

FOR OFFICIAL USE ONLY

JPRS L/9759

28 May 1981

Translation

PRODUCTION TECHNOLOGY OF MICROELECTRONIC DEVICES

By

Ideya Aleksandrovna Malysheva

FBIS

FOREIGN BROADCAST INFORMATION SERVICE

FOR OFFICIAL USE ONLY

NOTE

JPRS publications contain information primarily from foreign newspapers, periodicals and books, but also from news agency transmissions and broadcasts. Materials from foreign-language sources are translated; those from English-language sources are transcribed or reprinted, with the original phrasing and other characteristics retained.

Headlines, editorial reports, and material enclosed in brackets [] are supplied by JPRS. Processing indicators such as [Text] or [Excerpt] in the first line of each item, or following the last line of a brief, indicate how the original information was processed. Where no processing indicator is given, the information was summarized or extracted.

Unfamiliar names rendered phonetically or transliterated are enclosed in parentheses. Words or names preceded by a question mark and enclosed in parentheses were not clear in the original but have been supplied as appropriate in context. Other unattributed parenthetical notes within the body of an item originate with the source. Times within items are as given by source.

The contents of this publication in no way represent the policies, views or attitudes of the U.S. Government.

COPYRIGHT LAWS AND REGULATIONS GOVERNING OWNERSHIP OF MATERIALS REPRODUCED HEREIN REQUIRE THAT DISSEMINATION OF THIS PUBLICATION BE RESTRICTED FOR OFFICIAL USE ONLY.

FOR OFFICIAL USE ONLY

JPRS L/9759

28 May 1981

PRODUCTION TECHNOLOGY OF MICROELECTRONIC DEVICES

Moscow TEKHOLOGIYA PROIZVODSTVA MIKROELEKTRONNYKH USTROYSTV in Russian 1980 (signed to press 15 Feb 80) pp 1-448

[Book by Ideya Aleksandrovna Malysheva, Izdatel'stvo Energiya, approved by the USSR Ministry of the Electronic Industry as a textbook in the specialty of microelectronic circuitry production at the middle specialized schools, 15,000 copies, 448 pages, UDC 621.3.049.77.002(075)]

CONTENTS

Foreword	1
Introduction	3
Chapter 1. General Description of Microcircuit Production	6
1-1. Basic Concepts	6
1-2. Classification and General Description of Microcircuits	8
1-3. Development of Microcircuit Technology, Production Forms and Records	11
1-4. Basic Processes of IC Production Technology	13
Test Questions and Assignments	18
Chapter 2. General Requirements on Microcircuit Production	21
2-1. General Requirements on the Technological Process	21
2-2. Requirements on Cleanness of Air and Climatic Parameters	22
2-3. Requirements on the Production Gases and Water	27
2-4. Basic Principles of Electron Vacuum Hygiene	30
2-5. Basic Production Features of Microcircuits	31
Test Questions and Assignments	32
Chapter 3. Manufacture of Semiconductor Plates and Dielectric Substrates for Microcircuits	34
3-1. General Information About the Machining of Semiconductors and Dielectrics for IC	34
3-2. Abrasive Cutting	36
3-3. Grinding and Polishing Billets for IC Structures	41
3-4. Control of the Plates and Substrates After Machining	44
Test Questions and Assignments	46

- a -

[I - USSR - F FOUO]

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

Chapter 4. Chemical Treatment and Cleaning of the Surface of Semiconductor Plates and Substrates	49
4-1. General Information	49
4-2. Methods of Liquid Treatment of Plates and Substrates	52
4-3. Intensification of the Cleaning Processes	55
4-4. Standard Processes of Cleaning Plates and Substrates	57
4-5. Drying Cleaning of Plates and Substrates	59
4-6. Quality Control of Surface Cleanness of Plates and Substrates	63
Test Questions and Assignments	65
Chapter 5. Contact Photolithography	68
5-1. Application and Essence of the Photolithography Process	68
5-2. Formation of the Photoresistive Layer	76
5-3. Formation of a Photoresistive Mask	81
5-4. Obtaining the Configuration of the Elements	84
5-5. Photomask Production Technology	89
5-6. Types of Rejects and Quality Control	93
Test Questions and Assignments	94
Chapter 6. Obtaining the Configuration of IC Film Elements Using Free Masks	97
6-1. Free Mask Method	97
6-2. Free Mask Production Technology for Thin-Film IC	99
6-3. Stenciling Method	103
Test Questions and Assignments	105
Chapter 7. New Lithography Techniques	106
7-1. Contactless Photolithography	106
7-2. X-Ray Lithography	108
7-3. Electron Lithography	112
Test Questions and Assignments	116
Chapter 8. Methods of Obtaining Thin Films	118
8-1. Method of Thermovacuum Deposition	118
8-2. Ion Bombardment Sputtering	122
8-3. Thermal Oxidation	128
8-4. Film Deposition from the Vapor-Gas Phase	130
8-5. Anodic Electrolytic Oxidation	134
8-6. Deposition of Metals from the Electrolytes and Solutions	136
Test Questions and Assignments	137
Chapter 9. Epitaxial Growth of Semiconductor Layers	140
9-1. Fundamentals of the Epitaxy Processes	140
9-2. Methods of Epitaxial Growth from the Vapor-Gas Phase	144
9-3. Other Methods of Epitaxy	147
9-4. Heteroepitaxy	150

- b -

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

9-5. Local Epitaxy	153
9-6. Alloying of Epitaxial Layers	153
9-7. Defects in the Epitaxial Layers	155
9-8. Epitaxial Layer Control	157
Test Questions and Assignments	160
 Chapter 10. High-Temperature Diffusion	 163
10-1. Fundamentals of the Method of High-Temperature Diffusion	163
10-2. Characteristic Features of Diffusion in Planar Technology	168
10-3. Methods of Achieving Diffusion	169
10-4. Defects and Control of the Diffusion Structures	176
Test Questions and Assignments	177
 Chapter 11. Ion Alloying and Other Methods of Obtaining Semiconductor Elements	 179
11-1. Fundamentals of the Method of Ion Alloying	179
11-2. Distribution of the Admixture Concentration in the Ion-Alloyed Layers	182
11-3. Equipment for Implementing the Ion Alloying Process	187
11-4. Advantages and Disadvantages of Ion Alloying	189
11-5. Other Methods of Obtaining Semiconductor Elements	191
Test Questions and Assignments	195
 Chapter 12. Metal Coating of Silicon Structures	 197
12-1. General Information	197
12-2. Single-Layer Aluminum Coating	198
12-3. Multilayer Metallization	203
12-4. Multilevel Metal Coating	205
12-5. Metal Coating of Mounted Active Elements	208
12-6. Metallization Defects and Quality Control	211
Test Questions and Assignments	214
 Chapter 13. Technological Processes of Manufacturing the Structures of Bipolar Microcircuits	 216
13-1. Engineering Design Features of Bipolar Microcircuits	216
13-2. Manufacturing Technology of the Structures of Bipolar IC With Insulation by the p-n Junction	219
13-3. Manufacturing Technology of Structures of Bipolar IC With Dielectric Insulation	223
13-4. Manufacturing Technology of the Structures of Bipolar IC With Combined Insulation	228
13-5. Manufacturing Technology of the Structures of Compatible IC	233
Test Questions and Assignments	233

- c -

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

Chapter 14. Technological Processes of Making MDS-IC Structures	235
14-1. Structural-Technological Features of MDS-IC	235
14-2. Manufacturing Technology for Structures of Thin-Oxide p-Channel MDS-IC	237
14-3. Manufacturing Technology of MTOS-IC	239
14-4. Manufacturing Technology of the Structures of MDS-IC With Fixed Gates	240
14-5. Manufacturing Technology of CMDS-IC Structures	243
14-6. Ways of Improving the Quality of MDS-IC	247
14-7. Protection of Semiconductor Structures	251
14-8. Dielectric Film Control	254
Test Questions and Assignments	256
Chapter 15. Manufacturing Technology of the Structures for Thin-Film Microcircuits	258
15-1. General Information	258
15-2. Standard Circuits and Basic Steps in the Manufacture of the Structures of Thin-Film Microcircuits	260
15-3. Technology for Manufacturing the Structures of Thin-Film Microcircuits with the Application of Free Masks	263
15-4. Manufacturing Technology of the Structures of Thin-Film Microcircuits with the Application of Photolithography	266
15-5. Manufacturing Technology of Structures of Tantalum Thin-Film Microcircuits	268
15-6. Manufacturing Technology of the Structures of Thin-Film Microcircuits Using the Beam Processing	271
15-7. Adjustment of the Rated Values and Shielding of the Film Elements	273
15-8. Quality Control in the Production of Structures for Thin-Film Microcircuits	274
Test Questions and Assignments	277
Chapter 16. Manufacturing Technology of the Structure of Thick-Film Hybrid IC	279
16-1. General Information	279
16-2. Basic Steps in the Manufacturing Technology of the Passive Part of the Structures of Thick-Film Hybrid IC	282
16-3. Quality Control When Manufacturing Thick-Film Microcircuits	288
Test Questions and Assignments	289
Chapter 17. Assembly of Microcircuits	291
17-1. Separation of the Plates and Substrates with the Finished Structures	291
17-2. Basic Assembly Methods	296
17-3. Installation of Crystals and Plates	298
17-4. Wiring	300
17-5. Wireless Melting	307
17-6. Quality Control in the Welding Process	310
Test Questions and Assignments	311

FOR OFFICIAL USE ONLY

Chapter 18. Encapsulation of Microcircuits	313
18-1. Microcircuit Cases	313
18-2. Methods of Sealing in a Case [Encapsulation]	316
18-3. Various Types of Encapsulation	321
18-4. Caseless Encapsulation	324
18-5. Quality Control of Encapsulation	327
18-6. Finishing Operations of the Manufacture of Microcircuits	329
Test Questions and Assignments	330
Chapter 19. Insurance of Production Efficiency and Quality of Microcircuits	332
19-1. Basic Areas and Problems of Microelectronics in the Current Five-Year Period	332
19-2. Insurance of the Efficiency and Quality on the Modern Level of Development of Microelectronics	333
19-3. Nondestructive Control and Improvement of Technology	336
Conclusion	340
Test Questions and Assignments	343
Bibliography	344
Subject Index	346
Annotation	351

- e -

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

[Text] FOREWORD

At this time microelectronics plays the primary role in the electronics industry. As it affects other branches, it to a significant degree determines the scientific-technical and social progress of our country as a whole.

As a result of the constant attention on the part of the party and government a powerful scientific research and industrial base for microelectronics has been created in the Soviet Union. An important condition of the successful production activity of the enterprises in this branch is constant improvement of production technology which is possible in the presence of well-trained personnel. Accordingly, in a number of the special schools divisions have been set up for training process engineers in the production of microelectronic devices (specialty number 0658). The primary finishing course taught in these divisions is the subject which is the title of this textbook.

The subject of "production technology of microelectronic devices" has a number of specific features which unquestionably distinguish it both from the subjects in the general curriculum and from many of the subjects of the specialized curricula. This difference is determined by the enormous difference in the methods and processes used and also the exceptionally high rates of development of production. A continuous flow of information introduces the necessity for constant supplementing and revision in the development of engineering theory. Accordingly, when writing the textbook the author tried to discover the essence and the peculiarities of the basic technological methods and processes of manufacturing various groups of microcircuits. Attention was given to the theoretical production problems and not the special process formulas and conditions which can be extremely varied or can change as the production facilities and techniques are improved. When discussing the material in connection with other subjects of the special curriculum, the author has tried to point out the numerous interrelations of production, including the choice of materials, optimal process techniques and conditions, coordination of operations in the process cycle, the equipment used, production quality control and efficiency.

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

In Chapters 1 and 2, a general description is presented of the basic steps in the manufacture of microcircuits of various engineering designs, and the general production requirements and characteristic features are presented. The process of manufacturing microcircuit structures is discussed in Chapters 3 and 16. Chapters 17 and 18 familiarize the reader with the assembly and encapsulation of microcircuits. Chapter 19 investigates ways of improving the quality of the microcircuits and their production efficiency on the modern level. The conclusion considers the prospects for further development of microelectronics.

The author expresses her appreciation to docent of the Department of "Radio-electronic Production Technology" of the MATI Institute [Moscow Aviation Technological Institute], Candidate of Technical Sciences Yu. G. Obichkin, editor, professor of the Experimental Department of the MIREA Institute, Doctor of Technical Sciences Yu. A. Kontsevov, professor, Doctor of Technical Sciences, Ya. A. Fedotov, Candidate of Technical Sciences F. P. Press, docent of the Microelectronics Department of the MIET Institute, Candidate of Technical Sciences O. V. Mitrofanov, Microelectronics Department Instructor of the MIFI Institute Candidate of Technical Sciences O. S. Bulatov, reviewer, instructor of the STEP L. A. Levkova, instructors of the MTEP L. I. Konstantinova, M. G. Krutyakova and A. V. Gaykovich for valuable comments and assistance rendered when working on this book.

The author will be grateful to the readers for suggestions and comments on the book which should be sent to the following address: 113114, Moscow, Shlyuzovaya nab. 10, izd-vo "Energia."

Ideya Aleksandrovna Malysheva

FOR OFFICIAL USE ONLY

INTRODUCTION

The goal of the course in "Production Technology of Microelectronic Devices" is to study the principles of the production technology of the components on the basis of which modern microminiature radioelectronic and computer equipment is built.

The entire period of development of the component base of electronics can be broken down into four generations: discrete electronics based on electrovacuum tubes, discrete electronics based on semiconductor devices, integrated microelectronics based on integrated microcircuits (IC), and integrated microelectronics based on functional microdevices (FMU).

In the first generation the role of the active components was played by various electrovacuum tubes. Resistors, capacitors, inductance coils, transformers, connectors, switches and other discrete radio parts were used as passive elements. Radioelectronic equipment (REA) was assembled from individual, discrete components which were mechanically attached to special panels and electrically connected to each other by wires using soldering or welding techniques. Later, printed circuits were developed which are more reliable and insure complete reproducibility of the REA parameters and relative ease of automation of production. The complexity of the technology of electrovacuum devices, their short service life, significant size and weight and high energy consumption provided the incentive for building new active components -- semiconductor devices.

The second generation of the component base of electronics appeared with the invention of the first transistor in 1948 by American scientists Bardin and Brattein. The invention of the first transistor was also preceded by a great deal of work by Soviet scientists.

In 1900, Russian scientist and inventor of the radio A. S. Popov used a semiconductor device for the first time in the world to detect radio signals. In 1922 O. V. Losev was the first to amplify and generate electric signals with the help of a semiconductor. Systematic study of semiconductors on a broad scale in our country was started in the 1930's by a group of scientists headed by Academician A. F. Ioffe. The electrical conductivity of semiconductors and the influence of impurities on it were studied by I. V. Kurchatov, V.P. Zhuze, M. S. Sominskiy and a number of other scientists. B. I. Davydov and D. I. Blokhintsev developed the theory of rectifying a concentration $n-n^+$ junction. Many other Soviet scientists and developers also made a significant contribution to the theory and practice of creating semiconductor devices.

FOR OFFICIAL USE ONLY

The first transistor was a point-contact transistor; its p-n junctions were obtained at the point of contact with the semiconductor of two electrolytically pointed wires. However, the point contacts turned out to be unstable and the devices based on them had low mechanical strength. In 1949-1950 the first alloy transistors were developed in which the p-n junctions were obtained on the basis of interaction of the liquid phase of a consumable electrode containing the alloying element with a solid semiconductor. The alloy transistors were distinguished by large junction areas, low reproducibility of the parameters and impossibility of obtaining base regions less than 10-microns wide. In 1953 transistors with diffusion junctions were introduced, the parameters of which are more reproducible, and the base width can be decreased to 0.2-0.3 microns.

By comparison with electrovacuum devices, semiconductor devices have smaller dimensions and less weight. They consume less energy and have greater efficiency, longer service life and higher reliability. The lower energy consumption and high efficiency also made it possible to reduce the size of the passive radio parts significantly. For example, in the last decade the volume of resistors has been decreased by 50-75 times, and the volume of capacitors, by 60-70 times. The necessity for fast processing of a large quantity of information required that the electrode gaps in the active components be reduced. Therefore the improvement of semiconductor devices and passive radio parts in parallel with them proceeded along the path of microminiaturization. This promoted a transition to compact printed circuitry and the creation of small assemblies in the form of columnar, two-dimensional and cordwood microminiature modules. The general-purpose cordwood microcircuits are 10-15 mm high, they weigh 5-7 grams and have a packing density of 5-20 components/cm³.

The third generation arose from the development and introduction of planar technology into semiconductor production in 1957-1958. Planar technology combined with film technology made it possible to convert to REA based on theoretically new components -- integrated microcircuits. The integrated microcircuit is a microelectronic component which performs a defined function of conversion and processing of a signal and has high packing density of electrically connected elements (or elements and components) and (or) crystals which, from the point of view of test requirements, acceptance, delivery, operation and maintenance is considered as a unit whole.

Integrated microcircuits contain an enormous number of elements equivalent to the previously used discrete elements. This decreases the number of connections and simplifies the process of assembling of REA, it significantly improves the reliability of the REA, it decreases size and weight, improves characteristics, expands the functional possibilities and significantly lowers the cost. The appearance of IC [integrated microcircuits] played a decisive role in the development of electronics, establishing the basis for a new phase -- microelectronics -- which is a qualitatively new phase of microminiaturization. However, the complication of the functions performed by the IC is achieved by increasing the degree of integration, which leads to complication of their structural design and has physical limitations.

The component base of the fourth generation will be functional microdevices and assemblies in the structure of which it is difficult or impossible to isolate the elements or the components equivalent to the traditional discrete elements of the

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

first two generations. Functional microelectronics differs theoretically from all preceding generations. As a result of integration of various bulk and surface physical phenomena, the structural complexity barrier has been overcome in functional microdevices. This makes it possible to build more reliable, power-intensive and economical components for microelectronic equipment.

The component base of the electronics is developing at fast, continuously growing rates. Each of the indicated generations, appearing at a defined point in time, has continued to develop in the most justified directions. The development of the component phase from generation to generation has proceeded in the direction of functional complication of the elements, improvement of reliability and surface life, decreasing size, weight and cost, simplification of technology and improvement of the parameters of REA. The most improved base components at the present time are the IC. Many functional microdevices are still in the development stage. Therefore the primary goal of the course in "production technology of microelectronic devices" is to study the production technology of integrated microcircuits.

FOR OFFICIAL USE ONLY

CHAPTER 1. GENERAL DESCRIPTION OF MICROCIRCUIT PRODUCTION

1-1. Basic Concepts

The Term "Tekhnologiya [Technology, Production Process, Production Engineering]"

This term was derived from two Greek words: "τέχνη" meaning art, proficiency, skill, and "λόγος" meaning science. In microcircuit production the term is used to designate the methods, procedures and means of processing raw materials and intermediate products and to designate the processing operations themselves and combinations of them in technological production processes.

The term technology also refers to the descriptions consisting of the production flow charts, instructions, and so on.

Technology also designates the scientific discipline which deals with the discovery of the essence of laws of mechanical, physical, chemical and other phenomena to improve the existing production processes and to develop and introduce the most efficient, new technological processes into production.

"Tekhnologicheskaya operatsiya" [Production Operation].

Microcircuit production includes a defined sequence of series and parallel processing operations, on the performance of which the finished products are gradually obtained from the raw materials. The production operation is the basic component part of the overall production or technological process of manufacturing the product. According to All-Union State Standard 3.1109-73, the production operation is the complete part of the production process performed at one work place. It is characterized by purposeful alteration of the initial object (initial billet or during subsequent operations, initial intermediate product) during the process of performing the successive operating procedures -- conversions. The term "tekhnologicheskii perekhod" [production conversion] designates the completed part of the production operation characterized by constancy of the tools used and the surfaces formed by the processing or joined during assembly. During the performance of each production operation qualitative changes take place in the processed objects: a semiconductor ingot is converted to plates, the shape of the plates and their surfaces become geometrically more precise; then the plates become more finished, and so on.

FOR OFFICIAL USE ONLY

Basic and auxiliary production operations are distinguished, depending on the processed object. If objects are processed which, on completion of the production process, are converted to the required product, then the processing is considered to be part of the basic operations. The production operations performed on auxiliary objects are called auxiliary operations. These include, for example, the preparation of pickling agents, cleaning of equipment and drying of gases. The auxiliary operations are an inseparable part of production and the quality of the finished IC depends no less on the quality of the performance of the auxiliary operations than on the performance of the basic operations.

The IC production technology includes a large number of different operations with respect to their physical-chemical mechanism performed in a vacuum, and gases, liquids and in the air. The number of operations reaches thousands or more in a number of cases. If it is considered that large (LSI) and superlarge (SLSI) integrated microcircuits contain from 10^2 to 10^4 - 10^5 elements, it is clear that it is very complicated to manufacture defect-free, reliable microcircuits. An effective means of improving the quality and the percentage yield of good microcircuits is to check the absence of defects after performance of the process operations. Therefore the technological production process includes quality control operations. The number of quality control operations is determined by the type and the complexity of the IC.

"Tekhnologicheskii protsess" [Technological Process]

In accordance with All-Union State Standard 3.1109-73 the technological process is the "part of the production process containing the operations of changing and subsequent definition of the state of the object of production." The technological process of making microcircuits contains the optimal number of production operations arranged in a defined sequence and providing for economically substantiated production of microcircuits of the given structural design with given electrophysical parameters. Well developed and checked-out standard technological processes which provide for the operating reliability of the integrated circuits are used in mass and series production. These technological processes are characterized by unity of content and sequence of the majority of production operations and conversions for a group of products with common structural attributes.

"Tekhnologicheskii metod" [Production Method]

It is possible to solve the same production problem in the manufacture of IC by using various methods. For example, in order to obtain a p-n junction it is possible to perform the production operation by the methods of diffusion, epitaxy, ionic alloying and so on. The methods of performing the production operations are characterized by a defined essence of the phenomena occurring during the processing and the defined set of operating procedures. The methods of performing the technological production processes are characterized by a defined set of matched production operations, each of which is performed by a defined optimal method in the given production process.

The production methods are classified by different attributes. The following processing methods are distinguished in accordance with the division of production into production sections: machining, chemical processing, heat treatment, photolithographic processing, epitaxy, elionics (treatment by electron and ion beams), assembly, encapsulation, and so on.

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

In accordance with the purpose of the technological processes, methods of entry control, cutting, grinding, polishing, degreasing, pickling, washing, drying, obtaining thin films, and so on are distinguished. The separation into the methods of film and semiconductor technology is generally recognized. This division reflects the history of production development and emphasizes the specific nature of each production area. However, in this classification it is necessary to consider that almost all methods of thin film technology historically manifested earlier, are also used in semiconductor production. A special role is played by the method of stenciling (in order to obtain only thin-film components) and the free mask method (to obtain a configuration, given dimensions and mutual arrangement of thin-film components). The methods of obtaining rectifying p-n and concentration n-n⁺, p-p⁺ junctions are specific to semiconductor production. All the remaining methods -- machining to obtain plates and substrates, chemical treatment and finishing of the surfaces of the plates and substrates, obtaining the given configuration, dimensions and arrangement of the components, obtaining thin films, assembly, encapsulation -- were used for making all groups of microcircuits. It is true that they are not all used to an identical degree, and each method has its specific nature in each specific case of production.

Group and individual production methods are distinguished. In the group methods not one specimen, but an entire lot are subjected to simultaneous processing. The processing of the lot under identical process conditions permits the dispersion of the parameters from specimen to specimen to be decreased and the efficiency of the technological process to be increased. Automation of the group processes significantly lowers the cost of the IC.

1-2. Classification and General Description of Microcircuits

Classification of Microcircuits

The IC are divided into three groups in accordance with the engineering design execution (Figure 1-1): semiconductor, hybrid, other (film, vacuum, ceramic, and so on).

A semiconductor microcircuit is an integrated microcircuit, all the elements and interelement connections of which are made in the body and on the surface of the semiconductor. A semiconductor IC, all the elements of which, both active and passive, are in the body of the semiconductor plate is called monolithic. A semiconductor IC, the active elements of which are made in the body of the semiconductor and the film passive elements, on the dielectric film obtained on the surface of the semiconductor plate, is called compatible.

A hybrid microcircuit is an integrated circuit, which, in addition to the elements, contains elements and (or) crystals. A component is a part of an integrated circuit which realizes the functions of any electroradio element, which can be isolated as an independent product from the point of view of the test requirements, acceptance, delivery, operation and maintenance. A multicrystal integrated circuit is a special case of a hybrid integrated circuit. The crystal of the IC is part of the semiconductor plate, in the body and on the surface of which the elements of the semiconductor microcircuit, the interelement connections and contact areas are formed.

FOR OFFICIAL USE ONLY

A film microcircuit is an integrated circuit, all the elements and interelement connections of which are executed in the form of films. Depending on the thickness of the applied films, the film IC are divided into thin-film and thick-film IC. If the film thickness is comparable to the free path length of the electrons in them, the IC is called thin-film. In thick-film IC the film thickness greatly exceeds the free path length of the electron. In practice, thin-film IC in the majority of cases have films no more than 1 micron thick and, correspondingly, thick-film IC, more than 1 micron thick.

From the definitions it is clear that the primary structural attributes for dividing the IC into groups are the substrate material and the structural design of the elements. Semiconductor plates in the body of which it is possible to make active

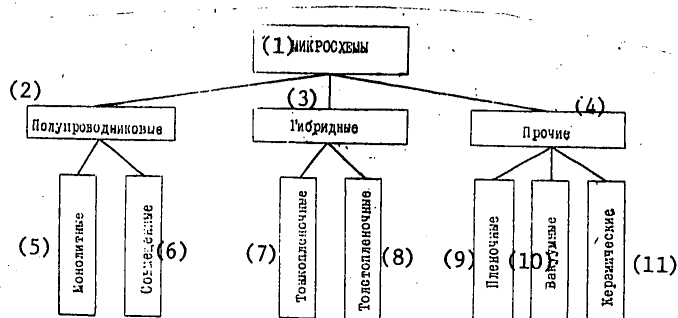


Figure 1-1. Classification of microcircuits by engineering design execution

Key:

- | | |
|-----------------------|---------------|
| 1. Microcircuits (IC) | 7. Thin-film |
| 2. Semiconductor | 8. Thick-film |
| 3. Hybrid | 9. Film |
| 4. Other | 10. Vacuum |
| 5. Monolithic | 11. Ceramic |
| 6. Compatible | |

elements or components, are provisionally called active, for they participate directly in the conversion and transmission of electric signals. Dielectric substrates are passive and, in addition, the functions of the structural base, on the surface of which the elements are formed or wired, can also play the role of heat transfer or part of the IC case. However, in some cases, for example, in film asymmetric strip IC in the microwave band, the dielectric substrate plays an active role.

The IC elements can be bulk or film. The active elements are primarily bulk, for not all film elements have the required characteristics as yet. The passive elements can also be both bulk and film.

FOR OFFICIAL USE ONLY

At the present time the monolithic and hybrid semiconductor IC have found the broadest application.

Characteristics of Semiconductor IC

The primary advantage of semiconductor IC is the possibility of obtaining high-quality active elements. However, bulk resistors and capacitors with a wide range of ratings and little dispersion of the parameters are difficult to obtain. This is explained by the fact that the parameters of passive elements are limited by the parameters of the semiconductor plate selected to obtain the required characteristics of the active elements. In addition, active and passive bulk elements are primarily made in one plate; therefore their parameters turn out to be inter-related. When developing IC additional difficulties also arise as a result of the strong temperature dependence of the resistances of the bulk resistors.

The cost of many operations in monolithic IC technology is comparatively high. The expenditures on materials and equipment require significant capital investments. This is why it is expedient to make monolithic IC only in mass production.

The deficiencies of monolithic IC have served as the reason for developing compatible IC for which the processes of creating active and passive elements are not simultaneous. Film resistors and capacitors have a wide range of ratings, little dispersion of the values of the parameters, low temperature coefficients, small stray couplings, and they are distinguished by high linearity and radiation resistance. However, the cost of compatible IC is higher than the cost of monolithic IC. Compatible IC are the most prospective for creating special microelectronic equipment.

However, in spite of the enumerated deficiencies, the monolithic IC have the highest degree of integration.

Characteristics of Hybrid Integrated Microcircuits (HIC)

The advantages of HIC consist in the possibility of combining the advantages of film passive elements with the high electrical properties of active semiconductor components. The active components are made in a separate production cycle on semiconductor plates in an independent case or caseless execution, and then they are mounted on the passive part of the IC. Sometimes inductance coils, capacitors, and so on are used as the mounted components. In the HIC it is possible to adjust the elements to the rating and replace the mounted components. The elements and components are on the dielectric substrate, which insures fewer stray couplings between them, a wider range of ratings, higher reproducibility and stability of the parameters. The passive part of the HIC can be made in the form of thin or thick films. The latter offer the possibility of obtaining resistors with high resistance and ratings, but it is difficult to obtain capacitors with high capacitances. Thick-film HIC technology is relatively simple; therefore their cost is appreciably lower.

The most serious deficiency of HIC is their comparatively low degree of integration. However, they require smaller capital investments for design and manufacture.

FOR OFFICIAL USE ONLY

1-3. Development of Microcircuit Technology, Production Forms and Records

Special Features of IC Development

The development of new IC differs significantly from the development of traditional electronic circuits. This is explained by the fact that the parameters of the individual elements and the operating characteristics of the IC are determined not only by the precision of the configuration, the dimensions of the elements and their mutual arrangement, but also the choice of the manufacturing process, where deviations from the given process conditions lead to deviations from the given topology. In addition, there is always an interrelation between the IC elements caused by their interaction and spurious effects which must be considered.

As a result of the interdependence of the structural design and technology in IC design the initial data, along with the external structural, electrical and operating data, are data reflecting the possibilities and limitations of technology. They determine the optimal design of the IC. Thus, when developing new IC the concepts of design development and technology development are closely interrelated, and the defining role goes to technology.

Technical Assignment for IC

The technical assignment (TA) determines the structural, electrical, operating and reliability requirements on IC which are formulated by the developers of the micro-electronic equipment -- the IC users. In turn, the IC developers estimate the possibility of satisfying these requirements, analyzing the information with respect to analogous developments and the modern state of the art.

Basic Steps in the Development of IC Manufacturing Technology

The development of new IC manufacturing technology takes place in several steps. The research is the preparatory step which determines the theoretical possibility of creating new IC based on calculation and design of topology, choice of initial materials and performance of experiments to select the possible methods of manufacture. Scientific research work (SRW) is performed to study various versions of technology and choose the optimal version. In the experimental design phase (EDW) the selection and practical implementation of the final optimal version of design and technology are realized, new equipment is manufactured, and all preparations are made to introduce the IC into experimental or series production. All of the production forms and records for the microcircuit must be ready at the time of introduction into production.

One of the basic goals when developing IC is maximum reduction of the elapsed time from compiling the technical assignment to introduction of the product into production. For acceleration of the entire IC development cycle, auxiliary production operations (PO) are performed in parallel with the basic steps.

IC Production Forms and Records

The production forms and records for IC consist of the technical specifications (Specs), the design and production documents.

FOR OFFICIAL USE ONLY

The technical specifications for IC make up the document reflecting the primary requirements on the output parameters and the conditions for storage, operation and maintenance of the IC. The technical specifications are divided into general (general Specs), special (special Specs), tentative, and so on. The general Specs establish the given requirements on all types of experimentally or mass-produced IC made by Soviet industry. The general Specs include the requirements on the electrical parameters, structural design, resistance to mechanical and climatic effects, reliability, service life and storability and also acceptance rules. The special specifications define the purpose of each type of IC and the series it belongs to, and they more precisely define the norms on the parameters, the test procedure and conditions; they establish the special and additional requirements. The tentative Specs are generated during the development process when the exact values of certain parameters are still unknown. The general and special Specs complement each other and are necessary for the various enterprises -- client, developer and manufacturer.

The design documents contain the basic set of documents, the set of documents for the basic parts of the IC (case, mounted components), operating instructions, log and certificate. The basic set of documents includes the following: functional, electrical, theoretical and wiring diagrams; drawings of the plate or crystal with the general view, section and topology; the assembly drawing of the crystal or (and) plate in the case and explanatory notes.

The production documents contain a detailed operation-by-operation description of the entire technological process of the manufacture of the IC and reflect the methods, means and procedures for the performance of each individual operation and the entire production process as a whole.

Depending on the production phase, the process documents are divided into categories: the process documents for experimentally produced IC which have the letter O affixed; the process documents for IC in series production to which the letter A is affixed; the process documents for mass-produced IC to which the letter B is affixed.

The set of production documents for series IC consists of the basic and auxiliary documents. The basic documents include the following: the flow chart for the production process, the operation charts of the production process, including production control and instructions. The auxiliary documents include the control flow chart, the list of standard and nonstandard equipment, normalized and special tools, equipment and an explanatory note.

The technological process flow chart establishes the sequence of the completed set of operations. The names of the production operations for making the part or assembling parts are listed in order of their execution with indication of the numbers of the operation flow charts.

The operation flow chart contains a description of the standard process operation with indication of conversions, processing conditions and data on the production equipment. The number of operation flow charts is equal to the number of types of operations in the production process.

FOR OFFICIAL USE ONLY

The production instructions contain a description of the operating procedures or the production process procedures for manufacture or repair of the product (including control and movement), the operating rules for the hardware, description of physical and chemical phenomena occurring during individual operations.

The control flow chart is the basic document which fixes the quality control of the product in the work place. The control flow chart determines the sequence of execution of control operations in the technological production process.

All of the enumerated documents are compiled on special forms strictly in accordance with the developed formats. Changes can be made in the documents only officially after approval of a special technological process alteration chart.

1-4. Basic Processes of IC Production Technology

General Diagram of IC Production Technology

The microelectronics industry produces a large number of IC differing with respect to structural design and performed functions. The IC of each type has its own production technology regulated by the set of production documents. In spite of this fact, it is possible to note general features which are inherent to the production technology of any microcircuit. The technology of any microcircuit includes the following basic processes: the manufacture and finishing of the billets for the IC (semiconductor plates and substrates); the manufacture of structures; assembly, encapsulation; testing, measurement; finishing operations (painting, marking, packing, and so on). After testing in the department of technical control (DTC) the finished microcircuits are sent to the finished production warehouse (Figure 1-2).

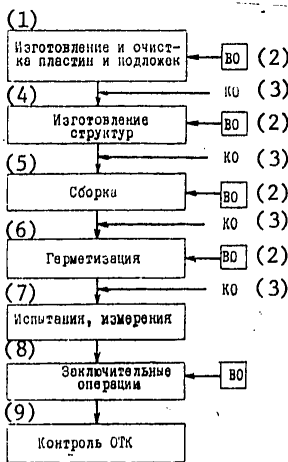


Figure 1-2. General diagram of the manufacture of IC.
 VO -- auxiliary operations; KO -- quality control operations

- Key:
- 1. Manufacture and finishing of plates and substrates; 2. VO; 3. KO;
 - 4. Manufacture of structures; 5. Assembly; 6. Encapsulation; 7. Testing, measurement; 8. Finishing operations; 9. DTC control

FOR OFFICIAL USE ONLY

Manufacture and Finishing of Plates and Substrates for IC

Among the semiconductor materials, silicon is most widely used for manufacturing IC. This is explained by the explicit advantages of silicon over the second best studied and assimilated material -- germanium. Silicon has greater width of the forbidden zone, which insures a broader range of operating temperatures, smaller back currents of the junctions and lower dependence of them on temperature, and it also permits resistors to be made with higher resistances which have slight dependence on the leakage current. The silicon junctions have higher breakdown voltages, and their breakdown comes at higher temperatures. The surface of silicon is more resistant to the presence of contamination. The density of silicon is approximately half that of germanium. Silicon is the most widespread element in nature. And, a very important thing, silicon is the only semiconductor, the natural thermal oxide of which has satisfactory masking properties. The diffusion coefficients of boron, phosphorus, arsenic and antimony in the oxide are significantly lower than in silicon, which permits diffusion to take place not along the entire surface area of the plate, but locally through openings in the oxide film. In addition to silicon, germanium, gallium arsenide and other double and triple semiconductor compounds are used for semiconductor microcircuits and mounted active elements for HIC.

The structure, properties and parameters of semiconductor plates determine the parameters and characteristics of the microcircuit. A semiconductor material of a defined type and defined crystallographic orientation is selected in accordance with the requirements on the IC parameters.

Substrates of glass, pyrocerams, photopyrocerams, ceramics and also multicrystalline sapphire and spinel substrates are used for HIC. Flexible polymer substrates of polyimide began to be used comparatively recently. The substrates for HIC must have high electrical resistance, low dielectric constant (for microwave IC, on the contrary, high), small dielectric losses, good thermal conductivity, high mechanical and electrical strength, thermal and chemical stability, and so on.

In addition to the defined requirements on the properties of materials, IC production imposes a number of requirements on the dimensions and shape of the plates and substrates and the condition of their surface. The thickness of the plates and substrates is determined by the resistance to mechanical loads in the manufacturing, operating or maintenance process. Frequently the thickness of the substrates is dictated by the necessity for obtaining defined functional electrical parameters. The area of the semiconductor plates is determined by the diameters of the ingots which reach 120 mm and more for silicon. The overall dimensions of the dielectric substrates are standardized; for the majority of materials they are 60x48 mm with a thickness from 0.5 to 2 mm. The process equipment and normalized series of types and sizes of microcircuit plates are oriented to these dimensions. The substrates and plates must not have bends, their working surfaces must be flat, parallel and have minimum sizes of microirregularities. The surface roughness must in the majority of cases be no less than class 13-14, which corresponds to a height of profile microirregularities from 0.100 to 0.025 microns. This arises from the rigid requirements imposed on the shape and sizes of the elements both in the horizontal and in the vertical directions and also the reproducibility of these dimensions. Semiconductor and thin-film technology impose especially high

FOR OFFICIAL USE ONLY

requirements on the precision of the surface treatment, for the microirregularities can lead to distortions of the sizes of the elements, holes in the film and short-circuiting of the elements. A necessary condition of obtaining IC with satisfactory and stable parameters is also a high degree of chemical purity and stability of the surface properties of the plates and substrates. The latter is especially important for semiconductor plates, for they are highly sensitive to contamination and surface migration. Multicrystalline plates and substrates must also have surface layers with minimal structural disturbances of the crystal lattice.

In accordance with the presented requirements, the preparation of the plates and substrates includes oriented or unoriented machining of the materials entering into production in the form of ingots and plates, chemical dimensional treatment and cleaning of the surface to remove contamination. In some cases when it is impossible to obtain the required height of surface microirregularities for dielectric substrates, the surface is primed with a layer of insulating material.

Manufacture of Microcircuit Structures

The manufacture of structures, expressed otherwise, the functional base of the IC, is the most important step in the technological process. As was noted earlier, the manufacturing technology of microcircuit structures is closely connected with their design. Two basic principles -- layering and localization -- provide the basis for the technological process, just as for the design. Given configurations, sizes and mutual arrangement of the IC elements are obtained in several steps in accordance with the number of layers of the microcircuit and the topology of each layer. Therefore exact matching of the images of each layer with each other is necessary. The technique for obtaining layer-by-layer images is based on the application of local processing. Depending on the method of localization, all of the processes of obtaining the images can be broken down into three groups: the free mask method, processes using contact masks, and local treatment without using masks.

The free mask method is used when forming film elements. The free, or interchangeable, mask is a plate with holes, the dimensions, configuration and mutual arrangement of which correspond to the topologic configuration of the formed film layer of the IC. A free mask shields the required sections of the substrate from the film being applied. The free mask is a multiple-use tool.

Processes using contact masks are universal. They can be used to obtain any microcircuit. A contact mask is a mask in the form of a relief layer with holes designed for one-time use. The contact mask is made directly on the substrate or the plate and is held on its surface by adhesive forces. The contact masks can be formed using optical lithography (contact photolithography, photolithography with clearance, projection photolithography), x-ray lithography, electron lithography, and so on. The best studied, technically developed and widely used method in microcircuit production is contact photolithography.

The local treatment without using masks can be carried out using a sharply focused electron beam, the switching on and off and displacement of which in accordance with a given program are computer-controlled.

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

Thus, in the manufacturing technology of the structures of any IC the defining methods are the methods of obtaining the images, that is, the given configurations, dimensions and mutual arrangement of the elements.

When making thin-film elements, the predominant methods of obtaining the films are thermovacuum deposition, gas discharge plasma spraying, anodic electrolytic and plasma oxidation, electrolytic cathode deposition. Thin-film elements are obtained predominantly by the stenciling method, that is, mechanical extrusion of special pastes through holes in a stencil.

Two-dimensional technology is the basis for producing semiconductor structures. It includes a defined sequence of group treatment, obtaining the masking film, photolithography, local treatment. The locally treated sections are located in a defined order and at a defined distance from each other, as a result of which a large number of structures are simultaneously obtained on each plate in one technological process.

The flow chart for the manufacture of the simplest two-dimensional structure (discrete diode structure) is presented in Figure 1-3. The prepared silicon plates are thermally oxidized to obtain a masking film of silicon dioxide. Then primary photolithography is carried out between "windows" in the masking film through which local diffusion of the acceptor admixture takes place. By secondary photolithography windows are obtained for the resistance contacts to the n and p-regions. Then an electrically conducting film is applied to insure good contact with the semiconductor, and tertiary photolithography is carried out to form the contact layout.

When making a discrete transistor, each of the above-indicated sequences -- masking, photolithography, local treatment -- forms a defined region of the structure: collector, base, emitter, contact. The recurrence interval of like operations is common to the manufacture of a discrete diode and transistor, just as other, more complex two-dimensional structures. With complication of the structure, the number of recurrences increases.

The recurrence interval of like operations permits the production of structures to be broken down into a comparatively small number of sections (mechanical, chemical, thermal, photolithographic, vacuum), where it is possible to make any two-dimensional structure beginning with the simplest discrete diodes and ending with complex microwave transistors and large integrated circuits (LSI). The transition to the production technology of a new two-dimensional structure frequently can be made only by changing the set of tools, accessories and the conditions of performing the production operations.

Two-dimensional technology is highly efficient, it insures small dispersion and good reproducibility of the parameters of the structures with comparatively small production expenditures.

As a result of the application of masking films the substrate surface turns out to be protected against external effects before the formation of the p-n junctions.

FOR OFFICIAL USE ONLY

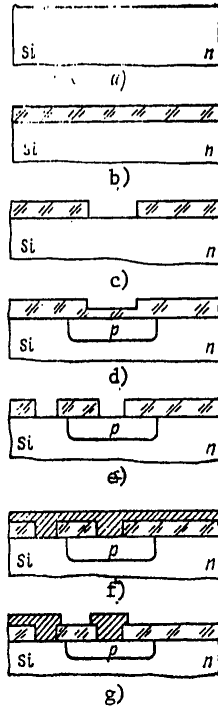


Figure 1-3. Chart of the manufacture of the structure of a two-dimensional discrete diode.
 a -- preparation of the plate; b -- thermal oxidation; c -- primary photolithography; d -- local diffusion; e -- secondary photolithography; f -- application of a metal film; g -- tertiary photolithography

In the finished structure the masking film remains and plays a passivating role, that is, it protects the structure from external effects. This explains the high stability of the parameters and reliability of the two-dimensional devices and IC.

Photolithography and local treatment permit p-n junctions of very small areas and complex configuration to be obtained with high precision. This increases the maximum frequencies, it increases the power and improves the other parameters of the IC and their reproducibility.

Assembly of Microcircuits

In the assembly section the plates and substrates with finished IC structures are separated into individual structures: crystals (semiconductor IC or semiconductor mounted components) or plates (film IC or the passive part of HIC). The crystals or plates are mounted on the mounting plane of the base of the case or lead-out frame, and the mounted components are also installed for HIC. Then the contact

FOR OFFICIAL USE ONLY

areas of the crystals or plates are electrically connected to the leads of the case or frame.

Encapsulation

Encapsulation is enclosure of the crystal or plate in a reliable shell which protects them from mechanical loads and environmental influences during operation, and also insures convenient, reliable installation of the microcircuit in the equipment.

With respect to engineering design attributes, the encapsulation is divided into case, caseless and combined. Case encapsulation provides for preliminary manufacture of the case, inside which the crystal or plate is enclosed. In this case there is always free space between the structure and the inside walls of the case. During caseless encapsulation there is no free space inside the finished IC, for the structure is entirely enclosed in the encapsulating material. In the case of combined encapsulation the fittings with the structure are placed in a specially made cover, and then all of the free space inside the cover is filled with the encapsulating material. When selecting the materials and the methods of encapsulation it is necessary to consider the possibility of the removal of the heat released during operation of the IC.

Tests and Measurement

Production tests and measurement of the IC parameters are performed after encapsulation in order to reject IC containing hidden defects. The production tests are divided into climatic, mechanical and electrical. During the climatic tests the IC are checked for alternate heating and cooling cycles, moisture, heat and cold resistance. Mechanical impact and vibration tests are run to check the mechanical strength of the structural design and the intracircuit connections of the IC. Current pre-burning and checkout of the stability of the IC and measurement of the IC parameters are carried out during the electrical tests.

Finishing Operations

The finishing operations are performed after the production tests and measurements of the encapsulating IC. The finishing operations include painting, lacquering, galvanic coatings, marking, preparation of leads for connection to the equipment and packing. The marking containing the provisional designation of the IC, the designation of the primary lead, the year and month of manufacture and the trademark of the enterprise and stamp of the Department of Technical Control must appear on each microcircuit produced.

Test Questions and Assignments

1. What basic elements make up the basis of modern microelectronics and what is the main goal of the course in "Production Technology of Microelectronic Devices"?
2. Give a comparative description of each generation of the element base of electronics.

FOR OFFICIAL USE ONLY

3. How do the reliability, functional complexity, parameters, service life, dimensions, weight, and cost of the basic elements of electronics vary from generation to generation?
4. Give the names of Soviet and foreign scientists who made a significant contribution to the development of microelectronics.
5. What is technology?
6. What is the role of the production operation in the IC manufacturing process?
7. What production operations do you know?
8. What is the production method and what methods do you know?
9. What is a technological production process? What technological production processes are called standard?
10. Into what groups are the IC divided with respect to engineering design execution?
11. Describe each IC group.
12. What distinguishes the bulk IC elements from film elements?
13. In which executions are active and passive elements most widely used and why?
14. Why is it necessary to consider the technological possibilities and limitations when designing IC?
15. What is an IC technical assignment?
16. Compare the operations in the basic phases of development of new IC technology.
17. What are the technical specifications for IC?
18. What documents are included in the design documents for IC?
19. What do the basic and auxiliary production documents for series IC include?
20. List the basic processes of the manufacturing technology of any IC.
21. Give a brief description of each process.
22. Explain the essence and the characteristic features of two-dimensional technology.
23. List the basic methods of film technology.
24. List the basic methods of two-dimensional technology.

FOR OFFICIAL USE ONLY

25. What is the purpose of the assembly and encapsulation processes?
26. What role do the tests and measurements of IC play?
27. What production operations are included in the finishing operations?

FOR OFFICIAL USE ONLY

CHAPTER 2. GENERAL REQUIREMENTS ON MICROCIRCUIT PRODUCTION

2-1. General Requirements on the Technological Process

Before the beginning of a detailed study of the basic steps of the technological production process of manufacturing microcircuits, let us consider the requirements on it.

1. The technological production process must provide reproducible production of IC of a given structural design, with electrical parameters and service life provided for by the technical specifications, independently of the manufacturer, the servicing operators, time of year, shift, and so on.
2. IC made in accordance with the technical forms and reports must retain the structural design, electrical parameters, external appearance and markings under the test, storage and operating conditions provided for by the technical specifications.
3. The technological process must be universal insofar as possible in order to permit the manufacture of a broad nomenclature of products by a united principle with minimum alterations and expenditures.
4. The equipment, accessories and tools used must be as cheap as possible, simple to manufacture and convenient to operate and maintain.
5. The basic structural elements (substrates, crystals, plates, case parts) must be unitized and standardized, for the manufacture of a lot of standardized parts requires lower economic expenditures than the manufacture of a nonstandard lot of the same size. In addition, the application of standard parts provides for continuous operation without partial modifications on like equipment, for example, for another size of crystal when dividing the semiconductor plates.
6. The materials used must be as cheap as possible, nonshort and nontoxic materials.
7. The technological process must be put together considering compatibility of the production operations and materials used. The production operations are considered to be compatible if the properties attained in the preceding operation or the parameters of the treated sample do not become worse as a result of the performance of a subsequent operation.

FOR OFFICIAL USE ONLY

The materials used are considered to be compatible if they do not enter into undesirable interactions with each other or with the environment during the production operation, storage or operation of the finished microcircuit.

8. When developing the technological process it is necessary to select group methods and the equipment which will insure mass production and the control of which is easily automated with the application of computers.

9. The production process and the equipment used must justify the material expenditures on the development and manufacture of the IC. For this purpose a high rate of execution of the production operations, reproducibility of the parameters in each operation and a high percentage yield of finished products are necessary.

10. The development of the production process must consider the requirements of safe operation and fire safety.

The presented requirements are to a great extent interrelated. They must be considered when developing new IC and also when modifying the production processes and equipment.

2-2. Requirements on Cleanness of Air and Climatic Parameters

Production Environments

Throughout the entire production process the treated specimens are in contact with one environment or another.

The production environment is the air, gas or other environment characterized by strictly defined parameters, the presence of which is a mandatory condition for performance of the given production operation. In accordance with the functions performed and the physical-chemical properties, the production environments are divided into five groups: air, gas, liquid, vacuum and plasma.

Air Environment Parameters

The improvement of the IC manufacturing technology connected with improving the degree of integration of the IC has raised the requirement not only on cleanness of the gases, liquid media and degree of vacuum, but also the air environment of the production facilities. The requirements on the air environment of the production facilities when building microcircuits are equivalent to the requirements imposed on other production environments. These requirements are determined by the nature of the production operations performed and the sanitary norms for the service personnel.

The following parameters of the air environment have the greatest influence on the production processes: temperature, moisture, dust and the content of gas and liquid vapor admixtures. For successful performance of the entire process in one production facility the parameters of the air environment must be maintained on the level required for the performance of the operations sensitive to these parameters. Insurance of high quality of the environment of an entire production facility is connected with large economic expenditures. Therefore the production process is divided into groups of operations requiring close values of the

FOR OFFICIAL USE ONLY

parameters of the air environment. Accordingly, the production is divided into a number of production sections occupying individual facilities.

Temperature and Humidity of the Air Environment

Temperature fluctuations change the linear dimensions of the fittings and the treatment units, the speed of the chemical reactions, the evaporation rate of the materials used, the readings of the monitoring and measuring instruments, the parameters of the structures and the finished IC. Therefore for each production operation sensitive to temperature fluctuations tolerances have been established for the deviations from the optimal air temperature.

A moist environment is undesirable when performing in practice all production operations. On being adsorbed on the various surfaces, the moisture and the materials dissolved in it lead to the formation of undesirable oxides and compounds that are difficult to dissolve. In the case of heat treatments, irregularities (erosion) can occur on such surfaces, having a negative effect on the parameters of the IC. An effort is made to keep the humidity of the air in the production facilities at a minimum, but no lower than the sanitary norm.

In accordance with the mandatory requirements on the production conditions of semiconductor devices and microcircuits the OST 11PO 050.001-74 establishes three categories of microclimate of the production facilities, temperature norms and relative humidity norms for which are presented in Table 2-1.

Table 2-1. Separation of Production Facilities into Categories with Respect to Temperature-Moisture Parameters

Category of microclimate	Temperature, °C		Relative humidity, %
	Winter	Summer	
I	21+1	23+1	45+5
II	20+2	23+2	45+15
III	According to SN 245-71		

For the majority of third-category facilities the sanitary norms SN 245-71 establish an admissible temperature of 17-22°C in the winter, and in the summer the temperature must not exceed 28°C.

Dust in the Air

In connection with the micron and submicron dimensions of the elements and their electrode gaps, the presence of mechanical aerosol particles in the air can have a significant negative effect on the quality and lower the percentage yield of usable IC. The mechanical particles can lead to scratches, breaks in the conducting tracks, electrical short-circuits of the elements; negative effects on film adhesion, and be the cause of nonuniform thickness of the films, through pores, punctures, and so on.

FOR OFFICIAL USE ONLY

If the dimensions of the IC elements are close to the dimensions of the dust particles, and the number of elements is large, then the dust in the air must be estimated by the number of dust particles per unit volume of air. In accordance with the established standard the dust content in the air of production facilities of microelectronics enterprises is estimated by the number of particles 0.5 micron in size per unit volume of air. All of the production facilities are divided into five classes of cleanness by the dust content of the air (Table 2-2).

Table 2-2. Classification of Production Facilities or Work Spaces by the Dust Content in the Air

Class of cleanness	Maximum number of particles 0.5 microns in diameter m^3 (particles per liter of air)
1	$4 \cdot 10^3$ (4)
2	$3.5 \cdot 10^4$ (35)
3	$3.5 \cdot 10^5$ (350)
4	$3.5 \cdot 10^6$ (3500)
5	According to SN 245-71

Insurance of Required Cleanness Classes and Microclimate Categories

It is desirable to place the enterprises of the microelectronic industry far from the large industrial cities in green zones. For the construction of industrial buildings and facilities it is necessary to use special wear-resistant materials which are easily cleaned and do not pollute the air. When supplying air to the facilities, it must go through a special filtration system. The production facilities must be equipped with an air conditioning system. In order to avoid penetration of outside air, the pressure inside the buildings must exceed atmospheric pressure somewhat. All of the industrial wiring must be concealed, and the content of the facilities and the use of specialized clothing are regulated.

In addition to water vapor and aerosols the air can also have other harmful impurities: carbon monoxide and carbon dioxide, hydrogen sulfide, the vapor of acids, bases, chromium, zinc, lead, and so on. In order to prevent the effect of these admixtures, responsible operations are performed in special plane spaces with somewhat excess air pressure with respect to the main facility.

The clean room (sealed zone) is a facility in which the production operations requiring special cleanness and careful maintenance of the temperature and humidity parameters of the air environment are performed and also strict observation of the production hygiene norms. Clean rooms have the following: a working facility where the equipment is placed and the production operations are performed; dressing rooms for the service personnel to get ready to work in the clean room; transfer and ventilation locks; facilities for treating the intake air. In the clean rooms it is important to insure not only defined temperature, humidity and dust content, but also a defined magnitude and direction of the air displacement velocity vector.

FOR OFFICIAL USE ONLY

Three types of aerodynamic layouts for the clean rooms are distinguished: convective, lamdnar and recirculating air flows.

In the clean rooms with convective air flows (Figure 2-1) the location of the entrance and exit registers promotes movement of the dust from its sources throughout the entire volume of the facility, and it complicates its removal through the exhaust opening. In such rooms it is impossible to obtain a dust content in the air of less than 350 particles/liter. At the same time such air flows promote equalization of the temperatures in various sections of the clean room.

In clean rooms with laminar air flow (Figure 2-2) it is possible to provide a dust content of less than 30 particles/liter. The dust from the sources cannot be spread throughout the room, for it is carried away by the general air flow to the exhaust openings. With vertical air flow the floor must not be taken up with heavy equipment. With horizontal air flow the cleanest operations are located against the wall through which the air is fed.

Clean rooms with recirculating air flow (Figure 2-3) have local VOPR-0,9 or VOPR-1,5 air cleaners which take in the air from the facility, clean it and direct it back into the facility.

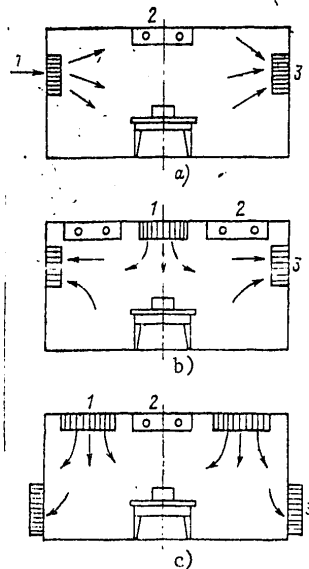


Figure 2-1. Layouts of clean rooms with convective air flows.
1 -- intake openings; 2 -- lighting; 3 -- exhaust openings

FOR OFFICIAL USE ONLY

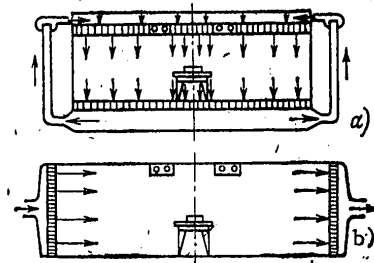


Figure 2-2. Layouts of clean rooms with laminar vertical (a) and horizontal (b) air flows

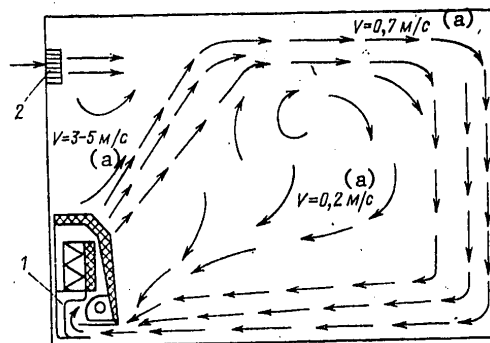


Figure 2-3. Layout of a clean room with recirculating air flow. 1 -- VOPR-09 or VOPR-1,5 air cleaners; 2 -- entrance of air from central air conditioner

Key:

a. m/sec

Local work spaces in the form of open and sealed boxes and enclosures are used for performing operations that require a cleaner air environment than in the facility or requiring a special gas environment. Open boxes (Figure 2-4) have an opening on the operator side which permits the products to be transferred directly from the facility into the work place and the operations to be performed without gloves. The cleanness of the environment in the open boxes corresponds to first class cleanliness. Sealed boxes and enclosures completely protect the processed specimens from the air environment of the facility. The specimens are transferred to the enclosures through locks by means of which the enclosures can be joined to the flow line. The specimens are handled in gloves through special openings. The most sensitive operations, for example, the mounting of the crystals, encapsulation, and so on are performed in the sealed enclosures.

FOR OFFICIAL USE ONLY

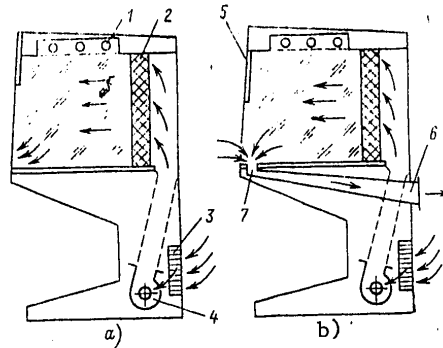


Figure 2-4. Open boxes with horizontal air flow for the performance of operations without harmful emissions (a) and with harmful emissions (b).

1 -- light fixtures; 2 -- filter; 3 -- air intake register with filter;
4 -- fan; 5 -- raisable glass shutter; 6 -- connecting pipe for removing the air; 7 -- slit for removing air

The monitoring of the air environment of the clean rooms and local work places permits deviations from the established norms to be detected and the required atmospheric parameters to be maintained.

Liquid thermometers, thermographs, resistance thermometers operating in combination with pen recorders are used to monitor the temperature.

The humidity is measured by psychrometers, hygrographs permitting the readings to be recorded on a strip chart or by continuous automated "Volna" instruments. The operating principle of the latter instruments is based on measuring the resonance frequency of the vibrations of a piezoelement coated with a film of hygroscopic material. On absorbing moisture the weight of the film increases which causes a proportional change in the vibration frequency of the piezoelement.

The dust content of the air is measured by sedimentation, filtration or optical methods, that is, by calculating the number of dust particles settling on 1 cm² of area of the clean silicon plate or microscope slide, by calculating the number of dust particles on the surface of the filter through which a defined volume of the air is passed, or by measuring the number of electric pulses obtained as a result of conversion of the signals scattered by the dust particles.

2-3. Requirements on the Production Gases and Water

Application of Gases

Gases are most frequently used as shielding atmospheres, reagents, carriers and ionized media in microcircuit production.

Gas shielding environments (nitrogen, argon, helium) are used to shield the products from oxidation and corrosion when performing operations connected with increased

FOR OFFICIAL USE ONLY

temperatures, for blowing the devices before and after performing the thermal processes and also to protect the specimens from the effect of the air in the production facilities during interoperations storage and transportation.

The gas reagents participate in the chemical reactions when alloying semiconductors (arsine AsH_3 , phosphine PH_3 , diborane B_2H_6), when pickling silicon plates (hydrogen chloride HCl , hydrogen bromide HBr , hydrogen sulfide H_2S , chlorine, and so on), when obtaining masking films (oxygen O_2 , silane SiH_4 , ammonia NH_3 , carbon dioxide CO_2), during reduction processes (hydrogen H_2), and so on.

The gas carriers (nitrogen, argon, oxygen, hydrogen and others) are used for transferring the vapors of sources of alloying impurities, water vapor and also the vapor of other initial compounds for growing semiconducting layers or various films to the working chambers of the devices.

Ionized environments, argon, nitrogen, oxygen are used to perform plasmothermal and plasmochemical processes and the processes of plasma spraying of a material.

Requirements on Gases

All of the process gases must, independently of their purpose, correspond to a defined degree of purity regulated by the branch standards and the technical specifications. The content of harmful impurities in the most widely used gases of first class purity must not exceed the admissible forms presented in Table 2-3 (OST 11050.003-75).

Table 2-3. Admissible Content of Impurities in the Process Gases

Gas	Impurity content (no more than)			
	Oxygen, % by volume	Water vapor, °C, by dew- point	Hydrogen, % by volume	Dust particles, particle/liter, 0.5 micron in size
Nitrogen	$5 \cdot 10^{-4}$	-65	$1 \cdot 10^{-3}$	4
Argon	$5 \cdot 10^{-4}$	-65	$1 \cdot 10^{-3}$	4
Hydrogen	$1 \cdot 10^{-5}$	-75	-	4
Oxygen	-	-85	$1 \cdot 10^{-3}$	4
Compressed air	-	-65	-	4

The CO , CO_2 , CH_4 , C_2H_2 and other impurities in the gases must be no more than $1 \cdot 10^{-5}$ to $1 \cdot 10^{-4}\%$ (by volume).

Insuring the Required Purity of Gases

Technical gases contain more impurities than required for IC production. Therefore, before the gases are supplied to the shops they are centrally purified of oxygen, hydrogen, mechanical particles, moisture, oil, and so on. If the degree of purity of the gases is insufficient after centralized purification, they are also purified directly before admission to the operating units (finish purification). Purification of the gases with respect to oxygen is realized by passing the gases through a membrane made of a palladium alloy with platinum which absorbs oxygen well or by

FOR OFFICIAL USE ONLY

binding the oxygen to hydrogen in the presence of a palladium catalyst. The dust is removed from the gases by filtering the gases through various filters (ceramic, lavsan, electrostatic, and so on).

Chemical and physical methods are used to dry the gases. In the chemical methods the water vapor enters into a chemical interaction with the formation of hydrates or other compounds. The physical methods include adsorption drying and drying by freezing. Adsorption drying is realized jointly with the removal of oxygen or hydrogen bypassing the gas through adsorbers filled with material with high adsorption capacity (silica gel, aluminogel or zeolites). The drying by freezing the water vapor, oil and other vapor is done directly before the work place where the gas is used by passing it through a coil in a container with liquid nitrogen. The condensed liquid precipitates on the coil walls in the form of solid crystals. In order to prevent transport of the crystals by the gas flow, a filter is required directly after the coil.

After centralized purification before supplying the gases to the shops, they are subjected to continuous automatic monitoring for the oxygen, hydrogen and water vapor content. At the exit from the finish purification units there is continuous automatic monitoring of the oxygen and water vapor contents in the gases. The periodicity of monitoring the impurities at the entrance to the process equipment and monitoring the dust content of the gases in all phases is established by the enterprise standards.

The maintenance of the pure gas networks has great significance for insuring a minimum amount of impurities in the gases. The pure gas networks must be sealed, noncorrosive, and they must have minimum length. In order to decrease the gas release from the inside surfaces, the lines are made of stainless materials. Before installing the pipes are degreased, they are blown out with compressed air or nitrogen, washed with hot water and again blown out. During operation the inside walls of the pipes are cleaned regularly.

Water Requirements and Insurance of Them

In microcircuit production water is used in large quantities to make various solutions, for washing the substrates, the finish structures, the case parts and also as a reagent and protective medium. Natural water contains a large number of mechanical parts (hydrosols), dissolved mineral salts, admixtures of copper, silver, gold, bacteria, and so on. Therefore water which has been purified of all pollution is used in IC production. The index of the water purity is its specific resistance. The resistance of water depends strongly on temperature; therefore the body of the specific resistance is given at 20°C. The natural specific resistance of ideally pure water is 25 mohms-cm. In microcircuit and semiconductor device production, first, second and third degrees of purity of water are distinguished, which correspond to specific resistances of 18, 10 and 1 mohm-cm (A, B and V type water, respectively). In addition to the specific resistance, the silicic acid, micro-particle, microorganism, copper, iron and so on content in the water is determined. For microcircuit production the tap water is subjected to preliminary and final purification. The preliminary purification of water to remove suspended and colloiddally-dissolved particles is carried out by the methods of distillation and sorption using special filters, reagent coagulation, electrocoagulation, and so on. The previously purified water has soluble salts and other admixtures. For final

FOR OFFICIAL USE ONLY

purification of water to remove soluble materials at the present time purification by ion exchange resins (deionization) and the method of inverse osmosis are used.

In order to obtain especially pure water with a specific resistance of 18-20 mohms-cm, the following purification systems are used: distillation-deionization; electrocoagulation-deionization; electrocoagulation-ultrafiltration-inverse osmosis; ultrafiltration-deionization. These systems also include preliminary filtration, for example, through activated charcoal to remove chlorine, organic pollutants, turbidity, large and fine solid particles.

The pure water is checked for specific resistance by an instrument with a bridge compensation electric circuit. The presence of silicic acid in the water is determined comparing the color of a special solution prepared on the basis of the checked water with the standard scale colors (the colorimetric method). Organic impurities are determined by their capacity to be oxidized on introduction of an oxidizing solution of potassium bichromate, concentrated sulfuric acid and crystalline silver sulfate into the water which plays the role of a catalyst. The suspended particles are analyzed by the filtration method, microphotography of the water sample, the optical method and other methods.

2-4. Basic Principles of Electron Vacuum Hygiene

The purity of the air in the production facilities that manufacture IC is on such a high level that the service personnel become sources of pollution. Therefore when possible the presence of service personnel is limited, and all of the personnel participating in the manufacture of integrated microcircuits are subject to the electron-vacuum hygiene (EVH) rules.

EVH is made up of the general hygiene of the production facility, individual hygiene of the process equipment and the personal hygiene of the service personnel.

Each enterprise has a special EVH service which realizes organization and control of the effective instructions for observing the EVH rules. The EVH service accounts for the state of the EVH in the enterprise subdivisions, it monitors the condition of the microclimate in the production sections, it develops the operating conditions and the conditions of maintenance of the facilities, the requirements on the work places, instruments, production forms and reports and also the requirements imposed on the people working in the production facilities.

In the clean rooms for maintaining the laminar state and, consequently, purity of the air flow, the equipment is arranged so that the spacing between work places will be no less than 1 to 1.2 meters, and the distance from the walls to the equipment, no less than 0.5 to 0.8 meters. In order to decrease the effect of the heat released by the equipment on the convective transport of dust particles the equipment is built into wall panels so that only the charging devices protrude into the clean room. This arrangement makes it possible to carry out preventive cleaning of the equipment outside the clean room.

In clean rooms, in addition to monitoring the dust content of the air atmosphere and microclimate, the conditions of labor are also monitored.

FOR OFFICIAL USE ONLY

The air of the clean rooms contains particles from hundredths to several tenths of a micron in size which are not trapped by the finish filters of the purifiers and are not monitored when analyzing the dust content. These particles can settle and accumulate in the facility. For removal of them periodic wet cleaning is needed, sometimes with the addition of 5% glycerine in order that the dust not separate from the surfaces. The water for wet cleaning of the floors, walls and ceilings must be pure, better deionized with a specific resistance of no less than 0.5 mohm-cm.

The sources of pollution from the personnel are primarily the skin of man and clothing. Water vapor, salt, fat and other materials are released through the skin of man. As a result of constant renewal of the outer layer of skin, the dead particles flake off. Depending on the type of activity the number of dust particles removed from human clothing can vary within broad limits. Therefore work done in the clean room is done in special work clothing made from material with minimum dust generation. Before going into the clean area, the workers go through intermediate areas, dressing rooms and blow-off locks. Observation of the personal EVH rules is a necessary condition of producing high-quality microcircuits.

2-5. Basic Production Features of Microcircuits

The production of IC differs to a great extent from other industrial production and has a number of specific features. Let us consider the basic ones of them:

1. Almost all the phenomena and processes known to science and practice are used in IC production: mechanical, physical, chemical, various types of treatments, including electronic, ionic, laser beams, various methods of measurement and control with the application of radioelectronic equipment, x-ray television, electron, laser microscopes and so on.
2. IC production has developed and is being improved at exceptionally fast rates, influencing, at the same time, the development of in practice all branches of the national economy.
3. IC production uses an enormous amount of materials of various properties, much of which is in a special class of materials "for semiconductor production" and must correspond to the requirements of exceptionally high purity. This defines the high requirements on the microclimate and purity of the production facilities, the production environments, fixtures, and so on.
4. When manufacturing IC broad use is made of group production methods. Deviations in the process conditions of even one operation can lead to rejection of an entire lot.
5. The processed units are distinguished by micron and submicron dimensions, which requires high precision and stability of the production treatments and conditions, for otherwise dispersion of the parameters is possible not only from substrate to substrate, but also within a single substrate. In addition, the complexity of handling micron size objects requires manual labor to be replaced by automated labor (machines).

FOR OFFICIAL USE ONLY

6. Many of the production processes of manufacturing IC are distinguished by comparative complexity of control, for the parameters of the elements obtained are simultaneously influenced by a large number of factors, the number of which are difficult to consider. The multifactor nature of the dependence of the parameters of the processed objects on the conditions of performance of the production operations imposes defined difficulties on the mathematical simulation of the technological processes and makes it complicated to write simple programs for computer control.

7. All of the operations of making IC are regulated and must be performed with observation of EVH and in exact correspondence to the production reports and forms of the microelectronic industry. The observation of the production discipline -- maintenance of all of the operating conditions, application of materials only of the needed types, the All-Union State Standards, technical specifications and so on -- is mandatory for everyone engaged in the development and manufacture of IC. Any alterations of the process can be made only after approval of the new documents.

8. Among the variety of materials used, the various types of energy carriers, gas media and reagents there are toxic, explosive and flammable materials. This requires observation of strict engineering safety rules specific to each work place.

9. Integration of knowledge of many materials, a broad class of phenomena occurring during the technological processing, the operating principles of complex (and sometimes unique) equipment, knowledge of the organization and economics of production, the physics of the operation of microcircuit elements and systems engineering is characteristic of the technologists. The development and manufacture of IC is not within the grasp of a single specialist -- this is the work of a collective. Therefore the technologist specializing in the production of microelectronic devices must not only have a defined set of skills, but must also be a production organizer.

Test Questions and Assignments

1. Enumerate the basic requirements imposed on the technological production process of manufacturing IC.
2. Briefly formulate the requirements on the production cycle.
3. What is a production environment?
4. What groups are the production environments divided into?
5. What determines the requirements on the air environment of the production facilities?
6. Why and by what attributes is production broken down into production sections?
7. What parameters characterize the quality of the air environment and why?
8. How many types of facilities are there with respect to temperature-humidity parameters and dust content?

FOR OFFICIAL USE ONLY

9. What is a clean room?
10. What types of aerodynamic systems for clean rooms do you know?
11. Compare clean rooms of different aerodynamic systems.
12. What is the role of local work spaces?
13. Compare the local work space known to us.
14. Explain the means of monitoring temperature, humidity and dust content in the air.
15. Briefly formulate the requirements imposed on process gases.
16. What is the role of the gases in IC production?
17. What measures are taken to insure the required purity of gases?
18. What is the role of water in IC production?
19. What distinguishes the final purity of the water from the preliminary purity?
20. Compare distilled water with deionized water.
21. What is the specific resistance of water used for the production of semiconductor devices and microcircuits?
22. How is the water purity quality controlled?
23. Enumerate the basic EVH principles.
24. What are the basic IC production features?

FOR OFFICIAL USE ONLY

CHAPTER 3. MANUFACTURE OF SEMICONDUCTOR PLATES AND DIELECTRIC SUBSTRATES FOR MICROCIRCUITS

3-1. General Information About the Machining of Semiconductors and Dielectrics for IC

Abrasive Machining

Semiconductor materials and monocrystalline dielectrics (sapphire and spinel) go to the machining production section in the form of ingots; amorphous and other dielectric materials (glass, pyroceram, polykor, ceramics) -- in the form of sheet billets. Semiconductors and dielectrics are distinguished by high hardness and brittleness; therefore the methods used for machining metals cannot be used to obtain billets (semiconductor plates and dielectric substrates) for the manufacture of IC structures. In order to obtain semiconductor plates and dielectric substrates, abrasives of various types are used. The essence of abrasive machining consists in the mechanical effect of a harder and less brittle material -- the abrasive -- on a less hard and more brittle material. The mechanical pressure from the tool is transmitted to the abrasive grains and from them to the machined material. The abrasive grains produce local microdestructions of the surface of the machined material, the released microparticles of which are removed from the machining zone.

Abrasive Materials

Basically synthetic abrasives are used in microcircuit production: diamonds, silicon carbide and electrocorundum.

Synthetic diamonds are not inferior to natural diamonds with respect to their mechanical properties. A diamond is the hardest of all known materials. The Mohs hardness of diamonds is 10. Among the abrasive materials diamonds are in a special class.

Silicon carbide (SiC carborundum) has different color depending on the amount of impurities -- from light green to black. In IC production most frequently green silicon carbide (KZ) is used. The Mohs hardness of silicon is 9.5-9.7.

White synthetic corundum is crystalline aluminum oxide (Al_2O_3) with different impurities (0.5-1.5%). Synthetic corundum is inferior to silicon carbide with respect to hardness, but the strength of the synthetic corundum is higher. The Mohs hardness is 9-9.2.

FOR OFFICIAL USE ONLY

Table 3-1. Characteristics of Powdered Abrasives and Diamonds

Group	Grain size		Grain size of basic fraction, microns
	According to All-Union State Standard 3647-71	According to All-Union State Standard 9206-70	
Abrasive grinding powders	12	-	160-125
	10	-	125-100
	8	-	100-80
	6	-	80-63
	5	-	63-50
	4	-	50-40
Abrasive micropowders	3	-	40-28
	M63	-	63-50
	M50	-	50-40
	M40	-	40-28
	M28	-	28-20
	M20	-	20-14
Abrasive fine micropowders	M14	-	14-10
	M10	-	10-7
	M7	-	7-5
Diamond micropowders	M5	-	5-3
	-	60/40	60-40
	-	40/28	40-28
	-	28/20	28-20
	-	20/14	20-14
	-	14/10	14-10
	-	10/7	10-7
	-	7/5	7-5
	-	5/3	5-3
	-	3/2	3-2
-	2/1	2-1	
-	1/0	2-1	

1 or less

Abrasive powders are divided into four groups depending on the grain size: grinding grain, grinding powder, micropowder and fine powder. The three last groups of powders are used in IC production. Powdered diamonds are divided into grinding grain and micropowders with respect to grain size, and with the micropowders are used in IC production. The designations for the grain sizes of abrasive and diamond powders with different grain size in the basic fraction are presented in Table 3-1.

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

The type of material and the grain size are part of the designation for abrasive or diamond powders. For example, EBMI4 is white synthetic corundum with basic fraction grain size from 14 to 10 microns; ASM 10/7 is synthetic diamond with basic fraction grain size from 10 to 7 microns.

Abrasive Tools

Abrasive materials are used in the free (suspension, paste) and bound (disc) state. Abrasive and diamond suspensions are mechanical suspensions of the corresponding powder in water, oil or another liquid. The pastes are basically prepared from diamond micropowders and surface-active materials which promote improved quality of machining the surface. The discs are of two types: simple metal and metal with diamond-containing cutting edge. Metal discs with diamond-cutting layer bound to their end or peripheral parts are used for grinding. The strength of the discs depends on their dimensions, the type of diamond powder, its concentration and the machining conditions. The binder which binds the diamond grains to the cutting part must provide for self-sharpening of the tools, that is, it must hold the diamond grains and at the same time not interfere with removal of the dulled grains from the diamond-containing layer.

Machining Operations

Machining includes the following operations: cutting, grinding, polishing. Crystallographic orientation which serves to determine the angle of deviation of the plane of the end of an ingot from the given crystallographic plane, is used before cutting monocrystalline materials. Cutting is used to separate the ingots and sheet billets into plates and substrates of the required size and also to separate semiconductor plates or substrates with finished structures of the microcircuits into individual crystals or plates. Before cutting into plates the ingots of non-uniform cross section are sized, and the large size ingots are laid out in bars. The separation of the sheet materials into substrate billets is done with an allowance for subsequent grinding of the ends.

The grinding and polishing are finishing operations, and they are used to insure accuracy of the dimensions and quality of the plate and substrate surfaces.

It is necessary to attach the samples in special holders to perform the machining operations. The fastening is most frequently done by gluing using a mastic based on ED-5 or ED-6 epoxy resin, BF adhesive, shellac, glyptal resins, and so on. During bonding it is necessary to insure continuity and uniformity of the thickness of the bonding layer, where the thickness of the bonding layer must be as thin as possible. The quality of attaching the ingots in sheet billets influences the planarity and parallelness of the sides of the plates and substrates, the thickness of the surface layer with disturbed crystal structure, and so on.

3-2. Abrasive Cutting

Bound Abrasive Cutting. Discs with inside and outside diamond cutting edge are used for bound abrasive cutting.

FOR OFFICIAL USE ONLY

Cutting by a disc with an inside cutting edge is basically used for cutting ingots into plates or bars (Figure 3-1). The cutting is distinguished by high output capacity (60-80 mm/min for silicon), and it insures good quality of machining (eighth to ninth class for silicon). The waste is small: with a thickness of the disc base of 0.1 mm the width of the silicon cut does not exceed 0.25 mm. The recovery of the material from the waste is comparatively easy.

The disc is stretched in the radial direction and it is fastened by the peripheral part on the head of the spindle of a cutting tool. The fastening around the periphery insures low vibration when the disc rotates at 3000 to 5000 rpm. Before beginning cutting the ingot glued to the mandrel is rotated using the machine tool rotary device so that the plan of the cut will be parallel to the given crystallographic plane.

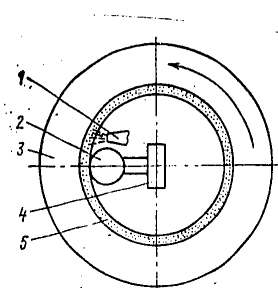


Figure 3-1. Diagram of a cut made by a disc with inside diamond-containing cutting edge.

1 -- nozzle supplying the cooling and lubricating fluid; 2 -- ingot; 3 -- disc base; 4 -- mandrel for fastening the ingot; 5 -- cutting edge of the disc

The diamond grains bound to the cutting edge press with large force against the machined surface when the disc is rotated forming scratches from which cracks penetrate deep into the ingot. The intersections of the set of cracks make punctures, and the particles of the material removed from the ingot. In addition to brittle fracture of the material, the protruberances from its surface are also cut away by the diamond grains.

A jet of liquid is fed to the cutting zone which serves to remove the heat generated during cutting, to remove the particles of cut-away material and the deteriorated diamond grains, to decrease friction, and to lessen the impact-vibration forces. On feeding the fluid to the microcracks of the machined material, as a result of capillary wedging it promotes rupture of the material. A 3 to 5% aqueous solution of calcined soda or NIIALMAZ fluid (0.6% trisodium phosphate, 0.3% borax, 0.25% calcined soda, 0.1% sodium nitrate and 98.75% water) is used as the cooling and lubricating fluid. The spent cooling and lubricating fluid with waste material is discharged by centrifugal force outside through the lateral openings in the drum.

FOR OFFICIAL USE ONLY

The cutting conditions -- the disc rpm, the ingot feed rate, the cooling and lubricating fluid consumption -- depend on the properties of the machined material, its dimensions, thickness of the cut plates, the requirements on finish and precision of the machining. Increasing the disc rpm increases the cutting rate, but vibrations at temperature rise and, consequently, the precision and finish of the machining diminish. With an increase in feed rate of the ingot, the precision and quality of the machining decrease. In addition, at high feed rates the disc bends which influences the shape of the cut-off plates or leads to breaking of them. The cooling and lubricating fluid consumption is selected so that it will exit through the side openings of the drum and the cut plates will not break under the pressure of the powerful jet when hitting the collector.

Thus, in order to insure high quality of the plates with maximum output capacity, the cutting conditions are made optimal.

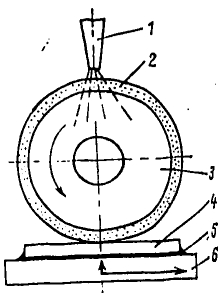


Figure 3-2. Diagram of disc cutting with outside diamond-containing cutting edge.

1 -- cooling and lubricating fluid feed nozzle; 2 -- cutting edge of the disc; 3 -- disc base; 4 -- cut plate; 5 -- bonding material; 6 -- mandrel for attachment of the plate

The recommended cutting conditions are as follows: linear velocity of the disc 17-22 m/sec corresponding to a tool spindle rpm of 4000 to 5000; ingot feed no higher than 30-40 mm/min for silicon and 40-50 mm/min for germanium and gallium arsenide; cooling and lubricating fluid consumption 2.5-4 liters/minute.

The cutting by discs with inside diamond cutting edge also has deficiencies: complexity of pulling up the disc, dependence of the quality and precision of machining on the fastening of the ingot, strengthened quality of the tool.

Cutting by disc with outside diamond cutting edge is realized in accordance with the diagram presented in Figure 3-2. The cutting mechanism is analogous to the disc cutting mechanism with inside cutting edge. The disc with the outside cutting edge is fastened to the machine tool spindle by its central part. This fastening does not insure high rigidity of the disc. Under the effect of cutting forces the disc can bend sharply in the transverse direction and vibrate in the axial direction, which promotes the formation of defects on the surface of the cut sample and also an increase in width of the cutting. In order to increase rigidity, the discs with outside cutting edge are made on a thicker base (0.3-0.7 mm). The

FOR OFFICIAL USE ONLY

minimum width of cut of silicon is about 0.45 mm, the machining of the surface corresponds to roughness class 6-7. It is inexpedient to use such discs to cut large-diameter ingots.

Discs with outside cutting edge are used to cut the substrates of the required dimensions from sheet billets of glass, pyroceram, ceramic, polycor and also to separate semiconductor plates and dielectric substrates with finished structures into crystals and plates. In this case the depth of the cut is small, and discs with a thinner base can be used. In order to increase the output capacity frequently several cutting discs separated by inserts are used.

Free Abrasive Cuttings

When cutting by means of a free abrasive the mechanical forces are transmitted to the suspension grains using a moving tool: disc, steel web, wire.

Cutting by a disc using abrasive suspension is less efficient, and the recovery of the materials from the waste is more complicated than when cutting by diamond discs. The basic advantage of the method is the better quality of the machining of the surface, for the abrasive micropowders are not as hard as diamonds.

The abrasive suspension is continuously fed to the working zone. With a ratio of liquid and solid phase of the suspension of 3:1 or 4:1, all of the abrasive grains are in the suspended state, and they are uniformly distributed in the cutting zone. Special cooling of the disc is not required in this case, for the heat is removed by the liquid part of the suspension. The disc delivers the abrasive grains to the cutting zone. Turning at high speed, it ejects the grain, and the grain hits the surface of the material with great force. Brittle rupture takes place, but in practice there is no cutting off of the protruberances of the machined surface.

The cutting by steel blades (Figure 3-3) fastened by means of inserts to a special holder and undergoing reciprocal motion, is comparatively rarely used, for it does not provide high output capacity or high quality of machining. The cutting speed depends on the speed of movement and length of stroke of the blades, the size, shape and hardness of the abrasive grains, the abrasive concentration in the suspension, the hardness of the machined material, the pressure of the blades of the ingot, the number and thickness of the steel blades. For cutting silicon ingots usually a suspension based on silicon carbide with M20 or M28 grain size is used; for cutting silicon plates, M10-M20.

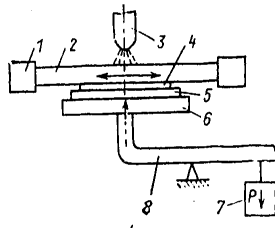


Figure 3-3. Diagram of steel blade cutting.
1 -- holder; 2 -- steel blade; 3 -- suspension feed nozzle; 4 -- cut plate; 5 -- insert; 6 -- table; 7 -- weight; 8 -- lever

FOR OFFICIAL USE ONLY

The rigidity of the blades is low. During the cutting process the blades wear, and their tension decreases. The paired inserts determining the thickness of the cut plate or crystal width are difficult to make in identically exact sizes.

Wire cutting is used to cut plates and substrates. The wires wear quickly. Therefore in addition to reciprocal movement the wire also is rewound. The method is distinguished by low efficiency and breaks in the wire. The advantages of the method include the following: better quality of cutting than when cutting with blades, the possibility of cutting without damaging the films applied to the blades and substrates; small width of the cut (0.08 to 0.2 mm); the possibility of rapid replacement of the worn wire.

Ultrasonic cutting is carried out using abrasive suspension and a tool that undergoes reciprocal vibrations at ultrasonic frequency. These vibrations are received from a magnetostrictive emitter (Figure 3-4), which is a core with a winding assembled from sheets of ferromagnetic material (nickel, permalloy, and permendur). AC voltage of ultrasonic frequency is fed to the core winding. The alternating current flowing through the core windings creates a variable magnetic field causing the magnetostriction phenomenon, that is, the conversion of the electromagnetic field oscillations to mechanical vibrations of the core. In order to increase the amplitude of the core vibrations and the energy concentration on the tool, a concentrator is attached to the core. The vibrations are transmitted from the concentrator to the tool. The suspension is continuously fed to the machining zone. The cutting speed depends on the frequency of the ultrasonic vibrations (the tool vibration amplitude) and also the parameters of the abrasive suspension and the machined material.

During ultrasonic cutting, the shape and size of the tool are copied on the machined sample, and as a result it is possible to cut out crystals of complex configuration, for example, circular ones, and also to perform embossing and obtaining lines, grooves, craters and holes which is impossible in other methods of mechanical cutting. During the cutting process the tool must be raised and lowered periodically, for as it penetrates into the machined material the exchange of abrasive suspension and removal of waste become complicated.

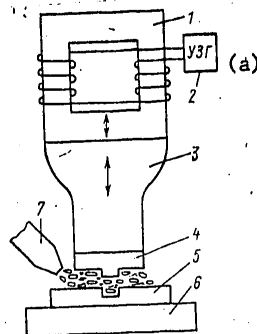


Figure 3-4. Ultrasonic cutting system.

1 -- magnetostrictive emitter; 2 -- ultrasonic voltage generator;
3 -- concentrator; 4 -- tool; 5 -- machined plate; 6 -- table;
7 -- abrasive suspension feed nozzle

Key:

a. ultrasonic generator

FOR OFFICIAL USE ONLY

Ultrasonic cutting is used to separate plates and substrates and to obtain shallow reliefs. Ultrasonic cutting is highly efficient, but it does not provide high quality machining.

3-3. Grinding and Polishing Billets for IC Structures

Grinding. After cutting, the plate and substrate billets have dispersion with respect to thickness, errors in shape and significant surface layer with disturbed structure with respect to thickness. In order to improve the precision and quality of machining the billets, the following finishing operations are performed -- grinding and polishing. Grinding is the finishing of the billets by machining on solid disc grinders made of cast iron, steel, glass and other materials by suspensions with abrasive grain size from 28 to 3 microns or using diamond grinding discs with grain size of 5 microns. The billets are ground in several steps with successive application of finer and finer abrasive grains and, correspondingly, with gradual improvement of the quality of surface machining. During the grinding of silicon, class 7 to 12 surface roughness is obtained.

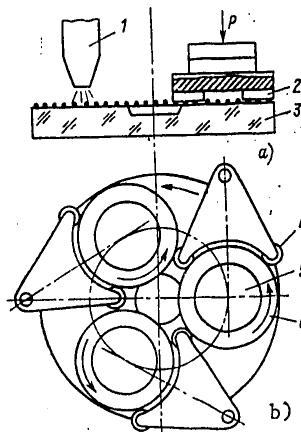


Figure 3-5. Diagram of one-way grinding of substrates (a) and the head location of the grinder (b).

1 -- suspension feed nozzle; 2 -- substrate; 3 -- grinder;
4 -- roller; 5 -- weight; 6 -- grinding head

Grinding with a free abrasive depending on the type of machine tool can be accomplished either one-sided or simultaneously from two sides.

In one-sided grinding (Figure 3-5) the billets are glued to special heads that move freely and are held against the surface of the grinder only by the rollers. In order to obtain a tighter fit of the heads against the grinder surface, weights are used. An abrasive suspension is fed to the machining zone from a batching unit. Under the effect of frictional forces, the motion of the grinder is transmitted to the grinding heads. An interlayer of abrasive suspension is formed between the surfaces of the machine billets and the grinder. During the rolling and turning, the suspension grains act on the machined surface, forming cracks. The chip size

FOR OFFICIAL USE ONLY

from the material and, consequently, the grinding speed depend on the size and hardness of the abrasive grains and the pressure on them from the grinder. During grinding the surface of the grinder and the grinding head become worn; therefore they are periodically checked, and if necessary the surfaces are polished. For one-sided grinding without bonding the billets are stacked in recesses in special separators. It is possible to grind billets of different thickness simultaneously by this method. Grinding without bonding provides better shape precision and high quality of surface machining, but as a result of splattering of the suspension it cannot be used at high grinder rpm.

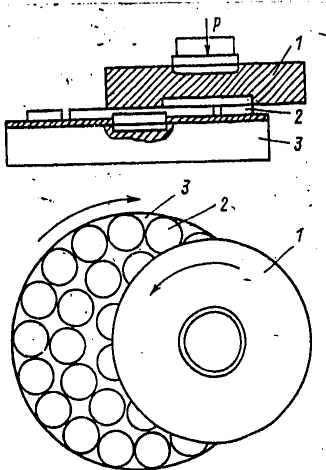


Figure 3-6. Diagram of diamond disc grinding.
1 -- diamond disc; 2 -- substrates; 3 -- grinding head

Two-sided grinding by a free abrasive is accomplished using two grinders. The lower grinder is usually stationary, and the upper grinder freely self-adjusts to the machined billets placed in separators. The separators are rotated by pinions around their own axes and the grinder axis. The billets entrained by the separators undergo complex movement with respect to the grinder surfaces. Two-sided grinding by a free abrasive is more efficient, it insures high surface machining precision and does not require bonding. The bending of the billets is decreased during two-sided grinding, for the residual mechanical stresses are more uniformly distributed.

Diamond disc grinding is performed in accordance with the diagram presented in Figure 3-6. The table to which the billets are fastened and the diamond discs are driven by different electric motors. A coolant is fed to the table surface. Bond abrasive grinding is the most efficient method insuring high roughness class of the surface, but as a result of the impact effect of the grains against the machined material it leaves a quite deep surface layer with disturbed structure. In addition, the rigid fastening of the table and grinder shafts leads to wedge shaping of the plates and substrates if the machine tool wears.

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

Polishing. The plate and substrate billets are polished using soft finishing polishing discs. Fabrics (cambric, velour, chamois, felt, and synthetic fabrics) are stretched on an ordinary grinding disc and fastened by a clamp for this purpose. Polishing is carried out in several steps, just as is grinding.

Preliminary polishing of silicon plates is realized by diamond suspension with grain size of no more than 3 microns using polishers based on fabrics with bulges. The fabric must not wrinkle when it is stretched on the disc. Polishing is accompanied by plastic deformations of the surface layers of the plates and high heat generation. In order to prevent burning of the liquid part of the suspension, softening of the adhesive and rupture of the plate, the polishing is done at low rpm of the polisher. The surface of the machined plates has a fine network of lines ("diamond background") arising under the effect of the sharp cutting edges of the diamond grain. In order to remove the "diamond background" and decrease the surface roughness, final fine polishing is carried out by mechanical, chemical-mechanical or chemical-dynamic methods.

Fine mechanical polishing is accomplished by soft polishing compounds based on aluminum, silicon, chromium, zirconium and other oxides with grain size of less than 1 micron using polishers with sleazy materials in which submicron powder grains are "submerged." This decreases the working surface of the grains and improves the quality of the surface machining of the plate.

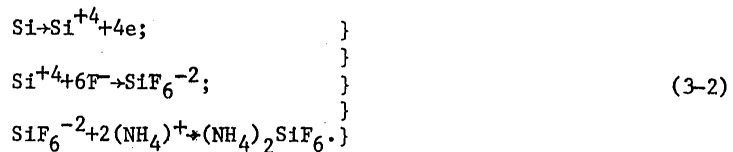
Chemical-mechanical polishing is distinguished by the fact that in addition to the ordinary abrasive mechanical effect the machined surface is subjected to the additional effect of chemical that intensified the process of removing the material.

Polishing of silicon with the application of aerosyl (silicon dioxide) or zirconium dioxide as the abrasive is carried out using water suspension with the addition of alkali (pH=10 to 11). At pH<9, the abrasive effect of the suspension predominates, and the quality of the polishing becomes worse. At pH>11, the silicon surface is aggravated. In order to insure more uniform distribution of the micropowder, small amounts of surface-active materials (ethylene glycol, liquid glass) are added to the suspension.

Chemical-dynamic polishing is a method in which the abrasive effect on the machined surface is entirely excluded. As an example it is possible to consider the polishing of silicon substrates by bivalent copper ions. A solution containing copper nitrate and ammonium fluoride is continuously fed to the chamois polisher. A layer of pure copper is precipitated from this solution on the machined surface



and the silicon dissolves simultaneously with the formation of water-soluble silicates:



FOR OFFICIAL USE ONLY

The copper layer is removed mechanically by the moving polisher, and the copper is again precipitated on the clean machined surface, and so on. The polishing of the silicon by copper ions is highly efficient (100 microns/hour), but it is distinguished by difficulty of removal of the copper residue from the surface of the plates.

3-4. Control of the Plates and Substrates After Machining

Defects and Criteria for Evaluating the Quality of Plates and Substrates

The surface condition of plates and substrates, variations in its shape and size have a significant influence on the subsequent technological operations, and to a significant degree they determine the quality and percentage yield of usable microcircuits.

The nonparallelness of the size of plates and substrates is estimated by difference in their thickness h_1-h_2 in a given length l (Figure 3-7, a). Nonplanarity Δh is the greatest distance from the points of the real surface to the ideally plane surface (Figure 3-7, b). The bending f is the greatest distance from the points of the actual profile in radial cross section to the corresponding flat surface of the adjacent profile (Figure 3-7, c).

The quality of surface machining is characterized by the depth of mechanically disturbed layer and roughness.

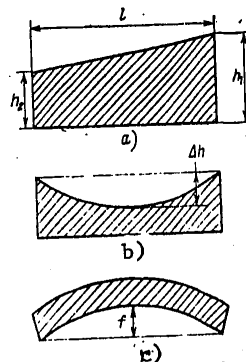


Figure 3-7. Deviations of the substrates from the precise form.
a -- nonparallelness of sides -- wedge shape; b -- nonplanarity;
c -- bending

A mechanically disturbed layer consists of three parts (Figure 3-8): the disturbed relief layer has randomly arranged protrusions, cracks and punctures; a cracked layer has individual noncrumbling punctures and microcracks running deep into the layer; the deformed layer has dislocation pileups, continuations of microcracks and the mechanical stress zones arranged around them.

FOR OFFICIAL USE ONLY

Roughness (Figure 3-9) can be estimated by the arithmetic mean deviation of the profile R_a , that is, the arithmetic mean absolute values of the profile deviations within the limits of the base line z :

$$R_a = \frac{1}{n} \sum_{i=1}^n |y_i| \quad (3-3)$$

or the height of the unevennesses in the profile R_z with respect to ten points:

$$R_z = \frac{1}{5} \left(\sum_{i=1}^5 H_{i \max} + \sum_{i=1}^5 H_{i \min} \right) \quad (3-4)$$

Depending on the sizes of the parameters R_a and R_z , the quality of the surface machining is estimated by the corresponding roughness class.

