NBSIR 81-2197

Semiconductor Technology for the Non-Technologist

Robert I. Scace

Electron Devics Division
Center for Electronics
and Electrical Engineering
U.S. Department of Commerce
National Bureau of Standards
Washisngton, DC 20234

January 1981

.S. DEPARTMENT OF COMMERCE
ATIONAL BUREAU OF STANDARDS

NBSIR 81-2197

SEMICONDUCTOR TECHNOLOGY FOR THE NON-TECHNOLOGIST

Robert I. Scace

Electron Devics Division
Center for Electronics
and Electrical Engineering
U.S. Department of Commerce
National Bureau of Standards
Washington, DC 20234

January 1981

U.S. DEPARTMENT OF COMMERCE, Malcolm Baldrige, *Secretary*
NATIONAL BUREAU OF STANDARDS, Ernest Ambler, *Director*

Table of Contents

Preface

As the semiconductor industry has grown in size, more and more people who are not semiconductor specialists have become involved with the industry. Some are clerical or other nontechnical employees, some are in other businesses such as equipment suppliers to the semiconductor industry, and some are in the communications media. In particular, the recent wave of studies of the industry has introduced many economists and government people to a highly technical subject which, at times, seems not to be well understood.

Yet some degree of understanding of the terms and concepts basic to semiconductors is needed by these people, whose specialties are in other fields. This text, which is intended to meet that need, was originally prepared in 1976 to be part of "A Report on the U.S. Semiconductor Industry" by the Department of Commerce. That report was issued in September 1979. Both before the issuance of the Commerce report and since, there have been requests for this material from people who were first getting acquainted with semiconductors. Their responses, and a continuing demand for copies, have suggested that independent publication would be useful.

The present text has been slightly revised to bring it up to date. The industry moves rapidly, and four years' time has brought changes. It is quite possible that future re-issues may occur. Suggestions from readers regarding possible improvements are welcome.

SEMICONDUCTOR TECHNOLOGY FOR THE NON-TECHNOLOGIST

Robert I. Scace
Electron Devices Division
National Bureau of Standards
Washington, DC 20234

Abstract

The properties of semiconductor materials, the methods of processing them, and the solid-state products made from them are described in terms intended to be understandable by the lay person. The semiconductor industry has grown at a rate of 21 percent per year compounded for the last twenty years, and its products have declined in unit cost by a factor of five in current dollars (a factor of ten in constant dollars) in the same period. This very satisfactory but anomalous behavior has attracted the interest of many who are not familiar with the technology of the industry, yet who need to have some understanding of it. This report is intended to help meet that need.

Key Words: Integrated circuits; semiconductors; semiconductor devices; semiconductor processes; semiconductor technology; silicon.

1. INTRODUCTION

This is a detailed discussion of semiconductor technology using nontechnical terms so readers not familiar with the subject can comprehend its complexity and breadth. While the knowledge underlying semiconductor technology is primarily in the fields of physics and electrical engineering, many of the processes used to manufacture semiconductor devices borrow heavily from the field of chemistry. A manufacturer of semiconductors will have on its professional staff physicists, chemists, metallurgists, ceramists, electrical engineers, mechanical engineers, industrial engineers, and possibly specialists in optics. Because of the great diversity of technical information required to make semiconductor devices, no one kind of professional employee is sufficient to do the entire job. Semiconductor products result from the team efforts of a variety of professionals.

2. BASIC MATERIALS

Semiconductor materials generally are crystalline substances of relatively simple physical and chemical composition. Because of this simplicity, good theoretical understanding of the properties of semiconductors has been developed. The theory is extensively used to understand the behavior of solid-state devices and to design new ones. While there are many semiconducting compounds and elements, only a few have commercial significance. Originally, because of the relative simplicity of processing it, germanium was the material most commonly used. From the middle 1950s to the middle 1960s, silicon gradually became the primary semiconducting material for several reasons. First, since silicon is much more abundant than germanium, it is cheaper. Second, germanium devices generally will not operate above 85°C, whereas silicon devices can operate up to 200°C. Third, and probably most important, silicon has a tenacious, electrically insulating, protective oxide. The oxide of germanium is not protective. The properties of silicon dioxide have important consequences in the processing of semiconductor devices, and they have made possible the batch processing necessary to make solid-state devices at low cost. Therefore, the following discussion will deal exclusively with silicon, except in special instances which will be noted.

The crystalline structure of germanium and silicon is basic to semiconducting action. Pure silicon at room temperature has the very high resistivity of 230,000 ohm-centimeters, which makes it a poor conductor of electricity. This arises because the atoms in the silicon crystal structure are arranged so that each atom has four other silicon atoms around it at equal spacings. This so-called diamond cubic arrangement allows each of the four outer electrons of the silicon atom to be bound to the four neighboring atoms. In turn the four neighboring atoms have their electrons similarly bound. Each interatomic bond in the crystal lattice therefore ties up two electrons, one from each of the atoms connected by any particular bond. Thus, all the available outer electrons in a perfect silicon crystal are bound together and are not available to conduct electricity, if the effects of temperature are ignored. At temperatures other than absolute zero, the atoms in a crystal vibrate about their nominal positions, and a few bonds may at any instant be disrupted. The electrons released this way do permit a slight amount of electrical conduction, and therefore the resistivity of silicon at room temperature is not infinitely high.

