивать число наблюдений (одновременно сокращая интервалы группировки аргумента) до тех пор, пока лежащая в основе закономерность не выявится с удовлетворяющей

нас точностью. Однако этот прямой путь обычно недоступен исследователю, который всегда имеет дело с ограниченным, а часто и не очень большим числом наблюдений. Чтобы в этих условиях обнаружить закономерность, отчетливо проявляющуюся лишь в массе наблюдений, необходимо уметь использовать *косвенные приемы анализа*.

Не следует, однако, полагать, что в результате их применения мы найдем действительную предельную теоретическую линию регрессии; мы можем получить лишь приближенную *оценку* последней, и для того, чтобы эта оценка была возможно лучшей, необходимо знать, какого рода информация может быть привлечена к делу, и уметь правильно распорядиться этой информацией.

Оценку, произведенную на основе ограниченного числа наблюдений, во избежание излишних терминов будем также называть теоретической линией регрессии, но без упоминания эпитета «предельная». Термин «предельная теоретическая линия регрессии» мы сохраняем лишь для действительной линии, к которой стремится эмпирическая линия регрессии при увеличении числа наблюдений.

Процесс нахождения теоретической линии регрессии заключается в выборе и обосновании типа кривой и расчете параметров ее уравнения.

Предельная теоретическая линия регрессии мыслится, как правило, в виде плавной кривой, выражающейся математическим уравнением того или иного типа. В таком же виде дается ее оценка, т. е. теоретическая линия регрессии. Процесс ее нахождения называется поэтому *выравниванием* эмпирической линии регрессии.

Запас кривых, который нам может предложить математический анализ, бесконечно разнообразен. Для того чтобы из этого запаса выбрать нужную кривую, необходимо прежде всего ориентироваться в важнейших типах функциональной зависимости. Отсылая читателя за указаниями к соответствующим руководствам и справочникам по математике¹, остановимся на семействе чаще всего используемых для рассматриваемой цели кривых, уравнения которых выражаются многочленами целых положительных степеней:

$$\begin{aligned}\bar{y}_x &= a + bx, \\ \bar{y}_x &= a + bx + cx^2, \\ y_x &= a + bx + cx^2 + dx^3\end{aligned}$$

и т. д.

(\bar{y}_x — ордината теоретической линии регрессии).

¹ См., например, И. Н. Бронштейн и К. А. Семендяев, Справочник по математике для инженеров и учащихся втузов. К. А. Семендяев, Эмпирические формулы, 1933.

Многочлен первой степени (прямая линия) выражает равномерный рост или падение функции. Многочлен второй степени (парабола второго порядка) выражает равнозамедленное или равноускоренное изменение y_x и может передать одну точку поворота функции (максимума или минимума). Многочлен третьей степени (парабола третьего порядка) может отразить две точки поворота функции, многочлен $s+1$ степени — s точек поворота (экстремальных точек). Чем выше степень многочлена и чем больше параметров содержит его уравнение, тем более сложные зависимости может передать соответствующая парабола (см. рис. 26).

Известно, что любая сколь угодно сложная непрерывная функция может быть с необходимой степенью точности воспроизведена многочленом s -го порядка, если выбрать s достаточно большим, правильно распорядиться параметрами уравнения и рассматривать вопрос в определенном, заранее заданном промежутке изменения аргумента¹.

Заметим также, что общий метод нахождения параметров уравнения регрессии (изложенный в следующей главе) технически наиболее просто осуществим в тех случаях, когда эти уравнения выражаются многочленами целых степеней. При использовании многочленов высокой степени особенно удобен метод последовательного повышения степени многочлена, теория которого создана великим русским математиком П. Л. Чебышевым².

Обладая столь совершенным инструментом отображения зависимости, каким является семейство многочленов, статистическая практика охотно к нему прибегает при построении линий регрессии. Может даже показаться, что мы не встречаем практической необходимости в использовании для этой цели каких-либо иных формул. Однако это не так. Прежде всего, как бы точно тот или иной многочлен не воспроизводил искомую зависимость в заданном интервале изменения аргумента, это само по себе не дает оснований полагать, что вне этого интервала данный многочлен будет соответствовать действительному ходу изменения функции. На рис. 33 приведены парабола второго порядка и прямая линия, которые в центральных интервалах близки друг к другу, а за их пределами расходятся.

Таким образом, с помощью многочленов достаточно высокого порядка мы можем воспроизвести значения произвольной функции *внутри* некоторого промежутка, т. е. решаем задачу приближенной *интерполяции*, но, вообще говоря, лишены возможности

¹ См. В. Л. Гончаров, Теория интерполирования и приближения функций, гл. II, Гостехиздат, 1954.

² А. К. Митропольский, Техника статистического исчисления, § 66, 1931.

В. С. Немчинов, Полиномы Чебышева и математическая статистика, М., 1946.

оценивать течение процесса вне заданного промежутка, т. е. не решаем задачи приближенной *экстраполяции*. Но дело не только в этом.

Сама по себе задача интерполяции представляет во многих случаях значительный интерес. Математическая формула, освобождающая эмпирический материал от случайных зигзагов и передающая в известных границах, хотя и приближенно, существенные черты искомой закономерности, является для научного и в особенности математического анализа гораздо более удобным инструментом, чем первоначальный ряд чисел или эмпирический график. Дело все же в том, что и задача приближенной интерполяции решается с помощью воспроизводящих функцию многочленов несовершенно и ее решение не всегда достигает цели по следующим причинам.

Мы увидим далее (§ 2 и 3 главы V), что способ расчета параметров уравнения искомой параболы основан на требовании максимальной близости ее к эмпирической линии регрессии. Чем выше степень параболы, тем больше параметров содержит ее уравнение, тем большей близости мы при этом достигаем. Но близость параболы к *эмпирической линии регрессии* не всегда означает близость этой кривой к *предельной теоретической линии регрессии*. Увеличивая порядок параболы и число параметров в ее уравнении, мы начинаем воспроизводить не только закономерное в ходе эмпирической линии, но и случайные зигзаги последней. Поэтому увеличение числа параметров выравнивающей кривой хотя и способствует лучшему приближению этой кривой к эмпирической линии регрессии, но одновременно может повести к удалению от действительной теоретической линии регрессии. Этим налагается предел увеличению точности выравнивания при помощи повышения степени многочленов, и отсюда возникает задача нахождения в каждом отдельном случае кривой, которая при меньшем числе параметров лучше передает закономерный ход линии регрессии по сравнению с параболой того или иного порядка.

Заметим, наконец, что безотносительно к задачам экстра- или интерполяции установление действительного типа зависимости в том или ином научном вопросе всегда способствует более глубокому проникновению в механизм действия факторов, обуславливающих эту зависимость, и поднимает теорию вопроса на более высокую ступень.

§ 4. ВЫБОР И ОБОСНОВАНИЕ ТИПА КРИВОЙ РЕГРЕССИИ

Для выбора и обоснования типа кривой регрессии нет универсального метода, пригодного «на все случаи жизни»: в каждом вопросе приходится избирать свои пути решения задачи. Вкратце характеризуя здесь эти пути и их разновидности, мы должны иметь в виду, что практически каждый из них используется не в чистом виде, а в сочетании с другими.

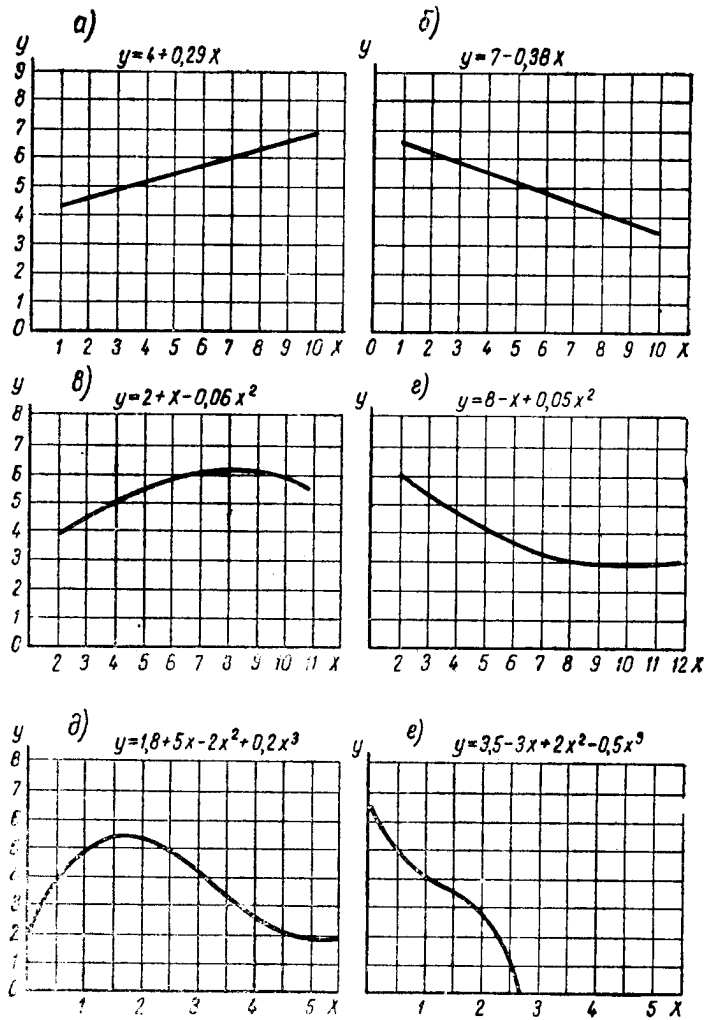
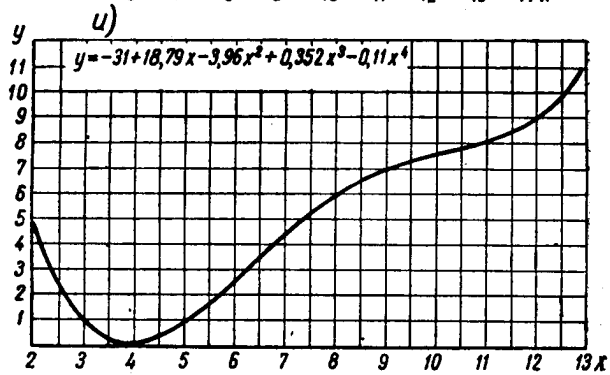
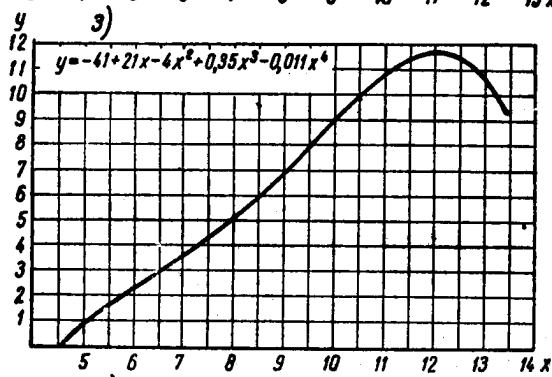
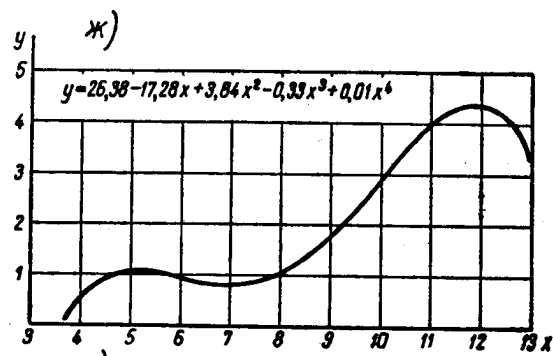


Рис. 26. Графики многочленов



целых степеней

Эмпирический путь

О типе теоретической кривой мы судим здесь по внешнему виду эмпирического графика регрессии. Этот путь основан на законе больших чисел. Он может дать удовлетворительные результаты при значительном количестве наблюдений, так как в этом случае эмпирический график воспроизводит характерные особенности теоретической линии. Мысленно устраняя те зигзаги эмпирического графика, которые можно предположить случайными, получаем представление о типе кривой, лежащей в его основе. Вопрос о том, какие зигзаги эмпирического графика считать случайными, исследователь обычно решает руководствуясь собственным опытом и интуицией, которые могут быть подкреплены соображениями, основанными на теории вероятностей. Если неизбежен выбор среди нескольких подходящих типов кривых, то предпочтение отдается тому типу, который зависит от меньшего числа параметров.

Эмпирический путь приводит, однако, при малом числе наблюдений к неясным результатам, так как наличие резких зигзагов эмпирической линии регрессии сильно затрудняет выявление закономерности.

На рис. 27 приводим две эмпирические линии регрессии упругости по размеру замка. Одна из этих линий построена для 200 поршневых колец на основе корреляционной табл. 26.

Другая линия построена на материале 30 поршневых колец, случайно выбранных из предыдущей совокупности. Количество наблюдений 30 (см. рис. 27,а) — оказалось недостаточным для того, чтобы уловить некоторое замедление в росте y_x , отчетливо наметившееся на рис. 27,б, где число наблюдений — 200.

Очевидно, что при малом числе наблюдений этим путем трудно уловить сколько-нибудь сложный тип зависимости, так как особенности последнего затушеваны случайными колебаниями. Привлечение дополнительных соображений становится здесь особенно желательным.

Теоретический путь

Излагая эмпирический путь исследования в своем «чистом виде», мы не интересовались вопросом о материальной природе признаков x и y , а исходили только из соображений, основанных на законе больших чисел. Между тем, если принять во внимание, что признаки x и y являются не абстрактными переменными, а величинами, заимствованными из конкретного вопроса, имеющего свою научную теорию, то станет ясным, что привлечение данных этой теории может оказать существенную, а иногда и решающую помощь при определении типа зависимости между y и x .

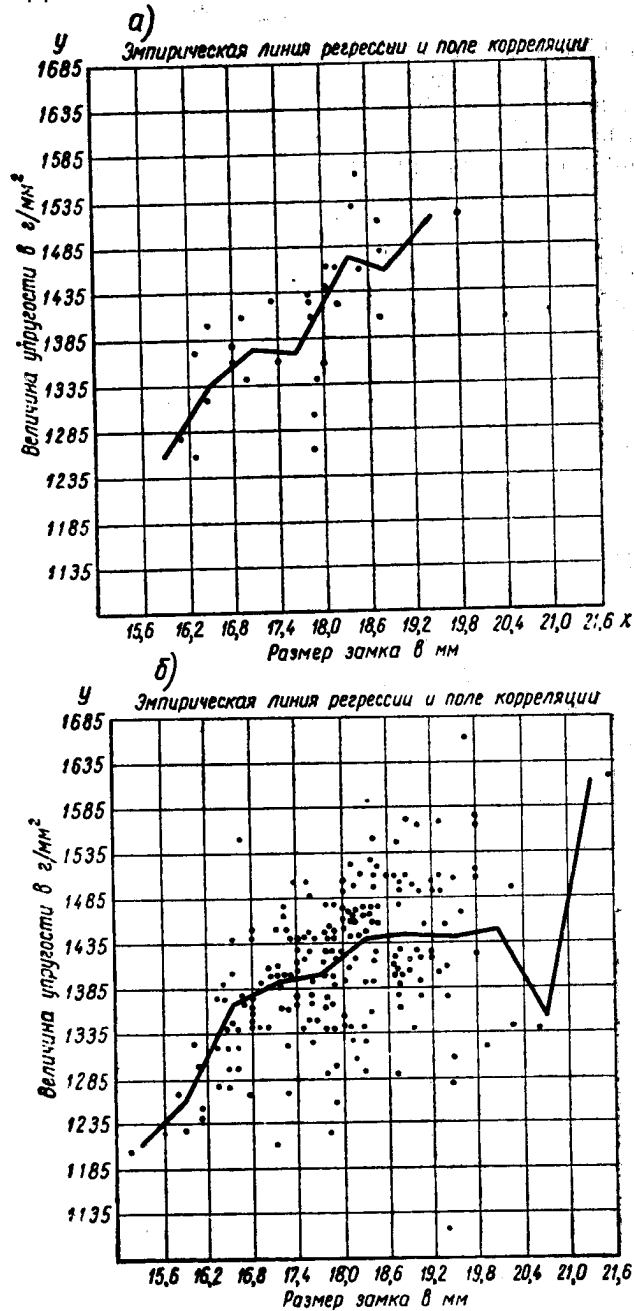


Рис. 27. а) Зависимость между упругостью и размером замка для 30 поршневых колец, б) Зависимость между упругостью и размером замка для 200 поршневых колец.

Итак, при установлении типа линии регрессии необходимо использовать теорию той конкретной науки, на базе которой возникает задача измерения связи между признаками y и x . Посмотрим, каковы же источники этой теории и какие пути ее использования возможны при решении поставленной задачи.

1. Опыт предыдущих исследований

Если подобного рода зависимости получались ранее и описаны в литературе, то мы учитываем опыт этих исследований. Так, например, если бы исследование зависимости между упругостью кольца и размером замка проводилось ранее различными авторами и если бы во всех случаях получаемые линии были близки к параболе второго порядка, то мы имели бы основания распространить этот тип кривой на наш материал и выдвинуть гипотезу квадратической зависимости. Разумеется, этот путь может быть использован лишь там, где проводились аналогичные исследования и где ставится задача проверки и конкретизации зависимости в новых условиях или для несколько отличного материала. Этот путь не применим в новых вопросах, где данная зависимость исследуется впервые.

2. Эксперимент

Этот путь заключается в искусственном сокращении случайных зигзагов эмпирической линии регрессии благодаря тому, что нам удастся в эксперименте закрепить или сократить изменчивость ряда факторов, вызывающих колеблемость нашей функции.

Так, например, если бы нам пришлось исследовать зависимость между объемом газа и давлением в условиях, когда случайные колебания температуры не устранены, мы получили бы график регрессии, осложненный случайными зигзагами, из-за которых при недостаточном числе наблюдений было бы трудно угадать действительный тип зависимости. Однако, опираясь на данные более точного эксперимента, где колебания температуры по возможности устранены, легко обнаружить гиперболическую зависимость между V и P .

Следует, однако, иметь в виду, что при закреплении или сокращении колеблемости побочных факторов, находящихся в корреляции с основным фактором, эмпирическая линия регрессии оказывается хотя и более закономерной, но систематически отличающейся от первоначальной линии. Лишь закрепление *независимых* от x и действующих по схеме (А—М—Н) факторов может привести к теоретически верному результату, отвечающему первоначальной постановке вопроса. Между тем без специальных расчетов, выводящих нас за рамки корреляции *двух переменных*, трудно установить, зависимы или независимы в данном процессе те или иные переменные, хотя об этом иногда можно делать более или менее обоснованные предположения. В этом — одна из слабых сторон рассматриваемого пути.

Второй слабой стороной является ограниченная возможность экспериментирования во многих вопросах и, в частности, в вопросах экономического характера. Между тем необходимость применения статистических методов измерения связи выступает с особой силой там, где экспериментальное устранение влияния побочных факторов затруднено или невозможно.

3. Логический анализ

В той или иной форме логический анализ материального объекта должен сопутствовать любому другому пути исследования типа зависимости. Значение и степень определенности выводов логического анализа обусловлены уровнем наших научных знаний в данном вопросе и характером имеющейся информации.

Форма, в которой выступает логический анализ зависимости, многообразна: от рассуждений, основанных на простом здравом смысле, до составления уравнений, связывающих функцию с рядом ее аргументов. Так, простой здравый смысл нам подсказывает, что с увеличением обеспеченности рабочей силой процент выполнения плана по выпуску продукции, при прочих равных условиях, должен расти. Поэтому такие «западения» эмпирической линии регрессии, как, например, на участке от $x=130$ до $x=140$ (см. рис. 23) мы расцениваем как случайные и считаем вправе отвлечься от них при исследовании закономерности.

На основе логического анализа материальной природы объекта мы составили уравнение (4.2.1), связывающее вес цилиндра с рядом его аргументов. Из этого уравнения следует, что если прочие аргументы варьируют независимо от радиуса, то вес цилиндра в среднем растёт пропорционально квадрату радиуса поперечного сечения.

В наиболее последовательном виде логический анализ направлен на выяснение общего характера зависимости между функцией y и ее аргументом x при «прочих равных», действующих по схеме (А—М—Н) условиях. Отсюда делается вывод, что этот тип зависимости должен лежать в основе линии регрессии y по x . Так, в упомянутом выше примере можно составить уравнение, связывающее процент выполнения плана предприятием по выпуску продукции (y) с процентом выполнения плана по числу человеко-дней (x)

$$y = 0,01 \cdot x \cdot v, \quad (4.4.1)$$

где v — процент выполнения плана по производительности труда (выработке на 1 человеко-день). Если v корреляционно не зависит от x , то линия регрессии y по x должна иметь вид $\bar{y}_x = cx$, где c — постоянное.

При составлении уравнений, связывающих функцию y с рядом ее аргументов, необходимо соблюдать условия схемы (А—М—Н)¹.

¹ Или более общей схемы Ф—А—М—Н (глава VI).

Этих условий, как известно, два: 1) аддитивное и мультипликативное действие «прочих» факторов и 2) корреляционная независимость «прочих» факторов от x . Большей частью нетрудно подобрать систему факторов, удовлетворяющую первому условию. Искусство анализа заключается в том, чтобы выбранная система факторов удовлетворяла и второму условию, т. е. требованию независимости. Практическое осуществление второго условия встречает, как правило, весьма значительные трудности. Так, например, легко составить равенство

$$y = vx,$$

где y — расход энергии на предприятии;
 x — объем выпущенной продукции;
 v — удельный расход энергии на единицу продукции.

Однако из этого равенства нельзя заключить о пропорциональности расхода энергии объему произведенной продукции, так как удельный расход энергии v обычно зависит от x . Несколько более удачной в данном случае является схема

$$y = u + vx,$$

где u — энергия, идущая на непроизводственные нужды;
 v — удельный расход энергии на производственные нужды предприятия.

Однако и при этой схеме коэффициенты u и v часто оказываются в корреляционной связи с x .

Следует иметь в виду, что прямая статистическая проверка независимости дополнительных факторов u, v, \dots от аргумента x , вообще говоря, исключена, так как u, v, \dots являются неучитываемыми признаками. Суждение о зависимости или независимости должно быть сделано на основе предварительных общих соображений о природе этих факторов. Разумеется, лишь крайне редко мы на основе этих предварительных соображений можем быть уверены в независимости u, v, \dots от x . Тем не менее во многих случаях мы можем вполне обоснованно предполагать слабую зависимость u, v, \dots от x , т. е. предполагать, что средние $\bar{u}_x, \bar{v}_x, \dots$ слабо изменяются с изменением x . В таких случаях задача нахождения регрессии y по x может получить приближенное решение. Укажем ход рассуждений, который приводит к решению задачи.

Пусть в ходе логического анализа установлено соотношение (4.2.3)

$$y = u + v\varphi(x) + w\psi(x) + \dots,$$

где $\varphi(x), \psi(x), \dots$ — известного вида функции x .

Следовательно, уравнение регрессии y по x должно иметь вид (4.2.4)

$$\bar{y}_x = \bar{u}_x + \bar{v}_x \varphi(x) + \bar{w}_x \psi(x) + \dots$$

Если есть основания полагать, что u, v, w, \dots корреляционно не зависят от x , то

$$\bar{u}_x = \bar{u}, \quad \bar{v}_x = \bar{v}, \quad \bar{w}_x = \bar{w}, \dots$$

и расчет теоретической линии регрессии сводится к вычислению средних $\bar{u}, \bar{v}, \bar{w}, \dots$ или, вообще говоря, к определению неизвестных параметров уравнения регрессии.

Так как u, v, w, \dots являются неучитываемыми признаками, то расчет $\bar{u}, \bar{v}, \bar{w}, \dots$ не может быть осуществлен прямым способом и определение параметров уравнения регрессии требует применения косвенных методов расчета. С таким косвенным методом расчета параметров — методом наименьших квадратов — мы познакомимся в следующем параграфе.

Если можно предполагать, что в схеме (4.2.3) u, v, w, \dots слабо зависят от x , то полагаем регрессии u, v, w, \dots по x приближенно прямолинейными

$$\bar{u}_x = a_1 + b_1 x; \quad \bar{v}_x = a_2 + b_2 x; \quad \bar{w}_x = a_3 + b_3 x; \dots$$

(a_i, b_i) неизвестны.

Подставляя эти выражения в формулу (4.2.4)

$$\bar{y}_x = \bar{u}_x + \bar{v}_x \varphi(x) + \bar{w}_x \psi(x) + \dots,$$

мы приходим к уравнению регрессии

$$\bar{y}_x = a + bx + c\varphi(x) + dx\varphi(x) + e\psi(x) + fx\psi(x) + \dots \quad (4.4.2)$$

с неизвестными параметрами, которые определяются способом, указанным в следующем параграфе.

Пример 1. В соотношении (4.4.1) мы не можем утверждать, что аргументы v и x независимы. Наоборот, часто имеются основания предполагать наличие тенденции, в силу которой предприятия, перевыполняющие план по числу человеко-дней, делают это за счет снижения процента выполнения плана по производительности труда, и, наоборот, повышение производительности труда ведет к экономии объема трудовых затрат. Следовательно, уравнение (4.4.1) не дает оснований считать y_x пропорциональным x .

Однако если имеются основания полагать, что указанная выше тенденция не резко выражена, то считаем приближенно регрессию v по x прямой

$$\bar{v}_x = a_1 + b_1 x,$$

где b_1 — малая отрицательная величина.

Подставляя это выражение в формулу (4.4.1), приходим к зависимости типа

$$\bar{y}_x = bx + cx^2, \quad (4.4.3)$$

где параметры b и c определяются способом, изложенным далее.

Рис. 23 обнаруживает слегка замедленный характер изменения ординаты эмпирической линии регрессии, что может служить подтверждением правильности наших предположений о типе кривой (парабола второго порядка с небольшим отрицательным коэффициентом при x^2).

Заметим, что в последнем примере мы могли бы осуществить и прямую статистическую проверку зависимости v от x , так как из соотношения $y = 0,01vx$ (4.4.1), где y и x — учитываемые признаки, мы легко определяем v , т. е. переводим этот аргумент в разряд учитываемых признаков.

В общем случае подобная операция, разумеется, неосуществима.

Пример 2. Из опыта известно, что при прочих равных условиях с увеличением количества изготавливаемой продукции сокращаются затраты предприятия на единицу продукции (себестоимость единицы). Это весьма важное экономическое положение имеет следующее обоснование. Различные виды затрат находятся в неодинаковом отношении к выпуску продукции.

Некоторые виды затрат незначительно меняются с увеличением выпуска. Таковы, например, расходы по амортизации и содержанию зданий, расходы по содержанию аппарата заводоуправления и т. д. Другие виды затрат, например затраты на сырье, инструмент, растут быстро, доходя до прямой пропорциональности выпуску продукции.

Для упрощения анализа можно грубо разделить все виды затрат на две группы: расходы *постоянные* относительно выпуска и расходы *пропорциональные* выпуску. При этом каждый вид затрат либо целиком относится к первой или ко второй группе (по принципу преобладания), либо подвергается с целью уточнения дальнейшему расчленению. Так, расходы по зарплате счетно-конторского персонала можно отнести к первой группе, расходы по зарплате производственных рабочих — ко второй группе затрат.

Таким образом, $z = v + ux$,
 где z — все затраты;
 x — выпуск продукции;
 v — затраты, постоянные относительно выпуска;
 u — переменные расходы на единицу продукции.

Отнеся все затраты к величине выпуска, получим соотношение

$$y = \frac{v}{x} + u, \quad (4.4.4)$$

где y — себестоимость единицы продукции.

Из этого соотношения видно, что себестоимость единицы с увеличением выпуска, при прочих равных условиях, должна изменяться по гиперболическому закону.

Однако равенство «прочих условий» практически нереально. В действительности эти «прочие условия» меняются от периода к периоду и от предприятия к предприятию, что находит свое выражение в изменчивости аргументов u и v . Эта изменчивость может находиться в известной корреляции с величиной выпуска.

Рост выпуска продукции, сопровождающийся, как правило, усовершенствованием технологии, улучшением технического оснащения и организации производства, влечет за собой сокращение как постоянных, так и переменных на единицу продукции расходов. Поэтому, переходя к уравнению регрессии y по x , мы записываем его в виде

$$\bar{y}_x = \bar{u}_x + \frac{\bar{v}_x}{x}, \quad (4.4.5)$$

где \bar{u}_x и \bar{v}_x — функции x .

Все же, если имеются основания предполагать слабую коррелированность u и v с x , то, полагая

$$\begin{aligned} \bar{u}_x &= a_1 + b_1x, \\ \bar{v}_x &= a_2 + b_2x, \end{aligned}$$

приходим к уравнению регрессии вида

$$\bar{y}_x = (a_1 + b_2) + b_1x + \frac{a_2}{x} \quad (4.4.6)$$

или

$$\bar{y}_x = a + cx + \frac{b}{x} \quad (4.4.6')$$

(коэффициент c ожидается малой и отрицательной величиной).

Пусть, например, нам известны конкретные данные о выпуске продукции и себестоимости единицы для предприятий, произво-

дящих однородную продукцию. В табл. 29 приведена группировка 44 предприятий по величине выпуска и указана средняя себестоимость в каждой группе предприятий.

Таблица 29

Выпуск продукции (в тыс. единиц)	Средняя себестоимость единицы	Число предприятий
До 1	16,50	6
1—2	13,75	6
2—3	13,31	8
3—4	12,50	7
4—5	13,52	4
5—6	12,75	4
6—7	12,30	3
7—8	12,83	2
8—9	12,28	2
9—10	12,34	2
		44

Данные этой таблицы изображены на рис. 28 в виде эмпирической линии регрессии.

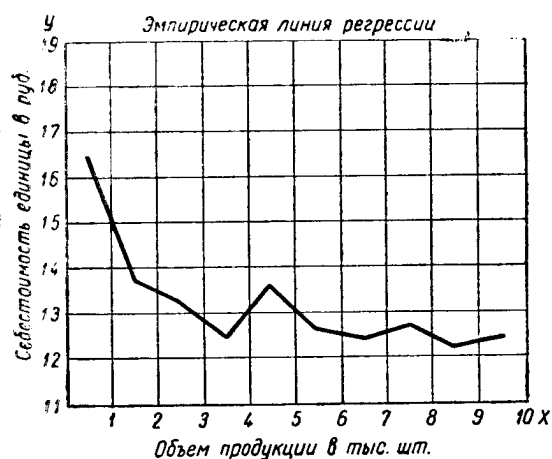


Рис. 28. Зависимость между себестоимостью единицы продукции и объемом продукции

Характер этой линии весьма напоминает гиперболу вида

$$\bar{y}_x = a + \frac{b}{x}.$$

Можно поэтому предположить, что коэффициент c в уравнении (4.4.6) для данного примера настолько мал, что им можно пренебречь и искать уравнение теоретической линии регрессии в виде

$$\bar{y}_x = a + \frac{b}{x}. \quad (4.4.7)$$

Параметры a и b этой зависимости мы находим в § 3 главы V и строим теоретическую линию регрессии (см. рис. 34).

Следует предостеречь от истолкования параметров a и b как средних значений u и v , т. е. как \bar{u} и \bar{v} . Подобное истолкование могло бы иметь место только при условии независимости u и v от x . В общем же случае смысл этих параметров уясняется из уравнения (4.4.6).

ГЛАВА V

РАСЧЕТ ПАРАМЕТРОВ УРАВНЕНИЯ РЕГРЕССИИ

§ 1. ПРЯМОЛИНЕЙНАЯ ЗАВИСИМОСТЬ. ДАННЫЕ НЕ СГРУППИРОВАНЫ

После того как выбран и обоснован тип уравнения регрессии, необходимо определить параметры, от которых зависит это уравнение. Процесс расчета параметров, составляющий содержание второго этапа работы, излагается в настоящем и последующих двух параграфах.

Пример. Предел выносливости стали не всегда удается установить путем непосредственного эксперимента в предоставляемый для этого на производстве срок. Поэтому желательна приближенная оценка предела выносливости на основании статистических характеристик металла. Одной из таких характеристик является предел прочности стали при кручении. Для установления зависимости между этими двумя признаками используем экспериментальные данные, полученные для 12 марок легированных сталей¹ (табл. 30).

Процесс установления зависимости разбивается на следующие стадии:

1. Прежде всего необходимо условиться, какой признак принимается в качестве функции и какой в качестве аргумента.

Так как целью исследования является установление возможности по величине предела прочности определять предел выносливости, то первый признак считаем аргументом x , второй — функцией y .

2. Данные табл. 30 необходимо привести в порядок — систематизировать. Для этой цели строим поле корреляции (рис. 29). Составление корреляционной таблицы опускаем ввиду малочисленности материала (12 наблюдений). Расположение точек в корреляционном поле свидетельствует о наличии отчетливо выраженной зависимости между признаками y и x . Точки, расположенные в левой части поля, оказываются ниже точек, расположенных

¹ С. Л. Жуков, К вопросу о связи предела выносливости со статическими характеристиками металлов, «Заводская лаборатория» № 11—12, 1945.

Таблица 30

**Предел прочности при кручении и предел
выносливости при изгибе для 12 марок
легированных сталей**

Марка стали	Предел проч- ности при кручении (в кг/мм ²)	Предел выно- сливости при изгибе (в кг/мм ²)
25 (У2)	51,1	25
45 (У4) нормализация	66,8	30
45 (У4) закалка	83,5	43
20ХНЗА (1)	80,8	43,6
37ХНЗА (1)	101,4	56,5
12ХНЗА (Х1Н)	109,3	58
20ХНЗА (Х2Н)	71,24	43
38ХМЮА	97,2	46
18ХНВА	109,0	62
25ХНФА	90,76	45
40ХНМА (4)	105,0	55
30ХГСА	89,0	44

справа: по мере увеличения абсцисс ординаты точек также увеличиваются.

3. Рассчитывается эмпирическая линия регрессии. Такой расчет требует значительного количества наблюдений и производится на основе корреляционной таблицы¹. Ввиду малочисленности экспериментального материала эту стадию опускаем.

4. Построение теоретической линии регрессии.

Характер расположения точек на корреляционном поле нам подсказывает, что теоретическая линия регрессии является прямой, т. е. выражается уравнением

$$\bar{y}_x = a + bx,$$

где a и b — неизвестные параметры.

Из бесчисленного множества прямых линий, которые можно провести на плоскости, следует выбрать одну, наилучшим образом соответствующую нашим экспериментальным данным. Каким образом это осуществить?

Легко видеть, что из трех различных прямых a , b , c , произвольно начерченных на корреляционном поле (рис. 30), прямая b лучше соответствует нашим эмпирическим данным, чем две

¹ Следует предостеречь против неправильного методического приема — последовательного соединения прямыми линиями точек корреляционного поля. Получающаяся при этом ломаная не является эмпирической линией регрессии. Последняя должна соединять точки, соответствующие средним значениям ординат, рассчитанным в каждом столбце таблицы.

остальные. Рассчитывая значения функции по прямой a или c , мы будем систематически завышать или занижать предел выносливости стали.

Найдем прямую, которая подходила бы к заданным точкам корреляционного поля ближе, чем какая-либо другая прямая. Близость точки к прямой измеряем отрезком ординаты, соеди-

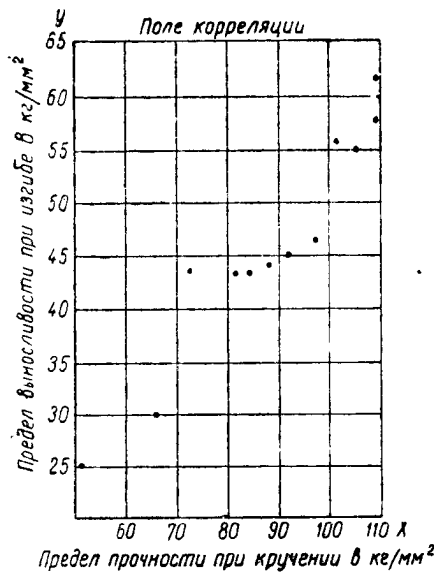


Рис. 29. Зависимость между пределом выносливости и пределом прочности для 12 марок стали.

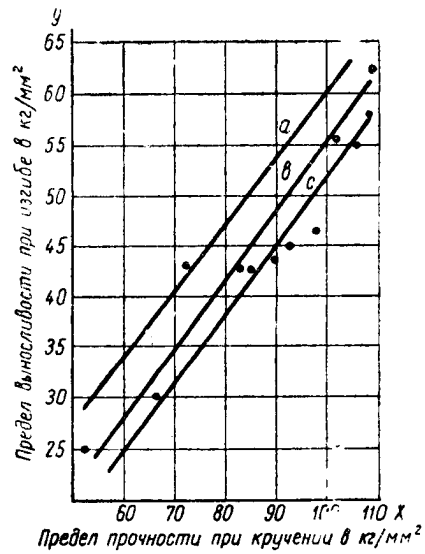


Рис. 30

няющим точку с этой прямой. На рис. 31 показан такой отрезок для точки A (искомая прямая проведена условно); он равен разности между фактической ординатой точки и соответствующей ординатой прямой линии

$$y - \bar{y}_x.$$

Отклонение точки от прямой (измеренное по ординате) может быть положительным или отрицательным в зависимости от того, расположена ли точка над прямой или ниже ее. Так как в данном случае нас не интересуют знаки отклонений, то возведем каждое из них в квадрат. образуем сумму квадратов отклонений для всех точек поля и потребуем, чтобы эта сумма была минимальной

$$\Sigma (y - \bar{y}_x)^2 = \min, \quad (5.1.1)$$

т. е. сумма квадратов отклонений фактических ординат от ординат, вычисленных по уравнению прямой, должна быть наименьшей.

Формула (5.1.1.) носит название *требования наименьших квадратов*, а способ отыскания параметров, основанный на этом требовании, называется *способом наименьших квадратов*.

Способ наименьших квадратов является одним из наиболее употребительных методов приближенной оценки по данным опыта неизвестных нам параметров уравнений. Наряду со способом наименьших квадратов существуют и другие методы, использующиеся для этой цели.

Показано, что при определенных условиях, часто с тем или иным приближением, осуществляющимся на практике, оценки параметров, получаемые по способу наименьших квадратов, являются наиболее точными в смысле теории вероятностей¹. Технически способ наименьших квадратов удобен в тех случаях, когда искомые параметры входят в уравнения связи линейным образом.

Исходя из требования (5.1.1), мы находим искомую прямую следующим образом.

Заменим в формуле (5.1.1) \bar{y}_x через $a + bx$ и обозначим левую часть формулы буквой f

$$f = \Sigma (y - a - bx)^2,$$

f — функция от неизвестных параметров a и b .

Значения a и b , удовлетворяющие минимуму функции f , найдутся, как известно, из уравнений:

$$\frac{\partial f}{\partial a} = 0, \quad \frac{\partial f}{\partial b} = 0. \quad (5.1.2)^2$$

Производя соответствующие выкладки, получаем

$$\frac{\partial f}{\partial a} = -2\Sigma(y - a - bx) = 0$$

или

$$\Sigma(y - a - bx) = 0;$$

откуда

$$\Sigma y = na + b\Sigma x$$

(n — число точек),

¹ См. [19].

² Уравнения (5.1.2) являются, как известно, лишь необходимыми условиями минимума. Однако для положительной квадратической функции f от параметров a , b эти условия являются и достаточными.

$$\frac{\partial f}{\partial b} = -2\Sigma(y - a - bx)x = 0$$

или

$$\Sigma(y - a - bx)x = 0;$$

откуда

$$\Sigma yx = a\Sigma x + b\Sigma x^2.$$

Мы пришли к системе двух уравнений первой степени относительно неизвестных параметров a и b

$$\left. \begin{aligned} \Sigma y &= na + b\Sigma x \\ \Sigma yx &= a\Sigma x + b\Sigma x^2 \end{aligned} \right\} \quad (5.1.3)$$

Система (5.1.3) называется *системой нормальных уравнений* по способу наименьших квадратов для определения параметров.

Суммы Σx , Σx^2 , Σy , Σyx подсчитываются по данным табл. 30; расчеты производятся при помощи табл. 31.

Таблица 31

x	y	x^2	xy	\bar{y}_x
51	25	2 601	1 275	24,7
67	30	4 489	2 010	33,9
84	43	7 056	3 612	43,7
81	44	6 561	3 564	42,0
101	57	10 201	5 757	53,5
109	58	11 881	6 322	58,1
71	43	5 041	3 053	36,3
97	46	9 409	4 462	51,2
109	62	11 881	6 758	58,1
91	45	8 281	4 095	47,8
105	55	11 025	5 775	55,8
89	44	7 921	3 916	46,6
$\Sigma = 1055$	552	96 347	50 599	—

Подставляя полученные величины в уравнения (5.1.3), решаем эти уравнения относительно неизвестных a и b

$$\begin{aligned} 552 &= 12a + 1055b & | : 12 \\ 50599 &= 1055a + 96347b & | : 1055 \\ \hline 46 &= a + 87,917b \\ 47,961 &= a + 91,324b \\ \hline 1,961 &= 3,407 \cdot b, \\ b &= \frac{1,961}{3,407} = 0,576, \end{aligned}$$

$$46 = a + 87,917 \cdot 0,576;$$

$$a = -50,64 + 46 = -4,64.$$

Отсюда уравнение искомой прямой

$$\bar{y}_x = -4,64 + 0,576x. \quad (5.1.4)$$

Придавая переменной x различные значения, получаем из уравнения (5.1.4) соответствующие значения \bar{y}_x . В столбце 5 табл. 31 указаны \bar{y}_x для всех x , записанных в столбце 1 таблицы.

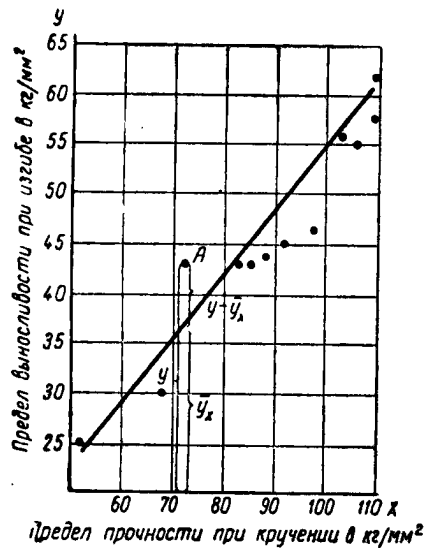


Рис. 31

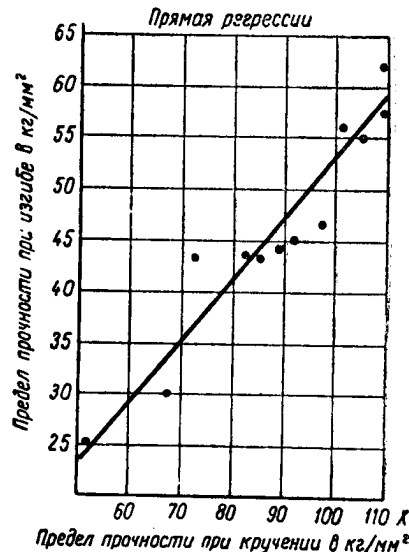


Рис. 32. Зависимость между пределом выносливости и пределом прочности для 12 марок стали

По этим данным чертим прямую регрессии в поле корреляции. Уравнение (5.1.4) или рис. 32 могут быть использованы для приближенной оценки предела выносливости стали, если имеются сведения о ее пределе прочности на кручение.

§ 2. ПРЯМОЛИНЕЙНАЯ ЗАВИСИМОСТЬ. ДАННЫЕ СГРУППИРОВАНЫ

Пример, рассмотренный в предыдущем параграфе, содержал небольшое количество данных ($n=12$). Результаты, полученные на таком материале, в значительной степени случайны и могут не повториться на другом однотипном материале. Более надежные результаты требуют массового материала, который обычно

подвергают группировке и сводят в корреляционную таблицу. Расчет теоретической линии регрессии производится на основе корреляционной таблицы. Для зависимости между валовой продукцией и числом человеко-дней (см. табл. 27) эмпирическая линия регрессии была рассчитана в § 1 главы IV (табл. 28, рис. 23).

Теоретическую линию регрессии можно положить в первом приближении прямой и искать ее уравнение в виде

$$\bar{y}_x = a + bx.$$

Параметры a и b определяются, как мы знаем, из уравнений (5. 1. 3)

$$\left. \begin{aligned} \Sigma y &= na + b\Sigma x \\ \Sigma yx &= a\Sigma x + b\Sigma x^2 \end{aligned} \right\}.$$

Однако в корреляционной таблице варианты x , y и их сочетания встречаются с определенными частотами. Поэтому все суммы в уравнениях (5. 1. 3): Σx , Σx^2 , Σy , Σxy удобнее вычислять во взвешенном виде. Чтобы указать на это, будем ставить точку над обозначением каждой суммы: $\dot{\Sigma}x$, $\dot{\Sigma}x^2$, $\dot{\Sigma}y$, $\dot{\Sigma}xy$. Каждое из этих выражений имеет, таким образом, двойной смысл: простой и взвешенной суммы.

В самом деле, пусть признак x принимает в корреляционной таблице значения (центры интервалов) x_1, x_2, \dots, x_s (s — число интервалов). Частоты этих значений обозначим соответственно через h_1, h_2, \dots, h_s . $\dot{\Sigma}x$ имеет, во-первых, смысл простой суммы Σx , подводящей итог значений x для всех элементов совокупности, т. е. *состоит из n член в x* . $\dot{\Sigma}x$ имеет, во-вторых, смысл взвешенной суммы $\Sigma h_i x_i$, подводящей итог по s различным значениям x_i , каждое из которых повторяется h_i раз. Эта сумма состоит, таким образом, из s членов вида $x_i h_i$. Аналогично для других сумм $\dot{\Sigma}x^2$, $\dot{\Sigma}y$, $\dot{\Sigma}xy$.

Представление каждой из этих сумм в простой или взвешенной форме определяется соображениями удобства. Числовые расчеты целесообразно всегда производить по формулам взвешенных сумм. Для теоретических выкладок мы будем часто прибегать к представлению этих сумм в простой форме.

Система нормальных уравнений запишется в виде

$$\left. \begin{aligned} \dot{\Sigma}y &= na + b\dot{\Sigma}x, \\ \dot{\Sigma}xy &= a\dot{\Sigma}x + b\dot{\Sigma}x^2. \end{aligned} \right\} \quad (5. 2. 1)$$

Вычисление сумм $\dot{\Sigma}x$, $\dot{\Sigma}x^2$, $\dot{\Sigma}y$, $\dot{\Sigma}xy$ проиллюстрируем на вспомогательном примере корреляционной табл. 32.

Таблица 32

Расчет теоретической линии регрессии

$$\bar{y}_x = a + bx$$

		Номера столбцов						
							1	2
$y \backslash x$		5	10	15	20	Итого l	ly	
50				2 ¹⁰⁰	1 ⁵⁰	3	150	
40			3 ¹²⁰	6 ²⁴⁰	6 ²⁴⁰	15	600	
30	3 ⁹⁰	8 ²⁴⁰	6 ¹⁸⁰	3 ⁹⁰	20	600		
20	5 ¹⁰⁰	4 ⁸⁰	1 ²⁰		10	200		
10	2 ²⁰				2	20		
Номера строк	1	Итого h	10	15	15	10	50= n	1570= $\Sigma ly = \dot{\Sigma}y$
	2	hx	50	150	225	200	625= $\Sigma hx = \dot{\Sigma}x$	
	3	hx^2	250	1500	3375	4000	9125= $\Sigma hx^2 = \dot{\Sigma}x^2$	
	4	Σmy	210	440	540	380	1570= $\Sigma ly = \dot{\Sigma}y$	
	5	$x\Sigma my$	1050	4400	8100	7600	21150= $\Sigma mxy = \dot{\Sigma}xy$	

$\dot{\Sigma}x$ рассчитывается как сумма произведений x на h , т. е. $\dot{\Sigma}x = \Sigma hx$. Результаты записываем в строке № 2. Аналогично $\dot{\Sigma}x^2 = \Sigma hx^2$ записываем в строке № 3. Обозначим буквой l

итоговые частоты y , буквой m — частоты клеток корреляционной таблицы (сочетаний x и y). $\dot{\Sigma}y$ рассчитывается как сумма произведений y на l

$$\dot{\Sigma}y = \Sigma ly.$$

Результат записываем в столбце № 2.

Одновременно $\dot{\Sigma}y$ может рассчитываться другим способом. Заметим, что

$$\dot{\Sigma}y = \dot{\Sigma}_1 y + \dot{\Sigma}_2 y + \dots + \dot{\Sigma}_s y = \Sigma_1 m y + \Sigma_2 m y + \dots + \Sigma_s m y,$$

где каждая из сумм $\dot{\Sigma}_i y = \Sigma_i m y$ подводит итоги произведений $m y$ для i -го столбца корреляционной таблицы. Произведения $m y$ записываем в верхних правых углах клеток таблицы, а их итоги $\Sigma_i m y$ — в строке № 4. Общий итог $\Sigma m y$ служит проверкой вычисления $\dot{\Sigma}y$, сделанного ранее другим способом. Наконец, для расчета $\Sigma x y$ заметим, что

$$\begin{aligned} \dot{\Sigma}x y &= \dot{\Sigma}_1 x y + \dot{\Sigma}_2 x y + \dots + \dot{\Sigma}_s x y = \\ &= x_1 \dot{\Sigma}_1 y + x_2 \dot{\Sigma}_2 y + \dots + x_s \dot{\Sigma}_s y = \\ &= x_1 \Sigma_1 m y + x_2 \Sigma_2 m y + \dots + x_s \Sigma_s m y. \end{aligned}$$

Следовательно, каждое $\Sigma_i m y$ мы умножаем на соответствующее x и складываем. Результат записываем в строке № 5.

Система нормальных уравнений для табл. 32

$$\left. \begin{aligned} 1570 &= 50a + 625b \\ 21150 &= 625a + 9125b \end{aligned} \right\}$$

Решая эту систему, получим $a = 16,875$, $b = 1,162$.
Откуда

$$\bar{y}_x = 16,875 + 1,162 \cdot x.$$

Вычисления значительно упрощаются, если воспользоваться заменой переменных

$$x' = \frac{x - C_x}{i_x}, \quad y' = \frac{y - C_y}{i_y},$$

где C_x и C_y — новые начала отсчета;
 i_x, i_y — интервалы по x и y .

Для табл. 32 примем $C_x = 10$, $C_y = 30$, $i_x = 5$, $i_y = 10$.

Отыскиваем параметры прямой $\bar{y}'_x = a' + b'x'$ с помощью системы нормальных уравнений

$$\begin{aligned} \Sigma y' &= na' + b'\Sigma x', \\ \Sigma x'y' &= a'\Sigma x' + b'\Sigma x'^2. \end{aligned} \quad (5. 2. 2)$$

Вычисления располагаем, как показано в табл. 33.

Таблица 33

Расчет теоретической линии регрессии

$$\bar{y}'_x = a' + b'x'$$

	x'	-1	0	1	2	Номера столбцов		
y'	y	5	10	15	20	Итого l	ly'	
2	50			2 ⁴	1 ²	3	6	
1	40		3 ³	6 ⁶	6 ⁶	15	15	
0	30	3 ⁰	8 ⁰	6 ⁰	3 ⁰	20	0	
-1	20	5 ⁻⁵	4 ⁻⁴	1 ⁻¹		10	-10	
-2	10	2 ⁻⁴				2	-4	
Номера строк	1	Итого h	10	15	15	10	50	7
	2	hx'	-10	0	15	20	25	-
	3	hx'^2	10	0	15	40	65	-
	4	$\Sigma_i m y'$	-9	-1	9	8	7	-
	5	$x' \Sigma_i m y'$	9	0	9	16	34	-

Система нормальных уравнений для табл. 33

$$\left. \begin{aligned} 7 &= 50a' + 25b' \\ 34 &= 25a' + 65b' \end{aligned} \right\}$$

Откуда

$$a' = -0,151, \quad b' = 0,581,$$

$$\bar{y}_{x'} = -0,151 + 0,581x'.$$

От параметров a' , b' остается перейти к параметрам a , b . Для этой цели подставим в уравнение $\bar{y}_{x'} = a' + b'x'$ выражения:

$$x' = \frac{x - C_x}{i_x}; \quad \bar{y}_{x'} = \frac{\bar{y}_x - C_y}{i_y}, \quad (5.2.3)$$

$$\frac{\bar{y}_x - C_y}{i_y} = a' + b' \frac{x - C_x}{i_x}.$$

После преобразования получаем

$$\bar{y}_x = \underbrace{\left[C_y + i_y a' - b' \frac{i_y}{i_x} C_x \right]}_a + \underbrace{b' \frac{i_y}{i_x}}_b x.$$

Отсюда формулы перехода от параметров a' , b' к параметрам a , b :

$$b = b' \frac{i_y}{i_x}, \quad a = C_y + i_y a' - b' C_x \frac{i_y}{i_x}. \quad (5.2.4)$$

Для табл. 33

$$b = 0,581 \cdot \frac{10}{5} = 1,162,$$

$$a = 30 - 10 \cdot 0,151 - 11,62 = 16,87.$$

Пришли к тем же значениям, которые были получены ранее другим путем.

Проведя аналогичные расчеты для зависимости между валовой продукцией и числом отработанных человеко-дней (см. табл. 27), мы получили уравнение регрессии

$$\bar{y}_x = 42,79 + 0,622x.$$

По этому уравнению строим прямую регрессии на поле корреляции (рис. 33).

Мы видим, что приросту выполнения плана по числу человеко-дней на 10% соответствует средний прирост выполнения плана по валовой продукции на 6,22%.

Отметим одно весьма важное обстоятельство. Заменим Σx через Σhx и Σx^2 — через Σhx^2 , как это было указано ранее.

Далее, выражения $\dot{\Sigma}y$ и $\dot{\Sigma}xy$ преобразуем следующим образом:

$$\begin{aligned} \dot{\Sigma}y &= \dot{\Sigma}_1y + \dot{\Sigma}_2y + \dot{\Sigma}_3y + \dots + \dot{\Sigma}_sy = \\ &= h_1 \frac{\dot{\Sigma}_1y}{h_1} + h_2 \frac{\dot{\Sigma}_2y}{h_2} + h_3 \frac{\dot{\Sigma}_3y}{h_3} + \dots + h_s \frac{\dot{\Sigma}_sy}{h_s} = \\ &= h_1 \bar{y}_1 + h_2 \bar{y}_2 + \dots + h_s \bar{y}_s = \Sigma h \bar{y}_i, \end{aligned}$$

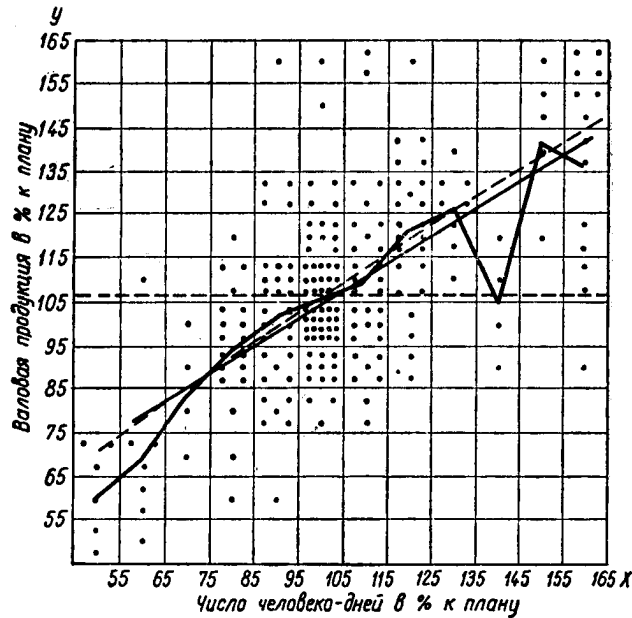


Рис. 33. Зависимость между валовой продукцией и числом отработанных человеко-дней (в % к плану) для 214 предприятий. Прямая и парабола регрессии 2-го порядка, эмпирическая линия регрессии и поле корреляции

где через $\bar{y}_1, \bar{y}_2, \dots, \bar{y}_s$, или короче \bar{y}_i , обозначены средние значения y в каждом столбце корреляционной таблицы (эмпирические средние)

$$\begin{aligned} \dot{\Sigma}xy &= x_1 \dot{\Sigma}_1y + x_2 \dot{\Sigma}_2y + \dots + x_s \dot{\Sigma}_sy = \\ &= h_1 x_1 \bar{y}_1 + h_2 x_2 \bar{y}_2 + \dots + h_s x_s \bar{y}_s = \Sigma h x \bar{y}_i. \end{aligned}$$

Уравнения (5.2.1) представятся в следующем виде:

$$\left. \begin{aligned} \Sigma h \bar{y}_i &= na + b \Sigma hx \\ \Sigma h x \bar{y}_i &= a \Sigma hx + b \Sigma hx^2 \end{aligned} \right\}. \quad (5.2.5)$$

Система нормальных уравнений (5.1.3) была получена из условия наибольшей близости искомой прямой к *индивидуальным* точкам корреляционного поля

$$\Sigma (y - \bar{y}_x)^2 = \min.$$

Уравнения же (5.1.4) получены из уравнений (5.1.3) тождественным преобразованием. Поэтому прямая регрессии, параметры которой найдены из уравнений (5.2.5), удовлетворяет тому же условию (5.1.1.).

Однако в то же время уравнения (5.2.5) можно получить из другого условия — условия наибольшей близости искомой прямой к *эмпирической линии регрессии*.

В самом деле, потребуем, чтобы *взвешенная сумма квадратов отклонений ординат эмпирической линии регрессии от ординат искомой прямой была наименьшей*

$$\Sigma h (\bar{y}_i - \bar{y}_x)^2 = \min. \quad (5.2.6)$$

Подставляя в формулу (5.2.6) вместо \bar{y}_x его выражение $a + bx$, обозначая левую часть формулы буквой f , дифференцируем частным образом f по a и по b и приравняем производные нулю

$$\frac{\partial f}{\partial a} = 0, \quad \frac{\partial f}{\partial b} = 0$$

или

$$\begin{aligned} -2\Sigma h (\bar{y}_i - a - bx) &= 0, \\ -2\Sigma hx (\bar{y}_i - a - bx) &= 0. \end{aligned}$$

Отсюда мы приходим к формулам (5.2.5):

$$\left. \begin{aligned} \Sigma h \bar{y}_i &= na + b\Sigma hx \\ \Sigma hx \bar{y}_i &= a\Sigma hx + b\Sigma hx^2 \end{aligned} \right\}$$

Следовательно, прямая, полученная из условия наибольшей близости к точкам корреляционного поля, оказалась в наибольшей близости и к эмпирической линии регрессии. Эта прямая выравнивает эмпирическую линию регрессии.

Условия (5.1.1.) и (5.2.6) эквивалентны.

В § 3 главы VII будет указан другой прием расчета прямой регрессии.

§ 3. КРИВОЛИНЕЙНАЯ ЗАВИСИМОСТЬ

Требование наименьших квадратов сохраняется и при расчете параметров криволинейной регрессии.

Пример. Зависимость между валовой продукцией и числом человеко-дней (в процентах к плану) рассматривалась в первом

приближении как прямолинейная (см. § 2). Однако по соображениям, указанным на стр. 90, здесь может быть принята квадратическая функция, т. е. парабола второго порядка ¹

$$\bar{y}_x = a + bx + cx^2. \quad (5.3.1)$$

Параметры a , b , c определяются из требования наименьших квадратов

$$f = \Sigma (y - \bar{y}_x)^2 = \Sigma (y - a - bx - cx^2)^2 = \min; \quad (5.3.2)$$

откуда

$$\frac{\partial f}{\partial a} = 0, \quad \frac{\partial f}{\partial b} = 0, \quad \frac{\partial f}{\partial c} = 0,$$

что дает после дифференцирования систему уравнений

$$\left. \begin{aligned} \Sigma y &= na + b\Sigma x + c\Sigma x^2 \\ \Sigma xy &= a\Sigma x + b\Sigma x^2 + c\Sigma x^3 \\ \Sigma x^2y &= a\Sigma x^2 + b\Sigma x^3 + c\Sigma x^4 \end{aligned} \right\}. \quad (5.3.3)$$

Решая систему относительно a , b , c , находим искомые значения параметров.

Когда материал сгруппирован и представлен в корреляционной таблице, суммы в уравнениях (5.3.3) вычисляются по взвешенным формулам и система (5.3.3) запишется так:

$$\left. \begin{aligned} \dot{\Sigma} y &= na + b\dot{\Sigma} x + c\dot{\Sigma} x^2 \\ \dot{\Sigma} xy &= a\dot{\Sigma} x + b\dot{\Sigma} x^2 + c\dot{\Sigma} x^3 \\ \dot{\Sigma} x^2y &= a\dot{\Sigma} x^2 + b\dot{\Sigma} x^3 + c\dot{\Sigma} x^4 \end{aligned} \right\}. \quad (5.3.4)$$

Проследим технику расчета параметров, используя данные табл. 32, но предполагая не линейную, а параболическую зависимость y от x .

Введем упрощающее преобразование (5.2.3)

$$x' = \frac{x - C_x}{i_x}, \quad y' = \frac{y - C_y}{i_y},$$

где $C_x = 10$; $C_y = 30$; $i_x = 5$; $i_y = 10$,

¹ Произведенный выше логический анализ привел к зависимости вида $\bar{y}_x = bx + cx^2$.

Свободный параметр a присоединен к уравнению по соображениям, связанным с техникой способа наименьших квадратов. При наличии свободного члена в уравнении регрессии осуществляется разложение дисперсии y на части, обусловленную и необусловленную влиянием фактора x , что важно при построении показателей тесноты зависимости (§ 2 главы VII).

Если логический анализ верен, то a окажется небольшим.

и будем искать зависимость

$$\bar{y}_{x'} = a' + b'x' + c'x'^2.$$

Параметры a' , b' , c' определяются из системы нормальных уравнений:

$$\left. \begin{aligned} \sum y' &= na' + b'\sum x' + c'\sum x'^2 \\ \sum x'y' &= a'\sum x' + b'\sum x'^2 + c'\sum x'^3 \\ \sum x'^2 y' &= a'\sum x'^2 + b'\sum x'^3 + c'\sum x'^4 \end{aligned} \right\}. \quad (5.3.5)$$

Расчет всех сумм, являющихся коэффициентами при неизвестных и свободными членами в уравнениях (5.3.5), производим в табл. 34.

Приходим к системе нормальных уравнений:

$$7 = 50a' + 25b' + 65c',$$

$$34 = 25a' + 65b' + 85c',$$

$$32 = 65a' + 85b' + 185c'.$$

Решая эту систему, получим

$$a' = -0,0233, \quad b' = 0,740, \quad c' = -0,159.$$

Следовательно, зависимость между $\bar{y}_{x'}$ и x' выражается уравнением

$$\bar{y}_{x'} = -0,0233 + 0,740x' - 0,159x'^2.$$

Произведя в этом уравнении замену переменных

$$x' = \frac{x - C_x}{i_x}, \quad \bar{y}_{x'} = \frac{\bar{y}_x - C_y}{i_y},$$

получим

$$\frac{\bar{y}_x - 30}{10} = -0,0233 + 0,740 \frac{x - 10}{5} - 0,159 \frac{(x - 10)^2}{25}.$$

Откуда

$$\bar{y}_x = 8,607 + 2,752x - 0,0636x^2.$$

Можно было бы воспользоваться и готовыми формулами перехода от параметров a' , b' , c' к параметрам a , b , c , выведя эти формулы приемом, указанным выше.

Таблица 34

Расчет теоретической линии регрессии

$$\bar{y}_x = a' + b'x' + c'x'^2$$

		x'	-1	0	1	2	Номера столбцов	
							1	2
y'	$y \backslash x$		5	10	15	20	Итого	ly'
2	50				2 ⁴	1 ²	3	6
1	40			3 ³	6 ⁶	6 ⁶	15	15
0	30	3 ⁰	8 ⁰	6 ⁰	3 ⁰		20	0
-1	20	5 ⁻⁵	4 ⁻⁴	1 ⁻¹			10	-10
-2	10	2 ⁻⁴					2	-4
Номера строк	1	Итого	10	15	15	10	50	$7 = \sum y'$
	2	hx'	-10	0	15	20		$25 = \sum x'$
	3	hx'^2	10	0	15	40		$65 = \sum x'^2$
	4	hx'^3	-10	0	15	80		$85 = \sum x'^3$
	5	hx'^4	10	0	15	160		$185 = \sum x'^4$
	6	$\sum_i m y'$	-9	-1	9	8		$7 = \sum y'$
	7	$x' \sum_i m y'$	9	0	9	16		$34 = \sum x' y'$
	8	$x'^2 \sum_i m y'$	-9	0	9	32		$32 = \sum x'^2 y'$

Подобным же образом было рассчитано уравнение регрессии для зависимости между валовой продукцией и числом отработанных человеко-дней

$$\bar{y}_x = 35,42 + 0,745x' - 0,000322x'^2$$

Графическое изображение этого уравнения дано на соответствующем поле корреляции (см. рис. 33). Теоретическая линия регрессии (парабола второго порядка) показывает слегка замедленный характер изменения ординаты и близко подходит к эмпирической линии регрессии.

Заметим, что система нормальных уравнений (5.3.4), так же как и при расчете параметров прямой линии, может быть представлена в виде

$$\left. \begin{aligned} \Sigma h \bar{y}_i &= na + b \Sigma hx + c \Sigma hx^2 \\ \Sigma hx \bar{y}_i &= a \Sigma hx + b \Sigma hx^2 + c \Sigma hx^3 \\ \Sigma hx^2 \bar{y}_i &= a \Sigma hx^2 + b \Sigma hx^3 + c \Sigma hx^4 \end{aligned} \right\}. \quad (5.3.6)$$

Таким образом, система (5.3.6) вытекает из условия (5.3.2). Однако уравнения (5.3.6) могут быть непосредственно получены и из условия

$$\Sigma (\bar{y}_i - \bar{y}_x)^2 h = \min.$$

Поэтому теоретическая линия регрессии оказывается в наибольшей близости не только к точкам поля, но и к эмпирической линии регрессии, т. е. решает задачу выравнивания последней.

Аналогичным путем определяются параметры уравнения регрессии, когда последнее выражается в виде *многочлена заданной степени k*

$$\bar{y}_x = a + a_1 x + a_2 x^2 + \dots + a_k x^k, \quad (5.3.7)$$

и в наиболее общем случае, когда \bar{y}_x предполагается *линейной функцией от параметров a, b, c*

$$\bar{y}_x = a + bp(x) + cq(x) + \dots, \quad (5.3.8)$$

где $p(x)$, $q(x)$ — любые функции от x , не зависящие от каких-либо неизвестных параметров. Так, из требования наименьших квадратов, записанного в виде

$$\Sigma (y - \bar{y}_x)^2 = \min,$$

мы приходим в случае (5.3.6) к системе нормальных уравнений

$$\left. \begin{aligned} \Sigma y &= na + b \Sigma p(x) + c \Sigma q(x) + \dots \\ \Sigma yp(x) &= a \Sigma p(x) + b \Sigma p^2(x) + c \Sigma q(x)p(x) + \dots \\ \Sigma yq(x) &= a \Sigma q(x) + b \Sigma p(x)q(x) + c \Sigma q^2(x) + \dots \end{aligned} \right\}. \quad (5.3.9)$$

Систему (5.3.9) решаем относительно неизвестных a, b, c, \dots , предварительно определив суммы:

$$\Sigma p(x), \quad \Sigma q(x), \quad \Sigma p^2(x), \quad \Sigma q(x) \cdot p(x), \dots$$

Полезно запомнить следующее правило составления уравнений (5.3.9), которое позволяет обойтись без дифференцирования функции

$$f = \Sigma (y - \bar{y}_x)^2.$$

Уравнение искомой линии регрессии (5.3.8)

$$\bar{y}_x = a + bp(x) + cq(x) + \dots$$

последовательно умножается на коэффициенты при параметрах, после чего берется сумма в обеих частях уравнения.

Так, коэффициент при параметре a равен 1; после суммирования (и отбросив значки при y) приходим к первому нормальному уравнению (5.3.9). Коэффициент при b равен $p(x)$; после умножения уравнения (5.3.8) на $p(x)$ и суммирования, отбросив значки при y , приходим ко второму нормальному уравнению и т. д.

В виде примера рассчитаем параметры уравнения регрессии себестоимости продукции по величине выпуска

$$\bar{y}_x = a + \frac{b}{x}$$

по данным табл. 29.

Система нормальных уравнений, представленная во взвешенном виде

$$\left. \begin{aligned} \Sigma h \bar{y}_i &= na + b \Sigma \frac{h}{x} \\ \Sigma h \frac{\bar{y}_i}{x} &= a \Sigma \frac{h}{x} + b \Sigma \frac{h}{x^2} \end{aligned} \right\} \quad (5.3.10)$$

Расчеты сумм производим в табл. 35. Следовательно,

$$\left. \begin{aligned} 592,36 &= 44a + 23,991b \\ 356,47 &= 23,991a + 29,005b \end{aligned} \right\}$$

Откуда

$$a = 12,318, \quad b = 2,101$$

и

$$\bar{y}_x = 12,318 + \frac{2,101}{x}.$$

Теоретическая линия регрессии изображена на рис. 34.

В случае, когда \bar{y}_x не является линейной функцией от

Таблица 35

x	\bar{y}_i	h	$\frac{h}{x}$	$\frac{h}{x^2}$	hy_i	$\frac{hy_i}{x}$
0,5	16,50	6	12,000	24,000	99,00	198,00
1,5	13,75	6	4,000	2,667	82,50	55,00
2,5	13,31	8	3,200	1,280	106,48	42,59
3,5	12,50	7	2,000	0,571	87,50	25,00
4,5	13,52	4	0,889	0,198	54,08	12,02
5,5	12,75	4	0,727	0,132	51,00	9,27
6,5	12,30	3	0,462	0,071	36,90	5,68
7,5	12,83	2	0,267	0,036	25,66	3,42
8,5	12,28	2	0,236	0,028	24,56	2,89
9,5	12,34	2	0,211	0,022	24,68	2,60
		44	23,991	29,005	592,36	356,47

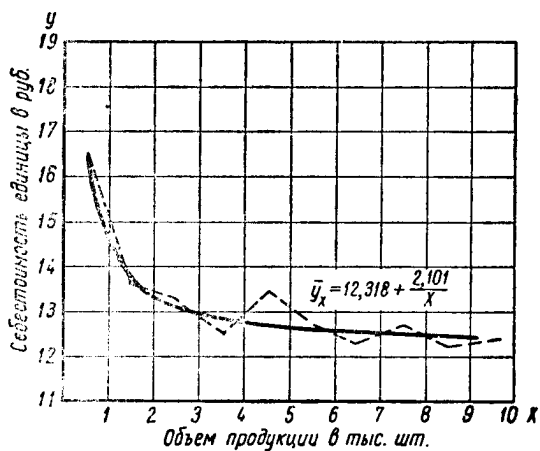


Рис. 34. Зависимость между себестоимостью единицы и объемом продукции. Теоретическая и эмпирическая линии регрессии

a, b, c, \dots , то для расчета этих параметров привлекаются специальные приемы, как-то: последовательные приближения¹, замена переменных. Последний прием рассматривается в следующей главе.

¹ Г. Л. Ритц и др., Математические методы в статистике, 1927. стр. 154.

ГЛАВА VI
ФУНКЦИОНАЛЬНАЯ КОРРЕЛЯЦИЯ

**§ 1. ФУНКЦИОНАЛЬНОЕ ПРЕОБРАЗОВАНИЕ ПЕРЕМЕННЫХ.
ЛИНИИ ФУНКЦИОНАЛЬНОЙ РЕГРЕССИИ**

В ряде случаев корреляционная зависимость оказывается более простой, если рассматривать не самые переменные, а некоторые функции этих переменных, например логарифмы, обратные величины и т. д. В таких случаях большую помощь при исследовании корреляционной зависимости оказывает предварительная замена переменных. Поясним примером.

Пусть на основе предварительного материального анализа вопроса между переменными y и x предполагается зависимость по схеме

$$y = \sqrt{u + ve^x}, \quad (6.1.1)$$

где u и v — неучитываемые аргументы, согласно предположению от x корреляционно не зависящие.

Заменой переменных

$$\left. \begin{aligned} z &= y^2 \\ t &= e^x \end{aligned} \right\} \quad (6.1.2)$$

приводим предыдущее уравнение к виду

$$z = u + vt. \quad (6.1.3)$$

Следовательно, уравнение предельной теоретической линии регрессии z по t должно иметь вид

$$\bar{z}_t = \bar{u} + \bar{v}t, \quad (6.1.4)$$

где \bar{u} , \bar{v} — неизвестные параметры, оцениваемые по опытным данным.

Пусть, например, даны результаты измерения величин x и y (столбцы 1 и 2 табл. 36).

Таблица 36

x	y	$t = e^x$	$z = y^2$	\bar{z}_t	t^2	$z \cdot t$	\bar{z}_t	\bar{y}_t	\bar{y}_x
1	2	3	4	5	6	7	8	9	10
0,7	1,71	2,01	2,92	} 3,56	4,04	5,87	} 4,27	1,89	2,07
0,7	1,90	2,01	3,60		4,04	7,24			
0,7	2,04	2,01	4,16		4,04	8,36			
1,0	2,09	2,72	4,36	} 5,29	7,40	11,86	} 5,09	2,31	2,26
1,0	2,28	2,72	5,20		7,40	14,14			
1,0	2,30	2,72	5,28		7,40	14,36			
1,0	2,51	2,72	6,32		7,40	17,19			
1,3	2,29	3,67	5,26	} 5,76	13,47	19,30	} 6,18	2,40	2,49
1,3	2,41	3,67	5,81		13,47	21,32			
1,3	2,39	3,67	5,70		13,47	20,92			
1,3	2,50	3,67	6,27		13,47	23,01			
1,6	2,70	4,95	7,30	} 7,84	24,50	36,13	} 7,65	2,80	2,77
1,6	2,81	4,95	7,88		24,50	39,01			
1,6	2,89	4,95	8,34		24,50	41,28			
1,9	3,07	6,68	9,41	} 9,84	44,62	62,86	} 9,64	3,14	3,10
1,9	3,20	6,68	10,27		44,62	68,60			
Итого		59,80	98,08		258,34	411,45			

По формулам (6.1.2) заменяем x и y через t и z (столбцы 3 и 4). По этим данным рассчитаем эмпирическую линию регрессии z по t ; в столбце 5 записываем средние арифметические значения z , соответствующие заданным значениям t . Так, при $t_1 = 2,01$ $z_{11} = 2,92$, $z_{21} = 3,60$, $z_{31} = 4,16$. Следовательно, $\bar{z}_1 = 3,56$.

При $t_2 = 2,72$, $z_{12} = 4,36$, $z_{22} = 5,20$, $z_{32} = 5,28$, $z_{42} = 6,32$, $\bar{z}_2 = 5,29$ и т. д.

Рассчитаем далее теоретическую линию регрессии z по t

$$\bar{z}_t = \bar{u} + \bar{v}t,$$

где оценки неизвестных \bar{u} , \bar{v} определяются из условия наименьших квадратов

$$\Sigma (z - \bar{z}_t)^2 = \min, \quad (6.1.5)$$

т. е. из системы нормальных уравнений

$$\left. \begin{aligned} \Sigma z &= n\bar{u} + \bar{v}\Sigma t, \\ \Sigma zt &= \bar{u}\Sigma t + \bar{v}\Sigma t^2 \end{aligned} \right\} \quad (6.1.6)$$

По данным столбцов 3, 4, 6 и 7 табл. 36 $\Sigma t = 59,80$, $\Sigma t^2 = 258,34$, $\Sigma z = 98,08$, $\Sigma zt = 411,45$. Подставляя эти суммы в уравнения (6. 1. 6), решаем их относительно \bar{u} , \bar{v} , получаем

$$\bar{u} = 1,31, \quad \bar{v} = 1,29 \quad \text{и} \quad \bar{z}_t = 1,31 + 1,29t.$$

Наносим эмпирическую и теоретическую линии регрессии z по t на поле корреляции (рис. 35).

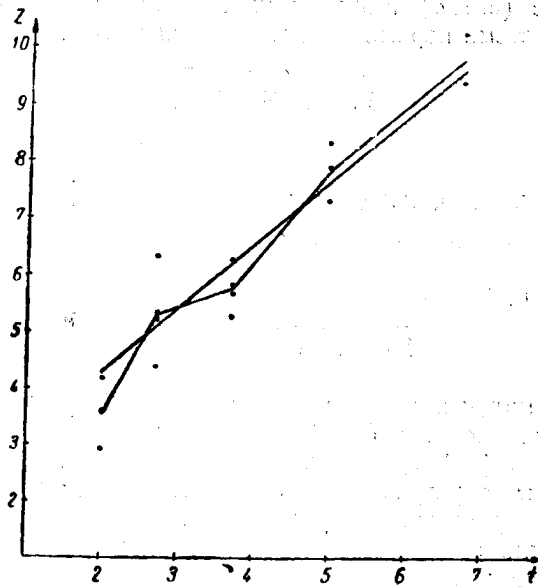


Рис. 35. Поле корреляции и эмпирическая и теоретическая линии регрессии z по t

Близость обеих линий служит подтверждением правильности положенной в основу расчетов схемы действия факторов (6. 1. 3), а следовательно, и схемы (6. 1. 1).

Восстановим теперь в полученных соотношениях прежнюю систему переменных. Формула эмпирической линии регрессии z по t

$$\bar{z}_i = \frac{\Sigma_i z}{h_i}$$

(где i — порядковый номер значения t или соответствующего значения x ; h_i — частота)

в прежней системе переменных запишется в виде

$$\tilde{y}_i = \sqrt{\frac{\Sigma_i y^2}{h_i}}, \quad (6. 1. 7)$$

где суммирование под знаком радикала производится по всем y^2 , соответствующим определенному значению x_i ; \tilde{y}_i — значение y , квадрат которого равен \tilde{z}_i .

В столбце 9 таблицы записываем значения \tilde{y}_i , рассчитанные для каждого x_i по формуле (6.1.7) или, что то же, по формуле $\tilde{y}_i = \sqrt{\tilde{z}_i}$.

Уравнение (6.1.4) теоретической линии регрессии z по t в прежней системе переменных запишется в виде

$$\tilde{y}_x = \sqrt{\bar{u} + \bar{v}e^x}, \quad (6.1.8)$$

где $x = \ln t$;

\tilde{y}_x — значение y , квадрат которого равен \tilde{z}_i ;
 \bar{u}, \bar{v} найдены из условия (6.1.5).

В нашем примере

$$\tilde{y}_x = \sqrt{1,31 + 1,29e^x}.$$

Обсудим полученные результаты и способ их получения.

Замена переменных (6.1.2) позволила нам привести более сложную схему (6.1.1) к простейшей схеме (6.1.3), для которой расчет уравнения регрессии по опытным данным и проверка его правильности представляют наименьшие трудности. В самом деле, преобразование зависимой переменной $z = y^2$ привело к схеме А—М—Н, для которой расчет параметров уравнения регрессии сводится, как известно, к решению системы нормальных уравнений *линейного* вида. В то же время преобразование аргумента $t = e^x$ привело к линейному виду уравнение регрессии, что особенно удобно для расчетов и графиков. В этом техническая сторона дела. Существует, однако, и принципиальная сторона, которая заключается в следующем.

В уравнении теоретической линии регрессии (6.1.4) мы закрепляем на среднем уровне факторы u и v , корреляционно не зависящие от x . То же, очевидно, и в уравнении (6.1.8). Следовательно, уравнение (6.1.8) отвечает на основной вопрос теории корреляции: «как изменялась бы функция y с изменением аргумента x , если бы прочие ее аргументы, корреляционно не зависящие от x , были закреплены на среднем уровне?».

Уравнение регрессии y по x : $\tilde{y}_x = \sqrt{a + be^x}$, в котором параметры a и b были бы рассчитаны из условия

$$\Sigma (\bar{y} - \tilde{y}_x)^2 = \min, \quad (6.1.9)$$

не дает правильного ответа на поставленный выше вопрос в условиях схемы (6.1.1).

Заметим, что \bar{y}_i не есть *средняя арифметическая* из значений y , соответствующая заданному x_i . Это средняя, определяемая формулой (6.1.7); она носит название *средней квадратической* и является разновидностью *функциональных средних*,

о которых речь будет идти ниже. Аналогично \bar{y}_x является выравненной средней квадратической. Поэтому линии, передающие изменение \bar{y}_i и \bar{y}_x с изменением x , не являются обычными линиями регрессии y по x ; мы назовем их *эмпирической и теоретической линиями функциональной регрессии y по x* .

Заметим также, что подобное принципиальное значение имеет лишь преобразование зависимой переменной $z=y^2$, преобразование аргумента $t=e^x$ имеет чисто техническое значение для упрощения графиков и расчетов.

Обобщим выводы из нашего примера.

Мы видим, что влияние изменчивости побочных факторов на зависимую переменную y в ряде случаев может быть исключено при помощи функционального преобразования зависимой переменной

$$z = \Phi(y) \quad (6.1.10)$$

и последующего расчета линии регрессии z по x .

Общая схема действия факторов в тех случаях, где подобный метод ведет к цели, выражается формулой

$$y = \Phi^{-1}[u + v\varphi(x) + w\psi(x) + \dots], \quad (6.1.11)$$

где u, v, w, \dots — аргументы, корреляционно не зависящие от x ;

Φ^{-1} — обозначение функциональной зависимости, обратной Φ .

Схему (6.1.11) будем называть *схемой Φ —А—М—Н*, Φ — ее *определяющей функцией*.

Преобразование (6.1.10) приводит уравнение (6.1.11) к виду

$$z = u + v\varphi(x) + w\psi(x) + \dots, \quad (6.1.12)$$

что соответствует схеме А—М—Н. Следовательно, уравнение предельной теоретической линии регрессии z по x должно иметь вид

$$\bar{z}_x = \bar{u} + \bar{v}\varphi(x) + \bar{w}\psi(x) + \dots \quad (6.1.13)$$

В отдельных случаях возможно дальнейшее упрощение уравнения (6.1.11) путем функционального преобразования аргумента x . Так, уравнение

$$z = u + v\varphi(x) \quad (6.1.14)$$

преобразованием $t = \varphi(x)$ приводится к линейному

$$z = u + vt.$$

Однако преобразование аргумента, как было показано на примере, принципиального значения не имеет, так как оно не изменяет окончательных результатов.

Для проверки правильности уравнения (6.1.13) и оценки параметров u, v, w, \dots по опытным данным рассчитываются эмпирическая и теоретическая линии регрессии z по x (или z по t).

Ординаты эмпирической линии регрессии рассчитываются по формуле

$$\bar{z}_i = \frac{\Sigma_i z}{h_i}, \quad (6.1.15)$$

где суммирование Σ_i производится для значений z , соответствующих данному значению x ; h_i — частота последнего. Оценка параметров $\bar{u}, \bar{v}, \bar{w}, \dots$ дается из условия наименьших квадратов

$$\Sigma (z - \bar{z}_x)^2 = \min, \quad (6.1.16)$$

т. е. из системы нормальных уравнений:

$$\left. \begin{aligned} \Sigma z &= n\bar{u} + \bar{v}\Sigma\varphi(x) + \bar{w}\Sigma\psi(x) + \dots \\ \Sigma z\varphi(x) &= \bar{u}\Sigma\varphi(x) + \bar{v}\Sigma[\varphi(x)]^2 + \bar{w}\Sigma\varphi(x)\psi(x) + \dots \\ \Sigma z\psi(x) &= \bar{u}\Sigma\psi(x) + \bar{v}\Sigma\varphi(x)\psi(x) + \bar{w}\Sigma[\psi(x)]^2 + \dots \end{aligned} \right\} \quad (6.1.17)$$

При удовлетворительном соответствии эмпирической и теоретической линий регрессии считаем схему (6.1.12), а следовательно, и схему (6.1.11) оправданными.

Совершая преобразования

$$\tilde{y}_i = \Phi^{-1}(\bar{z}_i)$$

и

$$\tilde{y}_x = \Phi^{-1}(\bar{z}_x),$$

обратные преобразованию (6.1.10), мы приходим к формуле эмпирической линии функциональной регрессии y по x

$$\tilde{y}_i = \Phi^{-1} \left[\frac{\sum_i \Phi(y)}{h_i} \right] \quad (6.1.18)$$

и уравнению теоретической линии функциональной регрессии y по x

$$\tilde{y}_x = \Phi^{-1} [\bar{u} + \bar{v}\varphi(x) + \bar{w}\psi(x) + \dots]. \quad (6.1.19)$$

Уравнение (6.1.19) отвечает на вопрос: как изменялась бы функция y с изменением аргумента x , если бы прочие ее аргументы, действующие по схеме Ф—А—М—Н, были закреплены на среднем уровне.

§ 2. ФУНКЦИОНАЛЬНЫЕ СРЕДНИЕ

Формула (6.1.18) предшествующего параграфа выражает при переменном x эмпирическую линию функциональной регрессии y по x , при постоянном x — функциональную среднюю значений y , соответствующих заданному значению x . В зависимости от вида определяющей функции Φ получаются те или иные разновидности функциональных средних. Рассмотрим некоторые наиболее интересные случаи.

В этом параграфе формулу функциональной средней будем писать, опуская индекс, в виде

$$\tilde{y} = \Phi^{-1} \left[\frac{\sum \Phi(y)}{h} \right]. \quad (6.2.1)$$

Расчеты различных средних будем производить на одном и том же числовом примере.

Таблица 37

Таблица значений y , из которых вычисляется функциональная средняя

y	5	2	1	4	4
-----	---	---	---	---	---

1. Пусть схема действия факторов имеет вид

$$y = u + v\varphi(x) + w\psi(x), \quad (6.2.2)$$

где u , v , w — корреляционно не зависят от x .

Это — схема А—М—Н, для которой при вычислении линии регрессии функционального преобразования зависимой переменной не требуется

$$\Phi(y) = y.$$

Следовательно, в формуле (6. 2. 1)

$$\tilde{y} = \frac{y_1 + y_2 + \dots + y_n}{n}.$$

Мы пришли к формуле средней арифметической: $\tilde{y} = \bar{y}$, которая также является разновидностью функциональной средней. По данным табл. 37

$$\bar{y} = \frac{5 + 2 + 1 + 4 + 4}{5} = 3,2.$$

2. Схема действия факторов имеет вид

$$y = \frac{1}{u + v\varphi(x) + w\psi(x)}, \quad (6. 2. 3)$$

u , v , w корреляционно не зависят от x .

Определяющая функция схемы $\Phi(y) = \frac{1}{y}$, следовательно, по формуле (6. 2. 1)

$$\tilde{y} = \frac{n}{\frac{1}{y_1} + \frac{1}{y_2} + \dots + \frac{1}{y_n}} \quad (6. 2. 4)$$

или по числовым данным таблицы

$$\tilde{y} = \frac{5}{\frac{1}{5} + \frac{1}{2} + \frac{1}{1} + \frac{1}{4} + \frac{1}{4}} = 2,27.$$

Формула (6. 2. 4) носит название формулы *средней гармонической*.

3. Пусть $y = a^{u+v\varphi(x)+w\psi(x)}$, u , v , w корреляционно не зависят от x .

Здесь $\Phi(y) = \lg y$, следовательно,

$$\tilde{y} = a^{\frac{\lg y_1 + \lg y_2 + \dots + \lg y_n}{n}}$$

или

$$\tilde{y} = \sqrt[n]{y_1 \cdot y_2 \cdot \dots \cdot y_n}. \quad (6. 2. 5)$$

По таблице

$$\tilde{y} = \sqrt[5]{5 \cdot 2 \cdot 1 \cdot 4 \cdot 4} = 2,76.$$

Формула (6.2.5) называется формулой *средней геометрической*.

4. Схема

$$y = u [\varphi(x)]^v [w]^\psi(x) \quad (6.2.6)$$

приводится к предыдущей путем преобразования

$$y = a^{\lg u + v \lg \varphi(x) + \lg w^\psi(x)},$$

если $\lg u$, v , $\lg w$ корреляционно не зависят от x .

Поэтому здесь также $\Phi(y) = \lg y$,

$$\tilde{y} = \sqrt[h]{y_1 \cdot y_2 \cdot \dots \cdot y_h} = 2,76.$$

5. Пусть

$$y = \sqrt[k]{u + v\varphi(x) + w\psi(x)}, \quad (6.2.7)$$

u , v , w корреляционно не зависят от x ;
 k — постоянное.

Определяющая функция схемы $\Phi(y) = y^k$. Следовательно,

$$\tilde{y} = \sqrt[k]{\frac{y_1^k + y_2^k + \dots + y_h^k}{h}}. \quad (6.2.8)$$

Это — формула *степенной средней* порядка k . В частности, при $k=1$ получаем *среднюю арифметическую*, при $k=2$ — *среднюю квадратическую*, при $k=3$ — *среднюю кубическую* и т. д. Для нашего примера кубическая средняя

$$\tilde{y} = \sqrt[3]{\frac{5^3 + 2^3 + 1^3 + 4^3 + 4^3}{5}} = 3,74.$$

Функциональные средние различных видов, вычисленные по одним и тем же числовым данным, различаются между собой по величине. Величина расхождения между ними зависит от характера числовых данных, однако *порядок возрастания*, в котором располагаются различные функциональные средние, рассчитанные по одним и тем же данным, во многих случаях *не зависит от этих числовых данных, а определяется видом функции Φ* . Это свойство носит название *сравнимости*, или *мажорантности* средних.

Следующий признак, установленный А. Я. Боярским в работе, посвященной обоснованию теории средних величин¹, дает достаточные условия сравнимости функциональных средних.

Признак мажорантности средних

Пусть $\Phi(y)$ и $\Phi_1(y)$ — две функции положительной переменной y , монотонно возрастающие и дважды дифференцируемые в некотором интервале значений y , и пусть в каждой точке этого интервала

$$\frac{\Phi''(y)}{\Phi'(y)} > \frac{\Phi_1''(y)}{\Phi_1'(y)}. \quad (6.2.9)$$

Тогда в этом интервале значений y

$$\tilde{y}_{(\Phi)} \geq \tilde{y}_{(\Phi_1)},$$

где индексы (Φ) и (Φ_1) указывают на вид функциональной средней.

Рассчитаем отношение $\frac{\Phi''(y)}{\Phi'(y)}$ в примерах, приведенных выше.

1. Для средней арифметической $\Phi(y) = y$ $\Phi'(y) = 1$, $\Phi''(y) = 0$,

$$\frac{\Phi''(y)}{\Phi'(y)} = 0. \quad (6.2.10)$$

2. Для средней гармонической $\Phi(y) = \frac{1}{y}$
 $\Phi'(y) = -\frac{1}{y^2}$, $\Phi''(y) = \frac{2}{y^3}$,

$$\frac{\Phi''(y)}{\Phi'(y)} = -\frac{2}{y}. \quad (6.2.11)$$

3. Для средней геометрической $\Phi(y) = \lg y$,

$$\frac{\Phi''(y)}{\Phi'(y)} = -\frac{1}{y}. \quad (6.2.12)$$

4. Для степенной средней $\Phi(y) = y^k$

$$\frac{\Phi''(y)}{\Phi'(y)} = \frac{k-1}{y}. \quad (6.2.13)$$

В частности, для средней квадратической

$$\frac{\Phi''(y)}{\Phi'(y)} = \frac{1}{y}, \quad (6.2.14)$$

¹ А. Я. Боярский, Средняя, «Вестник статистики» № 2, 1929.

для средней кубической

$$\frac{\Phi''(y)}{\Phi'(y)} = \frac{2}{y}. \quad (6.2.15)$$

Следовательно, при положительных значениях y мы должны иметь следующий порядок мажорантности средних: средняя гармоническая \leq средней геометрической \leq средней арифметической \leq средней квадратической \leq средней кубической, что подтверждается расчетом на нашем примере.

Из формул (6.2.12) и (6.2.13) видно, что величина критерия мажорантности $\frac{\Phi''(y)}{\Phi'(y)}$ для средней геометрической совпадает с величиной этого критерия для степенной средней, если показатель степени k положить равным нулю. Следовательно, с точки зрения мажорантности, средняя геометрическая играет роль степенной средней с показателем $k=0$. Это объясняется тем, что средняя геометрическая служит пределом выражения степенной средней при $k \rightarrow 0$.

Приводим доказательство этого предложения:

$$\ln \sqrt[k]{\frac{x_1^k + x_2^k + \dots + x_h^k}{h}} = \frac{\ln(x_1^k + x_2^k + \dots + x_h^k) - \ln h}{k}.$$

При $k=0$ последнее выражение представляет неопределенность вида $\frac{0}{0}$. Раскрывая эту неопределенность по правилу Лопиталю, получим

$$\begin{aligned} \lim_{k \rightarrow 0} \ln \sqrt[k]{\frac{x_1^k + x_2^k + \dots + x_h^k}{h}} &= \lim_{k \rightarrow 0} \frac{x_1^k \ln x_1 + x_2^k \ln x_2 + \dots + x_h^k \ln x_h}{x_1^k + x_2^k + \dots + x_h^k} = \\ &= \frac{\ln x_1 + \ln x_2 + \dots + \ln x_h}{h} = \ln \sqrt[h]{x_1 x_2 \dots x_h}. \end{aligned}$$

Следовательно,

$$\lim_{k \rightarrow 0} \sqrt[k]{\frac{x_1^k + x_2^k + \dots + x_h^k}{h}} = \sqrt[h]{x_1 x_2 \dots x_h},$$

что и требовалось доказать.

Функциональные средние имеют простую и взвешенную форму. Простая функциональная средняя определяется формулой (6.2.1). Если же каждое значение y встречается с весом m , то расчеты производятся по формуле взвешенной средней

$$\tilde{y} = \Phi^{-1} \left[\frac{\sum m \Phi(y)}{\sum m} \right]. \quad (6.2.16)$$

В частности, формула взвешенной средней гармонической

$$\tilde{y} = \frac{\Sigma m}{\Sigma \frac{m}{y}}, \quad (6.2.17)$$

взвешенной средней геометрической

$$\tilde{y} = \sqrt[\Pi]{\frac{\Sigma m}{\Pi y^m}}, \quad (6.2.18)$$

Π — знак произведения;

взвешенной степенной средней

$$\tilde{y} = \sqrt[\kappa]{\frac{\Sigma m y^{\kappa}}{\Sigma m}}. \quad (6.2.19)$$

Вычислим взвешенные средние по данным следующего примера.

Таблица 38

Варианты y	Веса m
1	5
2	10
3	25
4	15
5	5
Итого	60

Средняя гармоническая:

$$\tilde{y} = \frac{60}{\frac{5}{1} + \frac{10}{2} + \frac{25}{3} + \frac{15}{4} + \frac{5}{5}} = 2,60.$$

Средняя геометрическая:

$$\tilde{y} = \sqrt[60]{1^5 \cdot 2^{10} \cdot 3^{25} \cdot 4^{15} \cdot 5^5} = 2,87.$$

Средняя квадратическая:

$$\tilde{y} = \sqrt{\frac{1^5 \cdot 5 + 2^2 \cdot 10 + 3^2 \cdot 25 + 4^2 \cdot 15 + 5^2 \cdot 5}{60}} = 3,25.$$

Средняя кубическая:

$$\tilde{y} = \sqrt[3]{\frac{1^3 \cdot 5 + 2^3 \cdot 10 + 3^3 \cdot 25 + 4^3 \cdot 15 + 5^3 \cdot 5}{60}} = 3,39.$$

Для взвешенных функциональных средних сохраняется мажорантность простых функциональных средних.