Control of Plates and Substrates. The controllable parameters are the thickness of the substrates, planarity, nonparallelness of the sides (wedge shape), bending, thickness of the mechanically disturbed layer, roughness, the presence of scratches, chips and lines.

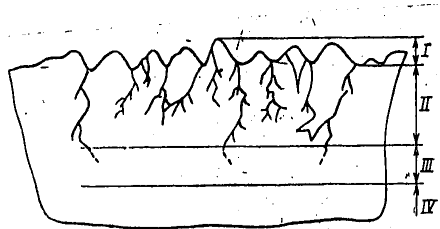


Figure 3-8. Mechanically disturbed layer of the plate surface. I -- relief layer; II -- cracked layer; III -- deformed layer; IV -- undisturbed structure of the plate

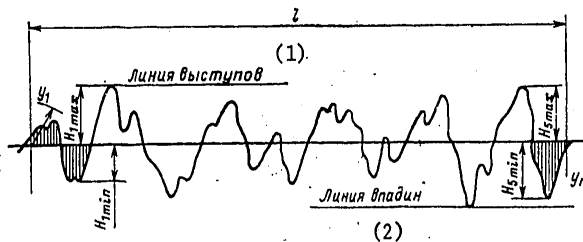


Figure 3-9. Profile of the rough surface of a substrate

Key:

1. protrusion line
2. depression line

FOR OFFICIAL USE ONLY

The thickness is measured by a clock type indicator which is fastened to a stand. The substrate is placed on the table. The indicator is adjusted so that its zero will correspond to the position of the measuring probe on the surface of the table. The thickness is measured at several points, and then the mean value is determined. The precision of the indicator measurement is 1 micron. Higher precision (0.5 micron) is provided by the IZV-2 optical length gauge. The clock indicator and length gauge can also be used to determine the wedge shape of the substrates.

The nonplanarity is determined by the plane-parallel standard glasses which are applied to the polished surface of the attached plate or substrate. In the sections where there is an air space, an interference pattern arises as a result of superposition of light beams reflected from the controlled surface of the adjacent surface of the standard glass, by which the nonplanarity is judged.

It is possible to measure the bending of plates and substrates by a toolmaker's straightedge during observation by a microscope. The size of the clearance between the straightedge and the bent plate can be measured with precision to 1.5 microns.

The thickness of the mechanically disturbed layer can be measured by various methods, the simplest of which is based on using the dependence of the speed of chemical pickling on the degree of the disturbance of the crystal. As the mechanically disturbed layer is pickled away, the pickling speed decreases. The time at which the pickling rate of the single crystal becomes constant signals removal of the entire disturbed layer.

The surface roughness within the limits of class 12-14 is determined using the MII-4 microinterferometer. At the locations of the microunevennesses, the interference bands are distorted, and it is possible to determine the magnitude of the irregularities by the degree of distortion.

The presence of scratches, lines, chips and traces of contamination of the surface can be detected by observation using the MBS-1 and MBS-2 stereoscopic microscopes or the MIM-7 metallographic microscope. The dimensions of the indicated defects can be measured using the scale of the eyepieces. The microscopes have a comparatively small field of view and do not permit the entire surface to be monitored, especially the defects such as the smooth irregularities with small height gradients, sags along the edges of the plates and the bends. These defects can be monitored by irradiation of the polished surfaces by parallel or diverging laser beams and subsequent analysis of the interference pattern.

Test Questions and Assignments

1. Repeat the questions from the adjacent topics: "Materials for IC substrates," "Materials for machining," "Requirements on IC substrates."
2. What is abrasive machining? What is the role of the tools during abrasive machining?
3. What abrasive materials are used for machining?
4. What explains the fact that the grain size of the powders in Table 3-1 is designated in accordance with different All-Union State Standards?

FOR OFFICIAL USE ONLY

5. Decipher the abrasive designations EBM20, KZM28, ASMZ/2.
6. Explain what the free and bound state of an abrasive is.
7. Which production operations belong to machining?
8. For what materials is preliminary crystallographic orientation required? Why?
9. What is the role of grinding and polishing during machining of billets?
10. What is the role of the quality of attaching the ingots and plates?
11. What are the advantages of cutting by a disc with inside diamond cutting edge? What are they caused by?
12. What is the cutting mechanism by a disc with inside diamond cutting edge?
13. What is the practical implementation of cutting by a disc with inside diamond cutting edge?
14. What is the role of the cooling and lubricating fluid in the cutting process?
15. Which parameters determine the disc cutting conditions?
16. From which arguments are the optimal values of the disc rpm, ingot feed rate and cooling and lubricating fluid flow rate selected?
17. What are the deficiencies of cutting by a disc with external cutting edge?
18. Compare the methods of cutting by a disc with external and internal cutting edges.
19. What are the differences in the cutting by free abrasive from cutting by bound abrasive?
20. Which tools are used for free abrasive cutting?
21. Why is the cutting by blades and wire rarely used in IC production for cutting ingots?
22. What is ultrasonic cutting?
23. What is the practical implementation of ultrasonic cutting?
24. What is the uniqueness of ultrasonic cutting by comparison with other methods?
25. Compare all the cutting techniques from the point of view of application, efficiency and quality.
26. What is mechanical grinding and polishing? Is there a theoretical difference between them?

FOR OFFICIAL USE ONLY

27. Why are mechanical grinding and polishing performed in several steps?
28. What is the practical implementation of free abrasive grinding?
29. Give a description of the method of diamond disc grinding.
30. What distinguishes a polisher from a grinder?
31. What is preliminary polishing and how is it done?
32. Which methods are used for final polishing?
33. Compare chemical-mechanical polishing with precision mechanical polishing.
34. How is chemical-mechanical polishing of silicon realized?
35. Compare the chemical-mechanical polishing of silicon by aerosyl with silicon polishing by copper ions.
36. What are the basic shape and surface defects of plates and substrates?
37. Enumerate the monitored substrate parameters after machining and explain how they are determined.
38. Enumerate the factors influencing the efficiency and quality of machining.
39. What are the deficiencies of the machining method?

FOR OFFICIAL USE ONLY

CHAPTER 4. CHEMICAL TREATMENT AND CLEANING OF THE SURFACE OF SEMICONDUCTOR PLATES AND SUBSTRATES

4-1. General Information

Concept of the "Clean Surface." The basic volume of cleaning operations in microcircuitry technology pertain to preparation of the surface of the plates (substrates) for the structure manufacturing process. Atoms of the plate (substrate) material located on the surface have many more unsaturated bonds than atoms in its body. This gives rise to high chemical activity on the surface. It is in practice impossible to obtain an ideally clean surface, that is, a surface without foreign impurities.

Under microcircuit production conditions the plates and substrates are in contact with various media, and it is impossible to completely shield them against the adsorption of impurities; therefore the concept of "clean surface" used has a relative nature. A surface which has an impurity concentration that does not prevent reproduction of the given values and stability of the microcircuit parameters is considered to be technologically clean. Even in the case of nonrigid requirements the impurity concentration on such a surface must not exceed 10^{-8} to 10^{-7} g/cm².

The processes of cleaning plates (substrates) are intended for the removal of contamination to a level corresponding to the technologically clean surface. The goals of cleaning monocrystalline plates (substrates) also include the removal of the surface layer with the structure disturbed during the machining process. The presence of a mechanically disturbed layer does not allow high-quality and reproducible semiconductor and film structures to be obtained on semiconductor plates, sapphire and spinel substrates.

Sources of Contamination of Plates and Substrates. The basic sources of contamination of the surfaces of plates and substrates are the following: abrasive and adhesive used during machining; dust in the air of the production facility; objects that have come in contact with the plates and substrates (equipment, tools, fittings, packaging, transportation and storage); production environments; organic and inorganic reagents, water, and so on.

Contamination of the plates and substrates is possible in practice in all operations of the manufacture of microcircuits. Therefore throughout the entire technological production process the cleaning of the surfaces of the plates and substrates is realized again and again. The most important processes are the processes of

FOR OFFICIAL USE ONLY

cleaning the surface after machining, before thermal processes in which the diffusion of the impurity atoms and the probability of the formation of undesirable chemical compounds and alloys before the application of various types of coatings, films, layers, and so on increase.

Types of Contamination. The technology of chemical treatment and cleaning of plates and substrates is determined in each step by the nature of the possible contamination and requirements imposed on the surface. It is expedient to classify the possible contamination on the surface of plates and substrates by their physical-chemical properties, for the latter basically determine the choice of methods of removing the contamination:

Organic contaminants are primarily greasy nonpolar contamination by adhesives, the oil from the machine tools, the operators' hands, and so on.

Water-soluble polar contaminants include salts, acids, pickling agent residues, fluxes, and so on.

Physical contaminants include dust particles, hair, abrasive and other foreign particles not chemically bound to the surface of the plates and substrates.

Contaminants chemically bound to the surface of the plates and substrates include oxide films and other compounds.

Gases adsorbed by the surface.

Many types of contaminants can be present on the surface of the plates and substrates simultaneously. The most difficult to remove are the organic and some contaminants chemically bound to the machined surface. Among the physical contaminants it is most difficult to remove small abrasive grains introduced into the surface layer. Among the water soluble polar contaminants it is difficult to remove the mobile metal ions which change the electrical conductivity of the surface causing current drift and the appearance of inversion semiconductor layers and at the same time having a harmful influence on the stability of the IC parameters.

Classification of the Cleaning Methods. From the point of view of the mechanism of the processes, all of the cleaning methods can be provisionally divided into physical and chemical methods (Figure 4-1). In the physical methods the contaminants are removed by solution, annealing and also treatment of the surface by ions of inert gases accelerated to high energy. In cases where it is impossible to remove the contaminants physically, chemical methods are used in which the contaminants located on the surface or in the surface layer are converted to new chemical compounds and then easily removed. The cleaning given by the removal of the surface layer of the plate and substrate is called pickling.

In accordance with the applied means, the cleaning can be divided into liquid and dry. It is very complicated to select a liquid medium which simultaneously removes all possible surface contaminants; therefore the liquid cleaning includes a number of successive operations. Water-insoluble greasy organic contaminants make the surface hydrophobic, that is, poorly wet by water and the majority of solvents. For uniform cleaning the surface of substrates must be converted to hydrophilic, that is, a state that is well wet by water. The operation of removing the greasy

FOR OFFICIAL USE ONLY

contaminants accompanied by conversion of the surface from the hydrophobic state to the hydrophilic state is called degreasing.

Dry cleaning, as a rule, is the finish cleaning, for it is done directly before performance of the subsequent technological processes, for example, before spraying on the films, oxidation, photolithography, and so on. Dry cleaning includes annealing, gas, ion and plasma-chemical pickling. These methods exclude the necessity for the application of expensive liquid reagents dangerous to work with and also the problems of the interoperations storage of the plates and substrates and cleaning of the waste water which are of little importance when using liquid cleaning means. In addition, the dry cleaning processes are more controllable and are easily automated.

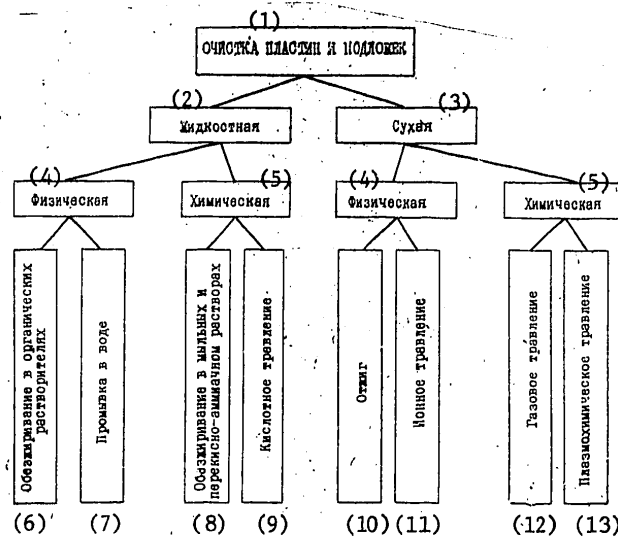


Figure 4-1. Classification of methods of cleaning plates and substrates.

Key:

- | | |
|--------------------------------------|---|
| 1. Cleaning of plates and substrates | 8. Degreasing in soap and ammonium peroxide solutions |
| 2. Liquid | 9. Acid pickling |
| 3. Dry | 10. Annealing |
| 4. Physical | 11. Ion pickling |
| 5. Chemical | 12. Gas pickling |
| 6. Degreasing in organic solvents | 13. Plasma-chemical pickling |
| 7. Washing in water | |

FOR OFFICIAL USE ONLY

4-2. Methods of Liquid Treatment of Plates and Substrates

Degreasing. Almost all greasy contaminants are effectively dissolved in organic solvents (carbon tetrachloride, benzene, toluene, isopropyl alcohol and others). When the specimens are submerged in the solvent, the grease molecules are gradually separated from their surface, which as a result of diffusion are uniformly distributed over the entire volume of the solvent. The separation of the grease molecules from the treated surface arises from their natural vibration movements and attraction on the part of the solvent molecules. Simultaneously with solution, the opposite process occurs -- adsorption of the grease molecules by the cleaned surface. In order to decrease the return contamination, the largest possible lot of solvent is used for the given number of plates or substrates. The process is carried out successively in several lots of fresh solvent. If the solvent is distilled, the degreasing is carried out in the last step in the solvent vapor. The solubility of the greases increases with an increase in temperature, and therefore the degreasing is carried out in hot or boiling organic solvents. For a given number of samples and a given lot of a specific solvent the controlled parameters of the degreasing process are the solvent temperature and the treatment time.

The plates and substrates are very effectively cleaned in organic solvents. Nevertheless, the application of organic solvents in production is undesirable. The grease molecules go into solution without chemical destruction and can again get on to the cleaned surfaces. Therefore multiple cleaning is required accompanied by large lots of solvent. As a result of the high requirements on the purity of the solvents themselves, they are subjected to multiple redistillation which also increases the cost of the process. Many of the solvents have low boiling points and high vapor pressure; therefore the degreasing is accompanied by large losses of solvent; in addition, organic solvents are frequently toxic and fire-hazardous.

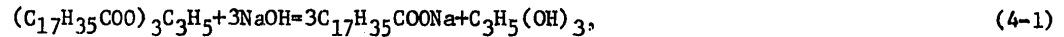
Freons which have now been widely introduced into production are advantageously distinguished from the above-presented organic solvents. Freon-113 -- a liquid with a boiling point of 47.6°C and a density of 1.57 g/cm³ -- is the most frequently used. Freon is incombustible and nontoxic; it provides high cleaning efficiency and makes it possible to do away not only with many of the organic solvents, but also the application of successive operations of chemical pickling and washing in deionized water.

Chemical degreasing is done in compounds that destroy the grease molecules and do not affect the treated material. The absence of grease molecules in the treated solution and, consequently, the absence of the probability of return contamination theoretically distinguish chemical degreasing.

In microcircuit production sometimes the surfaces are treated in soap solutions, which is used to convert the saponified greases to soaps which are water-soluble salts. The latter are removed from the surfaces of the plates and substrates by washing. The saponified greases include all of the vegetable and animal oils which are complex esters of glycerine and high-molecular organic acids (stearic, oleic, palmitic, and so on).

FOR OFFICIAL USE ONLY

For example, the process of saponification of stearin takes place by the equation



where $C_{17}H_{35}COONa$ is sodium stearate (soap); $C_3H_5(OH)_3$ is glycerine. Both materials are easily soluble in water.

At the present time hot (75-80°C) "universal" ammonia peroxide solution is used for chemical degreasing of silicon plates. It consists of an aqueous solution of a mixture of H_2O_2 Perhydrol and NH_4OH alkali. It removes saponified and unsaponified greases.

During degreasing Perhydrol decomposes with the release of atomic oxygen:



The release of atomic oxygen increases with an increase in temperature. Atomic oxygen oxidizes both organic and inorganic contaminants. The NH_4OH alkali accelerates the reaction of the composition of Perhydrol, and it also combines the compounds of certain metals of first and second group of the periodic table into well-soluble complexes.

By comparison with physical degreasing in organic solvents the mechanical degreasing is distinguished by lower toxicity, low cost of the reagents and less labor consumption of the processes.

Pickling. The pickling of the plates and substrates, as has already been noted, is accompanied by the removal of the surface layer together with which the contaminants available on the surface are also removed. The pickling is usually carried out after degreasing, for only in this case does the pickling agent wet the entire surface well and the upper layer is uniformly removed. The pickling is a mandatory production operation when preparing monocrystalline plates and substrates, for which it is done not only to clean the surface, but also to remove a layer with mechanically disturbed crystalline structure.

Acid pickling of the semiconductors proceeds in several steps in accordance with chemical theory: the diffusion of the reagent to the surface, adsorption of the reagent by the surface, surface chemical reactions, desorption of the reaction products and diffusion of them from the surface.

The pickling agents in which the diffusion stages are the slowest stages determining the overall pickling process are called polishing pickling agents. These pickling agents are insensitive to physical and chemical nonuniformities of the surface; they smooth out the roughness, leveling the microrelief. The pickling rate in polishing pickling agents essentially depends on the viscosity and mixing of the pickling agent and depends little on temperature.

The pickling agents in which the slowest reactions are the surface chemical reactions are called selective. The pickling rate in selective pickling agents

FOR OFFICIAL USE ONLY

depends on the temperature, the structure and the crystallographic orientation of the surface, and it does not depend on the viscosity or mixing of the pickling agent. The selected pickling agents with great difference in pickling rates and different crystallographic directions are called anisotropic.

Polishing pickling agents with high and stable pickling rate, low impoverishment and stability during storage are used to prepare the plates. Pickling agents based on a mixture of nitric acid and hydrofluoric acid have such properties for silicon and germanium.

In accordance with chemical theory, the surface chemical reactions take place in two stages during polishing pickling: oxidation of the surface layer and conversion of the oxide to an easily soluble salt.

The role of the oxidizing agent is played by nitric acid:



Hydrofluoric acid is a complex former which converts the silicon dioxide to silicon tetrafluoride:



The compositions of the pickling agents in the majority of cases are selected empirically. The composition of the pickling agent, in addition to the basic components, nitric acid and hydrofluoric acid, include various additives. For example, the most frequently introduced acetic acid retards the chemical reactions (4-3) and (4-4) and, consequently decreases the overall pickling rate. In the absence of acetic acid in a pickling agent the reactions proceed too rapidly, and the pickling process is difficult to control.

Let us note that the pickling is also used for dimensional treatment of semiconductor plates, for example, to bring the thickness to a given value, to obtain local depressions or through holes. Selected pickling of semiconductors is used to investigate surface defects, to discover the p-n junctions, for controlled formation of the depression of a defined shape. Selecting the defined surface orientation and also the orientation of the hole in the contact mask, it is possible to obtain completely defined shapes of depressions corresponding to them.

Washing the Plates and Substrates. During the manufacture of the microcircuits, the surfaces of the plates and substrates are washed several times. Especially pure deionized or distilled water is used for the washing. Careful washing of the silicon or dielectric substrates is necessary after degreasing them in soap solutions or ammonium peroxide solvent. Certain organic solvents remove only nonpolar greasy contaminants from the surface of the substrates; others can remove polar (ionic) contaminants also. Trichloroethylene, trichloroethane, perchloroethylene, trichlorotrifluoroethane, for example, remove only nonpolar molecules; therefore after degreasing subsequent washing of the polar contaminants in water is required.

FOR OFFICIAL USE ONLY

The forced method is used to wash silicon after pickling. During forced washing the pickling agent is not completely drained off so that the plates do not contact the air; deionized water is added to the bath, diluting the remains of the pickling agent and gradually forcing it out. Otherwise, that is, after pouring off the pickling agent completely or removing the plate, chemical reactions will continue at the points where the drops of pickling agent remain. As a result, the surface of the plates will not be uniformly smooth. During forced flushing it is necessary to consider the possibility of splattering of the acetic pickling agent.

4-3. Intensification of the Cleaning Processes

Classification of Intensification Techniques. In order to insure efficiency and quality of the treatment of the surface plates and substrates, various intensification procedures are used. For intensification the slower stages of the process (for example, the supply of fresh reagent to the treatment zone, the removal of the products of the chemical reactions from the treated surface) are accelerated, the desorption of the atoms or ions is insured, solution is accelerated, fast penetration and subsequent removal of the treating solutions from the microcracks, and so on are insured. The intensifying means can be divided into physical, chemical and combination. The first-mentioned include heating, boiling, treatment with a jet, hydrocirculation, flushing, hydromechanical cleaning, centrifuging, ultrasonic treatment, industrial-frequency vibrations, and plasma. Chemical means include surface-active materials, complex formers and catalysts.

Physical Methods of Intensification. Treatment during heating or boiling in organic solvents, acids, deionized water are carried out in exhaust hoods or enclosures using a quartz or fluoroplastic dish or bath. For example, the UTU-1 pickling unit has three baths: one for pickling with water heating, the other two for preliminary and final washing in running deionized water. A deficiency of the cleaning by heating and boiling is treatment in a constant volume of reagent which leads to repeated contamination of the substrates; to great losses of the reagents as a result of evaporation, and, in addition, the intensification of the cleaning processes is inadequate.

Jet cleaning can be carried out using pneumatic or centrifugal jets operating at a pressure of 4 to 5 atmospheres. The advantage of jet cleaning is continuous replacement of the reagent and acceleration of the removal of the reaction products as a result of the hydraulic effect of the jet. A deficiency of the method is the high rate of consumption of reagents. The reagent consumption rate can be lowered by circulating, filtering and regeneration of them, that is, utilizing a closed technological process. The effectiveness of the treatment can be improved by increasing the jet pressure.

Hydromechanical cleaning consists in mixing a solution or water by mixers, brushes, and the mechanical effect of them on the treated surface. All of this promotes renewal of the solution directly on the surface, it increases the wettability of the surface and mechanically knocks out the great molecules and other impurities, acting as mechanical "scrapers" on the cleaned surface. The deficiencies of hydromechanical cleaning include the possibility of return transfer of the contaminants from the brushes to the substrates, the probability of contamination as a result of wear of the brushes, the necessity for periodic careful cleaning of the brushes and mixers themselves.

FOR OFFICIAL USE ONLY

Centrifuging promotes mixing and renewal of the liquid near the treated surfaces. The substrates or plates are fastened in a holder which, in turn, is fastened to the spindle of the centrifuge turning at about 200 rpm. Increasing the centrifuging rpm improves the cleaning quality. Drying is possible at higher speeds (2400-2500 rpm). The centrifuged drying removes the drop moisture well, which is necessary for removal of the infinitesimal proportion of the solid sediment in deionized water with the water drops.

Ultrasonic cleaning is the most effective of the indicated physical methods. The bath of the ultrasonic unit is attached to the concentrator of the magnetostrictive emitter. The vibrations of the concentrator are transmitted to the bath walls, and from them, to the liquid medium. Elastic waves arise in the liquid medium (punching and rarefaction as a result of shifting of the liquid particles). At the rarefaction points in the liquid small bubbles appear which are filled with liquid vapor called cavitation bubbles. Under the effect of forces which try to return the shifted particles to the initial position, these bubbles collapse after brief existence. In the case of intense vibrations and collapse of the cavitation bubbles, shock waves arise communicating high accelerations to the liquid molecules. The liquid molecules collide with a force against the surface of the treated substrates and knock the contaminant particles off them. As a result of cavitation, the liquid is capable of penetrating into the deep pores, channels, depressions and cracks which remain uncleaned when using ordinary methods.

The effectiveness of ultrasonic cleaning depends on the frequency of the intensity of the ultrasonic vibrations, the location of the cleaned surface with respect to the concentrator, the temperature and pressure of the cleaned liquid vapor, and the cleaning time. Ultrasonic cleaning is appreciably more effective if the surfaces of the plates or substrates are arranged perpendicular to the direction of the ultrasonic wave propagation front.

However, as a result of strong mechanical effects, the ultrasonic cleaning must be carefully carried out for thin and brittle plates and substrates and also for finished structures in the precavitation mode, for a short period of time, locating the treated specimens far from the bottom of the bath. The vibrations of industrial frequency used, for example, in the RVKhO-GS60-1 group pickling unit are less dangerous for surface rupture.

Chemical Intensification. The introduction of catalysts, surface-active materials and complex formers into the cleaning process is one of the prospective ways of improving the quality of preparing plates and substrates and the effectiveness of cleaning processes, and in a number of cases it permits elimination of toxic and fire-hazardous organic solvents and also concentrated acids.

The complex formers form stable, complex compounds with the surface impurities or with the harmful products of chemical reactions, which go into solution and remain in solution. The complexes must have the smallest possible dissociation constants; otherwise the impurity found in the complex can be again adsorbed from solution by the treated surface. For example, copper or silver ions, the presence of which on the surface of the semiconductor substrates increases the recombination of the minority carriers and the back currents of the p-n junctions, and they are easily bound into complexes by ammonia:

FOR OFFICIAL USE ONLY



During pickling of the semiconductors potassium bichromate binds the reaction products into easily and quickly soluble complexes. Acetonitrile, ethylenediaminetetraacetic acid (trilon B), and so on have good complex-forming properties.

Combined Intensification. The combined methods of intensification are quite frequently used in practice: the treatment with a hot jet, ultrasonic treatment in solvents, treatment by a hot jet with centrifuging, pickling in a hot solvent with the application of complex-forming additives, and so on.

4-4. Standard Processes of Cleaning Plates and Substrates

Preliminary Cleaning of Plates and Substrates. It was noted earlier that it is in practice impossible to select a universal composition for wet cleaning of the surface of plates and substrates. In addition, any technological operations, including the cleaning operations themselves, can be sources of pollutants. When manufacturing the microcircuits, the billets for their structure are treated more than once, and each cleaning is complex, that is, it includes a number of operations for the removal of all possible contaminants. The nature of the treatment in each operation, the sequence of operations, the means and methods used are determined by the phase in which the cleaning is carried out, the material of the billet and the elements entering into the structure and also the requirements on the quality of cleaning.

After machining, the semiconductor plates and substrates are contaminated with polishing suspension, paste, adhesives and oil from the machine tools. The preliminary cleaning of the plates and substrates is carried out directly on a polishing disc using an aqueous solution of "Nega" or "Lotos" powder, and they are flushed with pure water. The water is removed by centrifuging. Then the discs are heated to melt the adhesive, and the plates or substrates are removed from the disc.

Final cleaning of the plates and substrates is realized by different paths, the basic operations of which are degreasing, polishing pickling, washing and drying.

Processes of Cleaning Silicon Plates. During the production cycle of manufacturing semiconductor IC structures, the substrates are cleaned more than once: after machining, before masking the surface and local treatment, and after photolithography. Let us consider one of the standard process flow charts for cleaning silicon plates.

The cleaning of silicon plates before the first thermal oxidation includes the following operations:

1. Degreasing in a hot (75-80°C) ammonia peroxide solution.
2. Washing in flowing deionized water to remove the products of the chemical reactions of the preceding treatment.

FOR OFFICIAL USE ONLY

3. Treatment in hot (90-100°C) concentrated nitric acid for oxidation of silicon to the required depth and partial removal of the metal ions.
4. Flushing in a flow of deionized water to remove the acid residues.
5. Hydrodynamic treatment of the plates with lime brushes in a jet of deionized water.
6. Drying of the plates in a centrifuge in a jet of purified dry air.
7. Treatment of the plates in a solution of hydrofluoric acid with acetone to remove the oxide film obtained during treatment in nitric acid, and together with it, also the surface contaminants.

Acetone is introduced into the solution to decrease the contamination of the silicon plates by the solid residue of hydrofluoric acid which is formed for an excess of chlorine ions in the solution as a result of chemical reactions:



Acetone has a low dielectric constant; therefore on addition of it, the dissociation of the hydrofluoric acid decreases. In addition, acetone forms complexes with certain cations.

8. Washing in a flow of deionized water until the specific resistance at the exit of the device becomes equal to the resistance of water and the entrance to the device.
9. Ultrasonic treatment in several lots of ammonia peroxide solution for more complete removal of organic and inorganic contaminants left in the microcracks and surface pores.
10. Washing in running deionized water.
11. Drying in a centrifuge.

Standard Process of Cleaning Substrates Made of Glass, Sapphire and Ceramic.
Good results are obtained when using the cleaning process including the following operations:

1. Ultrasonic washing in a solvent at room temperature.
2. Ultrasonic washing in a solvent heated to a temperature of 70°C and in solvent vapor.
3. Washing in running water.

FOR OFFICIAL USE ONLY

4. Boiling in a Perhydrol solution.
5. Washing in hot distilled water.
6. Boiling in distilled water.
7. Drying in a flow of pure nitrogen heated to a temperature of 110°C.

Standard Process of Cleaning Pyroceram Substrates. The pyroceram composition includes amorphous and crystalline phases. Different pickling rates of the different phases lead to the formation of relief on the substrate surface. In order to decrease the roughness, neutral and acid solutions are used which form slightly soluble compounds on the surface of the pyroceram preventing surface deterioration.

The standard process of cleaning pyroceram includes the following operations:

1. Degreasing by boiling in ammonia peroxide solution for 15-20 minutes.
2. Washing in running deionized or distilled water.
3. Washing by boiling in distilled water for 5-10 minutes.
4. Drying in isopropyl alcohol vapor (30 minutes) or in a flow of argon or nitrogen heated to a temperature of 320±30°C.

Freon Cleaning. Cleaning using freon is universal. It can be used for plates and substrates made of any material. Several standard technological processes have been developed for freon cleaning. For example:

1. Treatment with freon emulsion (water+freon+a surface-active material), in which inorganic salts are dissolved in the disperse water drops, and organic contaminants, in freon.
2. Treatment in a mixture of freon with other solvents (isopropyl alcohol, methanol, and methylene chloride) removes the remains of surface-active materials from the substrate surface.
3. Treatment with pure freon.
4. Treatment in freon vapor.

The treatment of silicon using freon-113 in the Soviet "Freon" unit accelerates the cleaning process by about 15 times and insures the same quality as during careful treatment by degreasing, acid pickling and washing.

4-5. Drying Cleaning of Plates and Substrates

Heat Treatment (Annealing). Heat treatment is used to remove the impurities adsorbed by the surface, for decomposition of surface contaminants and evaporation of volatiles. As a rule, annealing is carried out in vacuum and thermal units directly before carrying out the thermal processes. For example, when growing masking films on silicon, the gases and moisture are removed from the surface by

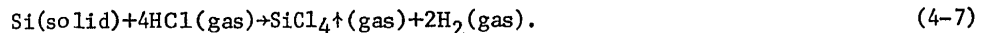
FOR OFFICIAL USE ONLY

heating the plates to the oxidizing temperature. When annealing semiconductor plates in a vacuum, moisture, carbon dioxide and light hydrocarbons are easily desorbed from their oxidized surface at a temperature of 400°C. The effectiveness of the cleaning increases with an increase in temperature, but the treatment temperature is limited by the melting point of the cleaned materials or diffusion of the alloying admixtures. Lower-temperature cleaning processes are used in these cases.

Gas Pickling. The essence of the gas pickling process consists in chemical interaction of the plate material with gaseous substances and the formation of easily removed volatile compounds when this happens. During gas pickling the contaminants are removed together with the surface layer of the plates.

Halogens, hydrogen halides, sulfur compounds and water vapor are used as the reagent gases for pickling silicon plates. Small amounts of these gases are added to the gas carrier (hydrogen or helium) and are transported to the chamber of the unit, in the temperature zone of which the plates are located.

The pickling of the silicon by hydrogen fluoride is widely used before growing the silicon layers on the plates:



Hydrogen fluoride vapor is delivered by hydrogen to the reaction chamber of the epitaxial growing unit where the silicon plates heated to a temperature of 1150-1250°C are located.

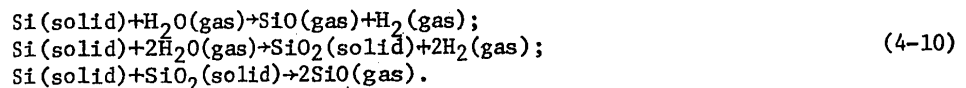
In silicon tetrachloride vapor the pickling of the silicon plates is accompanied by the reaction



In the case of chlorine pickling helium is used as the gas carrier. The pickling takes place at temperatures of about 1000°C and with a chlorine content in the helium of no more than 0.2%. On deviation from the optimal conditions, the pickling agent loses the polishing characteristics, and irregularities appear on the surface of the silicon. Pickling takes place in accordance with the reaction



During pickling in water vapor the following reactions take place:

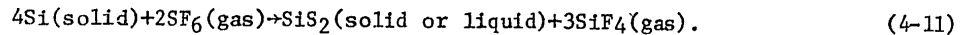


With a water vapor content in the hydrogen of more than 5·10⁻²%, a silicon dioxide film is formed on the surface of the silicon, and pickling stops.

FOR OFFICIAL USE ONLY

It is also possible to carry out the pickling of the silicon in hydrogen sulfide or sulfur hexafluoride vapor. During pickling in hydrogen sulfide, large rates of removal of the surface layers are obtained (to 15 microns/minute). However, the hydrogen sulfide is toxic. Sulfur hexafluoride, on the contrary, is not toxic and insures good quality of the silicon surface and also sapphire.

The pickling of the silicon is accompanied by the reaction



Gas pickling makes it possible to obtain cleaner surfaces than liquid pickling. However, in any case gas pickling has limited application as a result of high treatment temperatures and the necessity for using especially pure gases.

Ionic Pickling. Ionic pickling is the process of removing the contaminants together with the surface layer of treated material sprayed in a vacuum.

For spraying, the surface of the plates or substrates is bombarded with accelerated positive ions of inert gases. Most frequently argon is used for spraying, for it is cheap, plentiful and allows effective spraying. The accelerated ions penetrate the surface layer and, colliding with the atoms of the treated plate or substrate, transfer their energy to them. If the transmitted energy is sufficient, the atoms are shifted from the nodes and can transmit energy to other atoms. Thus, zones of shifted atoms -- radiation disturbances in the substrate structure -- are formed along the trajectory of motion of the ion. Spraying takes place if part of the shifted atoms reach the surface and if the energy of these atoms is of the work function of the substrate material.

From investigation of the mechanism of the effect of the accelerated ions on the substrate it is clear that the pickling begins with defined values of the energies sufficient for spraying. The silicon is pickled with ion current densities of more than 10 amps/m² and an ion energy of 1-10 kev. For very high energies the ions penetrate deeply into the treated substrate and spraying is not observed.

The effectiveness of the spraying and, consequently, pickling, is characterized by the spraying coefficient which is numerically equal to the number of atoms of cleaned substrates sprayed by one bombarding ion. The spraying coefficient S depends on the mass of treated material m_1 , the mass m_2 , energy E and angle of incidence θ of the ion and also the physical state of the surface:

$$S = k \frac{m_1 m_2}{\lambda(E) \cdot (m_1 + m_2)} E, \quad (4-12)$$

where k is the coefficient characterizing the surface state, $\lambda(E)$ is the mean free path length of the ion in the treated material, which depends on θ .

Depending on the structure of the devices and the method of generating the ions, plasma ion pickling and ion beam pickling are distinguished.

Plasma ion pickling is carried out in vacuum spray chambers. The plate or substrate holder is located in a gas discharge plasma. On feeding a negative

FOR OFFICIAL USE ONLY

potential to the holder, positive ions are drawn out of the plasma which are accelerated by an electric field and bombard the surface of the substrate, cleaning it. The pickling of poor-conducting or nonconducting substrates with constant potential on the holder is complicated and stops as a result of the accumulation of a positive ion charge on the substrates. Therefore high-frequency AC voltage is used to pickle them. With a negative potential the substrates are pickled; with a positive potential, electrons are drawn out of the plasma, and the accumulated positive ion charge is neutralized.

The rate of plasma ion pickling is regulated by varying the voltage on the electrodes, the gas-discharge current, the inert gas pressure and the time of performing the process. The pickling rate of silicon reaches 0.2 microns/minute, the substrate temperature is comparatively low (100-120°C) during pickling. A significant advantage of plasma ion pickling is inertialessness, for the pickling process stops immediately after removal of the potential from the substrates. A unique property of plasma ion pickling is the fact that its rate is directed along the normal to the cleaned surface. This permits treatment of strictly defined local sections.

Ion beam pickling is carried out in ion injection units -- ion beam accelerators (ILU) -- where the ion beam is formed by a special gas discharge source, a system of extracting, accelerating and focusing lenses. The beam is directed in the direction of a rotating inclined table, on the surface of which the plates or substrates are located. To compensate for the positive charge accumulated on the treated surfaces, a neutralizer -- a heated cathode emitting electrons -- is used.

During ion beam pickling the plates or substrates are located outside the plasma gas discharge gap. This permits regulation of the process parameters, the ion energy, the current density of the ion beam and angle of incidence of the ions on the surface of the plates and substrates independently of each other.

Ion pickling is universal. It is possible to clean the surface of any materials to remove impurities of any type. It is possible to use ion pickling to treat multilayer films with properties of the layers which are incompatible from the point of view of liquid chemical cleaning. Ion cleaning insures high quality without deep alterations of the treated surface layer, high precision of removal of the layers (± 0.03 microns) and excludes the interoperation expenditures of time, for it is possible to perform the subsequent operations (oxidation, film deposition) directly in the same vacuum chamber. Thus, ion pickling is used as finish cleaning of the plates and substrates before the process is performed in a vacuum and also for treatment of materials which are difficult to clean by other methods, for example, for substrates made of sapphire. The absence of lateral components of the pickling rate permits the application of ion pickling for precision dimensional local treatment.

Plasma Chemical Pickling. In contrast to plasma pickling, plasma chemical pickling is realized by ions of active gases instead of inert gases. During bombardment these active gases enter into chemical reactions with the treated material, forming volatiles. As a result of the electrical activation of the gases with the formation of reactive radicals, plasma chemical pickling can be carried out at significantly lower temperatures than ordinary gas pickling.

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

For pickling silicon, silicon dioxide or nitride, a high frequency plasma of carbon tetrafluoride CF_4 , sulfur hexafluoride SF_6 or freon-12 (difluorodichloromethane) CF_2Cl_2 are used.

On interaction of these gases with the plasma electrons, the fluorine radical F^- is formed which actively interacts with silicon:



The process parameters are the frequency, power, gas pressure in the reaction-discharge chamber.

Plasmochemical pickling, just as other methods of dry pickling insures good quality during local treatment.

Storage and Transportation Plates and Substrates. High-quality cleaning of semiconductor plates and substrates represents a complicated system for treatment with the application of various reagents and media. When developing the cleaning process, the real possibilities of bringing the surface to the maximum degree of cleanliness with minimum economic expenditures are taken into account, and a great deal of attention is given to the problems of the interoperation storage and transportation of the plates and substrates. The holders and packaging must be made of easily cleaned materials, and the working volume of the package must be small. For additional protection against atmospheric effects and convenience of transportation the packaging with the cleaned specimens is placed in a sealed stainless steel container. The time from completion of cleaning to the beginning of the next operation must be minimal, and it is regulated by the production forms and reports.

The protection of the surface by polymer films permits prevention of contamination of the cleaned plates and substrates during transportation and storage until the next operation. The KhSL polymer lacquers (chemically stable lacquer), PMMK polymer lacquer (polymethylmethacrylate), PAN (polynitriloacrylate), PVS (polyvinyl alcohol) and others are applied to the surface of the substrates. During polymerization the lacquers capture contamination on the surface in a solid film which, in addition, protects the surface from new contamination from the outside. The polymer films are easily and completely removed from the surface of the substrates by simple mechanical "stripping," which can be realized in the devices directly before performing the next operations, for example, before applying the films.

4-6. Quality Control of Surface Cleanness of Plates and Substrates

During cleaning it is necessary continuously to check the cleanness of the production environments, atmosphere, water, solvents, reagents and perform selective quality control of the surface. Direct and indirect methods of quality control of the cleanness of plates and substrates are distinguished. In the direct methods the contaminants are determined directly on the cleaned surfaces. In the indirect methods, for example, the properties of the Si-SiO₂ interface, the properties of the SiO₂ film, the parameters of the p-n junctions made on the given plates are investigated. The cleanness quality criterion in this case can be the breakdown voltage of the silicon dioxide film, the capacitance of the MOS-structure, back currents and breakdown voltages of the p-n junctions, and so on.

FOR OFFICIAL USE ONLY

The composition of the contaminants and their quantitative relations can be determined by spectral analysis, but this method is complicated, and therefore it is used in rare cases.

Under production conditions for rapid quality control of cleanness, simple methods are used which are easily carried out and not require complex equipment.

The method of investigating the specimens in the dark field of a microscope permits detection of many greasy, oxide and mechanical contaminants. The method is based on different light reflection by the clean surface of the plate or substrate and contaminated sections. In the case of oblique lighting of a clean surface the light rays reflected from it do not follow in the field of view of the observer, but those reflected from the contaminants do. The contaminants are visible as light spots on the dark field. The size of the contaminants and their quantity per unit area of the substrate can be determined, but the composition of the contaminants cannot be recognized. Microscopes are used for control in the dark field. The samples are placed in the microscope stage with the examined surface up, for example, using the MBI-11 microscope.

The water film rupture control method is based on the capacity of water to wet the surfaces of clean plates or substrates or plates or substrates contaminated with high-power substances and not to wet a surface contaminated with greases. The checked sample is lowered into a clean chemical beaker filled to the top with water, and it is taken out of the water in the vertical direction, converting to the horizontal position. The water dries uniformly off the clean surface without breaking into individual drops and patches. On the dirty surface the water film instantaneously breaks and forms individual drops, for during drying the water withdraws from the unwetted to the wetted sections, that is, from the dirty sections to the clean ones. The measure of the amount of contamination can be the time from removal of plate or substrate from the beaker filled with water to breaking of the water film.

The spraying technique consists in applying water to the dry investigated surface by spraying. The water immediately covers the clean substrate with a uniform continuous layer, and individual drops of water appear on the dirty surface. The spraying technique is 30 to 50 times more sensitive than the water film rupture method; it can be used to detect contamination to $1.6 \cdot 10^{-8}$ g/cm².

The "condensation figure" method is based on visual observation of the surfaces of the plates and substrates after moist air is blown over them. The moisture condensed on the surface forms so-called condensation figures. An insufficiently clean silicon surface has a frosted appearance, for it is coated with fine drops of moisture. This is a "gray condensation figure." In the case of good quality of cleaning of the substrate, a "black condensation figure" is observed, for the surface is covered by a solid thin layer of moisture.

The drop method is based on measuring the wetting angle of the surface by water or clock oil. By using the optical system of the device, a drop of water or clock oil applied to the investigated surface is projected on a screen. The wetting angle is measured by a protractor. For clean plate and substrate surfaces the wetting angle by oil must not exceed $1-3^\circ$.

FOR OFFICIAL USE ONLY

The scratch method is based on measuring the force required to move a metal needle over a surface of the substrate. The contaminants play the role of a lubricant. Clean surfaces are distinguished by exceptionally high friction coefficient.

The methods of direct estimation of the quality of cleanness of substrates are contaminated methods; therefore it is desirable to perform this check in enclosures with laminar flow of dust-free air. Only the control samples are checked.

Test Questions and Assignments

1. What is the role and what are the goals of chemical treatment and cleaning of the surfaces of plates and substrates when making IC structures?
2. What is the relativeness of the "clean surface" concept?
3. What surface of plates and substrates is considered to be technologically clean?
4. Enumerate the principal sources of contamination of plates and substrates.
5. How are the contaminants classified with respect to chemical properties?
6. What contaminants are most difficult to remove from the surfaces of plates and substrates?
7. What are the physical cleaning methods?
8. What are the chemical cleaning methods?
9. What is the pickling of a substrate?
10. What is the degreasing of a substrate?
11. What is the order of these operations in the technological cleaning process?
12. What is the essence of degreasing substrates in organic solvents?
13. What is the practical implementation of the method of degreasing in organic solvents?
14. What are the advantages and disadvantages of degreasing in organic solvents?
15. In what solutions is it possible to carry out chemical degreasing?
16. Compare which contaminants are to be removed by soap solutions and ammonia peroxide solution.
17. What is the role of the components of an ammonia peroxide solution?
18. Name the basic steps of chemical pickling.

FOR OFFICIAL USE ONLY

19. What are polishing and selective pickling agents?
20. Which factors does the pickling rate in a polishing pickling agent depend on? In a selective pickling agent?
21. How are the pickling rates of the crystallographic planes (111) and (100) distinguished in an anisotropic pickling agent?
22. What is the purpose of nitric acid, hydrofluoric acid and acetic acid during polishing pickling of silicon?
23. For what purposes is selective pickling used?
24. After which operations is the washing of the plates and substrates required?
25. What is forced washing?
26. What methods of intensification of cleaning are physical and which are chemical?
27. What is combined intensification of the processes of wet cleaning?
28. Which of the methods of intensification is the most effective when cleaning plates and substrates? What is its essence? Draw a diagram of an ultrasonic cleaning system by analogy with Figure 3-4.
29. Why does liquid cleaning involve not one, but several operations?
30. What is the role of acetone during the removal of oxide films from silicon plates?
31. Why is pickling not used during liquid cleaning of glass, sapphire and ceramics?
32. What are the characteristic aspects of cleaning pyrocerams?
33. What are the special features and advantages of freon cleaning?
34. Enumerate the dry cleaning methods.
35. What is the gas pickling mechanism?
36. Which gases are used for pickling silicon and sapphire?
37. What are the basic disadvantages of gas pickling?
38. What is the ion pickling mechanism?
39. What distinguishes plasma ion pickling from ion-beam pickling?
40. What are the advantages of ion-beam pickling over ion plasma pickling?

FOR OFFICIAL USE ONLY

41. What does universality of ion pickling consist of?
42. What distinguishes plasmochemical pickling from ion plasma pickling?
43. What is the role of protective polymer films used after cleaning the surface of plates and substrates?
44. What are the direct and indirect methods of cleaning quality control?
45. What is the procedure for determining the quality of cleaning by the simplest methods: investigation in the dark field with a microscope, water film rupture control, and so on?

FOR OFFICIAL USE ONLY

CHAPTER 5. CONTACT PHOTOLITHOGRAPHY

5-1. Application and Essence of the Photolithography Process

General Information About the Methods of Localization and Lithography.

The localization of processing in the production of the structures of modern microcircuits is realized by free and contact masks. The smaller dimensions of the IC elements, greater precision and better reproducibility are provided by contact masks.

In turn, in order to obtain contact and precision free masks, various lithographic processes are used. The term "lithography" is derived from the Greek words lithos meaning stone and grapho meaning write or draw. Lithography, which was invented in Germany in the second quarter of the 19th century by A. Senefeld, has become a widespread artistic medium. The production technology of microcircuits borrowed lithography from the polygraphic industry.

A layer of resistant material, that is, material which is resistant to successive technological aggressive effects, capable of irreversibly changing its properties under irradiation of a defined wave length and, above all, resistant to special compounds -- developers -- is formed on a surface in the lithographic processes. This resistant layer is irradiated locally using a specially made stencil. After processing in the developer, a resistive compact mass is obtained as a result of removal of the leftover local sections in the resistant layer. The contact mask is then used either to obtain a free mask or directly to obtain the microcircuit elements.

Depending on the wave length of the radiation used, optical, x-ray and electronic lithography are distinguished (see Figure 5-1). In accordance with the irradiation (exposure) techniques, optical lithography (photolithography) can be contact or contactless (photolithography on a microgap and projection photolithography). Electron lithography can be carried out by successive transfer of the stencil configuration to the resistive layer by a focused single electron beam or by simultaneous projection of the entire configuration of the stencil.

Application of Photolithography. The introduction of contact photolithography into semiconductor production in 1957 determined the further development of the element base of electronics and made it possible to convert from discrete elements to

FOR OFFICIAL USE ONLY

integral elements. Photolithography is an unacceptable process of planar technology. Windows are opened up in the masking films with its help, through which local diffusion takes place, and then the metalization configuration in the aluminum film is obtained.

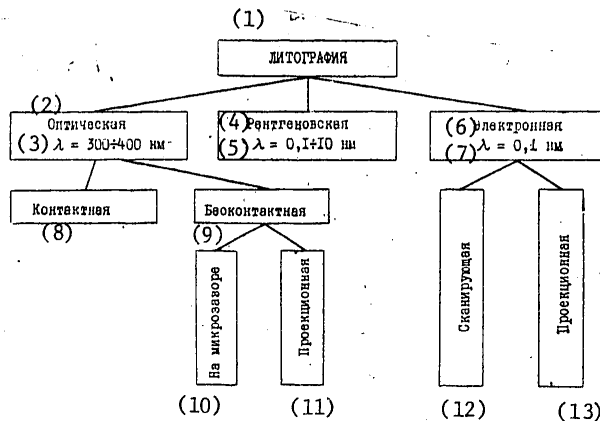


Figure 5-1. Classification of lithography methods

Key:

- | | |
|------------------------------|-------------------|
| 1. Lithography | 8. Contact |
| 2. Optical | 9. Contactless |
| 3. $\lambda=300$ to 400 nm | 10. On a microgap |
| 4. x-ray | 11. Projection |
| 5. $\lambda=0.1$ to 10 nm | 12. Scanning |
| 6. Electron | 13. Projection |
| 7. $\lambda=0.1$ nm | |

At the present time the areas of application of photolithography have expanded significantly. In modern microcircuit production photolithography is the most universal technological process. It permits reproducible and high-precision execution of complex figures with element sizes to 1 micron or less using various materials. Photolithography is used for the manufacture of semiconductor and film structures, to obtain all manner of grooves and depressions in semiconductor and other materials. Photolithography is used to make photographic stencils -- the tools for the photolithography process itself -- to obtain through holes in foil and making precision free masks or when making the lead frames or strips used for automated assembly and encapsulation of IC. The plates or substrates with finished structures are also divided by photolithography into individual crystals or plates, precision parts are made, and precision scales are created, and so on.

Essence of the Photolithography Process. Photolithography is a complex technological process based on using photochemical phenomena which occur in photoresistive layers during local actinic irradiation of them through a mask. Irradiation is actinic if the material of the photoresistive layer is sensitive to it and it causes irreversible changes in the properties of this material. The mask for making microcircuit structures is a plane-parallel plate (or flexible film) made

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

of material that is transparent to actinic radiation with a film figure opaque for actinic radiation applied to its working surface corresponding to the topology of one of the layers of the microcircuit structure and multiply repeated with a strictly defined step within the bounds of the working area of the plate (or film).

Optical borosilicate glass or polymer films which are easily worked and do not change properties under the effect of irradiation are used as the base for the masks. A silver halide photographic emulsion (emulsion mask), metals (metallized mask) and oxides or other materials transparent to visible light (transparent, colored mask) are used to form the figure.

Photoresists are used to form the photoresistive layer. The photoresists are complex polymer compounds which have photosensitive and film-forming components, solvents and special additives. The latter are introduced to improve the film formation conditions (diluent), to change the photosensitivity (sensitizer), to improve the adhesion of the photoresistive layer to the substrates, to improve the resistance to the effects of acids, alkalis, high temperatures, and so on. Solvents determine the stability of the properties of the finished photoresists, and they influence the application process and subsequent drying of the photoresistive layer.

At the present time when developing photoresists, primarily synthetic polymers are used. The presence in the structure of polymer molecules of the functional groups C=O, C=C, -N₂, -N₃, -NO, -NO₂ and so on causes absorption in the ultraviolet (UV) region of the spectrum.

The absorption of the irradiation energy takes place by quanta. Each absorbed quantum, depending on its energy, converts only one photosensitive molecule in the photoresistive layer to the electron-excited state or it breaks it into free radicals (photolysis). This is the so-called light stage of the photochemical process. For directional alteration of the physical-chemical properties of the irradiated sections of the photoresistive layer, the primary role for successive secondary chemical reactions is played by the formation of radicals in the light stage. The reactions occurring in the dark stage lead to structuring (crosslinking) or, on the contrary, to depolymerization of the molecular chains of the polymer. Accordingly, the resistance of the irradiated sections of the photoresistive layer to the effects of the developed increases or, on the contrary, decreases.

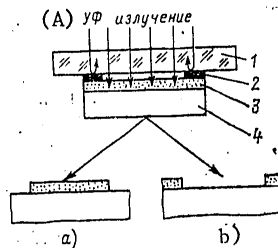


Figure 5-2. Transfer of images using negative (a) and positive (b) photoresists.

1 -- mask base; 2 -- opaque sections of the mask figure; 3 -- photoresistive layer; 4 -- substrate.

Key:

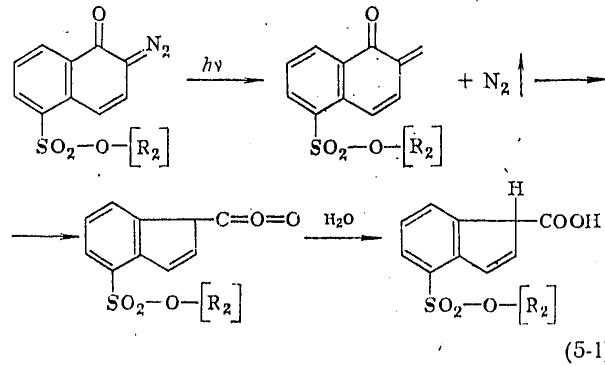
A. Ultraviolet radiation

FOR OFFICIAL USE ONLY

Positive and Negative Photoresists. Depending on the photochemical processes taking place in the photoresistive layer under the effect of actinic radiation and the direction of secondary dark reactions in accordance with the nature of the image transferred from the mask to the photoresistive layer, photoresists are divided into negative and positive. Photoresists, the local sections of which are removed in developers after the effect of radiation as a result of photodestruction, and the unirradiated sections of which remain on the substrate and form a photoresistive contact mask, are called positive photoresists.

Photoresists, local sections of which become resistant to the developer under the effect of radiation as a result of photostructuring and, in contrast to the unirradiated sections, remain on the substrate, forming a photoresistive mask, are called negative photoresists (Figure 5-2).

The positive photoresists are made on the basis of photosensitive complex esters of naphthoquinonediazides of sulfo acid and phenolformaldehyde resins which are film-forming polymers. A molecule of the positive polymer abbreviated NKhd, has the structure R_1-O-R_2 , where R_1 and R_2 are the light-sensitive and polymer parts of the ester, and O is the oxygen atom joining them. On absorption of the irradiation quanta, photolysis occurs. The NKhd molecule decomposes to nitrogen and an unstable radical which is converted to indenecarbene. Indenecarbene incorporates the moisture always existing in the photoresist, forming indenecarboxylic acid:



The local sections of depolymerized photoresists formed during irradiation, in contrast to the unirradiated sections, become hydrophilic and are easily removed in the developers with alkaline properties.

Negative photoresists are made on the basis of polyvinylcinnamate or on the basis of elastomers. Polyvinylcinnamate (PVC) is the complex ester of cinnamoylic acid and polyvinyl alcohol, and it has the general formula R_1-O-R_2 , where R_1 is the photosensitive cinnamoyl group, R_2 is the film-forming part of polyvinyl alcohol, and O is the oxygen atom joining them. The structure of polyvinylcinnamate is illustrated in Figure 5-3. On absorption of radiation quanta, the weakest chemical bonds $CH=CH$ in the light-sensitive parts of the molecule break. As a result of the freed bonds, photostructuring takes place, that is, crosslinking to a chemically stable trimer lattice.

FOR OFFICIAL USE ONLY

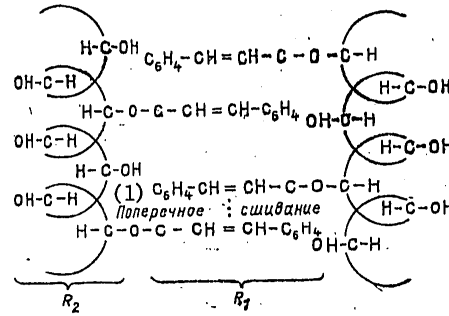


Figure 5-3. Structure of polyvinylcinnamate

Key:

1. Crosslinking

The negative photoresists based on elastomers are most frequently a mechanical mixture of cycloelastomer and bis-azide. The cycloelastomer is a polymer base and has weak photosensitivity. Under the effect of irradiation photodecomposition of the azides into nitrogen and nitrene takes place. Nitrene, which is in the excited state enters into chemical interaction with the elastomer; as a result, structuring of the polymer -- crosslinking of the linear polymers of the elastomer by the free radicals formed into a stable trimer lattice -- occurs.

Thus, as a result of the photochemical processes taking place in the photoresistive layers under the effect of actinic irradiation and subsequent treatment in developers photoresistive masks are formed on the surfaces of the substrates. As is obvious from Figure 5-2, a positive photoresist precisely transfers the figure of the mask: the openings in the photoresistive mask correspond to the sections of the mask transparent for radiation. A negative photoresist transfers the inverse, negative image of the mask: the opaque sections of the mask correspond to the openings in the photoresistive mask.

Basic Characteristics of Photoresists. Photoresists are evaluated by a number of characteristic parameters. Let us briefly consider the most important of them.

A type of photoresist (negative or positive) is selected depending on the specific application. A positive photoresist transfers unplanned through holes in the opaque part of the mask called punctures, in the form of punctures in the substrate, for example, in the silicon dioxide film. The negative photoresist, on the contrary, transfers in the form of islets. Whenever the punctures are dangerous, the negative photoresist is used, and whenever the islets are dangerous, a positive photoresist. The positive photoresist makes it possible to obtain smaller sizes of elements and clearer boundaries of the figure. The negative photoresists are more stable in the processes connected with electrolytic deposition of metals and deep pickling. In the technological production processes frequently combinations of positive and negative photoresists are used. This facilitates performance of some of the technological operations and permits the quality of the photolithography to be improved.

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

Photosensitivity characterizes the capacity of the photoresist irreversibly to change its resistance to the developer under the effect of irradiation.

The photosensitivity is estimated by the value of S which is inversely proportional to the exposure time H which is sufficient for crosslinking or depolymerization to the entire thickness of the photoresist and corresponds to the maximum rate of development of the image:

$$S = \frac{1}{H} = \frac{1}{Et}, \tag{5-2}$$

where E is the illumination; t is exposure time. The photosensitivity of the photoresist determines the efficiency of the photolithography process.

Integral and spectral photosensitivity are distinguished.

The photosensitivity to the effect of irradiation of all wave lengths absorbed by the given photoresist is called integral.

Photosensitivity to the effect of irradiation in a defined wave length is called spectral.

What irradiation are photoresists sensitive to? The answer to this question is given by the spectral characteristics. A photochemical reaction takes place if the radiation quantum energy is sufficient to break the bond of the photosensitive part of the photoresist molecule. In the spectral characteristic this corresponds to the appearance of maximum absorption.

From Figure 5-4 it is obvious that photoresists are sensitive to ultraviolet irradiation. The absorption spectrum of the positive photoresist based on NKhd is shifted in the direction of longer wave lengths by comparison with the PVC spectrum. The long-wave boundary of the NKhd spectrum is in the range of 460-480 nm, and the spectrum has three characteristic peaks on wave lengths of 280, 350, and 420 nm. PVC has maximum absorption in the wave length range of 280-300 nm.

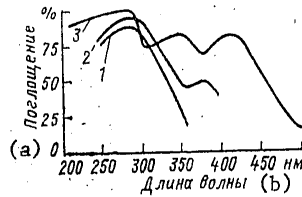


Figure 5-4. Photoresist absorption spectra
 1 -- TsVTs (Unsensitized); 2 -- PVC (sensitized); 3 -- NKhd.

Key:

- a. Absorption
- b. Wave length

The spectral characteristic of a positive photoresist agrees well with the characteristic of sources of ultraviolet radiation and transmission of the transparent sections of the mask, and the spectral characteristic of the negative photoresist does not agree; therefore their composition includes sensitizers that shift the spectrum to the longer wave part (350-370 nm).

FOR OFFICIAL USE ONLY

The photosensitivity depends primarily on the chemical composition of the components entering into the photoresist and also the conditions of formation of the photoresistive layer.

The resolution of the photoresist is the maximum number of lines of identical width separated by gaps equal to the width of the line which can be obtained in a photoresistive layer at a length of 1 mm after development of the figure:

$$R = \frac{1000}{2l} \quad (5-3)$$

where l is the width of the separately transferred line.

In practice the resolution frequently is characterized by the minimum line width in microns which can be transferred to the photoresistive layer.

The resolution depends on the molecular nature of the photoresist, the state of the coated surface, the adhesion of the photoresistive layer to the substrate, and the thickness of the photoresistive layer and its uniformity. Thin layers correspond to high resolutions. Positive photoresists, as a rule, have higher resolution.

The resistance of the photoresists to the effect of aggressive factors is determined by the stability of the photoresist itself, uniformity and defect-freeness of the applied photoresistive layer, the quality of the formed photoresistive contact mask and also, in the case of the effect of developers or pickling agents, the wettability of the surface of the substrate by them and the capacity to penetrate under the layer of photoresist on the edge of the figure.

The stability of the photoresist in the developers is estimated by the permissible development time in seconds.

In a formed photoresistive mask the stability of the photoresist is estimated by the number of defects transferred per unit area of substrate as a result of the effect of the aggressive factor. The admissible defect density for each type of photoresist is stipulated by the production forms and records.

The negative photoresists and, especially, photoresists based on elastomers are resistant to the effects of acids and bases. The positive photoresists have high resistance to acids, but they are unstable with respect to alkalis. The acid resistance of the positive photoresists based on NKHD can be increased by the introduction of functional halide groups through the resin molecule, and the alkali resistance can be increased by introducing epoxy groups.

The kinematic viscosity of the solutions of photoresists prepared for photolithography is estimated by the escape time of them between the upper and lower marks of the viscosimeter capillary. The viscosity has a significant effect on the conditions of application of the liquid photoresist to the substrate and the formation of the layer of required thickness. The viscosity of the photoresist depends strongly on the temperature and changes rapidly in time. It is possible to vary the viscosity by adding a diluent to the photoresist or varying the ratio of the dry photoresist components.

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

The surface tension of the photoresist determines the wetting of the substrate surface by it and has a significant influence on the adhesive binding of the photoresistive layer to the substrate. High adhesion is insured under the condition where the surface tension of the photoresist does not exceed some critical value. The critical surface tension of the photoresist is the maximum value of the surface tension for which the wetting tension does not depend on the surface state and the photoresist spontaneously spreads over the surface of the given substrate. For a number of Soviet positive photoresists the critical surface tension is $28 \cdot 10^{-7}$ joules/cm².

In practice the surface tension of the photoresists is reduced as a result of using a combination of different solvents.

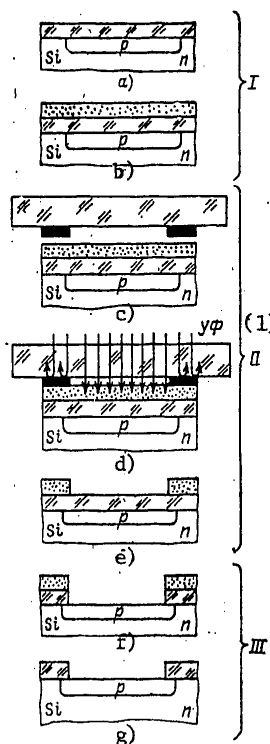


Figure 5-5. Diagram of the photolithography process with respect to a silicon dioxide film.

I -- phase of formation of the photoresistive layer; II -- phase of formation of the photoresistive mask; III -- phase of transfer of the image to the substrate; a -- preparation of the substrate; b -- application of photoresist and drying; c -- matching; d -- exposure; e -- development and heat treatment; f -- pickling; g -- removal of the photoresistive mask

Key:

- 1. ultraviolet

FOR OFFICIAL USE ONLY

The service life of a photoresist under defined storage and operating conditions is the most important quality parameter. The suitability of the photoresist is determined by the maintenance of stability of the properties within the limits given by the specifications. The stability of the photoresist properties is estimated by periodic measurement of the optical density which is proportional to the concentration of the photosensitive molecules. The stability depends on the purity of the initial products and the quality of synthesis of the photoresist, the storage and application conditions. The photoresists are stored in a dark container with a ground glass stopper at a temperature of about +15°C.

In Soviet industry broad use is made of the positive photoresist types FP-330, FP-383, FP-333, FP-334, FP-307, FP-309, FP-RN-7 having resolution from 400 to 500 lines/mm with a layer thickness of 1 micron. The most acid-resistant of these photoresists is FP-333, for which the transferred defect density does not exceed 0.2 mm^{-2} . The FN-3T, FN-5T, FN-11, FN-11K and FN-106 type negative photoresists have a resolution of 100 to 400 lines/mm for a photoresist layer thickness of 1 micron.

Basic Steps in the Photolithography Process. The entire photolithography process can be broken down into three basic steps: the formation of a photoresistive layer on the surface of the substrate, the formation of a photoresistive contact mask (transfer of the image from the mask to the photoresistive layer), and obtaining the configuration of elements using the resistive mask.

The sequence of the basic technological operations of the photolithography process with respect to the masking film of silicon dioxide in the planar process is presented in Figure 5-5.

5-2. Formation of the Photoresistive Layer

Preparation of Substrates. In photolithography the substrate is a material on which a photoresistive layer is formed. In IC production photolithography is carried out on a number of substrate materials. In the production of thin-film hybrid IC, for example, photolithography is carried out on metals, dielectrics and semiconductors. In the production of semiconductor IC photolithography is performed with respect to monocrystalline and polycrystalline silicon, silicon dioxide films, admixture-silicate glass, aluminum, and so on. The preparation of the substrates for the application of the photoresist is individual in each specific case, and it is determined by the material of the substrate, the process of obtaining it, the condition of the surface and the future purpose of the photoresistive mask.

For example, if the photoresistive mask is used to shield against pickling, the quality of transfer of the figure to the substrate depends on the adhesion of the mask to the substrate surface and the capacity of the pickling agent to penetrate the photoresist layer. The wetting angle of the surface of the substrate by water is taken as the wettability criterion and, consequently, the capacity to penetrate the substrate-photoresist interface. With a decrease in the hydrophilicity of the surface, favorable conditions are created for decreasing the erosion of the figure. The optimally prepared surface for photolithography is the surface which is well wet by the photoresist and poorly wet by water, that is, for which the following condition is satisfied:

FOR OFFICIAL USE ONLY

$$\begin{matrix} \theta_w \rightarrow 0^\circ, & \theta_R \rightarrow 180^\circ. \\ (1) & (2) \end{matrix} \quad (5-4)$$

Key: 1. ph; 2. w

Thus, in the planar technology in the process of pickling the silicon dioxide film for values of θ_w less than 35° the penetration of the pickling agent under the layer of photoresistive mask begins to increase, and the figure is distorted. For $\theta_w=49-63^\circ$ the distortion of the openings in the silicon dioxide film is minimal. For identical adhesion of the photoresistive mask to the substrate the distortion of the figure is lower for high wetting angles by water. Consequently, if the surface is hydrophilic, hydrophobizing treatment is required before applying the layer of photoresist.

It is expedient to perform photolithography with respect to the silicon dioxide layer directly after thermal oxidation of the silicon while the surface of the dioxide is hydrophobic. After prolonged storage of the substrate, the silicon dioxide surface will become hydrophilic; therefore additional heat treatment in an oxygen atmosphere is needed at a temperature of $900-1000^\circ\text{C}$ for 5-10 minutes or infrared drying (IKS). The temperature during infrared drying is about 400°C , which is more favorable for maintaining the parameters obtained in the preceding operations.

Hydrophobizing treatment of admixture-silicate films is also carried out by annealing in oxygen at a temperature of about 500°C , and aluminum films are annealed in an inert environment at a temperature of about 300°C .

In the manufacturing technology of thin-film structures if the substrates have been stored for a long time after obtaining the films, the treatment of the substrates usually reduces to ordinary degreasing ("brightening") in organic solvents and infrared drying.

Application of the Photoresist Layer. The operation of applying a photoresist must provide for the formation of defect-free, uniform photoresistive layers that are reproducible with respect to thickness with maximum adhesion to the substrates while maintaining initial characteristics of the applied photoresist. The photoresist can be applied by centrifuging, spraying, dipping, flooding and rolling.

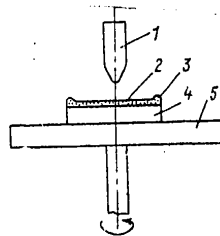


Figure 5-6. Application of the photoresist layer by centrifuging.
 1 -- batcher for feeding the photoresist; 2 -- photoresist layer;
 3 -- thickening of the layer around the periphery of the substrate;
 4 -- substrate; 5 -- centrifuged disc

FOR OFFICIAL USE ONLY

The centrifuging method (Figure 5-6) is used most frequently. The substrates are located on the surface of the centrifuged disc and are held during rotation by a vacuum clamp or using peripheral stops. Several drops of photoresist liquid are applied to the central part of the substrate surface before or after the centrifuge is switched on. On applying the photoresist to the stationary substrate the centrifuge is switched on instantaneously, for the viscosity of the photoresist very quickly changes as a result of evaporation of the solvents. During rotation of the centrifuge the liquid photoresist spreads over the surface of the substrate under the effect of centrifugal forces proportional to the centrifuging rpm. As a result of surface tension of the photoresist, forces of resistance to spreading arise which are proportional to the viscosity of the photoresist; therefore it is not completely thrown off the surface of the substrate. The dependence of the liquid layer of photoresist remaining on the surface on the centrifuge rpm and kinematic viscosity of the photoresist is defined by the expression

$$h_{\text{ж}} = A \sqrt{\nu/\omega}, \tag{5-5}$$

Key: 1. liquid

where $h_{\text{ж}}$ is a layer of liquid photoresist, microns; A is the empirical coefficient; ν is the kinematic viscosity, mm^2/sec ; ω is the centrifuge speed, rps.

During centrifuging, the thickness and quality of the layer depends on the viscosity of the photoresist, the rpm and nature of the acceleration of the centrifuge rotor, temperature and humidity of the environment, properties of the substrate surface. The centrifuging time has little influence on the layer thickness; a time of 20-30 seconds is sufficient for forming the layer.

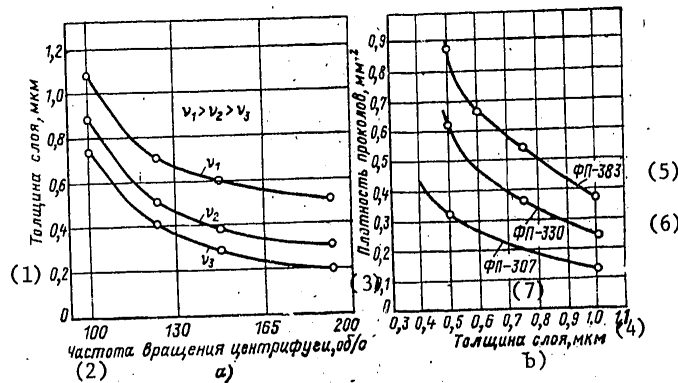


Figure 5-7. Thickness of the photoresistive layer as a function of the centrifuge rpm (a) and the density of the punctures as a function of the thickness of the photoresistive layer (b)

- Key:
- | | |
|---------------------------------------|-----------|
| 1. Layer thickness, microns | 5. FP-383 |
| 2. Centrifuge rps | 6. FP-330 |
| 3. Puncture density, mm^{-2} | 7. FP-307 |
| 4. Layer thickness, microns | |

FOR OFFICIAL USE ONLY

In practice, the centrifuging conditions are selected beginning with the requirement of obtaining a thin photoresistive layer with minimum number of defects. The experimental relations for the layer thickness as a function of centrifuge rps and puncture density as a function of layer thickness are presented in Figure 5-7. Usually high rps are used which insure minimum dispersion of the thicknesses with deviations of the centrifuge rps. The thicker layers correspond to lower density of punctures, but the image is more difficult to transfer exactly in the thick layers.

When centrifuging thickening -- a roll -- is always formed on the periphery of the substrate. The height and width of this roll depend on the photoresist viscosity, the centrifuge rps and the shape of the substrate. It is difficult to obtain uniform layers more than 2 microns thick by centrifuging. The dispersion with respect to thickness of the layers is about $\pm 10\%$. There are always mechanical stresses in the layer. Intake of foreign inclusions into the center of the rotating layer from the environment is possible.

A photoresist is applied by spraying from an air gun using compressed air or superheated freon. The molecular mass of the freon is higher than that of air, which makes it possible to lower the pressure required to disperse the jet of photoresist solution on exit from the nozzle. The substrates are placed on a table which undergoes reciprocal motion. The layer parameters depend on the air or freon pressure and temperature, the distance from the nozzle of the gun to the substrate and on the photoresist parameters. For improvement of adhesion of the photoresist it is possible to heat the substrates. It is possible to spray the layers on intricately shaped surfaces and obtain thick layers to 10 microns. The dispersion with respect to thickness of the layers does not exceed $\pm(15-20)\%$. The edge roll is absent. The process is distinguished by high efficiency and the possibility of automation. The basic problem when applying layers by spraying is entrainment of dust and other contaminants by the jet of dispersed photoresist.

The dipping method consists in simple submersion of the substrate in the photoresist and extraction from it at regulated speed of 6-150 mm/min. During slow extraction, thinner and more uniform layers are obtained. The dipping method is used for two-sided application of the layers, to obtain thick layers, and when the requirements on thickness uniformity are not too high. In order to eliminate wedge shaping of the layer, the substrate is rotated by 180° , and it is dipped again.

The flooding method consists in applying a batch of photoresist to the surface of the substrate and rotation of the substrate to spread it. The dispersion with respect to layer thickness reaches 30-40%.

Rolling requires the development of special photoresist -- dry film photoresist. The method permits application of uniform layers with a uniformity with respect to thickness of $\pm 5\%$.

The methods of dipping, flooding and rolling permit photoresistive layers to be obtained with minimum number of defects on substrates of any shape, and they are distinguished by high efficiency and the possibility of automation.

FOR OFFICIAL USE ONLY

The thickness of the photoresist film is selected optimal for compromise satisfaction of the requirements of maximum resolution, minimum puncture density, maximum adhesion of the photoresistive layer to the substrate and resistance to the aggressive factors. In the photolithographic processes in the Soviet Union and abroad, photoresistive layers are used which are from 0.5 to 1.5 micron thick, and the resolution is 800-200 lines/mm in this case.

After application, the photoresist layer is checked for the absence of mechanical inclusions, ruptures and runs.

Drying. During drying the solvent is removed, and a complex relaxation process of dense packing of the molecules takes place in the photoresist film, decreasing the internal stresses and increasing the adhesion of the formed dense photoresistive layer to the substrate. During drying it is necessary to remove the solvent completely, for it shields the photosensitive parts of the molecule during exposure of the photoresistive layer. It is possible to divide the removal of the solvent into two steps: diffusion from inside the layer to the layer-atmosphere interface and evaporation from the surface. If evaporation predominates over diffusion, the surface layer of the photoresist becomes dense sooner than the inside layer and prevents removal of the solvent from within. Internal stresses arise which weaken the layer and lead to rupture of it. For more uniform drying the photoresists are prepared using several solvents with different evaporation rates.

The basic parameters of the drying process are temperature and time. At low temperatures the adhesion of the photolayer to the substrate is poor, adhesion between the molecules of the polymer itself (cohesion) predominates. This explains the flaking off of the photolayer during exposure. For temperatures above critical temperatures, the same irreversible phenomena occur in the positive photoresist layer as during exposure.

The following methods of drying the photoresist layer are used: convective, infrared, and microwave. In the case of convective drying performed in thermostats and special furnaces the temperature gradients and rapid heating are dangerous, for a layer of increased density is formed on the surface which prevents uniform drying of the layer. Infrared drying realized, for example, in the unit for applying drying the ULS-1 from the lithographic line "Korund-C," proceeds from within, and the process of solvent removal is not complicated; therefore by comparison with convective drying, the infrared drying time is from 15-30 minutes to 5 minutes less. More uniform drying insures greater uniformity and fewer defects in the layer. Microwave drying is done in 200-400 watt furnaces operating at a frequency of 2.45 gigahertz. The drying time is reduced to seconds, which sharply increases the efficiency.

In practice frequently stepped drying is used with a gradual raising of the temperature. It is recommended that the drying be done in an inert environment, for oxidation of the photoresist molecules is possible in the air.

After drying the control is performed visually by the color of the photoresist layer which must not change or, which insures greater precision, by measuring the dielectric losses. Storage of the substrates after drying should not exceed 10 hours.

FOR OFFICIAL USE ONLY

5-3. Formation of a Photoresistive Mask

Matching. In accordance with the known (see §1-4) principles of localization and layering of the production processing in the manufacture of microcircuit structures photolithography is carried out several times. Beginning with the second photolithography it is necessary to match the mask figure with the figure on the substrate obtained as a result of the preceding photolithography. The matching is done in the same units as subsequent shielding, with ordinary nonactinic illumination in the presence of a clearance (10-15 microns) between the substrate and the mask. At the present time visual and automated photoelectric methods of matching are used.

The visual method of matching is realized using an optical microscope and special shift mechanism. The operator, simultaneously observing the substrate and mask figures through a microscope, sees to exact superposition of them. First the figures are matched by the approximate matching marks within the limits of quite large shifts of substrate. The role of these marks is played by absences of one or several frames on the mask which permit proper superposition of the corresponding rows and columns during matching. Then exact matching is carried out within the limits of essentially smaller substrate displacements using the matching figures (reference symbols) -- special topologic figures in the form of strokes, squares, rings, crosses with controlled clearance along the outline. The precision of visual matching depends on the microscope resolution, the smoothness and precision of the displacements and fixing of them, the type of matching attributes, and it is no less than ± 1 micron.

The automated photoelectric method of matching is more effective as compared to the visual methods determined by individual peculiarities of the operator's vision and manual manipulation when shifting the substrate. After preliminary approximate matching using an optical microscope, precise matching is accomplished using the photoelectric microscope which fixes the difference in illumination between the matching marks of the substrate and the mask and converts the information obtained to instructions to shift the table with the substrate. For automated photoelectric recording, matching symbols are used in the form of etched grooves on the substrates and opaque lines on the mask. The matching precision is no less than 0.05 to 0.1 micron.

One of the basic difficulties in insuring exact match is creation of mechanisms for smooth displacement of the substrate distances of less than 1 micron. The high matching precision can be achieved only with objective methods; therefore further improvement of them must take place in the direction of complete, all-around automation of the process.

Exposure. After matching, the substrate and the mask are put into tight contact with each other and the exposure operation takes place. The required contact force is created by a mechanical or vacuum clamp. The photoresists have a narrow spectral absorption region (350-450 nm) and relatively low photosensitivity. Therefore sources of ultraviolet radiation, mercury-quartz lamps DRSh-250, DRSh-500 are used which insure high illumination (to ten thousands of luxes). In order to match the absorption spectra of the photoresist and the source radiation, light filters are used. Parallelness of the radiation beam required to decrease the penumbras occurring in the presence of sections where the mask-substrate contact is not tight is provided by a system of condensers made of 1-5 lenses.

FOR OFFICIAL USE ONLY

The nonuniformity of illumination over the exposure field of 50-60 mm should not exceed 5-10%.

In the case of careful stabilization of the illumination and tight contact between the mask and the substrate in practice the exposure process is controlled by the irradiation time. The required exposure time is insured by an electromagnetic shutter, the opening and closing time of which is 0.05 to 0.1 second, which in an exposure time interval of (1 second to 2 minutes) insures good precision.

The optimal exposure time is selected experimentally using satellite substrates when developing a new technological process, when using a new set of masks or a new batch of photoresist, when replacing the radiation source and changing the thickness of the photolayer. It is possible correctly to select the exposure time only when considering the interrelation of the exposure and development conditions. This is explained by the fact that for an identical exposure H , with different ratios of the illumination E and the time t the photochemical processes take place differently, and the development speed is different.

For each satellite substrate the exposure takes place for a different time. Then, after development, by checking each satellite, the exposure time is determined which corresponds to the best quality of image transfer. Then the same thing is done, but the exposure time obtained is taken as the fixed parameter, and the development time as a variable parameter.

More precisely, the optimal exposure time is determined using a sensitometric test, as a result of which the characteristic curves are plotted. For negative photoresist the characteristic curve is the thickness of the layer crosslinked to a defined depth as a function of the exposure, and for positive photoresists based on naphthoquinonediazides, the speed of development and speed of reproduction of the figure as a function of the exposure.

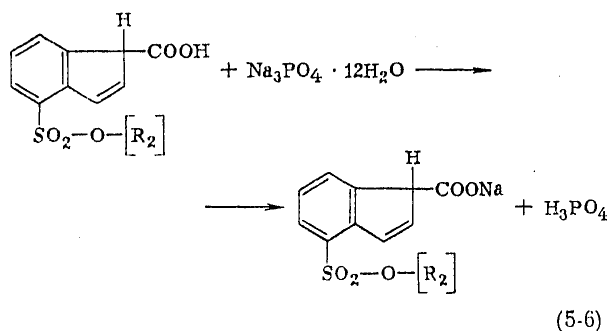
Development. Development is the process of removing the excess sections in the photoresistive layer in accordance with the local irradiation during exposure.

The development of negative photoresists is simple solution of the unirradiated sections in organic solvents: toluene, trichloroethylene, dioxane, and so on. The development must have good resolution and minimum effect on the irradiated sections of the photolayer. In the case of improperly selected exposure the irradiated sections swell sharply, which leads to distortion of the figure. Penetrating between the photoresist molecules, the solvent causes swelling of the layer. During formation of precision photoresistive masks, the swelling leads to shifts, interlocking and deformations of the elements of the figure. Underexposure of the negative photoresists leads to an increase in the swelling during development and, consequently, to a decrease in resolution of the photolithographic process. If exposure is carried out with optimal exposure time, overdevelopment of the negative photoresist is not dangerous; therefore the development processes are easily automated. In order to facilitate the removal of the unirradiated sections from the gaps between the irradiated sections, ultrasound, surface-active materials and shifting of the substrates or the developer are used. Development is most frequently carried out by spraying a solvent on the rotating substrates.

FOR OFFICIAL USE ONLY

The absence of alkali metal ions during development of negative photoresists is a valuable advantage, especially when making semiconductor structures.

The development of positive photoresists is accompanied by the removal of sections of the resistive layer irradiated during exposure. When developing a photoresist based on NKhD, the chemical reaction of the conversion of the indenecarboxylic acid obtained during exposure to a well-soluble salt which is then easily washed away takes place. After irradiation, the surface of the positive photoresist converts from hydrophobic to hydrophilic; therefore the irradiated sections of the photoresistive layer, in contrast to the unirradiated ones, are wet well by the developer. Aqueous alkali solutions (0.3-0.5% solution of caustic potassium, 1-2% solution of trisodiumphosphate) or organic alkalis -- ethanolamines -- are used as the developers. The development in a trisodiumphosphate solution is accompanied by the chemical reaction:



If the layer is not completely exposed, the solution turns a raspberry color, for part of the indenecarboxylic acid molecules are converted not to the sodium salt, but are connected to the undecomposed NKhD molecules, forming the pigment. If the NKhD molecules are decomposed completely to the entire depth of the photoresistive layer, the developer remains colorless.

A characteristic feature of the development of positive photoresists is practical absence of swelling of the unirradiated sections of the layer; therefore they have greater resolution and lower dependence of it on the thickness of the photoresistive layer by comparison with the negative photoresists. However, the positive photoresists are sensitive to overdevelopment. The selection of the optimal development time is made for the given values of the pH and temperature of the developer.

Insignificant changes in concentration of the developer have a strong influence on the accuracy of the image transfer. An effort is made to reduce the development time to a minimum, for otherwise destruction of the unilluminated sections of the photoresistive layer as a result of mechanical impurities, part of the decomposed NKhD molecules and also as a result of solution of the polymer component of the photoresist in the developer can take place. For standard positive developers the development time is 15-20 seconds.

FOR OFFICIAL USE ONLY

Development is realized by submersion in a solution, holding in the developer vapor or spraying on the rotating substrate. After development of the photoresists, the operation of careful washing of the substrates in running dionized water follows.

Heat Treatment (Secondary Drying). The heat treatment of the photoresistive layer is carried out for restoration of the swelling pattern and blending stability to the photoresistive mask with respect to subsequent aggressive effects. During heat treatment of the developed layers it is possible to not worry about the decomposition of the photoresist molecules; therefore in order to improve the protective properties of the photoresistive mask it is possible to use higher temperatures by comparison with the drying temperature after applying the layer of photoresist. In order that the edges of the photoresistive mask not be grown over, it is desirable to carry out the heat treatment, especially for thick layers, with a slow rise in temperature. Holding at maximum temperature must be short in order that the photoresistive mask not rupture.

5-4. Obtaining the Configuration of the Elements

Transfer of Images from the Photoresistive Mask to the IC Element Material. Depending on the type of element, its material, the process of obtaining it and other factors, several versions of the use of photoresistive mask are distinguished.

The direct photoresistive mask method (direct photolithography) provides for transfer of the photoresistive mask pattern by local removal of the excess sections of the substrate through the openings in the mask. A standard example of direct photolithography is photolithography with respect to the silicon dioxide film in planar technology (Figure 5-5). The pickling agent must interact only with the removed material. Better quality of transfer of the pattern is provided by buffered pickling agents, for it is easier to control the retarded pickling process and record the end of it. The pickling rate of silicon dioxide in the standard buffered pickling agent $\text{HF:NH}_4\text{F:H}_2\text{O}=1:3:7$ is 0.001 micron/sec. The ammonium fluoride dissociating in the solution binds the gaseous silicon tetrafluoride into a stable ion:



Pickling without the addition of a salt is accompanied by violent release of silicon tetrafluoride bubbles which have a hydromechanical effect on the photoresistive mask and promote its peeling.

The pickling is carried out in the devices of the photolithographic lines. Then the substrate is washed and dried in a jet of nitrogen heated to 40-50°C. The pickling is a responsible operation, for a reject after pickling is impossible to fix.

The basic deficiencies of the method are the following: the application of strong pickling agents which can act on other materials of the microcircuit structures,

FOR OFFICIAL USE ONLY

difficulty of selecting compatible photomask- film-substrate-pickling agent materials; difficult of obtaining the pattern in multilayer films. The method is used to obtain the layout in silicon dioxide films, in aluminum metallization films, in thin conducting films of chromium, copper, and so on.

The method of direct photolithography with a sublayer provides for obtaining an additional contact mask with the layout corresponding to the layout of the photoresistive mask. As the sublayer and, consequently, the material of the obtained contact mask, gold, molybdenum, silicon dioxide, and so on are used.

The method is used if the photoresistive mask does not withstand the pickling agent or when punctures carried over from the photoresistive mask to the substrate are especially dangerous.

The processes of photolithography with respect to titanium, silicon, and silicon nitride are examples of the first case (Figure 5-8). When transferring the pattern to titanium, a gold contact mask is used which, in contrast to the photoresistive mask is not destroyed in the pickling agent $H_2SO_4:HF:H_2O=65:1:30$.

When etching out the silicon mesa-structures on (111) plates in the pickling agent $HF:HNO_3:CH_3COOH=2:9:4$ (Figure 5-9) or during an anisotropic etching out of the V-grooves in the (100) plates, a double silicon dioxide and photoresist mask is used in the hot alkali solution (Figure 5-10).

Silicon nitride is pickled only in hydrofluoric acid which the photoresistive layers do not withstand; therefore a molybdenum sublayer is used.

For photolithography with respect to silicon dioxide in order to decrease the dimensions of the punctures, a molybdenum sublayer is used which has good adhesion to the dioxide. The punctures are transferred from the photomask first to the molybdenum and then from the molybdenum to the silicon dioxide. The size of the punctures is diminished, for the lateral etching of the oxide decreases with further masking.

Inverse photolithography (called "explosion" photolithography) provides for obtaining patterns in the films that are negative with respect to the photoresistive mask (Figure 5-11). The free sections of the substrate are cleaned off through the holes in the photoresistive mask in order to improve adhesion of the film, and above the photoresistive mask the film is applied to the entire surface. Then the photoresistive mask is removed. On being separated from the substrate, the photoresistive mask takes with it the part of the film applied to it. The film must remain only on the sections not protected by the photoresistive mask. A necessary condition of high-quality explosive photolithography is rupture of the film along the boundaries of the photoresistive mask pattern.

The method of inverse photolithography is used successfully for local electrolytic deposition of films. The inverse photoresistive mask is removed by a weak compound -- solvent -- which does not affect the materials entering into the structures of the microcircuits. With this method it is impossible to use high heating, for it leads to additional polarization or decomposition of the photoresist and twisting of the edge of the photomask pattern, which complicates obtaining a sharp edge of the layout of the film elements. Therefore secondary drying, as a rule, is not used.

FOR OFFICIAL USE ONLY

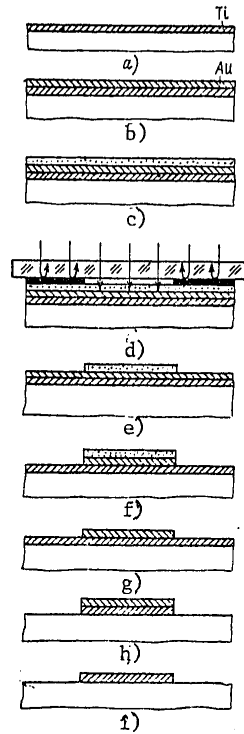


Figure 5-8. Diagram of the method of direct photolithography with a sublayer.
 a -- substrate with titanium film; b -- application of the sublayer (gold); c -- formation of the photoresistive layer; d -- exposure; e -- development; f -- pickling for transfer of the pattern to the sublayer; g -- removal of the photoresistive mask; h -- transfer of the pattern from the Au-mask to the titanium film; i -- removal of the contact mask

Inverse photolithography with a sublayer is free of the danger of heating. The contact mask with the pattern that is reversible with respect to the given pattern for the film element is obtained in a film of easily pickled material applied to the substrate, for example, copper, aluminum and bismuth oxide. The element film is applied over the contact mask. During pickling the film lying on the mask is removed together with the contact mask (see Figure 5-12).

In order to obtain the inverse contact mask, just as the direct one, it is possible to use organic materials which are removed by compounds less aggressive to the materials entering into the microcircuit structures.

FOR OFFICIAL USE ONLY

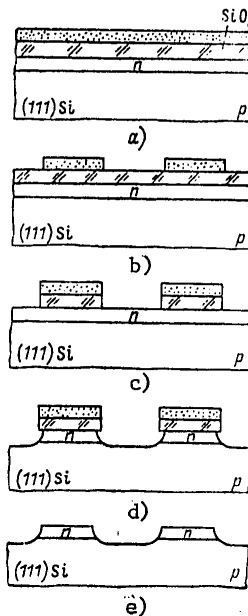


Figure 5-9. Diagram of the formation of mesa-structures.
 a -- application of a double layer of SiO_2 -photoresist; b -- formation of a photoresistive mask; c -- transfer of the pattern to the SiO_2 ; d -- plate pickling; e -- removal of the double contact mask

Removal of the Photoresistive Mask. In the finishing operations, as a rule, it is necessary to insure not only rupture and removal of the photoresistive mask playing its role, but also good cleaning of the surface to remove contamination introduced by the entire process of photolithography, for the surface state influences the quality of the subsequent operations. At the present time three methods of removing the photoresist are used: chemical destruction (decomposition) in sulfuric acid or solutions based on it, treatment in organic solvents, destruction by oxidation in oxygen or in oxygen-containing gas mixtures.

Chemical destruction is accompanied by the reaction of decomposition of the photoresist with the formation of less complex short molecules with small molecular mass which then are easily washed away with water. For acceleration of destruction, the concentrated sulfuric acid is heated to 160°C . The removal of the photoresist is accompanied by fast exhaustion of the acid; therefore it is consumed in large quantities.

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

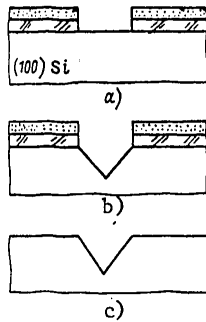


Figure 5-10. Diagram of obtaining V-grooves in (100) silicon. a -- formation of a double contact mask; b -- local pickling; c -- removal of the mask

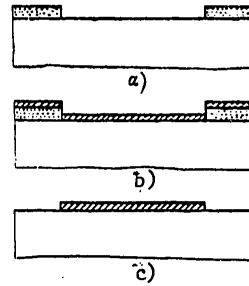


Figure 5-11. Diagram of the method of inverse photolithography (the "explosion" method). a -- substrate with photoresistive mask; b -- application of film; c -- removal of photoresistive mask

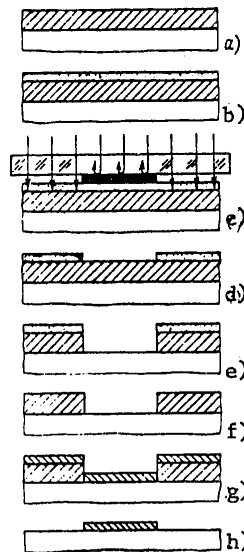


Figure 5-12. Diagram of the method of inverse photolithography with a sublayer. a -- substrate with a layer of contact mask material; b -- application and drying of the photoresist; c -- exposure; d -- development; e -- local pickling of the contact mask film; f -- removal of the photoresistive mask; g -- application of the element film; h -- removal of the contact mask (sublayer)

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

The better quality of removal even at lower temperatures (70-100°C) insures a mixture of concentrated sulfuric acid with 30% hydrogen peroxide in a volumetric ratio of 3:1. The method is not applicable for the removal of photoresist from the metallized substrate.

The treatment in organic solvents is successfully used to remove the photoresist from the metallized substrates. The substrates are held in solvents (acetone, methylethylketone, cellosolve, dimethylformamide). The quality of the process is improved on adding organic alkalis -- ethanolamines to the organic solvents. Then the swelling relief is removed by a mechanical tampon, and the treatment is repeated in fresh lots of the solvent.

The deficiencies of the liquid methods of removing photoresist are the multistage nature, labor consumption, uncontrolled contamination of the surface by impurities from the solutions, aggressiveness of the reagents, complexity of mechanization and automation.

Depending on the method of activation of the oxygen molecules, oxidation destruction is divided into heat treatment in an oxygen atmosphere and plasmochemical destruction.

Heat treatment in an oxygen atmosphere is realized at a temperature of 800°C at which the photoresist and the contamination are simultaneously destroyed and removed as a result of annealing and oxidation. Unfortunately, high temperatures lead to irreversible structural changes connected with oxidation, sublimation and burning in of the residual contamination. The oxygen of the air can be activated by ultraviolet radiation, which makes it possible to reduce the substrate treatment temperatures to 250°C. The process of removing the photoresistive mask can also be accelerated by introducing about 2% ozone into the air.

Plasmochemical destruction is treatment in a high-temperature high-frequency oxygen plasma at a pressure of $1.2 \cdot (10^2 \text{ to } 10^3)$ Pa.

In a high-frequency oxygen plasma the excited oxygen molecules, atomic oxygen and ozone are chemically active. Decomposition of photoresist in the oxygen plasma is of a chain nature; the resultant products with low molecular weight volatilize, and on being subjected to further oxidation, they decompose to the end products of carbon dioxide CO_2 , nitrogen oxide NO_2 and water H_2O . It is possible to use hydrogen, nitrogen and moisture additives as the catalyst accelerating the process of removing the photoresistive mask. Inorganic contaminants do not form volatile compounds during oxidation; for removal of them, carbon halides, for example, freon, are added to the plasma.

5-5. Photomask Production Technology

In order to obtain an exact pattern of the photomask, first an enlarge image of one module is drawn. A module is a single image of the IC elements or the set of them executed on the corresponding scale with respect to the dimensions of the elements of the topologic drawing of the given IC process layer. The image of the module is successively diminished and multiplied, that is, it is repeated a multiple number of times with respect to the working zone of the standard photomask.

FOR OFFICIAL USE ONLY

For series production of microcircuits using contact photolithography, as they are used the photomasks wear (defects accumulate). Therefore working copies are made from the standard photomasks which, in case of damage or wear, can be reproduced.

Let us consider the basic steps in the production technology of working photomasks.