If, however, elements from neighboring columns of the periodic table are introduced into the crystal lattice, this arrangement of electrons is disturbed. For instance, if a phosphorus atom, having five outer electrons, is substituted for a silicon atom in the crystal, four of the five electrons will be bound to the crystal lattice, but one will be available to move freely within the crystal. This makes the crystal *n*-type, meaning that the principal carriers of the electric current are negative charges (electrons). Conversely, if a boron atom, having three outer electrons, is substituted for a silicon atom in the crystal, the electrons can complete only three of the four necessary bonds. The absence of an electron in the one remaining bond represents a departure from the symmetrical bond arrangement of a perfect crystal. It is possible for an electron from a neighboring bond to move into this space and create a vacancy, having a net positive charge, in the bond left behind. This vacancy, with its positive charge, can propagate through the crystal in a manner quite similar to that of an electron, with the exception that the sign is positive instead of negative. This charge-carrying entity is called a hole, and crystals in which this (positive) charge transport mechanism predominates are referred to as *p*-type.

The elements that cause *p*-type or *n*-type conduction are referred to as dopants or impurities. They are impurities only in the sense that they represent a departure from pure silicon crystal. The silicon itself is quite pure as are the dopant elements. In fact, silicon is the purest raw material ever produced in tonnage quantities. Impurity levels of one part or less per billion are common in silicon used for semiconductor manufacture.

Dopant elements may be added to the silicon at several points in the semiconductor device manufacturing process. They can be added at the time that the single crystal of silicon is formed, or later by localized melting and recrystallization processes (alloying), by high temperature diffusion processes, during the growth of additional layers of silicon, by high energy bombardment (implantation), or by nuclear transmutation. As purchased, the doping of silicon is such that its resistivity usually falls between 0.001 and 1000 ohm-centimeters. Values of resistivity outside this range are achievable but not often used. The effects of dopant elements are quite significant, permitting silicon to have nearly any desired resistivity and both *p*- and *n*-type conductivity.

The boundary between a *p*-type region and an *n*-type region is called a *p-n* junction. The active portion of a semiconductor device consists of *p*- and *n*-type regions and the junctions between them. The processes for making such devices provide the capability to introduce dopant atoms in predetermined locations and amounts with precise control and reproducibility. The ways in which this is done will be discussed later.

Various external layers can be arranged to provide electrical insulation or connection to parts of the device. Of the insulators, silicon dioxide is most common partly because it is a byproduct of many process steps and is there anyway, and partly because of its excellent electrical insulation properties. In some cases, layers of aluminum oxide or silicon nitride are used. Where insulation thicknesses of more than one micrometer are needed, various special glass formulations can be deposited on the wafer surface.

The most common metals used for connection to the silicon are aluminum and gold, since both are readily available in very high purity, can be vacuum-evaporated easily, and form good electrical connections to silicon. For special purposes, very complex sequences of metal layers can be used, though at considerable extra cost. Even heavily doped silicon itself is used. The various metal films serve to connect the often minuscule regions of the semiconductor device together with each other, and to provide regions to which external connecting leads can be attached.

In general, all materials used in a semiconductor device--the silicon, its dopants, the coatings applied, and the metallization--must be extremely pure. Because elements other than dopants often adversely affect the electrical properties of the device if they find their way into the silicon, they must be excluded. The only way this can be done is to use materials of the highest purity. Uncontrolled elements can lead to uncontrolled effects in processing and to uncontrolled device electrical characteristics at the end of the process. The same stringent demands of purity are made of the various chemicals used in cleaning, etching, and otherwise processing the silicon from beginning to end of the process. One of the major technical achievements, not widely known outside the industry, towards making modern semiconductor devices possible was the achievement of suitable purity levels in the wide range of ancillary materials needed by the industry. Twenty-five years ago, such high purity materials were known only in the research laboratory. Today they are available off the shelf in most technically advanced parts of the world.

3. BASIC DEVICES

The *p-n* junction diode and the transistor, both elementary devices, are useful in themselves and in combination with one another can make up integrated circuits.

Figure 1 illustrates a *p-n* junction diode. It is divided into two portions by the *p-n* junction. In the upper portion the majority carrier of electrical current is electrons, while in the lower portion holes carry the current. If the upper electrode is biased negatively and the lower positively, then the charge carriers in those regions are repelled towards the *p-n* junction, and in crossing that junction find themselves in regions of the device in which they are not in equilibrium. An electron, for example, flowing downward past the *p-n* junction will combine with one of the holes with which it is surrounded and disappear. Other electrons will suffer the same fate. Conversely, holes flowing upward into the *n*-type region will combine with electrons there and disappear. This flow of electrons and holes across the *p-n* junction constitutes an electric current through the device and through the external circuit. This direction of polarity is called forward bias. If the upper electrode has a positive potential applied and the lower electrode a negative potential, the charges in the respective upper and lower regions are attracted away from the *p-n* junction leaving behind a region with no free electric charge carriers. This region acts as an effective insulator, and virtually no current exists. This direction of polarity is referred to as reverse bias. This diode or rectifier then acts as an electrical valve that permits the passage of current in one direction and obstructs current in the opposite direction.

Figure 2 is a schematic sketch of an *npn* transistor. (There are other kinds of transistors, too.) The upper *n*-region is biased rather strongly positively, and the upper *p-n* junction is therefore reverse biased. There is very little current in the circuit connecting the upper *n*-region with the lower *n*-region. If a positive potential is applied to the central *p*-region, holes pass from the *p*-region into the lower *n*-region and electrons pass from the lower *n*-region into the central *p*-region. A current therefore exists in the left-hand loop of the sketch. Although the figure does not show it, the *p*-region is extremely thin. Therefore, most of the electrons passing from the lower *n*-region into the central *p*-region do not have time to recombine with holes there, but they survive to reach the upper *p-n* junction. At this point they are attracted to the much larger positive potential on the other side of the junction and are swept across it. This constitutes a current in the right-hand loop of the sketch. If current in the left-hand loop is interrupted by breaking the circuit, the current in the right loop ceases. Similarly, an increase in current in the left-hand loop will lead to an increased current in the right-hand loop. Thus one current controls another, and this is the essential property of the transistor. The device illustrated is a junction transistor, so called because it contains two *p-n* junctions. Other transistors, such as the junction field-effect transistor and the metal-oxide-semiconductor field-effect transistor, accomplish a similar function although by different means.

