[redacted] 50X1

INFORMATION REPORT INFORMATION REPORT

CENTRAL INTELLIGENCE AGENCY

S-E-C-R-E-T PROCESSING COPY

[redacted] 50X1

COUNTRY	USSR	REPORT	[redacted]
SUBJECT	Control Theory and Nonlinear Mechanics (A REPORT BASED ON SOVIET PAPERS 1955-56)	DATE DISTR.	1 August 1958
		NO. PAGES	1
		REFERENCES	[redacted] 50X1-HUM
DATE OF INFO.	[redacted]		
PLACE & DATE ACQ.	[redacted]		

SOURCE EVALUATIONS ARE DEFINITIVE. APPRAISAL OF CONTENT IS TENTATIVE.

[redacted]

A [redacted] report [redacted] on "Control Theory and Nonlinear Mechanics" [redacted] The report is based on Soviet papers published on the subject during 1955 and 1956. 50X1-HUM

Distribution of Attachment:

[redacted] 50X1-HUM

OSI: Loan

[redacted] 50X1-HUM

S-E-C-R-E-T

[redacted] 50X1-HUM

STATE	X	ARMY	X	NAVY	X	AIR	X	FBI		AEC		OSI/EV	X		

(Note: Washington distribution indicated by "X"; Field distribution by "#".) 50X1-HUM

INFORMATION REPORT INFORMATION REPORT

[redacted]

[redacted]

50X1-HUM

Report On 1955-1956 Soviet Publications on Control Theory and Nonlinear Mechanics

[redacted]

50X1-HUM

Introduction

The following report refers to Soviet literature which appeared in 1955 and 1956. Although not all publications were available, it is assumed that the most important works are covered. Some were discussed in earlier reports and are thus not treated again here.

The report is divided as follows: The first part treats the works on control engineering. Then the results are compiled. The subsequent sections contain individual reports. A survey of the work is given by the bibliography, in which the publications connected with the 2nd Soviet Congress on Control Theory have been grouped together. The reports in this group are arranged as in the Congress report; they are cited by volume number and chapter (for example, Krug (1,V)). The other reports are listed alphabetically and are in the numerical order given in the references (for example, Gopp (16)). Section B of the first part deals with the position of nonlinear control engineering within the framework of nonlinear mechanics and acts as an introduction to the second part, which is devoted to nonlinear mechanics. It also contains a compilation of results, a bibliography, and individual reports.

[redacted]

50X1-HUM

The reporters have taken the liberty of treating only very briefly those partial problems which have already been treated in more detail in earlier reports. For this reason, the second method of Lyapunov and the Cypkin /Tsypkin/ theory of relay controllers have not been developed in detail. The present report, however, is in no place based directly on any previous report, and can thus be read independently.

Part I

A. Survey Of Publications On The Theory of Control Engineering

1. General Remarks

Although not exhaustive, the present report gives a reasonably complete picture of Soviet publications in the field during 1955 and 1956.

The proceedings of the 2d All-Soviet Congress On The Theory Of Automatic Control were published during the report period. The three-volume work contains 63 papers with discussions. Individual reports on some of the papers are given (see Section D1). A discussion of all the papers would have exceeded the scope of this work; nevertheless, the table of contents of the three volumes, which is given in full, should give an indication of what problems the Congress concerned itself with. It must be remembered that the Congress took place in 1953 and thus reflected the status of research work done up until 1952. No essential changes in the research areas have been noted since that time, except that perhaps the areas of "statistical methods" and "analog computers" have come more to the fore.

Since the Congress Report appeared in 1955, the number of works to be considered in the present report increased considerably. Even aside from the Congress Report itself, the number of works on the subject is larger than in previous years. In part this is due to the fact that Soviet publishers like to put out a whole series of articles on a single theme, many of which differ very little from one another.

The works of Tsypkin and Meyerov are an example. Increased Soviet interest in control engineering also resulted in increased publication. The most important journal on the subject, Avtomatika i Telemekhanika, appeared in 12 issues, rather than 6, during 1956; there are also other journals which were not available to this reporter. The quality of the publications has dropped somewhat, which is to be expected in view of the increase in volume.

The five papers which were read by the Soviet scientists at the International Conference on Control Engineering in Heidelberg are not, strictly speaking, to be considered USSR publications, since the Congress Report is to be published in Germany. They are nevertheless included here because they are so closely connected with the Soviet publications and also give a good indication of what subjects the Soviets chose to discuss at the International Congress. It should be mentioned here that the Soviets show a genuine desire to cooperate on an international level in matters pertaining to control engineering; the inclusion of foreign-language summaries in Avtomatika i Telemakhanika is evidence of this desire.

2. Mathematical Methods

Most of the works discussed in the report were limited to a presentation of the mathematical problem. A great deal of progress in theory can hardly be expected in the short period of only two years. Most of the works investigated known problems by means of known methods; only in a few cases were there new points of departure toward new developments. It must be remembered that many works are intended only to provide practical aids for actual application. The works are arranged according to linear and nonlinear problems, but it must be remembered that a strict division into these categories is not possible.

[redacted] 50X1-HUM

a) <u>Linear Methods</u>

The use of the term "linear methods" assumes that the behavior of a control system with respect to time can be described by a system of linear differential equations with constant coefficients or by means of a single differential equation equivalent to such a system and of a correspondingly higher order. This assumption is correct for most practical systems, at least in the first approximation. In general, the differential equation itself is not used, but rather the appropriate transmission function which is obtained as follows: let x_a and x_e be the output and input variables of the transmission system, and the differential equation which connects them have the form

$$D^n x_a + a_1 D^{n-1} x_a + \dots + a_n x_a = b_o D^m x_e + b_1 D^{m-1} x_e + \dots + b_m x_e,$$

where the operator D designates the derivative with respect to time, $D = \frac{d}{dt}$. If this equation is subjected to the Laplace transformation, a relationship is obtained (in the case of vanishing initial conditions) between the Laplace transform variables $\bar{x}_a$ and $\bar{x}_e$ in the form

$$\bar{x}_a = F(p)\,\bar{x}_e$$

F(p) is a rational function of the Laplace transform variables p with the denominator $p^n + a_1 p^{n-1} + \dots + a_n$ and the numerator $p^m + b_1 p^{m-1} + \dots + p_m$; it is called the transmission function or the operator of the element of a closed loop (transmission element) characterized by the differential equation. If a transmission system consists of several transmission elements, then a transmission function can be formed, according to fixed rules, from those of the individual elements, and since an open and a closed control loop can always be considered a transmission system, it is correct to speak of the transmission function of the closed and of the open control loop; the two have a simple relationship. If the open loop has the transmission function F(p), then that of the closed loop is

$$\frac{F(p)}{1 + F(p)}$$

A problem posed and solved by Kac (18) leads to a type of reversal of this relationship. He considers the transmission function of the controlled system as given and attempts to determine that of the controller in such a way that the transmission function of the total system has a given form. This problem can be solved by simple means; it is also related in a certain way to the problem of the quality of control.