The primary original is made on a special device -- coordinatograph. The primary original is a layout of the microcircuit module made on an enlarge (1000:1, 500:1, 200:1) scale intended for the manufacture of photomasks by the method of successive reducing and multiplication.

In order to make primary originals, substrates 600×600 to 1200×1200 mm² made of plate glass or polyester film 6-10 or 0.05-0.2 mm thick, respectively, are used.

A thin continuous vaseline layer is applied to the surface of the substrate, and then by multiple spraying, a layer of lacquer or nitro enamel 30-50 microns thick which dries in several hours at a temperature of 30-40°C.

After formation of the film it is cut along the outline of the pattern on the coordinatograph using a cutting tool with a diamond or tungsten carbide tip fastened in a rotary head. The displacement of the cutting tool with respect to a given trajectory and with the required precision is controlled by special mechanisms driven manually, by rotation of lead screws or automatically using a program unit. Before the operation of cutting the film, adjustment to optimal cutting depth is needed so that the substrate will not be cut through and the cutting tool will not be dulled if the original is made on a glass substrate. After cutting through the outline of the layout, the excess sections in the lacquer film are removed. The coordinatographs insure accuracy of cutting out the pattern within the limits of ± 50 microns.

An intermediate photomaster is made by the method of photography. The intermediate photomaster is an image of the original with element sizes, intermediate between the element sizes on the master and the corresponding dimensions given by the topologic drawing.

The intermediate photomaster can be obtained by copying the primary master using a reducing camera (Figure 5-13), which is a photographic device for precise photocopying of flat objects with 10-50fold reduction.

If the intermediate master is made with image reproduction, precise shifting of the holder with the photographic plate in two mutually perpendicular directions is provided for in the reducing camera. After exposure of the image of the primary master, the photographic plates are processed: they are developed, washed, fixed, go through secondary washing and drying. The tedious and multistage process of making a primary intermediate photomaster using a coordinatograph and reducing camera can be replaced by one process of successive photoprinting of the elements of the layout on the photographic plate which can be done by two methods:

Scanning of a focused light or electron beam computer-controlled by a given program;

FOR OFFICIAL USE ONLY

Photocomposition in which the entire pattern is divided into elementary rectangles with different ratio of sides and defined orientation of them with respect to angle; then these rectangles are projected successively on the required locations of the photoplate using the coordinate table and diaphragms.

In Soviet practice the method of photocomposition is more frequently used. Microphotocompositors EM-508, EM-519, EM-549 are distinguished by high precision and output capacity.

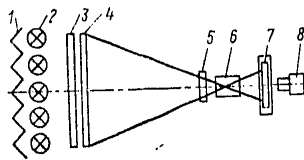


Figure 5-13. Diagram of a reducing camera.

- 1 -- reflecting shield; 2 -- light source; 3 -- dispersing glass;
- 4 -- primary master holder; 5 -- light filter; 6 -- objective;
- 7 -- holder with photographic plates; 8 -- microscope

The manufacture of the standard photomask is one of the responsible steps in the technological process. The standard photomask is the first photomask in the technological manufacturing process with element dimensions corresponding to the dimensions of the topologic drawing of the given technological layer. The standard photomask is designed for subsequent manufacture of the working photomasks.

The standard photomask is made by reducing the image of the intermediate master by the dimensions of the drawing of the working photomask and multiple repetition of this image for the defined working zone of the light sensitive plate. The basic parameters and compatibility of the set of photomasks are realized in this step.

At the present time basically the method of successive multiplication of a single image using precision photocopiers (step-repeater cameras) is used to make standard photomasks. The photocopier (Figure 5-14) consists of a projection optical exposure system, a precision coordination table and control panel. Before reproduction, a final check is made, foreign particles are removed and the intermediate photomaster is retouched. Then the intermediate photomaster is installed with the image down on the base frame, and it is fixed by the fast-drying adhesive.

The reproduction process consists of successive alternation of exposure and shifting of the photographic plate on the coordination table to the next coordinate position. The required coordinate positions are determined by the precision coordinate system, and they are given by a special reproduction program. In addition to the single-position photocopiers there are also multiposition photocopiers in which it is possible to make a number of photomasks on several photographic plates simultaneously (and with a high degree of matching).

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

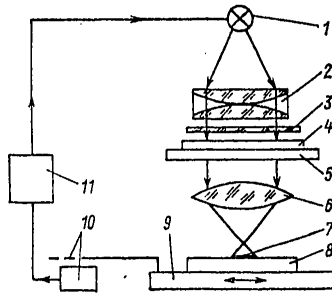


Figure 5-14. Diagram of successive reproduction using a single-position photocopier
 1 -- light source; 2 -- condenser; 3 -- light filter; 4 -- intermediate master; 5 -- base frame; 6 -- objective; 7 -- image on the photographic plate; 8 -- photographic plate; 9 -- coordination table; 10 -- measuring system; 11 -- exposure module control system

Then the photographic plates with the latent image of the photomask pattern go through further operations: development, washing, fixing and drying.

The manufacture of working photomasks is an ordinary photolithographic process (the standard photomask performs the role of an ordinary photomask), and it is used for reproducing the standard photomasks. A working photomask is a photomask designed directly for matching and exposure in photolithographic processes when making basic products.

The process of making working photomasks begins with preparation of the base -- the glass plate (or flexible polymer film). The plate is carefully cleaned and activated to improve adhesion of the film of the pattern material applied to it. Then a layer of chromium, iron-oxide, silicon monoxide, chalcogenide glass or others is applied to the surface of the glass base. Then a photoresistive layer is formed, exposure takes place, for example, in the EM-523 Soviet unit, and further operations are performed to transfer the pattern to the corresponding films.

For the manufacture of photomasks it is necessary to insure that a minimum amount of dust particles get on the billets, the surface of the optical objectives, the mechanical assemblies of the coordination tables. The basic equipment for making photomasks is placed in a class 2 cleaning facility. The operations of preparation, application and drying of the photoresist, the multiplication, photochemical treatment of the substrates are performed in the class 1 clean rooms with no more than four dust particles per liter of air.

A three-step system for making photomasks (the master -- intermediate master -- standard photomask) presented by us has been quite well worked out and satisfies the demands of microcircuit developers when a medium degree of integration and element sizes of no less than 2 microns are required. At the present time a two-step (master and standard photomask) and one-step (master) process flow charts

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

for making photomasks have been developed and are being introduced into production. In the two-step system the tedious operation of large-scale drawing and cutting out of the complex edges is excluded. For technical implementation of the single-step flow chart the most prospective are electron beam devices with multi-beam scanning or lasers with inertialess deflecting systems. The double and single-step systems make it possible to construct computer-controlled, completely automated production lines for the manufacture of photomasks.

5-6. Types of Rejects and Quality Control of Photolithography

The performance of the photolithographic process is accompanied by a number of control operations. The basic control steps are as follows:

Quality control of the photomasks;

Substrate surface control;

Photoresistive layer control;

Photoresistive mask control;

Control of the obtained pattern.

The rejection criteria and the admissible defects are regulated by the corresponding production control charts.

All disturbances of the quality of the image transfer during contact photolithography can be provisionally divided into local defects, inaccurate transfer of the dimensions of the pattern elements given by the photomask and inexact matching.

Local defects -- punctures (defects in the form of through holes or in the form of excess islets), cracks, scratches, foreign inclusions, ruptures of the conducting tracks, projections and depressions with respect to the boundaries of the pattern elements -- are individual. They do not pertain to all structures (modules) of the substrates, but only to individual ones, but they are very dangerous, for the appearance of even one local defect within the limits of the critical region will lead to rejection of the entire integrated circuit. The area of the module on which the excessive number of defects is inadmissible is called the critical region.

The causes for the appearance of local defects are as follows:

Natural contamination of the photoresist, the developer, the pickling agent, deionized water; external contamination from the atmosphere, from contact with packaging, fixtures, equipment and operators; defects in the photomasks -- punctures, the remains of opaque film on the transparent elements, chips in the glass, frosted spots, dust, dirt, photoresist residue; surface defects of the substrate, the presence of bumps on the film; defects of the photoresistive layer -- punctures, mechanical rupture as a result of the solid particles getting between the photomask and the photoresistive layer.

FOR OFFICIAL USE ONLY

Inexact transfer of the dimensions by the photomask, including unevenness of the edge, can occur basically from the following causes: inexact selection of the optimal methods and conditions of performing the production operations and also deviations from the given conditions; incorrect selection of the photoresist and (or) pickling agent for the substrate of the given type; unavailability of the clearance between the working surface of the photomask and the surface of the photoresistive layer and as a result of defects in the shape of the substrate or photomask, nonuniformity of thickness of the photolayer, nonuniformity of the films applied to the substrate, foreign particles between the photolayer and the photomask; imperfection of the equipment, and so on.

Inexact matching of the elements of the pattern can be obtained as a result of inexact matching of the photomasks of the set used, as a result of imperfection of the matching symbols or equipment, and in the case of visual matching, as a result of individual peculiarities and operator fatigue. The inexactness of the matching is estimated by using the NU-2E microscope (300-400 \times) by measuring the clearance between the boundaries of the matching patterns and calculation of the mismatch of their centers.

It is possible to improve the quality of photolithography only with a complex approach to this problem. A high yield of usable products of increased complexity can be obtained with simultaneous satisfaction of a number of conditions: the application of a united purification and filtration system (YeSOF) of the photoresists, water, all gases, air, and so on; periodic monitoring, proper organization of the cleaning of the photomasks, and as they wear out, replacement of them; the development of optimal technological processes and control of all photolithography steps; automation of the charging and transport operations.

The conversion to contactless photolithography is a radical solution to many of the problems when obtaining the configuration of microcircuit elements.

Test Questions and Assignments

1. What is the role of photolithography in IC production? List the examples of the application of photolithography known to us.
2. What components enter into the composition of a photoresist? What is their purpose?
3. What are negative photoresists? What are positive photoresists?
4. Which photochemical processes take place in the case of actinic irradiation in photoresist based on PVC, elastomers and NKd?
5. How do negative and positive photoresists transfer the image of the photomask?
6. List the basic characteristics of photoresists. Explain them.
7. Compare negative and positive photoresists from the point of view of resistance to acids and bases; from the point of view of resolution.

FOR OFFICIAL USE ONLY

8. What basic steps can the process of contact photolithography be divided into? Explain the production operations of each step for photolithography with respect to a silicon dioxide film in planar technology.
9. Formulate the basic requirement on the surface state of a substrate made for application of a photoresist layer. Explain it.
10. How are silicon plates prepared for application of a photoresist layer if they have been stored for a long time after thermal oxidation?
11. What is the essence of the methods and what is the technique for centrifuging, spraying, dipping, pouring and rolling? Compare these methods from the point of view of output capacity, the possibility of automation and quality of the applied layers.
12. How does the thickness of the photoresist layer depend on the centrifuge rpm and the viscosity of the photoresist? How does the density of the punctures depend on the thickness of the photoresist layer? How are the centrifuging conditions selected?
13. What is the mechanism and the technique for the photoresist layer drying process? What explains the preparation of photoresists on the basis of a combination of solvents rather than one solvent? What are the basic parameters of the drying process used on the photoresist layer and how are they selected?
14. Compare convective, infrared and microwave drying.
15. In what atmosphere is the photoresist layer dried?
16. Explain the necessity for the matching operation.
17. Which methods are used for matching? What is their essence and what is the technique?
18. What is the role of the exposure operation? Why are light filters and condensers needed for exposure?
19. Which light sources are used for exposure and why?
20. Name the parameters defining the exposure process. How are they selected?
21. What is the role of the exposure operation? What is the mechanism of the exposure of negative and positive photoresists? What is the technique for performing exposure operations?
22. What requirements are imposed on the developers?
23. What is indicated by coloring the alkaline developer a raspberry color?
24. What are the characteristic features of the exposure of negative and positive photoresists?

FOR OFFICIAL USE ONLY

25. Name the parameters of the development process. How are they selected?
26. Compare the purpose and the conditions of the operations involved in primary and secondary drying of a photoresist.
27. Compare the methods of obtaining the layout of the elements using photolithography: direct, direct with sublayer, inverse, inverse with sublayer.
28. Present examples of each of the methods of contact photolithography.
29. What is the purpose of the operation of removing the photoresistive mask?
30. What is the essence and the technique for performing the operation of removing the photoresistive layer by various methods? Compare these methods.
31. What is a photomask, what is the working zone of the photomask, a module? What materials are used to make photomasks and why?
32. Explain the three-step production process of making photomasks.
33. Give these characteristics of the two-step and one-step process flow charts for making photomasks.
34. What types of rejects are possible during photolithography and what are the causes of them?
35. What measures can be used to improve the quality of photolithography?

FOR OFFICIAL USE ONLY

CHAPTER 6. OBTAINING THE CONFIGURATION OF IC FILM ELEMENTS USING FREE MASKS

6-1. Free Mask Method

Classification of Methods of Obtaining Film Element Configurations. The required dimensions, configuration and mutual arrangement of thin and thick-film elements of microcircuits are obtained by using free masks. When using the free masks, the application of the film and obtaining the element layout are combined in a single production process. These processes, depending on the method of applying the film, are divided into the free mask method and stenciling. Both methods fail to provide high resolution; therefore when it is necessary to create precision thin-film and, sometimes, also thick-film elements, the method of contact photolithography investigated by us is used.

Essence and Special Features of the Free Mask Method. The free mask method consists in shielding the required sections of the IC substrate by the free mask from the flow of particles of the material deposited as the film. This method is used primarily when depositing the film elements of IC in vacuum devices.

The matching of the mask with a film pattern obtained on the substrate during the process of preceding deposition is either done in the air using special devices or directly in a vacuum chamber. This depends on the structural design of the intrachamber fittings of the vacuum devices.

The matching of the mask with the substrates using special devices in the air insures higher accuracy of obtaining the film pattern. The matching in an evacuated space is complicated, and it requires more expensive fixtures and the accuracy of matching is lower. In both cases it is necessary to consider the possibility of the deformation and expansion of the mask during heating in the deposition process, for it is located under the substrates.

Let us consider the basic features of the free mask method. The free masks are made in an independent auxiliary technological process, and when obtaining the film pattern on the substrates they are used multiply, just as other tools. Obtaining the pattern by means of a free mask differs from obtaining the pattern by contact mask in that the process is less labor-consuming and, consequently, less economically costly.

The free masks have a thickness appreciably greater than contact masks and, in addition, as a result of loose fit against the surface, sagging and mutual shifting of the mask in the substrate as a result of various TKLR during the deposition

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

process there is always a mask-substrate clearance that cannot be eliminated. This gives rise to lower resolution of the free mask method than in the case of photolithography.

The free masks make it possible to obtain a pattern that is only the inverse of the mask pattern. It is impossible to use free masks to obtain a closed or spiral pattern. The precision of the mask decreases as it is used.

In spite of the indicated deficiencies the free mask method is basic in the thin-film process of manufacturing microcircuits. When it is necessary to obtain a precision pattern with small dimensions of the film elements, photolithography is used.

Requirements on Free Masks. The requirement of obtaining a high-quality pattern under the film deposition conditions imposes defined requirements on the materials of free masks and their structural design.

The material for making masks must be machined to obtain flat smooth surfaces; it must have sufficient rigidity and elasticity that the mask will fit tightly against the surface of the substrate and not be deformed during heating. In addition, the mask material must clean well to remove buildups from preceding depositions, it must have low natural vapor pressure under deposition conditions, minimum gas release and not enter into undesirable interactions with the deposited material. Beryllium bronze, stainless steel, permalloy, molybdenum, tantalum, tungsten, invar and, among the nonconducting materials, graphite and photopyroceram, correspond to these requirements to the highest degree. Masks made of tantalum, molybdenum and tungsten are inert with respect to the deposited materials, and they have high mechanical properties. They clean well without wear. Invar has low TKLR and is used in the case of high requirements on the accuracy of the transfer of the element dimensions. However, when making monometal masks it is difficult to obtain high precision of reproduction of the hole sizes; therefore more accurate bimetal two and three-layer masks are used.

The mask thickness must not be too small so that the mask will remain sufficiently rigid and be deformed less, and it must not be too large so that the pattern will be transferred exactly. During deformations individual sections of the masks fit loosely against the substrate, which leads to a decrease or an increase in the sizes of the elements as a result of shading of the substrate by the deformed segment of the mask ("underdeposition") or as a result of the evaporated material getting under the mask ("overdeposition") (Figure 6-1).

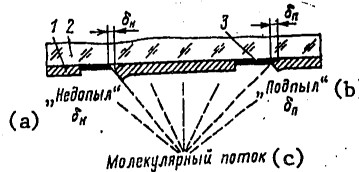


Figure 6-1. Distortion of the sizes of thin-film elements in the case of loose fitting of the free mask against a substrate.
 1 -- mask; 2 -- substrate; 3 -- deposited film

Key:
 a. underdeposition; b. overdeposition; c. molecular flux

FOR OFFICIAL USE ONLY

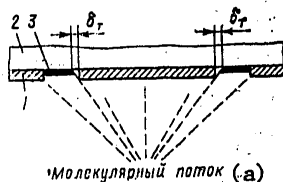


Figure 6-2. Distortion of the sizes of thin-film elements in the case of tight fit of the free mask against the substrate.

1 -- mask; 2 -- substrate; 3 -- deposited film; δ_T -- shadow width

Key:

a. molecular flux

The undeformed mask always gives a "shadow" on the edges of the layout, for not all of the particles of material arrive in the direction normal to the substrate (Figure 6-2). Therefore the thickness of the film deposited through the mask will be more uniform when using thin-mask. At the present time free masks 50-200 microns thick are used.

The precision of making the free mask layout is determined by the required accuracy of the rated values of the deposited elements of the microcircuits. Thus, in order to obtain resistors with a precision of the resistance ratings to $\pm 5\%$ and capacitors with accuracy of the ratings to $\pm 10\%$ requires precision of the hole sizes in the masks of ± 5 microns.

6-2. Free Mask Production Technology for Thin-Film IC

Mechanical Methods of Making Free Masks. In order to make monometal foil masks with simple pattern and relatively large hole sizes (no less than 0.3 mm), it is possible to use drilling, milling, cutouts, stamping and boring. However, these methods do not provide high precision and reproducibility of the results. The mask billet can be deformed sharply under the effect of the tools; therefore it is impossible to use billets that are too thin. The thickness of the metal form must be no less than 0.2 mm, and the spacing between holes, 1-2 mm.

Method of Electrospark (Electroerosion) Machining. Holes in metal billets are obtained using electrospark discharge which is created between two electrodes placed in a liquid dielectric medium (Figure 6-3). The free mask billet or set of billets made of 20-30 metal foil plates is one of the electrodes. The tool plays the role of the second electrode. A capacitor is charged from a voltage source through a current limiting resistance. When the maximum capacitance is reached, the capacitor discharges, that is, discharge occurs between the electrodes. After estimation of the electric strength of the electrode gap, the capacitor again begins to charge, and the charge-discharge cycle continues. With corresponding selection of the tool material and the pulse parameters the so-called polarity effect operates in which predominant melting and evaporation of one of the electrodes -- the mask billet -- take place.

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

As a result of the successive effect of a large number of discharges a hole is formed in the billet which repeats the shape of the tool. The waste is removed from the machine zone by forced pumping of the liquid through the spark gap. An extraordinary increase in energy of a single pulse leads to an increase in roughness of the holes; therefore in order to insure sufficient reproducibility with optimal pulse energy the pulse repetition frequency is increased.

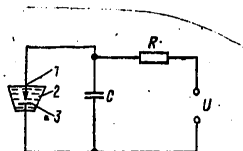


Figure 6-3. Diagram of electrospark machining.
1 -- tool electrode; 2 -- dielectric liquid; 3 -- free mask billet electrode

This method also fails to provide for obtaining masks with precision pattern and small hole sizes.

Manufacture of Free Masks by Photolithography. Monometal masks are made using electrolytic deposition of metal films or using through local pickling of the metal billet-foil. In both cases it is necessary to form a photoresistive mask. The pattern of the photomask used in the first case must provide for obtaining a photoresistive mask with layout that is the inverse of the layout of the manufactured free mask, and in the latter case, the direct pattern of the free mask. When making bimetal masks the electrolytic local deposition is used in combination with local chemical pickling.

The method of electrolytic deposition consists in local deposition of a nickel or copper layer on the surface of a polished steel plate shielded by a photoresistive mask. After formation of the photoresistive mask the steel plate is placed on the cathode of an electrolytic bath and a metal layer 0.05 to 1 micron thick is deposited. The nickel and copper films have low adhesion to a polished surface; therefore the mask obtained can be easily removed from the steel plate. It is possible to use the steel plate to obtain several free masks. In spite of the relative simplicity, the method is rarely used, for with small thickness insuring sufficient precision of pattern transfer, the strength of the free mask is inadequate. Increasing the thickness of the mask lowers the precision and reproducibility of the pattern. It is difficult to obtain masks with complex configuration of the pattern as a result of the possibility of rupture of the mask when it is separated from the steel plate. The efficiency of the production process is low as a result of the electrochemical deposition time.

The method of local chemical pickling is simpler and more efficient, and it is widely used to make monometal and bimetal free masks.

The flow chart for the manufacture of a monometal molybdenum free mask includes the formation of a photoresistive mask made of negative photoresists and local pickling of the molybdenum foil through openings in the photoresistive masks. The

FOR OFFICIAL USE ONLY

molybdenum is pickled electrolytically in a compound made up of $H_3PO_4:H_2SO_4:H_2O=4:1:4$ heated to a temperature of $160^\circ C$. The pickling proceeds not only in the direction of the normal to the foil surface; but also in the lateral directions, which leads to distortion of the holes and low precision of pattern transfer. On the average, the width of the side distortion region approaches the depth of pickling; therefore it is impossible to obtain holes that are wider than the foil thickness. The deep pickling does not provide reproducible results: the masks obtained have a divergence in the dimensions of the same elements of about 5-10 microns.

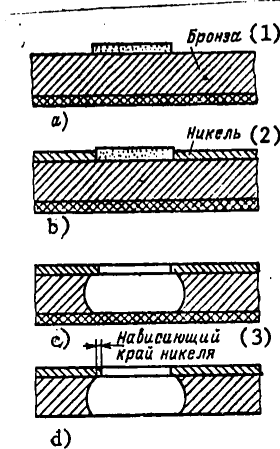


Figure 6-4. Diagram of the process of making bimetal, two-layer free masks.

a -- formation of the photoresistive mask and application of the protective coating to the back side of the substrate; b -- local electrolytic deposition nickel; c -- removal of the photoresistive mask and pickling of the mask base; d -- removal of the protective coating

Key:

1. Bronze
2. Nickel
3. Overhanging nickel edge

The process flow diagram for making a bimetal two-layer mask is presented in Figure 6-4. In the given diagram, in order to decrease the side distortions and, consequently, for more exact transfer of the image from the photomask, photolithography with a sublayer of nickel is used. At the end of the technological process the sublayer is not removed, but remains in the free mask and plays the basic role, being a type of mask, for it determines the shape and size of the elements during deposition. The plate billet made of beryllium bronze foil 100-150 microns thick in the free mask is the structural base insuring mechanical strength. A layer of nickel 7-10 microns thick is applied by electrochemical deposition, and it is held on the bronze base by the forces of adhesion. The holes in the base are pickled out using a mixture of chromium anhydride and sulfuric acid heated to $50-60^\circ C$ which does not act on the nickel.

FOR OFFICIAL USE ONLY

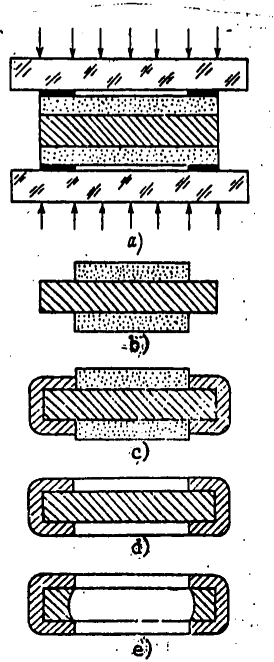


Figure 6-5. Flow chart of the manufacture of a three-layer bimetal free mask.
 a -- two-sided exposure; b -- development; c -- electrolytic application of nickel; d -- removal of the photoresistive mask; e -- two-sided pickling of the base

Bimetal free masks are distinguished by relative simplicity of the manufacturing process and quite high precision of transfer of the pattern from the photomask. Therefore the bimetal masks are the most widespread in the manufacture of film microcircuits. When depositing the films the substrate is placed on the nickel layer side; therefore the distortions of the mask base are not transferred to the pattern of the deposited elements. Nevertheless, they are undesirable, for weak overhanging edges of the nickel coating are formed which can be deformed, forming a gap with the substrate or even break off. The basic deficiency of two-layer bimetal masks is their strong deformation as a result of the differences in TKLR [thermal coefficient of linear expansion] of the nickel and the nickel base. In order to eliminate this deficiency preliminary heat treatment of the billets of the base and the finished masks is carried out or three-layer bimetal masks are used.

The plotted process flow diagram for making three-layer bimetal free masks is presented in Figure 6-5. The decrease in distortion of the base of masks is achieved by two-sided pickling. The total time for obtaining the holes is decreased and the size of the particles of the nickel layer overhanging the holes is decreased. Obtaining the "mirror" photomasks for two-sided exposure of photoresistive layers and also matching them before exposure presents special difficulty in the given technological process.

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

The manufacturing technology of three-layer bimetal masks is more complicated; therefore they are used when it is necessary to obtain a more precise and complex pattern, and when this is economically justified.

Manufacture of Free Masks by Beam Processing. Precision processing of a foil by a sharply focused electron, laser or ion beam permits a number of deficiencies of the photolithographic methods to be avoided. The processing by an electron or laser beam is based on their thermal effect on the billet material. When the corresponding temperature is reached, local evaporation takes place. Local removal of the material by using an ion beam is connected with the sputtering phenomenon. The rate of removal of the billet material can be regulated by varying the energy, the dosage and the duration of irradiation.

The ion and electron beams have quite high power and will permit us to obtain through holes in the foil up to 100 microns thick or more. The processing of the foil by laser pulses lasting $5 \cdot 10^{-8}$ to 10^{-7} seconds will permit removal of a layer about 1 micron thick from the surface in one clash and make it possible to obtain high-quality through holes.

The use of electron, laser and ion beams permits exclusion of the photomasks and chemical reagents from the technological process. The movement of the beams over the surface can be programmed and controlled by computer. These methods are highly prospective, for they make it possible to increase the efficiency of the technological processes, to improve the percentage yield of usable IC as a result of increasing the precision, and to decrease the sizes and size tolerances of the elements. The application of beam processing is still being delayed by the complexity and high cost of equipment.

6-3. Stenciling Method

Essence of the Method. The stenciling method used in microelectronics has ancient origins. Thousands of years ago the Egyptians decorated the walls of structures, tombs, and pottery by stenciling. The stencil is applied to the surface to be decorated, and paint is forced through the open parts of the stencil by a small board. Stenciling is also used at the present time: for applying patterns to fabrics in the textile industry (silk screening); in decorative applied art for reproducing graphics with simple patterns; and to manufacture small series of cards, and so on. In IC production stenciling has been borrowed from silk screening and the printed plates manufacturing technology.

The essence of the method consists in mechanical forcing of special pastes through openings in a free mask and subsequent heat treatment to give the film elements the required properties.

Production Technology of Free Masks for Thick-Film Technology. In order to obtain thick-film elements of given configuration, reticular, foil or combination masks are used, the manufacture of which is based on the use of the photolithography techniques.

For reticular masks (Figure 6-6) primarily a stainless steel or nylon screen is used. The stainless steel screen is more rigid and resistant to the effect of

FOR OFFICIAL USE ONLY

solvents and other components in the pastes. The screens are uniformly stretched by a special device and fastened to rigid frames made of aluminum alloy. The extra ends of the screen are cut off. Then the screen with the frame is carefully degreased in hot hydrogen peroxide, flushed in water and dried. A negative photoresist is applied to the working part of the prepared screens (by spraying, pouring or rolling), so that the screen mesh will be filled. After drying in the air the process of applying the photoresist is repeated. As a result of exposure and subsequent development at the required locations the photoresist is washed off, leaving the open sections of the screen -- "holes." In order to improve the resistance of the mask to wear, the photoresist is reinforced by drying at a temperature of 110°C.

Instead of photoresist it is possible to use pigment paper, for example, type VTU-115-56. The paper impregnated with photosensitive solution and the photomask prepared in advance are joined by the emulsion sides, they are tightly clamped together, and then the paper base is peeled off the pigment layer which remains on the photomask. Exposure, development and the pigment layer from the photomask are transferred to the screen. For this purpose the photomask is applied under load to the screen, it is dried, and after this, the photomask is separated. The masks obtained using pigment paper have high resolution, for the shielding of the pattern by the screen is excluded during exposure.

The manufacturing process of foil masks for thick-film microcircuit is the same as for thin-film microcircuits. The foil masks also can be bimetal and two or three-layer (with the application of two-sided pickling) based on beryllium bronze and electrochemically deposited nickel.

The combined masks are divided into composite and all-metal. For the composite masks, metal foil 25-125 microns thick is used in which the required pattern is pickled out using the method of photolithography. Then the foil is joined to the stainless steel screen. The primary difficulty when making composite masks is matching the screen mesh with the pattern lines in the foil.

When making all-metal masks, two-sided pickling of the molybdenum foil is used, the required pattern is etched out on one side, and the screen on the other. The production of these stencils is quite complex; therefore their cost is high.

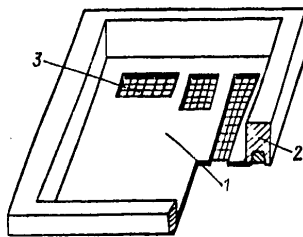


Figure 6-6. Reticular mask.
 1 -- solid part of the mask; 2 -- frame for attaching the screen;
 3 -- hole for paste to pass through

FOR OFFICIAL USE ONLY

Test Questions and Assignments

1. What methods are used to obtain the configuration of IC film elements?
2. What is the essence of the free mask method and in what case is it used?
3. How is the matching done to obtain the configuration of thin-film elements?
4. By comparing free and contact masks, the process for obtaining a layout using them, analyze the characteristic features, advantages and disadvantages of the free mask method.
5. What requirements are imposed on free masks?
6. Why have the methods of machining and electrospark processing of free masks not found broad industrial application for their manufacture? What is the essence of these methods?
7. Which masks and from what materials are masks made using photolithography?
8. What is the manufacturing process of monometal free masks using a photoresistive mask and electrolytic deposition of a metal film? Why is the metal film deposited on the surface of the photoresistive mask?
9. Which method of contact photolithography is used to make monometal masks for using local pickling (see §5-4)?
10. Which contact photolithographic method is used to make bimetal masks by local pickling? What requirements are imposed on the pickling agent of the mask base?
11. What are the basic deficiencies of two-layer, bi-metal masks and how can they be partially eliminated?
12. Compare the manufacturing technology of the three-layer bimetal mask with the manufacturing technology of the two-layer bimetal mask.
13. What is the mechanism for obtaining holes in foil when making free masks using electron, ion and laser beams?
14. What is the essence and application of the stenciling technique?
15. Which free masks are used in thick-film technology?
16. How can reticular masks be made?
17. What distinguishes composite masks from all-metal masks? Are they used frequently?
18. Compare all the methods of manufacturing free masks from the point of view of their advantages, disadvantages and practical application.

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

CHAPTER 7. NEW LITHOGRAPHY TECHNIQUES

7-1. Contactless Photolithography

Limitations of Contact Photolithography. Improvement of the functional complexity and, consequently, the degree of integration of IC faces the methods of obtaining the images with the problems of improving resolution, the achievement of maximum precision of matching the images and insuring minimum defect density in large working areas. When solving these problems contact photolithography encounters defined restrictions.

A significant limitation of contact photolithography is the unavoidability of mechanical damage to the working surfaces of the photomask and substrate, for when matching these surfaces are close to each other (10-15 microns), and during exposure they are tightly pressed against each other. As a result of mechanical wear of the film pattern, partial replacement of the photomasks is required, which makes it necessary to shut down the equipment and makes automation of the exposure process inexpedient.

On contact, the photomask presses dust particles, glass microparticles, and so on into the photoresistive layer. The photoresist is poured onto the photomask. In addition, any particles that are opaque for ultraviolet radiation getting between the photomask and photoresistive layer are also the cause of defects in the photoresistive mask.

Obtaining the tight complete contact between the photomask and the substrate is in practice an irresolvable problem as a result of bending of the plates (especially the epitaxial structures) and the substrates, nonidealness of the planeness of the contact surfaces, the presence of foreign particles between them, nonuniformity of the thickness of the various films and the photoresistive layer, and so on. The partial air gaps lead to intensification of the diffraction effects and cause additional expansion of the lines of the images obtained. However, as a result of the fact that the light refraction in the air is approximately half that in the photoresist, the transfer pattern is distorted still more. The cause of reduced resolution of the contact photolithography is also reflection of radiation from the substrate.

The maximum precision of matching in the case of contact photolithography is limited by the complexity of creating a system for fixing the transition from the "gap" position to the "contact" position; therefore making the transition from

FOR OFFICIAL USE ONLY

matching to exposure, a shift of the photomask relative to the substrate is possible. Errors can also appear as a result of the fact that the operator matches the patterns of the photomask in the substrate located in different planes.

In connection with the presented limitations of contact photolithography, further improvement of the processes of obtaining the configuration of the IC elements develops in the direction of applying contactless methods of exposure and decreasing the diffraction phenomena.

Photolithography on a Microgap. This method is based on the use of the effect of a double or multiple radiation source which is created in the exposure systems of special design. Ultraviolet beams hit the photomask and substrate inclined at identical angles to the common optical axis of the exposure system. As a result of inclination of the beams, the diffraction phenomena beyond the transparent sections of the photomask are eliminated or they are reduced to the minimum, the uniformity of irradiation of the substrate edges improves, the precision of the transfer of the pattern increases. As a result of exclusion of the diffraction phenomena, high resolution is achieved. Thus, for example, in the case of a layer of positive photoresist 1.8 microns thick it is possible to obtain a pattern element of less than 2 microns (with a photomask-substrate clearance of 10 microns) or less than 3.5 microns (with a clearance of 30 microns). The amount of clearance can be precisely given by a polyester film or peripheral strips on the photomask, for example, made of quartz. The industrial equipment for exposure on a microgap is appreciably more complicated than the contact exposure unit. At the same time the systems permit a reduction in exposure time to 2-5 seconds and insure uniformity of illumination of large-area substrates.

The absence of mechanical contact between the photomask and the substrate makes the service life of the expensive photomask in practice unlimited. Local damage to the photomask pattern by mechanical particles getting between the substrate and the mask is excluded. This completely eliminates the accumulation of photomask defects as they are used and rejection of photolithography as a result of the defects; therefore the percentage yield of usable IC also increases.

Projection Photolithography. Projection photolithography differs from contact photolithography by the technique used in the matching and exposure operations. The matching process is simplified, for special objectives are used to project the image of the photomask on the plane of the substrate or, vice versa, the substrate image is projected on the photomask plane and the operator observes the images of the substrate and the photomask in one plane. Accordingly, the problem of the depth of field of the objective is excluded along with the problem connected with it of precise establishment of the small substrate-photomask clearance. The matching time is decreased, and the precision is improved. After matching, the clearance between the substrate and the photomask remains, and the image of the photomask is projected on the substrate by an objective.

There are several optical systems for performing the projection photolithography process. The system presented in Figure 7-1 is most frequently used. This projection system provides for optimal illumination during matching and exposure. Only the objective and semitransparent mirror are located between the photomask and the substrate. The direct path of the beams from the light source through the objective to the photomask and then from the semitransparent mirror to the substrate completely

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

coincides with the return path of the beam observed using the microscope eyepiece. The operator matches in the plane of the photomask using a microscope; then he replaces the microscope by a source of ultraviolet radiation, condenser and light filter, and proceeds with exposure. A deficiency of this system is the necessity for rearrangement on making the transition from the matching operation to the exposure operation. In addition, an additional light source and condenser are required for matching.

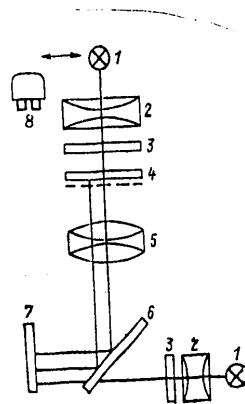


Figure 7-1. Optical system for projection photolithography.
 1 -- light source; 2 -- condenser; 3 -- light filter; 4 -- photo-mask; 5 -- objective; 6 -- semitransparent mirror; 7 -- substrate;
 8 -- microscope

The resolution of projection photolithography is higher, for diffraction of the radiation in the gap is excluded. The method of projection photolithography permits adaptation of the IC production process to a higher degree than the method of contact photolithography.

The basic technical difficulty in projection photolithography is complexity of developing high resolution objects for large image fields; therefore simultaneous projection of the complete pattern of the photomask on the substrate is not always possible. In the case of successive element-by-element projection of the image which can be carried out using photocopyers, the efficiency of photolithography drops sharply.

Projection photolithography imposes high requirements on the planarity of the substrate surface, the power and the monochromaticity of the ultraviolet source, and uniformity of thickness of the photoresistive layer.

In spite of the technical difficulties, projection photolithography is the most prospective optical method of obtaining the configuration of IC elements.

7-2. X-Ray Lithography

Limitations of Photolithography. The theoretical physical factor limiting the resolution of optical exposure systems are the light diffraction effects. The theoretically obtainable minimum line width λ_{\min} of the pattern as a function of

FOR OFFICIAL USE ONLY

the wave length λ of the irradiation used for exposure is defined in accordance with the Rayleigh number by the expression

$$l_{\min} = \frac{0,61\lambda}{n \sin \frac{\alpha}{2}}, \quad (7-1)$$

Key: 1. min

where n is the index of refraction of the medium between the objective and the image; α is the aperture angle of exit (the angle between the edge rays of a conical light beam exiting from the pupil of the objective in the direction of the image).

In real cases $n=1$; $\sin \alpha/2=0.95$; for $\lambda=400$ nm, $l_{\min}=0.26$ microns and for $\lambda=300$ nm, $l_{\min}=0.2$ micron. This is the theoretical limit for optical lithography. It has not been achieved in practice. The objectives have significant aberrations (distortions of the images), the photolayer has finite grain size, the patterns of the topologic layers of the microcircuits are complex combinations of elements, the diffraction patterns of which can be partially superposed on each other, and defects are introduced even in the photolithographic process itself.

In connection with what has been presented, further improvement of the technology of obtaining the configuration of IC elements is developing in the direction of introducing the methods of lithography using irradiation with shorter wave length by comparison with ultraviolet for exposure.

Principles of the Method of X-Ray Lithography. The basis for the method of x-ray lithography is interaction of characteristic x-radiation with x-ray resists leading to a change in their properties in the direction of a decrease or increase in resistance to the developers.

X-radiation is obtained by irradiation of the target by an accelerated electron flux. Depending on the nature of the interaction of the accelerated electrons with the atoms of the irradiated material, two types of x-radiation can occur: white or characteristic. White radiation is caused by breaking of the electrons on interaction of them with the electrons of the outer shells of the atoms of the irradiated material and subsequent transitions of the electrons of the external shells. Characteristic irradiation is caused by the interaction of accelerated electrons with the electrons of the internal shells of the atoms of the irradiated material, as a result of which the latter go from the internal shells to external or leave the atom. The electron transitions to the free internal shells of the atoms are accompanied by characteristic x-radiation. The difference in kinetic energies of the electrons of different internal shells is appreciably greater than external; therefore characteristic radiation has significantly shorter wave length than white radiation. X-rays with a wave length of 0.1-10 nm are used for x-ray lithography.

X-ray resists, just as photoresists, are divided into positive and negative. The former are depolymerized, and the latter are crosslinked under the effect of the x-rays. In photoresists ultraviolet rays are absorbed by photosensitive components

FOR OFFICIAL USE ONLY

of the molecules. X-rays are observed by the whole molecules, and electrons are knocked out from the internal shells of the atoms of the x-ray resist. The released electrons interact with the polymer and play the predominant role in the chemical conversions of depolymerization or crosslinking of the molecules. The resolution of the resist depends on the characteristic range of the electrons, that is, the distances which the free electrons travel in the polymer. Both types of resists (positive and negative) have identically high resolution.

In practice most frequently a positive resist based on polymethylmethacrylate is used (RMMA), which is distinguished by high stability of the properties, absence of sensitivity to ultraviolet radiation, resistance to the effect of acids (except hydrochloric acid).

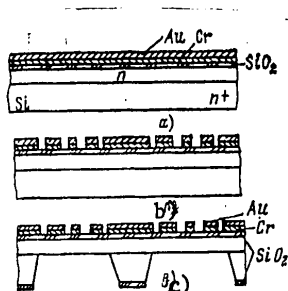


Figure 7-2. Diagram of the process of making silicon x-ray masks.
 a -- epitaxial growth of the n-Si layer, application of SiO₂ or Al₂O₃, Cr and Au films; b -- formation of the mask pattern;
 c -- local pickling of n⁺-Si

The masks for x-ray lithography must be made on a sufficiently transparent base for the x-radiation used. The material of the film pattern, on the contrary, must be opaque. For example, when using characteristic x-radiation of an aluminum target ($\lambda=0.83$ nm) silicon is used for the mask base, and for an opaque pattern, gold.

The production technology of patterns for x-ray lithography is a quite complex problem. A slightly alloyed layer of silicon about 3 microns thick is grown on a highly alloyed silicon plate. Films of silicon dioxide or aluminum oxide, chromium and gold (Figure 7-2) are applied over this layer. Then the pattern is formed in the double layer of chromium and gold. Then, using a contact mask made of silicon dioxide, the highly alloyed silicon is removed by pickling which does not affect the slightly alloyed layer. In order to obtain a flat, unbending mask of large area, the pickling is done not over the entire surface, but locally, removing the silicon in regions corresponding to the arrangement of the pattern of one micro-circuit or one semiconductor device. Thus, the initial silicon plate is the structural base lending rigidity to the mask; the role of the mask itself is played by a thin layer of slightly alloyed silicon with a pattern of gold applied to it.

X-ray Lithography Process. The order of the production operations for obtaining images is the same as in ordinary contact photolithography. Let us consider the basic ones of them.

FOR OFFICIAL USE ONLY

The application of the layer of x-ray resist RMMA (dissolved in methylethylketone) to the substrate can be done just as the application of the photoresist, for example, the method of centrifuging. The layer thickness is 0.1-0.5 micron.

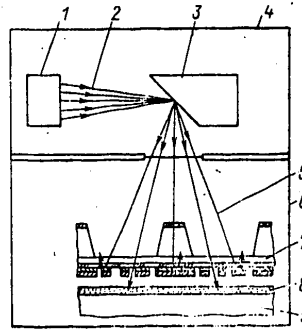


Figure 7-3. Schematic of x-ray exposure.

1 -- electron gun; 2 -- electron beam; 3 -- target; 4 -- x-ray formation chamber; 5 -- x-rays; 6 -- exposure chamber; 7 -- mask; 8 -- x-ray resist layer; 9 -- substrate

Matching is a serious technical problem. Several matching methods and devices are known. For example, matching by special depressions pickled out in the substrate and the mask, to the bottom of which matching marks made of a layer of gold or another material that absorbs x-rays well are applied. On the bottom of the substrate, directly under the matching marks, there is an x-ray detector which forms the mismatch signal. The signal is fed to the device that shifts the substrate in the plane parallel to the mask to an exact match.

X-ray exposure is done in the chamber, which in order to prevent attenuation of the x-rays is filled with helium or evacuated to a pressure of 1.33 Pa (Figure 7-3). The accelerated electron flux is aimed at the target, which serves as the source of the x-rays. The x-rays pass through a thin beryllium foil and transparent sections of the mask and they project its image on the layer of x-ray resist. The beryllium foil shields the x-ray resist from thermal radiation and secondary electrons which can cause polymerization of the x-ray resist. In order to decrease the erosion of the image projected on the substrate as a result of divergence of the angle of incidence of the x-rays, the distance from the target to the substrate is increased as much as possible. The clearance between the substrate and the mask is 3-10 microns. It is impossible to make the exposure without clearance, for the silicon mask is very brittle.

The exposure time $t = D / (\phi + \mu)$, where D is the required radiation dosage; ϕ is the incident x-ray flux; μ is the coefficient of absorption of the x-rays by the resist layer. The exposure time in x-ray lithography is from several seconds to tens of minutes. The RMMA film is exposed uniformly over the entire thickness, which insures that a sharp vertical profile of the edge of the image will be obtained.

The RMMA layer is exposed in a mixture of 40% methylisobutyl ketone and 60% isopropyl alcohol.

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

The x-ray lithography is distinguished by high resolution. A limitation of the resolution is insufficient absorption of x-rays by the mask pattern film. The absence of contact of the mask with the resist lowers the level of the defects and increases the service life of the masks. The contamination transmits x-rays and, consequently, is not transferred to the resist pattern. The reflections and dispersions of the radiation have no influence on the transfer of the image.

The basic deficiencies of x-ray lithography are large delay during exposure and the phenomenon of distortion of the image with respect to the pattern field (distortion), which is explained by the effect of the mechanical stresses occurring in the mask when making it.

7-3. Electron Lithography

Fundamentals of the Method of Electron Lithography. Electron beam exposure is done in vacuum devices, and it is based on the nonthermal interaction of accelerated electrons with electronoresist. Various polymers are used as the electronoresists, including photoresists. However, preference is given to special electronoresists insensitive to visible and ultraviolet radiation. The absence of light sensitivity of the electronoresists facilitates handling them in the process of manufacturing the microcircuits. During exposure the electronoresists must have no natural vapor pressure and must not form chemical compounds which contaminate the vacuum chamber of the device.

Electron bombardment causes excitation and ionization of the electronoresist molecules. The presence of the electric field of the electron increases the number of methods of rearranging the electronoresist molecules. Having high energy reserve, the electrons rupture almost all the chemical bonds on their path. Transverse crosslinking of the molecules takes place simultaneously. In each specific example usually some of these effects predominate. Accordingly, electronoresists, analogously to photoresists, are divided into negative and positive.

It has been demonstrated experimentally that the degree of crosslinking of negative electronoresists and the degree of depolymerization of positive electronoresists are directly proportional to the radiation dosage, that is, the magnitude of the electron charge per unit irradiated area. The crosslinking or depolymerization take place completely if the energy of the incident electrons is sufficient for their free path length to exceed the thickness of the electronoresist layer.

The exposure of the electronoresist can be accompanied by undesirable phenomena: contamination of the electronoresist with impurities made of the residual gases of the vacuum chamber of the device; accumulation of an electric charge which is the cause of distortion and worsening of the focus of the beam (this phenomenon is eliminated by the application of a thin transparent film of metal for the electrons, for example, aluminum); the occurrence of radiation effects in the substrate; random dispersion of part of the electrons, which has a negative effect on the verticalness of the pattern walls.

For electron beam exposure, devices are used with an accelerating voltage of 10^4 to $4 \cdot 10^4$ volts, which correspond to an electron wave length of 100-50 nm. In practice electronolithography is used to obtain images with element sizes of

FOR OFFICIAL USE ONLY

0.1 to 0.2 micron. The wave length of the electrons and the diameter of the electron beam decrease with an increase in the accelerating voltage. Obtaining electron beams with a wave length of less than 0.1 nm is considered technically achievable, that is, theoretically the resolution of electronolithography can reach values close to 10^{-4} micron.

At the present time two methods of electronolithography are used: treatment with a focused single beam (scanning electronolithography) and electron projection of the entire image (projection electronolithography).

Scanning Electronolithography. For exposure by a focused beam of electrons, scanning electron microscopes or specialized devices -- electron beam accelerators (ELU) -- are used. The scanning electron microscopes make it possible to obtain pattern lines about 0.1 micron wide in an area to 6 cm²; the line width obtained in the electron beam accelerators is about 1 micron in an area of $4 \cdot 10^{-2}$ cm² (the Japanese EBX-2B device).

The shifting, blocking and unblocking of the electron beam are realized using a remote copying device, a phototelegraphic unit or computer. In the first case (Figure 7-4) a mask made of glass with enlargement of the pattern is used as the program for controlling the electron beam. The light brightness is varied beyond the mask in accordance with its pattern. The photomultiplier located beyond the mask reads and intensifies this brightness signal. The intensified signal is transmitted to the electron beam accelerator for control of the electron beam in accordance with the image on the mask, but on a correspondingly reduced scale. When controlling an electron beam from a remote copier the exposure of the electrono-resist is realized by raster scanning (Figure 7-5, a), that is, the beam is moved line by line over the entire surface, switching on and off at the required locations. During electron beam control using a phototelegraphic set the pattern is transferred to the electrono-resist from the drawing paper. The beam is controlled from a computer more efficiently, for vector scanning is used (Figure 7-5, b). In this case the electron beam is shifted not over the entire surface of the electrono-resist layer, but it scans only the programmed sections, switching on at the points of transition from one element to another. Using the focused beam it is also possible to draw ("mill") the programmed outline. This decreases the total exposure time. In addition to shifting the beam, computer-controlled shifting of the table takes place on which the substrates are located. This increases the exposure area. Two methods of using the table are used: continuous and by the amount of one crystal or plate. The ELU-EVM [electron beam accelerator-computer] system permits exposure directly after development of the topology using a computer without making masks, which greatly reduces the time for the development of the new microcircuits and insures a minimum development-production cycle.

The matching is done using reference marks. When the electron beam falls on the edge of the mark, the signal from the reflected electrons changes and falls on the detector, transmitting information about the mismatch in the computer. The computer changes the shift of the beam, matching the created image with that previously obtained on the substrate. Usually the beam is deflected by small amounts, about 2 mm. The matching based on isolation of the signal from the reference mark is theoretically different from the photooptical methods of matching requiring microscopes. The precision of this matching is ± 0.5 micron.

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

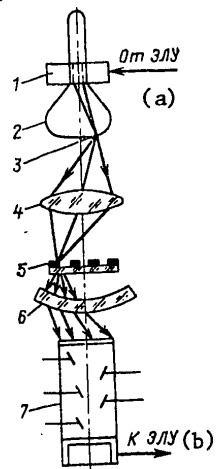


Figure 7-4. Diagram of the remote copying device for controlling an electron beam.

- 1 -- reflecting system; 2 -- kinescope;
- 3 -- light spot on the kinescope screen;
- 4 -- objective; 5 -- mask; 6 -- lens;
- 7 -- photomultiplier

Key:

- a. from the electron beam accelerator
- b. to the electron beam accelerator

The method of successive treatments by a focused single beam of electrons insures high resolution and precision of matching. Lines 0.1-0.4 micron wide are actually obtained.

The basic deficiency of the method of treatment with a focused electron beam is low output capacity as a result of the long exposure time, for when the beam diameter is decreased it is necessary to reduce its current and increase the number of scanning lines.

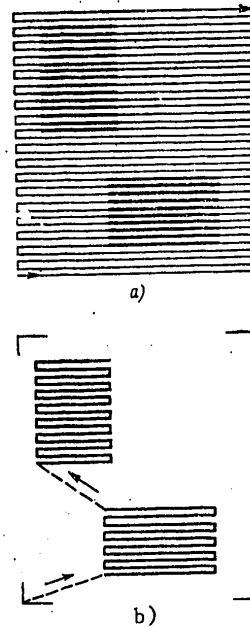


Figure 7-5. Diagram of the raster scanning of a beam (a) and vector (b) scanning of the beam on exposure of electronresists

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

Projection Electronolithography. The entire pattern of the mask is simultaneously transmitted to the electronresist (Figure 7-6). The basic element of the intra-chamber device of the vacuum unit is a three-layer photocathode which simultaneously plays the role of a source of electrons and the role of a mask. The photocathode is a polished quartz plate, on the surface of which the pattern is made on a 1:1 scale from a layer of titanium dioxide. A palladium film 4 nm thick is applied over the titanium dioxide over the entire area of the pattern. The titanium dioxide is opaque for ultraviolet radiation, and the palladium film has high photoemission properties. The photocathode on the quartz side is irradiated by ultraviolet, and on the back side from the sections coated only with the palladium film, photoelectrons are emitted. Then they are accelerated by the electron field, and by using the focusing system an image is projected from the cathode on the layer of electronresist in practice without distortion. In the structural design of the device provision is made for a deflecting system which permits the projected image to be shifted in the plane of the substrate and at the same time, matching to be provided. The mismatch signal of the matched patterns is picked up from the reference marks using detectors. The precision of the matching is ± 0.25 micron.

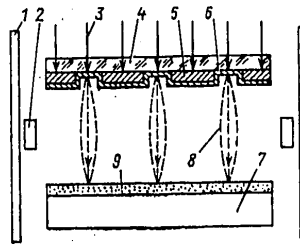


Figure 7-6. Projection electronolithography diagram.

1 -- deflecting system; 2 -- focusing system; 3 -- ultraviolet radiation; 4 -- quartz base of the photocathode mask; 5 -- titanium dioxide; 6 -- palladium film; 7 -- substrate; 8 -- electron flux; 9 -- electronresist layer

The method of electron projection of the entire image has good resolution (lines 1 micron wide are obtained on the working field 25 mm in diameter), great depth of field by comparison with the optical methods, reaching up to ± 50 microns. The service life of the photocathodes is greater than that of the photomasks. The efficiency of the process is comparable to the efficiency of photolithography.

The deficiencies of the method include the complexity of connecting the detectors for matching and complexity of making the precision photocathodes.

In conclusion let us note that the methods of contact photolithography and the free mask method widely used in production are not adaptable for fast rearrangement of the production of systems of a broad nomenclature without significant capital expenditures. This is connected with prolonged processes of making sets of photomasks or free masks. The method of electronolithography permits the expensive coordinatographs, devices for cutting out the enlarged images and other attachments

FOR OFFICIAL USE ONLY

for making the photomask masters to be eliminated. During electronolithography, changes in the configuration of the microcircuit structure can be produced by corrections in the program if a computer is used as the device for assignment and reading of the program. The total number of operations in the technological process of electronolithography is much smaller than in the optical processes. For example, the labor consumption of the manufacture of the masks for electronolithography is six times less than the labor consumption for the manufacture of photomasks. The technological process using electronolithography has the greatest advantage when creating special-application microcircuit, that is, circuits of large nomenclature and small circulation, for it permits significant reduction of the production preparation time as a result of the possibility of fast changing of the electron beam displacement program. Electronolithography is an easily automated technological process; therefore, in spite of the complexity of the equipment by comparison with optical and x-ray equipment, it is highly prospective for the manufacture of LSI with more than 10^5 elements.

Test Questions and Assignments

1. Repeat the basic operations of contact photolithography (Figure 5-5).
2. List the basic problems facing the technology of obtaining images of microcircuit elements.
3. What are the basic limitations of contact photolithography?
4. What methods of obtaining the configuration of the elements are included in the optical methods?
5. What is photolithography on a microgap and what are its advantages?
6. What is the essence of projection photolithography, what is the technique for carrying it out?
7. What is the basic technical difficulty of projection photolithography?
8. What are the basic advantages of projection photolithography by comparison with other optical methods of obtaining images?
9. What are the theoretical physical limitations of the optical methods of obtaining images?
10. What distinguishes the characteristic x-radiation from white? Which of them is used in x-ray diffraction and why?
11. How is x-radiation obtained for x-ray lithography?
12. What requirements are imposed on the materials and what is the structure of the masks for x-ray lithography?
13. Make up a flow chart for manufacturing the mask for x-ray lithography and check it using Figure 7-2.

FOR OFFICIAL USE ONLY

14. What are the requirements on the silicon pickling agent when making the mask? Why is pickling self-halting, leaving a layer of slightly alloyed silicon?
15. What is the mechanism of effect of x-rays on x-ray resists?
16. How is matching done in x-ray lithography?
17. Explain the x-ray exposure system (Figure 7-3).
18. Compare x-ray lithography with the optical methods of obtaining images.
19. What is the mechanism of effect of accelerated electrons on electronoresist?
20. What is scanning electronolithography, and in what devices is it carried out?
21. How is electron beam control realized?
22. What distinguishes vector scanning of a beam from raster scanning?
23. What is the basic advantage of the ELU-EVM [electron beam accelerator-computer] systems?
24. What is projection electronolithography and how is it carried out?
25. Explain the structural design of a photocathode (Figure 7-6) and the role of all of its elements. Compile a photocathode manufacturing system.
26. Why does the photocathode have long service life by comparison with photo-masks used in contact photolithography?
27. What are the basic features, advantages and disadvantages of the method of electronolithography by comparison with other methods of obtaining images in microcircuit production?

FOR OFFICIAL USE ONLY

CHAPTER 8. METHODS OF OBTAINING THIN FILMS

8-1. Method of Thermovacuum Deposition

Principles of the Method. The method is based on creating a directional vapor flow of the material and subsequent condensation of it on the surfaces of substrates having a temperature below the temperature of the vapor source. During condensation a film is formed from individual atoms or molecules of the vapor. The process of thermovacuum deposition can be broken down into four steps: the formation of the vapor, application of the vapor from the source to the substrate, condensation of the vapor on the substrates, formation of nucleating centers and film growth.

The vapor is formed by evaporation or sublimation. The materials are converted to vapor at any temperature above absolute zero, but in order to increase the intensity of vapor formation heating is required. With an increase in temperature the mean kinetic energy of the atoms rises and, consequently, the probability of rupture of the interatomic bonds increases. Atoms split off the surface and are propagated in free space, forming vapor.

The specific evaporation rate, equal to the number of grams of material evaporated per second from a surface area of 1 cm² is defined by the expression

$$G_{\text{max}} = 5.85 p_s \sqrt{\frac{M}{T}} \quad (8-1)$$

(1)

Key: 1. evap

where p_s is the saturated vapor pressure of the heated material, Pa; M is the molecular mass of the material, g/mole; T is the temperature of the evaporator, K.

The saturated vapor pressure depends strongly on temperature. The temperature increments every 5-10% above the evaporation temperature lead to an increase in saturated vapor pressure and, consequently, the evaporation rate by one order.

The propagation of the vapor from source to substrate is realized by diffusion and convective mass transfer which is primarily influenced by the degree of vacuum. In order to obtain high-quality films it is necessary that the atoms and molecules travel to the substrates without collisions with the molecules of the residual gases. This is possible under the condition where the free path length of the vapor particles λ is greater than the source-substrate distance.

FOR OFFICIAL USE ONLY

From the kinetic area of gases the mean free path length is determined by the expression

$$\bar{\lambda} = \frac{1}{\sqrt{2} \pi d^2 N}, \quad (8-2)$$

where $\bar{\lambda}$ is the mean free path length, cm; d is the effective diameter of the molecules, cm; N is the concentration of the molecules for given pressure and temperature, $1/\text{cm}^3$.

The concentration of the molecules and the pressure are related to each other by the expression

$$N = \frac{p}{kT}, \quad (8-3)$$

on substitution of which in expression (8-2), it is possible to determine the relation of the mean free path length to the pressure

$$\bar{\lambda} = \frac{kT}{\sqrt{2} \pi d^2 p}. \quad (8-4)$$

Beginning with a vacuum of $1.2 \cdot (10^{-2} \text{ to } 10^{-3})$ Pa, the mean free path length (47 cm to 4.7 meters) will become appreciably higher than the source-substrate distance (10-15 cm), and the probability of collision in the drift space with molecules of the residual gases is low. It is possible to consider that the atoms of the deposited material are propagated rectilinearly by a directional atomic flux, retaining their energy until they encounter the substrate.

The condensation of the vapor on the substrate depends on the temperature of the substrate and the density of the atomic flux. According to the modern theory of condensation, electrically neutral vapor particles, approaching the substrate surface, fall into the field of the forces of attraction generated by the instantaneous dipole moments of the surface atoms and molecules. At a very close distance the force of repulsion begins to act on the vapor particle. The vapor atoms, reaching the substrates, can instantaneously be repelled from it (elastic collision), adsorbed and after some time, repelled from the substrate (reevaporation), adsorbed and after brief migration on the surface, finally remain on it (condensation).

Condensation of the atoms takes place if the energy of the binding of them to the substrate atoms is greater than the mean energy of the substrate atoms; otherwise the atoms are repelled from the substrate.

The temperature above which all the atoms are repelled from the substrate and the film is not formed is called the critical condensation temperature. The critical temperature depends on the nature of the materials of the film and substrate and the state of the substrate surface. The critical density of the atomic flux for a given substrate temperature is the least density on which the atoms condense on the substrate.

The formation of the nucleating center takes place as a result of the atoms finding locations corresponding to the minimum free energy of the atom-substrate system.

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

The growth of the nucleating centers takes place as a result of connection of new atoms migrating over the surface or falling on the nucleating center directly from the source-substrate drift interval. As the vapor condenses, the nucleating centers grow; connecting bridges are formed between them, the nucleating centers coalesce into large islets. Then the stage of coalescence of the islets comes with the formation of a single grid. The grid becomes a solid film which begins to grow in thickness. In the given step the effect of the residual gases on the substrates and the growing film must be reduced to a minimum. This can be insured by increasing the degree of vacuum or increasing the vapor formation rate. However, the creation of supervacuum devices presents significant technological difficulties, and in addition, the evacuation time of the operating chambers is high.

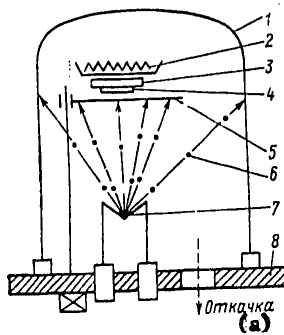


Figure 8-1. Diagram of the process of thermovacuum deposition. 1 -- vacuum chamber hood; 2 -- substrate heater; 3 -- substrate holder; 4 -- substrate; 5 -- spot valve; 6 -- articles of evaporated material; 7 -- evaporator with weighed sample of the film material; 8 -- supporting plate

Key:

a. evacuation

In microcircuit production frequently metal films are deposited on dielectric and semiconductor surfaces. For such combinations of materials of the condensate and substrate which have no chemical affinity, the reevaporation is high; adsorption and, consequently, the nucleating center formation and growth of thin films are complicated significantly.

Thermovacuum Deposition Technique. The diagram of the process of thermovacuum deposition is presented in Figure 8-1. The stationary and removable equipment of the hooded device is periodically cleaned to remove encrustations of the preceding depositions. Weighed samples of the deposited material are degreased, pickled to remove oxide films and contamination. Directly before deposition the evaporators and the weighed samples are annealed. Then the valve is opened and deposition on the substrate takes place.

The basic parameters of the process of thermovacuum deposition are the deposition rate of the film and the substrate temperature.

FOR OFFICIAL USE ONLY

The deposition rate of the film is directly proportional to the evaporation rate. The flux of atoms condensed on the substrate surface at fixed temperatures of the evaporator and substrate is defined by the expression

$$N_n = k \frac{A}{l^n} N_n \quad (8-5)$$

Key: 1. evap; 2. flux

where N_{evap} is the flow of evaporating particles; l is the distance from the source to the substrate; A is the coefficient taking into account the shape and the molecular-kinetic characteristics of the evaporator; k is the compensation coefficient -- the ratio of the number of molecules condensed on the surface of the substrate to the number of molecules impacting against it numerically equal to the ratio of the condensation rate and the evaporation rate.

Evaporation almost always takes place at temperatures exceeding the evaporation temperature of the material (the forced conditions) in order to insure quite high vaporization rates, and, as a result, to decrease the effect of the residual gases on the properties of the deposited films. At low evaporation rates the formation of loose, rough films is possible.

The temperature of the substrate is selected as optimal in order to insure condensation of the vapor and adhesion of the films to the substrates. The heating of the substrates is required for desorption of harmful materials (gases, moisture, oil from the pumps) which are the basic cause of poor adhesion. The substrate temperature also influences the structure of the deposited film and, consequently, its electrophysical parameters.

The uniformity of thickness of the films with respect to the area of the substrates when using the majority of simplest wire, strip and crucible evaporators is unsatisfactory. The film thickness is maximal at the substrate center, that is, in the section located directly above the evaporator, and it decreases to the periphery of the substrate. The uniformity of the thickness of the films can be increased as a result of increasing the distance between the evaporator and the substrate, but in this case the deposition rate decreases and, in addition, the operating chamber of the device has limited dimensions. At the present time the uniformity of the thickness of the films in a large lot of substrates is achieved by application of hooded devices insuring uniform rotation of the substrates fastened vertically to the generatrices of the cylinder, around the evaporators located along the central axis of the cylinder. The molecular flux from the evaporators is propagated through a stationary diaphragm. The operating principle of such evaporators consists in equalization of the average deposition rate at each point of the substrate surface.

The structure of the deposited film depends on the material, the state of the surface and the temperature of the substrates, the rate of deposition, and it can be amorphous, polycrystalline, fine-grained, polycrystalline, large-grained, and monocrystalline. The grain size of the metal films depends on the melting point of the metal. Metals with high melting point (tungsten, molybdenum, tantalum, platinum, and so on) form films with small grain sizes. The films become solid for comparatively small thicknesses. The low-temperature metals (zinc, cadmium, and so on) form large-grain films.

FOR OFFICIAL USE ONLY

At the contact boundaries of the grain structural defects are concentrated in the solid film. The presence of the boundaries itself introduces distortions into the crystal lattices of the grains. The cavities between the grains are frequently filled by amorphous contamination, oxides, and so on. Therefore the large-grain films have smaller concentration of the structural defects which explains the large stability of their electrophysical properties.

Thus, during thermovacuum deposition preference is given to increasing the deposition rate, and the film grain sizes increase as a result of an increase in the substrate temperature during deposition and annealing.

The films are annealed in vacuum devices directly after deposition at substrate temperatures which exceed the deposition temperatures somewhat. This is done to accelerate the structure and decrease the internal mechanical stresses of the films to increase their stability and improve adhesion to the substrate. During the annealing process the intergrain distances in the films decrease and as a consequence, the number of structural defects decreases. The resistance of the resistive and conducting films decreases in this case.

Advantages and Disadvantages of the Method. The method of thermovacuum deposition has been well mastered; it makes it possible to obtain any passive elements, metallization of the semiconductor structures; it is used when making photomasks, and so on. Using thermovacuum deposition it is possible to obtain films of metals, semiconductors and dielectrics. The method provides high growth rate of the films. The process is distinguished by a high degree of cleanness and it permits high quality films to be obtained in a high vacuum with comparatively low substrate temperatures. Comparatively easy automation of the process permits the creation of complex vacuum devices and computer-controlled complexes.

The disadvantages of the method include variation of the percentage ratio of the components with deposition of alloys and complex materials, and sufficient uniformity of thickness of the films on the large substrates, difficulty of obtaining films of refractory materials, high inertia when using the evaporators (after disconnecting the heating of the evaporators, the vapor formation continues; therefore the process of film deposition is stopped by using a mechanical slide valve), comparatively low adhesion of the films, short duration of preparations for the process and evacuation, and relative complexity of the equipment.

8-2. Ion Bombardment Sputtering

Principles of the Method. The methods of depositing films by sputtering the material in a gas discharge plasma, by comparison with thermovacuum deposition, expand the possibilities of obtaining films with given properties. They permit deposition of films of refractory materials, alloys, complex materials and materials with low vapor pressure.

The ion sputtering mechanism was briefly discussed above (§4-5) when describing ion pickling of substrates.

The material is sputtered in the form of neutral atoms or molecules, and about 1% of it is ionized by plasma electrons. Sputtering, in contrast to evaporation, does not depend on the vapor pressure of the deposited material. This permits

FOR OFFICIAL USE ONLY

application of films of refractory materials at relatively low temperatures. When obtaining alloy films, initially basically the component with the larger sputtering coefficient, for example, nickel in nichrome, is sputtered. The influx of the missing component is realized by diffusion; therefore equilibrium is quickly established, and the composition of the sputtered atom flux corresponds to the composition of the alloy. As a result, only the first few atomic layers of the film differ from the basic composition. The application of films of complex composition can be accomplished also by sputtering several different cathodes.

The particles are propagated in a gas discharge plasma. As a result of collisions with molecules of the residual gas, the directionality of the sputtered particle flux is lost, and the particles go to the surface of the substrates at different angles. Along with the particles of the sputtered material, molecules and ions of the residual gases also hit the substrates.

The process of formation of nucleating centers and the film formation on the surface of the substrates during ion sputtering differ significantly from this step in thermovacuum deposition. The sputtered material flux has a number of peculiarities: greater energy of the atoms reaching the substrate (by 1-2 orders), lower particle flux densities (by an order), random nature of the directions of impact of the atoms against the surface of the substrate, to a greater degree presence in the flux of the ions of the atoms along with the neutral atoms and also molecules of the residual gas.

All of this causes an increase in the substrate temperature, an increase in the tangential component of the velocity of the atom, that is, an increase in the migration rate along the surface, intensive desorption of the deposited molecules and residual gases.

It has been experimentally established that the films are deposited for any densities of the atomic flux and in a wide temperature range of the substrate, that is, during ion sputtering for the formation and growth of nucleating centers there is no critical temperature of the substrate or critical density of the sputtered material flux. Along with the mechanism of film growth from neutral particles, the charge mechanism growth operates here, for along with the atoms of the sputtered source, the ions have the inert gas and a relatively large number of source atoms ionized in the electrode gap hit the substrate. As a result of the additional electrostatic energy the charged particles increase the surface migration between the nucleating particles and accelerate their growth in the plane. The latter explains the fact that the films deposited by ion sputtering become solid at less thickness than the films obtained by thermovacuum deposition. The deposition rates of the films, as a rule, are lower than during thermovacuum deposition as a result of low flux densities. The film density is higher as a result of high energy of the deposited particles.

Cathode (physical and reactive) and ion-plasma sputtering are distinguished.

Physical Cathode Sputtering. In this form of sputtering the ion source is an independent glow discharge plasma which is created between the cathode and anode of a diode type vacuum chamber (Figure 8-2). The chamber is evacuated to maximum vacuum, and then argon is admitted to it to a pressure of $1.2 \cdot (10^{-1} \text{ to } 10^{-2})$ Pa.

FOR OFFICIAL USE ONLY

On feeding a voltage on the order of several kilovolts between the cathode and anode with a defined discharge current, glow discharge sets in which, by the nature of the potential distribution, can be divided into a number of regions in the spacing between the cathode and the anode: the dark cathode space, negative glow, dark anode space and positive column. The basic voltage drop applied between the cathode and the anode is in the dark cathode space adjacent to the cathode. In the positive column region of the discharge, the charged plasma particles, electrons and ions move diffusely. The ions incident in the vicinity of the cathode space are accelerated by the electric field and bombard the cathode. As a result of bombardment, the cathode sputtering and secondary electron emission take place. Thus, the cathode is simultaneously the source of material for obtaining the film and a source of electrons required to support glow discharge in the chamber. During movement of the electrons from the cathode to the anode, their energy in the cathode space increases and becomes sufficient for ionization of the gas molecules. The glow discharge is called independent, for the discharge itself insures cathode emission without auxiliary means. The positive columns performs the functions of the conductance section between the anode and cathode.

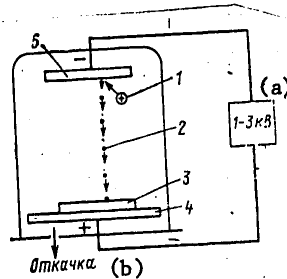


Figure 8-2. Diagram of the cathode sputtering process
 1 -- argon ion; 2 -- sputtering cathode particles; 3 -- substrate;
 4 -- anode; 5 -- cathode

Key:

- a. 3 kev
- b. evacuation

The amount of cathode materials sputtered per unit time Q_p is proportional to the voltage between the cathode and the anode U , the discharge current I , and it is inversely proportional to the product of the pressure in the vacuum chamber p times the distance l from the cathode to the surface of the substrate:

$$Q_p = k \frac{UI}{pl}, \quad (8-6)$$

where k is a constant.

With an increase in pressure p , the average free path length and also the dark cathode space length decrease. The movement of the sputtered particles from the cathode to the substrate basically is of a diffusion nature; therefore the displacement of the sputtered particles decreases. With an increase in the cathode-substrate spacing, the probability of collisions of the sputtered particles with the residual gases and, consequently, the probability of their return to the

FOR OFFICIAL USE ONLY

cathode, increases; therefore, in order to obtain the maximum growth rate of the film, the product $p\lambda$ must be minimal.

Let us estimate the effect of the discharge current I and the voltage U on the deposition rate considering the cathode sputtering coefficient S . The discharge current is determined by the number of ions incident on the sputtered cathode and the charge of the ions $I=qN$. The sputtering rate $V_p=NS=(I/q)S$. Increasing the voltage insignificantly increases the cathode sputtering coefficient, and increasing the discharge current leads to a noticeable increase in the sputtering rate. Therefore in order to increase the sputtering rate and, consequently, the growth rate of the films it is expedient to increase the discharge current and not the sputtering voltage.

The amount of material deposited on the substrate Q_0 is related to the amount of materials sputtered by the cathode Q_p by the expression

$$Q_0 = \frac{Q_p}{p^{\alpha} d^{\beta}}, \quad (8-7)$$

where $\alpha=0.06$ to 0.2 and $\beta=0.5$ to 2 are the coefficients which depend on the sputtering conditions.

The parameters of the cathode sputtering process are as follows: gas pressure in the vacuum chamber, voltage between the anode and cathode, discharge current, distance from the cathode to the surface of the substrates, sputtering time.

Optimal pressure (10^{-1}) Pa. At lower pressures the discharge current, and, consequently, the sputtering rate will be decreased. At higher pressures the increase in the number of collisions in the gas discharge gap also sharply decreases the sputtering rate. The voltage is maintained on a level of 1-3 kv. The greater the distance from the cathode to the substrates, the more frequently the sputtered atoms collide with the gas molecules and the smaller the number of them reaching the substrate. A distance between the cathode and anode which is 1.5 to 2 times greater than the width of the dark cathode space is optimal.

The advantages of cathode sputtering are as follows: low substrate temperatures during the film deposition process; greater film uniformity with respect to area of the substrates than in the case of thermovacuum deposition, for the surface area of the cathode is large by comparison with the area of the vapor source during deposition in a vacuum; inertialessness (sputtering begins on supplying the voltage to the electrodes, and it stops instantaneously on taking the voltage away); absence of the necessity for frequent checking of the source of the particles of the growing film (cathode); nonvariability of the stoichiometry of composition of the film by comparison with the cathode composition; high adhesion of the film to the substrates.

The basic deficiencies of the method of cathode sputtering are as follows: comparatively low deposition rates, contamination of the films by molecules of residual gases and more complex control of the technological process by comparison with thermovacuum deposition.

FOR OFFICIAL USE ONLY

In order to eliminate the contamination of the films, which has for a long time held up industrial application of the method, at the present time sputtering is used with shifting, sputtering of an asymmetric alternating current of industrial frequency, getter sputtering and combinations of these methods.

Cathode sputtering is widely used when obtaining dielectric films of SiO_2 , Al_2O_3 , Ta_2O_5 for thin-film capacitors, conducting Cu, Au films with a sublayer of Cr, Ti and Cr and NiCr films for thin-film resistors.

Sputtering of dielectric materials in diode systems on direct current is impossible, for the cathode surface is charged with positive ions, and further bombardment of the cathode is stopped. Therefore sputtering is carried out with a high-frequency AC voltage. In the case of a negative voltage halfwave, ordinary cathode sputtering takes place on the dielectric cathode; in the case of a positive voltage halfwave the positive charge accumulated on the cathode is neutralized by the electrons extracted from the plasma.

Reactive Cathode Sputtering. In contrast to ordinary physical cathode sputtering, reactive cathode sputtering is realized in a glow discharge of a mixture of inert and active gases. The particles of the sputtered cathode interact chemically with the active gas or they form solid solutions with it, and the new material reaches the substrates. In order that the process of formation of the substance of the applied film not take place on the cathode, which greatly complicates burning of the discharge, mixtures of argon with active gas content of no more than 10% are used.

In order to obtain oxide films, the sputtering takes place in an argon-oxygen plasma, nitrides in an argon-nitrogen plasma, carbides in an argon-carbon monoxide plasma or argon and methane. On introducing different active gases into the chamber, it is possible to obtain films of various compounds which in practice cannot be obtained by thermovacuum deposition. For example, ferrite magnetic films are obtained on sputtering of a nickel alloy with iron in a glow discharge plasma of argon and oxygen.

Reactive cathode sputtering makes it possible to obtain not only films of various compositions, but also it makes it possible to control the properties of the films, for example, the specific resistance of resistive films. Thus, when sputtering a tantalum cathode in an argon-nitrogen plasma, thin Ta_2N films are obtained, the resistance of which can be varied by the computation of the nitrogen introduced into the chamber. Reactive sputtering is widely used to obtain high-resistance resistors.

The basic technical difficulty of reactive cathode sputtering is exact matching of the active gas introduced into the chamber.

Ion Plasma Sputtering. Ion plasma sputtering is realized at low temperatures in a vacuum chamber $1.2 \cdot (10^{-2} \text{ to } 10^{-3})$ Pa. In order to maintain efficient concentration of the argon ions at lower pressure, the method of ionizing electrons is increased. For this purpose, arc discharge is used. Arc discharge is independent. To maintain arc discharge a thermoemission cathode is needed. In contrast to glow discharge, arc discharge burns at lower voltage between the anode and the cathode (100-300 volts). The discharge current is several amperes.

FOR OFFICIAL USE ONLY

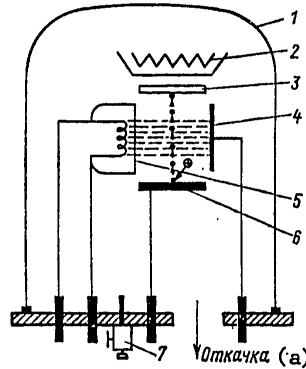


Figure 8-3. Diagram of the process of ion-plasma sputtering.
 1 -- hood of the vacuum chamber; 2 -- heater; 3 -- substrate;
 4 -- anode; 5 -- cathode; 6 -- target; 7 -- needle leak

Key:

a. Evacuation

A three-electrode vacuum system (Figure 8-3) is used to implement this method. An additional electrode-target is the source of particles of the material to be used for the deposited film. The substrates are located opposite the targets on carousels or a drum.

The chamber is evacuated to maximum vacuum, the cathode heating current is switched on, after heating of the cathode a voltage is applied between the cathode and anode, and an inert ionized gas is admitted to the chamber. Arc discharge is struck between the anode and the cathode. On feeding a small negative potential to the target or substrate it is possible to obtain ion cleaning. For sputtering, a negative potential of 200-1000 volts with respect to the anode is fed to the target. The electric field of the target extracts positive ions from the arc discharge plasma and accelerates them to energies of hundreds of electron volts. Bombarding the target surface, the ions sputter it. The target is usually placed at a distance of 2/3 of the anode-cathode distance from the cathode where the region of the passive discharge column is located. As a result of the thermoemission cathode it is possible independently to control the concentration and energy of the ions in the plasma, for the gas discharge circuits (anode-cathode) and sputtering circuit (target-substrate) are electrically decoupled from each other. It is possible to vary the ion concentration by varying the electron emission current or the accelerating voltage between the cathode and anode, and it is possible to vary the ion energy by variation of the target potential.

Ion-plasma sputtering by comparison with cathode sputtering is realized at higher vacuum; therefore the free path length and energy of the sputtered atoms are greater. Accordingly, the contamination of the films by molecules of residual and inert gases is less, the deposition rate of the films is higher, and it reaches several hundreds and even thousands of angstroms per minute. The film density is greater, the adhesion of the films to the surface of the substrate is better.

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

The deposition process is more easily controlled, the film thickness is regulated by the target potential and time. It is possible to carry out a large number of film deposition processes without changing the targets, which insures reproducibility of the properties of the deposited films. The process of film deposition with the given properties corresponds to the requirements of automation and can be used in the continuous cycle of creating the microcircuits.

The deficiencies of the ion-plasma sputtering are as follows: limited possibilities of reactive sputtering as a result of short service life of the thermoemission cathode in the presence of active gases, the cathode is an additional source of contamination; comparative complexity of the device and operation of the equipment.

Ion-plasma sputtering is widely used to make film passive elements, to obtain masking films on semiconductor plates, for deposition of semiconductor layers and magnetic films.

8-3. Thermal Oxidation

Principles of the Method. The process of oxidation, intensified by heating to high temperatures, is called thermal oxidation. The method of thermal oxidation is basic in planar technology for obtaining masking films on silicon and films of the gate oxide for MDS-structures.

Silicon has high affinity for oxygen. On a carefully cleaned silicon surface at room temperature a film of silicon dioxide 10-15 Å thick is formed instantaneously. This film grows at a rate of about 11-12 Å/day to a thickness of about 50-100 Å. Therefore thermal oxidation in any case takes place in the presence of a thin oxide film on the surface.

The process of obtaining a thermal oxide can be broken down into three steps: delivery of the oxidizing agent to the substrates and adsorption of them by the surface, diffusion of the oxidizing agent through the silicon dioxide film to the silicon surface, chemical interaction of the oxidizing agent with the silicon with the formation of oxide. Purified dry or wet oxygen is used as the oxidizing agent.

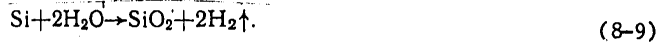
The oxidation rate is determined by the slow phase of diffusion penetration of the oxidizing agent through the growing film to the SiO₂-Si interface itself. The diffusion coefficients depend strongly on temperature. At low temperatures the diffusion coefficients and, consequently, the growth rate of the film, are small. It is possible to increase the growth rate either by increasing the pressure in the reaction chamber or increasing the process temperature. The creation of devices for oxidation at increased pressures is very complicated, and it is economically inexpedient. In practice the oxidation of the silicon takes place at low pressures, but at high temperatures (850-1350°C). The diffusion coefficient of water in silicon dioxide at the same temperature is appreciably greater than the diffusion coefficient of oxygen. This explains the high growth rates of the oxide in wet oxygen. However, the growth of films only in water vapor is not used as a result of poor quality of the oxide. Higher quality films are obtained in dry oxygen, but the growth rate of the films is too small.

FOR OFFICIAL USE ONLY

When using dry oxygen the following reaction takes place on the surface of the silicon



When using wet oxygen an additional reaction takes place



The hydrogen formed during the reaction quite rapidly diffuses from the Si-SiO₂ interface to the oxide surface.

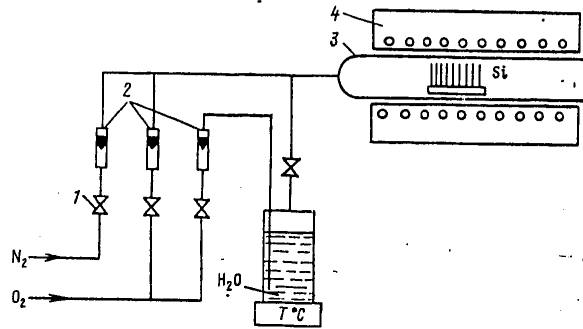


Figure 8-4. Diagram of the process of thermal oxidation of silicon. 1 -- valves; 2 -- rotameters; 3 -- quartz tube-reactor; 4 -- heating furnace

Silicon atoms are consumed to form the film; therefore the initial substrate surface goes into the body of the oxide during the oxidation process. This explains the absence of the problem of insuring high adhesion.

At temperatures of more than 1000°C the silicon oxidation is subject to a parabolic law:

$$x^2 = kt, \tag{8-10}$$

where x is the weight or the thickness of the oxide film; t is the oxidation time; k is the growth rate constant which depends on the type of oxidizing agent, its vapor pressure and the presence of impurities in the silicon plate and in the growing oxide film.

Process Equipment. The layout of the thermal oxidation device is presented in Figure 8-4. The reaction chamber of this device is made from a quartz tube which passes through the muffle of the heating furnace. The boats of fused quartz or high-resistance silicon filled with silicon plates prepared for oxidation are loaded in the reaction chamber heated to the required temperature. Dry or wet oxygen is passed through the chamber. In order to provide for feeding a strictly

FOR OFFICIAL USE ONLY

defined amount of water vapor to the chamber, evaporation takes place under strictly defined conditions. For this purpose a defined amount of water (batch) is poured into the sealed vessel -- batcher. The oxygen coming into the batcher goes through the water, capturing the water vapor. This batcher (bubbling type) is heated, as a rule, to a temperature of no more than 60°C, which provides for growth of higher quality films.

The control of the film growth process is realized by monitoring the substrate temperature, the oxygen flow rate, the water temperature and time.

The advantages of thermal oxidation of silicon are its technological nature, high quality of films obtained. The deficiencies are high temperature of the substrates during the film growth process which can lead to worsening of the properties of the structures previously obtained in the semiconducting plate, redistribution of the impurities and a change in properties of the silicon and also low growth rates.

8-4. Film Deposition from the Vapor-Gas Phase

Principles of the Method. Many processes of film deposition from the vapor-gas phase are relatively new and owe their appearance to the intense development of applied chemistry. The vapor of the initial compounds (halides, hydrides, carbonyls, organoelemental compounds) is delivered to the zone of the device where the substrates heated to the required temperatures are located. As a result of the chemical reaction, the material required for constructing the film is isolated in the solid phase, and the gaseous by-products are removed from the zone where the substrates are located. The entire process can thus be divided into the following stages: conversion of the initial compounds to the vapor state, transport of the vapor of the initial compounds to the hot substrates, chemical reaction, the formation of nucleating centers and film growth, removal of the gaseous reaction products and the nonreacting molecules of initial compounds from the reaction zone.

Depending on the chemical composition and the structure of the initial compound, the chemical reaction can take place both directly on the substrate and near it. In the former case first adsorption of the molecules of the initial compound by the surface of the substrate takes place, then chemical reaction and, finally, desorption of the by-products. In the second case the vapor of the initial compounds absorbs heat emitted by the substrate; chemical reaction takes place in the vapor gas phase. The atoms or clusters of the atoms formed near the surface diffuse to the substrates and are adsorbed by them.

The basis for the method is the various reactions of synthesis, substitution, hydrolysis and pyrolysis (thermal decomposition).

The processes of film deposition can be realized in a vacuum, in an inert or active medium. The application of electric, light, electron beam and other activations communicates additional energy to the reacting molecules, which permits acceleration of the chemical reaction with a decrease in the process temperature. For example, for activation of the processes of pyrolysis of organoelemental compounds, high-frequency plasma, electron and laser beams, ultraviolet illumination, the addition of oxygen or ozone to the reaction chamber, and so on are used.

FOR OFFICIAL USE ONLY

Process Equipment. Just as in thermal oxidation of silicon, the process takes place by the open tube method. A flow of gas containing a defined concentration of vapor of the initial compounds is passed through the reaction chamber of the device. In order to insure the supply of a strictly defined amount of the vapor the evaporation of the compounds also takes place under strictly established conditions. Along with the bubble type batchers, evaporation type batchers are used in cases where the initial products are easily evaporated. The film growth process is regulated by the percentage content of vapor of the initial compound in the reaction chamber, the substrate temperature, and the film deposition time. The percentage content of the vapor depends on the flow rate of the gas carrier through the batcher at defined temperature and the dosage of the initial compound. The regulation of the degree of saturation of the gas carrier by vapor of the initial compound is realized by variation of the gas flow velocity. The temperature and dosage are kept constant.

The advantages of the method of deposition from the vapor gas phase are as follows: the possibility of obtaining a large number of films of different composition and with different properties which can be alloyed during the growth process; the films can be deposited in practice on any substrates and also on complex relief surfaces; the substrate temperatures are comparatively low during the process of obtaining the films. The method is compatible with the operation of preliminary cleaning of the substrate surfaces by gas pickling. Obtaining thick films if necessary presents no difficulties. The application of an inert or active gas environment greatly simplifies the device, and in case of performing the processes in a vacuum usually a low degree of vacuum is required. The method of providing for obtaining films of satisfactory quality at comparatively high growth rates and activation of the process present no great difficulties.

The deficiencies of the method are considered to include the following: the necessity for obtaining especially pure initial compounds, the difficulty of controlling the film growth for processes with complex chemical reactions, the possibility of adsorption of the reaction by-products by the substrate surface (water, carbon, and so on).

Examples of Film Deposition from the Vapor-Gas Phase. Deposition from the vapor-gas phase is most widely used at the present time to obtain dielectric films and films of refractory metals.

The deposition of masking films of silicon dioxide differs from thermal oxidation by lower temperatures of the substrates during the growth process and the possibility of obtaining films not only on silicon, but also on any other substrates. By comparison with the vacuum methods, deposition from the vapor-gas phase is distinguished by simplicity of the processes and simplicity of the equipment. In the case of thermal vacuum evaporation of quartz very high temperatures are required and, consequently, special electron beam evaporators, and, in addition, the deposited films consist of a mixture of silicon dioxide and monoxide and silicon, and they do not have good masking and other electrophysical properties.

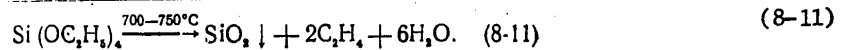
Deposition of the gas phase is at the present time finding broad application in the finishing stages of silicon planar technology when the application of thermal oxidation can lead to variations of the parameters of the already created regions of the semiconductor structures or in planar technology in semiconductors (germanium,

FOR OFFICIAL USE ONLY

gallium arsenide, and so on), the thermal oxides of which do not have masking properties.

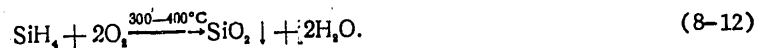
The method of pyrolytic deposition is most widespread after thermal oxidation. For pyrolysis, organosilicon compounds are used, from which most frequently tetraethoxysilane is used [the ester of ethyl orthosilicic acid $\text{Si}(\text{OC}_2\text{H}_5)_4$]. The gas carrier (argon, nitrogen or active oxygen) is saturated by tetraethoxysilane of vapor by bubbling and goes into the reaction zone of the quartz tube where the holder is located with the semiconductor plates.

The following decomposition reaction takes place in the substrate zone



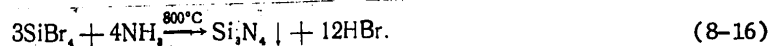
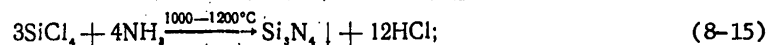
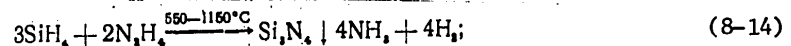
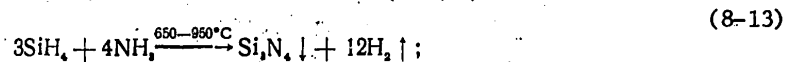
Silicon dioxide is deposited on plates; the remaining reaction products are carried away by the gas flow from the tube. The films obtained by pyrolysis are somewhat inferior to the thermally grown films with respect to their properties. The best result is provided by pyrolysis in a vacuum. Sometimes the films are compacted, for example, by annealing or introduction of phosphorus anhydride into them.

The oxidation of a gaseous monosilane compound proceeds at lower temperatures of 300-400°C:



Films obtained by the silane method are cleaner, for the organic radical and carbon are not formed as a result of chemical reactions.

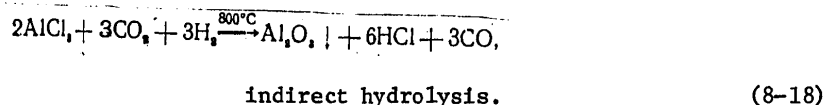
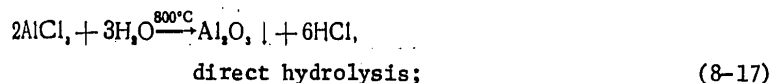
The deposition of silicon nitride Si_3N_4 films from the gas phase is realized as a result of chemical reactions of the interaction between the hydrides or halides of silicon and ammonia NH_3 or hydrazine N_2H_4 :



Silicon nitride films are the best studied after silicon dioxide films; they have the best dielectric and masking properties and can be used successfully in planar technology.

FOR OFFICIAL USE ONLY

The deposition of aluminum oxide films by hydrolysis of halide compounds of aluminum takes place at temperatures of 800°C and it is accompanied by chemical reactions:



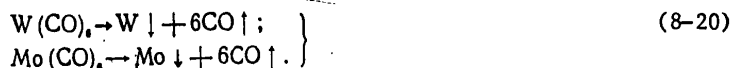
In order to obtain Al_2O_3 films at temperatures of 500°C most frequently the trimethylaluminum oxidation reaction is used:



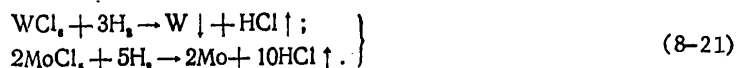
The Al_2O_3 films are used as the dielectric layer of thin-film capacitors, for the active elements of MDS-microcircuits with high radiation resistance as the gate insulator; they are also prospective for radiation-resistant bipolar structures.

The deposition of refractory metal films from the vapor gas phase is carried out most frequently by decomposition of carbonyls (the carbonyl method) or reduction of chlorides (the chloride method).

The decomposition of carbonyl takes place in a flow of hydrogen, nitrogen or argon at atmospheric pressure. The substrate temperature is about 250-320°C. The film deposition process is accompanied by the reactions:



Chloride reduction is carried out at substrate temperatures of 720-750°C:



Thin tungsten and molybdenum films are used to create rectifying Schottky barriers, for metallization, for the gates of MDS-structures and for contact masks.

The presented examples do not exhaust all of the possibilities of deposition from the vapor gas phase. Further development of applied chemistry, and in particular, the chemistry of organometal compounds (OMC) performed in the Soviet Union by Academician G. A. Razuvayem and his students, will permit films to be obtained for thin-film elements more simply than by the traditional vacuum methods. At the present time the OMC are already being used to obtain chromium, nichrome, copper, nickel and other films.

FOR OFFICIAL USE ONLY

8-5. Anodic Electrolytic Oxidation

Electrolytic oxidation is used to obtain films of semiconductor and metal oxides. Oxidation is realized on the anode of the electrolytic bath.

The mechanism of electrolytic anodic oxidation differs from the mechanism of thermal oxidation. For electrolytic oxidation the reaction of the interaction of the material with the ions of the oxidizing agent proceeds at the film-electrolyte interface.

Thus, for example, in tantalum the growth of the oxide film is accompanied by the reaction



As the oxide grows, the ions of the oxidizing material are delivered from the substrate-film ($\text{Ta-Ta}_2\text{O}_5$) interfaced to the surface under the effect of an electric field arising in the oxide on application of voltages between the electrodes of the bath. The greater part of the applied voltage goes to the oxide film; therefore high intensities of the electric field arise in it, leading to drift of the tantalum ions. Electrolytic oxidation can be carried out at room temperatures. The electrodes must be located strictly parallel to each other in order to insure uniform current density and, consequently, uniform film thickness over the surface of the substrate.

Anodic electrolytic oxidation can be carried out with a constant value of the current, with constant voltage and in the combined mode.

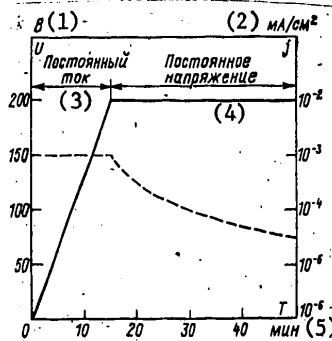


Figure 8-5. Conditions of combined electrolytic oxidation of tantalum.

Key:

1. Volts
2. milliamps/cm²
3. Direct current
4. DC voltage
5. minutes

FOR OFFICIAL USE ONLY

Oxidation with constant current can be carried out if the voltage between the electrodes is increased gradually. If the voltage cannot be increased, the current will diminish as the film grows, and the growth rate will decrease. Increasing the project can lead to breakdown of the film. Reducing the temperature of the electrolyte makes it possible to increase the breakdown voltage and increase the thickness of the grown film.

Oxidation with constant voltage is accompanied by a gradual decrease in the electric field intensity in the growing film. Correspondingly, the delivery of the oxidized ions to the surface of the film decreases, the ion current decreases, and the film thickness reaches its maximum value. Thus, the applied voltage limits the thickness of the grown film.

Combined electrolytic oxidation is considered to be optimal, for it permits us to obtain thicker films.

In Figure 8-5 conditions are presented for two stages of tantalum oxidation when obtaining the dielectric of the film capacitor. First oxidation is carried out with constant current density to achieve the maximum admissible voltage (the forming voltage). In this phase the oxide increases with constant rate in accordance with the increase in voltage required to maintain the ion drift through the film. In the second phase (at constant voltage) the controllability of the process improves, for the growth rate decreases quickly.