The two devices described above depend in their action upon the flow of electrons and holes under the influence of external currents and voltages. While

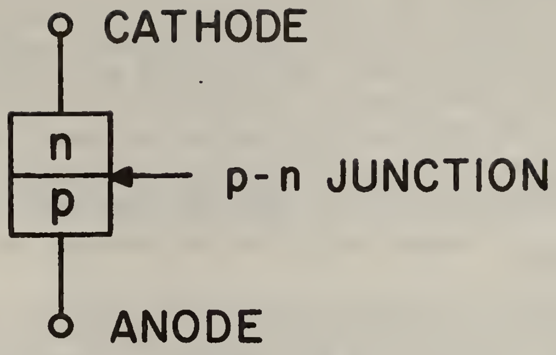

Figure 1. Elementary junction diode.

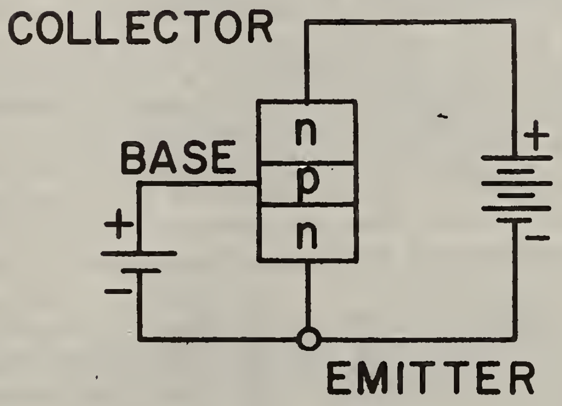

Figure 2. Elementary *npn* transistor.

the operation of semiconductor devices will not be discussed in detail, it should be kept in mind that all devices, from the simplest *p*-*n* diode to the most complex integrated circuit, operate using these elementary concepts. The art of making semiconductor devices is simply that of arranging the *n*-type and *p*-type regions, the junctions between them, and their interconnections so that useful electrical properties result.

4. PROCESSING TECHNOLOGY

Many techniques are used to manipulate the dopants and other parts of the semiconductor device structure so as to achieve useful results. The manufacture of a given kind of device may use only some of these processes, but they will all be briefly described to illustrate the wide range of technical disciplines necessary to make semiconductor devices. Any semiconductor manufacturing plant must be able to perform these processes in a controlled and reproducible fashion if it is to make devices with any degree of success.

4.1 Crystal Growth

Most of the metallic and ceramic objects with which we are familiar are composed of many tiny crystals of the material in question and are described as polycrystalline. Generally, these small crystals can be seen only under high power microscopes. In contrast, semiconductor devices are manufactured using a starting material that is a single crystal. This is required because the boundaries between crystallites in ordinary materials have different physical properties from the body of the crystal. If one attempted to make semiconductor devices from polycrystalline silicon, the diffusion processes and other techniques used to introduce the dopants into the silicon would result in dopant concentrations at the crystallite boundaries different from the concentrations in the interior of the crystallites and the desired degree of control over the structure would be impossible.

Semiconductor grade silicon is prepared in a classical chemical plant. Most manufacturers of semiconductor silicon are also substantial manufacturers of other chemicals. One of the intermediate materials in silicone* manufacture is particularly suited for the preparation of the semiconductor grade silicon. This is reduced to polycrystalline high purity silicon in the form of long slender rods. These rods are converted into single crystal silicon of the appropriate doping and physical dimensions in one of two ways.

In the Czochralski method, a quantity of silicon is melted in a quartz crucible under an atmosphere free from oxygen. A slender single-crystal bar (seed) of silicon is introduced into the melt from above and a portion of the melt begins to freeze onto the seed. If the temperature and other conditions are correct, the growth onto the seed crystal is a single crystal whose crystal structure is a continuation of that of the seed crystal. In this fashion the liquid silicon in the melt is converted into a cylindrical ingot of silicon without any intercrystalline boundaries. The doping of a Czochralski crystal is performed by adding appropriate dopant elements to the melt prior

---

*Distinct from silicon. Silicones are a class of chemicals similar to organic chemicals, but with a silicon-oxygen chain as a structural basis rather than a carbon-carbon chain.

to growth. In commercial practice, crystals up to 125 mm (5 inches) in diameter and up to 750 mm (30 inches) long are available. Larger crystals can be made. While the Czochralski process produces semiconductor silicon appropriate for manufacture of many kinds of devices, the silicon contains an appreciable amount of oxygen dissolved from the quartz crucible in which the melt was held. In addition, during the growth process some segregation of the doping element takes place. That is, the solubility of most dopants is greater in the melt than it is in the crystal. The concentration of dopant in the melt thus increases continuously during the growing process. The finished single crystal does not contain a uniform concentration of the doping element from one end to the other and contains oxygen that is troublesome in the manufacture of some devices, especially those required to operate at high voltages.

Some of these difficulties can be overcome by growth by the float-zone method. In this technique, a cylindrical rod of polycrystalline silicon is suspended in a chamber in an atmosphere free from oxygen.