The basic requirement of a technically feasible control loop is that of stable behavior. The stability behavior can be described conveniently with the aid of the transmission function. Stability exists when all the poles of the transmission function are located in the left half-plane of the complex variable p or, in other words, when the denominator of the transmission function, which is also called the characteristic polynomial of the corresponding differential equation, has only zeros in the left half-plane. Various (mathematically identical) stability criteria are known, which can be employed to solve the stability problem in any individual case. Theoretically, therefore, the problem is completely solved. The practical application of the criteria, to be sure, especially in high-order systems, is often rather difficult and laborious. For this reason, processes are suggested again and again, which are supposed to facilitate the computation, at least in special cases. Ostrovskiy (40) explains such methods for systems of the orders four, five, and six. Kislov (20) treats the stability problem for a control loop in which one transmission element contains two variable parameters and gives the construction of nomograms with which the stability ranges for the closed loop can be read off. A satisfactory stability criterion, which is much easier to apply than that of Hurwitz [Hurvits], is demonstrated by Dobronravov (13).

Some time ago, Ajzerman [Aytserman] and his pupils introduced the concept "structural stability." This concept can be defined as follows. Consider the time constants and the amplification factors of an arbitrary control loop to be arbitrarily variable. These parameters obviously can assume only positive values. In general, the system will be stable in the case of certain combinations of the parameter values, and unstable in the case of other combinations. If, however, even with an arbitrary choice of the variable parameter, it is not possible to make the system stable in any way at all, then it is designated "structurally unstable." For example, a system with the characteristic function $k(Tp - 1)$ is always unstable, no matter how the positive values T and k are chosen, whereas the system with the characteristic function $k_1(Tp - 1) + k_2$ is structurally stable, since the polynomial obtains a negative zero point in the case of suitable values of k_1 and k_2. This concept has proved to be of use for all sorts of theoretical and practical investigations and indicates the demand for criteria for structural stability or instability. Such criteria can actually be established according to the ratio of the total number of transmission elements to the numbers in which the various types of transmission elements occur. Some of these criteria are reported by Ajzerman [Aytserman] (2) and Gantmacher (1, VI); they involve propositions which can also be important for the practical synthesis of control loops. Up until now only simplest cases have been treated. (See Aytserman and Gantmacher, PMM 18, 103-122 [1954]. Further discussion of structural stability can also be found in the text by Aytserman, mentioned in the introduction.)

As already mentioned, it is possible to assign the transmission function of any complex control loop, if the transmission functions of the individual elements and the manner in which they are connected are known. The transmission function of a complex system depends on the parameters of the individual elements in a manner which is so difficult to survey that generally valid statements are difficult to arrive at. For this reason, the theory of multi-loop control systems is still in its infancy, except for basic fundamentals. Meerov /Meyerov/ has made some contributions to this theory. He is investigating multi-loop circuits with several controlled variables and assumes that the equations of motion have the form

$$k_i' M_i(p) \sum_{k=1}^{n} a_{ik} x_k + (D_i(p) M_i(p) + k_i) x_i = k_i' M_i(p) f_i \quad (i = 1, 2, \ldots, n).$$

x is the controlled variable, f the external effect, $D(p)$ and $M(p)$ the operators of the controlled system and the controller, k and k' amplification factors, and i refers to the values of the i^{th} partial system. The individual loops are thus coupled together in a definite manner. Meerov (33, 34; 1, VIII) then treats the following question: Under what assumptions does the total system remain stable, when the individual amplification factors k_i are allowed to increase arbitrarily? The necessary conditions refer to the location of the zeros of certain polynomials which are determined by the coefficients of the equations of motion. The order g_i of the polynomials $D_i(p)$ also plays a role: If individual values of g_i are greater than two, an instability occurs without fail when the coefficients k_i increase; this instability can be avoided by proper measures (introduction of differentiating effects or use of stabilizing elements).

A second problem posed and solved by Meerov (35) concerns the autonomy of the system under consideration: Is it (in cases of permanent stability) possible to make the individual partial loops of the system independent of one another by increasing the amplification factors? It is demonstrated that such a possibility exists when g_i is equal to or less than 2; otherwise, the system must be altered structurally in an appropriate way, just as in the case of the stability problem.

A favorite Soviet method of investigating stability is that of D-separation, developed by Nejmark /Neymark/; this method is also suitable for judging the quality of control. The basic idea of D-separation can be explained as follows: Let H(p) be the characteristic polynomial of a stable transmission system. Its zeros therefore all have negative real parts. If the zeros are shifted in any way in the complex plane, the stability of the transmission system will obviously disapper at the very moment the moveable zero positions reach the imaginary axis. If the polynomial H(p) (and thus its zeros) depends on certain parameters, then the "critical" parameter values, i.e., those at which the stability ceases to exist, are obtained with the aid of the equation $H(jw) = 0$. In this case, w is a real value. If this equation is separated into its real and imaginary parts, and if w is eliminated, a relationship between the parameters will be obtained, which can be interpreted geometrically as a limitation of the stability range in the space of the parameters and, correspondingly, is called the "limit of the D-separation." A knowledge of this affords the possibility of making statements on the stability behavior, especially on the dependence of stability on the parameters. The simplest conditions exist when the transmission function depends on only one single parameter, which is linear. In this case, the limit of the D-separation in the plane of this (complex) parameter can be established rather easily. It forms a curve closely related to the locus of the complex frequency response, as demonstrated by Meerov (2, III). Marjanovskij /Maryanovskiy/ (1, IX) used the method of D-separation in the investigation of a transmission system which is made up of n similar series-connected amplifiers with a transmission function of the order m and a feedback. He shows that the consideration of the characteristic equation of the total system, which is of the order $m \cdot n$, can be traced back to a system of n equations of the m^{th} order.