The choice of the electrolyte is an important event, for its composition and concentration have a strong influence on the film properties.

It is desirable to carry out anodic oxidation in nonaqueous alkalis, for the hydrogen ions are capable of destroying the films obtained. For example, the silicon and aluminum oxide films go into solution in accordance with the reactions:



Tantalum forms uniform oxide films in aqueous solutions of almost any electrolyte with the exception of hydrofluoric acid, which dissolves the oxide.

Anodic oxidation of silicon is realized in nonaqueous solutions of KNO_3 , NaNO_2 , NH_4NO_3 , HNO_3 , $\text{H}_4\text{P}_2\text{O}_7$ or NaOH in ethylene glycol, methylacetamide or tetrahydrofurfuryl alcohol. The method does not permit us to obtain silicon dioxide films more than 0.1-0.2 microns thick.

Anodic electrolytic oxidation of aluminum can be carried out in buffered solutions of salt of tartaric acid, citric acid, phosphoric acid and carbonic acid.

Thus, the thickness and properties of the oxide film obtained by electrolytic oxidation depend on the composition, concentration and temperature of the electrolyte and also the voltage on the electrodes, the current of film formation and the process time.

FOR OFFICIAL USE ONLY

The method of anodic electrolytic oxidation is distinguished by simplicity of the process, simplicity of equipment and good controllability of the growth rate. The method is widely used to obtain dielectric layers of Ta_2O_5 , Al_2O_3 , TiO_2 of thin-film capacitors.

8-6. Deposition of Metals from the Electrolytes and Solutions

Electrochemical (Galvanic) Deposition of Film. The basis for the electrolytic method of the deposition of metal films is the phenomenon of electrolysis in which cathode reduction of the cations takes place. The phenomenon of reduction of metal cations was investigated for the first time by Russian physicist and electrical engineer B. S. Yakobi, the works of which provided the basis for electroplating (the deposition of thin metal films) and galvanoplastics (the manufacture of embossed metal copies from subjects).

The electrolytic application of films is widely used in IC technology to create rectifying Schottky barriers, ohmic contacts with the active regions of the semiconductor structures, current carrying tracks, for the protection of metal parts and IC cases from corrosion, the manufacture of the free and contact masks, and so on. The method permits thick films to be obtained easily.

For film deposition the prepared substrates which play the role of the cathode are placed in an electrolytic bath. The salt of the deposited metal is introduced into the electrolyte. The anode is usually made of deposited metal. The process of film deposition is accompanied by anode solution, which prevents impoverishment of the electrolyte with respect to the metal ions. The surface area of the anode is approximately 1.5 times the surface area on which the film is deposited (cathode).

The relation between the mass of metal deposited on the cathode and the process parameters is expressed by the first and second laws of electrolysis:

1. The mass of deposited metal is proportional to the amount of electricity going through the solution.
2. The mass of material deposited on passing an identical amount of electricity through the solution is proportional to the electrochemical equivalent E .

Beginning with these laws, it is possible to write the expression for the film deposition rate:

$$V = jE\alpha/\rho, \quad (8-24)$$

where j is the current density; ρ is the density of the film material; α is the coefficient of effectiveness of the current to the ratio of the effective amount of deposited metal to the theoretical amount. Usually α has values from 0.5 to 1.

It is possible to control the film deposition by variation of the current density. With an increase in current density the deposition rate of the film increases, the solution near the cathode is quickly impoverished with respect to the metal ions. In this case the probability of cathode reduction of the hydrogen ions increases. The bubbles of hydrogen released on the cathode lower the quality of the film; it becomes spongelike and loose. In order to prevent this, the electrolyte is

FOR OFFICIAL USE ONLY

heated and mixed. The films are usually deposited with low current density. For example, gold films from alkaline electrolytes are deposited with current density of 0.002-0.003 amp/cm² and at a temperature of 60-70°C. The gold deposition rate is 1.5-3 microns/hr.

It is possible to carry out electrochemical deposition on DC and AC current. The higher quality films are obtained on AC current. This is explained by the fact that in the case of inverse polarity of the voltage when the cathode becomes the anode, the deposited film is dissolved, and first of all on the projections of the surface of the deposited film, for the current density is larger here. As a result, films of uniform thickness are obtained with a smooth surface. The deposition rate of the films on alternating current is approximately 3 times greater than on direct current. In order to improve the productivity and the quality of the films, ultrasound is used.

For preparation of electrolytes, especially pure (OSCh) and chemically pure (KhCh) reagents are used. In the finished electrolyte there should be no harmful impurities, for they have a negative effect on the adhesion of the film, they lead to peeling, blistering and nonuniformity of the thickness. Only deionized water is used for the electrolyte, for distilled water contains ions of various metals. Impurities can come out of the container for the electrolytes, from the baths and from the anodes. Electrolytes are prepared in ceramic, glass, vinylplastic, fluoroplastic and quartz containers. Anodes are made from especially pure metals, and they are carefully cleaned before use. The electrolyte must be stable, it must not release solid precipitates or gases.

The method of electrolytic deposition of films does not require complex equipment or prolonged preparation and for performing the process. However, it must be noted that out of 70 metal elements only 33 provide coatings, and of these 33 elements only 14 are used for industrial purposes: Al, Ag, Au, Cd, Co, Cr, Cu, Fe, Ni, Pb, Pt, Rh, Sn, Zn. On deposition of chromium, the current effectiveness coefficient usually is only 0.13.

Chemical Deposition of Films from Solutions. This method is distinguished by still greater simplicity of equipment and performance of process. It is sufficient to have an exhaust hood, electric stove, and a chemically and thermally stable container.

The chemical deposition of the metal films is based on catalytic reduction of the metal from solution without the application of an electric current. Most frequently chemical nickel and gold plating are used. For example, the deposition of nickel on n-type silicon is carried out from phosphorus-containing solution, which permits us to obtain contact resistance which is appreciably less than in the case of thermovacuum deposition of nickel.

Test Questions and Assignments

1. Into what steps can the process of thermovacuum deposition be divided? What is the mechanism of each step of thermovacuum deposition?
2. What determines the requirements on the degree of vacuum in the thermovacuum deposition process?

FOR OFFICIAL USE ONLY

3. Use Figure 8-1 to explain the procedure for carrying out the thermovacuum deposition process.
4. Which parameters of thermovacuum deposition are basic? From what arguments are they selected?
5. What are the purpose, basic advantages and disadvantages of thermovacuum deposition?
6. What has provided the incentive for the introduction of sputtering methods into production?
7. Compare the mechanism and the peculiarities of the process of ion sputtering with thermovacuum deposition.
8. Use Figure 8-2 to explain the process of cathode sputtering.
9. What distinguishes the reactive cathode sputtering from physical cathode sputtering?
10. Use Figure 8-3 to explain the process of ion-plasma sputtering.
11. Is it possible to deposit films by the method of ion-plasma sputtering in a plasma of active and inert gases?
12. Compare the processes of ion-plasma and cathode sputtering.
13. The methods of thermovacuum deposition, cathode (physical and reactive) and ion-plasma sputtering are used to obtain what films?
14. Compare the efficiency of the processes and quality of the films during thermovacuum deposition and ion sputtering.
15. What is the growth mechanism of films during thermal oxidation? How is it related to the process conditions? To adhesion of the films?
16. Compare the duration of the process and quality of the films during oxidation in dry and wet oxygen.
17. Explain the procedure for carrying out the thermal oxidation process (Figure 8-4).
18. What parameters are used to control the thermal oxidation process?
19. What are the basic advantages and disadvantages of thermal oxidation?
20. What are the basic steps of film deposition from the vapor gas phase?
21. What is the technique used for the processes of film deposition from the vapor gas phase?

FOR OFFICIAL USE ONLY

22. Explain the advantages of film deposition from the vapor gas phase in examples of using various chemical reactions.
23. What is the mechanism and the characteristic features of anodic, electrochemical oxidation? Compare with thermal oxidation.
24. How is the film growth process controlled during anodic electrolytic oxidation?
25. What is the application of the method of anodic electrolytic oxidation?
26. How is the process and the film quality controlled during electrolytic deposition?
27. What is the application of the method of electrolytic deposition of films?
28. Compile a classification of methods of obtaining thin films by various attributes.
29. Which of the investigated methods of obtaining thin films can be used when making thin film IC? Hybrid IC? Melted active components? Monolithic semiconductor IC? Compatible IC?
30. Which of the studied methods is used to obtain masking films in planar technology?

FOR OFFICIAL USE ONLY

CHAPTER 9. EPITAXIAL GROWTH OF SEMICONDUCTOR LAYERS

9-1. Fundamentals of the Epitaxy Processes

Epitaxy Processes. At the present time semiconductor elements and components of microcircuits are obtained using three basic methods: epitaxy, thermal diffusion and ionic alloying.

The term "epitaxy" is formed from the Greek words «επι» meaning on or over and «ταξις» meaning to arrange in order to designate the process of oriented growth of a layer of new material on the surface of a crystal substrate. Epitaxy is the process of growing layers with ordered crystal lattice by realization of the orienting effect of the substrate. The layers of new material grown with orientation and regularly continuing the crystal lattice of the substrate are called epitaxial layers.

The basic feature of epitaxy by comparison with diffusion and ion alloying is that the layers and local regions of opposite type conductivity or with different impurity concentration from the semiconductor plate are new formations of the initial surface. During the growth process the epitaxial layers are alloyed, that is, donor or acceptor alloys are introduced into them.

During diffusion and ionic alloying the rectifying or concentration junctions are obtained in the initial plate, changing its properties by alloying to the required depth. In contrast to epitaxy diffusion or ion alloying proceeds in the presence of substrate impurities. Therefore obtaining a region of opposite type of conductance is possible only with recompensation of the initial impurity.

The processes of epitaxial growth are analogous to obtaining thin films, and from the point of view of production technology could be investigated in Chapter 8. However, as a result of their exceptional role for semiconductor structures, we shall study these processes separately.

Epitaxial layers can be grown in a vacuum from the vapor-gas and liquid phases. From the point of view of the ratio of the materials of the semiconductor plates or substrates and layer and also the properties of the structures obtained epitaxy is divided into three types.

FOR OFFICIAL USE ONLY

Autoepitaxy is the process of oriented growth of crystalline material differing very insignificantly with respect to chemical composition from the substrate material. This offers the possibility of obtaining electron-hole junctions which are uniform with respect to composition (homogeneous); therefore the layer growth process is sometimes called homoepitaxy.

Heteroepitaxy is the process of oriented growth of material differing with respect to chemical composition from the substrate material. On growth of heteroepitaxial layers at the layer-substrate interface an electric junction is formed which is nonuniform with respect to composition (heterogeneous). The process of heteroepitaxy is possible for elements and complex materials not forming chemical compounds among each other.

Chemoepitaxy is the process of approximate growth, as a result of which the formation of a new phase -- the chemoepitaxial layer -- takes place with chemical interaction of the substrate material with the material of the growing layer. The layer material differs with respect to chemical composition both from the substrate material and from the material reaching the surface of the substrate. As a rule, the thickness of the chemoepitaxial layers is insignificant.

In IC production basically the first two types of epitaxy are used.

Epitaxial Structures. Semiconductor silicon billets -- structures of four types -- single layer, multilayer, with latent layers and heteroepitaxial layers -- are being made for the production of semiconductor devices and microcircuits. By analogy with the formation of types of silicon in types of epitaxial structures the following notation is used: K -- silicon; D, E -- hole and electron type of conductivity, respectively; B, S, M, F -- alloying elements, boron, antimony, arsenic and phosphorus.

The single-layer epitaxial structures are made from silicon monocrystalline plates with n-type conductivity during the process of growing the p-type autoepitaxial layer. The type of single-layer epitaxial structures is written in the form of a fraction with a numerical coefficient. For example,

$$80 \frac{8KDB-0.5}{200KES-0.01}$$

The numerical coefficient in front of the fraction designates the diameter of the epitaxial structure (80 mm), the first number in the numerator corresponds to the thickness of the epitaxial layer (8 microns), and in the denominator, the thickness of the plate (200 microns). The second numbers in the numerator and the denominator correspond to the specific resistance of the epitaxial layer (0.5 ohm-cm) and substrate (0.01 ohm-cm).

The multilayer epitaxial structures are made during the process of autoepitaxial growth of layers of different types of conductivity from two sides of the p-type silicon monocrystalline plate. The marking is analogous to the marking of single-layer structures, but it contains three levels: for the upper layer, the plate and lower layer, respectively.

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

The epitaxial structures with latent layers are made on silicon plates with p-type conductivity with local sections of low-resistance n⁺-silicon by growing the autoepitaxial layer of n-type conductivity alloyed with phosphorus. The notation from the epitaxial structures is included in the notation procedures: in the numerator is the thickness of the epitaxial layer in microns, the type of conductivity and the alloying elements of the epitaxial layer, its specific resistance in ohm-cm, thickness in microns, type of conductivity and alloying elements of the latent layer, its surface resistance in ohm/□ ; in the denominator is the total thickness of the epitaxial structure, type of conductivity and alloying element of the plate and its specific resistance.

An example of the notation for structures with latent layers alloyed with arsenic is as follows:

$$\frac{10KEFO.2/5 KEM12}{330/KDB10} .$$

Example of the notation for structures with latent layers alloyed with antimony:

$$\frac{10KEFO.5/4KES40}{330/KDB10} .$$

Heteroepitaxial structures are made by heteroepitaxial growth of the monocrystalline silicon layer of n or p type conductivity on a monocrystalline substrate.

An example of the marking of heteroepitaxial structures is as follows:

$$\frac{10KDB0.5}{60S250} ,$$

where the parameters of the layer are presented in the numerator: the thickness of 10 microns, material silicon, p-type, alloyed with boron, specific resistance of the layer 0.5 ohm-cm; in the denominator the first number designates the structure diameter 60 mm, the letter S designates the substrate material sapphire, and the second number, the substrate thickness 250 microns.

Mechanism of Layer Formation. Just as for the general case of film deposition, the process of epitaxy can be divided into the following steps: delivery of the atoms or molecules of the layer material to the surface of the plate or substrate and migration of them over the surface, the beginning of grouping of the particles of the layer material near the surface centers of crystallization and formation of the nucleating centers of the layer, the growth of individual nucleating centers to coalescence of them and the formation of a solid layer.

The material from which the layer is made can be delivered to the plate or the substrate directly from the source without going through chemical changes or as a result of chemical reactions taking place either near the surface or on the surface itself. The delivered atom moves over the surface of the plate or substrate until it attaches to it or again returns to the environment. If the atom on moving over the surface encounters a center of crystallization -- an energy-wise advantageous location corresponding to the absolute minimum free energy of the atom-substrate system -- it begins growth of a stable epitaxial nucleating

FOR OFFICIAL USE ONLY

center. If the attachment takes place at a location not corresponding to the absolute minimum free energy, the atom begins the growth of an unstable disoriented nucleating center. The basic initial crystallization centers are considered to be the monoatomic and multiatomic steps and the breaks in them always available on the surface of a single crystal and also the screw dislocations. When encountering a crystallization center the atom turns out to be bound with a minimum of two atoms.

When new lots of the material reach the surface, the atoms begin to combine into two-dimensional or three-dimensional nucleating centers, the size of which gradually increases. On being arranged near the monoatomic step, the atoms complete the monoatomic layer. Further growth of the layers is explained by the formation and growth of nucleating centers near the dislocation steps. As a result of completion, the dislocation step undergoes rotational motion, forming a terrace (Figure 9-1).

Near the screw dislocation several monolayers grow simultaneously. As the nucleating centers grow, they coalesce, and a solid epitaxial layer is formed.

Growth Rate of Epitaxial Layers. In accordance with the presented mechanism the growth rate is basically determined by the rate of delivery of the layout material to the surface of the plate or substrate, the rate of formation and completion of stable nucleating centers.

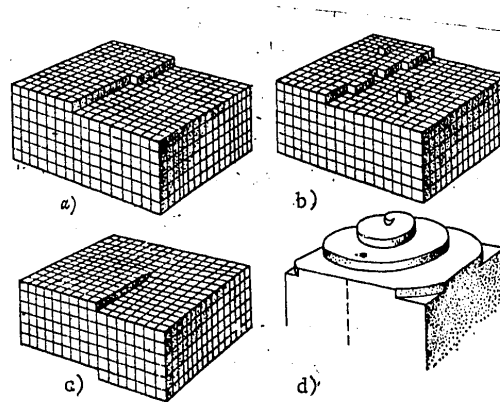


Figure 9-1. Growth of an epitaxial layer near surface defects
 a -- monoatomic step; b -- growth near a monoatomic step;
 c -- screw dislocation; d -- growth near a screw dislocation

The rate of delivery of the layer material is determined by the degree of supersaturation of the layer of the space surrounding the plate or substrate with the material. The rate of formation and growth of the nucleating centers with a defined number of delivered particles depends on the temperature and the number of steps and screw dislocations on the surface of the plate or substrate. With an increase in temperature, the rate of displacement of the atoms over the surface increases; therefore the probability of "favorable encounters" with the initial crystallization centers increases. This can be used to explain the existence of

FOR OFFICIAL USE ONLY

the so-called epitaxial temperature above which oriented multicrystalline epitaxial layers grow, and below which they are disoriented polycrystalline layers.

The number of steps on the surface depends on the crystallographic orientation and increases on making the transition from low-index to high-index planes. On the low-index planes the growth rates of the epitaxial layers increase in accordance with the series (111)<(100)<(110). In practice epitaxial growth tries to be reduced to movement of the atomic steps. The in practice inexhaustible source of atomic steps during growth of silicon, germanium and gallium arsenide in the case of autoepitaxy is the vicinal planes close to (111) where the disorientation angle exceeds 1°. The quality of the epitaxial layers is found to be better, and the growth rates sufficient if the growth takes place on slightly disoriented (1-5°) crystallographic planes.

In real processes the growth mechanism of the epitaxial layers is complicated by the effect of impurities present on the surface, the presence of various defects, and so on. Rotations and displacements of the growth centers, mismatching effect of the substrate or plate, the formation of the transition layer at the interface with the surface, and so on are also possible.

9-2. Methods of Epitaxial Growth from the Vapor-Gas Phase

Chloride Method. The basis for this method is the reversible chemical reaction of reduction of silicon tetrachloride vapor by pure hydrogen:



The diagram of the device for growing the layers is presented in Figure 9-2.

In order to improve the structure of the epitaxial layers before growing them the plates are annealed in a hydrogen atmosphere at a temperature of 1200°C. Then the plates are subjected to gas pickling; for this purpose dehydrated hydrogen chloride is admitted to the chamber. The reaction (9-1) proceeds from right to left.

For the beginning of growth of the epitaxial layers, a mixture of hydrogen with silicon tetrachloride vapor is admitted to the reaction chamber. The silicon tetrachloride is a liquid that boils at 57°C and has high vapor pressure which depends strongly on the temperature. In order to feed an exact quantity of silicon tetrachloride to the reaction chamber, it is evaporated in a batcher with strictly established conditions. The hydrogen participating in the reduction reaction is simultaneously a gas carrier. The degree of saturation of the hydrogen with silicon tetrachloride vapor can be regulated by varying the hydrogen flow rate and the temperature of the liquid silicon tetrachloride. Hydrogen must be well dried to remove the moisture, for the latter decomposes the silicon tetrachloride with the formation of silicic and hydrochloric acids:



FOR OFFICIAL USE ONLY

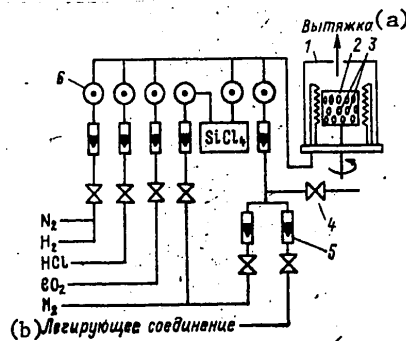


Figure 9-2. Diagram of the epitaxial growth unit.
 1 -- reaction chamber; 2 -- stand; 3 -- substrate; 4 -- gate valve; 5 -- rotameter; 6 -- needle valve

Key:

- a. Exhaust
- b. Alloying compound

The stand for the plates simultaneously plays the role of their heater. The stand is heated by the resistive, induction or radiant infrared method. At high temperatures the stand must not give off contaminants and must not enter into chemical interaction with the plates; therefore it is made of high-quality grades of glass graphite or graphite and it is coated with a layer of silicon carbide or silicon. In the case of induction heating it is possible to use stands made of high-resistance silicon coated with a layer of silicon dioxide. In order that the films be deposited only on the plates and the contaminants not get on the plates, the inside walls of the chamber, the latter are especially cooled.

The reproducibility of the parameters of the epitaxial layers from plate to plate and over the area of each plate is determined by the uniformity of the temperature distribution, the nature of the flow of vapor-gas mixture and uniformity of its distribution over the surface of each silicon plate. At a constant flow rate of the vapor-gas mixture and given method of heating the plate, identicalness of the growth conditions for each plate depends on the internal configuration of the chamber, the structure of the holder and location of the plates. Devices with horizontal and vertical arrangement of the chamber are used in production. For example, in the Soviet UNES-2P-V device there are two vertical chambers. Preference is given to the vertical chambers, for two undesirable factors can be excluded or reduced to a minimum in them: the variation of the concentration of the vapor-gas mixture and the temperature gradient in the position zone. In order to increase the output capacity of the units, the holders are made in the form of a cylinder with special recesses for the multilevel stacking of the plates.

Figure 9-2 shows the diagram of the device in which the vertical chamber is used. The vapor-gas mixture enters the reaction zone through a number of openings in quartz tubes which are located along the generatrices of the stand. In the hot zone at a temperature of 1150-1250°C the tetrachloride reduction reaction takes place (9-1). The silicon is deposited on the plates, hydrogen chloride is removed by a flow of hydrogen through the exit opening of the chamber. The reaction (9-1) can also proceed at significantly lower temperatures, but in this case it is more

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

complicated to obtain the monocrystalline layers. Higher temperatures are undesirable, for the growth rate of the layers in practice does not increase, and the probability of contaminants from the stand and from the walls of the chamber increases.

The dependence of the growth rate of the layer on the tetrachloride concentration in the vapor-gas mixture is presented in Figure 9-3. The increase in growth rate in the first section can be explained basically by an increase in delivery of material to the plate; a further decrease in the growth rate for concentrations of more than 0.1 mole is explained by incomplete reduction of the silicon tetrachloride and beginning pickling of the silicon by the tetrachloride:



Usually the growth of epitaxial silicon layers is realized for silicon tetrachloride concentrations in the hydrogen less than 0.1 mole for which the growth rate depends linearly on the tetrachloride content and, consequently, on the hydrogen flow rate. Under modern conditions the epitaxial growth rate of the silicon layers is 1-1.5 microns/min. The purity of the initial silicon tetrachloride determines the maximum specific resistance of the layers. The modern level of cleaning permits us to obtain silicon tetrachloride which provides specific resistance to hundreds of ohms per centimeter.

The basic deficiency of the chloride method is high temperatures of the process leading to diffusion of the impurities from the plates to the growing layer and also to autoalloying. The autoalloying is the process of penetration of impurities from the plate or the substrate to the growing epitaxial layer by transport from the side surfaces and from the back side of the plate or the substrate through the gas phase. In addition, reversibility of the tetrachloride reduction reaction requires high precision of maintaining the deposition conditions of the layer.

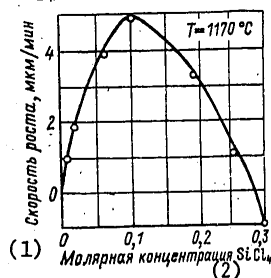


Figure 9-3. Epitaxial growth rate as a function of the silicon tetrachloride concentration

Key:

1. Growth rate, microns/min
2. Molar concentrations of SiCl₄

Hydride (Silane Method). The growth of the layers when using monosilane gas is accompanied by the irreversible reaction of thermal decomposition:



FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

The device for drying epitaxial layers by the silane method is similar with respect to structure to the devices used in the chloride method and, for precautionary measures when working with monosilane, it is equipped with a system for evacuation of the air and traces of moisture. Perfect (close to monocrystalline) layers are obtained at monosilane decomposition temperatures of 1000-1100°C, which is 100-150°C lower than when reducing silicon tetrachloride. This decreases the undesirable diffusion and autoalloying, which makes it possible to manufacture epitaxial structures with sharper boundaries of the junctions. The irreversibility of the monosilane decomposition reaction also promotes the obtaining of sharper concentration distribution of the alloying impurity at the plate-layer interface. Absence of hydrogen chloride or other harmful products among the decomposition products, in particular, chlorine, also is among the advantages of the silane method. The growth rate of the layers is higher than when reducing silicon tetrachloride.

The deficiencies of the method include self-ignitability and explosiveness of monosilane requiring special precautionary measures. Therefore monosilane is used in practice mixed with hydrogen. With a monosilane content of less than 5% the mixture is not spontaneously combustible. The safety rules for working with the mixture are the same as for working with pure hydrogen.

The processes of epitaxy from the vapor gas phase are controlled by smooth variation of the concentration of the reagents in the vapor gas mixture (the batcher temperature, the flow rate of the gas carrier), the temperature of the plates or substrates and time. These parameters can be monitored independently of each other, which facilitates the creation of devices with automated program control and makes the processes reproducible and economical.

The possibilities of the methods of epitaxy from the gas phase have still not been completely exhausted. When using a large variety of reagents it is possible to obtain layers not only of monoatomic, but also double and triple semiconductor compounds.

9-3. Other Methods of Epitaxy

Vacuum Epitaxy. The processes of obtaining epitaxial layers in vacuum devices are called vacuum epitaxy.

Thermal vacuum deposition is carried out in the devices insuring a high degree of vacuum. For example, in the Soviet superhigh vacuum, two-hood "vacuum in vacuum" UVN-84-R-1 device a vacuum of no less than $5 \cdot 10^{-5}$ Pa is maintained during the process of epitaxial growth. The melting point of the silicon is 1414°C, the evaporation temperature is 1342°C, that is, the silicon can be converted to the vapor state by evaporation or sublimation.

For evaporation the source is heated to a temperature several hundreds of degrees higher than the melting point. At such temperatures the vapor pressure is high, and it is possible to obtain high epitaxial growth rates reaching 4 microns/min. On evaporation from crucibles, defects are introduced into the epitaxial layers as a result of contamination of the melt by the crucible material and as a result of high chemical activity of the molten silicon. The pure epitaxial layers are grown by crucibleless methods of evaporation using local heating of a massive

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

sample of silicon by an electron beam (Figure 9-4). The molten drop makes contacts only with the solid silicon. The latter is placed on a cooled holder.

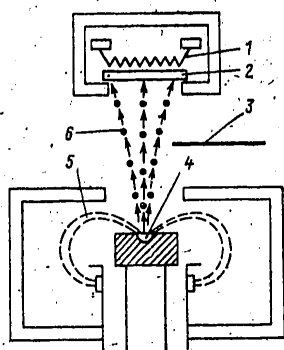


Figure 9-4. Diagram of the epitaxial growth of silicon using electron-beam evaporation.
 1 -- substrate heater; 2 -- substrate; 3 -- slide valve; 4 -- section of molten source; 5 -- electron beam; 6 -- particles of evaporated silicon

For sublimation the silicon source is heated to temperatures of 1200-1300°C by direct passage of an electric current, the induction method or an electron beam. The lower source temperatures correspond to lower growth rates of the epitaxial layers of 0.3 microns/min.

Vapor condensation is carried out on the plates heated to temperatures of 1000-1100°C. At temperatures of less than 1000°C, polycrystalline silicon layers are grown.

The creation and maintenance of a high vacuum during the layer deposition process, complexity of labor consumption of servicing the vacuum equipment, difficulty of obtaining epitaxial layers with n-type conductance and sufficiently highly alloyed layers with $N_n > 10^{17} \text{ cm}^{-3}$, difficulty of obtaining defect-free layers -- all of these prevent broad introduction of the method into production.

The method of thermal vacuum deposition is, nevertheless, of interest in connection with a number of its advantages by comparison with other methods: relatively low temperature of the plates in the process of growing the layers, the possibility of obtaining thin layers less than 1 micron thick, the possibility of performing a continuous cycle of operation in one unit, and the absence of explosive gases and toxic mixtures.

Ion sputtering is carried out without crucibles, the source of the layer material is the silicon cathode. The plate temperature is comparatively low (750-900°C). The basic deficiencies of the method are as follows: too low growth rates of no more than 400 Å/min and also capture of an inert gas in the growing layer. This is the reason for rare application of the method for epitaxial growth of silicon.

FOR OFFICIAL USE ONLY

Epitaxy from the Liquid Phase. The liquid phase epitaxy is primarily used to grow layers of double and triple semiconductor compounds widely used for optoelectronic devices. Growth takes place most frequently from solutions in molten metals.

The essence of the method consists in crystallization of the material dissolved in the molten metal on the surface of the plate. During the growth process of the layer the plate surface plays the role of inoculation. The metal solvents must have low melting point, low vapor pressure at operating temperatures and dissolve the semiconductor well from which the layer is constructed. For example, epitaxial layers of GaAs, GaP, $Ga_{1-x}Al_xAs$ are grown from solutions in gallium; CdSnP₂ layers are grown from solutions in tin; layers of $Ga_{1-x}In_xAs$ are grown from solutions in indium. The supersaturation of the solution with the material of the grown layer required for crystallization usually is created by gradual decrease in temperature.

Growth is carried out by the flow tube method if solutions are used which contain nonvolatile alloying additives (silicon, germanium, tin). The volatile alloying additives -- zinc, sulfur and tellurium -- lead to nonequilibrium growth conditions. In the cases of their use, epitaxy is carried out by the closed tube method. For this purpose the holder with the prepared plates and charge for obtaining the solution is placed in a quartz ampule. The ampule is evacuated to the maximum vacuum, it is filled with the formir gas, sealed and put in the temperature zone of the furnace to carry out the process.

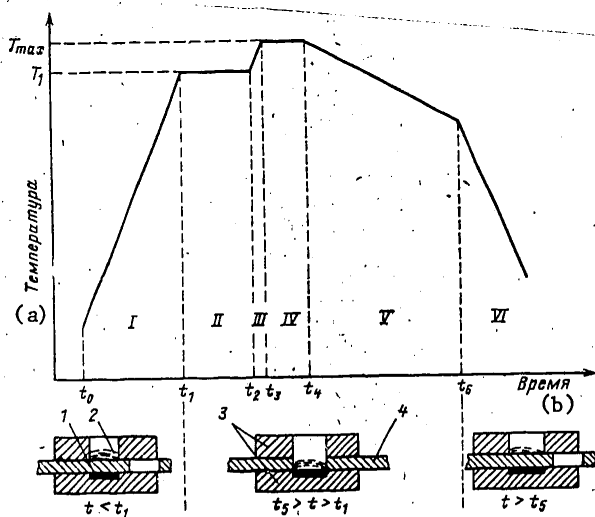


Figure 9-5. Temperature-time conditions of liquid phase epitaxy. 1 -- substrate; 2 -- solution-melt; 3 -- stationary part of the holder; 4 -- moving part of the holder

- Key:
 a. temperature
 b. Time

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

The characteristic temperature-time conditions of growth from a melt-solution is presented in Figure 9-5. When the holder is heated to a temperature of T_1 and held at this temperature for a time t_1-t_0 , a uniform equilibrium solution is formed; simultaneously the surfaces of the plates and the solution are cleaned to remove the oxides by reduction in a flow of hydrogen. At the time t_1 the solution is discharged. The solution wets the surface of the plates in a time $t_2-t_1=10$ to 30 minutes and with an increase in temperature by $5-10^\circ\text{C}$, in t_3-t_2 it dissolves the plates, removing contamination and defects from the surface. After a short holding time of t_4-t_3 at maximum temperature slow precision cooling is started. The solution goes from saturated to supersaturated state, which leads to crystallization of the excess semiconductor from the solution on the surface of the plates, that is, from the time t_4 growth of the epitaxial layer begins. After growth of the given layer of material at the time t_5 , by moving the slide valve of the holder, the surface of the plates is freed of the melt, and quickly cooled to room temperature. The excess solution is removed from the surface of the plates by boiling in acid. Thus, for example, gallium arsenide is removed by hydrochloric acid.

Epitaxy from cooled melt-solutions insures high growth rates of the layers, it permits growth of thick (more than 10 microns) layers and sharp p-n junctions. The deficiencies of this method are variation in composition of the solution during the growth process leading to variation of the composition of the growing layer and also complexity of process control when growing thin layers. For growth of sufficiently thin layers supersaturation is created by introducing a monocrystalline seed crystal of semiconductor into the solution before bringing the plates into contact with it.

The deficiencies of liquid-phase epitaxy include higher growth temperatures than in gas epitaxy, the necessity for removal of the excess solution from the plate surfaces after the end of the process, and the difficulty of creating highly efficient, noncontaminating holders.

9-4. Heteroepitaxy

Heteroepitaxy greatly expands the functional capabilities of the microelectronic devices. The oriented growth of a material on a different material with respect to composition can be realized in various combinations, from which the following are the most interesting for microelectronics: semiconductor-semiconductor, semiconductor-dielectric, semiconductor-metal.

When selecting the pair of substrate-layer materials it is necessary to consider the parameters of their crystal lattices, the thermal coefficient of linear expansion, melting points, mutual diffusion coefficients, vapor pressure at the process temperature, and so on. In the case of high lack of correspondence of the indicated parameters at the layer-substrate interface, mechanical stresses and dislocations occur. The transition reagent is distinguished by its properties from the properties of the containers on both sides of the substrate-layer interface, which influences the growth process of the layer and the parameters of the microcircuits.

In microcircuit production heteroepitaxy develops in two basic directions: growth of semiconductor heterojunctions and growth of semiconductor layers or semiconductor heterojunctions on dielectric substrates.

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

Growth of Heterojunctions. A heterojunction is a transition region occurring at the interface of two semiconductors of different chemical composition. In this case the semiconductors can be distinguished by width of the forbidden zone, crystal lattice and (necessarily), type of conductivity.

The heterojunctions are used to create rectifiers that operate on the basic carriers (CdSe-Se), tunnel diodes (GaAs-InP, CdS-ZnTe, GaAs-Ga_{1-x}In_xAs), transistors with wide-zone emitters (GaAs-Ge), injection lasers (GaAs-Ga_{1-x}Al_xAs), sources of spontaneous radiation (GaAs-GaP, CdSe-ZnTe, GaAs-InAs), and the corresponding microcircuits.

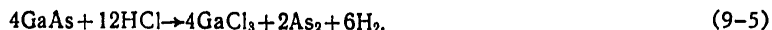
In order to obtain the heterojunctions, ordinary epitaxy methods are used: from the gas phase, vacuum epitaxy, epitaxy from a liquid phase. Let us briefly consider the epitaxy of the gallium arsenide layers on germanium plates.

The Ge-GaAs heterojunctions have a high degree of perfection as a result of the close parameters of the crystal lattices and the close values of the thermal coefficient of linear expansion of these semiconductors; they have been well studied and obtained primarily by growing gallium arsenide from the gas phase.

The structurally perfect epitaxial layers of gallium arsenide with a high degree of purity can be obtained as follows: transport of the GaAs by gaseous hydrochloric acid (the GaAs-HCl system), synthesis from Ga and As₃H with use of hydrochloric acid (the Ga-AsH₃-HCl system) as the gas carrier, synthesis from Ga and As using HCl gas carrier (the Ga-As-HCl system), synthesis in the Ga-AsCl₃-H₂ system.

The method of transporting GaAs by gaseous hydrogen chloride is based on using gas transport reversible reactions. The process is carried out by the flow tube method with the application of a two-zone furnace. Before deposition of the films the plates are annealed in a hydrogen current and they are pickled with hydrochloric acid. The source -- gallium arsenide -- is located in the current of hydrogen fluoride in the first zone with temperature of 850°C, and the plates, in the second zone with temperature of 700°C.

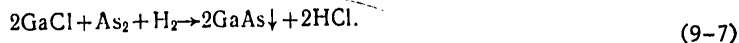
The following reaction proceeds in the source zone



As a result of the excess hydrogen, the gallium trichloride is reduced to the monochloride:



The products of the reactions (9-5) and (9-6) are transported to the plate zone where gallium arsenide is formed, which is also deposited in the form of an epitaxial layer:



After completion of the process the plates are cooled in a flow of hydrogen.

FOR OFFICIAL USE ONLY

The basic advantage of the method is its simplicity, but the necessity for high purity in the GaAs-source limits the application of the method.

In the last methods, the possibility appears for independent control of the Ga and the As concentrations in the gas phase and, consequently, the possibility of obtaining the GaAs layer with optimal ratio of the components. However, the use of Ga-As-HCl and Ga-AsH₃-HCl systems is held up by the absence of highly pure As, HCl, the complexity of obtaining highly pure arsine AsH₃ and its toxicity.

The Ga-AsCl₃-H₂ system has significant advantages over those investigated above, for it has the necessity for preliminary synthesis of the GaAs, and it is not necessary to obtain highly pure As and HCl. Arsenic trichloride AsCl₃ can be cleaned by rectification to a high degree of purity. Arsenic trichloride is poured into a bubbling type batcher, from which it is delivered by a flow of hydrogen. The container made of inert material (aluminum nitride or boron nitride) is filled with gallium and is placed in the first zone with a temperature of 850-900°C. The plates are located in the second zone with a temperature of 700-720°C.

Along with the layers made of double semiconductors at the present time a large number of heterojunctions of ternary semiconductor compounds are obtained. Thus, for example, the GaAs_{1-x}P_x layers on the germanium plates have improved structure and have interesting properties. By comparison with germanium, the gallium arsenide plates have less perfect structure, a smaller area, they are shorter and more expensive.

Heteroepitaxy on Dielectric Substrates. The manufacture of microcircuits on heteroepitaxial structures of dielectric and semiconductor solves the problem of insulating the elements, it insures improvement of the parameters, it improves the degree of integration and the percentage yield of usable products. At the present time heteroepitaxy of silicon on sapphire and spinel substrates has been the best studied.

Silicon is grown on sapphire by the methods of gas epitaxy and also the method of thermovacuum deposition. In contrast to cubic symmetry of silicon, sapphire has rhombohedral symmetry. As a result of the difference in structure of the silicon and sapphire crystals, the silicon layers have greater dislocation density and lower values of the mobility of the current carriers. Silicon with (111) orientation is grown on the (0001) and (1010) planes of sapphire. Silicon with (100) orientation is grown on sapphire with (1102) orientation. The requirements on the quality of polishing the sapphire substrates are extremely high.

Among the methods of gas epitaxy silane is most frequently used, for it permits growth of silicon with lower substrate temperature; in addition, there is no danger of pickling the sapphire inasmuch as hydrogen chloride is absent among the reaction products. When growing heteroepitaxial structures of "silicon on sapphire" (KNS) it is necessary to consider the possible autoalloying of the layers by aluminum from the substrates. Instead of pure silicon layers it is possible to obtain contaminated silicon layers with p-type conductivity. Increasing the thickness of the silicon layers leads to improvement of their electrical properties as a result of partial compensation of some of the growth defects.

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

Silicon on spinel by comparison with silicon on sapphire has additional advantages for IC. Spinel has cubic lattice which is closer to the structure of silicon. Spinel is less hard than sapphire, that is, it is easily processed. In addition, autoalloying during growth of the silicon layers on spinel is significantly less than during growth on sapphire. At the present time the manufacturing technology of spinel has been insufficiently developed, and spinel is still expensive.

9-5. Local Epitaxy

The localization of the processes of epitaxial growth can be insured by two procedures: local stimulation or local complication.

Local stimulation of chemical reactions and also the formation and growth of nucleating centers can be carried out, for example, using an electron or laser beam. The application of this method of localization is limited by the complexity of the formation and control of the beam displacement (especially electron) in vapor-gas media. In addition, nonuniformity of the current density with respect to the cross section of the electron beam leads to nonuniformity of a locally growing epitaxial layer.

Local complication of the epitaxial growth is carried out using contact masks. In this case it is necessary to create conditions under which the growth of the layer takes place only on the open sections of the substrate (selective epitaxial growth) or use contact masks which are then removed together with the deposition of a layer of semiconductor on their surface.

At the present time the selective local epitaxial growth using SiO₂-contact masks is most widely used. This is explained by the compatibility of the SiO₂-masks with epitaxial-planar technology and energy difference of the formation of the silicon nucleating centers on the Si-plates and on the surface of the SiO₂ promoting growth of the epitaxial layer in the first stage on the silicon. The selective growth in the chloride method is insured by quite high concentrations of silicon tetrachloride in the gas phase and the selection of the temperature range in which the formation of polycrystalline silicon on the mask does not take place. In addition, for complication of the deposition of the polycrystalline silicon on the surface of the SiO₂ masks, the chloride process is carried out in the presence of hydrogen chloride vapor. After bringing the temperature of the substrate to the operating temperatures, hydrogen chloride is admitted to the chamber of the device; on completion of the growth process, first the SiCl₄ feed is halted, then the HCl. Local growth is very sensitive to the composition and ratio of the H₂, HCl and SiCl₄ in the vapor gas phase and also the location of the substrate in the reaction chamber of the devices. It is also necessary to consider reduced chemical strength of silicon dioxide in the presence of hydrogen at increased temperatures. At the present time high temperatures are needed to obtain a smooth mirror surface of the local silicon sections.

9-6. Alloying of Epitaxial Layers

The alloying of layers in the case of epitaxy from the gas phase is realized using special compounds -- sources of alloying impurity. The source vapors are transported by the gas carrier to the temperature zone of the device where chemical

FOR OFFICIAL USE ONLY

reaction takes place with release of elementary alloying impurity which is deposited on the plates or substrates together with the material of the growing epitaxial layer. The following forms of sources of alloying impurity are distinguished.

Liquid sources (the halides PCl_3 , POCl_3 , BBr_3 and others) are poured into individual or common batcher with the source material of the layer. The halides of the alloying elements dissolve well in silicon tetrachloride; they are easily evaporated and saturate the passing flow of hydrogen. At the hot surfaces the vapors of the admixture sources are reduced by hydrogen:



Halides have high vapor pressure which depends sharply on temperature; therefore, insignificant temperature fluctuations can cause significant variations in concentration of the admixture in the growing layer. Recently when alloying with boron, high-boiling compounds have been used, the vapor pressure of which depends only slightly on the temperature. By using these compounds it is possible to obtain silicon layers with a broader range of specific resistances from 0.005 to 5 ohm-cm.

The gaseous sources -- hydrides, diborane B_2H_6 , arsine AsH_3 , phosphine PH_3 -- are distinguished by high toxicity, and therefore they are delivered to the production facility in tanks in a mixture with hydrogen or inert gases in concentrations of ($5 \cdot 10^{-4}$ to 1)%. The degree of alloying when using liquid and gaseous sources is regulated by the concentration of the existing compound with alloying gas flow and speed of the basic gas carrier flow.

When using diborane it is difficult to obtain a mixture of it with a content of less than 10%; the mixtures are unstable during storage; it is difficult to establish and exactly measure the small flow rates of the mixtures; therefore it is difficult to obtain silicon layers slightly alloyed with boron with specific resistance greater than 1 ohm-cm. Accordingly, high-boiling, liquid boron-containing compounds or boron tribromide is used.

Solid sources are used in the relatively new method of gas discharge spark alloying.

The gas discharge chamber using two connecting pipes is connected to the gas distribution system of the epitaxial growth device. Pure hydrogen or hydrogen saturated with silicon tetrachloride vapor is passed through the chamber. The electrodes are made of material containing the alloying element, LaP_6 , B_4C , AlB_2 or the alloy $\text{Sb}+(1-2)\%\text{As}$. When feeding a pulsed voltage between the electrodes, spark discharge is excited in the chamber. In the spark discharge chamber the electrode material is partially evaporated. The vapor is carried away by the passing hydrogen to the reaction zone of the device. There it decomposes with the release of a free alloying admixture. The concentration of the admixture can be adjusted by the distance between the electrodes, the voltage fed to the electrodes, the frequency and duration of the pulses and also the flow rate of the hydrogen through the gas discharge chamber.

Alloying during vacuum epitaxy is realized by an admixture which enters into the composition of the semiconductor -- the source of particles of the growing layer of material. The admixture atoms are delivered to the plates or substrates as a

FOR OFFICIAL USE ONLY

result of evaporation of the source or sputtering of the target. The admixture must be volatile (arsenic, antimony, phosphorus). The concentration of the alloying admixture in the growing layer can be varied within the broad range, varying the rate of formation of the vapor or the sputtering rate. Using somewhat differently alloyed sources of admixture, it is possible to grow multilayered epitaxial structures.

The alloying in liquid phase epitaxy is realized by an elemental admixture. The admixture is introduced into the melt-solution, from which it is encompassed by the crystallized material of the growing layer. For example, for alloying layers of gallium arsenide, silicon, germanium, tin, sulfur, zinc and tellurium are used. Silicon and germanium are amphotermal admixtures. Depending on the alloying conditions, the temperature of the beginning of crystallization, and the concentration of the alloying admixture in the melt, the silicon and germanium can replace either the gallium nodes and play the role of donors or the arsenic node and play the role of acceptors.

The possibilities of alloying during epitaxial growth are much broader than for other methods of obtaining alloyed layers of semiconductor, for example, during diffusion. A unique characteristic of epitaxy is the possibility of obtaining high-resistance layers of semiconductor on low-resistance plates. During epitaxy it is possible to obtain various distributions of the alloying admixtures, including uniform distribution or with a sharp concentration gradient at very short distance. It is possible to obtain multilayer structures in one growth cycle.

9-7. Defects in the Epitaxial Layers

In the epitaxial layers most frequently dislocations, regions of mechanical stresses, packing defects and growth defects occur. The defects limit the application of epitaxial structures for powerful and high-voltage semiconductor devices in which large areas of p-n junctions and thick epitaxial layers are used and specially for LSI and SLSI where rigid requirements are imposed on the defect density.

The dislocations in the epitaxial layers can occur at the substrate (plate)-layer boundary in the presence of mechanical disturbances or contamination of the surface as a result of crystallographic noncorrespondence of the materials, as a result of differences in the degree of alloying of the layer and the plate, as a result of point defects, and so on. The dislocations available on the surface are traced by the growing layer. The presence of various contaminants and noncorrespondence of the materials leads to the appearance of mechanical stresses. The plastic deformation of the material at high growth temperatures can remove the mechanical stresses with the formation of additional dislocations.

In order to decrease the density of the dislocations in the layers the following are required: selection of the substrate or plates with minimum density of the dislocations, careful preparation of the surface, selection of compatible materials of the layer and plate or substrate. In order to decrease the dislocations caused by various degrees of alloying, it is not recommended that plates with a specific resistance of less than 0.01 ohm-cm be used during the growth of high-resistance layers, for at the boundary with the layer too large mechanical stresses can arise, and, in addition, it is difficult to manufacture low-resistance, semiconductor ingots without dislocations.

FOR OFFICIAL USE ONLY

The sections of mechanical stresses around the admixtures are formed during alloying of epitaxial layers by admixtures with atom size differing from the dimensions of the semiconductor atoms. The stress depends on the ratio of the dimensions of the atoms and the concentrations of the impurities. The mechanical stresses lead to the appearance of dislocations and to bending of the structures. In order to decrease the bending when growing epitaxial layers the method of mixed alloying by admixtures causing deformation of different sign is used. It is possible to realize compensation of the deformations of silicon structures by simultaneous alloying by tin and phosphorus, tin and boron, antimony and phosphorus. The vapor of the impurities is selected so that the atomic radius of one will be less than the atomic radius of silicon, and the other, greater. For example, the atomic radius of tin is 1.40 Å, phosphorus 1.10 Å, and silicon 1.17 Å. The application of the method of shifted alloying permits growth of thick epitaxial layers with a minimum number of noncorrespondence defects.

The packing defects are regions with disturbance of the alternation of the atomic layers, on the side boundaries of which there is mismatch with the remaining part of the crystal lattice of the layer. On fitting of the layer during the growth process, for example, to an oxide islet, the atoms can begin growth of the mismatched layer of atoms, and as a result ordinary alternation of the atom layers can be disturbed. For example, in a perfect silicon crystal in the (111) direction the atoms are arranged in layers in a defined sequence ABCABCABC... The layers designated by different letters (A, B, C) are distinguished by mutual arrangement of the centers of the atoms. If during the growth process one of the layers is skipped, a subtraction packing defect arises, and if any layer is repeated twice, an interstitial packing defect arises. At the location of the formation the packing defect is of a point nature; as growth takes place the defect begins to occupy the entire region. Thus, points defects growing on the (111) plane are transformed as the layer grows to regions included inside a right tetrahedron, with crystal lattice not coinciding with the remaining part of the epitaxial layer. Basically the packing defects are initiated at the point-layer interface, and as growth takes place, they penetrate the entire epitaxial layer, emerging as the base of a tetrahedron, that is, equilateral triangle, on the surface (Figure 9-6). The height of the tetrahedron equal to the thickness of the epitaxial layer $h = a\sqrt{3}/3 = 0.816a$, where a is the side of the triangle. If the packing defect is initiated not on the plate-layer interface, the height of the tetrahedron will be less than the layer thickness. If the packing defects are initiated close to each other, then as growth takes place, superposition of defects can occur. If three defect layers are initiated in a row (microtwin), then twinning regions appear in the growth process.

The growth defects appear as follows: at locations on the surface of the plate or growing layer of larger foreign particles, for example, abrasive particles; in regions of accelerated growth or encounter of rapidly growing nucleating centers; in regions with high concentration of impurities when it exceeds the solubility limit; in the case of high supersaturation when the material is crystallized by conglomerates, that is, are clusters of atoms or molecules. During the growth process these defects are converted to holes, hills, pyramids and polycrystalline sections. A hole is a formation occurring as a result of failure of the epitaxial layer to grow over a section of the plate surface. At these locations the formation of a polycrystalline deposit is the most probable. A hill is a completely or partially disoriented monocrystalline formation in the form of a protrusion. The

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

growth pyramid occurs as a result of oriented growth on dislocations or other defects. Hills and pyramids protruding on both the surface of the layer are especially dangerous, for they have a negative effect on the lattice parameters and lead to rejects in subsequent technological operations.

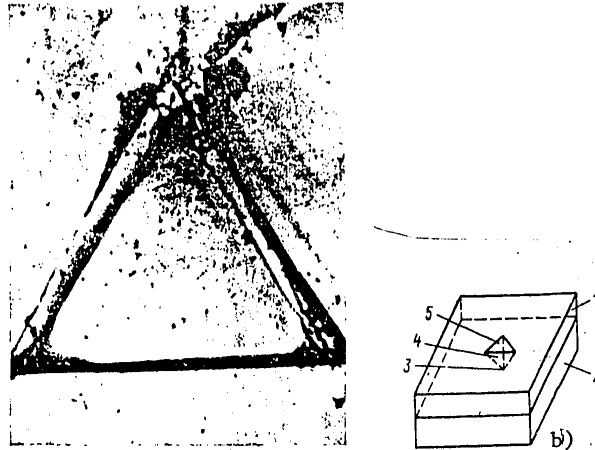


Figure 9-6. Packing defect of a silicon epitaxial layer.
 a -- microphotography of the defect on the (111) surface of silicon;
 b -- three-dimensional diagram of the defect; 1 -- epitaxial layer;
 2 -- (111); Si-plates; 3 -- initiation point of the defect on the
 plate surface; 4 -- defect region; 5 -- emergence of the defect at
 the surface of the epitaxial layer

Thus, all of the defects can be divided into those that take off from the substrate crystal and those occurring in the epitaxial layers in the growth process. The largest number of defects is connected with the quality of preparation of the surface of the substrates or plates, with alloying, with purity of the performance of the process and with crystallization conditions.

9-8. Epitaxial Layer Control

During the research and development of the technological processes of epitaxy, a large number of parameters are controlled, and a large number of various methods are used. During production most frequently the thickness, specific resistance of the epitaxial layer, the concentration distribution of the admixture with respect to layer thickness, and defect density are controlled. These layer parameters determine the breakdown voltages and back currents of the p-n junctions, the saturation resistances of the transistors, the loss resistances and the volt-farad characteristics of the structures.

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

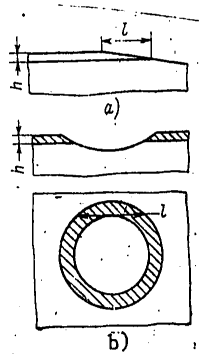


Figure 9-7. Oblique (a) and ball (b) microsections for determining the thickness of an epitaxial layer

The thickness of the layer very frequently is determined by the methods of coloring oblique or ball microsections. For this purpose the control sample is cut at a small angle of 1-5° or a small hole is cut out on its surface using a ball with a diameter $D=25$ to 150 mm. The coloring of the microsections in special solutions takes place as a result of the difference in electrode potentials of the p and n regions which gives rise to selective composition of copper from the solution on the p-region or selective oxidation of the n-region. In order to discover the interface of the concentration homojunctions $n-n^+$, $p-p^+$, selective pickling is used. It is possible by observation under a microscope to use the colored microsections to measure not the true thickness of the epitaxial layer h , but an essentially large value of l by the eyepiece scale (Figure 9-7). The thickness of the epitaxial layer is calculated by the formulas: $h=l \cdot \text{tg } \alpha$ for the oblique microsection, $h=l^2/4D$, for the ball microsection. The precision of the measurements is on the average $\pm 5-10\%$. The method of coloring the microsections is applicable for determining a broad range of thicknesses of the layers differing from the substrate by type of conductivity and degree of alloying.

The contactless, nondestructive method of infrared interferometry used for layers, the optical constants of which differ sharply from the optical constants of the substrates insures measurement precision of $\pm 5\%$.

The infrared beams are partially reflected from the air-layer interface, they pass through the epitaxial layer and are reflected from the layer-substrate interface. As a result of the application of two reflected beams, the interference pattern of alternation of dark (minima) and light (maxima) strips is observed. The thickness of the epitaxial layer is defined by the formula

$$h = \frac{\Delta N}{2n\Delta k} \quad (9-9)$$

where ΔN is the number of peaks or minima in the measured interval; n is the coefficient of refraction of the layer; $\Delta k = 1/\lambda_1 - 1/\lambda_2$ is the difference in values inverse to the wave lengths of the observed maxima or minima.

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

The method is distinguished by rapidity of measurement, but is applicable only for epitaxial layers from 2 to several tens of microns thick grown on highly alloyed substrates.

The specific resistance is measured by the sonde and volt-farad methods. For epitaxial layers, the ordinary four-sonde method is used on high-resistance substrates of opposite type conductivity. The specific resistance is calculated by the formula

$$\rho = 4,53 \frac{U}{I} h, \quad (9-10)$$

where U is the voltage drop between adjacent sondes, volts; I is the current passing through the epitaxial layer between the edge sondes, amps; h is the thickness of the layer, cm.

The specific resistance of the epitaxial layers on the high resistance substrates ($\rho_e/\rho_{Si} \geq 50$) of the same type of conductivity is measured by the free-sonde method (Figure 9-8). The specific resistance is determined from the equation

$$U_{mp(1)} = 68,39 \rho^{0,5360} + 21 \rho^{0,8772}, \quad (9-11)$$

Key: 1. breakdown

where $U_{breakdown}$ is the breakdown voltage between the sondes 2 and 3.

The breakdown voltage is measured using an oscillograph which has a large input impedance.

The volt-farad method of measuring the specific resistance is based on using the relation that relates the capacitance of the p-n junction to the voltage at the junction and the concentration of the admixture.

The concentration distribution of the admixture with respect to thickness of the epitaxial layer is determined by the method of differential electrical conductivity, that is, measurement of the electrical conductivity after successive removal of an insignificant part of the thickness of the epitaxial layer. On the surface of the epitaxial layer, the electrical conductivity σ_1 is measured. Then part of the epitaxial layer with a thickness Δ is removed, and the surface electrical conductivity σ_2 is again measured. The mean concentration of the admixture in the removed part of the layer is calculated:

$$N = \frac{1}{q\mu\Delta} (\sigma_1 - \sigma_2), \quad (9-12)$$

where q is the electron charge; μ is the mobility of the current carriers.

Then part of the epitaxial layer is again removed, and σ_3 is measured, and so on. The thick epitaxial layers can be gradually ground off. The removal of thin epitaxial layers is by cathode pickling or anode oxidation in an electrolyte with

FOR OFFICIAL USE ONLY

subsequent acid pickling of silicon dioxide. The precision of these methods of removing the silicon is 20-50 and 150-300 angstroms, respectively.

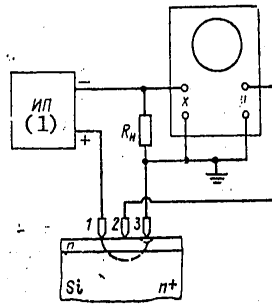


Figure 9-8. Diagram of measuring the specific resistance of an epitaxial layer by the three-sonde method

Key:

1. IP

The structural defects of the epitaxial layers are investigated using the transparent x-ray topography (the Long method), electron microscopy and visual observation under a microscope of the surface of the layer after discovery of the defects by selective pickling. The last method is the simplest and takes the least time. After pickling at the points of emergence of the packing defects at the surface, depending on the crystallographic orientation of the substrate, various pickling configurations are formed: triangles, tetrahedrons, trapezoids, identical or V-type lines and also combinations of pickling figures if the defects are superposed on each other. At the points where the dislocations emerge on the surface, dislocation pickling holes appear. The number of holes or pickling figures and also their size can be determined using the MIM-7 (200 power) and MMU-1 (190 power) metallographic microscopes. The compositions of the pickling agents for discovery of defects are selected as a function of the material of the epitaxial layer, the crystallographic orientation of the substrates, and so on. The defect density is determined by the number of defects per unit area of layer and, in accordance with the technical specifications, can vary within the limits from 0 to $5 \cdot 10^4 \text{ cm}^{-2}$.

Test Questions and Assignments

1. What is epitaxy and what is its purpose in IC production? Explain autoepitaxy, heteroepitaxy and chemoepitaxy.
2. Which epitaxial structures are used in IC production?
3. How are the types of epitaxial structures deciphered? Check yourself on the marking examples presented in §9-1.
4. What is the mechanism of epitaxial growth? What factors influence the growth rate?

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

5. Why is epitaxial growth carried out on slightly disoriented silicon plates?
6. What is the essence of the chloride method of growing silicon layers?
7. Using Figure 9-2 explain the order of the epitaxy process.
8. At what temperatures are monocrystalline layers grown? Polycrystalline layers? Why? At what temperatures do the layers not grow? What are the temperatures called at which the layers are not formed?
9. What chemical reaction provides the basis for the silane method of growing silicon layers? Compare it with reaction 9-1.
10. What distinguishes the techniques of epitaxial growth of silicon by the silane method from growth by the chloride method?
11. Compare the chloride and silane methods of growing epitaxial silicon?
12. What are the characteristic features, advantages and disadvantages of thermovacuum deposition of silicon?
13. What is the basic application of liquid phase epitaxy (ZhFF)?
14. What is the essence of the process of growing epitaxial layers from solutions in molten metals?
15. What is the technique for growing gallium arsenide layers from the liquid phase using Si, Ge, Sn-alloying admixtures? When using Zn, S, Te-alloying admixtures? Explain the structure of the holder and the temperature-time chart of liquid-phase epitaxy by Figure 9-5.
16. What are the advantages and disadvantages of liquid phase epitaxy?
17. What is the role of heteroepitaxy in the production of microelectronic devices? In what two basic directions is modern heteroepitaxy developing?
18. Give examples of heteroepitaxial semiconductor junctions and explain the essence and technique for obtaining them.
19. What are the peculiarities of silicon on sapphire epitaxy? What method is most frequently used to grow silicon on sapphire?
20. Explain why the p-Si-layers are obtained when growing on sapphire without alloying.
21. What are the advantages of epitaxy when growing silicon on spinel?
22. What is local stimulation and local complication of the processes of epitaxial growth? What method of localization is used in practice? Why?
23. How is alloying of the epitaxial layers realized?

FOR OFFICIAL USE ONLY

24. How is the degree of alloying of epitaxial layers controlled?
25. What distributions of the admixture in the epitaxial layers can be obtained in practice?
26. What is the unique feature of epitaxy by comparison with diffusion and ion alloying?
27. List the possible structural defects of epitaxial layers.
28. What are the causes of defectiveness of epitaxial layers and how can they be decreased?
29. What are the parameters of the epitaxial layers and why are they controlled during the production process?
30. How is the thickness of the epitaxial layers determined?
31. How is the specific resistance of the epitaxial layers determined?
32. How are the concentration distributions of the alloying impurities in epitaxial layers determined?
33. What defects of the epitaxial layers are controlled visually?

FOR OFFICIAL USE ONLY

CHAPTER 10. HIGH-TEMPERATURE DIFFUSION

10-1. Fundamentals of the Method of High-Temperature Diffusion

Application of the Method. Diffusion as an independent technological process is the most widely used method at the present time for introducing alloying impurities into semiconductor plates or into the epitaxial layers grown on them in order to obtain regions of opposite conductivity by comparison with the initial semiconductor or with lower electrical resistance. In the first case, for example, emitters, bases and insulating regions of the transistor are obtained; in the second case, n^+ , p^+ -latent regions decreasing the resistance of the collector body or the contact regions decreasing the injection of the minority carriers from the ohmic contacts and improving their quality.

When manufacturing high-speed structures with good pulse properties diffusion is used to introduce the admixtures forming deep levels in the forbidden zone of the semiconductor and decreasing the lifetime of the minority carriers of the current. For silicon gold and nickel are such admixtures.

Motive Force of Diffusion. Various nonceasing mechanical movements of the atoms are hidden behind the apparent external calm of solid states: vibrations, rotations, random translational movements. The translational movements are random as long as the solid state is an equilibrium corresponding to minimum free energy. On disturbance of the equilibrium, that is, on the appearance of any nonuniformities the translational movements of the atoms become directional.

Diffusion is the phenomenon of directional displacement of the particles of matter in the direction of their decreasing concentration. Diffusion is aimed at having the solid state avoid nonuniformity, giving it the possibility of releasing excess free energy and conversion to the equilibrium state. The diffusion takes place while various sections of the solid state have different concentration of particles. The motive force of diffusion is the concentration gradient of the atoms or molecules of the material. The greater the concentration gradient, the more intense the diffusion.

When manufacturing diffusion structures, increased concentrations of the alloying admixture are created on the surface of the semiconductor plates. The admixture begins to diffuse deeply.

FOR OFFICIAL USE ONLY

Diffusion Mechanism. In real semiconductors diffusion can be realized by three methods (Figure 10-1). In the volumetric method a simple exchange of places by two atoms or circular exchange with the participation of several atoms takes place. In the vacancy mechanism the diffusion is realized by successive jumps of the admixture atoms of the substitution from the nodes themselves to the vacancies (free nodes). The diffusion with internodal mechanism is realized as a result of successive transitions of the interstitial admixture from one interstitial node to another. The atoms of the admixture somehow are "forced" between the atoms at the nodes of the crystal lattice. The concepts of vacancy and internodal mechanisms of diffusion were introduced into physics by the Soviet scientist Ya. I. Frenkel'.

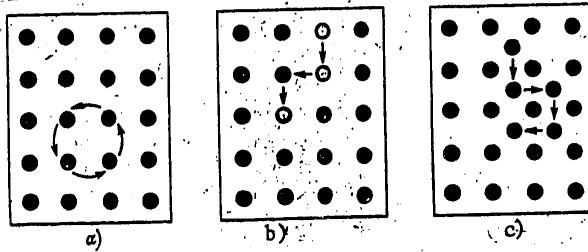


Figure 10-1. Diffusion mechanism in semiconductors.
a -- circular; b -- vacancy; c -- internodal

The diffusion of the basic alloying admixtures in germanium and silicon can be realized by all three mechanisms, but the most probable for them is the vacancy method determined by the presence in the crystal of Schottky and Frenkel' defects, that is, the presence of vacancies or paired vacancy-atom defects at the internode.

The formation of such defects is connected with the thermal vibrations of the lattice. With an increase in temperature, the number of vacancies increases in accordance with the expression

$$(1) \quad n_{\text{vak}} = n \cdot e^{-E/(kT)},$$

Key: 1. vacancy

where n_{vacancy} is the number of vacancies, cm^{-3} ; e is the base of the natural logarithm; E is energy required for the formation of the vacancy; k is the Boltzmann constant; T is the absolute temperature; n is the number of atoms of semiconductor at the nodes of the crystal lattice.

For the formation of one vacancy, an energy equal to several electron volts is required. At ordinary room temperatures the number of vacancies per unit volume of semiconductor is small; for 10^{15} to 10^{16} semiconductor atoms there is one vacancy. With an increase in temperature to $1000-1200^\circ\text{C}$ the number of vacancies becomes comparable to the number of semiconductor atoms. Under the effect of the thermal vibrations the admixture atoms can occupy the location of the next vacancy and thus move. The probability of transition of the admixture from the node occupied to it

FOR OFFICIAL USE ONLY

to the vacancy node depends on the probability of the presence of adjacent vacancies and the probability of overcoming the potential barrier when the atom moves to the vacancy location:

$$p \sim e^{-\Delta E / (kT)} \tag{10-1}$$

Here the value of ΔE is the so-called energy of activation of the diffusion process; it is made up of the energy required for formation of the vacancies and the energy of transition of the admixture atom from its position to the vacancy node. For vacancy diffusion in silicon $\Delta E=3.5$ to 4.3 electron volts.

Along with the process of diffusion of the alloying impurities, there is also diffusion of the atoms of the semiconductor itself -- self-diffusion. However, on the basis of the large chemical bond of the atoms themselves the self-diffusion process is insignificant by comparison with the diffusion of the alloying admixtures.

Many admixture atoms of the first, second, sixth, seventh and eighth groups of the periodic table in silicon occupy locations at the internodes, that is, they form solid interstitial solutions. The diffusion of these admixtures is realized by the internodal mechanism. The probability of internodal transitions of the atoms is appreciably higher than the probability of transitions from node to adjacent node. Therefore the diffusion of the interstitial admixtures takes place appreciably faster than the diffusion of the substitution admixtures.

Solubility of Admixtures. With an increase in the temperature, the solubility of the admixtures in the solid semiconductor increases, and after reaching the maximum solubility, it begins to decrease. The maximum solubility of the impurity in the semiconductor is the maximum possible amount of defined admixture per unit volume of the given semiconductor at a given temperature. The maximum solubilities of the alloying admixtures in the silicon are presented in Table 10-1.

Table 10-1. Maximum Solubility of Admixtures in Silicon

Admixture	Maximum solubility, cm^{-3}	Maximum solubility temperature, $^{\circ}\text{C}$
Aluminum	$10^{19}-10^{20}$	1150
Boron	$5 \cdot 10^{20}$	1200
Phosphorus	$1.3 \cdot 10^{21}$	1150
Gallium	$4 \cdot 10^{19}$	1250
Indium	10^{19}	1300
Antimony	$6 \cdot 10^{19}$	1300
Arsenic	$2 \cdot 10^{21}$	1150
Gold	10^{17}	1300

Using the diffusion method it is possible to introduce an admixture into the semiconductor to concentrations no greater than the maximum solubility for the given temperature or the maximum solubility for the temperature corresponding to it.

First Diffusion Law. The processes of diffusion transfer of the material in the semiconductors are described by the two Fick equations (laws). The thickness of

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

the diffusion layer usually is appreciably less than the plate area. If the diffusion proceeds with respect to the entire surface of the plate it is possible to consider that the diffusion is uniform, for the admixture diffuses primarily in the direction normal to the surface.

The first equation of one-dimensional diffusion determines the flow of impurity atoms from the region with increased concentration to the region with decreased concentration.

$$F = -D \frac{dN}{dx} \quad (10-2)$$

Here D is the diffusion coefficient numerically equal to the number of admixture atoms per unit time passing through a unit area normal to the diffusion direction with a concentration gradient of the admixture equal to one. The minus sign in the righthand side of the equation indicates that the diffusion shift of the atoms takes place from the region with higher concentration to the region with lower concentration.

Second Diffusion Law. The second equation of diffusion is derived from the first under the assumption that the diffusion coefficient does not depend on the concentration:

$$\frac{\partial N}{\partial t} = D \frac{\partial^2 N}{\partial x^2} \quad (10-3)$$

The second Fick law is the basic diffusion law. It defines the concentration of the admixture introduced into the semiconductor at any point in time at any distance from the surface for the given diffusion temperature. The temperature enters into the second equation not explicitly, but through the diffusion coefficient:

$$D = D_0 e^{-\Delta E / (kT)} \quad (10-4)$$

Here D_0 is a constant numerically equal to the diffusion coefficient at infinitely high temperatures; ΔE is the energy of activation of the diffusion process of the given admixture, that is, the energy required for the admixture atom to jump to the vacant node of the lattice. At ordinary room temperatures the diffusion in the solid states is not observed in practice. Diffusion processes in semiconductors take place at high temperatures of 800-900°C for germanium and 1000-1350°C for silicon.

Distribution of the Alloying Admixture with Respect to Depth of the Diffusion Layer. The solution of the basic equation for specific diffusion conditions determines the concentration of the admixtures at different depths for different duration of the process, and the function $N=f(x)$ is thus found for the given diffusion temperature.

The distribution of the impurity for uniform distribution from an infinite and constant source which insures constant replenishment of the admixture going into the semiconductor, is described by the equation of the complex function of the one's complement of the error integral:

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

$$N(x, t) = N_0 \operatorname{erfc} \frac{x}{2\sqrt{Dt}} \quad (10-5)$$

where N_0 is the admixture concentration on the surface of the plate; x is the depth of diffusion; D is the diffusion coefficient; t is the time for conducting the process; erfc is the provisional notation for the function of the one's complement of the error integral.

The concentration distribution graphs of the admixture are presented in Figure 10-2. The amount of admixture going to the surface is equal to the amount of admixture leaving the surface for the body of the plane. In the case of practical calculations the admixture distribution is described by the equation (10-5), the value of $x/2\sqrt{Dt}$ is defined for defined temperatures and diffusion time, and then, using the table, $\operatorname{erfc}(x/2\sqrt{Dt})$ is found.

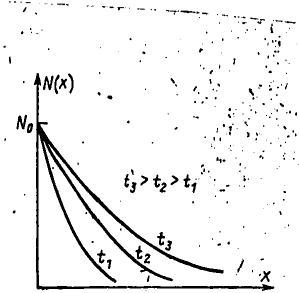


Figure 10-2. Admixture distribution during diffusion from an infinite constant source

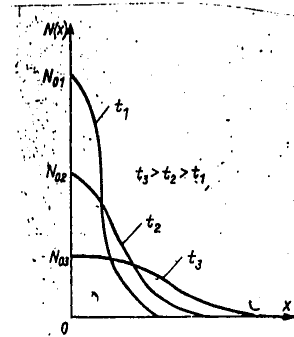


Figure 10-3. Admixture distribution during diffusion from a limited source

The admixture distribution for uniform diffusion from a limited source, for example, created in a thin surface layer of the plate and protected by the masking film from diffusion of the admixture into the surrounding space is described by the normal (gaussian) equation

$$N(x, t) = \frac{Q}{\sqrt{\pi Dt}} e^{-x^2/4(Dt)} \quad (10-6)$$

where Q is the total amount of the admixture in the semiconductor per cm^2 of surface.

The distribution graphs of the admixtures for this case are presented in Figure 10-3. As the admixture goes into the body of the plate the concentration of the admixture in the source is depleted, and influx from the outside is absent.

A characteristic feature of the presented distributions of the admixture corresponding to the complementary error function and the gaussian function is

FOR OFFICIAL USE ONLY

a monotonic decrease in the concentration of the admixture from the surface and to the semiconductor. The maximum concentration of the admixtures corresponds to the plate surface. The surface concentration with diffusion from an unlimited source is constant independently of the diffusion time, and with diffusion from the limited source, it decreases as the time increases.

The practical distributions of the admixtures, as a rule, do not correspond to the simplest theoretical cases. Very frequently the boundary conditions are distinguished or they are not maintained during the diffusion process and, in addition, the influence of various factors not considered by the calculation is felt on the diffusion processes. First of all, this is inconstancy of the diffusion coefficient which depends on the concentration of the introduced impurity, the presence in the plate of initial admixtures and structural defects. The dependence of diffusion coefficient on the concentration in practice is known very approximately, and the solution of the basic equation even in the case of uniform diffusion can be obtained in exceptional cases.

More complex cases of diffusion include multiple successive diffusion of a number of admixtures at various temperatures, for example, when creating a transistorized structure; the diffusion accompanied by evaporation; diffusion into the plate in the presence of a surface oxide; diffusion in the presence of a counterflow of impurities diffusing from the plate volume to the outside, and so on. Finding the solutions in these and other intermediate cases is a complicated problem.

From what has been indicated it is clear that the actual distributions of the admixtures differ from the calculated ones. At the same time when estimating the diffusion process in practice frequently the simplest solutions of the Fick equations are used, and the calculation results are checked and more precisely determined experimentally.

10-2. Characteristic Features of Diffusion in Planar Technology

Basic Equation for Local Diffusion. In the planar structure the dimensions of the openings in the masking film determining the dimensions of the alloyed regions are small and are comparable with the depth of diffusion; therefore in order to find the concentration distribution of the admixtures in the diffusion layer it is impossible to use the solution of the one-dimensional diffusion equation (10-3). The basic equations for local diffusion of the admixtures in the general case is three-dimensional:

$$\frac{\partial N}{\partial t} = D \left(\frac{\partial^2 N}{\partial x^2} + \frac{\partial^2 N}{\partial y^2} + \frac{\partial^2 N}{\partial z^2} \right) \quad (10-7)$$

The solutions of the three-dimensional Fick equation are complex and highly awkward. In practice these complex and awkward solutions can only be used on applying a computer.

Planar p-n Junction Front. In the case of local diffusion, the admixture penetrates into the semiconductor not by a planar front, but a front distorted at the edge of the opening. The distortion of the junction front at the edge of the opening in the masking film has great practical significance for the parameters of the planar structures. First, as a result of the effect of the edge distortion the density of the electric current with respect to the area of the p-n junction

FOR OFFICIAL USE ONLY

is nonuniform, and breakdown is possible for voltages that are less than in the case of the planar junction of the same area. Secondly, the actual area of the junction increases, which leads to an increase in the capacitance, variation in the resistance of the diffusion layer and to the possibility of electrical short-circuiting of the junctions in the nearby openings. The latter fact imposes a restriction on an increase in the degree of integration of the microcircuits.

Influence of the Masking Film. Local diffusion of the admixtures takes place under somewhat different conditions than ordinary one-dimensional diffusion which proceeds over the entire surface of the semiconductor plate. The presence of a masking film, as a result of the difference in values of the thermal coefficient of linear expansion of it and the semiconductor plates leads to the occurrence of mechanical stresses. In the masking film these stresses are tensile stresses. Mechanical stress causes the appearance of additional dislocations at the edge of the opening and bending with respect to the direction of the silicon surface. The intensified diffusion of the admixtures along these dislocations can lead to rejection of the planar structures.

It is very important that the p-n junction reaches the surface of the plates under the masking film, that is, it turns out to be protected from further production processing and external effects until its formation. This explains the high stability of the parameters of the planar structures.

Effect of Thermal Oxidation. It is necessary to consider that at high temperatures thermal oxidation of silicon takes place simultaneously with diffusion. The diffusion coefficients of the alloying admixtures in the oxide film are appreciably smaller than in silicon. Therefore the p-n junction formed in the presence of a surface oxide film has significantly less depth by comparison with the theoretically calculated film. In addition, as a result of differences in solubility in the oxide and in the silicon, redistribution of the admixture takes place in accordance with the distribution coefficient of the silicon oxide. For the most standard admixtures of boron and phosphorus the distribution coefficients are opposite, the surface layer of the silicon is impoverished with respect to boron (the distribution coefficient of the boron is greater than one), and in the case of phosphorus diffusion, on the contrary, it is enriched (the distribution coefficient of the phosphorus is less than one). All of this leads to complication of the diffusion process and to anomalous distribution of the admixture. In the case of diffusion of the admixture with the silicon oxide distribution coefficient less than one the maximum concentration shifts into the depths of the diffusion layer. Decreasing the surface concentration has a negative effect on the quality of the resistance contacts to the local regions.

10-3. Methods of Achieving Diffusion

Two-Stage Diffusion. In order to obtain reproducible parameters of the diffusion layers and comparatively small surface concentrations which are almost always required when making IC, the majority of the diffusion processes are conducted in two stages.

In the first stage, a defined amount of alloying admixture from an unlimited source is introduced into the thin surface layer of a semiconducting plate. At the same time, in the first, frequently called "buildup," the surface layer of increased concentration is created -- a source of admixture for the second step. The surface concentration of the admixture is large after the "buildup" phase.

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

The first stage usually is carried out quickly and at comparatively low temperatures. A layer of admixture silicate glass is formed on the surface of the silicon plates.

The second stage of diffusion is diffusion annealing sometimes called "dispersal," it is carried out with preliminary removal or without removal of the surface glass obtained in the first stage. The second stage, as a rule, is combined with the operation of thermal oxidation, that is, the process is carried out in an oxidizing environment. Diffusion annealing is carried out in the absence of a source of admixture used in the first stage. Here redistribution of the admixture introduced in the first stage to a defined depth takes place -- surface concentration of the admixture is lower as a result of going into the depths of the semiconductor and into the layer of growing oxide, for the latter penetrates deeper into the silicon plate. The temperature and the time of performing the second diffusion stage are higher than the first stage.

If the final depth of diffusion is appreciably greater than the depth of the diffusion layer obtained in the first stage and if the admixture does not go into the environment from the body of the plate, it is possible to consider that in the second stage diffusion proceeds from the limited surface source and the distribution of the admixture is gaussian. The amount of admixture per cm^2 of plate surface Q is determined from the conditions of the first diffusion stage:

$$Q = \int_0^{\infty} N(x) dx = N_0 \int_0^{\infty} \text{erfc} \frac{x}{2\sqrt{D_1 t_1}} dx = 2 \sqrt{\frac{D_1 t_1}{\pi}} N_0, \quad (10-8)$$

where D_1 is the diffusion coefficient; t_1 is time; N_0 is the surface concentration in the first stage of diffusion.

Two-stage diffusion permits easier control of the alloying of the layers and obtaining of reproducible parameters of the structure. In the second stage it is possible to eliminate the errors in the first stage of diffusion. In the buildup stage for masking surfaces it is possible to use thinner or lower quality, for example, silicon dioxide films grown by pyrolysis.

It is possible to carry out diffusion in a flow tube, in a sealed volume (in an ampule), in a semisealed volume and in a forevacuum. Solid, liquid and gaseous compounds are used as the sources of the admixture (diffusants). The procedure for carrying out diffusion and the diffusant are selected as a function of the properties of the semiconductor and the requirements of the diffusion structures imposed on the parameters:

Diffusion in the Flow Tube in a Gas Carrier Flow. The procedure is basic to the first diffusion stage. The silicon plates placed in special quartz, high-resistance silicon or platinum boats are loaded through the open exit opening of the quartz tube. The entrance opening of the tube is connected to the gas distribution system of the diffusion device. The gas carrier delivers the diffusant vapor to the surface of the plates. In order to prevent atmospheric air from reaching the tube and also to remove the spent gas, an exhaust system is placed over the exit opening of the tube which maintains the directionality of the gas flow. The quartz tube passes inside the muffle of a one-zone and two-zone furnace having temperature dispersion in the zone of no more than $\pm 0.5^\circ\text{C}$. The

FOR OFFICIAL USE ONLY

reproducibility of the parameters of the diffusion structures from plate to plate and over the area of each plate depends on the temperature distribution in the furnace zone, the number of plates and their location with respect to the gas flow, the type of diffusant, the cleanness of the process, and so on. An increase in the output capacity of the process is achieved by construction of long-zone furnaces, the application of multitube devices and the application of new types of diffusants making it possible to locate the plates more closely to each other.

The diffusion in the flow tube can be carried out by using solid, liquid and gaseous sources of the admixture. Solid sources, depending on their location with respect to the semiconductor plates, can be separate, parallel and surface.

The separate solid sources -- boron and especially phosphorus anhydride -- have high vapor pressure at the diffusion temperature; therefore the process is carried out in the two-zone units. The container with powdered source is placed in a low-temperature zone (950-1050)°C for B₂O₃ and (200-300)°C for P₂O₅, and silicon plates, in the high-temperature zone (1000-1300°C) (Figure 10-4).

The tube is first blown out by a mixture of inert gas and oxygen. After establishment of the temperature conditions, the boat with the plates is shifted to the operating zone and diffusion is carried out in the presence of oxygen. This promotes the formation of borosilicate nB₂O₃·mSiO₂ or phosphorus silicate nP₂O₅·mSiO₂ glass (BSS or FSS, respectively) on the surface. The admixture source molecules penetrate into the surface of the plates through the layer of liquid glass. Chemical reactions take place on the surface with the release of an elemental alloying impurity which diffuses into the semiconductor plate:



The glass layer protects the plate surface from evaporation, the incidence of foreign particles, the formation of erosion and, in addition, it absorbs the admixtures from the surface of the silicon.

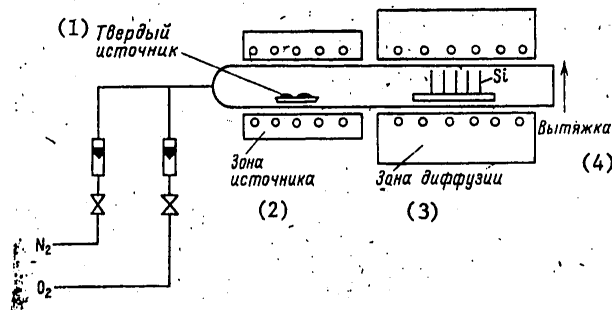


Figure 10-4. Diffusion in the gas carrier flow from a solid source

- Key:
- | | |
|-----------------|-------------------|
| 1. solid source | 3. diffusion zone |
| 2. source zone | 4. exhaust |

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

The deficiencies of diffusion from separate solid sources include the necessity for two-zone installation, insufficient reproducibility of the values of the surface concentration as a result of complexity of the exact maintenance of the concentration of the source vapor and as a result of variation of the evaporation area on spreading of the source which wets the container surface well. In addition, it is difficult to obtain low surface concentrations and it is impossible to carry out annealing in a pure gas atmosphere, for during the entire extent of the process the source vapors are in the gas phase.

Parallel sources of the admixture permit diffusion in the single zone unit. The role of the source can be performed by silicon or platinum plates with a layer of P_2O_5 or B_2O_3 applied to the surface; boron nitride plates previously oxidized in dry oxygen at a temperature of 1100-1150°C for the formation of boron anhydride on the surface of the layers; raw ceramic plates which include alloying oxide with high vapor pressure.

Diffusion from the parallel source is carried out with horizontal or vertical arrangement of the semiconductor plates. For vertical arrangement, a simpler holder is required, and greater productivity of the process is insured. Two plates of semiconductor in contact by the nonoperating sides are located between the source plates so that the spacing between the surfaces of the sources and the semiconductor will be no greater than $2\sqrt{Dt}$, where D is the diffusion coefficient of the impurity in the gas mixture; t is the diffusion time. Thus, it is possible to place up to 100 semiconductor plates in the temperature zone. The gas-carrier flux is normal to the surfaces of the plates; therefore the movement of the gas between the plates is very little. This promotes uniform transportation of vapor from source to silicon. Here high accuracy of regulating the speed of the gas flow is not required, and the depositions of the source vapors on the inside walls of the diffusion tube are insignificant. When insuring constant distance between the sources and surfaces of the semiconductor, high uniformity of the parameters of the diffusion layers is achieved over the plate area. The predominant application of boron nitride as the parallel source of oxidized plates is explained by high stability and nontoxicity of boron nitride and also the possibility of multiple use in the diffusion processes.

The surface sources, primarily alloyed oxides, sometimes metal and layers of alloyed polycrystalline silicon, also permit diffusion in the single-zone unit. The layers of alloyed oxides can be obtained by application from solutions, deposition from the vapor-gas phase, anodic electrolytic oxidation, reactive cathode sputtering in the presence of vapor of the compound containing the alloying elements. Thus, for example, for alloying the oxide with boron it is possible to use tripropylborate $B(OC_3H_7)_3$, for alloying the oxide by phosphorus it is possible to use trimethylphosphate $(CH_3)_3PO_4$, the vapor of which can be delivered from the batcher to the oxide growth zone using nitrogen. The B_2O_3 , P_2O_5 , H_3BO_3 , H_3PO_4 , $N_2B_4O_7 \cdot 10H_2O$ powder and others are dissolved in water or alcohol and are applied to the surface of the semiconductor plates by centrifuging. After drying in the air and heat treatment in an argon environment, a vitreous layer is formed on the surface.

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

During the confusion process, the oxide layers are in the liquid state; they draw off the harmful admixtures and metal ions located on the surface of the silicon well. Diffusion, as a rule, is carried out in an inert environment. In the case of using an oxidizing environment, it is possible to cut off the alloyed layer from the semiconductor surface by the thermal oxide. The diffusion from the alloyed oxides permits us to obtain a wide range of values of surface concentrations from 10^{16} to 10^{20} cm^{-3} with small dispersion for decreased erosion and defectiveness of the layers. On one semiconductor plate it is possible simultaneously to perform local diffusion of admixtures of different types. Alloying the surface oxide by two admixtures with different diffusion coefficients, it is possible simultaneously to obtain two diffusion layers.

It is possible to carry out diffusion of gold and nickel from the corresponding layers of the clean metals applied by thermovacuum deposition or electrolytic deposition. In the nickel layer applied electrolytically there is up to 5% atomic phosphorus, which permits it to be used as the source for obtaining n^+ -regions.

The diffusion from the surface sources provides for high uniformity of addition of the admixture to the surfaces of the semiconductor plates by comparison with diffusion from a separate source, the vapors of which are in the gas phase. The degree of delivery of the admixture to the semiconductor is regulated by the degree of alloying and the thickness of the layer applied to the surface. The low surface concentrations are more easily insured than when alloying from the gas phase. The film of the applied source protects the surface of the semiconductor from external effects and from evaporation into the surrounding space.

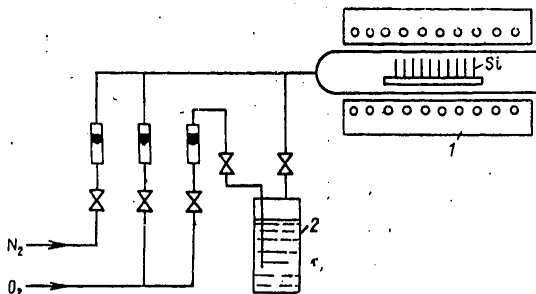


Figure 10-5. Diffusion in the gas carrier flow from a liquid source.
1 -- single-zone furnace; 2 -- liquid source

The liquid sources permit two-stage dilution of the source vapor by the gas flow passing through the batcher and the total flow going directly to the quartz tube (Figure 10-5).

In the reaction zone located in front of the diffusion zone, oxides of the corresponding alloying elements are formed



FOR OFFICIAL USE ONLY

or

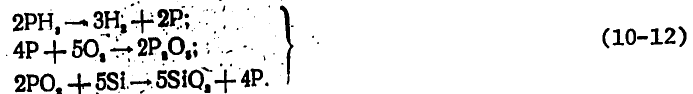


The reaction products are delivered by the gas flow to the silicon blades where the reactions (10-9) proceed. In order to prevent erosion, oxygen is added to the inert gas. Diffusion proceeds through the layer of liquid glass.

Liquid sources require sharper shifts by comparison with the separate solid sources, diffusion proceeds in the single-zone unit; curtailing the liquid source vapor feed, it is possible to carry out the second stage of diffusion in the presence of admixture-silicate glass on the surface. The range of values of the surface concentrations is wider than in the case of using separate solid sources.

The basic deficiency of diffusion from liquid sources is their toxicity.

Gaseous sources are fed from a bottle and they are mixed with nitrogen and oxygen before entering the quartz tube. The oxide of the alloying element is formed in the reaction zone, and an elementary admixture is released on the surface of the silicon plates. For example, the process of phosphorus diffusion when using phosphine is accompanied by the reactions:



The advantages and disadvantages of diffusion from gaseous sources are the same as from liquid sources.

Diffusion in Ampules. Semiconductor plates and the source are placed in individual boats in a quartz ampule, which after annealing to the maximum vacuum or after filling with inert gas is sealed and placed in a heated furnace. Depending on the evaporation temperature of the source, just as for diffusion from solid sources in a flow tube, a single zone or two-zone furnace is used. As the source of the admixture it is expedient to use powdered silicon alloyed with the required admixture. The concentration of the admixture on the surface of the plates cannot exceed the concentration of the admixture in the powder. The procedure is used for diffusion of various admixtures of boron, antimony, arsenic, phosphorus, gallium, and so on. The range of surface concentrations used for the first three admixtures is $2.8 \cdot 10^{17}$ to $1.3 \cdot 10^{20} \text{ cm}^{-3}$. Concentrations of up to $2.3 \cdot 10^{21} \text{ cm}^{-3}$ are reached for phosphorus.

The limitations of the method are as follows: technical difficulties when manufacturing, preparing and evacuating the ampule; one-time use of the ampules as a result of the necessity for breaking them after diffusion and also the possibility of contamination from the ampule walls.

The diffusion of the ampules insures good reproducibility of the parameters of the diffusion layers, it is less dangerous (by comparison with the slow tube

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

methods) when using volatile toxic admixtures. In cases where diffusion takes place to a larger lot of plates, the procedure is economically justified.

Diffusion in a Semisealed Volume (the Box Method). Semiconductor plates and the source of the admixture are placed in a container. The container with incompletely closed cover is placed in a heated quartz flow tube at the exit. After forcing air out of the container by an inert gas with which the quartz tube is blown, the cover is tightly closed, and the container is advanced into the temperature zone where the diffusion process takes place. On evaporation of the source vapor the entire volume of the container is filled, and it partially exits to the outside. The procedure can be used under the condition where the diffusion temperature does not exceed the evaporation temperature of the source too much. For example, it is impossible to carry out diffusion from phosphorus anhydride in a container. As the source of phosphorus, phosphorus-calcium glass $nP_2O_5 \cdot mCaO$ is used which has high stability, and at the diffusion temperature is not reduced to pure calcium.

In contrast to ordinary diffusion from separate solid sources in the flow tube the box method permits regulation of the surface concentration of the admixture within broader limits. The process is carried out in one temperature zone; it has little sensitivity to the gas flow velocity. The minimum flow velocity is limited by the probability that air will get into the container. There is no effect of turbulence of the gas flow.

Diffusion in a Forevacuum. For diffusion in a forevacuum, the application of a gas carrier is excluded and, consequently, the influence of its flow rate on the parameters of the diffusion layers. This provides good reproducibility of the results. When admitting the oxygen to the vacuum chamber it is possible to carry out the diffusion with oxidation. Alloyed layers are formed on the surface, from which the admixture is diffused into the plate. This provides for obtaining low surface concentrations and eliminates surface erosion. For the performance of diffusion rarefaction provided by a forevacuum pump is sufficient.

At the present time a large number of sources of alloying admixtures and a great variety of methods of diffusion providing for obtaining diffusion layers with high precision of reproduction of the parameters and thicknesses to tenths of a micron are used.

The flow tube method is used most widely in the production of microsecond at the present time. The diffusion process imposes high requirements on the sources of the admixtures, the cleanness of preparation of the semiconductor plates, fittings and quartz tube. The performance of diffusion of a defined admixture is realized in a specific device, in which the diffusor of another admixture is not carried out, and fittings from other devices are not used. In order to prevent impoverishment of the gas flow with admixture vapor as a result of absorption by the inside walls of the quartz tube, before diffusion, the latter is especially saturated with the admixture. The quartz fittings and the control thermocouples for measuring the temperature profile along the furnace and determining the boundaries of the temperature zone are stored in special quartz jackets.

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

10-4. Defects and Control of the Diffusion Structures

Defects of the Diffusion Layers. The basic defects are dislocations. The dislocations which are additional with respect to the initial semiconductor arise as a result of mechanical stresses of the crystal lattice as a result of differences in the dimensions of the semiconductor atoms and the introduced admixture. More improved layers are obtained for diffusion of the admixtures in which the atomic radii differ little from the silicon (As, Ga, Al).

For high diffusion temperatures undesirable processes of the formation of chemical compounds, alloys and subsequent evaporation of them can take place on the surface of the silicon. This leads to surface disturbances and the appearance of irregularities in the form of depressions and erosion. In order to eliminate erosion before the beginning of the diffusion process the silicon plates are easily oxidized, the thin oxide prevents the formation of erosion and in practice does not interfere with diffusion. The diffusion from the surface oxide layers is optimal from the point of view of maintaining a uniform, level surface. In addition to dislocations and erosion, the formation of sections of increased concentration of the admixtures is possible, between which and the semiconductor metallurgical interactions take place with the formation of a liquid phase and fusion of the diffusion layer. From the point of view of the appearance of the fusion zones, admixtures of iron, copper, gold, and so on are especially dangerous.

Control of the Diffusion Structures. The properties of the diffusion layers are examined on the control satellite plates. The control parameters are the depth of occurrence in the p-n junction, surface resistance or surface concentration of the admixture, concentration distribution of the admixture with respect to depth of the diffusion layer, and density of the defects of the diffusion layers.

The depth of occurrence of the junction is given first using the diffusion process parameters which can be theoretically calculated for the expressions for the admixture concentration distributions with respect to depth of the diffusion layer.

The depths of occurrence of the p-n junctions given using the diffusion conditions in practice are only approximately insured. For correction of the diffusion conditions the satellite plates are controlled by the methods of coloring oblique or spherical microsections with subsequent measurement under a microscope (see §9-8).

The surface resistance is determined by the four-sonde method. The current through the sondes must exceed a defined minimum value at which self-blocking of the p-n junction takes place and the current only passes through the diffusion layer, not branching into the semiconductor plate. The specific resistance of the diffusion layer is calculated by the formula (9-10). For measurements of the specific resistance the head with the sondes must be lowered smoothly, without shocks; during the measurement process it is necessary to check the condition of the sondes.

The distribution of the admixture with respect to depth of the diffusion layer, just as with respect to depth of the epitaxial layer, is determined by the method of differential electrical conductivity investigated in §9-8.

FOR OFFICIAL USE ONLY

The defects of the diffusion layers are investigated primarily in the stage of development of the microcircuits by the methods of x-ray topography and electron microscopy.

Under production conditions, the presence of erosion is checked both on the control satellite plates and on the operating silicon plates after local diffusion.

For observation of erosion it is possible to use any microscope with magnification of more than 200 power, for example, the MM-7, MBI-6, MI-4. The samples are examined in the dark field of the microscope with oblique illumination. The erosion corresponds to the presence of small light spots on the dark background.

Test Questions and Assignments

1. What is the purpose of using the method of thermal diffusion in the production of semiconductor devices and microcircuits?
2. Explain the terms indicating the possible mechanisms of diffusion in a solid state.
3. What concepts of vacancy and interstitial mechanisms of diffusion were introduced into physics?
4. What diffusion mechanism in silicon and germanium is the most probable for admixtures of groups III and V of the periodic table?
5. Why does diffusion of the interstitial admixtures proceed appreciably more rapidly than diffusion of the substitution admixtures?
6. What is the specific solubility of the admixture and the maximum solubility temperature of the admixture in a semiconductor? Explain your answer in examples from Table 10-1.
7. Write the first equation of diffusion and explain it.
8. Write the second equation of diffusion and explain it. Why is the second equation called the basic diffusion law?
9. For what values of the temperatures and why do diffusion processes take place?
10. What is the concentration distribution of the admixture with linear diffusion from an infinite source? What are its characteristic features?
11. What is the concentration distribution of the admixture with linear diffusion from a limited source? What are its characteristic features?
12. Why are the real admixture distributions different from the theoretical ones?
13. List the diffusion characteristics in planar technology and what causes them. How does it influence the parameters of the structures?