The bottom of the rod is melted with a high frequency induction coil, forming a liquid drop hanging from the bottom of the silicon bar. A seed crystal is introduced into this silicon drop from below, and by controlling the temperature of the molten region and the other thermal parameters of the apparatus, silicon is made to grow upon the seed as a single crystal. By raising the induction heating coil, the molten region of the bar is raised slowly, dissolving silicon from the polycrystalline rod above and growing it as single crystal onto the crystal ingot supported by the seed below. The molten region is held in place by a combination of surface tension and magnetic forces from the induction heating coil. Doping in this process is either by means of a small charge of concentrated dopant added to the initial melted drop before growth on the seed has begun, or by incorporating a gaseous compound of the dopant into the atmosphere around the growing crystal in a concentration such that it dissolves in the molten region at the same rate at which dopant atoms are withdrawn from the molten region by the growing single crystal. Float-zone silicon of 100-mm (4-inch) diameter is currently commercially available. This material is quite free from oxygen and can have a much more uniform distribution of doping species incorporated in it in comparison with Czochralski silicon.

Another technique for producing exceptionally uniformly doped *n*-type crystals of silicon employs neutron irradiation. If an undoped crystal is exposed to neutrons in a nuclear reactor, some of the silicon atoms will absorb a neutron and be transmuted into phosphorus atoms. Since this occurs at random, the newly created phosphorus (an *n*-type dopant) is distributed uniformly throughout the crystal. Such a degree of uniformity of doping is impossible by the methods described earlier. Local fluctuations in resistivity caused by nonuniform conventional doping often amount to >20 percent of the average. Since resistivity is one of the principal factors influencing the maximum blocking voltage of a *p-n* junction, these deviations are one cause of variations in the voltage ratings of devices made from conventionally doped silicon. This is a particular problem in large, high-voltage devices used for power control. Because of the large area of the silicon slices from which these devices are made, there is a high probability that there will be a lower than average resistivity region somewhere in the slice. This causes the breakdown voltage at that point to be less than at other places, thus limiting the voltage rating of the device to a lower value than it otherwise would be. The uniform resistivity of neutron-irradiated silicon allows this difficulty to be avoided. Devices made from this material can have a 15- to 20-percent higher maximum voltage rating without the necessity for making any changes in design and without affecting any of the design trade-offs that would otherwise have to be made.

4.2 Units of Measurement

At this point we must digress for a moment to become familiar with the terms of measurement to be used below. The semiconductor industry, being international in scope, uses metric measures to a great extent. Most Americans are not yet accustomed to thinking in these units, and they especially do not have an intuitive feeling for very small dimensions expressed in metric terms. This discussion is meant to relate metric linear measures to familiar objects.

A nickel is about 2 centimeters in diameter.
1 centimeter (cm) = 0.3937 inch

The diameter of paper clip wire is about 1 millimeter.
1 millimeter (mm) = 0.1 cm

A human hair is about 100 micrometers in diameter,
and cigarette paper is about 25 micrometers thick.
1 micrometer (μm) = 0.001 mm (one-thousandth mm)

The wavelength of green light is about 0.5 micrometers, and this dimension is close to that of the finest detail that can be seen using a high-powered light microscope.

There are no familiar objects in the size range from 1 to 25 micrometers simply for the reason that they are so small that they cannot be seen by the unaided eye; yet, this is just the region of dimensions of many of the parts of semiconductor devices as will be seen presently. Now, let us return to our discussion of silicon.

4.3 Mechanical Processes

Crystals obtained by either the Czochralski or float-zone method are usually ground to a predetermined diameter using diamond grinding equipment. Since uniform diameter of the crystal is important in subsequent processing operations, most semiconductor manufacturers buy their silicon to a specified diameter. At this point, the silicon has a market value ranging from 50 cents to $1 a gram ($230 to $460 a pound). Because of this high value, subsequent operations must be conducted so as to minimize the waste of silicon.

The crystal next is sawed into thin slices, using diamond-edged blades. The difficulty of this operation can be appreciated when one considers that silicon is as brittle as glass, and the slices may be as thin as 0.25 mm. Large-diameter thin slices are delicate, and the sawing operation must be conducted with great care. For many applications the slices need not be so thin, and for reasons of mechanical strength are more often in the thickness range of 0.5 to 0.8 mm. Slices are etched and lapped on both sides to remove the structural damage to the surface layers of the slice done during the sawing operation. For most transistor and integrated circuit manufacture, the slices are then polished to a mirror finish on one side. In this way all mechanical damage to the crystal on that side is removed, preparatory to further processing.

It is necessary at this point and at many subsequent points to clean the slices thoroughly. Mechanical operations are usually relatively dirty. The dirt includes small particles of silicon as well as other materials. These must all be removed from the silicon slice in preparation for later work. Cleaning involves solvent degreasing, heated detergent solutions under ultrasonic agitation, and physical wiping operations. In many cases the final cleaning step is a chemical etching operation, which removes some of the silicon from the surface.

4.4 Epitaxy

The cleaned, polished slice is now ready for growth of a layer of silicon by an epitaxial process. Epitaxy denotes a crystalline growth such that the atomic arrangement on the surface of the original slice (substrate) is continued into the growing layer. Growth is done at high temperature in an atmosphere containing a silicon compound that decomposes on the hot slice and leaves silicon there. A dopant compound is also present, which decomposes in a similar way to provide the desired doping of the growing silicon layer.