Gopp (16) presents an obvious modification of the method of D-separation. He defines the limit of the D_α-separation by means of the equation $H(\alpha + jw) = 0$; α is a real number (only negative α are of importance), w is a real parameter. With the D_α -separation, one can judge the degree of stability of a system which is defined as follows: If the zero points $p_1, \ldots, p_n$ of the characteristic polynomial are all in the left semiplane, then the real part of the zero point farthest to the right is designated the degree of stability. Popovskij /Popovskiy/ (1, VIII) determines, with D_α -separation, the degree of stability of an (idealized) system with two controlled variables, such as occur in practice in the case of the control of turbines and steam-boiler installations.

The latter works extend beyond the narrower scope of merely posing the problem, since they take the quality of the control into account, which is characteristic of the most recent developments in control engineering. Whereas the prime consideration formerly was on stability and the time response of the control action, the tendency now is to concentrate more on the control action itself. Naturally, the external interferences on the transmission or control system which play no part in the stability problem must likewise be taken into account.

From a mathematical point of view, this leads to the solution of a nonhomogeneous linear differential equation, which is simple enough theoretically, but which is not always so simple from the practical point of view. Moskvin (2, II) developed for this purpose an approximation method which makes use of matrix algebra. He substitutes for the differential equation an equivalent system of the first order of the form $x = Ax + f$ (A is a matrix with constant coefficients), represents the solution in closed form with the aid of the matrix e^{tA} and then derives an approximate expression. The method can, to a great extent, be put into schematic form and is thus well suited for treatment with computers. In the case of control problems, however, there will be many instances when the solution, in the case of arbitrary interference f, actually will not be carried out.

It is obviously sufficient to know the reactions of a linear system to certain pronounced interferences, which need be selected only so that any arbitrary interference can be produced from them by means of superimposition. The stationary reaction which occurs in the case of a harmonic effect of a given frequency w, or the reaction on an abrupt effect, is generally used. The synthesis of an arbitrary effect leads, in the first case, to its Fourier representation, and in the second case to the so-called Duhamel integral; thus it can be carried out according to known methods. With a sinusoidal input variable x_e = A sin wt, the output variable (in the stationary state) has the form

$$x_a = A\,F(jw)\sin wt \quad \text{or} \quad A\,\mathrm{Im}\,F(jw)e^{jw},$$

whereby F(p) is the transmission function. For the complex value F(jw) the Soviets use the designation "amplitude-phase characteristic," or "frequency characteristic." If the interference function is chosen as the unit surge which is defined by $z(t) = 0\ (t < 0)$, $z(t) = 1\ (t > 0)$, then the transfer function $\phi(t)$ is obtained as the reaction of the system. In Soviet terminology this is called the "time characteristic." If the symbol L is used for the Laplace transform, the relationship between F(p) and $\phi(t)$ can be represented by the formulas

$$L\left\{\phi(t)\right\} = \frac{F(p)}{p} \quad \text{and} \quad \phi(t) = L^{-1}\left\{\frac{F(p)}{p}\right\}$$

If the inverse of the Laplace transform is written as a complex integral, a simple connection between frequency response and transfer function (or between time characteristic and frequency characteristic) is obtained. Occasionally the impulse-transfer function k(t), thus the reaction of the system to a unit surge, is used in place of the function $\phi(t)$. It can be defined by the equations

$$L\left\{k(t)\right\} = f(p) \quad \text{and} \quad k(t) = L^{-1}\left\{F(p)\right\}$$

and, naturally, has sense only when F(p) is a genuinely fractional rational function.

[redacted]

The connection between $\phi(t)$ and $F(p)$ is used and investigated in various works. Kagan (18a) investigates the conditions under which the relationship

$$\phi(t) = \mathrm{Re}\,\frac{2}{\pi}\int_{0}^{\infty} F(jw)\,e^{jwt}\,\frac{dw}{tw}$$

is differentiated for t. Block (2;3) shows how conclusions concerning the transfer function $\phi(t)$ can be drawn from the behavior of the frequency response $F(jw)$. For example, he gives conditions for $F(jw)$ which guarantee a smooth response of ϕ (t). He further investigates the extent to which the real portion of $F(jw)$ alone, as compared with the portion of $F(jw)$ which stems from large w, participates in the behavior of $\phi(t)$; he further obtains estimates for the time from the beginning of the process to the first extreme of $\phi(t)$ and for similar values. Oveseevic (42) derives other dissimilar terms. He considers the transfer functions $\phi_1(t)$, $\phi_2(t)$ belonging to different transmission functions $F_1(p)$, $F_2(p)$ and estimates the expressions $[\phi_1 - \phi_2]$ or

$$\frac{1}{t}\int_{0}^{b}\left[\phi_1'(t) - \phi_2'(t)\right]^2 dt$$

with the aid of the expression

$$\int_{-\infty}^{+\infty}\left[F_1(jw) - F_2(jw)\right]^2 dw.$$

Voronov (2,II) gives a detailed survey of the various methods of determining the transfer function graphically from the locus of the frequency response, and vice versa. In most cases it involves graphical

[redacted]

integrations whereby the integral is approximated by means of a sum. Voronov also sketches methods of evaluating the parameters of the transmission system on the basis of experimentally determined transfer functions or frequency responses; these procedures are, to be sure, limited to low orders.

The use of the locus of the frequency response for determining, or at least for estimating, the transfer function is a recent development. Originally, these loci were used only to obtain information on the stability of the system with the aid of the Nyquist criterion. This criterion is also the subject of the brief note of Kuzovkov /Kutsovkov/ (26), who investigates the degenerate Nyquist diagram for transmission systems with one oscillating element (characteristic polynomial $T^2p^2 + 1$), as well as the note of Demcenko (12), which deals with the construction of the locus of a rational transmission function. More recently, the "logarithmic locus" is being used along with the actual Nyquist diagram. The logarithmic locus is obtained by a plotting of the logarithm of the value of $F(jw)$, multiplied by a suitable scale factor, as a function of w. This method of presentation is now appearing in the Soviet literature, and has been explained in detail in the textbooks of Aytserman and Popov. The note of Kuzovkov (26) also considers the logarithmic locus and its degeneration for a system with one oscillating element; in another work (27) he considers the connections between the limits of the D-separation and the logarithmic frequency characteristic, and gives algebraic-graphic criteria for the stability in the case of variable parameters.