FOR OFFICIAL USE ONLY

14. What explains the broad application of two-stage diffusion?
15. What procedures are used for high-temperature diffusion?
16. What sources of alloying admixture are used in diffusion processes on silicon?
17. What admixture sources can be used for performing diffusion by the flow tube method?
18. Explain the essence and the technique of the diffusion process from separate solid sources (Figure 10-4).
19. Compare diffusion from a parallel solid source with diffusion from a separate source and from surface sources.
20. Compare diffusion from a solid source with diffusion from a liquid and gas sources.
21. For what purpose and how is gold and nickel diffusion in silicon accomplished?
22. What chemical reactions accompany diffusion from solid, liquid and gaseous sources?
23. How is the degree of alloying regulated during diffusion from different sources?
24. Compare the sources of admixture diffusants with the admixture sources during epitaxy.
25. In what cases is it expedient to carry out diffusion with ampules and what is the technique? What sources of admixture are used here?
26. What is diffusion in a semisealed volume? What sources of admixture can be used in this diffusion? What are the advantages of this method of diffusion?
27. What are the advantages of diffusion in a forevacuum?
28. Which of the diffusion methods most frequently is used in IC production? Why?
29. Name the basic defects of diffusion layers, their causes and means of eliminating them.
30. What parameters of the diffusion structures are controlled to make IC? What is the method of controlling these parameters?
31. What are the deficiencies of the method of high temperature diffusion?

FOR OFFICIAL USE ONLY

CHAPTER 11. ION ALLOYING AND OTHER METHODS OF OBTAINING SEMICONDUCTOR ELEMENTS

11-1. Fundamentals of the Method of Ion Alloying.

Application of the Method. After high temperature diffusion ion alloying is the second basic industrial method of obtaining alloyed semiconductor layers when manufacturing the structures of microcircuits and mounted active elements for hybrid IC.

The method of ion alloying was used to make p-n junctions by V-Shockley in 1954. The idea of ion alloying was proposed in the USSR by M. M. Bredov. Intense study of ion alloying since the beginning of the 1960's in the Soviet Union and abroad has demonstrated the indisputable advantages of the new method and permitted it to go beyond the framework of laboratory research into industrial production. Soviet researchers and developers V. M. Struzhinskiy, P. V. Pavlov and many others made a significant contribution to the development and introduction of the method into Soviet production.

At the present time the method of ion alloying is used when making semiconductor devices with large junction area, solar cells and silicon detectors of nuclear radiation; varactor diodes, avalanche photodiodes; bipolar and MDS-microcircuits. For example, the method makes it possible to obtain bulk resistors of integrated microcircuits with rated resistances from 10 kilohms to 1 megohm.

Mechanism of Ion Injection. The process of ion alloying is carried out in two steps. First the ions are injected into the semiconductor plates; then it is annealed, during which the structure of the semiconductor disturbed by injection is restored, and the admixture ions occupy nodes of the crystal lattice.

The mechanism of ion injection differs theoretically from the diffusion mechanism. During ion injection the movement of the ions in the semiconductor is determined by their initial kinetic energy and not the concentration gradient as during diffusion. The diffusion process is caused by the effort of the system to return from the nonequilibrium state in which the sections exist with different concentration of the admixture, to the equilibrium state. Diffusion is a very slow process, and for practical realization of it, high temperatures are required. During diffusion the maximum concentrations of the introduced admixture cannot exceed the values of the maximum solubility in the semiconductor of given admixtures at given temperatures. Ion injection is not an equilibrium process; therefore the concentrations of the admixtures easily reach the maximum values, and in some

FOR OFFICIAL USE ONLY

cases can exceed them. In addition, ion injection is a fast process; it can be carried out at room temperature, it can be used for alloying with admixtures with low diffusion coefficient or low solubility in the solid phase and also for alloying semiconductors with low melting point (indium arsenide) or for semiconductors, the diffusion alloying of which requires very high temperatures (silicon carbide).

Depending on the direction of incidence of the ions, ion injection can take place by two mechanisms leading to different depth of penetration and distribution of the admixture.

Disoriented injection of the ions is observed in the angle between the direction of incidence of the ions on the surface of the semiconductor plate and the direction of the principal crystallographic axis exceeds the critical angle. In this case, in the body of the semiconductor ions move with transfer of significant energy randomly as in an amorphous target. The interaction of the ions with the encountered atoms of the plate takes place as an independent ion-atom pair collision. For each collision the transfer of energy to the plate atoms and inhibition of the moving ion take place until the ion finally stops. If the energy transferred by the ion exceeds the bond energy of the atoms in the lattice, the atoms shift and leave the lattice nodes, as a result of which Frenkel' defects are formed -- vacancies and atoms in the internodes. Inasmuch as the primary energy of the ion is large, on the path of its displacement inside the semiconductor, numerous shift defects are formed which merge into zones 50-100 Å in size (Figure 11-1). As the ions are introduced, the number and size of the zones with disordered crystal lattice increases and, if the zone density is high, then at a defined point in time complete merging of all the defect zones takes place, that is, the semiconductor goes from the crystalline state to the amorphous state.

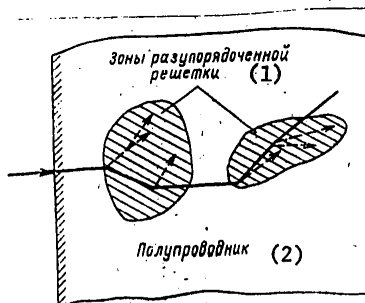


Figure 11-1. Trajectory of motion of injected ions.
 ----- deviations of the semiconductor atoms on collision with ions;
 ——— trajectory of motion of an ion

- Key:
1. disordered lattice zones
 2. semiconductor

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

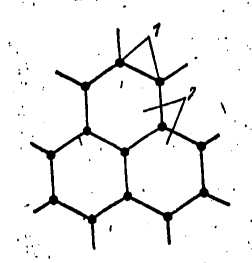


Figure 11-2. Channels in the (110) silicon plate.
1 -- silicon atom; 2 -- channels directed perpendicular to the plane of the pattern

The critical magnitude of the ion radiation dosage for which the semiconductor goes from the crystalline state to the amorphous state is called the amorphization dosage. The amorphization of the surface of the silicon plate under the effect of phosphorus ions takes place more intensely than under the effect of boron ions. The radiation dosage required for amorphization of the silicon surface is about $6.2 \cdot 10^{14} \text{ cm}^{-2}$.

The oriented introduction takes place if the direction of incidence of the ions on the surface of the semiconductor plate in practice coincides with one of the basic crystallographic directions [110], [100] or [111], that is, when the angle between the directions of incidence of the ions and the main crystallographic axis is less than critical. For approximate introduction the movement of the ions takes place with small energy losses and is of an ordered nature, for it is realized along the free channels (Figure 11-2). Channeling is penetration of the ions with respect to the internodes along the atomic series or planes forming the interaxial or interplanar channels. During channeling the ions experience sliding collisions for which the energy losses are insignificant, and the number of shifted ions is small. The majority of ions maintain the duration of their movement, and therefore the paths of the ions exceed by many times the paths of the unordered trajectories.

The depth of penetration of the ions on the channels depends on the crystallographic orientation of the channel and is maximal for the direction [110], for the density of the lattice atoms of the diamond type in this direction is less than in other directions. The depth of introduction varies in accordance with the series [110] > [100] > [111]. The depth of introduction along the direction [110] is approximately twice that in the direction [111].

If the direction of incidence of the initial beam of ions on the plate surface is deflected from the basic crystallographic direction and, consequently, from the channel axis, the ions go beyond the region of the channel and are shifted along disordered trajectories. Under real conditions, even if the angle of incidence is deflected from the channel axis by no more than the critical angle, clean channeling will not occur. This is explained by the fact that on introduction with an increase in radiation dosage part of the ions incident on the substrate will collide with the surface atoms. These ions do not hit the channels, but are deflected from the initial direction, and as a result of their disordered motion defective regions will be formed which, in turn, will influence the nature of

FOR OFFICIAL USE ONLY

introduction. For large radiation doses the concentration of the ions in the channel is large, and not all of them are under identical conditions. The energy of the ions moving along the central part of the channel is basically lost as a result of interaction with electrons. The losses of energy by the ions deflected from the center of the channel are also caused by interactions with the atomic nuclei. The deviations of the ions can also occur on interaction with the thermal vibrations of the semiconductor atoms, lattice defects, and so on. The influence of the additional causes of deviations of the ions from the center is greater, the farther the ion is located from the channel center. With an increase in the radiation dosage, the channeling becomes more noticeable, and for doses of 10^{15} cm^{-2} the channeling effect in practice disappears.

Annealing Radiation Defects and Activation of Impurities. The introduction of the ions leads to the appearance of radiation defects of the crystal lattice and does not insure active alloying. In order to insure active ion alloying it is necessary that the lattice of the semiconductor be disturbed to the minimum and that the ions of the introduced admixture fall on the nodes of the crystal lattice. Only in this case are the impurity atoms electrically active and can the role of the donors or acceptors be performed. The probability of direct incidence of the ions on the nodes is low. For restoration of the crystal lattice of the semiconductor and conversion of the ions to the electrically active state (from the internodes to the nodes) the semiconductor particles are subjected to annealing directly after ion injection. Most frequently the thermal annealing of the silicon in an argon atmosphere at temperatures of 600-800°C for 10-20 minutes is used. At these temperatures the shifted atoms of the lattice and the injected ions acquire sufficient mobility for conversion to the vacant nodes, as a result of which ordering of the crystal lattice of the semiconductor takes place. Sometimes annealing is matched with injection of the ions; for this purpose the injection is made into the heated plate. This procedure is called hot introduction. It is recommended that it be used for admixtures with a small diffusion coefficient, for otherwise it is quite difficult to control the profile of the admixture distribution.

11-2. Distribution of the Admixture Concentration in the Ion-Alloyed Layers

Factors Influencing the Depth of the Alloyed Layer. The analytical calculation of the complete distribution profile of the injected ions is very complicated and can hardly be carried out with acceptable degree of accuracy. Therefore the calculation is performed for a large number of assumptions separately for disoriented and oriented introduction.

The depth of the alloyed layer depends on the accelerating voltage, the crystallographic orientation of the surface of the plates, the angle of incidence of the ions on the surface of the plates, the temperature of the plates during injection, and the annealing temperature of the radiation defects.

Ion Distribution for Disoriented Injection. In practice injection is realized for an angle of incidence of the ions greater than critical. This is explained by the fact that in industrial devices no provision is made for accuracy of orientation required for channeling. In the case of disoriented injection the periodicity of the monocrystalline lattice is almost not felt, and movement of the ions is of a random nature. When the ions move they gradually lose their energy as a result of interaction with electrons (the direction of motion does

FOR OFFICIAL USE ONLY

not change), and as a result of collisions with the atomic nuclei (the directional motion changes). The trajectory of the motion of the ion has the form of a broken line (see Figure 11-1), each rectilinear section of which corresponds to the path of the ion to collision with the lattice atom. The total path length of the ion until it halts is felt from individual sections of the broken trajectory.

The average projection line of the total path on the direction of the initial ion velocity \bar{R}_π called the mean normal path has practical significance for determining the depth of injection of the ions.

Ion collisions with lattice atoms are of a probability nature; the energy lost for each collision is different. In accordance with the probability theory it is proposed that all of the set normal ion paths is grouped by a gaussian law with a mean total normal path \bar{R}_π and mean normal deviation ΔR_π .

The distribution of the introduced ions is determined by the gaussian law:

$$N_x = \frac{Q}{\sqrt{2\pi\Delta R_\pi^2}} e^{-\frac{(x-\bar{R}_\pi)^2}{2\Delta R_\pi^2}} \quad (11-1)$$

where x is the depth of the injected ions; Q is the radiation dosage. The values of \bar{R}_π and ΔR_π for ions of groups III and V of the periodic table injected into a silicon plate at various ion energies are presented in Table 11-1.

In Figure 11-3 the standard distribution curves for ions with low and high energies are presented. The characteristic features of these distributions are as follows: the distribution is subject to Gauss's law, the maximum distribution is not on the substrate surface as in the case of thermal diffusion, but at a distance from the surface equal to the mean normal path of the ions; with an increase in the radiation energy, the maximum distribution shifts into the depth of the substrate. This makes it possible to obtain a latent submerged layer with conductivity of the opposite type.

Ion Distribution with Oriented Injection

In Figure 11-4 the provisional diagram of the motion of ions incident on the surface of a plate at angles less than critical and also the distribution curve of the introduced ions corresponding to this diagram are presented in Figure 11-4.

The first maximum on the ion distribution curve is caused by deviation of the ions which are near the surface of the plate from the channeling direction as a result of interaction of part of the incident ions with the surface atoms of the semiconductor. The concentration distribution corresponds to gaussian distribution of the paths of the ions deflected from the surface. The second distribution maximum farther from the surface is caused by the ideally channeled ions. The calculations show that the distribution of ideally channeled ions is not gaussian. The average region of the distribution curve is caused by ions which get into the channels but not under ideal conditions. On traveling some path, they are dechanneled and introduced at depth in the interval between two maximums.

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

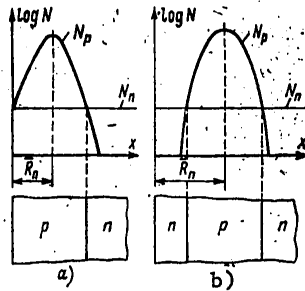


Figure 11-3. Ion distribution with low (a) and high (b) energies and disoriented injection

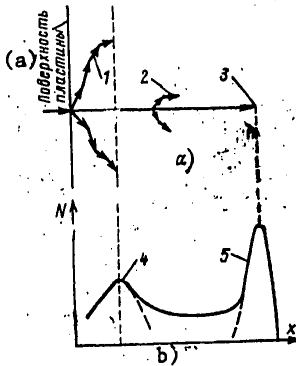


Figure 11-4. Provision diagram of motion (a) and the distribution curve of the ions (b) for oriented injection

1 -- deviations of the ions on collision with the surface atoms of the plate; 2 -- dechanneling of the ion; 3 -- ideally channeled ion; 4 -- gaussian distribution; 5 -- nongaussian distribution

Key:

a. Plate surface

With an increase in radiation dosage, the probability of scattering of the ions at the surface and further dechanneling increase, and therefore the channeling decreases, the second maximum on the distribution curve of the introduced ions decreases, and the first increases (Figure 11-5).

With higher irradiation dosage and at an initial ion energy that is not too high, the second distribution maximum in general can disappear against the background of the dechanneled ions.

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

Table 11-1. Values of \bar{R}_n and σ on a Silicon Plate, Å

Ускоряющее напряжение, кэВ		20	40	60	80	100	120	140	160	180	200
B	\bar{R}_n	784	1610	2438	3238	3977	4688	5374	6030	6654	7253
	σ	321	538	708	839	939	1025	1098	1160	1213	1259
N	\bar{R}_n	541	1116	1693	2270	2843	3400	3296	4430	4925	5409
	σ	229	396	532	649	751	837	908	969	1026	1076
Al	\bar{R}_n	294	566	853	1146	1446	1750	2055	2358	2662	2969
	σ	110	195	275	348	418	484	544	598	651	702
P	\bar{R}_n	260	490	732	980	1233	1491	1752	2014	2277	2538
	σ	94	164	231	295	354	411	466	518	566	610
Ga	\bar{R}_n	163	280	392	502	611	720	829	940	1053	1167
	σ	40	67	91	115	137	158	179	199	220	241
As	\bar{R}_n	159	271	377	481	584	686	789	891	995	1101
	σ	37	62	84	105	125	145	163	182	200	219
In	\bar{R}_n	141	231	312	390	464	538	611	684	756	827
	σ	25	40	53	66	77	89	100	111	121	132
Sb	\bar{R}_n	140	229	308	384	457	528	599	669	739	809
	σ	24	38	51	63	74	84	95	105	115	125

Key:

1. Accelerating voltage, kev

Practical Distribution of Injected Ions. The practical distribution curves of the introduced ions have characteristic features of the calculated curves and differ from them by the presence of a "loop" (Figure 11-6). With an increase in the radiation energy the maximum of the curves moves out, and the practical curves differ sharply from theoretical. The causes of these differences are as follows: diffusion, which with disordering of the crystal lattice is high even at low temperatures; incidence of ions not channeled at the beginning of the path in the channels; the influence of the defects occurring during ion injection, the nature of which can differ sharply from the ordinary crystal defects, and so on.

Effect of Annealing of the Distribution of Active Admixtures. The annealing of plates at temperatures of 500-700°C eliminates the radiation defects and promotes activation of the alloying admixtures. The probability of activation of the admixtures increases with an increase in temperature and the differences between the physical (which takes into account all the injected ions) and electrical (which considers only the electrically active ions incident on the nodes) distribution decrease. The annealing at temperatures of more than 700°C leads to variation of the physical distribution profile and, the more so, the higher the temperature. In Figure 11-7 distribution curves of the phosphorus atoms in the silicon after annealing are presented. The spreading of the distribution curves with an increase in the annealing temperature is explained by the influence of the diffusion of the admixtures.

FOR OFFICIAL USE ONLY

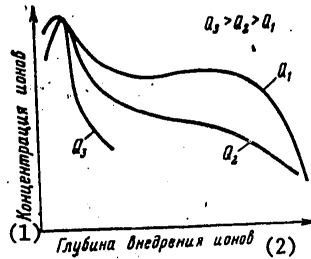


Figure 11-5. Influence of the radiation dosage Q on the ion distribution with oriented injection

- Key:
1. Ion concentration
 2. Depth of injection of the ions

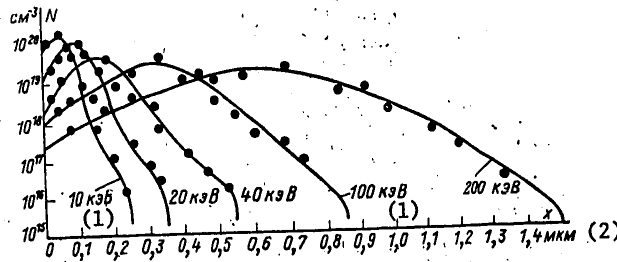


Figure 11-6. Experimental distribution profiles of electrically active boron introduced at different energies into silicon

- Key:
1. keV; 2. microns

The basic feature of ion alloying is the possibility of reproducibly obtaining of the given concentration of admixture at a given depth in practice in any area of the plates. This arises from the fact that it is possible to give the ion beam current with high precision. It is possible to obtain shallow uniformly alloyed layers and also sharp p-n junctions. The distributions of the admixtures can easily be controlled within broad limits, changing the radiation dosage, energy and angle of incidence of the ions.

In contrast to the diffusion method in which the thermal kinetic energy of the admixture atoms is tenths of an electron volt, during ion injection the accelerated ions have a kinetic energy of tens of kiloelectron-volts. The ions bombarding the substrate overcome the surface potential barriers several units of electron volts high caused by the presence of residual pollutants, without difficulty. Therefore the ion injection proceeds uniformly over the entire treated area of the surface of the plate independently of the nature of the distribution of these contaminants. The characteristics of the ion-alloyed layers turn out to be more reproducible than for diffusion.

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

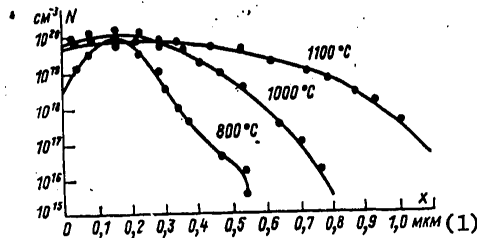


Figure 11-7. Influence of heat treatment on the distribution of the phosphorus introduced into the silicon

Key:

- 1. microns

The method of ion alloying makes it possible to introduce ions of different admixtures into the semiconductor plates and to obtain the required concentration distributions, including those which cannot be obtained by other methods of alloying.

11-3. Equipment for Implementing the Ion Alloying Process

Technical Implementation of the Method. Ionic injection is realized in differentially pumped vacuum devices, in which the degree of vacuum increases as the movement of the ions takes place from the source to the plates (Figure 11-8). The formation of the vapor of the material the composition of which includes an alloying admixture takes place in the ion source.

There is a great variety of structural designs of ion sources. In the source with arc discharge the chemical compound of the introduced element or pure element is placed in a heated crucible which is turned with its exit neck toward the wall of the gas discharge chamber. In the vapor of the ion source arc discharge is maintained. The ions from the arc discharge plasma are extracted through the set and accelerated by a system of extracting and accelerating lenses.

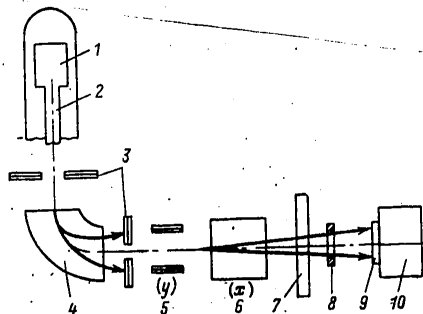


Figure 11-8. Diagram of the ion injection device

- 1 -- ion source; 2 -- accelerating and focusing system; 3 -- diaphragm;
- 4 -- electromagnetic analyzer; 5, 6 -- plate deviation; 7 -- scanning system (for the defocusing lenses); 8 -- ion current meter; 9 -- semiconductor plate; 10 -- plate holder

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

In the ion beam, in addition to the basic admixture there are also ions of other mixtures. In order to screen the ion there is an electromagnetic analyzer of the ions by mass. In the region of the analyzer where the ions move under the effect of the Lorents force around circular orbit, the radius of which depends on the mass m and the ion charge q :

$$r = \sqrt{2mU/(qB^2)}, \quad (11-2)$$

where U is the accelerating voltage, B is the magnetic induction of the field, in which the ion moves.

As a result of movement of the ions in circular orbits of different radii ion distribution by masses takes place in the electromagnetic analyzer. Thus, the beam of ions required for introduction is isolated.

Wide or sharp-focused ion beams are used for injection. In the first case the diameter of the ion beam is equal to or greater than the diameter of the machined plate. The current density distribution over the area of the plate must be uniform. In order to obtain a wide ion beam, defocusing lenses are used. In the latter case after separation of the ions with respect to masses, the system of focusing lenses and the scanning system of the ion beam over the surface of the plates are used. The scanning is necessary not only for uniform treatment of the surface of the plates, but also for local treatment of the required sections. In the first case, for local treatment it is necessary to use masking, and in the latter case, it is not mandatory.

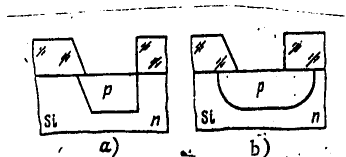


Figure 11-9. p-n junction front for ion (a) and diffusion (b) local allying

In contrast to local diffusion, during local ion allying independent selection of the impurity profiles in the vertical and horizontal directions is possible. For the wide-focus beam the distribution profile of the injected ions in the horizontal direction is formed using a mask; here the dimensions of the alloyed region reproduce the dimensions of the opening more precisely than during diffusion (Figure 11-9). The edge of the ion-alloyed junction coincides with the edge of the mask. This insures the high degree of integration of the microcircuits, but when the windows are opened under the resistance contacts, shunting of the junction by the metal films is possible. SiO_2 , Si_3N_4 , W, Mo, Au and other materials are used as the materials for the masking. The advantage of metal masks is the fact that the charge and the energy given up by the ion beam are carried off the surface of the plate. The thickness of the mask depends on the type of ion, its energy and braking capacity of the masking material. The standard thickness of the masks used in practice is 0.1-2 microns.

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

At the present time both ion-beam, wide-focused type devices and devices with focused ion beam are used in industry. Thus, for example, the wide-focus ion-beam "Vezuviy" device is used to introduce ions with an energy to 200 keV and insures alloying precision with respect to depth of 0.02 micron. The ILU-2 ion beam accelerator provides for the formation of boron and phosphorus ion beams to a diameter of 50 microns at energies of 150-250 keV for phosphorus and 50-150 keV for boron.

The ion beam is directed at the vapor of the receiving device. Semiconductor plates are placed on the coordinate table. The plates must be arranged at the exact angle to the direction of the ion. The ion injection units are equipped with special means of orientation directly in the chamber of the device. Provision is also made for the possibility of heating the plates during the ion injection process.

The basic parameters of the process of ion injection are as follows:

The ion energy $E=qnU$, where q is the elemental charge; n is the multiplicity of ionization of the ion; U is the accelerating voltage. Usually $E=10$ to 200 keV;

The ion beam current density which usually is $0.1-100$ microamps/cm²;

The radiation dosage Q determined by the product of the current density of the ion beam times the irradiation time $Q=jt$. Frequently the radiation dosage is measured in units of surface area of the introduced ions for calculations. When creating semiconductor devices and integrated microcircuits the radiation dosage range from $6 \cdot 10^{11}$ to $6 \cdot 10^{17}$ cm⁻² is used. The lower dosages are required to create base layers of bipolar transistors, thyristors, varactors, and high dosages for the creation of emitter layers and layers for the resistance contacts of the semiconducting structures;

The angle of inclination of the ion beam to the direction of the main crystallographic axis of the plate (is usually 7-8°).

Thus, in contrast to diffusion, during ion alloying the properties of the layers obtained basically are determined by the external electrical parameters of the process. It is easy to control the ion injection process by varying the accelerating voltage, the ion beam density, the angle of inclination of the beam, the irradiation time of the plates, and in the case of treatment by a focused beam, also its scanning speed over the surface of the plate.

Admixture Sources. When creating silicon structures for introduction, primarily phosphorus and boron ions are used. Elemental powdered red or crystalline phosphorus is used as the phosphorus source. The boron ions are obtained from chemical compounds -- boron halides. In contrast to solid sources submerged in a crucible, the halide vapor and gases are admitted to the gas discharge chamber of the ion source through a leak.

11-4. Advantages and Disadvantages of Ion Alloying

Ion alloying of semiconductor plates along with the advantages indicated in §11-1 has a number of additional advantages over the method of high-temperature diffusion:

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

1. Low temperature of the ion alloying process.

Thermal annealing is carried out at much lower temperatures than diffusion and, as a rule, discontinuously. This excludes worsening of the parameters of the initial plates and redistribution of the previously introduced admixtures. The low-temperature nature of the process of ion injection is especially important when alloying complex semiconductors containing easily volatile components.

2. The purity of the process, for ion injection is realized in a vacuum.

3. Less rigid requirements on the purity of the source of the alloying admixture, for on separation of the ions in the electromagnetic field of the analyzer of the device, isotopic purification of the alloying admixture is achieved.

4. High uniformity and reproducibility of the results. Ion alloying makes it possible reproducibly to obtain thin alloyed layers to tenths of a micron thick with the precision to 0.02 micron. This is possible by assignment of the ion energy.

5. The high regulation precision of the distribution profile of the introduced admixtures by exact maintenance of the ion beam parameters. During step regulation of the ion energy it is possible to form in practice any complex profile which is obtained as a result of superposition of individual profiles. In addition, it is possible to obtain latent alloyed regions.

6. The possibility of local treatment using an ion beam focused to the required size or using protective masks with a wide beam.

7. The possibility of alloying through the protective layer, for example, through a layer of silicon dioxide, permits regulation of the distribution of the admixture in the surface layer and also does not introduce contaminants at the Si-SiO₂ interface.

8. The possibility of realizing an entire set of operations in one unit: the cleaning of the surface, manufacture of active and passive elements, surface protection, and so on.

9. The high output capacity and ease of automation of the process caused by the fact that the parameters of the layers obtained are determined by the electrical parameters of the ion beam and not the properties of the plates as in the case of diffusion.

However, the ion alloying method is characterized by a number of deficiencies:

1. Comparative complexity and high cost of equipment, especially when it is necessary to obtain submicron sizes of IC elements.

2. Complexity of uniform alloying on a large area.

3. The complexity of reproducibly obtaining deep (more than 1 micron) alloyed layers. For the creation of deep layers it is possible to increase the ion path, using the channeling effect or increasing the ion energy. In the first case the

FOR OFFICIAL USE ONLY

accuracy of orientation of the crystal relative to the beam must be to 0.1° , which today is technically difficult to obtain.

The presented deficiencies will be overcome as the method of ion alloying is studied and investigated in practice.

The process of ion alloying is prospective for obtaining slightly alloyed local layers with an exact content of admixtures in the range from $1 \cdot 10^{15}$ to $5 \cdot 10^{17} \text{ cm}^{-3}$ and an exact concentration gradient of the admixtures when creating thin base regions of microwave bipolar transistors, to obtain regions with reversible concentration gradient of the admixtures for varicats with large capacitance of overlaps and also to obtain high speed MDS-IC. The method is also prospective for obtaining thin highly alloyed layers with admixture concentration from $5 \cdot 10^{20}$ to $2 \cdot 10^{21} \text{ cm}^{-3}$ when making various microwave structures, MDS-structures and photoconverters.

It is also necessary to know that the possibilities of ion injection are not limited to alloying semiconductor billets. Thus, for example, introducing the oxygen or nitrogen ions into the surface layer of silicon, oxide or nitride films are obtained which have good dielectric properties, stability of the parameters and mechanical strength. These films can be used to passivate the surface of the IC structures.

11-5. Other Methods of Obtaining Semiconductor Elements

Diffusion from Ion-Alloyed Layers. This process is carried out in two stages: in the first stage the method of ion alloying is used to create the surface layer, in the second stage the admixture is diffusely redistributed from this layer to the required depth. As a result of high accuracy of regulating the concentration of the introduced admixture, it is possible to obtain highly reproducible results and the high percentage yield of usable structures. In addition, the combination of the method of ion alloying and thermodiffusion permits alloying by admixtures, for example, hydrogen, which cannot be introduced into the semiconductor by diffusion alone. Ion implantation with diffusion dispersal makes it possible to obtain both shallow (0.3-0.5 micron and less) and deep (more than 1 micron) alloyed layers relatively easily.

Diffusion from Alloyed Polycrystalline Silicon. This comparatively new method of diffusion was developed to create emitter regions of transistors and microcircuits in the microwave range. Diffusion from polycrystalline silicon is carried out in two steps. First a layer of polycrystalline silicon is grown on the surface of the silicon plate, which is alloyed by ordinary methods during the growth process or after it, using ion implantation. Then redistribution of the alloying admixture from the surface source obtained in this way into the depths of the semiconductor plate takes place. Arsenic and phosphorus can be used as the alloying admixture. If the arsenic content in the polycrystalline silicon does not exceed 5%, the silicon-arsenic system in the temperature range of $700-1300^\circ\text{C}$ is in the solid phase, which promotes high stability of the admixture redistribution process and reproducibility of the parameters of the emitter regions. The phosphorus redistribution takes place at temperatures of $1000-1100^\circ\text{C}$. At a temperature of 1131°C the silicon-phosphorus system is converted to the liquid phase which promotes the appearance of erosion on the surface of the silicon plates.

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

Diffusion from the alloyed polycrystalline silicon layers makes it possible to obtain shallow (less than 0.3 micron) emitter regions with high surface concentration at comparatively low temperatures. The layer of the alloyed polycrystalline silicon remaining in the finished structure on the emitter permits prevention of short circuiting between the emitter regions and metallization and also improves adhesion of the metal electrodes.

Radiation-Stimulated Diffusion. In the case of ordinary diffusion the alloying processes of semiconductor plates are carried out at high temperatures: 1000 to 1350°C for silicon and 800-900°C for germanium. The duration of the processes sometimes reaches tens of hours.

A more effective means of increasing vacancies and, consequently, the diffusion rate is treatment of the semiconductor with accelerated electrons, protons (H^+) or other particles. In the treatment zone very high temperature gradients arise and regions are formed which are rich in vacancies and other defects. Additional vacancies increase the diffusion coefficient of the admixture by the amount ΔD proportional to the concentration of the additional vacancies, N_B and the diffusion coefficient of the vacancies D_B :

$$\Delta D \sim N_B D_B. \quad (11-3)$$

In order to obtain the p-n junctions using radiation-stimulated diffusion the surface of the semiconductor plate is subjected to local radiation treatment. Then the process is carried out over the entire area of the plate surface. The admixture diffuses predominantly into the irradiated sections. The depth of diffusion in the unirradiated part of the substrate is small. After removal of the layer of semiconductor with a thickness equal to the depth of diffusion in the unirradiated sections, the p-n junctions remain only at the points irradiated by a beam of accelerated particles.

The basic peculiarity of the method of alloying semiconductors using radiation-stimulated diffusion is low plate temperature during the process of creating the structures. For silicon this temperature can be 400-450°C. In this case there is no necessity for masking the surface. The method also insures high localization precision with small areas of the treated sections, high efficiency of the process, ease of automation and programming, high reproducibility of the structure parameters and the possibility of carrying out the cycle of operations in one unit.

Manufacture of p-n Junctions by Microfusion using Electron Beam. This method is based on using thermal effect on electrons accelerated to high energies on a solid state.

In contrast to the classical method of fusion the microfusion method using electron beam is based on the interaction of the liquid phase of semiconductor and the liquid phase of electrode material which is possible as a result of localization of the zone of thermal effect. This significantly increases the solubility of the semiconductor and, consequently, accelerates the process. In addition, there is no further problem with wetting the surface of the semiconductor, and the possibility of obtaining small junction areas is obtained.

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

In order to manufacture junctions a film of silicon dioxide is applied to the surface of the cleaned plate, and the alloying layer is applied over it. The plates prepared in this way are placed on the working table of the electron beam device which can be located in a vacuum chamber or outside it. In the latter case the electron beam is taken outside the vacuum chamber through a special iris. The processing conditions are selected so that the masking film will be destroyed and local fusion of the alloying layer and semiconductor will take place. After the beam is cut off during cooling the semiconductor is recrystallized, encompassing the atoms of the alloying admixture in the crystal lattice.

The duration of the effect of the electron beam does not exceed 1 second, there is no noticeable diffusion of the admixture from the melt into the solid phase in this time; therefore the distribution of the admixture in the p-n junction is obtained in stages.

The method of electron beam microfusion is characterized by the occurrence of very large temperature gradients in the treatment process which can lead to the appearance of a large number of defects in the junction region. The control of these defects permits the regions with small lifetime of nonequilibrium current carriers to be obtained reproducibly for the structures of high-speed devices. The processing is easily controlled by the accelerating voltage, current and time.

The basic advantages of the method are as follows: high accuracy of localization with small dimensions of the treated sections; high energy density of the electron beam (to 10^{13} watts/m²-sec) making it possible to reduce the treatment time to fractions of a second; high efficiency and ease of automation of the process. The method of microfusion using an electron beam is prospective for making structures of microcircuits in one vacuum unit with a continuous process.

Rectifying Metal-Semiconductor Contacts (Schottky Barriers). In 1938 Schottky developed a diffusion theory of the operation of a rectifying metal-semiconductor contact. The broad practical investigation of Schottky barriers which began in the 1960's demonstrated that the cause of the deviations of their volt-ampere characteristics from ideal is imperfections in the contact surfaces.

The process of manufacturing IC with the application of Schottky barriers is developing in the direction of finding materials and new structural designs of the Schottky barriers and combinations of them with other active elements of the IC.

The process of manufacturing the Schottky barrier includes the application of a metal film, annealing and formation of the Schottky barrier configuration. The choice of the metal-semiconductor pair providing for obtaining stable and reproducible characteristics of the Schottky barriers is one of the most important events of their manufacture. This problem is also connected with a second, no less important problem of selecting the optimal annealing conditions.

The Al-Si contacts are of great interest from the point of view of application in semiconductor IC. Aluminum with silicon of p-type conductivity or with highly alloyed silicon of the n⁺-type forms resistance contacts. Good Schottky barriers are formed at the interface with the slightly alloyed silicon. The height of the Al-n-Si potential barrier is very sensitive to the preparation of the silicon

FOR OFFICIAL USE ONLY

surface and the conditions of applying and annealing the aluminum. In order to obtain good contacts it is recommended that the metallurgical interaction between the silicon and aluminum with the formation of a liquid phase be reduced to a minimum. The aluminum dissolved in the silicon during cooling of the Al-Si system is captured in the recrystallizing silicon layer, promoting the formation of p-Si. In order to obtain Schottky barriers it is recommended that thin aluminum films be used and that high-temperature treatments not be applied. For n-Si the barrier height of an ideal contact is 0.65 to 0.7 ev.

In many cases the choice of the metal-semiconductor pair and annealing conditions reduces to intentional formation of intermediate phases on their interface which excludes the influence of the surface state on the characteristics of the Schottky barrier and improve their adhesion and stability. Thus, for example, when annealing Pt-Si, Pd-Si contacts, silicides of the corresponding metals are formed in their interfaces which increase the height of the potential barrier and improve the contact characteristics. The annealing of the Au-n-GaAs contact can lead to degradation of the Schottky barrier as a result of diffusion of gallium into the gold film and diffusion of gold into the gallium arsenide. The W-n-GaAs interface is metallurgically inert, therefore the annealing is not accompanied by the formation of intermediate phases and, consequently, is not critical to the contact properties.

In order to decrease the spurious parameters in the microwave range, the areas of the Schottky barriers must be small; therefore the configuration is formed using photolithography. In planar technology the area of the Schottky barrier is determined by the area of the opening in the masking film of silicon dioxide. In mesa technology (Figure 11-10) the metal film is applied to the entire area of the semiconducting plate and then through the contact mask obtained using photolithography the metal film and semiconductor around the required surface of the Schottky barrier are etched out.

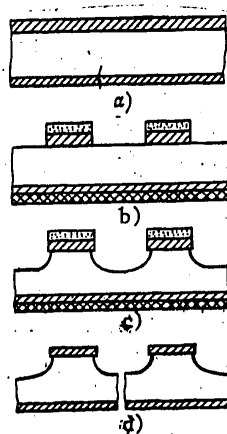


Figure 11-10. Diagram of making the Schottky mesa-diode
 a -- application of films for the barrier and ohmic contact; b -- application of a protective coating to the lower side of the plate, formation of the photoresistive mask and transfer of the pattern to the barrier films; c -- local pickling; d -- removal of the photoresistive mask and protective coating

FOR OFFICIAL USE ONLY

The most important features of the Schottky barrier are as follows: current transport is realized by the majority carriers; the accumulation of injected charge and, consequently, the time for resorption of it are absent; the voltage switching time is short; the Schottky barrier quickly reacts to changes in the external surface; the direct voltage drop is much less than in devices where the minority carriers participate in the current transport; the Schottky barrier has better noise characteristics compared to the p-n junction and is technologically more easily executed.

As a result of exceptional properties by comparison with p-n junctions at the present Schottky barriers are being widely studied and introduced into production. They are used to make microwave diodes, bipolar and field-type transistors, photodiode and solar cells. They are widely used in the logical microcircuits of diode-transistor, transistor-transistor logic, in memory microcircuits, and so on.

Test Questions and Assignments

1. What is the purpose of the ion alloying method in IC production?
2. Compare ion implantation with ion sputtering and ion pickling.
3. What is the mechanism for disoriented ion implantation (Figure 11-1)?
4. What is the mechanism of the channeling of ions?
5. What is the mechanism of ion injection under actual conditions for an angle of incidence of the ions of less than critical?
6. What explains the necessity for annealing after ion injection?
7. How are plates annealed in practice?
8. What are the plate temperatures during injection? During annealing?
9. Explain the nature and peculiarities of ion distribution during disoriented injection (Figure 11-3).
10. Explain the diagram of the motion of ions, nature and peculiarities of ion distribution in the case of oriented injection (Figure 11-4, 11-5).
11. Explain the differences in the practical distributions of the ion (Figure 11-6) from theoretical.
12. How does annealing influence ion distribution (Figure 11-7)?
13. Compare the admixture distributions during ion and diffusion alloying.
14. What is the technique for the ion alloying process?
15. List the basic parameters of the ion injection process.
16. How are the properties of ion-alloyed layers controlled? Compare with the diffusion layers.

FOR OFFICIAL USE ONLY

17. What initial materials are used as sources of admixtures during ion injection?
18. Compare the possibilities of ion and diffusion alloying.
19. How is the process of ion alloying localized?
20. What are the limitations of the method of ion alloying?
21. What is the method of diffusion from the ion-alloyed layers?
22. What is the role of local radiation of the plate by electrons during radiation-stimulated diffusion?
23. Compile a production flow chart for obtaining p-n junctions on a silicon plate using radiation-stimulated diffusion.
24. What is the physical essence, what are the advantages and prospects for microdiffusion in an electron beam? How are the depths of microdiffusion and the p-n junction area controlled? What is the admixture distribution in a microalloyed p-n junction?
25. How does the Schottky barrier differ from the p-n junction and what are its characteristic features?
26. What are the basic steps in the process of obtaining a Schottky barrier and how are they carried out?
27. What are the basic areas of development of Schottky barrier technology in semiconductor IC?

FOR OFFICIAL USE ONLY

CHAPTER 12. METAL COATING OF SILICON STRUCTURES

12-1. General Information

In the production of semiconductor microcircuits the term "metallization" means the application of ohmic contacts to active regions, conducting tracks for intracircuit, interelement connections and also contact areas for connecting the structure to the leads of the case or to the passive part of the hybrid IC.

Ohmic contacts are mandatory for any active elements; they provide for transmission of an electric current in a semiconductor and must not influence the properties and the parameters, that is, they must not be interference in the physical processes occurring in the active elements.

The following basic requirements are imposed on the ohmic contacts: the contact resistance must not depend on the direction of the current, that is, the effect of current rectification at the contact must be absent; the contact resistance must not depend on the current flowing or the applied voltage. The volt-ampere characteristic of an ideal ohmic contact must be strictly limited, that is, the relation between the current through the contact and the external voltage must be subject to Ohm's law in the entire section; the contact must have the minimum possible resistance; the thermal conductivity of the contact must be high in order to provide for the removal and dissipation of the released heat; the contact must be mechanically strong; the thermal coefficient of linear expansion of silicon and the material of the ohmic contact must be as close as possible.

In order that the contact be nonrectified, it is necessary to suppress injection from the contact. The noninjecting contact must not generate current carriers which are minority for the semiconductor, for this promotes impoverishment of the surface layer, and, consequently, the rectification effect. On the contrary, it is desirable to form an antibarrier layer for which in the surface layer the concentration of the free carriers of the semiconductor will be large and the layer thickness will be small, that is, the probability of tunneling will appear and the rectification effect will be eliminated.

According to Schottky theory in order to eliminate the rectification effect and, consequently, in order to obtain a contact with ideal volt-ampere characteristic it is necessary to select a metal with a work function that is less than for the n-type semiconductor or greater than for the p-type semiconductor. In practice this is correct in rare cases, for example, for cuprous and selenous oxide. Accordingly, for the majority of other semiconductors primary attention is given

FOR OFFICIAL USE ONLY

not so much to the choice of metal as to the treatment of the semiconductor surface, the process of obtaining the contact and the application of special production procedures.

The formation of the antibarrier layer is promoted by the following: an increase in the recombination rate of the majority carriers in the surface layer, for example, introduction of an admixture (gold) with large capture cross section; introduction of an admixture into the electrode material which determines the type of conductivity of the semiconductor; the creation of a strongly alloyed contact region of the semiconductor, for example, for p-type silicon the donor concentration near the ohmic contact with aluminum must be no less than $5 \cdot 10^{19} \text{ cm}^{-3}$. The latter method is more effective and the most frequently used. For creation of n^+ or p^+ -surface layers in n and p-type conductivity material, diffusion, ion or other methods of alloying are used.

The requirements on the metallurgical and other physical-chemical properties of the contacts in many cases are contradictory, and therefore it is more complicated to satisfy them than the requirements on the electrical parameters. For example, a contact with a semiconductor must be strong, and at the same time the electrode material must not penetrate deeply into the semiconductor, for the dimensions of the active regions in the vertical direction can be fractions of a micron.

The conducting track and contact areas must satisfy the following basic requirements: high electrical conductivity; good adhesion to the masking film; absence of chemical interactions in the metallization layers, at the interface with silicon with the formation of harmful compounds or at the interface with silicon dioxide; resistance to migration of the metal; absence of corrosion; high mechanical strength; technological nature of the methods of obtaining metallization and the methods of connecting the external leads; absence of influence on the bound charge in the dielectric; reliability in the manufacturing processes and during operation of the microcircuits.

It is expedient to carry out metallization for a specific microcircuit from one material and, corresponding, in one production cycle. Therefore one of the basic problems when developing the metallization technology is the choice of material which will satisfy the maximum number of the entire set of requirements imposed on metallization. The complexity of simultaneous satisfaction of a large number of requirements is the cause of failures of the microcircuits as a result of metallization effects, the cause of a large number of materials used and metallization systems.

All of the metallization systems used at the present time in the production of silicon microcircuits or mounted active elements can be divided into the following types: single-layer, multilayer, multilevel, bulk (bulk leads).

12-2. Single-Layer Aluminum Coating

Advantages of Aluminum Coating. Aluminum, to a higher degree than other metals, corresponds to the requirements imposed on metallization, and it is the basic material which is deposited in a single layer. Aluminum has high electrical conductivity, good adhesion to the silicon and silicon dioxide, low-ohmic contact

FOR OFFICIAL USE ONLY

with p and n⁺-type silicon. It has a sufficiently good technological nature for the application of thin films and obtaining the metallization pattern. Important advantages of aluminum also include low cost, ductility and therefore resistance to cyclic temperature variations, the absence of galvanic effects, technological nature when manufacturing capacitors in compatible IC. Microcircuits of aluminum coating are resistant to radiation.

The metallization is done by applying a film of aluminum to the plate, obtaining the given configuration and heat treatment.

Application of the Aluminum Film. After obtaining structures in the body of the silicon plate, a mask is created on its surface from the silicon dioxide film with openings corresponding to the arrangement of the contact regions with silicon. Then an aluminum film is applied to the entire surface.

The aluminum films are applied, as a rule, by thermovacuum deposition on a large lot of silicon plates in periodic or semicontinuous-action devices of the UVN (vacuum deposition devices) series. The melting point of the aluminum at 660°C, the evaporation temperature (at an aluminum vapor pressure of 1.3 Pa) is 996°C. On the surface of aluminum there is always an oxide film that is difficult to evaporate; therefore heating to a temperature of 1200 to 1400°C takes place gradually in order to prevent splashing and spattering of the metal through the cracks formed in the oxide film. Slow heating is more favorable for conversion of the oxide on interaction with the aluminum vapor to the easily volatile compound Al₂O.

Molten aluminum wets the majority of materials from which evaporators are made W, Ta, and Mo, well. It forms volatile alloys with them; therefore the deposited film is contaminated, and the evaporators can be destroyed. The required ratio between the weight of the weighed sample and the wetting area of the evaporator is selected so that the evaporator will not be destroyed. An increase in wetting area is achieved by using the structural designs of wire tungsten evaporators made of a set of tightly arranged wires, V or W-type evaporators made of wire 1.5 mm in diameter with an additional coil of fine wire. On interaction with molten aluminum the crucibles made of aluminum nitride AlN or titanium boride TiB₂ are less active. In order to decrease the film contamination the first evaporated lot with the oxides available on the surface of the aluminum sample and the contamination settles on a slide valve, and the last lot is left in the evaporator together with the contaminated alloys. The deposition of the aluminum using resistive evaporators is done in the Soviet UVN-71P-1 and UVN-71P-2 devices.

In order to decrease the alloy formation during the process of wetting the evaporator, instantaneous evaporation of the aluminum by discharge of small doses into a highly heated crucible is used. The basic deficiency of this method is complexity of realizing the precision batching discharge (bin) device and the possibility of splashing of the aluminum.

It is not possible completely to exclude contamination of the film by the material of the heated evaporators. Therefore recently the application of aluminum films has been carried out using electron beam evaporation. As a result of local melting by an electron beam the aluminum sample does not interact with the crucible material; therefore the evaporated aluminum is in practice uncontaminated.

FOR OFFICIAL USE ONLY

This increases the time of fail-safe operation of the aluminum metallization. There are a large number of variations of electron beam evaporators distinguished by power, diameter and angle of inclination of the beam. Thus, for example, in the UVN-73P-1 device an electron beam evaporator with ring cathode is used. Aluminum is evaporated from a powerful water-cooled copper crucible from the surface of the aluminum ingot or rod. With this autocrucible evaporation the liquid aluminum interacts only with solid aluminum, the purity of the film is determined by the purity of the initial material and partial pressure of the gas. For application of especially pure films, highly pure, type AV000 aluminum and oil-free pumping units are used which insure a degree of vacuum in the deposition process of $1.3 \cdot (10^{-5} \text{ to } 10^{-6}) \text{ Pa}$.

The aluminum evaporation conditions are selected so as to insure good adhesion of the metal coating, ordered structure of the film, minimum electric charge in the film and also minimum mechanical stresses of the Si-SiO₂-Al system. The larger the grain size and the more improved the crystal structure of the film, the greater its electrical conductivity and the less the electrodiffusion. The electrodiffusion is the phenomenon of displacement of the metal ions in the direction of the electron conduction flux, that is, in the direction of the positive potential observed for high densities of the electric current. This displacement is explained by the fact that the energy transferred by the conduction electrodes to the excited aluminum ion located at the node of the crystal lattice predominates over the energy of the electric field applied to the ends of the wire. The aluminum ions are removed by the electron flux and form "whiskers" and hills on the positive end of the wire. Near the negative end of the wire cavities form. Electrodiffusion is the most probable along the grain boundaries, that is, the larger the grain size, the less the electrodiffusion.

During electron beam evaporation the charged-particle concentration in the atomic flux of the aluminum vapor and, consequently, in the deposited film increases with an increase in the emission current of the cathode and the deposition rate.

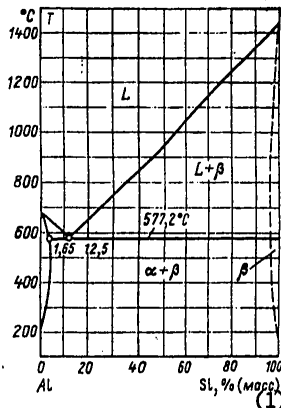


Figure 12-1. Diagram of the state of the Al-Si system

Key:
1. mass

FOR OFFICIAL USE ONLY

The temperature growth of the plates increases the film adhesion and the grain size. However, the temperature must not exceed 577°C, that is, the eutectic point of Al-Si (Figure 12-1), for diffusion of the surface of the plates can occur. In addition, at temperatures of more than 500°C, the aluminum reacts intensely with silicon dioxide. During the process of depositing the aluminum film the plates are usually heated to 200-400°C.

Increasing the evaporation rate, on the one hand, lowers the contamination of it with residual gases, and on the other hand, decreases the grain size, and during electron beam heating increases the charge density of the film. A deposition rate of 200-300 Å/sec is recommended for depositing aluminum.

The aluminum films are easily oxidized; therefore they must be cooled well before being taken out of the vacuum chamber.

Obtaining the Given Metallization Configuration. In the film deposited on the entire surface of the plate, the excess sections are removed using photolithography. It is recommended that FP-327 photoresists dissolved in methylethylketone be used, for the wetting angle of aluminum by this photoresist is less than 1°. A 5% solution of orthophosphoric acid or a mixture of it with acetic and nitric acid is used to pickle the aluminum. In order to decrease the side pickling under the mask and to increase the precision of the circuit layout in the metal coating, a sublayer of tungsten or aluminum 500-1000 angstroms is used. It is impossible to use acid to remove the photoresistive mask. The photoresist is removed in hot dimethylformamide. It is possible also to use the method of plasmochemical destruction to remove the photoresist.

Heat Treatment. In order to obtain a low-resistance stable contact and improve the adhesion of the Al to the silicon plate and to the silicon dioxide film, heat treatment in an inert gas atmosphere, for example, argon is used for several minutes at a temperature of 500-550°C. This temperature is below the hardening temperature of the Si-Al system; therefore the liquid phase is not formed. During heat treatment the silicon is dissolved in the aluminum and diffusion of silicon along the aluminum tracks takes place. The amount of silicon dissolved in the aluminum depends on the temperature (Figure 12-2), time, crystallographic orientation and type of conductivity of the silicon plate. The variation in depth of penetration of the Si into the Al with variation of the heat treatment conditions or with variation of the properties of the silicon plate leads to variation of the electrical properties of the contact. The greatest nonlinearity of the volt-ampere characteristics is observed for n-Si, for acceptor properties of the aluminum are manifested. If the concentration of the admixture in the n-Si is less than the maximum solubility of aluminum in silicon ($5 \cdot 10^{18} \text{ cm}^{-3}$), the electrical conductivity of the silicon changes to positive, and the p-n junction is obtained instead of the ohmic contact.

The adhesion of the aluminum film to the plate increases with an increase in the annealing temperature to 500°C; then with a further increase in the temperature the adhesion remains constant. The solution of the silicon in aluminum at too high temperatures leads to rupture of the aluminum film in the steps of the silicon dioxide film near the contact sites and manifestation of mechanical stresses as a result of difference in the thermal coefficient of layer expansion of aluminum ($\alpha_{Al} = 23.6 \cdot 10^{-6} \text{ deg}^{-1}$) and silicon ($\alpha_{Si} = 3.3 \cdot 10^{-6} \text{ deg}^{-1}$). The mechanical stresses promote the formation and growth of hills as a result of aluminum trans-

FOR OFFICIAL USE ONLY

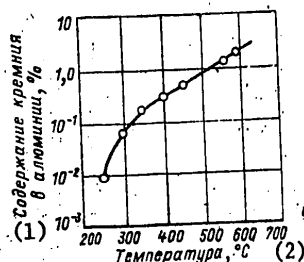


Figure 12-2. Solubility of silicon in aluminum (in the silicon phase)

Key:

1. silicon content in aluminum, %
2. temperature °C

In order to improve the reliability of the aluminum metallization and in order to obtain little-penetrating contacts, which is important to prevent shorting of the fine emitter junctions, about 2% silicon is introduced into the aluminum when depositing film. The metallization by Al-Si alloy can be obtained by evaporation of the aluminum and silicon from two separate electron beam evaporators. If we evaporate the weighed sample of the alloy, the film will contain very little silicon, for the vapor pressure of the silicon is much less than the vapor pressure of the aluminum.

Deficiencies of Aluminum Metallization. Aluminum metallization has a number of significant disadvantages:

Low Al-Si eutectic point (577°C);

High solubility of the silicon in aluminum in the solid phase (about 1% at T=500°C), when cooling the silicon is deposited along the grain boundaries and decreases the mechanical strength of the contact;

Low mechanical strength as a result of softness of the aluminum;

High thermal coefficient of linear expansion of aluminum by comparison with the thermal coefficient of linear expansion of silicon and silicon dioxide;

Chemical interaction with silicon dioxide which begins at 450°C and proceeds intensely at 500°C: $3SiO_2 + 4Al \rightarrow 2Al_2O_3 + 3Si$;

Significant electrodiffusion which appears for current densities of $5 \cdot 10^4$ amps/cm² and temperatures of 150°C, that is, under actual operating conditions of the microcircuits of medium and high power;

Low stability with respect to acids, bases and inclination toward corrosion;

The formation of brittle intermetallic compounds with gold Au₂Al and AuAl₂ at increased temperatures — a phenomenon known as "purple plague." On formation

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

of these compounds the aluminum diffuses into the gold faster than the gold into the aluminum; the cavities and cracks that are formed weaken the point of contact of the gold wire to the aluminum contact site still more, and the intermetallic couplings of the gold to the aluminum dissolve the silicon;

Difficulty of connecting leads to the aluminum film by ordinary soldering methods and the possibility of electrolytic deposition of aluminum.

Aluminum metallization is most advantageous for microcircuits operating on frequencies to 1 gigahertz which are not too powerful and are not designed for high reliability requirements.

12-3. Multilayer Metallization

Aluminum is the only material on the basis of which single-layer metallization of silicon microcircuits or components is carried out. However, the aluminum deficiencies limit the regions of application of the metallization. Therefore in many cases multilayer metallization is used, each layer of which performs an entirely defined function.

The contact layer -- the first with respect to order of application to the silicon plate -- must be made of metal which penetrates little into the silicon, does not enter into undesirable interactions with it, having low solubility and diffusion coefficient in silicon and at the same time good adhesion both to silicon and silicon dioxide, low transient resistance and the capacity for reduction of the oxide films.

Tungsten, molybdenum, chromium, nickel, aluminum, titanium and palladium can be used for the contact layer.

Tungsten has in practice identical thermal coefficient of linear expansion with silicon, good ohmic contact to the n^+ and p-silicon, significantly (by twofold) greater specific resistance by comparison with aluminum, high hardness excluding the scratching phenomenon, chemical inertia with respect to water and in an aqueous solution of hydrofluoric acid, high eutectic point with silicon (2165°C), absence of diffusion into the silicon and the highest energy of activation during self-diffusion of all metals. The tungsten films are primarily made by deposition from the vapor gas phase. The basic complexity in the process is the creation of the metallization pattern. It is impossible to connect the leads directly to the tungsten; therefore it is necessary to apply other methods to the contact sites over the tungsten.

Molybdenum, just as tungsten, provides for obtaining good ohmic contact with silicon of both types of conductivity, it has high electrical conductivity, it does not interact with aluminum, gold, silver, it pickles well during the photolithography process and corresponds to the majority of presented requirements on the contact layer. The molybdenum films are obtained by electron beam deposition, cathode sputtering or deposition from the vapor gas phase with subsequent burning in at a temperature of about 500°C. The deficiencies of the molybdenum films include their comparatively high porosity and also chemical activity leading to corrosion.

FOR OFFICIAL USE ONLY

Chromium is distinguished by very good adhesion to SiO₂ film. Just as aluminum, chromium actively reduces SiO₂. The chromium films can be applied by thermovacuum deposition and sputtering of the target. The chromium films have comparatively high mechanical stresses; therefore the thickness usually does not exceed 100 angstroms.

Nickel frequently is used as a result of its technological nature. It is easily applied from solutions and electrolytes. It can be applied by thermovacuum deposition, cathode sputtering, decomposition of carbonyl Ni(CO)₄ vapor at a temperature of 270-310°C.

With respect to properties and structure of the lattice, titanium has much in common with silicon and is frequently used as the contact sublayer. Just as other refractory metals, titanium has high adhesion to silicon, silicon dioxide and the admixture-silicate glasses. Titanium does not react chemically with silicon, it does not form intermetallic compounds. At a temperature of 400 and 500°C titanium interacts with the masking film:



The titanium dioxide film has low resistance, and therefore the total resistance of the contact is low. As a result of rapid oxidation in the air it is difficult to connect the lead to the film. Titanium films are obtained by thermovacuum deposition at a pressure no more than $1.3 \cdot 10^{-5}$ Pa or ion sputtering.

The conducting layer -- the last metallization layer with respect to order of application -- must have good electrical conductivity and provide a high-quality, reliable connection of the contact sites to the case leads. In cases where the conducting layer does not enter into undesirable interaction with the contact layer and has good adhesion to it, two-layer metallization systems are used. Copper, aluminum and gold are used for the conducting layers. Of the two-layer systems, the most stable and reliable metallization systems are nickel-silver, titanium-silver, molybdenum-gold, tungsten-aluminum, tungsten-gold, titanium-chromium, titanium-gold, chromium-gold. Let us consider some of these systems.

The titanium-gold metallization system is widely used for the metallization of high-frequency transistors and microcircuits. Gold has high electrical conductivity, it is resistant to corrosion, and it has little sensitivity to electrodiffusion. The gold films can be applied by thermovacuum deposition, cathode sputtering, and electrochemical deposition. During direct application to the substrate, the gold films form an alloy with silicon and have poor adhesion to silicon dioxide. Therefore the titanium film is applied first, and then the gold film. Two-layer Ti-Au metallization is distinguished by high reliability and good service life at increased temperatures.

The molybdenum-gold and tungsten-gold metallization systems are used for metallization of LSI. As a result of the refractory metals the metallization systems are more reliable during heat treatment. They are relatively stable to a temperature of the gold-silicon eutectic (370°C). However, prolonged holding at higher temperatures can lead to disastrous failures as a result of diffusion of the gold through the molybdenum into the silicon. By comparison with aluminum, these metallization systems are distinguished by higher resistance to scratching. The deficiencies of these systems include the greater cost by comparison with aluminum

FOR OFFICIAL USE ONLY

metallization, insufficiently high adhesion of molybdenum and tungsten to silicon dioxide and proneness to corrosion.

The two-layer systems are superior to aluminum metallization with respect to maximum current density and operating time per failure under analogous operating conditions. However, it is quite difficult to select the two-layer metallization system, for many two-layer systems are unstable under temperature variations.

The separating (barrier) layer is used in cases where it is complicated to select well-matched materials of the contact and connecting layers. The barrier layer permits the formation of intermetallic compounds between the upper and lower layers, for example, between the chromium and gold, titanium and gold, chromium and copper, and also prevents diffusion of the metal of one layer into the other, leading to worsening of the mechanical strength and a change in systems of the contact. Silver, gold, molybdenum, tungsten, nickel and palladium are used as the barrier layer. The latter has improved properties.

Let us consider examples of the metallization systems with more than three layers.

The chromium-silver-gold metallization system is used for microcircuits based on bipolar and MDS-transistors. The chromium provides good ohmic contact and high adhesion to silicon. The silver layer serves as a barrier to prevent interaction of the gold with the silicon and silicon dioxide and also for the formation of intermetallic compounds of chromium with gold increasing the contact resistance. Both of the upper layers (silver and gold) have high electrical conductivity, which permits an increase in the current density through the contact. The silver and gold are chemically low-active. Chromium and silver do not interact with silicon up to temperatures of 600°C.

The tantalum-gold-tantalum metallization system is distinguished by good adhesion to the masking film and to the silicon, higher stability with respect to corrosion and electrodiffusion and the possibility of operating at higher current densities than aluminum metallization. Tantalum prevents interaction of gold with silicon.

The chromium-titanium-platinum-gold metallization system provides good reproducible ohmic contact with the silicon and is designed for comparatively high current densities. The titanium layer has good adhesion to chromium and platinum and serves as a strong coupling. The platinum layer is used to improve the adhesion of the gold film to titanium and also as a barrier against the erosion of the gold coating. Gold has high electrical conductivity and provides good quality connection of the external leads.

In addition to the presented systems, Mo-Au-Mo, Mo-Cu-Mo, PtSi-Ti-Pt-Ti, PtSi-Pt-Au, PtSi-Mo-Au, PtSi-W-Au and other metallization systems are used.

12-4. Multilevel Metal Coating

Multilevel metallization systems are used to make LSI. Increasing the functional complexity of the microcircuits is connected with increasing the number of elements and interelement connections. In order to decrease the area occupied by the metallization, it is done on several levels separated from each other by insulating layers.

FOR OFFICIAL USE ONLY

The metallic layers of multilevel systems can be single or multilayer and must correspond to the same requirements as the layers of the single-level metallization systems. Especially high requirements are imposed on the absence of protruding films on the metal films, for they can lead to short circuits of the layers of different levels through the insulating dielectric.

The insulating dielectric films must have high-breakdown voltage, low dielectric constant and low losses, minimum chemical interaction with adjacent films, low level of mechanical stresses, low density of the bound electric charge, high chemical stability and technological nature when obtaining the films and creating the configurations. The presence of through microholes is inadmissible, for they can lead to short circuiting between the metallization levels.

Many series produced LSI are built on the basis of aluminum metallization with insulating layers of SiO_2 and Al_2O_3 .

Multilevel metallization with silicon dioxide as the insulating layer (Al- SiO_2 -Al) is well compatible with the manufacturing technology of semiconducting structures, but it does not insure the service life required at the present time for microcircuit as a result of electrodiffusion and ruptures of the metal coatings occurring primarily at these steps. The percentage yield of usable microcircuits is comparatively low as a result of short circuits between the aluminum layers of different, for during the silicon dioxide deposition process from the vapor-gas phase a large number of mounds appear on the aluminum film, and during deposition of the quartz, the insulating film has a large number of through microholes. However, the Al- SiO_2 -Al metallization is distinguished by simplicity; therefore the work on improvement of its quality is continuing.

Al- Al_2O_3 -Al metallization is simple to manufacture. The Al_2O_3 films are distinguished from the SiO_2 by high density, better insulating properties; they are a barrier for penetration of sodium ions, they have low sensitivity to atmospheric effects and water vapor, and they have higher radiation resistance.

The diagram of the manufacture of the two-level metallization system Al- Al_2O_3 -Al is presented in Figure 12-3. The aluminum film is deposited on a silicon plate with openings in the silicon dioxide film for the ohmic contacts. Then a photoresistive mask is formed over this film. The plate with the protective mask is placed on the anode of the electrolytic oxidation unit, and the unprotected sections of the aluminum are converted to oxide to the entire depth. Then the photoresistive mask is removed, and before deposition of the second aluminum layer, the surface of the plate is cleaned by ion pickling. After deposition of the second layer of aluminum, the process is repeated for the formation of a second metallization level. The sections of the anodized aluminum Al_2O_3 insulate the first and second levels of metallization, and the levels make contact at the required places. Then it is possible by the same process to obtain the third and subsequent metallization layers. The maximum temperature of obtaining metallization does not exceed 500°C .

In contrast to the Al- SiO_2 -Al system, in the Al- Al_2O_3 -Al system, the dielectric insulation is obtained at room temperature. The method permits us to obtain a smooth Al surface without protrusions and depressions. The insulating layer of Al_2O_3 , in contrast to SiO_2 does not have through holes. As a result of better

FOR OFFICIAL USE ONLY

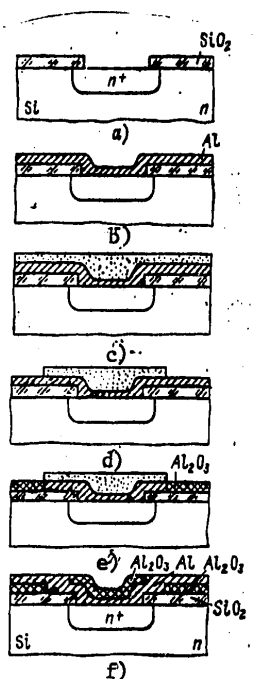


Figure 12-3. Diagram of the manufacture of two-level metallization Al-Al₂O₃-Al.

a -- photolithography for opening the holes for the ohmic contact;
 b -- deposition of aluminum; c, d -- formation of the photoresistive mask; e -- electrolytic oxidation of aluminum; f -- repetition of the process to obtain a second metallization level

dielectric properties and also the absence of protruding mounds it is possible to create a metal coating with less total thickness. Metallization is more reliable, the intersections of the conducting tracks and the connections of the levels are easily made.

Metallization with interlevel insulation made of polyimide is prospective for LSI with high packing density of the elements. After thermovacuum deposition of the first metallization level and obtaining a pattern in it by photolithography, a layer of polyimide is applied to the surface of the substrate. Then heat treatment is carried out for polymerization of it. Openings are made in the polyimide film by ion pickling. In order to obtain subsequent levels of metallization, the process is repeated. As a result of the flowability of the polyimide, it is possible to obtain stepless multilevel couplings and an ideal flat surface of the finished structure. The polyimide insulation is distinguished by very good physical, chemical and electrical properties which are maintained in the air in the temperature range from 260 to 420°C. Above 420°C, release of volatile components begins, and at temperature of about 485°C, destruction of the films takes place. In a vacuum the polyimide is stable to 500°C. By comparison with the silicon dioxide films, polyimide films are more resistant to moisture effects;

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

they have high adhesion to aluminum and other metals which are maintained after heating and cooling cycles and testing for thermal impact. The density of the through microholes in the polyimide films is comparable to the silicon dioxide films. The cracking of the polyimide films, in contrast to the silicon dioxide films is not observed. The polyimide films are distinguished by high radiation resistance and permit complete elimination of the failures of the microcircuits as a result of disturbances of the metal layers on the steps and intersections. At the present time five-level metallization systems based on polyimide insulation have already been obtained.

On manufacture of aluminum multilevel metallization it is difficult to obtain a good contact between the aluminum layers of adjacent levels, for aluminum has good affinity with the oxygen and in practice is always coated by an insulating oxide film. Therefore along with aluminum metallization, other systems are also widely used. The most difficult problem -- creation of a uniform continuous metallic coating -- is connected with the fact that the surface of the structure has a larger number of steps which are formed: next to the through openings, at the intersections of the metal tracks, on coincidence of the boundaries of the tracks on the first and second levels. The angle of inclination of the sides of the steps can vary from obtuse to acute. Therefore in a number of cases special production methods are used for smoothing the relief.

12-5. Metal Coating of Mounted Active Elements

When metallizing mounted active elements used for hybrid IC, consideration is given to the necessity for reliable connection of them to the contact sites of the microcircuit structures. The mounting elements are made with expanded contacts and also with bulk leads.

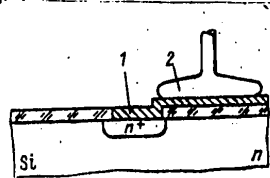


Figure 12-4. Expanded contact.
1 -- contact; 2 -- lead

The expanded contacts (Figure 12-4), in contrast to the ordinary ones, reach the surface of the masking film and at the same time permit separation of the point of contact of the lead from the active region of the mounted element. The expanded contacts are used for small areas of active elements and for shallow emitters when there is a probability of through penetration when connecting the lead.

The bulk leads simultaneously play the role of contacts to the active regions and the role of leads from the contacts, that is, the contact and lead are a united structural whole. Therefore such contacts are called bulk leads. The bulk leads can be made in the form of balls, protrusions (columns), bars (brackets). The bulk leads protrude above the common surface of the plate, the bar leads, in contrast to the balls and the protrusions, go beyond the boundaries of the crystal area.

FOR OFFICIAL USE ONLY

The bulk leads, just as the expanded contacts, are made by group methods in the planar process. In order to obtain bulk leads, basically deposition of metals from electrolytes and thermovacuum deposition are used.

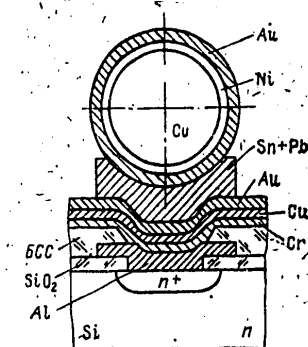


Figure 12-5. Ball bulk leads

The ball bulk leads were developed first. After formation of the aluminum contacts the plate is coated with a layer of borosilicate glass 1.5 microns thick (Figure 12-5), in which by using photolithography through openings are made over the contact sites. Then a multilayer chromium-copper-gold-solder (tin-lead) system is applied over the aluminum through the free molybdenum mask. The plates are placed in the recesses in the holder and copper balls coated with a double layer of nickel and gold are placed on the prepared contact sites. The holder is loaded in the chamber of the device, where the balls are soldered to the contact sites on heating in a hydrogen atmosphere. The manufacture of microballs and the assembly of the holders present some basic complications.

The leads in the form of columns are obtained by thermovacuum deposition of a thick solid layer (10-15 microns) of lead material (basically aluminum) with subsequent photolithographic treatment. A basic complexity is the choice of the photoresist which will withstand prolonged pickling of the entire thickness of the lead material. When growing columns from electrolyte first a current conducting coating is applied to realize quick growth of the leads, then a protective mask is formed, and electrolytic deposition of the leads is used, the protective mask and the current conducting coating are removed, and if necessary, the leads are tinned. The greatest difficulty in the technological process comes from insuring reliable protection against the electrolyte. When growing leads from acid electrolytes it is possible to use shielding by a positive photoresist. The alkaline electrolytes destroy the positive photoresists; therefore when growing leads from alkaline electrolytes, a protective mask made of borosilicate glass about 1 micron thick is used. When depositing leads made of electrolytes, the quality of the protective mask and especially the absence of through openings in it are very important. It is recommended that positive-negative photolithography be used. Simultaneous use of negative and positive photoresists permits the creation of in practice a defect-free photomask, for the photoresists are developed separately, each in a layer of developer. The mechanical defects of the layers and the defects transferred from the photomasks are not covered; therefore through punctures are absent. Initially a layer of positive photoresist is applied, and a photoresistive layer is formed. Then a negative photoresist is applied, and all of the

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

operations of photolithography are again performed, matching the image with the pattern of the positive photoresist. The complexity of positive-negative photolithography is paid for by the high quality of the photoresistive mask providing for deposition of thick metal projections.

The ball and column leads are used when making transistors, diode matrices and monolithic microcircuits used as mounted elements for hybrid IC.

With respect to their technology the ball leads do not differ theoretically from the manufacture of projection leads. Figure 12-6 shows a gold bar lead to a multilayer metallization system of PtSi-Ti-Pt. The same lead can also be used with Al-Ti-Mo-Au metallization.

In the first case the platinum is deposited using an electron beam or it is deposited by the plasma spraying method. Platinum silicide is formed at a temperature of 500-700°C in a vacuum chamber or in an inert environment (argon or nitrogen). The excess platinum is removed by "aqua regia," which does not interact with platinum silicide. The titanium and platinum 0.1-0.15 micron and about 0.2 micron thick, respectively, are applied in a continuous vacuum cycle. The gold bar lead is deposited electrolytically.

In the Al-Ti-Mo-Au system the thickness of the Al layer is about 1-1.5 micron. After photolithography with respect to the aluminum layer and heat treatment, the SiO₂ layer 0.5-0.6 micron thick is applied. Openings are made in the silicon dioxide layer, and the next layers of metallization of titanium 0.1 micron thick, molybdenum 0.2 micron thick, and gold 0.15 micron thick are applied successively. The gold bar leads 10-15 microns thick are deposited from electrolytes.

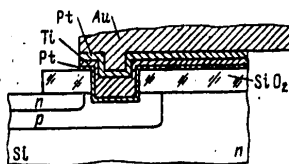


Figure 12-6. Bar lead

Along with the gold bar leads, leads made of aluminum 20-25 microns thick obtained by thermovacuum deposition with subsequent photolithography are widely used.

Devices with bar leads are the most prospective in multicrystalline LSI, in microwave microcircuits for frequencies of 10-40 gigahertz and in mounted diodes made of gallium arsenide with Schottky barriers. The bar design of the leads permits easy solution of the heat removal problem, and also reduction of the spurious impedances as a result of great width and constancy of the shape of the electrodes.

FOR OFFICIAL USE ONLY

12-6. Metallization Defects and Quality Control

The reliability of the operation of the microcircuits is determined to a significant degree by the metallization quality. Many failures of microcircuits are connected with metallization defects.

The basic types of failures of the microcircuits as a result of metallization defects are as follows: rupture of metallization on the steps, at the points of intersection of the conductors of different levels or at the points of interlevel contacts; short circuits of one or different metallization levels; ruptures in the regions of increased series resistance of the conductors.

The metallization ruptures can appear both as a result of the appearance of the properties of the metals and insufficient matching of them with the semiconductor and masking films. The low-quality performance of the photolithographic processes can be the cause of nonuniform transverse cross section of the current-conducting tracks. For multilevel metallization systems the causes of failures can be connected with a multilevel insulating layer, with the presence of mounds and other projections on the metallic films, with thermal interaction of the dielectric and the metal, and so on.

The basic mechanisms of rupture of the metallization: the formation of intermetallic compounds, corrosion, electrodiffusion and mechanical stresses.

In the technology of creating microcircuit metallization there are many problems, among which the primary ones are carefulness of preparing the surface, the selection of the materials for metallization and insulation, provision with high-quality photomasks, the selection of optimal methods and conditions of performing the process and also proper organization of the control operations.

When metallizing, the quality of the deposited metal film is controlled along with the strength of the coating, the thickness of the metal film, the specific tangent resistance, the volt-ampere characteristics of the ohmic contacts.

The quality of the deposited metal film is observed visually. The surface of the metal must be mirror reflecting, without spots, foreign particles and other defects distinguishable by the naked eye.

The strength of the coating is controlled directly after deposition by scraping the sharp pointed pincer at two or three points on the surface of the plate. The coating must not peel off.

The thickness of the metal film is measured by the method of multibeam interferometry. Before the measurement, part of the deposited film is removed by chemical pickling, and a step is obtained on the plate. The essence of the method consists in examining the interference pattern obtained on superposition of two coherent light beams reflected from the surfaces of the plate and the film. Alternating light and dark interference patterns with spacing L both on the surface of the film and on the surface of the plate are shifted with respect to each other near the step by an amount l (Figure 12-7). The shift depends on the height of the step, that is, the thickness of the metal film.

FOR OFFICIAL USE ONLY

Therefore the film thickness is calculated by the formula

(12-2)

where λ is the wave length of monochromatic light.

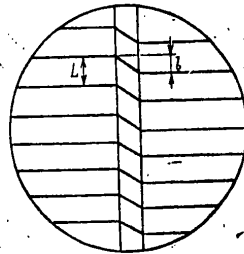


Figure 12-7. Variation of the film thickness with respect to shifting of the interference patterns on the step

The specific transient resistance of the ohmic contact ρ_k corresponds to the contact resistance of a unit area ($\rho_k = R_k S$), where R_k is the total resistance of the contact, and it does not depend on the contact area S . For measurement of the specific transient resistance, special test structures are used. One of them (Figure 12-8) contains five double contacts with the same width b , but with different lengths a , $2a$, $4a$ and different spacings between contacts ℓ , 2ℓ , 4ℓ . First the resistance R of the first three pairs of contacts 1, 2, 3 is measured. By the results of the measurements, the function $R=f(\ell)$ is plotted and, extrapolating to $\ell=0$, $2R_k$ is found, where R_k is the total resistance of the contacts (Figure 12-9). From the slope of the function $R=f(\ell)$, the surface resistance of the semiconductor $R_s = b \times dR/d\ell$ is found. Then the resistances (U/I) of the contacts 1, 4 and 5 are measured, and the total resistance of the contacts is calculated by the measurement results using the formula

$$R_k = \frac{1}{2} \left(\frac{U}{I} - \frac{R_s \ell}{b} \right) \quad (12-3)$$

If the values of R_k obtained for the structures 1, 4 and 5 coincide with each other, the specific resistance of the contact is calculated by the formula $\rho_k = b^2 R_k^2 / R_s$. If the contact resistances 1, 4 and 5 decrease with an increase in the contact length, then the specific resistances calculated by the formula $\rho_k = R_k S$, where $S = ab$, $2ab$, $4ab$ for the contacts 1, 4 and 5, respectively.

The volt-ampere characteristics of the ohmic contacts, just as the rectifying ones, are observed on the screen of the automatic volt-ampere characteristic recorder (Figure 12-10). When the sonde circuit is open, all of the voltage fed from the source falls on the horizontal deflecting plates, and a horizontal straight line is visible on the screen. When the sonde makes contact with the surface of the good ohmic contact, the basic voltage drop falls on R_H , from which it goes to the vertical deflecting plates, and a vertical straight line appears on the

FOR OFFICIAL USE ONLY

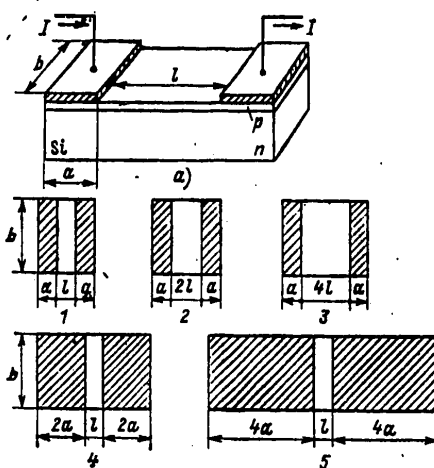


Figure 12-8. Diagram of the arrangement of contacts (a) and test structure (b) for measuring the specific transient resistance of ohmic contact.
1, 2, 3, 4, 5 -- pairs of contacts

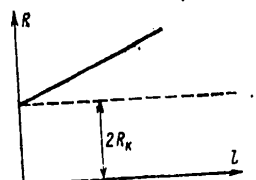


Figure 12-9. Specific transient resistance as a function of the distance between contacts
Key:
1. contact

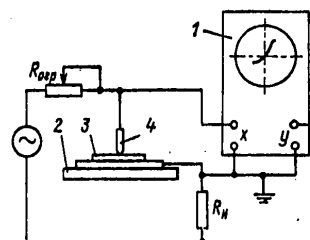


Figure 12-10. Automatic recorder for volt-ampere characteristics to record the static volt-ampere characteristics of contacts of the p-n junctions
1 -- oscillograph; 2 -- coordinate table; 3 -- checked sample; 4-- sonde
Key:
a. lim

screen. The volt-ampere characteristics of the rectifying contact have the form of a right angle, the location of which depends on the type of conductivity of the semiconductor.

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

Test Questions and Assignments

1. Repeat § 8-1.
2. What requirements are imposed on the ohmic contacts?
3. List production procedures for improving the quality of ohmic contacts.
4. What requirements are imposed on the conducting tracks and contact sites? What explains the large set of materials and metallization systems used?
5. List the basic advantages of aluminum coating.
6. What are the characteristic features of thermovacuum deposition of aluminum?
7. What is the technological process for obtaining an aluminum metallization pattern?
8. What is the purpose of heat treatment of metal coatings, and how is it done?
9. How is it possible to improve the quality of aluminum coatings?
10. List the deficiencies of aluminum coating.
11. What is the role of contact, conducting and barrier layers in multilayer multilayer metallization?
12. What requirements are imposed on metals for contact, conducting and barrier layers of metallization? Give examples of metals used for the corresponding layers.
13. What methods are used to obtain films of different metals for multilayer metallization? Give specific examples.
14. Give examples of the multilayer metallization systems used in practice. Explain the role of each layer and the process for obtaining it. Compare the specific multilayer metallization systems with aluminum metallization.
15. What are the difficulties of creating multilayer metallization systems?
16. What explains the application of multilevel metallization systems?
17. What requirements are imposed on the conducting part of multilayer metallization? The films of what materials are used for the conducting part of multilevel metallization systems?
18. What requirements are imposed on the insulating layers of multilevel metallization systems?
19. What systems of multilevel metallization are frequently used for LSI?
20. What is the process for obtaining Al-SiO₂-Al metallization; what are its advantages and disadvantages?

FOR OFFICIAL USE ONLY

21. What is the process for obtaining Al-Al₂O₃-Al metallization?
22. Compare Al-Al₂O₃-Al and Al-SiO₂-Al metallization.
23. What are the basic advantages of metallization with interlevel polyimide insulation?
24. How are mounted active components metal coated?
25. What is the process of metallization with ball leads? What is complicated about this process?
26. What is the process for obtaining protruding leads? What are the difficulties in Creating these Protrusions?
27. What are bar leads, what are their advantages?
28. List the basic types of metallization defects, their causes and mechanism of occurrence?
29. How is metallization film controlled for the presence of external defects, for coating strength?
30. How is the thickness of an aluminum film determined?
31. How is the transition resistance of a contact determined?
32. What is the volt-ampere characteristic method when determining the quality of an ohmic or rectifying contact?
33. Using Figure 12-10, explain how the volt-ampere characteristic of a good ohmic contact looks on the oscillograph screen? The p-n junction in n silicon? The p-n junction in p-silicon?

FOR OFFICIAL USE ONLY

CHAPTER 13. TECHNOLOGICAL PROCESSES OF MANUFACTURING THE STRUCTURES OF BIPOLAR MICROCIRCUITS

13-1. Engineering Design Features of Bipolar Microcircuits

Bipolar IC are made on monocrystalline silicon plates with a specific resistance from 0.5 ohm-cm to 100 ohms-cm. This type of semiconductor is a poor insulation; therefore the transistors and other elements of bipolar IC are made in individual regions -- pockets -- insulated from each other and from the common plate. All the contacts are taken out to one surface of the silicon plate. In the case of one-sided placement of the contacts, the path of the collector current is increased significantly. Therefore along with the problem of creating well-insulated pockets for the formation of microcircuit elements in them, we have the problem of decreasing the resistance to the collector current. The latter is necessary to decrease the saturation resistance of the collector and increase the speed of the transistor. Lowering the saturation resistance of the collector, it is necessary at the same time to maintain low capacitance and keep the breakdown voltage of the collector-base junction high. Accordingly, bipolar transistors, as a rule, are formed in n^+ -n pockets. In the part of the pocket adjacent to the surface with the contacts, there is a base-collector junction, and the strongly alloyed part of the pocket n^+ shunts the high-resistance part of the collector region.

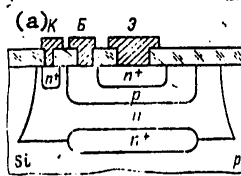


Figure 13-1. Structure of an integrated epitaxial-planar transistor with latent collector n^+ -layer.

Key: a. K = collector, B = base, E = emitter

The transistor is the basic structure of monolithic microcircuits (see Figure 13-1). All of the remaining elements -- the diodes, the resistors and capacitors -- can theoretically be made simultaneously with making the transistors.

Five different versions of diodes can be made on the basis of a bipolar transistor: the emitter-base diode with thread circuited collector-base junction (Figure 13-2,a);

FOR OFFICIAL USE ONLY

the emitter-based diode with open collector circuit (Figure 13-2,b); the collector-base diode with excluded emitter region (Figure 13-2,c); the emitter-base and collector-base junctions connected in parallel with short circuiting between the collector and the emitter (Figure 13-2,d); the base-collector diode with short-circuiting between the emitter and the base (Figure 13-2,e).

The diffusion resistors can be made by using any of the transistor regions: emitter, base, collector. For this purpose it is sufficient to create ohmic contacts for the corresponding regions. Most frequently in order to obtain resistors emitter and base regions are used. The resistors based on the emitter regions have low resistance (2-3 ohms/square). The resistance of the resistors based on the base regions is 100-300 ohms/square.

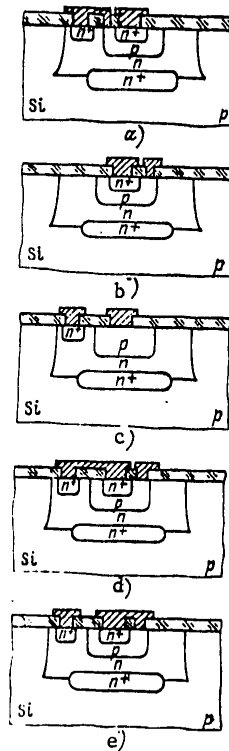


Figure 13-2. Five versions of diodes based on a monolithic transistor.

The capacitors in monolithic bipolar microcircuits are made on the basis of the barrier capacitance of the p-n junctions of the emitter-base, collector-base and collector-substrate.

The basis for the production of bipolar IC is planar technology. The standard planar technology includes a multiple sequence of the following operations: growth of a masking film of silicon dioxide on the surface of the silicon plate, photolithography to obtain the SiO₂-mask, local thermal diffusion through the openings in the mask.

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

In the overwhelming majority of microcircuits that are built, a silicon dioxide film 0.4-1 micron thick is used for masking. On the silicon plates, the oxide film is obtained by thermal oxidation; on germanium, gallium arsenide and other plates, the oxide film is obtained by high frequency ion-plasma sputtering of quartz or deposition from the vapor-gas phase (§ 8-2, 8-4).

During thermal oxidation of silicon in dry oxygen, a dense high-quality layer of oxide is obtained, but its growth rate is too small, and it is difficult to obtain a film of sufficient thickness. In wet oxygen the thermal oxide grows quickly, but the film is porous and does not correspond to the masking requirements. Therefore in practice in planar technology for masking the thermal oxidation of silicon takes place under combined conditions: first a thin, dense high-quality layer is grown in dry oxygen; then, a thick, looser, lower-quality layer in wet oxygen; then the process is repeated in dry oxygen. The high-quality oxide layers at the interface with the silicon surface and on the surface of the plate insure stability of the surface properties of the silicon and protection of the plate from external effects, respectively.

The silicon dioxide film masks only against boron during diffusion of the acceptor admixtures. Aluminum, indium and gallium have a high diffusion coefficient in the oxide film; therefore in those cases where it is desirable to carry out alloying by these admixtures to create the t-type diffusion regions, silicon nitride films are used as the masking films. When creating local diffusion n-regions phosphorus is used more frequently than other admixtures, for it has a larger diffusion coefficient than antimony and arsenic, and it has a greater maximum solubility in silicon than antimony.

As a rule, the donors diffuse more slowly and are dissolved better in the silicon than the acceptors. Therefore silicon n-p-n bipolar transistors are more widespread than p-n-p transistors. For creating the emitter regions, phosphorus is used, as a result of the high solubility of which high concentration of the admixture is insured (about 10^{21} cm^{-3}), and, consequently, a high emitter injection coefficient is obtained. Simultaneously with the emitter regions, n⁺-regions are obtained on the surface of the collector body, which decrease the resistance of the aluminum or other ohmic contacts.

Planar diffusion structures have a number of deficiencies. Thus, for example, a transistor with diffusion collector has nonuniform distribution of the admixture, the concentration of which decreases from the surface of the substrate inward. Accordingly, the series resistance of the collector is high. When obtaining the diffusion base it is necessary to introduce a large quantity of admixture in order to overcompensate for the admixture introduced during diffusion of the collector. This decreases the breakdown voltage of the base-collector junction.

Improvement of the parameters of the discrete semiconductor devices and microcircuits is achieved by improving the planar technology, replacement of thermal diffusion by new alloying methods, inclusion of active elements of Schottky barriers in the composition, the introduction of new masking materials, the application of deep local pickling of the silicon and, finally, what is very important, the introduction of epitaxial growing of semiconductor layers in the technological process. All of this not only improves the parameters, but also increases their reproducibility, increases the percentage yield of usable microcircuits and lowers their

FOR OFFICIAL USE ONLY

cost. Thus, the replacement of diffusion by epitaxy makes it possible to obtain collector regions with given distribution of the admixture and with quite wide range of specific resistances. The replacement of the SiO_2 -mask (or insulating layer) by Si_3N_4 or by SiO_2 - Si_3N_4 -mask (or an insulating layer) permits the use of a large set of alloying admixtures (boron, gallium, indium, aluminum, phosphorus, antimony, arsenic) during local diffusion with less thickness (0.1-0.15 micron) of the masking film, and it also makes it possible to obtain IC with more stable parameters, for Si_3N_4 has better dielectric and passivating properties. The application of deep anisotropic local pickling of the silicon plates oriented in the (100) plane, increases the degree of integration of the microcircuits, for it makes it possible to obtain insulated pockets surrounded by narrow V-grooves.

These and many other technological solutions have led to the development of a number of standard process cycles for the production of microcircuits. The majority of them include the epitaxy process, that is, basically epitaxial planar technology is used to make the bipolar microcircuits. One of the basic problems of the production technology of bipolar microcircuits is the creation of reliable insulation with maximum density of the arrangement of the elements.

The methods of insulation of the elements are divided into three groups: insulation by a back-biased p-n junction, insulation by the dielectric, combined insulation using the p-n junction and dielectric.

13-2. Manufacturing Technology of the Structures of Bipolar IC with Insulation by the p-n Junction

General Information. Many production processes are known for the manufacture of bipolar IC in which the insulation of the elements provided by a back biased p-n junction. All of these processes are based on the application of insulating diffusion. We shall consider the two flow diagrams for the manufacture of the structures of bipolar microcircuits most widely used in industrial production, in which the insulation is created using separate diffusion to the entire depth of the epitaxial layers. As a result of separate diffusion, pockets surrounded by the p-n junction are obtained in a p-type substrate. During operation of the microcircuit, the largest negative potential is fed to the substrate; therefore the p-n junction is blocked and the pocket with the transistor or other IC elements is insulated from the substrate and from other pockets. The manufacturing technology of the IC structure using a p-n junction as insulation is distinguished by a high technological nature; the production processes have been well assimilated.

Technological Process of the Manufacture of Structures of Bipolar IC by the Method of Separate Diffusion. At the present time this method is most widespread and is realized using epitaxial structures with latent collector regions.

The diagram of the process is presented in Figure 13-3 and can be divided into two steps: the manufacture of the epitaxial structure with latent collector n^+ -regions and the manufacture of the bipolar IC on this structure.

The manufacture of an epitaxial structure with latent collector n^+ -regions begins with thermal oxidation of the silicon plate and formation of a mask of silicon dioxide on the surface. Then local diffusion of the donor admixture is carried out in order to obtain highly alloyed latent collector n^+ regions. The alloying

FOR OFFICIAL USE ONLY

admixture for the latent collector regions must have high solubility in silicon with shallow depth of diffusion. At the same time, in order that the boundary of the latent collector layer not shift during subsequent heat treatments, the admixture must have the minimum possible diffusion coefficient in silicon. The surface concentration of the latent collector layer must not be too high, for this increases the autoalloying with subsequent epitaxy and mechanical stresses caused by noncorrespondence of the atomic radii of the silicon and the diffusing admixture. The latter leads to the appearance of dislocations. Accordingly, for alloying a silicon plate in order to obtain a latent n^+ -layer, antimony and arsenic are used, which have smaller diffusion coefficients than phosphorus. However, when using arsenic in latent layers, a large number of macrodefects are observed in the form of erosion, nodules, accumulations, fines, light spots and microrelief called "orange peel." Therefore for creating high-voltage bipolar microcircuits the latent layers are alloyed predominately with antimony. For low-voltage bipolar microcircuits it is possible to use epitaxial structures with latent layers alloyed with arsenic.

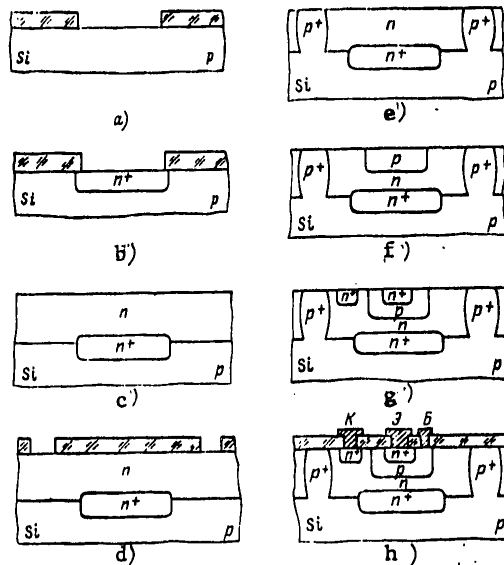


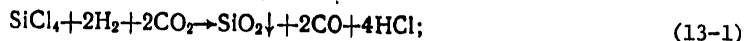
Figure 13-3. Diagram of the manufacture of an epitaxial-planar bipolar IC using separate diffusion. a -- thermal oxidation and first photolithography; b -- local diffusion; c -- epitaxy; d -- thermal oxidation and second photolithography; e -- separate diffusion; f -- formation of the base regions; g -- formation of emitter and contact regions of the collectors; h -- formation of the metal coating.

The antimony diffusion is carried out in a nitrogen or argon flux in a two-zone furnace. Antimonide glass in a quartz boat is placed in the first zone with a temperature of 650° C; silicon plates are placed in the second zone with a temperature of 1200° C. The antimonide glass is stable and has low toxicity. After the formation of the latent layers the oxide is removed and an epitaxial n-layer with a thickness of 12-15 microns and specific resistance of 0.1-10 ohm-cm is grown.

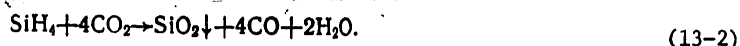
FOR OFFICIAL USE ONLY

In order to prevent contamination of the epitaxial structure during storage and transportation, silicon dioxide films are grown on its surface by methods that are compatible with epitaxial growing from the vapor-gas phase. The growth of the masking films is accompanied by chemical reactions:

with the chloride method of epitaxy



with the silane method of epitaxy



The silicon dioxide films are obtained at the same temperatures for which growth of the monocrystalline silicon epitaxial layers takes place. In the gas distribution system of the epitaxy unit there must be a carbon dioxide gas tube (Figure 9-2).

The manufacture of the structure of a bipolar IC includes the operation of thermal oxidation of the surface of the epitaxial structure, the process of photolithography for the formation of the SiO_2 -mask, diffusion of the acceptor admixture for the creation of the p^+ -regions separating the epitaxial layer and connected to the initial p-substrate. Thus, in the silicon substrate n^+ -pockets are obtained which are bounded by the p-n junctions. The microcircuit elements are formed in these pockets by the usual methods of planar technology.

For separate diffusion of boron, boron tribromide is used. The first diffusion (accumulation) phase is carried out at a temperature of 1150°C for 5-20 minutes. The redistribution of the boron to the depths of the epitaxial layer (dispersal) is carried out at $1200-1250^\circ\text{C}$ for 2-3 hours in a dry oxygen flux. Then the surface is oxidized, photolithography is used to make the openings for the base regions, and two-stage diffusion of the boron takes place. The approximate conditions are as follows: first step 950°C , 20-30 minutes; second step $1150-1200^\circ\text{C}$, 30-60 minutes. The diffusion distribution of the admixture in the base leads to the occurrence of an internal electric field that accelerates the movement of the minority carriers from the emitter to the collector and, as a result of this, increases the speed of the microcircuits.