Generally, epitaxial layers are from 2 to 20 micrometers thick and are the silicon layer in which the final devices will be formed. The function of the substrate is simply as a host surface for the epitaxial layer and as a thick mechanical support. The use of epitaxy permits the manufacturer to buy silicon in large quantities to relatively undemanding specifications (thus keeping the cost down), and to grow a layer tailor-made for the specific kind of product he makes.

4.5 Oxidation

As a first step to a diffusion operation, the slice is put in a high temperature furnace for growth of an oxide layer over its entire surface. The oxidation takes place at a precisely maintained temperature, usually between 900 to 1200°C, in a controlled atmosphere containing oxygen. Not only are the composition of the atmosphere and its purity maintained, but the temperature is also closely controlled. When diffusion became an important semiconductor process, manufacturers usually built their own furnaces because no suitable one could be bought. Typical present-day furnaces for semiconductor manufacture can maintain temperatures of 1000°C with a precision of one-half degree over a space exceeding 1 meter in length and 15 cm in diameter. Before the advent of semiconductor devices, there was no requirement for furnaces with such precise control. Most of the capital equipment used in semiconductor processing has similarly evolved from laboratory-made equipment to refined, commercially available equipment as new processes develop from laboratory scale to standard manufacturing operations. The extreme demands of cleanliness, precise dimensional control, or uniformity of temperature, unique to the semiconductor industry, have led to the creation of supporting industries manufacturing highly precise equipment for use in making semiconductor devices.

4.6 Lithography

To form a doped pattern in the slice of silicon by diffusion, openings must first be made in the layer of oxide created in the preceding step. A master pattern, prepared photographically on a glass plate, contains an image of the desired array of openings. A layer of photosensitive material (resist) is spread on the surface of the silicon wafer and dried. It is exposed to ultraviolet light through the master pattern to expose certain areas of the resist. The pattern is developed chemically, leaving areas of the slice covered with resist while other areas are clear. The slice is then etched to remove the silicon dioxide in the clear areas, following which the remaining resist is stripped away. The regions now free of oxide can be accessed by a diffusion process to add impurities to these areas.

The photographic plate containing the pattern that was used in the preceding step is generated itself by a complex sequence of operations. This may start with a precise master drawing of the pattern at 100 to 200 times its final size. This drawing is frequently generated by computer-controlled drafting equipment. Using specially designed cameras, the drawing is reduced in size, generally in two steps, to an image on a glass plate at the final size of the microcircuit. The single final image is used in a step-and-repeat camera to generate photographically a master mask containing many copies of the same pattern.

Most integrated circuit processes require a set of four to ten such patterns for the sequential steps necessary to fabricate the device. Each of the patterns in the set is reproduced in multiple. The resulting master set of plates, containing many images each, is used to reproduce working copies that are either used directly in manufacture or are the working masters to create the plates used in manufacture.

The foregoing procedure with its many optical steps is difficult to control to the degree of precision needed for making today's complex integrated circuits, and two newer methods are more widely used. The pattern generator technique employs a computer-controlled machine which produces tiny, intense pulses of light. The machine scans the photographic plate in a series of parallel lines, in much the same way as that in which a television picture is painted on the screen. The light pulses are produced as needed to expose the plate and build up, point by point, the desired image. Integrated circuit designs are produced with the aid of computers, and the data for controlling the pattern generator are produced by the computer design system on a magnetic tape. The image formed by the pattern generator can be the same size as the final circuit, or it can be few times larger. In the latter case, the final size image is produced on the silicon slice by optical reduction at the time the silicon is being processed.

The smallest images are produced with electron beams instead of by light. Because light has a wavelength of dimensions comparable to the dimensions of many parts of an integrated circuit, it is not possible to produce sharply defined images at the desired scale with light. Electrons which have been accelerated by a potential of 10 kilovolts or so have a much smaller wavelength than light, and are used in a machine which operates in much the same way as a pattern generator to draw tiny patterns on a photographic plate. For the most demanding imaging, at dimensions less than one micrometer in size, electron-beam equipment can be used to write the pattern directly on each slice. Special electron-sensitive resists are used. The process at present is much slower than the conventional optical one, and requires the use of a machine which costs well over a million dollars. The cost has so far limited the use of direct electron-beam writing to those circuits which must be made in this way if they are to be made at all.

In repeating photomasking operations, it is necessary to align the new pattern precisely with that already on the slice. This is done with the aid of a sophisticated mechanical apparatus called a photomask aligner. In making complex integrated circuits, it is necessary to align patterns to within typically 1 µm across the entire area of a 100 mm slice of silicon. Difficulties in this process will be appreciated when it is realized that linewidths of 3 µm are commonly used in making integrated circuits. A complex integrated circuit may contain thousands of lines of these dimensions. The patterns in sequential masks in a set thus must align with one another to within substantially less than that dimension, such that all the patterns in the sequence will be correctly aligned everywhere on a large slice of silicon. Further, any defects in the images will be reproduced on the silicon slice as defects in the pattern of resist. This leads to diffusion where it is not wanted or a lack of diffusion where it is wanted, and frequently to failure of that particular segment of the slice to produce a good final device. Scratches, dirt, and misalignment all contribute to losses of what would otherwise be good devices.

Such defects in pattern definition are a major cause of yield loss in device manufacture. In large, complex integrated circuits the yield may be as low as 5 percent. That is, only one good circuit is produced for every 20 that are started at the beginning of processing of the silicon slice. In some cases, duplicate photomasking operations are done, using two different photomasks having the same image patterns. Flaws in the pattern from random causes will not generally be in identical locations on both photomasks, and thus some of the defects produced by one mask will be corrected by the second. Of course, this doubles the labor expense of the operation.