Проф. Б. С. Ястремский показал¹, что между степенной средней \tilde{y}_k и геометрической средней \tilde{y}_0 существует соотношение

$$\tilde{y}_k = \tilde{y}_0 e^{\frac{k\bar{L}^2}{2}}, \quad (6.2.20)$$

где

$$\bar{L}^2 = \sigma_{\ln y}^2.$$

Соотношение (6.2.20) справедливо при наличии нормального распределения² для $\ln y$ и приближенно справедливо, если распределение $\ln y$ близко к нормальному.

В случае, когда коэффициент вариации

$$V_y = \frac{\sigma_y}{y}$$

невелик (например, меньше 10%), то вместо равенства (6.2.20) Б. С. Ястремский предлагает пользоваться более простым приближенным равенством

$$\tilde{y}_k \approx \tilde{y}_0 \left(1 + \frac{kV_y^2}{2}\right). \quad (6.2.21)$$

Формулы (6.2.20) и (6.2.21) можно использовать для оценки степенных средних с любым показателем k , если известно значение степенной средней при каком-либо одном значении k .

3. ЛИНИИ ФУНКЦИОНАЛЬНОЙ РЕГРЕССИИ. ФУНКЦИОНАЛЬНАЯ КОРРЕЛЯЦИЯ

Рассмотрим способ расчета линий функциональной регрессии в тех случаях, когда общая схема действия факторов нам заранее неизвестна и вопрос о целесообразности преобразования переменных решается лишь в процессе статистической обработки опытных данных.

¹ Б. С. Ястремский, Математическая статистика, Госстатиздат, 1956, стр. 114—117.

² О нормальном законе распределения см. главу XII, § 2.

Своеобразная трудность, которая здесь возникает, заключается в том, что при нахождении теоретической линии функциональной регрессии мы не можем опираться на график эмпирической линии функциональной регрессии, ибо для расчета последнего необходимо знать функциональное преобразование зависимой переменной. Выход из затруднения состоит в применении *способа ориентировочных линий регрессии*. Покажем его на конкретном примере.

Трудоемкость изготовления горячих штампов зависит от размеров штампа и его геометрической формы. Степень сложности геометрической формы штампа можно определить количеством поверхностей. Наблюдается, однако, некоторая зависимость между размерами штампа и числом его поверхностей: большие штампы, как правило, имеют большее число поверхностей; в силу этого показатель абсолютного числа поверхностей отражает не только сложность геометрической формы, но отчасти также и размеры штампа.

Для получения показателя сложности формы, не зависящего от размеров штампа, кажется естественным брать отношение числа поверхностей к объему (весу) штампа, т. е. относительное число поверхностей, приходящееся на единицу объема (веса). Однако последний показатель неудовлетворителен в силу того, что число поверхностей увеличивается не пропорционально объему (весу) штампа, а в значительно меньшей степени — более замедленно.

Чтобы правильно построить показатель сложности геометрической формы, не зависящий от размеров штампа, необходимо установить характер зависимости между числом поверхностей и объемом (весом) штампа. Рассчитаем эту зависимость на материале 100 горячих штампов, обработанных на копировальных станках с ручной подачей, по данным, предоставленным автору проф. Г. Б. Лурье (см. приложение I, табл. 2).

При нахождении зависимости ни один из элементов схемы Ф—А—М—Н (определяющая функции Φ , функции $\varphi(x)$, $\psi(x)$...) не предполагается известным.

Порядок определения зависимости.

1. По данным столбцов 2 и 4 табл. 2 приложения I составим поле корреляции, корреляционную таблицу и эмпирическую линию регрессии числа поверхностей по весу штампа (см. табл. 39 и рис. 36).

Линия регрессии рассчитана при неравных интервалах аргумента ввиду резкой асимметрии графика распределения по этому аргументу. Эмпирическая линия регрессии по внешнему виду напоминает график простой степенной зависимости с дробным показателем степени (рис. 37)

$$\bar{y}_x = ax^b. \quad (6.3.1)$$

Таблица 39

y'	$y \backslash x$	0,1252	0,6255	1,5005	2,5005	3,5005	5,0005	7,5005	10,5005	13,0005	Итого l
7	215,5		1 ⁷				1 ⁷				2
6	195,5										0
5	175,5		1 ⁵								1
4	155,5		1 ⁴								1
3	135,5		1 ³	1 ³	1 ³	3 ⁹	1 ³		1 ³		8
2	115,5		3 ⁶	4 ⁸	2 ⁴	1 ²	1 ²	1 ²			12
1	95,5		1 ¹	2 ²	7 ⁷	4 ⁴	1 ¹	1 ¹	2 ²	2 ²	20
0	75,5	2	7	4	3		1	1			18
-1	55,5	5 ⁻⁵	8 ⁻⁸	8 ⁻³	3 ⁻³	2 ⁻²	2 ⁻²		1 ⁻¹		24
-2	35,5	1 ⁻²	6 ⁻¹²	2 ⁻⁴	2 ⁻⁴	1 ⁻²	1 ⁻²				13
-3	15,5	1 ⁻³									1
Итого	h	9	29	16	18	11	8	3	4	2	100
	$\Sigma_i m y'$	-10	6	6	7	11	9	3	4	2	
	\bar{y}'_i	-1,11	0,21	0,38	0,39	1,00	1,12	1,00	1,00	1,00	
	\bar{y}_i	53,3	79,1	83,1	83,3	95,5	97,9	95,5	95,5	95,5	

Ориентируясь на эту формулу, выдвигаем предположение о том, что в основе зависимости лежит следующая схема действия факторов

$$y = ux^v, \quad (6.3.2)$$

которая является схемой $\Phi-A-M-N$

$$y = 10^{\lg u + v \lg x} \quad (6.3.3)$$

с определяющей функцией

$$\Phi(y) = \lg y,$$

если предположить, что $\lg u$ и v корреляционно не зависят от x .

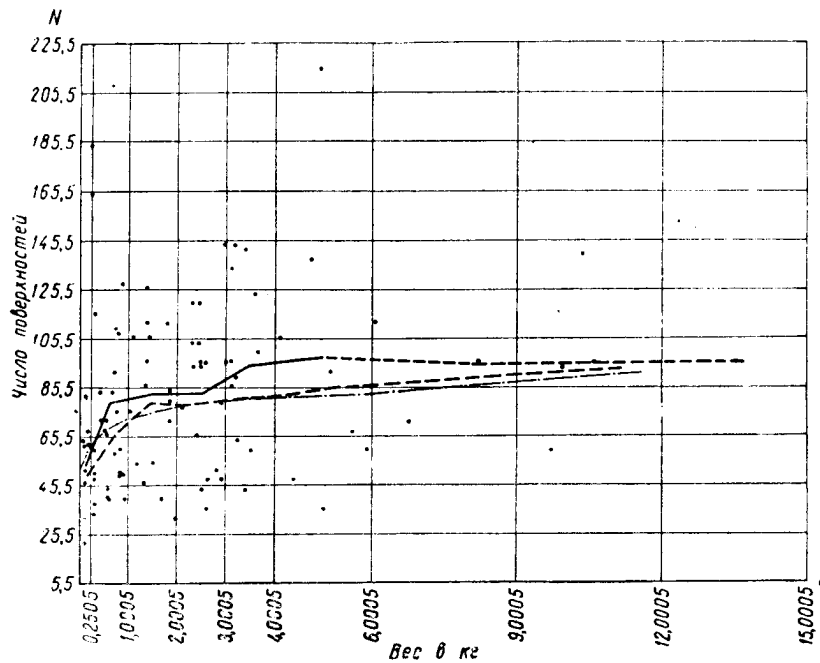


Рис. 36. Зависимость между числом поверхностей и весом штампа. Эмпирическая и теоретическая линия регрессии

2. Заменяем переменные

$$z = \lg y,$$

$$t = \lg x.$$

и представляем схему (6.3.3) в виде

$$z = \lg u + vt. \quad (6.3.4)$$

Следовательно, предельная теоретическая линия регрессии z по t должна иметь вид

$$\bar{z}_t = a + bt, \quad (6.3.5)$$

где $a = \overline{\lg u}$; $b = \bar{v}$.

Рассчитаем теоретическую линию регрессии z по t , для чего определяем параметры a и b в уравнении (6.3.5) из условия

$$\Sigma(z - \bar{z}_t)^2 = \min$$

и производим расчет эмпирической линии регрессии z по t .
 Для вычислений используем данные корреляционной табл. 40, которая составлена для логарифмов чисел x и y , записанных в столбцах 3, 5 таблицы 2 приложения I.

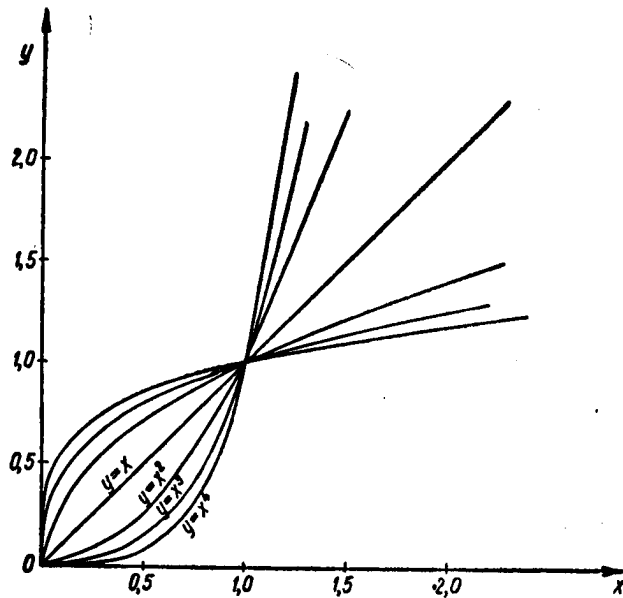


Рис. 37. Графики степенной функции

Уравнение регрессии в упрощенном масштабе

$$\bar{z}' = a' + b't'$$

Расчет параметров

$$\left. \begin{aligned} \dot{\Sigma} z' &= na' + b'\dot{\Sigma} t' \\ \dot{\Sigma} z't' &= a'\dot{\Sigma} t' + b'\dot{\Sigma} t'^2 \end{aligned} \right\}$$

или

$$\left. \begin{aligned} (1) \quad 17 &= 100a' + (-7)b' && : 100 \\ (2) \quad 93 &= -7a' + 299b' && : 7 \\ \hline (1)' \quad 0,17 &= a' - 0,07b' \\ (2)' \quad 13,29 &= -a' + 42,71b' \\ \hline (1)' + (2)' \quad 13,46 &= 42,64b' \end{aligned} \right\}$$

$$b' = \frac{13,46}{42,64} = 0,316.$$

Таблица 40

z'	t'	-5	-4	-3	-2	-1	0	1	2	3	Номера столбцов	
	t	(-1,505) - (-1,205)	(-1,205) - (-0,905)	(-0,905) - (-0,605)	(-0,605) - (-0,305)	(-0,305) - (-0,005)	(-0,005) - 0,295	0,295 - 0,595	0,595 - 0,895	0,895 - 1,195	1	2
											Итого t	
5	2,305-2,405					1 ⁵			1 ⁵		2	10
4	2,205 - 2,305				2 ⁸						2	8
3	2,105-2,205					1 ³		12 ⁴	3 ¹	3 ¹	7	21
2	2,005-2,105				1 ²	2 ⁴	10 ⁵	10 ⁵	2 ⁴		15	30
1	1,905-2,005			1 ¹	1 ¹	2 ²	5 ⁵	9 ⁹	1 ¹	3 ³	22	22
0	1,805-1,905	1		1	1	3	3	4	2		15	0
-1	1,705-1,805		-2 ²	-1 ¹	-1 ¹	-3 ³	-2 ²	-2 ²	-1 ¹	-1 ¹	13	-13
-2	1,605-1,7 5		-2 ²		-4 ⁴	-8 ⁸	-2 ²	-8 ⁸	-2 ²		13	-26
-3	1,505-1,605		-3 ³		-6 ⁶	-9 ⁹	-3 ³	-6 ⁶	-3 ³		10	-30
-4	1,405-1,505										0	0
-5	1,305-1 405			-5 ⁵							1	-5
1	Итого h	1	4	4	10	19	17	31	10	5	100	17
2	ht'	-5	-16	-12	-20	-19	0	30	20	15	-7	$= \Sigma t'$
3	ht'^2	25	64	36	40	19	0	30	40	45	299	$= \Sigma t'^2$
4	$\Sigma_i mz'$	0	-7	-5	0	-6	8	15	7	5	17	$= \Sigma z'$
5	$t' \Sigma_i mz'$	0	28	15	0	6	0	15	14	15	93	$= \Sigma z' t'$
6	\bar{z}'_i	0	-1,75	-1,25	0	-0,32	0,47	0,50	0,70	1	-	-
7	\bar{z}_i	-	1,68	1,70	1,86	1,82	1,90	1,90	1,92	1,96	-	-
8	\bar{z}_t	-	1,75	1,78	1,81	1,84	1,87	1,90	1,94	1,97	-	-
9	\bar{y}_i	-	47,9	50,1	72,4	66,1	79,4	80,3	83,2	91,2	-	-
10	\bar{y}_x	-	56,2	60,3	64,6	69,2	74,1	79,4	87,1	93,3	-	-

откуда

$$a' = 0,192.$$

По формулам (5. 2. 4)

$$a = 1,855 + 0,1 \cdot 0,192 - 0,316 \cdot \frac{0,1}{0,3} \cdot 0,145 = 1,859,$$

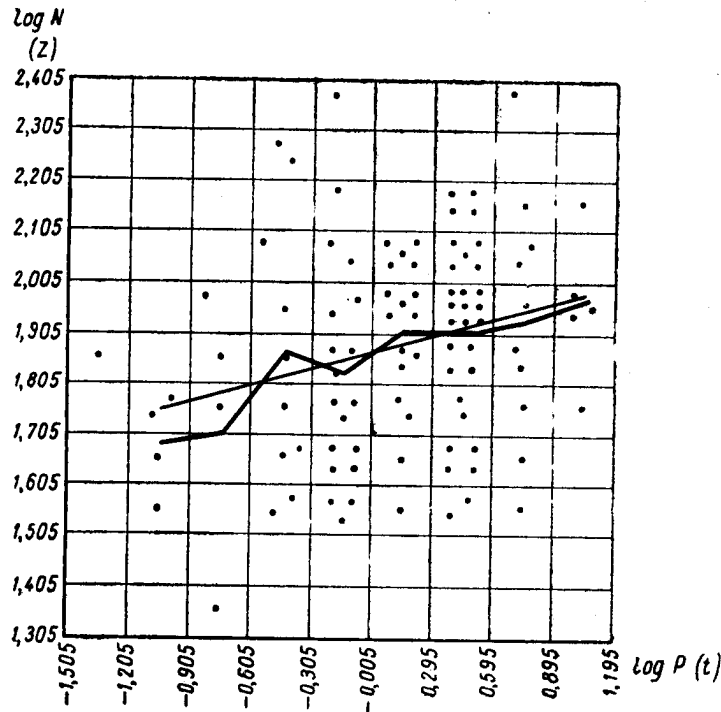


Рис. 38. Зависимость между числом поверхностей и весом штампа. Теоретическая и эмпирическая линии обобщенной регрессии

$$b = 0,316 \cdot \frac{0,1}{0,3} = 0,105.$$

Следовательно,

$$\bar{z}_t = 1,859 + 0,105t. \quad (6. 3. 5')$$

Теоретическая и эмпирическая линии регрессии z по t оказались близкими друг к другу, что подтверждает правильность сделанного нами предположения о схеме действия факторов (6. 3. 2).

3. Восстанавливая первоначальные переменные в формулах, выражающих теоретическую и эмпирическую линии регрессии z

9*

по t , получаем теоретическую и эмпирическую линии функциональной регрессии y по x . Уравнение теоретической линии функциональной регрессии получается потенцированием уравнения (6.3.5')

$$\tilde{y}_x = 72,28 \cdot x^{0,105}. \quad (6.3.6)$$

Значения \tilde{y}_i и \tilde{y}_x , полученные потенцированием значений \bar{z}_i и \bar{z}_x , записаны в строках 9 и 10 табл. 40.

Теоретическая и эмпирическая линии функциональной регрессии даны на рис. 36, где для сравнения нанесена и обычная эмпирическая линия регрессии.

Мы видим, что характер последней линии аналогичен характеру линии функциональной регрессии, но она имеет *систематическое смещение* в положительном направлении оси y . По этой причине график обычной регрессии дает нам лишь полезную ориентировку при определении типа уравнения функциональной регрессии [формула (6.3.1)], но не пригоден, однако, в силу систематического смещения для оценки параметров уравнения. Систематическое смещение обычной линии регрессии по сравнению с линиями функциональной регрессии объясняется тем, что первая осередняет индивидуальные значения y по формуле средней арифметической, в то время как линия функциональной регрессии осередняет эти значения по формуле функциональной средней, в данном примере по формуле средней геометрической. В силу же свойства мажорантности средняя геометрическая должна быть меньше средней арифметической.

Возвращаясь к формуле (6.3.6), мы видим, что число поверхностей растет в среднем пропорционально объему штампа, возведенному в степень 0,105. Отсюда для показателя сложности геометрической формы, не зависящего от размеров (объема) штампа, может быть предложена следующая формула

$$K = \frac{y}{x^{0,105}}, \quad (6.3.7)$$

где y — число поверхностей;
 x — вес штампа в кг.

Задача решена.

Функциональная корреляция

В связи с рассмотрением линий функциональной регрессии необходимо обобщить и понятие корреляционной зависимости.

О п р е д е л е н и е. Переменная y находится в функциональной корреляции вида Φ от x , если:

1) каждому значению аргумента x соответствует ряд распределения функции y ,

2) предельная теоретическая линия функциональной регрессии y по x с определяющей функцией Φ не параллельна оси x .

Если же предельная теоретическая линия функциональной регрессии y по x (с определяющей функцией Φ) параллельна оси абсцисс, то y не зависит от x в смысле корреляции (вида Φ). Так, например, y находится в логарифмической, квадратической, гармонической и т. д. корреляции от x , если предельные теоретические линии функциональной регрессии, рассчитанные по средним геометрическим, средним квадратическим, средним гармоническим и т. д. значениям \tilde{y}_x , не параллельны оси x .

Для переменной y может отсутствовать корреляционная зависимость от x вида Φ и в то же время существовать корреляционная зависимость другого вида — Φ_1 . Корреляционная зависимость любого вида является частным случаем зависимости статистической.

Рассмотрим, в виде примера, показатель сложности геометрической формы штампа K , определяемый формулой (6.3.7), и проверим, в каком смысле этот показатель оказался независимым от x .

Логарифмируем уравнение (6.3.7) и обозначим $\lg K$ через L $\lg y$ через z и $\lg x$ через t ; получим

$$L = z - 0,105t. \quad (6.3.8)$$

Уравнение регрессии $L - z$ по t

$$(L - z)_t = -0,105t \text{ или } \bar{L}_t = \bar{z}_t - 0,105t.$$

Но из (6.3.5')

$$\bar{L}_t = 1,859 + 0,105t - 0,105t = 1,859.$$

Итак, $\bar{L}_t = \text{const}$;

потенцируем,

$$\tilde{K}_x = \text{const}. \quad (6.3.9)$$

Следовательно, теоретическая линия регрессии K по x , образованная из средних геометрических значений K , параллельна оси x .

Логарифмическая корреляция y от x отсутствует.

ГЛАВА VII

ИЗМЕРЕНИЕ ТЕСНОТЫ СВЯЗИ

§ 1. ПРОБЛЕМА ТЕСНОТЫ СВЯЗИ. ЭМПИРИЧЕСКОЕ КОРРЕЛЯЦИОННОЕ ОТНОШЕНИЕ

В главе V излагались методы решения первой основной проблемы теории корреляции — проблемы установления формы связи. В математическом анализе при исследовании функциональных зависимостей мы встречаемся лишь с этой проблемой. Однако в корреляционной зависимости обнаруживается и другая сторона дела, которая ставит нас перед второй основной проблемой — проблемой измерения *тесноты* связи.

Поясним это на схематическом примере.

Рассмотрим два корреляционных поля a и b (рис. 39). Пусть линии регрессии y по x в обоих полях расположены одинаково. Однако точки поля b значительно ближе примыкают к линии регрессии, чем точки поля a . Отклонения точек поля от линии регрессии объясняются, очевидно, влиянием прочих факторов. Если бы y полностью определялся аргументом x , то все точки поля находились бы на линии регрессии; чем сильнее влияние прочих факторов, тем больше отклоняются от линии регрессии точки корреляционного поля.

Очевидно, что в случае b влияние аргумента x в меньшей степени осложняется действием прочих факторов, чем в случае a , и зависимость между y и x является более *тесной*. Степень тесноты связи — весьма важное свойство, определяющее научное и практическое значение данной корреляционной зависимости.

Если теснота связи велика, то, зная аргумент, мы по уравнению регрессии точно предсказываем значение функции, воздействуя на аргумент, если он является причиной изменения функции, можем управлять функцией. Если теснота связи мала, то влияние аргумента x на функцию обнаруживается лишь в *среднем* по уравнению регрессии. В каждом же отдельном наблюдении оно перекрывается действием побочных факторов.

Измерение степени тесноты корреляционной зависимости — вторая основная проблема теории корреляции. Рассмотрим пути ее решения.

Действие какого-либо фактора на функцию проявляется в изменчивости функции под влиянием изменчивости данного фак-

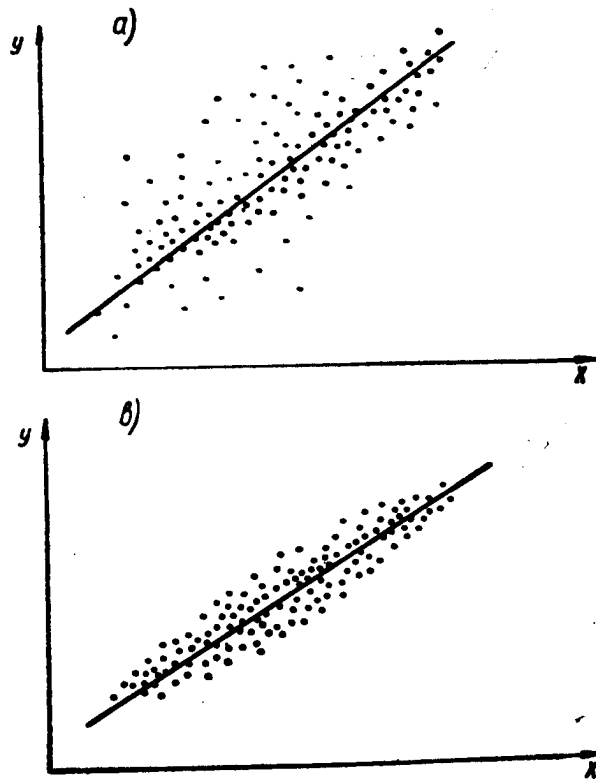


Рис. 39

тора. Поэтому показатель изменчивости функции — дисперсия — является одновременно показателем влияния на функцию изменчивости различных ее факторов. Так, в табл. 27 мы наблюдали значительную колеблемость предприятий по выполнению плана валовой продукции. Эта колеблемость объясняется различиями, которые имели место между предприятиями по многочисленным факторам выполнения плана: обеспеченности рабочей силой, использованию рабочего времени, организации производства и т. д.

Для того чтобы выяснить степень влияния какого-либо одного из этих факторов, следовало бы измерить колеблемость функции под влиянием данного фактора при условии, что прочие независимо от него действующие факторы закреплены. Цель

будет достигнута, если удастся разложить полную дисперсию нашей функции y на две части, из которых одна измеряет влияние данного фактора x , другая — влияние прочих независимо от x действующих факторов. На примере корреляционной табл. 27 рассмотрим метод такого разложения. Интересующим нас фактором x является процент выполнения плана по числу отработанных человеко-дней.

Заметим, что в корреляционной таблице вся совокупность наблюдений (предприятий) делится на частные совокупности (столбцы таблицы), в каждом из которых признак x закреплен на некотором уровне. Так, в первой части совокупности $x=50$, во второй — $x=60$ и т. д.

Рассмотрим среднюю арифметическую $\bar{y} = \frac{\dot{\Sigma}y}{n}$ и дисперсию

$$\sigma^2 = \frac{1}{n} \dot{\Sigma} (y - \bar{y})^2$$

для общей совокупности и аналогичные показатели

$$\bar{y}_i = \frac{\dot{\Sigma}_i y}{h_i} \text{ и } \sigma_i^2 = \frac{1}{h_i} \dot{\Sigma}_i (y - \bar{y}_i)^2$$

в каждой из частных совокупностей. Согласно известным свойствам (§ 7, гл. II)

$$\bar{y} = \frac{1}{n} \Sigma \bar{y}_i h_i, \quad (7.1.1)$$

$$\sigma^2 = \bar{\sigma}_i^2 + \bar{\delta}_i^2, \quad (7.1.2)$$

где $\bar{\sigma}_i^2 = \frac{1}{n} \Sigma \sigma_i^2 h_i$ — средняя из частных дисперсий,

$$\bar{\delta}_i^2 = \frac{1}{n} \Sigma \delta_i^2 h_i \text{ — дисперсия частных средних.}$$

Обратимся к анализу формулы (7.1.2).

Так как в каждой частной совокупности x закреплен, то дисперсия σ_i^2 не может быть объяснена изменчивостью x . Мы поэтому вправе предположить, что первый компонент формулы (7.1.2) $\bar{\sigma}_i^2$ измеряет ту часть колеблемости y , которая возникает благодаря изменчивости прочих факторов, не зависящих от x .

В самом деле, положим

$$y = u + \bar{y}_i, \quad (7.1.3)$$

где u — отклонение y от \bar{y}_i , и докажем, что:

- 1) фактор u корреляционно не зависит от x ,
- 2) дисперсия фактора u равна $\bar{\sigma}_i^2$.

Корреляционная независимость u от x очевидна, так как для каждого значения x_i среднее \bar{u}_i равно нулю в силу нулевого свойства средней арифметической и, следовательно, не меняется с изменением x .

Равенство $\sigma_u^2 = \sigma_i^2$ доказывается следующим образом:

$$\begin{aligned}\bar{u} &= \frac{1}{n} (\dot{\Sigma}_1 u + \dot{\Sigma}_2 u + \dots + \dot{\Sigma}_s u) = \\ &= \frac{1}{n} [\dot{\Sigma}_1 (y - \bar{y}_1) + \dot{\Sigma}_2 (y - \bar{y}_2) + \dots + \dot{\Sigma}_s (y - \bar{y}_s)] = \\ &= \frac{1}{n} [0 + 0 + \dots + 0] = 0, \\ \sigma_u^2 &= \frac{1}{n} \dot{\Sigma} (u - \bar{u})^2 = \frac{1}{n} \dot{\Sigma} u^2 = \frac{1}{n} (\dot{\Sigma}_1 u^2 + \dot{\Sigma}_2 u^2 + \dots + \dot{\Sigma}_s u^2) = \\ &= \frac{1}{n} [\dot{\Sigma}_1 (y - \bar{y}_1)^2 + \dots + \dot{\Sigma}_s (y - \bar{y}_s)^2] = \\ &= \frac{1}{n} [h_1 \sigma_1^2 + h_2 \sigma_2^2 + \dots + h_s \sigma_s^2] = \frac{1}{n} \Sigma h_i \sigma_i^2 = \sigma_i^2.\end{aligned}$$

Итак, мы доказали, что первый компонент формулы (7.1.2) σ_i^2 измеряет ту часть колеблемости функции y , которая определяется изменчивостью факторов, корреляционно не зависящих от x . Но так как полная дисперсия функции σ^2 возникает также и в связи с изменчивостью x , то можно предположить, что эта последняя отражена во втором компоненте формулы (7.1.2) δ_i^2 .

В самом деле, \bar{y}_i изменяется в основном под влиянием фактора x (в основном, но не полностью, так как при ограниченном числе наблюдений закономерное изменение \bar{y}_i нарушается случайными зигзагами, возникающими благодаря остаточному действию прочих факторов). Следовательно, δ_i^2 представляет колеблемость y в той ее части, которая в основном обусловлена влиянием фактора x .

Отметим следующее важное обстоятельство. Хотя фактор u корреляционно не зависит от x , он может находиться с последним в более общей статистической связи, в частности, дисперсия фактора u — σ_i^2 может систематически изменяться с изменением x . В этом случае величина $\sigma_i^2 = \frac{h_1 \sigma_1^2 + h_2 \sigma_2^2 + \dots + h_s \sigma_s^2}{h_1 + h_2 + \dots + h_s}$ будет, очевидно, зависеть от системы весов h_1, h_2, \dots, h_s и тем самым отражать и изменчивость фактора x . Следовательно, для того чтобы разложение (7.1.2) имело указанный выше физический смысл, необходимо, чтобы дисперсия σ_i^2 систематически не изменялась с изменением x , т. е. ее предельное значение было по-

стоянным относительно x . Это предположение и принимается в настоящем и следующем параграфах при истолковании физического смысла показателей тесноты зависимости. Графическую интерпретацию формулы (7.1.2) мы видим на рис. 23. На поле корреляции проведена эмпирическая линия регрессии и линия общей средней $\bar{y} = C$, параллельная оси абсцисс.

Общая дисперсия $\sigma^2 = \frac{1}{n} \Sigma (y - \bar{y})^2$ представляет средний квадрат отклонений (по ординате) точек корреляционного поля от линии $\bar{y} = C$.

$$\bar{\sigma}_i^2 = \frac{1}{n} \Sigma (y - \bar{y}_i)^2$$

представляет средний квадрат отклонений точек корреляционного поля от эмпирической линии регрессии.

$$\bar{\delta}_i^2 = \frac{1}{n} \Sigma (\bar{y}_i - \bar{y})^2$$

представляет средний квадрат отклонений эмпирической линии регрессии от линии $\bar{y} = C$. Рассмотрим показатель

$$\eta^2 = \frac{\bar{\delta}_i^2}{\sigma^2}. \quad (7.1.4)$$

η^2 (эта квадрат) отвечает на вопрос: *какая часть полной колеблемости y обусловлена изменчивостью аргумента x ?* Корень квадратный из величины η^2 называется *эмпирическим корреляционным отношением y по x* и обозначается η ; η служит *измерителем тесноты зависимости y от x* .

Очевидно, что η всегда заключено в пределах от 0 до 1

$$0 \leq \eta \leq 1.$$

Если $\eta = 0$, то $\bar{\delta}_i^2 = 0$ и, следовательно, отсутствует изменчивость частных средних \bar{y}_i . Линия регрессии параллельна оси абсцисс и совпадает с линией $\bar{y} = C$. В этом случае говорят об отсутствии корреляции y от x ¹.

Если $\eta = 1$, то $\bar{\sigma}_i^2 = 0$ и, следовательно, все точки корреляционного поля лежат на эмпирической линии регрессии. В этом случае говорят о наличии *точной, или функциональной, зависимости y от x* ².

¹ При условии безгранично большого числа наблюдений. Если число наблюдений невелико, то равенство $\eta = 0$ может случайно получиться и при наличии корреляции, и, наоборот, при отсутствии корреляции η случайно может оказаться > 0 .

² При наличии безгранично большого числа наблюдений. Если число наблюдений невелико, то равенство $\eta = 1$ может случайно получиться и при неточной зависимости.

Наконец, в общем случае, когда $0 < \eta < 1$, говорят о наличии более или менее тесной корреляционной зависимости y от x (рис. 40).

При вычислении η удобнее использовать замену переменного y через $y' = \frac{y - C_y}{i_y}$.

Так как $\frac{\overline{\delta_l^2}}{\sigma^2} = \frac{\overline{\delta_l'^2}}{\sigma'^2}$, где $\overline{\delta_l'^2}$ и σ'^2 рассчитаны для упрощенных

вариантов y' , то η может вычисляться по формуле

$$\eta = \sqrt{\frac{\overline{\delta_l'^2}}{\sigma'^2}} \quad (7. 1. 5)$$

Пример расчета η приводим для корреляционной табл. 30. Вычисления располагаем в табл. 41.

Таблица 41

Расчет эмпирического корреляционного отношения η

A		Номера столбцов							
		1	2	3					
y'	x	5	10	15	20	Итого l	ly'	ly'^2	
2	50			2 ⁴	1 ²	3	6	12	
1	40		3 ³	6 ⁶	6 ⁶	15	15	15	
0	30	3 ⁰	8 ⁰	6 ⁰	3 ⁰	20	0	0	
-1	20	5 ⁻⁵	4 ⁻⁴	1 ⁻¹		10	-10	10	
-2	10	2 ⁻⁴				2	-4	8	
Номера строк	1	Итого h	10	15	15	10	50	7	45
	2	$\Sigma ly'$	-9	-1	9	8	7		
	3	\bar{y}'_i	-0,90	-0,07	0,60	0,80			

$$\bar{y}' = \frac{7}{50} = 0,14, \quad \sigma'^2 = \frac{45}{50} - 0,14^2 = 0,900 - 0,0196 = 0,8804$$

Продолжение

В

x	h	\bar{y}_i'	$\bar{y}_i' - \bar{y}'$	$(\bar{y}_i' - \bar{y}')^2$	$(\bar{y}_i' - \bar{y}')^2 h$
5	10	-0,90	-1,04	1,08	10,80
10	15	-0,07	-0,21	0,04	0,60
15	15	0,60	0,46	0,21	3,15
20	10	0,80	0,66	0,44	4,40
	50				18,95

$$\bar{\delta}_i'^2 = \frac{18,95}{50} = 0,379, \quad \eta = \sqrt{\frac{0,379}{0,8804}} = \sqrt{0,4305} = 0,656.$$

Прежде всего находим показатели \bar{y}' , σ'^2 и \bar{y}_i' в разделе А табл. 41.

В разделе В рассчитываем показатель $\bar{\delta}_i'^2$ и затем η . Произведя аналогичные расчеты для зависимости между валовой продукцией и числом отработанных человеко-дней (см. табл. 27), получаем $\eta = 0,693$.

Так как $\eta^2 = 0,4796$, то отсюда заключаем, что дисперсия y , связанная с изменчивостью x , составляет 48% от полной дисперсии y . Можно, следовательно, сказать, что колеблемость рассматриваемых предприятий по проценту выполнения плана валовой продукции на 48% обусловлена различием этих предприятий по фактору x (процент выполнения плана по числу человеко-дней) и на 52% обусловлена влиянием прочих факторов, варьирующих независимо от x .

§ 2. ТЕОРЕТИЧЕСКОЕ КОРРЕЛЯЦИОННОЕ ОТНОШЕНИЕ

Метод разложения дисперсии, рассмотренный в предыдущем параграфе, представляет собой применение метода группировок к анализу дисперсии объединенной совокупности. Благодаря группировке совокупности по признаку x формула (7.1.2), расчлениющая полную дисперсию y на среднюю из частных дисперсий и дисперсию частных средних, приобретает новый смысл. Она решает задачу разложения дисперсии на части, обусловленную и необусловленную изменчивостью фактора x .

В пределе при неограниченно большом числе наблюдений формула (7.1.2) дает вполне точное и универсальное решение последней задачи в условиях, оговоренных в предыдущем параграфе.

Однако при ограниченном числе наблюдений наряду с неизбежно возникающими случайными погрешностями величин

$$\sigma^2, \bar{\sigma}_i^2, \bar{\delta}_i^2,$$

мы допускаем в формуле (7.1.2) также и систематическую ошибку, преувеличивая компонент $\bar{\delta}_i^2$ и преуменьшая компонент $\bar{\sigma}_i^2$. Эта ошибка возникает благодаря тому, что y_i при ограниченном числе наблюдений изменяется не только под влиянием фактора x , но обнаруживает также и случайные колебания, объясняющиеся остаточным действием прочих, не зависящих от x факторов. Эта случайная часть колеблемости y_i преувеличивает компонент $\bar{\delta}_i^2$ и одновременно преуменьшает компонент $\bar{\sigma}_i^2$, так как в последнем недоучтено влияние побочных факторов¹.

Поэтому эмпирическое корреляционное отношение систематически преувеличивает тесноту связи между y и x , преувеличивает тем сильнее, чем меньше наблюдений и чем больше столбцов в корреляционной таблице (с увеличением количества столбцов в каждом из них оказывается меньше наблюдений, а это ведет к более значительным случайным колебаниям y_i).

В разделе III книги будут указаны поправки, которые нужно внести в вычисление η для того,

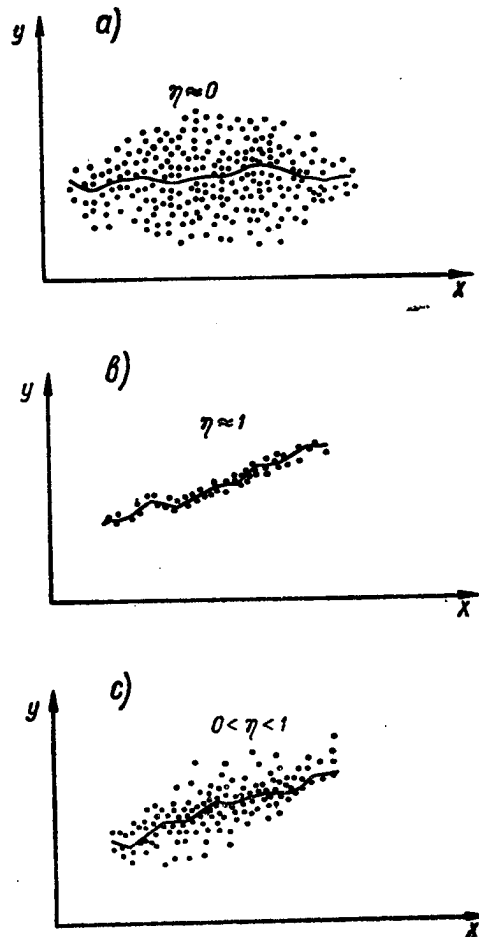


Рис. 40

¹ При ограниченном числе наблюдений величина σ^2 также имеет свою систематическую ошибку (см. § 3, глава XIII, раздел III), но она мала по сравнению с систематическими ошибками компонент $\bar{\sigma}_i^2$ и $\bar{\delta}_i^2$.

чтобы устранить систематическую ошибку, возникающую благодаря случайным зигзагам линии регрессии. Сейчас рассмотрим другой естественный прием, направленный к той же цели. Он заключается в замене эмпирической линии регрессии линией теоретической при расчете показателей σ_i^2 и δ_i^2 .

Пусть \bar{y}_x — частная средняя, вычисленная по уравнению регрессии $\bar{y}_x = f(x)$. Заменяя этой средней эмпирическую среднюю \bar{y}_i в формулах для σ_i^2 , $\bar{\sigma}_{iT}^2$, $\bar{\delta}_i^2$, получаем соответственно

$$\sigma_{iT}^2 = \frac{1}{h_i} \sum_i (y - \bar{y}_x)^2, \quad (7.2.1)$$

$$\bar{\sigma}_{iT}^2 = \frac{1}{n} \sum \sigma_{iT}^2 h_i = \frac{1}{n} \sum (y - \bar{y}_x)^2, \quad (7.2.2)$$

$$\bar{\delta}_{iT}^2 = \frac{1}{n} \sum \delta_{iT}^2 h_i = \frac{1}{n} \sum (\bar{y}_x - \bar{y})^2. \quad (7.2.3)$$

Графически $\bar{\sigma}_{iT}^2$ выражает средний квадрат отклонений точек корреляционного поля от теоретической линии регрессии, $\bar{\delta}_{iT}^2$ — средний квадрат отклонений теоретической линии регрессии от линии общей средней: $\bar{y} = \bar{C}$ (все отклонения измеряются по ординате).

Для показателей $\bar{\sigma}_{iT}^2$ и $\bar{\delta}_{iT}^2$ справедливо следующее предложение.

Если в уравнении регрессии \bar{y}_x является линейной функцией от параметров a, b, c, \dots , где a — «свободный» параметр,

$$\bar{y}_x = a + bp(x) + cq(x) + \dots, \quad (7.2.4)$$

то

$$\sigma^2 = \bar{\sigma}_{iT}^2 + \bar{\delta}_{iT}^2. \quad (7.2.5)$$

Доказательство

$$\begin{aligned} \sigma^2 &= \frac{1}{n} \sum (y - \bar{y})^2 = \frac{1}{n} \sum \{(y - \bar{y}_x) + (\bar{y}_x - \bar{y})\}^2 = \\ &= \frac{1}{n} \sum (y - \bar{y}_x)^2 + \frac{1}{n} \sum (\bar{y}_x - \bar{y})^2 + 2 \frac{1}{n} \sum (y - \bar{y}_x)(\bar{y}_x - \bar{y}). \end{aligned}$$

Но

$$\frac{1}{n} \sum (y - \bar{y}_x)^2 = \bar{\sigma}_{iT}^2; \quad \frac{1}{n} \sum (\bar{y}_x - \bar{y})^2 = \bar{\delta}_{iT}^2.$$

Докажем, что третий компонент $2 \frac{1}{n} \dot{\Sigma} (y - \bar{y}_x) (\bar{y}_x - \bar{y})$ при условии (7.2.4) тождественно равен нулю.

Для этой цели представим его в виде

$$\begin{aligned} & 2 \frac{1}{n} \dot{\Sigma} (y - \bar{y}_x) (\bar{y}_x - \bar{y}) = \\ & = 2 \frac{1}{n} \dot{\Sigma} (y - \bar{y}_x) \bar{y}_x - 2 \frac{1}{n} \bar{y} \cdot \dot{\Sigma} (y - \bar{y}_x). \end{aligned}$$

Далее, составим систему нормальных уравнений для определения параметров a, b, c, \dots

$$\frac{\partial f}{\partial a} = 0; \quad \frac{\partial f}{\partial b} = 0; \quad \frac{\partial f}{\partial c} = 0 \dots,$$

где $f = \dot{\Sigma} (y - \bar{y}_x)^2$.

Производя дифференцирование, получим

$$\left. \begin{aligned} \dot{\Sigma} (y - \bar{y}_x) &= 0 \\ \dot{\Sigma} (y - \bar{y}_x) p(x) &= 0 \\ \dot{\Sigma} (y - \bar{y}_x) q(x) &= 0 \end{aligned} \right\} \quad (7.2.6)$$

Но из первого уравнения системы (7.2.6) следует, что

$$2 \frac{1}{n} \bar{y} \dot{\Sigma} (y - \bar{y}_x) = 0.$$

Далее умножаем первое из уравнений (7.2.6) на a , второе — на b , третье — на c и т. д.; складывая, получаем

$$\dot{\Sigma} (y - \bar{y}_x) [a + bp(x) + cq(x) + \dots] = 0$$

или

$$\dot{\Sigma} (y - \bar{y}_x) \bar{y}_x = 0.$$

Отсюда

$$2 \frac{1}{n} \dot{\Sigma} (y - \bar{y}_x) (\bar{y}_x - \bar{y}) = 0,$$

что и требовалось доказать.

Итак, равенство (7.2.5) осуществляет, подобно равенству (7.1.2), разложение дисперсии y на два компонента, из которых первый измеряет влияние факторов, действующих независимо от аргумента x , второй — влияние аргумента x .

Вместо показателя η мы можем теперь использовать показатель

$$\eta_T = \sqrt{\frac{\overline{\delta_{iT}^2}}{\sigma^2}}, \quad (7.2.7)$$

называемый *теоретическим корреляционным отношением y по x* и имеющий аналогичный с η смысл измерителя тесноты зависимости y от x . Очевидно, что

$$0 \leq \eta_T \leq 1. \quad (7.2.8)$$

Когда $\eta_T = 0$, то $\overline{\delta_{iT}^2} = 0$ и теоретическая линия регрессии параллельна оси абсцисс. При весьма большом числе наблюдений в этом случае говорят об отсутствии корреляции между y и x .

Когда $\eta_T = 1$, то $\overline{\sigma_{iT}^2} = 0$ и все точки поля лежат на теоретической линии регрессии. При большом числе наблюдений в этом случае считают зависимость точной — функциональной.

Наконец, при $0 < \eta_T < 1$ говорят о наличии корреляционной зависимости между y и x , более или менее тесной (рис. 41).

Легко видеть, что в равенстве (7.2.5) дисперсия перераспределяется между двумя компонентами по сравнению с равенством (7.1.2) следующим образом: первый из компонентов увеличивается и соответственно уменьшается второй. Действительно, в каждом столбце корреляционной таблицы

$$\sigma_i^2 \leq \sigma_{iT}^2$$

согласно минимальному свойству средней арифметической. Следовательно,

$$\overline{\sigma_i^2} \leq \overline{\sigma_{iT}^2}, \quad (7.2.9)$$

а в силу равенств (7.1.2) и (7.2.5)

$$\overline{\delta_i^2} \geq \overline{\delta_{iT}^2}, \quad (7.2.9')^1$$

¹ Средний квадрат отклонений точек поля от эмпирической линии регрессии меньше, чем от какой-либо другой линии, и, в частности, меньше, чем от теоретической линии регрессии. На первый взгляд это может показаться странным, ибо теоретическая линия регрессии — прямая, парабола и т. д. — также получается на основе требования наименьших квадратов. Однако легко видеть, что каждая из этих линий удовлетворяет этому требованию лишь по сравнению с другими линиями *того же самого типа*. Так, например, теоретическая прямая регрессии является «лучшей» из всех возможных прямых, теоретическая парабола регрессии является «лучшей» параболой и т. д. Между тем эмпирическая линия регрессии осуществляет требование **наименьших квадратов по сравнению с любой** иной линией.

Отсюда

$$\eta_T \leq \eta \quad (7.2.10)$$

или теоретическое корреляционное отношение меньше или, в крайнем случае, равно эмпирическому корреляционному отношению.

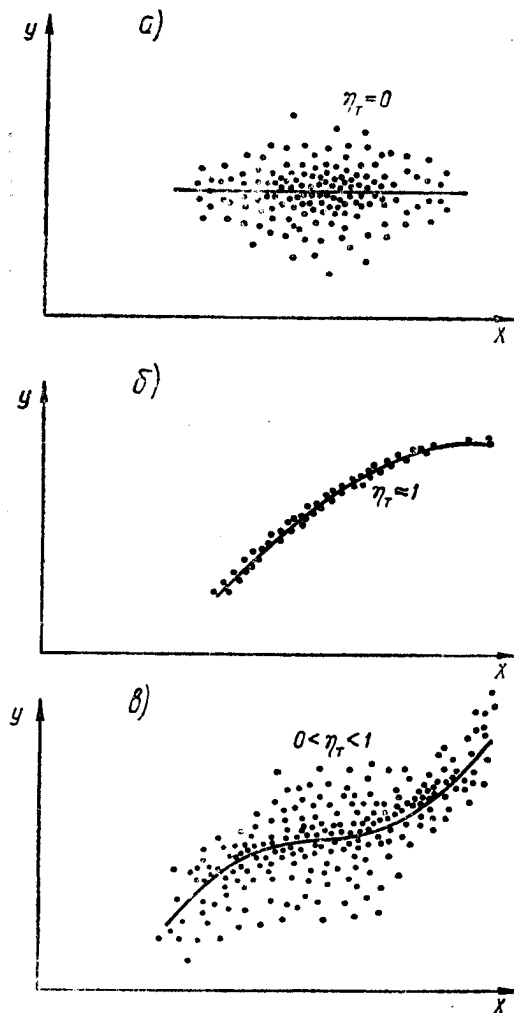


Рис. 41

Неравенство (7.2.10) показывает, что η_T в какой-то мере исправляет систематическую ошибку, свойственную эмпирическому корреляционному отношению. Однако устраняется ли полностью эта ошибка, когда мы переходим от эмпирического к теоретическому корреляционному отношению, покажет дальнейшее исследование.

10 я. И. Лукомский

ретическому корреляционному отношению? На это можно дать **отрицательный ответ.**

В самом деле, использованная нами при определении компонентов $\overline{\sigma}_{IT}^2$ и $\overline{\delta}_{IT}^2$ теоретическая линия регрессии не является предельной теоретической линией, а получена на ограниченном количестве наблюдений косвенным путем. Она также, хотя и в меньшей степени, чем эмпирическая линия регрессии, подвержена влиянию случайных факторов. Это влияние проявляется не в колебаниях тех или иных участков кривой, как это имеет место у эмпирической линии регрессии; оно сказывается на положении всей линии, которая может случайно отклоняться от своего предельного положения. Случайная колеблемость \overline{y}_x приводит к некоторому систематическому завышению η_T .

Суть дела заключается в том, что теоретическая линия регрессии, рассчитанная по способу наименьших квадратов, *ближе подходит* к тем точкам поля, на основе которых она рассчитана, чем *действительная предельная линия регрессии.* Поэтому η_T имеет тенденцию к преувеличению тесноты связи, хотя и не в такой мере, как η . Разумеется, это справедливо при условии, что тип уравнения регрессии выбран правильно и \overline{y}_x лишь случайно отклоняется от предельных $\overline{Y}(x)$.

Если же тип уравнения регрессии выбран неправильно, если, например, он слишком упрощает форму связи (выбор прямолинейного изменения вместо параболы и т. п.), то теоретическая линия регрессии, вычисленная по способу наименьших квадратов, может оказаться более удаленной от тех точек, на которых она рассчитана, чем предельная теоретическая линия регрессии. Это может привести к систематическому *занижению* тесноты зависимости в показателе η_T , *занижению тем более значительному, чем сильнее упрощает выбранный тип уравнения регрессии действительную зависимость.*

Поправки, которые следует внести в вычисление η_T для устранения систематических ошибок, будут рассмотрены в следующем **разделе книги.**

Приводим в качестве примера расчет η_T для корреляционной табл. 33.

η_T можно вычислять в упрощенных координатах по формуле

$$\eta_T = \sqrt{\frac{\overline{\delta}_{IT}^2}{\sigma^2}}$$

Теоретическую линию регрессии мы рассчитывали для этой таблицы в двух вариантах: в виде прямой и в виде параболы второго порядка. Поэтому и расчет η_T можно здесь осуществить в двух вариантах применительно к каждому типу теоретической линии регрессии. Расчет показан в табл. 42 и 43.

Расчет теоретического корреляционного отношения η_T

Уравнение регрессии (в упрощенных координатах)

$$\bar{y}'_{x'} = -0,151 + 0,581x' \text{ (см табл. 33).}$$

Таблица 42

x'	h	$\bar{y}'_{x'}$	$\bar{y}'_{x'} - \bar{y}'$	$(\bar{y}'_{x'} - \bar{y}')^2$	$(\bar{y}'_{x'} - \bar{y}')^2 h$
-1	10	-0,732	-0,872	0,76	7,6
0	15	-0,151	-0,291	0,08	1,2
1	15	0,430	0,290	0,08	1,2
2	10	1,011	0,871	0,76	7,6
	50				17,6

$$\bar{\delta}_{iT}^2 = \frac{17,6}{50} = 0,3520,$$

 \bar{y}' и σ'^2 берем из табл. 41.

Отсюда

$$\eta_T = \sqrt{\frac{0,3520}{0,8804}} = 0,632.$$

Уравнение регрессии (в упрощенных координатах)

$$\bar{y}'_{x'} = -0,0233 + 0,740x' - 0,159x'^2 \text{ (см. табл. 34).}$$

Таблица 43

x'	h	$\bar{y}'_{x'}$	$\bar{y}'_{x'} - \bar{y}'$	$(\bar{y}'_{x'} - \bar{y}')^2$	$(\bar{y}'_{x'} - \bar{y}')^2 h$
-1	10	-0,922	-1,062	1,13	11,3
0	15	-0,023	-0,163	0,03	0,45
1	15	0,558	0,418	0,17	2,55
2	10	0,821	0,681	0,46	4,6
	50				18,90

$$\bar{\delta}_{iT}^2 = \frac{18,90}{50} = 0,378,$$

 \bar{y}' и σ'^2 берем из табл. 41.

Откуда

$$\eta_T = \sqrt{\frac{0,3780}{0,8804}} = 0,655.$$

Для зависимости между валовой продукцией и числом человеко-дней теоретические корреляционные отношения, рассчитанные аналогичным путем, оказались равными:

- 1) применительно к прямой регрессии $\eta_T = 0,653$;
- 2) применительно к параболе регрессии второго порядка $\eta_T = 0,691$.

Оба корреляционных отношения практически мало отличаются от вычисленного ранее эмпирического корреляционного отношения.

§ 3. КОЭФФИЦИЕНТ КОРРЕЛЯЦИИ

Рассмотрим наиболее важный для практики и для теории случай, когда теоретическая линия регрессии выражается уравнением первой степени, т. е. случай линейной зависимости

$$\bar{y}_x = a + bx.$$

Угловым коэффициентом уравнения регрессии b , или, как его называют, *коэффициент регрессии y по x* , показывает, на сколько единиц в среднем изменяется y , когда x увеличивается на одну единицу.

Чем больше коэффициент регрессии, тем более значительно изменение среднего y , соответствующее тому же изменению x . Этот смысл коэффициента регрессии делает его на первый взгляд подходящим измерителем силы зависимости y от x .

Однако возникают следующие возражения. Численное значение коэффициента регрессии b зависит от того, что принято за единицу измерения по x и по y . Если единица измерения x велика, то увеличению x на единицу соответствует большее изменение среднего y , т. е. b велико. Если единица измерения y велика, то соответствующее изменение y выражается меньшим количеством единиц и, следовательно, b мало. Например, для уравнения регрессии веса человеческого тела по его росту мы получим в 100 раз меньший коэффициент b , измеряя рост в сантиметрах, по сравнению с тем, если его измерять в метрах. Наоборот, мы получим в 1000 раз больший коэффициент регрессии, если вес будем измерять в граммах вместо килограммов.

Так как обычно мы рассматриваем зависимости между признаками разной природы, имеющими различные единицы измерения, то, очевидно, что коэффициент b является величиной, зависящей от произвола в выборе единиц измерения переменных, и поэтому не может служить универсальным показателем силы связи. Однако если бы удалось найти такую стандартную систему единиц измерения, в которой данные по различным признакам оказались сравнимы между собой, то влияние произвола в выборе единицы измерения на коэффициент регрессии было бы устранено. Статистика знает такую систему единиц. Эта система ис-

пользует в качестве единицы измерения признака его среднее квадратическое отклонение σ (при этом начало отсчета обычно перемещается в точку средней арифметической). Мы вправе полагать, что в стандартизованной системе коэффициент регрессии окажется подходящим измерителем силы связи. Исследуем этот вопрос.

Коэффициент регрессии в стандартизованной системе единиц или, короче, стандартизованный коэффициент регрессии, отвечает на вопрос: на сколько сигм изменяется в среднем y , когда x увеличивается на одну сигму?

Стандартизованный коэффициент регрессии y по x обычно называется *коэффициентом корреляции* между y и x и обозначается буквой r_{yx} или просто r .

Очевидно, мы получим r , если умножим b на отношение единиц (сигм): это вытекает из определения r и b .

Итак,

$$r = b \frac{\sigma_x}{\sigma_y}. \quad (7.3.1)$$

Формула (7.3.1) — основная исходная формула коэффициента корреляции.

Покажем теперь, что коэффициент корреляции в случае прямолинейной регрессии действительно измеряет силу зависимости. Для этой цели рассмотрим формулу теоретического корреляционного отношения

$$\eta_T^2 = \frac{\overline{\delta_{IT}^2}}{\sigma^2} = \frac{1}{n} \frac{\sum (\bar{y}_x - \bar{y})^2}{\sigma^2}.$$

В этой формуле заменим \bar{y}_x и \bar{y} их выражениями:

$$\bar{y}_x = a + bx,$$

$$\bar{y} = a + b\bar{x};$$

последнее равенство вытекает из первого нормального уравнения системы (5.1.3):

$$\sum y = na + b\sum x$$

после деления на n .

Получаем:

$$\eta_T^2 = \frac{\frac{1}{n} \sum b^2 (x - \bar{x})^2}{\sigma^2} = \frac{b^2 \frac{1}{n} \sum (x - \bar{x})^2}{\sigma^2} = b^2 \frac{\sigma_x^2}{\sigma_y^2}$$

(σ^2 , стоящей в знаменателе, приспан знак y для отличия от σ_x^2 , стоящей в числителе).

Сопоставляя с формулой (7.3.1), получаем

$$\eta_T^2 = r^2 \quad (7.3.2)$$

или

$$\eta_T = |r|, \quad (7.3.2')$$

т. е. коэффициент корреляции по абсолютной величине равен теоретическому корреляционному отношению (в случае прямолинейной регрессии).

Отсюда следует, что коэффициент корреляции можно считать измерителем *силы связи* y с x , если понятие силы связи, до сих пор нами не вполне определенное, считать равнозначным *понятию тесноты зависимости*.

Из равенства (7.3.2') заключаем, что абсолютная величина коэффициента корреляции способна изменяться так же, как и η_T в пределах от 0 до 1.

Однако в отличие от η_T коэффициент корреляции может принимать как положительные, так и отрицательные значения. Знак при r совпадает со знаком коэффициента регрессии b ; это вытекает из уравнения (7.3.1).

Итак,

$$-1 \leq r \leq +1, \quad (7.3.3)$$

при $r = +1$ или $r = -1$ все точки корреляционного поля лежат на прямой регрессии. В этом случае имеет место строгая пропорциональность в изменении y и x (при безгранично большом числе наблюдений).

При r положительном \bar{y}_x растет с увеличением x — имеет место положительная корреляция более или менее тесная; при r отрицательном \bar{y}_x падает с ростом x — имеет место отрицательная корреляция. Наконец, при $r = 0$ прямая регрессии параллельна оси x , и прямолинейная корреляция отсутствует (если число наблюдений безгранично велико).

Приводим ряд корреляционных таблиц, соответствующих последовательному нарастанию силы корреляционной зависимости (см. табл. 44)¹. Пользуясь этими таблицами как системой шаблонов, можно приближенно на глаз определить значение коэффициента корреляции для таблиц, встречающихся на практике.

Различные модификации формулы коэффициента корреляции

Коэффициент корреляции r определяется формулой (7.3.1). Полезно также запомнить и некоторые модификации (видоизме-

¹ Заимствованы из книги А. К. Митропольского, Техника статистического исчисления, 1931, стр. 372.

Таблица 44

Теснота связи
 $r = 0$

II \ I		II										Σ	
		0	1	2	3	4	5	6	7	8	9		10
10						1							1
9				2	2	2	2	2					10
8			2	5	9	12	9	5	2				44
7		2	5	14	24	27	24	14	5	2			117
6		2	9	24	42	51	42	24	9	2			205
5	1	2	12	27	51	60	51	27	12	2	1		246
4		2	9	24	42	51	42	24	9	2			205
3		2	5	14	24	27	24	14	5	2			117
2			2	5	9	12	9	5	2				44
1				2	2	2	2	2					10
0						1							1
Σ		1	10	44	117	205	246	205	117	44	10	1	1000

$r = 0,2$

II \ I		II										Σ	
		0	1	2	3	4	5	6	7	8	9		10
10							1						1
9					1	1	2	3	2	1			10
8			1	3	7	10	11	8	3	3			44
7		1	3	9	20	29	27	18	8	2			117
6		1	7	20	39	51	45	27	11	3	1		205
5		2	10	29	51	62	51	29	10	2		1	246
4	1	3	11	27	45	51	39	20	7	1			205
3		2	8	18	27	29	20	9	3	1			117
2		1	3	8	11	10	7	3	1				44
1				1	2	3	2	1	1				10
0					1		1						1
Σ		1	10	44	117	205	246	205	117	44	10	1	1000

$r = 0,4$

II \ I		II										Σ	
		0	1	2	3	4	5	6	7	8	9		10
10								1					1
9						1	3	3	2	1			10
8				1	4	9	12	10	6	2			44
7			1	5	15	27	32	23	10	3		1	117
6			4	15	36	53	50	32	12	3			205
5		1	9	27	53	66	53	27	9	1			246
4		3	12	32	50	53	36	15	4				205
3	1	3	10	23	32	27	15	5	1				117
2		2	6	10	12	9	4	1					44
1		1	2	3	3	1							10
0				1									1
Σ		1	10	44	117	205	246	205	117	44	10	1	1000

Продолжение

$r = 0,6$

II												Σ
I \ II	0	1	2	3	4	5	6	7	8	9	10	
10									1			1
9						1	2	3	3	1		10
8				2	8	5	11	14	9	3	1	44
7				8	31	24	36	30	14	3	1	117
6			1	8	31	56	60	36	11	2		205
5		1	5	24	56	74	56	24	5	1		246
4		2	11	36	60	56	31	8				205
3		3	14	30	36	24	8	2				117
2	1	3	9	14	11	5	1					44
1		1	3	3	2	1						10
0			1									1
Σ	1	10	44	117	205	246	205	117	44	10	1	1000

$r = 0,8$

II												Σ
I \ II	0	1	2	3	4	5	6	7	8	9	10	
10										1		1
9								2	4	3	1	10
8							7	18	15	4		44
7						13	41	43	18	2		117
6					17	62	78	41	7			205
5				13	62	96	62	13				246
4			7	41	78	62	17					205
3		2	18	43	41	13						117
2		4	15	18	7							44
1	1	3	4	2								10
0		1										1
Σ	1	10	44	117	205	246	205	117	44	10	1	1000

$r = 1,0$

II												Σ
I \ II	0	1	2	3	4	5	6	7	8	9	10	
10											1	1
9										10		10
8									44			44
7								117				117
6							205					205
5						246						246
4					205							205
3				117								117
2			44									44
1		10										10
0	1											1
Σ	1	10	44	117	205	246	205	117	44	10	1	1000

нения) формулы (7.3.1), которые получаются тождественным преобразованием последней.

Определим коэффициент b из системы нормальных уравнений (5.1.3)

$$\left. \begin{aligned} \Sigma y &= na + b \Sigma x \\ \Sigma xy &= a \Sigma x + b \Sigma x^2 \end{aligned} \right\}$$

которая после деления на n приводится к виду

$$\left. \begin{aligned} \bar{y} &= a + b\bar{x} \\ \overline{yx} &= a\bar{x} + b \frac{\Sigma x^2}{n} \end{aligned} \right\} \quad (7.3.4)$$

где $\overline{yx} = \frac{\Sigma yx}{n}$.

Решая систему (7.3.4) относительно b , получим

$$b = \frac{\begin{vmatrix} 1 & \bar{y} \\ \bar{x} & \overline{yx} \end{vmatrix}}{\begin{vmatrix} 1 & \bar{x} \\ \bar{x} & \frac{\Sigma x^2}{n} \end{vmatrix}} = \frac{\bar{y}\bar{x} - \overline{yx}}{\frac{\Sigma x^2}{n} - (\bar{x})^2} = \frac{\bar{y}\bar{x} - \overline{yx}}{\sigma_x^2}. \quad (7.3.5)$$

Подставляя это выражение в формулу (7.3.1), получаем первую модификацию формулы r

$$r = \frac{\bar{y}\bar{x} - \overline{yx}}{\sigma_x \sigma_y}. \quad (7.3.6)$$

Из (7.3.6) легко получить и следующее, второе видоизменение формулы r

$$r = \frac{\frac{1}{n} \Sigma (x - \bar{x})(y - \bar{y})}{\sigma_x \sigma_y}. \quad (7.3.7)$$

В самом деле,

$$\begin{aligned} \frac{1}{n} \Sigma (x - \bar{x})(y - \bar{y}) &= \frac{1}{n} \Sigma xy - \frac{1}{n} \Sigma x\bar{y} - \\ &\quad - \frac{1}{n} \Sigma x\bar{y} + \frac{1}{n} \Sigma \bar{x}\bar{y} = \overline{xy} - \bar{x} \frac{1}{n} \Sigma y - \\ &\quad - \bar{y} \frac{1}{n} \Sigma x + \frac{1}{n} n\bar{x}\bar{y} = \overline{xy} - \bar{x}\bar{y} - \bar{y}\bar{x} + \bar{x}\bar{y} = \\ &= \overline{xy} - \bar{x}\bar{y}, \end{aligned}$$

из чего следует справедливость формулы (7.3.7).

Формулы (7.3.1), (7.3.6), (7.3.7) представляют в основном теоретический интерес и будут в дальнейшем нами использованы для целей теоретического анализа.

Вычислять же коэффициент корреляции удобнее по формуле

$$r = \frac{n\sum xy - \sum x \sum y}{\sqrt{n\sum x^2 - (\sum x)^2} \sqrt{n\sum y^2 - (\sum y)^2}}, \quad (7.3.8)$$

которая получается из формулы (7.3.6) путем умножения числителя и знаменателя на n^2 .

Рассчитаем коэффициент корреляции по формуле (7.3.8) для примера на стр. 94.

Таблица 45

Расчет коэффициента корреляции

x	y	x^2	y^2	xy
51	25	2 601	625	1 275
67	30	4 489	900	2 010
84	43	7 056	1 849	3 612
81	44	6 561	1 936	3 564
101	57	10 201	3 249	5 757
109	58	11 881	3 364	6 322
71	43	5 041	1 849	3 053
97	46	9 409	2 116	4 462
109	62	11 881	3 844	6 758
91	45	8 281	2 025	4 095
105	55	11 025	3 025	5 775
89	44	7 921	1 936	3 916
$\Sigma = 1055$	552	96 347	26 718	50 599

Отсюда

$$r = \frac{12 \cdot 50 599 - 1055 \cdot 552}{\sqrt{12 \cdot 96 347 - 1055^2} \sqrt{12 \cdot 26 718 - 552^2}} =$$

$$= \frac{24 828}{\sqrt{43 139} \sqrt{15 912}} = \frac{24 828}{26203,7} = 0,948.$$

Коэффициент корреляции между пределом выносливости и пределом прочности стали оказался весьма высоким. Однако число наблюдений ($n=12$) слишком мало для того, чтобы этот результат можно было считать вполне надежным.

Когда число наблюдений велико, материал подвергается группировке и заносится в корреляционную таблицу. В этом случае применяется преобразование (5.2.4)

$$x' = \frac{x - C_x}{i_x}; \quad y' = \frac{y - C_y}{i_y},$$

и коэффициент корреляции вычисляется между переменными y' и x' по формуле

$$r_{y'x'} = \frac{n\dot{\Sigma}x'y' - \dot{\Sigma}x'\dot{\Sigma}y'}{\sqrt{n\dot{\Sigma}x'^2 - (\dot{\Sigma}x')^2} \sqrt{n\dot{\Sigma}y'^2 - (\dot{\Sigma}y')^2}}. \quad (7.3.8')$$

Преобразование (5.2.4) не отражается на величине коэффициента корреляции, поэтому

$$r_{yx} = r_{y'x'}. \quad (7.3.9)$$

В самом деле, из уравнений (5.2.4) получаем

$$x = x'i_x + C_x, \quad y = y'i_y + C_y;$$

отсюда

$$\bar{x} = \bar{x}'i_x + C_x, \quad \bar{y} = \bar{y}'i_y + C_y; \quad (7.3.10)$$

$$\sigma_x = i_x\sigma'_x, \quad \sigma_y = i_y\sigma'_y; \quad (7.3.11)$$

$$\frac{1}{n} \Sigma (x - \bar{x})(y - \bar{y}) = i_x i_y \frac{1}{n} \Sigma (x' - \bar{x}')(y' - \bar{y}').$$

Следовательно,

$$\begin{aligned} r_{yx} &= \frac{\frac{1}{n} \Sigma (x - \bar{x})(y - \bar{y})}{\sigma_x \sigma_y} = \frac{i_y i_x \frac{1}{n} \Sigma (x' - \bar{x}')(y' - \bar{y}')}{i_x i_y \sigma'_x \sigma'_y} = \\ &= \frac{\frac{1}{n} \Sigma (x' - \bar{x}')(y' - \bar{y}')}{\sigma'_x \sigma'_y} = r_{y'x'}, \end{aligned}$$

что и требовалось доказать.

Расчет r на примере табл. 33 показан в табл. 46.

$$\begin{aligned} r &= \frac{n\dot{\Sigma}x'y' - \dot{\Sigma}x'\dot{\Sigma}y'}{\sqrt{n\dot{\Sigma}x'^2 - (\dot{\Sigma}x')^2} \sqrt{n\dot{\Sigma}y'^2 - (\dot{\Sigma}y')^2}} = \\ &= \frac{50 \cdot 34 - 25 \cdot 7}{\sqrt{50 \cdot 65 - (25)^2} \sqrt{50 \cdot 45 - (7)^2}} = 0,635. \end{aligned}$$

Для зависимости между валовой продукцией и числом отработанных человеко-дней (см. табл. 27) коэффициент корреляции, рассчитанный аналогичным путем, оказался равным 0,653.

Таблица 46

Расчет коэффициента корреляции r

		x'				Номера столбцов			
		-1	0	1	2	1	2	3	
y'	y	5	10	15	20	Итого l	ly'	ly'^2	
	x								
2	50			2	4	2	3	6	12
1	40		3	3	6	6	15	15	15
0	30	3	0	8	0	0	20	0	0
-1	20	5	-5	4	-4	-1	10	-10	10
-2	10	2	-4				2	-4	8
Номера строк	1	Итого h	10	15	15	10	50	7	45
	2	hx'	-10	0	15	20	25		
	3	hx'^2	10	0	15	40	65		
	4	$\Sigma_i ty'$	-9	-1	9	8	7		
	5	$x' \Sigma_i ty'$	9	0	9	16	34		

Это значение совпадает, как и следовало ожидать, с величиной η_T , рассчитанной ранее (см. стр. 148), в предположении прямолинейной регрессии.

Коэффициент корреляции в случае криволинейной зависимости

При измерении тесноты связи мы отвечаем на следующий основной вопрос: в какой мере соблюдается строгая функциональная зависимость в изменении переменных y и x ? Ответом служит вычисление корреляционного отношения. Если η или $\eta_T = 1$, то зависимость функциональная; чем ближе η к нулю¹, тем дальше переменные y и x уходят от функциональной зависимости.

¹ При наличии безгранично большого числа наблюдений.

Во многих случаях мы можем, однако, удовлетвориться более узкой постановкой вопроса, а именно мы спрашиваем: *в какой мере соблюдается строгая пропорциональность в изменении переменных y и x ?*

На этот вопрос отвечает вычисление коэффициента корреляции r . Если $r=1$ или $r=-1$, то заключают (при безгранично большом количестве наблюдений) о строгой пропорциональности в изменении y и x ¹. Чем ближе r к нулю, тем дальше переменные y и x от строго пропорционального изменения. В этом смысле коэффициент корреляции называют мерой пропорционального изменения или мерой так называемой *спрямленной зависимости* переменных. Этот свой смысл коэффициент корреляции сохраняет независимо от того, прямолинейна или криволинейна действительная теоретическая линия регрессии y по x .

Однако в первом случае коэффициент корреляции является одновременно и правильным измерителем тесноты корреляционной зависимости. Во втором же случае величина коэффициента корреляции может оказаться *значительно меньше* теоретического

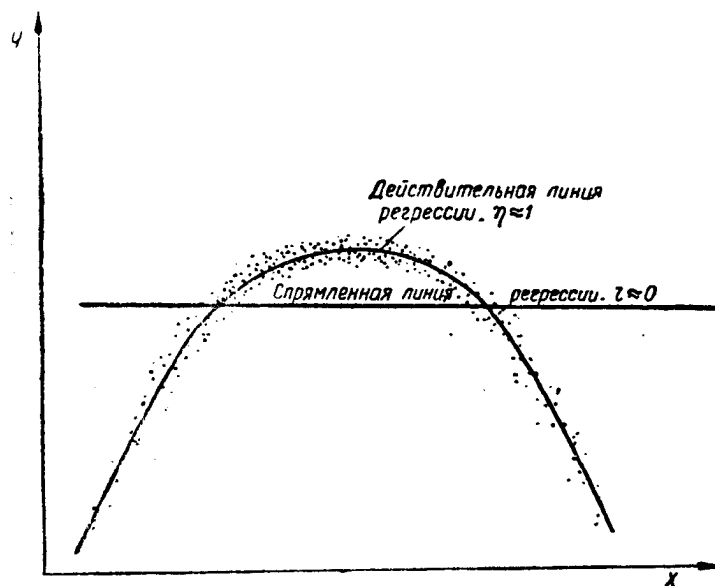


Рис. 42

корреляционного отношения, рассчитанного в предположении криволинейной зависимости, если последняя сильно искажается в результате ее замены прямой линией. Возможны даже случаи, когда при спрямлении криволинейных зависимостей коэффициент

¹ Речь идет, разумеется, не о пропорциональности y и x , а о пропорциональности их изменений (приращений) Δy и Δx .

корреляции оказывается близким к 0, в то время как корреляционное отношение близко к 1 (рис. 42).

Если, однако, теоретическую линию регрессии можно предполагать не слишком уклоняющейся от прямого направления, например, если она имеет монотонно возрастающий или монотонно убывающий характер, можно пользоваться коэффициентом корреляции для предварительной ориентировки в силе зависимости.

Итак, в случае криволинейной зависимости вычисление коэффициента корреляции может преследовать две цели:

- 1) характеризовать степень приближения корреляционной зависимости к функциональной прямолинейной зависимости. Эта цель во многих теоретических вопросах имеет важное значение;
- 2) дать ориентировочное представление о тесноте корреляционной зависимости.

Эта цель имеет серьезное практическое значение, так как вычисление коэффициента корреляции значительно проще, чем вычисление η или η_T . По этой причине измерение связи между переменными обычно начинают с вычисления коэффициента корреляции даже до нахождения линии регрессии.

Вычисление прямой регрессии с помощью таблицы основных показателей

Коэффициент корреляции принадлежит к числу основных статистических показателей.

При изучении материала по одному варьирующему признаку (x) основными показателями являются \bar{x} и σ_x .

При изучении же материала по двум признакам (x) и (y) таблица основных показателей содержит следующие показатели:

Таблица основных показателей

\bar{x}	\bar{y}
σ_x	σ_y
r_{yx}	

Эти показатели наряду с эмпирической линией регрессии рассчитываются в первую очередь, после этого переходят к построению теоретической линии регрессии и к более точному измерению тесноты связи при помощи корреляционных отношений.

Имея таблицу основных показателей, легко рассчитать параметры прямой регрессии y по x

$$\bar{y}_x = a + bx.$$

Это осуществляется с помощью формул

$$b = r \frac{\sigma_y}{\sigma_x}, \quad (7.3.12)$$

$$a = \bar{y} - b\bar{x}. \quad (7.3.13)$$

Формула (7.3.12) вытекает из формулы (7.3.1).

Формула (7.3.13) — из первого нормального уравнения $\Sigma y = na + b\Sigma x$ после деления на n .

Применим эти формулы к расчету прямой регрессии для корреляционной табл. 27.

Основные показатели мы получаем из расчетной табл. 41.

$$\bar{x} = 104,16, \quad \bar{y} = 107,62,$$

$$\sigma_x = 24,265, \quad \sigma_y = 23,128,$$

$$r_{yx} = 0,653.$$

Отсюда

$$b = 0,653 \frac{23,128}{24,265} = 0,622,$$

$$a = 107,62 - 0,622 \cdot 104,16 = 42,79,$$

$$\bar{y}_x = 42,79 + 0,622x,$$

что совпадает с уравнением, вычисленным ранее другим способом.

ГЛАВА VIII

СОПРЯЖЕННЫЕ ПОКАЗАТЕЛИ КОРРЕЛЯЦИИ. СИММЕТРИЧНЫЕ СХЕМЫ ДЕЙСТВИЯ ФАКТОРОВ

§ 1. СОПРЯЖЕННЫЕ ПОКАЗАТЕЛИ КОРРЕЛЯЦИИ

При исследовании связи между двумя переменными мы рассматривали одну из них как функцию, другую — как аргумент. Выбор функции и аргумента определяется целевой установкой исследования. Если изменение одной из величин является причиной изменения другой, то, естественно, принять первую в качестве аргумента, вторую — в качестве функции, ибо таким образом непосредственно можно проследить влияние причины на следствие. Однако нередки случаи, когда обе находящиеся в корреляции переменные могут рассматриваться одновременно и как причины и как следствия по отношению друг к другу. Так, например, известно, что увеличение объема продукции является фактором, способствующим снижению себестоимости изделия (вследствие относительного сокращения так называемых постоянных или слабо меняющихся расходов). Но снижение себестоимости, в свою очередь, является весьма существенной предпосылкой увеличения производственной программы предприятия. Таким образом, корреляция (обратная) между объемом продукции и уровнем себестоимости имеет двустороннюю причинную направленность.

Во многих случаях ни одна из коррелированных переменных не может рассматриваться как причина изменения другой, а корреляция возникает на базе общих причин, совместно влияющих на обе переменные.

Так, например, показатели твердости стали по Бринеллю тесно связаны с результатами испытаний по Роквеллу. Однако между теми и другими не существует причинной зависимости, а корреляция возникает на базе свойств материала, одновременно проявляющихся в тех и других испытаниях. Точно так же корреляция, наблюдающаяся между ростом братьев, не вытекает из

причинной зависимости между этими признаками, а является выражением наличия общих факторов наследственности и среды, одновременно действующих на рост братьев. Или корреляция между ценами на двух колхозных рынках может в некоторой степени объясняться непосредственным взаимным влиянием этих рынков, но в основном порождается общими экономическими условиями, определяющими рыночные цены.

Далее заметим, что в процессах, направляемых к определенной цели, интересна такая постановка вопроса, при которой следствие является аргументом, а причина — функцией. Так, например, между выпуском продукции и количеством потребленного сырья на предприятии имеется прямая корреляционная зависимость. Сырье является при этом материальной причиной возникновения продукции. Однако нетрудно понять, что экономически не количество сырья определяет выпуск продукции, а, наоборот, производственная программа по продукции определяет потребность в сырье. Поэтому здесь интересна постановка вопроса, при которой объем продукции служит аргументом, а количество сырья — функцией.

Мы видим, таким образом, что во многих случаях коррелированные переменные равноправны с точки зрения их выбора в качестве аргумента или функции. Поэтому, наряду с вычислением уравнений регрессии, корреляционных отношений и других показателей, характеризующих форму и тесноту связи y по x , представляет также интерес и вычисление аналогичных показателей, характеризующих форму и тесноту связи x по y . Такие показатели мы называем *взаимно-сопряженными*. На примере корреляционной табл. 26 рассчитаем сопряженные эмпирические и теоретические линии регрессии (прямые и параболы второго порядка).

В табл. 47 производим расчет ординат эмпирических линий регрессии и находим все суммы, необходимые для составления систем нормальных уравнений и расчета параметров.

$$\begin{aligned} 1. \bar{y}_{x'} &= a' + b'x', \\ 3 &= 200a' + 98b', \\ 258 &= 98a' + 664b'. \end{aligned}$$

Отсюда $a' = -0,186$, $b' = 0,416$.

По формулам (5.2.1) или из формулы

$$\frac{\bar{y}_x - 1410}{50} = a' + b' \cdot \frac{x - 17,7}{0,6}$$

получаем

$$\bar{y}_x = 787,1 + 34,666x. \quad (I)$$

$$2. \bar{y}_x = a' + b'x' + c'x'^2,$$

$$3 = 200a' + 98b' + 664c',$$

$$258 = 98a' + 664b' + 1052c',$$

$$80 = 664a' + 1052b' + 6508c'.$$

Отсюда

$$a' = -0,00394, \quad b' = 0,4961, \quad c' = 0,0675.$$

Из формулы

$$\frac{\bar{y}_x - 1410}{50} = a' + b' \frac{x - 17,7}{0,6} + c' \frac{x - 17,7^2}{0,6}$$

получаем

$$\bar{y}_x = -2258,645 + 373,217x - 9,375x^2. \quad (II)$$

$$3. \bar{x}_y = a' + b'y',$$

$$98 = 200a' + 3b',$$

$$258 = 3a' + 553b'.$$

Отсюда

$$a' = 0,481, \quad b' = 0,464.$$

По формулам (5. 2. 1) или из формулы

$$\frac{\bar{x}_y - 17,7}{0,6} = a' + b' \cdot \frac{y - 1410}{50}$$

находим:

$$\bar{x}_y = 10,15 + 0,005586y. \quad (Ia)$$

$$4. \bar{x}_y = a' + b'y' + c'y'^2,$$

$$98 = 200a' + 3b' + 553c',$$

$$258 = 3a' + 553b' - 243c',$$

$$366 = 553a' - 243b' + 5365c'.$$

Отсюда

$$a' = 0,32527, \quad b' = 0,4898, \quad c' = 0,05687.$$

Из уравнения

$$\frac{\bar{x}_y - 17,7}{0,6} = a' + b' \frac{y - 1410}{50} + c' \cdot \left| \frac{y - 1410}{50} \right|^2$$

получаем

$$\bar{x}_y = 36,74 - 0,0326y + 0,0001364y^2. \quad (II a)$$

Ординаты теоретических линий указаны в табл. 47. По этим ординатам строим график сопряженных линий регрессии (рис. 43 а, б).

Обратимся к анализу этих линий.

Прежде всего заметим, что уже самый факт существования различающихся между собой сопряженных линий регрессии является особенностью корреляционной зависимости. Графическое изображение функциональной зависимости не меняется от того, относительно которой из двух переменных решено уравнение, т. е. какая переменная считается аргументом и какая — функцией. Сопряженные уравнения получают здесь одно из другого тождественным преобразованием.

Между тем для корреляционной зависимости выбор аргумента и функции меняет график регрессии: сопряженные уравнения не могут быть получены одно из другого тождественным преобразованием. Они существенно различны или, как говорят, *необратимы*. Если мы сопоставим приращения абсциссы и ординаты на сопряженных линиях регрессии, то увидим, что y изменяется по отношению к x значительно быстрее на второй линии (x по y), чем на первой (y по x) (см. рис. 43).

Таким образом, мы видим, что изменение признака *замедляется*, когда он выступает в корреляционной зависимости на правах функции по сравнению с изменением его на правах аргумента. Это явление может быть названо *законом регрессии*.

В законе регрессии получает объяснение термин: «*линии и уравнения регрессии*»¹.

Сущность этого своеобразного закона уясняется из следующих соображений. Следует, во-первых, иметь в виду, что сопряженные линии регрессии различаются между собой по смыслу.

С внешней стороны очевидно, что переменные x и y неравноправны в уравнениях сопряженных линий. В первой линии (y по x) x является индивидуальным значением признака, в то время как y выступает в виде соответствующей ему групповой средней (y_x).

Во второй линии (x по y) y является индивидуальным значением признака, а x — групповой средней (x_y).

Расхождение между сопряженными линиями регрессии оказывается, с формальной стороны, возможным потому, что их уравнения связывают различные переменные: первое уравнение — y_x и x , второе — x_y и y . Однако

¹ Закон регрессии был впервые обнаружен английскими биометриками при изучении методами корреляции вопросов наследственности. Однако биометрики дали этому закону неправильное и случайное истолкование. В связи с этим, предложенный ими и вошедший во всеобщее употребление термин «уравнение регрессии» признавался неудачным. В нашей литературе были предложения о замене этого термина другими, как, например: «уравнение прогрессии», «корреляционное уравнение» и др. Мы полагаем, что при правильном понимании закона регрессии нет необходимости менять термин.

РАСЧЕТ СОПРЯЖЕННЫХ ПОКАЗАТЕЛЕЙ

	x'	-4	-3	-2	-1	0	1	2	3	4	5	6
y'	x	15,3	15,9	16,5	17,1	17,7	18,3	18,9	19,5	20,1	20,7	21,3
	y											
5	1660								3) (5) 1			
4	1610											6) (4) 1
3	1560			-2) (3) 1			1) (3) 1	6) (9) 3	6) (6) 2	4) (3) 1		
2	1510					6) (3) 3	8) (16) 8	16) (16) 8	15) (10) 5	8) (4) 2		
1	1460			-4) (2) 2	-6) (6) 6	13) (13) 13	17) (17) 17	6) (3) 3	6) (2) 2			
0	1410			-4) (2) 2	-10) (10) 10	13	7) (7) 7	26) (26) 13	9) (9) 3	4) (4) 1		
-1	1360			-20) (20) 10	-7) (7) 7	-7) (7) 7	-12) (12) 12	4) (4) 4	-4) (4) 4	8) (8) 4	-4) (4) 1	-1) (1) 1
-2	1310			-6) (6) 2	-4) (4) 5	-10) (10) 5	-1) (1) 1	-2) (2) 1	3) (3) 3	-6) (6) 4	-8) (8) 2	-2) (2) 1
-3	1260			-9) (9) 3	-9) (9) 4	-4) (4) 2	-6) (6) 2	-6) (6) 1	-3) (3) 1	3) (3) 1	-3) (3) 1	
-4	1210	-4) (4) 1	-4) (4) 1	-3) (3) 1	-4) (4) 1	-1) (1) 1	-4) (4) 1					
-5	1160											
-6	1110								3) (3) 1	-6) (6) 1		
Итого h		1	6	22	27	46	41	32	17	6	1	1
$\Sigma y'y'$		-4	-17	-21	-13	-6	24	22	11	4	-1	4
$\overline{y'_i}$		-4	-2,83	-0,96	-0,48	-0,13	0,59	0,69	0,65	0,67	-1	4
$x' \Sigma_i m y'$		16	51	42	13	0	24	44	33	16	-5	24
$x'^2 \Sigma_i m y'$		-64	-153	-84	-13	0	24	88	99	64	-25	144
$h x'$		-4	-18	-44	-27	0	41	64	51	24	5	6
$h x'^2$		16	54	88	27	0	41	128	153	96	25	36
$h x'^3$		-64	-162	-176	-27	0	41	256	459	384	125	216
$h x'^4$		256	486	352	27	0	41	512	1377	1536	625	1296
$\overline{y_i}$		1210	1268,5	1362,0	1386,0	1403,5	1439,5	1444,5	1442,5	1443,5	1360	1610
$\overline{y_x^I}$		1317,49	1338,29	1359,09	1379,89	1400,69	1421,49	1442,29	1463,08	1483,88	1504,69	1525,48
$\overline{y_x^{II}}$		1239,35	1305,41	1347,09	1382,02	1410,20	1431,63	1446,31	1454,24	1455,42	1454,14	1437,53

Таблица 47

УПРУГОСТЬ — РАЗМЕР ЗАМКА

Итого l	$\Sigma_j m x'$	$\frac{1}{x_j}$	$y' \Sigma_j m x'$	$y'^2 \Sigma_j m x'$	ly'	ly'^2	ly'^3	ly'^4	\bar{x}_j	$\frac{1}{x_y}$	$\frac{11}{x_y}$
1	3	3	15	75	5	25	125	625	19,50	19,40	20,20
1	6	6	24	96	4	16	64	256	21,30	19,12	19,60
8	15	1,87	45	135	24	72	216	648	18,82	18,85	19,06
26	47	1,80	94	188	52	104	208	416	18,78	18,57	18,60
43	19	0,44	19	19	43	43	43	43	17,96	18,29	18,21
49	32	0,65	0	0	0	0	0	0	18,09	18,01	17,88
40	-3	-0,08	3	-3	-40	40	-40	40	17,65	17,73	17,62
18	-4	-0,22	8	-16	-36	72	-144	288	17,57	17,45	17,43
9	-12	-1,33	36	-108	-27	81	-243	729	16,90	17,17	17,32
4	-8	-2	32	-128	-16	64	-256	1024	16,50	16,89	17,26
0	0	0	0	0	0	0	0	0	17,70	16,61	17,27
1	3	3	-18	108	-6	36	-216	1296	19,50	16,33	17,35
200 = n	98 =		258 =	366 =	3 =	553 =	-243 =	5365 =			
$3 = \Sigma y'$	$= \Sigma x'$		$= \Sigma x' y'$	$= \Sigma x' y'^2$	$= \Sigma y'$	$= \Sigma y'^2$	$= \Sigma y'^3$	$= \Sigma y'^4$			
Примечание. В левых верхних углах клеток таблицы стоят произведения частоты клетки на соответствующие x' . В правых верхних углах — произведения частоты клетки на y' .											
258 = $\Sigma x' y'$											
80 = $\Sigma x'^2 y'$											
98 = $\Sigma x'$											
664 = $\Sigma x'^2$											
1052 = $\Sigma x'^3$											
6508 = $\Sigma x'^4$											

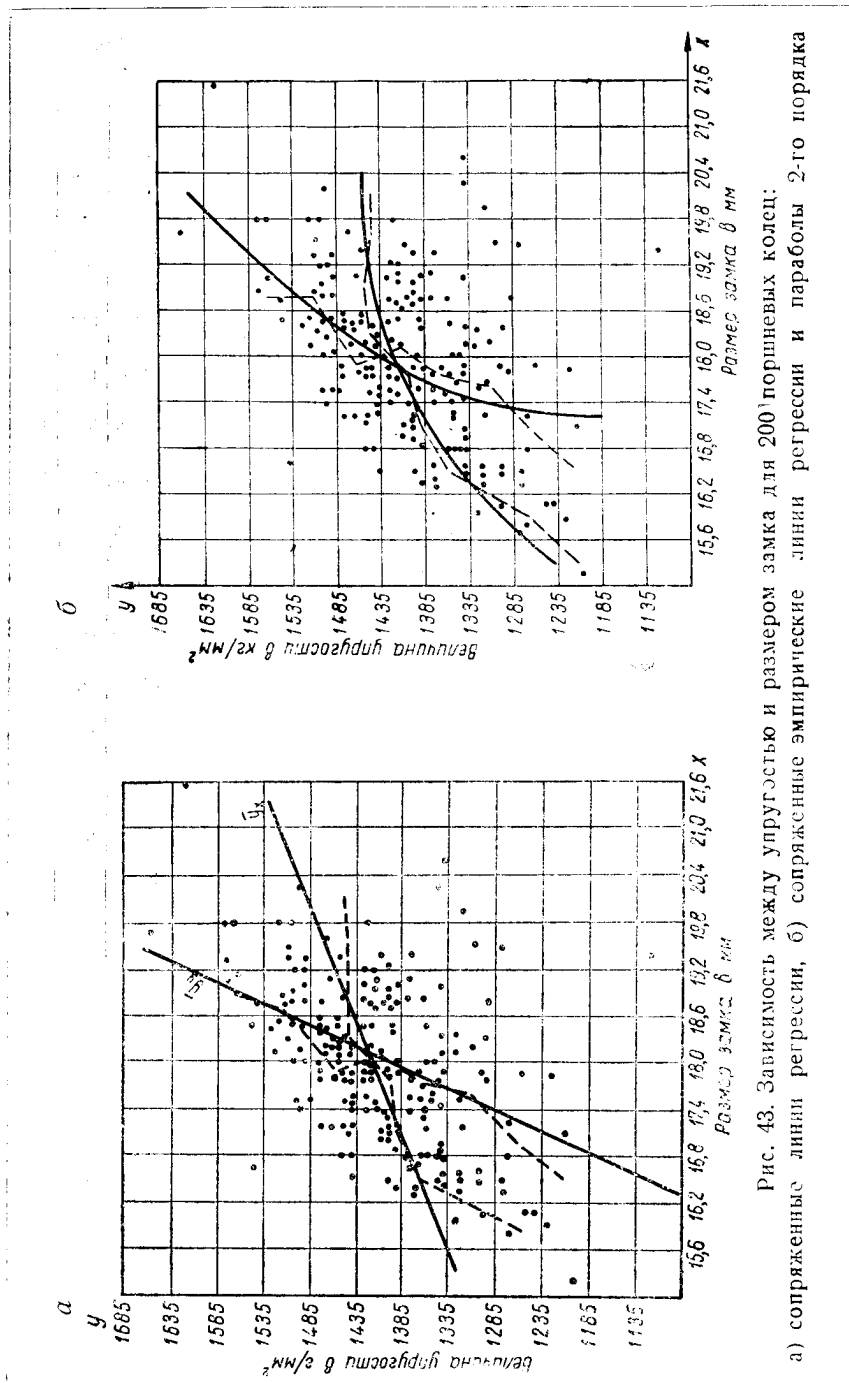


Рис. 43. Зависимость между упругостью и размером замка для 200³ поршневых колец:
а) сопряженные линии регрессии, б) сопряженные эмпирические линии регрессии 2-го порядка

подобное формальное объяснение не раскрывает внутренних причин расхождения.

По существу же *оба уравнения* регрессии связывают групповые *средние значения* x и y , но первое уравнение получено на базе *группировки материала по признаку* x . По этой причине в первом уравнении групповые средние по признаку x совпадают с индивидуальными значениями x . Второе уравнение получено на базе группировки материала по признаку y . Поэтому групповые средние по признаку y здесь совпадают с индивидуальными значениями y .

Перемена признака группировки ведет к изменению соотношения между групповыми средними.

Для лучшего уяснения этого важного положения заметим, что метод регрессии является частным случаем более общего приема, известного под названием метода *параллельных рядов*. Он состоит в следующем:

- а) производится группировка материала по некоторому признаку z ;
- б) в каждой группе рассчитываются средние значения по x и по y ;
- в) ряды этих средних сопоставляются между собой для уяснения зависимости между y и x .

Так, например, можно получить зависимость между ростом и весом школьников, если: а) разгруппировать материал по возрасту, б) в каждой возрастной группе рассчитать средний рост и средний вес, в) сопоставить между собой два ряда средних, следя за параллельным изменением этих показателей.

При достаточно большом числе наблюдений можно считать групповые средние по x и по y точными функциями z :

$$\left. \begin{aligned} \bar{x}_z &= \varphi(z) \\ \bar{y}_z &= \psi(z) \end{aligned} \right\} \quad (8.1.1)$$

Уравнения (8.1.1) являются уравнениями регрессии x по z и y по z . В них выражается зависимость между \bar{y}_z и \bar{x}_z в так называемой параметрической форме.

Представляя эту зависимость в явной форме (там, где это возможно), получаем

$$\bar{y}_z = f(\bar{x}_z) \quad (8.1.2)$$

или

$$\bar{x}_z = f^{-1}(\bar{y}_z), \quad (8.1.2')$$

f^{-1} —функция, обратная f , следовательно, уравнения (8.1.2) и (8.1.2') *обратимы*, т. е. получаются одно из другого тождественным преобразованием.

Так, например, из двух уравнений регрессии — роста по возрасту и веса по возрасту — можно получить зависимость среднего роста от среднего веса и обратно — среднего веса от среднего роста.

Очевидно, что если полагать в основу группировки различные признаки z , то зависимость между y_z и x_z может существенно изменяться.

В самом деле, уравнения (8.1.1) в схеме (А—М—Н) показывают, как изменялись бы признаки x и y в связи с изменением аргумента z , если бы прочие независимо от z действующие факторы были закреплены.

Следовательно, уравнения (8.1.2) и (8.1.2') показывают, какова была бы зависимость между y и x , если бы она складывалась под исключительным влиянием изменения фактора z (если бы прочие независимо от z действующие факторы не изменялись). Но поскольку соотношение изменений x и y не постоянно, а зависит, вообще говоря, от природы признака z , под влиянием которого эти изменения происходят, то ясно, что зависимость между y_z и x_z может оказаться иной при разных z .

Так, зависимость между ростом и весом, если бы она сложилась исключительно в связи с изменением возраста, выглядела бы иначе по сравнению с тем, как если бы эта зависимость сложилась при постоянном возрасте в связи с изменением других факторов, например режима питания (изменение возраста может сильнее отразиться на росте, режим питания — на весе).