Phosphorus is used as the alloying admixture for the formation of the emitter regions. The n^+ -regions for the ohmic contacts of the collectors are obtained simultaneously. The process for manufacturing the structures of the microcircuits ends with metallization. On the oxidized surface of the silicon openings are made for the ohmic contacts; the method of thermovacuum deposition is used to apply an aluminum film which is annealed, then a metallization pattern is created using photolithography.

On the modern level of development of microelectronics the possibilities of this method are limited as a result of a number of deficiencies: the presence of large leakage currents; a large area of insulating p-n junctions and, consequently, the stray coupling capacitance; comparatively low breakdown voltages; the presence of spurious p-n-p transistors between the base and the insulating p^+ -layers; low radiation resistance; low degree of integration as a result of the large area occupied by one microcircuit element.

FOR OFFICIAL USE ONLY

Technological Process for the Manufacture of the Structure of a Bipolar IC by the Method of Collector Insulating Diffusion. The flow diagram of the technological process is presented in Figure 13-4. As a result of replacing separate p-diffusion through the thick n-epitaxial layer by collector n⁺-diffusion through a thin p-epitaxial layer, the method permits a significant increase in the packing density of the elements.

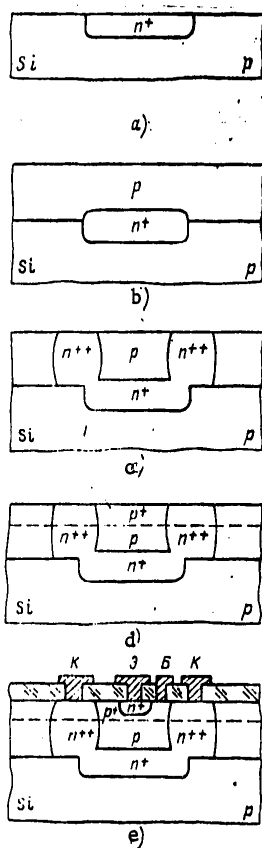


Figure 13-4. Flow diagram for the manufacture of bipolar IC using collector separate diffusion. a -- local diffusion; b -- epitaxy; c -- collector separate diffusion; d -- base p⁺-diffusion; e -- formation of the emitters and metalization.

In the initial silicon plates of p-type conductivity, n⁺-latent local regions are created, then an epitaxial layer of p-silicon, 1-2 microns thick, is grown. Then the plates are thermally oxidized and by using photolithography in the masking film, openings are obtained through which collector diffusion takes place until joining the latent layer. Thus, insulated n⁺-p-pockets are obtained. The base p⁺-diffusion is carried out without the application of the SiO₂-mask which simplifies the technological process. A layer with increased concentration of the

FOR OFFICIAL USE ONLY

acceptor admixture p^+ in the upper part of the base excludes the possible inversion of p-Si as a result of a positive electric charge in the silicon dioxide which attracts the electrons and repulses the holes of the surface layer of the silicon plate. The depth of the base p^+ -layer is 0.3-0.5 micron. Then the emitters and metallization are formed.

The method of collector insulating diffusion finds broad application, for it is distinguished by great simplicity by comparison with the preceding diagram (the number of photolithographies is decreased from 6 to 5). The collector region has high admixture concentration; therefore there is no necessity for special diffusion of gold or other admixtures that decrease the lifetime of the minority carriers when creating high-speed microcircuits.

The deficiencies of the microcircuits with collector insulating diffusion are as follows: the presence of an epitaxial base in which the drift of the carriers from the emitter to the collector is decreased, which leads to a decrease in speed of the microcircuits; the use of thin epitaxial layers limits the breakdown voltage of the collector-base (7-10 volts) as a result of propagation of volumetric charge in the base region.

The methods of insulation by back-biased p-n junction continue to be improved. In addition to the presented, widely used method of insulating separate and collector insulating diffusion, several methods have been developed, for example, the base insulating diffusion method, the method of three photomasks, the method of self-insulation using an n-cavity. All of them are distinguished by simplicity of technology, a decrease in area occupied by one element, and they can be used when making low-power microcircuits with moderate speed.

13-3. Manufacturing Technology of Structures of Bipolar IC with Dielectric Insulation

Advantages of Dielectric Insulation. By comparison with insulation by a back-biased p-n junction, dielectric insulation has significant advantages: significantly lower leakage currents, which in practice eliminates the problem of decoupling with respect to direct current; higher breakdown voltages; higher radiation resistance; higher operating frequencies in the linear microcircuits and higher speed in digital microcircuits; a smaller number of heat treatments, which decreases the undesirable effect on silicon, insures improvement of the parameters with sufficiently higher percentage yield of usable microcircuits.

Technological Process for Manufacturing the Structure of a Bipolar IC with the Application of Polycrystalline Silicon (the Öpik Process). The flow diagram for the manufacture of a microcircuit structure is presented in Figure 13-5. A highly alloyed n^+ -layer is grown on a silicon plate of n-type conductivity oriented in the (111) plane. The purpose of the n^+ -layer in the finished structure is analogous to the latent regions. Then by using thermal oxidation and photolithography, a mask of silicon dioxide is created on the surface of the n^+ -epitaxial layer. The mask is used for local pickling of the silicon to a depth of about 25 microns. After removal of the mask, a layer of silicon dioxide 1-1.5 microns thick is deposited on the entire surface of the plate from the vapor-gas phase. Deposition from the vapor-gas phase in the given case is preferable to thermal oxidation, for the large mechanical stresses occurring in thick thermal oxides lead to bending of the

FOR OFFICIAL USE ONLY

plates. Then a layer of high-resistance, polycrystalline silicon about 175 microns thick is grown on the surface of the silicon dioxide. Before growth of a polycrystal, the oxide surface is subjected to special chemical treatment for deformation of the centers of crystallization. In order to obtain insulated pockets the structure is ground off on the initial monocrystalline n-silicon plate side to complete removal of the n-"bridges" joining the pockets.

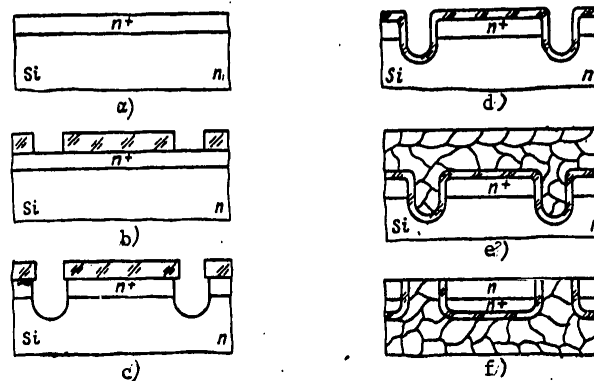


Figure 13-5. Diagram of the manufacture of bipolar IC with the application of polycrystalline silicon (the Opik process). a -- epitaxy; b -- thermal oxidation and photolithography; c -- local pickling; d -- removal of the SiO_2 -mask, deposition of the SiO_2 -film; e -- growth of polycrystalline silicon; f -- grinding from the initial plate side.

The pockets obtained in the polycrystalline silicon are insulated from each other by a layer of SiO_2 dielectric. In order to improve the insulation, sometimes the SiO_2 layer is replaced by a Si_3N_4 layer, double layers of SiO_2 - Si_3N_4 or SiO_2 - SiC . Silicon nitride and carbide have better dielectric properties, which decreases the leakage currents and the stray capacitances between the monocrystalline and polycrystalline silicon still more, and it increases the breakdown voltages. The silicon carbide is distinguished by high hardness, which facilitates mechanical removal of the excess n-type silicon, for it permits easy fixing of the grinding time to the level of the carbide layer. Transistors and other elements are formed in the pockets obtained.

Manufacture of the Structures of Complementary IC with the Application of Polycrystalline Silicon. It is possible to use insulation with the help of polycrystalline silicon when manufacturing complementary bipolar microcircuits, that is, microcircuits containing n-p-n and p-n-p transistors. Local diffusion of the donor admixture is carried out to the initial silicon plate with p-type conductivity (Figure 13-6); the grooves are pickled out locally around the n-regions. The SiO_2 film is grown on the relief surface obtained, and then the layer of high-resistance polycrystalline silicon. The excess part of the initial p-plate is removed by precision mechanical grinding, and as a result a structure is obtained containing n and p-pockets made of monocrystalline silicon insulated by dielectric and polycrystalline silicon. Then transistors and other elements are formed in these pockets by the usual methods of planar technology.

FOR OFFICIAL USE ONLY

The methods of insulation with the application of polycrystalline silicon are relatively complicated and tedious. The basic complexity consists in the necessity for precision machining. The deficiencies of the method also includes a comparatively low packing density of the microcircuit elements.

Manufacture of Structures of Bipolar IC with Glass, Pyroceram or Ceramic Insulation. In contrast to polycrystalline silicon, glass, pyroceram and ceramic have thermal coefficients of linear expansion close to that of silicon. When realizing the insulation, an auxiliary plate is used (Figure 13-7). In the n^+n epitaxial monocrystalline silicon structure microcircuit elements are formed, and then a silicon dioxide film is grown over the entire surface. After photolithography and local pickling, the mesa-regions with the circuit elements protruding above the surface are obtained. Then an auxiliary (technological) plate is glued on the mesa-region side, and the opposite side is ground off until separate elements are obtained held by the auxiliary plate. The billet obtained is pressed into glass, pyroceram or ceramic. Then the auxiliary plate is removed.

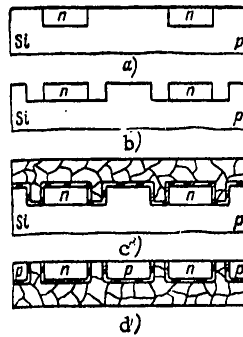


Figure 13-6. Diagram of the manufacture of complementary bipolar IC with the application of polycrystalline silicon. a -- local diffusion; b -- local pickling; c -- growth of SiO_2 and polycrystalline silicon; d -- grinding on the initial plate side.

Manufacture of Structures of Bipolar IC with Air Gap Insulation. The essence of the insulation consists in obtaining the semiconductor elements of IC separated from each other on the sides by air gaps and maintained on a united structural base. Let us consider three engineering design methods of finding solutions for this type of insulation.

The first method (Figure 13-8) consists in creating elements in the n^+n single-layer epitaxial structure that are not insulated from each other and which after gluing by liquid glass to the face of the plate of the dielectric base are separated by local pickling to the metallization layer. Before pickling, part of the epitaxial structure on the n^+ -region side is ground off. Thus, on the dielectric base, semiconductor mesa-regions are obtained with IC elements formed in them, connected by metallization and insulated from each other by air gaps. Then these mesa regions are covered with a protective dielectric film from the top.

FOR OFFICIAL USE ONLY

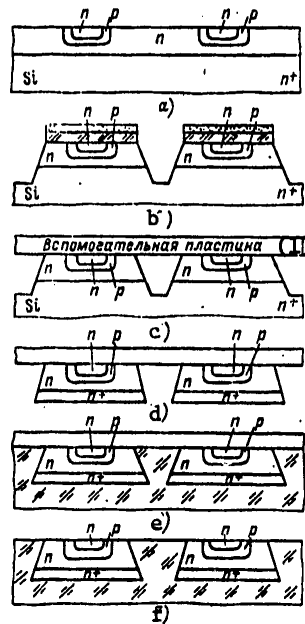


Figure 13-7. Manufacture of bipolar IC with insulation by glass, pyroceram or ceramic. a -- formation of elements; b -- obtaining the mesa; c -- gluing on the auxiliary plate; d -- grinding; e -- pressing of the elements into glass, pyroceram or ceramic; f -- removal of the auxiliary plate.

Key: 1. auxiliary plate

The lateral air insulation is somewhat inferior to solid insulation made of glass, pyroceram or ceramic with respect to breakdown voltages. The advantage of the given insulation by air gaps is the absence of high-temperature processes. However, precision grinding, exact matching of the photomask with the back side of the structure and complex deep pickling are required.

The second procedure usually called the bar lead method provides for a side air insulation in which the role of the supporting structural base is played not by a dielectric, but massive bar leads, which, in addition, electrically connect the microcircuit elements to each other (Figure 13-9).

Microcircuit elements are formed in a plate of monocrystalline n-Si. Then the plate is protected by a double dielectric layer of SiO_2 - Si_3N_4 through which photolithography is conducted to make the openings for the contacts. After creation of the platinum silicide contacts, layers of titanium and platinum are deposited, and gold is deposited electrolytically. The pattern of the bar leads is formed using photolithography. On the opposite side of the silicon plate the excess regions between the elements are removed by local pickling. The production procedure is also complicated.

FOR OFFICIAL USE ONLY

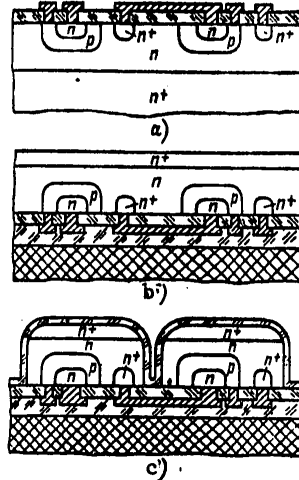


Figure 13-8. Manufacture of bipolar IC with insulation by air gaps. a -- formation of elements; b -- gluing to the dielectric plate; c -- separation of the elements by local pickling, application of a protective film.

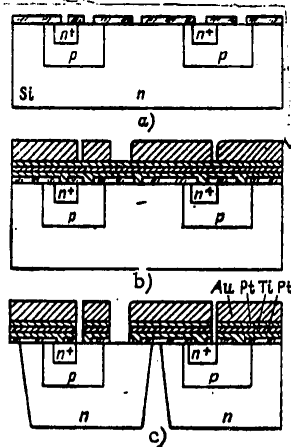


Figure 13-9. Manufacture of bipolar IC with insulation by air gaps using bar leads. a -- formation of elements; b -- manufacture of bar leads; c -- local pickling.

The third procedure -- the "silicon on sapphire (or spinel)" technology includes a smaller number of operations than the standard technological processes on silicon plates. The yield of usable microcircuits is also increased, for in contrast to silicon the sapphire substrate is not deformed during high-temperature treatment. In contrast to silicon, the sapphire and spinel substrates permit secondary treatment, that is, correction of a reject.

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

The procedure makes use of the technique of heteroepitaxy of silicon layers on sapphire or spinel substrates (see § 9-4). The silicon layers are deposited from the vapor-gas phase by decomposition of silicon monosilane. The chloride procedure is not used, for the hydrogen chloride released during reduction of silicon tetrachloride pickles the sapphire substrates. Then by using photolithography, local pickling is carried out. "Islets" of monocrystalline silicon remain on the substrate, in which bipolar transistors are formed. The base and emitter regions can be obtained by diffusion, ion alloying or preliminary epitaxy.

The air insulation between the lateral sides of the elements and practical absence of capacitive coupling of the elements through the substrate increases the frequency range of the microcircuits. The packing density of the elements of the microcircuits made on heteroepitaxial "silicon on sapphire" structures is higher than for the standard monolithic analogs.

The development of "silicon on spinel" technology is limited by the insufficient development of the process for obtaining the spinel substrates and the corresponding high cost of them.

The deficiencies of insulation using spinel and sapphire are of a temporary nature. The method is prospective for linear high frequency and microwave integrated microcircuits for which the basic requirement is the requirement of small leakage currents and high speed and also for optical integrated microcircuits. It is possible to grow layers of gallium arsenide and phosphide and other complex semiconductor compounds on sapphire and spinel.

The dielectric insulation has a more complex technology, greater area occupied by the elements (approximately 80% more than for insulation by the p-n junctions), but the manufacturing technology of the IC structures with dielectric insulation is being constantly improved, and as a result of the explicit advantages by comparison with insulation by the p-n junctions, is finding broader application in microcircuit production.

13-4. Manufacturing Technology of the Structures of Bipolar IC with Combined Insulation

Isoplanar Process. Combined insulation is realized using a number of methods in which the insulation at the base of the elements is insured by back-biased p-n junctions, and the side insulation, using a dielectric. Two versions of the isoplanar process are known.

The "Isoplanar-I" process (Figure 13-10) begins with growth of the silicon nitride film on the surface of the epitaxial structure with latent n^+ -collector regions. Then in the silicon nitride film photolithography is used to open up the openings through which local pickling of the epitaxial n-layer to a depth of about 60% of the total thickness is carried out. During the process of subsequent thermal oxidation the grooves are grown over by silicon dioxide (the oxide thickness must correspond to the thickness of the n-epitaxial layer). As a result of protection by the silicon nitride film, the entire surface of the plates is not oxidized, but only the side insulating dielectric layer bounding the n^+ -n pockets is formed.

FOR OFFICIAL USE ONLY

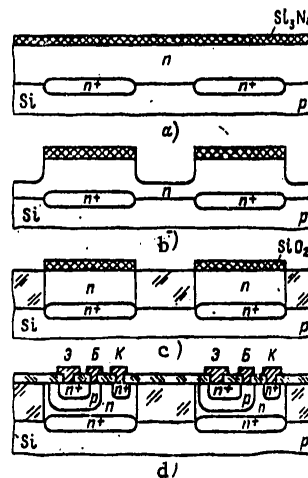


Figure 13-10. Diagram of the "Izoplanar-I" process of making bipolar IC circuits with combined insulation. a -- deposition of the silicon nitride film; b -- photolithography and local pickling; c -- local thermal oxidation; d -- removal of the mask and formation of elements.

After removal of the silicon nitride film, the surface is masked by the silicon dioxide film and then by ordinary methods the elements of the microcircuit are formed in the insulated pockets, and the metal coating is completed.

The "Izoplanar-II" process (Figure 13-11) is distinguished by the fact that in the obtained structure the emitter regions of the transistors have their side in the insulating layer, and the contact collector n^+ -regions are formed in independent pockets joined to the emitter-base regions by latent n^+ -regions. When manufacturing bipolar microcircuits by the "Izoplanar-II" technology the error in matching when obtaining emitter openings is not so significant as in the "Izoplanar-I" structure, for the opening is shifted to the oxide region, diffusion into which does not occur. As a result of a decrease in the area of the sidewalls of the emitter it is possible to obtain transistors with higher amplification coefficients. The base diffusion can take place throughout the entire pocket area, not forming the base openings, which simplifies the process.

Isoplanar technology combines the advantages of planar and mesa-technology; it makes it possible to avoid nonuniformities of the electric field on the periphery of the planar p-n junction, to reduce the stray capacitances between the active regions of the structure and improve the quality of insulation and packing density of the elements.

Manufacture of Structures of Bipolar IC using Vertical Anisotropic Pickling with the Application of Polycrystalline Silicon (the V-Method or the "Polyplanar" Method). The diagram of the process is presented in Figure 13-12. In the epitaxial structure with latent n^+ -regions made on a silicon plate of p-type conductivity with crystallographic orientation (100), local diffusion p-regions are created. Then the SiO_2 -mask is obtained, through the openings of which local processing is carried out in

FOR OFFICIAL USE ONLY

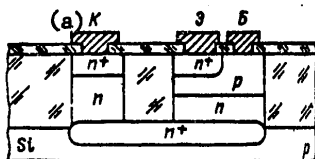


Figure 13-11. "Izoplanar-II" structure.

Key: a. K = collector
 E = emitter
 B = base

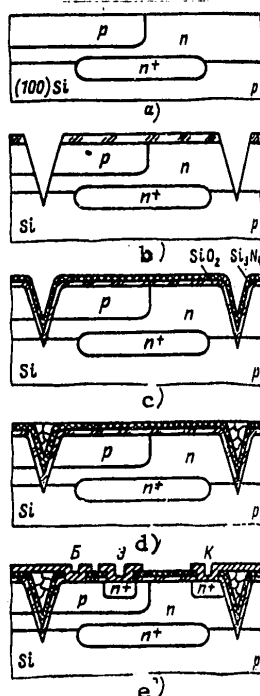


Figure 13-12. Manufacture of bipolar IC by the method of vertical anisotropic pickling with the application of polycrystalline silicon. a -- local diffusion of the acceptor admixture; b -- formation of the SiO₂-mask and local anisotropic pickling; c -- removal of the SiO₂-mask and protection by an SiO₂-Si₃N₄ layer; d -- growth of V-grooves by polycrystalline silicon; E -- formation of the IC elements.

hydrazine, ethylenediamine, caustic potassium or another anisotropic pickling agent. As a result of the great difference in pickling rates of silicon in the [111] and [100] directions, V-grooves are formed bounded by the walls oriented in the [111] planes. The process of obtaining the V-grooves is easily controlled, for the pickling is self-halting at the point of intersection of the sides of the grooves. If the pickling is halted early, then the bottom of the grooves will be flat and

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

oriented in the (100) plane. The area occupied by the V-grooves on the surface of the plates is comparatively small.

The relief surface obtained after pickling is protected by a double layer of SiO_2 - Si_3N_4 ; polycrystalline silicon is grown over it, the excess part of which is removed by polishing to the Si_3N_4 layer. The polishing process completes the formation of the pockets insulated from each other and from the substrate in which the bipolar transistors and other elements are made. Then the metal coating is formed. The given process makes it possible to increase the density of the arrangement of the elements, the radiation resistance and decrease the stray capacitances. The basic difficulty is precision polishing to remove the polycrystalline silicon.

Manufacture of the Structures of Bipolar IC by the Method of Vertical Anisotropic Pickling with Insulation by a Dielectric and Air Interlayer. In contrast to the preceding method, filling of the V-grooves with polycrystalline silicon is excluded (Figure 13-13). The surface of the V-grooves is coated by a triple dielectric layer SiO_2 - Si_3N_4 - SiO_2 . The basic complexity consists in creating reliable metallization with respect to the relief surface. Usually three-layer metallization Ti-Pt-Au is used.

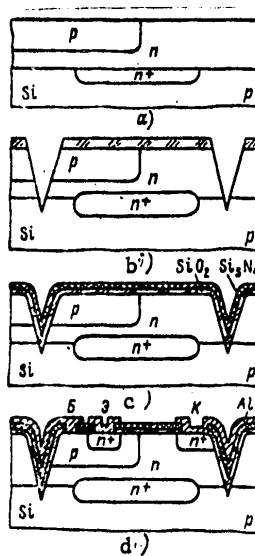


Figure 13-13. Diagram of the manufacture of bipolar IC by the method of vertical anisotropic pickling with insulation by the dielectric and the air interlayer. a -- local diffusion of the acceptor admixture; b -- local anisotropic pickling; c -- protection by SiO_2 - Si_3N_4 film; d -- formation of the IC elements.

By comparison with insulation by a p-n junction, the given technology permits a decrease in the area occupied by one element in the spacing between elements and also an increase in speed and percentage yield of the microcircuits.

FOR OFFICIAL USE ONLY

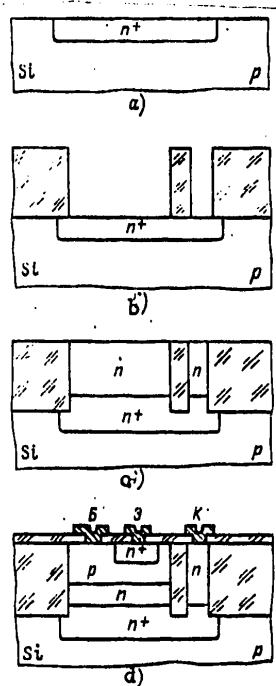


Figure 13-14. Diagram of the manufacture of bipolar IC using local epitaxy. a -- local diffusion; b -- pyrolytic deposition of silicon dioxide, photolithography; c -- local epitaxy; d -- formation of the IC elements.

Manufacture of the Structures of Bipolar IC by the Method of Local Epitaxy. Side insulation of the IC elements obtained by this technology is realized using a thick layer of silicon dioxide (Figure 13-14). A layer of silicon dioxide 1.5-2.5 microns thick is deposited pyrolytically on the surface of the p-type silicon plate of n^+ local regions. Openings are made in the oxide, and local epitaxial growth of the n-Si takes place, then transistors and other elements are formed in the n^+ -n-pockets obtained in this way. The manufacture of the IC structure ends with metal coating. The basic difficulty consists in the necessity for performing exact local epitaxy. The method is prospective for high frequency microcircuits.

None of the insulation methods can be considered universal or to solve all of the problems which arise during the manufacture of microcircuits. When selecting the method and material of the insulation it is necessary to consider specific functions performed by the system, the requirements on its parameters and operating conditions. The necessity for insulation of the elements of the bipolar microcircuits limits the density of the arrangement of the elements, the speed, and it also complicates their manufacturing technology. The total number of technological production operations is very large.

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

13-5. Manufacturing Technology of the Structures of Compatible IC

In contrast to the monolithic microcircuits in which both active and passive elements are made in the body of the semiconductor plate, in the compatible microcircuits the active elements are made in the semiconductor plate, and all or part of the passive elements are made on its surface in the form of films. The role of the insulation between the body active and surface passive elements is played by a dielectric film.

The combination of semiconductor and film technology makes it possible to combine the advantages of semiconductor active elements (by comparison with film elements) and film passive elements (by comparison with body elements) in one microcircuit. The film passive elements have a broader range of ratings and better operating characteristics. The body active elements have significantly better properties and permit a high degree of integration of the microcircuits to be obtained. However, the technological process of manufacturing the structures of compatible IC is more complicated and expensive.

Compatible technology is used to make micropowerful and superfast semiconductor integrated circuits.

Test Questions and Assignments

1. Repeat the physics of the operation of a bipolar transistor.
2. Repeat § 1-4.
3. What technological processes form the basis for planar technology and what properties of planar structures do they give rise to?
4. What are the engineering design peculiarities of bipolar microcircuits?
5. List the technological production methods of obtaining the SiO_2 masking film. Which of them is used most frequently and why?
6. Under what conditions is the masking film of SiO_2 grown and why?
7. What admixtures are used for local diffusion when making bipolar structures and why?
8. What is epitaxial-planar technology and what are the advantages of the corresponding IC?
9. Repeat the marking of epitaxial structures with latent n^+ -collector regions (§ 9-1).
10. Repeat the basic operations of contact photolithography (§ 5-2 to 5-4).
11. Compare the developed flow chart for the manufacture of an epitaxial n-p structure with latent n^+ -regions.
12. How is the protection of the epitaxial structures made when using the chloride and silane methods of epitaxy?

FOR OFFICIAL USE ONLY

13. Explain the technological process of manufacturing a bipolar IC by the method of separate diffusion (Figure 13-3).
14. What are the characteristic features of the process for making bipolar IC by the method of collector insulating diffusion (Figure 13-4)?
15. Compare the methods of separate and collector insulation diffusions.
16. What are the basic advantages of dielectric insulation of IC elements?
17. Compare the technological process for making bipolar IC with the application of a dielectric and polycrystalline silicon (the Opik process). Check the flow diagram that you compiled by Figure 13-5. Explain the basic features of this technology.
18. Compare the structures of transistors in the diagrams in Figures 13-5 and 13-3.
19. How is it possible to make a complementary IC with the application of a dielectric and polycrystalline silicon?
20. What is a mesa structure, how is it obtained and what is its role in the process of making IC with glass, pyroceram and ceramic insulation?
21. What methods are used to insulate by air gaps?
22. What are the advantages and disadvantages of each of the methods of air gap insulation?
23. What is combined insulation and what are its advantages?
24. What is the role of the Si_3N_4 -mask in the isoplanar process?
25. Compare the "Isoplanar-I" and "Isoplanar-II" structures (Figures 13-10, 13-11).
26. Explain the flow chart for the "polyplanar" process (Figure 13-12). What are the peculiarities and the disadvantages of this process?
27. What are the disadvantages of the IC structures with side insulation by a dielectric and V air gaps (Figure 13-13)?
28. Explain the peculiarities of the process of manufacturing bipolar IC by the local epitaxy method (Figure 13-14).
29. Give the general characteristic of the process of making bipolar microcircuits.
30. What is the compatible technology of making microcircuits; in what case is it used?

FOR OFFICIAL USE ONLY

CHAPTER 14. TECHNOLOGICAL PROCESSES OF MAKING MDS-IC STRUCTURES

14-1. Structural-Technological Features of MDS-IC

Characteristic Features of the Structural Design of MDS-IC. Metal-dielectric-semiconductor (MDS) transistors or, in the special case, metal-oxide-semiconductor (MOS) transistors are in the class of field devices, the operating principle of which is based on using the electric field effect. With respect to engineering design execution, they are made from transistors with builtin channels and transistors with induced channels. The builtin channel is created by technologically alloying the semiconductor, for example, by diffusion. The induced channel is induced in the surface layer of the semiconductor as a result of the effect of a transverse electric field. The MDS transistors with induced channels are most widely used in the IC composition, for they make it possible most simply to match the circuits with each other and realize universal logical functions.

The structure of the MDS-transistor with induced p-channel is shown in Figure 14-1. The source and discharge are made in the form of local, strongly alloyed p⁺-regions in a relatively high-resistance silicon plate with n-type conductivity. On feeding a voltage of any polarity between the source and the discharge, the p-n junctions of the source and discharge turn out to be included in opposite directions; therefore the source-discharge circuit is open. When feeding a negative potential to the controlling electrode-gap located over the thin layer of dielectric between the source and discharge, impoverishment with respect to the majority carriers takes place in the surface layer of the semiconductor, and with sufficient potential as a result of inversion, the conducting p-channel appears which connects the source and discharge regions. A further increase in negative potential on the gate enriches the channel with the majority carriers -- holes -- the conductivity between the source and the discharge increases. Thus, the current flowing through the channel between the source and the discharge can be modulated by the potential on the gate. In the analogous MDS-transistor on a p-plate with source and discharge regions of the n⁺-type conductivity on application of a positive potential to the gate, an n-channel is induced. The majority of microcircuits produced are made from transistors with induced p-channels and with a silicon dioxide film as the dielectric under the gate. The use of primarily p-channel transistors, in spite of their slower speed by comparison with the n-channels, is explained by the fact that in the process of thermal growing of the oxide for the gate on the surface of the p-silicon, a thin inversion n-layer arises. The distribution coefficient of boron at the Si-SiO₂ boundary is less than one, that is, the boron has less solubility in silicon and not in the oxide; therefore in the process of thermal oxidation it goes from the substrate to the growing oxide. The formation of the inversion n-layer is also promoted by the presence in the oxide of a positive charge.

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

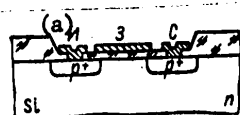


Figure 14-1. Structure of MDS-transistor with induced p-channel.

Key: a. I = source, Z = gate, S = discharge

The MDS-IC are distinguished by high noiseproofness. The input resistance of the MDS-transistor is determined by the resistance of the oxide, and therefore it is high. When working with digital equipment, large load resistances are needed; therefore the IC functions are made with a delay. The speed of the digital IC on bipolar transistors is therefore higher than on the MDS-transistors. In contrast to the bipolar transistors, the current in the MDS-transistors is transported only by the majority carriers. Accordingly, the MDS-IC operating as amplifiers have higher operating frequencies.

The MDS-IC are made from structures with channels of one type of conductivity and with channels of opposite types of conductivity -- with mutually complementing types of conductivity (complementary microcircuits -- CMDS-IC). The CMDS-IC, which are somewhat inferior to the better bipolar circuits with respect to speed, are superior to them with respect to intake power.

Characteristic Features of the Manufacturing Technology of MDS-IC Structures. The MDS-IC are made in the overwhelming majority of cases by planar technology. As a result of the simpler structural design and absence of a specially created insulation of the elements, the total number of operations in the production cycle decreases significantly by comparison with the bipolar microcircuits. This decreases the labor consumption, increases the percentage yield of usable microcircuits and decreases their cost. When making MDS-IC with channels of one type of conductivity, only one diffusion process is used to obtain the source and discharge regions at the same time as for bipolar IC it is necessary to create the collector, base and emitter regions and carry out insulating diffusion (4 diffusion processes). When manufacturing the CMDS-IC, the diffusion is carried out three times: to obtain the pockets, to create the source and discharge regions in the initial plate and to obtain the sources and discharges in the pockets.

The most responsible events in the manufacturing technology of MDS-IC are the creation of an oxide for the gate, matching of the gate to the channel and obtaining structures with short channel length.

The oxide for the gate is a component part of the active element of the IC, and more rigid requirements are imposed on it than the requirements on the masking films. The oxide must be electrically strong ($E \geq 10^6$ volts/cm) and have minimum number of charges and maximum stability.

The gate must be exactly matched to the channel. If the gate length is less than the channel length, the MDS-transistor will not operate, for the source-discharge circuit will be open. If the gate greatly overlaps the source and discharge regions, the speed of the IC will be diminished as a result of the stray capacitances.

FOR OFFICIAL USE ONLY

The channel length determines the drift time of the current carriers from the source to the discharge and consequently, the speed of the MDS-IC.

The peculiarities of MDS-IC production will be investigated in more detail when studying specific technological processes. The use of MDS structures permits microcircuits to be obtained with a high degree of integration, with high functional possibilities with simpler structural design, technology and lower cost than bipolar microcircuits.

14-2. Manufacturing Technology for Structures of Thin-Oxide p-Channel MDS-IC

The standard flow diagram for the manufacture of thin oxide MDS-IC with p-type channels is presented in Figure 14-2. On the comparatively high-resistance silicon plate (2.5-10 ohm-cm), a masking film about 0.6 microns thick is grown. The admixture concentration in the plate is selected low to decrease its influence on the inversion of the surface layer. The upper limit of the thickness of the masking oxide is limited by the resolution of photolithography when opening the windows in the source and discharge regions arranged in a row, and the lower bound, by the stray couplings between the metal coating and substrate. Thermal oxidation is carried out in the combined mode: dry, wet, dry oxygen. Then the usual contact method is used for photolithography. The protective photoresistive mask is removed, and the plates are transferred to the diffusion section. In order to obtain low-resistance source and discharge regions the boron diffusion is carried out two stages. The requirements on the precision of performing the diffusion process are less rigid than when making bipolar IC, for the MDS-IC are less sensitive to reproducibility of the surface resistance and depth of diffusion. The second photolithography is carried out to make the openings for the thin gate oxide.

The creation of the oxide for the gate, just as was noted, is a responsible operation. The quality of the oxide and the oxide-silicon interface have a strong influence on the operating stability of the IC, the threshold voltage and the mobility of the minority carriers of the current in the channel. On application of a potential to the gate under the effect of the transverse electric field, drift of the oxygen vacancies, ions, alkali metals and hydrogen through the oxide is possible. In order to decrease the admixtures in the oxide, sterile conditions of performing the operations of cleaning the silicon plates and growing the oxide are required. Special production procedures are used in practice to reduce the charge in the oxide and the density of the surface states at the oxide-silicon interface to a minimum. The creation of the oxide for the gates is in two phases: first the thin oxide is grown, then it is stabilized.

The thin oxide (0.1-0.15 micron) is grown by thermal oxidation in dry oxygen. The presence of moisture will lead to the presence of hydroxyl groups strongly bound to the oxide and adsorbed molecules. The OH⁻ ions lead to crystallization of the oxide, to a large number of vacancies, an increase in defects, ruptures of the chemical bonds, instability of the oxide-silicon boundary and growth of the positive charge in the oxide. Thermal oxidation is carried out at temperatures of 1150-1200° C. The holder with the plates is cooled quickly after oxidation. In the case of slow cooling crystallization of the oxide is possible leading to worsening of its stability and dielectric properties. For oxidation of the silicon in the dry oxygen, large mechanical stresses occur in the oxide film, for the removal of which annealing is carried out in a nitrogen atmosphere at a temperature of about 900° C.

FOR OFFICIAL USE ONLY

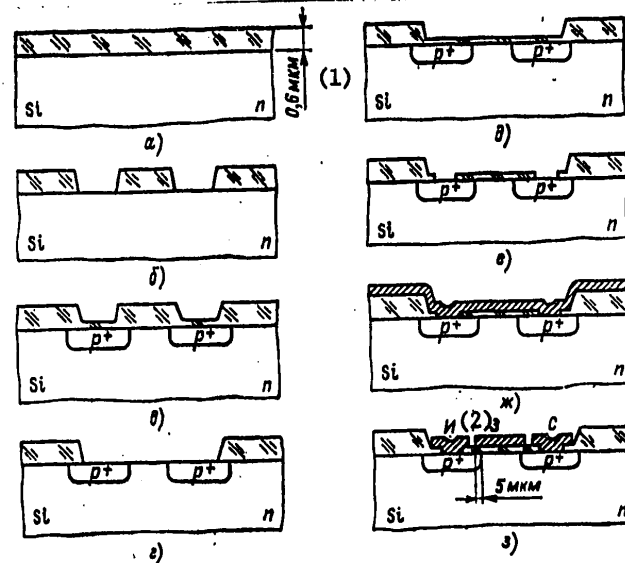


Figure 14-2. Diagram of the manufacture of thin-oxide MDS-IC with p-type channels. a -- growth of the masking oxide; b -- first photolithography; c -- diffusion of the source and discharge regions; d -- second lithography; e -- growth of the thin gate oxides; f -- third photolithography; g -- deposition of aluminum; h -- fourth photolithography.

Key: h. microns 2. I = source, Z = gate, S discharge

Additional low-temperature treatment at 450° C with wet nitrogen promotes a decrease in the density of the surface states as a result of filling them with the hydrogen ions penetrating the oxide to the surface of the silicon. The hydrogen ions are formed during dissociation of the water on the surface of the oxide. Low-temperature annealing can be carried after deposition of aluminum, for the latter is a catalyst for dissociation of the water.

The second step in the creation of the oxide for the gates -- stabilizing processing -- is carried by applying a thin film (100-200 angstroms) of phosphorus silicate glass to the surface of the thermally grown oxide. The glass is obtained when treating oxidized plates in P_2O_5 vapor at a temperature of 1000° C. In order to obtain the vapor usually a liquid source of $POCl_3$ or P_2O_5 is usually used. In the liquid states the gas absorbs sodium ions, and so on, from the oxide. In the solid state after cooling the silicon plates the glass prevents the penetration of hydrogen ions and admixtures from the gate into the oxide and also decomposition of the water on the surface and the formation of free vacancies in the oxide. However, it is necessary to consider that with an increase in the phosphorus concentration in the glass, the charge increases; therefore it is necessary to select the optimal stabilization conditions. It is possible to lower the phosphorus concentration in the glass by annealing in nitrogen at a temperature of 900° C. The glass with low phosphorus content is moisture proof and, in addition, it is distinguished by sufficiently high density, which insures a clearer pattern during photolithography with less side pickling distortion of the double layer of $SiO_2-nSiO_2 \cdot mP_2O_5$.

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

The third photolithography is carried out to open the windows for the ohmic contacts of the source and discharge regions. After removal of the photoresist, an aluminum film is applied to the entire surface of the plates which after the fourth photolithography plays the role of gates and metallization of the MDS-IC.

In the investigated technology of p-channel MDS-IC, the thickness of the oxide between the metal coating and the source and discharge regions in the final structure of the IC is comparable to the thickness of the oxide for the gates, and the thickness of the oxide on the n-Si is appreciably greater. In the presence of punctures in the thin oxide, short circuits are possible between the metallization and the p⁺-regions. Large stray capacitances between the metallization and the p⁺-regions have a negative effect on the characteristics of the MDS-IC. The difference in thicknesses of the oxide over the p⁺-regions and n-silicon complicates obtaining of a photoresistive mask and correspondingly lowers the quality of the metallization pattern. As a result of incomplete removal of the aluminum between the closely located metallization lines, shorting cross connections appear. To complement this, the regions of the channels and the gates are formed in practice during the first, third and fourth photolithography; therefore, in spite of not considering possible inaccuracies in reproduction of the photomask patterns, it is very difficult to obtain coverage of the channels by gates of less than 5 microns.

14-3. Manufacturing Technology of MTOS-IC

MTOS (metal-thick oxide-semiconductor) is the process providing for obtaining an oxide film of great thickness over the active p⁺-regions and over the surface of the initial n-silicon plate in the final structure of the microcircuit.

In MTOS-technology the processes for obtaining oxide masking against diffusion and oxide under the metallization of the source and discharge regions are separate. This decreases the capacitance between the metallization and the p⁺-regions and also the capacitance of the coverage of the channel by the gate and, in addition, facilitates photolithography with respect to the aluminum film. Accordingly, the thick oxide technology is predominant in MDS-IC production.

The basic steps in the MTOS-technology are presented in Figure 14-3. A masking film of about 0.2 microns thick is grown on the surface of the silicon. Then photolithography is carried out to obtain the SiO₂-mask required for local diffusion of the boron. The local diffusion and growth of the thick oxide can be carried out simultaneously or separately. In the second case the oxide is deposited from the vapor-gas phase. The thickness of the final oxide over the active regions of the sources and discharges is about 1 micron; over the initial surface of the n-silicon plate 1.5 microns. The second photolithography and subsequent pickling make it possible completely to eliminate the oxide from the surface over the channel regions. Then comes growth of the thin oxide for the gates. Then subsequent operations are carried out just as in ordinary thin-oxide technology of manufacturing p-channel MDS-IC: third photolithography for opening the windows of the source and discharge regions, thermovacuum deposition of the aluminum film, fourth lithography for obtaining the metallization pattern and gates. The most complicated in the MTOS-technology are performance of the photolithography to open the window for the thin oxide and subsequent pickling of the thick oxide. In order to protect the aluminum coating from scratches, a passivating film is deposited on the entire surface of the finished structure. Then windows are opened in this layer by using

FOR OFFICIAL USE ONLY

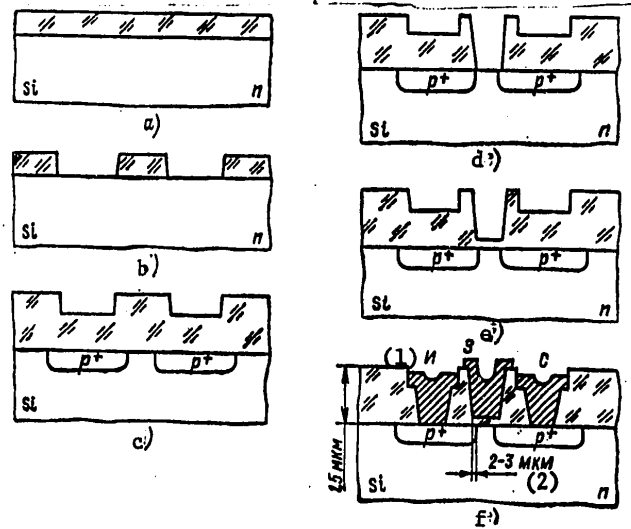


Figure 14-3. Diagram of the manufacture of MTOS-IC. a -- growing a masking oxide; b -- first photolithography; c -- local diffusion, growing a thick oxide; d -- second photolithography; e -- growing a thin gate oxide; f -- third photolithography, formation of gates and metallization.

Key: 1. I = source, Z = gate and S = discharge
2. microns

photolithography at the locations of the contact sites. As a result of the creation of a thick oxide for metallization the parameters of the microcircuits are improved significantly. However, the problem of obtaining the minimum coverage of the regions of the sources and discharges by the gates in MTOS-technology is not solved, for matching the channels to the gates here, just as in ordinary thin-oxide technology, is carried out using photolithography and it is complicated as a result of indeterminacy of the location of the lateral boundaries of the p-n junctions of the sources and discharges and also as a result of the necessity for pickling the thick oxide. With a gate oxide thickness of 0.1 micron in practice the coverage of the channel by the gate will be 2-3 microns.

14-4. Manufacturing Technology of the Structures of MDS-IC with Fixed Gates

The manufacturing technology of MDS-IC with fixed gates completely eliminates the probability of obtaining channels with length less than the distance between the p⁺-regions, and it significantly decreases the coverage by the gates of the source and discharge regions. At the present time three methods of fixing the gates are most frequently used: using shallow diffusion, by ion alloying and combined using diffusion and ion alloying.

The flow diagram for the manufacture of MDS-IC with fixed silicon gates using shallow diffusion of the source and discharge regions is presented in Figure 14-4. A film of thermal oxide is grown on the surface of the silicon plate, and the first photolithography is carried out to obtain the openings under the source, the

FOR OFFICIAL USE ONLY

discharge and the gate of each MDS-transistor. Then a thin (0.1 micron) oxide is grown under the gates, over which a layer of highly alloyed polycrystalline silicon about 0.5 microns thick is grown. During the process of second photolithography the windows are opened in the double layer of silicon dioxide and polycrystalline silicon for local diffusion of boron to obtain the p^+ regions of the sources and gates. The diffusion depth is 0.5-1 micron. The less depth of diffusion, the less propagation of boron in the lateral directions and, consequently, the less coverage of the sources and discharges by the gates. After diffusion, a silicon dioxide film is deposited on the entire surface of the silicon plate to increase the total thickness of the oxide. Then photolithography is carried out to open the windows under the sources, discharges and the contact area of the gates, an aluminum film is applied, and the microcircuit layout pattern is formed using the last photolithography.

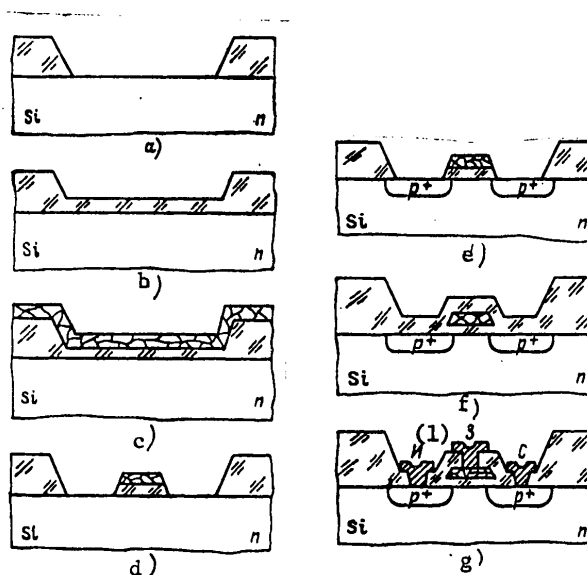


Figure 14-4. Diagram of the manufacture of MDS-IC with fixed silicon gates. a -- growing the masking oxide and first photolithography; b -- growing the gate oxide; c -- growing the polycrystalline silicon; d -- second photolithography; e -- diffusion of the source and discharge region; f -- deposition of oxide film; g -- third photolithography, formation of the metal coating.

Key: 1. I = source, Z = gate, S = discharge

The MDS-IC with silicon gates have high speed by comparison with ordinary microcircuits obtained using MTOS-technology. As a result of decreasing the tolerances on the matching, the size of one microcircuit element is decreased, and the degree of integration is increased by approximately 2.5 times. As a result of replacement of the aluminum gates by silicon, the threshold voltage of the transistors is decreased, for the work function of the polycrystalline silicon is less than for aluminum. The operating stability of the microcircuits increases, for the

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

polycrystalline silicon has a smaller amount of admixtures causing drift of the electric charge in the oxide and at the boundary with the silicon dam. When using polycrystalline silicon in the IC structure, less mechanical stress occurs. In addition, the application of polycrystalline silicon gates makes it possible to create both MDS and bipolar transistors on one plate, for after obtaining the gates it is possible to carry out the high-temperature processes of thermal oxidation, diffusion, and so on.

The most responsible events in the MDS-IC manufacturing technology with silicon gates are growth of the polycrystalline silicon and the photolithography process with respect to the double silicon dioxide and polycrystalline silicon film. During high-temperature growth of a polycrystal, deoxidation of the silicon dioxide takes place, which leads to breakdown of the oxide under the gates during operation of the microcircuits. With photolithography with respect to the double layer, pickling of the oxide takes place. These deficiencies are excluded when the silicon gate is replaced by a metal gate.

The flow diagram for the manufacture of the MDS-IC structure with metal gates using ion alloying of the source and discharge regions is analogous to the diagram in Figure 14-4. The metal corresponding to the technological requirements and having the smallest possible work function by comparison with silicon is selected for the gate: aluminum, vanadium, molybdenum, tungsten. Of the presented metals aluminum has the largest work function. Ion injection is carried out without opening the windows in the thin oxide. Subsequent annealing is mandatory to remove the radiation defects and for electric activation of the admixture. The ion energy is selected so that the gate electrodes and the thick oxide will play the role of masks during ion injection. During ion alloying, the penetration of the admixture in the lateral direction is in practice absent by comparison with diffusion alloying; therefore the coverage of the source and discharge regions by the gates is less than in microcircuits with silicon gates with source and discharge regions obtained by shallow diffusion. Ion alloying makes it possible to use metal masks, and photolithography with respect to the metals has been better developed than with respect to polycrystalline silicon. The thin oxide is not subject to high temperature effects during application of the gate film or during ion alloying, just as occurs when growing polycrystalline silicon and during diffusion.

The diffusion-ion process of manufacturing MDS-IC structures with fixed gates permits lower values of the resistances of the source and discharge regions to be obtained. The flow diagram for the manufacture of MDS-IC is to a great extent analogous to the ordinary MTOS technology, but it is distinguished by the fact that after the application of the metal film the gate electrodes are formed in such a way that their length will be less than the channel length (Figure 14-5, a). After forming the gates, the ions are injected from the same admixture as was used to alloy the diffusion p^+ -regions. The ion energies, just as in the preceding case, are selected sufficient only to penetrate the thin layer of oxide. Thus, ion-alloyed sections of the sources and discharges are obtained, as a result of which the channels are in practice ideally matched with the gates of the microcircuit.

The combined method of fixing the gates permits a significant decrease in channel length and, accordingly, an increase in the degree of integration, still higher speed of the digital and frequency limit of the amplifying microcircuits.

FOR OFFICIAL USE ONLY

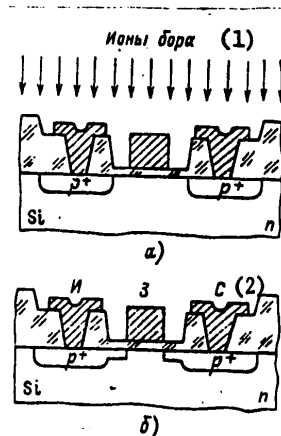


Figure 14-5. Diagram of the diffusion-ion MDS process. a -- MDS-structure after local diffusion; b -- MDS-structure after ion alloying.

Key: 1. boron ions 2. I = source, Z = gate, S = discharge

14-5. Manufacturing Technology of CMDS-IC Structures

The complementary microcircuits based on MDS-structures are distinguished by high noiseproofness, negligibly small static intake power and high speed. For manufacture of them in one plate it is necessary to create n and p-regions that are well insulated from each other with optimal specific resistance to insure good parameters of the MDS-transistors with channels of opposite types of conductivity.

Insurance of optimal specific resistance in the p-regions in which the n-channel transistors are executed is one of the responsible technological events in the creation of the CMDS-IC based on n-type silicon plates. In order to decrease the probability of inversion of the p-silicon at the surface as a result of redistribution of the boron when growing a silicon dioxide film and as a result of the effect of positive charge of the oxide it is necessary that the boron concentration in the p-region be as large as possible. On the other hand, increasing the concentration decreases the breakdown voltage of the source-discharge. The boron concentration of 10^{17} to $5 \cdot 10^{16}$ cm^{-3} corresponding to a specific resistance of the p-region of 0.4-0.8 ohm-cm is optimal.

The CMDS-IC can be manufactured by four methods: by creating p-pockets in the n-silicon using local diffusion, by the creation of p-pockets in the n-silicon using epitaxy, by assembling the MDS-structures made from individual semiconductor plates of different type of conductivity (bicrystalline MDS-IC) using the "silicon on sapphire (or spinel)" technology.

The manufacturing technology of CMDS-IC structures using diffusion is the simplest and well mastered. The standard flow chart is presented in Figure 14-6. After thermal oxidation of the plates the first photolithography is carried out for opening windows under the p-pocket region. Then the p-pockets are created in the boron diffusion process. The second lithography is carried out for opening the

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

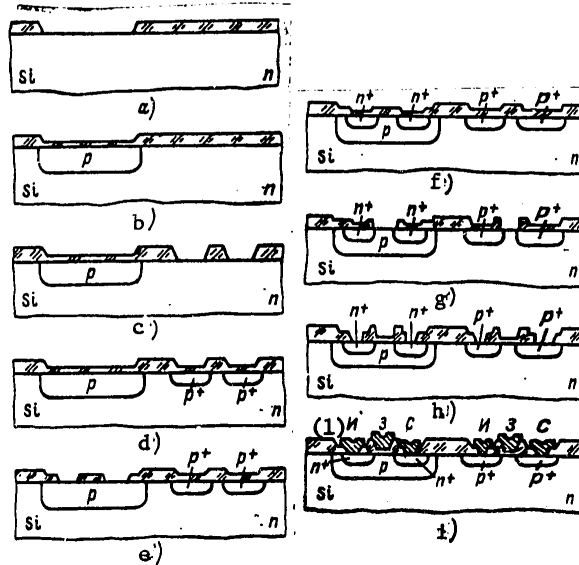


Figure 14-6. Diagram of the manufacture of CMDS-IC using diffusion. a -- thermal oxidation, first photolithography; b -- local diffusion; c -- second photolithography; d -- local diffusion; e -- third photolithography; f -- local diffusion; g -- fourth photolithography; h -- growing of a thin gate oxide and fifth photolithography; i -- forming the gates and metal coating.

Key: 1. I = source, Z = gate, S = discharge

windows under the sources and discharges of the p-channel transistors. After the second boron diffusion, third photolithography is carried out to obtain windows under diffusion of the source and discharge regions of the n-channel transistors. They are obtained using phosphorus diffusion. Then fourth photolithography is carried out under the thin gate oxide. A thin oxide is grown. Fifth photolithography is carried out to obtain windows over the source and discharge regions, the thermovacuum deposition of aluminum film and sixth photolithography to form the gates and metal coating.

By comparison with the process of manufacturing MDS-IC based on transistors with channels of one type of conductivity, here it is necessary to obtain diffusion pockets with low concentration of the acceptor admixture. When using silicon dioxide as the masking film, out of the acceptor admixtures it is possible to use only boron, and boron has high maximum solubility of 10^{21} cm^{-3} in silicon; therefore the creation of the p-pocket presents a defined problem. Boron diffusion is carried out in two stages: diffusion introduction of the boron is realized in a forevacuum, and redistribution of the boron in an oxygen atmosphere.

The manufacturing technology of CMDS-IC structures using epitaxy permits easy obtaining of any concentration distribution with respect to the pocket depth. In

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

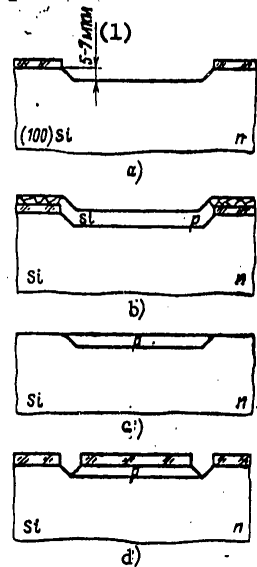


Figure 14-7. Formation of the p-pocket in an n-Si plate when manufacturing CMDS-IC using shallow anisotropy pickling and local epitaxy. a -- local anisotropic pickling; b -- epitaxial growth over the grooves; c -- removal of polycrystalline silicon; d -- local anisotropic pickling.

Key: 1. microns

order to obtain p-pockets in silicon plates of the n-type, equipment is used for local anisotropic pickling and subsequent local epitaxy. In contrast to the deep anisotropic pickling through narrow windows to obtain V-grooves in the planar-epitaxial process of manufacturing CMDS-IC used for combined insulation of the elements of bipolar microcircuits, anisotropic pickling is done to a shallow depth and through wide windows.

The composition of the pickling agent and the pickling conditions are selected so that the bottom of the groove will correspond exactly to the crystallographic plane (110). The diagram of obtaining a p-pocket is presented in Figure 14-7. After thermal oxidation, photolithography and local anisotropic liquid or dry pickling are carried out. During epitaxial growth over the grooves in the vapor-gas system of $\text{SiCl}_4\text{-H}_2\text{-B}_2\text{H}_6$ or in the $\text{SiH}_4\text{-HCl-H}_2\text{-B}_2\text{H}_6$ system it is expedient to use anisotropic gas pickling by hydrogen chloride, for the pickling and epitaxy processes are performed in a continuous cycle in one unit. The local epitaxy rate noticeably influences the filling profile of the grooves and the formation of the ridge protruding over the surface of the plates around the periphery of the window. A flatter plate surface is obtained after growth over the pockets during gas pickling of the grooves in HCl at a temperature of more than 1200° C and for epitaxy in the $\text{SiH}_4\text{-HCl-H}_2\text{-B}_2\text{H}_6$ system with a growth rate of 2-3 microns/min. After epitaxial growth the silicon surface is masked by the silicon dioxide, and local anisotropic pickling is carried out through narrow windows in the mask in order to isolate the p-pockets.

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

The epitaxial process of manufacturing CMDS-IC has not as yet been completely developed; therefore the cost of the microcircuits is still high.

Bicrystalline CMDS-IC: on the silicon plates of different type of conductivity MDS-structures with p-channels are made separately, and the MDS-structures with n-channels separately with respect to one of the processes investigated for single-channel microcircuits, but with the execution of ball leads on certain structures and corresponding contact sites on other structures. Then the wiring is done, that is, the ball leads are connected to the contact sites.

The "silicon on sapphire" CMDS-IC permit a three to fivefold increase in speed by comparison with monolithic microcircuits, which is basically connected with reducing the stray capacitances. The sapphire substrates provide high packing density of the elements and high radiation strength. These advantages are realized most completely when manufacturing LSI. Just as the monolithic MDS-IC, the microcircuits on sapphire substrates can be made with metal and silicon gates.

The flow chart for the manufacture of CMDS-IC structures with aluminum gates is presented in Figure 14-8. Silicon layers with (100) orientation are grown on the surface of the monocrystalline sapphire substrate with crystallographic orientation of (1102). Then a contact mask of silicon dioxide is formed on the surface of this layer, and local anisotropic pickling is carried out to complete separation of silicon into islets. The SiO_2 contact mask is removed, and a new mask is formed to perform local diffusion or ion injection to overcompensate for the admixture in the open silicon islets. After removal of the second SiO_2 -mask, silicon islets of p and n-type conductivity remain on the surface of the sapphire. Then the masking film of silicon dioxide is grown, and by means of photolithography the oxide is left only on the central parts of the silicon islets. A phosphorus-alloyed oxide (FSS) is applied to the islets of p-type conductivity. A boron alloyed oxide (BSS) is applied to the entire substrate surface. As a result of diffusion from the surface alloyed oxides, source and discharge regions of p and n channel transistors are obtained. The oxides are removed and then, just as in ordinary technology, the gate oxide and the oxide under the metal coating are obtained, the aluminum film is deposited, and the gates and contacts to the sources and discharges are formed.

In order to obtain the source and discharge regions, ion alloying or diffusion from the ion-alloyed layers are used.

The flow diagram for the manufacture of CMDS-IC structures with silicon gates is presented in Figure 14-9. Just as in the preceding case, the silicon islets of opposite type conductivity are formed on a sapphire substrate. Then a film of silicon dioxide and a layer of polycrystalline silicon are grown over the entire surface of the substrate. Using photolithography, the gate oxide and gate are formed in each islet. Then alloyed oxides are applied to the islets, from which diffusion of the phosphorus and boron take place, respectively, to obtain sources and discharges of the n and p-channel transistors. The alloyed oxides with the islets are removed, silicon dioxide is deposited over the entire surface of the substrate, photolithography is carried out, an aluminum film is deposited and a metal coating is formed using the last photolithography.

The "silicon on sapphire" technology for CMDS-IC with silicon gates permits us to obtain high accuracy of matching the transformers and shorter channel length.

FOR OFFICIAL USE ONLY

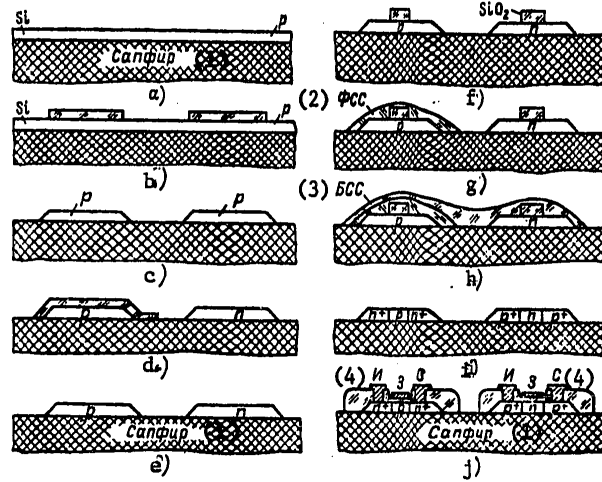


Figure 14-8. Diagram of the manufacture of CMDS-IC with aluminum gates on a sapphire substrate. a -- heteroepitaxy; b -- shaping the protective mask; c -- local anisotropic pickling; d -- masking and diffusion of phosphorus (or ion alloying); e -- removal of the protective mask; f -- masking; g -- application of FSS [phosphorus-alloyed oxide]; h -- application of BSS [boron-alloyed oxide]; i -- diffusion and removal of glass; j -- formation of the gate oxide, gates and metal coating.

- Key: 1. sapphire 2. FSS [phosphorus-alloyed oxide]
 3. BSS [boron-alloyed oxide]
 4. I = source, Z = gate, S = discharge

14-6. Ways of Improving the Quality of MDS-IC

The MDS-IC produced at the present time are superior to the bipolar IC with respect to degree of integration, specific intake power, noiseproofness, radiation resistance, and cost, but the majority of them are inferior to the bipolar IC without speed. Increasing the speed of the MDS-IC without increasing the intake power can be achieved by reducing the stray capacitances of the MDS-structures and decreasing the channel length and threshold voltage.

In the production methods widely used in practice at the present time, the parameters of the MDS-structures are improved by careful cleaning of the surface of the semiconductor plates before growing the gate oxide, optimization of the methods and the conditions of obtaining it, use of a thick oxide film metal coating, the application of the production methods of fixing the gates using a decrease in the overlap of the source and discharge regions by them, the application of especially pure materials for the gates, for example, silicon instead of aluminum, the execution of the microcircuits on silicon plates oriented in the (100) plane and also on sapphire substrates. In order to decrease the instability of the parameters of the MDS-structures under the effect of an electric field and temperature connected with the drift of the electric charge through the silicon dioxide layer, special processing -- stabilization of the oxide -- is needed. Reducing the positive charge

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

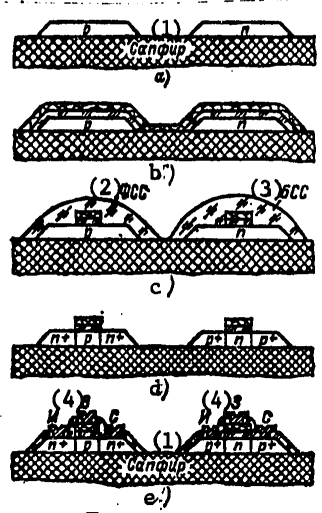


Figure 14-9. Diagram of the manufacture of CMDS-IC with silicon gates on a sapphire substrate. a -- obtaining the silicon islets; b -- thermal oxidation and growth of polycrystalline silicon; c -- photolithography, application of FSS [phosphorus-alloyed oxide], BSS [boron-alloyed oxide]; d -- diffusion and removal of glass, formation of the metal coating.

- Key:
- 1. sapphire
 - 2. FSS [phosphorus-alloyed oxide]
 - 3. BSS [boron-alloyed oxide]
 - 4. I = source, Z = gate, S = discharge

in the oxide increases the stability of the parameters and decreases the threshold voltage of the MDS-transistors. The processing of the plate surface before growing the gate oxide is very important. The application of hydrofluoric acid in the composition of the pickling agent leads to a sharp increase in porosity of the grown oxide. This is explained by the presence of admixtures in the oxygen. The mixtures adsorbed by the silicon surface prevent normal growth of the oxide and promote the appearance of pores. In addition, fluorine is introduced into the silicon lattice, forming trivalent silicon which during subsequent thermal oxidation introduces an additional positive charge. The stabilization by phosphorus silicate glass (FSS) used in the ordinary technology of p-channel MDS-IC has a number of deficiencies, the basic ones of which are the following: high hygroscopicity of the FSS and the possibility of removing it during treatment of the structures in water or pickling agents containing water and hydrofluoric acid; low adhesion of the FSS to the photoresists. Other procedures are also used to stabilize the gate oxide.

The stabilization of the oxide with the help of hydrogen chloride is achieved by treatment of the silicon plates with hydrogen chloride or growth of the oxide in the presence of its vapor.

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

During the treatment of the silicon plates in hydrogen chloride vapor, mobile sodium ions Na^+ are bound by the chlorine ions Cl^- . Hydrogen interacts with the atoms of heavy metals in the body of the silicon.

On oxidation of silicon in hydrochloric acid vapor at temperatures of more 1000°C , the chlorine ions penetrate into the body of the oxide and are concentrated on the SiO_2 -Si interface. At temperatures below 1000°C , the stabilizing effect of the chlorine disappears, which is explained by a sharp decrease of its coefficient of diffusion into the oxide. An increased chlorine concentration at the SiO_2 -Si interface is explained by the presence of a large number of free silicon bonds. The chlorine ions interact with broken or weakened silicon bonds, forming bound Si-Cl vapor having ionic nature. As a result of significant asymmetry of the Si-Cl bonds, coulomb interaction and capture of the sodium ions take place with the formation of a stationary ternary system of $\text{Si-Cl}^- \text{-Na}^+$. On stabilization of the oxide with the help of hydrogen chloride vapor, it is important to select the conditions under which the chlorine traps capture all of the positive free charge.

The stabilization of the oxide by zinc is carried out by thermovacuum deposition of a Zn film on the surface of the silicon dioxide and subsequent diffusion of it to the oxide at a temperature of $1000\text{-}1200^\circ\text{C}$. The negative zinc ions neutralize the positive oxide charge. In contrast to the previously used diffusion of gold for compensation of the positive charge the zinc has a number of advantages: higher degree of purity, low evaporation point, the possibility of diffusion in a wide temperature range.

The stabilization of the oxide by injection of phosphorus ions is done at room temperature with ion energies providing for penetration of the oxide into the surface layer, but insufficient for penetration to the Si-SiO₂ boundary and to the surface layer of the silicon plate. On irradiation by phosphorus, the molecular bonds of the oxide are destroyed and new molecular bonds are formed capturing the admixtures and insuring tighter packing of the oxide. The oxygen ion released here is the trap for the positive sodium ions. Annealing at temperatures of $400\text{-}500^\circ\text{C}$ eliminates the radiation defects of the oxide and decreases the positive volumetric charge.

The application of new gate dielectrics, silicon nitride, silicon oxynitride, aluminum, zinc, titanium and other oxides having better dielectric properties and stability by comparison with silicon dioxide, improves the quality of the MDS-IC. The silicon oxynitride films are more easily treated by comparison with the silicon nitride films when making a pattern using photolithography. For the gates based on silicon oxynitride usually molybdenum is used. Silicon oxynitride can be obtained by reactive cathode sputtering of silicon in a mixture of $95\% \text{N}_2 + 5\% \text{NO}_2$. The application of double dielectrics $\text{SiO}_2\text{-Si}_3\text{N}_4$, $\text{SiO}_2\text{-Al}_2\text{O}_3$ and others increases the speed, decreases the threshold voltage of the MDS-transistors and permits creation of the MDS-IC which are not inferior with respect to reliability to the bipolar IC. Thus, for example, the application of a double gate dielectric $\text{SiO}_2\text{-Si}_3\text{N}_4$ with thicker SiO_2 layer in MNOS (metal-silicon nitride-oxide-semiconductor)-structures increases the specific capacitance of the gates as a result of high dielectric permeability of the silicon nitrite (about 7) and insures stability of the result of a decrease in the drift of the alkali ions.

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

A further decrease in the length or an increase in the channel width for the MDS-structures with large output powers is realized as a result of searches for new engineering design solutions. Thus, for example, the channel length can be decreased still more using the method of double diffusion and also when using V-MDS-technology with the application of vertical anisotropic pickling.

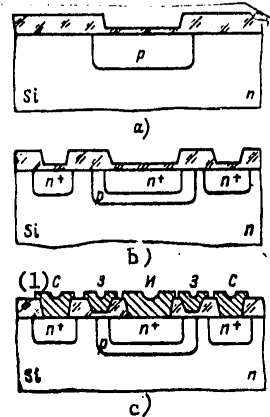


Figure 14-10. Diagram of the manufacture of MDS-IC by double diffusion. a -- local diffusion of the p-admixture; b -- local diffusion of the n-admixture; c -- growth of the gate oxide and formation of the gates and metal coating.

Key: 1. S = discharge, Z = gate, I = source,

The flow chart for the manufacture of MDS-IC with the application of double diffusion is shown in Figure 14-10. A mask of thermal oxide is formed on the surface of the silicon n-plate for local diffusion of the acceptor admixture to a depth of 0.5-3 microns. On formation of the second SiO_2 -mask, a window of smaller size is opened up under the sources, and additional windows under the discharges. During the process of local diffusion of the donor admixture, the source, discharge and channel regions are obtained. The channels are formed between the n^+ -sources and the initial silicon plate. Then, just as in ordinary technology, a thin oxide is grown under the gates, the windows are opened under the metal coating, the aluminum film is applied, and the metal coating pattern is formed.

The flow chart for the manufacture of p-channel V-MDS-IC is presented in Figure 14-11. The windows are opened in the thermal oxide, and diffusion of the acceptor admixture is carried out. Then using the local anisotropic pickling of the V-grooves, the p^+ -regions are divided into two parts, sources and discharges. The gate oxide is grown, and the subsequent operations are performed just as in ordinary p-channel technology.

The channel length is determined by the depth of the groove which, in turn, depends on the width of the window on the oxide. The width of the window is controlled is controlled well, and the depth of the V-groove is self-monitored, for the pickling halts at the point of intersection of the metal surfaces of the grooves

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

forming an angle of 55° with horizontal. The angle at the groove apex is 70° . As a result, the channel length is a highly reproducible parameter. The anisotropic pickling technique permits the manufacture of MDS-IC with channel length of 2-5 microns.

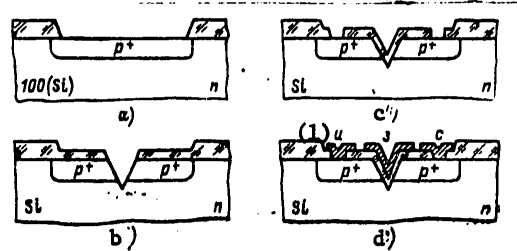


Figure 14-11. Diagram of the manufacture of MDS-IC with V-grooves. a -- local diffusion; b -- local anisotropic pickling; c -- growth of the gate oxide and photolithography; d -- formation of gates and metal coating.

Key: 1. I = sources, Z = gate, S = discharge

14-7. Protection of Semiconductor Structures

The surface of the semiconductor unprotected from external effects is distinguished by high sensitivity to any changes in the environment, temperature, moisture, radiation, and so on. Instability of the surface properties leads to instability of the parameters of the semiconductor structures. Therefore along with the problem of cleaning the surface in the production of microcircuits the problem of protecting the semiconductor structures is solved simultaneously.

The applied protection means -- various dielectric films and coatings -- must have good adhesion, they must not enter into undesirable interactions with the materials of the protected elements, they must be elastic and have good dielectric properties in order to insure mechanical stability and insulation of the elements from each other, they must have good passivating properties. By passivating properties we mean nonpermeability for moisture, various gases, vapor, radiation and also resistance to variations in temperature, pressure and other parameters of the external environment.

The protection of semiconductor planar and epitaxial-planar structures of IC is realized primarily by the masking film of silicon dioxide. It protects the required sections of the surface of the semiconductor plates during the technological treatment, and then after the last photolithography remains in the finished structure for protection against external effects. However, the passivating properties of the silicon dioxide do not fully correspond to modern requirements: the films transmit oxygen, the ions of alkali metals; at temperatures of 400°C and higher, ion conductivity is observed in the films, an additional charge is created on the Si-SiO₂ interface. All of this leads to instability of the parameters, poor resistance to external effects and, in particular, to poor radiation resistance.

Many of these deficiencies are eliminated when replacing the silicon dioxide films by higher quality silicon nitride Si₃N₄ films, the oxide films of the transition metals Al₂O₃, Ta₂O₅, TiO₂, Nb₂O₅, carbides, and so on.

FOR OFFICIAL USE ONLY

The silicon nitride films which have the following advantages over silicon dioxide are the most investigated after silicon dioxide:

1. Silicon is protected most effectively also for significantly smaller thicknesses (0.1 micron) from diffusion of phosphorus, boron, arsenic and also gallium and aluminum.
2. The surface of the silicon is best protected from the penetration of oxygen and alkali metal ions. The effective protection from oxygen diffusion is used for local oxidation of the silicon.
3. In the Si_3N_4 films ion conductivity at room temperature and increased temperatures to 400°C and also for high electric fields is absent.
4. The silicon nitride is chemically stable during heat treatments in a hydrogen environment to $1200\text{--}1250^\circ\text{C}$. It is not reduced, it does not form chemical compounds with the silicon similarly to the silicon dioxide film. Therefore the dimensions of the windows in the Si_3N_4 are better maintained after heat treatment than the dimensions of the windows in the SiO_2 .
5. Silicon nitride has higher dielectric permeability.

By comparison with silicon dioxide, silicon nitride has the following deficiencies:

1. High chemical strength with respect to the pickling agents complicates the process of local pickling; therefore the development of special photolithography procedures with respect to the Si_3N_4 films is required.
2. The Si_3N_4 is distinguished by high rigidity (high value of the Young's modulus), which is the cause of cracking of the layers during high temperature treatment.

In order to use the advantages of the Si_3N_4 when considering its deficiencies in practice double dielectric coatings are used: first a thin layer of SiO_2 is applied to the silicon surface, and a Si_3N_4 layer to it.

The methods of obtaining the masking films of silicon nitride do not differ theoretically, and to a great extent they resemble the methods of obtaining silicon dioxide films. As was noted in § 8-4, the silicon nitride films of good quality can be obtained by deposition from the vapor-gas phase as a result of chemical reactions between the hydrides or halides of silicon and nitrogen compounds (ammonia NH_3 or hydrazine N_2H_4). The silicon nitride films can also be obtained using reactive sputtering of a silicon cathode in glow discharge in a mixture of nitrogen with argon.

As a result of the above-indicated properties the silicon nitride films can be successfully used for passivation of the surface in the case of caseless encapsulation.

Among the oxides of the transition metals, the most investigated are the Al_2O_3 films. In order to obtain Al_2O_3 films it is possible to use the methods of anodic oxidation of aluminum, reactive sputtering of aluminum in an oxygen plasma, laser evaporation of the Al_2O_3 and sputtering of sapphire Al_2O_3 by an electron

FOR OFFICIAL USE ONLY

beam. The Al_2O_3 films are distinguished by high density, good insulating properties, high masking and passivating capacity. The Al_2O_3 films are a good barrier for the penetration of sodium ions; they have low sensitivity to atmospheric effects on water vapor; they have high radiation resistance (for example, an irradiation by electrons significantly fewer positive charges arise in them than in SiO_2). A deficiency of the Al_2O_3 films is relatively high density of the surface states and also electric instability of the $\text{Si-Al}_2\text{O}_3$ interface. Improvement of the properties of $\text{Si-Al}_2\text{O}_3$ interface is achieved by applying an SiO_2 sublayer 200 angstroms thick: $\text{Si-SiO}_2\text{-Al}_2\text{O}_3$.

The films of oxides of rare earth elements of the La_2O_3 type (where La is Nd, Sm, Eu, Gd, Dy, Ho, Er, Tm, Yb, Y) and also CeO can be obtained by direct deposition of the metallic film of rare earth elements with subsequent oxidation in the air or in flow of dry oxygen at temperatures not exceeding 600°C . The films are distinguished by high continuity; they have a specific resistance of 10^{12} to 10^{15} ohm-cm and high values of the dielectric constant $\epsilon_r = 7$ to 20; they have sufficient mechanical strength, good adhesion to the silicon, glass, aluminum, and they correspond to the majority of the requirements imposed on masking and passivating coatings.

At the present time basically silicon dioxide films are used for masking the series produced microcircuits. Masking with the help of these films does not provide complete protection of the structures; therefore additional passivation and protection of the surface are carried out.

Additional protection of the surface from the influences of the environment, possible contamination and damage during the assembly operations and also for stabilization of the surface properties and, consequently, the parameters of the structures is insured by all possible additional silicon nitride films, glass, lacquers, enamels and compounds.

The protection by glass promotes stabilization of the parameters of the structures, for the glass finds all possible free ions. Borosilicate, aluminosilicate, lead silicate, chalcogenide glass and glass of complex composition are most widely used for protection. The glass includes various additives to improve adhesion, to match the thermal coefficient of linear expansion of the glass to the semiconductor plate, and so on. For the application of glass coatings, the methods of thermovacuum deposition, cathode sputtering, centrifuging and sputtering by atomization with subsequent heat treatment, dipping in molten glass, and so on are used. During thermovacuum deposition the glass in the form of granules or powder is evaporated on the heated plates. The method permits us to obtain comparatively thin, uniform, pure films in a single cycle with the performance of other production operations. It is possible to carry out local deposition of the glass. The process is easily automated.

For centrifuging and spraying a suspension of finely-dispersed powder in a mixture of organic liquids (benzine, isopropyl alcohol, butyl alcohol, benzene, and so on) is prepared. During subsequent heat treatment, the layer of suspension is melted, and a continuous protective film of glass is formed.

Protection by enveloping polymers, lacquers, enamels, compounds and plastics is used as additional passivation of semiconductor planar structures; as insulation

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

during dense wiring of the IC to prevent short circuits between individual wire leads, between the wire leads and the passive part of the hybrid IC; for mechanical attachment of the microconnection points. The enveloping polymers are applied to the protective location by the batched drop method, spraying, or dipping, then drying is carried out to remove the easily volatile components and harden the polymer to obtain a continuous solid film.

14-8. Dielectric Film Control

When manufacturing semiconductor structures, the surface quality of the plates, the masking film, photolithography, diffusion, metal and insulating layers and the finished structures are controlled. In this section we shall briefly discuss the basic parameters of dielectric films and methods of controlling them.

In semiconductor structures the dielectric films are used as masking, protective, gate and insulating films.

The composition of the dielectric films can be determined using infrared spectroscopy. The infrared absorption spectra can be recorded on an automatic twin-beam UR-20 spectrophotometer built by the Karl Zeiss Company (German Democratic Republic). The absorption spectrum of silicon dioxide is characterized by maximum absorption on a wavelength of $\lambda = 9.3$ microns. The absorption spectrum of the silicon nitride has a sharp absorption band with peak in the wavelength range of $\lambda = 11.6$ to 12.1 microns. The presence of a peak is connected with absorption on the Si-N bonds. If there are admixtures of oxygen, carbon, hydrogen and so on in the silicon nitride layers, then additional peaks will appear in the infrared absorption spectrum. For example, absorption on the Si-O bonds gives a maximum at $\lambda = 9.4$ microns, absorption on the Si-H bonds gives a peak at $\lambda = 2.9$ microns, and on the Si-C bonds, at $\lambda = 12.4$ microns, and so on. Comparing the intensity of the various absorption bands in the infrared spectra, it is possible to determine the relative content of the elements entering into the film composition.

The structure of the films is investigated using an electron microscope or the x-ray method. As a rule, amorphous films have the best dielectric properties.

Strokes, lines and depressions on the film surface are observed visually using microscopes with high magnification (MBI-6, MIM-7). The number of defects must not exceed 10^4 cm^{-2} .

The film density can be determined directly by the increase in mass of the sample after applying the film. Usually in production the methods of indirect quality evaluation are used: by the pickling rate of the silicon dioxide film in a standard P-pickling agent or by the coefficient of optical refraction. In the former case the film of known thickness is pickled in a compound made up of 15 parts by volume of 49% HF + 10 parts by volume of 70% HNO₃ + 30 parts by volume H₂O. The dense, thermally grown films of silicon dioxide have a pickling rate of about 2 Å/sec. High pickling rates correspond to less dense films.

There is a linear relation between the density and the index of optical refraction

$$\rho = A + Bn, \quad (14-1)$$

FOR OFFICIAL USE ONLY

where ρ is the density, g/cm^3 ; A and B are constants; n is the coefficient of refraction measured on a wavelength of $\lambda = 5460$ angstroms.

Dense thermally grown silicon dioxide films have an index of refraction $n = 1.46$.

The film thickness is measured ellipsometrically, by the interference methods with respect to the reflection intensity, by the color of the films, and so on.

In the simple case a plate with silicon dioxide film is submerged on a slope in a beaker with hydrofluoric acid, and then it is slowly removed. A wedge is obtained during this pickling process. When investigating the pickling wedge under a microscope, color rings are observed. Counting the number of rings and noting the color of the last ring bordering the silicon dioxide on the unpickled film side, the film thickness is determined by a special table.

In order to measure the thickness and the index of refraction of the dielectric films, the ellipsometry method is most prospective. The LEM-1 laser ellipsometric microscope based on the LG-55 industrial gas laser permits measurement of dielectric films 10-10,000 angstroms thick with accuracy of ± 10 angstroms in the entire range and also their index of refraction.

The density of the electrically conducting defects is determined using the method of electrography developed by S. P. Sinitza. Photographic paper wet with water and the Si-SiO₂ sample are squeezed between two graphite electrodes. A voltage of about 10-15 volts is fed between the electrodes. The electric current passing through the system acts on the photopaper emulsion, causing darkening of it at the locations of the microdefects.

The amount of adhesion is determined using mechanical tests (for example, by the scratch test) which are performed until the film peels off the surface of the semiconductor plate.

The electric strength of the films is determined by the breakdown voltage of the electric field at a defined current.

The dielectric constant is determined by the capacitance of the MDS-structure. The voltage on the resistance connected in the measuring device in series to the capacitance is proportional to the capacitance; therefore it is possible to calibrate the measuring scale in capacitance units.

The condition of the dielectric-semiconductor interface is estimated by the concentration of charge carriers under the dielectric, surface potential, stationary charge in the dielectric, density of the surface states, stability of the stationary charge, and relaxation time of the surface states. The majority of parameters of the dielectric-semiconductor interface can be controlled by measuring the volt-farad characteristics of the MDS-structures on the ZhKM 3.416.001 automatic industrial device. The values of the parameters are determined by the nomograms calculated on a computer.

The passivating properties of the films are estimated by the degree of degradation (worsening) of the p-n junctions after the effect of a moist atmosphere on them. The back currents of the junctions on the breakdown voltages are the criterion for estimating the passivation quality.

FOR OFFICIAL USE ONLY

Test Questions and Assignments

1. Repeat the physics of the operation of MDS-transistors.
2. Compare the structural designs of MDS and bipolar transistors.
3. What explains the predominant use of p-channel transistors in the MDS-IC?
4. What explains the absence of the necessity for making special insulation for the MDS-IC elements?
5. Which IC elements can be made on the basis of the MDS-transistor structure?
6. What requirements are imposed on the silicon dioxide film in the MDS-structure?
7. What are the basic characteristic features and problems of the manufacturing technology of MDS-IC?
8. Using Figure 14-2, explain the basic steps in the manufacturing technology of thin oxide p-channel MDS-IC.
9. What is the process for obtaining a masking oxide, a gate oxide?
10. What are the basic deficiencies of the thin oxide MDS-IC technology?
11. What distinguishes the structural design of the MTOS-IC from thin-oxide MOS-IC?
12. Explain the basic steps of MTOS-technology (Figure 14-3). What step in MTOS-technology is the most complicated? Why?
13. Compare the coverage of the channel by the gate obtained in thin and thick-oxide technology.
14. Using Figure 14-4, relate the manufacturing process for MDS-IC with fixed silicon gates.
15. What are the advantages of silicon gates over aluminum gates?
16. Compare the flow diagram for the manufacture of MDS-IC with fixed metal gates using ion alloying. What are the characteristic features of the process?
17. What is the essence of the diffusion-ion method of fixing the gate. What are the advantages of the diffusion-ion technology of MDS-IC with fixed gates?
18. What are the advantages of CMDS-IC? Explain the flow diagram for the manufacture of diffusion CMDS-IC (Figure 14-6).
19. What are the basic advantages of the manufacturing technology of CMDS-IC using epitaxy? What is the essence and the technique for obtaining epitaxial pockets?
20. What is the orientation of silicon plates for planar-epitaxial CMDS-IC? Why?

FOR OFFICIAL USE ONLY

21. What are bicrystalline CMDS-IC? How are they made?
22. Repeat § 10-3 (diffusion from surface sources); § 9-4 (heteroepitaxy on dielectric substrates).
23. What are the advantages of CMDS-IC on sapphire substrates? Explain the flow diagram for manufacturing CMDS-IC with aluminum gates on sapphire substrates (Figure 14-8).
24. Explain the flow diagram for the manufacture of CMDS-IC with silicon gates on sapphire substrates (Figure 14-9).
25. List and explain the methods of improving the quality of MDS-IC used in practice.
26. Why is it not recommended that pickling agents contain hydrofluoric acid be used when cleaning plates before growing a thin oxide?
27. What are the methods of stabilizing the oxide by hydrogen chloride?
28. What is achieved when using new gate dielectrics? What are the dielectrics and why should they be used?
29. What parameters of MDS-IC are influenced by the length and width of the channel? How can the channel length be decreased?
30. Explain the flow diagram for the manufacture of MDS-IC using double diffusion (Figure 14-10).
31. Explain the flow diagram for manufacturing MDS-IC using vertical anisotropic pickling of the V-grooves (Figure 14-11).
32. Compare the diagrams in Figures 14-10 and 14-11 from the point of view of reproducibility of the channel length.
33. What are the deficiencies of protection using silicon dioxide films?
34. The films of what materials are prospective for masking and protecting semiconductor structures?
35. What explains the broad industrial application of SiO_2 protective films?
36. What is complementary passivation and complementary protection of structures and how are they executed?
37. List the controlled parameters and methods of controlling dielectric films.

FOR OFFICIAL USE ONLY

CHAPTER 15. MANUFACTURING TECHNOLOGY OF THE STRUCTURES FOR THIN-FILM MICROCIRCUITS

15-1. General Information

In this chapter we shall consider the manufacturing process for passive thin-film resistive and resistive-capacitive structures.

The substrates for passive thin-film structures must be made from materials with low electrical conductivity, low dielectric constant, high thermal conductivity and heat resistance, mechanical strength, stability of the dimensions, shape and chemical composition during the technological processing, with good compatibility with vacuum processes, and so on. There are no ideal materials that satisfy all of the imposed requirements. Substrates made of glass, pyroceram and ceramic correspond best to the basic requirements. The basic deficiencies of glass substrates are as follows: low thermal conductivity and mechanical strength. Substrates made of ceramic have comparatively high surface roughness to which the parameters of the thin-film resistors, capacitors and interelement connections are sensitive. This explains the broadest application of pyroceram substrates in Soviet industry. In cases where it is necessary to insure a higher level of heat removal, high mechanical strength and structural rigidity, anodized aluminum substrates are used.

Resistors are the most widespread film elements. The resistor film material must provide for obtaining a broad range of resistance ratings with stable properties and low thermal coefficient of resistance (TKS). Quite small values of the thermal coefficient of resistance of resistors can be obtained with film thicknesses of no less than 0.01 micron. However, for such thicknesses the majority of metals do not insure high values of the specific resistance of the film. This limits the application of many materials for obtaining thin-film resistors. Four groups of materials are distinguished which satisfy the requirements imposed on thin-film resistors. The first group includes pure metals: chromium, tantalum, titanium, rhenium, tungsten and certain others. The second group of resistive materials includes alloys and compounds: nichrome, metal silicide alloys (MLT-2, MLT-3, RS-3710, RS-3001) and tantalum nitride. The third group is made up of cermets -- mixtures of dielectrics and metals: chromium and silicon monoxide, gold and silicon monoxide, chromium and magnesium fluoride, and so on. The fourth group of materials includes the highly alloyed semiconductors: stannates based on tin dioxide with antimony trioxide, indium trioxide, and zinc oxide additives; silicon with chromium, iron or nickel additives, and so on.

The properties of the materials most frequently used to make thin-film resistors are presented in Table 15-1.

FOR OFFICIAL USE ONLY

Table 15-1. Materials for thin-film resistors

Material	Notation (composition)	Specific surface resistance, ohms/□	Thermal coeffi- cient of resistance, 10^{-6} , 1/°C	Method of of obtain- ing the films
Chromium	Cr	200-300	100	Thermovacuum deposition
Tantalum	Ta	200-500	+100-200	Cathode sput- tering, ther- movacuum deposition
Tungsten	W	100-5000	100	Thermovacuum deposition, pre- cipitation from PGF
Rhenium	Re	300-10,000	0-200	Thermovacuum deposition
Nichrome	NiCr (Ni--80% and Cr--20%)	To 300	+50-500	Thermovacuum deposition
Tantalum nitride	Ta ₂ N	50-500	+50-200	Reactive cathode sputtering
Metal sili- cide alloys	MLT (43.6%--Si 17.6%--Cr, 14.1%--Fe 24.7%--W) RC--3710 (37.9%--Cr 9.4%--Ni 52.7%--Si)	350-500 1000-30,000	+100-200 50	Thermovacuum deposition Cathode sputter- ing, thermovacuum deposition
Cermet	(50-90%)Cr (50-90%)SiO ₄	1000-2000	+75	Thermovacuum deposition

Thin-film wires and contact areas must have high electrical conductivity, low transient resistance between a conducting layer and other elements of the microcircuit, high adhesion to the substrate and good compatibility with other materials entering into the microcircuitry composition and also insure high-quality connection of the external leads of the microcircuit or the mounted crystals or components. For the conducting layers, copper, aluminum, nickel, silver and gold are used. All of these metals, with the exception of aluminum, have poor adhesion to the dielectric substrates; therefore on deposition of them an additional sublayer of chromium, tantalum, molybdenum, and so on are obtained.

The properties of some of the wire materials are presented in Table 15-2.

In the case of a purely resistive passive part of the microcircuit basically copper is used for commutation. In microcircuits with film capacitors, aluminum is used which simultaneously serves as the capacitor plates. Gold is used to make microcircuits that operate under harsher operating conditions. As the upper layer providing for connection and shielding of the wires and contact areas frequently nickel is used, sometimes a layer of solder. The thickness of the solder can be 10-15 microns, which sharply decreases the commutation resistance.

FOR OFFICIAL USE ONLY

Table 15-2. Electrophysical properties of the basic materials of thin-film conductors and contact areas.

Material (thickness, microns)		Specific surface resistance, ohms/□	TKC,* 10 ⁻⁶ 1/°C	Method of obtaining wire films
Wire	Sublayer			
Copper (0.6-0.8)	Chromium, titanium, nichrome (0.01-0.03)	0.02-0.04	4300	Thermovacuum deposition
Aluminum (0.25-0.35)	Chromium, titanium, nichrome (0.04-0.05)	0.1-0.2	4200	The same
Gold (0.6-0.8)	Chromium(0.01-0.03)	0.03-0.04	3900	Thermovacuum deposition. Cathode sputtering

* Thermal coefficient of resistance

At the intersections in the case of multilevel metal plating the wires are insulated from each other by a thin film of chalcogenide glass, and so on. During formation of the multilevel metal plating of hybrid LSI, silicon dioxide, aluminum oxide, polyimide, and so on are used as the interlevel dielectric insulation.

Film capacitors are formed on the basis of three films; on the lower plate, the dielectric and the upper plate. The basic element of the film capacitor is the dielectric. The requirements imposed on the dielectric material for thin-film capacitors (high dielectric constant, low dielectric losses, high breakdown voltage, absence of through pores in the films, minimum hygroscopicity, and so on) correspond to a high degree to silicon monoxide SiO. Along with the silicon monoxide, antimony trisulfide Sb₂S₃, germanium monoxide GeO, silicon dioxide SiO₂, the oxides of the transition metals Al₂O₃, Ta₂O₅, borosilicate and aluminosilicate glass, and so on are used. The properties of some dielectric materials and methods of obtaining films are presented in Table 15-3.

Materials having high migration to the dielectric layer, for example, copper or gold, are not suitable for capacitor plates, for they form a large number of shorts between the plates. Metals with high evaporation temperature, for example, chromium and nickel, are also unsuitable. The atoms of these metals have high energies on evaporation, and therefore they are introduced deeply into the dielectric, lowering its electric strength.

Most frequently aluminum is used for the plates of thin-film capacitors. Aluminum has low evaporation temperature and comparatively low mobility of the atoms. On its surface there is always an oxide film which prevents penetration of the atoms of the plate material into the dielectric layer and plays the role of shielding.

15-2. Standard Circuits and Basic Steps in the Manufacture of the Structures of Thin-Film Microcircuits

Standard Schemes for the Manufacture of the Structures of Thin-Film Microcircuits. The formation of the configuration and mutual arrangement of thin-film elements is realized using free masks contact photolithography and electron beam. When using free masks the application of the film and formation of the configuration of the elements are combined, that is, they are carried out simultaneously in one process. When using photolithography or an electron beam, these processes are separated in

FOR OFFICIAL USE ONLY

time: first a continuous thin-film is applied to the entire surface, then the excess sections in this film are removed. Accordingly, the flow charts for the manufacture of the structures of thin-film microcircuits are distinguished:

By thermovacuum deposition through free masks;

Ion-plasma spraying through free masks;

Thermovacuum deposition with the application of photolithography;

Cathode or ion-plasma sputtering with the application of photolithography;

Ion-plasma spraying and electrolytic oxidation using photolithographic treatment;

Thermovacuum deposition with the application of an electron beam.

The free metal masks for obtaining element configurations can be used only in combination with the methods of thermovacuum deposition and ion-plasma spraying. When obtaining thin films by the method of cathode sputtering, metallic free masks cannot be used, for they distort the electric field between the cathode and the anode. As a result, the energy of the argon ions or ions of another ionizing gas bombarding the cathode decreases sharply. This leads to a significant decrease in the deposition rate of the films or to complete cessation of the process. The application of free masks made of dielectric materials that do not distort the electric field, for example, masks made of photopyroceram, still is limited by the complexity of the manufacturing technology. In the case of application of the method of cathode sputtering to obtain thin films, the configuration of the elements is obtained by photolithographic treatment.

Basic Steps in the Manufacture of the Structures of Thin-Film Microcircuits. In all of the flow charts for making the film part of the microcircuits it is possible to isolate the following basic steps:

The manufacture and preparation of the free masks or photomasks;

Preparation of the equipment, fittings and materials to obtain the films;

Manufacture of the substrate;

Obtaining the films (with simultaneous obtaining of the layout in the case of using free masks);

Obtaining the configuration of film elements for cases of using photolithography or cathode-ray treatment;

Adjustment of the rated values of the film elements;

Shielding.

The manufacturing technology of free masks was investigated in § 6-2, and photomasks in § 5-6. The free masks and the photomasks are multiple-use tools. Therefore they must be prepared each time before carrying out the processes. The free masks are

FOR OFFICIAL USE ONLY

Table 15-3. Properties of some dielectric materials of thin-film capacitors

Materials	Dielectric constant ϵ_r	Dielectric loss tangent, % (for $f = 1$ kilohertz)	Electric strength, kv/mm	TKE*, 10^{-4} $1/^\circ\text{C}$	Specific capacitance, picofarads/ cm^2	Method of obtaining
SiO	5-6	0,5-1	10-100	2-5	10^4	Thermovacuum deposition
GeO	10-12	0,2-0,5	10-50	2-5	$10^4-2 \cdot 10^4$	The same
SiO ₂	4	0,5	$5 \cdot (10^4-10^5)$	2	$2 \cdot 10^4$	Ion-plasma spraying, reactive spraying
Sb ₂ S ₃	18-20	0,5-1	T ₀ 200	5-6	$2 \cdot 10^4$	Thermovacuum deposition
Al ₂ O ₃	8	0,3-1	T ₀ 500	3-5	$3 \cdot 10^4-4 \cdot 10^4$	Reactive sputtering, anodic oxidation
Ta ₂ O ₅	20-22	0,3-1	T ₀ 500	2-3	$5 \cdot 10^4-10^5$	The same
SiO ₂ ·B ₂ O ₃	4	0,1-0,3	T ₀ 100	100	$10^4-2 \cdot 10^4$	Thermovacuum deposition
SiO ₂ ·Al ₂ O ₃	4-5	0,2-0,3	T ₀ 100	100	$10^4-5 \cdot 10^4$	The same

* Thermal coefficient of resistance

cleaned to remove contamination and traces of the preceding film deposition, which unavoidably gets on the surface of the mask and leads to clogging of the openings. During the operating process the masks gradually lose their precision, and therefore after a defined time they must be replaced by new ones. Thus, for example, bimetal masks withstand about a 100 film deposition cycles.