The most advanced optical photolithographic printers use a projection method, rather than the simpler contact printing procedure. Masks have a short life in contact printing since they are exposed to the hazards of contamination or scratching. Projection printing allows the mask to be better protected. To reduce the incidence of defects in the patterns on the silicon, masks for projection printing are rigorously inspected and defects are repaired before the masks are used. This is a costly procedure, but the long life of masks in a projection printer and the reduction of yield loss makes it worthwhile.

There are serious optical problems related to the photomasking process. The lenses used for reduction of original drawings or for projection printing are specially designed for this purpose and are very expensive. Such lenses must be capable of forming an image without distortion and with excellent edge definition. In the final images of the photomask, the dimensions are so small that accurate measurement is difficult. Lines on these masks are commonly 3 µm wide (only 6 wavelengths of green light), and sometimes only 1 µm. They do not appear well defined in measuring microscopes because of their size is so close to the dimensions of the light waves used to observe them. The problem of verifying that each of the thousands of lines and spaces on a photomask is of the proper size and in the right place is staggering.

4.7 Diffusion

Following the creation of open areas in the oxide coating, and the removal of the photoresist, slices can be diffused with a dopant. This is a batch process, in which a number of slices are inserted in a furnace at a precise temperature, usually between 900 and 1200°C, in a closely controlled atmosphere. The atmosphere contains a quantity of the dopant to be diffused into the silicon surface to dope the regions that are free from oxide. The concentration of the dopant in the silicon is related directly to its concentration in the atmosphere, and the depth to which it penetrates is related to the square root of the time that the slices are exposed to the atmosphere in the furnace. Diffusion times may range from a few seconds to many hours depending on the depth of diffused layer to be produced. High purity sources of diffusing atoms may be either gaseous, liquid, or solid.

4.8 Ion Implantation

Ion implantation permits the placement of dopant atoms into the slice at room temperature, rather than at the high temperatures required in diffusion. In ion implantation, a high-energy stream of dopant atoms is directed toward the silicon slice in a high vacuum. They penetrate into the surface of the target slice to a depth depending upon their kinetic energy. The number of atoms implanted in a given volume of the target slice can be very precisely controlled, as can the depth of penetration. Furthermore, distributions of dopant atoms not possible to achieve by other means can be produced by ion implantation. This technique has become a production process only during the last decade, but it is now routinely used in the manufacture of high performance integrated circuits. For example, the circuits for pocket calculators require the use of ion implantation to permit their operation at low power levels and thus prolong battery life.

For most products it is necessary to repeat the preceding steps of oxidation, photomasking, and diffusion or implantation a number of times to create the desired structure. Simple transistors require two or more diffusion steps plus a masking operation to expose the surface regions where electrical contact is required. After the metal connecting layer is deposited as described below, a further photomasking and etching operation permits delineation of the desired pattern of metal.

4.9 Passivation

Frequently, it is necessary to provide an inorganic coating on the semiconductor slice to protect it from exterior influences. Such layers are called passivation layers, and they range from a 1-µm thick layer of silicon dioxide to relatively thick deposits of special glasses. The objective is to render the semiconductor chip inert to attack by external chemical species. In former times, this protection was provided by a hermetic package. Today, however, large numbers of devices are packaged in less expensive nonhermetic assemblies. It is necessary to arrange for the protection of the chip in spite of the nonhermeticity of the package. One can regard passivation as a technique for shrinking the hermetic package down directly onto the surface of the semiconductor chip. Particularly in power devices, passivation is a necessary step to provide the low surface leakage necessary for high-voltage junctions. Many power device processes leave the intersection of the *p-n* junction and the surface of the device unprotected by any material. Considering the fact that the dielectric breakdown strength within a *p-n* junction may be as high as 200,000 volts per centimeter, and that most gases, including air, break down at around 2000 volts per centimeter, the need for passivation of the edge of the *p-n* junction becomes obvious. Regardless of the kind of device, a *p-n* junction exposed to the outside world will not have stable reverse-biased electrical characteristics. In fact, many *p-n* junctions that are protected by a thin silicon dioxide layer are still not stable over long periods of time due to diffusion of mobile ions through the silicon dioxide overcoat. Passivation of the device requires the isolation of the junction from these influences by the addition of inert surface layers. For low-power devices, silicon nitride and aluminum oxide have been used with considerable success in protecting oxide-coated junctions from diffusion of external species. Higher voltage devices generally employ thicker coatings of glass applied in a variety of ways.

4.10 Metallization

Silicon is not a metal. It is not possible to make electrical connection to silicon in the ways commonly used to make connections to metals. Yet external connections are required, and are usually provided by the evaporation in vacuum of thin metal films onto selected areas of the surface of the semiconductor chip. In discrete devices, these provide metallic surfaces to which further metal members may be attached. In integrated circuits, these evaporated layers also provide the interconnections between regions of the surface that need to be connected together electrically. In many cases these interconnection runs never are connected to the outside world at all. Various metals and multimetal systems are used for these purposes. Nearly every metal in the periodic table has been used in a production semiconductor device for making electrical connection to the silicon. However, the most common are aluminum, gold, and multilayer systems involving platinum, titanium, chromium, nickel, silver, or copper.