Подтвердим эти соображения математическим расчетом. Зависимость между средними \bar{y}_z и \bar{x}_z в спрямленном виде может быть выражена формулой $\bar{y}_z - \bar{y} = k(\bar{x}_z - \bar{x})$ или тождественной формулой $(\bar{x}_z - \bar{x}) = k_1(\bar{y}_z - \bar{y})$, где $k_1 = \frac{1}{k}$.

Коэффициенты k и k_1 получаются из уравнений прямых регрессии y по z и x по z путем деления одного уравнения на другое

$$\bar{y}_z - \bar{y} = r_{yz} \frac{\sigma_y}{\sigma_z} (z - \bar{z}), \quad (8.1.3)$$

$$\bar{x}_z - \bar{x} = r_{xz} \frac{\sigma_x}{\sigma_z} (z - \bar{z}). \quad (8.1.4)$$

Отсюда

$$\text{и} \quad \left. \begin{aligned} k &= \frac{\bar{y}_z - \bar{y}}{\bar{x}_z - \bar{x}} = \frac{r_{yz} \frac{\sigma_y}{\sigma_z}}{r_{xz} \frac{\sigma_x}{\sigma_z}} \\ k_1 &= \frac{1}{k} = \frac{r_{xz} \frac{\sigma_x}{\sigma_z}}{r_{yz} \frac{\sigma_y}{\sigma_z}} \end{aligned} \right\} \quad (8.1.5)$$

Из последних выражений можно сделать вывод, что чем больше теснота спрямленной зависимости признака y от признака группировки z и чем слабее теснота спрямленной зависимости x от z , тем больше коэффициент k и меньше k_1 .

Уравнения регрессии y по x или x по y можно рассматривать как соотношения между x_z и y_z для того частного случая, когда группировка материала производится по признаку x либо по признаку y . И так как при перемене признака группировки соотношение между x_z и y_z видоизменяется, то отсюда понятно расхождение между сопряженными линиями регрессии.

В первом случае абсолютная величина коэффициента корреляции между x и z достигает максимальной величины ($|r_{xz}| = 1$) и заведомо больше абсолютной величины коэффициента корреляции между y и z , во втором случае — наоборот¹.

Заметим, что понятию аргумента так же, как и понятию функции, можно придать в статистическом исследовании двойкий смысл. В первом смысле слова, мы называем, как это обычно принято в математике, аргументом ту из переменных, значения которой нами задаются, т. е. рассматриваются в любом по желанию порядке.

Функцией же мы называем переменную, значения которой рассматриваются так, как это определено порядком рассмотрения значений аргу-

¹ В случае если z единственный фактор, обуславливающий корреляцию между x и y (см. след. параграф), соотношение (8.1.5) превращается в уравнение так называемой единой линии связи, предложенное А. А. Конюсом (см. статью в сборнике «Вопросы статистического измерения связей между явлениями», Госпланиздат, 1950).

мента. Аргумент и функцию, в этом смысле слова, будем называть *формальными*.

Во втором смысле слова, аргументом является тот признак, по которому производится группировка материала. В отличие от формального аргумента этот признак будем называть *статистическим аргументом*. Иначе его называют *статистическим подлежащим*. *Статистической функцией* будем называть признак, по которому вычисляются средние показатели¹ для каждой группы подлежащего. Эти средние носят название *статистического сказуемого*.

Понятия статистического и формального аргумента (или функции) не совпадают. В одном и том же исследовании признак одновременно может быть формальным аргументом и статистической функцией и наоборот — формальной функцией и статистическим аргументом. В корреляционном исследовании признак x будет формальным аргументом, если мы по средним значениям x_z определяем средние значения y_z . Он будет статистическим аргументом лишь в том случае, когда по нему производится группировка, т. е. каждая средняя x_z вычисляется из *одинаковых смежных значений x* .

Рассмотрим пример. В процессе изготовления поршневых колец упругие свойства кольца регулируют путем изменения размера замка. Поэтому решают вопрос о размерах замка, ведущих к желаемым значениям величины упругости. При такой постановке вопроса упругость (y) является формальным аргументом, а размеры замка (x) — формальной функцией. Поэтому следует отправляться от уравнения

$$\bar{x}_z = f(\bar{y}_z). \quad (8.1.6)$$

По какому же признаку (z) для этого уравнения следует производить группировку?

Практически мы непосредственно распоряжаемся размерами замка. Следовательно, партию колец мы можем группировать или комплектовать только по размеру замка, отбирая нужные нам размеры. Отсюда следует, что в уравнении $x_z = f_1(\bar{y}_z)$ статистическим аргументом должен быть *размер замка x* .

Итак, оценка требуемого размера замка должна производиться по *первому сопряженному уравнению регрессии*

$$\bar{y}_x = f(x),$$

решая его относительно x (в нашем случае по уравнению $\bar{y}_x = 787,1 + 34,666x$ или $\bar{y}_x = -2258,645 + 373,217x - 9,375x^2$, решенному относительно x).

По этому уравнению мы находим *размер замка (x)*, ведущий к желаемому среднему значению упругости (\bar{y}_x).

Использовать для аналогичной цели сопряженные уравнения

$$\bar{x}_y = 10,15 + 0,005586y$$

или

$$\bar{x}_y = 36,74 - 0,03264y + 0,00001364y^2$$

мы не можем, так как практически неосуществима группировка материала по упругости в условиях, когда последняя непосредственно не определяется, а регулируется размером замка.

¹ Или другие обобщающие показатели.

Расхождение между формальным и статистическим аргументом мы получим и в примере зависимости между объемом продукции и количеством потребного сырья и в других случаях практического использования уравнения связи в целях воздействия на течение функции.

Обратимся к рассмотрению сопряженных показателей тесноты связи. Очевидно, что коэффициент корреляции r между признаками x и y не зависит от того, на месте ли аргумента или на месте функции оказалась та или иная переменная. Это вытекает из формул (7.3.5), (7.3.6), (7.3.7) для коэффициента корреляции, которые вполне симметричны относительно переменных x и y .

Итак, $r_{yx} = r_{xy} = r$, т. е. сопряженные коэффициенты корреляции равны. Ранее мы определили коэффициент корреляции между y и x как показатель, отвечающий на вопрос: на сколько сигм в среднем изменяется y при увеличении x на одну сигму? Сопряженный коэффициент корреляции отвечает на вопрос: на сколько сигм в среднем изменяется x при увеличении y на одну сигму? Из равенства сопряженных коэффициентов корреляции следует, что наряду с формулой (7.3.1)

$$r = b \frac{\sigma_x}{\sigma_y}$$

справедлива также и формула

$$r = b_1 \frac{\sigma_y}{\sigma_x}, \quad (7.3.1')$$

где b_1 — коэффициент регрессии x по y .

Неравенство $|r| \leq 1$ показывает, что *среднее изменение функции меньше изменения аргумента*, если оба признака выражены в сравнимых единицах (сигмах).

В этом другая форма выражения закона *регрессии*.

Из формул (7.3.1) и (7.3.1') следует, что

$$r^2 = bb_1 \quad (8.1.7)$$

или *произведение сопряженных коэффициентов регрессии равно квадрату коэффициента корреляции*.

В случае, когда $r = \pm 1$, т. е. когда имеет место точная пропорциональность изменения x и y , мы получаем из формулы (8.1.7) $b = \frac{1}{b_1}$. Это значит, что при $r = \pm 1$ сопряженные прямые регрессии совпадают. Чем меньше абсолютная величина коэффициента корреляции, тем больше расхождение между сопряженными прямыми регрессии, тем в более резкой форме проявляется закон регрессии. При $r = 0$ сопряженные прямые регрессии взаимно-перпендикулярны, так как из формулы (7.3.1) и (7.3.1') следует, что в этом случае b и b_1 равны нулю (рис. 44).

Заметим также, что сопряженные коэффициенты регрессии

b и b_1 имеют знак, совпадающий со знаком при r [как это следует из формул (7.3.1) и (7.3.1')]. Поэтому b и b_1 всегда имеют одинаковые знаки: они либо оба положительны, либо оба отрицательны.

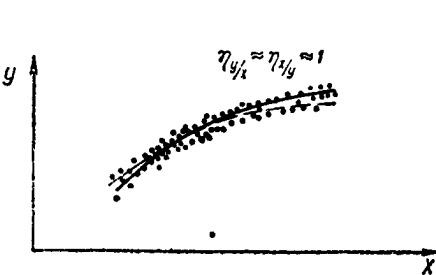
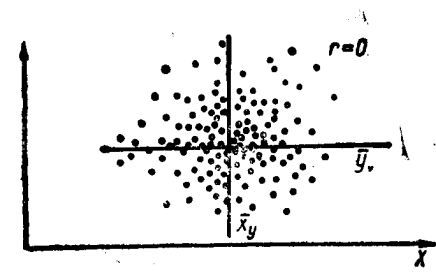
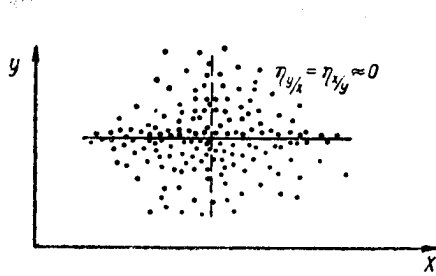
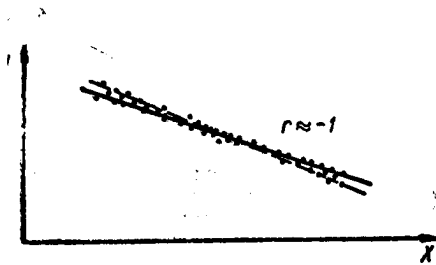
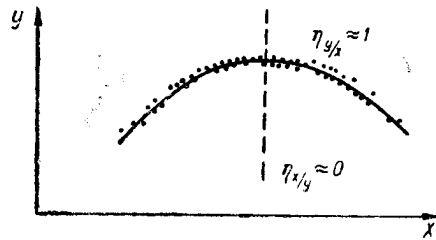
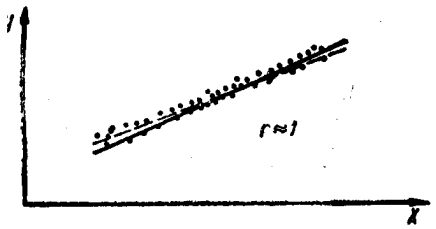


Рис. 44

Рис. 45

Если коэффициент корреляции не зависит от того, какая из переменных стоит на месте функции и какая на месте аргумента, то корреляционные отношения y по x и x по y (теоретические и эмпирические) являются в случае криволинейной зависимости различными показателями чаще всего $\eta_{y/x}$ и $\eta_{x/y}$ близки друг к другу; однако встречаются случаи и большого расхождения между ними. Не исключена даже возможность близости к 1 одного из сопряженных корреляционных отношений и одновременно близости к 0 — другого. Подобный случай изображен на рис. 45.

Необходимо, следовательно, различать *тесноту связи y по x и x по y* .

При $\eta_{y/x} = \eta_{x/y} = 0$ сопряженные линии регрессии превращаются во взаимно-перпендикулярные прямые. При $\eta_{y/x} = \eta_{x/y} = 1$ имеет место совпадение обеих линий регрессий.

Приводим расчет сопряженных корреляционных отношений, эмпирических и теоретических, для зависимости между упругостью и размером замка поршневых колец (см. табл. 48).

Сопряженные корреляционные отношения оказались близки между собой

$$\eta_{y/x} = 0,517, \quad \eta_{x/y} = 0,557,$$

Сопряженные теоретические корреляционные отношения $\eta_{T_{y/x}}$ и $\eta_{T_{x/y}}$, рассчитанные в предположении прямолинейной зависимости, должны быть равны коэффициенту корреляции $r = 0,440$. Сопряженные теоретические корреляционные отношения, рассчитанные в предположении криволинейной зависимости, также оказались близкими по величине

$$\eta_{T_{y/x}} = 0,475, \quad \eta_{T_{x/y}} = 0,471.$$

Заметим, что выравнивание сопряженных линий регрессии по параболам второго порядка в нашем примере производилось на том основании, что эти кривые на данном ограниченном участке изменения аргументов приблизительно воспроизводят ход изменения функций. Во всей же плоскости (x, y) сопряженные линии регрессии, как показано О. В. Сармановым¹, не могут одновременно выражаться многочленами степени выше, чем 1-й, так как порядки роста сопряженных линий регрессии должны быть связаны между собой так же, как порядки роста взаимно обратных функций $y = \varphi(x)$ и $x = \varphi^{-1}(y)$.

§ 2. КОЭФФИЦИЕНТ КОРРЕЛЯЦИИ И СИММЕТРИЧНЫЕ СХЕМЫ ДЕЙСТВИЯ ФАКТОРОВ

В общем случае корреляция между двумя переменными x и y возникает в силу наличия факторов, одновременно действующих на обе переменные (общих факторов).

Пусть

$$\left. \begin{aligned} x &= \varphi(u, v) \\ y &= \psi(u, w) \end{aligned} \right\}, \quad (8.2.1)$$

т. е. x и y являются некоторыми функциями от факторов u, v, w .

Фактор u является общим, факторы v, w — специфическими. Система (8.2.1) является частным случаем более общей системы

$$\left. \begin{aligned} x &= \varphi(u_1, u_2, \dots, u_s, v_1, v_2, \dots, v_k) \\ y &= \psi(u_1, u_2, \dots, u_s, w_1, w_2, \dots, w_l) \end{aligned} \right\}, \quad (8.2.2)$$

¹ О. В. Сарманов. О прямолинейной корреляции. Диссертация, 1948. О. В. Сарманов. О порядке роста линий регрессии. Труды второго Всесоюзного совещания по математической статистике, Ташкент, 1949.

Таблица 48а

Упругость — размер замка

Расчет корреляционных отношений $r_{y/x}$ и $r_{T/y/x}$

x	h	\bar{y}_i	$(\bar{y}_i - \bar{y})$	$(\bar{y}_i - \bar{y})^2$	$(\bar{y}_i - \bar{y})^2 h$	\bar{y}_x	$\bar{y}_x - \bar{y}$	$(\bar{y}_x - \bar{y})^2$	$(\bar{y}_x - \bar{y})^2 h$
21,3	1	1610,0	199,25	39700,6	39700,6	1437,53	26,78	717,2	717,2
20,7	1	1360,0	-50,75	2575,6	2575,6	1454,14	43,39	1882,69	1882,69
20,1	6	1443,5	32,75	1072,6	6435,6	1455,42	44,67	1995,4	11972,4
19,5	17	1442,5	31,75	1008,1	17137,0	1454,24	43,49	1891,4	32153,8
18,9	32	1444,5	33,75	1139,1	36451,2	1446,31	35,56	1264,5	40464,0
18,3	41	1439,5	28,75	826,6	33890,6	1431,63	20,88	435,97	17874,7
17,7	46	1403,5	-7,25	52,6	2419,6	1410,20	-0,55	0,3	13,8
17,1	27	1386,0	-24,75	612,6	16540,2	1382,02	-28,73	825,4	22285,8
16,5	22	1362,0	-48,75	2376,6	52285,2	1347,09	-63,66	4052,6	89157,2
15,9	6	1268,5	-142,25	20235,0	121410,0	1305,41	-105,34	11096,5	66579,0
15,3	1	1210,0	-200,75	40300,6	40300,6	1239,35	-171,40	29377,96	29377,96
200					369146,2	312478,55			

$$\bar{y} = 1410,75,$$

$$\sigma_y^2 = 6913,92,$$

$$\bar{\delta}_i^2 = \frac{369146,2}{200} = 1845,73,$$

$$r_{y/x}^2 = \frac{1845,73}{6913,92} = 0,2670,$$

$$r_{y/x} = 0,517,$$

$$\bar{y}_x = -2258,6 + 373,22x - 9,375x^2.$$

$$\bar{\delta}_{iT}^2 = \frac{312478,55}{200} = 1562,39,$$

$$r_{T/y/x}^2 = \frac{1562,39}{6913,92} = 0,2260,$$

$$r_{T/y/x} = 0,475,$$

Упругость — размер закла
 Расчет корреляционных отношений $r_{x/y}$ и $r_{T_{x/y}}$

y	l	\bar{x}_j	$\bar{x}_j - \bar{x}$	$(\bar{x}_j - \bar{x})^2$	$(\bar{x}_j - \bar{x})^2 l$	\bar{x}_y	$\bar{x}_y - \bar{x}$	$(\bar{x}_y - \bar{x})^2$	$(\bar{x}_y - \bar{x})^2 l$
1660	1	19,50	1,51	2,280	2,280	20,20	2,21	4,884	4,884
1610	1	21,30	3,31	10,956	10,956	19,60	1,61	2,592	2,592
1560	8	18,82	0,83	0,689	6,512	19,06	1,07	1,145	9,160
1510	26	18,78	0,79	0,624	16,224	18,60	0,61	0,372	9,672
1460	43	17,96	-0,03	0,0009	0,039	18,21	0,22	0,048	2,064
1410	49	18,09	0,10	0,010	0,490	17,88	-0,11	0,012	0,588
1360	40	17,65	-0,34	0,116	4,640	17,62	-0,37	0,137	5,480
1310	18	17,57	-0,42	0,176	3,168	17,43	-0,56	0,314	5,652
1260	9	16,90	-1,09	1,188	10,692	17,32	-0,67	0,449	4,041
1210	4	16,50	-1,49	2,220	8,880	17,26	-0,73	0,533	2,132
1160	0	17,70	-0,29	0,084	0,000	17,27	-0,72	0,518	0,000
1110	1	19,50	1,51	2,280	2,280	17,35	-0,64	0,410	0,410
200					65,161				46,675

$$\bar{x} = 17,99,$$

$$\sigma_x^2 = 1,11,$$

$$\bar{\delta}_j^2 = \frac{65,161}{200} = 0,326,$$

$$r_{x/y}^2 = \frac{0,326}{1,11} = 0,2937,$$

$$r_{x/y} = 0,542,$$

$$\bar{x}_y = 36,74 - 0,0326y + 0,00001364y^2,$$

$$\bar{\delta}_{JT}^2 = \frac{46,675}{200} = 0,233,$$

$$r_{T_{x/y}}^2 = \frac{0,233}{1,11} = 0,2099,$$

$$r_{T_{x/y}} = 0,458,$$

в которой переменные x и y представлены в виде функций от ряда факторов, часть из которых (u_1, u_2, \dots, u_s) являются *общими*, другие — *специфическими*. Такого рода системы будем называть *симметричными* схемами действия факторов.

Покажем на простейшей симметричной схеме, как при определенных условиях между x и y возникает корреляция, обусловленная наличием общих факторов.

1) Пусть

$$\left. \begin{aligned} x &= u + v \\ y &= u + w \end{aligned} \right\} \quad (8.2.3)$$

Каждая из переменных x и y складывается в виде суммы двух факторов, один из которых является общим. Факторы u, v, w, \dots будем считать линейно независимыми, т. е. коэффициенты корреляции между ними при неограниченно большом числе наблюдений равны нулю¹. Все дальнейшие расчеты будем производить в предположении, что число наблюдений достаточно велико и коэффициенты корреляции между факторами можно приближенно приравнять нулю.

Схему (8.2.3), представляющую, очевидно, частный случай схемы (8.2.1), назовем симметричной схемой аддитивного и независимого действия факторов или, короче, схемой (А—Н)с.

Для простоты исследования положим, что все факторы имеют одинаковую дисперсию, т. е. $\sigma_u = \sigma_v = \sigma_w$. Это ограничение в дальнейшем будет снято.

Рассчитаем коэффициент корреляции между x и y

$$r_{xy} = \frac{\frac{1}{n} \Sigma (x - \bar{x})(y - \bar{y})}{\sigma_x \sigma_y}.$$

Но

$$\bar{x} = \frac{\Sigma x}{n} = \frac{\Sigma (u + v)}{n} = \frac{\Sigma u}{n} + \frac{\Sigma v}{n} = \bar{u} + \bar{v}$$

и по аналогии $\bar{y} = \bar{u} + \bar{w}$. Отсюда

$$\begin{aligned} \frac{1}{n} \Sigma (x - \bar{x})(y - \bar{y}) &= \frac{1}{n} \Sigma [(u - \bar{u}) + (v - \bar{v})] [(u - \bar{u}) + \\ &+ (w - \bar{w})] = \frac{1}{n} \Sigma (u - \bar{u})^2 + \frac{1}{n} \Sigma (u - \bar{u})(v - \bar{v}) + \\ &+ \frac{1}{n} \Sigma (u - \bar{u})(w - \bar{w}) + \frac{1}{n} \Sigma (v - \bar{v})(w - \bar{w}) = \sigma_u^2 + \\ &+ r_{uv} \sigma_u \sigma_v + r_{uw} \sigma_u \sigma_w + r_{vw} \sigma_v \sigma_w = \sigma_u^2, \end{aligned}$$

¹ Мы не требуем здесь полной корреляционной независимости факторов, а лишь их линейной независимости. Линейно (точнее, прямолинейно) независимые переменные не являются обязательно корреляционно независимыми, так как при $r = 0$ может отличаться от нуля. Однако корреляционная независимость переменных влечет за собой и их линейную независимость, так как при $\eta = 0$ r также равен 0.

так как

$$r_{uv} = r_{uw} = r_{vw} = 0$$

в силу линейной независимости факторов.

Далее

$$\begin{aligned} \sigma_x &= \sqrt{\frac{1}{n} \Sigma (x - \bar{x})^2} = \sqrt{\frac{1}{n} \Sigma [(u - \bar{u}) + (v - \bar{v})]^2} = \\ &= \sqrt{\frac{1}{n} \Sigma (u - \bar{u})^2 + \frac{1}{n} \Sigma (v - \bar{v})^2 + 2 \frac{1}{n} \Sigma (u - \bar{u})(v - \bar{v})} = \\ &= \sqrt{\sigma_u^2 + \sigma_v^2}. \end{aligned}$$

Аналогично

$$\sigma_y = \sqrt{\sigma_u^2 + \sigma_w^2}.$$

Отсюда

$$r_{xy} = \frac{\sigma_u^2}{\sqrt{\sigma_u^2 + \sigma_v^2} \sqrt{\sigma_u^2 + \sigma_w^2}}. \quad (8.2.4)$$

Но так как согласно условию $\sigma_u = \sigma_v = \sigma_w$,

$$\text{то} \quad r_{xy} = \frac{1}{\sqrt{2} \sqrt{2}} = \frac{1}{2}.$$

Коэффициент корреляции между x и y оказался равным доле общих факторов в составе x и y , т. е. отношению числа общих факторов к числу всех факторов, входящих в x и y .

2) Рассмотрим более общую схему (А—Н)с, когда x и y имеют в своем составе несколько линейно независимых общих факторов u_1, u_2, \dots, u_s и ряд специфических линейно независимых факторов: $v_1, v_2, \dots, v_k; w_1, w_2, \dots, w_l$.

Пусть

$$\left. \begin{aligned} x &= u_1 + u_2 + \dots + u_s + v_1 + v_2 + \dots + v_k \\ y &= u_1 + u_2 + \dots + u_s + w_1 + w_2 + \dots + w_l \end{aligned} \right\}, \quad (8.2.5)$$

где

$$r_{u_i u_j} = r_{u_i v_j} = r_{u_i w_j} = r_{v_i w_j} = r_{v_i v_j} = 0,$$

и приближенно считаем $\sigma_{u_i} = \sigma_{u_j} = \sigma_{w_k}$.

Легко показать, что в этом случае

$$\begin{aligned}
 r_{xy} &= \frac{\frac{1}{n} \Sigma (x - \bar{x})(y - \bar{y})}{\sigma_x \sigma_y} = \\
 &= \frac{\sigma_{u_1}^2 + \sigma_{u_2}^2 + \dots + \sigma_{u_s}^2}{\sqrt{\sigma_{u_1}^2 + \dots + \sigma_{u_s}^2 + \sigma_{v_1}^2 + \dots + \sigma_{v_k}^2} \times} \\
 &\rightarrow \frac{\sigma_{u_1}^2 + \sigma_{u_2}^2 + \dots + \sigma_{u_s}^2}{\times \sqrt{\sigma_{u_1}^2 + \dots + \sigma_{u_s}^2 + \sigma_{w_1}^2 + \dots + \sigma_{w_l}^2}} \quad (8.2.6)
 \end{aligned}$$

и в силу равенства дисперсий

$$r_{xy} = \frac{s}{\sqrt{s+k} \sqrt{s+l}} = \sqrt{\frac{s}{s+k} \cdot \frac{s}{s+l}}; \quad (8.2.7)$$

$\frac{s}{s+k}$ является долей общих факторов в составе x ,

$\frac{s}{s+l}$ — долей общих факторов в составе y .

Формула \sqrt{ab} называется, как известно, средней геометрической чисел a, b .

Таким образом, коэффициент корреляции равен *средней геометрической из двух чисел: доли общих факторов в составе x и доли общих факторов в составе y* .

Схема (8.2.3) получается из схемы (8.2.5) при $s = k = l = 1$.

3) Рассмотрим еще более общую схему (А—Н)с, при которой каждый фактор входит с определенным весом (коэффициентом пропорциональности) в состав переменных x и y , т. е. обе переменные можно рассматривать как взвешенные суммы или линейные комбинации этих факторов

$$\left. \begin{aligned}
 x &= a + b_1 u_1 + \dots + b_s u_s + c_1 v_1 + \dots + c_k v_k \\
 y &= a' + b'_1 u_1 + \dots + b'_s u_s + c'_1 w_1 + \dots + c'_l w_l
 \end{aligned} \right\} \quad (8.2.8)$$

При сделанных ранее допущениях (линейная некоррелированность, равенство дисперсий факторов) легко получить для схемы (8.2.8) следующее выражение коэффициента корреляции:

$$\begin{aligned}
 r_{xy} &= \frac{b_1 b'_1 + \dots + b_s b'_s}{\sqrt{b_1^2 + \dots + b_s^2 + c_1^2 + \dots + c_k^2} \times} \\
 &\rightarrow \frac{b_1 b'_1 + \dots + b_s b'_s}{\times \sqrt{b_1'^2 + \dots + b_s'^2 + c_1'^2 + \dots + c_l'^2}} \quad (8.2.9)
 \end{aligned}$$

Из формулы (8.2.9) следует, что r_{xy} тем больше, чем значительнее весовые коэффициенты при общих факторах по сравнению с весовыми коэффициентами при специфических факторах. Это справедливо при условии, что соответственные коэффициенты (b_i и b'_i) имеют одинаковые знаки, т. е. что общие факторы влияют на признаки x и y в том же направлении.

Так, например, для схемы

$$x = 2u + v,$$

$$y = 3u + w,$$

$$r_{xy} = \frac{2 \cdot 3}{\sqrt{2^2 + 1^2} \sqrt{3^2 + 1^2}} = \frac{6}{\sqrt{5} \sqrt{10}} = 0,85.$$

Значение r_{xy} оказалось больше, чем в схеме (8.2.3), так как общий фактор u входит здесь с относительно большими весами b и b' .

Если же некоторые соответственные коэффициенты (b_i и b'_i) имеют разные знаки, т. е. некоторые общие факторы противоположным образом влияют на признак x и признак y , то величина r_{xy} уменьшается и может оказаться равной 0, отрицательной и даже близкой к -1 .

Так, например, для схемы

$$x = 2u + v,$$

$$y = -3u + w,$$

$$r_{xy} = \frac{-2 \cdot 3}{\sqrt{2^2 + 1^2} \sqrt{3^2 + 1^2}} = -0,85.$$

В схеме

$$x = u + v,$$

$$y = u - v,$$

оба фактора u и v являются общими, и схема не содержит специфических факторов. Но так как фактор u влияет на переменные x и y в том же направлении, а фактор v — в противоположных направлениях, то в результате r_{xy} оказывается равным нулю:

$$r_{xy} = \frac{1 - 1}{\sqrt{1^2 + 1^2} \sqrt{1^2 + 1^2}} = 0.$$

4) В предыдущих схемах предполагалось равенство дисперсий для всех факторов. Это условие всегда может быть реализовано, если измерять каждый из факторов в стандартизованных единицах — сигмах. Переход от натуральных единиц измерения к стандартизованным единицам должен, очевидно, отразиться на весовых коэффициентах. В самом деле, пусть, например,

$$\left. \begin{aligned} x &= bu + cv \\ y &= b'u + c'w \end{aligned} \right\}, \quad (8.2.10)$$

где v, u, w выражены в своих натуральных единицах. Обозначим буквами U, V, W переменные u, v, w , когда последние выражены в стандартизованных единицах — сигмах. Очевидно, что

$$U = \frac{u}{\sigma_u}, \quad V = \frac{v}{\sigma_v}, \quad W = \frac{w}{\sigma_w}.$$

Но тогда

$$\left. \begin{aligned} x &= b\sigma_u U + c\sigma_v V \\ y &= b'\sigma_u U + c'\sigma_w W \end{aligned} \right\} \quad (8.2.10')$$

или иначе

$$\left. \begin{aligned} x &= BU + CV \\ y &= B'U + C'W \end{aligned} \right\}, \quad (8.2.10'')$$

где

$$\begin{aligned} B &= b\sigma_u; & C &= c\sigma_v; \\ B' &= b'\sigma_u; & C' &= c'\sigma_w. \end{aligned}$$

В схеме (8.2.10'') соблюдается равенство дисперсий, так как $\sigma_U = \sigma_V = \sigma_W = 1$. Следовательно, применяя формулу (8.2.9), получим

$$r_{xy} = \frac{B \cdot B'}{\sqrt{B^2 + C^2} \sqrt{B'^2 + C'^2}}.$$

Отсюда окончательно

$$r_{xy} = \frac{b \cdot b' \sigma_u^2}{\sqrt{b^2 \sigma_u^2 + c^2 \sigma_v^2} \sqrt{b'^2 \sigma_u^2 + c'^2 \sigma_w^2}}. \quad (8.2.11)$$

Формула (8.2.11) легко обобщается на случай многих факторов. Пусть

$$\left. \begin{aligned} x &= a + b_1 u_1 + \dots + b_s u_s + c_1 v_1 + \dots + c_k v_k \\ y &= a' + b'_1 u_1 + \dots + b'_s u_s + c'_1 w_1 + \dots + c'_l w_l \end{aligned} \right\} \quad (8.2.8')$$

является схемой (А—Н)с, в которой мы сейчас не предполагаем равенства дисперсий факторов, как это имело место для схемы (8.2.8).

Тогда вполне аналогично формуле (8.2.11) мы получаем формулу

$$\begin{aligned} r_{xy} &= \frac{b_1 b'_1 \sigma_{u_1}^2 + \dots + b_s b'_s \sigma_{u_s}^2}{\sqrt{b_1^2 \sigma_{u_1}^2 + \dots + b_s^2 \sigma_{u_s}^2 + c_1^2 \sigma_{v_1}^2 + \dots + c_k^2 \sigma_{v_k}^2} \times} \\ &\rightarrow \frac{b_1 b'_1 \sigma_{u_1}^2 + \dots + b_s b'_s \sigma_{u_s}^2}{\times \sqrt{b_1'^2 \sigma_{u_1}^2 + \dots + b_s'^2 \sigma_{u_s}^2 + c_1'^2 \sigma_{w_1}^2 + \dots + c_l'^2 \sigma_{w_l}^2}}. \end{aligned} \quad (8.2.12)$$

5) Приведенные выше схемы, несмотря на их частный характер (аддитивное и независимое действие факторов), дают все же возможность в ряде случаев понять происхождение величины коэффициента корреляции. В некоторых случаях эти схемы могут быть использованы для грубого определения величины коэффициента корреляции на основе предварительных соображений о природе зависимости без данных статистического наблюдения.

Пример 1. Для веса валика g и боковой поверхности s существуют зависимости

$$\left. \begin{aligned} g &= \pi r^2 h \gamma (1 + \varepsilon) \\ s &= 2\pi r h \end{aligned} \right\},$$

где r — радиус поперечного сечения;

h — высота;

γ — удельный вес материала;

$(1 + \varepsilon)$ — коэффициент искажения цилиндрической формы.

Логарифмируя эти равенства, приходим к соотношениям между логарифмами

$$\ln g = \ln \pi + 2 \ln r + \ln h + \ln \gamma + \ln (1 + \varepsilon),$$

$$\ln s = \ln 2\pi + \ln r + \ln h$$

или в других обозначениях

$$x = \ln \pi + 2u_1 + u_2 + v_1 + v_2,$$

$$y = \ln 2\pi + u_1 + u_2.$$

Примем следующие ориентировочные допущения:

а) аргументы u_1 , u_2 , v_1 , v_2 линейно не коррелируют между собой;

$$\text{б) } \sigma_{u_1} \approx \sigma_{u_2} \approx \sigma_{v_1} \approx \sigma_{v_2}.$$

На основе этих ориентировочных допущений мы можем грубо оценить значение коэффициента корреляции между x и y , не прибегая к обработке статистических данных:

$$r_{xy} = \frac{2 + 1}{\sqrt{4 + 1 + 1} \sqrt{1 + 1}} = \frac{3}{\sqrt{12}} = 0,867.$$

Итак, коэффициент корреляции между логарифмами веса и боковой поверхности цилиндра мы при сделанных выше допущениях ожидаем получить (при достаточно большом количестве наблюдений) равным 0,87.

Если вместо допущения (б) принять допущение

$$\sigma_{u_1} : \sigma_{u_2} : \sigma_{v_1} : \sigma_{v_2} = c_1 : c_2 : c_3 : c_4,$$

где c_1, c_2, c_3, c_4 — заранее предполагаемые коэффициенты пропорциональности, то r_{xy} рассчитаем по формуле

$$r_{xy} = \frac{2c_1^2 + c_2^2}{\sqrt{4c_1^2 + c_2^2 + c_3^2} \sqrt{c_1^2 + c_2^2}}.$$

6) Общие схемы действия факторов

$$\left. \begin{aligned} x &= \varphi(u, v) \\ y &= \psi(u, w) \end{aligned} \right\} \quad (8.2.1)$$

и

$$\left. \begin{aligned} x &= \varphi(u_1 \dots u_s, v_1 \dots v_k) \\ y &= \psi(u_1 \dots u_s, w_1 \dots w_l) \end{aligned} \right\}, \quad (8.2.2)$$

где φ и ψ — функции от линейно некоррелированных аргументов, часто могут быть приближенно сведены к схемам (А—Н)с. Так, если φ и ψ монотонно изменяющиеся аналитические функции своих аргументов, то обычно в довольно широких пределах можно удовлетвориться линейной частью разложения этих функций в ряды Тейлора

$$\varphi(u, v) = \varphi_0 + \frac{u - u_0}{1} \left[\frac{\partial \varphi}{\partial u} \right]_0 + \frac{v - v_0}{1} \left[\frac{\partial \varphi}{\partial v} \right]_0 + \dots,$$

$$\psi(u, w) = \psi_0 + \frac{u - u_0}{1} \left[\frac{\partial \psi}{\partial u} \right]_0 + \frac{w - w_0}{1} \left[\frac{\partial \psi}{\partial w} \right]_0 + \dots,$$

что приводит к схеме (А—Н)с.

Это обстоятельство делает схемы (А—Н)с полезными для грубой ориентировки и при неаддитивных схемах действия факторов.

Пример 2. Между оценками успеваемости, которые получают студенты по различным предметам экзаменационной сессии, существует корреляционная зависимость. Студенты, получившие лучшие оценки по одним предметам, получают, как правило, более высокие отметки и по другим предметам. И, наоборот, худшим оценкам по одним предметам соответствуют худшие оценки и по другим предметам.

Наличие такой корреляции нельзя объяснить причинной зависимостью между оценками. Оценки по иностранному языку не стоят, например, в причинной зависимости с оценками по сопротивлению материалов. Тем не менее наличие корреляции между оценками по этим двум предметам вполне закономерно и, даже не производя наблюдений, можно ожидать ее заранее. Объяснение этой корреляции может быть дано на основе теории общих факторов. В самом деле, можно с несомненностью утверждать, что такие качества студентов, как трудолюбие, способности, умение организовать свою работу и другие, являются факто-

рами, определяющими у одних студентов более высокие оценки, у других — худшие результаты по многим предметам одновременно. Наличие таких общих факторов и должно вызвать корреляцию между оценками по различным предметам. Однако нельзя заранее предсказать величину этой корреляции, так как не знаем ни полного списка общих и специфических взаимно-независимых факторов корреляции¹, ни удельных весов, выясняющих значимость каждого фактора, и, наконец, не известны показатели дисперсии этих факторов в данном коллективе студентов².

Величина коэффициентов корреляции может быть получена лишь путем расчета на конкретном материале.

Приводим начало таблицы, в которой даны результаты экзаменационной сессии по предметам II курса института для 40 студентов.

Таблица 49

Результаты экзаменационной сессии студентов II курса

Номера студентов по порядку	Оценка по предметам					
	Математика	Теоретическая механика	Сопромат	Электротехника	Основы марксизма-ленинизма	Иностранный язык
1	4	4	4	5	3	5
2	3	4	3	3	4	4
3	5	3	—	5	5	5
4	5	5	3	4	5	4
5	3	3	3	3	3	5
6	3	3	3	3	3	4
7	3	4	3	5	3	3

По этой таблице нами были исчислены коэффициенты корреляции между всеми предметами попарно. Приводим таблицу этих коэффициентов.

Анализ табл. 50 показывает, что коэффициенты корреляции колеблются вокруг среднего уровня $r=0,33$.

Выделив в одну группу предметы физико-технического направления, убеждаемся в том, что средний коэффициент корреляции между предметами входящими в данную группу, оказался заметно выше, чем средний коэффициент корреляции между предметами физико-технического направления и другими предметами сессии

$$r_A = \frac{0,33 + 0,39 + 0,33 + 0,43 + 0,31 + 0,50}{6} = 0,38,$$

¹ Перечисленные выше факторы не являются взаимно-независимыми.

² Заметим также, что схема (A—H)c должна здесь рассматриваться лишь в качестве весьма грубого приближения.

$$r_B = \frac{0,29 + 0,29 + 0,28 + 0,13 + 0,35 + 0,31 + 0,29 + 0,31}{8} = 0,28.$$

Это понятно, так как величина коэффициентов корреляции в первой группе обусловлена действием *дополнительного общего фактора* — родством этих предметов и наличием у студентов большей или меньшей склонности к физико-техническим дисциплинам.

Таблица 50

	Математика	Теоретическая механика	Сопромат	Электротехника	Основы марксизма-ленинизма	Иностранный язык	Средний
Математика	×	0,33	0,39	0,33	0,29	0,29	0,33
Теоретическая механика	0,33	×	0,43	0,31	0,28	0,13	0,30
Сопромат	0,39	0,43	×	0,50	0,35	0,31	0,40
Электротехника	0,33	0,31	0,50	×	0,29	0,31	0,35
Основы марксизма-ленинизма	0,29	0,28	0,35	0,29	×	0,33	0,31
Иностранный язык	0,29	0,13	0,31	0,31	0,33	×	0,27
							0,33

ГЛАВА IX

КОРРЕЛЯЦИЯ ТРЕХ ПЕРЕМЕННЫХ

§ 1. УРАВНЕНИЕ МНОЖЕСТВЕННОЙ РЕГРЕССИИ
ДЛЯ ТРЕХ ПЕРЕМЕННЫХ

В настоящей главе мы рассматриваем задачи корреляционного анализа, возникающие при изучении статистического материала по трем варьирующим признакам.

Исходный материал представляет собой совокупность наблюдений над элементами одного и того же вида по трем признакам x , y , z одновременно. Таковы, например, данные табл. 49, которая содержит примерную запись результатов наблюдений над элементами совокупности (студентами) по 6 варьирующим признакам (оценки по различным предметам сессии).

В качестве второго примера таблицы, содержащей данные наблюдений по трем признакам, рассмотрим табл. 51, в которой указаны механические характеристики (предел выносливости, предел прочности, текучести) для 55 образцов различных марок стали¹.

При изучении корреляции трех переменных одна из них рассматривается как функция, две другие — как аргументы. Вместо двумерной корреляционной таблицы для систематизации данных наблюдений здесь используется трехмерная корреляционная таблица, в которой каждой комбинации вариантов по трем признакам соответствует своя частота (см. табл. 52).

Подобно тому как двумерная корреляционная таблица графически изображалась в виде поля корреляции, данные трехмерной корреляционной таблицы можно себе представить на графике в виде корреляционного пространства. По оси x и оси y указаны интервалы аргументов, по оси z — интервалы функции. Частота каждой комбинации (x, y, z) выражается соответствующим коли-

¹ См. С. Л. Жуков, Вид связи предела выносливости с характеристиками прочности при растяжении, «Заводская лаборатория», № 1, 1946, стр. 105.

Таблица 51

Марки стали	Предел прочности на разрыв (в кг/мм ²) у	Предел текучести (в кг/мм ²) х	Предел выносливости при симметричном изгибе (в кг/мм ²) z	Марки стали	Предел прочности на разрыв (в кг/мм ²) у	Предел текучести (в кг/мм ²) х	Предел выносливости при симметричном изгибе (в кг/мм ²) z
25	51,6	33,0	25,0	30ХМА	143,0	138,0	61,5
45	93,0	70,4	43,0	30ХМА	107,0	94,0	48,0
75	69,2	40,0	30,0	38ХА	86,0	69,0	38,0
45РС	80,8	66,4	40,0	3ХТВ	101,0	—	42,0
12Г1А	60,0	41,0	29,5	12ХН3А	141,0	112,0	64,0
12Г2А	145,0	123,0	55,0	12ХН3А	61,6	47,5	36,0
12Г2А	123,0	113,0	49,0	20ХН3А	96,0	87,0	43,6
12Г2А	80,0	72,0	43,0	20ХН3А	84,0	79,0	43,0
12Г2А	85,0	42,0	36,0	20ХН3А	76,4	71,1	32,0
25ХГСА	162,0	135,0	54,5	37Н3А	126,5	121,0	56,5
25ХГСА	146,0	132,0	51,5	18ХНВА	150,0	25,0	63,0
25ХГСА	80,9	82,0	41,0	18ХНВА	139,0	119,0	62,0
30ХГСА	180,0	153,0	71,5	18ХНВА	137,5	114,0	62,0
30ХГСА	154,0	137,0	65,0	18ХНВА	124,6	112,1	51,0
30ХГСА	90,0	85,0	48,0	18ХНВА	116,5	77,0	56,0
30ХГСА	193,0	155,0	60,0	25ХНВА	98,0	84,0	45,0
(плавка 11008)				25ХНВА	112,0	96,0	50,0
30ХГСА	178,0	154,0	59,0	25ХНВА	108,6	98,0	50,0
(плавка 11008)				40ХНМА	212,0	172,0	73,0
30ХГСА	158,0	142,0	54,0	40ХАМА (Б)	102,0	86,0	55,0
(плавка 11008)				40ХНМА	110,0	106,0	58,0
30ХГСА	165,6	134,2	63,0	40ХНМА	104,0	92,0	45,0
(плавка 1445)				40ХНМА	119,0	113,0	55,0
30ХГСА	116,8	91,7	48,0	20ХНФА	102,0	98,0	48,0
(плавка 1445)				20ХНФА	107,5	103,0	50,0
30ХГСА	74,9	57,9	36,0	38ХМЮА	103,0	85,0	46,0
(плавка 1445)				ЭИ-273	98,3	83,0	46,0
30ХГСА	69,2	44,1	32,0	ЭИ-274	143,0	105,0	65,0
(плавка 1445)				ЭИ-275	114,0	100,0	63,0
30ХГСА	69,2	44,1	32,0	ЭИ-312	95,0	50,0	44,5
(плавка 1445)				ЭИ-310	90,0	45,0	42,0
30ХМА	160,0	145,0	62,0	ЭИ-69	72,0	40,0	45,0

Таблица 52

Трехмерная корреляционная таблица

	y_1				y_2				y_l										
	x_1	x_2	...	x_s	x_1	x_2	...	x_s	x_1	x_2	...	x_s	x_1	x_2	...	x_s	
z_1																			
z_2																			
...																			
z_k																			

чеством точек, размещенных в пространстве, в координатной ячейке (x, y, z) .

В каждой интервальной клетке (x_i, y_j) мы рассчитаем среднее значение аппликаты z_{ij} для всех точек, оказавшихся в соответствующем параллелепипеде (рис. 46). Отмечаем значком (*) вершину каждой аппликаты. Геометрическое место этих значков

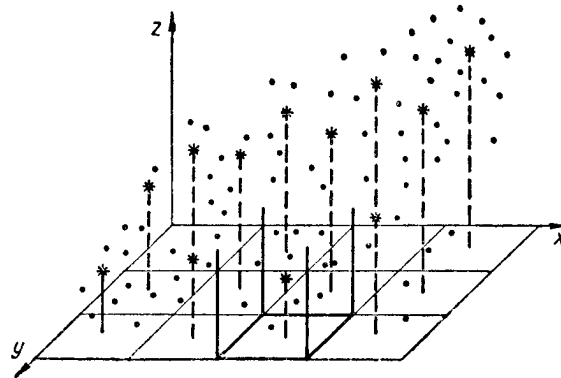


Рис. 46

дает представление об эмпирической поверхности регрессии z по x и y , если вообразить последнюю в виде гибкой материи, натянутой на наши значки.

С увеличением количества наблюдений и одновременном, но не столь быстром увеличении числа интервалов группировки (по x и y) эмпирическая поверхность регрессии выравнивается и стремится к предельной теоретической поверхности регрессии (рис. 47).

Нахождение теоретической поверхности регрессии, выражающей зависимость между средним значением функции \bar{z}_{xy} и аргументами x и y , составляет содержание настоящего параграфа. Разумеется, так же как и в случае двух переменных, решение задачи должно быть дано на ограниченном количестве наблюдений

ний косвенными методами. Мы ограничимся исследованием того частного, но наиболее важного случая, когда теоретическая поверхность регрессии является плоскостью (рис. 48).

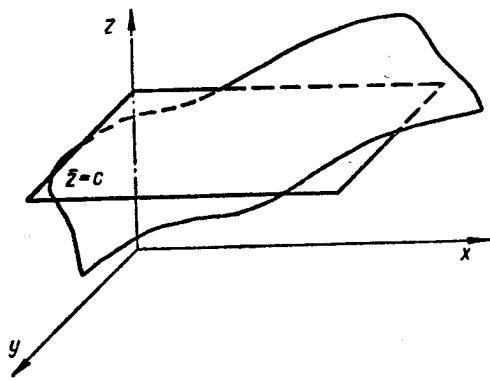


Рис. 47

Пусть

$$\bar{z}_{xy} = a + bx + cy \tag{9.1.1}$$

уравнение плоскости регрессии, где a, b, c — неизвестные параметры, требующие определения.

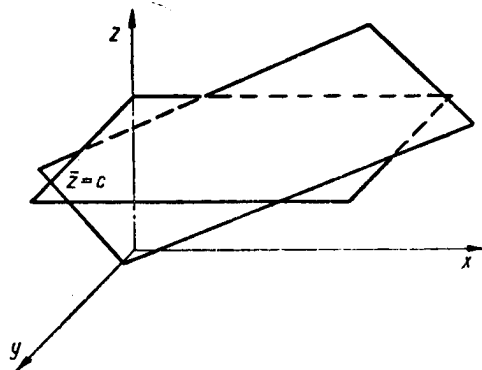


Рис. 48

Расчет параметров a, b, c производим способом наименьших квадратов. Потребуем, чтобы сумма квадратов отклонений фактических аппликат z от аппликат \bar{z}_{xy} , вычисленных по уравнению регрессии, была наименьшей.

$$\Sigma (z - \bar{z}_{xy})^2 = \min. \tag{9.1.2}$$

Левую часть выражения (9.1.2) обозначим буквой f . f есть функция от неизвестных параметров a, b, c .

Минимум этой функции мы найдем из уравнений:

$$\frac{\partial f}{\partial a} = 0, \quad \frac{\partial f}{\partial b} = 0, \quad \frac{\partial f}{\partial c} = 0. \quad (9.1.3)$$

Произведя дифференцирование, напомним систему (9.1.3) в окончательном виде

$$\left. \begin{aligned} \Sigma z &= na + b\Sigma x + c\Sigma y \\ \Sigma zx &= a\Sigma x + b\Sigma x^2 + c\Sigma yx \\ \Sigma zy &= a\Sigma y + b\Sigma xy + c\Sigma y^2 \end{aligned} \right\}. \quad (9.1.4)$$

Систему (9.1.4) решаем относительно неизвестных a, b, c следующим образом. Разделим каждое уравнение на n , получим

$$\left. \begin{aligned} \bar{z} &= a + b\bar{x} + c\bar{y} \\ \bar{zx} &= a\bar{x} + b\frac{\Sigma x^2}{n} + c\bar{y}\bar{x} \\ \bar{zy} &= a\bar{y} + b\bar{y}\bar{x} + c\frac{\Sigma y^2}{n} \end{aligned} \right\}. \quad (9.1.4')$$

Умножаем верхнее уравнение системы (9.1.4') на \bar{x} и вычитаем из второго

$$\bar{zx} - \bar{z} \cdot \bar{x} = b\left(\frac{\Sigma x^2}{n} - \bar{x}^2\right) + c(\bar{y}\bar{x} - \bar{y} \cdot \bar{x}). \quad (9.1.5)$$

Аналогично, умножая верхнее уравнение (9.1.4') на \bar{y} и вычитая из нижнего, получим

$$\bar{zy} - \bar{z} \cdot \bar{y} = b(\bar{y}\bar{x} - \bar{y} \cdot \bar{x}) + c\left(\frac{\Sigma y^2}{n} - \bar{y}^2\right). \quad (9.1.5')$$

Замечая, что

$$\frac{\Sigma x^2}{n} - \bar{x}^2 = \sigma_x^2, \quad \frac{\Sigma y^2}{n} - \bar{y}^2 = \sigma_y^2,$$

и используя формулу (7.3.5), запишем уравнения (9.1.5) и (9.1.5') в виде

$$\left. \begin{aligned} r_{zx}\sigma_z\sigma_x &= b\sigma_x^2 + cr_{yx}\sigma_y\sigma_x \\ r_{zy}\sigma_z\sigma_y &= br_{yx}\sigma_y\sigma_x + c\sigma_y^2 \end{aligned} \right\} \quad (9.1.6)$$

или

$$\left. \begin{aligned} r_{zx}\sigma_z &= b\sigma_x + cr_{yx}\sigma_y \\ r_{zy}\sigma_z &= br_{yx}\sigma_x + c\sigma_y \end{aligned} \right\}. \quad (9.1.7)$$

Решая уравнения (9.1.7) относительно b и c , приходим к следующим формулам:

$$\left. \begin{aligned} b &= \frac{\sigma_z}{\sigma_x} \cdot \frac{r_{zx} - r_{zy}r_{xy}}{1 - r_{xy}^2} \\ c &= \frac{\sigma_z}{\sigma_y} \cdot \frac{r_{zy} - r_{zx}r_{xy}}{1 - r_{xy}^2} \end{aligned} \right\} \quad (9.1.8)$$

Из первого уравнения системы (9.1.4') находим, что

$$a = \bar{z} - b\bar{x} - c\bar{y}, \quad (9.1.9)$$

где b , c уже определены формулами (9.1.8).

Посмотрим, что же нового для уяснения зависимости между величинами вносит уравнение множественной регрессии по сравнению с уравнениями парной регрессии. Сопоставим одноименные коэффициенты в уравнениях парной и множественной регрессии

$$\bar{z}_x = a + bx, \quad (I)$$

$$\bar{z}_{xy} = a + bx + cy. \quad (II)$$

Коэффициент b в уравнении (I) рассчитывается по формуле:

$$b = \frac{\sigma_z}{\sigma_x} r_{zx}.$$

Он отвечает на вопрос, на сколько единиц в среднем изменяется z при увеличении x на одну единицу. Или иначе, на сколько единиц изменился бы z с изменением x на одну единицу, если бы прочие действующие по схеме (А—М—Н) факторы z остались бы на одном и том же среднем уровне.

Коэффициент b в уравнении (II) рассчитывается по формуле (9.1.8). Он отвечает на вопрос, на сколько единиц изменяется в среднем z с увеличением x на одну единицу *при постоянном y* . Или иначе, на сколько единиц изменился бы z с увеличением x на единицу, если бы y не изменялся и прочие действующие по схеме (А—М—Н) факторы z оставались на одном и том же среднем уровне.

Таким образом, уравнение множественной регрессии позволяет элиминировать (устранить) не только влияние тех факторов, которые действуют по схеме (А—М—Н), как это имеет место в уравнении парной регрессии, но и *влияние фактора y , корреляционно связанного с x* .

В этом и заключается то существенно новое, что вносит уравнение множественной регрессии в исследование связи между величинами.

Пример 1. Для исследования зависимости между производительностью труда, возрастом и стажем рабочего были собраны сведения о большой группе рабочих одной и той же специальности. Приводим таблицу основных показателей этой группы по трем признакам: производительности труда — z , возрасту — x , стажу — y .

Таблица 53

$\bar{z} = 70$ единиц	$\sigma_z = 8$ единиц	$r_{zx} = 0,20$
$\bar{x} = 40$ лет	$\sigma_x = 5$ лет	$r_{zy} = 0,40$
$\bar{y} = 8$ лет	$\sigma_y = 4$ года	$r_{xy} = 0,80$

По данным этой таблицы рассчитаем в порядке приближения уравнения спрямленной регрессии:

а) Уравнение парной регрессии производительности труда по возрасту

$$\bar{z}_x = a + bx;$$

$$b = r_{zx} \frac{\sigma_z}{\sigma_x} = 0,20 \cdot \frac{8}{5} = 0,32;$$

$$a = \bar{z} - b\bar{x} = 70 - 0,32 \cdot 40 = 57,2;$$

откуда

$$\bar{z}_x = 57,2 + 0,32x. \quad (\alpha)$$

в) Уравнение парной регрессии производительности труда по стажу

$$\bar{z}_y = a + cy;$$

$$c = \frac{8}{4} \cdot 0,40 = 0,80;$$

$$a = 70 - 0,80 \cdot 8 = 63,6;$$

откуда

$$\bar{z}_y = 63,6 + 0,80y. \quad (\beta)$$

с) Уравнение множественной регрессии производительности труда по возрасту и стажу одновременно

$$\bar{z}_{xy} = a + bx + cy;$$

$$b = \frac{8}{5} \cdot \frac{0,20 - 0,40 \cdot 0,80}{1 - 0,80^2} = \frac{8}{5} \cdot \frac{(-0,12)}{0,36} = -0,53;$$

$$c = \frac{8}{4} \cdot \frac{0,40 - 0,20 \cdot 0,80}{1 - 0,80^2} = \frac{8}{4} \cdot \frac{0,24}{0,36} = 1,33;$$

$$a = 70 + 0,53 \cdot 40 - 1,33 \cdot 8 = 80,56;$$

$$\bar{z}_{xy} = 80,56 - 0,53x + 1,33y. \quad (\gamma)$$

Сопоставляя полученные уравнения, мы видим, что коэффициенты регрессии существенно изменились при переходе от парной корреляции к уравнению множественной регрессии. Так, коэффициент при стаже вырос от значения 0,80 до значения 1,33; коэффициент при возрасте b упал и превратился в отрицательный — от значения 0,32 до значения $(-0,53)$. Обратим внимание на последнее расхождение. Пользуясь уравнением (α) , находим, что с увеличением возраста рабочего на один год производительность труда растет в среднем на 0,32 единицы. Пользуясь уравнением (γ) , можно заключить, что с увеличением возраста на один год производительность падает в среднем на 0,53 единицы.

В чем же разгадка этого противоречия? Укажем, что коэффициенты регрессии в уравнениях (α) и (γ) отвечают различной постановке вопроса. В случае (γ) коэффициент регрессии показывает скорость изменения z по x при постоянном y , т. е. показывает, как изменялась бы производительность труда с увеличением возраста, *если бы стаж оставался на том же уровне*. В случае (α) *последнее условие не ставится*.

Чтобы понять смысл этого условия, обратимся к анализу природы действующих факторов. Возраст, если рассматривать его как показатель работоспособности организма, должен, по-видимому, оказывать следующее влияние на производительность труда рабочего: с увеличением возраста производительность труда растет, достигает максимума, а затем постепенно снижается¹. Оптимальное значение возраста не является чем-то неизменным. Можно, по-видимому, утверждать, что для профессий физического труда максимум работоспособности достигается в более раннем возрасте, чем для профессий умственного труда. В нашей профессиональной группе средний возраст рабочих оказался равным 40 годам (см. табл. 53). Есть основания полагать, что оптимальное значение возраста в данной профессии значительно меньше среднего. Поэтому на нашем материале следовало бы ожидать отрицательной корреляции между производительностью труда и возрастом, если рассматривать последний лишь как показатель работоспособности организма. Однако с увеличением возраста увеличивается стаж, а следовательно, производственный

¹ Отсюда следует, что схемы прямолинейного изменения переменных, положенные в основу уравнений (α) и (γ) , не могут рассматриваться в качестве правильного отображения искомой зависимости. Мы, однако, ограничимся рассмотрением вопроса в упрощенном виде, имея в виду, что получаемые коэффициенты регрессии выражают не истинные скорости изменения переменных, а лишь средние значения этих скоростей в определенной области изменения переменных.

опыт и квалификация рабочего. Между возрастом и производственным стажем наблюдается тесная корреляция:

$$r_{xy} = 0,89.$$

Увеличение производственного стажа, сопутствующее увеличению возраста, компенсирует отрицательную роль последнего как показателя работоспособности организма. Этим обусловлена небольшая положительная величина коэффициента корреляции между производительностью труда и возрастом ($r_{zx} = 0,20$) и этим же объясняется небольшой положительный коэффициент регрессии для возраста в уравнении (а) : $b = 0,32$.

В самом деле, уравнение (а) отвечает на вопрос, как в среднем изменяется производительность труда с увеличением возраста на один год. Но, увеличивая возраст, мы тем самым увеличиваем и стаж рабочего. Из двух случайно взятых рабочих более старший по возрасту обладает в среднем и большим производственным опытом. Последнее обстоятельство заставляет ожидать от него в среднем и несколько более высокой производительности труда.

Уравнение (γ) отвечает на аналогичный вопрос и в то же время требует закрепления производственного стажа. Но из двух рабочих, обладающих одним и тем же производственным стажем, от младшего по возрасту ожидается более высокая производительность труда. Это запечатлено в виде отрицательного коэффициента b для возраста в уравнении множественной регрессии. Так объясняется кажущееся противоречие между уравнениями (а) и (γ).

Аналогично устанавливаются причины, в силу которых коэффициенты регрессии производительности труда по стажу в уравнениях (β) и (γ) сильно расходятся между собой.

Связывая производительность труда со стажем уравнением (β), мы не делаем никаких оговорок об изменении возраста. Но с увеличением стажа в среднем увеличивается возраст рабочего в силу тесной корреляции между этими двумя признаками. Поэтому увеличению стажа на один год соответствует по уравнению (β) значительно меньший прирост производительности труда, нежели это имеет место по уравнению (γ), где возраст считается закрепленным. В этом причина расхождения коэффициентов регрессии.

Пример 2. При изучении механических свойств стали представляет интерес нахождение формулы, с помощью которой можно определять предел выносливости на основе статических характеристик металла. Для этой цели найдем по данным табл. 51 уравнение множественной регрессии, связывающее предел выносливости (z) с пределом прочности на разрыв (y) и пределом текучести стали (x). Прежде всего рассчитаем таблицу основных показателей:

Таблица 54

$\bar{z} = 49,72$	$\sigma_z = 11,01$	$r_{zx} = 0,814$
$\bar{x} = 93,89$	$\sigma_x = 35,73$	$r_{zy} = 0,881$
$\bar{y} = 114,47$	$\sigma_y = 35,62$	$r_{xy} = 0,896$
	$n = 55$	

С помощью табл. 54 вычисляем коэффициенты уравнения множественной регрессии по формулам (9.1.8), (9.1.9):

$$b = \frac{11,01}{35,73} \cdot \frac{0,814 - 0,881 \cdot 0,896}{1 - 0,896^2} = 0,039;$$

$$c = \frac{11,01}{35,62} \cdot \frac{0,881 - 0,814 \cdot 0,896}{1 - 0,896^2} = 0,238;$$

$$a = 49,72 - 0,039 \cdot 93,89 - 0,238 \cdot 114,47 = 18,82;$$

$$z_{xy} = 18,82 + 0,039x + 0,238y. \quad (A)$$

Уравнение (A) позволяет рассчитывать предел выносливости стали по данным о пределе текучести и пределе прочности.

В работе С. Л. Жукова¹ для той же цели предложена другая формула

$$\bar{z}_{xy} = 20,5 + 0,138(x + y). \quad (B)$$

При выводе формулы (B) методом *парной* корреляции z с суммой $x+y$ автор исходил из предположения о равенстве весовых коэффициентов при x и y .

В формуле (A), представляющей наилучшую (в смысле метода наименьших квадратов) линейную комбинацию x и y , вес при x оказался намного меньше, чем вес при y . Поэтому формула (B) должна уступать в точности формуле (A), если сравнивать их пригодность на материале табл. 51².

Пример 3. По данным табл. 55 найдем уравнение множественной регрессии, связывающее упругость поршневого кольца с двумя ее факторами одновременно: размером замка в свободном состоянии и магнитной проницаемостью кольца. Приводим табл. 55 основных показателей, рассчитанных по данным табл. 4 приложения I.

¹ С. Л. Жуков, Вид связи предела выносливости с характеристиками прочности при растяжении, «Заводская лаборатория» № 1, 1946, стр. 105.

² В таблицу, приведенную в статье, по-видимому, вкралась ошибка. В строке 35 сверху значение предела текучести (x), вероятно, должно быть 125, вместо указанного 25. Если учесть это исправление, то r_{zx} сблизится с r_{zy} и веса при x и y окажутся близкими друг к другу.

Таблица 55

$\bar{z} = 1410,75$	$\sigma_z = 83,14$	$r_{xz} = 0,440$
$\bar{x} = 17,99$	$\sigma_x = 1,05$	$r_{zy} = 0,305$
$\bar{y} = 7,48$	$\sigma_y = 2,72$	$r_{xy} = 0,099$
	$n = 200$	

Рассчитаем параметры уравнения множественной регрессии

$$b = \frac{83,14}{1,05} \cdot \frac{0,440 - 0,305 \cdot 0,099}{1 - 0,099^2} = 32,70;$$

$$c = \frac{83,14}{2,72} \cdot \frac{0,305 - 0,440 \cdot 0,099}{1 - 0,099^2} = 8,07;$$

$$a = 1410,75 - 32,70 \cdot 17,99 - 8,07 \cdot 7,48 = 762,12$$

отсюда

$$\bar{z}_{xy} = 762,12 + 32,70x + 8,07y. \quad (\lambda)$$

Следует, однако, иметь в виду, что прямолинейный тип зависимости, положенной в основу уравнения (λ) , может быть здесь принят лишь в качестве первого и грубого приближения. Более детального исследования этой зависимости, основанного на схемах криволинейного изменения переменных, не производим.

§ 2. ПОКАЗАТЕЛИ ТЕСНОТЫ СВЯЗИ

Рассмотрим корреляционное пространство переменных x, y, z . Построим эмпирическую поверхность регрессии и плоскость $z=c$, параллельную координатной плоскости xy на уровне средней величины z (см. рис. 47).

Дисперсия σ_z^2 измеряет колеблемость точек корреляционного пространства вокруг плоскости $z=c$. Эту дисперсию можно разложить на два компонента:

$$\sigma_z^2 = \sigma_{ij}^2 + \delta_{ij}^2, \quad (9.2.1)$$

$$\sigma_{ij}^2 = \frac{1}{n} \sum (z - \bar{z}_{ij})^2 = \frac{1}{n} \sum z_{ij}^2 m \quad (9.2.2)$$

— измеряет колеблемость точек корреляционного пространства вокруг эмпирической поверхности регрессии;

$$\delta_{ij}^2 = \frac{1}{n} \sum (\bar{z}_{ij} - \bar{z})^2 = \frac{1}{n} \sum \delta_{ij}^2 m \quad (9.2.3)$$

— измеряет колеблемость эмпирической поверхности регрессии вокруг плоскости $z=c$.

В формулах (9. 2. 1), (9. 2. 2), (9. 2. 3) показатели \bar{z}_{ij} , σ_{ij} , δ_{ij} сопровождаются двумя индексами i, j , так как значения этих показателей отнесены к координатам (x, y) и суммирование производится по двум направлениям.

Вывод формулы (9. 2. 1), так же как и формулы (7. 1. 2) для случая двух переменных, опирается на теорему о дисперсии для объединенной совокупности, доказанную в предыдущем разделе книги.

Так же, как и в случае корреляции двух переменных, можно считать, что при безгранично большом числе наблюдений компонент $\bar{\delta}_{ij}^2$ отражает влияние факторов x и y , а компонент $\bar{\sigma}_{ij}^2$ влияние прочих, независимо от x и y действующих факторов.

При малом числе наблюдений случайные зигзаги эмпирической поверхности регрессии систематически преувеличивают компонент $\bar{\delta}_{ij}^2$ и преуменьшают компонент σ_{ij}^2 .

Заменяя эмпирическую поверхность регрессии поверхностью теоретической, мы можем воспользоваться вместо равенства (9. 2. 1) формулой

$$\sigma_z^2 = \bar{\sigma}_{ijT}^2 + \bar{\delta}_{ijT}^2, \quad (9. 2. 4)$$

где

$$\bar{\sigma}_{ijT}^2 = \frac{1}{n} \sum (z - \bar{z}_{xy})^2 = \frac{1}{n} \sum \sigma_{ijT}^2 m_{ij} \quad (9. 2. 5)$$

и

$$\bar{\delta}_{ijT}^2 = \frac{1}{n} \sum (\bar{z}_{xy} - \bar{z})^2. \quad (9. 2. 6)$$

Формула (9. 2. 4) ведет к более правильному распределению дисперсии между двумя компонентами (сравни § 2 гл. VII).

Показатели

$$\eta_{z|xy} = \sqrt{\frac{\bar{\delta}_{ij}^2}{\sigma_z^2}} \quad (9. 2. 7)$$

и

$$\eta_{z|xyT} = \sqrt{\frac{\bar{\delta}_{ijT}^2}{\sigma_z^2}} \quad (9. 2. 8)$$

называются эмпирическим и теоретическим множественными корреляционными отношениями z по x и y . Они служат измерителями тесноты зависимости функции z от ее аргументов x и y одновременно.

Как и в случае корреляции двух переменных, можно утверждать, что

$$0 \leq \eta_T \leq \eta \leq 1. \quad (9. 2. 9)$$

При $\eta=0$ или $\eta_T=0$ соответствующая (эмпирическая или теоретическая) поверхность регрессии параллельна координатной плоскости (x, y) и совпадает с плоскостями $z=c$. В этом случае говорят об отсутствии корреляции между z и x, y .

При $\eta=1$ или $\eta_T=1$, $\bar{\sigma}_{ij}^2=0$ или $\bar{\sigma}_{ijT}^2=0$ все точки корреляционного пространства лежат на соответствующей поверхности регрессии. В этом случае говорят о наличии точной (функциональной) зависимости между z и x, y .

В случае, когда $0<\eta<1$ или $0<\eta_T<1$, говорят о наличии корреляционной зависимости более или менее тесной. Разумеется, все эти заключения надежны лишь в той мере, в какой они опираются на большое количество наблюдений.

Коэффициент множественной корреляции для трех переменных

Рассмотрим случай, имеющий особо важное значение в статистической теории и практике, когда теоретическая поверхность регрессии является плоскостью (см. рис. 48).

В этом случае теоретическое корреляционное отношение $\eta_{z/xyT}$ носит название *коэффициента множественной корреляции* и обозначается буквой R .

Вывод формулы R

$$R^2 = \frac{\bar{\sigma}_{ijT}^2}{\sigma_z^2} = \frac{\frac{1}{n} \dot{\Sigma} (\bar{z}_{xy} - \bar{z})^2}{\sigma_z^2};$$

при этом

$$\bar{z}_{xy} = a + bx + cy \text{ (уравнение плоскости)}$$

$$\bar{z} = a + b\bar{x} + c\bar{y} \text{ [см. формулы (9. 1. 4')].}$$

Следовательно,

$$\begin{aligned} R^2 &= \frac{\frac{1}{n} \dot{\Sigma} [b(x - \bar{x}) + c(y - \bar{y})]^2}{\sigma_z^2} = \\ &= \left\{ \frac{1}{n} \dot{\Sigma} b^2 (x - \bar{x})^2 + \frac{1}{n} \dot{\Sigma} c^2 (y - \bar{y})^2 + \right. \\ &\quad \left. + 2 \frac{1}{n} \dot{\Sigma} bc (x - \bar{x}) (y - \bar{y}) \right\} \frac{1}{\sigma_z^2}. \end{aligned}$$

Но

$$\frac{1}{n} \dot{\Sigma} b^2 (x - \bar{x})^2 = b^2 \frac{1}{n} \dot{\Sigma} (x - \bar{x})^2 = b^2 \sigma_x^2,$$

$$\frac{1}{n} \dot{\Sigma} c^2 (y - \bar{y})^2 = c^2 \frac{1}{n} \dot{\Sigma} (y - \bar{y})^2 = c^2 \sigma_y^2,$$

$$2 \frac{1}{n} \sum bc (x - \bar{x})(y - \bar{y}) = 2bc \frac{1}{n} \sum (x - \bar{x})(y - \bar{y}) = \\ = 2bcr_{xy}\sigma_x\sigma_y.$$

Откуда

$$R^2 = \frac{b^2\sigma_x^2 + c^2\sigma_y^2 + 2bcr_{xy}\sigma_x\sigma_y}{\sigma_z^2}.$$

Но [см. формулы (9.1.8)]

$$b = \frac{r_{zx} - r_{zy}r_{xy}}{1 - r_{xy}^2} \frac{\sigma_z}{\sigma_x} = \beta \frac{\sigma_z}{\sigma_x}, \\ c = \frac{r_{zy} - r_{zx}r_{xy}}{1 - r_{xy}^2} \frac{\sigma_z}{\sigma_y} = \gamma \frac{\sigma_z}{\sigma_y},$$

где β и γ — условные обозначения:

$$\left. \begin{aligned} \beta &= \frac{r_{zx} - r_{zy}r_{xy}}{1 - r_{xy}^2} \\ \gamma &= \frac{r_{zy} - r_{zx}r_{xy}}{1 - r_{xy}^2} \end{aligned} \right\}. \quad (9.2.10)$$

Откуда

$$R^2 = \beta^2 + \gamma^2 + 2\beta\gamma r_{xy}. \quad (9.2.11)$$

Представим это выражение в виде

$$R^2 = \beta(\beta + \gamma r_{xy}) + \gamma(\gamma + \beta r_{xy}).$$

Из (9.2.10) вытекает

$$\beta + \gamma r_{xy} = \frac{r_{zx} - r_{zy}r_{xy}}{1 - r_{xy}^2} + \frac{r_{zy} - r_{zx}r_{xy}}{1 - r_{xy}^2} r_{xy} = \\ = \frac{r_{zx}(1 - r_{xy}^2)}{1 - r_{xy}^2} = r_{zx};$$

аналогично

$$\gamma + \beta r_{xy} = r_{zy}.$$

Откуда

$$R^2 = \beta r_{zx} + \gamma r_{zy}. \quad (9.2.12)$$

Формула (9.2.12) имеет важное теоретическое значение, которое выяснится в следующем параграфе.

Заменяя в этой формуле β и γ их выражениями (9.2.10), получаем

$$\begin{aligned} R^2 &= \frac{r_{zx} - r_{zy}r_{xy}}{1 - r_{xy}^2} r_{zx} + \frac{r_{zy} - r_{zx}r_{xy}}{1 - r_{xy}^2} r_{zy} = \\ &= \frac{r_{zx}^2 + r_{zy}^2 - 2r_{zx}r_{zy}r_{xy}}{1 - r_{xy}^2} \end{aligned}$$

Итак,

$$R = \sqrt{\frac{r_{zx}^2 + r_{zy}^2 - 2r_{zx}r_{zy}r_{xy}}{1 - r_{xy}^2}}. \quad (9.2.13)$$

Коэффициент множественной корреляции всегда положителен и заключается в пределах от 0 до 1.

Приведем примеры вычисления R .

Пример 1. По табл. 53 рассчитаем коэффициент множественной корреляции между производительностью труда, возрастом и стажем рабочего:

$$\begin{aligned} R_{z,xy} &= \sqrt{\frac{0,20^2 + 0,40^2 - 2 \cdot 0,20 \cdot 0,40 \cdot 0,80}{1 - 0,80^2}} = \\ &= \sqrt{\frac{0,072}{0,36}} = \sqrt{0,20} = 0,447. \end{aligned}$$

Пример 2. По данным табл. 54 рассчитаем коэффициент множественной корреляции между пределом выносливости, пределом прочности на разрыв и пределом текучести стали:

$$R_{z,xy} = \sqrt{\frac{0,814^2 + 0,881^2 - 2 \cdot 0,814 \cdot 0,881 \cdot 0,896}{1 - 0,896^2}} = 0,887.$$

Пример 3. По данным табл. 55 рассчитаем коэффициент множественной корреляции между упругостью, размером замка и магнитной проницаемостью поршневых колец:

$$R_{z,xy} = \sqrt{\frac{0,440^2 + 0,305^2 - 2 \cdot 0,410 \cdot 0,305 \cdot 0,099}{1 - 0,099^2}} = 0,509.$$

Сопоставление коэффициентов множественной и парной корреляции

Между коэффициентом множественной корреляции $R_{z,xy}$ и каждым из двух коэффициентов парной корреляции r_{zx} и r_{zy} существует следующее соотношение.

Каждый из коэффициентов r_{zx} и r_{zy} не может по абсолютной величине превосходить коэффициент $R_{z.xy}$

$$\left. \begin{aligned} R_{z.xy} &\geq |r_{zx}| \\ R_{z.xy} &\geq |r_{zy}| \end{aligned} \right\} \quad (9.2.14)$$

Это понятно из общих логических соображений. Включая в уравнение регрессии дополнительный аргумент, мы более полно узнаем функцию. Схематически (при $r > 0$) это уясняется на рис. 49, где кружками обозначены факторы, определяющие значения признаков x, y, z , а площадь пересечения этих кружков символически обозначает наличие общих факторов и, следовательно, величину коэффициента корреляции.

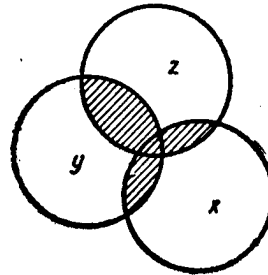


Рис. 49

Очевидно, что оба кружка x и y отсекают вместе от кружка z большую площадь, чем каждый из них в отдельности.

Приводим математическое доказательство соотношений (9.2.14)

$$\begin{aligned} R_{z.xy}^2 - r_{zx}^2 &= \frac{r_{zx}^2 + r_{zy}^2 - 2r_{zx}r_{zy}r_{xy}}{1 - r_{xy}^2} - r_{zx}^2 = \\ &= \frac{r_{zx}^2 + r_{zy}^2 - 2r_{zx}r_{zy}r_{xy} - r_{zx}^2 + r_{zx}^2 r_{zy}^2}{1 - r_{xy}^2} = \\ &= \frac{(r_{zy} - r_{zx}r_{xy})^2}{1 - r_{xy}^2} \geq 0. \end{aligned}$$

Следовательно, $R_{z.xy}^2 \geq r_{zx}^2$, что доказывает первое из соотношений (9.2.14). Аналогично доказывается второе соотношение.

Коэффициент корреляции как показатель точности оценки

Уравнение регрессии может быть использовано для расчета (оценки) значения функции по данным значениям ее аргументов. Так, в случае парной регрессии $\bar{z}_x = a + bx$ мы по данному значению x рассчитываем \bar{z}_x , который используется для оценки индивидуального значения z . В случае множественной регрессии $\bar{z}_{xy} = a + bx + cy$ по данным значениям аргументов x и y рассчитывают \bar{z}_{xy} , который служит для оценки индивидуального z . При таком способе оценки (распространения средних данных на индивидуальное значение) всегда допускают

ошибку, средний размер которой зависит от коэффициентов корреляции r_{zx} и $R_{z, xy}$.

Рассмотрим вопрос о точности оценки на конкретном примере. Поступила партия поршневых колец, удовлетворяющая стандартам упругости (z): $\bar{z} = 1410$, $\sigma_z = 83$. Из партии взято наудачу одно кольцо. Требуется дать оценку упругости для этого кольца.

Очевидно, мы можем судить об упругости взятого кольца только по средним данным партии (стандарту). Однако при этом мы допустим ошибку, средний размер которой выражается стандартом σ_z . Следовательно, оценка может быть дана в виде

$$z = \bar{z} \pm \sigma_z = 1410 \pm 83.$$

Пусть теперь для взятого кольца нам известен размер замка $x = 19,0$. Для определения упругости этого кольца мы можем воспользоваться уравнением регрессии $\bar{z}_x = 787,1 + 34,666 x$, если это уравнение указано в системе стандартов. При $x = 19,0$ получаем $\bar{z}_x = 1445,7$. Это значение будет служить оценкой упругости взятого кольца. Однако при таких оценках допускается ошибка, средний размер которой равен σ_{iT} .

Показатель σ_{iT} может быть выражен через коэффициент корреляции r_{zx} . В самом деле, из формулы

$$r_{zx}^2 = 1 - \frac{\sigma_{iT}^2}{\sigma_z^2}$$

получаем

$$\sigma_{iT} = \sigma_z \sqrt{1 - r_{zx}^2}. \quad (9. 2. 15)$$

В нашем случае $r_{zx} = 0,44$, отсюда

$$\sigma_{iT} = 83 \sqrt{1 - 0,44^2} = 74,7.$$

Следовательно, окончательная оценка упругости взятого кольца может быть дана в виде $z = \bar{z}_x \pm \sigma_{iT} = 1445,7 \pm 74,7$. Оценка стала более гибкой, ибо она меняется в зависимости от размера замка в кольце, и точность этой оценки увеличилась, так как

$$\sigma_{iT} < \sigma_z.$$

Точность оценки увеличивается тем сильнее, чем больше коэффициент корреляции между оцениваемым признаком (z) и его аргументом (x).

Пусть теперь для взятого кольца определена, кроме размера замка, также и магнитная проницаемость $y = 14,0$.

Оценку упругости можно дать на основе уравнения множественной регрессии

$$\bar{z}_{xy} = 762,12 + 32,70x + 8,07y,$$

если это уравнение указано в системе стандартов.

Подставляя в уравнение $x=19,0$ и $y=14,0$, получаем $\bar{z}_{xy} = 1496$. Это значение используем в качестве новой оценки упругости взятого кольца. Однако при этой оценке допускается ошибка, средний размер которой равен $\bar{\sigma}_{ijT}$.

$\bar{\sigma}_{ijT}$ рассчитывается по формуле

$$\bar{\sigma}_{ijT} = \sigma_z \sqrt{1 - R_{z, xy}^2}, \quad (9.2.16)$$

которая выводится из формулы

$$R^2 = 1 - \frac{\bar{\sigma}_{ijT}^2}{\sigma_z^2}.$$

В нашем случае

$$\bar{\sigma}_{ijT} = 83 \sqrt{1 - 0,51^2} = 71,4.$$

Оценку упругости взятого кольца даем в виде

$$z = \bar{z}_{xy} \pm \bar{\sigma}_{ijT} = 1496 \pm 71,4.$$

Точность новой оценки выше, чем точность предыдущих оценок, так как $\bar{\sigma}_{ijT} < \bar{\sigma}_{iT}$. Увеличение точности оценки тем сильнее, чем больше отличается коэффициент множественной корреляции $R_{z, xy}$ от коэффициента парной корреляции r_{zx} .

ГЛАВА X

КОРРЕЛЯЦИЯ МНОГИХ ПЕРЕМЕННЫХ

При изучении корреляции многих переменных ограничимся наиболее важным для приложений случаем прямолинейной корреляции и рассмотрим также некоторые способы приведения корреляции к прямолинейной.

§ 1. УРАВНЕНИЕ МНОЖЕСТВЕННОЙ РЕГРЕССИИ ДЛЯ p ПЕРЕМЕННЫХ

Уравнение прямолинейной регрессии p переменных имеет следующий вид:

$$\bar{x}_{1,2,3\dots p} = b_1 + b_2x_2 + b_3x_3 + \dots + b_px_p, \quad (10. 1. 1)$$

где $\bar{x}_{1,2,3\dots p}$ — среднее значение функции x_1 , соответствующее заданным значениям аргументов x_2, x_3, \dots, x_p .

При расчете параметров этого уравнения используем метод, изложенный в предыдущей главе, в несколько измененном виде: все переменные и соотношения между ними будем выражать в стандартизованном масштабе. В этом масштабе за начало отсчета для каждой переменной принимается значение средней арифметической, а за единицу измерения — величина среднего квадратического отклонения. В стандартизованном масштабе упрощаются линейные соотношения между переменными, что особенно важно при изучении зависимости многих переменных. Изложение иллюстрируем примером.

Коэрцитивная сила, т. е. величина магнитного напряжения, которая требуется для устранения остаточного магнетизма, является важнейшей характеристикой электротехнического железа типа «Армко». Слаботочная промышленность и заводы, изготавливающие специальные приборы, заинтересованы в снижении коэрцитивной силы с целью улучшения магнитных свойств железа «Армко».

Для выяснения факторов, влияющих на величину коэрцитивной силы по предложению Центральной лаборатории завода «Серп и молот» Н. К. Лебедевой, Р. В. Катковой и Л. И. Крыловой была проведена статистическая обработка лабораторных данных испытаний текущих плавок железа «Армко».

Приводим в виде примера некоторые результаты этой работы.

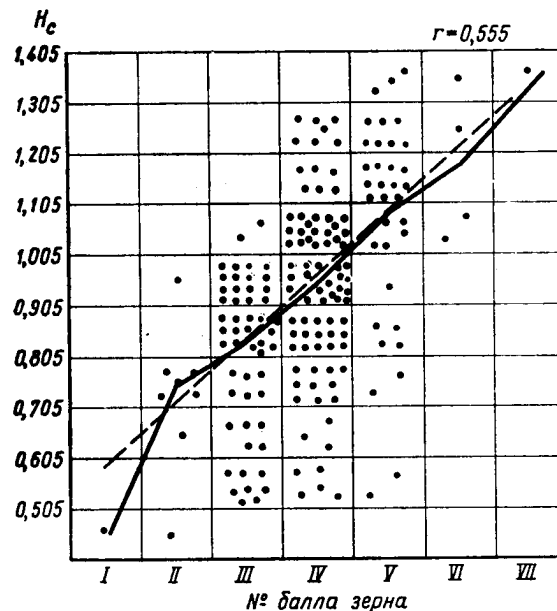


Рис. 50. Зависимость между коэрцитивной силой и баллом зерна. Поле корреляции и линии регрессии

Прежде всего изучалась корреляционная зависимость между показателями коэрцитивной силы и величиной зерна микроструктуры.

Приводим поле корреляции и линии регрессии коэрцитивной силы по баллу зерна (чем больше балл, тем мельче зерно) для 190 образцов железа «Армко».

Приведенные данные указывают на наличие положительной зависимости между коэрцитивной силой (x_1) и баллом зерна (x_2), которая имеет отчетливо выраженный прямолинейный характер.

Уравнение теоретической линии регрессии x_1 по x_2 (в натуральном масштабе):

$$\bar{x}_{1.2} = 0,38 + 0,128x_2.$$

Коэффициент корреляции $r_{12} = 0,555$.

Величина коэффициента корреляции (не слишком близкая к 0 и к 1) подтверждает наличие зависимости и одновременно указывает на то, что наряду с величиной зерна имеются и другие факторы, не зависящие или частично зависящие от величины зерна, которые также влияют на величину коэрцитивной силы. К таким факторам принадлежит химический состав железа «Армко».

Влияние химического состава на коэрцитивную силу проверялось для всех химических элементов, процентное содержание которых определялось для каждого образца. Приводим два поля корреляции, где корреляционная зависимость оказалась наиболее отчетливо выраженной.

Рассчитанные по этим данным линии регрессии и коэффициенты корреляции подтверждают наличие зависимости между коэрцитивной силой и содержанием углерода и серы. Эти зависимости также имеют прямолинейный характер

$$\bar{x}_{1.3} = 0,70 + 15,43x_3, \quad r_{13} = 0,301,$$

$$\bar{x}_{1.4} = 0,665 + 9,25x_4, \quad r_{14} = 0,372.$$

Задача заключалась в том, чтобы изучить совместное влияние трех аргументов — величины зерна, содержания углерода, содержания серы — на величину коэрцитивной силы и выявить влияние каждого из них при постоянных значениях других.

Решение этой задачи, как известно, затрудняется тем обстоятельством, что аргументы, влияние которых исследуется, могут находиться между собой в известной корреляции, и, следовательно, каждая из парных зависимостей осложняется неявным влиянием других аргументов. Рассмотрим поэтому зависимости между аргументами.

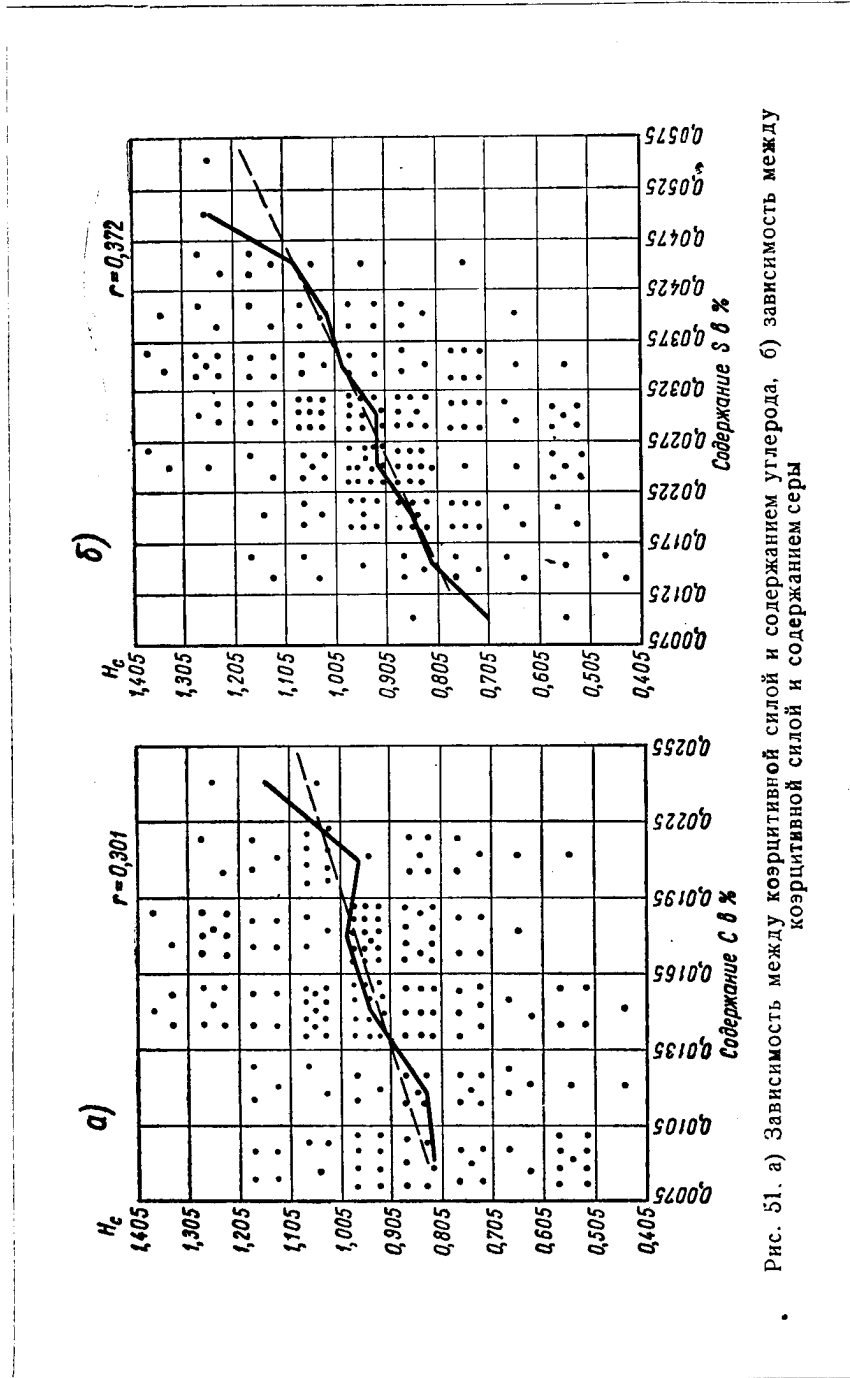
На рис. 52 приведены корреляционные поля и линии регрессии для зависимостей: балл зерна — содержание углерода, балл зерна — содержание серы, содержание углерода — содержание серы, и указаны коэффициенты корреляции. Последняя зависимость оказалась наиболее значительной ($r_{34} = 0,334$). Она, по видимому, имеет несколько криволинейный характер и лишь приближенно может быть выражена уравнением прямой линии

$$\bar{x}_{3.4} = 0,0104 + 0,162x_4.$$

Первые две зависимости малозначительны ($r_{23} = 0,029$, $r_{24} = 0,179$) с прямолинейными регрессиями

$$\bar{x}_{2.3} = 3,87 + 6,48x_3,$$

$$\bar{x}_{2.4} = 3,32 + 19,39x_4.$$



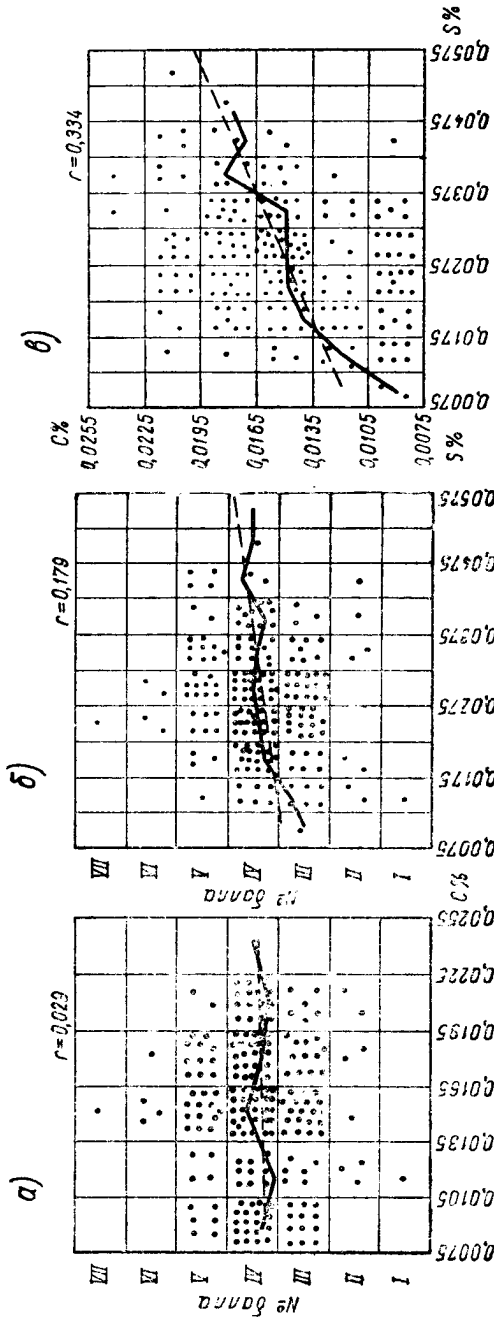


Рис. 52. а) Зависимость между баллом зерна и содержанием углерода, б) зависимость между баллом зерна и содержанием серы, в) зависимость между содержанием углерода и содержанием серы

Коэффициенты корреляции для всех зависимостей помещены в сводной табл. 56.

Таблица 56

Таблица коэффициентов корреляции

	Коэрцитивная сила x_1	Балл зерна x_2	Содержание углерода x_3	Содержание серы x_4
Коэрцитивная сила	*	0,555	0,301	0,372
Балл зерна	0,555	*	0,029	0,179
Содержание углерода	0,301	0,029	*	0,334
Содержание серы	0,372	0,179	0,334	*

Переходим к расчету уравнения множественной регрессии. Прежде всего возникает вопрос, допустимо ли здесь предположение о линейности множественной регрессии. Поскольку металлургия в настоящее время не дает еще указаний о форме такой зависимости, мы вынуждены ограничиться анализом опытных данных.

Основные парные зависимости коэрцитивной силы от величины зерна, содержания углерода, содержания серы носят отчетливо прямолинейный характер; взаимосвязи между этими факторами также можно считать прямолинейными, за исключением зависимости между содержанием углерода и содержанием серы, для которой линия регрессии имеет слегка искривленный характер.

Далее следует заметить, что наши аргументы изменяются в основном в сравнительно узких пределах, разрешаемых техническими условиями. Из математического анализа известно, что в достаточно малых промежутках изменения аргументов криволинейные зависимости могут быть с достаточной точностью воспроизведены зависимостями прямолинейными. Поэтому можно считать, что предположение о приближенной прямолинейности множественной регрессии (коэрцитивная сила — балл зерна — содержание углерода — содержание серы) имеет достаточные основания.

Выразим предварительно все переменные и зависимости между ними в стандартизованном масштабе. Формула перевода

$$t_x = \frac{x - \bar{x}}{\sigma_x}, \quad (10. 1. 2)$$

где x — значение признака в натуральном масштабе;

t_x — соответствующее значение в стандартизованном масштабе.

Аналогично определяются t_y , t_z .

Функции

$$f = \frac{1}{n} \Sigma [t_1 - \bar{t}_{1.23\dots p}]^2 = \frac{1}{n} \Sigma [t_1 - \beta_2 t_2 - \beta_3 t_3 - \dots - \beta_p t_p]^2$$

и заметим, что

$$\frac{1}{n} \Sigma t_i^2 = 1 \quad \text{и} \quad \frac{1}{n} \Sigma t_i t_j = r_{ij}.$$

Легко видеть, что свободный член в стандартизованном уравнении прямолинейной множественной регрессии отсутствует. В самом деле, если написать уравнение множественной регрессии со свободным членом

$$\bar{t}_{1.23\dots p} = \beta_1 + \beta_2 t_2 + \beta_3 t_3 + \dots + \beta_p t_p,$$

то в силу условия (10.1.6) к системе (10.1.7) добавится уравнение

$$\Sigma [t_1 - \beta_1 - \beta_2 t_2 - \beta_3 t_3 - \dots - \beta_p t_p] = 0.$$

Откуда $\beta_1 = 0$.

Решая систему (10.1.7), мы получим искомые значения коэффициентов $\beta_2, \beta_3, \dots, \beta_p$.

Решение может быть записано в виде

$$\beta_i = \frac{\Delta_i}{\Delta}, \quad i = 2, 3, \dots, p,$$

где

$$\Delta = \begin{vmatrix} 1 & r_{32} & r_{42} & \dots & r_{p2} \\ r_{23} & 1 & r_{43} & \dots & r_{p3} \\ \dots & \dots & \dots & \dots & \dots \\ r_{2p} & r_{3p} & r_{4p} & \dots & 1 \end{vmatrix} \quad (10.1.8)$$

определитель системы (10.1.7), Δ_i получается из определителя Δ заменой в нем соответствующего столбца столбцом свободных членов системы (10.1.7).