During the process of contact exposure in practice a photoresist is always poured on the working surface of the photomasks; it is also possible for contamination to get on the working surface from the atmosphere; therefore proper organization of the cleaning of the photomasks before each new photolithographic process is necessary. The masks are put in a clean container with acetone; then they are wiped with a tampon of lint-free fabric wet in acetone; then they are held for 1-1,5 minutes in potassium dichromate, they are washed in deionized water and dried with a jet of purified air. Good cleaning qualities are achieved when treating the masks in an oxygen plasma.

The preparation of the equipment, the fixtures and initial materials to obtain the films is an important step and takes up a significant part of the time and especially in the thermovacuum deposition. Before each thermovacuum deposition cycle, the surfaces of the parts of the vacuum devices and the hood are cleaned to remove materials that have been left from preceding depositions. The parts are cleaned using a scalpel, an electric vacuum cleaner and then coarse calico wet in ethyl alcohol. The cleaning is stopped when there are no more traces of dirt on the

FOR OFFICIAL USE ONLY

fabric. Careful cleaning of the removable parts of the hooded structure by an emery cloth, washing in benzine and ethylalcohol are carried out periodically once a week in a special facility. The preparation of the evaporators and also the sample of evaporated material includes the operations of degreasing, pickling, washing and drying. The formulas for the degreasing solutions and pickling agents and also the treatment conditions are determined by the material of the evaporator or sample. It is expedient to prepare the evaporators and the samples directly before charging in the vacuum chamber.

After cleaning the evaporators and the samples there should be no spots or raised places on their surfaces. The prepared evaporators and samples are stored in sealed packaging or moistureproof cabinet in a dry facility.

The manufacture of the substrates includes machining (see Chapter 3) and cleaning of the surface. The basic operations of the processes of cleaning glass, sapphire, ceramic and pyroceram substrates are presented in § 4-4. It is possible to add that as a result of the presence of the remains of ion exchange resins in deionized water which quickly decompose into component parts, preference is given to distilled or bidistilled water. When treating the pyroceram substrates it is necessary to maintain the pH of the ammonia peroxide solution on a constant level of 6.8-7. Regulation is realized by adding ammonia to the solution by drops to the required pH. The pH is checked using indicator paper with a color scale and digital designations of the pH or using special pH-meters.

A protective polymer film is poured on the clean substrate surface. After drying, adhesive tape is applied to the film which is fastened by the protruding edge to the lever in the vacuum chamber. On turning the lever, the polymer film separates, baring the clean substrate surface. Finish cleaning of the substrates by the "dry" methods, annealing, ion bombardment, and so on is carried out in a vacuum chamber.

Thin films are obtained for the passive elements and the interelement connections most frequently by thermovacuum deposition and ion sputtering. Recently, they have also started to use electrolytic oxidation, deposition from the vapor-gas phase and from electrolytes and other methods investigated in Chapter 8.

The films are applied in the following order: first the resistive film is applied, then conducting and the contact areas, then the lower capacitor plate film, the dielectric, the upper capacitor plate film, and the protective film, last. Depending on the microcircuit nomenclature, some films can be absent, and others can be obtained together (for example, contact areas and capacitor plates).

The next steps in the manufacture of thin-film microcircuits will be investigated in detail below.

15-3. Technology for Manufacturing the Structures of Thin-Film Microcircuits with the Application of Free Masks

Separate and Continuous Methods of Applying Films. When manufacturing microcircuits by the vacuum method the films can be applied by separate, continuous and combination methods.

In the separate method, each film is applied in a separate vacuum unit. Free masks are combined with the substrate in the air and then the assembled combination unit is installed in the chamber of the device on a carousel or substrate drum.

FOR OFFICIAL USE ONLY

Then weighed samples of the evaporated material are placed in the evaporator or on the evaporator carousel. Then the device is sealed, the air is pumped out and the substrates are heated. The film deposition takes place successively on one or several substrates, or simultaneously on all of the substrates installed on the turning carousel. After cooling of the substrates air is admitted to the chamber. The substrates are combined with the masks of the next layer and then the deposition of the new layer is accomplished in the same sequence. Thus, after deposition of each layer of the microcircuit, the vacuum chamber is unsealed.

During separate deposition, high accuracy of matching the patterns of the microcircuit layers with each other is achieved, for it is easier to handle them in the air than in the vacuum chamber. The interoperation monitoring is possible here; mutual contamination with the deposited materials is excluded, the hooded devices are simpler by comparison with the structures for continuous deposition. The deficiencies of the separate method include the possibility of contamination of the films during transport to the next vacuum chambers and the duration of the overall process as a result of multiple evacuation and unsealing.

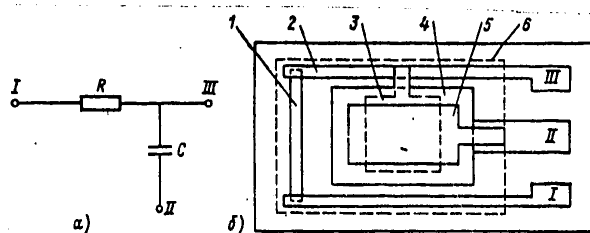


Figure 15-1. Circuit diagram (a) and topology (b) of a thin film RC-IC. 1 -- resistor (chromium); 2 -- conducting tracks and contact areas (copper with a chromium sublayer); 3 -- lower capacitor plate (aluminum); 4 -- capacitor dielectric (silicon monoxide); 5 -- output capacitor plate (aluminum); 6 -- shielding layer (silicon monoxide).

In the continuous method the substrates and masks are placed on separate carousels of the hooded unit. An evaporator is placed in each of the evaporation positions, and the corresponding mask on them. The matching of the masks with substrates and charging of the evaporators with the weighed samples are realized in the evacuated operating chamber using a mechanical device for matching and special manipulators. This permits several different films to be deposited with only one evacuation and one unsealing of the vacuum chamber of the unit. Here the interaction of a separate film with the air is excluded. The total deposition process time decreases. However, the continuous method requires significantly more complex hooded equipment, the operation of which under increased temperature conditions and a high vacuum is difficult. In order to decrease the mutual contamination of the films, reliable shielding of the substrates and evaporators is required, and, in addition, it is difficult to check the films.

Technological Process of Manufacturing a Resistive-Capacitive Thin-Film Microcircuit (Figure 15-1). When manufacturing RC-microcircuits the films are deposited in the majority of cases by a combined method. The resistive and conducting films are applied in one vacuum cycle (Figure 15-2). After installation of all of the attachments in the vacuum chamber; evaporators, weighed samples, free masks,

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

substrates, control samples -- "indicators" -- and checking the reliability of the setting, the system is evacuated. Directly before deposition of the film the finish cleaning of the evaporators and the weighed samples by annealing are carried out (in this case the evaporation behind a shutter).

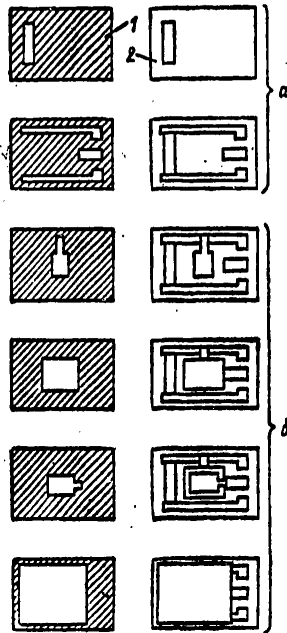


Figure 15-2. Flow chart for the manufacture of a thin-film RC-IC structure. a -- deposition of resistors, wires and contact areas; b -- deposition of condenser layers and protective film, 1 -- free mask; 2 -- substrate.

Chromium is one of the easily sublimable materials. Its vapor pressure of 1.33 Pa is reached at a temperature 500° C below its melting point. A resistive tungsten evaporator coated with a galvanic layer of chromium or a rod made of pressed and sintered chromium powder uniformly heated by radiation from a tantalum cylindrical evaporator is used for chromium sublimation. During deposition the chromium actively interacts with the residual oxygen, which makes it possible to obtain high specific resistances of the films. The deposition of the chromium is carried out on substrates heated to temperatures of 200-250° C. The substrate temperature is regulated by varying the heating element current. When the required substrate temperature is reached, the chromium evaporator is brought to the evaporation position and a voltage is fed to it. When heating the evaporator the vacuum in the chamber becomes worse, for gases are released from the surfaces of the evaporator and the weighed sample of chromium. After completion of the gas release and the restoration of the vacuum, the shutter is opened, and the chromium film is deposited on the substrate and the "indicator." When the required film thickness is reached, the shutter is closed, the next substrate is moved to the evaporation position, and the process continues to deposit the film on all of the substrates.

FOR OFFICIAL USE ONLY

In contrast to chromium, copper is evaporated from the liquid state. It in practice does not interact with refractory materials, among which molybdenum is preferred for evaporators, for it is machined well and conducts heat well. In order to prevent spattering of the copper, it is heated in stages: first it is melted, then the current through the evaporator is smoothly increased to the working value, the shutter is opened, and deposition on the substrate takes place. After deposition the substrate temperature is lowered to 50-60° C, the vacuum line of the working chamber is closed, and air is admitted to it.

The deposition of the capacitor layers and the protective layer takes place in another device and frequently in a continuous vacuum cycle. This structure of the process permits monitoring of the parameters of the resistors and adjustment of the ratings after deposition of the wiring and the contact areas.

Four positions of the vacuum chamber are charged with weighed samples in accordance with the sequence of film deposition: aluminum, germanium monoxide, aluminum and silicon monoxide. The substrates fastened to the substrate holders are installed in recesses in the matching device, and masks for each evaporator, respectively, are installed under the substrates above the evaporation positions. Then the chamber is sealed, and it is evacuated to a pressure of about $1.3 \cdot 10^{-3}$ Pa. The substrate temperature is brought to the working temperature, and it is maintained on a constant level during all of the depositions. The substrate is matched with the mask of the lower capacitor plate and the current begins to be smoothly increased through the evaporator to a temperature of 1000° C corresponding to the temperature of yellow heat of aluminum at which the surface film of Al_2O_3 decomposes and evaporation begins. The shutter is opened and the aluminum is deposited successively on all of the substrates charged in the chamber.

Then the first substrate from the lot is moved to the second position, it is matched with the second mask and the heating of the germanium monoxide evaporator is started. Thus, the deposition of all of the capacitor layers on all of the substrates takes place in one device.

The protective layer of silicon monoxide usually is evaporated from tantalum or molybdenum evaporators. Silicon monoxide is distinguished by high adsorption properties; therefore before deposition on the substrates it is carefully degreased to complete restoration of the vacuum in the working chamber.

As has already been noted, the deposition through free masks can be used also for ion-plasma spraying of films. In cases where it is impossible to use free masks or where it is necessary to obtain a sharper, finer pattern, photolithography is used.

15-4. Manufacturing Technology of the Structures of Thin-Film Microcircuits with the Application of Photolithography

Photolithography is most frequently used when manufacturing precision resistive microcircuits with complex layout. In order to obtain a contact mask, a photorealist or another material is used which is more resistant to the subsequent technological effects (photolithography with a sublayer). The method of direct photolithography provides for the following sequence of the process of formation of the configuration of the elements: the application of a continuous film of the material

FOR OFFICIAL USE ONLY

of the thin-film element; the formation of a contact mask; pickling through openings in the contact mask to remove the excess sections of the film. The method of inverse photolithography is realized in practice in the following sequence: formation of contact masks on the substrates, the application of the film element material, removal of contact masks. On application of a film to a substrate with formed contact mask two versions are possible. In the first version the film is deposited both on the contact mask and on the segments of the substrate free of it. In the second version the film is deposited only on the sections of the substrate free of the mask. The first version is realized during thermovacuum deposition, ion sputtering, deposition from the vapor-gas phase, and the second version, during electrolytic deposition of electrically conducting materials.

Thus, during direct photolithography the configuration of the elements is shaped as a result of the application of local pickling; during inverse photolithography, as a result of local application of the film. In addition to localness a necessary condition of the formation of the element configuration is also observation of the principle of selectiveness. The principle of selectiveness in film technology consists in the fact that the pickling agent must act only on the film, sections of which must be removed, or on application to a substrate with a contact mask the film must only be deposited in the holes. If the film is also deposited on the contact mask, it is necessary to provide for breaking the pattern around the outline of the opening in the mask and it is necessary to insure good adhesion of the film deposited in the opening to the substrate so that during subsequent removal of the mask the excess sections of the film will also be removed with it. Correspondingly, in the literature frequently the methods themselves are called selective pickling and selective deposition.

Flow Diagram for the Manufacture of Structures of Thin-Film R-Microcircuits Using Selective Pickling (see Figure 15-3). This flow diagram is used in cases where it is necessary to obtain a complex pattern of interelement connections and high precision of the resistor ratings. In this layout the photolithographic process is used twice: first for the formation of the pattern of contact areas and then for the formation of the resistor layout. During selective pickling of chromium, a gold contact mask is used which is obtained during the process of primary photolithography. In connection with the duration, the technological process is used when it is economically justified by high electrophysical parameters and a high percentage yield of good microcircuits.

The system for manufacturing thin-film R-microcircuits using a free mask and selective pickling is presented in Figure 15-4. The free mask is used to obtain the layout of the wiring and contact areas, and selective pickling, for the resistors. Photolithography is used once to obtain a protective photoresistive mask required for selective pickling. This flow chart is simpler by comparison with the preceding one, for the complex photolithographic process for formation of the interelement connections is excluded. However, the method does not allow for obtaining resistive microcircuits with complex precision system of interelement connections.

A characteristic feature of the presented flow charts for the manufacture of R-IC structures is also the application of films of multicomponent composition. During ordinary evaporation of such materials fractionation, decomposition, and chemical interaction of them leading to the obtaining of films of the composition differing from the evaporated material are possible. Therefore special methods of evaporation

FOR OFFICIAL USE ONLY

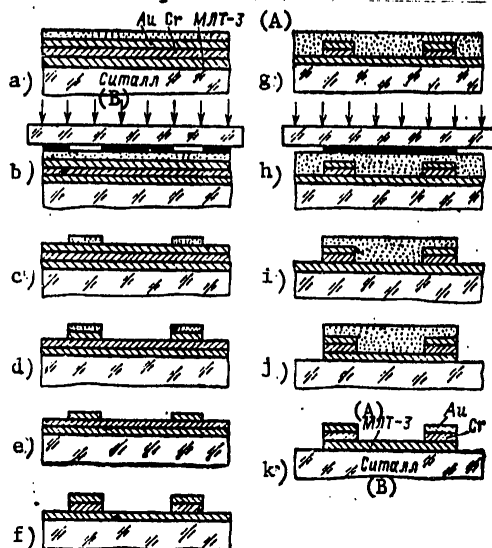


Figure 15-3. Flow diagram of the manufacture of thin-film R-IC using selective pickling (double photolithography). a -- application of continuous films MLT-3, Cr, Au and formation of the photoresistive layer; b -- exposure; c -- development; d -- selective local pickling of the gold film; e -- removal of the photoresistive mask; f -- selective pickling of the chromium film; g -- formation of the photoresistive layer; h -- exposure; i -- development; j -- selective local pickling MLT-3; k -- removal of the photoresistive mask,

Key: A, MLT-3 B, pyroceram

are used; evaporation from several evaporators; explosion evaporation (discrete evaporation); cathode ray evaporation. The maintenance of the stoichiometric composition of multicomponent materials is also provided for during cathode sputtering of them.

Nichrome, for example, is distinguished by great difference in chromium and nickel vapor pressures at the same temperatures. Thus, at 1300° C the chromium vapor pressure exceeds the nickel vapor pressure by eight times. In addition, nichrome has a high chemical activity with respect to many of the crucible materials and residual oxygen. The critical thickness of 0.01 micron has been defined for nichrome, below which the film parameters are unstable. During thermovacuum deposition, the evaporation of the nichrome is carried out by the explosion method. Nichrome wire is unwound by a transport mechanism from a coil, it passes through a guide tube and then is brought into contact with an evaporator heated to 1620° C. The evaporation of the wire from the nichrome melt can also be realized by an electron beam.

15-5, Manufacturing Technology of Structures of Tantalum Thin-Film Microcircuits

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

large area of getter film by spraying on a water-cooled shield. It is complicated to obtain tantalum resistors in that the tantalum films can exist in three versions: α -Ta -- this is usually coarse crystalline body-centered lattice with comparatively low specific resistance from 20 to 50 $\mu\text{ohms-cm}$; β -Ta with fine-crystalline tetragonal lattice and specific resistance of 160-200 $\mu\text{ohms-cm}$ and porous tantalum with low density and specific resistance approximately equal to 3200-7000 $\mu\text{ohms-cm}$, consisting of α and β -Ta grains. Entirely precise conditions of obtaining pure tantalum films of the given structure have still not been fully defined. Therefore certain admixtures, for example, oxygen, nitrogen and gold, are introduced into the tantalum to give the films the required properties.

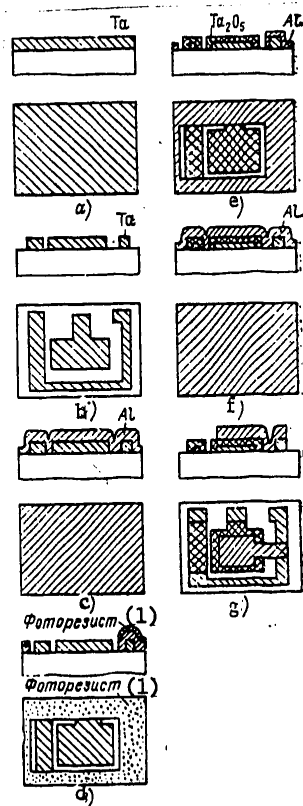


Figure 15-5. Flow diagram for the manufacture of thin-film tantalum RC-IC, a -- application of tantalum film; b -- first photolithography; c -- deposition of aluminum film; d -- second photolithography; e -- electrolytic oxidation of tantalum and removal of photoresist; f -- deposition of aluminum film; g -- third photolithography.

Key: 1, photoresist

Oxide films of tantalum are obtained by the methods of electrolytic or plasma anodic oxidation. Reactive cathode sputtering is not used, for the films have low electric

FOR OFFICIAL USE ONLY

strength and large leakage currents. Films obtained by electrolytic and plasma anodizing have a dense, amorphous structure; they have mechanical and chemical strength, low dielectric losses and high breakdown voltage.

The tantalum nitride films are obtained by reactive sputtering of a tantalum cathode in an argon plasma with the addition of nitrogen. Depending on the partial pressure of the nitrogen it is possible to obtain tantalum nitride with different structure and different electrophysical properties TaN and Ta_2N . In practice the sputtering conditions are selected so as to obtain Ta_2N , for resistors made of such films are distinguished by high stability.

For the lower capacitor plates and sometimes also for the resistors β -Ta is used; α -Ta films are not used as a result of poor adhesion to the substrates and high mechanical stresses. The oxide films of tantalum Ta_2O_5 are used as the dielectric of capacitors, insulation of the intersections during multilevel commutation and also as a protective layer for resistors.

Tantalum Technology of the Manufacture of RC-Microcircuits. The process flow chart for the manufacture of an RC-microcircuit is presented in Figure 15-5. A continuous film of tantalum is deposited on a pyroceram or glass substrate by spraying, and photolithography is used to form the layout of the lower capacitor plates and wiring; then thermal deposition of a continuous aluminum film is carried out, in which a conducting track is obtained by means of secondary photolithography for the electrical connection of all sections of the tantalum obtained after primary photolithography to each other. Without removing the photoresistive mask, the substrates are charged in an electrolytic bath as the anode, and electrolytic oxidation of all sections of the tantalum connected by the aluminum track is carried out. The tantalum oxide film obtained during the oxidation process is a means of adjusting the resistance ratings, it protects the tantalum resistors and it plays the role of the capacitor dielectric. After oxidation, the photoresistive mask is removed, and a continuous aluminum film is applied. As a result of the tertiary photolithography, the upper capacitor plates are obtained.

Thus, tantalum technology permits one material to be used for the resistors, inter-element connections and capacitors. This simplifies the process of manufacturing the microcircuits and makes it possible to obtain high packing density of the elements with high stability and precision of their parameters.

15-6. Manufacturing Technology of the Structures of Thin-Film Microcircuits Using the Beam Processing

At the present time electron, laser and ion beams have begun to be used in the technological processes of manufacturing thin-film structures. Obtaining the configuration of thin-film elements is based on the nonthermal or thermal effects of beams on the material of the contact masks or on directly on the thin-film material.

Contact masks can be made of electronresists (by electron lithography) or materials that evaporate easily during beam treatment and then are easily removed from the surface of the substrates. The optimal thickness of the masks is 0.5-5 microns. Contact masks that are too thin cannot withstand treatment in pickling agents when transferring the pattern to the film; thick mass undesirable as a result of lowering the resolution.

FOR OFFICIAL USE ONLY

Obtaining thin film structures by electron lithography differs from the photolithography process only by the application of the electron resists and accelerated electrons for exposure.

In order to obtain a contact mask, for example, a metal contact mask, pulse treatment is used with the help of the thermal effect of an electron beam. For complete removal of the required section of the film during one pulse it is necessary that the depth of penetration of the electrons exceed the thickness of the removed film by more than 1.5 times.

In order to obtain contact masks using a laser beam, primarily solid-state lasers are used.

The application of a laser beam is prospective as a result of the possibility of local treatment by a sharply focused beam; high energy density at the focal point sufficient to evaporate in practice any material; the possibility of short-term effect (10^{-3} to 10^{-6} sec) in which the laser beam does not change the properties of the surrounding sections of the treated surface; shallow depth of penetration of the laser radiation into the metals which permits thin films to be treated without destroying the lower layers or the substrate.

The obtaining of contact masks by an ion beam consists in sputtering of the film material by the focused beam of ions accelerated to high energies.

For local removal of thin films, ion pickling through contact masks has begun to be widely used at the present time. During pickling to greater depth the photoresistive mask is unsuitable as a result of flowability under the effect of bombardment and the quite high characteristics pickling speed. It is recommended that titanium, vanadium, chromium, manganese, aluminum, tantalum and also carbon be used as the contact mask material. These materials have low ion pickling rate and are easily removed after obtaining the film pattern.

By maskless treatment using sharply focused beams it is possible to evaporate not only the material of the shielding contact mask, but by the corresponding increase in power, even the thinnest film ("milling" of the film). During direct treatment of the films it is necessary to consider that in contrast to a laser beam, the electron beam penetrates deeply, and the depth of the ions usually is an order higher than the thicknesses of the films used; therefore for direct electron-beam "milling" of thin films it is possible to use only substrates capable of dispersing the heat released in them well, for example, substrates made of aluminum ceramic, the thermal conductivity of which exceeds the thermal conductivity of glass and pyroceram by more than an order.

The contact mask method is applicable for obtaining the configuration of elements on films obtained by various methods. When using treatment by electron, laser and ion beams, the contact mask methods permit elimination of the expensive photomasks and also chemical reagents and the transition to completely automated processes with the application of computers. The methods of direct treatment of the film material by the beams when the circuit will be traced by the beam without using a contact mask are especially advanced. At the present time combined methods are most widely used in producing film microcircuits. These methods combine electron lithography and direct "milling" by an electron beam.

FOR OFFICIAL USE ONLY

15-7. Adjustment of the Rated Values and Shielding of the Film Elements

The possibility of bringing the parameters to the rated values with the required precision after making the elements is an important advantage of film technology. The application of special adjustment techniques permits the dispersion of the parameters caused by nonuniformity of the conditions during the technological treatments to be decreased and it makes it possible to obtain precision film elements. At the present time the methods of adjusting the resistors have been better worked out than the methods of adjusting capacitors.

The adjustment of the elements can be provided for by the topology and carried out by cutting or soldering special jumpers. This fitting insures a discrete change in the parameters in the direction of increasing or decreasing the rated value.

In contrast to the topologic discrete fitting, the fitting realized by the technological procedures without noticeable variation of the topology is smooth. Smooth adjustment is achieved by varying the transverse cross section of the resistor or the area of the upper capacitor plate by removing part of the film and also as a result of varying the film structure.

It is possible to remove part of the film by abrasive machining, evaporation using electrospark, laser or electron beam treatment and also by electrochemical or ionic pickling. The machining method does not provide the required precision and finish. Electrochemical pickling is a low-efficiency process. Most frequently the methods of electrospark and laser adjustment are used which insure high precision and reproducibility. Adjustment by varying the configuration of the elements requires a shift of the instrument with respect to the film. With a high degree of integration of the microcircuits complex precision electromechanical units are required to shift the instrument. All of the methods, except electrochemical and ionic pickling, are individual, for only one element is simultaneously adjusted. The deficiencies of adjustment by decreasing the transverse cross section of the resistors include the following: the one-way nature of the adjustment only in the direction of increasing the resistance, the possibility of burning out the resistor during its operation as a result of overheating at the point of least cross-section, unacceptability for adjusting multilayer microcircuits as a result of the necessity for direct access to the resistor.

The variation in the film structure leading to the required variation of its specific resistance can be achieved during heat treatment, anodic oxidation, pulse current tuning, and so on.

During heat treatment and pulse current tuning the resistance of the resistor film changes as a result of annealing of the defects and oxidation of both the film surface and its intergrain boundaries. Here the specific resistance of the film can be varied both in the direction of a decrease and in the direction of an increase.

The adjustment of the resistors by anodization is used only for rectifying metals; its advantage is simultaneous obtaining of the shielding coating, for example, tantalum oxide Ta_2O_5 .

Along with pickling, heat treatment, anodizing, the group methods of adjustment include the photochemical method. The photochemical method consists in the fact that the deviation of the resistances of the resistors from the rated value detected during the measurements can be decreased by additional photolithography with the

FOR OFFICIAL USE ONLY

application of photomasks with the pattern taking into account the obtained resistance deviation. This method requires a set of special adjusting photomasks. The photochemical method makes it possible to vary the resistances both in the direction of an increase and a decrease.

The shielding of the thin-film elements is provided by silicon monoxide films, aluminum and tantalum oxides, photoresists, an additional layer of solder, polymer films, for example, an epoxy resin film.

15-8. Quality Control in the Production of Structures for Thin-Film Microcircuits

Quality Control in the Film Application Process. During the production of thin-film microcircuits, the degree of vacuum, the partial pressure of the residual gases in the working chambers of the devices, the substrate and evaporator temperatures, the thickness and the growth rate of the films, and the parameters of the deposited elements are controlled.

The vacuum is controlled by means of thermovapor and ionization vacuum meters. The thermovapor vacuummeters measure pressure from $1,3 \cdot 10^3$ to $0,13$ Pa during operation of a forevacuum pump. Their operating principle is based on the use of the thermoconductivity of a gas as a function of pressure in the volume of the measuring thermocouple tube (LT-2 or LT-4M) equal to the pressure in the vacuum chamber. The filament of the thermocouple tube is made of metal with low evaporative capacity (W, Pt); therefore its temperature is determined by the voltage fed to it and the thermal conductivity of the gas. A thermocouple is welded to the middle of the filament. During evacuation of the air the thermoconductivity diminishes, and the emf grows on the outer ends of the thermocouple. Ionization vacuummeters are used to measure the pressure from $0,13$ to $0,6 \cdot 10^{-5}$ Pa during operation of the high-vacuum pumps. Their operating principle is based on using the phenomenon of ionization of the residual gases. The ionization of the residual gases takes place under the effect of the electrons emitted by the incandescent cathode and accelerated by the electric grid field in a three-electrode tube connected to the evacuated chamber. The ion current on the tube collector is proportional to the gas pressure.

Since an electric signal is obtained at the output of the manometer, they are very convenient for automation of the control operations of the vacuum processes.

The partial pressure and composition of the residual gases are determined by mass spectrometers (for example, KM-1), the operating principle of which is based on the interaction of the moving charged particles with electromagnetic field. In the chamber of the mass spectrometer the residual gases are ionized by electrons emitted by the incandescent cathode. In an electromagnetic field between two parallel plates the ions of the residual gases move in a helix and fall on the collector of the instrument. The intensity of the electromagnetic field and its frequency are selected so that the ions with a defined mass will hit the collector, that is, the current will be proportional to the concentration of the particles with given mass and, consequently, the given component of the residual gas. As a result, the spectrogram of the composition of the residual gases is drawn on the recorder tape.

The substrate temperature is controlled by wire thermocouples and thermistors. The operating principle of the thermocouples is based on using the effect of the

FOR OFFICIAL USE ONLY

occurrence of a thermal emf on heating the junction of two heterogeneous wires. The emf measured on the free "cold" ends of the wires depends on the temperature difference of the free and the soldered ends of the wires and their materials. For checking the temperature of the substrates, primarily chromel-copel thermocouples are used (to 600° C). The free ends of the thermocouple are connected to millivoltmeters, regulating automatic bridges or potentiometers. Thermistors which are nonlinear resistors, are built into the substrate holders and are connected to the measuring devices.

The thickness and growth rate of the film is monitored during the process of thermovacuum deposition most frequently by resistive, ionization and resonance-frequency methods. The resistive method is applicable only for films made of electrically conducting materials, and the second two methods, for films made of any materials.

The resistive method is based on direct measurement of the resistance of the monitored sample -- "indicator," placed in the vacuum chamber along with the substrate. The deposition of the film proceeds simultaneously on the working substrates and on the "indicator." The contact areas of the "indicator" are connected by means of spring contacts to an external ohmmeter with the corresponding measurement range providing the required accuracy. The current through the "indicator" must not exceed the maximum value, above which the conditions of formation of the controlled film change.

If the specific resistance of the film remains unchanged during deposition, then the film resistance is defined by the nonrelation

$$R = \frac{\rho b}{hl}, \quad (15-1)$$

where R is the film resistance, ohms; ρ is the specific resistance of the film material, ohm-cm; l and b are the length and width of the film on the controlled sample, cm; h is the film thickness, cm.

Measuring the resistance of the "indicator" in this way, it is possible to determine the thickness and the growth rate of the film, for the deposition times are known. The deposition process is stopped by automatic inclusion of the shutter on reaching the given resistance and, consequently, thickness of the film. The method is used in the KS-1 and KS-2 measuring devices, the measurement accuracy of the thickness by which is no less than $\pm(5-10)\%$.

The ionization method is used in Soviet meters that measure the deposition rate and thickness of the applied films ISTI-1, ISTI-2 (the ionization speed and thickness meter). The ionization method is based on the principle used in an ionization vacuum meter. The ionization sensor is located along side the substrate. A rotating shield isolates part of the flow of evaporated material, and it passes in pulses to the inside volume of the sensor. Along with the flow of evaporated material, the randomly mixing molecules of residual gases continuously pass through the openings of the shield. As a result of ionization under the effect of the electrons emitted by the cathode and accelerated by the anode on the collector, an ion current appears. This current can be divided into two components; a pulse current defined by the flow of evaporated material and characterizing the growth rate and, consequently, thickness of the film; a direct current determined by the pressure in the vacuum chamber of the device.

FOR OFFICIAL USE ONLY

In order to determine the film thickness it is necessary to integrate the deposition rate in the process time limits. For this purpose, current integrators or experimental graphical relations are used.

The ionization method permits continuous monitoring of a number of films without using special control samples. Thus, for example, for the ISTI-1 meter, the total thickness of the applied film without changing the sensor is 100 microns. The sensor is distinguished by low inertia, which permits automatic regulation of the deposition rate using feedback with the evaporator. On deposition of metallic films the method is distinguished by high precision (to $\pm 1\%$) and broad measurement limits (from units to thousands of angstroms per second). However, the measuring circuit is relatively awkward, for each deposited material, its own calibration curve is required; the volume of the sensor is quickly contaminated by the material of the deposited film, and measurements of the dielectric films are difficult.

The resonance-frequency method is based on using the piezoeffect of quartz. A quartz plate cut out of a single crystal in a defined crystallographic plane has the property of changing its dimensions in the electric field cycle (they compress and extend). The quartz plate included in the ac circuit experiences elastic mechanical vibrations. If the electric field frequency is close to the frequency of the natural mechanical vibrations of the quartz f_0 , an electromechanical resonance is observed in which the current amplitude and the amplitude of the natural mechanical vibrations of quartz are maximal.

The variation in mass of the quartz Δm leads to variation in the resonance frequency Δf :

$$\frac{\Delta f}{f_0} = \frac{\Delta m}{m_0}, \quad (15-2)$$

where m_0 is the initial mass of the quartz,

The quartz plate is attached in a water-cooled holder and placed in the vacuum chamber of the device along with the substrates. During the process of growing the film the mass of the quartz resonator increases, which leads to a decrease in its resonance frequency. If the density of the applied film d is known, the thickness h can be determined from the expression

$$h = \frac{\Delta m}{dS}, \quad (15-3)$$

where S is the area of the film. From equations (15-2) and (15-3):

$$h = \frac{m_0}{f_0 d S} \Delta f. \quad (15-4)$$

Thus, measuring Δf , it is possible to determine the thickness of the applied film. However, in practice the film density is not constant; therefore for each deposited material it is necessary to construct a calibration curve.

By using one quartz plate it is possible to measure 20-40 films 1000 Å thick each, that is, the total thickness of the applied film without changing the quartz crystal will be 2-4 microns. The applied layer is removed by pickling, and then the quartz

FOR OFFICIAL USE ONLY

sensor can again be used for measurements. The quartz sensors are universal; they can be used to measure films of any composition and multilayer films.

The resonance-frequency method is used in the Soviet KIT-1 and KIT-2 instruments (control measurement of thickness). When the film mass reaches the required value, the measuring circuit of the instrument feeds the signal to a device that drives the shutter which cuts off the molecular flux. The deposition process stops.

Structure Quality Control. In order to reject structures with clearly expressed defects, the operations of visual control are performed by means of the MBS-1 or other microscopes. The basic forms of explicit defects are as follows: low-quality films (scaling, blistering, humps, scratches, runs, drops of evaporated material, and so on) and deviations from the topology (a shift of the layers with respect to each other, and distortions in the pattern). The visual monitoring makes it possible to exclude an explicit reject from the technological process and at the same time decreases the nonreproducible expenditures of time.

An implicit reject of structures is discovered by measuring the electrical parameters of the elements: the resistances of the resistors, the wiring and the insulation, the capacitance and the dielectric loss tangent of the capacitors. The rejection of structures by parameters not corresponding to the technical specifications excludes the nonproductive expenditures on assembly, encapsulation and other finishing operations of the technological process.

Test Questions and Assignments

1. Repeat § 1-2.
2. What requirements are imposed on the substrate materials and the thin-film elements?
3. What materials are used for thin-film resistors, wiring, capacitors, and by means of what methods can the films of these materials be theoretically obtained?
4. Enumerate the possible flow diagrams for the manufacture of thin-film structures in accordance with the methods of obtaining the films and the configuration of the elements.
5. Give a brief description of the basic steps in the manufacture of structures of thin-film resistive and resistive-capacitive microcircuits.
6. Compare the separate and continuous methods of thermovacuum deposition of films.
7. Using Figure 15-2, explain the technological process of manufacturing a resistive-capacitive thin-film microcircuit.
8. What are the characteristic features and the techniques of thermovacuum deposition of chromium?
9. Using Figure 15-3 explain the technological process of manufacturing a resistive thin-film microcircuit. What are the peculiarities of this process?
10. Compare the technological processes of manufacturing resistive thin-film microcircuits presented in Figures 15-3 and 15-4.

FOR OFFICIAL USE ONLY

11. What are the characteristic features of thermovacuum deposition of films with multicomponent composition, for example, nichrome?
12. Explain the essence of the special methods of evaporation of multicomponent materials.
13. Why is tantalum technology especially isolated among the processes of manufacturing thin-film microcircuits?
14. Explain the technological process of making resistive-capacitive tantalum microcircuits (Figure 15-5).
15. What are the essence and the characteristic features of obtaining contact masks using electron, laser and ion beams?
16. By analogy with Figure 15-3, compile a production flow diagram for making a resistive thin-film microcircuit using projection electron lithography.
17. What requirements are imposed on the materials of the contact masks during transfer of the pattern from the mask to the material of the films by ion pickling? What materials must completely correspond to these requirements?
18. What is electron beam "milling" and what is its advantage?
19. What is smooth and discrete adjustment of the film element ratings?
20. List the production methods of adjustment, by means of which part of the thin film is removed. What is their essence? Which of these methods are group methods?
21. By which adjustment methods is it possible to change the film structure? What is the essence of the adjustment processes?
22. What is the photochemical method of adjustment and what does it consist of?
23. Why is shielding of the thin-film elements necessary? How and by means of which materials is it accomplished?
24. By means of what instruments is the vacuum measured? What is the operating principle of these instruments?
25. How is the partial pressure and composition of the residual gases in a vacuum chamber determined?
26. How is the substrate temperature measured during the film deposition process?
27. What does the resistive method of measuring the film thickness and growth rate consist of? For films made of which materials is it applicable?
28. What is the ionization method of measuring the thickness and growth rate of films and what are its characteristic features?
29. Soviet meters of the thickness of the films KIT-1 and KIT-2 operate on the basis of what method? What is the essence of this method?
30. In what way are explicit and implicit rejects of thin-film structures controlled?

FOR OFFICIAL USE ONLY

CHAPTER 16, MANUFACTURING TECHNOLOGY OF THE STRUCTURE OF THICK-FILM HYBRID IC

16-1. General Information

The rapid development of thick-film technology in recent times has led to the fact that in a number of cases it provides element characteristics close to the characteristics of thin-film IC.

A general diagram of the manufacturing technology of thick-film hybrid microcircuits is presented in Figure 16-1,

Complex, pastelike compounds are used in thick-film technology, the manufacture of which is a responsible event.

The pastes are suspensions of filler powder and glass in any organic binder or solvent. The filler is the base for the paste and lends the films the required resistive, conducting or dielectric properties.

Metals and alloys are used as the fillers of conducting pastes. The fillers of the pastes must have extremely low chemical activity at high heat treatment temperatures in an oxidizing environment and on contact with chemically active glass, and they must also be susceptible and stable (insoluble) to the effect of solder used for solder wiring. This explains the application of noble metals as fillers: gold, silver, gold-palladium, gold-platinum, silver-platinum and other alloys. A comparative estimate of the conducting pastes based on various fillers is presented in Table 16-1.

As is obvious from the table, the most inexpensive, highly conducting pastes having good adhesion to the substrates are the pastes based on silver, but they are subject to corrosion and migration. Pastes based on Pt-Au alloy provide minimum solution in the solder, and therefore they guarantee high reliability of the microcircuits with a large number of soldered connections, but they are the most expensive. Pastes based on Au-Pd alloy are less expensive. Pure gold cannot be used for contact areas, for when soldering the external leads it very quickly dissolves in the solder. The specific resistance and conducting films based on noble metals will be 0.3-0.01 Ohms/ \square .

Along with the fillers presented in tables 16-1 for conducting pastes they have started to use nonnoble metals: copper, nickel, aluminum, tungsten and molybdenum. They provide not only lower cost of the paste, but in a number of cases better

FOR OFFICIAL USE ONLY

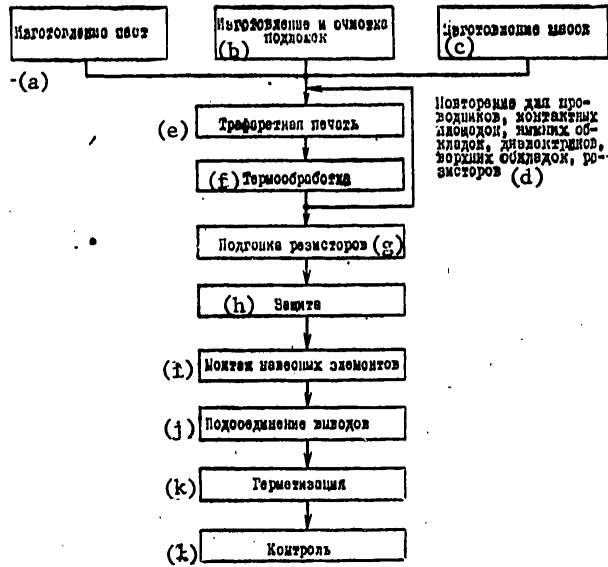


Figure 16-1. General diagram of the manufacture of thick-film hybrid IC.

- Key:
- a. preparation of pastes
 - b. manufacture and purification of substrates
 - c. manufacture of masks
 - d. repetition for wiring, contact areas, lower plates, dielectrics, upper plates and resistors
 - e. stenciling
 - f. heat treatment
 - g. adjustment of resistors
 - h. shielding
 - i. installation of mounted elements
 - j. connection of leads
 - k. encapsulation
 - l. control

Table 16-1, Table 16-1, Properties of conducting pastes¹

Filler	Cost	Electrical conductivity	Adhesion	Resistance to molten solder	Contact resistance	Migration
Au	4	3	4	5	5	2
Pt—Au	5	5	3	3	1	1
Au—Pd	3	4	3	2	3	2
Ag—Pd	2	2	2	3	2	3
Ag	1	1	1	5	4	5

¹An estimate is made on a five-point system; 1 is the highest point.

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

parameters and stability at high temperatures. Copper, for example, is the only metal with high electrical conductivity to which it is possible to connect the external leads both by welding and by soldering. In addition, copper has good adhesion to aluminum-ceramic substrates, high thermal conductivity, resistance to leaching out, to radiation and it has good properties in the microwave range.

For the manufacture of resistive pastes, the fillers can be silver, gold, palladium, platinum, rhenium, the oxides of thallium, ruthenium, rhenium and palladium, and also various compounds; silver-palladium-palladium oxide, silver-ruthenium oxide, bismuth-ruthenium, ruthenium-iridium, platinum-iridium oxide, and so on. The thick-film resistors have rated values of the resistances from one to 10^9 ohms, a specific resistance from 1 to 10^7 ohms/cm² and a wide range of values of the thermal coefficient of resistance.

Mixtures of ceramic materials and fluxes and also glass and ferroelectric materials are used as the fillers for dielectric pastes for capacitors. For example, pastes based on the compound made up of barium titanate, titanium oxide, aluminum oxide and low-melting glass have a dielectric constant from 10 to 2000.

Dielectrics for the interlevel insulation of the wiring are made on the basis of glass-ceramic materials. By comparison with the dielectrics for capacitors, they have lower dielectric constants.

As the permanent binder in the paste there is glass which is not removed after formation of the film but remains in the finished film element. In a conducting paste the glass is used to keep the filler grains at the contact and to insure adhesion of the thick film to the substrate. When selecting the glass composition it is necessary to consider the dependence of its viscosity on the temperature, the wetting of the substrate, chemical activity and coefficient of thermal expansion. These properties of glass influence the heat treatment conditions, the formation of mechanical bonds between the grains of the metal filler, the specific resistance of the film and the processes of joining of the leads to the contact sites. The glass composition to a significant degree determines the stability of the resistor parameters. In a dielectric used for interlevel insulation of wiring, the permanent binder is simultaneously the basic functional component. The glass must not introduce defects into the capacitor dielectrics leading to the occurrence of short circuits between the plates. Low-melting lead-borosilicate, zinc-borosilicate and cadmium glass are used as the constant binder.

A broad selection of materials can be used as the organic binders and solvents for the pastes: ethyl cellulose, wax, lanolin, vaseline, cyclohexanol, liquid resins and organic solvents.

The organic binders and solvents are introduced to insure uniform distribution of the components during the process of preparing the paste, to obtain defined consistency and to lend the paste optimal viscosity. After applying the paste to the substrate the organic binders are completely removed during the heat treatment. In the case of incomplete removal of organic binders in the dielectric layer, for example, there will be carbon which will improve the electrical conductivity sharply.

Special additives are being introduced into the composition of pastes to improve adhesion, solderability and lend the pastes thixotropicity and so on.

FOR OFFICIAL USE ONLY

Thixotropicity is the capacity of the pastes to increase their fluidity under the effect of the mechanical pressure and then after cessation of the pressure effect, to thicken up again. In order to lend the pastes thixotropic properties, high-molecular compounds are introduced into them, for example, furane carboxylic or terephthalic acid.

The substrates for thick-film microcircuits are made of ceramic based on aluminum oxide, beryllium oxide and barium titanate. The alumoceramic substrates are the most widespread. Beryllium ceramics are used to make powerful microcircuits as a result of their high thermal conductivity. Ceramics based on barium titanate are distinguished by high dielectric constant. The application of ceramic substrates for thick-film microcircuits arises from the requirements of the resistance to high-temperature treatment of the film after the application of the pastes (about 1000°C) and also high mechanical strength and resistance to wear. The condition of the substrate surface has a significant influence on the process of applying the films and formation of the thick-film elements.

16-2. Basic Steps in the Manufacturing Technology of the Passive Part of the Structures of Thick-Film Hybrid IC.

Manufacture of Pastes. The process of manufacturing the pastes includes two steps: the manufacture of the components and preparation of homogeneous mixtures of them.

The manufacture of the components consists in obtaining finely disperse powders of the fillers and glass of defined granulometric composition. The finely disperse silver and palladium powders are obtained by chemical reduction from the corresponding solutions at increased temperatures. The silver or palladium obtained by reduction is carefully washed with distilled water. The superfine powders of metals for resistive pastes can be obtained using electrolysis of the solutions so that the cathode deposit will be obtained in the form of a powder.

Glass enters into the composition of the pastes also in the form of a finely disperse powder. After founding of the glass, fritting takes place -- the molten glass is poured into water for cracking or it is sprayed with compressed air. Then it is ground in vibration mills. The process can be intensified by adding a liquid that wets the particles subjected to grinding well. The liquid, penetrating into the microcracks of the particles, has a wedging effect and promotes finer particles. Grinding is also used to obtain finely disperse powders of ceramic fillers for dielectric pastes.

After drying, the powder is screened, is packed in a container and transferred to the operation of obtaining homogeneous mixtures.

The preparation of homogeneous mixtures of micropowders of fillers, glass and organic binders is carried out by mixing in a special paste mixer. During mixing it is necessary to insure local exhaust of air and avoid high speeds, for introduction of air into the paste is possible.

After mixing, the viscosity of the paste is checked. The finished pastes are packed into tightly closed containers. As they are used and during storage the solvents evaporate; therefore it is necessary to correct the viscosity of the paste by adding solvents and by additional mixing.

FOR OFFICIAL USE ONLY

Stenciling. Stenciling is done when manufacturing microcircuits on special automatic printers or semiautomatic machines by two methods: contactless or contact.

In the contactless method the reticular mask is located at a fixed distance from the substrate. The substrate is attached by a mechanical holder or vacuum clamp.

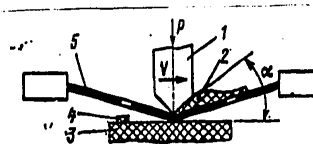


Figure 16-2. Diagram of the method of contactless stenciling. 1 -- doctor; 2 -- paste; 3 -- substrate; 4 -- paste impression ; 5 -- mask; V -- speed of the doctor; α -- angle of encounter of the doctor with the surface of the substrate.

Batched feed of the pastes through the holes in the mask is realized using a special blade -- a doctor (see Figure 16-2). The blade of the doctor is made of other and plastic materials: rubber, elastomer, polyurethane, neoprene, and vitona which are compatible with the pastes used. The blade of the doctor has a thickness of approximately 6 mm and a square or pointed tip. It is important to select the proper angle of encounter of the doctor with the substrate surface so as to insure an optimal force for punching and cutting the paste.

Moving over the surface of the mask, the doctor pushes a wave of the paste ahead of itself and simultaneously applies it to the mask. The mask is bent and touches the substrate. Under the effect of the arising hydromechanical pressure the paste fills the openings of the grid at the points of contact with the substrate, and it is forced through. As the doctor leaves, the mask is straightened, leaving the pattern of applied paste on the substrate. A necessary condition of obtaining a high-quality impression with a sharp edge of the pattern is the use of pastes with good thixotropic properties.

The dimensions of the mask are approximately four times the dimensions of the substrate; therefore during the application of the paste, the working part of the doctor does not reach the edges of the mask. This explains the reversible deformation and uniform bending of the mask in the paste application process.

The hydromechanical pressure on the paste of given viscosity depends on the displacement rate of the doctor, the pressure of the doctor on the mask, the angle of encounter of the doctor with the substrate surface, the gap between the mask and substrate and the mask tangent.

Using these factors, the stenciling process is also regulated:

Increasing the weight of displacement of the doctor increases the hydromechanical pressure and thickness of the applied layer; usually the doctor moves over the mask at a speed of 50-200 mm/sec;

Increasing the pressure of the doctor on the mask increases the thickness of the paste layer; the pressure influences the uniformity of the layer thickness and the

FOR OFFICIAL USE ONLY

quality of the edge of the pattern. Too high a doctor pressure leads to the formation of teeth along the edges of the pattern and can cause rupture of the stencil. Usually the pressure is $7 \cdot 10^3$ newtons/m²;

The angle of encounter of the doctor with the surface of the substrate influences the pressure and, consequently, the thickness of the layer. The smaller the angle of encounter, the greater the pressure and the thickness of the film. Depending on the properties of the paste and the total area of the openings in the stencil, the angle of encounter of the doctor with the mask surface varies from 4 to 60°.

Increasing the clearance between the stencil and the substrate leads to an increase in thickness of the layer. It is recommended that a clearance of 0.4-0.5 mm be used.

The method of contact stenciling in which the substrate is fastened directly to a foil mask makes it possible to apply thinner lines and elements of smaller area. The dispersion of the parameters of the elements will decrease in this case.

With high viscosity of the pastes, in order to facilitate the extrusion of them, sometimes the mask is heated. The separation of the mask from its substrate is done best by vertical motion without slipping.

With the contact method the paste can be applied by spraying, using a paste sprayer, which is attached vertically above the table moving in the horizontal direction with the substrate located on it.

After stenciling the substrates must be held for sometime at normal temperature so that "setting" of the paste can take place. During setting, the unevenness of the edge of the applied films will disappear, the thickness of the impression will become more uniform. The setting time depends on the type of paste, and it is selected experimentally.

Heat Treatment of Paste. Heat treatment is carried out with the formation of solid films with the required properties. Heat treatment includes two operations: drying and annealing.

Drying is necessary for gradual removal of volatile components from the paste. During fast removal of the solvent bulging appears in the film, cavities and pits can occur. Drying is carried out at temperatures of 80-150° C for 5-15 minutes in furnaces or under infrared lamps. The infrared radiation penetrates deep into the paste layer and dries it uniformly without the formation of a surface crust that interferes with drying.

Annealing of the paste is carried out in continuous-action traveling ovens with quartz muffle. Before entering the furnace frequently an additional attachment is installed for drying so that the entire heat treatment process will be performed on one conveyor belt in one cycle. The layout of a traveling oven and the time temperature conditions are presented in Figure 16-3. The substrates placed on the conveyor belt go into the furnace muffle. Dry air under pressure is blown into the furnace opposite to the direction of the substrates. The air with the gases released during heat treatment exit through the opening in the muffle to the exhaust ventilation system. In order to get the air from the production facility into the muffle, at the exit of the oven an air curtain is set up by means of a fan.

The process of annealing the pastes is carried out in two steps. In the first step the organic binder is burned out of the paste at temperatures of 300-350° C.

FOR OFFICIAL USE ONLY

The visers are burned out with a slow rise in temperature so that through holes will not occur in the films, which is especially undesirable for dielectric pastes as a result of the possibility of forming crusts around these defects. Temperatures higher than 300° C are undesirable, for carbonization processes are possible. The removal of the organic binder during drying must be complete, for the remains of organic materials in the subsequent step of heat treatment lead to the formation of carbides in the conducting paste on the surface of the particles of metallic filler and, consequently, to an increase in the specific resistance of the conductors. A counterflow of dry air removes the products of organic gas generation from the temperature zone and excludes penetration of them into the next zone of the furnace where the second heat treatment step takes place.

The second step is sintering of the pastes at significantly higher temperatures of 700-1000° C (depending on the type of paste), and it is more critical with respect to temperature fluctuations not only along the oven, but also in the transverse cross section. The second step is most responsible, for the chemical reactions between the paste components, the paste and substrate, the paste and the environment inside the furnace muffle, between the components of different pastes occurring during sintering determine the basic electrophysical properties of thick-film elements.

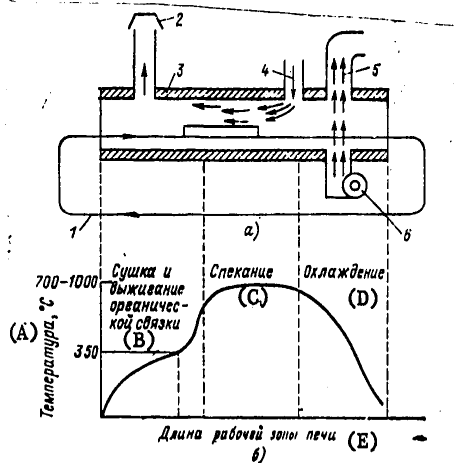


Figure 16-3. Diagram of a traveling oven for heat treatment of pastes (a) and graph of the temperature distribution along the oven (b).
 1 -- conveyor belt; 2 -- ventilation system; 3 -- quartz muffle (working zones); 4 -- blowing with dry air; 5 -- air curtain; 6 -- fan.

Key: A. temperature, °C
 B. drying and burning out organic binder
 C. sintering
 D. cooling
 E. length of the working zone of the furnace

During heat treatment the particles of the metal fillers are sintered and form larger grains. With an increase in the treatment temperature as a result of better sintering and improvement of the quality of the contact between the particles the specific resistance of the conducting pastes decreases. At higher temperatures

FOR OFFICIAL USE ONLY

melting takes place and as a result of agglomeration of the particles and the formation of a melt specific resistance rises.

The metal fillers of the pastes can enter into a chemical interaction with the permanent binder, forming oxides. These oxides influence the nature of the contacts between the filler particles, and the specific resistance of the conducting pastes, as a rule, increases.

The environment inside the furnace muffle can vary from oxidizing to reducing. During sintering in an oxidizing environment, the resistance of the conducting pastes increases and the solderability becomes worse as a result of the formation of oxides on the surface of the filler particles. The reducing environment in practice does not change the filler properties. In the air environment usually used for heat treatment, along with oxides, nitrides are formed on the surface of the filler particles which also increase the film resistance. In an air environment at temperatures above 875° C the following reaction takes place in a resistive paste based on palladium oxide



In resistive pastes based on palladium, reactions of its interaction with glass oxides are possible:



By comparison with a conducting paste, in a resistive paste there are a larger number of components, correspondingly larger number of probable chemical reactions. Therefore it is necessary to maintain the heat treatment conditions with high precision.

As a result of the chemical interaction between the paste components and the substrate material, as a rule, good adhesion of thick films to the substrates is insured.

Thus, the sintering of the paste is a complex process accompanied by chemical reactions. Proper selection of the composition of the environment and the sintering conditions permits control of the reactions occurring in the paste and make it possible to obtain microcircuit elements with the required properties.

After sintering the substrates are cooled and removed from the oven conveyor belt.

The process of heat treatment of the pastes plays the decisive role in insuring the basic parameters of the thick-film elements: conductivity and the capacity for solderability of the wires, specific capacitance of the dielectrics, specific resistance, thermal coefficient of resistance and noise of the resistors, time stability of all parameters. The stenciling and heat treatment procedures have special significance. Usually the pastes are stenciled and subjected to heat treatment in the following sequence: wiring, contact sites, lower capacitor plates, dielectrics, upper plates and resistors. This decreases the effect of the repeated heat treatments on the resistor properties, for the heat treatment procedure in this case corresponds to a decrease in temperature. After the resistive paste it is

FOR OFFICIAL USE ONLY

possible to apply only dielectric protective films and subject them to heat treatment.

Adjustment of Thick-Film Elements. After heat treatment the resistors have resistance dispersion of about $\pm 10-20\%$; therefore in order to insure the small tolerances on the rated values of the resistances it is necessary to adjust them. The adjustment of the thick-film resistors and capacitors theoretically does not differ from thin-film, and it is achieved either by varying the configuration of the elements or by varying the properties of the films.

Air-abrasive adjustment consists in the fact that the air jet bearing the abrasive particles, for example, aluminum oxide, cuts off part of the resistor or the upper capacitor plate. The width of the cut can be regulated by varying the flow of abrasive particles, the speed of the air jet, and the air pressure. However, it is difficult to obtain an abrasive jet with small cross section; therefore the method of abrasive-air adjustment can be used only for resistors with no less than $3-4 \text{ mm}^2$ of area.

Air-abrasive adjustment of the capacitors is a more complicated operation. Usually not all of the capacitor material is removed to the substrate, but only the required part of the upper plate of the capacitor, without removing the dielectric. This is necessary to maintain the electric strength and the reliability of the capacitors. For automation of the adjustment process it is necessary to select an abrasive material which has the property of selective effect only on the material of the upper electrodes but not on the dielectric.

The air-abrasive adjustment has a number of deficiencies; low precision, inapplicability to small elements, complexity of removal of abrasive from the substrates after adjustment, awkwardness of the systems for simultaneous adjustment of several elements, the possibility of mechanical damage by a number of arranged elements, and low output capacity (tens of millimeters per minute). In spite of this fact, it is comparatively widely used for adjusting thick-film elements.

The laser method of adjustment is distinguished by high efficiency, and it insures high precision. Just as the adjustment of thin-film resistors, laser adjustment of thick-film resistors is realized by a series of pulses. The intervals between the pulses are used to measure the resistances of the resistors, for the plasma formed during laser treatment lowers the true value of the resistances. As a result of a number of significant advantages by comparison with air-abrasive adjustment, the laser adjustment, in spite of the high cost, plays the leading role at the present time.

The process of adjusting the resistors and the capacitors is accompanied not only by variation of the required parameters, but also variation of the thermal coefficient of resistance, the noise level, power and moisture resistance. The variation of the thermal coefficient of resistance is connected with the nonuniformity of the composition of the thick film with respect to the area of the resistor. The noise level of the resistors varies as a result of the formation of unevenness of the edge of the film during adjustment. During abrasive-air adjustments the mechanical disturbances of the upper dense layer of films leads to worsening of the moisture resistance of the elements. After laser adjustment, as a result of melting of the edges, the film elements are less subject to the effects of atmosphere and moisture.

FOR OFFICIAL USE ONLY

The modern methods insure an adjustment precision for the rated values of the resistors of $\pm 1\%$ and less.

Shielding of the Thick-Film Elements. Glazes which are also applied by the stenciling method are used to shield the thick-film resistors and contact sites from moisture and the ambient gas environment. Low-temperature glass with a high lead content is most widely used. In order to increase the resistance of the glaze to chemical effects, refractory compounds are added to the glass. The regulation of the TKLR of the glass is realized by introducing zirconium oxide. The process of heat treating the shielding coating is carried out at temperatures of no more than 500° C. At higher temperatures, changes in the resistor parameters are observed. Properly annealed shielding films are semitransparent and have a frosted surface. Different colored shielding pastes are used for convenience in sorting various structures during assembly operations.

16-3. Quality Control When Manufacturing Thick-Film Microcircuits

Appearing at first glance to be simple, thick-film technology is highly sensitive to the accuracy of observation of the conditions of performing all of the technological process operations and, especially, the application of pastes and heat treatment and also stability of the paste properties, the composition of the gas medium in the muffle of the oven, and so on.

During the process of manufacturing thick-film circuits, the pastes are checked for correspondence of the certificate data to the rated values obtained in practice. Test structures are used for this quality control. In addition, the viscosity of the paste is monitored by a viscosimeter.

During stenciling the precision of superposing the device and fixing the mask position with respect to the table for the substrates has important significance. Decreasing the mask-substrate gap leads to an increase in resistances of the resistors, for in this case the mask returns to the initial position more slowly after the doctor has left, and more paste remains on the grid, and the thickness of the impression decreases. The dispersion of the mask-substrate distance can be caused by scattering of the thickness and curvature of the substrates; therefore in order to decrease the dispersion of the resistances and the capacitances it is necessary to check the substrates.

During stenciling the tension of the mask grid, the pressure and angle of inclination of the doctor and the displacement rate of the doctor are also checked.

The quality of the performance of the stenciling operation is checked visually. The basic forms of defects in the thick-film impression, their causes and ways of preventing them are presented in Table 16-2.

The adhesion of thick films to substrates is estimated by the force of separation of the copper wire soldered to the control area of the film. The separation of the wire from the film is realized with a defined rate on a special attachment.

The erosion resistance of the conducting films is estimated by the time during which complete or possible solution of the film in molten solder takes place.

FOR OFFICIAL USE ONLY

Table 16-2. Defects during stenciling

Defect	Cause	Prevention
Cerrated edges	Too much pressure on the doctor	Decrease pressure
Broken line	Sharp step on the substrate	Change the inclination of the doctor
Nonuniform thickness of impression	Poor installation of the mask	Set the mask more precisely
Incomplete line width	Large mask-substrate spacing	Decrease the distance
Regular distortion of the shape of the impression	Insufficient grid tension	Replace the mask

As has already been noted, a necessary condition of obtaining minimum deviations from the rated values of the element parameters is exact maintenance of the temperature and composition of the gas environment during the heat treatment process. It is necessary periodically to check the temperature profiles of the furnace, the accuracy of installation of the thermocouples, the moisture and composition of the environment, the speed of movement of the conveyor belt.

The parameters of thick-film elements depend on a large number of factors acting simultaneously. Even small deviations of the temperature or composition of the gas atmosphere can lead to a great deal of dispersion of the parameters. Obtaining the thick-film elements with small tolerances is connected with great difficulties.

Test Questions and Assignments

1. What components are the pastes made of and what is their significance?
2. What materials are used as conductive, resistive and dielectric fillers for the paste? What are their properties?
3. What is the role of glass organic binders, solvents, special additives in the paste composition?
4. What is thixotropy and what is its role in thick-film technology?
5. Substrates of any materials are used for the thick-film microcircuits? Is it possible to apply substrates made of glass and pyroceram? Why?
6. How are the finely disperse filler powders obtained?
7. What is fritting and what is it used for?
8. What is the technique for preparing homogeneous mixtures of paste components?
9. Which masks are used for stenciling? What is the technique for making them?
10. What are the contact and contactless methods of stenciling, what masks are used for them?

FOR OFFICIAL USE ONLY

11. Explain the technique for stenciling by the contactless method (Figure 16-2).
12. Why is the working part of the reticular mask approximately four times smaller than the overall area of the mask?
13. What factors influence the quality and efficiency of contactless stenciling and how?
14. What is the technique for contact stenciling?
15. What is the role of the "setting" of the pastes after stenciling?
16. What is the purpose of heat treatment of the paste?
17. Explain the temperature graph and the role of each heat treatment step (Figure 16-3).
18. How is it possible to regulate the parameters of the elements obtained in the phase of sintering of the paste?
19. What must the sequence be for the application and heat treatment of pastes?
20. How are the rated values of thick-film elements adjusted?
21. Compare laser adjustment with abrasive air adjustment.
22. What materials are used to shield thick-film microcircuits?
23. What is the technique for applying shielding coatings?
24. List the control operations used when making thick-film microcircuits.
25. Analyze the data in Table 16-2.
26. Draw a general conclusion with respect to the technological process of obtaining thick-film IC.

FOR OFFICIAL USE ONLY

CHAPTER 17. ASSEMBLY OF MICROCIRCUITS

17-1. Separation of the Plates and Substrates with the Finished Structures

Basis Assembly Operations. The assembly of the microcircuits includes approximate separation of the plates and substrates with the finished structures into crystals or plates; mounting the crystals and the suspended components on the passive part of the hybrid IC; mounting the crystals or plates on the base of the cases, the seats for the lead frames and for additional substrates; realization of the inter-element connections and the connections of the contact sites of the integrated structures with external leads of the microcircuits.

The assembly technology at the present time lags behind the manufacturing technology of the structures, for the majority of the presented operations are performed by individual methods.

The basic features of the assembly operations are as follows: very small dimensions of the connected objects, complexity of matching them connected with this and also high requirements on the quality of the mechanical and electrical contacts under the condition of shallow penetration into the connected regions of the structures (for example, in the emitter region).

The assembly of the microcircuit begins with the operation of separation of the plates and substrates with the finished structures. The basic features of the operations of separation of the substrates are as follows: the necessity for exact orientation of the cutting lines between the sites occupied by single structures; insurance of integrality of the elements and metallization of the microcircuits as a result of reduction of mechanical stresses after the cutting line to a minimum; maintenance of the orientation of the separated structures and insurance of reproducible precision of their dimensions and geometric shape, which is highly important for automated welding; obtaining the minimum width of the cutting line to increase the yield of the structures with given substrate area.

The separation of the substrates and plates into plates and crystals is accomplished by abrasive cutting, scribing and by using a laser beam and chemical pickling.

Abrasive Cutting. The essence of the processes of abrasive cutting was described in Chapter 3. In order to separate the plates and substrates with finished structures, sets of wires, panels, steel discs and abrasive suspension or discs with external diamond-containing cutting edges are used. The abrasive methods of cutting are universal, for they make it possible to separate the substrates from any materials.

FOR OFFICIAL USE ONLY

The cutting is done after preliminary orientation using special optical systems. In order to decrease the harmful effect of the abrasive suspension the plates and substrates are applied approximately to transparent plates (with the lines plotted on their grid) on the active side.

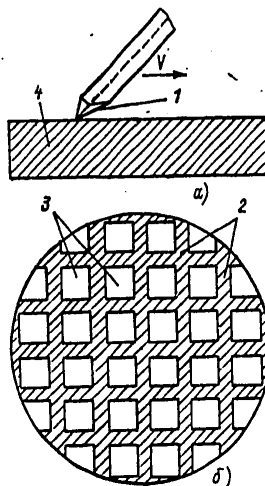


Figure 17-1. Scribing by a diamond cutter. a -- application of lines; b -- plate with lines; 1 -- cutting edge of a knife; 2 -- cutting lines; 3 -- individual crystals; 4 -- weight.

The advantages of abrasive separation include the following: high yield of good crystals or plates reaching 98-100%; proper geometric shape of the crystals or plates with strictly vertical side faces which is important for automation of assembly; reproducible accuracy of the dimensions.

Scribing. The process of obtaining separate structures in this case is carried out in two stages. First using a cutting tool with diamond working part the lines are made (the scribing is done). A diamond cutting tool is moved at a defined speed over the surface and transmits the required force to the separated material (see Figure 17-1). After the cutting tool passes a line is left on the surface. Along the lines in the body of the material mechanical stresses appear, and deformations and microcracks occur which weaken the strength of the plate or the substrate. After the working stroke the cutter is raised, the table with the separated sample is returned (the idle stroke), a second line is drawn with proper spacing, and so on. Then the table is rotated by 90° , and the scribing is continued until the complete line grid is obtained.

The second stage -- separation of the plates or substrates by breaking along points weakened by lines -- is carried out either using the spring-loaded roller on an elastic, thick support (rubber) (Figure 17-2) or using a hemisphere. In order to maintain the mutual position of the structures after separation, the plate or the substrate is coated on the back side with fast-drying KhSL lacquer, and it is placed between two flexible transparent films made of polyethylene or fluoroplastic.

FOR OFFICIAL USE ONLY

When using the roller it is necessary to match the direction of its rolling to the direction of the lines. The plates and substrates are first divided into strips and then, on changing direction of motion of the roller, into crystals or plates. On separation by the roller, breaking is possible not along the lines as a result of nonparallelness of motion of the roller and the direction of the lines, but also as a result of unfavorable distribution of the mechanical load with respect to the separated surface.

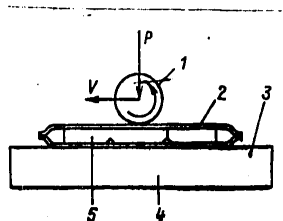


Figure 17-2. Separation of scribed plates using a spring-loaded roller.
1 -- roller; 2 -- polyethylene film; 3 -- support; 4 -- separation line;
5 -- plate.

Separation in a hemisphere (Figure 17-3) takes place simultaneously with respect to all lines as a result of compression of the plate or substrate with respect to the spherical surface with the help of an elastic diaphragm that stretches well. The pressure from the diaphragm is transferred to the separated substrate sample by graphical means or compressed air.

Separation must be carried out directly after obtaining the lines, in order to decrease the crack propagation from the mouth of the lines in the lateral directions.

The quality of scribing depends on the shape of the diamond tip, the slope, the pressure and the speed of displacement of the cutting tool, the properties of the separated material, and for single crystals, the crystallographic orientation and direction of the working stroke of the cutting tool.

For scribing cutting tools are used with working parts in the form of three-sided and four-sided pyramids, a four-sided truncated pyramid and also four-sided truncated pyramid with rounded cutting edges and cutting tools with curvilinear cutting edge. The finish of the working surfaces must be no less than class 12 roughness. The proper choice of the shape of the working part and the slope of the cutting tool insure the decrease in forces transmitted in the lateral directions and lower dependence of the results of cutting on the load vibrations and other factors. The load depends on the separated material, its thickness and the degree of wear of the cutting tool. In Table 17-1 values of the loads on the cutting tool are presented for semiproductive plates.

When scribing with a worn tool it is necessary to increase the load. In order to decrease the wear of the tool when separating plates with silicon dioxide or glass films applied to their working side free tracks without coatings are made in the latter, usually with the help of photolithography.

FOR OFFICIAL USE ONLY

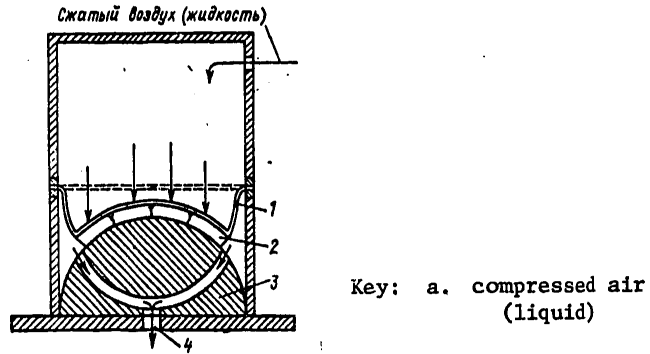


Figure 17-3. Separation of the scribed plates in a hemisphere. 1 -- flexible diaphragm; 2 -- plate; 3 -- hemisphere; 4 -- air exit.

Table 17-1. Load on the cutting tool during scribing

Plate material	Plate thickness, microns	Load, N
Silicon	Less than 200	0.4-0.6
	300-400	0.8-1.2
Germanium	Less than 200	0.2-0.4
	300	0.4-0.6
Gallium arsenide	300-400	0.4-0.6

The speed at which the cutting tool moves is selected considering the efficiency of the process. If the speeds are too high, the quality of separation becomes worse. In order to keep the chips of separated material from getting into the active region of the structures, the scribing is done at moderate speeds and light loads when the line is formed as a result of plastic flow.

The crystallographic orientation of the plates and the orientation of the lines influence the shape of the crystal. The most favorable separation planes for silicon and germanium are the cleavage planes having (111) orientation. These planes intersect each other at angles of $70^{\circ}53'$. If the direction of the lines is not parallel to the cleavage plane or other crystallographic planes, then the separation line of the plate can be zig-zag. In the case of oriented scribing (111) of the silicon plates, the lines are perpendicular to the (110) plane, and a second row is parallel to the (110) plane (Figure 17-4). Here it is possible to obtain a crystal having a lateral cross section -- parallelogram. The acute angles between the sidewalls and the upper and lower surfaces are $70^{\circ}53'$ for separation takes place along the cleavage planes. The front and rear walls of the crystal are perpendicular to the base, for separation takes place along the (110) plane.

The basic advantages of the scribed method are as follows: small width of the lines and absence of the saw cut which saves on the material consumption and increases the efficiency of using the area of the plates and substrates; the absence of an abrasive suspension and adhesives; simplicity and high efficiency of the process.

FOR OFFICIAL USE ONLY

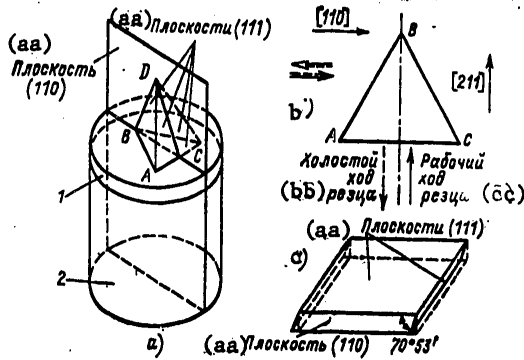


Figure 17-4. Influence of the crystallographic orientation of the plate (a) and direction of the cutting (b) on the shape of the crystal (c). 1 -- semiconductor plate; 2 -- ingot.

Key: aa. plane
 bb. idle of the cutting tool
 cc. working stroke of the cutting tool

The basic deficiencies are as follows; low accuracy of the geometric dimensions of the crystals obtained; dependence of the quality of separation on the ratio of the crystal dimensions and thickness of the separated plate. The minimum size of crystals a and thickness of the plate h are related by the expression $a = kh$. For silicon $k = 4$; for germanium $k = 32$.

Laser Separation of Plates and Substrates. Separation by laser radiation is among the contactless methods in which there is no mechanical effect on the machined material. Separation can be carried out either with preliminary obtaining of lines (laser scribing), or by passage through the entire thickness of the material (laser cutting).

The formation of the lines takes place as a result of evaporation of material by the high-power focused laser beam. In through cutting there is also fusion.

The application of laser scribing permits a fourfold to fivefold increase in efficiency of the process by comparison with diamond scribing. As a result of the great depth of the lines (40-50 microns) the percentage yield of good structures after breaking increases. Using a laser beam it is possible to cut through the oxide and the metal layer. As a result of absence of mechanical effect there are no microchips or microcracks. The fusion of the material along the edges of the line decreases the probability of peeling of the film coatings. The crystallographic orientation has no influence on the quality of separation or shape of the crystals. The crystals obtained have in practice vertical lateral surfaces which greatly facilitates automatic assembly.

The line dimensions are determined by the diameter of the laser beam, the speed of displacement of the plate or the substrate relative to the laser beam, power, frequency and duration of the radiation pulses.

FOR OFFICIAL USE ONLY

The deficiencies of laser separation include the following: high cost and complexity of the equipment, contamination of the surface of the structures by products of evaporation and melting, the occurrence of a deformed zone in the silicon with structural disturbances and a defective zone, in which changes in the properties of the material take place and under the effect of laser treatment.

Chemical Separation. Separation by through chemical pickling is used comparatively rarely for semiconductor structures and provides the corresponding preliminary masking of the surface for which the photolithography process is undertaken.

17-2. Basic Assembly Methods

Soldering. Three basic methods of joining parts are used for assembly: soldering, welding and bonding.

Soldering is the process of joining two parts in the solid state using the molten material — solder. The solder must wet the joined surfaces well, spread, filling the entire space between them.

When heating the solder to the melting point, the following take place simultaneously between the solder and the joined parts: solution of the joined materials in the liquid solder, diffusion of the solder into the joined materials with the formation of a solid solution, chemical interaction of the solder with the materials of the joined parts with the formation of intermetallic compounds. During cooling the solder crystallizes on the surfaces of the joined parts, entering into a strong metallic bond with them.

The solder must not change its properties at the operating temperatures of the IC (25°C for silicon and 85°C for germanium). The temperature of obtaining the soldered joint must be as low as possible in order not to have negative effects on the parameters of the finished products.

Depending on the melting point of the solders, low temperature solder (to 450°C) and high-temperature solder (above 450°C) are distinguished.

Soft or low-temperature solders include tin alloys with lead (POS-40: 40% Sn+60% Pb; POS-61: 61% Sn+38.2% Pb+0.8% Sb), tin with bismuth (POVi-05: 99.6 to 99.4% Sn+0.4 to 0.6% Bi). The solid or high-temperature solders include alloys based on silver (PSr-45: 45% Ag+30% Cu+25% Zn; PSr-72: 72% Ag+28% Cu). Basically low-temperature solders are used for microcircuits.

The quality of joining by solder is determined by the quality of cleaning the surface of the joined parts and the solder to remove contamination and oxides and also the choice of the design of the joint.

In order to improve the wetting fluxes are used which remove the surface oxides during the soldering process, prevent new oxidation and lower the surface tension of the solder. When manufacturing microcircuits, acid-free anticorrosive fluxes based on colophony, types FKSp, FPET, FKTS and based on zinc chloride FKHTs, are used. However, during soldering the fluxes introduce contamination; therefore an effort is made to use them in rare cases. In the majority of cases the soldering is done without flux, but in a reducing (hydrogen, formigas) or inert environment (argon, krypton and helium). During soldering hydrogen forces other gases out of

FOR OFFICIAL USE ONLY

the soldered joint, and then is easily removed itself during degassing of the microcircuits. Hydrogen must be purified of moisture (dew point no more than -50 to -60° C) and oxygen (no more than 0.003 to 0.005%). Hydrogen is explosion hazardous; therefore when soldering in large spaces -- furnaces -- formirgas is used (a mixture of 85% nitrogen and 15% hydrogen).

The parts are butt and lap joined and also joined by combination methods. The mechanical strength of the joints determined by the method and the conditions of soldering, the preparation of the surfaces, the strength of the solder in the joint, the strength of binding of the solder to the material of the joined parts, the presence of intermetallic compounds in the joint, the strength of the joined materials in the joining zone after cooling. The strength of the butt joints is lower than lap joints.

The advantages of soldering are simplicity of the process, the absence of fusion, relatively low heating of the joined parts, the possibility of joining parts of complex configuration.

In order to perform the operations of assembly of the microcircuits, the soldering is done in hydrogen furnaces; by an electric soldering iron; by heat transfer from a tool that is pulse-heated by an electric current; ultrasound; electrical resistance as a result of joule heat released at the point of joining the parts; submersion in solder; radiation.

Welding. Welding is gradually displacing soldering for the assembly of microcircuits. This is explained by the high quality of welded joints and the slight effect on the parameters of the structures. Welding is the process of joining two parts without the participation of solder as a result of bringing them to the distance of atomic effect. During welding the following states of the surface materials of the joined parts are possible: plastic deformation, fusion and plastic deformation, fusion and subsequent crystallization.

Welding can be done with heating or without heating, in the presence or absence of a compressive force and also with simultaneous effect of heating and compressive force.

In order to perform the operations of assembly of the microcircuits most frequently the following welding techniques are used: thermocompression, indirect pulse heating, ultrasonic, double electrode, laser spot, electron beam.

Bonding. Bonded joints do not require complex equipment, they are easy to make, but they do not always provide good contact quality. This explains their application primarily for microcircuits operating under nonharsh operating conditions.

Current nonconducting and current conducting (contactols) adhesives are distinguished. In order to obtain the joints, the adhesive is applied in batches to the surfaces, they are bonded to contact and in the absence (or without) compressive forces the glue hardens. During hardening the glue shrinks. The mechanical stresses occurring in this case provide for drawing the parts together and a tight mechanical joint of the parts. The current nonconducting glues insure high mechanical strength of the joints, which increases in the presence of an external compressive force. However, in this case the probability arises of transmission of an electric current through the metal contact spots of the parts, by tunnel

FOR OFFICIAL USE ONLY

infiltration of the electrons through the thin interlayer of glue and also through the conducting channels obtained by diffusion of the metal through the defective sections. The joints using contactols have high electrical conductivity but less mechanical strength. An increase in mechanical strength is achieved by decreasing the metal filler content.

17-3. Installation of Crystals and Plates

Direct Mounting Method. The direct method, that is, with the working surface up, is used to mount structures of the microcircuits based on cases, the seats of the lead frames (strips) or for additional contact substrates. The mounting of the mounted active elements on the passive part of hybrid microcircuits can also be carried out directly.

The basic requirements on the mounting operations are as follows: insurance of high mechanical strength of the joints, good heat removal from the structure and in a number of cases, good electrical conductivity. The temperatures and the compressive forces when doing the mounting must not be too high so as not to disturb the previously obtained joints, not have a negative influence on the structure parameters, and not destroy their mechanical integrality. At the same time they must be sufficient for strong joining. The elements of the microcircuits take up only the surface part of the crystals or are located on the surface; therefore the operations of direct mounting are not critical to the depth of penetration of the connecting joint.

Electrical insulating connections are made using current nonconducting adhesives, glass or special compounds.

The joints by adhesives and compounds are quite broadly used for mounting, for they are distinguished by simplicity of the process, low hardening temperatures and sufficient mechanical strength and reliability. By bonding it is possible to join various materials of different thickness. The bonded joints simplify the structural design, they increase the weight, and they save with respect to the consumption of expensive metals.

The bonded surfaces must be carefully degreased and dried well to complete removal of the solvents, for later when the glued joint hardens the remains of the solvent will lead to the appearance of porosity and mechanical stresses that lower the strength of the joint. Sometimes vacuum annealing is used to clean the surface of the bases of the case or the lead frames.

The thickness of the layer of glue applied to the joined parts must be small, for with an increase in thickness the strength of the joint decreases. The glued mounting is done in holders when the spreading of the glue is small and an external compressive force is required or it is done without a holder when the glue spreads well over the surface and external pressure is not required. The heat treatment of the glue, as a rule, is carried out in two steps: first for complete removal of the solvent, then at a higher temperature, for hardening. The methods and conditions of heat treatment are determined by the composition of the glue and the structural design of the joined parts. Epoxy resin ED-5, glue VK-2 (a solution of organosilicon resin in an organic solvent with finely disperse asbestos), the glues K-400, K-350-61, VK-4, VK-8, VK-32-200, and so on are used at the present time for mounting crystals and plates. These adhesives provide good strength at

FOR OFFICIAL USE ONLY

temperatures to 300° C, which permits subsequent joining of the leads without rupturing the glued joint.

The deficiencies of glued joints include complexity of repair (replacement of structures) and poor heat removal.

Joining by glass provides good matching of the TKLR, it permits simultaneous joining of the structure and leads to the frame. The basic deficiencies of joining by glass are the high temperatures of the process (about 500° C) and poor thermal conductivity of the glass. It is recommended that the mounting with glass be used for the structures of small IC.

The electrically conducting joints are basically used for mounting crystals on the base of metal cases. The insulation of the crystal is not required, for it always has low potential during the operation of the microcircuit.

The mounting of the silicon crystals can be done by electric soldering and in some cases by welding or joining by contactols.

Eutectic soldering is the most widespread method of mounting silicon crystals. Most frequently eutectic soldering is done using gold-silicon or gold-germanium solders having a melting point of 370 and 356° C, respectively. Let us remember that alloys are called eutectic in which crystallization of the components takes place simultaneously over the entire volume at the lowest eutectic temperature for the given system. During cooling of the solder at a time corresponding to the eutectic temperature, a complete discontinuous transition takes place in both alloy components from the liquid to the solid phase. The hardening is accompanied by the release of fine crystals well bound to each other and this provides for a high-quality soldered joint.

Eutectic soldering can be done using solder inserts (or balls) or by direct contact joining of a silicon crystal to the seat to which a layer of gold 6-9 microns thick has been applied in advance. At the present time contact soldering is most widely used. The nonoperating side of the semiconductor substrate is also coated with a gold film or gold with alloying admixtures corresponding to the electrical conductivity of the crystal in order to improve the wetting. The gold film is melted, and then the plate is divided into crystals. By using a vacuum capturing capillary and a matching system the crystal is approximately installed in the seat. Heated inert gas (nitrogen, argon) or formigas is fed to the joining point. The optimal conditions of eutectic soldering of a silicon crystal are as follows: temperature 390-420° C, time 3-5 seconds, pressure $(3-5) \cdot 10^7$ newtons/m². For intensification of the soldering process, a compressive vertical force and horizontal ultrasonic vibrations, vibrations of industrial frequency or mechanical vibrations with a frequency of 4-8 hertz are used. This provides reliable contact of the joined parts, the eutectic forms quickly and uniformly over the entire area.

Eutectic soldering is widely used for automated mounting of crystals on metal strips or cranes, the seats of which are coated with gold. Here group placement of the crystals in the stenciling holder, the holes in which correspond to the location of the seats of the strip or frame will be used. The crystals are brought into tight contact with the strip or frame by special clamps of the upper cover of the holder. Then the holder is put in the furnace and the soldering takes place.

FOR OFFICIAL USE ONLY

For mounting silicon crystals sometimes gold-tin solder (80% Au + 20% Sn) is used with a eutectic temperature of 280° C or POS-61 is used with a eutectic temperature of 185° C. The advantage of these soft solders — low soldering temperature — can turn out to be insufficient if further operations are carried out at a higher temperature. The basic advantage of mounting by low-temperature soft solders is ease of dismantling the rejected crystals.

The mounting using contactols, as was already noted, is distinguished by simplicity of the process. For connecting the mounted active elements to the contact areas of the passive part of the hybrid IC, the contactols K-I, K-II, K-III, VK-20T and so on are used.

"Rotated Crystal" Method. Crystals with rigid three-dimensional leads entering into the composition of the hybrid microcircuits or microassemblies are mounted on the passive part of the working side of the crystal at the bottom. Thus, it is possible to mount the crystals with expanded contacts. In this case the joining protrusions are executed on the contact sites of the substrates. Using volumetric leads or protrusions on the substrates, both the crystal and all the leads are simultaneously connected. The basic difficulties of the installation by the rotated crystal method are reduction of the difference in height of the leads protruding above the crystal or substrate to a minimum and matching them with the contact sites. The difference in height leads to the necessity for creating sufficient deformations for the most protruding balls or columns in order to provide for contact with the lowest protrusions. Here the deformations must not exceed the admissible norms, for the mechanical stresses occurring after installation lead to potential failures of the microcircuits. The matching operation is performed using quite complex optical devices with stereomicroscopes or infrared microscopes with optoelectronic image converters that shift the radiation to the visible part of the spectrum. The majority of semiconductors transmit infrared beams, and metals are nontransparent for these beams; therefore against the substrate background the metal sections are obvious in the form of black spots. In some devices the image is projected on a screen to facilitate matching.

The matching of the rotating crystals with the bar leads is simpler to do, for they go beyond the crystal limits.

The process of joining the volumetric leads will be investigated in § 17-5.

17-4. Wiring

Final Installation Operations. After joining the plates and crystals we have the operations that complete all of the electrical connections between the elements and also between the structures and the external leads of the finished microcircuits. These include the connections of the ohmic contacts of the active suspended elements with the film contact sites of the passive part of the hybrid IC, the contact sites of the IC structures with external leads of the cases, contact sites with contact sites (or volumetric leads) of the monocrystalline IC. All of these joints are made with soldering, welding and bonding (§ 17-2).

Depending on the means used for making the joints, wire and wireless mounting are distinguished. In the case of wire mounting, basically gold and aluminum wire are used. The wireless mounting includes the following: joining the crystals to the volumetric leads, assembly on frame, a strip or flexible carrier (the "spider" lead method).

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

Special Features of Wiring. A basic feature of the joints made using gold or aluminum circular wires is a great difference in thickness of the joined parts. As a rule, the wire diameter is appreciably greater than the film thickness of the contact areas and appreciably less than the diameter of the case leads. The microcontacting processes are therefore highly critical, for the formation of a connecting weld with a thickness comparable to the film thickness of the contact site is possible. The wiring is individual, each connection is made separately. Wiring is one of the weakest points in microcircuit production. Unreliability of the individual wiring assembly is the cause for the greater part of failures of microcircuits. Nevertheless, up to the present time the majority of connections have been made using wires. The greatest reliability of the wiring connections is insured by welding and soldering. Let us consider the varieties of wire connections using these methods.

Thermocompression Welding. Microcontacting by thermocompression welding is done with simultaneous effect of temperature and pressure on the joined parts for a defined time interval. A necessary condition of the formation of a strong joint is plastic deformation in the contact zone. For this purpose the joined wires are selected from the soft plastic metals: gold, aluminum, silver, and so on.

The connection temperature during thermal compression must not exceed the temperature of formation of the eutectic of the joined materials, and it is usually close to the annealing point of the more plastic metal. It is most expedient to heat by a hot working tool, for localization of the heat release in the welding zone is insured.

The pressure is transmitted to the joined materials through the welding torch which is in the shape of a wedge (needle), capillary or "bird's beak" (Figure 17-5). The pressure must provide no less than 30% deformation of the joined material, but no more than 60%. On application of pressure the contacting takes place first at individual protrusion points of the joined surfaces. The deformation of the protrusions promotes an increase in the contact zones and approach of the joined surfaces. Here forcing of the adsorbed gases and surface contamination out of the welding zone begins. In the case of sufficient approach, the physical interaction of the atoms of the joined materials begins as a result of Van der Waals forces. The stronger chemical bonds occur as a result of heating and plastic deformation. Plastic deformations lead to effective rupture of the surface oxide films, which are the basic obstacle to chemical interaction. On rupture of the oxides, sections of the clean surface are denuded, on which the strongest chemical bonds occur. The temperature and pressure are interrelated; they must be selected so as to insure optimal deformations, closeness of the joined parts and activation of the surface atoms to the formation of chemical bonds.

The optimal conditions of thermal compression for each pair of welded parts are selected experimentally beginning with the requirement of obtaining maximum strength of the joint. For this purpose the relations are defined for the strength as a function of temperature and pressure, respectively, for pressure and time constants or temperature and time and also dependence of the strength on time for constant temperature and pressure.

The duration of the thermal compression process usually is fractions to tens of seconds depending on the quality of preparing the joined parts, the properties of

FOR OFFICIAL USE ONLY

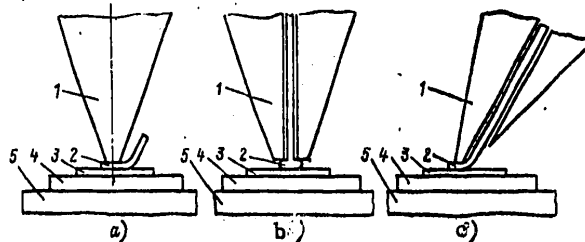


Figure 17-5. Thermal compression tool, a -- wedge; b -- capillary; c -- "bird's beak"; 1 -- tool; 2 -- joined wire; 3 -- joined area; 4 -- substrate; 5 -- welder table.

the joined materials, temperature and pressure. When the optimal time is exceeded, a lowering of the strength of the thermal compression joint is observed. This is explained by the fact that for the creation of a strong contact "fresh" bonds are needed; if the pressure is applied longer, then the process of rupture of the formed bonds begins.

The thermal compression conditions recommended for certain pairs of joined parts are presented in Table 17-2.

Table 17-2. Thermal compression conditions

Welded parts	Temperature, °C	Pressure, $\times 10^7$ newtons/m ²	Lead deformation, %	Time, seconds
Silicon and gold wire	350	14-15	60	5-10
Aluminum film and gold wire	350	10-11	50	0.5-3
Gold film and gold wire	320-340	7-10	50	1-5
Silicon and aluminum wire	450	7	60	10
Aluminum film and aluminum wire	400	6-7	60	1-3
Gold film and aluminum wire	320	6-7	60	1-3
Aluminum film and silver wire	400	18-19	--	5-7
Gold film and silver wire	350	18-19	--	5-7

In order to carry out thermal compression welding the IC structure or holder with IC structures is fastened to the work table of the device. Before welding using the MBS-2 microscope or a special projector and manipulators for displacements, the joined parts are matched. The welding is carried out in the air, but in some cases to protect the structures from oxidation nitrogen, formigas or argon is fed to the welding zone through the work table. The procedure for displacement and separation of the joined wire is determined by the structural design of the welding tool and unit.

It is possible to have butt and lap thermal compression. Butt thermal compression is carried out by a capillary (Figure 17-5,b) with preliminary formation of a ball on the end of the joined wire using a hydrogen burner. Thus, only the gold wires are connected, for the formation of a ball on aluminum and other materials is difficult as a result of their oxidation. When transferring the pressure and

FOR OFFICIAL USE ONLY

heating the gold is split to sizes approximately twice the diameter. After the formation of the joint capillary is raised, it is taken to the case lead (or another part) and the second joint is made lap (without a ball). The wire can be broken by cutting, pulling after making the second loop connection or using a hydrogen burner with subsequent cutting off of the "tail."

Lap thermal compression is carried out by means of a needle or a tool in the form of a "bird's beak" (Figure 17-5,a,c). During thermal compression the wire is fed by a needle using an auxiliary capillary.

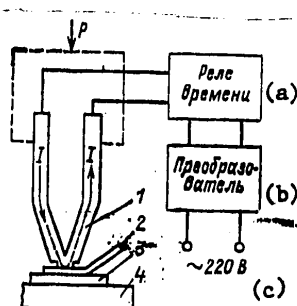


Figure 17-6. Diagram of microwelding by the indirect pulse heating.
1 -- welding tool; 2 -- connected wiring; 3 -- contact site; 4 -- substrate.

Key: a. time relay b. converter c. ~220 volts

The quality of welding is determined to a significant degree by the quality of the welding tool; the end of the tool must be flat and parallel to the contact surface of the site. When working on thermocompression units it is necessary to preserve the tool from pollution and contact with harder materials.

The advantages of microcontacting using thermal compression includes simplicity of the welders, long service life of the tool (to several hundreds of thousands of welds), easy control of the process, stability and low sensitivity to small deviations of the welding conditions. The output capacity of the Soviet units is 180-800 welds/hour.

Indirect Pulse Heating Welding. This method is distinguished from thermal compression welding by the fact that the heating takes place by the direct passage of an electric current through the working tool and only at the time of joining of the parts (Figure 17-6). The released heat, as a result of the structural design of the tool, is concentrated in its lower working part. This permits more exact regulation of the magnitude and the duration of heating of the compared parts. The heating temperature of the tool depends on the fed voltage and the duration of the welding pulse. The conditions of welding by indirect pulse heating are characterized by a primary voltage, pulse duration, the free tool pressure. The duration of the welding process by indirect pulse heating is less than the thermal compression process, for the concentration of heat at the tip of the tool insures more effective heating of the welding spot. This permits an increase in the welding efficiency. The pressure of the welding tool is selected as a function of the diameter or thickness of the welded wire and the ductility of the joined materials. The recommended welding conditions are presented in Table 17-3.

FOR OFFICIAL USE ONLY

Table 17-3. Welding conditions by indirect pulse heating

Welded parts	Wire diameter, microns	Primary voltage, volts	Pulse duration, sec	Compressive force $\times 10^2$, newtons
Aluminum wire and gold film on pyroceram	100	110	0.2	150
Aluminum wire and gold kovar	100	120	0.2	120
Aluminum wire and metalized silicon	30	120	0.2	60
Gold wire and gold film on pyroceram	50	105	0.7	80

At the present time there are a large number of microwelding units that use indirect pulse feeding: "Kontakt-3A," SKIN-1, MKS-02, USP-01, EM-408A, EM-425A, EM-440, EM-441, MS-3R2-2. The USP-01, EM-425A and EM-441 welders with semiautomatic control have an output capacity of 900, 800 and 1600 welds/hour, respectively. The automated EM-440 welder has a capacity of 2000 welds/hour.

Ultrasonic Welding. As a result of the displacement of oxide films and contamination from the weld-affected zone and activation of the surface atoms, longitudinal ultrasonic vibrations intensify the process of joining and essentially improve its quality and reproducibility.

The weld is made using normal pressure and longitudinal vibrations of the tool with an ultrasonic frequency of 20-60 kilohertz. The joined surfaces are subjected to harsh shearing forces, as a result of which the microirregularities are sheared off, the oxide films are ruptured, and they are forced into the gaps between the microprotrusions. The heating as a result of friction not exceeding 30-50% of the melting point of the joined materials, leads to plastic deformation and promotes the occurrence of a direct metallic bond and a strong joint without structural changes in the material in the weld-affected zone.

The welding process parameters for some of the Soviet welders are presented in Table 17-4.

Table 17-4. Characteristics of ultrasonic welders

Welder	Output capacity, welds/hr	Lead diameter, microns	Operating frequency, kilohertz	Welding time, seconds	Force of compression of parts $\times 10^3$, newtons
UZP-03	600	20-50	75±2	0.05-0.5	10-100
EM-424A	800	25-60	66±6.6	0.08-3.6	10-120
MS-41P3-3	1000	20-50	66	0.05-1.5	20-150
UZP-02	2500	24-40	70±5	---	10-20
NPV-2	2400	24-40	70±5	---	10-20

Ultrasonic welding makes it possible to weld parts from different materials, including dielectrics that differ sharply with respect to thickness for low requirements

FOR OFFICIAL USE ONLY

on the quality of preparing the surfaces. In practice ultrasonic welding is carried out most frequently in combination with indirect pulse heating.