As remarked above, both the use of the D-separation and the construction and evaluation of the transfer function are closely associated with the question of the quality of the control. An exact definition of the concepts "control quality" and "optimal control action" is very difficult to give. One relatively simple definition uses the "quadratic control surface:" the process $x(t)$ is optimal when the integral $\int_0 (x(t)^2 dt$ becomes a minimum. This definition is frequently used, and can be evaluated mathematically, at least in the case of low-order systems. It does not, to be sure, consider the fact that any practical process is subject to certain "secondary conditions" as a result of limitations on response and velocity, etc. For this reason, Feldbaum (14; 2,V) uses another explanation: He calls a control action optimal when it takes place in a minimum period of time and, to be sure, with the secondary condition that a certain expression, formed linearly with the aid of the values $x, \dot{x}, \ldots, x^{(n)}$ remains below a fixed limit. As Feldmann /Feldbaum/ shows, the evaluation of this condition can be carried out most conveniently by following the trajectories of the equations in the phase-space, and permits a designation of the trajectories of an "optimal" process. It is further shown that the optimal system, in this sense, can be obtained only through nonlinear correction terms (2,V). The Feldbaum definition of the control quality is used in a somewhat specialized form by Lerner (2,V). He is interested primarily in fast-acting control systems and gives a series of methods of arriving at the optimal process for actual engineering applications (instrumentation). He, like Feldbaum, employs the method of the phase-plane or phase-space and likewise considers nonlinear correction terms.

The just-mentioned definitions for the quality of control are not very well suited for practice. For this reason, more convenient means have been sought. Among these belongs the already-mentioned degree of stability which, as Meerov (36) explains, can be determined with sufficient accuracy with the aid of the limit of the D-separation. Kalis (1,IX) uses the time in which the process decays to half its value as a measure for the quality of an aperiodic process; in the case of oscillatory processes, he uses the ratio of two successive extreme values of the same sign, and shows how these values can be determined graphically. Ocasionally, an attempt is made to determine the control quality through the maximum value of [redacted] of the closed loop. Meerov (36) shows that this value is not always suitable for an adequate designation of the control action. This applies more or less for all those definitions of control quality which do not take the total course of the control action into account. The degree of stability, for example, depends only on one zero point (or a pair of zero points) of the characteristic polynomial, whereas the control action is influenced, naturally, by all zero points. The methods which are based on the estimations of the transfer function (Bloch (2;3); see above) all take the total /control/ action into account.

Generally, the control quality is investigated in connection with the synthesis of control loops. Originally, the mathematical treatment of control problems was essentially confined to the analysis of the control loop, i.e., limited to the setting up and discussion of the equations of motion. Recently, the synthesis has gained in importance. Even in the case of the design of the control system an attempt is made to attain certain desired dynamic properties, one of which, naturally, is stability. The mathematical task is to determine the transmission function of a control loop in such a way that the control action is stable and "optimal." Once the transmission function is determined, the computed control loop must be realized physically. Bloch (2,III) and Fateyev (2,IV) describe a series of procedures for the determination of the parameters of the system on the basis of specifications for the control quality. General requirements, such as maximal degree of stability, minimal control time, etc., are discussed. One simple synthesis problem is treated by Kac (18a) (see above), who attempts to determine the transmission function of the controller on the basis of certain conditions. Sokolov (2,III) points out that the synthesis should not be carried out solely from mathematical points of view, but that the possibility of realizing computed transmission function in practice and the design characteristics of the components must be taken into account; he gives several examples. Krasovskiy (2,III) treats a special problem of synthesis: In a multi-loop system, in which the transfer functions of all the elements except one are known, the unknown function must be determined in such a way that the total system will have predetermined properties. For this purpose he uses methods for the approximate calculation of the transfer function, when the transmission function is known. Lerner (2,V) and Feldbaum (2,V) treat those cases where the limitations of coordinates are taken into account in the synthesis of practical control loops. Solodovnikov (2,IV) points out that the synthesis of correction elements for the improvement of the control quality is mathematically equivalent to a problem from the Tchebycheff approximation theory; this involves the approximation of a given function $f(x)$ in an interval $a \leq x \leq b$ through a rational function $\frac{B(x)}{A(x)}$ with fixed degree of numerator and denominator is such a way that the deviation

$$\max \left[f(x) - \frac{B(x)}{A(x)} \right] \quad (a \leq x \leq b) \text{ becomes a minimum.}$$

The definitions of the control quality mentioned up until now and the corresponding problems of synthesis refer to an arbitrary interference and make no assumptions regarding the external effect. The tendency recently, however, has been not to give a general explanation of the concept of optimal action, but to explain it in regard to those interferences and effects which are considered typical for the particular relay system. The concept "optimal" is, in such cases, not defined absolutely, but in reference to the "environment" of the relay system; it no longer depends solely on the left side of the equations of motion, but also on the interference function which appears on the right side. The problem of synthesis is altered likewise, and now represents the attempt to compose the control loop in the light of the typical interference functions. Kulebakin (2,3) describes the principle of invariance as an effective principle for the realization of loops which are optimal for certain interferences. He explains as follows: The attempt is frequently made to compensate rapidly the interferences effecting the system by measuring not only the control deviation, but also the interference value directly. This latter measurement is fed into the control loop through a suitable transmission system, which causes an additional influence (so-called "imposition of interference values") on the control action. Kulebakin speaks of invariance when the influence takes place in such a way that the control value becomes independent of the interference value. Conditions can be established for this, and even carried out in instrumentation, obviously only for a certain class of interferences (for example, those which are characterized by a differential equation). Naturally, the principle can only be carried out approximately in practice. Similar investigations were carried out by Ivachneko (2,3), who also gives methods for computing the parameters.

The concept "typical interference" is an idealization. The interferences which actually occur in practice do not obey any devisable mathematical law, by which they can be predicted. They cannot be considered known until they have actually occurred. The concept "typical" interference can thus be interpreted only on the basis of a statistical approach, wherein the interference is to be looked upon as a value, the time behavior of which can be determined statistically through a large number of observations; it is therefore a probability function. Thus, even the coordinates of the control system, especially the control variable and the control deviation itself, now have the character of probability functions. Thus the theory of linear communications systems must be expanded, if the mathematical treatment of control loops which effect the probability values is to be built up. At least it would be well to elaborate the theory in the case of a number of problems of synthesis.