Так, для нашего примера система нормальных уравнений

$$\left. \begin{aligned} 0,555 &= \beta_2 + \beta_3 \cdot 0,029 + \beta_4 \cdot 0,179 \\ 0,301 &= \beta_2 \cdot 0,029 + \beta_3 + \beta_4 \cdot 0,334 \\ 0,372 &= \beta_2 \cdot 0,179 + \beta_3 \cdot 0,334 + \beta_4 \end{aligned} \right\}$$

Решение системы

$$\beta_2 = \frac{\Delta_2}{\Delta}, \quad \beta_3 = \frac{\Delta_3}{\Delta}, \quad \beta_4 = \frac{\Delta_4}{\Delta},$$

$$\begin{aligned}
 \text{где } \Delta &= \begin{vmatrix} 1 & 0,029 & 0,179 \\ 0,029 & 1 & 0,334 \\ 0,179 & 0,334 & 1 \end{vmatrix} = \\
 &= \begin{vmatrix} 1 & 0,334 \\ 0,334 & 1 \end{vmatrix} - 0,029 \begin{vmatrix} 0,029 & 0,334 \\ 0,179 & 1 \end{vmatrix} + \\
 &+ 0,179 \begin{vmatrix} 0,029 & 1 \\ 0,179 & 0,334 \end{vmatrix} = (1 - 0,334^2) - \\
 &- 0,029(0,029 - 0,334 \cdot 0,179) + 0,179(0,029 \cdot 0,334 - 0,179) = \\
 &= (1 - 0,112) - 0,029 \times (0,029 - 0,060) + 0,179(0,010 - \\
 &- 0,179) = 0,888 + 0,029 \times 0,031 - 0,179 \cdot 0,169 = 0,888 + \\
 &+ 0,001 - 0,030 = 0,859;
 \end{aligned}$$

$$\begin{aligned}
 \Delta_2 &= \begin{vmatrix} 0,555 & 0,029 & 0,179 \\ 0,301 & 1 & 0,334 \\ 0,372 & 0,334 & 1 \end{vmatrix} = 0,555 \begin{vmatrix} 1 & 0,334 \\ 0,334 & 1 \end{vmatrix} - \\
 &- 0,029 \begin{vmatrix} 0,301 & 0,334 \\ 0,372 & 1 \end{vmatrix} + 0,179 \begin{vmatrix} 0,301 & 1 \\ 0,372 & 0,334 \end{vmatrix} = \\
 &= 0,555(1 - 0,334^2) - 0,029(0,301 - 0,334 \cdot 0,372) + \\
 &+ 0,179(0,301 \cdot 0,334 - 0,372) = 0,440.
 \end{aligned}$$

Аналогичным образом рассчитываем

$$\begin{aligned}
 \Delta_3 &= \begin{vmatrix} 1 & 0,555 & 0,179 \\ 0,029 & 0,301 & 0,334 \\ 0,179 & 0,372 & 1 \end{vmatrix} = 0,186; \\
 \Delta_4 &= \begin{vmatrix} 1 & 0,029 & 0,555 \\ 0,029 & 1 & 0,301 \\ 0,179 & 0,334 & 0,372 \end{vmatrix} = 0,179.
 \end{aligned}$$

Отсюда

$$\begin{aligned}
 \beta_2 &= \frac{0,440}{0,859} = 0,512, \\
 \beta_3 &= \frac{0,186}{0,859} = 0,217, \\
 \beta_4 &= \frac{0,179}{0,859} = 0,208.
 \end{aligned}$$

Проверяем решение подстановкой в систему (10. 1. 7)

$$\begin{aligned}
 0,555 &= 0,512 + 0,217 \cdot 0,029 + 0,208 \cdot 0,179 = \\
 &= 0,512 + 0,006 + 0,037 = 0,555;
 \end{aligned}$$

$$\begin{aligned}
 0,301 &= 0,512 \cdot 0,029 + 0,217 + 0,208 \cdot 0,334 = \\
 &= 0,0148 + 0,217 + 0,0695 = 0,301;
 \end{aligned}$$

$$\begin{aligned} 0,372 &= 0,512 \cdot 0,179 + 0,217 \cdot 0,334 + 0,208 = \\ &= 0,092 + 0,072 + 0,208 = 0,372. \end{aligned}$$

Система удовлетворяется.

Формулы (10.1.8) полезны в теоретических исследованиях. Однако вычисление определителей не является простейшим путем численного решения системы (10.1.7).

Более простым является способ последовательного исключения неизвестных, который мы показываем на том же примере:

$$\begin{aligned} \text{I. (1)} \quad 0,555 &= \beta_2 + \beta_3 \cdot 0,029 + \beta_4 \cdot 0,179; \\ (2) \quad 0,301 &= \beta_2 \cdot 0,029 + \beta_3 + \beta_4 \cdot 0,334; \\ (3) \quad 0,372 &= \beta_2 \cdot 0,179 + \beta_3 \cdot 0,334 + \beta_4. \end{aligned}$$

II. (1), (2), (3) делим на коэффициенты при β_2 , получаем

$$\begin{aligned} (1)' \quad 0,555 &= \beta_2 + \beta_3 \cdot 0,029 + \beta_4 \cdot 0,179; \\ (2)' \quad 10,380 &= \beta_2 + \beta_3 \cdot 34,483 + \beta_4 \cdot 11,517; \\ (3)' \quad 2,078 &= \beta_2 + \beta_3 \cdot 1,866 + \beta_4 \cdot 5,587. \end{aligned}$$

III. Вычитаем из уравнения (2)' уравнение (1)' и из уравнения (3)' уравнение (1)'

$$\begin{aligned} (2)' - (1)' \quad 9,825 &= \beta_3 \cdot 34,454 + \beta_4 \cdot 11,338; \\ (3)' - (1)' \quad 1,523 &= \beta_3 \cdot 1,837 + \beta_4 \cdot 5,408. \end{aligned}$$

IV. Делим предыдущие уравнения на коэффициенты при β_3 , получаем

$$\begin{aligned} (1)'' \quad 0,285 &= \beta_3 + \beta_4 \cdot 0,329; \\ (2)'' \quad 0,829 &= \beta_3 + \beta_4 \cdot 2,944. \end{aligned}$$

V. Вычитаем из уравнения (2)'' уравнение (1)''

$$\begin{aligned} (2)'' - (1)'' \quad 0,544 &= \beta_4 \cdot 2,615; \\ \beta_4 &= \frac{0,544}{2,615} = 0,208. \end{aligned}$$

Подстановкой β_4 в (1)'' получаем

$$\beta_3 = 0,285 - 0,208 \cdot 0,329 = 0,217.$$

Подстановкой β_4 и β_3 в (1)' получаем

$$\beta_2 = 0,555 - 0,217 \cdot 0,029 - 0,208 \cdot 0,179 = 0,512.$$

Решение проверяем подстановкой в систему (10.1.7).

При большом количестве аргументов целесообразно к уравнению множественной регрессии, охватывающему все аргументы, подойти не сразу, а постепенно увеличивая число включаемых в уравнение аргументов. При этом надлежит отправляться от уравнения парной регрессии по ведущему аргументу и последовательно присоединять к этому уравнению новые аргументы в порядке их значимости. Следя за изменением коэффициентов β_i после каждого присоединения нового аргумента, мы сможем лучше ориентироваться в происхождении окончательных результатов и более правильно истолковать их смысл.

Метод последовательного включения аргументов мы используем в разделе IV, в примере исследования влияния химического состава на ударную вязкость стали 30ХГСА.

Коэффициенты $\beta_2, \beta_3, \dots, \beta_p$ показывают, на какую часть сигмы изменилось бы среднее значение функции, если бы соответствующий аргумент увеличился на сигму, а прочие аргументы остались без изменения. Они выражают, таким образом, скорость изменения среднего значения функции по каждому из аргументов при постоянных значениях прочих аргументов. Поскольку все переменные выражены при этом в сравнимых единицах измерения (сигмах), то коэффициенты $\beta_2, \beta_3, \dots, \beta_p$ показывают сравнительную силу влияния изменения каждой переменной на изменение функционального признака.

В нашем примере уравнение множественной регрессии в стандартизованном масштабе оказалось

$$\bar{t}_{1.234} = 0,512t_2 + 0,217t_3 + 0,208t_4.$$

Следовательно, наибольшее влияние на коэрцитивную силу имело изменение величины зерна. С уменьшением балла зерна на одну сигму (при постоянном содержании С и S) коэрцитивная сила уменьшается в среднем примерно на 0,5 сигмы. Значительно меньшее и приблизительно равное влияние на изменение коэрцитивной силы имело содержание С и S. С изменением каждого из этих факторов на одну сигму (при прочих постоянных условиях) коэрцитивная сила изменяется лишь на 0,2 сигмы. Для практического использования эти зависимости должны быть выражены в натуральном масштабе.

Уравнение множественной регрессии в натуральном масштабе имеет вид

$$\bar{x}_{1.23\dots p} = b_1 + b_2x_2 + b_3x_3 + \dots + b_px_p.$$

Коэффициенты $b_1, b_2, b_3, \dots, b_p$ этого уравнения получаются по формулам

$$\left. \begin{aligned} b_i &= \beta_i \frac{\sigma_1}{\sigma_i}, \quad i = 2, 3, \dots, p \\ b_1 &= \bar{x}_1 - b_2\bar{x}_2 - b_3\bar{x}_3 - \dots - b_p\bar{x}_p \end{aligned} \right\}, \quad (10.1.9)$$

где \bar{x}_i и σ_i — соответственно средние арифметические и средние

квадратические отклонения соответствующих признаков. Те же значения коэффициентов b_i можно было бы получить и непосредственно из условия

$$\Sigma (x_1 - b_1 - b_2 x_2 - b_3 x_3 - \dots - b_p x_p)^2 = \min, \quad (10.1.10)$$

так как легко показать, что условия (10.1.6) и (10.1.10) математически эквивалентны.

Приводим таблицу \bar{x}_i и σ_i , рассчитанных по данным нашего примера.

Таблица 57

Признаки	Показатели	\bar{x}_i	σ_i
1. Коэффициентная сила . . .		0,929	0,205
2. Балл зерна		3,87	0,893
3. Содержание углерода . .		0,0150	0,0040
4. Содержание серы		0,0285	0,0082

Отсюда

$$b_2 = \beta_2 \frac{\sigma_1}{\sigma_2} = 0,512 \cdot \frac{0,205}{0,893} = 0,118;$$

$$b_3 = \beta_3 \frac{\sigma_1}{\sigma_3} = 0,217 \cdot \frac{0,205}{0,0040} = 11,12;$$

$$b_4 = \beta_4 \frac{\sigma_1}{\sigma_4} = 0,208 \cdot \frac{0,205}{0,0082} = 5,200;$$

$$b_1 = \bar{x}_1 - b_2 \bar{x}_2 - b_3 \bar{x}_3 - b_4 \bar{x}_4 = 0,929 - 0,118 \cdot 3,87 - \\ - 11,12 \cdot 0,015 - 5,2 \cdot 0,0285 = 0,157$$

и уравнение регрессии в натуральном масштабе

$$\bar{x}_{1.234} = 0,157 + 0,118x_2 + 11,12x_3 + 5,20x_4.$$

§ 2. УРАВНЕНИЕ ЧИСТОЙ РЕГРЕССИИ

Уравнение чистой регрессии отвечает на вопрос: как изменилось бы среднее значение функции x_1 с изменением ряда аргументов x_2, x_3, \dots, x_k , если бы другие учитываемые аргументы $x_{k+1}, x_{k+2}, \dots, x_p$ были закреплены на среднем уровне?

Уравнение чистой регрессии получается из уравнения множественной регрессии, если в последнем на место значений $x_{k+1}, x_{k+2}, \dots, x_p$ подставить средние $\bar{x}_{k+1}, \bar{x}_{k+2}, \dots, \bar{x}_p$.

Очевидно, что эта операция отражается лишь на свободном члене уравнения множественной линейной регрессии. Свободный член уравнения чистой линейной регрессии можно рассчитать по формуле

$$\bar{x}_{1(k+1, k+2, \dots, p)} = b_1 + b_{k+1}\bar{x}_{k+1} + b_{k+2}\bar{x}_{k+2} + \dots + b_p\bar{x}_p, \quad (10.2.1)$$

а уравнение чистой регрессии пишется следующим образом

$$\bar{x}_{1.23\dots k(k+1, k+2, \dots, p)} = b_1(k+1, k+2, \dots, p) + b_2x_2 + b_3x_3 + \dots + b_kx_k, \quad (10.2.2)$$

где b_2, b_3, \dots, b_k — коэффициенты, непосредственно взятые из уравнения множественной регрессии x_1 по переменным x_2, x_3, \dots, x_p .

Проще всего уравнение чистой регрессии записать в стандартизованном масштабе; для этого в уравнении множественной регрессии нужно просто опустить члены с закрепляемыми переменными:

$$\bar{t}_{1.23\dots k(k+1, k+2, \dots, p)} = \beta_2t_2 + \beta_3t_3 + \dots + \beta_kt_k. \quad (10.2.3)$$

Рассчитаем по данным нашего примера чистую линию регрессии коэрцитивной силы по баллу зерна при закрепленном содержании углерода и серы.

В стандартизованном масштабе

$$\bar{t}_{1.2(34)} = 0,512t_2.$$

В натуральном масштабе

$$\begin{aligned} \bar{x}_{1.2(34)} &= (0,157 + 11,12 \cdot 0,0150 + 5,20 \cdot 0,0285) + 0,118x_2 = \\ &= 0,472 + 0,118x_2. \end{aligned}$$

Сравним полученное уравнение чистой регрессии с уравнением простой регрессии для тех же признаков

$$\bar{x}_{1.2(34)} = 0,472 + 0,118x_2$$

— уравнение чистой регрессии,

$$\bar{x}_{1.2} = 0,430 + 0,127x_2$$

— уравнение простой регрессии.

В первом уравнении по сравнению со вторым уменьшился коэффициент регрессии и увеличился свободный член. Этому явлению можно дать следующее объяснение. В уравнении обыкновенной регрессии влияние аргумента x_2 на функцию x_1 было осложнено неявным влиянием корреляционно связанных с x_2 аргументов x_3 и x_4 .

Для плавков с большим размером зерна мы имели, в силу этой корреляции, и несколько большее содержание С и S. Поэтому в обыкновенной регрессии влияние размера зерна на коэрцитив-

ную силу несколько преувеличено за счет сопутствующего влияния C и S . В уравнении чистой регрессии мы освободились от этого влияния, закрепив содержание C и S на среднем уровне. В результате закрепления увеличился постоянный член уравнения и уменьшился коэффициент регрессии.

Уравнение чистой регрессии имеет смысл вычислять тогда, когда исключаемые переменные не сильно связаны с переменными, остающимися в уравнении. Когда же эта зависимость велика, целесообразно рассчитывать так называемые уравнения *частных регрессий*, где исключаемые переменные закрепляются не на средних, а на других уровнях, выбранных в соответствии с интересующими нас участками изменения основных переменных, оставшихся в уравнении. В этом случае уравнение чистой регрессии может утратить свое реальное содержание.

§ 3. КОЭФФИЦИЕНТ МНОЖЕСТВЕННОЙ КОРРЕЛЯЦИИ

Теоретическое множественное корреляционное отношение $\gamma_{1.23\dots pT}$ определяется по формуле

$$\gamma_{1.23\dots pT}^2 = \frac{\overline{\delta_{1.23\dots pT}^2}}{\sigma_1^2}, \quad (10.3.1)$$

где $\overline{\delta_{1.23\dots pT}^2} = \frac{1}{n} \Sigma [\bar{x}_{1.23\dots p} - \bar{x}_1]^2$; $\sigma_1^2 = \frac{1}{n} \Sigma (x_1 - \bar{x}_1)^2$.

В случае прямолинейной корреляции это отношение называется коэффициентом множественной корреляции $R_{1.23\dots p}$.

Таким образом, коэффициент множественной корреляции $R_{1.23\dots p}$ характеризует тесноту связи переменной x_1 с совокупностью переменных x_2, x_3, \dots, x_p в случае прямолинейной корреляции. Можно показать, что $R_{1.23\dots p}$ выражается формулой

$$R_{1.23\dots p} = \sqrt{\beta_2 r_{12} + \beta_3 r_{13} + \dots + \beta_p r_{1p}}, \quad (10.3.2)$$

которая аналогична формуле (9.2.12) для случая трех переменных. Приводим вывод формулы (10.3.2).

Выражая переменные в стандартизованном масштабе, получим

$$R_{1.23\dots p}^2 = \frac{1}{n} \Sigma [\bar{t}_{1.23\dots p}]^2, \quad (10.3.3)$$

где $\bar{t}_{1.23\dots p} = \beta_2 t_2 + \beta_3 t_3 + \dots + \beta_p t_p$.

Формулу (10.3.3) запишем в виде

$$\begin{aligned} R_{1.23\dots p}^2 &= \frac{1}{n} \Sigma \bar{t}_{1.23\dots p} (\beta_2 t_2 + \beta_3 t_3 + \dots + \beta_p t_p) = \beta_2 \frac{1}{n} \Sigma \bar{t}_{1.23\dots p} t_2 + \\ &+ \beta_3 \frac{1}{n} \Sigma \bar{t}_{1.23\dots p} t_3 + \dots + \beta_p \frac{1}{n} \Sigma \bar{t}_{1.23\dots p} t_p. \end{aligned} \quad (10.3.4)$$

Проведем по этой формуле расчет $R_{1,234}$ для нашего примера:

$$\begin{aligned} \Delta^* &= \begin{vmatrix} 0,555 & 0,301 & 0,372 & 0 \\ 1 & 0,029 & 0,179 & 0,555 \\ 0,029 & 1 & 0,334 & 0,301 \\ 0,179 & 0,334 & 1 & 0,372 \end{vmatrix} = \\ &= 0,555 \begin{vmatrix} 0,029 & 0,179 & 0,555 \\ 1 & 0,334 & 0,301 \\ 0,334 & 1 & 0,372 \end{vmatrix} - 0,301 \begin{vmatrix} 1 & 0,179 & 0,555 \\ 0,029 & 0,334 & 0,301 \\ 0,179 & 1 & 0,372 \end{vmatrix} + \\ &+ 0,372 \begin{vmatrix} 1 & 0,029 & 0,555 \\ 0,029 & 1 & 0,301 \\ 0,179 & 0,334 & 0,372 \end{vmatrix} = 0,555 \cdot 0,440 - \\ &- 0,301 \cdot (-0,186) + 0,372 \cdot 0,179 = 0,367, \\ \Delta &= \begin{vmatrix} 1 & 0,029 & 0,179 \\ 0,029 & 1 & 0,334 \\ 0,179 & 0,334 & 1 \end{vmatrix} = 0,859, \end{aligned}$$

$$R_{1,234} = \sqrt{\frac{0,367}{0,859}} = \sqrt{0,427} = 0,653,$$

что совпадает с полученным ранее значением.

§ 4. ПРИВЕДЕНИЕ МНОЖЕСТВЕННОЙ КОРРЕЛЯЦИИ К ЛИНЕЙНОМУ ВИДУ ПУТЕМ ФУНКЦИОНАЛЬНОГО ПРЕОБРАЗОВАНИЯ ПЕРЕМЕННЫХ

В ряде случаев множественную корреляцию можно привести к линейному виду путем соответствующего преобразования или надлежащего выбора переменных. Часто используется логарифмическое преобразование, при котором значения коррелируемых переменных заменяются их логарифмами. Это преобразование ведет к цели, если признаки связаны соотношением вида

$$x_1 = u x_2^v \cdot x_3^w \cdot x_4^z \dots, \quad (10.4.1)$$

где u, v, w, z, \dots — дополнительные факторы схемы, корреляционно не зависящие от $\lg x_2, \lg x_3, \lg x_4, \dots$

Тогда после логарифмирования формулы (10.4.1) и замены переменных $\lg x_1 = y_1, \lg x_2 = y_2, \lg x_3 = y_3, \lg x_4 = y_4, \lg u = a_1$ получаем

$$y_1 = a_1 + v y_2 + w y_3 + z y_4 + \dots, \quad (10.4.2)$$

где u, v, w, z, \dots корреляционно не зависят от y_2, y_3, y_4, \dots

Осередняя уравнение (10. 4. 2), получаем

$$\bar{y}_{1.234} = a + by_2 + cy_3 + dy_4 + \dots, \quad (10. 4. 3)$$

где a, b, c, d, \dots — постоянные, соответствующие средним значениям факторов u, v, w, z, \dots и определяемые по способу наименьших квадратов. Практически этот случай распознается, если в двойных логарифмических координатах парные зависимости признаков $x_1, x_2, x_3, x_4, \dots$ имеют прямолинейный характер или если линии регрессии между логарифмами этих признаков прямолинейны.

Более общая схема действия факторов, выражаемая уравнением

$$x_1 = ux_2^v w^{x_3} \dots, \quad (10. 4. 4)$$

где дополнительные факторы схемы u, v, w, \dots статистически не зависят от x_2, x_3, \dots , также приводится к линейной схеме (А — М — Н) путем логарифмирования уравнения (10. 4. 4) и замены переменных

$$\lg x_1 = y_1, \quad \lg x_2 = y_2, \quad x_3 = y_3, \quad \lg u = u_1, \quad v = v_1, \quad \lg w = w_1, \dots$$

$$y_1 = u_1 + v_1 y_2 + w_1 y_3 + \dots$$

Практически этот случай распознается, если парные зависимости между x_1, x_2, x_3, \dots в координатных сетках с логарифмическим масштабом по осям x_1 и x_2 и натуральным по оси x_3 имеют прямолинейный характер.

Пример. Установим по данным табл. 2 приложения I зависимость между трудоемкостью изготовления горячего штампа, весом штампа и числом его поверхностей. Используем для решения задачи логарифмический масштаб выражения признаков, уже оправдавший себя при изучении зависимости между весом и числом поверхностей штампа (см. § 3, гл. VI).

Рассмотрим поля корреляции и линии регрессии между логарифмами этих признаков (рис. 36 и 53). Все три линии регрессии имеют в порядке грубого приближения прямолинейный характер. Следовательно, можно предположить в порядке такого приближения схему действия факторов (10. 4. 1) и искать по методу наименьших квадратов значения параметров a, b, c в уравнении

$$\bar{y}_{1.23} = a + by_2 + cy_3.$$

В стандартизованном масштабе получаем

$$\bar{t}_{1.23} = 0,477t_2 + 0,322t_3 \quad (10. 4. 5)$$

с коэффициентом множественной регрессии $R_{1.23} = 0,647$.

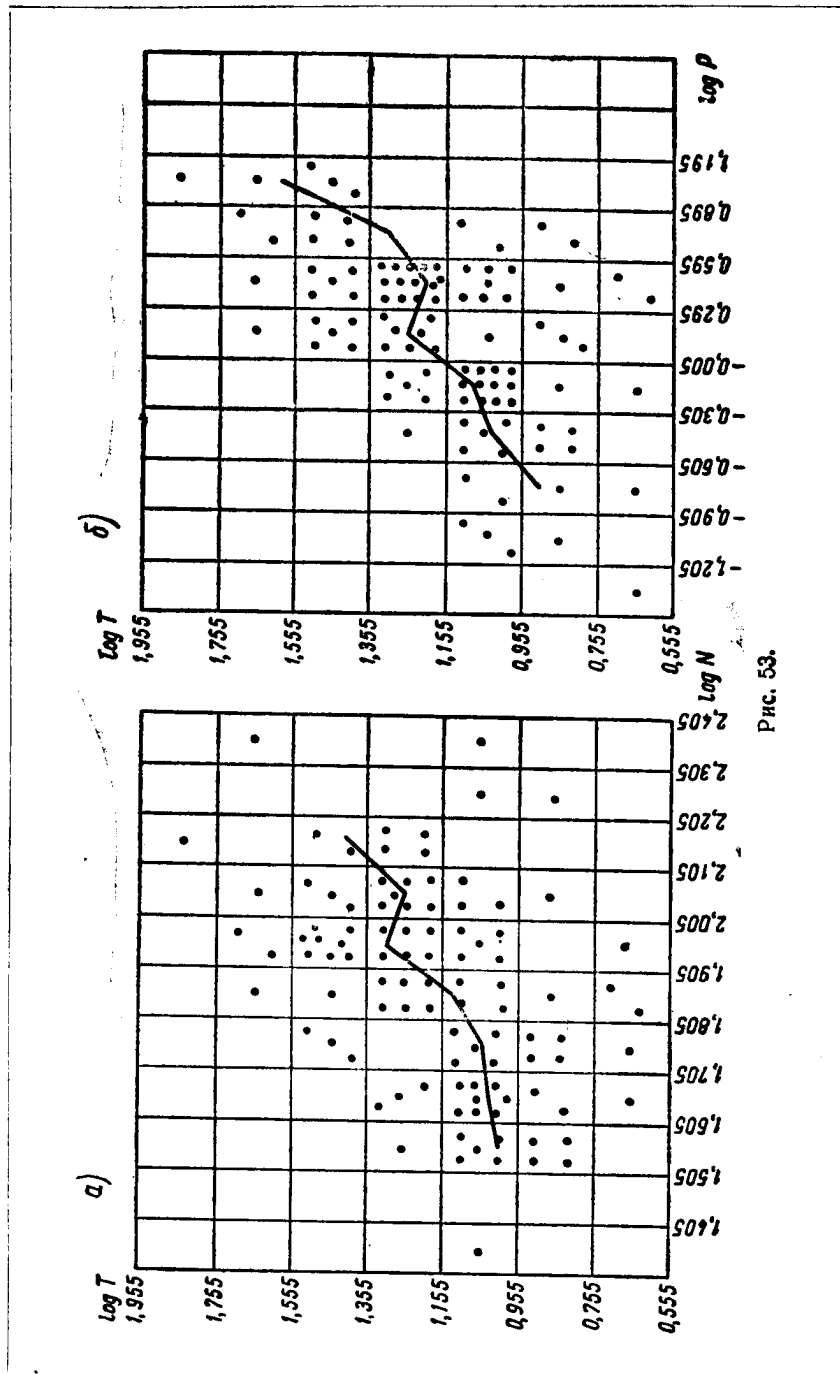


Рис. 53.

В натуральном масштабе

$$\bar{y}_{1.23} = 0,388 + 0,231y_2 + 0,402y_3 \quad (10.4.6)$$

или

$$\tilde{x}_{1.23} = 0,388x_2^{0,231}x_3^{0,402}. \quad (10.4.7)$$

Из уравнения (10.4.5) следует вывод, что вес штампа сильнее влияет на трудоемкость его изготовления, нежели число поверхностей.

Уравнение (10.4.7) может быть использовано для приближенной оценки трудоемкости изготовления горячего штампа рассматриваемого вида по данному весу и данному числу поверхностей. Заметим, что вместо соотношения (10.4.7) можно найти соотношение между трудоемкостью изготовления штампа, его весом и показателем сложности геометрической формы K , не зависящим от веса, что теоретически представляется интересным (см. стр. 132). Для этой цели следует составить уравнение

$$\bar{y}_{1.24} = a_1 + b_1y_2 + c_1y_4, \quad (10.4.8)$$

где $y_1 = \lg x_1$, $y_2 = \lg x_2$, $y_4 = \lg x_4$, а x_4 — другое обозначение показателя K .

Соответствующее уравнение в стандартизованном масштабе

$$\bar{t}_{1.24} = \beta_2t_2 + \beta_4t_4 \quad (10.4.9)$$

находится весьма просто, так как в силу некоррелированности y_4 и y_2 коэффициенты β_2 и β_4 равны коэффициентам корреляции между y_1 и аргументами y_2 и y_4 , т. е. $\beta_2 = r_{y_1y_2}$, $\beta_4 = r_{y_1y_4}$. Отсюда, потенцируя (10.4.8), получающееся известным образом из (10.4.9), приходим к искомому соотношению

$$\tilde{x}_{1.24} = a_1y_2^{b_1}y_4^{c_1}. \quad (10.4.10)$$

Однако из соотношения (10.4.7) можно сразу составить и уравнение (10.4.10), пользуясь формулой (6.3.7).

В самом деле, из уравнений (10.4.7)

$$\tilde{x}_{1.23} = 0,388x_2^{0,231} \cdot x_3^{0,402}$$

и (6.3.7)

$$x_4 = x_2^{-0,105}x_3,$$

исключая переменную x_3 , получаем

$$\tilde{x}_{1.23} = 0,388x_2^{0,231}x_2^{0,105} \cdot 0,402x_4^{0,402}$$

или

$$\tilde{x}_{1.23} = 0,388x_2^{0,273}x_4^{0,402}. \quad (10.4.10')$$

Итак, уравнение (10.4.10') представляет зависимость между трудоемкостью изготовления штампа x_1 , его весом x_2 и показателем сложности геометрической формы x_4 . Последний рассчитывается по формуле $x_4 = \frac{x_3}{x_2^{0,105}}$ и не зависит от веса (в смысле отсутствия корреляции между логарифмами этих величин).

Следует заметить, что получаемые из уравнения функциональной регрессии (10.4.7) величины $\tilde{x}_{1.23}$ являются выравненными функциональными средними определенного вида, в данном случае выравненными средними геометрическими. Поэтому там, где по условиям производственной задачи требуется использовать средние другого вида, необходимо вводить поправки на смещение средних. Так, например, если требуется определить суммарную трудоемкость изготовления некоторого числа h штампов (с определенным весом и числом поверхностей), то было бы неверно рассчитывать эту трудоемкость по формуле $h\tilde{x}_{1.23}$, так как это привело бы к преуменьшению потребности в трудовых затратах. Для получения правильных данных число штампов следует умножать на среднюю арифметическую, а не на среднюю геометрическую трудоемкость изготовления одного штампа. В таких случаях необходимо уметь оценивать по величине средней одного вида (в данном случае средней геометрической $\tilde{x}_{1.23}$) величину средней другого вида (средней арифметической $\tilde{x}_{1.23}$). Для этой цели существенную помощь оказывают формулы (6.2.20) и (6.2.21), предложенные Б. С. Ястремским.

§ 5. ПРИВЕДЕНИЕ МНОЖЕСТВЕННОЙ КОРРЕЛЯЦИИ К ЛИНЕЙНОМУ ВИДУ ПУТЕМ УМНОЖЕНИЯ ПЕРЕМЕННЫХ НА ФУНКЦИИ АРГУМЕНТОВ

В некоторых случаях удается получить линейное соотношение между переменными x_1, x_2, \dots, x_p , умноженными на некоторые специально для этой цели подобранные функции от аргументов x_2, \dots, x_p . Это соотношение может служить для приближенного нахождения корреляционной зависимости между x_1 и x_2, x_3, \dots, x_p .

где x_2, x_3, \dots, x_p — значения, определяемые из системы (10. 5. 1). Следовательно,

$$\bar{x}_{y_1, x_2, x_3, \dots, x_p} = \frac{\bar{u}_{y_1, x_2, x_3, \dots, x_p}}{\chi_1(x_2, x_3, \dots, x_p)} + \bar{u}_{y_2, x_2, x_3, \dots, x_p} x_2 \frac{\chi_2(x_2, x_3, \dots, x_p)}{\chi_1(x_2, x_3, \dots, x_p)} + \dots + \bar{u}_{y_p, x_2, x_3, \dots, x_p} x_p \frac{\chi_p(x_2, x_3, \dots, x_p)}{\chi_1(x_2, x_3, \dots, x_p)} \dots \quad (10. 5. 8)$$

и в пределе при $n \rightarrow \infty$

$$\bar{x}_{0, x_2, x_3, \dots, x_p} = \frac{\bar{u}_{01}}{\chi_1(x_2, \dots, x_p)} + \bar{u}_{02} x_2 \frac{\chi_2(x_2, \dots, x_p)}{\chi_1(x_2, \dots, x_p)} + \dots \quad (10. 5. 9)$$

Уравнения (10. 5. 8) и (10. 5. 9) выражают эмпирическую и предельную теоретическую регрессии переменной x_1 по аргументам x_2, x_3, \dots, x_p . Следовательно, рассчитав эмпирическую регрессию переменной y_1 по аргументам y_2, y_3, \dots, y_p и восстанавливая прежние переменные при помощи преобразования (10. 5. 1), (10. 5. 6) и (10. 5. 7), получаем эмпирическую регрессию переменной x_1 по аргументам x_2, x_3, \dots, x_p . Поскольку это положение справедливо при любом числе наблюдений n , то оно сохраняется и для предельных теоретических уравнений регрессии (10. 5. 4) и (10. 5. 9). Однако не совсем так обстоит дело для уравнения регрессии (10. 5. 5), где параметры b_1, b_2, \dots, b_p рассчитаны по методу наименьших квадратов. Значения этих параметров, получаемые из условия

$$\Sigma (y_1 - b_1 - b_2 y_2 - b_3 y_3 - \dots - b_p y_p)^2 = \min, \quad (10. 5. 10)$$

не будут совпадать со значениями параметров b'_1, b'_2, \dots, b'_p уравнения

$$\bar{x}_{y_1, x_2, x_3, \dots, x_p} = \frac{b'_1}{\chi_1(x_2, \dots, x_p)} + b'_2 x_2 \frac{\chi_2(x_2, \dots, x_p)}{\chi_1(x_2, \dots, x_p)} + \dots + b'_p x_p \frac{\chi_p(x_2, \dots, x_p)}{\chi_1(x_2, \dots, x_p)}, \quad (10. 5. 11)$$

если последние определяются из условия

$$\Sigma \left(x_1 - \frac{b'_1}{\chi_1} - b'_2 x_2 \frac{\chi_2}{\chi_1} - \dots - b'_p x_p \frac{\chi_p}{\chi_1} \right)^2 = \min, \quad (10. 5. 12)$$

так как условия (10. 5. 10) и (10. 5. 12) не эквивалентны.

Однако с увеличением числа наблюдений обе регрессии выражаемые уравнениями (10. 5. 5) и (10. 5. 11), сближаются между собой и стремятся при $n \rightarrow \infty$ одной и той же пре-

дельной теоретической регрессии (10.5.9). Поэтому при больших значениях n допустима замена уравнения (10.5.11) уравнением (10.5.5).

Итак, если нам удастся обнаружить систему переменных y_1, y_2, \dots, y_p , связанных между собой линейной регрессией

$$\bar{y}_{1.23\dots p} = b_1 + b_2 y_2 + b_3 y_3 + \dots + b_p y_p \quad (10.5.5)$$

и отличающихся от первоначальной системы переменных x_1, x_2, \dots, x_p множителями $\chi_1, \chi_2, \dots, \chi_p$, представляющими собой известные функции аргументов x_2, \dots, x_p , то из уравнения (10.5.5) при больших числах n можно получить приближенное уравнение регрессии x_1 по x_2, x_3, \dots, x_p по формуле

$$\begin{aligned} \bar{x}_{1.x_2, x_3, \dots, x_p} = & b_1 + b_2 x_2 \frac{\chi_2(x_2, \dots, x_p)}{\chi_1(x_2, \dots, x_p)} + b_3 x_3 \frac{\chi_3(x_2, \dots, x_p)}{\chi_1(x_2, \dots, x_p)} + \\ & + \dots + b_p x_p \frac{\chi_p(x_2, \dots, x_p)}{\chi_1(x_2, \dots, x_p)}. \end{aligned}$$

При этом расчет эмпирической регрессии по формуле (10.5.7)

$$\bar{x}_{1.x_2, x_3, \dots, x_p} = \frac{\bar{y}_{1.23\dots p}}{\chi_1(x_2, \dots, x_p)}$$

будет вполне точным при любом n .

Применение этого правила помогает в некоторых случаях считать уравнение криволинейной множественной регрессии, приводя его к уравнению прямолинейной регрессии между надлежаще выбранными переменными.

Изложенный выше метод применяется в главе XVI при нахождении зависимости между показателями хозяйственной деятельности строительных организаций. Было установлено, что зависимости между объемными показателями хозяйственной деятельности являются прямолинейными, в связи с чем эти показатели были положены в основу исследования. Из уравнения множественной регрессии между этими показателями, применяя правило, сформулированное в настоящем параграфе, можно получить уравнения регрессии между относительными (качественными) показателями хозяйственной деятельности, как это будет сделано в § 5 главы XVI.

РАЗДЕЛ III
ОШИБКИ СТАТИСТИЧЕСКИХ ПОКАЗАТЕЛЕЙ

ГЛАВА XI
ВЕРоятНОСТЬ

§ 1. ПОНЯТИЕ ВЕРоятНОСТИ

При формулировке положений или понятий теории корреляции нам неоднократно приходилось прибегать к закону больших чисел. Мы неоднократно опирались на факт сокращения случайной колеблемости средних показателей, мер дисперсии и корреляции, сокращения случайных изломов эмпирической линии регрессии и графиков распределения при увеличении числа наблюдений. На законе больших чисел основано определение понятий теоретической линии регрессии и корреляционной зависимости.

В настоящем разделе мы познакомимся с количественным выражением процесса сокращения колеблемости основных показателей статистического материала по мере увеличения числа наблюдений. Мы также увидим, как это количественное выражение используется для оценки надежности статистических выводов и заключений.

Вопросы о том, как складывается статистическая закономерность по мере увеличения количества случайных факторов, участвующих в процессе, изучаются методами теории вероятностей. В частности, эти методы используются и для освещения интересующего нас процесса уточнения статистических показателей по мере накопления материала.

Применительно к рассматриваемым нами вопросам остановимся на некоторых понятиях и выводах теории вероятностей.

В науке и в практической деятельности мы часто, отправляясь от комплекса известных нам фактов, даем оценку возможности осуществления каких-либо событий. Событие мы называем *возможным*, если его появление (или непоявление) полностью не обусловлено комплексом известных нам фактов. Так, например, мы говорим о возможности получения брака при обработке деталей на данном станке, о возможности для данного предприятия не только выполнить, но и перевыполнить назначенный ему план, о возможности хорошей летной погоды в день рейса, о воз-

возможности крупного выигрыша по наличной облигации в очередном тираже государственного займа и т. д.

Сравнивая между собой возможности появления и непоявления данного события, а также возможности появления различных событий, мы часто имеем основания утверждать, что появление данного события более или менее возможно, чем его непоявление, или что одно из событий более возможно, чем другое.

Таким образом, возможность события может быть большей или меньшей, или, можно сказать, имеет определенную степень сравнения. Так, можно утверждать, что возможность появления брака в очередной детали, обработанной на данном станке, меньше, чем возможность получения годной детали, или что возможность невыполнения плана на данном предприятии значительно меньше, чем возможность его перевыполнения. Переходя улицу в оживленном месте, мы ощущаем большую опасность несчастного случая попасть под машину, нежели в глухом переулке. Степень возможности данного события не является чем-то неизменным, она зависит от того комплекса условий (С), который в данном вопросе принимается во внимание, считается известным; с изменением этого комплекса степень возможности события меняется. Так, в последнем примере мы можем легко уменьшить возможность несчастного случая при переходе оживленной улицы, проявляя особую бдительность и соблюдая правила уличного движения. Степень возможности получения брака на станке уменьшится, если известно, что станок хорошо настроен и находится в исправном состоянии. Возможность хорошей летной погоды в назначенный день рейса будет различной для нас и для метеорологической станции, которая располагает по сравнению с нами большей совокупностью сведений о факторах, обуславливающих эту погоду, и т. д. На первый взгляд может показаться, что степень возможности является субъективным понятием, зависящим от комплекса фактов, на которые опирается тот или иной субъект, и от умения последнего делать из этих фактов правильные выводы. Однако в действительности обе эти предпосылки нет никакой необходимости связывать с тем или иным субъектом. Им можно приписывать вполне объективный характер. В отдельных случаях это положение достаточно очевидно. Так, извлекая из урны, содержащей 3 черных и 8 белых шаров, наудачу один шар, мы имеем вполне объективную уверенность в том, что извлечь белый шар более возможно, чем черный.

В других случаях принципиальное существование степени объективной возможности является для нас руководящим принципом, вытекающим из материалистического мировоззрения.

Сделаем следующий шаг в направлении уточнения рассматриваемого понятия. Предположим, что мера объективной возможности явления при наличии комплекса условий (С) выражается числом. В этом случае степень возможности будет называться *вероятностью события*.

Таким образом, вероятностью события называется степень объективной возможности данного события, выраженная количественно, или, иначе говоря, — *мера объективной возможности данного события*. Следует условиться о шкале измерения вероятностей. Эта шкала ограничена двумя критическими точками: 0 и 1. Невозможному событию приписывается вероятность 0, достоверному — вероятность 1. Всем промежуточным степеням возможности соответствуют числа, заключенные между 0 и 1. Можно дать конкретную интерпретацию величины вероятности, пригодную в отдельных случаях для прямого подсчета вероятности. Назовем шансом, *благоприятствующим* событию А, такое другое событие В, появление которого неизбежно влечет за собой появление А. Шансом, *неблагоприятствующим* событию А, назовем событие Д, появление которого исключает появление А. Так, например, в урне, содержащей некоторое количество черных и белых шаров, наличие каждого белого шара является шансом, благоприятствующим выходу белого шара при выборке, наличие каждого черного шара — шансом, неблагоприятствующим появлению белого шара.

Вероятность определяют как отношение числа M шансов, благоприятствующих данному событию, к числу N всех шансов:

$$P = \frac{M}{N}. \quad (11.1.1)$$

При этом шансы должны удовлетворять следующим трем условиям: 1) они должны быть *равновозможными*, т. е. появление одного из них не должно быть более возможно, чем появление другого; 2) они должны *взаимно исключать друг друга*, т. е. появление одного из них должно исключать появление другого; 3) один из этих шансов обязательно должен осуществиться.

Так, если в урне находятся 10 черных и 20 белых шаров, то вероятность выхода белого шара при случайной выборке составит $\frac{20}{30} = \frac{2}{3}$, вероятность выхода черного шара $\frac{10}{30} = \frac{1}{3}$.

Если бросается монета, то имеются два равновозможных шанса — выпадеть решеткой или гербом, поэтому вероятность выпадения герба или решетки равна $\frac{1}{2}$.

Если бросается игральная кость, то выпадение четного числа очков имеет вероятность $\frac{1}{2}$, так как число сторон с четным количеством очков равно числу сторон с нечетным числом очков, причем выпадение кости любой стороной одинаково возможно.

Итак, мы видим, что в некоторых случаях можно исчислить вероятность, основываясь на вышеприведенной интерпретации этого понятия. Однако в громадном большинстве случаев при-

менения теории вероятностей к практическим и научным вопросам данная интерпретация ничего не может дать для прямого исчисления вероятности, так как мы не в состоянии различить в комплексе условий (С) тех равновозможных, исключающих друг друга шансов, которых требует определение. Так, имеет смысл говорить о вероятности получения брака в очередной детали, обрабатываемой на данном станке; однако никто не в состоянии усмотреть равновозможные шансы, из которых эту вероятность можно было бы построить. В страховом деле говорят о вероятности дожития до определенного возраста, однако и здесь не имеет смысла искать и выводить эту вероятность из каких-либо равновозможных шансов и т. д. Во всех таких случаях вероятности событий мы находим косвенным путем — на основе статистики.

В случае косвенного определения вероятности важно иметь уверенность в том, что вероятности рассматриваемых явлений принципиально существуют, подобно тому, как мы, измеряя косвенным путем расстояния до недоступных точек, твердо знаем, что эти расстояния действительно существуют.

Вопрос о предпосылках, при наличии которых можно заранее утверждать существование вероятности рассматриваемых явлений, не получил еще в науке надлежащего разрешения¹. Решение этого вопроса будет иметь огромное методологическое значение, так как позволит очертить область применений и использования теории вероятностей как орудия познания.

В ряде случаев удается установить существование вероятностей, непосредственно исследуя поведение статистических показателей в процессе накопления материала. Об этом — в следующем параграфе.

§ 2. ВЕРОЯТНОСТИ И ПРОЦЕССЫ НАКОПЛЕНИЯ МАТЕРИАЛА

Рассмотрим обычную схему эксперимента или наблюдения, которая используется в научном исследовании. Пусть нас интересует некоторое событие А, появление которого обусловлено разнообразными причинами или факторами, из которых одни ему благоприятствуют, другие не благоприятствуют. Желая изучить действие некоторых причин на появление события А в условиях, когда другие причины не могут быть закреплены, подвергаем наблюдению или воспроизводим комплекс изучаемых условий (назовем его комплексом С) несколько раз, следя за появлением данного события. Мы часто ожидаем, что при многократном повторении комплекса (С) в появлении события А обнаружится некоторая закономерность, несмотря на изменчивость прочих условий появления события, не входящих в комплекс (С).

¹ См. статью А. Я. Хинчина в сборнике «Философские вопросы современной физики», АН СССР, 1952.

Простейшая закономерность, которую в ряде таких случаев действительно удается обнаружить, заключается в том, что относительная частота появления события A в сериях из ряда последовательных повторений комплекса (C) не претерпевает систематических изменений, а устойчиво колеблется вокруг некоторого числового уровня, не зависящего от случайных повышений или понижений частот в отдельных сериях. С увеличением числа повторений комплекса (C) в каждой серии уровень относительных частот не меняется, но колеблемость частот вокруг этого уровня неограниченно сокращается. Этот уровень характеризует меру объективной возможности появления события A при определенном процессе повторения комплекса условий (C) и соответствует *постоянной вероятности* события A .

Подобные закономерности, которые в дальнейшем будем называть закономерностями вида (α), обычно наблюдаются в тех случаях, когда в комплексе постоянных условий (C) закреплены основные факторы изучаемого явления, а множество остальных факторов, остающихся переменными, не обнаруживает заметных и систематических изменений от опыта к опыту по какому-либо определенному закону.

Пусть, например, нас интересует возможность появления определенного дефекта A в стали при двух различных вариантах технологического процесса плавки. Для этого производим в экспериментальном порядке два ряда плавок (C) и (C_1), причем основные условия процесса плавления для каждого ряда плавок сохраняются по возможности одинаковыми. Если изменение прочих условий, способных повлиять на результат плавки, не носит систематического характера и если результаты предыдущих плавок не влияют на последующие, то имеются предпосылки для закономерности (α) в процессе накопления плавок (C) и (C_1) при постоянных вероятностях P и P_1 появления дефекта A . По вероятностям P и P_1 можно сравнивать качество вариантов C и C_1 тех-

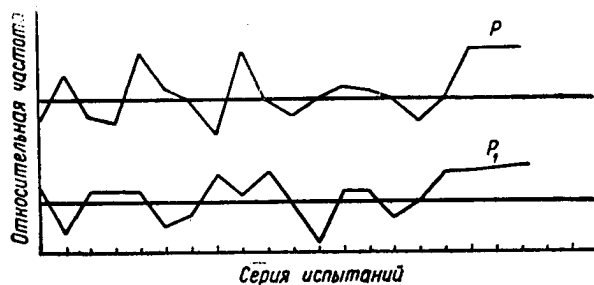


Рис. 54

нологического процесса (рис. 54). Вероятности P и P_1 оцениваются приближенно по относительным частотам появления A в ря-

дах экспериментальных плавков. Чем длиннее ряд, тем точнее оценка.

Наряду с подобными простейшими закономерностями в процессе накопления материала существуют и другие, более сложные. Во многих процессах возможность наступления того или иного события меняется с течением времени — вероятность является переменной величиной, в результате чего относительная частота колеблется вокруг уровня, изменяющегося по ходу процесса.

Так, если накопление числа опытных плавков происходит не в порядке специально поставленного эксперимента, а в порядке последовательного присоединения плавков, получаемых в естественных условиях по ходу производственного процесса, то может оказаться, что некоторые из условий, оставшиеся незакрепленными, претерпевают довольно значительные систематические изменения от плавки к плавке, что создает предпосылки для схемы переменной вероятности. В этом случае сопоставление качества двух вариантов процесса (C) и (C_1) может быть при известных условиях произведено путем сравнения двух кривых, представляющих изменение вероятности P и P_1 во времени. Кри-

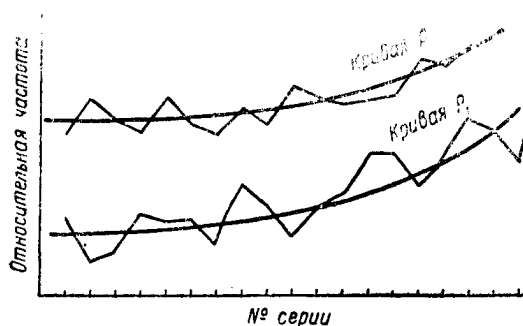


Рис. 55

вые изменения P и P_1 могут быть приближенно получены путем выравнивания опытных графиков, тем точнее, чем больше опытов.

В ряде случаев наблюдаются такие закономерности, при которых случайное появление или непоявление события в предыдущем испытании комплекса (C) оказывает влияние на формирование уровня вероятности в дальнейших испытаниях, повышая или понижая его. В таких случаях говорят о рядах взаимосвязанных испытаний комплекса (C). Примеры таких рядов дают нам многие производственные процессы.

Подъемы и падения на графиках изменения относительных частот образуют в этих случаях неправильные волны случайного характера. При наличии положительной зависимости между

смежными испытаниями волны оказываются длиннее, при наличии отрицательной зависимости — короче.

Следует подчеркнуть, что поставленная выше задача — найти количественное выражение для сокращения колеблемости статистических показателей по мере увеличения числа наблюдений — может получить решение лишь для таких процессов накопления материала, в которых наблюдаются определенные закономерности появления интересующих нас событий, управляемые вероятностями. Там, где таких закономерностей нет, задача не может получить никакого решения. И, наоборот, для различных закономерностей задача получает и различные решения. Так, при положительной зависимости между результатами испытаний процесс сокращения колеблемости относительных частот протекает медленнее, чем в случае независимости, и наоборот, при наличии отрицательной зависимости этот процесс ускоряется.

Примеры процессов, в основе которых лежат четко выраженные вероятности появления событий, дает нам выборочный метод исследования, широко используемый в народном хозяйстве, в частности, в вопросах контроля качества продукции. Задача этого метода состоит в том, чтобы устанавливать распределения большой массы единиц по тем или иным признакам с достаточной точностью путем выборочной проверки, организованной на началах теории вероятностей.

Рассмотрим, каким образом на основе расчета вероятностей создаются условия, обеспечивающие решение поставленных задач.

Существуют различные способы выборочного исследования, которым соответствуют различные вероятностные схемы накопления материала. Рассмотрим так называемый простой случайный отбор, характеризующийся постоянством вероятностей и независимостью испытаний, следовательно, отвечающий процессу накопления материала (α). Этот способ заключается в следующем.

Пусть совокупность, из которой производится выборка, именуемая *генеральной совокупностью*, содержит N деталей, из которых M — дефектные. Относительная частота или доля дефектных деталей в генеральной совокупности равна $\frac{M}{N}$. Из генеральной совокупности последовательно извлекаются в случайном порядке n деталей, причем после каждого извлечения детали фиксируется ее годность или дефектность, а деталь *возвращается обратно* в генеральную совокупность и участвует в последующих испытаниях.

Организация отбора должна быть такова, чтобы обеспечить при каждом испытании для каждой детали, находящейся в генеральной совокупности, *равные шансы* попасть в выборку. Это обеспечивается тщательным перемешиванием деталей или другим эквивалентным способом, создающим равновозможность вы-

хода. При расчете вероятностей здесь используется схема равно-возможных шансов. В отдельном испытании должна быть вынута одна из N находящихся в генеральной совокупности деталей. Следовательно, имеется M благоприятствующих и $N - M$ неблагоприятствующих шансов появления в отдельном испытании интересующего нас события — дефектной детали; наступление одного из них исключает возможность наступления другого в том же испытании; они равновозможны согласно условиям организации отбора.

Следовательно, вероятность появления события в определенном испытании равна $\frac{M}{N}$ и совпадает с относительной частотой этого события в генеральной совокупности. Наша выборка организована таким образом, что *в ней возникает вероятность появления дефектной детали, количественно совпадающая с относительной частотой этой детали в генеральной совокупности*. Определение вероятности способом, изложенным выше, через равно-возможные и несовместные шансы будет практически пригодным в той мере, в какой удастся обеспечить равные шансы элементарных исходов путем перемешивания или другими эквивалентными приемами.

Широкая практика показывает, что это действительно так, и что вероятности, определенные подобным образом, служат центрами, вокруг которых колеблются относительные частоты события при повторении испытаний методом простого случайного отбора. Но вероятность появления дефектной детали при соблюдении равновозможности выхода каждой детали, находящейся в генеральной совокупности, равна относительной частоте дефектных деталей в генеральной совокупности. Поэтому выборочная относительная частота будет колебаться вокруг генеральной относительной частоты, как вокруг своего центра, и при этом будет, как правило, тем меньше от нее отличаться, чем больше число испытаний. В этом основная идея отбора, организованного на началах теории вероятностей. Важен здесь не только самый факт стремления выборочной относительной частоты к генеральной, но и то обстоятельство, что в процессе накопления материала осуществляется закономерность, которая позволяет оценить величину ожидаемых погрешностей при любом числе испытаний. Эти оценки, с которыми мы в дальнейшем познакомимся, и могут быть положены в основу инженерных и экономических расчетов, обосновывающих сравнительную эффективность различных систем отбора.

Процесс накопления статистического материала определяет структуру последнего. Статистическая совокупность является материалом простейшей статистической структуры. Она отвечает процессу (α) накопления материала при постоянной вероятности событий и независимых испытаниях.

Наряду со статистическими совокупностями математическая статистика рассматривает образования других структур — статистический комплекс, статистическую последовательность (временной ряд) и др., которым соответствуют переменные вероятности или взаимосвязанные испытания.

В этой книге изучается материал простейшей статистической структуры — статистической совокупности.

§ 3. СВОЙСТВА ВЕРОЯТНОСТЕЙ

Основные свойства и операции над вероятностями в современной математической науке формулируются в виде системы целесообразно выбранных допущений или аксиом. Общепринятая в настоящее время система аксиом была предложена акад. А. Н. Колмогоровым в 1933 г.¹ Для наших целей методически удобно излагать свойства вероятностей, опираясь на какую-либо конкретную физическую интерпретацию этого понятия. Такую интерпретацию может дать схема равновероятных шансов, изложенная в § 1, или же схема повторения событий при многократном воспроизведении комплекса условий (С) процессом (α), изложенная в § 2. Мы изберем здесь этот последний путь ознакомления со свойствами вероятностей.

Вероятность проявляется как средняя характеристика относительной частоты появления события при испытании некоторого комплекса условий (С). Но так как средняя всегда величина того же вида, что и осредняемые значения, то вероятность должна обладать свойствами относительной частоты. Рассмотрим эти свойства. Пояснения, которые при этом даются, не следует считать доказательствами, а лишь естественной мотивировкой свойств, которыми характеризуется понятие вероятности.

Рассмотрим два события А и В, которые ожидаются в процессе накопления (α) с вероятностями P_A и P_B . Нижеследующая таблица дает числовой пример совместного распределения событий А и В в 1000 произведенных испытаний.

1. Таблица абсолютных частот

А \ В	В		Итого
	Наличие	Отсутствие	
Наличие	50	250	300
Отсутствие	200	500	700
Итого	250	750	1000

¹ А. Н. Колмогоров. Основные понятия теории вероятностей, 1936 (на немецком яз., изд. 1933 г.).

Примем за основание общее число наблюдений $n=1000$ и рассчитаем отношения абсолютных частот к этому основанию. Получаем следующую таблицу относительных частот.

II. Таблица относительных частот

	В	\bar{B}	Итого
А	$p_{AB} = 0,05$	$p_{A\bar{B}} = 0,25$	$p_A = 0,30$
\bar{A}	$p_{\bar{A}B} = 0,20$	$p_{\bar{A}\bar{B}} = 0,50$	$p_{\bar{A}} = 0,70$
Итого	$p_B = 0,25$	$p_{\bar{B}} = 0,75$	1,00

Здесь \bar{A} , \bar{B} означают непооявление A и B ;

p_A , p_B , $p_{\bar{A}}$, $p_{\bar{B}}$ — относительные частоты событий A , B , \bar{A} , \bar{B} ;

p_{AB} , $p_{\bar{A}B}$, $p_{A\bar{B}}$, $p_{\bar{A}\bar{B}}$ — относительные частоты совмещения событий A и B , \bar{A} и \bar{B} и т. д.

Обозначим через $A+B$ событие, состоящее в появлении какого-либо из двух событий A и B (либо обоих событий одновременно). Относительная частота события $A+B$ равна, очевидно, сумме относительных частот событий AB , $A\bar{B}$, $\bar{A}B$:

$$p_{A+B} = p_{AB} + p_{A\bar{B}} + p_{\bar{A}B} = 0,05 + 0,25 + 0,20 = 0,50 \quad (11.3.1)$$

или

$$p_{A+B} = p_A + p_B - p_{AB} = 0,30 + 0,25 - 0,05 = 0,50.$$

Аналогичные соотношения имеют место для вероятностей соответствующих событий, приближенными значениями которых являются относительные частоты:

$$P_{A+B} = P_{AB} + P_{A\bar{B}} + P_{\bar{A}B}$$

или

$$P_{A+B} = P_A + P_B - P_{AB}. \quad (11.3.2)$$

Последняя формула читается следующим образом: *вероятность появления какого-либо одного из двух событий A и B равна сумме вероятностей этих событий минус вероятность их совмещения.*

Соотношение (11.3.2) называется *правилом сложения вероятностей*. В частности, если события A и B несовместимы, то $P_{AB} = 0$, и формула (11.3.2) принимает вид

$$P_{A+B} = P_A + P_B. \quad (11.3.3)$$

Вероятность появления какого-либо из двух несовместимых событий A и B равна сумме вероятностей этих событий. Разумеется, аналогичные формулы можно составить и для комбинации любых двух других событий таблицы:

$$P_{A+B}, P_{\bar{A}+\bar{B}}, P_{A+\bar{A}}$$

и т. д., например,

$$P_{A+\bar{A}} = P_A + P_{\bar{A}} - P_{A\bar{A}} = P_A + P_{\bar{A}} = 1,$$

так как $P_{A\bar{A}} = 0$.

Аналогичное правило можно сформулировать для любого числа несовместимых событий. При анализе табл. I нас может интересовать вопрос: как часто появляется (или как вероятно) одно из двух событий A или B в тех случаях, когда происходит другое событие, или же в тех случаях, когда другое событие не происходит? Этой цели служит вычисление так называемых *условных* относительных частот и соответственно *условных* вероятностей события, которые вычисляются при *условии*, что *другое событие произошло*, или же, наоборот, при *условии*, что *оно не произошло*.

Условные относительные частоты обозначаются: $p_{A|B}$, $p_{A|\bar{B}}$, $p_{\bar{A}|B}$ и т. д., впереди черты пишется событие, частота которого определяется, позади черты — условие, которое при этом закрепляется.

Для вычисления $p_{A|B}$ нужно абсолютную частоту m_{AB} делить на абсолютную частоту события B , для вычисления $p_{B|\bar{A}}$ нужно $m_{\bar{A}B}$ делить на $m_{\bar{A}}$ и т. д., как это видно в табл. III и IV.

Таблица III

	B	\bar{B}
A	$p_{A B} = 0,20$	$p_{A \bar{B}} = 0,33$
\bar{A}	$p_{\bar{A} B} = 0,80$	$p_{\bar{A} \bar{B}} = 0,67$
Итого . .	1,00	1,00

Между безусловными и условными относительными частотами существуют соотношения, которые легко проверить по табл. II—IV (с точностью до второго знака после запятой):

$$p_A = p_{A|B}p_B + p_{A|\bar{B}}p_{\bar{B}}$$

(и т. п. для $p_{\bar{A}}$, p_B , $p_{\bar{B}}$),

Таблица IV

	B	\bar{B}	Итого
A	$P_{B A} = 0,17$	$P_{\bar{B} A} = 0,83$	1,00
\bar{A}	$P_{B \bar{A}} = 0,29$	$P_{\bar{B} \bar{A}} = 0,71$	1,00

$$P_{AB} = P_A P_{B|A} = P_B P_{A|B}$$

(и т. п. для $P_{\bar{A}\bar{B}}$, $P_{A\bar{B}}$, $P_{\bar{A}B}$).

Аналогичные соотношения должны иметь место между соответствующими вероятностями:

$$P_A = P_{A|B} P_B + P_{A|\bar{B}} P_{\bar{B}} \quad (11.3.4)$$

и т. п. для

$$P_{AB} = P_A P_{B|A} = P_B P_{A|B} \quad (11.3.5)$$

и т. п. для

$$P_{\bar{A}\bar{B}}, P_{\bar{A}B}, P_{A\bar{B}}.$$

Соотношение (11.3.4) называется свойством полной вероятности. Оно читается: *полная или безусловная вероятность события равна средней арифметической из условных вероятностей этого события, взвешенных по вероятностям условий.*

Соотношение (11.3.5) называется *правилом умножения вероятностей*. Оно читается: *вероятность совмещения двух событий равна произведению вероятности одного события на условную вероятность другого при условии, что первое событие произошло.*

Если условные вероятности $P_{A|B}$ и $P_{A|\bar{B}}$ равны, то событие A не зависит от события B. Чем сильнее отличается $P_{A|B}$ от $P_{A|\bar{B}}$, тем сильнее зависимость A от B.

Из равенства (11.3.4) видно, что в первом случае

$$P_{A|B} = P_{A|\bar{B}} = P_A, \quad (11.3.6)$$

из равенства (11.3.5)

$$P_{AB} = P_A P_B, \quad (11.3.7)$$

из равенств (11.3.5) и (11.3.4)

$$P_{B|A} = P_{B|\bar{A}} = P_B. \quad (11.3.8)$$

Равенство (11.3.8) означает, что если A не зависит от B , то и B не зависит от A , т. е. свойство независимости событий обратимо.

Равенство (11.3.7) читается: вероятность совмещения двух независимых событий равна произведению вероятностей этих событий. Аналогичным образом формулируется правило умножения вероятностей для любого числа независимых событий.

Сложение и умножение вероятностей принадлежат к числу основных операций, при помощи которых по вероятностям одних событий рассчитывают вероятности других событий, с ними связанных.

Задача. Найти вероятность того, что в n испытаниях, произведенных по схеме процесса (α), событие A , имеющее вероятность p_0 , появится $0, 1, 2, \dots, n$ раз.

Решение. Вероятность $P(0)$ того, что в n испытаниях событие A не появится вовсе, равна вероятности n -кратного появления события \bar{A} . Следовательно, по свойству умножения

$$P(0) = (1 - p_0)^n$$

или, обозначая $(1 - p_0)$ через q_0 , $P(0) = q_0^n$.

Вероятность $P(1)$ того, что событие A появится один раз в n испытаниях, найдется следующим образом. Определим сперва вероятность того, что A появится в первом испытании и не появится в следующих. Эта вероятность по свойству умножения равна:

$$P_{\underbrace{1 \ 00 \dots 0}_{n-1}} = p_0 q_0^{n-1}.$$

Очевидно, той же величине равна вероятность того, что событие A появится во втором, третьем и т. д. испытаниях и не появится в остальных. Но все эти случаи являются несовместимыми путями однократного появления события A . По правилу сложения искомая вероятность $P(1)$ является суммой всех этих вероятностей, а следовательно,

$$P(1) = n p_0 q_0^{n-1}.$$

Вероятность $P(2)$ того, что событие появится два раза в n испытаниях, равна сумме вероятностей появления этого события на двух определенных порядковых местах (например, 1 и 2, 1 и 3 и т. д.). Но каждая такая вероятность равна по правилу умножения

$$p_0^2 q_0^{n-2}.$$

Следовательно,

$$P(2) = \frac{n(n-1)}{1 \cdot 2} p_0^2 q_0^{n-2}$$

и т. д.

В общем случае m -кратного появления события

$$P(m) = C_n^m p_0^m q_0^{n-m}, \quad (11.3.9)$$

где C_n^m — число сочетаний из n элементов по m .

Формула (11.3.9) называется *биномиальным законом распределения частоты* появления события в n испытаниях, произведенных по схеме процесса (α).

Рассчитаем вероятность $P(m)$ для конкретного случая, когда $p_0 = 0,2$ и $n = 6$.

Таблица 58

m	$P(m)$
0	$0,8^6 = 0,2621$
1	$6 \cdot 0,2 \cdot 0,8^5 = 0,3932$
2	$15 \cdot 0,2^2 \cdot 0,8^4 = 0,2458$
3	$20 \cdot 0,2^3 \cdot 0,8^3 = 0,0819$
4	$15 \cdot 0,2^4 \cdot 0,8^2 = 0,0154$
5	$6 \cdot 0,2^5 \cdot 0,8 = 0,0015$
6	$0,2^6 = 0,0001$
	<hr/> 1,0000

ГЛАВА XII

ЗАКОНЫ РАСПРЕДЕЛЕНИЯ И КОРРЕЛЯЦИИ

§ 1. ОБЩИЕ СВЕДЕНИЯ

Наряду с событиями, для которых в испытаниях возможны лишь два исхода — появление или непоявление, будем рассматривать многовариантные события или *признаки*, для которых при каждом испытании в условиях комплекса (C) возможен один из нескольких различных исходов A_1, A_2, \dots, A_k . В принятой нами схеме накопления испытаний процессом (α) этим вариантам будут соответствовать постоянные вероятности P_1, P_2, \dots, P_k .

Формула или правило, которое для каждого варианта A_1, A_2, \dots, A_k дает соответствующую вероятность P_1, P_2, \dots, P_k , выражает закон распределения признака A . Если число вариантов признака конечно и невелико, то они могут быть записаны вместе с соответствующими вероятностями в таблице.

Вероятности P_1, P_2, \dots, P_k находят либо на основе теоретического анализа явления, либо приближенно по относительным частотам вариантов A_1, A_2, \dots, A_k в достаточно длинном ряду повторений комплекса условий

(C). Пример вероятностей, найденных первым путем, дает закон распределения частоты m появления события в n испытаниях, выведенный в конце предыдущей главы. Примеры вероятностей, найденные вторым путем, дают нам табл. 2, 20 и другие таблицы раздела I. Для всех этих таблиц следует считать, что материал накоплен процессом (α), и рассматривать относительные частоты как приближенные значения вероятностей соответствующих вариантов.

Таблица 59

Варианты	Вероятность
A_1	P_1
A_2	P_2
⋮	⋮
A_k	P_k
Итого . .	$\sum P = 1$

Законы распределения количественных признаков изображаются графически в виде полигонов, гистограмм и плавных кривых подобно рядам распределения. Место частот здесь занимают вероятности. При дискретной вариации вероятности изображаются ординатами соответствующих вариантов (рис. 56); при не-

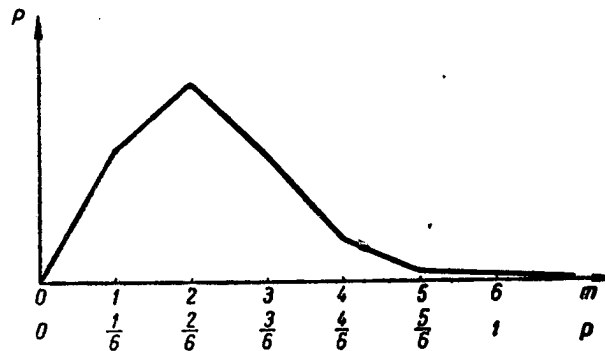


Рис. 56

прерывной — вероятности относятся не к отдельным вариантам, а к интервалам значений признака и изображаются в виде площадей криволинейных трапеций, построенных на этих интервалах (рис. 57).

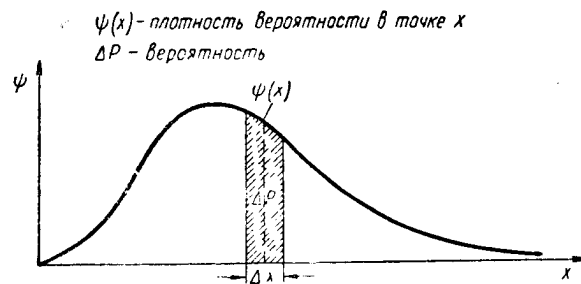


Рис. 57

Относя вероятность к единице измерения признака, т. е. деля вероятность ΔP на величину интервала Δx , получаем *плотность вероятности* (среднюю плотность на интервале Δx) — понятие, аналогичное плотности частоты.

Плотность вероятности в точке x представляет собой производную

$$\psi(x) = \frac{dP}{dx} = \lim_{\Delta x \rightarrow \infty} \frac{\Delta P}{\Delta x} \quad (12.1.1)$$

и изображается (там, где она существует) в виде ординаты теоретической кривой распределения, восстановленной в точке x (рис. 57).

Для теоретических кривых распределения среднее значение \bar{x}_0 и среднее квадратическое отклонение σ_0 определяются при дискретной вариации по формулам

$$\bar{x}_0 = \Sigma xP, \quad (12.1.2)$$

$$\sigma_0 = \sqrt{\Sigma (x - \bar{x}_0)^2 P}, \quad (12.1.3)$$

при непрерывной вариации по формулам

$$\bar{x}_0 = \int_{-\infty}^{+\infty} x\psi(x) dx, \quad (12.1.4)$$

$$\sigma_0 = \sqrt{\int_{-\infty}^{+\infty} (x - \bar{x}_0)^2 \psi(x) dx}, \quad (12.1.5)$$

где $\psi(x)$ — плотность вероятности в точке x , $\int_{-\infty}^{+\infty} \psi(x) dx = 1$.

\bar{x}_0 и σ_0 называются теоретическими показателями, так как при их вычислении варианты признака взвешиваются не по относительным частотам, а по вероятностям, которые можно рассматривать как теоретические значения относительных частот, или по плотностям вероятностей.

Теоретическую среднюю \bar{x}_0 принято называть *математическим ожиданием* случайной переменной x и обозначать в этом случае через $M(x)$. Этим термином в дальнейшем мы будем часто пользоваться.

Построение теоретических кривых распределения по опытным данным составляет задачу математического выравнивания рядов распределения, которая была поставлена в главе II раздела I и рассматривается в следующем параграфе.

Наряду с законами, лежащими в основе статистических рядов распределения, теория вероятностей изучает законы, лежащие в основе многомерных статистических распределений — корреляционных таблиц двух и большего числа измерений. Такие законы называются многомерными законами распределения, или законами корреляции.

Многомерный закон распределения определяется правилом или формулой, которая в случае дискретных распределений для каждой комбинации значений признаков $x, y, z \dots$ дает вероят-

ность этой комбинации, а в случае непрерывных распределений — плотность вероятности. При этом плотность вероятности в точке $(x, y, z \dots)$ определяется так

$$\psi(x, y, z, \dots) = \lim_{\substack{\Delta x \rightarrow 0 \\ \Delta y \rightarrow 0 \\ \Delta z \rightarrow 0 \\ \dots}} \frac{\Delta P}{\Delta x \Delta y \Delta z \dots},$$

где $(\Delta x \Delta y \Delta z \dots)$ — объем бесконечно малого элемента области, на которую распространяется вероятность ΔP .

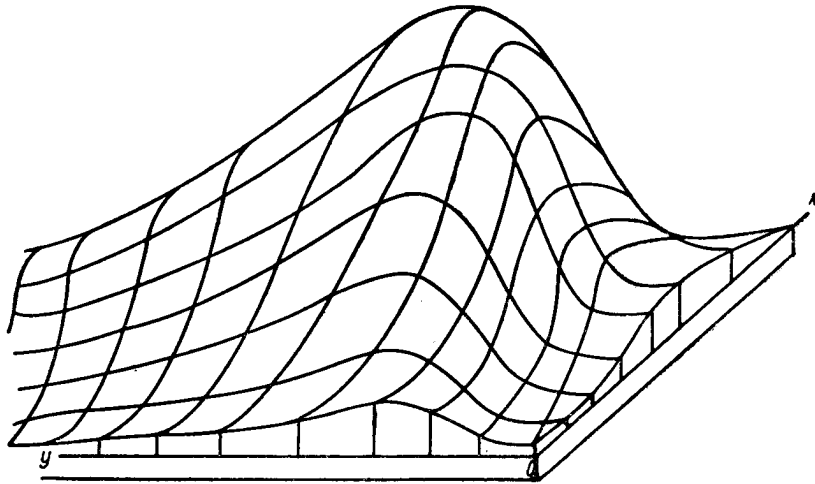


Рис. 58. Поверхность распределения

Двумерный закон распределения непрерывного типа изображается в декартовой системе координат в виде поверхности распределения (рис. 58), где аппликаты выражают плотности вероятностей соответствующих точек координатной плоскости (xOy) . Сечения этой поверхности плоскостями, параллельными координатной плоскости yOz , дают теоретические кривые распределения по признаку y , соответствующие заданным значениям x . Эти кривые являются графическим изображением так называемых условных законов распределения y при заданных x .

Проекции центров тяжести этих кривых на плоскость xOy называются условными математическими ожиданиями y ; геометрическое место последних является теоретической линией регрессии y по x . Аналогично сечения поверхности распределения плоскостями, параллельными координатной плоскости xOz , образуют теоретические кривые распределения x по признаку y , соответствующие заданным значениям y и выражающие условные

законы распределения x . Проекции центров тяжести этих кривых на плоскость xOy являются изображениями условных математических ожиданий x ; их геометрическое место — теоретическая линия регрессии x по y .

Аналогично определяются условные теоретические дисперсии $\sigma_{oy.x}^2$ и $\sigma_{ox.y}^2$, т. е. дисперсии условных законов распределения y и x , и показатели, характеризующие тесноту корреляционной зависимости, как-то: теоретические корреляционные отношения, теоретический коэффициент корреляции. Все эти показатели определяются формулами, вполне аналогичными формулам подсчета соответствующих статистических показателей, с тем различием, что весами при вариантах служат не частоты, а вероятности (или плотности вероятностей), а конечное суммирование в случае непрерывных законов заменяется интегрированием. Так, например, для теоретического коэффициента корреляции имеем формулу:

$$r_0 = \frac{M\{(x - \bar{x}_0)(y - \bar{y}_0)\}}{\sigma_{0x}\sigma_{0y}},$$

где в случае непрерывной вариации

$$M\{(x - \bar{x}_0)(y - \bar{y}_0)\} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x - \bar{x}_0)(y - \bar{y}_0) \varphi(x, y) dx dy.$$

При накоплении материала процессом (α) все рассматриваемые нами теоретические показатели можно считать предельными значениями соответствующих статистических показателей подобно тому, как вероятности можно считать предельными значениями относительных частот. Теоретические показатели обозначаются теми же символами, что и соответствующие им статистические, но сопровождаются индексом 0.

§ 2. НОРМАЛЬНЫЙ ЗАКОН РАСПРЕДЕЛЕНИЯ

Наиболее значительные приложения в практике статистического анализа и основное значение для теории математической статистики имеет так называемый нормальный закон распределения. Многочисленные ряды распределения опытного материала на производстве обнаруживают в большинстве случаев черты нормального закона. Теория вероятностей освещает этот факт, исследуя условия, при которых возникает нормальный закон распределения.

Суть дела заключается в том, что нормальный закон распределения является *предельным* законом, к которому приближаются многие другие законы распределения при определенных условиях. Если интересующий нас признак можно рассматривать как результат суммарного действия ряда факторов, каждый из которых мало связан с большинством остальных, и влияние каж-

ного фактора на результат намного перекрывается суммарным влиянием всех остальных факторов, то при большом числе таких факторов закон распределения признака становится близким к нормальному¹.

Перечисленные условия имеют, по-видимому, достаточно общий характер, чем объясняется широкое распространение нормального закона². Особую важность представляет то обстоятельство, что нормальному закону распределения приближенно подчиняются случайные ошибки важнейших статистических показателей при достаточном числе наблюдений, накопленных процессом (n). Отсюда необходимо подробно остановиться на свойствах этого закона и его использовании.

Нормальный закон распределения выражается формулой

$$\varphi(x) = \frac{1}{\sqrt{2\pi b}} e^{-\frac{1}{2} \left(\frac{x-a}{b}\right)^2}, \quad (12.2.1)$$

где x — вариант признака;

$\varphi(x)$ — соответствующая плотность вероятности;

a, b — параметры нормального закона;

e, π — известные константы математического анализа.

Исследование формулы (12.2.1) показывает, что кривая нормального закона — так называемая нормальная кривая распределения — симметрична относительно вертикальной оси, проходящей через точку $x=a$, в которой имеет место максимум кривой. Кривая имеет две точки перегиба, находящиеся на расстоянии $\pm b$ от центра a . Кривая в обе стороны асимптотически сближается с осью x (рис. 59).

Легко показать, что параметры a и b могут быть выражены следующим образом:

$$a = \int_{-\infty}^{+\infty} x\varphi(x) dx,$$

$$b = \sqrt{\int_{-\infty}^{+\infty} (x-a)^2 \varphi(x) dx}.$$

Следовательно, $a = \bar{x}_0$ и $b = \sigma_0$, т. е. параметры a и b являются соответственно теоретической средней арифметической величи-

¹ Г. Крамер, Математические методы статистики, гл. 17 и 18, 1948, А. Я. Хинчин, Предельные законы для сумм независимых случайных величин, 1938.

² О значении нормального закона распределения при обработке статистических данных см. Б. С. Ястремский, Математическая статистика, 1957.

ной признака и теоретическим средним квадратическим отклонением признака от его средней величины. Поэтому формула (12.2.1) будет в дальнейшем изображаться в виде

$$\varphi(x) = \frac{1}{\sqrt{2\pi}\sigma_0} e^{-\frac{1}{2}\left(\frac{x-\bar{x}_0}{\sigma_0}\right)^2}. \quad (12.2.2)$$

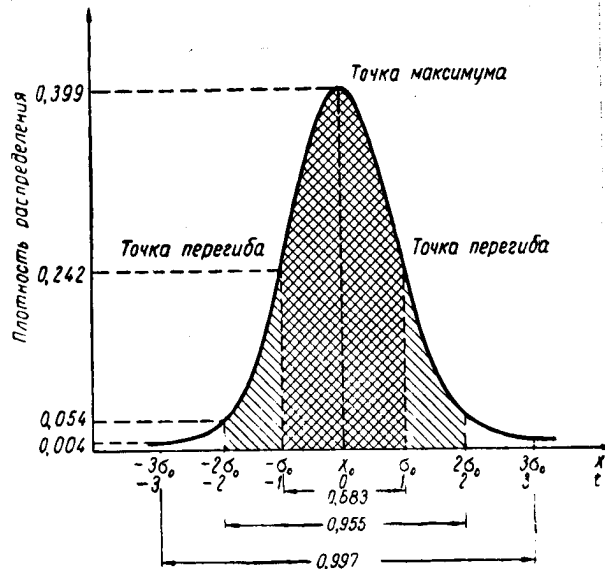


Рис. 59. Нормальная кривая распределения

С изменением параметра \bar{x}_0 кривая, сохраняя свою форму, перемещается по оси x (рис. 60). С изменением параметра σ_0 кривая изменяет свое рассеяние (рис. 61).

В частном случае, когда $\bar{x}_0 = 0$, $\sigma_0 = 1$, уравнение нормальной кривой принимает простейший вид

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}. \quad (12.2.3)$$

Уравнение (12.2.3) называется стандартным уравнением нормальной кривой. Общее уравнение (12.2.2) всегда можно привести к стандартной форме (12.2.3) путем замены аргумента x переменным t

$$t = \frac{x - \bar{x}_0}{\sigma_0}, \quad (12.2.4)$$

что означает стандартизацию масштаба x .

Напомним физический смысл переменной t :
 t показывает, на сколько сигм (σ_0) отклоняется данное значение x от своей средней \bar{x}_0 . Следовательно, новое начало отсчета совпадает с прежней \bar{x}_0 и новая единица измерения равна прежней σ_0 .

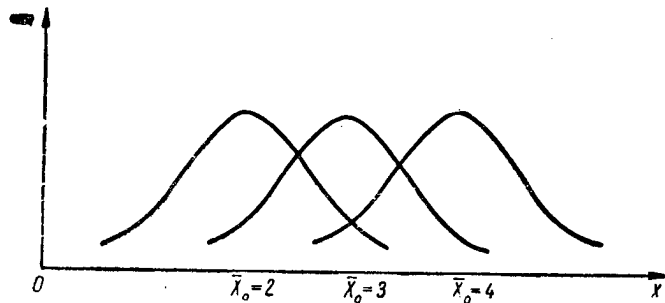


Рис. 60

Преобразование (12. 2. 4) позволяет написать уравнение (12. 2. 2) в виде

$$\varphi(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}, \quad (12. 2. 5)$$

где $\varphi(t)$ — плотность вероятности на единицу измерения признака t , которая в σ_0 раз больше единицы измерения x

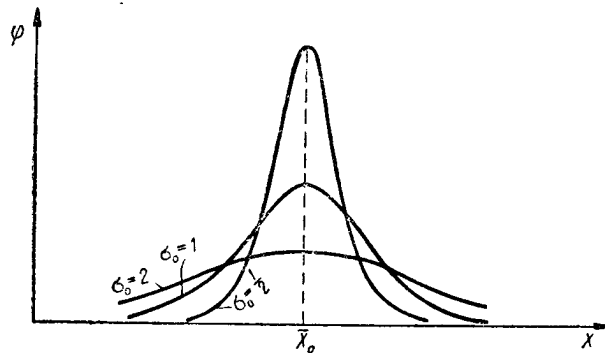


Рис. 61

(см. рис. 59). Следовательно, $\varphi(t) = \sigma_0 \varphi(x)$ и σ_0 в знаменателе правой части формулы (12. 2. 2) сокращаются. Вычисление $\varphi(t)$ по формуле (12. 2. 5) заменяется таблицами этой функции, которые даны в приложении II, табл. 2.

Расчет нормальной кривой распределения применительно к опытным данным производится так называемым методом мо-

ментов, который состоит в том, что показатели \bar{x}_0 и σ_0 нормального закона приравниваются соответственным показателям \bar{x} и σ нашего опытного материала. Моментами распределения (одномерного) порядка k относительно C называются показатели вида: $M_k = \frac{\sum m(x-C)^k}{\sum m}$ или $M_{0k} = M\{(x-C)^k\}$. Для многомерных распределений моментами называются показатели вида: $M_{k, l, \dots} = \frac{\sum m(x-C_1)^k(y-C_2)^l \dots}{\sum m}$ или соответственно $M_{0k, l, \dots} = M\{(x-C_1)^k(y-C_2)^l \dots\}$. \bar{x} , σ и \bar{x}_0 , σ_0 являются разновидностями моментов¹. Следовательно, подставляя в формуле (11.2.2) на место неизвестных \bar{x}_0 и σ_0 величины \bar{x} и σ , рассчитанные по опытным данным, получаем уравнение нормальной кривой². Для построения нормальной кривой остается задать ряд значений x , получить по уравнению (12.2.2) соответствующие $\varphi(x)$ и по полученным точкам построить кривую. Этот процесс сильно облегчается, если применить замену переменной (11.2.4) и использовать таблицы функции $\varphi(t)$.

Процесс построения нормальной кривой проиллюстрируем на примере распределения поршневых колец по величине упругости.

1. Прежде всего рассчитаем параметры \bar{x} и σ . Расчет производим в столбцах 3, 4, 5 табл. 60

Таблица 60

x	m	x'	$x'm$	x'^2m	t	$\varphi(t)$	$\varphi(x)1000$	Y°/∞	$S(x_1, x_2) = \frac{1}{\varphi(x)} \cdot t$
1110	1	-6	-6	36	-3,62	0,0006	0,007	0,1	0,0004
1160	0	-5	0	0	-3,02	0,0042	0,051	0	0,0025
1210	4	-4	-16	64	-2,42	0,0213	0,257	0,4	0,0128
1260	9	-3	-27	81	-1,82	0,0761	0,917	0,9	0,0459
1310	18	-2	-36	72	-1,21	0,1919	2,312	1,8	0,1156
1360	40	-1	-40	40	-0,61	0,3312	3,990	4,0	0,1995
1410	49	0	0	0	-0,01	0,3970	4,783	4,9	0,2392
1460	43	1	43	43	0,59	0,3352	4,038	4,3	0,2019
1510	26	2	52	104	1,20	0,1942	2,310	2,6	0,1170
1560	8	3	24	72	1,80	0,0790	0,952	0,8	0,0476
1610	1	4	4	16	2,40	0,0224	0,270	0,1	0,0135
1660	1	5	5	25	3,00	0,0044	0,053	0,1	0,0026
Итого	200	—	3	553	—	—	—	—	—

$$\bar{x}' = \frac{3}{200} = 0,015; \quad \bar{x} = 1410 + 0,015 \cdot 50 = 1410,75;$$

¹ О моментах см. [6], [7], [8], [10], [11], [12].

² Это уравнение не является уравнением действительной теоретической кривой, а лишь наиболее подходящей оценкой уравнения, которую можно сделать на основании ограниченного числа испытаний.

$$\sigma' = \sqrt{\frac{553}{200} - 0,015^2} = 1,66; \sigma = 50 \cdot 1,66 = 83,00.$$

2. Находим значения t , соответствующие центрам интервалов (т. е. значениям x , указанным в столбце 1 таблицы), по формуле $t = \frac{x - \bar{x}}{\sigma}$; результаты записываем в столбце 6 таблицы; разность между последовательными значениями t постоянна и равна $\frac{i}{\sigma}$, где i — длина интервала.

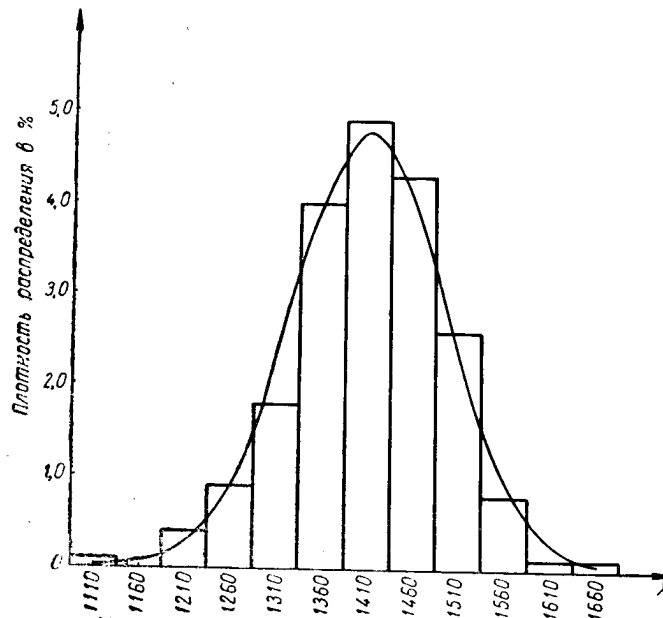


Рис. 62. Гистограмма и нормальная кривая распределения 200 поршневых колец по величине упругости

3. По таблице функции $\varphi(t)$ находим для каждого t значение $\varphi(t)$. Результаты пишем в столбце 7 таблицы.

Кроме указанных в таблице, для построения нормальной кривой важно отметить еще три значения $\varphi(t)$:

при $t = 0$ $\varphi(t) = 0,399$ (точка максимума)

$$(x = \bar{x});$$

при $t = \pm 1$ $\varphi(t) = 0,242$ (точки перегиба)

$$(x = \bar{x} \pm \sigma).$$

4. Рассчитываем ординаты нормальной кривой по формуле

$$\varphi(x) = \frac{\varphi(t)}{\sigma}. \quad (12.2.6)$$

Эти ординаты записываем в столбце 8 табл. 60.

5. Строим нормальную кривую по ординатам $\varphi(t)$, восстановленным в центрах интервалов (столбец 8 таблицы) и в дополнительных трех точках — точке максимума и точках перегиба (рис. 62).

6. На том же рисунке помещаем опытный график распределения — полигон или гистограмму. Для согласования масштабов ординаты опытного графика выражаются через плотности относительных частот ряда распределения Y , для чего абсолютные частоты интервалов m делятся на ni ,

$$Y = \frac{m}{in}, \quad (12. 2. 7)$$

где n — объем совокупности;

i — длина интервала.

Значения Y записаны в столбце 9 табл. 60.

Пользуясь нормальной кривой распределения, можно оценить вероятности значений признака более точно, чем это делается непосредственно по частотам ряда распределения, в частности оценить вероятный брак продукции. Для этой цели необходимо уметь рассчитывать площади криволинейных трапеций, ограниченные сверху отрезками нормальной кривой и имеющие основанием соответствующие интервалы оси абсцисс.

Площадь трапеции измеряет вероятность для x попасть в соответствующий интервал значений x или, что то же, вероятность для t оказаться в соответствующем интервале значений t (рис. 63).

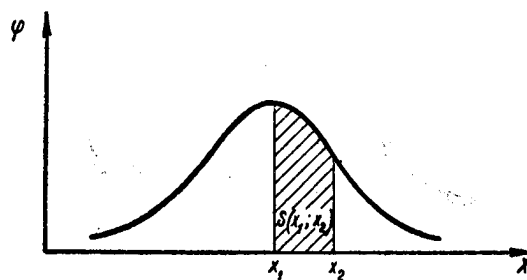


Рис. 63

Эта площадь выражается интегралами

$$S(x_1, x_2) = \int_{x_1}^{x_2} \varphi(x) dx = \int_{t_1}^{t_2} \varphi(t) dt,$$

где $t = \frac{x - \bar{x}}{\sigma}$, $x_2 \geq x_1$.

В случае, когда интересующие нас интервалы значений невелики, то расчет площадей приближенно осуществляется по формуле

$$S(x_1, x_2) = \varphi(x) \cdot i, \quad (12.2.8)$$

где x — центр интервала (x_1, x_2) ; i — величина интервала. Соответствующие числа для нашего примера указаны в столбце 10 таблицы.

В общем случае расчет площадей производится следующим образом.

Рассмотрим функцию

$$\Phi(t) = \int_{-t}^{+t} \varphi(t) dt. \quad (12.2.9)$$

Функция $\Phi(t)$ при $t \geq 0$ измеряет площадь нормальной кривой на отрезках $(-t, +t)$, концы которых расположены симметрично относительно центра (рис. 64).

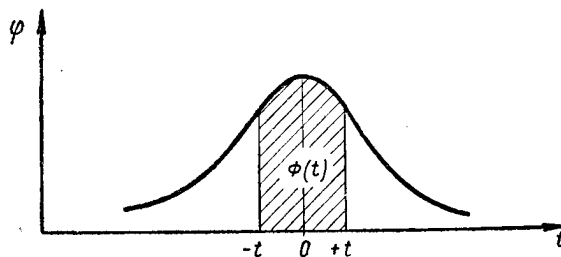


Рис. 64

Значения функции $\Phi(t)$ даны в расчетной табл. 2 приложения II. С помощью этой таблицы можно определять площадь кривой $S(t_1, t_2)$ на интервале с произвольно расположенными концами t_1 и t_2 ($t_2 \geq t_1$) по формуле

$$S(t_1, t_2) = \frac{\Phi(t_2) - \Phi(t_1)}{2}. \quad (12.2.10)$$

Для примера рассчитаем по формуле (12.2.10) площади кривой для интервалов табл. 61. Отмечаем границы интервалов по x в столбце 1 и соответствующие границы интервалов по t в столбце 2 табл. 61. Соответствующие значения $\Phi(t)$ указаны в столбце 3, а $S(t_1, t_2)$ — в столбце 4. При расчете следует иметь в виду нечетность функции $\Phi(t)$, т. е. что $\Phi(-t) = -\Phi(t)$.

Таблица 61

x_{zp}	t_{zp}	$\Phi(t_{zp})$	$S(t_1, t_2)$
1085	-3,92	-0,9999	0,0004
1135	-3,32	-0,9991	0,0028
1185	-2,72	-0,9935	0,0138
1235	-2,12	-0,9660	0,0472
1285	-1,52	-0,8715	0,1172
1335	-0,91	-0,6372	0,1969
1385	-0,31	-0,2434	0,2358
1435	0,29	0,2282	0,1992
1485	0,89	0,6255	0,1200
1535	1,50	0,8664	0,0490
1585	2,10	0,9643	0,0144
1635	2,70	0,9931	0,0030
1685	3,30	0,9990	

Полученные оценки вероятностей близки к оценкам, рассчитанным по формуле (12. 2. 8) (столбец 10 табл. 60).

В качестве другого примера рассмотрим следующую задачу. Для размера детали x после механической обработки задано поле допуска $20\mu \div 40\mu$. Известно, что размер x подчиняется нормальному закону распределения и что для некоторой партии деталей $\bar{x} = 33\mu$, $\sigma = 4\mu$. Определить вероятность брака по завышенному и заниженному размеру — „брак справа“ и „брак слева“ (рис. 65).

Решение. 1. Границы поля допуска $x_n = 20$ и $x_n = 40$ выражаем в масштабе t :
для левой границы

$$t_n = \frac{x_n - \bar{x}}{\sigma} = \frac{20 - 33}{4} = -3,25,$$

для правой границы

$$t_n = \frac{x_n - \bar{x}}{\sigma} = \frac{40 - 33}{4} = 1,75.$$

2. Применяем формулу (12. 2. 10) для определения площади правого „хвоста“ кривой, считая $\Phi(\infty) = 1$,

$$S(t_n, \infty) = \frac{1 - \Phi(t_n)}{2} = \frac{1 - \Phi(1,75)}{2} = \frac{1 - 0,9199}{2} = \frac{0,0801}{2} = 0,0400.$$

То же, для левого „хвоста“, считая $\Phi(-\infty) = -1$,

$$\begin{aligned} S(-\infty, t_n) &= \frac{\Phi(t_n) + 1}{2} = \frac{1 + \Phi(-3,25)}{2} = \\ &= \frac{1 - \Phi(3,25)}{2} = \frac{1 - 0,99885}{2} = \frac{0,00115}{2} \approx 0,0006. \end{aligned}$$

Следовательно, вероятный процент брака справа

$$B_n = 100 \cdot S(t_n, \infty) = 4,00\%;$$

вероятный % брака слева

$$B_s = 100 \cdot S(-\infty, t_s) = 0,06\%.$$

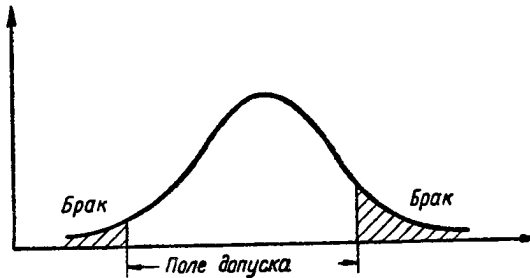


Рис. 65

Общий процент вероятного брака

$$B = B_n + B_s = 4,00 + 0,06 = 4,06\%.$$

§ 3. НОРМАЛЬНАЯ КОРРЕЛЯЦИЯ

Наряду с одномерными нормальными законами распределения в теории вероятностей рассматриваются нормальные законы двух и большего числа измерений — так называемые законы нормальной корреляции. Нормальные законы играют важную роль среди многомерных законов распределения. Как и в одномерном случае, закону нормальной корреляции приближенно подчиняются погрешности ряда важнейших статистических показателей при достаточно большом числе наблюдений по схеме процесса (а).

Двумерный нормальный закон распределения выражается формулой:

$$\varphi(x, y) = \frac{1}{2\pi\sqrt{1-r_0^2}\sigma_{0x}\sigma_{0y}} e^{-\frac{1}{2(1-r_0^2)}\left[\left(\frac{x-\bar{x}_0}{\sigma_{0x}}\right)^2 + \left(\frac{y-\bar{y}_0}{\sigma_{0y}}\right)^2 - 2\frac{x-\bar{x}_0}{\sigma_{0x}}\frac{y-\bar{y}_0}{\sigma_{0y}}r_0\right]}, \quad (12.3.1)$$

где $\varphi(x, y)$ — плотность вероятности в точке x, y .