4.11 Testing

Once the processing of the slice has reached a point at which metal contact regions are available, the devices can be electrically tested to some extent. It is not possible to carry out the full range of electrical tests, because the device is not yet mounted in a package that permits removal of substantial amounts of heat. However, low-power characteristics can be measured on discrete devices and the functioning of most integrated circuits can be checked. Testing at this stage is one of the most valuable tools in controlling cost, since defective parts can be weeded out and discarded without incurring further cost in their processing. At this point, a slice of silicon will typically contain a very large number of apparently identical semiconductor devices in it. The testing apparatus has an array of precisely positioned, needle-like probes, which are arranged to descend on the contact pads that make electrical connections to each individual device in turn. In most cases in the interests of speed and accuracy the testing regime is conducted by computer-controlled equipment. When a defective device is detected, the usual practice is to mark it with a dot of identifying ink. This permits that particular device to be discarded after the slice has been separated into individual parts.

4.12 Dicing

The separation of individual dice* from the slice may be accomplished by either scribing the surface of the wafer with a diamond-pointed tool, or by engraving the surface in the same pattern by the use of a high-powered laser beam, and subsequently cracking the slice along these marks. Both operations are referred to as scribing, the diamond-point-scribing process being the older. Laser scribing provides improved control of the breaking process that follows. Alternately, the dice may be cut apart with a thin diamond-edged saw blade. In many cases, this scribing operation is performed at the plant location in which previous processing has been carried out, but the wafers are not broken until they are delivered to the plant where the dice are to be assembled into completely packaged devices, which may be halfway around the world. In this way the operations requiring the highest degrees of skill and equipment investment are confined to the processing plant, while operations requiring lesser skills are at the assembly location.

4.13 Assembly

The regularity of the pattern of dice in each slice is valuable to the assembler of devices. If the slice is broken and the dice become displaced from their initial positions, this regularity is lost. At assembly, each dice must be identified, oriented and put in the proper position for further assembly. This costs money. If the position and orientation of each die can be maintained, cost can be saved. There is a variety of techniques in use for mounting a scribed wafer on a thin plastic membrane and then stretching the membrane uniformly in two directions. The individual dice follow the stretched membrane and retain their relative position and orientation. During subsequent assembly operations, the devices can be picked from the membrane and positioned with a minimum of additional effort and cost.

The assembly of transistors and integrated circuits into finished devices generally involves mounting each individual die into a package. Often a metallic bond is created, using a layer of gold between the silicon chip and the metal surface of the package. The package and chip are raised in temperature until the gold fuses with the silicon and forms a strong bond. A more recently developed technique uses electrically conducting epoxy cements.

---

*Dice, chips, and pellets are the same. These terms refer to a segment of the wafer which contains usually one device or integrated circuit and which will be assembled into a package.

In devices that must dissipate a significant amount of power, other techniques are employed. Generally, the part to which the device is to be attached is made of copper whose high thermal conductivity is important in extracting heat from the semiconductor chip. Bonding of this type of semiconductor device is done using a variety of metallurgical methods, but rarely an epoxy or other material with poor thermal conductivity.

In very large devices, the silicon may be bonded between two pieces of a metal such as tungsten. Tungsten has a thermal expansion coefficient similar to that of silicon and has good thermal and electrical conductivity. Its high elastic modulus provides a rigid backup for the silicon, protecting the silicon from external mechanical forces. In hockey-puck shaped packages, this tungsten-silicon-tungsten sandwich is assembled between copper end plates, but not bonded to them. The flexible diaphragm-like end plates of the package then permit the copper surfaces to be clamped by external pressure to the silicon-tungsten sandwich to provide good electrical and thermal paths to the device. The absence of a solid metallic bond permits differential expansion to take place between the tungsten and the copper, which have differing thermal expansion coefficients, without the risk of deforming a metallic joint and consequent fatigue failure after many heating and cooling cycles.

In low-power devices, once the die is mounted in the package, it is necessary to connect the individual metal contact patterns to the external connections of the package. The usual approach employs fine gold wire and the thermocompression bond technique. In this procedure, the end of the gold wire is melted into a spherical droplet, which is then solidified before being pressed against a metallic area on the silicon chip. A combination of temperature not as high as the melting point and high physical pressure serves to weld the wire ball to the metallization on the surface of the silicon. The wire is then led from this point to the external package connection where it is welded in a similar fashion. These wires range in diameter from 50 to 150 μm.

Once the electrical connections are complete, it is necessary to enclose the device so as to protect it mechanically from the shocks of the external environment. Silicon is brittle and will not tolerate physical abuse. Therefore, the package in which it is contained must protect it. In many cases the packages are molded of plastic materials, while in other cases the package is a metal and ceramic assembly. Semiconductor dice are also sold without any packaging at all. These are used by sophisticated customers who mount and connect these devices to other circuits of their own manufacture. With reasonable care it is possible to test and ship unpackaged dice to another location in large quantities without damage.

4.14 Final Testing

Final electrical testing of semiconductor devices is done mostly by computer-controlled equipment. The number and complexity of tests are large, and the precision with which they must be conducted is high. Furthermore, given the large amount of data obtained during a test sequence on a single integrated circuit and the complex logic that governs decisions regarding the fate of the device under test, the manufacturer quickly decides to use a computer for control of the test sequence.

In most cases it is not possible for a semiconductor process line to deliver a part that is characterized by a single type number. The number of serial steps in the process and the inevitable variability of the product mean that the output distribution from a product line will be wide enough that it must be subclassified into different parts. This sorting may take place on the basis of electrical characteristics, response of the device to temperature extremes, or other factors. Devices with a high degree of uniformity and that are operable over wide temperature ranges command premium prices. Other devices, perhaps made on the identical line at the same time, which have less desirable characteristics but which still function adequately, are sold for lesser prices. Additional sampling tests are done to explore the performance of the device in high stress environments or over long periods of time, to assure stability and long life. The testing costs incurred in qualifying devices for very high reliability service, such as in cardiac pacemakers or in satellites, may exceed the cost of manufacturing the device many times over.