[redacted]

The theory of transmission and control systems which are subject to probability effects has been built up within the last ten years. It is based primarily on the investigations of N. Wiener and his associates. In the USSR, the theory was introduced by Solodovnikov, who, in keeping with the American works, carried out his own investigations and even wrote a textbook. He caused a number of Soviet "control mathematicians" to concern themselves with statistical methods. The work here is generally limited to a consideration of stationary probability functions. Such a function m(t) can be characterized by its autocorrelation function

$$R_m(\tau) = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{+T} m(t+\tau)m(t)dt$$

The functional value of $R_m(\tau)$ for $\tau = 0$ is also designated as its mean square value

$$\bar{m} = \lim \frac{1}{2\pi T} \int_{-T}^{+T} (x(t))^2 dt$$

In place of the autocorrelation function, the so-called spectral density $S_m(w)$ can be given. There exist the relationships

$$R(\tau) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} S(w)e^{jwt}dw, \quad S(w) = \int_{-\infty}^{+\infty} R(\tau)e^{-jw\tau} d\tau$$

Let the transmission system be given (see above) by its transmission function F(p) or its impulse-transfer function k(t). Let the input variable be the probability function m(t), and the corresponding output variable x(t). Between the spectral densities of these two functions there then exists the relationship

[redacted]

$$S_x(w) = \left[F(jw) \right]^2 S_m(w)$$

thus an analog of the known formula for the frequency response. For the correlations

$$R_x(\tau) = \int_{-\infty}^{+\infty} k(u) \int_{-\infty}^{+\infty} R_m(\tau + u - v) k(v)\, dv\, du$$

applies. The spectral density $S_\varepsilon(w)$ of the deviation $\varepsilon(t) = m(t) - x(t)$ is determined through the formula

$$S_\varepsilon(w) = \left[1 - F(jw) \right]^2 S_m(w).$$

If, besides m(t), there is another interference value n(t), which is independent of m(t), then we have the sum

$$S_\varepsilon(w) = \left[1 - F(jw) \right]^2 S_m(w) + \left[F(iw) \right]^2 S_m(w)$$

The mean square error can be computed from the equation

$$\overline{\varepsilon^2} = \frac{1}{2\pi} \int_{-\infty}^{+\infty} S_\varepsilon(w)\, dw$$

It plays a role in the definition of the control quality of a system which is subject to probability effects. If $\mu(t)$ designates that previously fixed time function which the system is supposed to simulate (thus the command variable), and if the actual input signal (which still contains secondary effects) is designated by $\phi(t)$, and the output variable by x(t), then the system can be called optimal when the expression

$$\overline{\varepsilon^2} = \lim \frac{1}{2T} \int_{-T}^{+T} (\mu(t) - x(t))^2 dt \quad \text{is a minimum}$$

[redacted]

This "condition of the minimal means square error" can be converted mathematically and leads to an integral equation for the impulse-transfer function k(t). It can also be put in the form

$$R_{\mu\varphi}(\tau) - \int_0^\infty R_\phi(\tau - u)k(u)du = 0 \quad (\tau > 0);$$

here R_ϕ is the autocorrelation of the input signal and $R_{\mu\varphi}$ is the cross correlation of the probability functions μ and ϕ.

The criterion of the minimal mean square error is used by Kurakin (25) to determine the optimal transmission function of a linear differentiator under the assumption that the input variable is a probability function of given spectral density superimposed by a white noise. (White noise is a probability function, the values of which are completely independent of one another; their autocorrelation is a δ-function, i.e., zero for all $\tau \neq 0$.) Kurakin gives an explicit expression for the transmission function. Solodovnikov and Batkov (45) have a more general approach. They consider systems shich are optimal not only in the case of an effect of a given form, but which can adapt to a variable effect in such a way that the operating state at any time is as close to the optimal as possible. Mathematically, such a system, which is called "self-adjusting" by its authors, can be characterized under somewhat general conditions by means of an integral equation.

The above-mentioned connection between the spectral densities of the input and output variables, which is completely analogous to the formula of the frequency response in cases of continuous effects, leads naturally to the use of the locus process even in the "statistical theory," provided only the spectral densities are introduced. Idelson (2,VII) uses the logarithmic frequency characteristic; his system is

[redacted]

effected by an input variable, one consitituent of which is a white noise, and the problem is to eliminate the noise as much as possible by means of an appropriate filter without impairing the required stability and reproduction quality. Cypkin [Tsypkin] (11) treats a pulse-control system (see below) which is influenced by a probability variable. Here, too, the introduction of the statistical values in the analysis brings no new viewpoints. Kazakov (19) shows that even nonlinear transmission systems can be made accessible to the statistical method with the aid of a suitable linearization formula; the method is treated briefly below.

The systems with lag time, which are described mathematically by differential equations of finite differences, have a preferred position among linear systems. Approximation methods are used for the most part for the investigation of stability. Lag-time problems have scarcely been treated during the report period. We might mention only the work of Mjasnikov [Myasnikov] (1,IX), who investigates with graphic methods a linear system of the third order (in addition to a nonlinear system). Occasionally, lag-time problems appear in other articles, but without special emphasis.

b) Pulsed systems

Formally speaking, it would be justified to include the so-called pulsed systems under the linear systems; a special treatment for such systems, however, is to be recommended. These systems involve a discontinuous transmission and control. Their characteristic component is the pulsing element, which transmits pulses in equidistant time intervals. According to the design of the system, either the amplitude or the width of the pulses is modified by the input variable. The effect of this type of influence is that all the variables do not vary continuously, but only at discrete points in time, and are thus defined for discrete arguments. The proper facilities for a mathematical treatment of such transmission systems are, therefore, not differential equations, but rather equations of finite differences. Several years ago, Cypkin (Tsypkin) pointed out that the analysis of a pulsed control system can be carried out in exactly the same manner as that of a continuous system; the chief aid is the "discrete" Laplace transform. An analog of the transmission function or the frequency response is obtained, and a relationship of the same form as in the continuous case is obtained between the transmission functions of the open and of the closed pulsed control loop; a stability criterion can be derived which is based on the course of the locus. One necessary condition is that the so-called "linear" part of the control loop be controlled, i.e., that the system be built of all the transmission elements with the exception of the pulsing element. In the works of Tsypkin which fall within the report period, this formalism is taken for granted. He treats several individual problems. In his work (6) he points out that his theory applies not only for square pulses (for which it was devised), but also for pulses of a more general form. In his notes (8) and (2,VI) he assumes that the pulse width is a nonlinear function of the input variable, and shows how the corresponding nonlinear equations of finite differences can be looked upon as recursion formulas and solved by a numerical method; the method can be illustrated to a great extent graphically. Moreover, the equations of finite differences are always nonlinear in the case of modified pulse widths, even when the modification itself is not linear. In the note (10) the transfer function of a pulsed relay system is used to introduce the concept of the statistical error, which can serve as a designation of the control action. A different characterization is given in the note (11), where the discrete analog of the control deviation and, in the case of the effect of probability functions, of the mean square error is taken into account. This work also considers questions of synthesis; the author is thinking primarily of pulse corrections, i.e., the introduction of additional pulse trains in order to improve the quality of the control.