Как видно, двумерный нормальный закон распределения определяется пятью параметрами: $\bar{x}_0, \bar{y}_0, \sigma_{0x}, \sigma_{0y}, r_0$.

Преобразование переменных:

$$t_x = \frac{x-\bar{x}_0}{\sigma_{0x}}; \quad t_y = \frac{y-\bar{y}_0}{\sigma_{0y}}; \quad \varphi(t_x, t_y) = \varphi(x, y)\sigma_{0x}\sigma_{0y} \quad (12.3.2)$$

приводит уравнение нормальной корреляции (12.3.1) к стандартному виду

$$\varphi(t_x, t_y) = \frac{1}{2\pi\sqrt{1-r_0^2}} e^{-\frac{1}{2(1-r_0^2)} [t_x^2 + t_y^2 - 2t_x t_y r_0]}, \quad (12.3.3)$$

где $\varphi(t_x, t_y)$ — плотность вероятности в точке t_x, t_y .

Уравнение (12.3.1) в декартовой системе координат изображается так называемой нормальной поверхностью распределения.

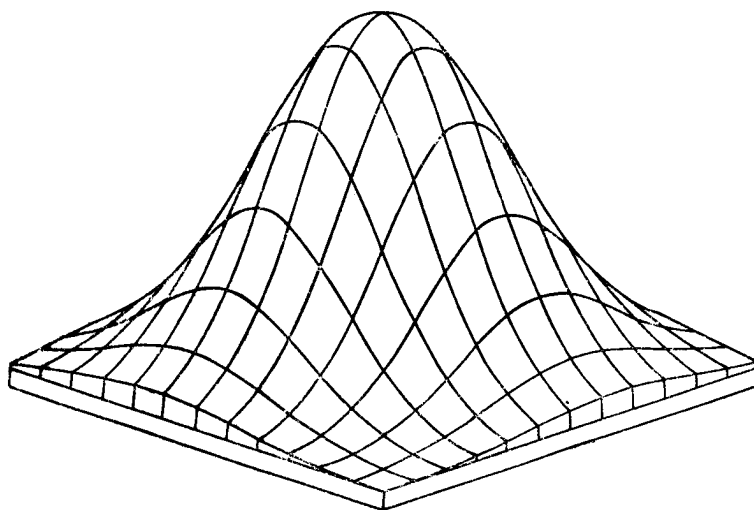


Рис. 66. Нормальная поверхность распределения

Исследование формулы (12.3.1) позволяет сделать следующие выводы:

1. Сечения нормальной поверхности распределения плоскостями, параллельными координатным плоскостям ($xO\varphi$) и ($yO\varphi$), дают нормальные кривые распределения. Следовательно, условные законы распределения по признакам x и y являются *нормальными*. Безусловные законы распределения по x и y также являются нормальными.

2. Условные математические ожидания \bar{y}_{0x} и \bar{x}_{0y} для этих законов рассчитываются по формулам:

$$\left. \begin{aligned} \bar{y}_{0x} &= \bar{y}_0 + r_0 \frac{\sigma_{0y}}{\sigma_{0x}} (x - \bar{x}_0) \\ \bar{x}_{0y} &= \bar{x}_0 + r_0 \frac{\sigma_{0x}}{\sigma_{0y}} (y - \bar{y}_0) \end{aligned} \right\}. \quad (12.3.4)$$

Следовательно, линии регрессии y по x и x по y прямые.

3. Условные дисперсии $\sigma_{oy.x}^2$ и $\sigma_{ox.y}^2$ рассчитываются по формулам:

$$\left. \begin{aligned} \sigma_{oy.x}^2 &= \sigma_{oy}^2 (1 - r_0^2) \\ \sigma_{ox.y}^2 &= \sigma_{ox}^2 (1 - r_0^2) \end{aligned} \right\} \quad (12.3.5)$$

Следовательно, условные дисперсии y по x и x по y постоянны.

4. Сечения нормальной поверхности распределения плоскостями, параллельными координатной плоскости xOy , образуют в проекции на эту плоскость семейство концентрических эллипсов, выражаемых уравнением:

$$\frac{1}{1 - r_0^2} \left\{ \left(\frac{x - \bar{x}_0}{\sigma_{ox}} \right)^2 + \left(\frac{y - \bar{y}_0}{\sigma_{oy}} \right)^2 - 2 \left(\frac{x - \bar{x}_0}{\sigma_{ox}} \right) \left(\frac{y - \bar{y}_0}{\sigma_{oy}} \right) r_0 \right\} = \chi^2, \quad (12.3.6)$$

где χ — параметр эллипса. Эти эллипсы называются корреляционными.

На контуре корреляционного эллипса плотность вероятности постоянна и равна

$$\varphi(x, y) = \frac{1}{2\pi\sigma_{ox}\sigma_{oy}\sqrt{1-r_0^2}} e^{-\frac{\chi^2}{2}}; \quad (12.3.7)$$

В каждой точке, находящейся внутри эллипса, плотность вероятности больше и в каждой точке, находящейся вне эллипса, плотность вероятности меньше, чем на контуре. Суммарная

вероятность точек, лежащих внутри эллипса, равна $1 - e^{-\frac{\chi^2}{2}}$; назовем ее *основной вероятностью* корреляционного эллипса и обозначим через $Q(\chi)$. Суммарная вероятность точек, лежащих вне эллипса, равна $e^{-\frac{\chi^2}{2}}$.

С увеличением параметра χ плотность вероятности $\varphi(x, y)$ на контуре эллипса убывает, а его основная вероятность $Q(\chi)$ — растет. Основная вероятность $Q(\chi)$ корреляционного эллипса больше суммарной вероятности точек, лежащих внутри любой другой равновеликой эллипсу фигуры корреляционного поля. И наоборот, при заданной суммарной вероятности точек наименьшая площадь ограничивается контуром корреляционного эллипса. Эти особенности делают корреляционный эллипс удобным для графической характеристики двумерного распределения, а также полезным при решении ряда практических задач.

Корреляционный эллипс проще всего строить, когда переменные выражены в стандартизованном масштабе. В этом случае эллипс называется *стандартным* корреляционным эллипсом. Его уравнение:

$$\frac{1}{1 - r_0^2} \{ t_x^2 + t_y^2 - 2t_x t_y r_0 \} = \chi^2. \quad (12.3.8)$$

В стандартизованном масштабе центр эллипса находится в начале координат ($t_x=0, t_y=0$), оси эллипса направлены вдоль биссектрис координатных углов: 1-го и 3-го углов — первая ось, 2-го и 4-го — вторая ось.

Длина первой оси $d_1=2\chi\sqrt{1+r_0}$; координаты ее концов:

$$A_1\left(\chi\sqrt{\frac{1+r_0}{2}}; \chi\sqrt{\frac{1+r_0}{2}}\right),$$

$$A_2\left(-\chi\sqrt{\frac{1+r_0}{2}}; -\chi\sqrt{\frac{1+r_0}{2}}\right).$$

Длина второй оси $d_2=2\chi\sqrt{1-r_0}$; координаты ее концов:

$$B_1\left(\chi\sqrt{\frac{1-r_0}{2}}; -\chi\sqrt{\frac{1-r_0}{2}}\right),$$

$$B_2\left(-\chi\sqrt{\frac{1-r_0}{2}}; \chi\sqrt{\frac{1-r_0}{2}}\right).$$

При $r_0 > 0$ первая ось является большой осью эллипса, вторая — малой; фокусы эллипса:

$$F_1(\chi\sqrt{r_0}; \chi\sqrt{r_0}), F_2(-\chi\sqrt{r_0}; -\chi\sqrt{r_0}).$$

При $r_0 < 0$ первая ось является малой осью эллипса, вторая — большой; фокусы эллипса:

$$F_1(\chi\sqrt{-r_0}; -\chi\sqrt{-r_0}), F_2(-\chi\sqrt{-r_0}; \chi\sqrt{-r_0}).$$

По этим элементам легко построить эллипс.

Чем больше $|r_0|$, тем сильнее вытянут стандартный эллипс вдоль большой оси; при $r_0=0$ стандартный эллипс превращается в окружность радиуса χ .

Расчет показателей нормальной корреляции применительно к опытным данным

Показатели нормальной корреляции рассчитываются применительно к опытным данным по методу моментов. Теоретические показатели $\bar{x}_0, \sigma_{0x}, \bar{y}_0, \sigma_{0y}, r_0$ в расчетных формулах заменяются соответствующими показателями опытного материала. С их помощью рассчитываются нормальная поверхность распределения и теоретические частоты интервальных клеток корреляционного поля, а также корреляционные эллипсы.

Покажем ход этих расчетов на примере корреляционной зависимости между двумя механическими характеристиками 209 плавков стали 30ХГСА — пределом прочности (σ_B) и пределом текучести (σ_S). Значения этих характеристик по каждой плавке были получены как средние арифметические из результатов испытаний двух образцов.

Приводим поле корреляции между σ_B и σ_S , построенное в стандартизованном масштабе (рис. 67), и соответствующую корреляционную таблицу (табл. 62).

Таблица 62

$y \backslash x$	102,25	106,25	110,25	114,25	118,25	122,25	126,25	130,25	134,25	Итого l
127,75									1	1
122,75							2	6	2	10
117,75						19	20	2		41
112,75				1	26	31	8			66
107,75			4	21	26	3				54
102,75			8	15	1	1				25
97,75		1	5	2						8
92,75	2	2								4
Итого h	2	3	17	39	53	54	30	8	3	n 209

Прежде всего рассчитаем показатели \bar{x} , σ_x , \bar{y} , σ_y , r . Проведя расчеты по корреляционной таблице, получим:

$$\bar{x} = 119,40; \sigma_x = 5,84; \bar{y} = 110,83; \sigma_y = 6,46; r = 0,882.$$

Уравнения регрессии y по x и x по y :

$$\bar{y}_x = -5,58 + 0,975x,$$

$$\bar{x}_y = 31,07 + 0,797y.$$

Условные дисперсии y по x и x по y :

$$\sigma_{y.x}^2 = \sigma_y^2(1 - r^2) = 9,26; \sigma_{y.x} = 3,04,$$

$$\sigma_{x.y}^2 = \sigma_x^2(1 - r^2) = 7,57; \sigma_{x.y} = 2,75.$$

На поле корреляции наносим линии регрессии.

Рассчитаем и построим стандартный корреляционный эллипс, для чего:

1. Выбираем основную вероятность корреляционного эллипса $Q(\chi)$. Эта вероятность может быть фиксирована на любом уровне,

однако для общей графической характеристики поля рассеяния ее целесообразно принять близкой к 1. В нашем примере примем $Q(\chi) = 0,980$.

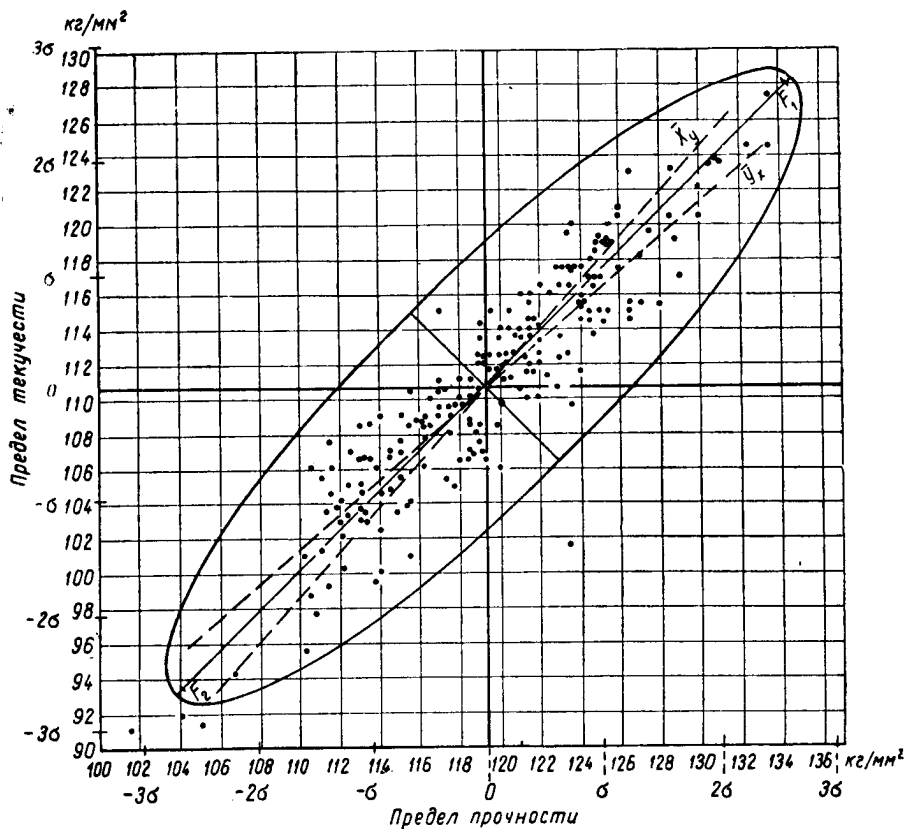


Рис. 67. Корреляционный эллипс

2. По основной вероятности $Q(\chi)$ находим значение χ по формуле $\chi = \sqrt{-2 \ln[1 - Q(\chi)]}$ или, что проще, пользуясь таблицей 2 приложения II. Для этого против значения

$$\varphi(t) = \frac{1 - Q(\chi)}{\sqrt{2\pi}}$$

находим $t = \chi$. В нашем случае

$$\varphi(t) = \frac{1 - 0,980}{\sqrt{6,284}} = 0,0080, \quad \chi = t = 2,80.$$

3. Так как $r > 0$, то определяем положение концов большой оси по формулам:

$$A_1 \left\{ \chi \sqrt{\frac{1+r}{2}}; \chi \sqrt{\frac{1+r}{2}} \right\} = A_1 \left\{ 2,80 \sqrt{\frac{1+0,882}{2}}; \right. \\ \left. 2,80 \sqrt{\frac{1+0,882}{2}} \right\} = A_1 \{2,72; 2,72\};$$

$$A_2 \left\{ -\chi \sqrt{\frac{1+r}{2}}; -\chi \sqrt{\frac{1+r}{2}} \right\} = A_2 \{-2,72; -2,72\}.$$

Определяем координаты концов малой оси.

$$B_1 \left\{ \chi \sqrt{\frac{1-r}{2}}; -\chi \sqrt{\frac{1-r}{2}} \right\} = B_1 \{0,672; -0,672\}.$$

$$B_2 \left\{ -\chi \sqrt{\frac{1-r}{2}}; \chi \sqrt{\frac{1-r}{2}} \right\} = B_2 \{-0,672; 0,672\}.$$

Положение фокусов эллипса:

$$F_1 \{ \chi \sqrt{r}; \chi \sqrt{r} \} = F_1 \{2,63; 2,63\},$$

$$F_2 \{ -\chi \sqrt{r}; -\chi \sqrt{r} \} = F_2 \{-2,63; -2,63\}.$$

По этим данным на поле корреляции строим эллипс (см. рис. 61). Легко проверить по чертежу, что длина большой оси

$$d_1 = 2\chi \sqrt{1+r} = 7,67$$

и длина малой оси

$$d_2 = 2\chi \sqrt{1-r} = 1,92.$$

4. Подсчитаем по чертежу число точек корреляционного поля, оказавшихся за пределами эллипса. Если число таких точек мало отличается от теоретически ожидаемого числа $n[1 - Q(\chi)]$, то это служит подтверждением правильности теоретических расчетов. В нашем случае за пределами эллипса оказались фактически 4 точки. Теоретически ожидаемое по закону нормальной корреляции число таких точек составляет $209(1 - 0,980) = 4,18$. Соответствие в этом случае полное. Некоторые практические приложения корреляционного эллипса указаны в § 3 гл. XIII.

Рассчитаем нормальную поверхность распределения и теоретические частоты корреляционной таблицы.

Пусть требуется рассчитать теоретическую частоту m_T для интервальной клетки ($x=114,25$; $y=107,75$), где x и y — координаты центра клетки. Для этой цели:

1. Выражаем координаты центра клетки в стандартизованном масштабе

$$t_x = \frac{x - \bar{x}}{\sigma_x} = \frac{114,25 - 119,40}{5,84} = -0,882,$$

$$t_y = \frac{y - \bar{y}}{\sigma_y} = \frac{107,75 - 110,83}{6,46} = -0,477.$$

2. Вычисляем $\chi^2(x, y)$ по формуле

$$\chi^2(x, y) = \frac{1}{1-r^2} \{t_x^2 + t_y^2 - 2t_x t_y r\} = 1,189,$$

откуда

$$\chi(x, y) = 1,09.$$

3. По табл. 2 приложения II, подставляя вместо t значение χ , находим $\varphi(\chi) = 0,2203$.

4. Находим плотность вероятности в точке ($t_x = -0,882$, $t_y = -0,477$):

$$\varphi(t_x, t_y) = \frac{1}{\sqrt{2\pi}\sqrt{1-r^2}} \varphi(\chi) = 0,186.$$

5. Находим приближенную частоту клетки по формуле:

$$m_T \approx \varphi(t_x, t_y) \frac{i_x i_y}{\sigma_x \sigma_y} n. \quad (12.3.9)$$

Формула (12.3.9) тем точнее, чем меньше i_x и i_y .

В нашем случае

$$m_T \approx 0,186 \frac{4 \cdot 5}{5,84 \cdot 6,46} \cdot 209 = 20,6.$$

Фактическая частота клетки — 21.

Теоретические частоты m_T служат для выравнивания эмпирических частот корреляционной таблицы.

Приводим таблицу теоретических частот, заменяющих эмпирические частоты табл. 62.

Таблица 62а

$x \backslash y$	102,25	106,25	110,25	114,25	118,25	122,25	126,25	130,25	134,25	138,25	Итого
132,75									0,1		0,1
127,75							0,1	0,8	1,1	0,2	2,2
122,75						0,4	4,3	6,0	1,1		11,8
117,75					1,1	13,3	18,7	3,2	0,1		36,4
112,75			0,2	1,9	22,0	32,0	5,5	0,1			61,7
107,75		0,2	1,7	20,6	29,9	5,2	0,1				57,7
102,75		0,9	10,6	15,4	2,7	0,1					29,7
97,75	0,2	3,0	4,3	0,8							8,3
92,75	0,5	0,7	0,1								1,3
Итого	0,7	4,8	16,9	38,7	55,7	51,0	28,7	10,1	2,4	0,2	209,2

Формула (12.3.9) может быть использована для приближенного расчета теоретической частоты любого прямоугольника на плоскости x, y с достаточно малыми сторонами, параллельными осям координат; в этом случае i_x и i_y обозначают длины сторон, а (x, y) координаты центра прямоугольника; таким образом, формула (12.3.9) имеет *интерполяционное значение*.

Например, найдем теоретическую частоту прямоугольника с центром в точке $(120,0; 110,0)$ и сторонами, параллельными осям, причем $i_x=2, i_y=3$.

$$t_x = \frac{120,0 - 119,40}{5,84} = 0,103,$$

$$t_y = \frac{110,0 - 110,83}{6,46} = -0,128,$$

$$\chi^2(x, y) = 2,266, \quad \chi = 1,505, \quad \varphi(\chi) = 0,1286,$$

$$\varphi(t_x, t_y) = 0,1089,$$

$$m_T = 0,1089 \frac{2 \cdot 3 \cdot 209}{5,84 \cdot 6,46} = 3,01.$$

ГЛАВА XIII

СТАТИСТИЧЕСКИЕ ПОКАЗАТЕЛИ В ПРОЦЕССЕ
НАКОПЛЕНИЯ МАТЕРИАЛА§ 1. ОТНОСИТЕЛЬНАЯ ЧАСТОТА В ПРОЦЕССЕ НАКОПЛЕНИЯ
МАТЕРИАЛА

Пусть в процессе (α) мы следим за появлением события А, вероятность которого равна p_0 . Обобщающим показателем статистического материала, на основе которого судят о вероятности p_0 , служит относительная частота появления события в n испытаниях:

$$p = \frac{m}{n}$$

Уже говорилось, что относительные частоты колеблются вокруг вероятности и тем теснее к ней примыкают, чем больше число наблюдений. При помощи правил теории вероятностей можно математически обосновать и уточнить это положение. В то же время математическая теория позволяет определять, какова вероятность тех или иных отклонений частоты какого-либо события от его вероятности при заданном числе наблюдений.

Исходя из постоянной вероятности p_0 появления события А и независимости испытаний, чем *математически характеризуется процесс накопления* (α), и при помощи правил сложения и умножения вероятностей в § 3 главы XI была установлена формула (11. 3. 9)

$$P(m) = C_n p_0^m q_0^{n-m},$$

дающая вероятность $P(m)$ абсолютной частоты m появления события А в n испытаниях. Заменяя в формуле (11. 3. 9.) m через np , получаем закон распределения относительной частоты p в следующем виде

$$P(p) = C_n^{np} p_0^{np} q_0^{nq}, \quad (13. 1. 1)$$

где $q = 1 - p$.

В частности, для примера где $p_0=0,2$, $n=6$, закон распределения p представлен в столбцах 1 и 2 табл. 63.

Таблица 63

p	P	pP	$p-p_0$	$(p-p_0)^2$	$(p-p_0)^2P$
0	0,2621	0	-0,2000	0,0400	0,01048
$\frac{1}{6}=0,1667$	0,3932	0,0655	-0,0333	0,0011	0,00043
$\frac{2}{6}=0,3333$	0,2458	0,0819	0,1333	0,0178	0,00438
$\frac{3}{6}=0,5000$	0,0819	0,0410	0,3000	0,0900	0,00737
$\frac{4}{6}=0,6667$	0,0154	0,0103	0,4667	0,2178	0,00335
$\frac{5}{6}=0,8333$	0,0015	0,0012	0,6333	0,4011	0,00060
1	0,0001	0,0001	0,8000	0,6400	0,00006
Итого	1,0000	0,2000			0,02667

График этого закона изображен на рис. 56, где по оси абсцисс откладываются значения p , по оси ординат — вероятности $P(p)$. Из формулы (11.3.9) можно вывести следующие основные положения.

Теорема I (о центре колеблемости p)

Средняя арифметическая из всех возможных значений p , взвешенных по своим вероятностям, совпадает с p_0 , или, иначе, математическое ожидание p равно p_0

$$M(p) = p_0. \quad (13.1.2)$$

Доказательство

$$\begin{aligned} M(p) &= \sum_{p=0}^1 pP(p) = \sum_{m=0}^n \frac{m}{n} \frac{n!}{m!(n-m)!} p_0^m q_0^{n-m} = \\ &= p_0 \sum_{m=1}^n \frac{(n-1)!}{(m-1)!(n-1-(m-1))!} p_0^{m-1} q_0^{(n-1)-(m-1)} = \\ &= p_0(p_0 + q_0)^{n-1} = p_0. \end{aligned}$$

Проверим теорему I по данным табл. 63. Вычисления располагаем в столбце 3 таблицы. Мы видим, что $Mp = \sum pP = 0,2000$, что совпадает со значением p_0 .

Теорема II (о размере колеблемости p)

Среднее квадратическое отклонение величины p , взвешенное по вероятностям значений p , обратно пропорционально корню квадратному из числа наблюдений n

$$\sigma_{0p} = \sqrt{\frac{p \cdot q_0}{n}}. \quad (13.1.3)$$

Доказательство

$$\begin{aligned} \sigma_{0p}^2 &= \sum_{p=0}^1 (p - p_0)^2 P(p) = \sum_{p=0}^1 p^2 P(p) - \\ &- 2 \sum_{p=0}^1 p p_0 P(p) + \sum_{p=0}^1 p_0^2 P(p) = \sum_{p=0}^1 p^2 P(p) - \\ &- 2 p_0 \sum_{p=0}^1 p P(p) + p_0^2 \sum_{p=0}^1 P(p) = \sum_{p=0}^1 p^2 P(p) - \\ &- 2 p_0^2 + p_0^2 = \sum_{p=0}^1 p^2 P(p) - p_0^2. \end{aligned}$$

Но

$$\begin{aligned} \sum_{p=0}^1 p^2 P(p) &= \sum_{m=0}^n \left(\frac{m}{n}\right)^2 \frac{n!}{m!(n-m)!} p_0^m q_0^{n-m} = \\ &= \sum_{m=2}^n \frac{m(m-1)}{n^2} \frac{n!}{m!(n-m)!} p_0^m q_0^{n-m} + \sum_{m=1}^n \frac{m}{n^2} \frac{n!}{m!(n-m)!} p_0^m q_0^{n-m} = \\ &= \frac{n-1}{n} p_0^2 \sum_{m=2=0}^{m=2=n-2} \frac{(n-2)!}{(m-2)! [(n-2)-(m-2)]} p_0^{m-2} q_0^{[(n-2)-(m-2)]} + \\ &+ \frac{1}{n} \sum_{m=1}^n \frac{m}{n} \frac{n!}{m!(n-m)!} p_0^m q_0^{n-m} = \\ &= \frac{n-1}{n} p_0^2 (p_0 + q_0)^{n-2} + \frac{p_0}{n} = \frac{n-1}{n} p_0^2 + \frac{p_0}{n}. \end{aligned}$$

Следовательно,

$$\sigma_{0p}^2 = \frac{n-1}{n} p_0^2 + \frac{p_0}{n} - p_0^2 = \frac{p_0 - p_0^2}{n} = \frac{p_0 q_0}{n},$$

что и требовалось доказать.

Проверка этого соотношения для нашего числового примера проведена в столбцах 4—6 таблицы.

По этим данным

$$\sigma_{0p}^2 = \Sigma (p - p_0)^2 P = 0,02667,$$

что совпадает с

$$\frac{p_0 q_0}{n} = \frac{0,2 \cdot 0,8}{6} = 0,02667.$$

Теорема III

$$P \{ |p - p_0| > \varepsilon \} \xrightarrow{n \rightarrow \infty} 0 \quad (13. 1. 4)$$

Вероятность отклонений p от p_0 , превосходящих по абсолютной величине любое малое положительное число ε , может быть сделана сколь угодно малой при n достаточно большом.

Теорема IV

$$P \left\{ t_1 \leq \frac{p - p_0}{\sigma_{0p}} \leq t_2 \right\} \xrightarrow{n \rightarrow \infty} \frac{1}{\sqrt{2\pi}} \int_{t_1}^{t_2} e^{-\frac{t^2}{2}} dt \quad (13. 1. 5)$$

Закон распределения величины $\frac{p - p_0}{\sigma_{0p}}$ при $n \rightarrow \infty$ стремится к стандартному нормальному закону распределения.

Доказательства последних двух теорем читатель найдет в курсах теории вероятностей [17—23].

Остановимся на смысле этих теорем.

Теорема I подтверждает истолкование вероятности как центра, вокруг которого колеблются относительные частоты. Она показывает, что этот центр должен определяться по формуле средней арифметической, соответствующим образом взвешенной.

Теорема II устанавливает закон, по которому сокращается средняя квадратическая ошибка σ_{0p} относительной частоты по мере увеличения числа наблюдений по схеме процесса (α).

Теорема III подтверждает истолкование вероятности как предела, к которому стремится относительная частота по мере увеличения числа наблюдений. Она раскрывает характер предельного перехода, который существенно отличается от предельного перехода, рассматриваемого в курсах математического анализа. В математическом анализе переменная величина, стремящаяся к пределу, принимает для любого значения аргумента строго опре-

деленное значение. В нашем же случае относительная частота p для любого значения n имеет строго определенную кривую распределения вероятностей, которая с увеличением n все более стягивается к центру — вероятности p_0 .

Теорема III может быть записана следующим образом:

$$p_0 = \lim_{B \quad n \rightarrow \infty} p, \quad (13.1.6)$$

что читается: вероятность p_0 является при n , стремящемся к бесконечности, *пределом по вероятности* относительной частоты p . Теорема III установлена Яковом Бернулли.

Теорема IV означает, что кривая распределения показателя p с увеличением n все более сближается с нормальными кривыми распределения, имеющими фиксированный центр p_0 и переменную дисперсию $\sigma_{0p}^2 = \frac{p_0(1-p_0)}{n}$. Следовательно, расчеты вероятностей

уклонений относительной частоты p от вероятности p_0 можно производить при больших n , пользуясь таблицами функции $\Phi(t)$ нормального закона распределения. Это обстоятельство значительно облегчает расчеты погрешностей. Необходимое для этого число n зависит от исходной вероятности p_0 . Чем ближе p_0 к 0 или к 1, тем больше должно быть n для достижения заданной точности расчетов при замене биномиального закона распределения нормальным. Теорема IV установлена Муавром и Лапласом.

§ 2. ОЦЕНКА ПОГРЕШНОСТЕЙ ОТНОСИТЕЛЬНОЙ ЧАСТОТЫ

Отклонение $p - p_0$ называется погрешностью или ошибкой относительной частоты. Чем больше величина допускаемых погрешностей, тем меньше точность опыта. Исчерпывающую характеристику возможных погрешностей дает закон распределения p . Однако для практических целей обычно используется более конкретная, хотя и менее общая, характеристика погрешности — так называемый *интервал рассеяния* показателя p , гарантированный с определенной близкой к 1 вероятностью (рис. 68).

Интервалом рассеяния показателя p назовем интервал, который содержит наиболее близкие к p_0 значения p и в котором значение p ожидается с заданной близкой к 1 вероятностью. Согласно теореме IV этот интервал при больших n должен располагаться симметрично относительно точки p_0 .

Величина интервала рассеяния $2\Delta_p$ зависит от вероятности P , с которой гарантируется нахождение показателя p в интервале рассеяния; чем ближе P к 1, тем больше интервал рассеяния (при неизменной точности опыта).

P называется *гарантийной вероятностью*. Точность опыта определяется двумя параметрами: интервалом рассеяния $2\Delta_p$ и гарантийной вероятностью P :

$$T = T[2\Delta_p, P].$$

Точность опыта тем выше, чем больше гарантийная вероятность при том же интервале рассеяния или чем меньше интервал рассеяния при той же гарантийной вероятности.

Гарантийная вероятность связана с интервалом рассеяния соотношением, вытекающим из теорем предыдущего параграфа. Введем так называемый *гарантийный коэффициент* t , который

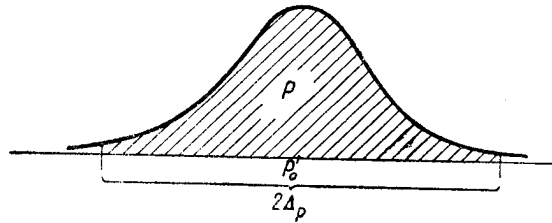


Рис. 68

показывает, сколько раз средняя квадратическая ошибка показателя p укладывается в половине интервала рассеяния

$$t = \frac{\Delta p}{\sigma_{0p}}. \quad (13.2.1)$$

Тогда согласно теореме IV

$$P \approx \Phi(t) \quad (13.2.2)$$

при больших n .

Но σ_{0p} определяется по формуле (13.1.3)

$$\sigma_{0p} = \sqrt{\frac{p_0(1-p_0)}{n}},$$

откуда видно, что при увеличении n t увеличивается и, следовательно, растет гарантийная вероятность при заданном интервале рассеяния. И, наоборот, задавая гарантийную вероятность $\Phi(t)$, тем самым фиксируем t , а из соотношения (13.2.1) видно, что интервал рассеяния $2\Delta p$ с увеличением n должен при этом сокращаться. Следовательно, точность опыта определяется, в конечном счете, числом наблюдений n . Увеличение числа наблюдений повышает точность опыта, причем это повышение точности может проявляться как в сокращении интервала рассеяния, так и в повышении гарантийной вероятности. Обычно при установлении параметров точности опыта гарантийную вероятность фиксируют на некотором принятом в таких случаях уровне, близком к 1, и,

исходя из этого уровня, определяют величину интервала рассеяния по формуле

$$2\Delta_p = 2t \sqrt{\frac{p_0(1-p_0)}{n}}. \quad (13.2.3)$$

В качестве уровня гарантийной вероятности принимают чаще всего одно из следующих чисел: 0,9973; 0,999; 0,990; 0,950.

Пример 1. Бросается 400 раз монета. Определить интервал рассеяния относительной частоты выпадения «герба» при гарантии $P=0,9973$.

Решение. $P \approx \Phi(t) = 0,9973$, откуда по таблицам функции $\Phi(t)$ $t=3$.

Величина интервала рассеяния показателя p составляет

$$2\Delta_p = 2t \sqrt{\frac{p_0(1-p_0)}{n}} = 2 \cdot 3 \sqrt{\frac{\frac{1}{2} \cdot \frac{1}{2}}{400}} = 0,15.$$

Следовательно, с вероятностью 0,9973 можно гарантировать, что p окажется в пределах от $p_0 - \Delta_p$ до $p_0 + \Delta_p$, т. е. от $0,500 - 0,075 = 0,425$ до $0,500 + 0,075 = 0,575$. А, следовательно, абсолютное число выпадений «герба» с той же вероятностью ожидается в пределах от $0,425 \cdot 400 = 170$ до $0,575 \cdot 400 = 230$.

Мы использовали здесь формулу

$$P \{ p_0 - t\sigma_{op} \leq p \leq p_0 + t\sigma_{op} \} \approx \Phi(t), \quad (13.2.4)$$

где $\sigma_{op} = \sqrt{\frac{p_0(1-p_0)}{n}}$,

которая установлена (теорема IV) при условии, что p_0 фиксировано, а p является неизвестной случайной переменной величиной.

В большинстве случаев на практике имеет место обратная постановка вопроса: по полученному в опыте значению p требуется указать пределы, в которых возможно ожидать p_0 . Существует правило, согласно которому для образования подобных пределов используется при больших значениях n и p , не слишком близком к 0 или 1, формула

$$p - t \sqrt{\frac{p(1-p)}{n}} \leq p_0 \leq p + t \sqrt{\frac{p(1-p)}{n}}, \quad (13.2.5)$$

где t выбирается в соответствии с принятой гарантией.

Имеются различные точки зрения по вопросу о способах обоснования этого правила. Не имея здесь возможности останавливаться на этом вопросе, отсылаем читателя к соответствующей

литературе¹. Рассмотрим, как применяется это правило к конкретным задачам.

Независимо от способа истолкования формулы (13.2.5), величина $2\Delta'_p = 2t \sqrt{\frac{p(1-p)}{n}}$ является в известном смысле приближенной величиной интервала рассеяния показателя p при гарантийной вероятности $P = \Phi(t)$. Этой терминологией пользуемся в дальнейшем.

Пример 2. Для выяснения доли дефектных изделий в партии было взято на выборку [процессом (α)] 225 деталей. Среди них оказалось 45 деталей дефектных. Определить границы, в которых ожидается неизвестная доля брака в партии при гарантийной вероятности 0,9973.

Решение. $\Phi(t) = 0,9973$, следовательно, $t = 3$. Относительная частота дефектных изделий в выборке $p = \frac{45}{225} = 0,20$. Величина поля рассеяния (приближенно)

$$2\Delta'_p = 6 \sqrt{\frac{0,20 \cdot 0,80}{225}} = 0,16.$$

Следовательно, $0,20 - 0,08 \leq p_0 \leq 0,20 + 0,08$ или $0,12 \leq p_0 \leq 0,28$ при гарантии 0,9973.

Один из основных вопросов, который возникает при проведении любого экспериментального исследования и любого исследования, опирающегося на статистический анализ опытных данных, — это вопрос о числе наблюдений, которые необходимо произвести для получения надежных выводов из материала. Очевидно, что ответ зависит от того, с какой точностью необходимо получить характеристики материала.

Если интересующей нас характеристикой является вероятность какого-либо события в процессе (α) , то задача решается следующим образом. Задается точность, с которой требуется определить вероятность p_0 , т. е. задается интервал рассеяния относительной частоты $2\Delta_p$ и гарантийная вероятность $P \approx \Phi(t)$.

Из формулы $\Delta_p = t \sqrt{\frac{p_0(1-p_0)}{n}}$ определяем необходимое n

$$n = \frac{p_0(1-p_0)}{\left(\frac{\Delta_p}{t}\right)^2}. \quad (13.2.6)$$

¹ А. Н. Колмогоров, «Известия Академии наук СССР» № 1—2, сер. мат., 6, 1942, стр. 3—32.

С. Н. Бернштейн, «Известия Академии наук СССР», сер. мат., 5, 1941, стр. 85—94, а также [7] гл. 34 и [17] гл. 11.

В правой части формулы (13. 2. 6) гарантийный коэффициент t определяется по таблице функции $\Phi(t)$ исходя из заданной гарантийной вероятности P . Неизвестное p_0 может быть оценено ориентировочно по данным предварительных исследований. Здесь разрешается весьма грубая степень приближения, так как нас обычно интересует ориентировочно лишь порядок величины n .

Пример 3. Требуется найти число испытаний процессом (α), необходимое для определения вероятности появления определенного дефекта, с точностью $2\Delta_p = 0,06$, $P = 0,990$, если ориентировочно эта вероятность находится на уровне $0,10—0,20$.

Решение. Из $P \approx \Phi(t) = 0,990$ находим по таблицам $t = 2,58$. p_0 для формулы (13. 2. 6) ориентировочно принимаем равным $\frac{0,10 + 0,20}{2} = 0,15$.

Подставляя в формулу (13. 2. 6), находим

$$n = \frac{0,15 \cdot 0,85}{\left(\frac{0,03}{2,58}\right)^2} = 962 \approx 1000.$$

§ 3. СРЕДНЯЯ АРИФМЕТИЧЕСКАЯ И СРЕДНЕЕ КВАДРАТИЧЕСКОЕ ОТКЛОНЕНИЕ В ПРОЦЕССЕ НАКОПЛЕНИЯ МАТЕРИАЛА

Мы познакомились с закономерностями поведения относительной частоты при накоплении материала по схеме процесса (α) и выяснили, как эти закономерности используются для оценки точности опыта. Аналогичные закономерности имеют место и для других статистических показателей. Рассмотрим эти закономерности для основных характеристик распределения — средней арифметической и среднего квадратического отклонения — и для показателей корреляции (в следующем параграфе).

Пусть испытания признака x производятся процессом (α). Это значит, что закон распределения x остается неизменным от испытания к испытанию и не зависит от результатов уже произведенных испытаний. Испытания обычно производятся для выяснения тех или иных недостающих нам особенностей закона распределения x , в частности, для оценки среднего значения x_0 и среднего квадратического отклонения σ_0 этого закона.

Обозначим через x_1, x_2, \dots, x_n возможные результаты 1-го, 2-го, ... n -го испытания. Эти результаты можно рассматривать как независимые друг от друга случайные величины (признаки), подчиненные одному и тому же закону распределения.

Для оценки по данным опыта неизвестных характеристик x_0 и σ_0 пользуемся соответствующими показателями статистического материала, рассчитанными по результатам испытаний:

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n}$$

и

$$\sigma = \sqrt{\frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{n}}$$

Показатели \bar{x} и σ являются функциями случайных величин x_1, x_2, \dots, x_n

$$P = f(x_1, x_2, \dots, x_n),$$

где P — обозначение статистического показателя. Следовательно, P само является в процессе (α) случайной величиной, подчиняющейся определенному закону распределения. Законы распределения \bar{x} и σ , изучаемые методами теории вероятностей, используются при оценке возможных погрешностей этих показателей, т. е. отклонений \bar{x} и σ от x_0 и σ_0 . Основное значение имеет установленный теорией факт, что законы распределения важнейших статистических показателей (в том числе всех показателей, алгебраически выражающихся через моменты) при большом числе наблюдений по схеме процесса (α) близки к нормальным законам. Следовательно, при значительном количестве наблюдений можно рассматривать законы распределения этих показателей как нормальные и пользоваться для расчета погрешностей таблицами функции $\Phi(t)$. При малом числе наблюдений (так называемая область «малых выборок») для оценки погрешностей приходится иметь дело с точными законами распределения статистических характеристик, отличающимися от нормального закона. Однако область «малых выборок» является весьма специфической как с точки зрения теории, так и с точки зрения круга приложений и в данной работе не рассматривается (о малых выборках см., например, [6], [7], [11], [12]):

$$\text{Итак, пусть } \bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n}, \text{ где } x_1, x_2, \dots, x_n —$$

независимые случайные величины, подчиненные одному и тому же закону распределения со средней x_0 и средним квадратическим отклонением σ_0 . Обозначим через $M(\bar{x})$ и $\sigma_{\bar{x}}$ соответственно математическое ожидание и среднее квадратическое отклонение для закона распределения \bar{x} .

Закон распределения средней арифметической характеризуется следующими теоремами.

Теорема I (о центре колеблемости \bar{x})

$$M(\bar{x}) = \bar{x}_0. \quad (13.3.1)$$

Теорема II (о размере колеблемости)

$$\sigma_{\bar{x}} = \frac{\sigma_0}{\sqrt{n}}. \quad (13.3.2)$$

Теорема III

$$\lim_{B, n \rightarrow \infty} \bar{x} = \bar{x}_0. \quad (13.3.3)$$

Теорема IV.

Закон распределения \bar{x} асимптотически нормален¹.

Поясним вкратце смысл этих теорем.

Теорема I утверждает, что центр, вокруг которого колеблется \bar{x} , совпадает с центром колеблемости x . Следовательно, принимая \bar{x} за приближенную оценку x_0 , не делаем при этом систематической погрешности.

По теореме II средний размер случайных погрешностей показателя \bar{x} , т. е. случайных отклонений \bar{x} от x_0 , прямо пропорционален средней погрешности индивидуальных значений x и обратно пропорционален корню квадратному из числа наблюдений.

Теорема III утверждает, что показатель \bar{x} с увеличением n стремится к x_0 как к своему пределу по вероятности. Следовательно, при достаточно большом числе наблюдений вероятность отклонений \bar{x} от x_0 , превышающих по абсолютной величине любое малое положительное число ϵ , может быть сделана сколь угодно малой.

Для общего случая, когда взаимно независимые величины x_1, x_2, \dots подчинены различным законам распределения, имеющим конечные дисперсии $\sigma_{01}^2, \sigma_{02}^2, \dots$, не превосходящие одного и того же постоянного числа C , теорема III была установлена П. Л. Чебышевым. В частном случае, который здесь рассматривается, когда x_1, x_2, \dots подчинены одному и тому же закону распределения, А. Я. Хинчин показал, что для справедливости теоремы достаточно лишь существование конечного \bar{x}_0 .

По теореме IV \bar{x} относится к числу показателей, имеющих при большом n закон распределения, близкий к нормальному (каков бы ни был исходный закон распределения \bar{x} , имеющий конечные \bar{x}_0 и σ_0). Следовательно, закон распределения \bar{x} практически при достаточно большом числе наблюдений полностью определяется значениями параметров \bar{x}_0, σ_0 и числом n .

¹ Распределение показателя \bar{x} называется асимптотически нормальным, если закон распределения величины $\frac{\bar{x} - M(\bar{x})}{\sigma_{\bar{x}}}$ стремится при $n \rightarrow \infty$ к стандартному нормальному закону распределения. В теореме IV § 1 этой главы утверждалась асимптотическая нормальность распределения относительной частоты p .

В общем случае, когда взаимно независимые величины x_1, x_2, \dots подчинены различным законам распределения, теорема IV в весьма общих условиях была впервые установлена А. М. Ляпуновым. Для рассматриваемого здесь частного случая, когда x_1, x_2, \dots подчинены одному и тому же закону распределения, Линдбергом и П. Леви было показано, что для справедливости теоремы достаточно лишь существование конечных x_0 и σ_0 .

Возможные погрешности показателя σ (отклонения σ от σ_0) характеризуются законом распределения σ , для которого имеют место следующие теоремы (через $M(\sigma)$ и σ_0 , обозначены средняя арифметическая и среднее квадратическое отклонение для закона распределения σ).

Теорема I (о центре колеблемости)

$$M(\sigma) = b_n \sigma_0, \quad (13.3.4)$$

где

$$b_n = \frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)} \sqrt{\frac{2}{n}} \approx 1 - \frac{3}{4n} \quad (13.3.5)$$

(где Γ — гамма-функция).

Теорема II (о размере колеблемости)

$$\sigma_{0\sigma} \approx \frac{\sigma_0}{\sqrt{2n}}. \quad (13.3.6)$$

Теоремы I и II справедливы в предположении, что переменная x подчиняется нормальному закону распределения. Если это условие не соблюдается, то формулы (13.3.4) и (13.3.6) неточны.

Теорема III

$$\lim_{n \rightarrow \infty} \sigma = \sigma_0. \quad (13.3.7)$$

Теорема IV

Закон распределения σ асимптотически нормален.

Смысл теорем II, III, IV вполне аналогичен содержанию соответствующих теорем, характеризующих закон распределения x или закон распределения p .

Остановимся на теореме I, имеющей некоторые принципиальные особенности. Она означает, что центр, вокруг которого колеблется σ , не совпадает с σ_0 , а несколько меньше последнего (так как множитель b_n меньше 1). Следовательно, оценивая σ_0 по σ , мы допускаем систематическое преуменьшение величины σ_0 .

Разность $M(\sigma) - \sigma_0$ называется *систематической погрешностью* или смещением величины σ . Это смещение исправляется тем, что показатель σ , полученный по данным опыта, делят на число b_n . Полученное значение

$$\hat{\sigma} = \frac{\sigma}{b_n} \quad (13.3.8)$$

называется *несмещенной оценкой* σ_0 или, иначе, σ -й, скорректированной на систематическую погрешность.

Существует другой способ корректировки σ , основанный на учете смещения σ^2 . Дело в том, что среднее значение σ^2 связано с σ_0^2 зависимостью

$$M(\sigma^2) = \sigma_0^2 \frac{n-1}{n}, \quad (13.3.9)$$

которая не только проще, но и универсальнее, чем зависимость (13.3.4) между средним значением σ и σ_0 , так как она справедлива при любом законе распределения x .

Из (13.3.9) вытекает следующий корректив σ^2 на систематическую погрешность

$$\hat{\sigma}^2 = \frac{n}{n-1} \sigma^2 = \frac{\sum (x - \bar{x})^2}{n-1}. \quad (13.3.10)$$

$$\hat{\sigma} = \sqrt{\frac{\sum (x - \bar{x})^2}{n-1}} \quad (13.3.11)$$

будем называть *квадратически несмещенной оценкой* показателя σ_0 .

Рассмотрим пример. Пусть в процессе (α) получены следующие результаты испытаний x : 2, 5, 1, 2.

$$\bar{x} = \frac{2+5+1+2}{4} = 2,5;$$

$$\begin{aligned} \sigma &= \sqrt{\frac{(2-2,5)^2 + (5-2,5)^2 + (1-2,5)^2 + (2-2,5)^2}{4}} = \\ &= \sqrt{\frac{0,25 + 6,25 + 2,25 + 0,25}{4}} = \sqrt{\frac{9}{4}} = 1,5. \end{aligned}$$

Несмещенная оценка при нормальном распределении x :

$$\hat{\sigma} \approx \frac{1,5}{1 - \frac{3}{4 \cdot 4}} = 1,85.$$

Квадратически несмещенная оценка

$$\hat{\sigma} = \sqrt{\frac{9}{3}} = \sqrt{3} = 1,73.$$

Следует заметить, что корректировка σ на систематическую погрешность имеет практическое значение лишь при малом числе наблюдений. При сколько-нибудь значительном n смещение σ настолько невелико, что корректировка является излишней.

Мы остановились на вопросах корректировки σ лишь по принципиальным соображениям, но, как увидим в следующем параграфе, существуют показатели, для которых корректировка на систематическую погрешность имеет практическое значение не только при малом числе наблюдений.

Величина случайных погрешностей показателей \bar{x} и σ характеризуется интервалом рассеяния данного показателя, гарантированным с определенной близкой к 1 вероятностью, аналогично случаю относительной частоты, разобранным в предыдущем параграфе.

Для показателя \bar{x} величина интервала рассеяния при большом n определяется по формуле

$$2\Delta_{\bar{x}} = \frac{2t\sigma_0}{\sqrt{n}}, \quad (13.3.12)$$

где t определяется из уравнения

$$\Phi(t) = P, \quad (13.3.13)$$

P — заданное, близкое к 1 значение гарантийной вероятности.

Если σ_0 не известно, то величина интервала рассеяния находится по приближенной формуле

$$2\Delta_{\bar{x}} \approx \frac{2t\sigma}{\sqrt{n}}. \quad (13.3.14)$$

Для показателя σ величина интервала рассеяния при большом n и нормальном законе распределения x определяется по формуле

$$2\Delta_{\sigma} = \frac{2t\sigma_0}{\sqrt{2n}}, \quad (13.3.15)$$

где t — корень уравнения

$$\Phi(t) = P \quad (13.3.16)$$

и приближенно, в случае, когда σ_0 неизвестно,

$$2\Delta_{\sigma} \approx \frac{2t\sigma}{\sqrt{2n}}. \quad (13.3.17)$$

Оценка неизвестных параметров \bar{x}_0 и σ_0 при большом n производится по правилу, изложенному в предыдущем параграфе, при аналогичной оценке параметра ρ_0 .

При гарантии $P = \Phi(t)$ для \bar{x}_0 имеем

$$\bar{x} - t \frac{\sigma}{\sqrt{n}} \leq \bar{x}_0 \leq \bar{x} + t \frac{\sigma}{\sqrt{n}}. \quad (13.3.18)$$

Для σ_0

$$\sigma - \frac{t\sigma}{\sqrt{2n}} \leq \sigma_0 \leq \sigma + \frac{t\sigma}{\sqrt{2n}}. \quad (13.3.19)$$

Пример 1. Были произведены расчеты \bar{x} и σ по данным 220 испытаний ударной вязкости стали 30ХГСА. $\bar{x} = 6,35$, $\sigma = 1,4$. Требуется указать пределы, в которых ожидаются теоретические значения \bar{x}_0 и σ_0 этих показателей, если считать, что процесс накопления материала соответствует процессу (α) .

Решение. Применяя формулы (13.3.18) и (13.3.19), найдем при гарантии $\Phi(t) = 0,990$

$$6,35 - 2,58 \frac{1,4}{\sqrt{220}} \leq \bar{x}_0 \leq 6,35 + 2,58 \frac{1,4}{\sqrt{220}}$$

или

$$6,11 \leq \bar{x}_0 \leq 6,59$$

и

$$1,4 - 2,58 \frac{1,4}{\sqrt{440}} \leq \sigma_0 \leq 1,4 + 2,58 \frac{1,4}{\sqrt{440}}$$

или

$$1,23 \leq \sigma_0 \leq 1,57.$$

При установлении необходимого числа наблюдений для оценки \bar{x}_0 и σ_0 исходят из заданной точности определения интересующего нас показателя так же, как это делалось в аналогичном случае для ρ_0 в предыдущем параграфе. Получаем следующие формулы для определения n :

1) при оценке \bar{x}_0

$$n = \frac{\sigma_0^2}{\left(\frac{\Delta \bar{x}}{t}\right)^2}; \quad (13.3.20)$$

2) при оценке σ_0

$$n = \frac{\sigma_0^2}{2 \left(\frac{\Delta_\sigma}{t} \right)^2}, \quad (13.3.21)$$

где $\Delta_{\bar{x}}$, $\Phi(t)$ и $2\Delta_\sigma$, $\Phi(t)$ — исходные параметры точности опыта. σ_0 находится ориентировочно по данным предварительных исследований.

Пример 2. Считая результаты, полученные в предыдущем примере, ориентировочными, указать, сколько необходимо сделать испытаний по схеме процесса (α) , чтобы при гарантии 0,997 получить интервалы рассеяния для \bar{x} и σ соответственно

$$2\Delta_{\bar{x}} = 0,30 \text{ и } 2\Delta_\sigma = 0,20.$$

Решение. Считая, что ориентировочно $\sigma_0 = 1,4$, получим для оценки \bar{x}_0 :

$$n = \frac{1,4^2}{\left(\frac{0,15}{3} \right)^2} = \frac{1,96}{0,0025} \approx 800,$$

для оценки σ_0

$$n = \frac{1,4^2}{2 \left(\frac{0,15}{3} \right)^2} \approx 400.$$

Следовательно, необходимо произвести около 800 испытаний.

Применение корреляционного эллипса к оценке параметров

В примере первом были построены по формулам (13.3.18) и (13.3.9) интервалы рассеяния для показателей \bar{x} и σ , гарантируемые с вероятностью 0,990. Эти интервалы были определены независимо друг от друга. Между тем наши оценки параметров могут быть улучшены, если рассматривать оба параметра не изолированно, а совместно. Будем исходить из допущения, что признак x подчиняется нормальному закону распределения, и опираться при этом на следующие два предложения [7].

а) Совместный закон распределения показателей \bar{x} и σ с увеличением n в процессе накопления (α) приближается к двумерному нормальному закону. Этот закон и будет положен в основу дальнейших расчетов.

б) Показатели \bar{x} и σ в процессе накопления (α) варьируют независимо друг от друга и, следовательно, теоретический коэффициент корреляции между ними равен нулю: $r_{\sigma\bar{x}} = 0$.

Рассмотрим в стандартизованном масштабе поле возможного рассеяния показателей \bar{x} и σ при фиксированном и достаточно большом n . На этом поле определенным ранее интервалам рассеяния \bar{x} и σ соответствуют полосы, параллельные осям. Они определяют область совместного рассеяния \bar{x} и σ в виде прямоугольника, а при равенстве гарантийных вероятностей для \bar{x} и σ — квадрата. Легко показать, что в силу независимости \bar{x} и σ

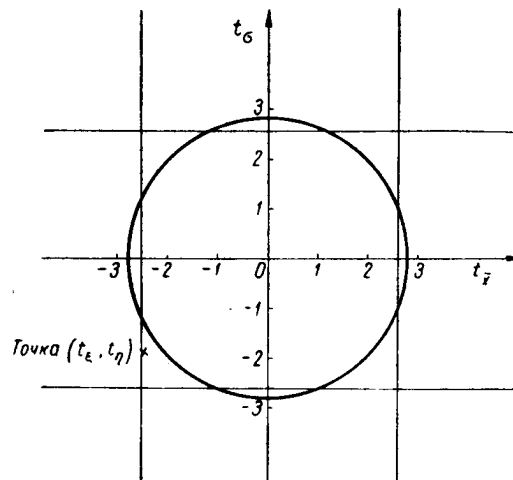


Рис. 69. Область совместного рассеяния показателей \bar{x} и σ , гарантированная с вероятностью 0,980

вероятность точке поля оказаться в очерченном прямоугольнике (квадрате) равна произведению вероятностей для соответствующих полос, т. е. $0,990 \cdot 0,990 = 0,980$. Следовательно, очерченный квадрат представляет собой область совместного рассеяния параметров \bar{x} и σ , гарантированную с вероятностью 0,980. Нетрудно однако убедиться в том, что это не лучшим образом определенная область, так как существуют точки, лежащие вне этой области и обладающие большей плотностью вероятности, чем некоторые точки, находящиеся внутри области (сравнительно мало вероятны точки области, находящиеся в углах квадрата). Поэтому можно с той же гарантийной вероятностью 0,980 определить область совместного рассеяния \bar{x} и σ , площадь которой меньше площади квадрата, если выбрать эту область надлежащим образом. Наиболее целесообразным является определение области по корреляционному эллипсу, который в данном случае (при

$t_{0\bar{x}\sigma} = 0$) в стандартизованном масштабе представляет собой окружность

$$t_x^2 + t_\sigma^2 = \chi^2, \quad (13.3.22)$$

где $Q(\chi) = 1 - e^{-\frac{\chi^2}{2}} = 0,980$.

Находим радиус $\chi = 2,80$ и проводим окружность (рис. 69). Эта окружность и является наилучшей областью совместного рассеяния параметров x и σ , которая может быть гарантирована с вероятностью 0,980. Ее площадь $S_o = \pi\chi^2 = 24,6$. В то же время площадь квадрата, построенного с той же гарантийной вероятностью 0,980, составляет

$$S_{\square} = (2 \cdot 2,58)^2 = 26,6.$$

Сокращение размеров области рассеяния означает повышение точности оценки.

Для того чтобы испытать, согласуется ли с опытными данными та или иная комбинация значений неизвестных параметров $\bar{x}_0 = \xi$, $\sigma_0 = \eta$, где ξ , η — испытываемые значения параметров при заданной гарантии $Q(\chi)$, следует:

1. Найти $t_\xi = \frac{\xi - \bar{x}}{\sigma_{0\bar{x}}}$ и $t_\eta = \frac{\eta - \sigma}{\sigma_{0\sigma}}$.

2. Если $t_\xi^2 + t_\eta^2$ окажется меньше чем χ^2 , то точка (ξ, η) окажется внутри гарантированной области рассеяния (окружности) и комбинация значений (ξ, η) допустима.

Если $t_\xi^2 + t_\eta^2$ окажется больше чем χ^2 , то точка (ξ, η) окажется вне гарантируемой области рассеяния и комбинация значений (ξ, η) не согласуется с опытными данными.

В виде примера испытаем значения $\bar{x}_0 = \xi = 6,11$, $\sigma_0 = \eta = 1,28$ для данных примера 1 (стр. 275). Здесь

$$\bar{x} = 6,35, \quad \sigma_{0\bar{x}} \approx \frac{1,4}{220} = 0,095, \quad \sigma = 1,4,$$

$$\sigma_{0\sigma} \approx \frac{1,4}{440} = 0,067, \quad Q(\chi) = 0,980, \quad \chi = 2,80, \quad \chi^2 = 7,84.$$

Находим

$$t_\xi = \frac{6,11 - 6,35}{0,095} = -2,52,$$

$$t_\eta = \frac{1,28 - 1,40}{0,067} = -1,79,$$

$$t_\xi^2 + t_\eta^2 = 6,35 + 3,20 = 9,55.$$

Так как $t_{\xi}^2 + t_{\eta}^2$ оказалось больше чем χ^2 , то комбинация $\xi=6,11$, $\eta=1,28$ не может быть принята. На рис. 69 видно, что точка (ξ, η) лежит вне окружности рассеяния, что и подтверждается нашим расчетом. Любопытно отметить, что каждое в отдельности значение $\xi=6,11$ и $\eta=1,28$ является допустимым значением при заданной гарантии 0,990. Недопустимой оказывается лишь комбинация этих значений.

§ 4. ПОКАЗАТЕЛИ КОРРЕЛЯЦИИ В ПРОЦЕССЕ НАКОПЛЕНИЯ МАТЕРИАЛА

Ознакомление с основными элементами теории вероятностей дает возможность уточнить некоторые понятия, которые рассматривались при изучении корреляции в разделе II книги, и научиться оценивать надежность получаемых опытным путем показателей корреляционной зависимости.

Рассмотрим совместное распределение материала по двум или нескольким варьирующим признакам x, y, z, \dots . Предположим, что материал накапливается по схеме процесса (α) , следовательно, совместный закон распределения вероятностей различных возможных значений x, y, z, \dots остается неизменным во всех испытаниях и не зависит от результатов уже произведенных испытаний. Если этот закон таков, что различным значениям признака y соответствуют различные значения условного математического ожидания признака x , то это означает, что признак x связан корреляционной зависимостью с y . Рассмотрение зависимости условного математического ожидания x одновременно от значений ряда других признаков y, z, \dots приводит к понятию множественной корреляции.

Чем больше число наблюдений, тем ближе к 1 вероятность того, что средние из опытных значений признака x , соответствующие тем или иным значениям признака y , будут сколь угодно мало отличаться от своих математических ожиданий. Поэтому действительная теоретическая линия регрессии, определяющая величины математических ожиданий x , соответствующих различным значениям y , действительно будет «предельной линией», к которой в смысле теории вероятностей стремятся как непосредственно получаемая из опытных данных эмпирическая линия регрессии, так и вычисляемая на основе конечного числа наблюдений «теоретическая линия регрессии».

Коэффициенты корреляции, регрессии и другие показатели, определяемые законом распределения вероятностей, будем обозначать общим символом P_0 и называть в дальнейшем *теоретическими* показателями. Аналогичные статистические показатели, рассчитываемые по материалам конечного числа испытаний для оценки показателей P_0 , обозначаем символом P и называем *опытными* или *эмпирическими* показателями.

Для оценки надежности тех выводов, которые можно делать

из опытных показателей корреляции, важно определять возможные величины отклонений $\Pi - \Pi_0$, называемых погрешностями или ошибками соответствующих показателей.

О законах распределения основных показателей корреляции Π (коэффициентов корреляции, коэффициентов регрессии, корреляционных отношений) известны следующие положения ([7] гл. 27, 28, 29, [2] гл. 6).

1. С увеличением числа испытаний n законы распределения показателей Π стягиваются к своим центрам, т. е. к математическим ожиданиям этих показателей, которые, в свою очередь, стремятся к пределам Π_0

$$\lim_{n \rightarrow \infty} M(\Pi) = \Pi_0. \quad (13.4.1)$$

Следовательно, Π_0 являются пределами по вероятности показателей Π

$$\lim_{N \rightarrow \infty} \Pi = \Pi_0. \quad (13.4.2)$$

2. Если рассматривать погрешность $\Pi - \Pi_0$ как сумму случайной погрешности $\Pi - M(\Pi)$ и систематической погрешности $M(\Pi) - \Pi_0$, то случайные погрешности показателей Π в среднем, как правило, сокращаются по закону обратной пропорциональности корню квадратному из числа испытаний; систематические погрешности у этих показателей либо отсутствуют, либо сокращаются, как правило, по закону обратной пропорциональности числу наблюдений, т. е. значительно быстрее, нежели случайные погрешности.

3. Законы распределения рассматриваемых показателей асимптотически нормальны; поэтому при значительных числах наблюдений оценка погрешности Π может производиться на основе нормального закона распределения, который, как известно, определяется двумя параметрами: $M(\Pi)$ и σ_{Π} . Исключение в некоторых случаях представляет совокупный коэффициент корреляции и корреляционное отношение (при $R_0=0$ и $\eta_0=0$).

Приводим формулы σ_{Π} и способ корректировки на систематическую погрешность для ряда основных показателей корреляции.

Коэффициенты корреляции

Пусть

$$\bar{x}_{1.23\dots p} = b_1 + b_2 x_2 + b_3 x_3 + \dots + b_p x_p \quad (13.4.3)$$

уравнение прямолинейной регрессии p переменных, где параметры b_1, b_2, \dots, b_p определяются по методу наименьших квадратов, как это изложено в разделе II.

Обозначим через

$$\sigma_{1.23\dots p} = \sqrt{\frac{\sum [x_1 - \bar{x}_{1.23\dots p}]^2}{n}} \quad (13.4.4)$$

среднее квадратическое отклонение x_1 от поверхности регрессии (13.4.3). Величина $\sigma_{1.23\dots p}$ играет важную роль при оценке надежности показателей корреляции.

$\sigma_{1.23\dots p}$ имеет систематическую ошибку в сторону преуменьшения, тем более значительную, чем больше параметров в уравнении регрессии определяется по опытным данным. Для исправления этих погрешностей вводятся коррективы

$$\hat{\sigma}_{1.23\dots p} = \sigma_{1.23\dots p} \sqrt{\frac{n}{n-p}} = \sqrt{\frac{\sum (x_1 - \bar{x}_{1.23\dots p})^2}{n-p}}, \quad (13.4.5)$$

где p — число параметров в уравнении регрессии, определяемых по данным опыта методом наименьших квадратов. Для прямолинейной зависимости p совпадает с числом переменных.

Коэффициент множественной корреляции $R_{1.23\dots p}$ выражается формулой

$$R_{1.23\dots p} = \sqrt{1 - \frac{\sigma_{1.23\dots p}^2}{\sigma_1^2}}, \quad (13.4.6)$$

которая получается из формулы (10.3.2) в случае прямолинейной зависимости.

Исправляя в формуле (13.4.6) величины $\sigma_{1.23\dots p}$ и σ_1 на систематическую погрешность, исправляем приближенно и $R_{1.23\dots p}$.

Формула коррекции

$$\hat{R}_{1.23\dots p} = \sqrt{1 - \frac{\hat{\sigma}_{1.23\dots p}^2}{\hat{\sigma}_1^2}} \quad (13.4.7)$$

или

$$\hat{R}_{1.23\dots p} = \sqrt{1 - (1 - R_{1.23\dots p}^2) \frac{n-1}{n-p}}, \quad (13.4.8)$$

где $\hat{R}_{1.23\dots p}$ — скорректированное значение $R_{1.23\dots p}$.

Этот метод корректирования R , хотя и не является точным, но он прост и практически себя оправдывает.

Корректирование R , как видно из самой формулы (13.4.8), имеет смысл применять лишь в тех случаях, когда число параметров p не очень мало по сравнению с числом наблюдений n , т. е. для корреляции многих переменных при не очень больших числах n .

Формула (13.4.8) представляет известный интерес в качестве критерия, на основании которого можно судить о целесообразности дополнительного включения в уравнение множественной регрессии тех или иных аргументов. В самом деле, при дополнительном включении в уравнение регрессии новых аргументов коэффициент множественной корреляции R , как правило, увеличивается. Однако это увеличение R может быть иллюзорным и являться лишь следствием появления в уравнении регрессии большего числа параметров, определяемых по данным опыта. Легко сообразить, что если бы, например, число таких параметров p оказалось равным числу наблюдений n , то рассчитанная по методу наименьших квадратов поверхность регрессии прошла бы через все n опытных точек. Следовательно, в этом случае $R_{1.23\dots p}$ окажется равной 1, а $R_{1.23\dots p}$ равным 0, хотя в действительности связь могла быть и слабой, что сразу обнаружилось бы при увеличении числа наблюдений. Чем ближе p к n , тем сильнее преувеличивается теснота связи. Это преувеличение и исправляется по формуле (13.4.8).

Средняя квадратическая погрешность коэффициента корреляции определяется по формуле

$$\sigma_{0R} = \frac{1 - R_0^2}{\sqrt{n}}, \quad (13.4.9)$$

где R_0 — теоретическое значение коэффициента корреляции, заменяемое обычно через R или \hat{R} . Подобная замена допустима, когда n значительно, а R не очень близко к 1.

Для коэффициента парной корреляции

$$r = \frac{\frac{1}{n} \sum (x - \bar{x})(y - \bar{y})}{\sigma_x \sigma_y}$$

корректировка на систематическую погрешность обычно не производится; σ_{0r} определяется по формуле

$$\sigma_{0r} = \frac{1 - r_0^2}{\sqrt{n}}, \quad (13.4.10)$$

где r_0 — теоретическое значение коэффициента корреляции — заменяется приближенно через r , если n велико, а r не близко к 1.

При наличии нормальной корреляции между переменными формулы (13.4.9) и (13.4.10) являются весьма точными. Они используются в качестве приближенных формул расчета, независимо от тех или иных предположений о характере корреляции.

Пример 1. В главе X рассматривалась зависимость коэрцитивной силы от балла зерна, содержания углерода и содержания серы на материале 190 испытаний железа «Армко». Коэффициенты корреляции между коэрцитивной силой и каждым из трех перечисленных выше аргументов составляли соответственно: $r_{12}=0,555$, $r_{13}=0,301$, $r_{14}=0,372$. Рассчитаем средние квадратические ошибки этих показателей:

$$\sigma_{0r_{12}} = \frac{1 - 0,555^2}{\sqrt{190}} = 0,050,$$

$$\sigma_{0r_{13}} = \frac{1 - 0,301^2}{\sqrt{190}} = 0,066,$$

$$\sigma_{0r_{14}} = \frac{1 - 0,372^2}{\sqrt{190}} = 0,063.$$

Следовательно, применяя формулу оценки для r_0 при больших n : $r - t\sigma_{0r} \leq r_0 \leq r + t\sigma_{0r}$ и приняв, например, гарантийную вероятность $P \approx \Phi(t) = 0,990$, получим $t = 2,58$ и границы для r_0 .

$$0,426 \leq r_0 \leq 0,684,$$

$$0,131 \leq r_0 \leq 0,471,$$

$$0,210 \leq r_0 \leq 0,534.$$

Величина $r - t\sigma_{0r}$ называется *гарантированным минимумом*, величина $r + t\sigma_{0r}$ — *гарантированным максимумом* коэффициента r_0 .

Если гарантированный минимум r_0 положителен или же гарантированный максимум отрицателен, то коэффициент корреляции имеет *достоверный знак*, в первом случае — положительный, во втором — отрицательный; в этих случаях говорят, что *зависимость достоверна*. Очевидно, что достоверность зависимости определяется отношением $\frac{|r|}{\sigma_{0r}}$: если это отношение превосходит коэффициент t , соответствующий принятой гарантии $\Phi(t)$, то зависимость достоверна, так как число 0 не содержится в интервале рассеяния показателя r , гарантированном с вероятностью $\Phi(t)$. В противном случае зависимость не достоверна. Для гарантии $\Phi(t) = 0,990$, которую мы в таких случаях рекомендуем, число $t \approx 2,58 \approx 2,6$. Отсюда правило: зависимость достоверна, если отношение $\frac{|r|}{\sigma_{0r}} \geq 2,6$.

В нашем примере наличие зависимостей практически достоверно, так как

$$\frac{|r_{12}|}{\sigma_{0r_{12}}} = \frac{0,555}{0,050} = 11,1,$$

$$\frac{|r_{13}|}{\sigma_{0r_{13}}} = \frac{0,301}{0,066} = 4,5,$$

$$\frac{|r_{14}|}{\sigma_{0r_{14}}} = \frac{0,372}{0,063} = 6,0,$$

что значительно превышает во всех трех случаях значение 2,6.

Коэффициент множественной корреляции между коэрцитивной силой и совокупностью трех аргументов составлял

$$R_{1.234} = 0,653.$$

Коррекция этого коэффициента на систематическую погрешность дает

$$\hat{R}_{1.234} = \sqrt{1 - (1 - 0,653^2) \frac{190 - 1}{190 - 4}} = 0,646.$$

Средняя квадратическая ошибка коэффициента R исчисляется по формуле

$$\sigma_{0R_{1.234}} = \frac{1 - 0,653^2}{\sqrt{190}} = 0,042$$

либо по формуле

$$\sigma_{0R_{1.234}} = \frac{1 - 0,646^2}{\sqrt{190}} = 0,042.$$

Отношение

$$\frac{\hat{R}_{1.234}}{\sigma_{0R_{1.234}}} = \frac{0,646}{0,042} = 15,4.$$

Зависимость достоверна.

Коэффициенты регрессии в процессе накопления материала

Коэффициенты регрессии не корректируются на систематические ошибки.

Средние квадратические ошибки коэффициентов регрессии определяются следующим образом.

Для парной регрессии $\bar{y}_x = a + bx$

$$\sigma_{ob} = \frac{\sigma_{oy.x}}{\sigma_{ox}\sqrt{n}} = \frac{\sigma_{oy}\sqrt{1-r_{oxy}^2}}{\sigma_{ox}\sqrt{n}}. \quad (13.4.11)$$

Из формулы видно, что коэффициенты регрессии определяются тем точнее, чем меньше рассеяние функционального признака y и чем больше рассеяние аргумента x , и чем теснее зависимость между y и x .

При больших значениях n предельные теоретические значения показателей в формуле (13.4.11) могут быть заменены соответствующими показателями, вычисленными по результатам испытаний

$$\sigma_{ob} = \frac{\sigma_{y.x}}{\sigma_x\sqrt{n}} = \frac{\sigma_y\sqrt{1-r_{xy}^2}}{\sigma_x\sqrt{n}}. \quad (13.4.12)$$

Пример. Уравнение регрессии коэрцитивной силы по баллу зерна (глава X, стр. 203)

$$\bar{x}_{1.2} = 0,38 + 0,128x_2,$$

причем

$$\sigma_1 = 0,205, \quad \sigma_2 = 0,893, \quad r_{12} = 0,555, \quad n = 190.$$

Найдем среднюю квадратическую ошибку коэффициента регрессии:

$$\sigma_{ob} = \frac{0,205}{0,893\sqrt{190}} \sqrt{1-0,555^2} = 0,014.$$

При гарантии $P = \Phi(t)$ $b - t\sigma_{ob} \leq b_0 \leq b + t\sigma_{ob}$. Если принять $\Phi(t) = 0,990$, то знак (+ или -) коэффициента регрессии следует считать статистически достоверным при $\frac{|b|}{\sigma_{ob}} > 2,6$.

В нашем случае $\frac{|b|}{\sigma_{ob}} = \frac{0,127}{0,014} = 9,0$. Положительная зависимость достоверна.

Для регрессии трех переменных $\bar{z}_{xy} = a + bx + cy$

$$\sigma_{ob} = \frac{\sigma_{oz.xy}}{\sigma_{ox.y}\sqrt{n}} = \frac{\sigma_{oz}\sqrt{1-R_{oz.xy}^2}}{\sigma_{ox}\sqrt{1-r_{oxy}^2}\sqrt{n}} \quad (13.4.13)$$

и

$$\sigma_{oc} = \frac{\sigma_{oz.xy}}{\sigma_{oy.x}\sqrt{n}} = \frac{\sigma_{oz}\sqrt{1-R_{oz.xy}^2}}{\sigma_{oy}\sqrt{1-r_{oxy}^2}\sqrt{n}} \quad (13.4.14)$$

и после замены при больших n теоретических показателей их опытными значениями получаем приближенно

$$\sigma_{0b} = \frac{\sigma_{z \cdot xy}}{\sigma_{xy} \sqrt{n}} = \frac{\sigma_z \sqrt{1 - R_{z \cdot xy}^2}}{\sigma_x \sqrt{1 - r_{xy}^2} \sqrt{n}} \quad (13.4.15)$$

и

$$\sigma_{0c} = \frac{\sigma_{z \cdot xy}}{\sigma_{y \cdot x} \sqrt{n}} = \frac{\sigma_z \sqrt{1 - R_{z \cdot xy}^2}}{\sigma_y \sqrt{1 - r_{xy}^2} \sqrt{n}}. \quad (13.4.16)$$

Пример. В главе IX зависимость между упругостью, размером замка и магнитной проницаемостью на материале 200 поршневых колец выражалась уравнением регрессии

$$\bar{z}_{xy} = 761,97 + 32,70x + 8,09y,$$

причем

$$\sigma_z = 83,14, \quad r_{xz} = 0,439;$$

$$\sigma_x = 1,05, \quad r_{yz} = 0,305;$$

$$\sigma_y = 2,72, \quad r_{xy} = 0,099.$$

Рассчитаем средние квадратические ошибки коэффициентов регрессии.

$$\sigma_{0b} = \frac{83,14}{1,05} \frac{\sqrt{1 - 0,512^2}}{\sqrt{1 - 0,099^2} \sqrt{200}} = 4,83,$$

$$\sigma_{0c} = \frac{83,14}{2,72} \frac{\sqrt{1 - 0,512^2}}{\sqrt{1 - 0,099^2} \sqrt{200}} = 1,87.$$

Отношение

$$\frac{|b|}{\sigma_{0b}} = \frac{32,70}{4,83} = 6,8,$$

$$\frac{|c|}{\sigma_{0c}} = \frac{8,09}{1,87} = 4,3.$$

Знаки коэффициентов достоверны.

Для регрессии многих переменных

$$\bar{x}_{1.23\dots p} = b_1 + b_2 x_2 + b_3 x_3 + \dots + b_p x_p,$$

$$\sigma_{0b_2} = \frac{\sigma_{01.23\dots p}}{\sigma_{02.34\dots p} \sqrt{n}} = \frac{\sigma_{01} \sqrt{1 - R_{01.23\dots p}^2}}{\sigma_{02} \sqrt{1 - R_{02.3\dots p}^2} \sqrt{n}} \quad (13.4.17)$$

и при большом n приближенно

$$\sigma_{0b_2} = \frac{\sigma_{1.23\dots p}}{\sigma_{2.34\dots p} \sqrt{n}} = \frac{\sigma_1 \sqrt{1 - R_{1.23\dots p}^2}}{\sigma_2 \sqrt{1 - R_{2.3\dots p}^2} \sqrt{n}}. \quad (13.4.18)$$

Аналогичные формулы для $\sigma_{0b_1}, \sigma_{0b_4}, \dots, \sigma_{0b_p}$.

Пример. Уравнение регрессии коэрцитивной силы по баллу зерна, содержанию углерода и содержанию серы для 190 испытаний железа «Армко» (глава X, стр. 213)

$$\bar{x}_{1.234} = 0,157 + 0,118x_2 + 11,12x_3 + 5,20x_4,$$

причем

$$\begin{aligned} \sigma_1 &= 0,205, & r_{23} &= 0,029; \\ \sigma_2 &= 0,893, & r_{24} &= 0,179; \\ \sigma_3 &= 0,0040, & r_{34} &= 0,334; \\ \sigma_4 &= 0,0082, & R_{1.234} &= 0,653. \end{aligned}$$

Рассчитаем средние квадратические ошибки коэффициентов регрессии. Предварительно необходимо рассчитать коэффициенты

$$R_{2.34}^2, R_{3.24}^2, R_{4.23}^2$$

$$\begin{aligned} R_{2.34}^2 &= \frac{r_{23}^2 + r_{24}^2 - 2r_{23}r_{24}r_{34}}{1 - r_{34}^2} = \\ &= \frac{0,029^2 + 0,179^2 - 2 \cdot 0,029 \cdot 0,179 \cdot 0,334}{1 - 0,334^2} = 0,0330, \end{aligned}$$

$$\begin{aligned} R_{3.24}^2 &= \frac{r_{32}^2 + r_{34}^2 - 2r_{32}r_{34}r_{24}}{1 - r_{24}^2} = \\ &= \frac{0,029^2 + 0,334^2 - 2 \cdot 0,029 \cdot 0,334 \cdot 0,179}{1 - 0,179^2} = 0,1125, \end{aligned}$$

$$\begin{aligned} R_{4.23}^2 &= \frac{r_{42}^2 + r_{43}^2 - 2r_{42}r_{43}r_{23}}{1 - r_{23}^2} = \\ &= \frac{0,179^2 + 0,334^2 - 2 \cdot 0,179 \cdot 0,334 \cdot 0,029}{1 - 0,029^2} = 0,1402. \end{aligned}$$

Далее

$$\sigma_{0b_2} = \frac{0,205 \sqrt{1 - 0,653^2}}{0,893 \sqrt{1 - 0,033} \sqrt{190}} = 0,0128,$$

$$\sigma_{0b_3} = \frac{0,205 \sqrt{1 - 0,653^2}}{0,004 \sqrt{1 - 0,1125} \sqrt{190}} = 2,993,$$

$$\sigma_{0b_4} = \frac{0,205 \sqrt{1 - 0,653^2}}{0,0082 \sqrt{1 - 0,1402} \sqrt{190}} = 1,476.$$

Отношения

$$\frac{|b_2|}{\sigma_{0b_2}} = \frac{0,118}{0,013} = 9,0,$$

$$\frac{|b_3|}{\sigma_{0b_3}} = \frac{11,12}{2,99} = 3,7,$$

$$\frac{|b_4|}{\sigma_{0b_4}} = \frac{5,20}{1,48} = 3,6$$

во всех трех случаях больше, чем 2,6. Следовательно, знаки коэффициентов достоверны.

Наличие нормальной корреляции между переменными является и здесь достаточным, хотя и не необходимым условием использования формул (13.4.11) и т. д. до (13.4.18). На практике эти формулы применяют в качестве приближенных, не считаясь со строгим соблюдением условия нормальной корреляции.

Теоретическое корреляционное отношение ¹

Пусть $\bar{y}_x = a + bp(x) + cq(x) + \dots$ — уравнение регрессии y по x , линейное относительно параметров a, b, c, \dots , определяемых по опытным данным методом наименьших квадратов.

Теоретическое корреляционное отношение y по x выражается, как известно, формулой

$$\eta_{T_{y/x}} = \sqrt{1 - \frac{\sigma_{T_{y,x}}^2}{\sigma_y^2}}.$$

Заменяя в этой формуле показатели σ_y^2 и $\sigma_{T_{y,x}}^2$ их значениями, скорректированными по формулам

¹ Термин «теоретический показатель» в разделе III книги всегда имеет смысл предельного теоретического показателя или показателя, исчисляемого посредством вероятностей. В данном случае этот термин употреблен в другом смысле в соответствии с понятиями раздела II. Он означает здесь, что корреляционное отношение исчисляется относительно линии регрессии, выражающейся математической формулой, параметры которой рассчитаны по опытным данным.

$$\left. \begin{aligned} \hat{\sigma}_y^2 &= \frac{n}{n-1} \sigma_y^2, \\ \hat{\sigma}_{Ty, x}^2 &= \frac{n}{n-p} \sigma_{y, x}^2, \end{aligned} \right\}, \quad (13.4.19)$$

получаем формулу приближенной коррекции показателя $\eta_{y/x}$, вполне аналогичную формуле (13.4.8)

$$\hat{\eta}_{Ty/x} = \sqrt{1 - \left(1 - \eta_{Ty/x}^2\right) \frac{n-1}{n-p}}, \quad (13.4.20)$$

где p — число параметров уравнения регрессии.

Формула (13.4.20) сохраняет свой вид и для множественного теоретического корреляционного отношения $\eta_{T1(m)}$ если уравнение регрессии линейно относительно параметров b_1, b_2, \dots, b_p .

$$\hat{\eta}_{T1(m)} = \sqrt{1 - \left[1 - \eta_{T1(m)}^2\right] \frac{n-1}{n-p}}, \quad (13.4.21)$$

где p — число параметров уравнения регрессии, определяемых по данным опыта.

Средняя квадратическая ошибка теоретического корреляционного отношения (парного или множественного) вычисляется по формуле

$$\sigma_{\eta_T} = \frac{1 - \eta_0^2}{\sqrt{n}}, \quad (13.4.22)$$

в которой η_0 обычно заменяется через η_T или $\hat{\eta}_T$ (если n велико и η_T не близко к 1).

Пример. Для зависимости между упругостью и размером замка $\eta_{Ty/x} = 0,475$, $p = 3$, $n = 200$.

Следовательно,

$$\begin{aligned} \hat{\eta}_{Ty/x} &= \sqrt{(1 - 0,226) \frac{200-1}{200-3}} = \sqrt{0,2182} = 0,467, \\ \sigma_{\eta_T} &= \frac{1 - 0,2182}{\sqrt{200}} = 0,055. \end{aligned}$$

Эмпирическое корреляционное отношение

Эмпирическое корреляционное отношение (множественное или парное) определяется формулой

$$\eta_{1(m)} = \sqrt{1 - \frac{\sigma_{1(m)}}{\sigma_1^2}}, \quad (13.4.23)$$

где

$$\sigma_{1(m)} = \sqrt{\frac{\sum (x_i - \bar{x}_{1(m)})^2}{n}} \quad (13.4.24)$$

среднее квадратическое отклонение x_1 относительно эмпирической поверхности множественной регрессии; $x_{1(m)}$ — частные средние x_1 , соответствующие заданным значениям аргументов.

Показатель $\sigma_{1(m)}^2$ обладает систематической погрешностью в сторону преуменьшения, размер которой зависит от числа групп k , на которые разбивается материал для вычисления частных средних $\bar{x}_{1(m)}$. Чем больше число k , чем оно ближе к числу наблюдений n , тем сильнее преуменьшение. Если представить себе крайний случай, при котором $k=n$ и, следовательно, в каждой группе имеется лишь одно наблюдение, то $\sigma_{1(m)}^2$ будет равняться нулю, каково бы ни было предельное теоретическое значение этого показателя.

Коррекция показателя $\sigma_{1(m)}^2$ на систематическую погрешность может быть произведена по формуле

$$\hat{\sigma}_{1(m)}^2 = \sigma_{1(m)}^2 \frac{n}{n-k} = \frac{\sum [x_i - \bar{x}_{1(m)}]^2}{n-k}. \quad (13.4.25)$$

Заменяя в формуле (13.4.25) $\sigma_{1(m)}^2$ и σ_1^2 их скорректированными значениями, получаем приближенную коррекцию эмпирического корреляционного отношения

$$\hat{\eta}_{1(m)} = \sqrt{1 - \frac{\hat{\sigma}_{1(m)}^2}{\hat{\sigma}_1^2}} = \sqrt{1 - (1 - \eta_{1(m)}^2) \frac{n-1}{n-k}}, \quad (13.4.26)$$

где k — число групп, на которые разбивается материал.

Средняя квадратическая ошибка эмпирического корреляционного отношения (парного или множественного) вычисляется по формуле

$$\sigma_{0\eta} = \frac{1 - \eta_0^2}{\sqrt{n}}, \quad (13.4.27)$$

где η_0 обычно заменяют через η или $\hat{\eta}$, если n велико и η не близко к 1.

Пример. Для зависимости между упругостью и размером замка поршневых колец

$$\eta_{y/x} = 0,517, \quad n = 200, \quad k = 11.$$

Следовательно,

$$\hat{\eta} = \sqrt{1 - (1 - 0,517^2) \frac{200 - 1}{200 - 11}} = 0,478,$$
$$\sigma_{0\eta} = \frac{1 - 0,478^2}{\sqrt{200}} = 0,055.$$

Различные примеры использования теории ошибок статистических показателей читатель найдет в курсах математической статистики. См., напр., [6].

РАЗДЕЛ IV

**КОМПЛЕКСНЫЕ ПРИМЕРЫ ИСПОЛЬЗОВАНИЯ
КОРРЕЛЯЦИОННОГО АНАЛИЗА НА ПРОИЗВОДСТВЕ¹**

ГЛАВА XIV

**ВЛИЯНИЕ ХИМИЧЕСКОГО СОСТАВА НА УДАРНУЮ
ВЯЗКОСТЬ СТАЛИ²**

§ 1. Ударная вязкость стали зависит от технологического процесса выплавки и передела стали, ее химического состава и термической обработки. Химический состав стали, как известно, оказывает существенное влияние на ее ударную вязкость. Обычно при экспериментальном исследовании влияния содержания какого-либо химического элемента на величину ударной вязкости стали не удается сохранить содержание других элементов на постоянном уровне. Поэтому влияние изучаемого химического элемента искажается влиянием других элементов.

Обработка опытных данных и данных текущего контроля с помощью методов математической статистики дает возможность установить степень влияния каждого из химических элементов на величину ударной вязкости при закреплении на постоянном уровне содержания других химических элементов.

Задачей настоящего исследования является установление влияния содержания каждого из химических элементов на ударную вязкость стали 30ХГСА путем корреляционного анализа материалов текущего контроля плавок. Решение этой задачи позволило бы более обоснованно подойти к установлению технических условий (ГОСТ 4543—48) на содержание химических элементов в стали 30ХГСА с целью получения наилучших свойств последней.

Материалом исследования явились данные контрольных испытаний 220 плавок стали, проведенных заводом «Серп и молот». Были использованы следующие данные протоколов испытаний

¹ В четвертом разделе значок «о» в обозначениях теоретических показателей для упрощения записи опускается.

² Я. И. Лукомский, Н. К. Лебедева, Р. В. Каткова, «Стандартизация» № 1 и 2, 1954.

каждой плавки: номер печи, размер и профиль прокатанных прутков, химический состав стали каждой плавки, способ приготовления образца (поперек или вдоль волокна), результаты испытаний каждой плавки на ударную вязкость по двум образцам. Все плавки были проконтролированы на ударную вязкость на продольных образцах и 116 из них — на поперечных образцах.

Как выше указывалось, предварительным условием каждого корреляционного исследования является, во-первых, проверка однородности подлежащего обработке материала и отграничение неоднородных групп, которые должны быть исследованы раздельно; во-вторых, установление области изменения признаков и характера их изменчивости. Решение указанных вопросов и являлось содержанием первого этапа работы.

§ 2. По существующим условиям приемки стали 30ХГСА каждая партия-плавка контролируется по ударной вязкости на двух образцах типа Менаже, прошедших одинаковую термическую обработку (закалка 890°, отпуск 510°, охлаждение в масле). За показатель ударной вязкости плавки принято среднее значение ее для двух образцов данной плавки

$$a_{k \text{ сред}} = \frac{a_{k_1} + a_{k_2}}{2}.$$

В дальнейшем изложении под величиной ударной вязкости плавки будем понимать $a_{k \text{ сред}}$.

Для проверки однородности материала было проведено сопоставление показателей ударной вязкости стали, выплавленной в различных однотипных основных мартеновских печах средней емкости. Распределение четырех групп плавков по величине ударной вязкости, отвечающее четырем различным печам (рис. 70), показывает, что все печи дают продукцию, однородную по величине ударной вязкости. На рис. 70 приведены также основные параметры рядов распределения ударной вязкости для каждой из четырех печей: средний показатель

$$\bar{x} = \frac{\sum xm}{\sum m}$$

и среднее квадратическое отклонение

$$\sigma = \sqrt{\frac{\sum (x - \bar{x})^2 m}{\sum m}}.$$

Оценка наибольшего расхождения значений средних арифметических с помощью критерия

$$t_{1,2} = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{\sigma_{x_1}^2 + \sigma_{x_2}^2}}$$

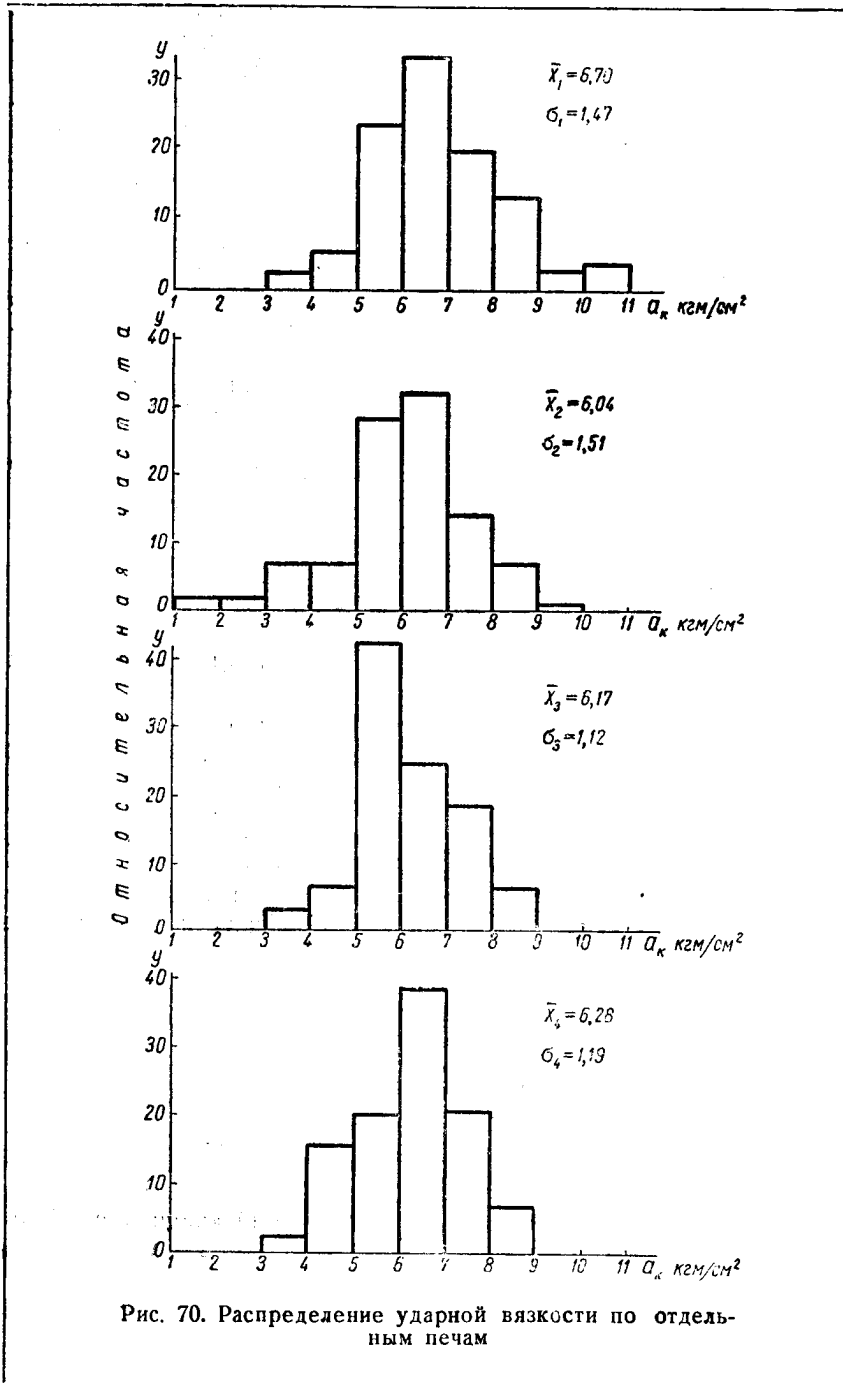


Рис. 70. Распределение ударной вязкости по отдельным печам

и средних квадратических отклонений с помощью критерия ¹

$$t_{2,3} = \frac{|\sigma_2 - \sigma_3|}{\sqrt{\sigma_{\sigma_2}^2 + \sigma_{\sigma_3}^2}}$$

показывает, что основные параметры рядов распределения, характеризующих качество выплавленной в каждой из печей стали, отличаются друг от друга в пределах случайных отклонений ($\mu < 2,6$).

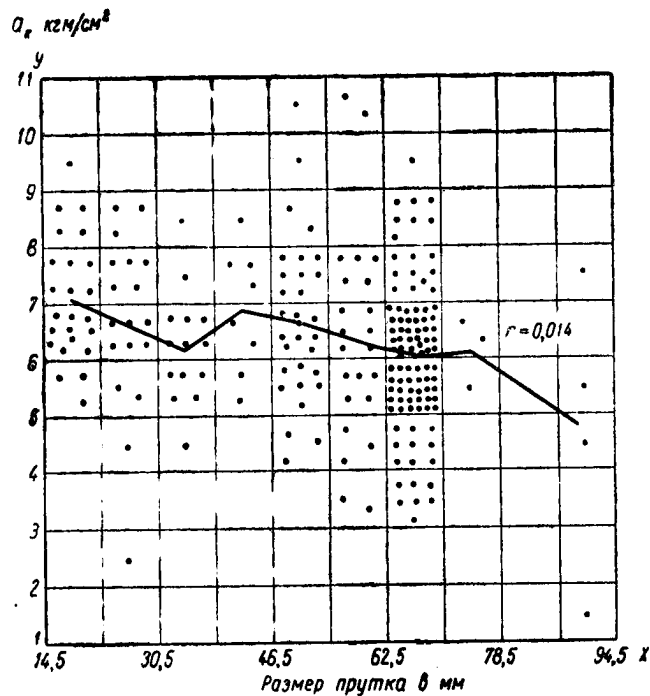


Рис. 71. Зависимость между ударной вязкостью и размером прутка

Не обнаружено также сколько-нибудь существенной зависимости между величиной ударной вязкости и размером прутка. Об этом свидетельствует рис. 71, на котором представлено поле корреляции между размером прутка и величиной ударной вязкости. Среднее значение ударной вязкости в каждой из групп сорта остается практически постоянным. Намечается лишь, на первый взгляд, небольшая тенденция к уменьшению ударной вязкости

¹ См., напр., В. И. Романовский, Применение математической статистики в опытном деле, 1947, гл. 4 и 5.