Double Electrode Welding. This procedure is a version of resistance welding. The joint is made by passing a high-density electric current pulse through the welding zone and simultaneously chemical compression of the welded parts. The electrodes through which the electric current is fed are located on one side of the welded parts (Figure 17-7). On passage of the electric current the basic part of the joule heat is released in the region of maximum resistance -- at the point of contact of the joined parts. The contacted protrusions of the joined surfaces heat rapidly, they are brought together under the effect of compression, and the contact zone of the parts begins to expand. The contact resistance and current density diminish in this case. The welded surfaces are subject to plastic deformation and melt. The current feed is stopped, and the melt region, not extending to the nearest sections, begins to be cooled, forming the weld. With careful cleaning of the surfaces of the welded parts and the electrodes, a high-quality weld can be achieved which is not inferior to the basic material with respect to strength characteristics.

The welding conditions are determined by the energy and duration of the welding pulse, the pressure on the electrodes and time. For example, the Soviet USDYe welder is characterized by a pulse energy of 4.8-5.2 watt-seconds. The pulse duration is 0.02 to 1 second, the load on the instrument is 0.5-20 newtons with weldable wire diameters from 30 to 150 microns.

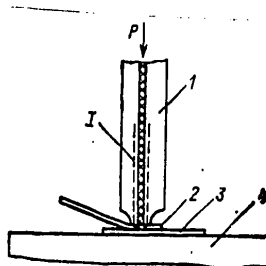


Figure 17-7. Microwelding with dual electrode.
1 -- tool; 2 -- wire; 3 -- contact site; 4 -- substrate.

The heating time and pressure on the joined parts during double electron welding is less than for indirect pulse heating welding. This decreases the harmful effect of the tool on the adjacent regions of the structures.

The electrode material must have high electrical and thermal conductivity with high mechanical strength. For the uncombined electrodes, bonds, tungsten and special alloys are used; for the combined electrodes, the working part is made of copper alloy with tungsten. The roughness of the working surfaces of the electrodes must be no less than class 9-10.

The basic advantages of the method are as follows: a small zone and short time for the thermal effect, high output capacity (to 1200 welds/hour).

FOR OFFICIAL USE ONLY

The deficiencies of the dual electrode welding are as follows; the possibility of current leakage and thermal impact.

Laser Spot Welding. This method reduces to welding by fusion with subsequent crystallization of the fused metal. The welding is done using a laser beam focused to a diameter of 0.25-1 mm with a specific power of 10^5 to 10^6 watts/cm². As a result of nonuniform temperature distribution inside the welding zone the metal boils, a high pressure region is created, and a thermal explosion can occur. In order to prevent spattering of the metal from the welding zone, it is necessary to select radiation that is not too high-powered. In practice, the welding is frequently done by a diverging laser beam, that is, the weld-affected zone is located below the focal plane. The elimination of metal splashes is one of the basic production problems during laser beam welding. The pulse duration of the laser irradiation τ must be less than the time required for the beginning of fusion of the film material and more than the time required for fusion of the wire connected to the film contact site: $t_{\text{melt}} > \tau > t_{\text{wire}}$.

By using a laser beam it is possible to do the welding near the elements of the microcircuits, glass or ceramic insulators. By a laser beam it is possible to perform remote welding in a vacuum, in high pressure chambers and in other environments. The welded connections obtained using a laser beam have smaller (by approximately an order) transient resistance by comparison with the joints obtained by pressure welding.

Table 17-5. Laser spot welders

Welder	Maximum depth of fusion, mm	Welded spot diameter, mm	Duration, part/minute	Maximum radiation energy, joules	Pulse duration, milliseconds	Pulse repetition frequency, hertz
SLS-10-1	0.3	0.4-1.5	30	8	2-4	0.1-0.5
"Kvant"	0.5	0.4-1.5	60	15(30)	4	0.1-1
"Kvant-16"	0.7	0.4-1.5	30	30	6-7	0.1-0.3

The application of laser welding has been held up for a long time as a result of high cost and deficiencies of the lasers. The characteristics of some Soviet glass laser welders are presented in Table 17-5.

Cathode Ray Welding. Analogously to laser welding, cathode ray welding is accompanied by fusion and subsequent crystallization during cooling after the pulse effect of the beam of accelerated electrons. Welding is realized in vacuum units by an electron beam focused to a diameter of 0.5 mm with a specific power to 2000 kilowatts/cm². The fusion of the contact region under the cathode ray spot takes place as a result of conversion of the kinetic energy of the electrons to thermal energy.

The basic advantages of the method are the following: the possibility of welding many metals and alloys (nickel-nickel, nickel-copper, nickel-steel, nickel-kovar, copper-molybdenum, steel-kovar, kovar-steel, kovar-kovar, and so on); the small thermal effect zone permitting welding during glass insulators and a metal-ceramic junction; the possibility of controlling the beam displacement and automation of the assembly processes; degassing of the weld in a vacuum,

FOR OFFICIAL USE ONLY

The high cost of equipment and duration of the evacuation process limit the application of the method.

Wire Joints Using Microsolder. The process for connecting wires with solder is highly critical with respect to the choice of the solder material, temperature and time of soldering. The latter must be as small as possible, and the solder must not noticeably dissolve the film material and form brittle intermetallic compounds. Soldering, more rarely than welding, is used for wire assembly of semiconductor microcircuits, for it does not provide sufficient stability of the results, and reliability, and it has comparatively low output capacity. On the contrary, soldering is more frequently used for the assembly of the hybrid IC.

In order to increase the strength of the joints, the area of the soldered contact is increased, the wires are fastened by special clamps or the wire joints are made passing through the substrates (Figure 17-8).

The leads are first tinned, and then for soldering to the contact sites, a "split electrode" is used along with a jet of hot inert gas, and an "infrared soldering iron" (soldering by focused infrared radiation). Soldering "by infrared soldering gun" which is among the contactless methods, insures relatively precise localization of the joining point and comparatively high output capacity. For welding the wires to the thick-film contact sites the latter are tinned by a solder wave or the stenciling method. In the former case the liquid solder continuously fed by a pump to a nozzle set at an angle to the horizon, forms water, through the crest of which the thick-film contact sites pass. The tinning by the solder wave is distinguished by high output capacity, it prevents contamination with slag and flux residue, for the process is performed in a moving jet with continuous renewal of the solder. The soldering technology of wires to thick-film contact sites is less critical and provides for obtaining contacts with high strength.

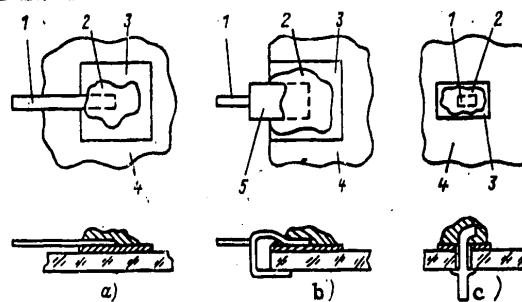


Figure 17-8. Microsoldering of wire leads to thin-film contact sites, a -- surface connection; b -- connection with clamp; c -- connection of a through lead; 1 -- lead; 2 -- solder; 3 -- contact site; 4 -- substrate; 5 -- clamp.

17-5. Wireless Melting

Connecting the Crystals to the Bulk Leads. The bulk leads of the mounted active elements or semiconductor mounted microcircuits combine both the ohmic contact and the lead from it. In § 13-3 it was noted that structures with leads in the form of balls or columns are mounted by the rotated crystal method. The structures

FOR OFFICIAL USE ONLY

with bar leads can also be mounted with working surface up. The development of the structures with bulk leads made it possible sharply to reduce the number of unreliable wire connections, to improve the output capacity and automate the assembly.

After matching with previously grayed contact sites, the bulk leads in the form of balls or columns coated with soft solder are joined by soldering. The surface tension forces of the molten solder attract the crystal exactly to the location. The short circuiting of the leads by the solder does not occur, for the solder does not wet the passivated surface of the crystal and the untinned sections of the substrate on which the crystal is mounted. The solder fills all of the cavities, it eliminates the irregularities, forming a soldered joint on cooling. The crystals connected by the soft solder are easily demounted, large mechanical stresses do not occur in the soldered joint. For connecting the solid solder leads usually thermal compression, welding by indirect pulse heating, ultrasonic welding or combined methods are used. The thermal compression is used primarily for making gold-gold joints. The joining by the rotated crystal method is possible also when using current-conducting adhesives. It is expedient to join the crystals to the bulk leads using the holder method of assembly. Thus, for example, the holder method of assembly used in the EM-432 ultrasonic welder provides an output capacity of 4000 welds/hour.

The bar leads are basically connected by microwelding with ultrasonic intensification. Sometimes soldering is used. In contrast to the ball and column leads, the bar leads remove the heat better from the crystal, they are more easily matched with the contact sites. The compressive force during welding can be applied directly to the leads and not to the crystal. However, crystals with bar leads occupy a large area of the substrate on which they are mounted. In order to decrease the substrate area used by the bar leads, they are arranged in a comb (Figure 17-9).

"Spider" Lead Method. The "spider" leads, that is, planar leads made of metal foil arranged in the radial direction in accordance with the location of the contact sites to which they will be connected, can make up a united whole with the metal frame or strip, or they can be manufactured on a flexible dielectric carrier.

The lead frames or strips are made of thin (25-75 microns) copper, Kovar or aluminum foil using photolithography and through local pickling. The metal frames or strips can be used to obtain the external leads of the microcircuit case from the contact sites of the structures and also for installation of the mounted elements. In the former case (Figure 17-10) after simultaneous connection of all of the leads to all of the contact sites it is possible to encapsulate the structure and cut off the excessive part of the frame. The end of the "spider" lead free after it is cut off is the finished microcircuit lead. In order to lend rigidity to the structural element, the structure is first mounted on an auxiliary dielectric substrate.

In the latter case (Figure 17-11) the free ends of the "spider" leads after making the joint with the contact sites of the mounted element and cutting the frame are shaped and connected to the contact sites of the passive part of the microcircuit.

The flexible dielectric carrier — strip with edge perforation (for automatic feed) — is made of polyimide, lavsan, polyester or another polymer 80-120 microns thick. Metal foil, for example, aluminum foil, type A-7, 40 microns thick, is rolled onto

FOR OFFICIAL USE ONLY

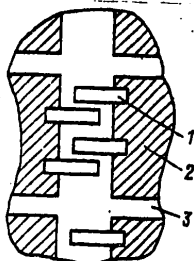


Figure 17-9. Comb location of the bar leads
1--bar lead, 2--crystal; 3--free part of the substrate

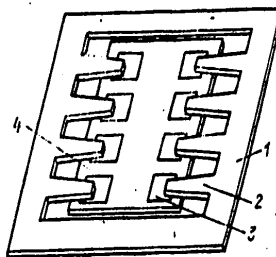


Figure 17-10. Wireless installation using "spiderless" leads
1--lead frame (tape);
2--"spider" leads
3--contact area
4--microcircuit structure

this strip using a nondrying glue or special adhesive. Then by means of photolithography with continuous rewinding of the strip frames are formed with the "spider" lead pattern. The group method is used to join the contact sites of the crystal to the internal leads of each frame, in the center of which there is a hole. In order to prevent electric short circuits between the leads and the crystals it is possible to use Al_2O_3 -insulating rings which are obtained by local electrolytic oxidation of aluminum (Figure 17-12). The leads located on the flexible carrier are insulated from each other; therefore the parameters of the crystals can be controlled directly in the technological process. After installation of the crystals,

FOR OFFICIAL USE ONLY

FOR OFFICIAL USE ONLY

the outer ends of the spider leads are simultaneously connected to all of the contact sites of the passive part of the hybrid IC.

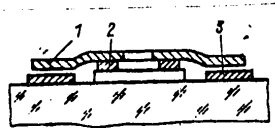


Figure 17-11. Installation of a mounted crystal using a lead frame. 1 -- frame with leads; 2 -- contact site of crystal; 3 -- contact site of the passive part of the hybrid IC.

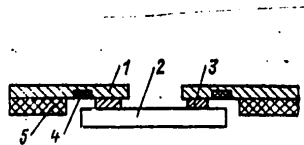


Figure 17-12. Mounting on a flexible carrier. 1 -- Al-lead-"spider"; 2 -- crystal; 3 -- contact site of the crystal; 4 -- Al_2O_3 -insulating ring; 5 -- carrier dielectric.

The "spider" leads are joined by the investigated methods of microwelding. The difference from connecting the wire leads consists only in the structure of the working part of the welder.

The wireless methods of installation are relatively new, and they are not yet fully developed with respect to structural engineering. In spite of the temporary difficulties wireless mounting is still the only path to complete automation of the assembly and encapsulation processes.

17-6. Quality Control in the Welding Process

The most effective methods of quality control of joints are metallographic analysis and mechanical strength testing.

In order to check the mechanical strength of the joints there are many attachments and devices and also test methods. For example, when shear testing, the structure with the connected leads is subjected to stretching by a force acting parallel to the surface of the substrate. If the strength of the joint is no less than 70% of the strength of the wire used, the joint is considered to have high quality. The joints are rupture tested by multiple bending of the lead at an angle of 30, 45 and 90° with respect to the substrate surface (the UKPM-1 device).

FOR OFFICIAL USE ONLY

The strength of bonded joints is determined by rupture tests. The rupture strength of a bonded joint must be no less than $(125-150) \cdot 10^5$ newtons/m².

The metallographic analysis consists in examination of transverse or oblique macro-sections of the welds, and it permits discovery of their internal structure and detection of sections not wet during soldering, through fusion, microcracks, pits, pores, intermetallic inclusions, and traces of fusion of the solder with respect to the grain boundaries.

X-ray defectoscopy using a diverging beam permits detection of internal defects and offers sufficient information about the reliability of the joints. In contrast to the metallographic analysis this method is nondestructive.

Visual monitoring makes it possible to detect breaks, short circuits, shifts of the welding or soldering zone, pores, microcracks and deformations. The relative deformation of the connected lead is determined by the width of the welding spot. With good quality of the soldered joint the solder flows out of the clearance between the joined parts, forming a good appearing solid fillet.

Controlling the electrical parameters of the microcircuits permits estimation of the correctness of the selection of the assembly conditions.

Test Questions and Assignments

1. What technological operations pertain to the assembly of IC and what are their peculiarities?
2. Repeat § 3-2.
3. What are the advantages and the deficiencies of abrasive cutting of plates and substrates with finished structures?
4. What is the essence of scribing and what is the technique for applying lines?
5. Compare the methods of separation of the structures after applying lines. How is it possible to maintain orientation of the crystals and plates after separation?
6. What determines the scribing efficiency?
7. What is the characteristic feature of separation of monocrystalline plates and substrates? What are the peculiarities of the scribing of planar structures?
8. What are the advantages and disadvantages of scribing by a diamond cutting tool?
9. Compare laser scribing with diamond scribing.
10. What is chemical separation, what are the difficulties with it?
11. What is soldering and what are its mechanisms, advantages and deficiencies?
12. How is it possible to improve the quality of a soldered joint when assembling IC?

FOR OFFICIAL USE ONLY

13. What methods are used for soldering when assembling IC?
14. What do welding and soldering have in common and how do they differ which can occur on the surfaces of welded parts?
15. What welding procedures are used for IC assembly?
16. Compare the properties of the joints obtained using current conducting and current nonconducting joints. What are the peculiarities of the glued joints by comparison with soldered and welded ones?
17. What requirements are imposed on the installation operations?
18. What is the essence and what is the technique for direct mounting using the electrical insulating compounds?
19. What is the essence and what is the technique for eutectic soldering of silicon crystals?
20. What is installation by the "rotated crystal" method and what are its basic difficulties?
21. What are the characteristic features of wire mounting?
22. What is the thermal compression procedure, what technique implements it and how is the optimal condition selected?
23. Compare welding by an indirect pulse heating with thermal compression welding.
24. What are the basic advantages of ultrasonic welding?
25. What is dual electrode welding, what is its essence and technique? What are the basic advantages of this welding over welding by indirect pulse heating and thermal compression?
26. What microwelding procedures are contactless? What do laser and electron beam have in common and what are their differences?
27. Give a brief description of microsoldering during wire installation.
28. What methods of wireless installation are used when assembling IC and what is their essence?
29. What is the process of joining the crystals to the bulk leads by the "rotated crystal" method?
30. What is the method of "spider" leads and what are its advantages? Compare the wiring on a flexible carrier and on metal lead frame or strip.
31. What is the basic advantage of wireless installation?
32. How is the quality of the joints checked after assembly?

FOR OFFICIAL USE ONLY

CHAPTER 18. ENCAPSULATION OF MICROCIRCUITS

18-1. Microcircuit Cases

General Information on Encapsulaton. The problems of encapsulation are as follows: insurance of reliable connection of the microcircuits to the equipment, protection of the structures from all types of external operating loads and also external aesthetic appearance.

With respect to structural engineering attributes it is possible to divide encapsulation into three types: case, caseless and combination.

Case encapsulation provides for preliminary manufacture of the part (cover, insulators, leads, flange, solder inserts) and subassemblies of the cases (the case base). In the case of caseless encapsulation the process of manufacturing the case is matched with the encapsulation process. The structure of the microcircuit turns out to be included in the shell material and represents a united whole with it. Combination encapsulation is done by manufacturing the cover, in the free space of which the structure connected to the lead holder is placed, and then the sealing compound is poured over the structure through the free entrance of the cover.

Types of Cases. Standardized cases are used to encapsulate microcircuits. This simplifies the manufacture, it permits mechanization and automation of the processes of assembly and encapsulation, it lowers the cost of the microcircuits and also simplifies the construction of the equipment.

The classification of cases by external structural appearance is presented in Table 18-1. The overall and connecting dimensions of each type of case are strictly standardized.

Depending on the materials used to make the cases, the latter are divided into glass, ceramic, cermet, metal-glass and plastic. The glass and ceramic cases differ from metal-glass and cermet cases in that only the leads enter into their structure made of metal parts.

Manufacture of the Parts and Subassemblies of the Cases. In the production sections the following are realized in a defined sequence (Figure 18-1): entrance control of the materials, obtaining the billets, stamping the parts, obtaining the junctions, application of the galvanic coatings, output control.

The metal parts of the cases are basically made by the method of cold stamping. Before stamping the strips, tapes, bars, tubes and wire made of kovar, copper,

FOR OFFICIAL USE ONLY

Table 18-1. Types of cases

Type	Form of projection of the case body on the base plane	Location of the projection of the leads on the base plane	Location of the leads with respect to the base plane
1	Rectangular	Within the limits of projection of the case body	Perpendicular
2	Rectangular	Beyond the limits of projection of the case body	The same
3	Circular	Within the limits of projection of the case body with respect to a circle	The same
4	Rectangular	Beyond the limits of projection of the case body	Parallel

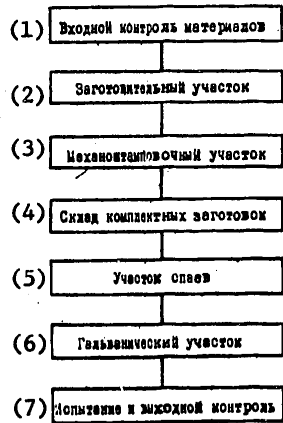


Figure 18-1. Basic production sections for the manufacture of case parts and subassemblies.

- Key:
1. input control of materials
 2. billeting section
 3. mechanical stamping section
 4. billet make-up storage area
 5. junction section
 6. galvanizing section
 7. testing and output control

steel and nickel are laid out, degreased, annealed in hydrogen furnaces and pickled. The tapes and strips are laid out by roller shears into billets of the required sizes. From the obtained billets the parts are made on mechanical stamps, the operating tool of which is a punch and die. The obtained part is forced through the

FOR OFFICIAL USE ONLY

opening in the die into the receiver for the finished parts. Cold stamping permits manufacture of parts of complex shape, it is distinguished by high output capacity and is easily automated. It is possible to perform the following operations by cold stamping: cutting out, trimming, punching holes, drawing, die forging, upsetting. Cutting out is a complete separation of part of the material with respect to a closed outline from the total mass. The less the punch-die clearance, the less rough the cutout surface and the fewer burrs on it. Trimming and punching are operations analogous to cutting out. Drawing is an operation of making hollow parts of closed outline from a two-dimensional billet open on one end (for example, the case cover). In order to prevent the formation of wrinkles when drawing the billet by the punch into the die, the billet is clamped to the surface of the die. The thickness of the billet is maintained only at the center of the bottom of the part; the material thins out at the points of transition to the walls, and at the open edge it gets somewhat thicker. Die forging is used to make flanges and bases for the cases. During volumetric stamping redistribution of the metal takes place with a decrease or preservation of the height of the billet.

The cleaning of the parts and removal of burrs after cold stamping is done by tumbling in drums with a tumbling mix (metal needles with nonresonance wood shavings), chemical pickling or abrasive grinding.

The metal-glass subassemblies which provide reliable electrical insulation are obtained by soldering. Before soldering the glass insulators (beads and tablets) are pickled, they are washed in running water and dried.

Good adhesion of the glass to the metal parts is achieved if the machining of the latter corresponds to roughness class 5-7. Before soldering the metal parts also go through a preparation cycle: degreasing, pickling, washing, drying, and the degassing in hydrogen.

After mechanized assembly the joined parts in special graphite holders are transferred to the traveling oven. The approximate time-temperature conditions of joining the kovar to S49-2 glass are presented in Figure 18-2. The parts are heated in section I, the glass is softened and the kovar is oxidized by oxygen added to the nitrogen flow. In section II at maximum temperature of 900-1000° C the glass is soldered to the kovar in a nitrogen atmosphere. In section III the oxide film is reduced on kovar surfaces free of glass in a hydrogen atmosphere. At a temperature of 575° C, annealing of the obtained junction begins to remove the mechanical stresses (section IV). Then the holders are cooled to room temperature (section V), and the finished subassemblies are taken out of the holders. During soldering, stability of the gas regime is important. Hydrogen should not get into the soldering zone (section II), for the reduction of the oxides on the surface of the metal parts has a negative effect on the quality of the joint. With an oxide joint as a result of diffusion of the surface of the oxide layer of metal into the softened glass, an intermediate layer is formed which lowers the mechanical stresses and increases the strength of the junction. It is important to insure optimal thickness of the oxide, for with small thickness the strength of the junction is less, and with great thickness, the seal.

The case parts made of vacuum-type ceramic are primarily made by the method of hot pressure casting of paraffin ceramic mass (slip) in metal molds on special casting

FOR OFFICIAL USE ONLY

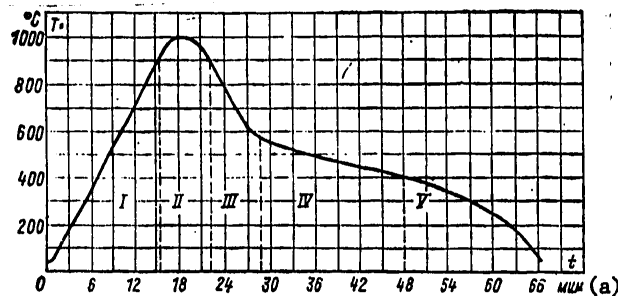


Figure 18-2. Time-temperature conditions of joining kovar to S49-2 glass.

Key: a. minutes

machines. The paraffin is added to lend the ceramic mass plasticity during the hot pouring process. The pressed ceramic parts are annealed in a bank of adsorbed powder (alumina). After cleaning off the alumina using a jet of compressed air, final annealing of the parts takes place to lend the ceramic the required physical-chemical properties. The annealed parts are checked for porosity and correspondence to the given dimensions, they are ground with diamond discs, washed in hot water, dried, heated in the muffle furnaces and again checked for the presence of chips, cracks and dark spots.

For joining the ceramic parts to metal parts, for example, leads with a frame for a flat ceramic case, the ceramic is metallized using a molybdenum-manganese paste or metal foil 30-50 microns thick. The metallization is carried out by brushing on a paste, spraying from an air gun through a free mask, stenciling, cutting out from foil with subsequent gluing. A layer of nickel or copper 3-5 microns thick is electrolytically applied for metallization, and then it is burned in. The second method of joining the metal to ceramic is through a layer of soldering glass which is applied in the form of a suspension to the joined surfaces and then sintered at a temperature of 400-500° C.

The galvanic coatings of the case parts are needed to obtain smooth surfaces, for protection against corrosion, to insure high quality of their joints during encapsulation. Before the application of the galvanic coatings the parts are carefully degreased, pickled and washed. The surface of the parts after pickling is again quickly covered with oxide; therefore the pickling operation is carried out directly before putting the parts in the electrolytic bath. The following operations are performed in the galvanizing section: chemical nickel plating of the kovar parts, electrolytic nickel plating of copper and steel parts, copper plating of insulators, gold plating of knives, and so on.

18-2. Methods of Sealing in a Case [Encapsulation]

Methods and Procedures for Encapsulating Microcircuits. The same methods are used to encapsulate microcircuits as to assemble them: soldering, welding and bonding (Figure 18-3). The methods of soldering and welding are most widely used, for they make it possible to obtain vacuum-tight sealed joints. Bonding is one of the

FOR OFFICIAL USE ONLY

simplest and most economical methods, but it does not allow for obtaining sealed joints. In many cases preference is given to soldering when encapsulating.

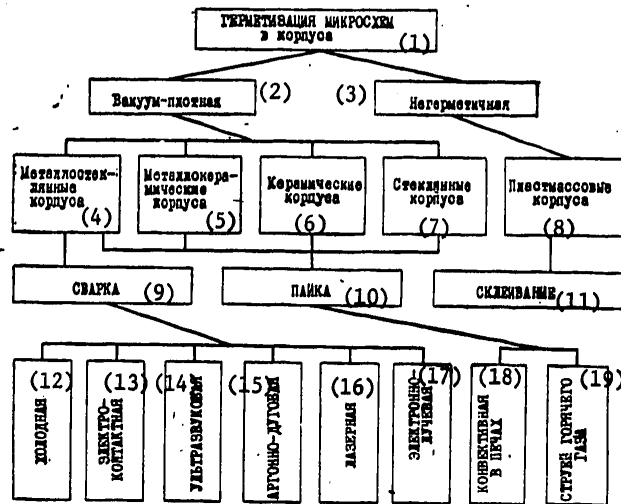


Figure 18-3. Classification of methods and means of encapsulating IC.

- Key:
- | | |
|-----------------------------------|----------------------------|
| 1. encapsulation of microcircuits | 11. bonding |
| 2. vacuum-tight | 12. cold |
| 3. unsealed | 13. resistance |
| 4. metal-glass cases | 14. ultrasonic |
| 5. cermet cases | 15. argon-arc |
| 6. ceramic cases | 16. laser |
| 7. glass cases | 17. electron beam |
| 8. plastic cases | 18. convective in furnaces |
| 9. welding | 19. hot gas jet |
| 10. soldering | |

The structural designs of many of the widely used microcircuit cases with location of the projection of the leads beyond the limits of the projection of the case body, including with parallel arrangement with respect to the base plane (see Table 18-1), and with the presence of glass or ceramic insulators directly under the sealing zone, the application of welding and pressure are made impossible. In addition, wide use is made of ceramic cases, the joints are made by metallization when sealing. During soldering the metallic, metallized or glass-coated surfaces of the case parts are connected to the sealing system with the help of solder and flux, the role of the flux can be replaced by hydrogen, inert gas or special additives to the solder. During soldering the entire microcircuit is heated to temperatures of 200-350° C, and the presence of a flux can have a negative effect on its electrical parameters. The advantages of the solder include the absence of significant compressive forces and special tools. Soldering is done by a hot gas jet or convective heating of the holder with the assembled parts in furnaces. The soldering procedures used for encapsulation in practice do not differ from the soldering techniques during assembly; we shall consider them in § 18-3, in examples of sealing specific cases. In this section we shall consider the soldering procedures used for encapsulation.

FOR OFFICIAL USE ONLY

Cold Welding. This is pressure welding without heating realized by joint plastic deformation of the surfaces of the joined parts. Careful preliminary cleaning of the parts is insufficient to insure strong metal bonds. A necessary condition of obtaining a sealed welded joint is the presence of an oxide film or galvanic coating on the joined surfaces having great hardness and brittleness by comparison with the material of the parts. In the majority of cases after careful cleaning and annealing the parts are nickel plated or chrome plated. The nickel or chrome coatings 3-9 microns thick protect the clean surfaces of the parts reliably before welding.

When encapsulating microcircuits one-way cold lap welding around the perimeter is used. The base of the case is placed in a lower hollow punch to the protrusion and it is covered with the case cover, aligned by means of the upper punch (Figure 18-4). Under the effect of the compressive forces of the punches the surface film of nickel or chromium cracks and it is forced out of the welding zone; the base material of each part is bared. As a result of bringing the welded surfaces together to the distance of effect of interatomic forces, a common electron cloud is formed, a metal bond arises between the surface atoms. On increasing the pressure, the bond zone grows, and a strong sealed joint is formed.

The force required to execute cold welding depends on many factors and can vary within broad limits. The pressure must be no less than 3×10^9 N/cm². The cold welding conditions usually are determined by the degree of deformation:

$$k = \frac{2H-t}{2H} \cdot 100\% \quad (18-1)$$

where $2H$ is the total thickness of the welded parts; t is the thickness of the welded joint.

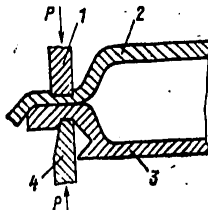


Figure 18-4. Diagram of two-way cold welding of a IC case. 1 -- upper flat punch; 2 -- cover made of soft material (copper); 3 -- base of the case made of harder material (kovar); 4 -- lower trapezoidal punch; P -- compressive force.

Thus, for high-quality welding it is necessary to insure the following: cleanness of the surfaces and the presence of hard, brittle, sufficiently thick films of nickel, chromium or other film; precision assembly of the parts; sufficient deformation of the welded materials (75-80%); selection of plastic materials for welding.

The punches are made of hard, high-quality alloy steels Kh12M, Kh12F1. The roughness of the working surfaces of the punch must correspond to classes 9-11. During operation it is necessary to see that there are no dents, chips or burrs on the punches. The punches should be wiped with a dry coarse calico regularly.

FOR OFFICIAL USE ONLY

The required degree of deformation of the welded parts is given by the restriction of the travel of the punches by the ends of a special sleeve.

The advantage of cold welding is the following: absence of heating, gas releases and splattering of the metal and also simplicity of the welding equipment. The deficiencies include an increase in the perimeter of the outside contour of the case and significant deformations of the weld-affected zone, difficulty in reliable connection of the thin-walled parts, nonuniformity of plastic deformations with respect to the perimeter of the rectangular case, the possibility of disturbance of vacuum tightness of the weld as a result of incomplete fusion or undercutting on the part of the softer material, a limited selection of materials both with respect to thickness and with respect to properties. It is recommended that the parts made of 29NK alloy or 47MD alloy with MB copper or M-1 copper be joined.

Electrocontact Welding. When encapsulating microcircuits most frequently capacitor spot or roll welding is used. The welding is done at the time of discharge of the capacitor bank to the primary winding of the welding transformer. The secondary winding of the transformer is connected to the electrodes that play the role of the welding tool. Sealing by spot welding is done using electrodes, the shape of which repeats the shape of the case perimeter (Figure 18-5). The base of the sealed case is inserted into an opening in the lower electrode, and the cover of the case is placed on top. When pressing on the pedal of the welder, the upper electrode is lowered, and in one current pulse welding is accomplished with respect to the entire perimeter of the case. A special sleeve is used to align the welding rods. In the intervals between welding, the capacitors are charged from an ac network through a rectifier.

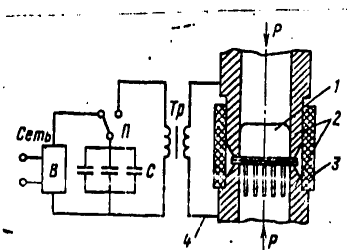


Figure 18-5. System for spot capacitor welding. 1 -- case; 2 -- electrodes; 3 -- sleeve; 4 -- electric network of the welder; P -- compressive force.

The optimal welding conditions are selected by regulating the capacitance of the capacitor bank, the transformation coefficient of the welding transformer and the force of compression of the electrodes. The quality of the weld depends on the quality of preparing the joined surfaces, the shape and area of the contact, the proper choice of the pair of joined materials, cleanness of the machining of the joined surfaces which must be no less than class 5 and the working surfaces of the electrodes which must be no less than class 9-10.

The method is widely used to seal small, round metal glass hinges. The application of the given procedure for sealing large-perimeter cases has been delayed for some time as a result of absence of powerful capacitor banks. At the present time

FOR OFFICIAL USE ONLY

Soviet industry has developed powerful welders. For example, the ShchYaM 1.124.001 makes it possible to weld rectangular cases from 55 to 100 mm around the perimeter.

Roll welding, in contrast to spot annular welding, is welding by moving electrodes. The electrodes supplying the current have the shape of rolls, and during welding, they are rolled over the perimeter of the case at a short distance from each other on the same surface. The welding is argon shielded. During welding it is necessary to see that there are no defects on the electrodes: dents, chips, erosion, and so on. The electrodes are periodically cleaned to remove contamination. The method provides for obtaining a tight weld, it is used to seal metal-glass and cermet rectangular cases up to 75 mm around the perimeter.

Sealing by resistance welding is used to join parts made of 29NK kovar or 47ND alloy to parts made of NP-2 nickel, 0.8 KP, 10 or Kh18N9T steel. Parts made of nickel and stainless steel are welded without a coating. The parts made of kovar are shifted on the carriage along a guide rail and the cases assembled into holders are welded. In the Soviet welders USKM-2, USKM-3, USKM-4, a lock of about 100 cases is welded simultaneously.

The basic disadvantage of the method is high thermal effect on the welded parts and the necessity for increasing the welding bead connected with this.

Laser Ion Welding. A new type of welding has become possible as a result of the development of YAG lasers, which in contrast to the ruby and neodymium glass lasers previously used for spot welding, have higher pulse repetition frequency and radiation power. During pulsed laser welding, a long weld is formed by superposition of the points on each other with some overlap $k = \ell/d$, where ℓ is the overlap; d is the diameter of the welded spot. The welding speed is defined by the formula

$$V = fd(1-k), \quad (18-2)$$

where f is the pulse repetition frequency.

The laser welding of nickel, kovar, stainless steel, titanium and other materials is done on the "Kvant-12" device which provides a speed of not <150 mm/minute, a weld width of 0.25-1 mm with a fusion depth to 0.3 mm. The maximum pulse frequency is 20 hertz, the maximum average radiation power is 30 watts. The welding process is carried out in a jet of shielding gas -- argon. The welded parts are shifted relative to the beam.

Laser welding is used to seal rectangular metal cases of microcircuits. During welding the laser beam is directed at the joint of the edges of the base and cover.

The welding conditions are characterized by the following parameters: the radiation energy, the pulse repetition frequency, the rate of displacement of the table with the case, the diameter of the beam on the welded surface, and the argon flow rate.

The advantages are as follows: high quality of the weld, uniform and low temperature along the weld and, what is especially important, at the glass insulator (as a result of high localization and briefness of the heating); high yield of usable microcircuits to 98%, and high efficiency; the possibility of easy focusing of the

FOR OFFICIAL USE ONLY

beam by ordinary optical systems, and the possibility of welding through transparent shells.

The basic limitation of the method connected with instability of the electrical characteristics of the lasers is the time limitation. At the present time welders have been developed, for example, the Kvant-12 and Kvant-17, which insure high quality, efficiency and percentage yield of usable products.

Cathode Ray Welding. At the present time cathode ray welding is used for welding metal-glass cases. The beam is moved along the welding line by the deflecting system of the welder. In order to prevent strong heating of the case during welding, it is cooled with silicone oil which conducts heat well and copper heat transfer -- the mandrel.

18-3. Various Types of Encapsulation

Encapsulation in Metal-Glass Cases. Metal-glass cases and flat rectangular cases (Figure 18-6) are used to encapsulate semiconductor microcircuits. Square and rectangular cases with leads which are perpendicular with respect to the base plane (Figure 18-7) are used to seal hybrid thin-film microcircuits.

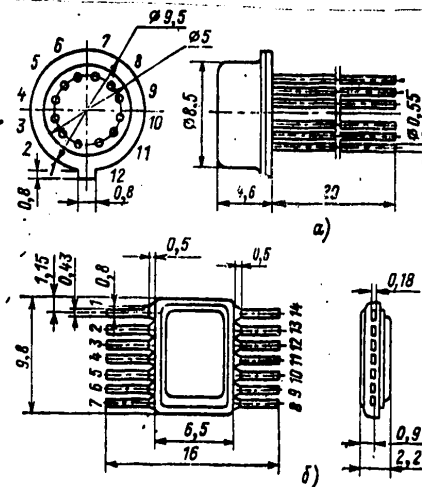


Figure 18-6. Metal-glass circular (a) and rectangular flat (b) cases.

The circular cases (type TO) are distinguished by high reliability checked during prolonged storage and operation. The cases are borrowed from transistor production. The case technology has been well-developed; therefore the cost is relatively low. The cases are distinguished by high rigidity of the structural elements and as a result of good electromagnetic shielding of the structures they are used well for linear microcircuits. The basic deficiencies of the cases are as follows: a limited number of leads, low filling coefficient of the case volume.

Rectangular flat cases can have two-way and four-way arrangement of the leads. The seal is made on the lead frame from a flat sheet of metal. The basic advantages of

FOR OFFICIAL USE ONLY

the cases are as follows: light weight and small height, high filling coefficient of the space, and a large number of leads. The basic disadvantages are as follows: high cost and damage to the flexible strip leads requiring special mandrels for transportation and storage before mounting in the microelectronic equipment.

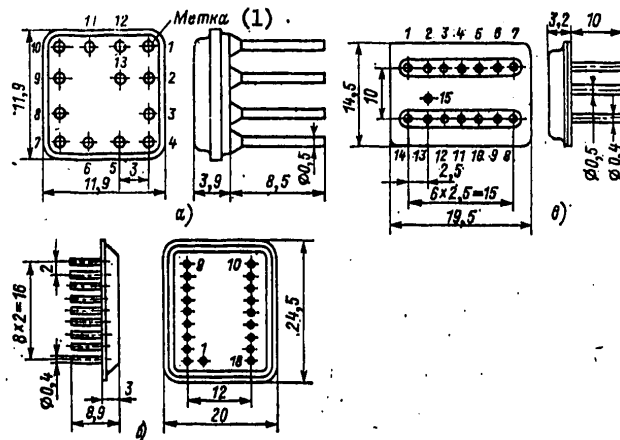


Figure 18-7. Square (a) and rectangular (b, c) cases with leads perpendicular to the base plane.

Key: 1. mark

Square and rectangular cases with perpendicular arrangement of the pin leads are distinguished from the flat cases by high rigidity of the leads.

The metal-glass cases are sealed by the previously described welding procedures. For circular cases basically cold and resistance welding are used; for square and rectangular cases, roll, laser, cathode ray, ultrasonic and argon-arc welding are used. The enumerated types of welding provide for obtaining a reliable, strong weld and a vacuum type seal.

Sealing by soldering is used primarily for microcircuits in rectangular and square cases. It is possible to do the soldering using solder inserts, solder or glass fritts applied to the points at which the parts are joined. The soldering is done in holders in traveling ovens in a helium or hydrogen atmosphere.

The solder inserts are placed between the base of the case and the cover is clamped on the top using a spring. The holder is assembled in an enclosure and it is transferred from it through a lock to the oven conveyor. In mass production wide use is made of the POS-61 eutectic solder. By comparison with other tin-lead solders, POS-61 solder has a minimum melting point and the least porosity. However, the use of the POS-61 solder requires the application of a flux. The soldering without a solder insert provides for preliminary application of the solder to the cover of the case. The base of the case is first clad with gold at the joining point. The heating is done by a jet of hot inert gas. The molten layer of solder wets the surface of the material of a case base well as a result of the gold, and

FOR OFFICIAL USE ONLY

it forms a sealed, soldered joint. The application of the gold, silver or palladium coating of the parts insures high quality of joining without using a flux. The basic difficulty consists in local application of the gold to the point of formation of the soldered joint.

The glass frit seal is made analogously to soldering without the solder insert. The low-melting glass powder which is carefully mixed with solvent and binder is applied to the joined sections of the parts. For preparation of the fritt, it is necessary to use glass with a thermal coefficient of linear expansion close to the thermal coefficient of linear expansion of the joined metals. After drying the applied layer, the case is assembled. During heating, the parts are soldered through the glass insert.

Encapsulation in Glass, Ceramic and Cermet Cases. The shaped glass cases (Figure 18-8) are primarily used for general-purpose semiconductor microcircuits. The case provides good electrical insulation; it is distinguished by low cost, but it has poor thermoconductivity and low strength. Therefore the glass cases are made in a flat, small execution. The cases are sealed by soldering through a layer of low-melting glass or metallization.

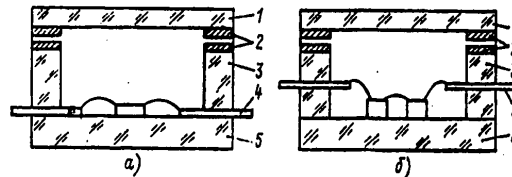


Figure 18-8. Standard structural designs of glass cases (a, b).
1 -- cover; 2 -- metallization; 3 -- walls of the case; 4 -- leads;
5 -- base.

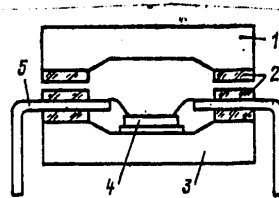


Figure 18-9. Ceramic case with two-row arrangement of the leads (DIP). 1 -- cover; 2 -- junction point; 3 -- base; 4 -- IC structure; 5 -- lead.

The ceramic cases, analogously to gas cases, provide high quality of insulation with sealing of several structures in one case. For manufacture of the cases, aluminum ceramic or ceramic based on beryllium oxide is used. The joining of the ceramic parts of the case to each other realized through a layer of solder glass.

Ceramic cases with two-row arrangement of the leads (DIP type) (Figure 18-9) are finding broad application for encapsulation of LSI. After installation of the structure around the perimeter of the base and the cover of the case, a glass

FOR OFFICIAL USE ONLY

suspension is applied by stenciling. After drying the applied layer, the case is assembled and sealed. The case is distinguished by high strength and low cost.

Ceramic cases with up to 50 leads are used to encapsulate thick-film hybrid microcircuits of second and third degree integration (Figure 18-10). A characteristic feature of such cases is the fact that the upper and lower covers simultaneously play the role of bearing substrates for thick-film elements and mounted active caseless transistors. The covers are glued to the outline frame and are sealed at the ends by compound. The covers and the frame are made of 22KhS ceramic. The external leads of the IC are located in the same plane as a result of forming the leads of the upper substrate.

For cermet cases, it is important to select the materials of the metal cover and ceramic base matched with respect to the thermal coefficient of linear expansion. When using a polycor base, it is possible to use copper, nickel, kovar and titanium covers. The sealing is accomplished by soldering the cover with the base through a solder frame based on silver. The soldering point on the ceramic base is metallized in advance using a manganese-molybdenum paste with subsequent galvanic coating of nickel or copper.

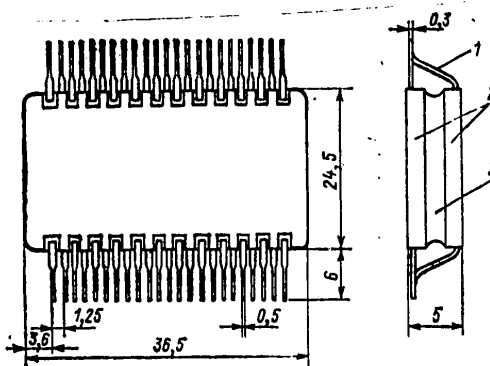


Figure 18-10. Ceramic case for thick-film hybrid IC.
1 -- lead; 2 -- upper and lower substrate covers;
3 -- outline frame.

Encapsulation in Plastic Cases (Figure 18-11). The parts of the plastic case are made of K-124-38 molding powder or polystyrene by hot pressing. The cover is joined to the base of the case (reinforced by leads) by glue. On setting of the glue, during hardening the joined parts are pulled together and tightly clamped.

The glued joint has the best technological nature and is the most economical form of sealing microcircuits for household electronic equipment and computers that operate under normal operating conditions.

18-4. Caseless Encapsulation

Caseless Encapsulation Technology. For caseless encapsulation, the methods of pressure casting, vacuum casting, lacquering and surface coating by dipping, spraying

FOR OFFICIAL USE ONLY

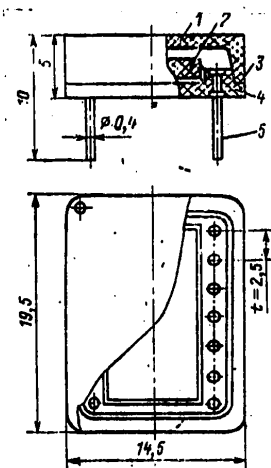


Figure 18-11. Plastic case. 1 -- cover;
2 -- IC structure; 3 -- gluing point; 4 --
base; 5 -- lead.

and the "drop method" are used. The method of pressure casting is most widely used, for it is characterized by ease of automation and the highest efficiency. The encapsulation process is carried out on transfer molding machines; therefore it is frequently called transfer casting. The method is based on the capacity of plastics to melt and flow under the effect of temperature and pressure, filling all of the cavities in a special mold. The plastic materials must have the largest possible fluidity at the lowest possible temperature and pressure and also poor adhesion to the walls of the mold.

For encapsulation by transfer molding, the lead frame, strip or flexible carrier with connected structures is placed in the working recesses of the lower part of the pressure mold having a charging chamber for pressing and a gating system (system of channels) connecting the charging chamber to the operating recesses. The lower part of the pressure mold is covered, and it is placed between two heating plates of a two-way hydraulic press, during the operation of which the upper and lower parts of the mold are brought into contact, and the molding material is extruded by the sliding transfer plunger. The molding material goes into a highly viscous state under the effect of temperature and pressure, and it goes from the charging chamber through the gating system to the working recesses. The molding material fills all of the cavities of the working recesses and completely engulfs the microcircuit reinforcing. After relieving the pressure and cooling the pressure mold, the molding material is polymerized, and it takes the shape of the inside volume of the working recesses. After extraction of the microcircuits from the pressure mold, the fin and burrs are removed, the excess edges of the lead strip are trimmed off, and the leads are bent on a special die.

The plastic shell of the finished microcircuit has a clearly expressed form, the distances between the leads are determined by the configuration of the lead frame. It is possible simultaneously to seal a lock of up to hundreds of microcircuits in a single pressure mold.

FOR OFFICIAL USE ONLY

Transfer molding is distinguished by high efficiency, it insures good external appearance and defined geometric shapes of the finished microcircuits. However, the sealing process requires large capital expenditures, and therefore it is used for large-series production.

The pouring in a vacuum has less output capacity, but it does not require expensive presses. The assembled structures are placed in the inside volumes of special molds which are moved along the operating section of the automatic casting machine. The carefully mixed sealing compound is simultaneously poured into all the molds by a plunger type batcher. The molds are evacuated to remove air bubbles from the compound. After pouring, the compound is hardened. The time-temperature conditions of hardening are determined by the composition and the ratio of the compound components. After hardening of the compound the microcircuits are extracted from the molds and they are separated from the excess part of the lead strip in the die.

When using the vacuum pouring method, the amount of manual labor is still high, and the efficiency is low. The quality and the external appearance of the microcircuits and the mounted components are insufficiently satisfactory. The method is economically justified only when encapsulating small lots of products.

The envelopment type encapsulation consists in applying a drop of evacuated sealant to a rotating structure. When the compound spreads as a result of the surface tension forces, a spherical shape of the shell is obtained, the configuration of which depends on the shape of the sealed structure.

The process is easily carried out, it does not require complex equipment, and with intermediate drying it permits the application of multilayer shells. However, the method of envelopment encapsulation has a number of deficiencies: the necessity for the application of easily volatile solvents, the duration of the drying process in the air before final hardening at increased temperatures, the complexity of obtaining uniform thickness of the coating. Therefore this type of encapsulation usually is used for sealing mounted active components for hybrid microcircuits.

The spraying and application of a batched drop to the working side of the structure and also dipping are used to seal active components or microcircuits which will then be finally encapsulated in a common shell.

The advantages of caseless encapsulation include the following: simplicity and low cost of the technological processes, reduction of the number of operations, nonshortness and low cost of materials, the possibility of automation with continuity of the assembly and sealing processes.

The microcircuits encapsulated in plastic, by comparison with the devices and cases have higher reliability with respect to mechanical loads, which is explained by the absence of internal free space (for moving the leads). Therefore they are widely used in motor transportation, in the drives of forging hammers, and so on.

When operating under conditions with moderate relative humidity of no more than 80% and moderate temperature from -20° C to $+100^{\circ}$ C the microcircuits encapsulated in plastic are not inferior with respect to reliability to their analogs in sealed cases.

FOR OFFICIAL USE ONLY

The microcircuits sealed in plastic are characterized by insufficient moisture resistance as a result of the moisture permeability of the materials themselves and especially their interfaces with the metal leads. The latter is explained by the sharp difference of their electrophysical properties, structure, composition and chemical bonds.

The basic types of failures during storage and operation of plastic encapsulated microcircuits are the following: ruptures of the metallization and leads, shorts which shunt the circuits. The cause of these types of failures is the processes of the dissociation of salts present in the sealants. Dissociation can lead to corrosion, the formation of undesirable compounds and breaks of the electric circuit. The interelectrode transport of the metal ions can lead to shortcircuits. Ruptures of the leads can occur as a result of mismatch of the thermal coefficients of linear expansion of the leads and the plastics and corresponding alternation of their tension and compression during the heating and cooling cycles.

When encapsulating in plastics, care in mixing and evacuation of them is very important, for the slightest nonuniformity of the composition or the presence of gas bubbles can lead to the appearance of leaks, pores, microcracks which become deeper during the heating and cooling cycles and can lead to a loss of seal.

The thermal properties of the microcircuits in plastic shells can be improved only by increasing the operating temperature of the sealant.

18-5. Quality Control of Encapsulation

During the encapsulation process, input control of the materials, intermediate products, the parts and subassemblies, the sealing weld, testing of the finished microcircuits and exit control are exercised.

The parts are controlled by visual inspection after cold stamping. The basic forms of rejects after cold stamping and their causes are presented in Table 18-2.

The size of the parts is measured by universal measuring instruments: slide calipers, a micrometer, display and optical instrument — instrument microscope.

The planarity of the surfaces of the parts is checked by the light slit method using a toolmaker's straight edge. The eye of man is capable of catching a clearance of 0.003-0.004 mm.

The subassemblies of the parts are controlled for accuracy of the dimensions, quality of machining of the surfaces, galvanic coatings and junctions.

For monitoring the soldered, welded and glued sealing joints, the same methods are used as for assembly (see § 17-5).

The seal is checked twice: after manufacturing the base of the case with insulated leads and after encapsulating the microcircuits. The encapsulation of the junction of the leads with the material of the base or the encapsulation of the microcircuit in a case is characterized by the rate of admission of helium. For finished microcircuits, the seal criterion is the rate of admission of helium (cm^3/sec) with a pressure difference outside and inside the case of 10^5 Pa. High-quality cases have an admission rate that does not exceed 10^{-8} cm^3/sec .

FOR OFFICIAL USE ONLY

Table 18-2. Rejects during cold stamping of kovar

Type of reject	Cause
Dark color of the part	Poor degreasing and annealing
Breaking of the material	Improper annealing of the billet
Burrs	Improper die-punch clearance; poor alignment; poor sharpening of the cutting edges of the punch and dies
Scratches, longitudinal lines	Contaminated strips, insufficient polishing of the drawing die of the female dies and punches
Breaking of the edges of the parts, the formation of folds	Weak clamping of the material in the drawing die
Breaking of the bottom of the part, crack formation	Extraordinary clamping of the material in the drawing die
Reject for dimensions	Errors in the structural design, in the manufacture or during repair of the dies; wear of the working parts of the dies; rejection of the material for thickness; violation of the technological process

The base of the case is checked for seal by using special attachments which permit the creation of a volume which is closed on the checked part using vacuum seals.

There are many methods of checking the seal. Most frequently the mass-spectrometric, vacuum-liquid and moisture methods are used.

The mass-spectrometric method is based on indication of the helium atoms leaking through the leaks in individual subassemblies or the sealed cases. The application of helium for leak detection is explained by the fact that it is the most mobile gas and has high penetrability. Helium is introduced into the microcircuit case either during sealing or by prolonged holding of the sealed microcircuits in special sealed chambers (bombs) filled after preliminary evacuation with helium to a pressure of $(3-5) \cdot 10^3$ Pa. During the holding time (3-48 hours) helium penetrates into the cases of the microcircuits having leaks. The microcircuits are removed from the bomb and are placed in a beaker from which the air is pumped out to maximum vacuum and which is connected to the sensitive element of a leak detector. The PTI-6, PTI-7 and PTI-10 helium leak detectors are used for indication of the helium.

The mass-spectrometric method is distinguished by high sensitivity. The deficiencies of the method include the following: low output capacity (100-200 pieces/hour), complexity of servicing the units.

The vacuum-liquid method is based on recording air bubbles which leak out through the leaks in the case into a liquid above which a vacuum has been created. The microcircuits are submerged in a glass chamber with oil or kerosene and air is pumped out of it. If there is a leak in the case, then as a result of the difference in pressures inside and outside the case the gas will leak outside in the form of a stream of small bubbles. Thus, the location of the leak is detected by visual observation. The method is simple, operative, and more efficient (to 700 pieces/hour), but it is less sensitive, and therefore it permits detection of only gross leaks.

FOR OFFICIAL USE ONLY

The moisture method of control is the simplest, the most reliable and permits simultaneous control of, in addition to the seal, the resistance of the case coatings to the effect of increased moisture. The microcircuits are held in heat and moisture chambers for several days under increased moisture conditions (95-98%) at a temperature of about 40° C. The reject criterion of the microcircuits is worsening of electrical parameters of the unsealed microcircuits as a result of penetration of moisture into the cases. However, in the heat and moisture chambers only the microcircuits with gross leaks are rejected. In addition, the chamber does not permit operative detection of leaks in the microcircuits with well-protected structures. The penetration of moisture into the case of such microcircuits is detected significantly later when a failure occurs, for example, as a result of corrosion of the intermetallic joints.

18-6. Finishing Operations of the Manufacture of Microcircuits

Before performing the technological finishing operations after assembly and encapsulation, the microcircuits go through climatic, electrical and mechanical tests. The necessity for these tests arises from the presence in some microcircuits of latent defects as a result of imperfections of the structural design, the use of materials or conditions not corresponding to the technical specifications or the production reports and forms, and so on. During the process testing the microcircuits with latent defects are rejected; at the same time reliability of the microcircuits passing the tests under operating conditions is guaranteed.

After going through all of the phases of the production testing and measurement of the parameters, the final production operations are performed, and the microcircuits can be shipped to the users.

The finishing of technological operations include the operations of external forming providing for protection of the microcircuits from external effects, convenience of connection of the microcircuits to the equipment and also providing information about the type of microcircuit and external aesthetic appearance.

In order to obtain galvanic coatings, frequently electrolytes containing tin and bismuth are used. Galvanic coatings of microcircuits operating under harsh operating conditions are made from tin-cadmium or tin-nickel alloys.

The cases are prepared for painting by careful degreasing in boiling organic solvents, drying to complete removal of the solvent. Then a thin layer of primer is applied, it is dried and the paint is put on. Sometimes the cases are painted without primer. It is possible to use brushes, dipping, jet sprinkling and spraying in the painting operations. The painting is done in special holders which cover the places not to be painted and mechanically hold the microcircuits. The painted microcircuits are dried. The paint and varnish coatings must be uniform after drying, without runs, bubbles, blisters, scratches or foreign inclusions. The microcircuits that are rejected after drying are returned to remove the paint and varnish, and they are repainted. A mark is applied to the well-painted surface using manual printing attachments, semiautomatic or automatic machines. The BM, 4M, SN, KN, EP-572, MA-514 and other marking paints have good adhesion to the painted surface and insure clear marking. The marking symbols of the microcircuits sealed in plastic are obtained in the sealing process (the marking symbols are knocked out on the molds for pressing or pouring).

FOR OFFICIAL USE ONLY

After drying the marking symbols the microcircuit cases are lacquered. The lacquers also protect the cases from corrosion and they increase the stability of the markings.

For convenience of connecting the microcircuits to the equipment the leads of the cases are straightened and tinned. Most frequently the tinning is done by dipping in solder after preliminary treatment in a solution of colophony or zinc chloride. For tinning primarily tin-lead solders are used. When producing the microcircuits in large series the tinning is done automatically by a solder wave.

The microcircuits that are ready for shipment to the user are placed in special packaging, the structural design of which provides for preservation of the microcircuits during transportation. Usually each microcircuit is placed in a separate recess in the packaging or a pocket of polyethylene tape which then is put in plastic or cardboard boxes. Between the insert or the tape and the cover of the box inserts of porolon or corrugated cardboard are used to prevent shifting. The boxes are taped up with designation of the type and number of microcircuits.

Test Questions and Assignments

1. What are the basic problems of microcircuit encapsulation?
2. Explain what case, caseless and combination encapsulation of microcircuits are.
3. Which cases are used to encapsulate microcircuits; what are the basic steps in manufacture of them?
4. What is the cold stamping method and what is it used for?
5. What is the manufacturing technology of metal-glass assemblies of cases? Explain the temperature-time conditions of obtaining a kovar junction with glass (Figure 18-2).
6. What is the hot casting method when obtaining ceramic parts of cases?
7. How are cermet subassemblies of cases made?
8. What is the purpose and the technology for obtaining galvanic coatings when making parts of cases?
9. What are the methods and means used for encapsulation of microcircuits in cases?
10. What is the essence, the technique, the peculiarities, advantages and disadvantages of sealing the cases by cold welding?
11. What distinguishes resistance welding from cold welding?
12. What do capacitor spot welding and roller welding have in common and what are the differences?
13. How is argon-arc welding done. What are its advantages?
14. What are the basic features and techniques for encapsulation using laser welding?

FOR OFFICIAL USE ONLY

15. What is the essence, the equipment and the characteristics of encapsulation by electron beam welding?
16. Compare the welding methods from the point of view of insuring quality and high efficiency of encapsulation.
17. What are the peculiarities of metal-glass cases and how are they sealed?
18. What are the characteristic features of glass, ceramic and cermet cases and how are they sealed?
19. For which cases is glued sealing used, and what are its characteristic features?
20. What requirements are imposed on the materials for caseless encapsulation?
21. What explains the broadest application of transfer molding for encapsulation in plastic? In what cases is it expedient to use transfer molding?
22. What is the essence and the technique of the transfer molding process?
23. What is the technique for encapsulation in plastic by vacuum molding? Compare this method with transfer molding.
24. What are the methods of encapsulation by envelopment, spraying, and a batched drop? Compare them.
25. Compare encapsulation in a case and in plastic.
26. Enumerate the basic steps of control during the process of encapsulating microcircuits. What is the procedure for performing the basic control operations?
27. What is the role and the purpose of the production testing and measurement of microcircuits?
28. What operations in the manufacture of microcircuits are among the finishing operations, what is their role and how are they performed?

FOR OFFICIAL USE ONLY

CHAPTER 19. INSURANCE OF PRODUCTION EFFICIENCY AND QUALITY OF MICROCIRCUITS

19-1. Basic Areas and Problems of Microelectronics in the Current Five-Year Period

Five-Year Period of Quality and Efficiency. The party and the Soviet government have always given a great deal of attention to the problem of improving production efficiency and product quality. "A systematic improvement of production quality," is noted in the CPSU Program, "is a mandatory requirement of the development of the economy. The production quality of Soviet enterprises must be appreciably higher than at the best capitalist enterprises."

In the Ninth Five-Year Plan the leading development of production providing for improvement of the quality of the produced product was noted. L. I. Brezhnev considered electronics among the branches which are a "catalyst of scientific and technical progress."

The Tenth Five-Year Plan is a five-year plan of quality and high efficiency in the name of further growth of the economy and national welfare. The primary goal of the Tenth Five-Year Plan is successive implementation of the course of the CPSU to raise the material and cultural standard of living of the people based on dynamic development of social production and improvement of its efficiency, acceleration of scientific and technical progress, growth of the productivity of labor, comprehensive improvement of the quality of work in all branches of the national economy.

The basic areas of improvement of efficiency are as follows:

1. Improvement of the production structure based on scientific and technical achievements.
2. Reduction of the material capacity of production as a result of economy of raw material, materials, energy, improvement of engineering and production technology.
3. Efficient use of manpower by reducing the work time losses, creation of a flexible employment system and rhythmic work, and advanced training of the workers.
4. Improvement of production quality.

According to the All-Union State Standard 15467-70 quality is the "set of properties of a product giving rise to its suitability for satisfying defined needs in accordance with its purpose."

FOR OFFICIAL USE ONLY

By improving production quality, expenditures are reduced, and the production volume is increased, that is, growth of cost benefits takes place. At the 25th Party Congress the problem of improving quality was stated in a new way. We are talking about comprehensive improvement of quality in all branches of the national economy. At the present time the criteria and the content itself of the work quality are defined by the set of all aspects of labor activity which provide for achievement of the highest production efficiency.

"High quality," noted L. I. Brezhnev at the 25th CPSU Congress, "is saving of the manpower and material resources, growth of export capabilities, and in the final analysis, better, more complete satisfaction of the demands of the society. This is why the entire mechanism of planning and control, the entire system of material and moral incentives, the efforts of the engineers and designers, and the proficiency of the workers must be aimed at improvement of production quality."

Thus, improvement of quality is the key problem of improving the efficiency of the national economy in the Tenth Five-Year Plan.

Basic Areas and Problems of Microelectronics in the Tenth Five-Year Plan. The electronics industry is one of the largest base branches of the national economy. It actively influences the development of other branches and determines technical and social progress of our country as a whole. The party and government are giving the most serious attention to the problems of improving efficiency, quality and reliability of the products of electronic engineering.

The board of the Ministry of the Electronics Industry noted in March 1976, the basic areas of development of microelectronics in the Tenth Five-Year Plan:

1. Further improvement of the degree of integration, the creation of microprocesses and high-capacity semiconductor memories and also linear microcircuits to satisfy the needs of household radio electronic devices; creation of computer controlled means of automating the technological production processes used for electronics engineering products by laser radiation and electron beams.
2. Further improvement of the operating efficiency of the scientific and design organizations, reduction of the length of the "research-production" cycle and introduction of the methods of machine design in all subbranches.
3. The development and approval of the plans for introducing an all-around quality control system, insurance of the development and introduction of standards with respect to operative quality control during the production process and standard production operation and section control systems.

The basic goals of modern microelectronics are: an increase in production volume and growth rates of the product output, further improvement of quality and reliability and also functional complication of the products.

19-2. Insurance of the Efficiency and Quality on the Modern Level of Development of Microelectronics

Complex Approach to the Solution of the Problem of Improving the Quality of the Products and Production Efficiency. The problem of improving quality and production efficiency of the microelectronics enterprises is a complicated, multilevel,

FOR OFFICIAL USE ONLY

complex problem and can be solved with successful simultaneous system solution of the entire set of measures. The most important of these measures are as follows: improvement of control, planning and organization of production; improvement of the design phase of microelectronic devices -- the creation of completely automated design systems using a computer; standardization of software and hardware; conversion of individual machine design systems to a united procedural base; improvement of production technology; improvement of equipment, creation of automated technological process control systems; improvement of the quality of the initial materials and introduction of new materials; improvement of the control system, control operations and introduction of automated quality control systems into production; improvement of the level of training and the qualifications of the workers in the microelectronics branch; economical substantiation of the measures taken.

Program-Goal Planning and Organizational Measures. In the 25th CPSU Congress it was stated that the institutes of the USSR Academy of Sciences jointly with the ministries and departments, by authority of the Central Committee of the CPSU and the government have prepared a draft for a complex scientific and technical progress program and its social-economic consequences in 1976-1990.

The program-goal method is an effective means of improving planning and control of scientific and technical development of the branch. It permits improvement of the efficiency of the use of resources and reduction of the times for introduction of the results of scientific and technical developments into production.

The complex goal programs make up a system of scientific research, experimental design, production, economical and organizational organizations and measures, aimed at achievement of the stated specific goal and realized under a united direction.

Depending on the purpose the programs are distinguished by content: improvement of new products produced or created, creation of new advanced technological processes, creation of new equipment, creation of new materials, development of branch complex quality programs, reliability, standardization, and so on, the performance of scientific research work.

The basic characteristics of complex goal planning are scientific substantiation of the choice of the goal and means of achieving it, orientation of all of the operations and measures at the final result -- program goal -- unity of planning of the "research-production" cycle with a significant decrease in duration and optimization of each program.

Program-goal planning provides for generality of the long range and current plans. The long-range plans are the basis for development of the five-year and annual plans and will permit preparation of the required technical base and resources for a specific planned event in accordance with the developed scientific and technical strategy, that is, the long-range plans provide for future development.

In accordance with the program-goal planning, organizational measures are performed with respect to provision for the implementation of these programs. Improvement of efficiency, quality and reliability of the technological processes is achieved by constant improvement of the organizational structure of technological process control. For establishment of the optimal structure of control for a defined period, it is necessary to compile forecasts, define the goals, select the criteria and stimuli, determine the duties and insure the required level of qualification of the workers, choose the form of organizational structure, and so on.

FOR OFFICIAL USE ONLY

Statistical Control and Automated Technological Process Control Systems. The statistical method is one of the objective methods of gathering, processing and analyzing information received during measurements and tests. The statistical control is selective control based on probability theory and mathematical statistics which permit not only monitoring, but also active control of the quality in the production process. It permits quantitative evaluation of the interrelations between the parameters of the manufactured microcircuits and the factors of the technological process, prevention of the appearance of rejects in the case of clear manifestation of causes. The introduction of statistical control has made it possible to implement automated technological process control systems (ASU TP) based on computers.

The technological process systems -- the ASU TP -- are systems for purposeful adjustment of the parameters of the technological process to insure given product parameters. The object of control here is the equipment by means of which the production operations are performed.

The introduction of the ASU TP permits the following:

Conduct of the process with maximum efficiency, for the deviations of the process conditions, the properties of the initial materials, deviations of the ambient parameters from the norms are considered automatically, and input from the operators with their subjective influences on the course of the production process is excluded;

Process control, for the conditions of the equipment are adjusted in a timely manner, operations are redistributed, and so on;

Automatic control under conditions that are harmful to man.

The ASU TP is used both for individual technological process operations and for production as a whole, on the level of the production lines, sections and shops.

As a result of the ASU TP, high cost benefit is obtained, the labor resources are used efficiently, the overall level of organization and culture of production is raised, the quality of production is improved, and the problem of short personnel is solved.

Quality Control. High quality of production cannot be insured as a result of intensification of control and elimination of organizational deficiencies alone. At the present time the problem of improving the quality of microelectronic products is being solved in a new way, by quality control.

Quality control provides for a system of uniform mutually coordinated organizational, engineering and economic enterprises which, taken together, insure improvement of the quality in the development, production, testing and operating phases of the products. The quality of the products is standardized and planned. The purpose of control is the required, planned improvement in quality. Quality control is a continuous process, for on the basis of analyzing the operating characteristics, new requirements on the product are developed, and the control process is repeated on a higher quality level.

The structural model of a quality control system for electronics industry products is presented in Figure 19-1. The basis for this model is formed of the basic

FOR OFFICIAL USE ONLY

steps of the formation and maintenance of quality: development, production, application, operation and maintenance with the corresponding ties among them. Quality control is realized in this model by three loops: control and test loop, analysis loop and operative regulation, planning and control loop.

The all-around systems approach to the solution of the problem of improving quality encompasses many spheres of production and is a component part of the overall enterprise control system. Thus, the quality control system is an enterprise control subsystem.

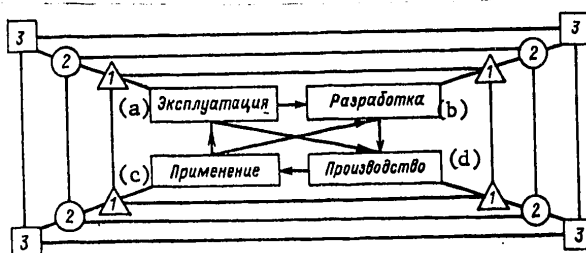


Figure 19-1. Structural model of the quality control system for products of the electronics industry. 1 -- monitoring and testing; 2 -- analysis and operative regulation; 3 -- planning and control.

Key: a. operation and maintenance
b. development
c. application
d. production

19-3. Nondestructive Control and Improvement of Technology

Nondestructive Control. Nondestructive control is a prospective area of technical diagnosis promoting improvement of the quality of mass produced products without demonstrating the continuity of the technological process.

Nondestructive control is the name given the performance of any measure that permits estimation of the attribute, parameter or quality index of a product without having a negative effect on the properties characteristic of it at the time of control.

The credit for the development of the theory of nondestructive control techniques goes to B. M. Sotskiy, and for introduction, to Yu. A. Kontsevoy, and other scientists. The physical basis of nondestructive control is made up of studies of the physical characteristics of materials and detection of imperfections in their structure. In microcircuit production, the methods of nondestructive control based on the application of electromagnetic vibrations with different wavelength is the most widespread. The visible part of the spectrum is used for visual control, by means of which surface defects are detected.

The methods of interferometry and laser ellipsometry are effective optical methods. By using microscopes or microinterferometers (MIM-7, MII-4, MII-10), it is possible

FOR OFFICIAL USE ONLY

to examine the comparatively small surface areas of the samples. A large field of view provides for examination of the surface by coherent laser radiation.

Lasers are used for quality control of a polished surface, for controlling a plane surface (the UKP-2 laser television interferometer), to measure the thickness and index of refraction of oxide and nitride films on silicon and germanium surfaces (the LEM-2, LEM-3M laser ellipsometric microscopes). The ellipsometric method is based on using the phenomenon of elliptic polarization of light, which arises as a result of the complex nature of the index of refraction of the semiconductor. A plane-polarized monochromatic light beam incident on the surface of the semiconductor is converted after reflection to elliptically polarized as a result of the phase difference of the parallel beam and the reflected beam normal to the surface (the end of the amplitude vector of a reflected electric field describes an ellipse in the plane normal to the direction of the light beam). If there is a thin dielectric film on the surface of the semiconductor plate, an additional phase shift takes place between the beams reflected from the surface of the semiconductor and from the upper surface of the dielectric film. This phase shift depends on the thickness of the film and its coefficient of refraction. The differences in film thickness and the optical constants during ellipsometric microscopy are converted to differences in brightness of the examined specimen. The absolute values of the film thicknesses and optical constants are determined by extinguishing of the beam reflected from the surface on the ellipsometric pattern. By using a vidicon it is possible to reproduce the ellipsometric pattern on an enlarged scale on the screen of the videocontrol device of a television set. Laser ellipsometric microscopes make it possible to monitor oxide layers in openings, traces of contamination after cleaning the surface of a superconductor or glass for photomasks; investigation of the kinetics, the nature of growth of oxide films and the variation and thickness and index of refraction under various effects; control of various metallographic micro-sections.

In contrast to optical methods, x-ray methods permit detection of internal defects: mechanical damage, foreign inclusions, breaks in the leads, sagging, excess weight of the thermocompression balls, cavities between the lower surface of the structure and the base of the case, noncoaxialness of the parts during assembly, and so on. The x-ray television microscopes (MTR-3I, MTR-4) insure high resolution and contrast sensitivity. The samples can be examined in an ordinary work facility, for complete protection against x-radiation is provided in these microscopes.

The methods of electron microscopy are distinguished from the optical and x-ray methods by higher resolution, and they can be used to check the configuration, mutual arrangement and volumetric properties of microcircuit structures with precision technology. As a result of the shorter wavelength of the electron beams used (about 0.037 angstroms for an accelerating voltage of 100 kv) by comparison with optical (about 5000 angstroms) and x-ray (about 1.54 angstroms) beams, the resolution of electron microscopy increases by approximately 1000 times.

The methods of nondestructive control are relatively new, and many of them require highly complex equipment and are still used only in the IC development phase.

Improvement of Technology. Advance technology provides for constant improvement of the quality of products with high cost benefit of production. These basic production criteria are determined by a number of interrelated criteria, in the main ones of which are the following:

FOR OFFICIAL USE ONLY

1. Universality of the technological methods and processes used expressed in the fact that it is possible to perform a large number of operations of different purposes with their help. For example, by using thermovacuum deposition it is possible to obtain thin-film elements of hybrid IC and metallization of semiconducting structures; with the help of photolithography it is possible to obtain a pattern of microcircuits without groups, and also to manufacture free masks, lead frames, and so on.
2. The standardization of the structural elements used (substrates, case parts, fittings, and so on).
3. Standardization of materials used.
4. Standardization of the production processes achieved with high indexes of the first three criteria and permitting in practice the manufacture of a broad nomenclature of products by a united technology (with insignificant variations in conditions or replacement of equipment and materials). For example, epitaxial-planar technology is standardized.
5. Low number of operations.
6. High speed of performance of operations entering into the production process.
7. Continuity of the performance of a number of production operations in the production process with simultaneous group processing.
8. Reproducibility of the parameters in each operation and high percentage yield of usable products.
9. Technological nature of the structural elements of the intermediate products corresponding to the requirements of automated production.
10. Formalization, that is, the possibility of mathematical simulation of the production operations, which is needed to control the technological processes by computer.
11. Adaptability or competitiveness of the technological processes permitting rapid rearrangement under conditions of continuous improvement of production to make new types of products without significant capital expenditures.
12. Economicalness (low consumption of raw material, fuel and materials).
13. Safety with respect to environmental protection.

The improvement of the manufacturing technology of microcircuits is made by improving the methods and processes which have already become traditional by the development and introduction of new methods and new technological solutions into production during the creation of production processes, by the introduction of new materials, reduction of the material consumption of the products of microelectronics as a result of reducing the consumption norms, economy, selection of the optimal structural designs, and so on.

FOR OFFICIAL USE ONLY

Test Questions and Assignments

1. What is the role of the party and the Soviet government in the development of microelectronics in our country?
2. What are the basic areas and the goals of microelectronics in the current five year plan?
3. What is the complex overall approach to the solution of the problem of improving production efficiency and quality?
4. What is program goal planning?
5. What is the role of statistical control in quality control and production efficiency? What is an automated technological process control system (ASU TP)?
6. Explain the structural model of quality control presented in Figure 19-1.
7. What is nondestructive control and what are its basic methods?
8. Explain the essence of the basic criteria of evaluating the progressiveness of technology.
9. Name the basic ways to improve technology.

FOR OFFICIAL USE ONLY

CONCLUSION

A distinguishing feature of microelectronics is the development and mass production of microcircuits in practice for all types of radioelectronic and computer equipment. The proportion of the production of microcircuits in the overall volume of the electronic products will increase continuously. Silicon will remain the most widely used material for semiconductor microcircuits for the next 10 or 15 years. Intensified development of powerful microwave microcircuits is expected. They will displace electrovacuum microwave devices. Gallium arsenide and indium phosphide will be used to manufacture microwave microcircuits.

A characteristic trend in modern microelectronics is the increase in degree of integration and improvement of the operating characteristics of microcircuits. In 1977, the degree of integration of IC increased by an order by comparison with 1976; the capacity of the semiconductor memories increased by more than tenfold. The rapid growth of the degree of integration caused the appearance of a new area -- the creation of microprocessors. The latter embodied the modern achievements of semiconductor technology, experience in the creation of computers and control machines and systems, and also software. A microprocessor is a part of a microcomputer that converts the data constructed from LSI and making up a functionally complete product consisting of a logical arithmetic circuit, registers and control circuits. The appearance of microprocessors will in turn stimulate further improvement of the LSI structural design and technology.

With respect to engineering design execution modern LSI can be divided into four basic groups.

The first group of LSI is made up of an ordinary semiconductor crystal with high packing density of the elements and with a system of fixed, that is, previously given photomasks, interelement connections.

The second group of LSI is an undivided semiconductor plate which under various metallization conditions can combine a large number of functionally different microcircuits (cells). For metallization of such LSI it is impossible to use the method of fixed interelement connections, for it presupposes 100% yield of all of the cells on the plate. Therefore other methods are used.

The method of selective connections consists in the fact that first all the cells for functioning are checked, then by machine design, a set of photomasks is made. In this case, for the measurement of each cell it is necessary to make contact sites, and they occupy the free area of the plate.

FOR OFFICIAL USE ONLY

The second method consists in creating redundant cells and two levels of separate standard metallization for all of the interelement connections inside the cells and for all connections between cells. After checking for functioning, it is necessary to break the connections with defective cells. Here, by comparison with the preceding method, the number of measurements when checking for functioning is decreased.

The third method -- transfer of contact sites -- provides for the manufacture of standard interelement connections within the cells and between cells. However, for the second connections, contact sites separately located on the dielectric are used. Then after the manufacture of the photomask, the usable cells are connected to these contact sites by short jumpers.

In connection with the complexity of carrying out the technological processes on the large-area plates and the great difficulties of creating the intracircuit connections for the LSI of the second group, there is still no united opinion of developers regarding the optimal plate area.

The third group of LSI includes the semiconductor, multicrystalline LSI. In one multilead case individual crystals are mounted and they are connected to each other by means of bulk or wire leads. Structurally, these circuits are simpler and more flexible, but they require a large number of external connections and the creation of complex cases.

The fourth group of LSI is made up of the large hybrid integrated microcircuits (hybrid LSI), which contain a large number of mounted crystals on a dielectric substrate with passive film part of the microcircuit. The hybrid LSI are simpler in design and manufacture, they are more flexible during the course of development, for the problems of heat transfer and stray couplings are more simply solved for them. However, they are inferior to the semiconductor LS with respect to packing density.

The LSI structures are manufactured basically by means of the known technological processes, but they require serious improvement. Improvement in the degree of microcircuit integration is achieved by automation and introduction of the machine design of the topology into the technological process, industrial transition to new methods of formation of the element configuration (including maskless), the introduction of ionic pickling for transferring the layout to the elements, replacement of high temperature diffusion by ion diffusion and other methods of alloying, using optimal structural designs of the elements (for example, MDS and elements with Schottky barriers), the application of multilevel metallization systems on new materials, and so on. According to modern predictions in approximately 10 years the minimum line width in industrial microcircuits will reach 0.1 micron. The functional density of the IC will reach more than 10^6 elements for a crystal several square millimeters in area.

A characteristic feature of the LSI, just as the microcircuits of a lesser degree of integration is high disintegration in the encapsulation step, for the volume occupied by the encapsulated microcircuit is appreciably greater than the volume of its structure. The degree of disintegration on transition from the LSI to the next structural level of microelectronic devices increases by more than two orders. In order to decrease the degree of disintegration of the LSI, they are made in the caseless execution, and then they are mounted on a common substrate with connections. Then the substrate is encapsulated in a flat case. Thus, it is possible to create hybrid LSI of no less than fifth degree integration.

FOR OFFICIAL USE ONLY

The LSI are the last phase of the development of classical microcircuits for which it is possible to isolate regions (elements) equivalent to the previously used discrete devices (see the introduction). In spite of the fact that the IC are the basis of the element base of modern microelectronics, their capabilities are limited and have energy and thermal barriers, structural and functional limits.

The further development of the element base of electronics -- functional microelectronics -- will proceed along the path of integration of new physical phenomena. Physical phenomena which make up the base of functional microelectronics are in the stage of intensive study by the scientists of many countries. When creating functional devices, microcircuits and junctions, not only are electrical processes in a solid state used, but also the processes accompanied by mechanical, thermal, radiation and magnetic phenomena and also electrochemical processes and processes occurring in animate nature. Let us briefly consider the basic areas of functional microelectronics.

Optoelectronics is based on the conversion of optical signals that carry information to electric signals and, vice versa, electric signals to optical signals. The technology of optoelectronics devices is to a great extent analogous to the technology of ordinary microcircuits.

Devices based on the Gunn effect observed in some semiconductors with special structure of the conductivity zone (GaAs) permit powerful microwave vibrations to be obtained. On the basis of the Gunn effect it is possible to create devices which are different with respect to the functions performed.

Acoustoelectronics is based on the phenomena of the interaction of an electron flux with sound waves in a solid state (electron-phonon interactions). Piezoelectric frequency converters, ultrasonic generators and amplifiers, circuits for pulse broadening and compression for radar, delay lines, filters, phase shifters, directional couplers, attenuators and amplifiers for the microwave band have the greatest practical significance. The manufacturing technology of acoustoelectric devices is analogous to the technology of classical integrated microcircuits; it includes the epitaxial growth of semiconductor layers, the application of thin films and electro-nolithography.

Cryoelectronics is developing on the basis of the study of phenomena occurring in solid states at low temperatures and also on the basis of the achievements of cryogenic engineering, that is, the engineering of the creation of cooling systems. The use of low temperatures will permit the limit of the capabilities of radioelectronics obtainable at ordinary temperatures to be overcome, first, as a result of improving the parameters and lowering the noise, and, secondly, as a result of theoretically new phenomena. The latter are also used to create functional micro-devices.

The most interesting phenomenon is the phenomenon of superconductivity which occurs only at low temperatures. A superconductor has zero specific resistance and is characterized by zero magnetic induction. Under the effect of a magnetic field, electric current or temperature of the specific resistance of a superconductor varies from zero to a defined value. Thus, for very small expenditures of energy ("a nondamping current") superconducting materials can operate as switching devices and memories in the computer. Cryogenic elements do not require special heat transfer equipment, for they generate very little heat; they are distinguished by high speed.

FOR OFFICIAL USE ONLY

The cryotronic device based on the Josephson effect (tunnel cryotron) can be executed in the form of two semiconductors separated by very thin insulator film. The current passes between the semiconductors at zero voltage. The current switching in a tunnel cryotron is realized in a tenth of a second.

At the present time the development of cryoelectronics is being delayed by the technical difficulties of manufacturing the cooling systems.

Amorphous semiconductors are prospective for creating devices having a memory effect. The mechanism of nonthermal switching in relatively thin amorphous films with a thickness of less than 1 micron was investigated for the first time by Soviet scientist B. T. Kolomiys, et al. in 1969.

Magnetolectronics uses new magnetic materials (weak ferromagnetics and magnetic semiconductors). A distinguishing feature of weak ferromagnetics is small saturation magnetization by comparison with classical magnetic materials. This permits control of the movement of magnetic nonuniformities -- "bubbles" -- in two and three measurements by weak magnetic fields, and it makes it possible at the same time to realize the storage, displacement and processing functions on large volumes of information. Information storage is realized without feed, and the displacement of the "bubbles" with small dispersion of the power. A number of new materials -- magnetic semiconductors -- having the characteristic features of magnetic materials and semiconductors, will permit the creation of devices with large functional flexibility.

Many functional devices are in the research and development phase. Therefore along with improvement of the existing technological processes the problem arises of creating a new technological base for industrial assimilation of the devices of functional electronics.

Test Questions and Assignments

1. What are the trends in the further development of microelectronics?
2. Give a description of the engineering design features of modern LSI.
3. What gives rise to the necessity for transition from integrated microcircuits to functional microdevices?
4. Give a brief description of the basic areas of functional microelectronics.
5. Read the foreword, introduction, conclusion and table of contents carefully again.
6. Construct a structural-logical diagram of the subject "Production Technology of Microelectronic Devices." Check yourself by the table of contents.
7. Compare the table of production operations for the manufacture of IC with indication of possible methods and equipment used.

FOR OFFICIAL USE ONLY

BIBLIOGRAPHY

1. L. N. Kolesov, VVEDENIYE V INZHENERIUY MIKROELEKTRONIKU (Introduction to Engineering Microelectronics), Moscow, Sovetskoye radio, 1974.
2. Ye. Z. Mazel', F. P. Press, PLANARNAYA TEKHOLOGIYA KREMINEVYKH PRIBOROV (Planar Technology of Silicon Devices), Moscow, Energiya, 1974.
3. I. D. Goloto, B. P. Dokuchayev, G. D. Kolmogorov, CHISTOTA V PRONZVODSTVE POLUPROVODNIKOVYKH PRIBOROV I INTEGRAL'NYKH SKEM (Cleanness and the Production of Semiconductor Devices in Integrated Circuits), Moscow, Energiya, 1975.
4. O. I. Bochkin, V. A. Bruk, S. N. Inkiforova-Denisova, MEKHANICHESKAYA OBRABOTKA POLUPROVODNIKOVYKH MATERIALOV (Machining of Semiconductor Materials), Moscow, Vysshaya skola, 1977.
5. Ya. A. Uray, VVEDENIYE V KHIMIYU POLUPROVODNIKOV (Introduction to Semiconductor Chemistry), Moscow, Vysshaya Shkola, 1975.
6. B. D. Luft, A. M. Shustina, OCHISTKA DETALEY ELEKTRONNYKH PRIBOROV (Cleanness of the Parts of Electronic Devices), Moscow, Energiya, 1968.
7. V. D. Gimpel'son, Yu. A. Radionov, TOKOPLENOCHNYE MIKROSKHEMY DLYA PRIBOROSTROYENIYA I VYCHISLITEL'NOY TEKHNIKI (Thin-Film Microcircuits for Instrument Making and Computer Engineering), Moscow, Mashinostroyeniye, 1976.
8. Yu. S. Bokov, V. S. Korsakov, V. P. Lavrishchev, A. V. Lubashevskaya, V. I. Mishachev, VVEDENIYE V FOTOLITOGRAFIYU (Introduction to Photolithography), Moscow, Energiya, 1977.
9. FOTOLITOGRAFIYA I OPTIKA (Photolithography and Optics), Edited by Ya. A. Fedotov, G. Pol', Moscow, Sovetskoye radio, Berlin, Tekhnika, 1974.
10. B. S. Danilin, POLUCHENIYE TONKOPLENOCHNYKH ELEMENTOV MIKROSKHEM (Obtaining Thin-Film Elements of Microcircuits), Moscow, Energiya, 1977.
11. G. L. Zhironov, TEKHOLOGIYA GIBRIDNYKH INTEGRAL'NYKH MIKROSKHEM (Technology of Hybrid Integrated Microcircuits), Kiev, Vishcha shkola, 1976.
12. D. Hamer, J. Briggers, TEKHOLOGIYA TOLSTOPLENOCHNYKH GIBRIDNYKH INTEGRAL'NYKH SKHEM (Technology of Thin-Film Hybrid Integrated Circuits), Moscow, Mir, 1975.
13. TEKHOLOGIYA TONKIKH PLENOK. SPRAVOCHNIK. TOM 1 (Thin Film Technology. A Reference. Vol 1). Edited by L. Maysse, R. Gleng, Translated from the English, Edited by M. I. Yelinson, G. G. Smolko, Moscow, Sovetskoye radio, 1977.
14. TEKHOLOGIYA TONKIKH PLENOK. SPRAVOCHNIK. TOM 2 (Thin Film Technology. A Reference. Vol 2), Edited by L. Maysse, R. Gleng, Translated from the English Edited by M. I. Yelinson, G. G. Smolko, Moscow, Sovetskoye radio, 1977.
15. A. I. Kurnosov, MATERIALY DLYA POLUPROVODNIKOVYKH PRIBOROV I INTEGRAL'NYKH SKEM (Materials for Semiconductor Devices and Integrated Circuits), Moscow, Vysshaya shkola, 1976.

FOR OFFICIAL USE ONLY

16. A. I. Kurnosov, V. V. Yudin, TEKHNOLOGIYA PROIZVODSTVA POLUPROVODNIKOVYKH PRIBOROV (Production Technology of Semiconductor Devices), Moscow, Vysshaya shkola, 1974.
17. V. V. Novikov, TEORETICHESKIYE OSNOVY MIKROELEKTRONIKI (Theoretical Principles of Microelectronics), Moscow, Vysshaya shkola, 1972.
18. V. N. Chernyayev, TEKHNOLOGIYA PROIZVODSTVA INTEGRAL'NYKH MIKROSKHEM (Production Technology of Integrated Microcircuits), Moscow, Energiya, 1977.
19. TEKHNOLOGIYA IONNOGO LEGIROVANIYA (Ion Alloying Technology), Edited by S. Namba, Translated from Japanese by V. F. Ovcharov, Edited by Professor P. V. Pavlov, Moscow, Sovetskoye radio, 1974.
20. A. I. Kurnosov, ZASHCHITA I GERMETIZATSIYA POLUPROVODNIKOVYKH PRIBOROV I INTEGRAL'NYKH SKEM (Protection and Encapsulation of Semiconductor Devices and Integrated Circuits), Moscow, Vysshaya shkola, 1978.
21. K. A. Valiyev, A. N. Karmazinskiy, M. A. Korolev, TSIFROVYYE INTEGRAL'NYE SKEMY NA MDP-TRANZISTORAKH (Digital Integrated Circuits Based on MDS-Transistors), Moscow, Sovetskoye radio, 1971.
22. A. Yu. Ber, F. Ye. Minsker, SBORKA POLUPROVODNIKOVYKH PRIBOROV I INTEGRAL'NYKH MIKROSKHEM (Assembly of Semiconductor Devices and Integrated Microcircuits), Moscow, Vysshaya shkola, 1977.
23. V. M. Val'kov, ASU TP V PROIZVODSTVE IZDELIY ELEKTRONNOY TEKHNIKI (Automated Technological Process Control Systems in the Production of the Products of Electronic Engineering), Moscow, Sovetskoye radio, 1974.
24. Ya. A. Fedotov, POLUPROVODNIKOVAYA ELEKTRONIKA GOD 2001-I (Semiconductor Electronics of the Year 2001), Moscow, Sovetskoye radio, 1975.
25. V. V. Shebanin A. A. Tykhin, V. I. Tomov, INTEGRAL'NYE FUNKTSIONAL'NYE UZLY DLYA ZAPOMINAYUSHCHIKH USTROYSTV (Integrated Functional Subassemblies for Memories), Moscow, Sovetskoye radio, 1976.

FOR OFFICIAL USE ONLY

SUBJECT INDEX

A

Autoalloying 181, 274
Autoepitaxy 173, 274
Adhesion 251, 256, 259, 364
Acoustoelectronics 434
Aluminum 242, 247, 250, 251, 271

G

Encapsulation 25, 395, 406-413
Heteroepitaxy 173, 186, 189

D

Defects 55, 116, 193, 196, 200, 220, 226, 263
Diffusion 201, 304
-- base 277
-- in ampoules 217
-- in semien encapsulated space 218
-- in a flow tube 212
-- in a forevacuum 218
-- double 312
-- two-stage 210, 276
-- from ion-alloyed layers 238
-- insulating 273, 276-278
-- from polycrystalline silicon 238
-- local 203, 273
-- shallow 300
-- interstitial impurities 204
-- radiation stimulated 239

Z

Protection 86, 313, 314, 316, 344, 362

I

Alloying admixture sources:
gaseous 192, 217
liquid 191, 216

FOR OFFICIAL USE ONLY

with ion introduction 236
solid 192, 212

K

Channeling 225
Contacts:
 rectifying 241
 low-penetrating 252
 ohmic 245
 expanded 260
Case 406, 409-411
Cryoelectronics 435

L

Alloying 191, 192, 194, 222, 234, 236, 275, 301

M

Magnetolectronics 436
Mask:
 contact 22, 340
 free 22, 121-129, 328
Copper 333
Mesa structure 107
Method 11
 -- vacuum-liquid 416
 -- group 12
 -- hybrid 181
 -- individual 12
 -- interferometry 197, 427
 -- studies in the microscope dark field 80
 -- cathode sputtering 153, 254, 258
 -- laser ellipsometry 427
 -- mass spectrometric 416
 -- microsection dying 197, 220
 -- deposition from the vapor-gas phase 161, 253, 254, 258
 -- "spider" leads 390
 -- "inverted crystal" 378
 -- pyrolytic depositon 164
 -- x-ray 428
 -- thermovacuum depositon 145, 182, 248, 254
 -- stenciling 127
 -- "condensation figure" 81
 -- electron microscopy 428
Mechanically disturbed layer 56, 58
Microsoldering 388
Microprocessor 431

N

Nonplanarity 56, 58

FOR OFFICIAL USE ONLY

Nickel 254
Nitride:
 silicon 164, 165, 271, 277, 314, 315
 tantalum 339
Nichrome 337

O

Production operation 9
-- auxiliary 10
-- finishing 25, 379, 389, 417
-- machining 46
-- degreasing 62, 64
-- painting 418
-- basic 10
-- assembly 366
Optoelectronics 434
Annealing:
 plates 231
 films 151, 251
 radiation defects 226

P

Soldering 372, 377, 388, 389
Paste 45, 351, 353, 361
Paste, manufacture 355
-- sintering 358
Films:
 silicon dioxide 163
 protective 351, 362
 masking 163, 271, 272
 multicomponent position 336
 aluminum oxide 165, 339
 polyimide 259
 refractor metals 165, 166, 233, 254
 chromium 254, 332
Sublayer 251, 254
Pickling 251
Washing 67, 68
Development 93, 103, 104

R

Technology development 15, 16
Admixture distribution 199, 206, 207
Admixture solubility 204, 205
Cutting 46-50, 271, 367
X-ray resists 135

S

Welding 374
-- argon-arc 404

FOR OFFICIAL USE ONLY

- indirect pulse heating 383, 384
- laser 386, 404
- roller 404
- double electrode 385
- thermocompression 380
- spot 403
- ultrasonic 384
- cold 401
- resistance 403
- electron beam 387, 405
- Bonding 374
- Scribing 367, 368, 371
- Layer:
 - contact 253
 - conducting 254
 - separating 256
- latent collector 175, 273, 275
- Polymerization 87
- Sublimation 183
- Drying 69, 73, 96, 105

T

- Thermocompression 381-383
- Heat treatment 74, 75, 105, 251, 255, 358
- Technical specifications 17
- Technical assignment 16
- Process conversion 9
- Technological process 10
 - standard 11, 73
- Technology 9
 - diffusion-ion 302
 - "silicon on sapphire" 283, 306
 - planar 23, 209, 270, 271
 - planar-epitaxial 305
 - advanced 429
- Thixotropicity 354
- Pickling 62, 66
 - anisotropic 272
 - gas 63, 75
 - ion beam 78
 - oxygen 66
 - local 243, 277
 - plasma 63, 74
 - selective 335, 337

U

- Control:
 - quality 425, 426
 - technological processes 425

FOR OFFICIAL USE ONLY

F

Photolysis 87, 88
Photolithography 84, 85, 130, 261, 301, 334
-- contactless 130
-- double 336
-- in a microgap 132
-- inverse 107, 108, 334
-- single 337
-- projection 132-134
-- direct 105, 334
Photoresists 86, 87-91, 94, 103, 104
Photomasks 86, 111, 113, 115
Fritting 355, 408

Ts

Centrifuging 69, 96

Sh

Roughness 56, 58
Stamping 396, 414

E

EVH 39, 41
Exposure 101, 102
Electrodifffusion 239, 250, 257, 264
Electronoresists 139
Epitaxial:
 layers 193, 197-199
 structures 174, 175, 273, 275
Erosion 30, 219, 229, 255, 275
Effectiveness 420-422

FOR OFFICIAL USE ONLY

ANNOTATION

A general description of the production of microelectronic devices is presented, and the general requirements on the production of these devices are formulated. Basic production methods and processes such as machining, cleaning the surface of the substrates, photolithography, the free mask method, x-ray and electron lithography, methods of obtaining films, epitaxy, diffusion and ion alloying are described. The standard technological processes of manufacturing bipolar and MDS microcircuits, thin-film and thick-film microcircuits are analyzed. A study is made of the processes of assembly, encapsulation, the problems of quality control and efficiency and also trends in the further development of microelectronics.

The book is intended as a text for students at the technical high schools studying the specialty of "Microelectronic Production."

COPYRIGHT: Izdatel'stvo Energiya, 1980
[8144/0993-10845]

10845
CSO: 8144/0993

END

FOR OFFICIAL USE ONLY