50X1-HUM

50X1-HUM

Closely allied to the pulsed control systems is the use of digital computers, with which Tsypkin (9;11) is more closely concerned. Such an instrument converts the input variable, which is represented by a series of numbers, into an output variable, likewise represented by a series of numbers. The relationship between the input variable and the output variable, that is to say the connection between the two series of numbers, is expressed by a linear (or nonlinear) differential equation with constant coefficients; such a differential equation is formed in the same way as the equations of finite differences which occur in the case of pulsed control systems. For this reason, any digital computer can be assigned a "transmission function," obviously a transmission function for a discontinuous transmission, and then a control loop, into which a digital computer is incorporated, can be treated, theoretically, the same as a pulsed system, thus with the discrete Laplace transform and the formulas set up with it. Tsypkin (9) gives a report on the various possibilities of using digital devices in control systems and the circuits for the individual applications (element providing set-point adjustment, comparator or final control element). He shows how the quality of control can be improved with the aid of a digital device -- the instrument then works as a pulse corrector -- and how the influence of a lag time can be eliminated. Such a device is even useful when probability effects occur (Tsypkin (11)).

Relay-Control Systems

The simplest nonlinear control systems are the relay control systems, also called two-position action controllers or three-position action controllers. Along with the linear part, they contain the relay, the coordinates of which can assume only two or three values. For this reason, the control process is discontinuous. A transmission system with relays can be described mathematically by means of differential equations with sectionwise constant coefficients. In most of the practically important cases the differential equations

50X1-HUM

themselves are linear. The nonlinearity of the process thus rests on the fact that the description is given in various sectio[redacted] variable by means of equations with different constants. Other types of nonlinearities also have the same property, for example, linear systems in which an influence is exerted by dry friction or backlash. Thus the same means can be employed to analyze such systems as are used with the relay systems. In contrast to continuous controls, the relay systems have, in many cases, the advantage of greater simplicity and lower cost. They are especially suited for use where the processes to be controlled are relatively insensitive to sudden changes of power output. Thus the control of thermal processes represents a wide area of application for the relay systems. Kampe-Nemm's monograph explains the use of relay systems for temperature control. He investigates (18 b) certain methods of improving temperature controls, for example, the introduction of differentiating effects or the use of additional relays which change the release times.

Several ways of analyzing a relay system are given. The most obvious is the direct integration of the (linear) equations of motion in the individual sections, within which the coefficients at all times have the same value. The initial values of the unknown functions in a section are, in such a case, equal to the output values from the preceding section. This construction can be interpreted geometrically in the phase plane or in the phase space; to be sure, it gets out of hand when there are a rather large number of degrees of freedom. The method of point-transformation, the basic concept of which is supposed to be explained by means of a system which can be discussed in the phase plane, has been derived from the "integration in sections." Since the process in a relay control system always has the character of an undamped oscillation, the phase trajectories can, through the right choice of coordinates, be made to move around the null point. For this reason they will again and again strike a radial line through the null point and thereby produce a topological mapping of the radial line, which will be unequivocal and continuo[redacted] systems. When this mapping is mastered, all the essential statements regarding the course of the motion can be made. The same applies for systems of a higher order; the mappings of planes or hyperplanes must then be observed.

Alekseyev (1) investigates a temperature control with lag time with the aid of such point transformations. The occurrence of the lag time causes the motion to be described through differential equations of finite differences. The exact analysis of the process becomes very difficult; Questions of practical importances can be answered, however, with one simple point transformation which must be carried out with regard to the lag time in a multilayer phase plane. Those points which overlap themselves in the mapping determine periodic solutions of the equations of motion, and thus, simple or complex natural oscillations of the physical system. In simple cases, the formulas which define them can be set up explicitly; in complicated cases, at least the dependence of the natural oscillations on the parameter values can be investigated. Petrov and Rytkovskiy (43) treat a control system with lag time in an analogous way, i.e., in a multi-layer phase plane. They discuss the periodic solutions and show that, under certain conditions, a so-called temporary "slippage state" can occur, during which the relay does not release, but moves back and forth between the null point and the upper or lower maximum deflection. A study of such slippage states is of great importance for many practical problems. Dolgolenko (1, VII) has given conditions for its occurrence and estimates of its maximum duration.

Neymark (38) presents a very detailed theory of the relay systems. His chief aid is likewise the point transformation; he uses the analytical presentation of the variables of the system with the aid of the transfer function of the linear part. He studies primarily the various types of natural oscillations and their stability. In order to discuss stability, he forms the equations of the first approximation

[redacted]

and traces the problem back to a purely algebraic question, namely to the location of the zero positions of a certain polynomial. A shorter work of Neymark (39) is devoted to relay systems with lag time.

Another method of treating relay systems has been developed by Tsypkin. It makes use of the "characteristic of the relay system," which can easily be constructed, if the locus of the linear part and the characteristic curve of the relay element are known. With the aid of the relay characteristic, graphical procedures can be used to obtain information on the existence of periodic solutions and their properties, for example their stability behavior, and a study can be made of forced oscillations which can occur in relay systems as a result of nonlinearity. Tsypkin (5) gives a brief survey of his method, which he explained in detail in the monograph. Korolev (24) is making use of the characteristic to study a relay system with a retarded follow-up. He will have to elaborate the theory of Tsypkin somewhat.