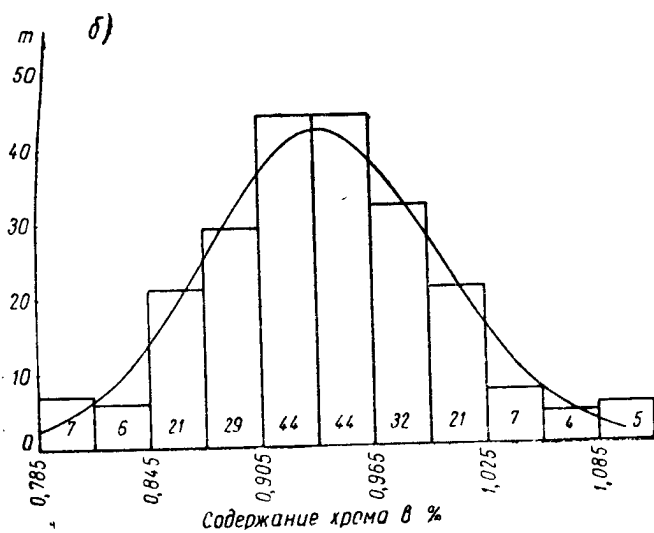
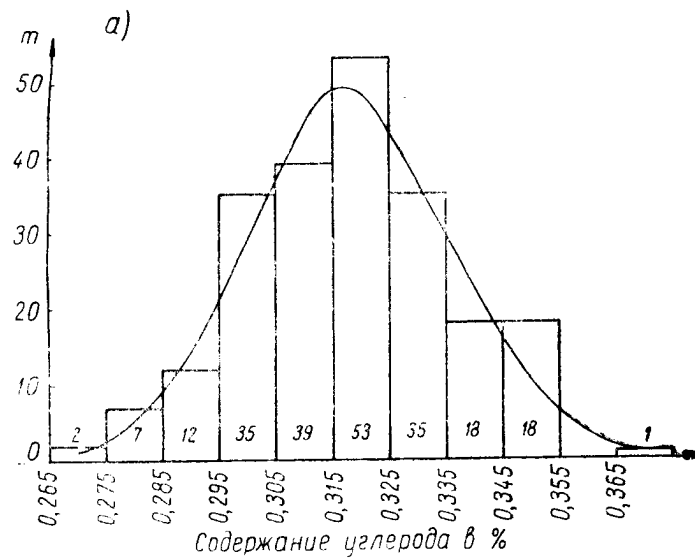
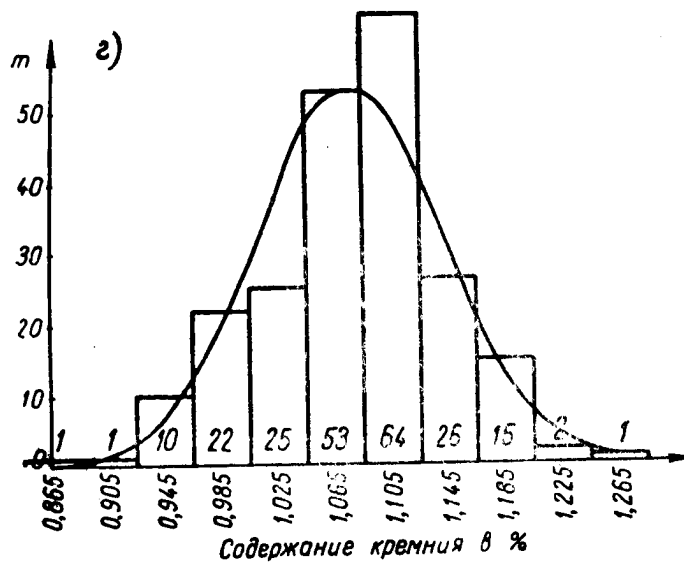
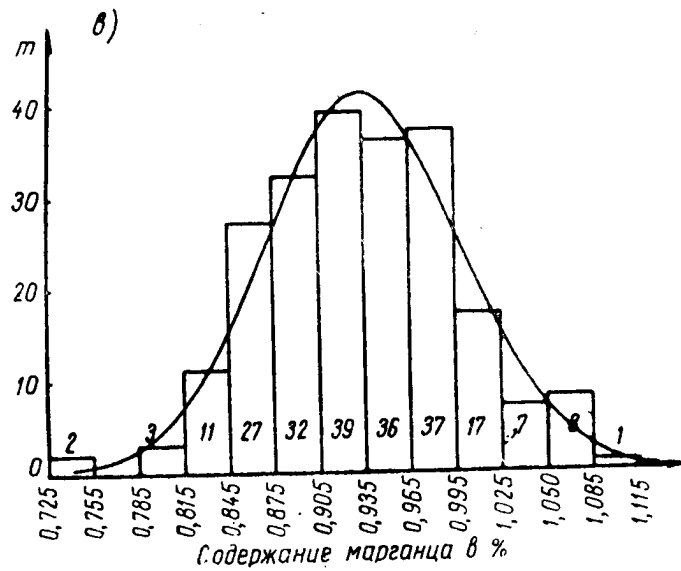


Рис. 72. Распределение 220 плавков стали 30ХГСА по содержанию в) марганца,



химических элементов: а) углерода, б) хрома, г) кремния

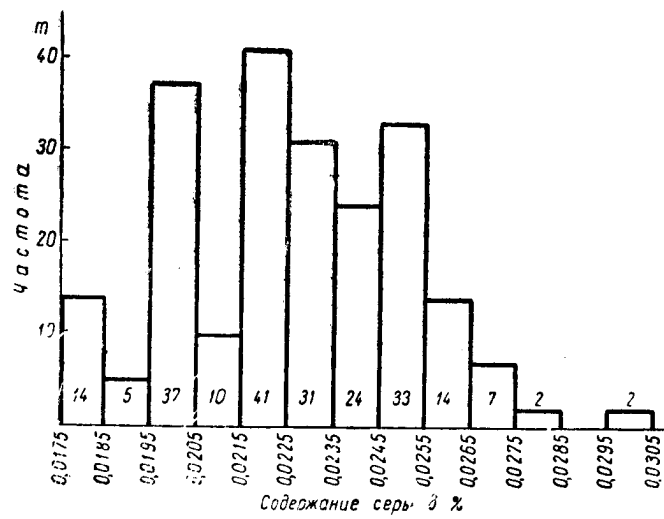
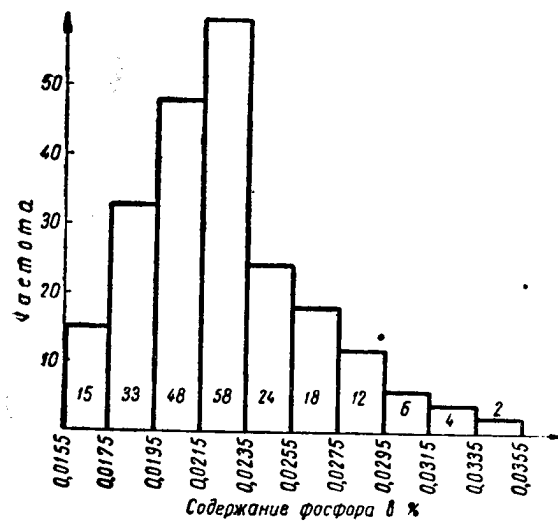
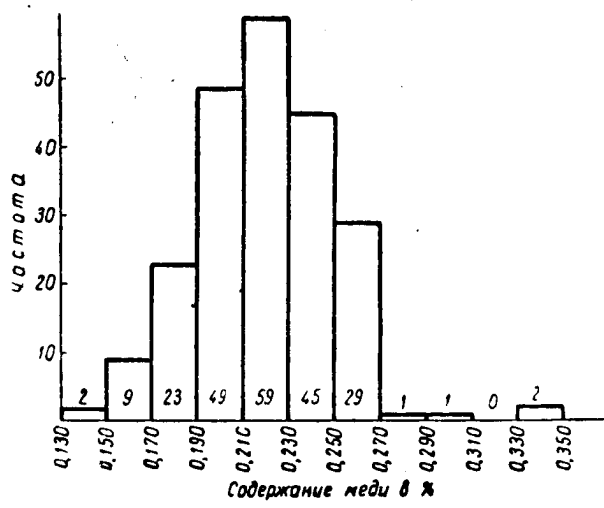
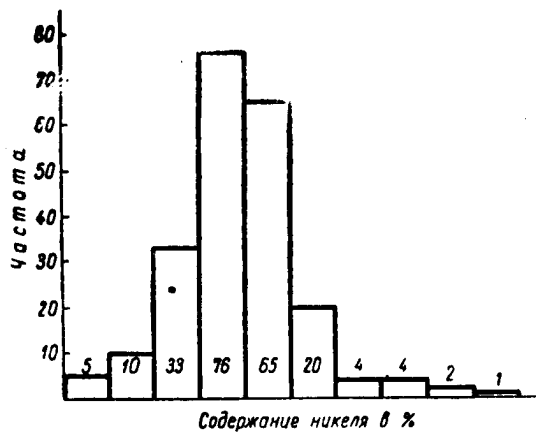


Рис. 73. Распределение плавок по содержанию:



а) фосфора, б) серы, в) никеля, г) меди

для относительно больших размеров прутков. Однако коэффициент корреляции r , характеризующий тесноту связи между размером сорта и величиной ударной вязкости, оказывается практически равным нулю ($r=0,014$ при величине случайной квадратической ошибки $\sigma_r=0,067$).

На основании проведенного изучения однородности исследуемого материала был сделан вывод, что при установлении искомым зависимостей между величиной ударной вязкости и химическим составом стали различия в размерах прутков и индивидуальные особенности плавильных агрегатов не нарушают однородности материала.

Ниже приводятся результаты статистического изучения изменчивости химических элементов 220 плавок стали 30ХГСА. На рис. 72 и 73 представлены ряды распределения плавок по содержанию химических элементов. В табл. 64 сведены статистические

Таблица 64

Химические элементы / Статистические показатели рядов	$\bar{x} \pm \sigma_{\bar{x}}$	$\sigma \pm \sigma_{\sigma}$	$A \pm \sigma_A$
Углерод (C)	0,32 ± 0,001	0,018 ± 0,001	0,02 ± 0,17
Хром (Cr)	0,94 ± 0,004	0,062 ± 0,003	0,17 ± 0,17
Марганец (Mn)	0,93 ± 0,004	0,061 ± 0,003	0,004 ± 0,17
Кремний (Si)	1,08 ± 0,04	0,065 ± 0,003	0,28 ± 0,17
Фосфор (P)	0,0224 ± 0,0003	0,0038 ± 0,0002	—0,78 ± 0,17
Сера (S)	0,0227 ± 0,0002	0,0025 ± 0,0001	0,77 ± 0,17
Никель (Ni)	0,162 ± 0,002	0,028 ± 0,001	0,58 ± 0,17
Медь (Cu)	0,219 ± 0,002	0,031 ± 0,001	0,32 ± 0,17

Продолжение

Химические элементы / Статистические показатели рядов	$E \pm \sigma_E$	$V \pm \sigma_V$	Содержание в процентах по ГОСТ 4543—48
Углерод (C)	0,22 ± 0,33	5,74 ± 0,27	0,28—0,35
Хром (Cr)	0,10 ± 0,33	6,66 ± 0,32	0,80—1,10
Марганец (Mn)	0,014 ± 0,33	6,87 ± 0,33	0,80—1,10
Кремний (Si)	0,15 ± 0,33	6,05 ± 0,29	0,90—1,20
Фосфор (P)	0,45 ± 0,33	16,93 ± 0,83	≤ 0,035
Сера (S)	0,03 ± 0,33	10,95 ± 0,53	≤ 0,03
Никель (Ni)	1,92 ± 0,33	17,0 ± 0,83	≤ 0,40
Медь (Cu)	1,31 ± 0,33	13,91 ± 0,68	≤ 0,25

показатели, характеризующие эти ряды распределения: средняя арифметическая \bar{x} ; среднее квадратическое отклонение σ ; показатель асимметрии

$$A = \frac{\mu_3}{\sigma^3},$$

где

$$\mu_3 = \frac{\sum (x - \bar{x})^3 m}{\sum m};$$

показатель эксцесса¹

$$E = \frac{\mu_4}{\sigma^4} - 3,$$

где $\mu_4 = \frac{\sum (x - \bar{x})^4 m}{\sum m}$; коэффициент вариации $V = 100 \frac{\sigma}{\bar{x}}$ и средние квадратические ошибки этих показателей

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}, \quad \sigma_{\sigma} = \frac{\sigma}{\sqrt{2n}}, \quad \sigma_A = \sqrt{\frac{6}{n}},$$

$$\sigma_E = \sqrt{\frac{24}{n}}, \quad \sigma_V = \frac{V}{\sqrt{2n}} \sqrt{1 + 2 \left(\frac{V}{100}\right)^2}.$$

В таблице указаны также допустимые стандартом пределы содержания химических элементов стали.

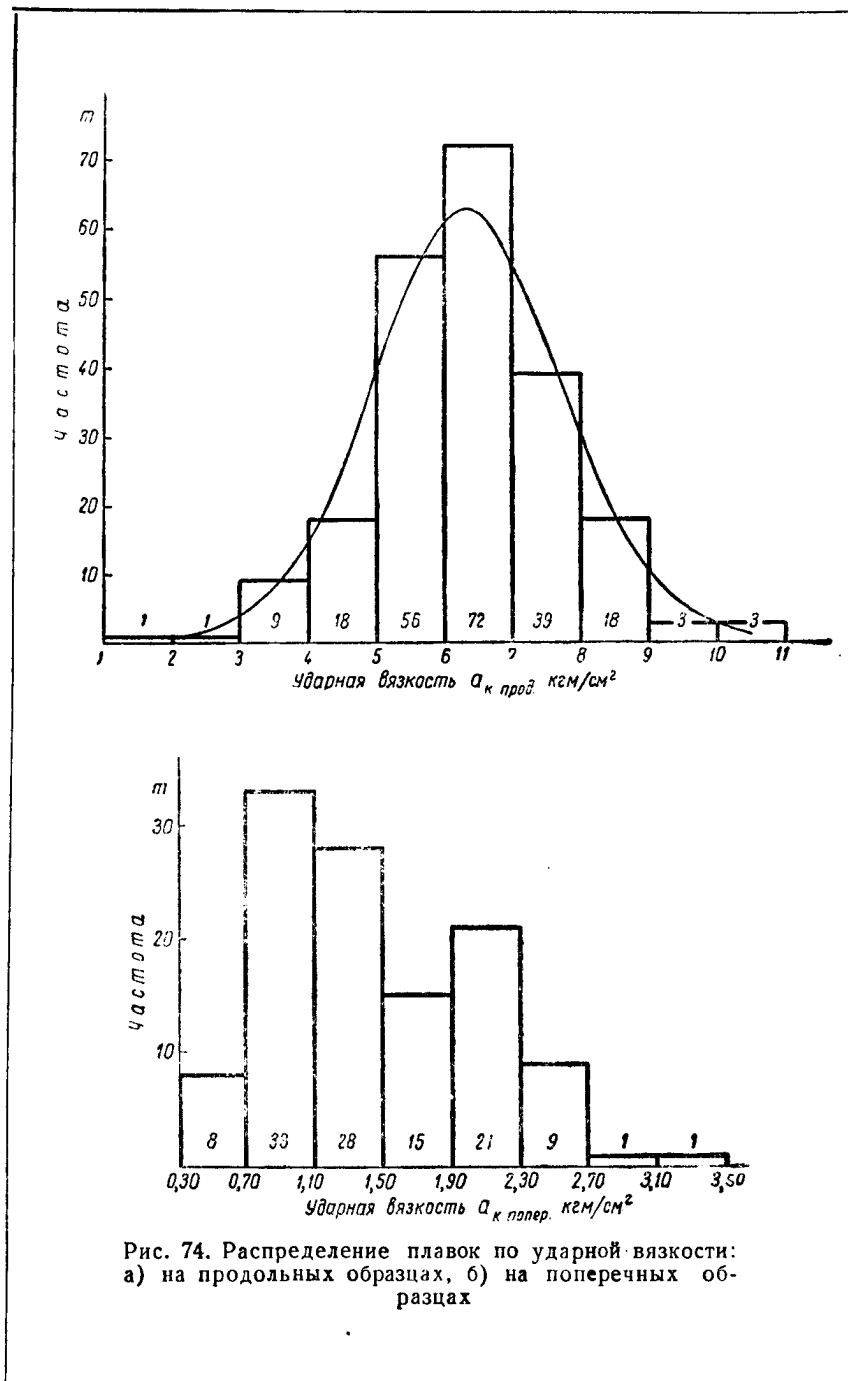
На основании проведенного исследования изменчивости химических элементов были сделаны следующие выводы:

1. Распределения плавок по содержанию углерода, хрома, марганца, кремния, для которых установлены двусторонние допуски, характеризуются кривыми, удовлетворяющими нормальному закону распределения; показатели асимметрии и эксцесса для этих кривых незначительны, их значения находятся в пределах случайных ошибок этих показателей. Колеблемость процентного содержания этих элементов соответствует установленным техническими условиями допускам. Средние значения процентного содержания примерно соответствуют серединам полей допусков. Коэффициенты вариации находятся на уровне 6—7% от средних значений признаков.

2. Распределения остальных химических элементов: фосфора, серы, никеля, меди, для которых установлены односторонние допуски, отличаются от нормального закона. Коэффициенты вариации содержания этих элементов находятся на уровне 11—17% от средних значений.

Распределение плавок по содержанию меди оказалось с «отрезанным» правым краем, что свидетельствует о неправильном

¹ О показателях асимметрии и эксцесса см. [6], [7], [8], [10], [11], [12].



отнесении некоторого числа плавок (при химическом анализе) с процентным содержанием меди (превосходящим предельную величину 0,25) к значениям 0,25. Максимальное содержание никеля, серы и фосфора не превышает установленных техническими условиями допусков.

Изменчивость ударной вязкости изучалась отдельно для испытаний на продольных и поперечных образцах. На рис. 74 приведены соответствующие кривые распределения, а в табл. 65 даны основные параметры этих распределений.

Для ударной вязкости в продольных образцах ($a_{к прод}$) была установлена кривая распределения, несколько уклоняющаяся от нормальной формы, с высокой вершиной ($E=0,83$). Это отклонение можно объяснить неправильным отнесением лаборантами части замеров с a_k менее 5 кгм/см² в интервал от 5 до 6 кгм/см², т. е. в поле допуска, как это отчетливо видно по кривой распределения. Кривая значительно выходит за допускаемые пределы.

Таблица 65

Образцы	Показатели рядов распределения ударной вязкости				
	$\bar{x} \pm \sigma_{\bar{x}}$	$\sigma \pm \sigma_{\sigma}$	$A \pm \sigma_A$	$E \pm \sigma_E$	$V \pm \sigma_V$
Продольные	$6,35 \pm 0,10$	$1,4 \pm 0,07$	$0,02 \pm 0,17$	$0,83 \pm 0,33$	$22,03 \pm 1,10$
Поперечные	$1,45 \pm 0,06$	$0,61 \pm 0,04$	$0,50 \pm 0,23$	$-0,48 \pm 0,45$	$41,99 \pm 3,2$

Для плавок, испытанных на поперечных образцах, имеет место резкое несоответствие поля рассеяния кривой распределения техническим условиям.

Кривая распределения ударной вязкости в поперечных образцах ($a_{к попереч}$) имеет вид, характерный для распределения так называемых существенно положительных величин, т. е. величин с односторонней вариацией.

§ 3. Конечной целью исследования являлось установление комплексной зависимости между ударной вязкостью и содержанием нескольких химических элементов. Для решения этой задачи необходимо было предварительно изучить парные зависимости между ударной вязкостью и содержанием каждого химического элемента в отдельности, а также между содержанием различных химических элементов.

Для исследования корреляционной зависимости между величиной ударной вязкости и содержанием каждого из химических элементов производилось построение полей корреляции (рис. 75 и 76).

Исследование формы связи производилось путем построения и анализа эмпирической линии регрессии.

Из рис. 75 и 76 видно, что линии регрессии $a_{к прод}$ и $a_{к попереч}$ по содержанию каждого из элементов имеют приближенно прямой характер; на рисунках приведены теоретические линии регрессии и их уравнения.

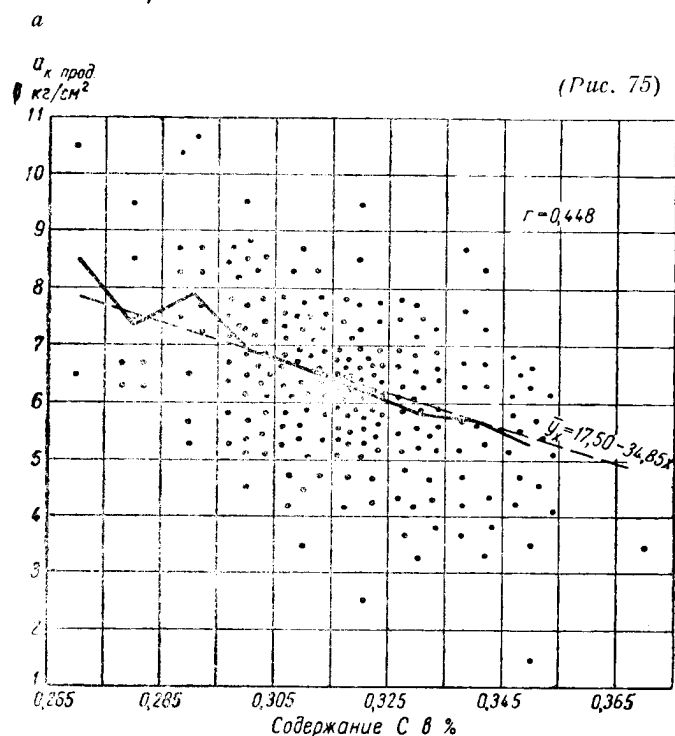
В качестве показателя тесноты связи определялся коэффициент корреляции r

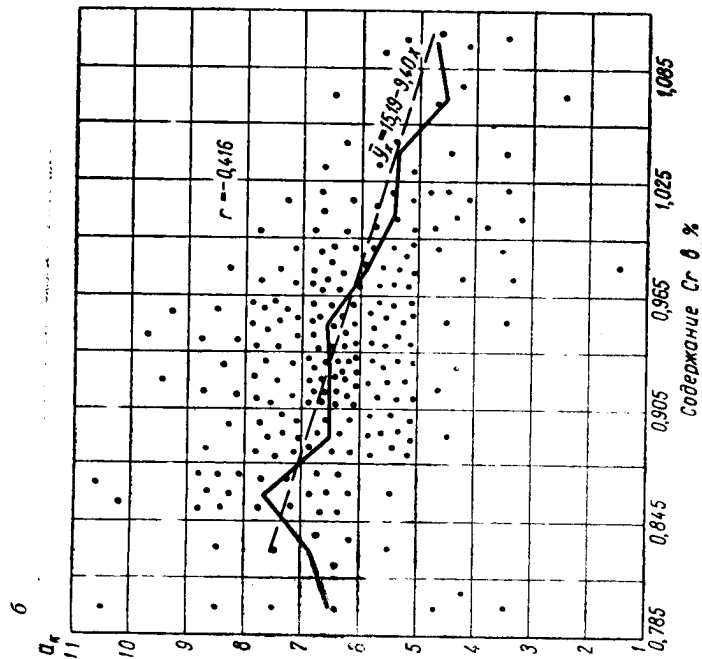
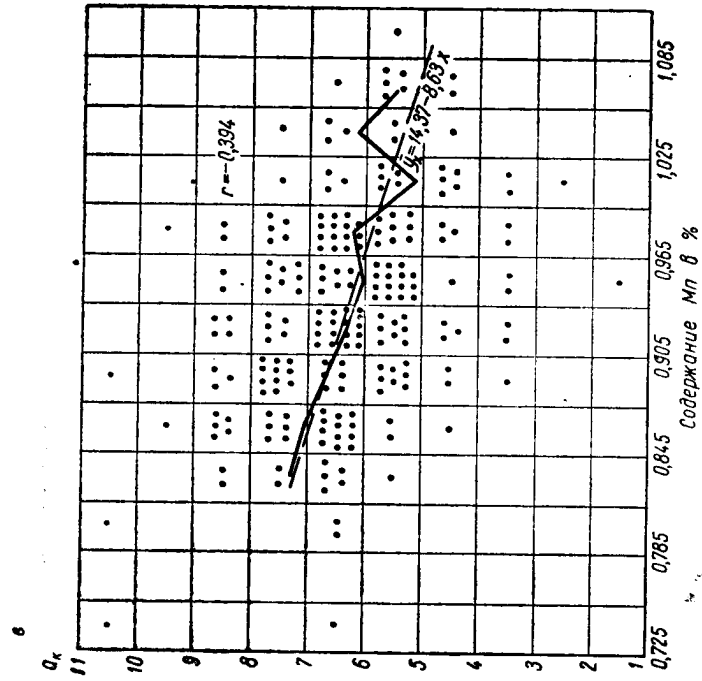
$$r = b \frac{\sigma_x}{\sigma_y}$$

Случайная средняя квадратическая ошибка коэффициента корреляции определялась по формуле

$$\sigma_r = \frac{1 - r^2}{\sqrt{n}}$$

Критерий надежности коэффициента корреляции рассчитывался как $\mu = \frac{|r|}{\sigma_r}$. Если $\mu > 2,6$, связь между признаками счита-





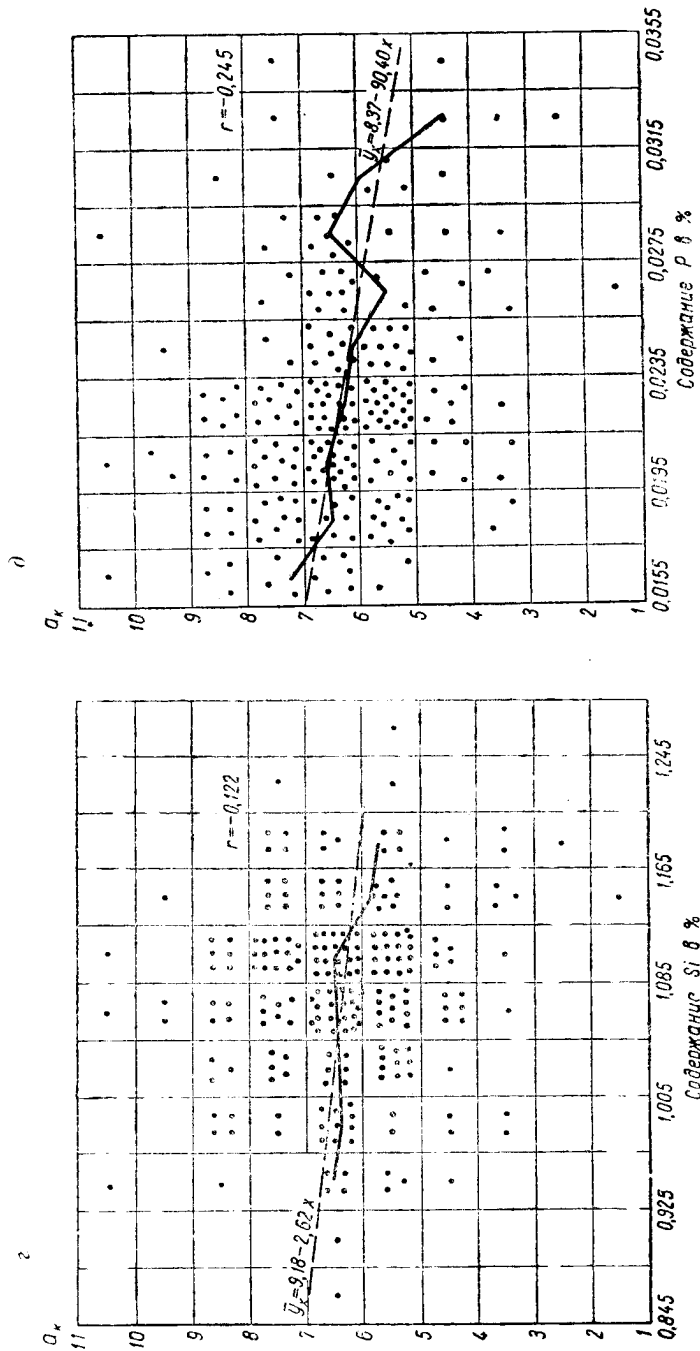
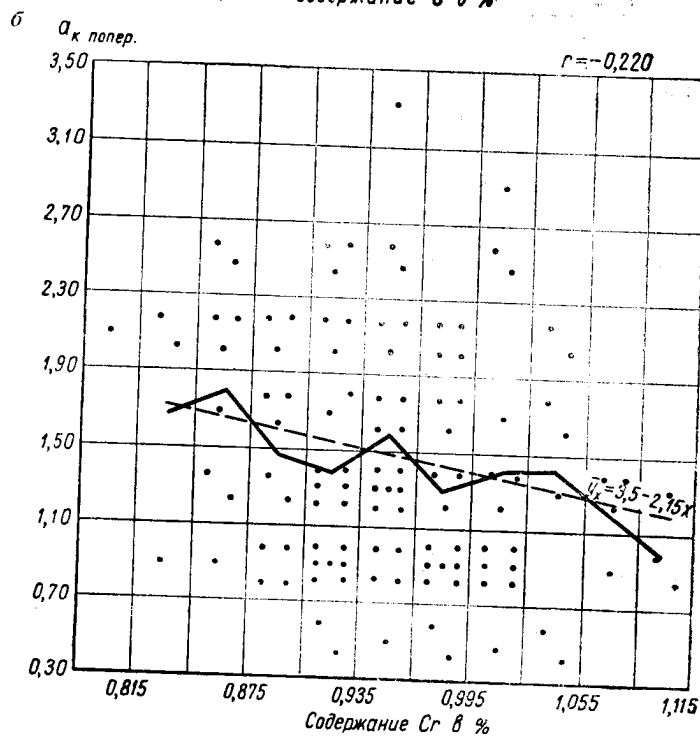
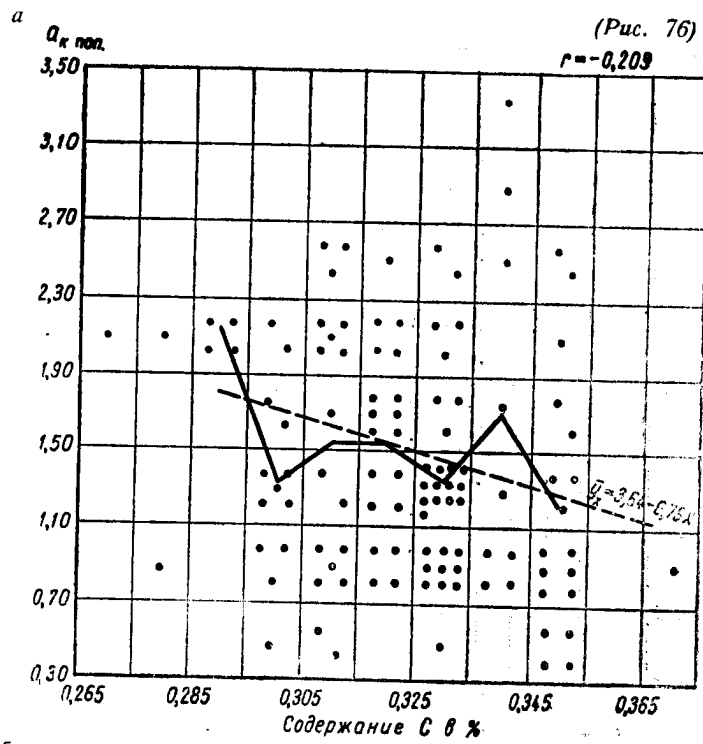


Рис. 75. Зависимость a_k от содержания химических элементов:
а) углерода, б) хрома, в) марганца, г) кремния, д) фосфора



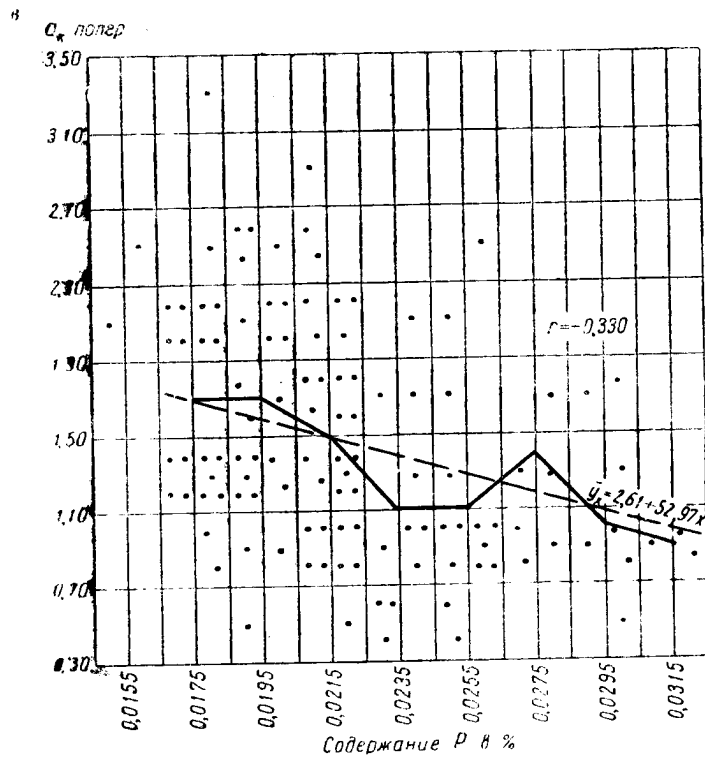


Рис. 76. Зависимость a_k попер от содержания химических элементов: а) углерода, б) хрома, в) фосфора

лась не случайной. В табл. 66 приведены коэффициенты корреляции между ударной вязкостью и содержанием каждого из химических элементов.

В результате проведенного исследования установлена существенная отрицательная зависимость между ударной вязкостью на продольных образцах и содержанием следующих химических элементов: углерода, хрома, марганца, фосфора. Не вполне достоверной оказалась зависимость между ударной вязкостью и содержанием кремния. Недостоверными являются зависимости ударной вязкости от серы, никеля и меди.

Для ударной вязкости на поперечных образцах обнаружена лишь одна существенная отрицательная связь с содержанием фосфора. Зависимости между ударной вязкостью и хромом, ударной вязкостью и углеродом не вполне надежны.

Взаимосвязи между содержанием различных химических элементов в стали 30ХГСА исследовались аналогичным образом. В табл. 67 сведены коэффициенты корреляции между содержанием химических элементов.

Таблица 66
Коэффициенты корреляции между содержанием химических элементов и величиной ударной вязкости

Химические элементы	Ударная вязкость	
	в продольных образцах	в поперечных образцах
C	-0,448	-0,209
Cr	-0,416	-0,220
Mn	-0,394	-0,137
P	-0,245	-0,330
Si	-0,122	-0,016
Cu	-0,044	-0,159
S	-0,037	-0,134
Ni	-0,015	-0,012

Примечание. Достоверные зависимости подчеркнуты.

Таблица 67
Коэффициенты корреляции между содержанием химических элементов

	C	Cr	Mn	P	Si	Cu	S	Ni
C	*	<u>0,293</u>	<u>0,430</u>	-0,051	<u>0,182</u>	-0,014	<u>0,024</u>	-0,03
Cr	<u>0,293</u>	*	<u>0,357</u>	<u>0,228</u>	<u>0,038</u>	0,03	<u>0,040</u>	<u>0,06</u>
Mn	<u>0,430</u>	<u>0,357</u>	*	<u>0,109</u>	<u>0,187</u>	-0,084	<u>0,117</u>	-0,02
P	-0,051	<u>0,228</u>	<u>0,109</u>	*	<u>-0,241</u>	0,043	<u>0,047</u>	<u>0,06</u>
Si	<u>0,182</u>	<u>0,038</u>	<u>0,187</u>	<u>-0,241</u>	*	-0,064	<u>0,130</u>	-0,07
Cu	-0,014	0,03	<u>0,084</u>	<u>0,043</u>	<u>0,064</u>	*	<u>0,104</u>	-0,11
S	<u>0,024</u>	<u>0,04</u>	<u>0,117</u>	<u>0,047</u>	<u>0,130</u>	<u>0,104</u>	*	<u>0,07</u>
Ni	-0,037	<u>0,06</u>	-0,029	<u>0,067</u>	-0,07	-0,112	<u>0,07</u>	*

Исследование показало, что имеют место существенные положительные зависимости между содержанием углерода и хрома ($r=0,293$), углерода и марганца ($r=0,430$), хрома и марганца ($r=0,357$), хрома и фосфора ($r=0,228$), марганца и кремния ($r=0,187$), углерода и кремния ($r=0,182$) и отрицательная зависимость между фосфором и кремнием ($r=-0,241$). Остальные коэффициенты корреляции статистически не достоверны

$$\left(\frac{|r|}{\sigma_r} < 2,6\right).$$

Линии регрессии для всех этих зависимостей оказались приближенно прямолинейными.

В приведенных ранее парных зависимостях между ударной вязкостью и содержанием химических элементов влияние содержания каждого элемента на ударную вязкость искажалось неявным влиянием сопутствующих элементов. О характере этих сопутствующих влияний говорят таблицы коэффициентов корреляции (см. табл. 67 и 68).

Например, каждый из двух химических элементов — углерод и хром — отрицательно влияет на величину ударной вязкости. Между названными химическими элементами обнаружена существенная положительная зависимость ($r=0,293$). При установлении парной зависимости между содержанием углерода (С) и ударной вязкостью (a_k) корреляционно связанный с углеродом хром и другие химические элементы оказывают неявное влияние на силу связи между a_k и С, и, наоборот, зависимость между содержанием хрома и величиной ударной вязкости искажается влиянием углерода и других корреляционно связанных с хромом химических элементов.

Задача, следовательно, состояла в том, чтобы освободиться от искажающего влияния сопутствующих факторов и получить искомые зависимости в «чистом» виде, т. е. при закреплении содержания прочих элементов на постоянном уровне. Вторая задача состоит в определении зависимости ударной вязкости от комбинации химических элементов.

Обе эти задачи решаются нахождением уравнения множественной регрессии.

Наиболее существенно на ударную вязкость влияет содержание С и Сг. Поэтому в первую очередь была изучена зависимость между ударной вязкостью и комбинацией этих элементов, имеющая вид

$$\bar{x}_{1,23} = b_1 + b_2x_2 + b_3x_3,$$

где x_2, x_3 — содержание С и Сг;

$\bar{x}_{1,23}$ — соответствующее им среднее значение a_k ;
 b_1, b_2, b_3 — коэффициенты уравнения регрессии, подлежащие определению.

Выбор линейной формы уравнения множественной регрессии обусловлен тем, что парные зависимости между a_k и С, a_k и Сг, а также между С и Сг имеют приближенно прямолинейный характер.

Для установления степени влияния каждого из двух химических элементов (С и Сг) в отдельности на величину ударной вязкости применялось приведение переменных к стандартизованному масштабу по формуле

$$t = \frac{x - \bar{x}}{\sigma}.$$

Уравнение множественной регрессии в сопоставимых единицах имеет вид:

$\bar{t}_{1.23} = \beta_2 t_2 + \beta_3 t_3$, где искомые параметры β_2 и β_3 являются корнями системы уравнений:

$$r_{21} = \beta_2 + \beta_3 r_{23},$$

$$r_{31} = \beta_2 r_{23} + \beta_3,$$

где r_{21} , r_{31} , r_{23} — коэффициенты корреляции между переменными x_1 , x_2 , x_3 .

Подставляя значения коэффициентов корреляции для С и Сг из табл. 66 и 67 и решая систему уравнений

$$\left. \begin{aligned} -0,448 &= \beta_2 + \beta_3 \cdot 0,293 \\ -0,416 &= \beta_2 \cdot 0,293 + \beta_3 \end{aligned} \right\}$$

получим $\beta_2 = -0,357$ и $\beta_3 = -0,312$.

Таким образом, было установлено, что влияние комбинации содержания углерода и хрома на величину ударной вязкости выражается уравнением

$$\bar{t}_{1.23} = -0,357 t_2 - 0,312 t_3.$$

Коэффициенты при t_2 и t_3 показывают, как менялась бы ударная вязкость в зависимости от содержания одного из рассматриваемых элементов, если бы содержание второго элемента было закреплено на постоянном уровне. Мы видим, что с увеличением содержания углерода на его среднее квадратическое отклонение (σ) ударная вязкость уменьшается в среднем на 0,357 своего среднего квадратического отклонения, а с увеличением содержания хрома на σ ударная вязкость уменьшается соответственно на 0,312 σ . Следовательно, степень влияния углерода при постоянном хроме несколько больше, чем хрома при постоянном углероде.

Методика расчета уравнения множественной регрессии с большим количеством переменных остается той же. Так, зависимость величины ударной вязкости от пяти химических элементов (С, Сг, Мп, Р и Si) выражается в натуральном масштабе следующим уравнением:

$$\bar{x}_{1.234} = b_1 + b_2 x_2 + b_3 x_3 + b_4 x_4 + b_5 x_5 + b_6 x_6$$

или в стандартизованном масштабе

$$\bar{t}_{1.234} = \beta_2 t_2 + \beta_3 t_3 + \beta_4 t_4 + \beta_5 t_5 + \beta_6 t_6.$$

Параметры $\beta_2, \beta_3, \beta_4, \beta_5, \beta_6$ находятся из системы линейных уравнений

$$\left. \begin{aligned} r_{21} &= \beta_2 + \beta_3 r_{23} + \beta_4 r_{24} + \beta_5 r_{25} + \beta_6 r_{26} \\ r_{31} &= \beta_2 r_{32} + \beta_3 + \beta_4 r_{34} + \beta_5 r_{35} + \beta_6 r_{36} \\ r_{41} &= \beta_2 r_{42} + \beta_3 r_{43} + \beta_4 + \beta_5 r_{45} + \beta_6 r_{46} \\ r_{51} &= \beta_2 r_{52} + \beta_3 r_{53} + \beta_4 r_{54} + \beta_5 + \beta_6 r_{56} \\ r_{61} &= \beta_2 r_{62} + \beta_3 r_{63} + \beta_4 r_{64} + \beta_5 r_{65} + \beta_6 \end{aligned} \right\}.$$

На величину коэффициента корреляции влияет точность измерения сопоставляемых признаков, т. е. их воспроизводимость при повторных испытаниях. Степень воспроизводимости или надежность показателя ударной вязкости может быть охарактеризована коэффициентом корреляции между показателями ударной вязкости двух одновременно испытываемых образцов, подвергавшихся одной и той же термической обработке.

При помощи этого коэффициента корреляции r рассчитывается показатель надежности ρ по формуле

$$\rho = \sqrt{\frac{2r}{1+r}}.$$

Приводим вывод формулы для показателя ρ . Пусть ударная вязкость испытывается на l образцах (в нашем случае $l=2$). Обозначим через r теоретический (предельный) коэффициент корреляции между ударной вязкостью каждой пары образцов, через r_{xy} — теоретический коэффициент корреляции между ударной вязкостью какого-либо образца x и содержанием химического элемента y , и через $r_{\bar{xy}}$ — теоретический коэффициент корреляции между средней ударной вязкостью x на l образцах и содержанием химического элемента y .

Можно доказать, что

$$r_{\bar{xy}} = r_{xy} \sqrt{\frac{l}{1+(l-1)r}}.$$

Отсюда

$$\lim_{l \rightarrow \infty} r_{\bar{xy}} = \frac{r_{xy}}{\sqrt{r}}.$$

Следовательно, корректирующий коэффициент

$$\rho = \frac{r_{\bar{xy}}}{\lim_{l \rightarrow \infty} r_{xy}} = \sqrt{\frac{lr}{1+(l-1)r}}.$$

При $l=2$ $\rho = \sqrt{\frac{2r}{1+r}}.$

Корректируя путем деления на ρ коэффициенты корреляции ударной вязкости с содержанием химических элементов, получаем представление о действительных коэффициентах корреляции, которые имели бы место, если бы ударная вязкость определялась как средняя не из двух, а из бесконечного множества одновременно взятых образцов.

В нашем случае коэффициенты r и ρ оказались: для ударной вязкости в продольных образцах $r=0,853$, $\rho=0,960$; для поперечных образцов $r=0,736$, $\rho=0,921$.

Аналогичным образом корректируются параметры β_i в уравнениях множественной регрессии по формуле

$$\beta'_i = \frac{\beta_i}{\rho}$$

В табл. 68 приведены скорректированные коэффициенты корреляции между ударной вязкостью и содержанием химических элементов.

Таблица 68

Химические элементы	Ударная вязкость	
	в продольных образцах	в поперечных образцах
C	-0,467	-0,227
Cr	-0,433	-0,239
Mn	-0,410	-0,149
P	-0,255	-0,358
Si	-0,127	-0,017
Cu	-0,046	-0,173
S	-0,039	-0,134
Ni	-0,016	-0,013

В табл. 69 приводится сводка полученных уравнений множественной регрессии с поправкой и без поправки на надежность определения ударной вязкости для комбинации двух-пяти элементов (C, Cr, Mn, P, Si). По этой таблице можно проследить, как уточняются коэффициенты регрессии при элементах по мере того, как в уравнение множественной регрессии включаются новые элементы, и, следовательно, влияние каждого из них на ударную вязкость «очищается» от искажающего влияния сопутствующих элементов.

Теснота связи ударной вязкости с содержанием всех элементов, участвующих в уравнении множественной регрессии, измеряется совокупным коэффициентом корреляции $R_{1.23...k}$.

Совокупные коэффициенты корреляции (некорректированные), данные в табл. 69, определялись по формуле

$$R_{1.23...k} = \sqrt{\beta_2 r_{12} + \beta_3 r_{13} + \dots + \beta_k r_{1k}},$$

где k — количество изучаемых химических элементов.

Таблица 69

Количество элементов	Уравнения множественной регрессии a_k прод по химическим элементам	Совокупный коэффициент корреляции, исправленный на число параметров R'	То же, с поправкой на надежность R''
Один	$\bar{t}_1 = -0,448t_2$ $\bar{t}_1 = -0,467t_2$	0,448	0,467
Два	$\bar{t}_1 = -0,357t_2 - 0,312t_3$ $\bar{t}_1 = -0,372t_2 - 0,325t_3$	0,532	0,554
Три	$\bar{t}_1 = -0,245t_2 - 0,269t_3 - 0,171t_4$ $\bar{t}_1 = -0,255t_2 - 0,280t_3 - 0,178t_4$	0,549	0,572
Четыре	$\bar{t}_1 = -0,327t_2 - 0,222t_3 - 0,154t_4 - 0,189t_5$ $\bar{t}_1 = -0,341t_2 - 0,231t_3 - 0,160t_4 - 0,197t_5$	0,576	0,600
Пять	$\bar{t}_1 = -0,320t_2 - 0,220t_3 - 0,139t_4 - 0,216t_5 - 0,081t_6$ $\bar{t}_1 = -0,333t_2 - 0,229t_3 - 0,145t_4 - 0,225t_5 - 0,084t_6$	0,581	0,605
Пять (в натуральном масштабе)	$\bar{x}_1 = 25,5 - 25,0C - 4,97Cr - 3,04Mn - 79,7P - 1,74Si$		
	Уравнение множественной регрессии a_k попереч по химическим элементам		
Один	$\bar{t}_1 = -0,209t_2$ $\bar{t}_1 = -0,227t_2$	0,209	0,227
Два	$\bar{t}_1 = -0,158t_2 - 0,174t_3$ $\bar{t}_1 = -0,172t_2 - 0,189t_3$	0,235	0,255
Три	$\bar{t}_1 = -0,047t_2 - 0,095t_3 - 0,311t_4$ $\bar{t}_1 = -0,051t_2 - 0,103t_3 - 0,326t_4$	0,333	0,362

Примечания. 1. В уравнениях условно даны:

t_1 вместо a_k , t_4 вместо Mn,
 t_2 „ C, t_5 „ P,
 t_3 „ Cr, t_6 „ Si.

2. Вторая строка каждого из уравнений дается с поправкой на надежность.

Коэффициент корреляции между a_k и комбинацией С и Сг получается

$$R_{1.23} = \sqrt{(-0,357) \cdot (-0,448) + (-0,312) \cdot (-0,416)} = 0,538.$$

Коэффициент корреляции корректировался на число параметров уравнения регрессии (в натуральном масштабе) по формуле

$$\hat{R}^2 = 1 - (1 - R^2) \left(\frac{n-1}{n-m} \right),$$

где n — число наблюдений;
 m — число параметров.

Уравнение регрессии a_k по содержанию С и Сг включает три параметра: b_1, b_2, b_3 , следовательно,

$$\hat{R}^2 = 1 - (1 - 0,538^2) \left(\frac{220-2}{220-3} \right) = 0,283;$$

$$\hat{R} = \sqrt{0,283} = 0,532.$$

Далее совокупный коэффициент корреляции корректировался на надежность определения ударной вязкости по формуле

$$\hat{R} = \frac{\hat{R}}{\rho} = \frac{0,532}{0,960} = 0,554.$$

Исследование уравнения множественной регрессии ударной вязкости по содержанию элементов С, Сг, Мп, Р, Si показывает, что наибольшее отрицательное влияние на ударную вязкость в продольных образцах оказывает содержание углерода, затем содержание хрома и фосфора, потом марганца и незначительное влияние оказывает кремний. Эти выводы справедливы в том диапазоне изменения содержания каждого из элементов, который выявлен на кривых распределения.

Совокупный коэффициент множественной корреляции между ударной вязкостью на продольных образцах и комбинацией углерода, хрома, марганца, фосфора и кремния равен 0,605. Сера, никель и медь не вводились в уравнение множественной регрессии ввиду их незначительных связей с ударной вязкостью и другими элементами.

На ударную вязкость в поперечных образцах значительное отрицательное влияние оказывает содержание фосфора. Другие химические элементы также оказывают отрицательное влияние на ударную вязкость, однако степень их влияния значительно меньше. Это может быть объяснено тем, что ударная вязкость в об-

разцах с поперечным направлением волокна в большей степени определяется факторами технологического процесса, в частности дендритной ликвацией, раскислением стали и др.

Совокупный коэффициент корреляции между величиной ударной вязкости в поперечных образцах и комбинацией химических элементов (С, Сг и Р) невелик, $R=0,362$.

Уравнение множественной регрессии, связывающее ударную вязкость в продольных образцах с комбинацией элементов С, Сг, Мп, Р, Si может быть использовано для построения контрольного показателя ударной вязкости по химическому составу. Для этой цели переменные последнего уравнения выражаются в натуральном масштабе (содержание химических элементов в процентах и величина ударной вязкости в $кгм/см^2$) и соответственно пересчитываются коэффициенты регрессии при переменных по формулам:

$$b_i = \frac{\sigma_1}{\sigma_i} \beta_i, \quad b_1 = \bar{x}_{1.23\dots k} - \Sigma b_i \bar{x}_i.$$

Таким образом, получаем уравнение множественной регрессии в виде

$$\bar{x}_{1.23456} = 25,5 - 25,0 С - 4,97 Сг - 3,04 Мп - 79,7 Р - 1,74 Si.$$

Подставив в эту формулу данные химического анализа конкретной плавки, можно рассчитать ожидаемое значение ударной вязкости в продольных образцах. Так, например, для плавки № 1 расчет по вышеприведенной формуле дает значение ударной вязкости, равное $a'_k = 6,1$, фактическое же значение — 6,6. Для плавки № 3 (табл. 1, прилож. 1) расчет дает $a'_k = 5,5$, фактическая $a_k = 4,2$. Для плавки № 10 по расчету $a'_k = 6,7$, фактически $a_k = 6,2$ и т. д.

Показатель a'_k дополнительно контролирует правильность определения величины ударной вязкости данной плавки. При наличии резких расхождений между a_k и a'_k следует величину ударной вязкости проверять на большем количестве образцов.

Таким образом, корреляционный анализ материалов текущего контроля плавки стали 30ХГСА позволяет сделать следующие выводы:

1. Имеет место существенное влияние химического состава стали 30ХГСА на величину ударной вязкости продольных образцов. Совокупный коэффициент корреляции содержания химических элементов с величиной ударной вязкости оказался равным $R=0,605$.

2. Установлено существенное и отрицательное влияние на ударную вязкость продольных образцов содержания следующих химических элементов: углерода, хрома, фосфора, марганца и кремния. По степени влияния на ударную вязкость элементы химического состава (в пределах фактической изменчивости их

содержания) располагаются следующим образом: наибольшее отрицательное влияние имеет углерод, затем хром, фосфор, менее существенно влияние марганца и кремния и практически не оказывают влияния медь, сера и никель.

3. Наблюдается существенное и отрицательное влияние содержания фосфора на ударную вязкость стали при испытании на поперечных образцах.

4. Результаты исследования показывают, что для получения высоких показателей ударной вязкости стали 30ХГСА необходимо держать углерод, фосфор, марганец, хром, кремний по возможности на нижних пределах поля допуска.

5. В результате исследования дана методика построения контрольного показателя ударной вязкости в продольных образцах, позволяющая для каждой плавки по содержанию в ней углерода, хрома, марганца, фосфора и кремния рассчитывать вероятное значение ударной вязкости. Этот показатель может быть использован для назначения плавки по заказам, а также для корректирования лабораторной методики определения ударной вязкости плавки стали 30ХГСА.

ГЛАВА XV
КОРРЕЛЯЦИОННЫЙ АНАЛИЗ ТЕХНОЛОГИЧЕСКИХ
ЦЕПЕЙ

§ 1. ПРЕДВАРИТЕЛЬНЫЕ ЗАМЕЧАНИЯ

Важное применение теория корреляции получает при исследовании точности технологических процессов в вопросах анализа технологических цепей.

Технологической цепью мы называем ряд последовательных операций, которые совершаются над изделием для получения определенных размеров, формы, чистоты поверхности и других параметров изделия с необходимой точностью.

В процессе обработки изделия его параметры постепенно уточняются, приближаясь к требуемым значениям. Об изменении точности на каждой операции судят по кривым распределения погрешностей, составленным после ряда операций как это показано в разделе I (см. гл. II, § 10, пример 3).

Количественная оценка изменения точности дается сопоставлением средних квадратических отклонений или средних величин погрешностей (для существенно положительных величин) после каждой операции. Этой цели служат коэффициенты $K_{\sigma} = \frac{\sigma_2}{\sigma_1}$

и $K_{\bar{x}} = \frac{\bar{x}_2}{\bar{x}_1}$, представляющие собой отношения соответствующих

показателей после и до проведения операции. Однако эти критерии являются лишь констатационными. Они не раскрывают механизма сокращения погрешностей и не пригодны для оценки точности, получаемой на данной операции в условиях изменяющейся точности исходного материала. Было бы неправильным рассматривать K_{σ} как коэффициент пропорциональности в зависимости, связывающей рассеяние размеров после операции с рассеянием до операции, и использовать его в этой роли при оценке предпологаемой точности после операции, если бы исходная точность изменилась — стала большей или меньшей. Так, в табл. 25 при-

мера \bar{z} среднее квадратическое отклонение биений после токарной обточки составляло 6,92, после следующей операции — предварительной шлифовки — 1,87. Следовательно, коэффициент сокращения $K_{\sigma} = \frac{\sigma_2}{\sigma_1} = \frac{1,87}{6,92} = 0,270$. Но если бы на операции токарной обточки удалось сократить σ_1 до 5 единиц, то было бы неверным определять возможное σ_2 по формуле $\sigma_2 = 0,270 \cdot 5 = 1,350$, так как с увеличением или с уменьшением рассеяния на предыдущей операции сам коэффициент K_{σ} , вообще говоря, меняется и уравнение $K_{\sigma} = \frac{\sigma_2}{\sigma_1}$ оказывается непригодным для подобного расчета.

Аналогично для показателя $K_{\bar{x}}$. Между тем при проектировании и отладке точности технологических процессов для обоснования пооперационных допусков, а также при разработке методик статистического контроля качества продукции важно знать, какое влияние на последующую операцию имеет увеличение или уменьшение точности обработки на предыдущих операциях. Для решения подобных вопросов необходимо иметь показатели изменения точности не констатационного, а аналитического характера, отражающие механизм сокращения погрешностей. Такие показатели (способы построения которых рассматриваются здесь) могут быть установлены путем исследования корреляционных связей между погрешностями, возникающими в процессе обработки в технологической цепи следующих друг за другом операций.

§ 2. ОБЩАЯ ПОСТАНОВКА И ОБЩИЙ МЕТОД РЕШЕНИЯ ЗАДАЧИ

Рассмотрим отдельную операцию в составе технологического процесса, которую будем называть звеном технологической цепи.

Обозначим через $\varphi(x)$ закон распределения интересующего нас параметра x до операции и через $\psi(y)$ — закон распределения соответствующего параметра y после операции¹. Первый закон характеризует *входное качество* продукции, второй — *выходное качество*. Пусть далее мы изменяем входное качество продукции, следовательно, подвергаем изменению закон распределения $\varphi(x)$. Так, например, мы можем сократить дисперсию закона распределения $\varphi(x)$ за счет улучшения работы на предшествующей операции или отсева на контроле полученного брака. Спрашивается, каким образом в силу этого обстоятельства изменится закон распределения $\psi(y)$?

Для того чтобы задача приобрела должную определенность и могла получить решение, необходимо охарактеризовать процесс изменения входного качества некоторыми дополнительными условиями. Математически эти условия будут выражаться в том, чтобы трансформация закона распределения $\varphi(x)$ не затрагивала

¹ $\varphi(x)$ и $\psi(y)$ означают вероятности или плотности вероятностей соответствующих значений x и y в зависимости от того, в дискретной или непрерывной форме заданы законы распределения.

условных законов распределения y , соответствующих заданным значениям x , т. е. чтобы $\psi(y|x)$ оставались при этом неизменными. Подобное требование является вполне естественным.

В самом деле, естественно допустить, что трансформация закона распределения x не затрагивает качества исходного материала *при заданных значениях x* и не затрагивает технологического процесса на данной операции. Она лишь означает замену материала с одним *набором значений параметра x* материалом с другим, *несколько отличным от первого набором значений x* . А это предположение и приводит к сохранению условных законов распределения $\psi(y|x)$, хотя безусловный закон распределения $\psi(y)$ претерпевает при этом изменения. Разумеется, принятые допущения не всегда могут реализоваться на практике. Так, например, в случае больших изменений закона распределения $\varphi(x)$ может оказаться необходимым изменить технологический процесс на данной операции, чтобы приспособить его к новому, изменившемуся качеству исходного материала. В этом случае условные законы распределения $\psi(y|x)$ могут быть нарушены. Разумеется также, что если при одних и тех же значениях параметра x изменяется качество исходного материала по другим параметрам, влияющим на ход производственного процесса, то условные законы распределения $\psi(y|x)$ также могут измениться. На подобные случаи приводимая ниже теория не распространяется¹. В принятых предположениях мы можем выразить искомый закон распределения $\psi(y)$ в зависимости от исходного закона распределения $\varphi(x)$ и других элементов, независимых от возможной трансформации закона $\varphi(x)$, а именно от условных вероятностей $\psi(y|x)$.

Рассмотрим корреляционную таблицу между параметром входного качества x и параметром выходного качества y , где для каждой комбинации значений x , y указана ее вероятность $\chi(x, y)$. Эти вероятности можно получить приближенно, как относительные частоты повторений каждой комбинации x и y , если провести при стабильных условиях через операцию достаточно большое число деталей. Детали для этой цели подвергаются клеймению и для каждой детали отмечается показатель x до операции и показатель y — после операции; затем подсчитывается число повторений каждой комбинации значений (x, y) и полученные числа (абсолютные частоты комбинаций (x, y)) делятся на общее число деталей. Эти числа непосредственно или после некоторой матема-

¹ Заметим также, что для некоторых процессов и параметров условные законы распределения $\psi(y|x)$ и безусловные $\psi(y)$ могут быть легко смещены на одну и ту же величину c соответствующей настройкой процесса. Для этих случаев вероятности ψ правильнее писать как $\psi(y+c|x)$ и $\psi(y+c)$, где c — компонент, определяемый настройкой процесса. Мы не будем пользоваться этими обозначениями, но будем подразумевать, что для таких случаев мы вправе смещать условные законы распределения на одну и ту же величину c , задаваясь определенным уровнем настройки процесса.

тической обработки (интерполяции, выравнивания) заменяют в порядке приближения значения вероятностей.

Таблица 70

y_j	x_i	1	2	3	4	5	6	7	Итого $\psi(y_j)$
4,0								0,010	0,010
3,5						0,010	0,012	0,010	0,032
3,0			0,012	0,011	0,024	0,040	0,036	0,020	0,131
2,5			0,033	0,033	0,060	0,060	0,036	0,010	0,211
2,0	0,010	0,030	0,055	0,072	0,050	0,024			0,241
1,5	0,010	0,036	0,066	0,060	0,030	0,012			0,214
1,0	0,020	0,030	0,044	0,021	0,010				0,128
0,5	0,010	0,012	0,011						0,033
Итого $\varphi(x_i)$		0,050	0,120	0,220	0,240	0,200	0,120	0,050	1,000

Табл. 70 имеет чисто методическое назначение, в ней представлены округленные и условные данные. Заглавная и итоговая строки выражают безусловный закон распределения параметра входного качества x . В заглавной строке представлены возможные варианты x , в итоговой — вероятности $\varphi(x)$ этих вариантов. Заглавный и итоговый столбцы таблицы выражают безусловный закон распределения параметра выходного качества y . В заглавном столбце приведены варианты y , в итоговом — вероятности $\psi(y)$ этих значений. Числа внутри таблицы представляют собой вероятности $\chi(x, y)$ соответствующих комбинаций x и y . Если разделить каждую такую вероятность на вероятность $\varphi(x)$ значения x , помещенную в итоге соответствующего столбца таблицы, то получим $\psi(y|x)$ — условные вероятности y в предположении, что x осуществились.

$$\psi(y|x) = \frac{\chi(x,y)}{\varphi(x)}. \quad (15.2.1)$$

Продлав соответствующую операцию для каждой $\chi(x, y)$ табл. 70, приходим к таблице 71, где во внутренних столбцах указаны $\psi(y|x)$ — условные законы распределения y , соответствующие заданным значениям x .

Из рассмотрения таблицы видно, что с увеличением x условные законы распределения y перемещаются в сторону больших значений y . Следовательно, имеет место корреляционная зависимость между y и x — погрешности входного качества в некоторой мере влияют на погрешности выходного качества. Посмотрим, как эту зависимость можно использовать для решения поставленной задачи: оценки возможного изменения выходного качества в связи с заданным изменением входного качества продукции.

Таблица 71

$y_j \backslash x_i$	1	2	3	4	5	6	7	Итого $\psi(y_j)$
4,0					0,05	0,10	0,20	0,010
3,5					0,20	0,30	0,20	0,032
3,0			0,05	0,10	0,20	0,30	0,40	0,131
2,5		0,10	0,15	0,25	0,30	0,30	0,20	0,211
2,0	0,20	0,25	0,25	0,30	0,25	0,20		0,241
1,5	0,20	0,30	0,30	0,25	0,15	0,10		0,214
1,0	0,40	0,25	0,20	0,10	0,05			0,128
0,5	0,20	0,10	0,05					0,033
Итого	1,00	1,00	1,00	1,00	1,00	1,00	1,00	—
Итого $\varphi(x_i)$	0,050	0,120	0,220	0,240	0,200	0,120	0,050	1,000

Заметим прежде всего, что между вероятностями $\varphi(x)$, $\psi(y|x)$ и $\psi(y)$ существуют следующие соотношения, которые вытекают из способа образования $\psi(y|x)$:

$$\sum_{i=1}^s \varphi(x_i) \psi(y_j|x_i) = \psi(y_j); \quad j=1, 2, \dots, t, \quad (15.2.2)$$

где s — число различных значений x ,
 t — число различных значений y .

Проверим эти соотношения на числах табл. 71:

$$\begin{aligned} 0,050 \cdot 0,20 &= 0,010 \\ 0,200 \cdot 0,05 + 0,120 \cdot 0,10 + 0,050 \cdot 0,20 &= 0,032 \\ 0,220 \cdot 0,50 + 0,240 \cdot 0,10 + 0,200 \cdot 0,20 + 0,120 \cdot 0,30 + 0,050 \cdot 0,40 &= \\ = 0,131 \\ 0,120 \cdot 0,10 + 0,220 \cdot 0,15 + 0,240 \cdot 0,25 + 0,200 \cdot 0,30 + 0,120 \cdot 0,30 + \\ + 0,050 \cdot 0,20 &= 0,211 \\ 0,050 \cdot 0,20 + 0,120 \cdot 0,25 + 0,220 \cdot 0,25 + 0,240 \cdot 0,30 + 0,200 \cdot 0,25 + \\ + 0,120 \cdot 0,20 &= 0,241 \\ 0,050 \cdot 0,20 + 0,120 \cdot 0,30 + 0,220 \cdot 0,30 + 0,240 \cdot 0,25 + 0,200 \cdot 0,15 + \\ + 0,120 \cdot 0,10 &= 0,214 \\ 0,050 \cdot 0,40 + 0,120 \cdot 0,25 + 0,220 \cdot 0,20 + 0,240 \cdot 0,10 + 0,200 \cdot 0,05 &= \\ = 0,128 \\ 0,050 \cdot 0,20 + 0,120 \cdot 0,10 + 0,220 \cdot 0,05 &= 0,033 \end{aligned}$$

Соотношения (15.2.2) выполняются.

Соотношения (15.2.2) можно записать в матричной форме:

$$\|\varphi(x)\| \cdot \|\psi(y|x)\| = \|\psi(y)\|, \quad (15.2.3)$$

которая означает, что таблица чисел $\psi(y)$ получается из таблицы чисел $\psi(y|x)$ и таблицы чисел $\varphi(x)$ по правилу произведения матриц.

$\varphi(x)$ будем называть вероятностями входного качества, $\psi(y)$ — вероятностями выходного качества, $\psi(y|x)$ — вероятностями перехода.

Возвратимся теперь к исходной задаче. Пусть требуется изменить качество входного материала, трансформируя закон распределения $\varphi(x)$, и оценить протекающую отсюда трансформацию $\psi(y)$.

Если процесс трансформации $\varphi(x)$ протекает в условиях, оговоренных в § 2, т. е. он не затрагивает технологического процесса на данной операции и входного качества продукции при заданных значениях параметра x , то сохраняются неизменными условные законы распределения $\psi(y|x)$.

Таблица 72

$y_j \backslash x_i$	1	2	3	4	Итого $\psi(y_j)$
3,0		0,10	0,05	0,10	0,035
2,5		0,25	0,15	0,25	0,125
2,0	0,20	0,25	0,25	0,30	0,250
1,5	0,20	0,30	0,30	0,25	0,270
1,0	0,40	0,25	0,20	0,10	0,235
0,5	0,20	0,10	0,05		0,085
Итого $\varphi(x_i)$	0,200	0,300	0,300	0,200	1,000

Следовательно, соотношения (15.2.2) можно рассматривать как уравнения, из которых определяется закон распределения выходного качества $\psi(y)$, когда закон распределения входного качества $\varphi(x)$ подвергается трансформации в оговоренных выше условиях.

Пример. Улучшаем входное качество материала, заменив закон распределения $\varphi(x)$ в табл. 71 другим законом $\varphi(x)$, приведенным в табл. 72.

Предполагая условные законы $\psi(y|x)$ неизменившимися, рассчитаем по формулам (15.2.2) новые значения $\psi(y)$. Закон распределения выходного качества несколько сдвинулся в сторону меньших значений y , и рассеяние y сократилось.

Зная матрицу начальных вероятностей $\|\varphi(x)\|$ и матрицы вероятностей перехода для ряда последовательных операций, можно получить матрицу вероятностей выходного качества для конечного звена по формуле

$$\|\varphi(x)\| \cdot \|\psi_1(y|x)\| \cdot \|\psi_2(z|y)\| \dots \|\psi_n(v|u)\| = \|\psi_n(v)\|. \quad (15.2.4)$$

которая получается последовательным применением формулы (15.2.2). Формула (15.2.4) справедлива при условии, что погрешности, возникающие на предшествующих операциях, передаются на последующие операции *лишь через все промежуточные операции*. Такую цепь будем называть *простой* технологической цепью. Ее математической моделью служит простая цепь Маркова.

Изложенный в настоящем параграфе метод решения задачи является хотя и общим, но практически мало удобным. В самом деле, для решения задачи этим методом требуется иметь матрицу достаточно надежно определенных вероятностей $\psi(y|x)$ для всех рассматриваемых комбинаций значений x_i, y_j .

А для этого требуется выполнить большое количество наблюдений, заполняющих все интересующие нас клетки корреляционной таблицы, что на практике представляет затруднения. Замена корреляционной таблицы после надлежащей математической обработки (интерполяции, выравнивания) аналитической формой задания законов распределения значительно сокращает потребное количество наблюдений, приводя его к практически вполне приемлемым числам. В этих случаях законы распределения $\varphi(x), \psi(y), \psi(y|x)$ могут выражаться в виде непрерывных функций от аргументов, причем величины $\varphi(x), \psi(y), \psi(y|x)$ обозначают плотности вероятностей для соответствующих значений аргументов.

Система (15.2.2) заменяется в этих случаях интегральным соотношением

$$\psi(y) = \int_{-\infty}^{+\infty} \psi(y|x) \varphi(x) dx, \quad (15.2.5)$$

из которого определяется функция $\psi(y)$, соответствующая измененной функции $\varphi(x)$.

Замена корреляционных таблиц законами распределения, заданными в аналитическом виде (т. е. в виде формул), является весьма перспективным методом. Однако в настоящее время мы еще не располагаем достаточно удобными для практических целей способами такого выражения, так как многие распределения, встречающиеся на практике, в особенности двумерные, еще недостаточно исследованы. Поэтому ниже предлагаем другие методы, которые решают поставленную задачу с меньшей точностью, но практически достаточно удобны.

§ 3. СЛУЧАЙ ПРЯМОЛИНЕЙНОЙ РЕГРЕССИИ И ПОСТОЯННОГО РАССЕЯНИЯ

Практические результаты могут быть достигнуты более быстрым и в ряде отношений поучительным способом, если рассматривать не законы распределения параметров, а их важнейшие ха-

рактеристики, выражаемые моментами первых двух порядков, а именно — среднюю величину и дисперсию (или среднее квадратическое отклонение). Задача, следовательно, состоит в том, чтобы выяснить, как меняется средняя величина \bar{y} и дисперсия σ_y^2 по параметру y выходного качества продукции с изменением средней величины \bar{x} и дисперсии σ_x^2 по параметру x входного качества. Разумеется, поставленная задача, так же как и более общая задача, рассмотренная в предыдущем параграфе, будет иметь определенный смысл и решение, если процесс изменения x и σ_x^2 охарактеризовать некоторыми дополнительными условиями. Как и в предыдущем параграфе, будем предполагать, что изменение x и σ_x^2 не затрагивает технологического процесса на данной операции и не изменяет входного качества продукции по другим параметрам при заданных x , следовательно, не отражается на условных законах распределения y при заданных x , т. е. на $\psi(y|x)$.

Для установления необходимых соотношений воспользуемся корреляционной зависимостью, существующей между показателями входного и выходного качества.

В настоящей главе рассмотрим простейший случай, когда регрессия y по x прямолинейна, а условная дисперсия y относительно линии регрессии постоянна.

Пусть

$$\bar{y}_x = a + bx \quad (15.3.1)$$

— уравнение регрессии y по x .

Из теории корреляции известно, что прямая регрессии, выражаемая уравнением (15.3.1), должна проходить через центр тяжести корреляционного поля (\bar{x}, \bar{y}) , т. е. существует тождество

$$\boxed{\bar{y} = a + b\bar{x}} \quad (15.3.2)$$

Далее известна следующая формула разложения дисперсии:

$$\sigma_y^2 = \sigma_{y,x}^2 + \delta_{y,x}^2 \quad (15.3.3)$$

где $\sigma_{y,x}^2$ — средняя дисперсия y относительно линии регрессии, $\delta_{y,x}^2$ — дисперсия линии регрессии относительно \bar{y} .

$$\sigma_{y,x}^2 = \sum_{i=1}^s \sum_{j=1}^t (y_{ij} - \bar{y}_{x_i})^2 \chi(x_i, y_j), \quad (15.3.4)$$

$$\delta_{y,x}^2 = \sum_{i=1}^s (\bar{y}_{x_i} - \bar{y})^2 \varphi(x_i). \quad (15.3.5)$$

Для непрерывных законов распределения формулы (15.3.4) и (15.3.5) пишутся в интегральной форме:

$$\overline{\sigma}_{y,x}^2 = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (y - \bar{y}_x)^2 \chi(x, y) dx dy, \quad (15.3.4')$$

$$\overline{\delta}_{y,x}^2 = \int_{-\infty}^{+\infty} (\bar{y}_x - \bar{y})^2 \varphi(x) dx, \quad (15.3.5')$$

где χ и φ означают плотности вероятностей в соответствующих точках. В дальнейшем мы будем пользоваться главным образом символикой, относящейся к дискретным или выраженным в дискретной форме законам распределения.

Легко видеть, что в рассматриваемом нами случае (15.3.1)

$$\overline{\delta}_{y,x}^2 = b^2 \sigma_x^2. \quad (15.3.6)$$

Далее, при постоянстве условной дисперсии y относительно линии регрессии y по x

$$\sigma_{y,x_i}^2 = \sum_{j=1}^t (y_{ij} - \bar{y}_{x_i})^2 \psi(y_j | x_i) = \text{const} \quad (15.3.7)$$

она должна совпадать со средней дисперсией y относительно линии регрессии, т. е.

$$\overline{\sigma}_{y,x}^2 = \sigma_{y,x}^2. \quad (15.3.8)$$

Следовательно,

$$\boxed{\sigma_y^2 = \sigma_{y,x}^2 + b^2 \sigma_x^2}. \quad (15.3.9)$$

Из формул (15.3.2) и (15.3.9) вытекает:

1. Средняя величина выходной погрешности y складывается из части $b\bar{x}$, пропорциональной средней величине входной погрешности x , и части a , постоянной относительно x .

2. Дисперсия выходной погрешности σ_y^2 складывается из части $b^2 \sigma_x^2$, пропорциональной дисперсии входной погрешности σ_x^2 , и части $\sigma_{y,x}^2$, постоянной относительно σ_x^2 .

Однако эти выводы нуждаются в обосновании. Дело в том, что формулы (15.3.2) и (15.3.9) представляют собой соотношения, справедливые относительно величин \bar{x} , \bar{y} , σ_x^2 , σ_y^2 , a , b , $\sigma_{y,x}^2$, определенных на данном конкретном материале (процессе). Из-

менить какие-либо из этих величин, например \bar{x} и σ_x^2 , можно лишь изменив тот материал, для которого они определены. Но в таком случае, вообще говоря, меняются и другие, участвующие в этих соотношениях величины и, следовательно, выводы 1 и 2 не будут верными, так как они сформулированы в предположении, что величины a , b , $\sigma_{y,x}^2$ постоянны относительно x и σ_x^2 . Однако если мы примем во внимание, что показатели a , b , $\sigma_{y,x}^2$ являются характеристиками *условных законов распределения y* , соответствующих заданным значениям x , которые в силу соображений, указанных в начале параграфа, не подвергаются изменению с изменением x и σ_x^2 , то убедимся в правильности выводов 1 и 2. Следовательно, соотношения (15.3.2) и (15.3.9) решают поставленную в начале параграфа задачу.

§ 4. О ФИЗИЧЕСКОМ СМЫСЛЕ ФОРМУЛ

Обратимся к выяснению физического смысла величин, входящих в формулы (15.3.2) и (15.3.9). Хотя приводимая ниже интерпретация и не имеет обязательного значения, она может оказаться полезной для лучшего понимания соотношений (15.3.2) и (15.3.9).

Условимся прежде всего о следующих терминах и обозначениях. Полной погрешностью какого-либо параметра x или y будем называть разность между значением этого параметра и требуемым (или номинальным) значением, за которое для параметров первого рода¹ естественно принять середину поля допуска, а для существенно положительных величин — начало отсчета шкалы — значение 0. Для параметров первого рода условимся отклонения $x - C_x$ и $y - C_y$ обозначать просто через x и y , считая, что начало отсчета этих признаков переносится в точку (C_x, C_y) . В таком виде они входят в формулы (15.3.2) и (15.3.9). Эти формулы, следовательно, связывают между собой характеристики *полных погрешностей параметров x , y* .

Заметим далее, что в процессе исправления погрешностей любая операция, помимо исправления, вносит новые погрешности, полностью или частично не зависящие от предшествующих погрешностей. Поэтому естественно рассматривать погрешности выходного качества как сумму погрешностей, в известной мере воспроизводящих погрешности входного качества и не зависящих от них. Посредством уравнения регрессии y по x удается практически произвести подобное разложение погрешностей и использовать его для построения системы аналитических показателей, входящих в уравнения (15.3.2) и (15.3.9).

Рассмотрим поля корреляции (x, y) с прямолинейной регрес-

¹ Параметрами первого рода мы называем параметры, имеющие двусторонние погрешности и ограниченные двусторонними допусками.

сией y по x — одно для параметров первого рода (рис. 77 а), другое — для существенно положительных величин (рис. 77 б).

Очевидно, что полная погрешность выходного параметра какой-либо детали y складывается из трех частей: $y - \bar{y}_x$, $b x$, a , обозначенных соответственно на рис. 77. Эти три компонента полной погрешности получают следующее истолкование.

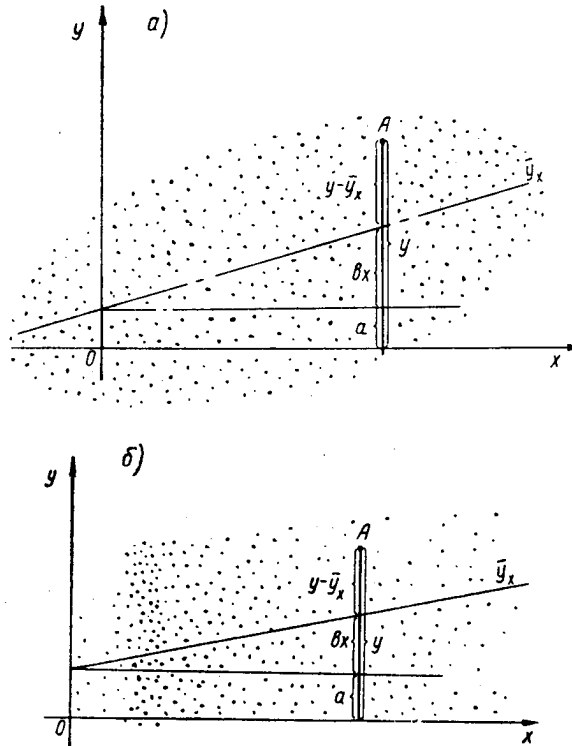


Рис. 77

Компонент $b x$ представляет собой часть выходной погрешности, пропорциональную входной погрешности x . Он называется *перенесенной из предыдущей операции* частью погрешности y , а коэффициент b можно назвать *коэффициентом переноса* (передачи, копирования, наследования) погрешностей.

При $b=0$ линия регрессии параллельна оси x , *переноса не существует*, имевшиеся ранее погрешности *полностью исправляются*. При $b=1$ имеет место полный перенос прежних погрешностей, *исправление отсутствует*. При $0 < b < 1$ имеет место частичный перенос и частичное (неполное) исправление прежних погрешностей. Таким образом, коэффициент переноса b показывает, какая часть прежней погрешности *в среднем* переносится на выходное качество. Величина $1-b$, которую можно назвать *коэффици-*

циентом исправления, показывает, какая часть прежней погрешности *в среднем исправляется* в ходе операции.

Формально возможны и случаи, когда $b > 1$ и когда $b < 0$. Первый можно назвать случаем *переноса с усилением*, второй — случаем *сверхисправления*. Остается пока открытым вопрос, встречаются ли такие случаи на практике и при каких условиях.

Второй компонент $y - \bar{y}_x$ является частью выходной погрешности, которая корреляционно не зависит от величины входной погрешности x . Эта погрешность образуется вновь на данной операции и не имеет систематического характера, так как средняя величина всех таких погрешностей (с учетом знаков) равна нулю. Ее можно назвать *собственной случайной погрешностью операции*. Размеры собственных случайных погрешностей характеризуются дисперсией $\sigma_{y,x}^2$. $\sigma_{y,x}$ — *среднеквадратическая собственная случайная погрешность операции*.

Третий компонент a является *постоянной (положительной или отрицательной) собственной погрешностью операции*. Она увеличивает или уменьшает (смотря по знаку a) все выходные погрешности на одну и ту же величину. Для параметров первого рода постоянная погрешность a часто может быть отрегулирована настройкой процесса на соответствующий размер. В таких случаях соотношение (15.3.2) можно использовать не для определения \bar{y} по данным x , a и b , а для определения неизвестного смещения a , которое необходимо для того, чтобы компенсировать отклонение $b\bar{x}$; последнее возникает на выходе операции благодаря смещению центра группирования x на входе и явлению переноса погрешностей. Из формулы (15.3.2) следует, например, что для совмещения центра группирования \bar{y} с серединой поля допуска (т. е. с нулем) необходимо компонент a рассчитывать по формуле

$$a = -b\bar{x}. \quad (15.4.1)$$

Для существенно положительных величин компонент a не исправляется настройкой процесса на размер. Он является таким же дефектом точности процесса, как и другие компоненты погрешности, и его устранение требует ряда мероприятий, связанных с изменением процесса обработки детали. Здесь выясняется существенное различие между показателями \bar{y} (и соответственно x) для параметров первого рода и для существенно положительных величин.

Для параметров первого рода \bar{y} обычно рассматривается как показатель средней настроенности процесса на правильный размер. Основной характеристикой точности служит дисперсия погрешности — σ_y^2 , и основное значение для оценки выходной точности по входным данным имеет соотношение (15.3.9), оценивающее эту дисперсию.

Для существенно положительных величин \bar{y} нельзя рассматривать как показатель настроенности процесса. Он является, наряду с σ_y^2 показателем точности и в сочетании с σ_y^2 дает более полное представление о точности процесса. Пути такого сочетания указаны в параграфе 6.

§ 5. ПРИМЕР РАСЧЕТА ПОКАЗАТЕЛЕЙ

Показатели, входящие в формулы (15.3.2) и (15.3.9), рассчитываются по данным большого числа наблюдений, проведенных по схеме процесса (α), следующим образом. Параметры a и b рассчитываются по способу наименьших квадратов. Оценку показателя $\sigma_{y,x}^2$ производим по формуле:

$$\sigma_{y,x}^2 = \frac{1}{n} \sum_{i=1}^s \sum_{j=1}^t (\bar{y}_{ij} - \bar{y}_x)^2 m_{ij}, \quad (15.5.1)$$

где m_{ij} — частота комбинации (x_i, y_j) ,
 n — общее число наблюдений,
 $\bar{y}_{x_i} = a + bx_i$ — ординаты линии регрессии.

Пример. Для выяснения вопроса о том, в какой мере операция хонингования исправляет погрешности геометрической формы детали, в лаборатории резания ЗИЛ Н. И. Гороецкий исследовал нецилиндричность деталей до и после этой операции. Показатель нецилиндричности определяется как разность между наибольшим и наименьшим диаметрами детали в ряде надлежащим образом выбранных сечений и относится к существенно положительным величинам. Результаты измерений систематизированы в корреляционной таблице (табл. 73).

Рассчитанная по таблице эмпирическая линия регрессии y по x имеет отчетливо выраженный прямолинейный характер. Прямая регрессии, рассчитанная по способу наименьших квадратов, указана на рисунке (рис. 78). Приводим таблицу важнейших показателей точности операции (табл. 74).

Проверим соотношения (15.3.2) и (15.3.9).

$$(15.3.2) — 1,145 + 0,148 \cdot 12,62 = 3,012$$

$$(15.3.9) — 1,637 + 0,022 \cdot 34,66 = 2,396$$

Соотношения удовлетворяются.

Решим по данным предыдущего примера следующую задачу. Пусть удалось сократить средний размер входного качества x на 20%, а σ_x на 30%. Требуется оценить, в какой степени сократятся аналогичные показатели y и σ_y после операции хонингования. Исходя из допущения, что $\sigma_{y,x}$, a и b не изменились, подставим в формулы (15.3.2) и (15.3.9) прежние значения этих пока-

зателей, взятые из табл. 74, и новые значения \bar{x} и σ_x . Получим $\bar{y} = 2,64$, $\sigma_y = 1,417$. Следовательно, \bar{y} сократится на 12,35%, σ_y — на 8,46%.

Таблица 74

1. Средняя входная погрешность	$\bar{x} = 12,62$
2. Среднее квадратическое отклонение входных погрешностей	$\sigma_x = 5,887$
3. Дисперсия входной погрешности	$\sigma_x^2 = 34,66$
4. Коэффициент переноса погрешностей	$b = 0,148$
5. Постоянная собственная погрешность операции	$a = 1,145$
6. Среднеквадратическая собственная случайная погрешность операции	$\bar{\sigma}_{y,x} = 1,279$
7. Дисперсия собственных случайных погрешностей	$\sigma_{y,x}^2 = 1,637$
8. Средняя выходная погрешность	$\bar{y} = 3,012$
9. Среднее квадратическое отклонение выходных погрешностей	$\sigma_y = 1,548$
10. Дисперсия выходных погрешностей	$\sigma_y^2 = 2,396$

В предыдущем параграфе отмечалось, что для существенно положительных параметров качества средний размер параметра \bar{y} (или \bar{x}) и его дисперсия σ_y^2 (или σ_x^2) должны быть использованы для построения обобщающего показателя точности. В качестве такого показателя естественно принять величину интервала рассеяния погрешностей, гарантируемого с фиксированной, достаточно близкой к единице вероятностью, например, с вероятностью 0,999 или 0,9973. Методика определения этого показателя требует ознакомления с законами распределения существенно положительных величин (см. ниже).

§ 6. РАСПРЕДЕЛЕНИЕ СУЩЕСТВЕННО ПОЛОЖИТЕЛЬНЫХ ВЕЛИЧИН

Существенно положительные величины, характеризующие качество машиностроительной продукции (эксцентриситеты, биения, несоосности и др.), в большинстве случаев определяются разностью между наибольшим и наименьшим значениями какого-либо параметра, распределенного по нормальному закону. Обнаружено, что законы распределения таких величин близки к закону распределения размахов отклонений в случайных выборках из нормально распределенной совокупности¹. В связи с этим закон распределения размахов приобретает важное значение в прак-

¹ Я. И. Лукомский, Статистический анализ и контроль существенно положительных величин, характеризующих качество продукции, ж. «Стандартизация» № 1 и 2, 1955.

x'		-8	-7	-6	-5	-4	-3	-2	-1	0	1	2	3
y'	x	3	4	5	6	7	8	9	10	11	12	13	14
	y												
4	7												
3	6								1	3			
2	5									1	2	2	2
1	4			2			2	1	1	1	1	1	1
0	3			1	1	4	2	3	1	1			
-1	2	-1		-1	-3	-2	-3	-1	-1	-3	-3	-1	
		1		1	3	2	3	1	1	3	3	1	
-2	1	-2		-2	-2	-4	-12	-2	-2	-4	-2		-2
		1		1	1	2	6	1	1	2	1		1
-2.75	0				-2.75								
					1								
Итого h		2	0	5	6	8	13	6	5	8	6	2	2
	hx'	-16	0	-30	-30	-32	-39	-12	-5	0	6	4	6
	hx'^2	128	0	180	150	128	117	24	5	0	6	8	18
	$\Sigma_i my'$	-3	0	-1	-7.75	-6	-13	-2	1	-4	-2	1	-1
	$x' \Sigma_i my'^2$	24	0	6	38.75	24	39	4	-1	0	-2	2	-3
	\bar{y}'_i	-1.50	-	-0.20	-1.29	-0.75	-1.00	-0.33	0.20	-0.50	-0.33	0.50	-0.50
	\bar{y}_i	1.50	-	2.80	1.71	2.25	2.00	2.67	3.20	2.50	2.67	3.50	2.50

тике статистического анализа и контроля качества машиностроительной продукции. Ниже излагаются свойства этого закона и методика обработки опытных данных, следующих этому закону. Закон распределения размахов в выборках из нормально распределенной совокупности выражается формулой

$$\psi_\nu(x; a) = \frac{\nu(\nu-1)}{a^\nu (2\tau)^{\frac{\nu}{2}}} e^{-\frac{x^2}{4a^2}} \int_{-\infty}^{+\infty} e^{-\frac{\tau^2}{a^2}} \times \left[\int_{\tau-\frac{x}{2}}^{\tau+\frac{x}{2}} e^{-\frac{t^2}{2a^2}} dt \right]^{\nu-2} d\tau, \quad (15.6.1)$$

где $\psi_\nu(x; a)$ — плотность распределения в точке x ;
 ν и a — параметры закона распределения.

Таблица 73

4	5	6	7	8	9	10	11	12	13	14	15	16	Итого		
15	16	17	18	19	20	21	22	23	24	25	26	27	l	ly'	ly''
		4											1	4	16
	1												5	15	45
	3				3		3					3			
2		4		6	2		2			2		2	13	26	52
1		2		3	1		1			1		1	20	20	20
		1		2	3	2		1	2						
1	2	1			1								18	0	0
-1	-1			-1	-1								23	-23	23
1	1			1	1										
		-2		-2									19	-38	76
		1		1									1	-2,75	7,56
3	4	6	2	6	6	2	2	1	2	1	0	2	100	1,25	239,56
12	20	36	14	48	54	20	22	12	26	14	0	32	162		
48	100	216	98	384	486	200	242	144	338	196	0	512	3728		
1	2	7	-3	7	8	2	5	1	2	2	0	5	1,25		
4	10	42	-21	56	72	20	55	12	26	28	0	80	515,75		
0,33	0,50	1,17	-1,50	1,17	1,33	1,00	2,50	1,00	1,00	2,00	-	2,50			
3,33	3,50	4,17	1,50	4,17	4,33	4,00	5,50	4,00	4,00	5,00	-	5,50			

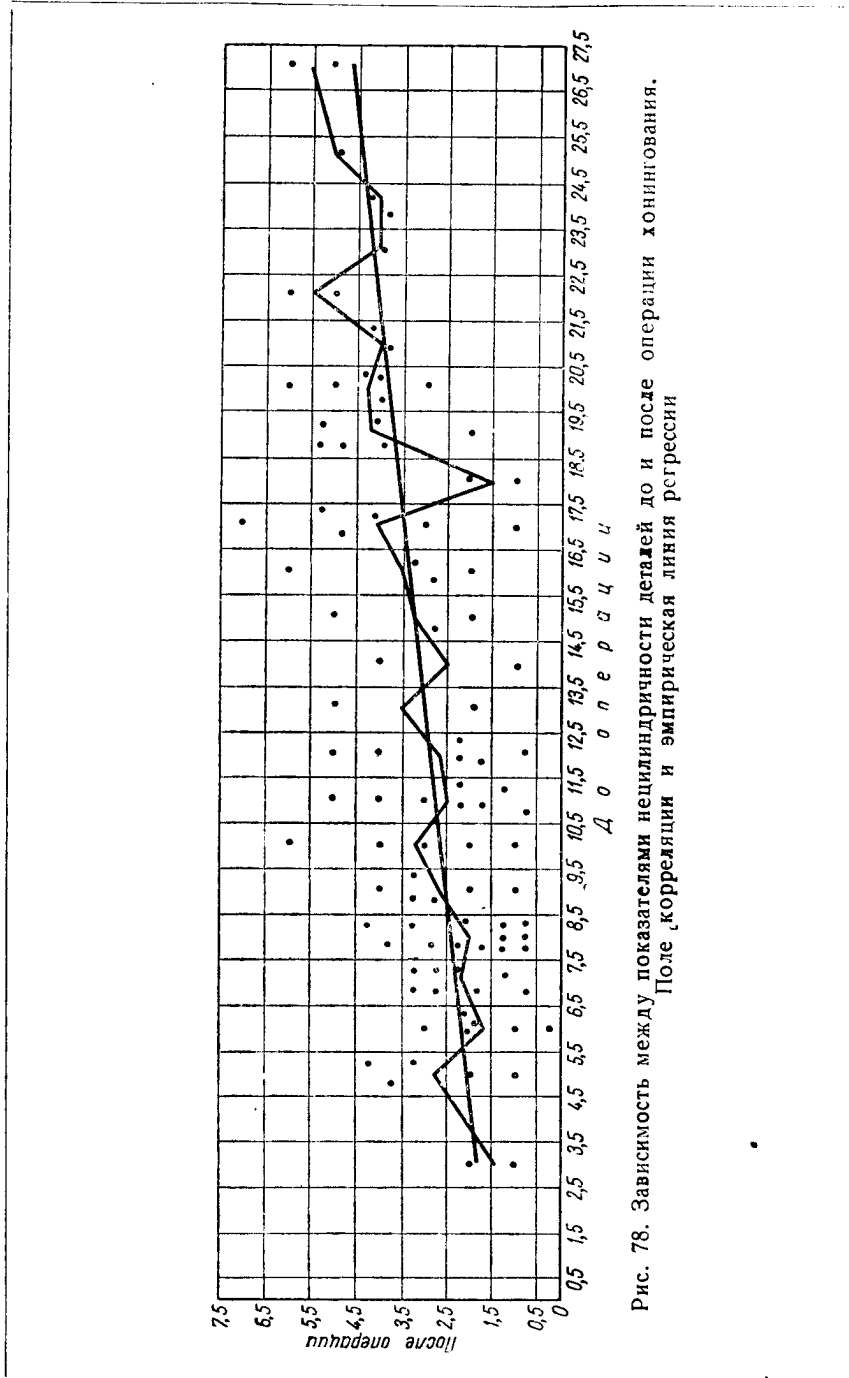
Заменой переменных

$$\left. \begin{aligned} \omega &= \frac{x}{a} \\ \psi_v(\omega; 1) &= a\psi_v(x; a) \end{aligned} \right\} \quad (15.6.2)$$

уравнение (15.6.1) приводится к нормированному виду с параметром $a = 1$.

Закон распределения (15.6.1) будем называть законом $a\omega$ и при $a = 1$ — законом ω .

Параметр ν определяет порядок закона распределения $a\omega$ (в случайных выборках он равен объему выборки n). При $\nu = 2$



кривая $a\omega$ представляет собой половину нормальной кривой. С увеличением ν кривая смещается в положительном направлении оси x , рассеяние кривой медленно сокращается. Любопытен характер изменения формы кривой $a\omega$. В области небольших зна-

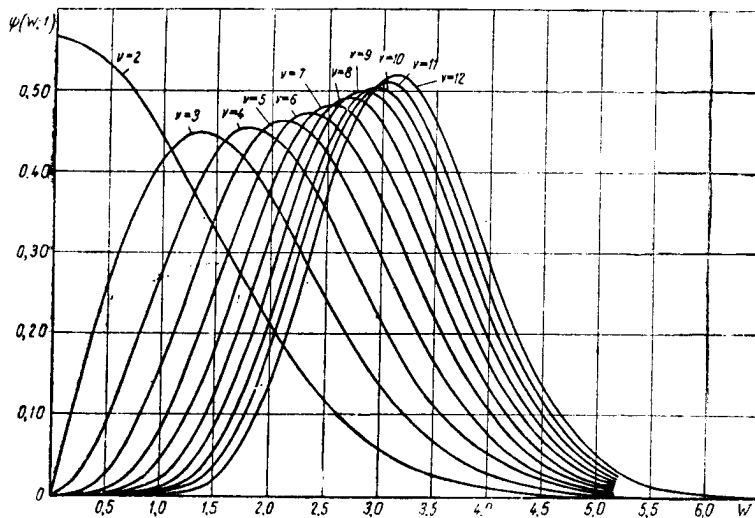


Рис. 79. Кривые распределения $a\omega$ при $a = 1$

чений ν , практически используемых для выравнивания распределений существенно положительных величин, с увеличением ν кривая становится более симметричной, затем при более значительных ν асимметрия и эксцесс кривой медленно увеличиваются. Это видно из рис. 79 и табл. 75¹, в которой приведены основные параметры, характеризующие положение, рассеяние и форму кривых распределения для целых значений ν при $a=1$: ω_ν — средние арифметические; σ_ν — средние квадратические отклонения; $V_\nu = \frac{\sigma_\nu}{\omega_\nu}$ — коэффициенты вариации; A_ν — показатели асимметрии; E_ν — показатели эксцесса.

С увеличением параметра a кривая распределения равномерно растягивается по оси абсцисс и сжимается по оси ординат таким образом, что вся площадь, ограниченная кривой и осью абсцисс, сохраняется неизменной и равной 1. При этом все безразмерные

¹ Таблица составлена по данным Л. Типпета (Biometrika, 1925) и Э. Пирсона (Biometrika, 1926), некоторые числа восстановлены по графикам.

характеристики распределения, как-то: A_v , E_v , V_v остаются без изменения. Размерные же показатели распределения \bar{w}_v и σ_v изменяются пропорционально величине a .

Таблица 75

v	\bar{w}_v	σ_v	V_v	A_v	E_v
2	1,1284	0,8525	0,756	0,9953	0,8692
3	1,6926	0,8884	0,525	0,6461	0,2864
4	2,0588	0,8798	0,427	0,5230	0,1884
5	2,3259	0,8641	0,371	0,4655	0,1693
6	2,5344	0,8480	0,335	0,4350	0,1698
7	2,7044	0,833	0,308	0,41	0,18
8	2,8472	0,820	0,288	0,40	0,19
9	2,9700	0,808	0,272	0,39	0,20
10	3,0775	0,797	0,259	0,39	0,22
11	3,1729	0,787	0,248	0,39	0,23
12	3,2585	0,778	0,239	0,39	0,24
20	3,735	0,729	0,195	0,40	0,26

Ниже излагается методика выравнивания опытных кривых распределения. В табл. 76 приводится в качестве примера ряд распределения колец по параметру «непараллельность торцов» после шлифования отверстия (см. рис. 80). В графе 1 таблицы указаны границы, в графе 2 — центры интервалов, получаемых при округлении значений признака в процессе измерения. Отметим, что длина интервала, который начинается в точке 0, как видно из таблицы, должна быть вдвое меньше длины каждого из последующих интервалов. В графе 3 приведены частоты соответствующих интервалов.

По данным таблицы рассчитываются средняя арифметическая, среднее квадратическое отклонение, коэффициент вариации. В нашем примере $\bar{x}=6,950$, $\sigma=2,255$, $V=0,324$. В табл. 75 находим значение V_v , наиболее близкое к рассчитанному значению 0,324. Это значение соответствует $v=6$. Поэтому в основу выравнивания опытного распределения принимается шестой порядок теоретического распределения (15.6.1).

Для определения параметра a находим предварительно a' и a'' по формулам:

$$a' = \frac{\bar{x}}{w_v}, \quad (15.6.3)$$

$$a'' = \frac{\sigma}{\sigma_v}. \quad (15.6.4)$$

Принимаем

$$a = \frac{a' + a''}{2}. \quad (15.6.5)$$

Таблица 76

Границы интервалов x_{zp}	Центры интервалов x	Частоты m	$w_{zp} = \frac{x_{zp}}{a}$	$P_v(w_{zp}; 1)$	Вероятности для интервалов ΔP_v	Теоретические частоты $m_1 = \Delta P_v \cdot \Sigma m$
0	0,5	0	0,000	0,0000	0,0001	0,03
1	2	7	0,370	0,0001	0,0300	8,91
3	4	46	1,111	0,0301	0,1904	56,55
5	6	111	1,852	0,2205	0,3350	99,49
7	8	78	2,592	0,5555	0,2725	80,93
9	10	45	3,333	0,8280	0,1260	37,42
11	12	8	4,073	0,9540	0,0373	11,08
13	14	1	4,814	0,9913	0,0075	2,23
15	16	1	5,554	0,9988	0,0011	0,33
17			6,295	0,9999	0,0001	0,03
—	—	297	—	—	—	297,00

В нашем примере $a' = 2,742$, $a'' = 2,659$, $a = 2,700$.

Для построения интерполяционной кривой следует использовать табл. 77, содержащую значения $\psi_v(w; 1)$ для некоторого набора значений w . При нанесении интерполяционной кривой на опытный график значения w умножаются на коэффициент $j \frac{\Sigma m}{a}$, где j — величина интервала. При построении опытного графика частоты интервалов рассматривают как средние плотности распределения на них, поэтому частота начального интервала, который вдвое короче других, должна быть удвоена. На рис. 80 представлены интерполяционные кривые $a\omega$ вместе с опытными графиками. Хорошее соответствие этих кривых очевидно.

Для определения вероятностей и теоретических частот распределения $a\omega$ можно воспользоваться табл. 78 кумулятивной функции распределения w :

$$P_v(w; 1) = \int_0^w \psi_v(t; 1) dt, \quad (15.6.6)$$

Таблица $\psi_y(w; 1)$ (среднее значение на интервале 0,05)

Таблица 77

w	2	3	4	5	6	7	8	9	10	11	12
	$\psi_y(w; 1)$										
0,000	0,564	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
0,225	0,556	0,122	0,020	0,002	0,000	0,000	0,000	0,000	0,000	0,000	0,000
0,475	0,532	0,246	0,082	0,022	0,006	0,002	0,000	0,000	0,000	0,000	0,000
0,725	0,494	0,346	0,170	0,072	0,028	0,010	0,004	0,002	0,002	0,000	0,000
0,975	0,444	0,412	0,270	0,150	0,076	0,038	0,018	0,008	0,004	0,002	0,000
1,225	0,386	0,446	0,360	0,246	0,156	0,094	0,054	0,030	0,016	0,010	0,004
1,475	0,328	0,444	0,422	0,342	0,254	0,178	0,122	0,080	0,050	0,032	0,012
1,725	0,268	0,418	0,452	0,416	0,352	0,280	0,216	0,160	0,118	0,084	0,058
1,975	0,212	0,370	0,444	0,458	0,428	0,378	0,320	0,264	0,214	0,170	0,132
2,225	0,164	0,312	0,412	0,460	0,470	0,450	0,416	0,374	0,326	0,280	0,238
2,475	0,122	0,252	0,356	0,428	0,466	0,482	0,476	0,458	0,426	0,392	0,356
2,725	0,088	0,194	0,292	0,372	0,432	0,470	0,492	0,496	0,490	0,478	0,456
2,975	0,062	0,144	0,226	0,304	0,368	0,422	0,462	0,488	0,506	0,512	0,514
3,225	0,042	0,102	0,170	0,234	0,296	0,352	0,398	0,438	0,470	0,496	0,514
3,475	0,028	0,070	0,118	0,172	0,224	0,274	0,322	0,364	0,404	0,436	0,466
3,725	0,018	0,046	0,082	0,120	0,162	0,202	0,242	0,282	0,320	0,352	0,386
3,975	0,010	0,028	0,052	0,080	0,108	0,144	0,172	0,204	0,234	0,266	0,296
4,225	0,008	0,018	0,034	0,050	0,070	0,092	0,114	0,138	0,164	0,188	0,210
4,475	0,002	0,010	0,020	0,032	0,044	0,060	0,074	0,090	0,106	0,124	0,142
4,725	0,002	0,006	0,010	0,018	0,026	0,034	0,046	0,054	0,066	0,076	0,090
4,975	0,002	0,002	0,006	0,010	0,016	0,020	0,026	0,032	0,038	0,046	0,052
5,225		0,002	0,004	0,006	0,008	0,010	0,014	0,018	0,022	0,026	0,030
5,475			0,002	0,004	0,004	0,006	0,008	0,010	0,012	0,014	0,016
5,725				0,002	0,004	0,004	0,004	0,004	0,008	0,008	0,008
5,975					0,002	0,002	0,002	0,004	0,004	0,004	0,004
6,225								0,002	0,002	0,002	0,002
Мода . . .	0,000	1,375	1,800	2,125	2,333	2,525	2,700	2,800	2,925	3,025	3,100
Модальная ордината	0,564	0,450	0,452	0,462	0,474	0,482	0,492	0,500	0,506	0,514	0,520

дающей вероятность значений $x \leq a\omega$. В графах 4—7 табл. 76 указан порядок расчета вероятностей и теоретических частот для интервалов x нашего примера.

*
*

На основе изложенной выше методики можно исчислять обобщающие показатели точности до и после операции по данным \bar{x} и σ_x и соответственно \bar{y} и σ_y . В качестве такого показателя принимается интервал рассеяния погрешностей, имеющий начало в

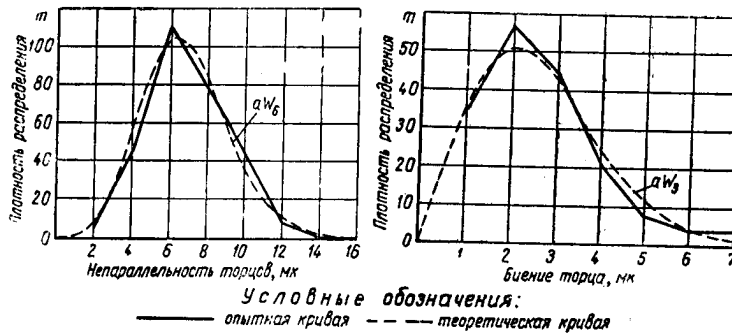


Рис. 80. Опытные и теоретические кривые распределения погрешностей

нулевой точке и гарантированный с вероятностью, фиксированной на уровне, достаточно близком к 1, например с вероятностью 0,9973.

Излагаем ниже методику расчета такого показателя по данным примера табл. 74.

Рассчитываем сперва коэффициенты вариации $V_x = \frac{\sigma_x}{\bar{x}}$ и $V_y = \frac{\sigma_y}{\bar{y}}$, а по коэффициентам вариации, полагая в основу расчетов семейство теоретических кривых распределения $a\omega$, — параметры ν и a .

Получаем:

$$V_x = \frac{5,89}{12,62} = 0,467, \quad V_y = \frac{1,548}{3,012} = 0,514.$$

Значение V_y близко к табличному $V_3 = 0,525$, следовательно, принимаем $\nu = 3$, откуда a_y оказывается равным 1,761. Пользуясь табл. 78, находим ω_y^* как корень уравнения

$$P_\nu(\omega_y^*; 1) = \int_0^{\omega_y^*} \psi_\nu(t; 1) dt = 0,9973.$$

Таблица вероят

w	v	2	3	4	5	6	7	8	9	10
0,00		0,0000	0,0000							
0,10		,0564	,0028	,0001						
0,20		,1125	,0110	,0010	,0001					
0,30		,1680	,0245	,0034	,0004	0,0000				
0,40		,2227	,0431	,0079	,0014	,0002	0,0000			
0,50		0,2763	0,0663	0,0152	0,0033	0,0007	0,0002	0,0000		
0,60		,3286	,0944	,0257	,0068	,0017	,0004	,0001	0,0000	
0,70		,3794	,1263	,0398	,0121	,0036	,0011	,0003	,0001	0,0000
0,80		,4284	,1616	,0578	,0200	,0068	,0023	,0003	,0002	,0001
0,90		,4755	,2000	,0797	,0309	,0117	,0044	,0016	,0006	,0002
1,00		0,5205	0,2407	0,1057	0,0450	0,0188	0,0078	0,0032	0,0013	0,0005
1,10		,5633	,2833	,1355	,0629	,0287	,0129	,0057	,0025	,0011
1,20		,6039	,3272	,1688	,0847	,0417	,0203	,0098	,0047	,0022
1,30		,6420	,3719	,2054	,1104	,0583	,0304	,0157	,0080	,0041
1,40		,6778	,4168	,2448	,1400	,0767	,0437	,0240	,0131	,0071
1,50		0,7112	0,4614	0,2865	0,1733	0,1031	0,0606	0,0353	0,0204	0,0117
1,60		,7421	,5053	,3299	,2011	,1316	,0814	,0499	,0304	,0184
1,70		,7707	,5481	,3745	,2498	,1639	,1064	,0684	,0437	,0277
1,80		,7969	,5894	,4197	,2920	,2000	,1355	,0910	,0607	,0403
1,90		,8209	,6290	,4649	,3361	,2394	,1686	,1178	,0818	,0565
2,00		0,8427	0,6655	0,5096	0,3816	0,2816	0,2056	0,1489	0,1072	0,0768
2,10		,8624	,7019	,5534	,4277	,3260	,2460	,1842	,1371	,1011
2,20		,8802	,7349	,5957	,4739	,3720	,2893	,2232	,1712	,1307
2,30		,8961	,7655	,6363	,5196	,4190	,3348	,2656	,2095	,1645
2,40		,9103	,7937	,6748	,5643	,4663	,3820	,3107	,2514	,2025
2,50		0,9229	0,8195	0,7110	0,6075	0,5132	0,4300	0,3579	0,2964	0,2443
2,60		,9340	,8429	,7448	,6487	,5592	,4782	,4064	,3437	,2894
2,70		,9438	,8640	,7759	,6877	,6036	,5259	,4555	,3927	,3372
2,80		,9523	,8828	,8045	,7242	,6461	,5725	,5044	,4425	,3867
2,90		,9597	,8996	,8304	,7580	,6863	,6174	,5525	,4923	,4372
3,00		0,9661	0,9145	0,8537	0,7891	0,7239	0,6601	0,5991	0,5415	0,4878
3,10		,9716	,9275	,8746	,8174	,7587	,7003	,6436	,5892	,5378
3,20		,9763	,9388	,8931	,8429	,7905	,7377	,6856	,6350	,5864
3,30		,9804	,9487	,9095	,8657	,8194	,7721	,7248	,6782	,6329
3,40		,9838	,9572	,9237	,8859	,8454	,8034	,7609	,7186	,6769
3,50		0,9867	0,9644	0,9361	0,9037	0,8685	0,8316	0,7939	0,7558	0,7180

Перепечатана в сокращенном виде из ж. Биометрика. т. XXXII, стр. 301

Таблица 78

ностей Р, (w; 1)

11	12	13	14	15	16	17	18	19	20
.0001									
0,0002	0,0001	0,0000							
,0005	,0002	,0001	0,0000						
,0010	,0005	,0002	,0001	0,0000					
,0021	,0010	,0005	,0003	,0001	,0001	0,0000			
,0038	,0021	,0011	,0006	,0003	,0002	,0001	,0001		
0,0067	0,0038	0,0022	0,0012	0,0007	0,0004	0,0002	0,0001	0,0001	0,0000
,0111	,0067	,0040	,0024	,0014	,0008	,0005	,0003	,0002	,0001
,0175	,0111	,0069	,0043	,0027	,0017	,0011	,0007	,0004	,0003
,0266	,0175	,0115	,0075	,0049	,0032	,0021	,0013	,0009	,0006
,0388	,0266	,0182	,0124	,0084	,0057	,0039	,0026	,0018	,0012
0,0548	0,0389	0,0276	0,0195	0,0137	0,0097	0,0068	0,0048	0,0033	0,0023
,0749	,0550	,0403	,0295	,0215	,0056	,0114	,0082	,0060	,0043
,0994	,0753	,0564	,0429	,0323	,0242	,0181	,0136	,0102	,0076
,1226	,1003	,0779	,0605	,0468	,0361	,0279	,0215	,0165	,0127
,1625	,1300	,1037	,0825	,0655	,0519	,0411	,0325	,0256	,0202
0,2007	0,1644	0,1342	0,1094	0,0890	0,0722	0,0586	0,0474	0,0383	0,0309
,2429	,2033	,1697	,1413	,1174	,0974	,0807	,0668	,0552	,0455
,2885	,2462	,2096	,1780	,1509	,1278	,1080	,0911	,0768	,0647
,3368	,2926	,2536	,2194	,1894	,1632	,1405	,1208	,1037	,0889
,3870	,3417	,3011	,2647	,2324	,2006	,1782	,1558	,1360	,1186
0,4382	0,3927	0,3512	0,3134	0,2792	0,2484	0,2207	0,1959	0,1737	0,1538
,4895	,4446	,4029	,3645	,3292	,2970	,2675	,2407	,2164	,1943
,5401	,4965	,4555	,4171	,3814	,3483	,3177	,2895	,2635	,2396
,5893	,5475	,5078	,4703	,4348	,4016	,3704	,3413	,3142	,2890
,6363	,5970	,5592	,5230	,4885	,4557	,4246	,3953	,3677	,3417
0,6807	0,6442	0,6087	0,5744	0,5413	0,5096	0,4792	0,4502	0,4226	0,3964

w \ v	2	3	4	5	6	7	8	9	10
3,60	,9891	,9706	,9468	,9192	,8889	,8568	,8236	,7898	,7558
3,70	,9911	,9759	,9559	,9326	,9067	,8790	,8501	,8204	,7902
3,80	,9928	,9803	,9637	,9441	,9222	,8985	,8736	,8477	,8212
3,90	,9942	,9839	,9703	,9540	,9355	,9155	,8941	,8718	,8488
4,00	0,9953	0,9870	0,9758	0,9623	0,9469	0,9300	0,9120	0,8919	0,8731
4,10	,9963	,9895	,9804	,9693	,9566	,9425	,9273	,9112	,8943
4,20	,9970	,9916	,9842	,9752	,9647	,9530	,9404	,9269	,9126
4,30	,9976	,9933	,9874	,9800	,9715	,9619	,9514	,9402	,9283
4,40	,9981	,9947	,9899	,9840	,9771	,9692	,9607	,9514	,9416
4,50	0,9985	0,9958	0,9920	0,9873	0,9817	0,9754	0,9684	0,9608	0,9527
4,60	,9989	,9967	,9937	,9899	,9855	,9804	,9747	,9686	,9620
4,70	,9991	,9974	,9951	,9921	,9885	,9845	,9799	,9750	,9696
4,80	,9993	,9980	,9962	,9938	,9910	,9878	,9842	,9802	,9759
4,90	,9995	,9985	,9970	,9952	,9930	,9904	,9876	,9844	,9810
5,00	0,9996	0,9988	0,9977	0,9963	0,9946	0,9926	0,9903	0,9878	0,9851
5,10	,9997	,9991	,9982	,9971	,9958	,9942	,9925	,9906	,9884
5,20	,9998	,9993	,9986	,9978	,9968	,9956	,9942	,9927	,9911
5,30	,9998	,9995	,9990	,9983	,9975	,9966	,9956	,9944	,9931
5,40	,9999	,9996	,9992	,9987	,9981	,9974	,9966	,9957	,9948
5,50	0,9999	0,9997	0,9994	0,9991	0,9986	0,9981	0,9975	0,9968	0,9960
5,60	,9999	,9998	,9996	,9993	,9989	,9985	,9981	,9976	,9970
5,70	,9999	,9998	,9997	,9995	,9992	,9989	,9986	,9982	,9977
5,80		,9999	,9998	,9996	,9994	,9992	,9989	,9986	,9983
5,90		,9999	,9998	,9997	,9996	,9994	,9992	,9990	,9988
6,00		0,9999	0,9999	0,9998	0,9997	0,9996	0,9994	0,9993	0,9991
6,10		,9999	,9999	,9998	,9998	,9997	,9996	,9995	,9993
6,20			,9999	,9999	,9998	,9998	,9997	,9996	,9995
6,30			,9999	,9999	,9999	,9998	,9998	,9997	,9996
6,40				,9999	,9999	,9999	,9998	,9998	,9997
6,50					0,9999	0,9999	0,9999	0,9999	0,9998
6,60					1,0000	,9999	,9999	,9999	,9999
6,70						1,0000	0,9999	,9999	,9999
6,80								0,9999	,9999
6,90									0,9999
7,00									
7,10									
7,20									

Продолжение

11	12	13	14	15	16	17	18	19	20
,7220	,6886	,6558	,6237	,5926	,5624	,5332	,5051	,4781	,4522
,7000	,7298	,6998	,6704	,6414	,6132	,5856	,5588	,5329	,5078
,7944	,7675	,7406	,7138	,6873	,6613	,6357	,6106	,5861	,5622
,8254	,8016	,7777	,7537	,7298	,7061	,6827	,6596	,6369	,6145
0,8528	0,8321	0,8111	0,7899	0,7686	0,7474	0,7263	0,7053	0,6845	0,6640
,8769	,8590	,8408	,8223	,8036	,7848	,7660	,7472	,7285	,7099
,8978	,8826	,8669	,8509	,8347	,8183	,8018	,7852	,7686	,7520
,9159	,9029	,8896	,8760	,8620	,8479	,8336	,8191	,8046	,7899
,9312	,9204	,9092	,8976	,8858	,8737	,8614	,8490	,8364	,8237
0,9441	0,9352	0,9258	0,9162	0,9062	0,8960	0,8856	0,8750	0,8643	0,8534
,9550	,9476	,9399	,9319	,9236	,9151	,9064	,8975	,8884	,8791
,9640	,9579	,9516	,9451	,9383	,9312	,9240	,9165	,9090	,9012
,9713	,9665	,9614	,9560	,9505	,9447	,9387	,9326	,9264	,9199
,9774	,9735	,9694	,9650	,9605	,9559	,9510	,9460	,9409	,9356
0,9822	0,9791	0,9759	0,9724	0,9688	0,9650	0,9611	0,9571	0,9529	0,9486
,9861	,9837	,9811	,9784	,9755	,9725	,9694	,9661	,9628	,9593
,9893	,9874	,9853	,9832	,9809	,9785	,9760	,9735	,9708	,9681
,9917	,9903	,9887	,9870	,9852	,9833	,9814	,9794	,9773	,9751
,9937	,9925	,9913	,9900	,9886	,9872	,9856	,9841	,9824	,9807
0,9952	0,9943	0,9934	0,9924	0,9913	0,9902	0,9890	0,9878	0,9865	0,9852
,9964	,9957	,9950	,9942	,9934	,9925	,9916	,9907	,9897	,9887
,9973	,9968	,9962	,9956	,9950	,9944	,9937	,9929	,9922	,9914
,9980	,9976	,9972	,9967	,9963	,9958	,9952	,9947	,9941	,9935
,9985	,9982	,9979	,9976	,9972	,9968	,9964	,9960	,9956	,9952
0,9989	0,9987	0,9984	0,9982	0,9979	0,9977	0,9974	0,9971	0,9967	0,9964
,9992	,9990	,9989	,9987	,9985	,9983	,9981	,9978	,9976	,9973
,9994	,9993	,9992	,9990	,9989	,9987	,9986	,9984	,9982	,9980
,9996	,9995	,9994	,9993	,9992	,9991	,9990	,9988	,9987	,9986
,9997	,9996	,9996	,9995	,9994	,9993	,9992	,9992	,9991	,9990
0,9998	0,9997	0,9997	0,9996	0,9996	0,9995	0,9995	0,9994	0,9993	0,9993
,9998	,9998	,9998	,9997	,9997	,9997	,9996	,9996	,9995	,9995
,9999	,9999	,9998	,9998	,9998	,9998	,9997	,9997	,9997	,9996
,9999	,9999	,9999	,9999	,9998	,9998	,9998	,9998	,9998	,9997
,9999	,9999	,9999	,9999	,9999	,9999	,9999	,9998	,9998	,9998
	1,0000	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999
		1,0000	1,0000	0,9999	,9999	,9999	,9999	,9999	,9999
						1,0000	1,0000	0,9999	0,9999

При $\nu = 3$ получаем $\omega_y^* = 4,69$. Наконец искомый показатель L_y находим по формуле:

$$L_y = a_y \omega_y^*;$$

откуда

$$L_y = 8,25.$$

При помощи табл. 77 строим теоретическую кривую распределения деталей после операции (рис. 81) и сопоставляем ее с

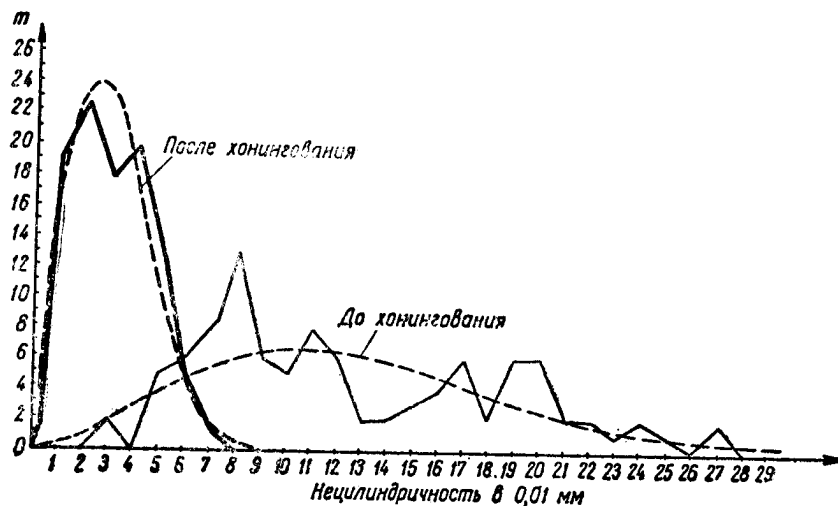


Рис. 81. Опытные графики и теоретические кривые распределения нецилиндричности деталей до и после операции хонингования

опытным графиком, построенным по соответствующим итогам корреляционной табл. 73. Хорошее соответствие опытного графика теоретической кривой подтверждает правильность проведенного здесь расчета показателя L_y .

Аналогично должен рассчитываться показатель точности до операции — L_x . Однако здесь дело осложняется тем, что фактический коэффициент вариации недостаточно близок ни к одному из табличных значений V , а занимает промежуточное положение между $V_3 = 0,525$ и $V_4 = 0,427$. Следовательно, нельзя ожидать достаточно хорошей близости между опытным графиком распределения и теоретической кривой $a\omega$, построенными в предположении $\nu_x = 3$ или $\nu_x = 4$, что и подтверждается проведенным нами специальным расчетом. Естественно искать теоретическую кривую с индексом V имеющим промежуточное значение между 3 и 4. Мы подошли к построению такой кривой следующим весьма приближенным, но простым и оправдывающим себя на практике

способом — элементы искомой теоретической кривой, как-то: ординаты, чистоты интервалов, а также параметр a_x и величина w_x^* , рассчитаны как средние арифметические из соответствующих элементов кривых, полученных для индексов $V_x=3$ и $V_x=4$; условно новой кривой был присвоен индекс $V_x=3,5$. Получили соответственно: $a_x=6,73$, $w_x^*=4,815$ и $L_x=32,40$.

На рис. 81 сопоставлены опытный график распределения деталей до операции с полученной теоретической кривой. Соответствие оказывается вполне удовлетворительным, что подтвердило правильность примененной методики расчета L_x .

Сопоставляя значения $L_x=32,40$ и $L_y=8,25$, мы устанавливаем, что в результате операции интервал рассеяния кривой распределения, гарантированный с вероятностью 0,9973, сократился на 74,5%.

В задаче на стр. 330 расчет показателей L после сокращения x на 20% и σ_x на 30% дал значения $L_x=25,08$ и $L_y=7,398$. Следовательно, в то время как L_x сократился на 22,6%, L_y сократился лишь на 10,3%.

§ 7. О ПОПРАВКАХ НА НЕНАДЕЖНОСТЬ ИЗМЕРЕНИЯ И НА ГРУППИРОВКУ ДАННЫХ

Процесс исправления погрешностей проявляется в некоррелированности показателей входного и выходного качества. Мы видели, что чем сильнее исправление, тем меньше коэффициент регрессии y по x . Однако ослабление корреляции и уменьшение коэффициента регрессии может происходить также и по другим причинам, не связанным с процессом погашения погрешностей. К этим причинам в первую очередь принадлежат ненадежность измерений и группировка данных. Во избежание неверных заключений необходимо показатели, характеризующие процесс исправления погрешностей, *корректировать* на ненадежность измерений и группировку данных. Ниже покажем как это делается.

Случайные погрешности, возникающие в процессе измерения, можно разделить на две группы:

1. *Погрешности округления*, заключающиеся в том, что смежные значения признака округляются в процессе измерения к центру интервала, в который они включаются. Величина интервала округления заранее определена ценой деления прибора, и ее половина составляет максимальную погрешность округления.

2. Ко второй группе относятся *ошибки воспроизведения*, которые заключаются в том, что при повторных измерениях одной и той же величины центр интервала округления обнаруживает некоторую колеблемость.

Характеристикой ошибок воспроизведения может служить дисперсия и среднее квадратическое отклонение таких погрешностей, полученных в результате многократных повторений изме-

рения. Ошибки округления по своему характеру родственны *ошибкам группировки*, которые также заключаются в том, что значения признака, попавшие в интервалы группировки, заменяются при дальнейших расчетах центрами этих интервалов. Поэтому поправки на группировку и поправки на округление аналогичны друг другу.

Поскольку интервал группировки состоит из нескольких интервалов округления, то при расчете поправок на группировку можно не вносить поправки на округление значений. Последние необходимо вносить лишь в тех случаях, когда последующей группировки материала не производится.

Известно, что в результате случайных погрешностей измерения, а также последующей группировки данных некоторые показатели статистического материала систематически преувеличиваются или систематически преуменьшаются. В таких случаях необходимо в них вносить соответствующие поправки, освобождающие эти показатели от систематической погрешности.

Первыми в ряду таких показателей являются дисперсии признака: σ_x^2 — дисперсия входного качества и σ_y^2 — дисперсия выходного качества. Оба эти показателя систематически преувеличиваются в результате случайных погрешностей измерения и группировки.

Для устранения этих погрешностей следует использовать формулы:

$$\left. \begin{aligned} \hat{\sigma}_x^2 &= \sigma_x^2 - \frac{i_x^2}{12} - \sigma_{\varepsilon(x)}^2 \\ \hat{\sigma}_y^2 &= \sigma_y^2 - \frac{i_y^2}{12} - \sigma_{\varepsilon(y)}^2 \end{aligned} \right\}, \quad (15.7.1)$$

где $\hat{\sigma}_x^2$ и $\hat{\sigma}_y^2$ обозначают скорректированные дисперсии входного и выходного качества,

i_x и i_y — величины соответствующих интервалов группировки, а если группировки не производится — то величины соответствующих интервалов округления,

$\sigma_{\varepsilon(x)}^2$ и $\sigma_{\varepsilon(y)}^2$ — дисперсии случайных погрешностей воспроизведения, которые следует рассчитывать, многократно измеряя одну и ту же величину x или y , заменяя ее центром интервала округления, если группировка не производится, и вычисляя дисперсию полученных данных.

Формулы (15.7.1) основаны:

- 1) на предположении, что ошибки воспроизведения $\varepsilon(x)$ или $\varepsilon(y)$ не зависят от измеряемых величин x и y ,
- 2) на применении поправок Шеппарда¹ к округленным в процессе группировки или измерения значениям.

¹ См. [7], гл. 27.

Формулы (15.7.1) являются основой для внесения поправок во все остальные показатели.

Коэффициент регрессии b исчисляется по формуле

$$b = \frac{\mu_{xy}}{\sigma_x^2}, \quad (15.7.2)$$

где μ_{xy} — ковариация y по x , которая не корректируется на погрешность измерения и на группировку.

Следовательно, скорректированное значение b мы получим, корректируя знаменатель формулы (15.7.2), т. е. по формуле:

$$\hat{b} = \frac{\mu_{xy}}{\hat{\sigma}_x^2} \quad (15.7.3)$$

или

$$\hat{b} = b \frac{1}{1 - \frac{i_x^2}{12\sigma_x^2} - \frac{\sigma_{\epsilon(x)}^2}{\sigma_x^2}}, \quad (15.7.4)$$

где \hat{b} — скорректированное значение b .

Отсюда видно, что значение b после корректировки *увеличивается*.

Из уравнения (15.3.2) $\bar{y} = a + b\bar{x}$, мы получаем скорректированное значение a по формуле

$$\hat{a} = \bar{y} - \hat{b}\bar{x} \quad (15.7.5)$$

Формула (15.7.5) основана на том обстоятельстве, что средние величины \bar{x} и \bar{y} не имеют систематических погрешностей и, следовательно, не корректируются. Из формулы (15.7.5) видно, что в результате корректировки величина a *уменьшается*, следовательно, до корректировки она была *преувеличена*.

Из формулы (15.3.9) $\sigma_y^2 = \sigma_{y.x}^2 + b^2\sigma_x^2$ можно получить корректировку для показателя $\sigma_{y.x}^2$ по формуле

$$\hat{\sigma}_{y.x}^2 = \hat{\sigma}_y^2 - \hat{b}^2\hat{\sigma}_x^2. \quad (15.7.6)$$

Так как

$$\hat{b}^2\hat{\sigma}_x^2 = \frac{\mu_{xy}^2}{\hat{\sigma}_x^2}, \quad (15.7.7)$$

то из формулы (15.7.6) видно, что показатель $\hat{\sigma}_{y,x}^2$ после коррекции *уменьшается* тем сильнее, чем больше сокращается после коррекции σ_y^2 и σ_x^2 . Следовательно, до коррекции величина $\sigma_{y,x}^2$ была преувеличена.

§ 8. РАСЧЕТ ПРОСТОЙ ТЕХНОЛОГИЧЕСКОЙ ЦЕПИ

Зная показатели $a_i, b_i, \sigma_{i, i-1}^2$ для ряда m последовательных операций, образующих *простую* технологическую цепь и показатели исходного качества \bar{x}_0 и σ_0^2 , мы можем, последовательно применяя формулы (15.3.2) и (15.3.9), оценить показатели \bar{x}_m и σ_m^2 выходного качества после конечной m -й операции по формулам:

$$\bar{x}_m = \{[(b_1\bar{x}_0 + a_1) b_2 + a_2] b_3 + a_3\} \dots \quad (15.8.1)$$

или

$$\bar{x}_m = A_0\bar{x}_0 + A_1a_1 + A_2a_2 + \dots + A_m a_m, \quad (15.8.2)$$

где

$$A_0 = \prod_{i=1}^m b_i, \quad A_1 = \prod_{i=2}^m b_i, \quad A_2 = \prod_{i=3}^m b_i, \dots, \quad A_{m-1} = b_m, \quad A_m = 1.$$

$$\sigma_m^2 = \{[(b_1^2\sigma_0^2 + \sigma_{1,0}^2) b_2^2 + \sigma_{2,1}^2] b_3^2 + \sigma_{3,2}^2\} \dots \quad (15.8.3)$$

или

$$\sigma_m^2 = B_0\sigma_0^2 + B_1\sigma_{1,0}^2 + B_2\sigma_{2,1}^2 + \dots + B_m\sigma_{m, m-1}^2, \quad (15.8.4)$$

где

$$B_0 = \prod_{i=1}^m b_i^2, \quad B_1 = \prod_{i=2}^m b_i^2, \quad B_2 = \prod_{i=3}^m b_i^2, \dots, \quad B_{m-1} = b_m^2, \quad B_m = 1.$$

По-прежнему предполагается прямолинейность регрессии $\bar{x}_{i(i-1)} = a_i + b_i x_{i-1}$ и постоянство условной дисперсии $\sigma_{i, i-1}^2$ для всех звеньев технологической цепи.

Остановимся на физическом смысле формул (15.3.2) и (15.3.9).

Показатели выходного качества в конечном звене \bar{x}_m и σ_m^2 представлены в виде суммы $m+1$ компонентов. Компонент $B_0\sigma_0^2$ является произведением дисперсии погрешностей, имевшихся в начале процесса на квадраты коэффициентов переноса этих погрешностей вдоль технологической цепи от звена к звену. Каж-

дый из следующих компонентов $B_1\sigma_{1.0}^2, B_2\sigma_{2.1}^2, \dots$ является произведением дисперсии собственных случайных погрешностей, дополнительно возникающих на 1, 2, ... m -м звеньях цепи, умноженной на квадраты коэффициентов переноса этих погрешностей на последующие операции вплоть до конечного звена цепи.

Таким образом, компонент $B_i\sigma_{i, i-1}^2$ измеряет ту часть дисперсии σ_m^2 , которая создается случайными погрешностями, поступающими на конечное звено с i -го звена, но не в порядке переноса с $(i-1)$ -го звена.

Аналогичный смысл имеют для существенно положительных величин компоненты формулы (15. 8. 2) в применении к \bar{x}_m . Числа $A_i a_i$ показывают, какая часть \bar{x}_m создается постоянными погрешностями, поступающими с i -го звена на конечное звено, но не в порядке переноса с $(i-1)$ -го звена. Для параметров первого рода формула (15. 8. 2) практического значения, по-видимому, не имеет, так как показатели a_i регулируются настройкой.

Для выяснения эффекта от повторного применения одной и той же операции может быть использован следующий метод (при условии, что припуски достаточны и вообще технологический процесс остается неизменным).

Обозначим через u_i и v_i сокращение дисперсии и средней (предполагаются существенно положительные величины) после повторения операции:

$$u_i = \sigma_{i-1}^2 - \sigma_i^2,$$

$$v_i = \bar{x}_{i-1} - \bar{x}_i,$$

$$i = 1, 2, \dots, m.$$

Применяя формулы (15. 3. 2) и (15. 3. 9) и считая, что

$$b_1 = b_2 = \dots = b_m,$$

$$a_1 = a_2 = \dots = a_m,$$

$$\sigma_{1.0}^2 = \sigma_{2.1}^2 = \dots = \sigma_{m, m-1}^2,$$

получаем методом редукции

$$u_m = \sigma_{m-1}^2 - \sigma_m^2 = b^2 (\sigma_{m-2}^2 - \sigma_{m-1}^2) = b^4 (\sigma_{m-3}^2 - \sigma_{m-2}^2),$$

$$v_m = \bar{x}_{m-1} - \bar{x}_m = b (\bar{x}_{m-2} - \bar{x}_{m-1}) = b^2 (\bar{x}_{m-3} - \bar{x}_{m-2}).$$

Следовательно,

$$u_m = b^{2(m-1)} u_1, \quad (15. 8. 5)$$

$$v_m = b^{m-1} v_1, \quad (15. 8. 6)$$

где u_1 , v_1 — величины сокращения дисперсии и средней после первого прохода. Очевидно, что данная операция вообще ведет к сокращению погрешностей данного вида, если $u_1 > 0$, $v_1 > 0$, т. е. величина исправления имевшихся до операции погрешностей перекрывает величину собственных погрешностей данной операции.

При повторении данной операции выигрыш в точности *сокращается в геометрической прогрессии*, как это видно из формул (15.8.5) и (15.8.6), если b является, как обычно, положительной правильной дробью.

ГЛАВА XVI

ПРИМЕНЕНИЕ МЕТОДОВ КОРРЕЛЯЦИИ К АНАЛИЗУ ФАКТОРОВ, ВЛИЯЮЩИХ НА НАКЛАДНЫЕ РАСХОДЫ

§ 1. Для планирования накладных расходов в строительстве при разработке соответствующих нормативов большое значение имеет правильная оценка значимости факторов, влияющих на уровень накладных расходов. Такая оценка может быть получена на основе анализа работы строительных организаций за прошлый период.

Обобщая данные отчетности многих организаций интересующей нас отрасли строительства, при помощи методов теории корреляции можно решать задачи, представляющие пока почти непреодолимые затруднения для других методов экономического анализа, — задачи разделения влияния многих факторов, действующих совместно во взаимной связи друг с другом.

Настоящая глава составлена на основе материалов такого исследования, выполненного сотрудниками института ВНИИом-промжилстрой гг. А. В. Заруба, И. С. Пальма, Е. Г. Эльгорт при участии автора по данным отчетности за 1951—1954 гг. 113 шахтостроительных организаций. Более полное освещение экономической стороны и строительной специфики исследования дано в статье гг. Заруба, Пальма и Эльгорт, публикуемой в Ученых записках по статистике Академии наук СССР.

Накладные расходы строительных организаций включают, как известно, в основном следующие статьи затрат: административно-хозяйственные расходы, дополнительная зарплата, коммунальные расходы, износ малоценного и быстроизнашивающегося инвентаря и инструмента, прочие расходы.

На фактический уровень накладных расходов наибольшее влияние оказывают следующие факторы: объем выполненных работ, численность рабочих, выработка на одного рабочего, средняя зарплата и фонд зарплаты.

Хорошо известно общее направление влияния каждого из этих факторов на уровень накладных расходов. Однако эти сведения совершенно недостаточны для выяснения роли каждого фактора в общем комплексе причин, обуславливающих уровень накладных расходов на том или ином строительном предприятии. Лишь в результате статистической обработки опытного материала намечаются те количественные пропорции и соотношения, которые после экономического осмысливания могут быть положены в основу инженерно-экономических расчетов. Для выяснения этих соотношений из годовой отчетности строительных предприятий были взяты следующие показатели:

- 1) величина фактических накладных расходов (в млн. руб.),
- 2) объем фактически выполненных капитальных работ (в сметной стоимости в млн. руб.),
- 3) численность строительно-монтажных рабочих (среднесписочное число),
- 4) годовой фонд зарплаты строительно-монтажных рабочих (в млн. руб.).

Эти данные составляют круг исходных показателей нашего исследования. Показатели средней выработки на одного рабочего и средней годовой зарплаты в исследование непосредственно не вводились, так как они являются производными. Их влияние на уровень накладных расходов определялось на основе зависимостей, полученных для исходных показателей.

Причины, по которым к кругу основных показателей были отнесены объемные, а не относительные («качественные») показатели хозяйственной деятельности, будут раскрыты в процессе обработки материала.

Материал 113 годовых отчетов строек, включенных в обработку, распределялся по годам следующим образом:

за 1951 г. — 27	годовых отчетов
за 1952 г. — 29	„ „
за 1953 г. — 30	„ „
за 1954 г. — 27	„ „

Всего 113 годовых отчетов

Ввиду малочисленности материала, представленного за отдельные годы, нецелесообразно было производить его разработку по отдельным годам, и весь материал рассматривался в целом. При этом, естественно, возникает вопрос, в какой мере данные за различные годы однородны — не происходило ли систематического изменения исследуемых соотношений на протяжении рассматриваемого промежутка времени. Это обстоятельство подвергалось специальной проверке методом, изложенным ниже. Проверка не обнаружила существенных различий между материалами раз-

ных лет с точки зрения их соответствия общей закономерности, установленной в результате исследования. Следовательно, объединение материала было оправдано.

Следующий вопрос возникает в связи с тем обстоятельством, что некоторые номера строек представлены отчетами за несколько лет и соответственно фигурируют в материале два, три и четыре раза. Поскольку между данными одних и тех же строек за ряд лет можно ожидать некоторой корреляции, т. е. такие данные должны отличаться между собой меньше, чем данные различных строек, то наш материал, состоящий из 113 годовых отчетов строек, в действительности не представляет собой 113 «независимых испытаний». Его надежность эквивалентна надежности материала, получаемого на несколько меньшем количестве «независимых испытаний». Следовательно, подобное явление, не ставя под сомнение правомерность объединения материала различных лет, ведет к некоторому снижению надежности получаемых соотношений. Поэтому расчет случайных погрешностей показателей, произведенный в § 6 по формулам, справедливым в условиях процесса накопления (α), в нашем случае приводит к некоторому преуменьшению величины случайных погрешностей.

§ 2. Задачей исследования являлось установление зависимости между величиной накладных расходов и следующими основными факторами: объемом выполненных работ, численностью рабочих и фондом зарплаты (здесь будут исследованы парные зависимости между этими признаками, в следующем параграфе — множественная корреляция). Приводим корреляционные поля и линии регрессии между накладными расходами и каждым из трех рассматриваемых факторов.

При рассмотрении этих материалов обращает на себя внимание четкий прямолинейный характер всех трех линий регрессии. Точки корреляционного поля, изображающие данные отдельных строек, кучно группируются вокруг линии регрессии. С большой четкостью обнаруживается характерный факт: *соотношения показателей, характеризующих различные стороны хозяйственной деятельности отдельных предприятий данной отрасли, как бы специфичны ни были индивидуальные условия их работы, образуют поле рассеяния, на котором при достаточно большом числе предприятий вырисовываются общие закономерности.*

Эти закономерности получают точную количественную характеристику при помощи методов математической статистики.

Для зависимостей, изображенных на рис. 82, 83, 84, получены следующие уравнения регрессии:

$$\bar{x}_{1.2} = 1,241 + 0,153x_2, \quad (16.2.1)$$

$$\bar{x}_{1.3} = 0,817 + 0,00429x_3, \quad (16.2.2)$$

$$\bar{x}_{1.4} = 1,152 + 0,375x_4. \quad (16.2.3)$$

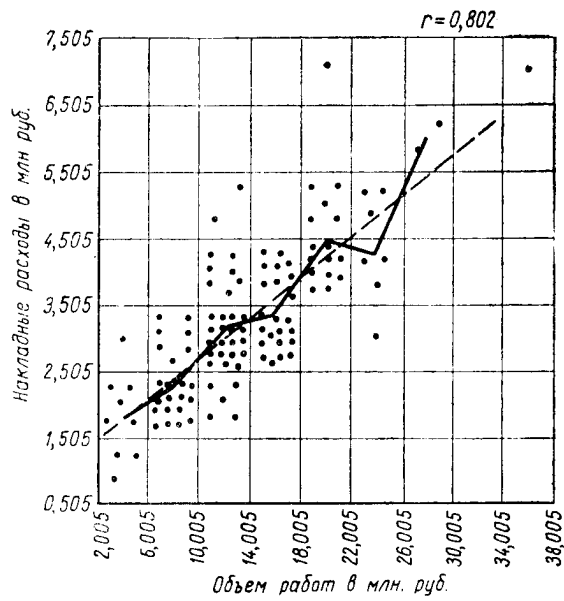


Рис. 82. Зависимость между величиной накладных расходов и объемом работ

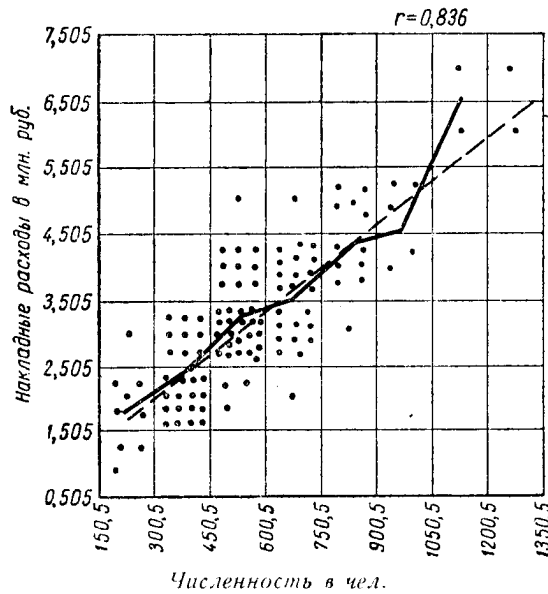


Рис. 83. Зависимость между величиной накладных расходов и численностью рабочих

Соответствующие прямые регрессии нанесены на рис. 82, 83, 84.

Коэффициенты корреляции накладных расходов с каждым фактором оказались:

с объемом работ $r_{12} = 0,802$,

с численностью рабочих $r_{13} = 0,836$,

с фондом зарплаты $r_{14} = 0,833$.

Значения коэффициентов свидетельствуют о наличии тесной зависимости накладных расходов от изучаемых факторов, причем зависимость от фонда зарплаты и численности рабочих заметно более тесная, чем от объема работ.

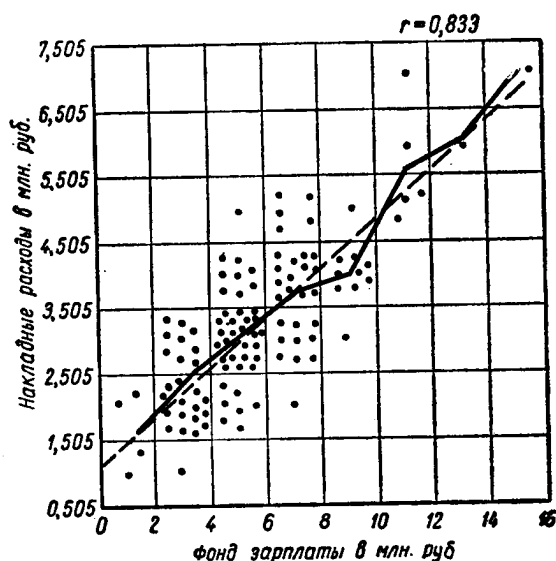
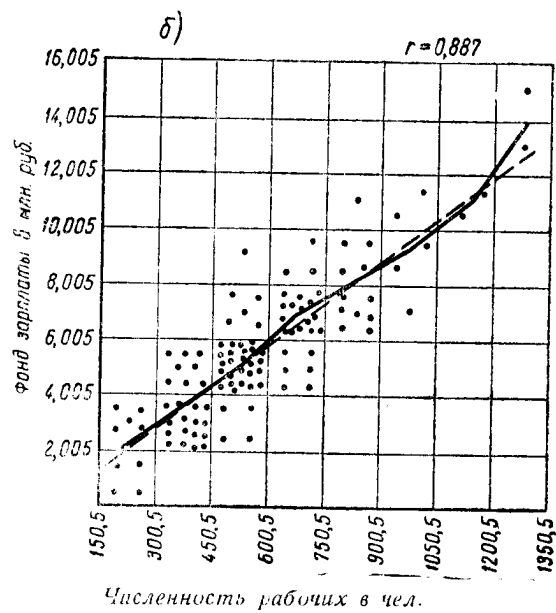
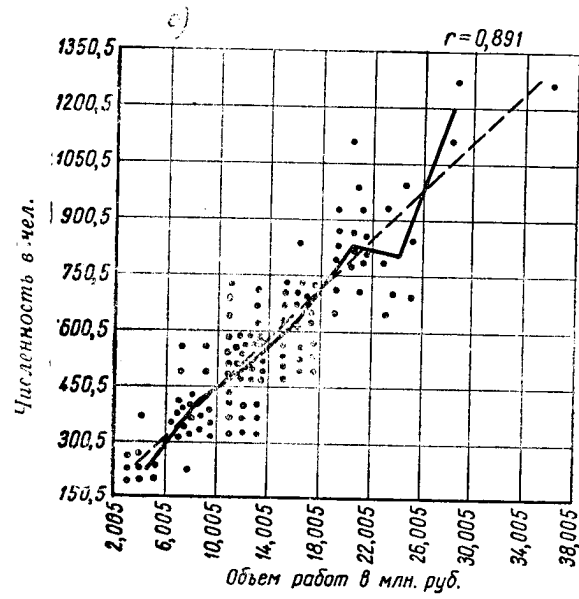


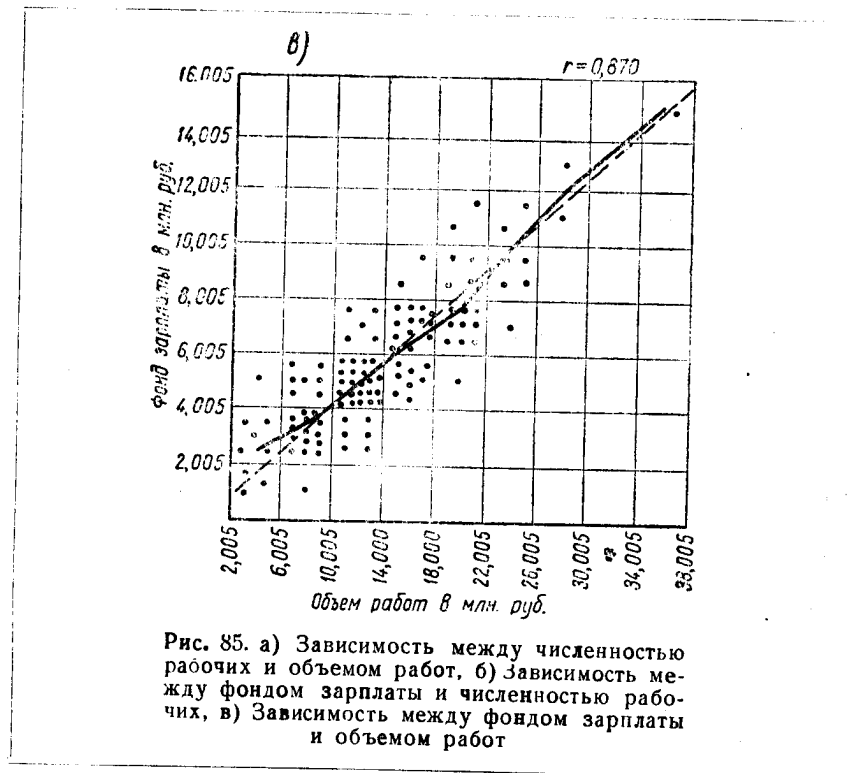
Рис. 84. Зависимость между величиной накладных расходов и фондом зарплаты

Обнаружены следующие пропорции в изменении интересующих нас показателей:

1) С увеличением объема работ на 1 млн. руб. средние накладные расходы стройки $\bar{x}_{1.2}$ увеличиваются на 0,153 млн.; в то же время существует постоянная часть средних накладных расходов $\bar{x}_{1.2}$, не изменяющаяся с изменением объема работ, которая выражается числом 1,241 млн. руб.

2) С увеличением численности строительного-монтажных рабочих на одного рабочего средние накладные расходы $\bar{x}_{1.3}$ увеличиваются на 4290 руб. Постоянная часть средних накладных расходов $\bar{x}_{1.3}$, не изменяющаяся с изменением числа рабочих, составляет 0,817 млн. руб.





3) С увеличением фонда зарплаты на 1 млн. руб. средние расходы \bar{x}_1 увеличиваются на 375 тыс. руб. Постоянная относительно фонда зарплаты часть средних накладных расходов составляет 1,152 млн. руб.

Однако в этих пропорциях отражается влияние не только тех факторов, которые включены в уравнения, но и тех, которые им неявно сопутствуют. Так, например, увеличение объема работ обычно сопровождается увеличением числа рабочих и фонда зарплаты, что отражается на накладных расходах; увеличение числа рабочих сопровождается ростом фонда зарплаты и объема работ, что также оказывает влияние на накладные расходы.

Для того чтобы выявить влияние каждого фактора в «чистом» виде, необходимо отвлечься от сопутствующего влияния других факторов; это достигается методами множественной корреляции.

Переходим к соотношениям между факторами, которые важно установить для получения в дальнейшем множественных связей. На рис. 85 (а, б, в) представлены поля корреляции и линии регрессии между каждой парой факторов. Прямолинейный харак-

тер всех рассматриваемых зависимостей очевиден. Приводим соответствующие уравнения регрессии:

$$\bar{x}_{3.2} = 121,9 + 32,97x_2, \quad (16.2.4)$$

$$\bar{x}_{4.3} = -0,122 + 0,0101x_3, \quad (16.2.5)$$

$$\bar{x}_{4.2} = 0,628 + 0,368x_2. \quad (16.2.6)$$

Коэффициенты корреляции между каждой парой факторов оказались:

$$r_{23} = 0,891, \quad r_{24} = 0,870, \quad r_{34} = 0,887.$$

Эти коэффициенты свидетельствуют о весьма тесной зависимости, существующей между факторами.

Приводим также таблицу, содержащую значения средних арифметических и средних квадратических отклонений для всех рассматриваемых признаков.

Таблица 79

Показатели	\bar{x}_i	σ_i
Факторы		
1. Накладные расходы	3,288	1,14
2. Объем выполненных работ	13,77	6,00
3. Численность рабочих . . .	575,9	222,0
4. Фонд зарплаты	5,695	2,536

§ 3. В связи с тем, что все парные зависимости между признаками оказались прямолинейными, мы вправе предположить, что и зависимости множественные окажутся также прямолинейными. При составлении уравнений множественной регрессии использован сигмальный (стандартизованный) масштаб представления переменных.

Приводим уравнения множественной регрессии, связывающие величину накладных расходов с двумя факторами одновременно, и окончательное уравнение, дающее зависимость накладных рас-

ходов от всех трех факторов. Все уравнения выражены в стандартизованном масштабе:

$$\bar{t}_{1.23} = 0,277t_2 + 0,589t_3, \quad (16.3.1)$$

$$\bar{t}_{1.24} = 0,320t_2 + 0,554t_4, \quad (16.3.2)$$

$$\bar{t}_{1.34} = 0,456t_3 + 0,429t_4, \quad (16.3.3)$$

$$\bar{t}_{1.234} = 0,132t_2 + 0,384t_3 + 0,377t_4. \quad (16.3.4)$$

Соответствующие коэффициенты множественной корреляции:

$$R_{1.23} = 0,845 \text{ с коррекцией на систематическую погрешность } \hat{R}_{1.23} = 0,841$$

$$R_{1.24} = 0,847 \quad . \quad . \quad . \quad . \quad \hat{R}_{1.24} = 0,844$$

$$R_{1.34} = 0,859 \quad . \quad . \quad . \quad . \quad \hat{R}_{1.34} = 0,856$$

$$R_{1.234} = 0,867 \quad . \quad . \quad . \quad . \quad \hat{R}_{1.234} = 0,856$$

Из приведенных данных можно сделать следующие выводы. На уровень накладных расходов в наибольшей и в примерно равной степени влияют факторы численности рабочих и фонда зарплаты; фактор объема капитальных работ при учете первых двух факторов имеет сравнительно небольшое влияние на величину накладных расходов; учет этого фактора не повышает множественный коэффициент корреляции между накладными расходами и совокупностью факторов $\hat{R}_{1.234} = \hat{R}_{1.34}$. Следовательно, при определении величины накладных расходов можно было бы ограничиться учетом численности рабочей силы и фонда зарплаты. Это явление объясняется, по-видимому, тем, что экономический объем работ влияет на накладные расходы в основном не непосредственно, а через число рабочих и фонд зарплаты, как это можно проследить из подробного рассмотрения конкретных статей накладных расходов. Сила же влияния факторов, действующих на функцию непосредственно, должна быть, очевидно, больше, чем сила влияния факторов, более удаленных и действующих на функцию через посредство первых: очевидно, также, что при учете непосредственно действующих на функцию факторов учет более удаленных факторов не повышает точность оценки.

Преобразуя переменные, представим уравнения регрессии в натуральном масштабе:

$$\bar{x}_{1.23} = 0,819 + 0,0526x_2 + 0,00303x_3, \quad (16.3.5)$$

$$\bar{x}_{1.24} = 1,033 + 0,0508x_2 + 0,249x_4, \quad (16.3.6)$$

$$\bar{x}_{1.34} = 0,841 + 0,00234x_3 + 0,193x_4, \quad (16.3.7)$$

$$\bar{x}_{1.234} = 0,841 + 0,0250x_2 + 0,00197x_3 + 0,170x_4. \quad (16.3.8)$$

Согласно уравнению (16.3.8) при постоянном числе рабочих и постоянном фонде зарплаты увеличение объема работ на 1 млн. руб. влечет за собой увеличение накладных расходов в среднем на 25 тыс. руб. Если сопоставить эту величину с величиной аналогичного показателя в 153 тыс. руб. из уравнения парной регрессии (16.2.1), то станет очевидным, что столь резкое сокращение нормы накладных расходов на 1 млн. руб. объема работ вызвано закреплением числа рабочих и фонда зарплаты, т. е. основных факторов, через которые увеличение объема работ влияет на величину накладных расходов.

Из уравнения (16.3.7) видно, что с увеличением числа рабочих на 1 человека при постоянном фонде зарплаты накладные расходы увеличиваются в среднем на 2340 руб., а с увеличением фонда зарплаты на 1 млн. при постоянном количестве рабочих накладные расходы увеличиваются в среднем на 193 руб. Эти величины значительно меньше аналогичных показателей из уравнений парной регрессии (16.2.2), (16.2.3) — 4290 руб. и 375 тыс. руб. и объясняются тем, что в уравнениях парной регрессии изменение числа рабочих сопровождается неявным для уравнения изменением, оно связано с числом рабочих, фондом зарплаты, а изменение фонда зарплаты, учитываемое в другом уравнении, сопровождается неявным изменением связанного с ним числа рабочих. По этой причине в уравнениях парной регрессии изменения накладных расходов, вызываемые каждым из двух факторов, значительно больше, чем в уравнениях множественной регрессии.

Закрепляя в уравнении (16.3.8) дополнительно объем капитальных работ, мы, естественно, сокращаем коэффициенты при факторах числа рабочих и фонда зарплаты. При закрепленном объеме работ и фонде зарплаты увеличение числа рабочих на 1 человека увеличивает накладные расходы в среднем на 1970 руб., а увеличение фонда зарплаты на 1 млн. руб. при закрепленном объеме продукции и числе рабочих увеличивает накладные расходы в среднем на 170 тыс. руб.

Заметим также, что каждое уравнение регрессии определяет постоянную часть средних накладных расходов, не зависящую от изменения факторов, включаемых в уравнение. Так, по уравнению (16.3.8) постоянная часть средних накладных расходов, не изменяющаяся с увеличением объема работ, числа рабочих и фонда зарплаты, составляет 841 тыс. руб.

§ 4. Уравнение (16.3.8) может быть использовано для ориентировочной оценки величины накладных расходов по данному объему капитальных работ, числу рабочих и фонду зарплаты. Подставляя в уравнение на место переменных x_2 , x_3 , x_4 интересующие нас конкретные значения, получаем оценку накладных расходов. Так, например, для стройки с показателями

$x_2 = 10,0$, $x_3 = 497$, $x_4 = 4,324$ по уравнению (16.3.8) они составляют

$$\bar{x}_{1.234} = 0,841 + 0,025 \cdot 10,0 + 0,00197 \cdot 497 + 0,170 \cdot 4,324 = 2,805$$

Фактическая величина накладных расходов по данной стройке была 2,5 млн. руб. Для другой стройки с показателями

$$x_2 = 26,2, \quad x_3 = 1100, \quad x_4 = 11,000,$$

$$\bar{x}_{1.234} = 0,841 + 0,025 \cdot 26,2 + 0,00197 \cdot 1100 + 0,170 \cdot 11,0 = 5,533.$$

Фактическая величина была 5,9.

Подобные оценки были рассчитаны по всем 113 годовым отчетам строек. Результаты сопоставления этих оценок с фактическими накладными расходами даны на рис. 86 в виде поля корреляции. Через z и z' обозначены соответственно фактические накладные расходы и их оценка по уравнению.

Исследуя закономерности этого поля, обнаруживаем, что линия регрессии z по z' очень близка к биссектрисе координатного угла и выражается уравнением

$$\bar{z}' = 0,230 + 0,948z. \quad (16.4.1)$$

Следовательно, оценка z' воспроизводит фактические значения z без систематической погрешности. Точки корреляционного поля кучно группируются вокруг линии регрессии при коэффициенте корреляции $r_{zz'} = 0,870$, отклонения фактических накладных расходов от их оценки по уравнению в большинстве случаев невелики, не имеют систематической направленности и их величина измеряется средней квадратической ошибкой оценки

$$\sigma_{zz'} = \sigma_z \sqrt{1 - r^2} = 1,10 \sqrt{1 - 0,870^2} = 0,542.$$

Заметим, что теоретически должны иметь место следующие соотношения, на основе которых линия регрессии z по z' , коэффициент корреляции $r_{zz'}$ и средняя квадратическая погрешность оценки $\sigma_{zz'}$ могут быть рассчитаны заранее без фактического подсчета показателей z' и построения поля корреляции

$$\bar{z}' = z', \quad r_{zz'} = R_{1.234},$$

$$\sigma_{zz'} = \sigma_z \sqrt{1 - R_{1.234}^2}$$

Некоторые неточности в этих формулах могут возникнуть по причине группировки данных¹.

Показатель z' может оказать серьезную помощь в анализе хозяйственной деятельности строек. Изучение причин, вызываю-

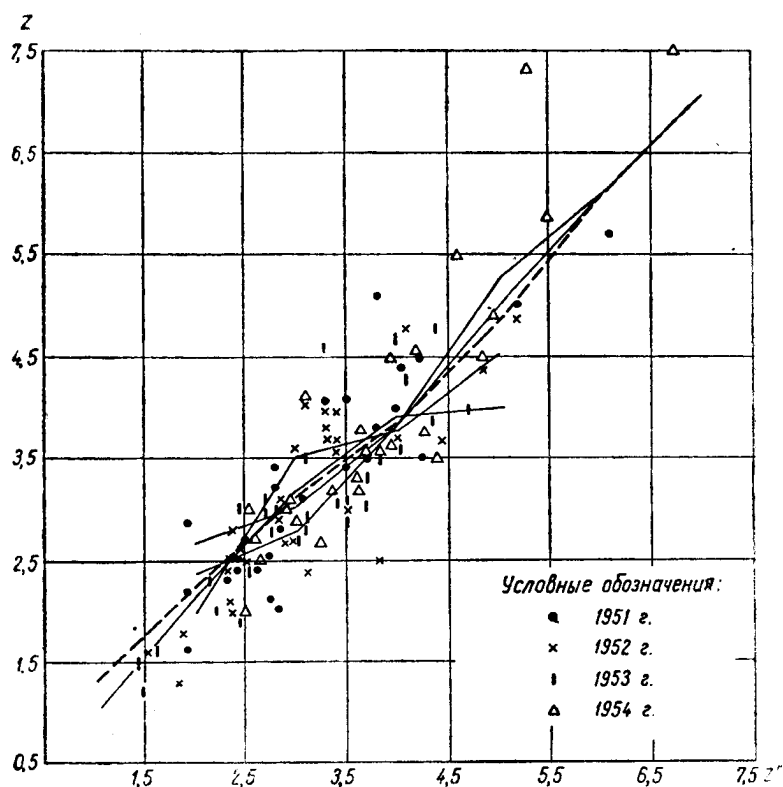


Рис. 86. Зависимость между величиной фактических накладных расходов и рассчитанных по уравнению

щих большие отклонения z от z' в отдельных стройках, поможет вскрыть резервы для сокращения накладных расходов.

Остановимся на вопросе о том, в какой мере данные различных лет соответствуют полученной закономерности. На корреляционном поле (рис. 86) точки, относящиеся к различным годам, отмечены разными значками. В пределах каждого года были рассчитаны линии регрессии, которые нанесены на рисунке.

¹ Следует применять поправки Шеппарда на группировку (см. стр. 346). Они учтены в уравнении (16.4.1).

Мы видим, что все эти линии несущественно отличаются между собой и практически совпадают с общей линией регрессии. Следовательно, можно считать, что данные различных лет имеют в основе одну и ту же линию регрессии, и закономерность, выражаемая уравнением (16.3.8), сохраняется на протяжении рассматриваемых четырех лет. Это обстоятельство оправдывает объединение материала различных лет при его статистической обработке.

§ 5. Из исследованных соотношений между объемными показателями хозяйственной деятельности строек нетрудно получить зависимости и между важнейшими относительными показателями.

Так, деля обе части уравнения (16.3.8) на x_2 , можно привести уравнение к виду

$$\left(\frac{x_1}{x_2}\right)_{234} = \frac{0,841}{x_2} + 0,0250 + 0,00197 \frac{x_3}{x_2} + 0,170 \frac{\frac{x_4}{x_2}}{\frac{x_3}{x_2}}. \quad (16.5.1)$$

Далее обозначим через u_1 так называемые «удельные» накладные расходы, т. е. отношение накладных расходов к объему капитальных работ

$$u_1 = \frac{x_1}{x_2};$$

через u_2 — среднюю выработку на одного списочного рабочего, т. е. отношение объема работ к числу рабочих

$$u_2 = \frac{x_2}{x_3};$$

через u_3 — среднюю зарплату, т. е. отношение фонда зарплат к числу рабочих

$$u_3 = \frac{x_4}{x_3}.$$

Тогда уравнение (16.5.1), записанное в виде

$$\bar{u}_{1, x_1, u_2, u_3} = 0,0250 + \frac{0,841}{x_2} + \frac{0,00197}{u_2} + 0,170 \frac{u_3}{u_2}. \quad (16.5.2)$$

будет давать соотношение между удельными накладными расходами, объемом работ, средней выработкой и средней зарплатой.

Таким образом, средний уровень удельных накладных расходов складывается из четырех частей: постоянной части, части, изменяющейся обратно пропорционально объему работ, части

изменяющейся обратно пропорционально производительности труда рабочего, и, наконец, части, пропорциональной отношению зарплаты к производительности труда.

Можно задаться целью получить соотношение между показателями u_1 , x_2 , u_2 , u_3 непосредственно, приняв эти показатели в качестве основных; показатели x_1 , x_3 , x_4 играли бы в этом случае роль производных. Однако, как легко видеть, техника получения множественных взаимосвязей для первой группы переменных была бы много сложнее, так как эти переменные образуют между собой криволинейные зависимости гиперболического типа. В частности, было бы затруднительно обнаружить в этой системе переменных форму множественной зависимости (16.5.1) с четвертым членом вида $\frac{u_3}{u_2}$.

Выбор системы переменных x_1 , x_2 , x_3 , x_4 , образующих между собой прямолинейные зависимости, облегчает нам нахождение множественных зависимостей. Он позволяет применить рассмотренный ранее прием сведения криволинейных связей к прямолинейным при исследовании корреляции многих переменных.

§ 6. Полученные уравнения могут оказать значительную помощь при планировании накладных расходов. Разумеется, соотношения, установленные на материалах прошедшего периода, нельзя безоговорочно распространять на будущее время. В них должны вноситься коррективы, связанные с новыми условиями и задачами планируемого периода.

Необходимо иметь в виду также и то обстоятельство, что соотношения, установленные на материале 113 единиц годовых отчетов, обладают заметными случайными погрешностями, размеры которых, рассчитанные по формулам § 4 главы XIII, приводятся ниже (табл. 80).

Анализ этой таблицы позволяет внести следующие существенные уточнения в выводы предыдущих параграфов. Все коэффициенты регрессии обладают значительными средними квадратическими ошибками. Относительные размеры этих ошибок увеличиваются с увеличением числа переменных в уравнении регрессии и достигают наибольшей величины для коэффициентов последнего уравнения (16.3.8), как показывает таблица относительных ошибок коэффициентов регрессии (табл. 81).

Отсюда видно, что выводы, сделанные в предыдущих параграфах из уравнения (16.3.8), не могут претендовать на высокую точность в определении числовых коэффициентов. В особенности это относится к коэффициенту регрессии при объеме капитальных работ, величину которого следует считать несущественной. Уточнить эти коэффициенты можно было бы путем увеличения числа годовых отчетов, взятых для обработки. Автор такими сведениями не располагал.

Таблица 80
Средние квадратические ошибки коэффициентов
регрессии и корреляции

Уравнение	Коэффициент	Средняя квадратическая ошибка коэффициента
(16.2.1)	$b = 0,153$	$\sigma_b = 0,011$
(16.2.1)	$r_{12} = 0,802$	$\sigma_{r_{12}} = 0,034$
(16.2.2)	$c = 0,00129$	$\sigma_c = 0,00026$
(16.2.2)	$r_{13} = 0,836$	$\sigma_{r_{13}} = 0,028$
(16.2.3)	$d = 0,375$	$\sigma_d = 0,023$
(16.2.3)	$r_{14} = 0,833$	$\sigma_{r_{14}} = 0,029$
(16.3.5)	$b = 0,0536$	$\sigma_b = 0,0211$
(16.3.5)	$c = 0,00303$	$\sigma_c = 0,00057$
(16.3.5)	$R_{1.23} = 0,845$	$\sigma_{R_{1.23}} = 0,027$
(16.3.6)	$b = 0,0608$	$\sigma_b = 0,0192$
(16.3.6)	$d = 0,241$	$\sigma_d = 0,045$
(16.3.6)	$R_{1.24} = 0,847$	$\sigma_{R_{1.24}} = 0,027$
(16.3.7)	$c = 0,00234$	$\sigma_c = 0,00054$
(16.3.7)	$d = 0,193$	$\sigma_d = 0,047$
(16.3.7)	$R_{1.34} = 0,859$	$\sigma_{R_{1.34}} = 0,025$
(16.3.8)	$b = 0,0250$	$\sigma_b = 0,0211$
(16.3.8)	$c = 0,00197$	$\sigma_c = 0,00061$
(16.3.8)	$d = 0,170$	$\sigma_d = 0,049$
(16.3.8)	$R_{1.234} = 0,867$	$\sigma_{R_{1.234}} = 0,023$

Таблица 81

Уравнение	Относительные ошибки коэффициентов (в процентах)		
	$V_b = 100 \frac{\sigma_b}{b}$	$V_c = 100 \frac{\sigma_c}{c}$	$V_d = 100 \frac{\sigma_d}{d}$
С двумя переменными	(16.2.1)	7,2	
	(16.2.2)		6,2
	(16.2.3)		8,0
С тремя переменными	(16.3.5)	39,4	18,8
	(16.3.6)	31,6	
	(16.3.7)		23,1
С четырьмя переменными	(16.3.8)	84,4	31,0

Приложение 1
ТАБЛИЦЫ ОПЫТНЫХ ДАННЫХ

Таблица 1

№ п/п	Номер печи	Сечение штанги в мм	Ударная вязкость a_k кгм/см ² (два испы- тания)	Содержание химических элементов (в процентах)								Среднее зна- чение удар- ной вязкости a_{kcp}
				C	P	Mn	S	Si	Cr	Ni	Cu	
1	2	3	4	5	6	7	8	9	10	11	12	13
1	4	22	6,24— 7,05	0,31	0,026	0,99	0,022	0,98	0,98	0,14	0,19	6,645
2	2	65	6,86— 6,42	0,33	0,022	0,96	0,022	1,06	0,93	0,12	0,24	6,64
3	1	60	3,91— 4,50	0,33	0,029	0,91	0,021	1,02	0,99	0,14	0,21	4,205
4	2	70	9,72 - 9,07	0,30	0,020	0,86	0,021	1,14	0,95	0,15	0,17	9,395
5	3	20	9,07— 7,64	0,34	0,017	0,98	0,018	1,02	0,98	0,15	0,22	8,37
6	1	50	5,00— 5,45	0,31	0,020	0,97	0,019	1,07	0,88	0,13	0,25	5,225
7	1	70	5,65— 5,65	0,30	0,024	1,00	0,022	1,12	0,94	0,15	0,19	5,65
8	2	45	7,94— 8,62	0,30	0,020	0,92	0,025	0,98	0,95	0,15	0,17	8,28
9	4	70	5,85— 6,50	0,31	0,020	0,93	0,025	0,94	0,80	0,18	0,21	6,175
10	1	60	6,50— 5,90	0,32	0,026	0,89	0,022	0,93	0,88	0,15	0,21	6,20
11	1	60	9,27—11,20	0,27	0,028	0,74	0,024	0,93	0,80	0,15	0,20	10,235
12	1	60	6,00— 6,05	0,31	0,026	0,92	0,025	0,93	1,03	0,15	0,22	6,025
13	1	65	6,72— 8,26	0,32	0,022	0,89	0,023	1,03	0,87	0,15	0,20	7,49
14	1	18	7,56— 6,85	0,30	0,027	1,02	0,024	1,03	0,96	0,13	0,22	7,205
15	3	70	5,46— 5,68	0,32	0,023	1,06	0,025	1,10	0,96	0,16	0,20	5,57
16	3	70	6,00— 5,94	0,32	0,018	1,07	0,023	1,09	0,90	0,18	0,18	5,97
17	3	70	5,98— 5,88	0,33	0,023	1,07	0,023	1,02	1,00	0,15	0,18	5,93
18	3	60	5,35— 5,05	0,33	0,022	1,06	0,020	1,03	1,05	0,14	0,16	5,20
19	2	50	4,35— 5,06	0,32	0,027	1,06	0,024	1,08	0,92	0,15	0,18	4,705
20	1	50	6,36— 7,24	0,32	0,020	0,93	0,022	1,07	0,89	0,16	0,15	6,80
21	1	65	8,32— 8,05	0,30	0,021	0,89	0,023	0,98	0,91	0,15	0,18	8,185
22	1	50	6,85— 6,36	0,28	0,021	0,97	0,021	1,12	0,93	0,17	0,16	6,605
23	3	65	3,17— 3,17	0,34	0,020	1,00	0,022	1,18	0,95	0,17	0,21	3,17
24	1	65	8,06— 7,38	0,30	0,023	0,85	0,025	1,10	0,95	0,16	0,21	7,72
25	1	60	6,58— 6,00	0,31	0,023	0,97	0,024	1,16	0,96	0,17	0,20	6,29
26	1	50	7,55— 7,07	0,31	0,028	0,94	0,022	1,16	0,95	0,17	0,14	7,31
27	1	22	7,28— 7,09	0,31	0,027	0,98	0,018	1,12	0,97	0,19	0,20	7,185
28	2	50	5,85— 5,85	0,31	0,021	0,98	0,018	1,16	0,99	0,16	0,19	5,85
29	2	65	5,46— 5,30	0,31	0,025	0,96	0,022	1,18	0,95	0,16	0,19	5,38
30	1	17	7,36— 7,62	0,31	0,023	0,94	0,024	1,01	0,94	0,16	0,17	7,49
31	1	22	9,40— 9,40	0,32	0,021	0,97	0,020	1,08	0,94	0,16	0,15	9,40
32	1	18	6,28— 6,55	0,31	0,028	0,84	0,022	1,10	0,93	0,19	0,20	6,415
33	1	65	3,16— 2,98	0,33	0,023	0,95	0,019	1,00	0,99	0,15	0,18	3,07
34	1	24	7,08— 7,57	0,31	0,028	0,88	0,018	1,12	1,00	0,16	0,20	7,325
35	1	22	8,75— 8,60	0,30	0,021	0,92	0,026	1,06	0,85	0,16	0,20	8,675
36	1	65	6,95— 6,20	0,30	0,027	0,93	0,020	0,98	0,94	0,15	0,21	6,575
37	1	60	6,72— 5,50	0,32	0,020	0,86	0,022	1,10	0,92	0,15	0,18	6,11
38	1	60	5,95— 5,70	0,31	0,022	0,96	0,018	1,08	0,91	0,17	0,18	5,825
39	1	65	6,67— 6,85	0,32	0,025	1,02	0,025	1,04	0,98	0,16	0,21	6,76
40	1	24	9,50— 8,01	0,30	0,018	0,87	0,022	1,12	0,86	0,16	0,17	8,77
41	1	65	7,27— 6,70	0,31	0,026	0,82	0,020	1,02	0,83	0,18	0,21	6,985
42	1	26	6,28— 6,01	0,28	0,020	0,92	0,021	1,10	0,93	0,18	0,22	6,145

Продолжение

№ п/п	Номер печи	Сечение штанги в мм	Ударная вязкость a_k кгм/см ² (два испытания)	Содержание химических элементов (в процентах)								Среднее значение ударной вязкости a_k ср
				C	P	Mn	S	Si	Cr	Ni	Cu	
1	2	3	4	5	6	7	8	9	10	11	12	13
43	1	22	8,32— 8,32	0,32	0,018	0,91	0,020	1,07	0,91	0,17	0,19	8,32
44	1	50	8,84— 8,58	0,30	0,022	0,90	0,023	1,02	0,80	0,18	0,23	8,71
45	1	65	6,35— 7,80	0,32	0,020	0,92	0,024	1,11	0,92	0,16	0,19	7,075
46	1	65	10,70— 9,63	0,29	0,021	0,90	0,022	1,10	0,85	0,18	0,25	10,165
47	3	65	7,65— 7,37	0,29	0,018	0,75	0,022	0,94	0,86	0,17	0,21	6,51
48	3	48	5,32— 5,22	0,32	0,025	0,93	0,026	1,02	0,96	0,19	0,19	5,27
49	1	45	7,50— 7,40	0,30	0,018	0,95	0,026	1,06	0,95	0,16	0,22	7,45
50	1	70	8,50— 8,65	0,29	0,019	0,84	0,027	1,09	0,91	0,15	0,22	8,575
51	4	65	7,07— 7,07	0,29	0,021	0,90	0,021	1,02	0,93	0,16	0,20	7,07
52	3	60	6,35— 6,00	0,30	0,028	0,93	0,025	1,10	0,92	0,15	0,22	6,175
53	1	70	3,74— 3,94	0,31	0,028	0,89	0,028	0,98	1,00	0,17	0,20	3,84
54	1	50	7,08— 7,37	0,30	0,033	0,90	0,020	1,00	0,88	0,18	0,16	7,225
55	2	65	6,85— 6,05	0,32	0,020	0,88	0,027	1,12	0,95	0,16	0,23	6,45
56	1	65	7,33— 7,72	0,32	0,022	0,89	0,026	1,12	0,92	0,16	0,18	7,525
57	1	65	8,13— 8,13	0,30	0,023	0,85	0,023	1,03	0,83	0,15	0,20	8,13
58	1	75	6,63— 7,20	0,32	0,023	0,92	0,018	1,00	0,94	0,18	0,17	6,915
59	1	70	5,00— 4,56	0,32	0,023	0,93	0,022	0,98	1,00	0,16	0,20	4,78
60	1	70	5,21— 5,44	0,29	0,019	0,95	0,025	1,12	0,89	0,16	0,17	5,325
61	1	65	6,25— 6,75	0,32	0,019	0,85	0,025	1,08	1,01	0,17	0,25	6,50
62	1	50	7,02— 7,40	0,31	0,023	0,94	0,020	1,10	0,99	0,17	0,17	7,21
63	3	70	7,25— 7,00	0,33	0,022	0,89	0,025	1,11	0,96	0,16	0,19	7,125
64	3	60	5,66— 5,43	0,35	0,023	0,94	0,026	1,07	0,95	0,17	0,18	5,545
65	3	70	5,20— 5,46	0,32	0,023	0,90	0,025	1,02	0,94	0,18	0,30	5,33
66	4	65	6,97— 6,97	0,32	0,020	0,96	0,025	1,07	1,04	0,15	0,25	6,97
67	2	50	6,58— 5,43	0,30	0,023	0,81	0,027	1,08	0,87	0,17	0,25	6,005
68	1	38	6,14— 5,10	0,32	0,026	1,00	0,022	1,02	0,92	0,18	0,21	5,62
69	2	25	5,83— 7,56	0,33	0,018	0,89	0,026	1,09	0,96	0,19	0,22	6,695
70	1	65	6,38— 5,18	0,31	0,019	0,89	0,022	1,02	0,94	0,21	0,20	5,78
71	1	65	6,04— 7,18	0,32	0,023	0,92	0,023	0,98	0,98	0,18	0,23	6,61
72	1	65	6,60— 6,88	0,31	0,023	0,86	0,025	1,06	0,86	0,18	0,20	6,74
73	4	60	4,18— 4,58	0,31	0,034	0,98	0,023	1,00	0,99	0,16	0,20	4,38
74	1	25	8,06— 8,58	0,29	0,030	0,86	0,024	1,08	0,87	0,20	0,17	8,32
75	1	25	5,28— 5,58	0,34	0,024	1,03	0,020	1,14	0,95	0,17	0,18	5,43
76	4	21	8,37— 8,37	0,31	0,021	0,99	0,027	1,12	0,86	0,17	0,17	8,37
77	2	65	6,84— 3,90	0,32	0,022	1,10	0,025	1,08	1,03	0,17	0,21	5,37
78	1	22	5,66— 6,35	0,32	0,022	1,05	0,018	1,07	0,90	0,18	0,22	6,005
79	1	32	8,05— 7,31	0,31	0,023	0,96	0,018	1,07	0,93	0,17	0,23	7,68
80	1	38	6,59— 6,59	0,30	0,020	0,84	0,022	0,98	0,86	0,17	0,20	6,59
81	1	50	9,80— 10,80	0,29	0,016	0,81	0,020	1,06	0,86	0,17	0,20	10,30
82	1	50	8,06— 8,58	0,31	0,018	0,85	0,019	1,12	0,85	0,17	0,21	8,32
83	2	65	6,48— 6,39	0,35	0,017	0,80	0,020	0,99	0,93	0,17	0,16	6,435
84	4	65	6,84— 6,96	0,30	0,017	0,87	0,018	1,02	0,92	0,17	0,16	6,90
85	4	19	7,15— 6,80	0,32	0,024	0,91	0,018	1,06	1,06	0,17	0,16	6,975
86	4	65	5,45— 5,45	0,32	0,023	0,92	0,022	1,07	0,94	0,19	0,18	5,45
87	2	65	5,46— 6,36	0,33	0,018	0,96	0,023	1,11	0,89	0,16	0,21	5,91

Продолжение

№ п.п.	Номер печи		Ударная вязкость a_k кгм/см ² (два испытания)	Содержание химических элементов (в процентах)									Среднее значение ударной вязкости a_k ср.
	2	3		С	Р	Мп	S	Si	Cr	Ni	Cu		
1	2	3	4	5	6	7	8	9	10	11	12	13	
88	4	65	6,20—6,90	0,35	0,024	0,99	0,023	1,07	0,92	0,17	0,21	6,55	
89	4	27	5,90—6,68	0,32	0,022	0,97	0,020	1,00	0,85	0,16	0,21	6,29	
90	4	50	7,34—8,35	0,31	0,018	0,87	0,022	1,11	0,92	0,17	0,20	7,845	
91	4	50	7,40—7,97	0,28	0,035	0,95	0,023	1,07	0,96	0,16	0,24	7,685	
92	4	65	6,50—6,35	0,31	0,028	0,93	0,025	1,10	0,97	0,16	0,22	6,425	
93	2	22	7,25—6,60	0,27	0,023	0,95	0,022	1,06	0,91	0,17	0,20	6,925	
94	1	70	5,28—5,89	0,30	0,022	0,91	0,020	1,02	0,90	0,17	0,28	5,585	
95	1	65	7,08—6,60	0,28	0,025	0,94	0,020	1,04	0,94	0,17	0,23	6,84	
96	2	65	3,81—3,81	0,37	0,021	0,99	0,020	1,10	1,00	0,17	0,22	3,81	
97	2	70	5,90—7,80	0,35	0,021	1,03	0,022	1,07	0,95	0,16	0,23	6,85	
98	4	70	6,13—6,35	0,33	0,021	0,98	0,020	1,07	0,83	0,21	0,19	6,24	
99	4	60	4,71—4,75	0,34	0,021	0,97	0,023	1,11	0,80	0,16	0,19	4,73	
100	4	18	7,08—7,08	0,34	0,02	0,89	0,022	1,17	0,89	0,17	0,23	7,08	
101	2	50	5,60—6,13	0,35	0,023	1,06	0,023	1,00	0,99	0,22	0,22	5,865	
102	2	24	6,21—6,35	0,32	0,021	1,02	0,025	1,12	1,00	0,18	0,17	6,28	
103	3	65	4,78—5,02	0,33	0,022	1,01	0,022	1,12	1,00	0,18	0,21	4,90	
104	3	40	6,58—6,14	0,31	0,027	0,97	0,022	1,12	0,91	0,18	0,20	6,36	
105	3	16	7,48—7,86	0,33	0,021	0,95	0,022	1,12	0,85	0,19	0,25	7,67	
106	3	38	5,85—6,05	0,33	0,021	0,95	0,020	1,10	0,88	0,16	0,19	5,95	
107	1	50	5,97—6,21	0,34	0,025	0,93	0,020	1,10	0,97	0,17	0,20	6,00	
108	1	70	5,27—5,27	0,35	0,017	0,98	0,025	1,18	0,92	0,18	0,19	5,27	
109	1	70	7,45—5,00	0,33	0,022	0,99	0,027	1,12	0,94	0,25	0,22	6,225	
110	3	50	5,66—5,60	0,35	0,022	1,03	0,022	1,12	0,93	0,17	0,22	5,63	
111	4	60	5,22—5,50	0,35	0,024	0,88	0,020	1,27	0,88	0,19	0,23	5,36	
112	3	38	5,15—5,10	0,33	0,023	0,95	0,018	0,96	0,97	0,19	0,23	5,125	
113	4	33	6,52—6,82	0,35	0,019	0,91	0,023	1,12	0,97	0,23	0,24	6,67	
114	4	38	5,54—6,02	0,34	0,023	0,97	0,020	1,08	0,98	0,15	0,23	5,78	
115	3	17	6,92—7,58	0,33	0,017	0,87	0,019	1,09	0,88	0,16	0,23	7,25	
116	1	65	6,14—6,60	0,34	0,019	0,87	0,020	1,06	0,91	0,19	0,24	6,37	
117	3	52	4,78—4,35	0,32	0,021	0,87	0,024	1,06	0,81	0,18	0,23	4,565	
118	1	70	5,17—5,57	0,33	0,021	0,97	0,022	1,07	1,01	0,14	0,19	5,37	
119	1	65	5,90—6,36	0,32	0,025	0,96	0,023	1,12	0,90	0,17	0,21	6,13	
120	4	65	7,75—9,52	0,30	0,023	0,85	0,026	1,00	0,90	0,17	0,21	8,635	
121	1	65	5,12—5,36	0,30	0,025	0,88	0,026	1,12	1,00	0,20	0,21	5,24	
122	2	20	5,22—5,07	0,32	0,021	0,95	0,023	1,12	1,00	0,13	0,25	5,145	
123	2	50	5,90—5,84	0,32	0,020	0,96	0,020	1,12	0,97	0,11	0,23	5,87	
124	2	70	6,18—5,90	0,30	0,020	0,94	0,024	1,02	0,93	0,19	0,23	6,04	
125	2	75	5,20—6,35	0,32	0,023	0,98	0,019	1,02	1,01	0,14	0,25	5,775	
126	2	24	2,44—2,26	0,32	0,033	1,00	0,026	1,20	1,08	0,20	0,25	2,35	
127	2	50	5,44—5,36	0,32	0,024	0,95	0,023	1,10	0,99	0,16	0,22	5,40	
128	2	70	3,16—3,02	0,33	0,026	0,96	0,024	1,16	1,04	0,19	0,25	3,09	
129	1	70	5,87—5,87	0,35	0,019	0,91	0,024	1,06	1,04	0,20	0,24	5,87	
130	1	65	5,60—5,94	0,32	0,019	0,85	0,022	1,14	0,99	0,16	0,23	5,77	
131	4	28	7,07—7,32	0,31	0,021	0,88	0,022	1,02	0,90	0,20	0,23	7,195	
132	1	48	9,10—9,35	0,28	0,025	0,84	0,020	1,08	0,91	0,19	0,25	9,225	

Продолжение

№ п/п	Номер печи	Сечение штанги в мм	Ударная вязкость a_k кгм/см ² (два испы- тания)	Содержание химических элементов (в процентах)								Среднее зна- чение удар- ной вязкости a_k ср
				C	P	Mn	S	Si	Cr	Ni	Cu	
1	2	3	4	5	6	7	8	9	10	11	12	13
133	3	28	7,55-7,10	0,33	0,025	0,85	0,025	1,13	0,96	0,19	0,23	7,325
134	4	65	4,15-2,80	0,33	0,018	0,92	0,022	1,15	0,80	0,18	0,24	3,475
135	2	38	6,70-6,70	0,32	0,022	0,93	0,023	1,06	0,88	0,16	0,26	6,70
136	4	28	7,60-7,57	0,31	0,024	0,90	0,021	1,00	0,84	0,16	0,25	7,585
137	1	45	7,60-8,08	0,29	0,022	0,97	0,026	1,08	0,95	0,18	0,25	7,84
138	3	65	6,14-6,14	0,33	0,027	0,89	0,025	1,00	0,96	0,16	0,25	6,14
139	2	60	7,88-7,88	0,29	0,020	0,84	0,023	1,12	0,80	0,23	0,25	7,88
140	4	20	5,73-5,00	0,29	0,030	0,87	0,022	1,08	0,84	0,16	0,24	5,365
141	4	90	4,80-3,74	0,35	0,022	0,92	0,022	1,15	1,09	0,15	0,23	4,27
142	1	28	5,31-4,94	0,31	0,023	0,88	0,024	0,96	1,09	0,28	0,25	5,12
143	1	90	5,88-5,88	0,30	0,021	0,88	0,025	1,00	0,94	0,25	0,25	5,88
144	4	90	7,10-7,10	0,31	0,020	0,89	0,025	1,09	0,90	0,15	0,23	7,10
145	4	70	4,78-4,15	0,34	0,025	0,90	0,020	1,08	0,90	0,13	0,25	4,465
146	1	24	5,45-7,10	0,33	0,017	0,96	0,025	1,12	0,92	0,13	0,25	6,275
147	4	70	4,78-3,94	0,30	0,021	0,89	0,020	1,10	0,94	0,13	0,22	4,36
148	4	70	5,90-5,23	0,31	0,030	0,95	0,024	0,96	0,95	0,18	0,22	5,565
149	4	70	5,10-4,60	0,32	0,022	1,00	0,022	1,08	1,00	0,09	0,20	4,85
150	4	32	6,10-6,10	0,33	0,028	0,9	0,023	1,08	0,93	0,15	0,24	6,10
151	1	60	4,23-4,23	0,34	0,027	0,97	0,020	1,08	1,07	0,18	0,21	4,23
152	1	50	7,65-7,65	0,32	0,017	0,88	0,023	1,12	0,89	0,17	0,23	7,65
153	3	32	8,60-8,15	0,31	0,023	0,88	0,018	0,98	0,95	0,10	0,25	8,375
154	3	65	5,22-5,22	0,32	0,027	1,02	0,025	1,12	1,10	0,23	0,23	5,22
155	2	65	5,22-5,25	0,33	0,029	0,93	0,023	1,08	0,93	0,20	0,24	5,235
156	3	65	5,45-5,45	0,30	0,030	0,98	0,025	1,14	0,94	0,19	0,25	5,45
157	3	70	6,95-6,95	0,32	0,021	1,01	0,021	1,10	1,00	0,18	0,25	6,95
158	2	70	5,72-6,90	0,32	0,022	0,88	0,020	1,09	0,99	0,22	0,24	6,31
159	2	70	6,12-4,90	0,30	0,020	0,90	0,022	1,10	0,90	0,16	0,25	5,51
160	1	50	7,55-7,10	0,30	0,020	0,93	0,023	1,18	1,00	0,16	0,24	7,325
161	2	27	7,10-8,06	0,30	0,016	0,82	0,022	1,04	0,89	0,15	0,23	7,58
162	2	60	6,12-6,94	0,34	0,020	1,06	0,021	1,18	0,85	0,16	0,34	6,53
163	2	48	6,13-5,90	0,33	0,017	0,93	0,021	1,12	0,92	0,18	0,20	6,015
164	2	48	6,60-5,83	0,33	0,017	0,97	0,020	1,20	0,82	0,14	0,22	6,215
165	2	36	5,46-5,09	0,34	0,018	1,00	0,018	1,14	0,97	0,12	0,20	5,275
166	2	42	7,82-7,34	0,32	0,022	0,98	0,023	1,08	0,86	0,11	0,24	7,58
167	2	65	5,3-5,66	0,33	0,017	0,88	0,030	1,18	0,89	0,13	0,24	5,695
168	2	28	6,80-6,32	0,33	0,022	1,04	0,025	1,16	0,90	0,12	0,20	5,56
169	2	90	1,50-1,97	0,35	0,027	0,94	0,024	1,15	0,97	0,13	0,15	1,735
170	2	65	5,00-6,84	0,34	0,022	0,97	0,023	1,12	0,86	0,14	0,14	5,92
171	2	65	7,25-7,25	0,33	0,019	0,98	0,025	1,15	0,88	0,13	0,22	7,25
172	3	42	6,17-6,84	0,35	0,019	0,99	0,022	1,16	0,89	0,16	0,20	6,505
173	3	42	5,53-6,05	0,32	0,023	0,96	0,026	1,16	0,91	0,16	0,24	5,79
174	2	28	8,23-8,38	0,34	0,023	0,85	0,023	1,08	0,94	0,13	0,25	8,305
175	2	63	7,20-6,94	0,32	0,019	0,91	0,022	1,20	0,88	0,16	0,23	7,07
176	2	65	8,05-9,60	0,29	0,022	0,96	0,022	1,12	0,85	0,16	0,25	8,825