A special form of relay-control systems are the so-called vibration controllers. They are based on the following phenomenon: If natural oscillations of very high frequency and small amplitude occur in a relay system, the system behaves toward a slowly varying input variable in a manner very similar to that of a linear system. We therefore find the expression "linearization of vibration." Pospelov (1, VII) portrays a series of technical realizations of this principle and various applications. Bernshteyn (1, VII) emphasized the use of the principle in the control of electrical machines, especially in the voltage regulation of small generators.

Finally, let us mention the works of V. V. Petrov (1, VII) and Fufayev (1,7). They are concerned with transmission systems or control loops containing two transmission elements with relay-like characteristic, which considerably complicates the picture. Nevertheless, the point-transformation method or the method of trajectories in the phase space can be used successfully to obtain general information on the possible

[redacted]

periodic solutions and their stability. Tal' (46) employs the point-transformation method in the treatment of a control loop without relays, but with dry friction.

Other Nonlinear Systems

Nonlinear transmission systems are those in which the relationship between input variable and output variable is expressed by means of a nonlinear function or a nonlinear differential equation. No general, universally applicable methods, such as can be given for linear systems, can be given here, however. Nonlinear problems require special methods from case to case; thus the survey of works on nonlinear control engineering presents a picture which is anything but uniform. Some of the work can be grouped, but it must be considered within the wider scope of nonlinear mechanics.

As mentioned already, some nonlinear systems, such as systems with dry friction, can be described, just like the relay systems, by differential equations with coefficients which are constant within steps only. The motions within them have the character of damped vibrations, and the position of rest is one possible position of equilibrium. Tal' (46) investigates the influence of dry friction on the operation of a centrifugal tachometer and gives conditions for the stability of the position of rest, which naturally depend on the constants of the friction and the friction of rest. He uses the point-transformation method.

Maslennikov (32) likewise treats a system with sectionwise constant coefficients; he poses the problem in an entirely different way, however. The equations of motion have the form $\dot{x} + ay = \phi$, $\dot{y} = kx\,\phi(x)$, where $\phi(x) = +1$ or -1. The problem is to select the function $\phi(x)$, i.e., the law according to which the control takes place, in such a way that the process occurs in a proscribed manner. It is thus a matter of the influence of a nonlinearity on the control action. The author gives formulas for the solution, but does not touch on any practical applications. A work by Ostrovskiy (41) belongs in the same category. It de [redacted] which is described by an

equation in the form.

[redacted] 50X1-HUM

$$x + F(x,x) + x = 0$$

[redacted] 50X1-HUM

The nonlinear function F is to be chosen (under certain secondary conditions) in such a way that the control action is optimal. It is shown that F must be equal to $k\dot{x}$; the constant k has two different values.

If the characteristic curves of the nonlinear transmission elements can not be composed at the outset from linear segments, it is often possible to use such a representation at least as an approximation. Kagan (2, II) studies and compares certain methods of replacing a non-linear characteristic curve with straight-line sections and shows that the transfer process can be constructed with the methods applicable in the linear case, for example, with the frequency method (Voronov (2,II), see above). A method developed by Bashkirov (2,II) is based on the same idea, thus on the replacement of the curve by a polygon series, but is to be carried out essentially in a graphical way. This permits the approximate construction of the transfer process in a system which is described by one or several differential equations of the type

$$\ddot{x} + a\dot{x} + bx = F(x, \dot{x}, t),$$

and thus takes interference functions into account. The method is simple and can be applied with quite good accuracy to current conditions.

Naumov (37) suggests another approximation-construction for the transfer process. It is based on the rewriting of the nonlinear differential equation into an integral equation of the Volterra type and the replacing of the integral by a suitable approximate expression. The integral equation then goes over into a finite system of algebraic equations and thus becomes accessible to treatment. Tsypkin (8) touches on similar recursively solvable nonlinear equations in posing certain problems in the theory of pulsed systems.)

[redacted] 50X1-HUM

[redacted] 50X1-HUM

The treatment of the transfer processes is, in most cases, quite laborious, even when limited to approxim [redacted] 50X1-HUM therefore, this treatment has been circumvented, and mere general statements on the connections between the time behavior of the system and its parameters, especially the nonlinearities, have been considered sufficient. The tendency has been to linearize immediately the stability behavior, in order to study the influence which the nonlinearity exerts on the stability behavior. Krinekiy (1,V) replaces the nonlinear right side F(x) of an equation having a linear left side by a linear function hx with suitable h, and investigates the resultant linear system for control quality, using the degree of stability (see above). Krug (1,V) works with the "description function," i.e., he uses the "harmonic linearization" or the "harmonic balance." He then investigates the linearized system with the D-separation and obtains information on natural oscillations and their stability. Kislov (20) shows how the linearized system can be discussed with the aid of certain nomograms; he is interested primarily in the stability ranges of the parameters (see above).

Tsypkin (6) gives a brief account on the calculation of the description function. Korolev (24) and Kagan (2,II) give examples to show that the method of harmonic linearization by no means always leads to useful results and thus must be used with a certain caution. Levitan (31) attempts to attain linearization by first replacing the nonlinear transmission term with a lag-time term and then using a differential equation of a higher order in place of the differential equation of finite differences. The method might perhaps lead to useful results only in special cases. Myasnikov (1, IX) considers a nonlinear system with lag time and linearizes it with the aid of the description function.

[redacted] 50X1-HUM

[redacted]

The second method of Lyapunov, used so frequently in former years, has been applied only in very few cases during the report period. The most important contribution, the monograph of Letov (28), actually does not belong in the report period, since it represents a compilation of earlier works of the author. Letov's presentation is centered around several modifications of a problem, stemming probably from Lur'ye, which can be expressed in general form as follows: Let a control loop be described by the equations

$$\dot{x}_i = \sum_{k=1}^{n} a_{ik} x_k - b_i y, \quad \dot{y} = f(s), \quad s = \sum_{\alpha=1}^{n} p_\alpha x_\alpha - ry;$$

$f(s)$ is the characteristic of the servomotor and s the command signal for the motor; r designates the feedback. The parameter p_α of the controller is supposed to be fixed in such a way that the system is stable with any input variables and especially with any form of the nonlinear function $f(s)$, insofar as the latter satisfies only the (obvious) conditions $f(0) = 0$, $sf(s) > 0$ for $s \neq 0$. In the solution of this problem the equation system is first converted into a "cononical form" by means of a linear transformation (for which there are several possibilities); then, in addition to the cononical equations, an appropriate Lyapunov function is constructed in various ways, according to whether the linear portion of the control loop is itself stable or not. The analysis of the conditions which the Lyapunov function has to satisfy leads to algebraic equations for the parameters. For general explanations, Letov uses a control loop, the linear portion of which has the form

$$\ddot{x} + a_1\dot{x} + a_2 x + a_3 y = 0,$$

without going into its technical significance. (The corresponding equation system has also been used as an example by other Soviet autho[redacted] interpreted as an automatic pilot for aircraft.) Numerical examples are not given.