Продолжение

№ п.п.	Номер печи	Сечение штанги в мм	Ударная вязкость a_k кгм/см ² (два испытания)	Содержание химических элементов (в процентах)									Среднее значение ударной вязкости a_k ср
				C	P	Mn	S	Si	Cr	Ni	Cu		
1	2	3	4	5	6	7	8	9	10	11	12	13	
177	2	65	7,81—7,32	0,31	0,020	0,97	0,024	1,15	0,99	0,18	0,23	7,565	
178	4	60	7,62—7,25	0,30	0,018	0,91	0,020	1,06	0,87	0,16	0,20	7,435	
179	4	38	5,88—7,08	0,33	0,029	0,85	0,025	0,90	0,99	0,16	0,23	6,48	
180	4	65	5,08—5,05	0,33	0,018	0,83	0,024	1,16	0,93	0,14	0,20	5,065	
181	2	65	7,10—6,85	0,34	0,023	0,98	0,020	1,08	0,98	0,15	0,22	6,975	
182	2	60	7,00—8,35	0,33	0,020	0,94	0,023	1,10	0,97	0,15	0,19	7,675	
183	4	42	5,13—5,13	0,33	0,023	1,00	0,025	1,07	1,00	0,16	0,19	5,13	
184	4	65	6,10—5,94	0,31	0,022	0,87	0,023	1,07	0,93	0,16	0,24	6,02	
185	4	70	7,00—6,85	0,32	0,025	0,85	0,022	0,98	0,92	0,15	0,21	6,925	
186	4	32	6,12—6,83	0,32	0,029	0,87	0,028	1,16	0,97	0,16	0,34	6,475	
187	1	65	6,50—6,67	0,28	0,026	0,82	0,025	1,10	0,97	0,13	0,22	6,585	
188	4	20	6,94—5,04	0,32	0,019	0,95	0,020	1,11	0,93	0,18	0,22	5,99	
189	1	48	7,95—5,77	0,30	0,020	0,82	0,020	1,07	0,96	0,13	0,22	6,86	
190	2	75	6,70—6,42	0,33	0,022	0,91	0,023	1,07	1,01	0,15	0,25	6,56	
191	4	19	6,38—5,90	0,31	0,019	0,98	0,022	1,20	1,01	0,17	0,25	6,14	
192	1	70	6,58—6,33	0,30	0,018	0,87	0,024	1,07	0,86	0,13	0,22	6,455	
193	1	19	6,35—5,90	0,30	0,025	0,98	0,024	1,03	0,94	0,18	0,25	6,125	
194	1	22	6,36—5,88	0,30	0,023	0,91	0,020	1,07	0,94	0,12	0,22	6,12	
195	1	32	3,91—4,45	0,33	0,027	1,06	0,020	1,14	1,10	0,16	0,24	4,18	
196	2	65	3,26—4,23	0,34	0,026	0,93	0,030	1,20	1,09	0,13	0,22	3,745	
197	3	65	6,37—6,37	0,31	0,025	0,87	0,020	0,98	0,95	0,14	0,21	6,37	
198	1	70	5,65—5,75	0,30	0,022	0,93	0,025	1,17	0,93	0,15	0,23	5,70	
199	3	50	7,32—7,32	0,31	0,019	0,86	0,025	1,16	0,92	0,14	0,23	7,32	
200	1	70	4,45—3,65	0,32	0,023	1,00	0,025	1,18	1,00	0,24	0,21	4,05	
201	3	70	5,25—5,25	0,30	0,024	0,93	0,020	1,12	0,90	0,12	0,20	5,25	
202	4	60	7,54—8,24	0,30	0,018	0,87	0,023	1,18	0,96	0,13	0,22	7,89	
203	2	60	3,24—4,56	0,35	0,032	1,00	0,024	1,08	1,00	0,14	0,21	3,90	
204	1	70	4,43—4,33	0,35	0,032	1,00	0,024	1,08	1,00	0,14	0,21	4,38	
205	1	60	3,83—2,81	0,34	0,019	0,98	0,027	1,15	0,98	0,10	0,20	3,32	
206	2	60	7,22—6,85	0,32	0,019	0,91	0,024	1,22	0,95	0,13	0,21	7,035	
207	2	50	3,74—4,80	0,35	0,025	1,04	0,021	0,96	1,04	0,16	0,24	4,27	
208	2	50	5,81—7,10	0,32	0,030	0,86	0,018	0,88	0,91	0,14	0,24	6,455	
209	2	70	6,95—5,05	0,31	0,023	0,91	0,024	1,24	0,92	0,14	0,22	6,00	
210	3	38	6,67—6,43	0,35	0,020	0,98	0,024	1,16	0,92	0,14	0,19	6,55	
211	1	60	5,94—5,74	0,32	0,024	0,96	0,026	1,02	0,97	0,18	0,18	5,84	
212	2	65	5,16—5,79	0,31	0,018	0,98	0,024	1,12	0,98	0,12	0,22	5,475	
213	3	18	7,17—6,13	0,31	0,017	0,91	0,023	1,12	0,95	0,18	0,21	6,65	
214	4	26	4,42—4,83	0,35	0,031	1,02	0,020	1,12	0,92	0,10	0,22	4,625	
215	2	70	4,06—4,98	0,33	0,020	0,94	0,025	1,12	1,06	0,12	0,22	4,52	
216	4	70	6,13—5,97	0,32	0,020	0,97	0,022	1,13	0,83	0,12	0,20	6,05	
217	4	70	8,83—9,06	0,28	0,017	0,93	0,027	1,12	0,86	0,14	0,22	8,945	
218	2	70	9,32—8,50	0,29	0,017	0,84	0,026	0,96	0,90	0,13	0,20	8,91	
219	1	65	5,22—5,33	0,32	0,024	0,96	0,026	1,02	0,97	0,18	0,18	5,275	
220	3	65	7,10—7,10	0,34	0,018	1,05	0,020	1,06	0,92	0,10	0,20	7,10	

Таблица 2

№ п/п	Вес штампа P (кг)	$\log P$	Число поверхно- стей N	$\log N$	Трудоем- кость $T_{шт}$ (в часах)	$\log T$
1	2	3	4	5	6	7
1	0,566	-0,25	44	1,64	12,0	1,08
2	0,818	-0,09	60	1,78	13,0	1,11
3	0,539	-0,27	72	1,86	10,4	1,02
4	1,342	0,13	126	2,10	16,0	1,20
5	0,120	-0,92	52	1,72	11,2	1,05
6	1,78	0,25	112	2,05	34,4	1,53
7	1,153	0,06	54	1,73	9,0	0,954
8	0,585	-0,23	40	1,60	10,0	1,00
9	0,253	-0,597	164	2,21	13,3	1,12
10	0,187	-0,73	62	1,79	9,0	0,954
11	1,5	0,18	56	1,75	24,0	1,38
12	0,742	-0,13	110	2,04	20,8	1,32
13	0,841	-0,08	50	1,698	10,0	1,00
14	0,94	-0,03	40	1,60	7,5	0,88
15	0,55	-0,26	66	1,82	18,0	1,25
16	0,08	-1,10	62	1,79	10,0	1,00
17	0,115	-0,94	32	1,505	11,0	1,04
18	0,04	-1,40	64	1,81	4,3	0,63
19	0,28	-0,55	116	2,06	20,0	1,30
20	0,107	-0,97	46	1,66	8,4	0,92
21	0,20	-0,70	68	1,83	5,0	0,70
22	0,314	-0,50	50	1,698	11,2	1,05
23	1,32	0,12	46	1,66	11,0	1,04
24	0,774	-0,11	56	1,75	3,6	0,56
25	0,841	-0,08	50	1,698	10,0	1,00
26	0,675	-0,17	84	1,92	12,6	1,10
27	1,844	0,26	84	1,92	27,0	1,43
28	1,066	0,03	106	2,02	9,0	0,954
29	1,85	0,27	72	1,86	15,2	1,18
30	1,85	0,27	80	1,903	15,2	1,18
31	1,38	0,14	96	1,98	19,6	1,29
32	0,722	-0,14	58	1,76	12,0	1,08
33	0,268	-0,57	60	1,78	10,0	1,00

Продолжение

№ п.п.	Вес штампа P (кг)	$\log P$	Число поверхностей N	$\log N$	Трудоем- кость $T_{шт}$ (в часах)	$\log T$
1	2	3	4	5	6	7
34	0,724	-0,14	92	1,96	10,0	1,09
35	0,420	-0,38	84	1,92	11,5	1,06
36	0,307	-0,51	38	1,58	6,4	0,81
37	0,270	-0,57	48	1,68	6,4	0,81
38	3,68	0,57	100	2,00	22,4	1,35
39	4,11	0,62	106	2,02	30,0	1,48
40	0,662	-0,18	208	2,32	13,6	1,13
41	1,025	0,01	76	1,88	18,0	1,26
42	0,56	-0,25	40	1,60	20,8	1,32
43	0,128	-0,89	82	1,91	10,0	1,00
44	0,85	-0,08	50	1,698	18,0	1,26
45	1,685	0,23	40	1,60	8,4	0,92
46	0,134	-0,87	22	1,34	12,0	1,08
47	0,253	0,597	184	2,26	9,0	0,954
48	2,42	0,38	66	1,82	21,6	1,33
49	3,12	0,49	96	1,98	20,8	1,32
50	2,32	0,37	104	2,02	21,6	1,33
51	2,472	0,39	104	2,02	12,6	1,10
52	2,5	0,40	96	1,98	18,0	1,26
53	3,21	0,51	90	1,95	29,0	1,46
54	3,4	0,53	142	2,15	28,0	1,45
55	3,6	0,56	124	2,09	20,48	1,31
56	10,34	1,02	140	2,14	58,8	1,77
57	3,246	0,51	64	1,81	9,8	0,99
58	10,0	1,00	94	1,97	26,6	1,42
59	5,03	0,70	36	1,56	12,0	1,08
60	2,82	0,45	52	1,72	8,4	0,92
61	2,3	0,36	120	2,08	20,6	1,31
62	2,5	0,40	94	1,97	28,0	1,45
63	4,76	0,68	138	2,14	25,5	1,41
64	6,8	0,83	72	1,86	7,0	0,845
65	5,6	0,75	68	1,83	25,2	1,40
66	4,66	0,67	212	2,33	36,0	1,56

№ п/п	Вес штампа P (кг)	$\log P$	Число поверхно- стей N	$\log N$	Трудоем- кость $T_{ит.}$ (в часах)	$\log T$
1	2	3	4	5	6	7
67	2,953	0,47	48	1,68	10,64	1,03
68	3,18	0,50	144	2,16	20,5	1,31
69	3,00	0,48	96	1,98	9,8	0,99
70	2,61	0,42	36	1,56	11,2	1,05
71	3,12	0,49	134	2,13	22,0	1,34
72	2,35	0,37	94	1,97	23,8	1,38
73	10,83	0,04	96	1,98	33,6	1,53
74	6,04	0,78	112	2,05	46,2*	1,66
75	2,96	0,47	80	1,903	15,5	1,19
76	2,65	0,42	48	1,68	17,0	1,23
77	3,12	0,49	86	1,93	17,6	1,25
78	3,515	0,55	60	1,78	24,5	1,39
79	13,72	0,14	96	1,98	50,0	1,70
80	2,08	0,32	78	1,89	54,0	1,73
81	5,9	0,77	60	1,78	8,4	0,92
82	3,41	0,53	44	1,64	19,6	1,29
83	2,0	0,30	32	1,51	12,6	1,10
84	13,65	1,14	96	1,98	54,0	1,73
85	9,793	0,99	60	1,72	34,6	1,54
86	4,38	0,64	48	1,68	13,3	1,12
87	0,85	-0,08	128	2,12	16,0	1,20
88	0,539	-0,27	68	1,83	10,4	1,02
89	5,15	0,71	92	1,96	27,0	1,43
90	8,2	0,91	96	1,98	30,8	1,49
91	0,30	-0,52	34	1,53	7,0	0,85
92	1,384	0,14	112	2,05	19,0	1,28
93	1,422	0,15	106	2,03	33,4	1,52
94	1,34	0,13	86	1,93	20,0	1,30
95	0,486	-0,32	72	1,86	10,0	1,00
96	2,472	0,39	120	2,08	12,6	1,10
97	2,594	0,41	96	1,98	5,6	0,748
98	2,52	0,40	44	1,64	5,6	0,748
99	2,994	0,48	144	2,16	15,1	1,18
100	0,77	-0,11	108	2,03	14,0	1,15

Таблица 3

№ п/п	Порядковый номер предыдущего года	Объем работ (млн. руб.)	Численность рабочих (человек)	Фонд зарплаты рабочих (млн. руб.)	Накладные расходы (млн. руб.)
1951 г.					
1	—	26,9	1276	12,250	5,7
2	—	24,5	975	10,627	5,0
3	—	18,4	869	6,865	4,5
4	—	18,1	757	6,964	4,0
5	—	18,1	740	7,622	4,4
6	—	17,9	699	6,291	3,5
7	—	15,7	840	7,980	3,5
8	—	14,2	744	6,770	3,8
9	—	13,3	725	7,105	5,1
10	—	15,0	670	5,762	3,4
11	—	14,7	622	6,096	4,1
12	—	13,3	566	6,056	4,1
13	—	14,6	518	4,921	3,1
14	—	11,7	510	4,131	2,8
15	—	10,6	452	4,384	2,1
16	—	10,0	447	4,157	2,5
17	—	9,0	497	4,324	2,0
18	—	9,5	428	4,023	2,4
19	—	7,0	381	3,315	2,3
20	—	9,1	385	3,619	2,4
21	—	6,8	412	3,461	2,5
22	—	5,5	293	2,139	2,2
23	—	5,1	284	2,244	1,6
24	—	12,2	514	3,928	3,4
25	—	11,0	407	3,337	2,7
26	—	9,3	577	3,676	3,2
27	—	5,9	265	2,120	2,9
1952 г.					
28	—	25,9	977	10,649	4,8
29	—	23,5	724	6,806	3,7
30	6	19,8	983	9,24	4,4
31	2	18,8	828	8,860	3,7
32	4	19,1	766	7,354	4,8
33	24	18,8	615	5,289	3,7
34	—	17,4	583	5,830	3,6
35	15	14,1	591	6,265	4,0
36	—	13,8	593	5,396	3,8

Продолжение

№ п/п	Порядковый номер предыдущего года	Объем работ (млн. руб.)	Численность рабочих (человек)	Фонд зарплаты рабочих (млн. руб.)	Накладные расходы (млн. руб.)
37	—	13,7	611	5,194	3,7
38	14	13,8	562	4,608	4,1
39	20	13,9	488	5,856	2,4
40	—	10,6	740	7,326	2,5
41	—	11,9	685	5,754	3,5
42	—	12,1	582	5,820	4,0
43	—	11,2	508	4,267	3,1
44	—	10,8	509	4,581	2,7
45	—	11,7	499	5,190	3,6
46	—	11,8	483	4,830	2,7
47	—	9,8	502	4,518	2,9
48	—	2,8	169	1,318	1,6
49	—	5,9	250	2,150	1,3
50	—	8,7	359	3,482	2,5
51	—	7,6	363	3,231	2,1
52	—	7,3	373	3,059	2,4
53	—	7,9	387	3,212	2,0
54	3	8,9	395	3,634	2,5
55	—	5,4	253	2,125	1,8
56	10	10,2	376	3,008	2,8
1953 г.					
57	29	25,1	861	9,213	4,0
58	34	22,7	725	8,990	3,9
59	33	20,3	793	6,265	4,7
60	40	19,9	907	7,347	4,8
61	32	18,2	760	7,524	4,3
62	37	17,3	738	6,642	3,5
63	36	16,5	634	5,833	3,0
64	39	17,0	683	8,009	3,6
65	—	17,1	624	7,051	3,3
66	20	16,2	593	6,404	2,9
67	—	17,3	606	5,575	3,1
68	—	16,3	507	5,019	2,8
69	35	12,9	629	7,485	3,5
70	48	13,8	565	5,820	4,6
71	44	10,1	588	5,116	3,5
72	45	10,9	515	5,510	2,9
73	46	11,4	500	5,200	2,7
74	54	11,3	450	4,455	2,8

Продолжение

№ п/п	Порядковый номер предыдущего года	Объем работ (млн. руб.)	Численность рабочих (человек)	Фонд зарплаты рабочих (млн. руб.)	Накладные расходы (млн. руб.)
75	43	8,7	510	4,488	3,0
76	50	10,0	432	4,968	2,9
77	49	5,2	419	4,022	2,4
78	—	7,4	259	1,570	1,6
79	41	2,2	162	1,442	1,2
80	47	2,6	151	1,329	1,5
81	—	13,9	448	2,150	1,9
82	42	7,6	316	2,812	2,3
83	55	8,0	326	3,130	2,0
84	—	13,2	541	2,867	3,1
85	38	11,0	475	3,752	3,0
86	52	8,2	373	3,842	3,0
1954 г.					
87	—	37,5	1299	14,159	7,5
88	60	26,2	1100	11,000	5,9
89	67	25,9	794	9,131	5,5
90	68	23,2	729	8,821	3,5
91	43	20,1	1067	10,883	7,3
92	62	20,5	929	9,847	4,5
93	64	21,4	824	11,454	4,9
94	—	20,3	694	7,495	3,7
95	58	20,0	801	8,010	3,8
96	59	19,6	788	7,722	4,6
97	81	17,1	604	6,886	3,2
98	63	15,1	731	7,529	4,5
99	85	15,2	678	6,509	3,8
100	66	15,6	563	7,432	3,3
101	86	15,2	562	8,149	3,6
102	—	15,2	553	5,475	2,7
103	74	16,3	490	5,194	4,1
104	73	14,1	563	6,193	3,2
105	—	12,3	749	7,265	3,6
106	69	11,2	486	5,492	2,9
107	72	11,1	453	5,391	3,0
108	—	10,8	435	4,220	2,5
109	76	11,0	428	5,916	3,1
110	83	10,4	377	4,411	2,7
111	70	7,6	389	4,084	3,0
112	100	9,8	399	3,671	2,0
113	82	9,4	365	3,431	2,6

Таблица 4

№ п/п	Упру- гость по Тоledo	Замок в свободном состоянии (в мм)	Магнит- ная про- ницае- мость	№ п/п	Упру- гость по Тоledo	Замок в свободном состоянии (в мм)	Магнит- ная про- ницае- мость
1	1450	18,10	9	34	1500	17,40	8
2	1450	19,10	8	35	1430	18,80	6
3	1460	18,10	8	36	1500	20,20	4
4	1660	19,60	9	37	1460	18,50	9
5	1575	18,90	9	38	1340	16,50	4
6	1460	18,50	5	39	1450	17,10	10
7	1520	18,10	9	40	1440	16,50	10
8	1480	17,80	9	41	1480	17,20	11
9	1440	16,80	9	42	1450	17,50	9
10	1420	17,90	8	43	1350	18,00	4
11	1520	18,50	10	44	1480	18,30	8
12	1420	18,50	—	45	1580	19,30	6
13	1380	19,10	5	46	1360	17,30	5
14	1380	17,40	6	47	1490	18,00	9
15	1380	17,60	10	48	1300	16,50	4
16	1500	18,80	5	49	1280	16,10	5
17	1480	18,80	3	50	1390	16,40	10
18	1340	16,90	6	51	1510	18,80	11
19	1380	16,80	6	52	1320	16,50	6
20	1340	16,50	8	53	1340	17,90	8
21	1410	17,20	11	54	1430	17,60	10
22	1360	16,80	5	55	1230	15,90	4
23	1550	18,50	10	56	1350	16,80	10
24	1480	18,50	7	57	1340	16,80	10
25	1340	17,00	6	58	1330	18,30	2
26	1510	18,80	11	59	1340	16,80	9
27	1400	17,00	9	60	1450	17,80	9
28	1440	17,50	5	61	1290	17,80	3
29	1460	18,20	5	62	1360	17,40	11
30	1400	17,80	4	63	1500	19,20	8
31	1580	19,80	11	64	1400	17,80	5
32	1400	18,80	3	65	1430	17,80	9
33	1480	18,00	9	66	1430	17,80	4

Продолжение

№ п/п	Упру- ость по Тоledo	Замок в свободном состоянии (в мм)	Магнит- ная про- ницае- мость	№ п/п	Упру- гость по Тоledo	Замок в свободном состоянии (в мм)	Магнит- ная про- ницае- мость
67	1370	16,30	6	100	1470	18,30	10
68	1390	17,10	6	101	1370	16,30	9
69	1480	18,50	11	102	1390	18,00	4
70	1120	19,40	11	103	1430	19,20	5
71	1510	19,50	8	104	1460	18,00	11
72	1420	18,00	5	105	1260	17,10	3
73	1410	17,00	10	106	1360	17,30	4
74	1460	18,10	10	107	1390	18,40	4
75	1210	15,20	9	108	1430	19,00	5
76	1410	17,40	11	109	1490	19,30	7
77	1250	16,30	9	110	1340	18,30	6
78	1340	16,50	8	111	1400	18,80	6
79	1390	16,90	10	112	1490	18,80	7
80	1410	17,40	10	113	1390	17,60	5
81	1500	18,40	9	114	1310	19,50	9
82	1510	19,30	8	115	1440	17,40	8
83	1440	17,50	11	116	1300	16,50	4
84	1430	18,00	10	117	1350	18,00	11
85	1400	17,20	8	118	1380	16,60	5
86	1460	18,00	10	119	1450	17,10	9
87	1260	17,90	5	120	1350	17,60	11
88	1420	17,60	4	121	1430	17,40	10
89	1360	18,40	2	122	1410	17,10	11
90	1420	18,80	8	123	1440	17,30	10
91	1490	19,20	9	124	1280	16,50	5
92	1330	18,10	6	125	1470	19,60	6
93	1340	18,10	8	126	1350	17,60	8
94	1430	19,20	8	127	1370	17,90	7
95	1310	18,20	3	128	1450	17,80	9
96	1430	18,50	4	129	1530	18,40	9
97	1240	16,10	5	130	1450	17,80	7
98	1470	18,40	6	131	1370	16,60	7
99	1450	17,80	7	132	1440	18,80	8

Продолжение

№ п/п	Упру- гость по Тоledo	Замок в свободном состоянии (в мм)	Магнит- ная про- ницае- мость	№ п/п	Упру- гость по Тоledo	Замок в свободном состоянии (в мм)	Магнит- ная про- ницае- мость
133	1340	17,80	6	167	1390	18,80	4
134	1400	17,10	7	168	1380	17,80	3
135	1370	17,60	8	169	1330	17,60	8
136	1270	16,80	8	170	1380	17,80	8
137	1320	16,30	5	171	1390	19,00	2,5
138	1380	17,80	6	172	1570	19,00	5
139	1320	17,80	8	173	1420	19,80	14
140	1320	16,50	5	174	1500	18,90	8
141	1270	15,80	9	175	1410	19,40	9
142	1320	19,00	10	176	1500	18,00	7
143	1440	18,20	10	177	1450	18,90	8
144	1450	17,80	5	178	1400	18,80	10
145	1390	18,40	2	179	1470	18,10	5
146	1390	17,90	5	180	1400	19,30	12
147	1470	18,10	11	181	1550	18,80	8
148	1500	17,50	11	182	1490	19,00	10
149	1330	16,00	6	183	1480	17,80	10
150	1450	16,80	11	184	1340	20,60	3
151	1300	17,90	6	185	1440	18,10	4
152	1350	17,80	3,5	186	1400	18,80	8
153	1515	19,80	15,5	187	1300	18,30	5
154	1400	18,90	5	188	1500	18,50	8
155	1500	18,80	9,5	189	1380	19,40	9
156	1415	19,00	3	190	1280	19,60	4
157	1580	19,80	8	191	1220	17,20	3
158	1510	19,80	15	192	1300	16,10	6
159	1425	18,80	10	193	1540	16,60	10
160	1435	17,50	13	194	1390	17,60	14
161	1620	21,60	10	195	1480	17,80	10
162	1230	17,80	3	196	1340	19,00	12
163	1460	19,40	8	197	1510	18,30	5
164	1410	18,10	10	198	1340	20,30	5
165	1420	18,20	8	199	1290	18,80	9
166	1510	18,60	6,5	200	1400	16,50	4

Приложение II
ТАБЛИЦЫ НОРМАЛЬНОГО ЗАКОНА РАСПРЕДЕЛЕНИЯ

Таблица I

Таблица значений функции $\Phi(t) = \frac{1}{\sqrt{2\pi}} \int_{-t}^{+t} e^{-\frac{1}{2}t^2} dt$

<i>t</i>	$\Phi(t)$	<i>t</i>	$\Phi(t)$	<i>t</i>	$\Phi(t)$	<i>t</i>	$\Phi(t)$
0,00	0,00080	0,30	0,23582	0,60	0,45149	0,90	0,63188
01	00798	31	24344	61	45814	91	63718
02	01596	32	25103	62	46474	92	64243
03	02393	33	25860	63	47131	93	64763
04	03191	34	26614	64	47783	94	65278
05	03988	35	27366	65	48431	95	65789
06	04784	36	28115	66	49075	96	66294
07	05581	37	28862	67	49714	97	66795
08	06376	38	29605	68	50350	98	67291
09	07171	39	30346	69	50981	99	67783
0,10	0,07966	0,40	0,31084	0,70	0,51607	1,00	0,68269
11	08759	41	31819	71	52230	01	68750
12	09552	42	32552	72	52848	02	69227
13	10343	43	33280	73	53461	03	69699
14	11134	44	34006	74	54070	04	70166
15	11924	45	34729	75	54675	05	70628
16	12712	46	35448	76	55275	06	71086
17	13499	47	36164	77	55870	07	71538
18	14285	48	36877	78	56461	08	71986
19	15069	49	37587	79	57047	09	72429
0,20	0,15852	0,50	0,38292	0,80	0,57629	1,10	0,72867
21	16633	51	38995	81	58206	11	73300
22	17433	52	39694	82	58778	12	73729
23	18191	53	40389	83	59346	13	74152
24	18967	54	41080	84	59909	14	74571
25	19741	55	41768	85	60468	15	74986
26	20514	56	42452	86	61021	16	75395
27	21284	57	43132	87	61570	17	75800
28	22052	58	43809	88	62114	18	76200
29	22818	59	44481	89	62653	19	76595

Продолжение

t	$\Phi(t)$	t	$\Phi(t)$	t	$\Phi(t)$	t	$\Phi(t)$
1,20	0,76986	1,54	0,87644	1,88	0,93989	2,22	0,97358
21	77372	55	87886	89	94124	23	97425
22	77754	56	88124	1,90	0,94257	24	97491
23	78130	57	88358	91	94387	25	97555
24	78502	58	88589	92	94514	26	97618
25	78870	59	88817	93	94639	27	97679
26	79233	1,60	0,89040	94	94762	28	97739
27	79592	61	89260	95	94882	29	97798
28	79945	62	89477	96	95000	2,30	0,97855
29	80295	63	89690	97	95116	31	97911
1,30	0,80640	64	89899	98	95230	32	97966
31	80980	65	90106	99	95341	33	98019
32	81316	66	90309	2,00	0,95450	34	98072
33	81648	67	90508	01	95557	35	98123
34	81975	68	90704	02	95662	36	98172
35	82298	69	90897	03	95764	37	98221
36	82617	1,70	0,91087	04	95865	38	98269
37	82931	71	91273	05	95964	39	98315
38	83241	72	91457	06	96060	2,40	0,98360
39	83547	73	91637	07	96155	41	98405
1,40	0,83849	74	91814	08	96247	42	98448
41	84146	75	91988	09	96338	43	98490
42	84439	76	92159	2,10	0,96427	44	98531
43	84728	77	92327	11	96514	45	98571
44	85013	78	92492	12	96599	46	98611
45	85294	79	92655	13	96683	47	98649
46	85571	1,80	0,92814	14	96765	48	98686
47	85844	81	92970	15	96844	49	98723
48	86113	82	93124	16	96923	2,50	0,98758
49	86378	83	93275	17	96999	51	98793
1,50	0,86639	84	93423	18	97074	52	98826
51	86696	85	93569	19	97148	53	98859
52	87149	86	93711	2,20	0,97219	54	98891
53	87398	87	93852	21	97289	55	98923

Продолжение

t	$\Phi(t)$	t	$\Phi(t)$	t	$\Phi(t)$	t	$\Phi(t)$
2,56	0,98953	2,90	0,99627	3,25	0,99885	3,60	0,99968
57	98983	91	99639	26	99889	61	99969
58	99012	92	99650	27	99892	62	99971
59	99040	93	99661	28	99896	63	99972
2,60	0,99068	94	99672	29	99900	64	99973
61	99095	95	99682	3,30	0,99903	65	99974
62	99121	96	99692	31	99907	66	99975
63	99146	97	99702	32	99910	67	99976
64	99171	98	99712	33	99913	68	99977
65	99195	99	99721	34	99916	69	99978
66	99219	3,00	0,99730	35	99919	3,70	0,99978
67	99241	01	99739	36	99922	71	99979
68	99263	02	99747	37	99925	72	99980
69	99285	03	99755	38	99928	73	99981
2,70	0,99307	04	99763	39	99930	74	99982
71	99327	05	99771	3,40	0,99933	75	99982
72	99347	06	99779	41	99935	76	99983
73	99367	07	99786	42	99937	77	99984
74	99386	08	99793	43	99940	78	99984
75	99404	09	99800	44	99942	79	99985
76	99422	3,10	0,99806	45	99944	3,80	0,99986
77	99439	11	99813	46	99946	81	99986
78	99456	12	99819	47	99948	82	99987
79	99473	13	99825	48	99950	83	99987
2,80	0,99489	14	99831	49	99952	84	99988
81	99505	15	99837	3,50	0,99953	85	99988
82	99520	16	99842	51	99955	86	99989
83	99535	17	99848	52	99957	87	99989
84	99549	18	99853	53	99958	88	99990
85	99563	19	99858	54	99960	89	99990
86	99576	3,20	0,99863	55	99961	3,90	0,99990
87	99590	21	99867	56	99963	91	99991
88	99602	22	99872	57	99964	92	99991
89	99615	23	99876	58	99966	93	99992
		24	99880	59	99967	94	99992
						95	99992
						96	99992
						97	99993
						98	99993
						99	99993

Таблица 2

Таблица значений функции $\varphi(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2}$

t	0	1	2	3	4	5	6	7	8	9
0,0	0,3989	3989	3989	3988	3986	3984	3982	3980	3977	3973
0,1	3970	3955	3961	3956	3951	3945	3939	3932	3925	3918
0,2	3910	3902	3894	3885	3876	3867	3857	3847	3836	3825
0,3	3814	3802	3790	3778	3765	3752	3739	3726	3712	3697
0,4	3683	3668	3656	3637	3621	3605	3589	3572	3555	3538
0,5	3521	3503	3485	3467	3448	3429	3410	3391	3372	3352
0,6	3332	3312	3292	3271	3251	3230	3209	3187	3166	3144
0,7	3123	3101	3079	3056	3034	3011	2989	2966	2943	2920
0,8	2897	2874	2850	2827	2803	2780	2756	2732	2709	2685
0,9	2661	2637	2613	2589	2565	2541	2516	2492	2468	2444
1,0	0,2420	2396	2371	2347	2323	2299	2275	2251	2227	2203
1,1	2179	2155	2131	2107	2083	2059	2036	2012	1989	1965
1,2	1942	1919	1895	1872	1849	1826	1804	1781	1758	1736
1,3	1714	1691	1669	1647	1626	1604	1582	1561	1539	1518
1,4	1497	1476	1456	1435	1415	1394	1374	1354	1334	1315
1,5	1295	1276	1257	1238	1219	1200	1182	1163	1145	1127
1,6	1109	1092	1074	1057	1040	1023	1006	989	973	957
1,7	0940	0925	0909	0893	0878	0863	0848	0833	0818	0804
1,8	0790	0775	0761	0748	0734	0721	0707	0694	0681	0669
1,9	0656	0644	0632	0620	0608	0596	0584	0573	0562	0551
2,0	0,0540	0529	0519	0508	0498	0488	0478	0468	0459	0449
2,1	0440	0431	0422	0413	0404	0396	0387	0379	0371	0363
2,2	0355	0347	0339	0332	0325	0317	0310	0303	0297	0290
2,3	0283	0277	0270	0264	0258	0252	0246	0241	0235	0229
2,4	0224	0219	0213	0208	0203	0198	0194	0189	0184	0180
2,5	0175	0171	0167	0163	0158	0154	0154	0147	0143	0139
2,6	0136	0132	0129	0126	0122	0119	0116	0113	0110	0107
2,7	0104	0101	0099	0096	0093	0091	0088	0086	0084	0081
2,8	0079	0077	0075	0073	0071	0069	0067	0065	0063	0061
2,9	0060	0058	0056	0055	0053	0051	0050	0048	0047	0046
3,0	0,0044	0043	0042	0040	0039	0038	0037	0036	0035	0034
3,1	0033	0032	0031	0030	0029	0028	0027	0026	0025	0025
3,2	0024	0023	0022	0022	0021	0020	0020	0019	0018	0018
3,3	0017	0017	0016	0016	0015	0015	0014	0014	0013	0013
3,4	0012	0012	0012	0011	0011	0010	0010	0010	0009	0009
3,5	0009	0008	0008	0008	0008	0007	0007	0007	0007	0006
3,6	0006	0006	0006	0005	0005	0005	0005	0005	0005	0004
3,7	0004	0004	0004	0004	0004	0004	0003	0003	0003	0003
3,8	0003	0003	0003	0003	0003	0002	0002	0002	0002	0002
3,9	0002	0002	0002	0002	0002	0002	0002	0002	0001	0001

ЛИТЕРАТУРА

На русском языке имеются следующие монографии по теории корреляции:

1. Е. Е. Слуцкий, Теория корреляции и элементы учения о кривых распределения, Киев, 1912.
2. А. А. Чупров, Основные проблемы теории корреляции, М., 1926.
3. В. И. Романовский, Элементы теории корреляции, Ташкент, 1928.
4. Я. И. Лукомский, Теория корреляции. Основы статистического изучения связи между величинами, М., 1948.

Большие разделы теории корреляции содержатся в следующих курсах математической статистики:

5. А. Я. Боярский, В. Н. Старовский, В. И. Хотимский, Б. С. Ястремский, Курс математической статистики, 1-е изд., 1930, 2-е изд., 1931.

6. В. И. Романовский, Математическая статистика, 1938.

7. Г. Крамер, Математические методы статистики, 1948.

8. А. К. Митропольский, Статистическое исчисление, вып. 1—4, Л., 1952—1954, а также в книге

9. В. С. Немчинов, Сельскохозяйственная статистика с основами общей теории, М., 1945.

Изложение вопросов теории корреляции в связи с техническими приложениями читатель найдет в книгах:

10. А. М. Длин, Математическая статистика в технике, 2-е изд., 1953, 3-е изд. (в печати).

11. И. В. Дунин-Барковский и Н. В. Смирнов, Теория вероятностей и математическая статистика в технике (общая часть), М., 1955.

12. А. Хальд, Математическая статистика с техническими приложениями, М., 1955.

Из иностранной литературы рекомендуются:

13. M. Ezekiel, Methods of Correlation Analysis, 1 изд., 1930, 2 изд., 1941, и курсы математической статистики с большими разделами теории корреляции.

14. M. G. Kendall, The Advanced Theory of Statistics, т. 1, 1943, т. II, 1948.

15. G. U. Yule and M. G. Kendall, An Introduction to the Theory of Statistics. 14 изд., 1950.

Популярное изложение элементов теории вероятностей читатель найдет в книге

16. Б. В. Гнеденко и А. Я. Хинчин, Элементарное введение в теорию вероятностей, 5 изд., Москва, 1957.

Систематическое изложение теории вероятностей дано в книгах:

17. Б. В. Гнеденко, Курс теории вероятностей, 2 изд., М., 1954.

18. С. Н. Бернштейн, Теория вероятностей, 4 изд., М.—Л., 1946.

19. А. А. Марков, Исчисление вероятностей, 4 изд., М., 1924.

20. Г. П. Боев, Теория вероятностей, М.—Л., 1950.

21. В. И. Гливенко, Курс теории вероятностей, М.—Л., 1939.

22. В. Л. Гончаров, Теория вероятностей, М.—Л., 1939.

23. Торнтон Фрай, Теория вероятностей для инженеров, М.—Л., 1934.

СОДЕРЖАНИЕ

Предисловие	3
-----------------------	---

РАЗДЕЛ I

РЯДЫ РАСПРЕДЕЛЕНИЯ

Глава I. Построение и графическое изображение рядов распределения	7
§ 1. Построение рядов распределения	7
§ 2. Графическое изображение рядов распределения	14
Глава II. Анализ рядов распределения	19
§ 1. Проблема измерения массового уровня	19
§ 2. Средняя арифметическая	21
§ 3. Медиана	23
§ 4. Мода	25
§ 5. Сопоставление средней арифметической, медианы и моды	26
§ 6. Проблема измерения колеблемости. Среднее абсолютное отклонение, среднее квадратическое отклонение, размах и коэффициент вариации	31
§ 7. Упрощенный способ вычисления \bar{x} и σ	36
§ 8. Нулевое и минимальное свойства средней арифметической	41
§ 9. Средняя арифметическая и дисперсия для объединенной совокупности	43
§ 10. Примеры использования рядов распределения на производстве	46
§ 11. Проблема формы кривой распределения	56

РАЗДЕЛ II

ТЕОРИЯ КОРРЕЛЯЦИИ

Глава III.	59
§ 1. Задачи статистического измерения связи	59
§ 2. Корреляционная зависимость	62
Глава IV. Исследование формы связи	70
§ 1. Эмпирическая линия регрессии	70
§ 2. Истолкование линии регрессии	74
§ 3. Теоретическая линия регрессии	78
§ 4. Выбор и обоснование типа кривой регрессии	81

Глава V. Расчет параметров уравнения регрессии	94
§ 1. Прямолинейная зависимость. Данные не сгруппированы	94
§ 2. Прямолинейная зависимость. Данные сгруппированы	99
§ 3. Криволинейная зависимость	106
Глава VI. Функциональная корреляция	113
§ 1. Функциональное преобразование переменных. Линии функциональной регрессии	113
§ 2. Функциональные средние	119
§ 3. Линии функциональной регрессии. Функциональная корреляция	125
Глава VII. Измерение тесноты связи	134
§ 1. Проблема тесноты связи. Эмпирическое корреляционное отношение	134
§ 2. Теоретическое корреляционное отношение	140
§ 3. Коэффициент корреляции	148
Глава VIII. Сопряженные показатели корреляции. Симметричные схемы действия факторов	160
§ 1. Сопряженные показатели корреляции	160
§ 2. Коэффициент корреляции и симметричные схемы действия факторов	172
Глава IX. Корреляция трех переменных	184
§ 1. Уравнение множественной регрессии для трех переменных	184
§ 2. Показатели тесноты связи	194
Глава X. Корреляция многих переменных	202
§ 1. Уравнение множественной регрессии для p переменных	202
§ 2. Уравнение чистой регрессии	213
§ 3. Коэффициент множественной корреляции	215
§ 4. Приведение множественной корреляции к линейному виду путем функционального преобразования переменных	217
§ 5. Приведение множественной корреляции к линейному виду путем умножения переменных на функции аргументов	221

РАЗДЕЛ III

ОШИБКИ СТАТИСТИЧЕСКИХ ПОКАЗАТЕЛЕЙ

Глава XI. Вероятность	225
§ 1. Понятие вероятности	225
§ 2. Вероятности и процессы накопления материала	228
§ 3. Свойства вероятностей	233
Глава XII. Законы распределения и корреляции	239
§ 1. Общие сведения	239
§ 2. Нормальный закон распределения	243
§ 3. Нормальная корреляция	252
Глава XIII. Статистические показатели в процессе накопления материала	261
§ 1. Относительная частота в процессе накопления материала	261
§ 2. Оценка погрешностей относительной частоты	265
§ 3. Средняя арифметическая и среднее квадратическое отклонение в процессе накопления материала	269
§ 4. Показатели корреляции в процессе накопления материала	279

РАЗДЕЛ IV

КОМПЛЕКСНЫЕ ПРИМЕРЫ ИСПОЛЬЗОВАНИЯ
КОРРЕЛЯЦИОННОГО АНАЛИЗА НА ПРОИЗВОДСТВЕ

Глава XIV. Влияние химического состава на ударную вязкость стали	292
---	------------

Глава XV. Корреляционный анализ технологических цепей . . .	318
§ 1. Предварительные замечания	318
§ 2. Общая постановка и общий метод решения задачи	319
§ 3. Случай прямолинейной регрессии и постоянного рассеяния	324
§ 4. О физическом смысле формул	327
§ 5. Пример расчета показателей	330
§ 6. Распределение существенно положительных величин	331
§ 7. О поправках на ненадежность измерений и на группировку данных	345
§ 8. Расчет простой технологической цепи	348
Глава XVI. Применение методов корреляции к анализу факторов, влияющих на накладные расходы	351
Приложение I. Таблицы опытных данных	366
Приложение II. Таблицы нормального закона распределения	380
Литература	384

Яков Ильич Лукомский

Теория корреляции и ее применение к анализу производства

Редакторы *Е. М. Шенцис, А. Л. Лурье*
Техн. редактор *В. А. Виноградова*

Художник *Н. А. Зарин*
Корректор *А. Н. Теряева*

Сдано в набор 8/VI 1957 г. Подписано к печати 1/III 1958 г.
Формат 60×92¹/₁₆ Объем 24,25 печ. л. Уч.-изд. л. 22,75
А-00119 Тираж 4000 Цена 12 р. 90 к. Заказ 2380

Типография № 5 Углетехиздата. Москва, Южно-портовый 1-й пр., д. 17

Стр.	Строка	Напечатано	Должно быть
23	12 снизу	23, 20, 25	24, 23, 20, 25
36	6 и 7 снизу	x и σ	\bar{x} и σ
172	9 сверху	0,557	0,542
172	16 "	0,471	0,458
175	3 снизу		пропущено η
264	формула (13.1.5)	$e^{-\frac{1}{2} dt}$	$e^{-\frac{t^2}{2} dt}$
270	5 снизу	x .	\bar{x}
272	формула (13.3.5)	$\left(\frac{n-1}{2}\right)$	$\left(\frac{n-1}{2}\right)$
278	7 сверху	x и σ	\bar{x} и σ
380	столбец 2 8 снизу	17 433	17 413
383	столбец 1 2 сверху	3955	3965
383	столбец 2 5 сверху	3656	3653
383	столбец 6 15 снизу	0154	0151
383	столбец 7 17 сверху	0939	0989

Academician S. A. Lebedev

**CERTAIN WORKS
IN THE SPHERE
OF COMPUTING
TECHNIQUE**

1 9 5 6

INSTITUTE OF EXACT MECHANICS AND COMPUTING TECHNIQUE
ACADEMY OF SCIENCES OF THE U.S.S.R.

Academician S. A. Lebedev

CERTAIN WORKS
IN THE SPHERE
OF COMPUTING
TECHNIQUE

1 9 5 6

CERTAIN WORKS IN THE SPHERE OF COMPUTING TECHNIQUE

One of the principal works carried out by the Institute is the high speed computer (BESM) on which I and my collaborators have already reported at a number of conferences.

I wish to recall certain basic parameters of the machine.

The BESM is a digital electronic machine for the solution of laborious problems in physics, mechanics, astronomy, engineering, etc. In designing the machine emphasis was laid on convenience of programming and simplicity of operation.

A binary number system with a floating point was selected. Calculations are made, as a rule, with normalized numbers. The mantissa of the number is represented by 32 binary positions; then, there is one position for the sign of the number, 5 positions for the characteristic of the number and one position for the sign of the characteristic.

The machine has a three-address system. The code of each address consists of 11 bits, the operation part -- of 5 bits, and the sixth operation position is for result normalization interlocking. Thus, provision has been made in the machine for 31 instructions.

The standard cycle of the machine includes selection of two numbers from the memory, carrying the predetermined operation with these numbers, sending the result to the memory and selection of a new instruction from the memory. This cycle takes 77 microsecs.

Addition and subtraction are carried out in 77 to 182 microsecs, depending on the necessity of equalizing characteristics or normalizing results. Multiplication takes 270 microsecs, and division 288 microsecs.

When solving complex problems on the machine, the average operation speed is from 7,000 to 8,000 three-address operations per second including reference to the magnetic drum and the magnetic tapes.

All operations are carried out by a single universal parallel action arithmetical unit.

Special cathode-ray tubes are employed for the memory. The capacity of the memory is 1023 numbers.

In order to extend the sphere of solvable problems to include such as require large storage capacities, the machine is provided with a magnetic drum and a magnetic tape storage devices.

The magnetic drum has a capacity of 5120 numbers.

The magnetic tape storage consists of 4 series action magnetophones. One tape can hold an order of 30,000 numbers, making a total of about 120,000 numbers for the 4 magnetophones.



Fig.1 General View of the BESM.

The numbers and instructions are put into the machine from a punched tape in the form of a series code. The reading rate is 20 numbers per second.

The results of the calculations are withdrawn from the machine by recording on magnetic tape and then printed on a motion picture film inde-

pendently, on a special photoprinting device. The printing device works at a rate of 200 numbers per second.

Besides the above photoprinting unit there is an electromechanical printing unit, controlled directly by the machine. The printing rate is 1.5 numbers per second. This printing unit is used when the material to be withdrawn is small in volume compared to the calculations, and to print control values for checking the progress of the calculations.

The general view of the machine is shown on Fig.1. The BESM has been in operation since 1952.

The machine operates 24 hours per day, part of the time being spent on checking. The useful operating time of the machine is 72 per cent, the time spent on checking is 20 per cent, and error losses total eight per cent.

During the time of the operation the machine has solved many laborious and complex mathematical problems in various fields of science and technology, as well as certain logical problems, for instance experiments in translation of scientific texts from English into Russian.

The Institute conducts further work in elaborating the principles of designing electronic computers and creating more perfect units. I shall dwell on some of these works.

A dynamic trigger circuit with an accumulating capacitor has been worked out for the temporary storage of codes in separate units of the machine (suggested by Master of Science P. P. Golovistikov). The circuit of the dynamic trigger and the diagram of operation are shown on Figs.2 and 3.

When the trigger state corresponds to code "0", there is a big negative voltage on the grid of the valve which is backed up by the bias voltage ($-V_{CM}$) coupled with the grid by the resistance R_2 . The valve is reliably closed and the synchronising pulses (Γ) applied to the cathode do not pass through the valve. The bias value ($-V_{CM}$) and the Γ pulse amplitude are selected so that when the cathode voltage decreases at the expense of the Γ pulses, the grid voltage with relation to the cathode would not surpass the cut-off voltage of the valve. When the trigger is in this state there would be no pulses on the output winding of the transformer which corresponds to code "0".

When a positive pulse is applied to the input Y"I" the capacitance "C" is charged through the D"I" diode to -10v. The winding of the transformer which gives the Y"I" pulse is coupled with the base voltage $-V_{CM}$ so that

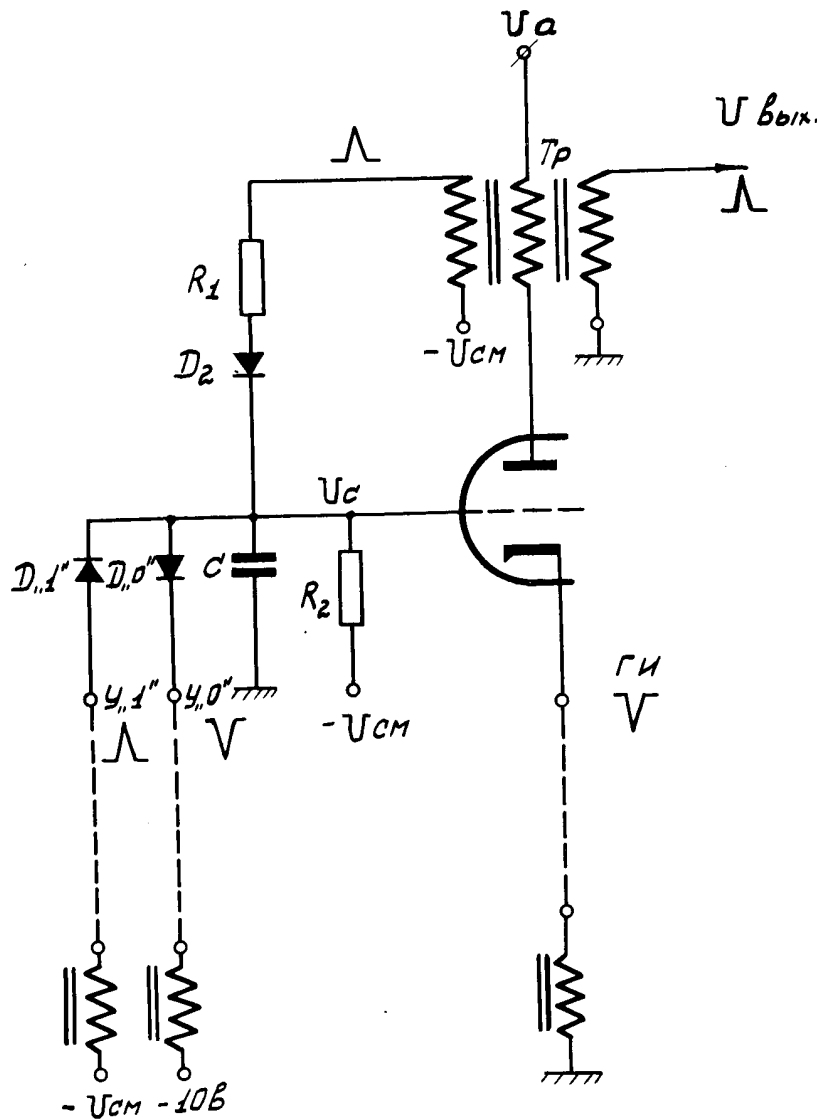


Fig. 2. 1) The scheme of the triode dynamic trigger; 2) output.

there would be no charging of the capacitance through diode D"I" when the trigger is in a state corresponding to code "O". Therefore the Y"I" pulse begins the voltage equalling $-V_{CM}$. The surpassing of the Y"I" signal over and above -10v will entail only a short time increase in voltage on the capacitance "C" since the excessive voltage is discharged through the D"O" diode which is coupled through the winding of the transformer which creates

the Y"O" signal with the reference voltage, $-10v$. Thus, after the passage of the Y"I" pulse the voltage on the valve grid will be $-10v$. The capacitance "C" maintaining this voltage will be gradually discharging through R_2 resistance. The time constant is taken sufficiently great so that in the

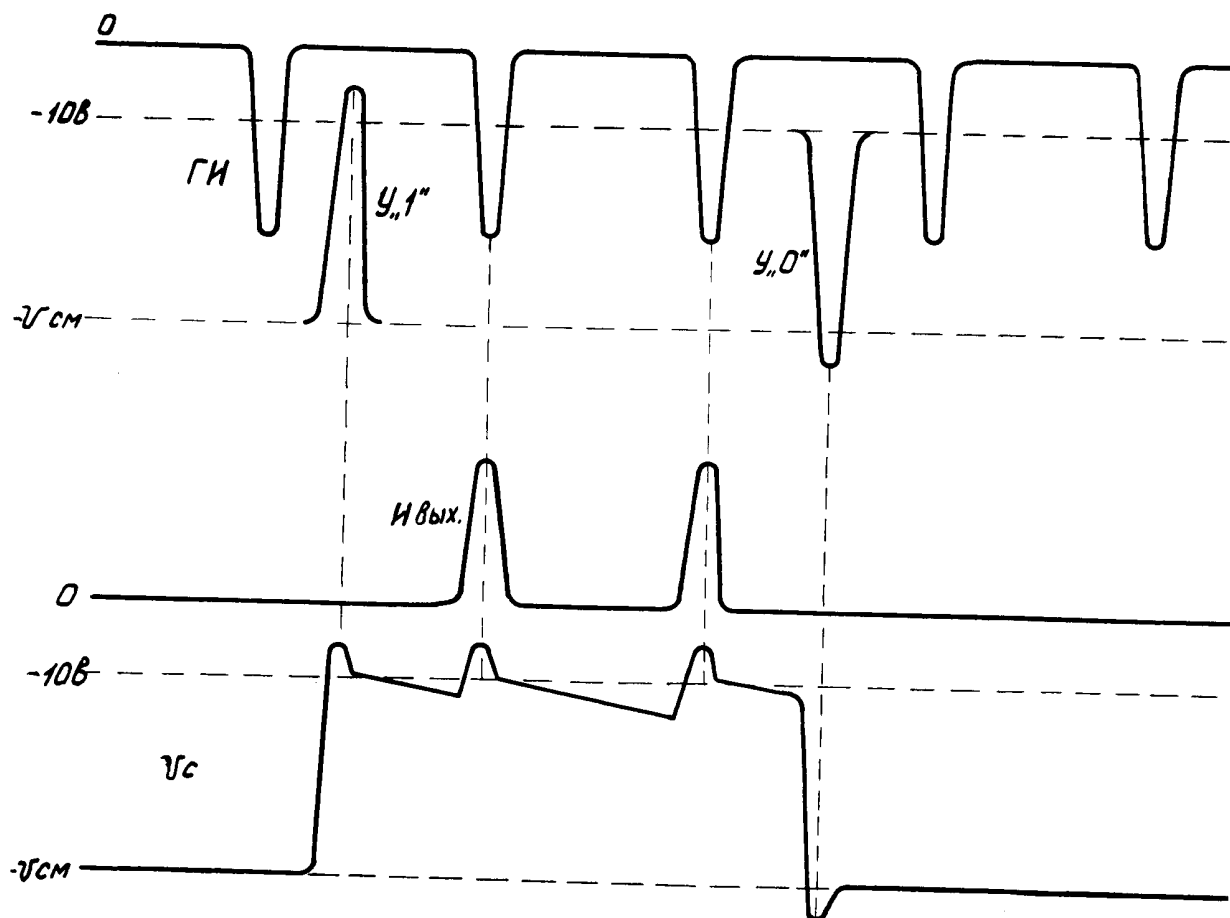


Fig. 3. The diagram of the dynamic trigger operation.

time between the two synchronising pulses, the capacitor discharge was insignificant. When the grid voltage is in the order of $-10v$ the decrease in the cathode voltage, at the expense of the synchronising impulses ($\Gamma И$), opens the valve and the synchronising impulses pass through it and the output transformer (T_p). Pulses arriving to the external circuit appear on the outer winding of the transformer. From the other winding of the transformer, positive pulses are applied to the valve grid through the resistance R_1 and the diode D_2 . These pulses maintain the charge of the capacitance. The winding is coupled to the reference voltage $-V_M$, so that when the trigger state corresponds to code "O" there would be no charging of

the capacitance through the D_2 diode. The surpassing of the voltage on the grid over and above $-10v$ is taken off by the $D''O$ diode just as at the operation of the $Y''I$ signal. Thanks to the sub-charging of the capacitor, the given trigger state may be maintained indefinitely.

To switch the trigger into a state corresponding to "O" code, a negative pulse is supplied to the $Y''O$ input. This pulse discharges the "C" capacitance through the $D''O$ diode to a voltage equalling V_{CM} . The surplus capacitor discharge below the $-V_{CM}$ voltage is taken up by a subcharge through the $D''I$ and $D''2$ diodes which are coupled with the reference voltage V_{CM} by corresponding windings of the transformers.

Besides this circuit, we have worked out a series of its modifications.

The dynamic trigger circuits with a storage capacitor have a number of important advantages compared with the trigger circuits with potential couplings. The most important ones are the following.

The circuits are very simple and operate on a coarse amplifier principle. They have a sufficiently high input resistance and a very low output resistance, which helps to bring them into agreement with simple logical circuits on the diodes.

The circuits operate on one valve only and do not present particularly great demands for the spread of its parameters and also for the spread of the values of the parameter's details.

Since the circuit operates in a voltage pulse condition with a great duty-ratio it is not difficult to obtain, with a small supply power expenditure, output signals with great front slopes and consequently a greater speed of circuits on small power, economic receiving-amplifying valves.

Compared with the known circuits of delay line dynamic triggers, or the elements of the type that were employed in the IBM-701 machine, the circuits worked out by us are favourably distinguished by the fact that they permit to perform a number of operations non-synchronously with synchronising pulses. In the main this is determined by the fact that the voltage on the dynamic trigger accumulating capacity may be employed to control the pulse-potential gate. Thanks to this the high speed of the dynamic triggers is determined not by the synchronising pulse frequency, but by the duration signals setting the "O" and the "I".

Fig. 4 gives an example of a simplified adder circuit with employment of dynamic triggers. The withdrawal of the code is performed simultaneously with the synchronising pulses (Γ). When the trigger is in an "O" code state there would be no signal on its output and the non-coincidence gate (1) will be open while the coincidence gate (3) will be closed. The code signal will arrive to the Y"I" input and will set the trigger into the "I" code position. If the trigger was in the "I" code position, its outgoing signal would close the non-coincidence gate (1) and open the coincidence gate (3). The code signal will come to the Y"O" input and will set the trigger into the "O" code position. Besides the signal will be applied through the delay line (3) to the ripple-through carry unit.

The impulse will pass via the above unit through the impulse-potential gates (4) which are controlled by the trigger accumulating capacitor voltage. When the trigger is in a state of "I" code the through carry pulse passes through the gate (4), is applied to the next position and establishes the overrun position into a state of "O" code. If the trigger is in a state of "O" code then the ripple-through carry pulse will not pass through the gate (4), the signal will be absent on its output, the non-coincidence gate (2) will be open and the ripple-through carry pulse will be applied to the Y"I" input of the trigger establishing it into the state of code "I".

To increase the speed of modern computers the multiplication time is of great importance. Our Institute has worked out a number of methods reducing the multiplication time. I shall dwell on one of them (suggested by senior engineer, V.S. Burtsev). According to this method the adder is supplied not by all the positions of the multiplicand at once on every adding tact, but in succession beginning with the first one.

The difference in the time of supplying two neighbouring positions equals the delay time in the ripple-through carry gate of one position. At the same time the partial product shift on the adder begins without waiting for the reception of the result of adding of all positions, immediately upon the formation of the result in the second position of the adder. Similarly to the adding, this shift is performed in turns, one position after another beginning with the smallest.

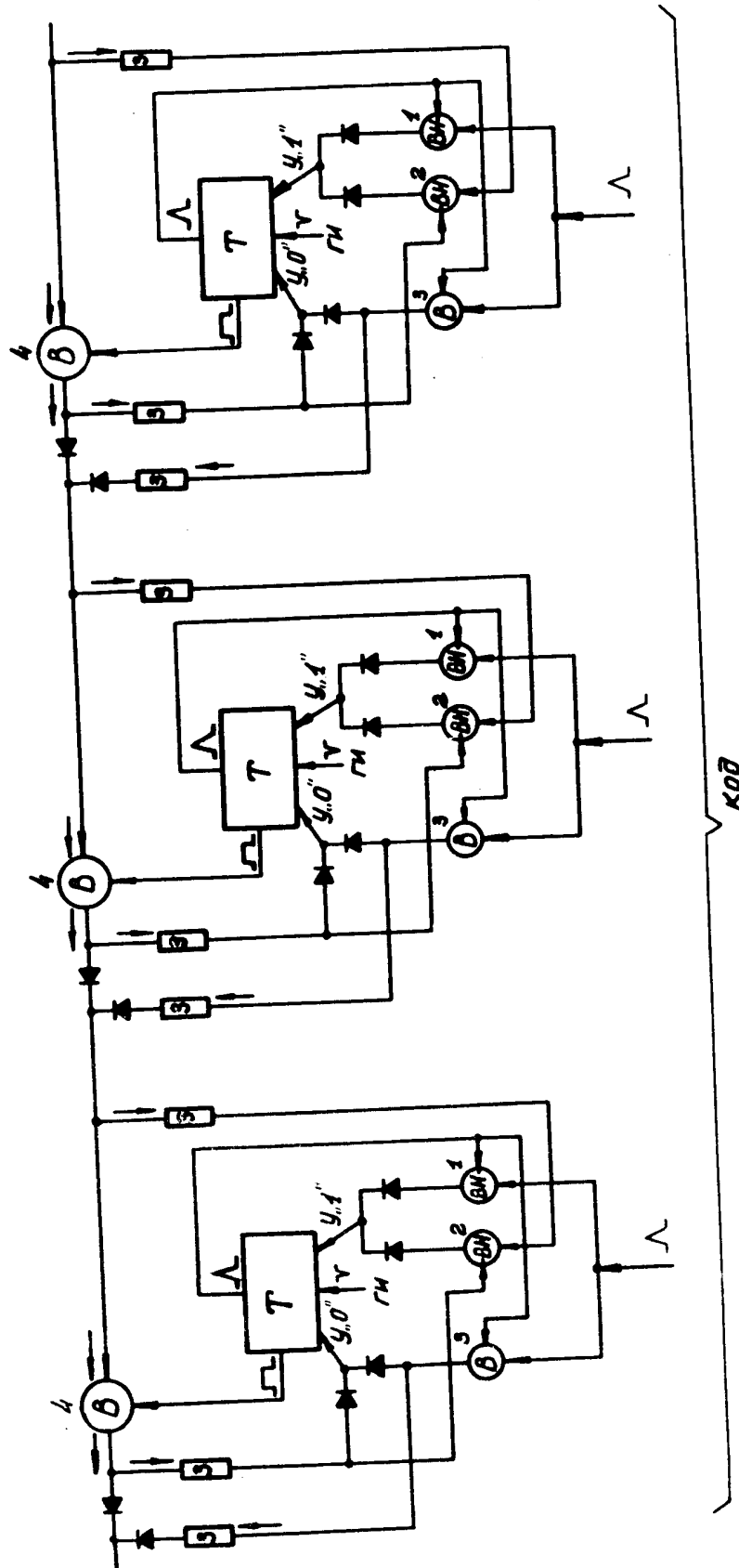


Fig. 4. 1) Simplified circuit of the address; 2) code.

Under such a regime of adding it proved to be possible to avoid two operations of the adder trigger per one adding tact. The time of one adding tact in this case is determined: 1. By the transfer time from the first position to the second - $\frac{T_n}{N}$, and 2. By the operational time of the second position of the adder - T_3 . Using a single tact circuit shift and one operation of the basic element per one shift will give us the multiplication time as being equal to

$$t_{\text{mult}} = \left(\frac{T_n}{N} + T_3\right) N + T_3 N + T_n + T_3.$$

In the usual multiplication circuit, the multiplication time will be

$$t_{\text{mult}} = (T_n + 2T_3)N + 2T_3N.$$

In these formulas:

- N is the number of binary positions;
- T_3 the adder's trigger operation time;
- T_n the ripple-through carry maximum time.

Fig. 5 shows the circuit of an arithmetic unit performing step-by-step adding with a single trigger operation.

Here the principle of operation is as follows: supplying the code to the adder of every position is conducted by individual control lines 1, 2, ..., 30, which are tappings of the delay line, its length being T_n . The second operation of the base element is avoided by closing its input by means of an "pulse-no pulse" circuit in case of coincidence of the code and carry pulses.

Fig. 6 shows a circuit guaranteeing a consequent shift of positions with a single operation of the trigger.

The circuit employs the same sequence shift principle as in adding. The double operation of the trigger is non-existent thanks to the elimination of the "Y" chain during the shift and the introduction of a paraphase carry chain.

Combining this method with others worked out by the Institute, the multiplication time may be substantially reduced without any noticeable increase in the instruments.

The Institute followed a new line in working out a memory made up of ferrite cores with a rectangular hysteresis loop. Instead of a usual matrix system we have worked out a system which we style as "Z type" (suggested by Prof. L. I. Gutenmacher). In this system the ampere-windings of the read

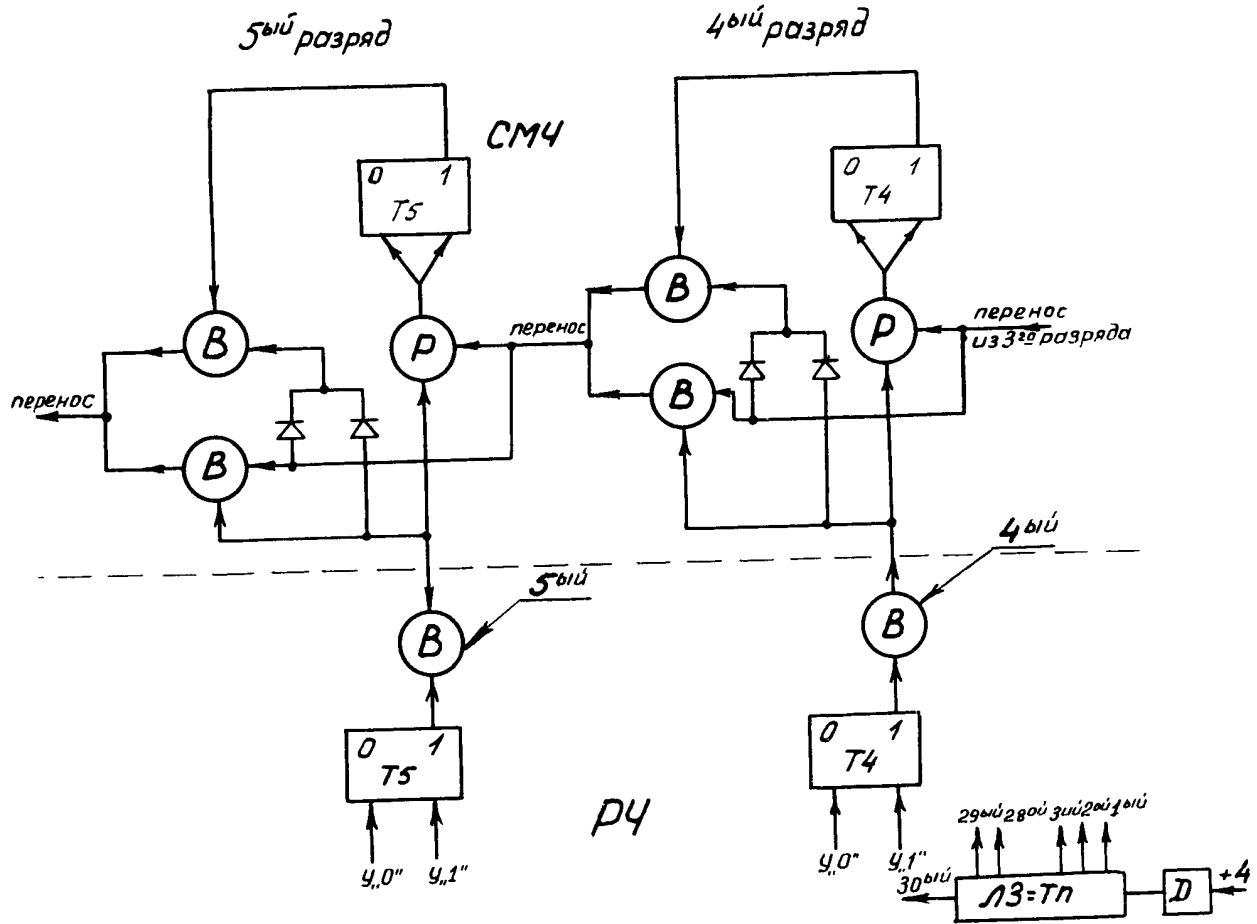


Fig. 5. 1) 5th position; 2) 4th position; 3) carry; 4) carry from the third position.

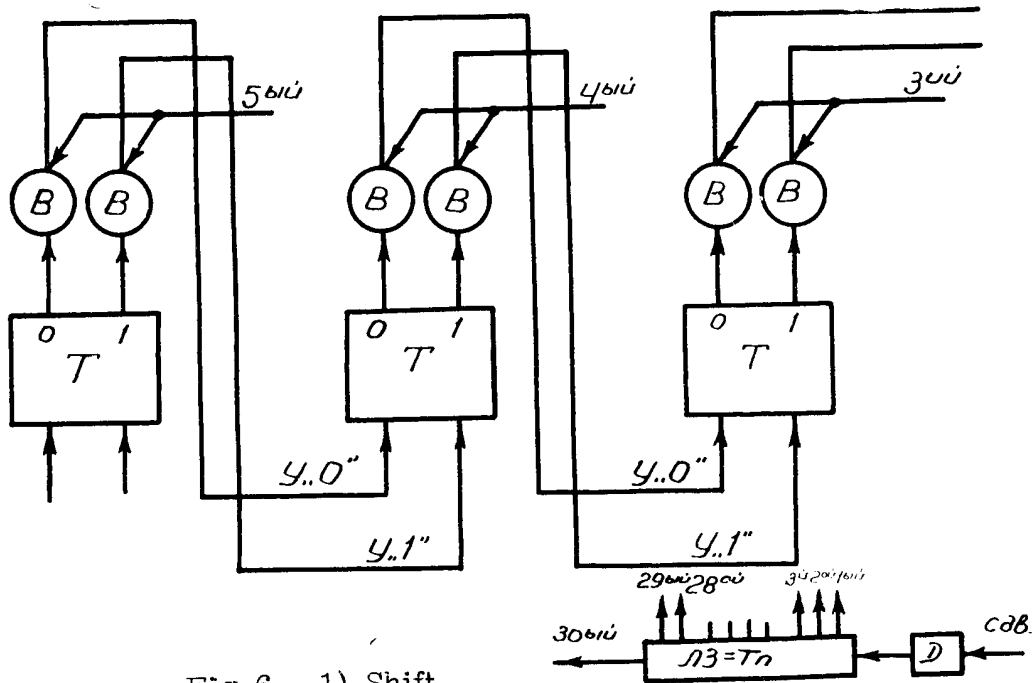


Fig. 6. 1) Shift.

off are applied simultaneously to all the positions of one number selected at a given moment. Apparently, in such units it is necessary to provide an individual gate for every number. These gates are executed in the form of coordinate transformers.

Let us analyse the circuit of number selection of the 16-number Z-type memory shown on Fig. 7. Here the registers PA_x and PA_y and accor-

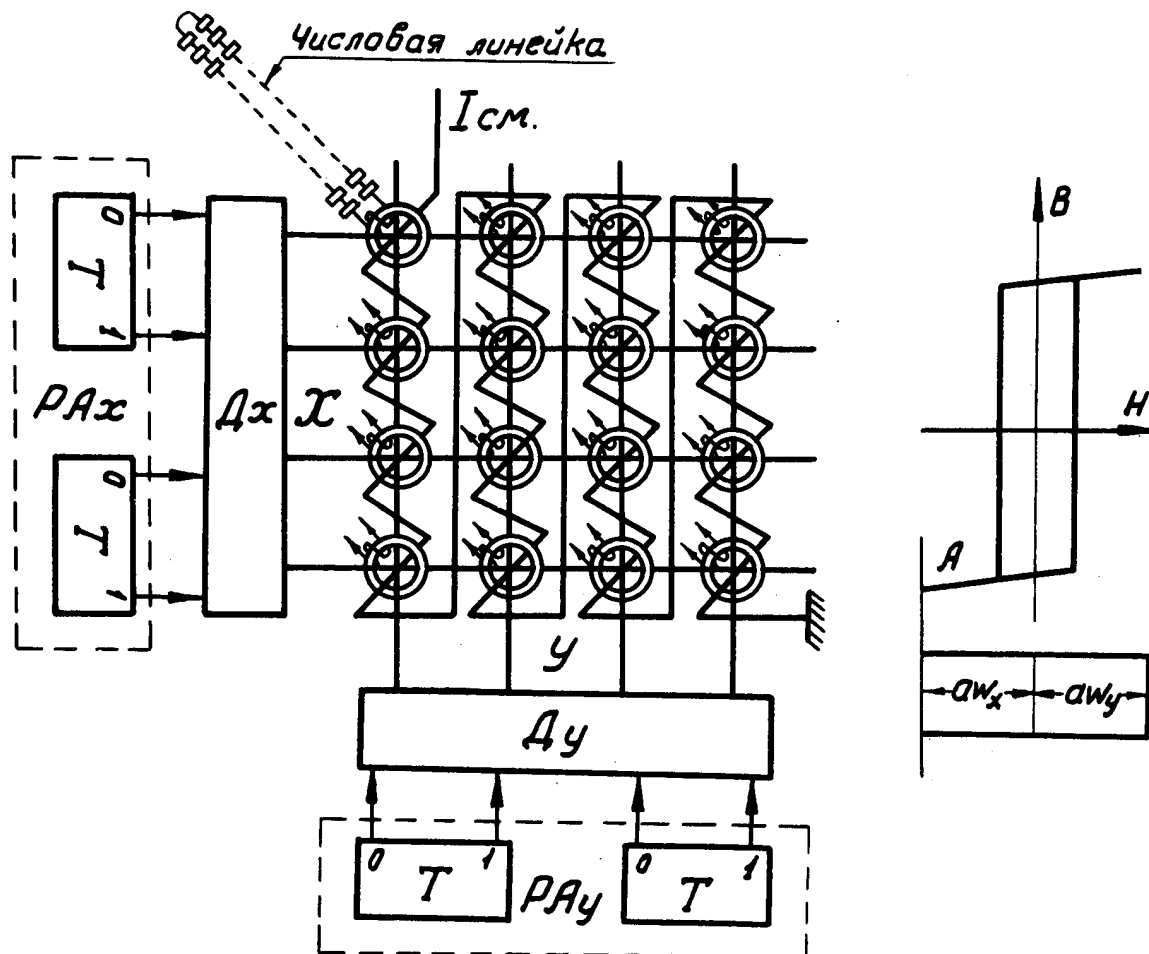


Fig. 7. 1) Digit line.

dingly the D_x and D_y decoders are employed for the transforming of the binary code address into current pulses along the selected wires, in the direction of X and Y. A common winding passes through all the coordinate transformers. The direct current flowing through this winding creates the magnetising ampere-windings, aw_{cm} , and thereby determines the operation point A. Thus every coordinate transformer is a gate which operates at a coincidence of the aw_x and aw_y ampere-windings. When the address code is

applied, the D_x and D_y decoders excite one wire each and on their intersection is a single selected gate which forms the current pulse for the memory cores of the given number in its output winding.

Those coordinate transformers which are influenced only by the aw_x or aw_y ampere-windings (we shall style them half-selected) will supply to their digit lines the disturbance current caused by the deflection of the hysteresis loopshape from that of rectangular.

To eliminate the disturbance currents from the half-selected transformers we use a disturbance compensation method worked out by professor L.I. Guttenmacher, which employs an additional compensation transformer. Fig. 8 shows a circuit of a coordinate transformer helping to

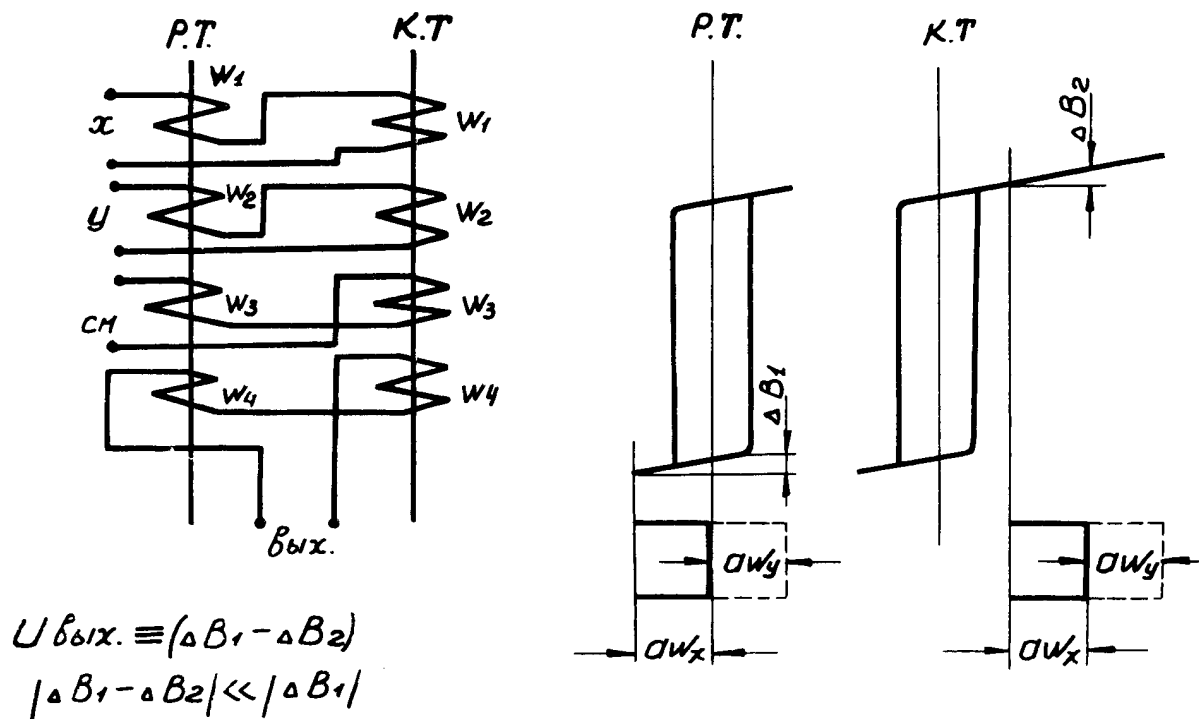


Fig. 8. Output.

understand the idea of this method. Two cores, working and compensating, have four windings each. The windings w_1 and w_2 of each core are coupled uniformly and serve for selection of the coordinate transformer. The windings of w_3 are coupled in a counterphase and the permanent bias current flows through them determining the working points of both cores. The output windings are also coupled into the counterphase. If affecting the coordinate transformer of the ampere-windings, only on one

coordinate on the output winding of the half-selected transformer there appears a signal equal to the signal difference of the working and the compensation cores. Now, if we select the winding data and the working points of both cores in an appropriate way, the difference of signals on the output of the half-selected coordinate transformer may be made very small and the disturbance would be practically eliminated.

Every coordinate transformer serves a line of memory cores (Fig. 9). It is known that the resistance offered by the core to the current which flows through its windings, changes depending on the remagnetizing or not of the core.

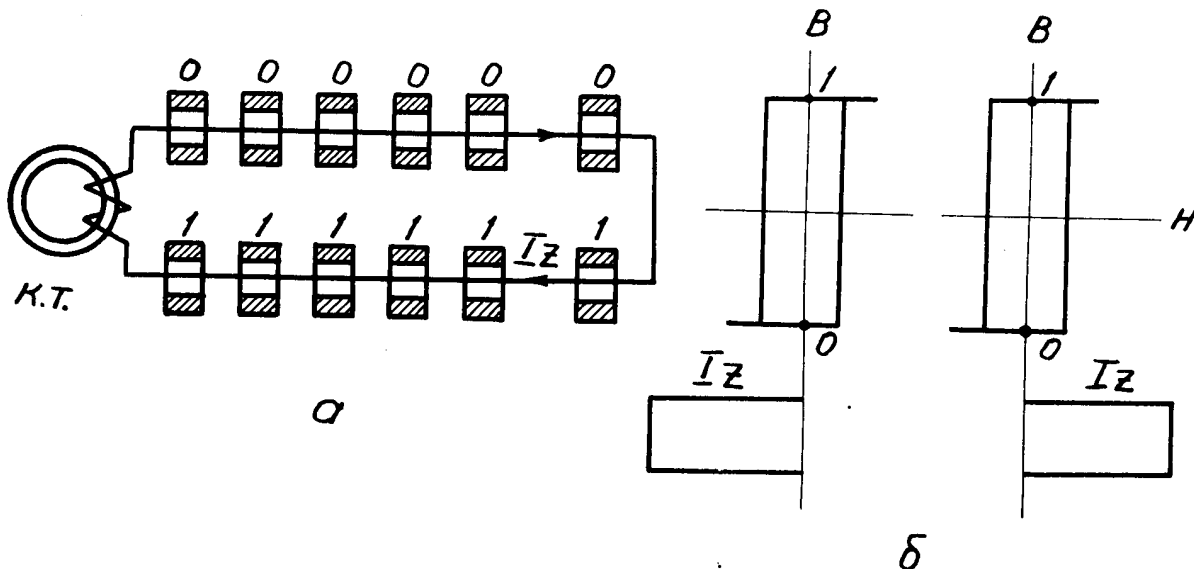


Fig. 9.

Therefore, the problem of stabilising the current which comes from the coordinate transformer into the digit line arises in the units of the Z type.

This problem was solved by two cores for every position - a working core and a stabilising one. During the read-off in any position one core is remagnetizing in a major cycle, while the other core, in a minor cycle maintaining the load stability.

The scheme of one memory line and the diagram of its operation are shown on Fig. 10. Every pair of cores besides the winding z has another two absolutely identical windings. One of these windings serves for writing the codes into the given position, while the other is used as a readout. Both windings are common for all the memory cores in the given position.

The coordinate transformer supplies to its line two half waves of the I_{z1} and I_{z2} currents. (Up to now we have been considering only the first half wave). The I_{z1} current has a big amplitude and performs the switching of one of the cores of the pair, i.e. the readout. The I_{z2} current is of a different polarity than the I_{z1} and is used for writing together with the current flowing through the write winding.

In order to perform the writing of the "I" code, a positive current pulse is applied on the write winding simultaneously with the I_{z2} current. For writing the "O" code it is necessary to supply a negative polarity pulse current through the write winding. Let us consider the diagram shown on Fig.10. It presents the sequence of the writing and readout currents.

After the I_{z1} pulse current at a moment t_1 the first core will be in a state of "O" while the second, in the state of "I". At a time moment t_2 , together with the I_{z2} current both cores are affected by the writing code "I" current (on Fig.10 it is etched).

Its polarity is such that its absolute value will add up with the I_{z2} current in the first core and will perform the writing of the "I" code in the first core. In the second core the writing current "I" will be subtracted from the I_{z2} current and the summed ampere-windings will be unable to change the state of the second core. If it is considered that during the switching from the "O" state into the "I" state every core induces a positive polarity EMF then during the writing "I" the first core will give a big positive pulse and the second, a small negative pulse. The sum of the EMF's of both cores will ensure a big positive pulse on the output winding. During the next reading at a time moment t_3 there will be a switching of the first core from the "I" state into the "O" state, while the second core will not change its "I" state. At this time the first core will give out a big negative signal while the second core will give out a small positive signal. As a result there will appear a big negative pulse on the output. During the writing of the "O" code during a time moment t_4 , the pulse current of the "O" writing will have a negative polarity and will be subtracted from the I_{z2} current of the first core. The summed ampere-windings in this process will be unable to change the magnetic state of

the first core while their summed ampere windings of the second core will transfer it into the "O" state. Apparently a big negative pulse will appear then at the output winding. Finally, in the readout of the "O" code the first core will not change its state, while the second will take the state of "I". At this moment there will appear a big positive pulse on the output. Regarding the pulse table on Fig.10, we see one of the most peculiar features of the given system, namely: the reading signals of the "I" code differ from the reading signals of the "O" code not by the amplitude, as is the case of matrix systems but by the sign.

As distinct from the matrix systems where the reliability of the readout is determined by the ratio $\frac{U_{"I" \text{ min}}}{U_{"O" \text{ max}}}$, in the given system the problem of readout is solved in a very simple manner. To bring this about it is sufficient to render the output chains sensitive to the pulses of only one polarity. It is easy to notice that the second core, besides the stabilisation load, performs in this case the disturbance compensation in reading out "O" which appears in the matrix circuits of the memory units with the utilisation of one core only.

The worked out system favourably differs from the usual matrix circuits by smaller requirements for the characteristics of the ferrite cores and their identity, and it does not require great stabilisation of the readout signals. The reliability of the operation improves greatly thanks to the absence of disturbances during the reading, a great selection speed and a greater amplitude of the output signal.

The memory of such a type for 1023 numbers is manufactured by the Institute and is adjusted as such. At present its adjustment is being made in operation jointly with the BESM, as a high speed operational memory.

The photoprinter employed in the BESM for the withdrawing of the obtained results at a rate of 200 numbers per second, has proved to be a reliable and highly efficient unit. Many millions of figures have been printed by it. However, the development of the film and the printing on paper delay the obtaining of the results which creates certain hindrances in the operation. Therefore the Institute is developing an electromechanical printing machine with a high operational speed.

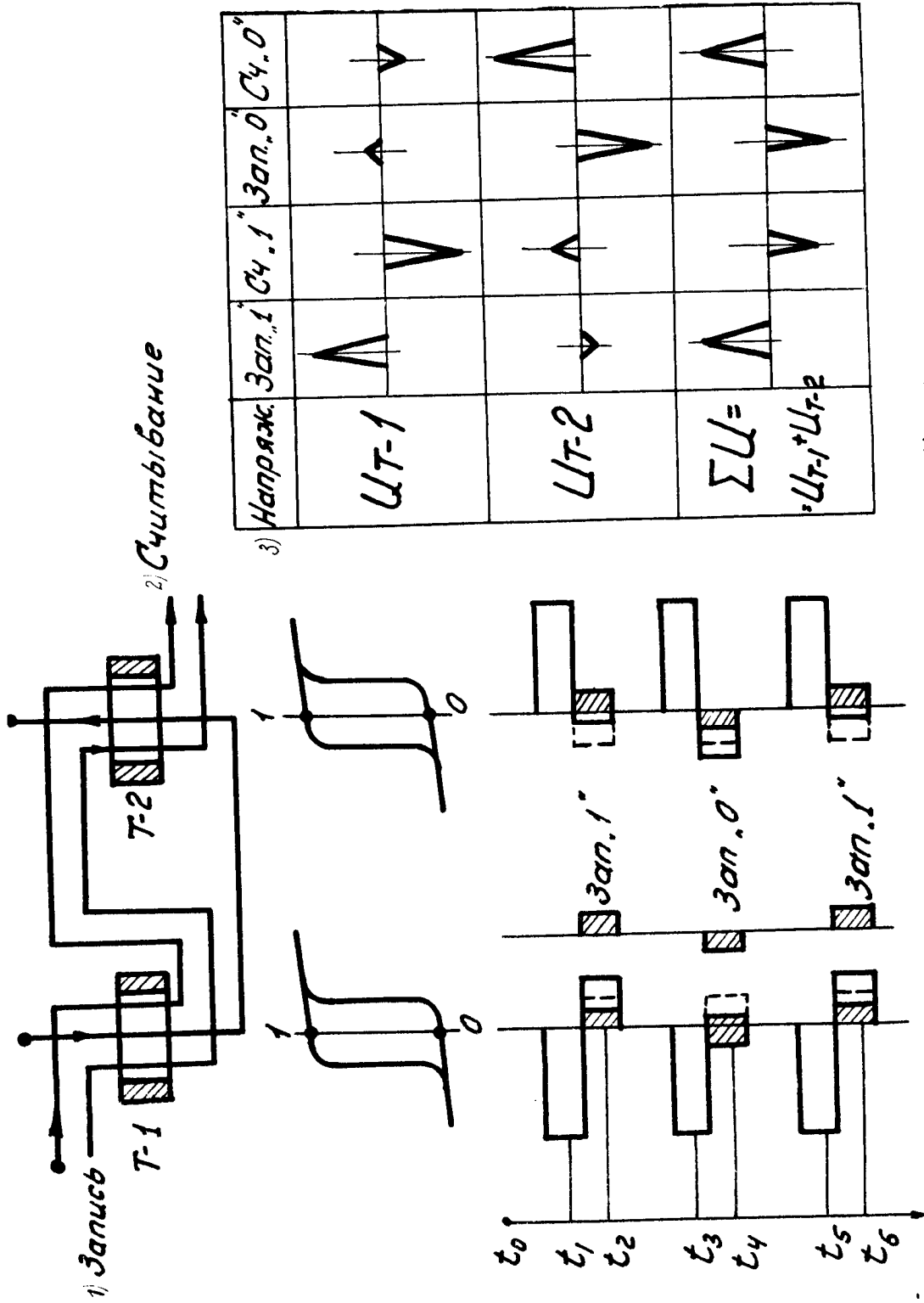


Fig. 10. 1) Writing; 2) readout; 3) voltage.

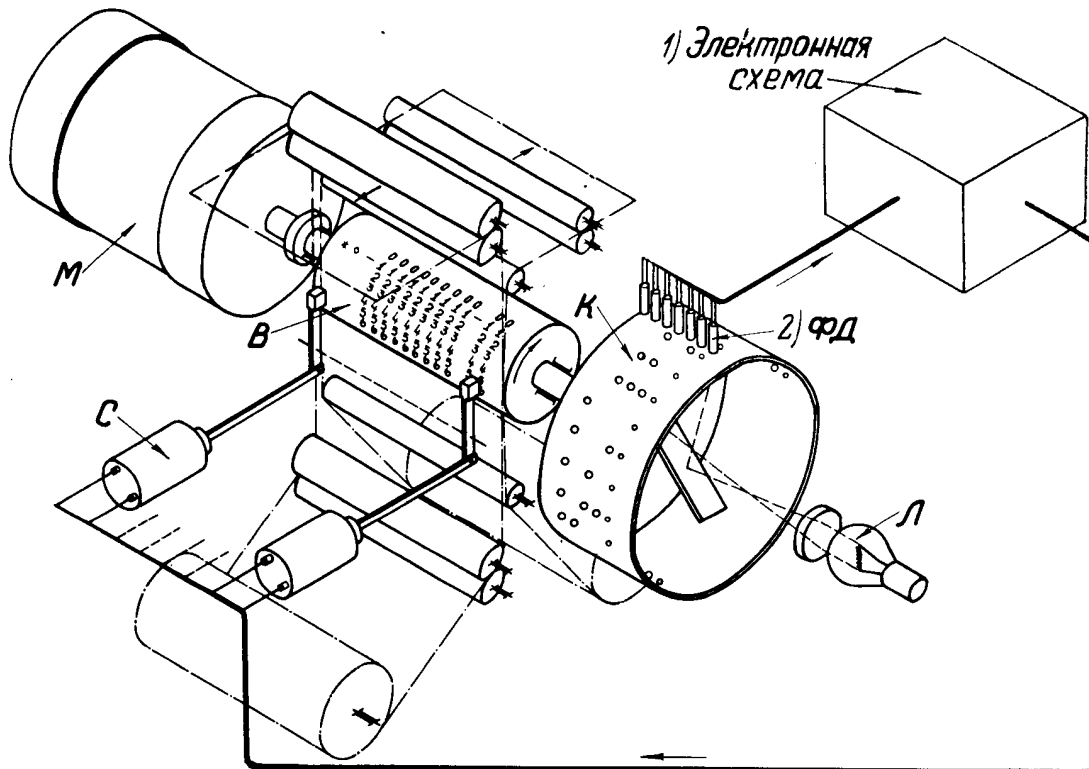


Fig. 11. 1) Electronic circuit; 2) FD.

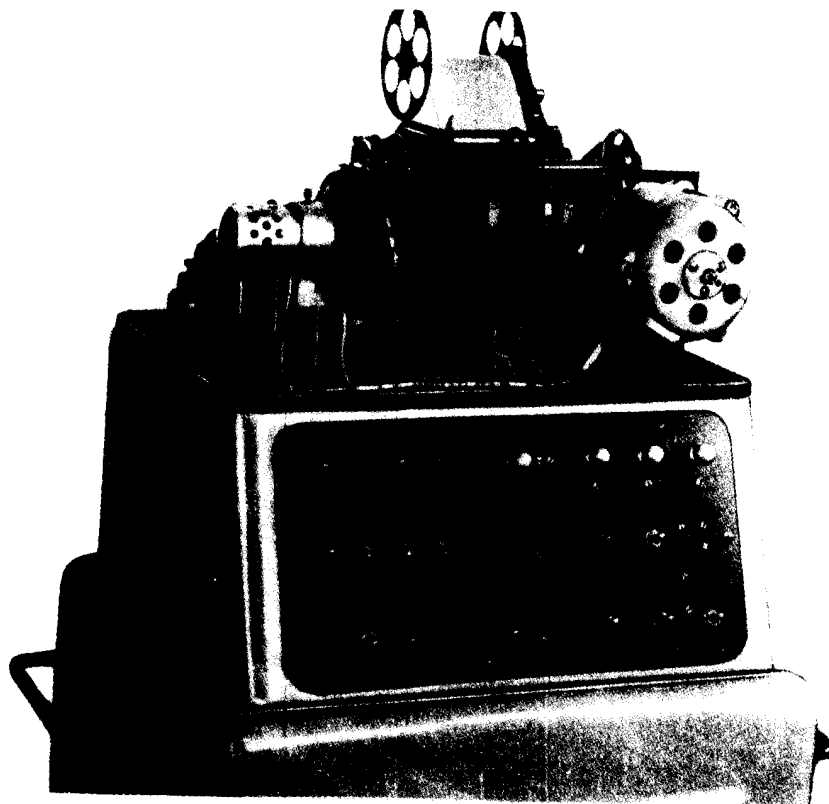


Fig. 12. The electromechanical printing machine.

Fig.11 shows the mechanical scheme of this machine. The motor M with a speed of 20 revolutions per second rotates the printing drum B on the circumference of which there are digits from 0 to 15 for every position. On the same axis there is mounted a cup-shape commutator K with holes corresponding to the digits from 0 to 15 in the binary system.

Before the drum, the supplying mechanism pulls up the paper and the carbon of the same size in a start-stop method. The lamp J lights the germanium photo-diodes FD at a moment when the holes in the commutator and the corresponding lines of digits on the printing drum come to the keys. The photo-diodes send the impulses of the code of the digits and the synchronising impulses to the electron circuit. The high speed solenoids C, which are controlled by the electronic circuit put in motion the keys which hit through the carbon on the paper and print the corresponding digits.

One third part of the printing drum is not occupied by digits. While the unoccupied by digits part of the drum B passes, the paper and the carbon are supplied one step forward by the controlled mechanism. In the time of one turn all the digits of one number are printed (the paper is motionless) and one step is made by the paper and the carbon (there is no printing this time). At present a sample of such a machine (Fig.12) is being checked in joint operation with the BESM. Fig.13 shows a sample of number printing on this machine at a rate of 20 lines per second.

```

- 1 5 1 5 1 5 1 5 - 1 1
- 1 5 1 5 1 5 1 5 - 1 1
- 1 5 1 5 1 5 1 5 - 1 1
- 1 5 1 5 1 5 1 5 - 1 1
- 1 5 1 5 1 5 1 5 - 1 1
- 1 5 1 5 1 5 1 5 - 1 1
- 1 5 1 5 1 5 1 5 - 1 1

```

Fig.13. A sample of printing done by the machine.

Besides the works presented in this report, the Institute conducts a number of other researches and experiments directed towards a further development of the computing technique and its application.

Sanitized Copy Approved for Release 2010/06/16 : CIA-RDP80T00246A045100330001-4

Page Denied

25X1

Sanitized Copy Approved for Release 2010/06/16 : CIA-RDP80T00246A045100330001-4