A slight elaboration of Lur'ye's problem is given in a brief account (29) by Letov, in which the command signal s for the motor has an additional feedback term

$$s = p_1x_1 + \dots + p_nx_n - ry - N\dot{y}$$

A study is made of the influence of the variable N; the method used is the same. In the work (30) (and in the last chapter of 28) Letov concerns himself for the first time with nonstationary control loops. He considers the coefficients a_{ik} in the above equations as functions of time. As far as the rest is concerned, the treatment of the problem is the same. Formally, the method is the same; naturally, the Lyapunov theory must be drawn upon for nonautonomous systems. The analysis of the results becomes correspondingly more difficult. It is not yet possible to tell whether this elaboration of the presentation of the problem will be of importance in practice, especially since control loops with the time-dependent parameters have been treated only very little up until now.

In addition to the problem of Lur'ye, the Lyapunov method of presenting problems of control engineering is also used in the case of the Aytserman problem, which has been the basis of several articles during recent years. Generally, it is a case of a system with several nonlinearities which is described in the form

$$\dot{x}_i = \sum_{k=1}^{n} a_{ik}\, x_k + b_i f_i(x_1, \dots, x_n)$$

For the functions f_i, linear computations

$$\sum_{k=1}^{n} c'_{ik}\, x_k \leq f_i(x_1, \dots, x_n) \leq \sum_{k=1}^{n} c''_{ik}\, x_k$$

should be given in such a way that the system is stable in the case of any input variables, if only the nonlinear functions satisfy these computations. In general, the problem can [redacted] second method of Lyapunov, but leads to somewhat laborious computations. Razumichin /Ratsmikhin/ (44) investigates a special case, where the functions f_i agree for all i and depend on a linear connection of the x_k, and, through several deft manipulations, arrives to the point where the solution can be obtained by means of a quadratic equation.

Certain attempts to use the second method for an investigation of the quality of the control action (Letov (28), Plishkin, Avtomatika i Telemekhanika, 16, 19-26 (1955)) will only be mentioned here, since they have been reported previously.

If only the symmetrical natural oscillations of a relay system, i.e., the symmetrical solutions of the corresponding differential equations, are to be computed, a method developed earlier by Lur'ye, which leads to transcendental equations for the length of the periods and the initial values of the periodic solution can be used; the stability problem can be treated by going over to a "contiguous" solution and discussing the resultant system of equations of the first approximation. Trojckij /Troykiy/ (43) extends this method to relay systems with two servomotors. He is consistent in using matrix algebra to make the quite cumbersome mathematical treatment reasonably clear. Formally, the results correspond to those already obtained by Lur'ye and others in the case of the servomotor. Whether or not the practical utilization of these results can be recommended in regard to a saving of time seems questionable.

Kazakov treats a nonlinear problem of a special form (19). He investigates a nonlinear transmission system

$$Y(t) = f(X(t))$$

under the assumption that the input variable X(t) is a probability function. In order to arrive at a linearization, he considers the

[redacted]

variable X to be divided into two terms of a sum, $X = m_x + X_o$, the first of which represents the mean value (the mathematical expectation) of X, whereas the second is a probability function with the mean value 0. The equation for the approximate linearization is

$$Y(t) = f(X(t)) \approx k_o m_x + k_1 X_o \, .$$

The author gives various ways of calculating the constants k_o, k_1, which, naturally, assume that the nonlinear function is known. This "statistical linearization" has all sorts of advantages and could facilitate the treatment of nonlinear transmission systems within the framework of "statistical" control theory.

<u>The Analog Method</u>

For some time the USSR has been interested in the investigation of control engineering problems with the aid of analog computers, preferably electronic computers. The amount of literature on analog computers has increased in proportion with this interest. It is well known that electrical circuits, the time behavior of which is described directly by the given differential equation, can be formed with satisfactory accuracy for any linear differential equation with constant coefficients. This is possible, up to a certain point, even for differential equations with variable coefficients and for nonlinear equations. The most important component for such an analog system is the amplifier, a diode with an impedance. Depending on the switching and the type of impedance, it can be used as amplifier, integrator, or differentiator, i.e., the input voltage, which can be considered the input variable, is amplified, integrated, or differentiated. In the treatment of a concrete problem, the differential equation must first be simulated through the required components in a suitable circuit, and then the individual amplification factors must be chosen so that the model circuit reproduces the correct parameter values. In addition, of course, the accuracy of the reproduction, the errors, etc., must also be known.

The work of Kogan (22) gives a detailed view of the basic principles in the case of the solution of differential equations with the aid of analog computers and the mathematical and instrument skills necessary for such a method. Talancev /Talantsev/ (47) compares various possibilities for the switching of diodes and impedances and suggests modern circuits which, according to his data, afford considerable advantages. They have been used in the EMU-5 which was set up in 1955 at the Institute of Automatics and Telemechanics in Moscow. Gurov and his associates (17) give a report on this installation. Kogan (3,II) reports on earlier computers of the institute (EMU-1 to EMU-4). These instruments were built individually. G. M. Petrov (3,II) gives a survey of the analog computers which have been built in the USSR and on the experiences obtained with them.

The treatment of nonlinear differential equations with analog computers is of special interest, particularly for the purposes of control engineering. Kogan (21;23) gives a series of circuits for the realization of nonlinear relationships. He differentiates between the "typical" nonlinearities and the rest. Systems with typical nonlinearities are those which are represented by differential equations with section-wise constant coefficients, thus relay systems, systems with dry friction or backlash, etc. They can be simulated rather easily, for example through diode pairs in push-pull. Either approximations through polgyon series or special functional transformations, as described by Vitenberg (3,III), must be used in the case of nonlinearities with constant characteristics. Even components which simulate lag times are used (Gurov (3,III)).

- E N D -

50X1-HUM

Page Denied

Next 195 Page(s) In Document Denied