4.15 Process Control

The subject of process control deserves separate discussion. The reader can appreciate from the foregoing description that semiconductor device manufacture is a long and delicate chain of events with more opportunity for failure than for success. Disaster in the form of catastrophically low yield is always lurking nearby, and the process engineer is the person who must control the many variables and keep the factory running.

Each of the processes described above produces a change in the silicon wafer, be it a pattern in photoresist, a diffused layer, or a deposit of contact metal. The majority of such changes are irreversible. If the process has been wrongly done, expensive scrap has been created.

The key to control is measurement. Too often the variables one needs to measure are in such tiny areas that they are not accessible to measurement, or the measurement is destructive. On the other hand, if the measurement is not done and an error in processing has been made, the money spent in further work on the material is wasted.

One way to deal with this problem is to provide on each silicon wafer a small area devoted to special test patterns for process control purposes. These patterns range from simple rectangular regions for measurement of sheet resistance of diffused layers, or dot patterns for checking alignment of photoresist processes, to transistor structures typical of those buried deep within an integrated circuit. Measurements routinely made on these structures can tell the process engineer much about the day-to-day reproducibility of processes and can help him pinpoint the troublesome process when disaster strikes. This information more than adequately compensates for the small area of each wafer used to carry the test structures and thus not available for salable devices.

Semiconductor manufacturing processes are by no means thoroughly understood, largely because of the lack of adequate measurement methods. Most process steps involve high purity materials, thin layers, small dimensions, low concentrations, or other extremes of conditions which make measurements difficult at best. What is needed are simple, reproducible, nondestructive measurement techniques that can be done routinely in a factory environment by technically unsophisticated people. There is a large gap between the need and what is possible at present. The cost of this unsatisfactory measurement and control situation is yield loss, which in total has been estimated at about $2 billion annually for the U.S. semiconductor industry. Not all of this loss can be avoided solely by improved measurement--people will always make mistakes--but a significant amount certainly could be.

5. PRODUCT DEVELOPMENT

Successful development of a new semiconductor product requires a confluence of several factors. First, there must be a need established for the product. If there is no need, there is no market, and the development is an exercise. While this seems obvious, it has not always been recognized. The need implies a set of performance specifications for the new product if it is to be successfully applied.

Given a set of characteristics, a device structure can be designed to have at least a first approximation of these characteristics. While designs for discrete devices can frequently be done without recourse to computer systems, the complexity of today's integrated circuits and the very large number of individual devices that go together to make up a single integrated circuit require the use of computer assistance. Modern integrated circuits may contain more than 100,000 individual transistors in a single device; the number of devices overtaxes human abilities to position and interconnect them correctly. Since most of these devices are exact repetitions of other devices within the same integrated circuit, differing only in location and perhaps orientation, computer aid is natural. In addition, since the final set of masks to produce the product requires registration of many levels of extreme complexity, the ability of the computer to remember what exists in the other levels aids immeasurably in creating the proper form for the one being worked on at the moment. Thus registration errors between masks can be reduced.

In addition, the particular characteristics of the processes in use by each manufacturer are different. These properties lead to requirements, called design rules, which apply in the preparation of the photomasks, and which are different for different manufacturers--even for different process sequences within the same manufacturer's house. In general, processes in use by different firms, even to produce devices with essentially identical properties, are sufficiently different that the masks used by manufacturer A will not work in manufacturer B's environment.

When the initial design is complete, and the masks are available, processing of silicon can begin. Initial attempts to build the product are usually done in a development laboratory. The products of this effort are tested and the results compared with the initial specifications for the product. Any design corrections are made, and the fabrication process is repeated. Once the design to produce the new product is established, and the process is considered well enough controlled to give satisfactory yields, manufacture of the product in the factory can begin. The transfer from laboratory to factory is fraught with difficulties, because what was previously done by highly skilled people must now be done by relatively less skilled factory workers who do not in general understand the details of what they are doing. If the new product uses the same sequence of processes as other products already being manufactured, less difficulty is usually encountered. If, however, an entirely new process sequence is required, the road to successful manufacture of a new product is studded with obstacles.

Initial yields of complex integrated circuits are frequently below five percent. This is obviously not satisfactory. Through continuous refinement these yields improve, leading to a final product cost that decreases sharply with time. The semiconductor industry's record of product cost reduction is probably unmatched.

Semiconductor manufacturers have always provided substantial technical assistance to their customers in the use of their products. In the beginning, when semiconductors were replacing vacuum tubes in old applications, the sharply different characteristics of semiconductors created considerable difficulty in application. As a part of marketing operations, engineering assistance was extended to customers to help them to apply semiconductors correctly. With the proliferation of types of semiconductor products, this application assistance has continued to be required.

6. PRODUCT TYPES

In the following paragraphs each of the major kinds of semiconductor devices will be described briefly.

6.1 Diodes

Signal diodes are elementary one-junction small rectifiers, by definition limited to forward average currents of 100 millamperes or less. Their reverse voltage rating is generally 200 volts or less. Typically, these devices are packaged in small glass envelopes with axial leads. They are mature products; that is, their design and manufacturing processes are stable and their consumption is steady or decreasing with time. Many of the circuit functions formerly done by signal diodes are now performed by integrated circuits.

An important class of signal diodes is made using a slightly different junction structure than has been described up until now. If a layer of a properly selected metal is evaporated onto a silicon slice, it is found that the metal-silicon interface possesses rectifying characteristics. This is termed a Schottky barrier. Devices utilizing this effect are called Schottky diodes; they have a very fast recovery time from the forward-conducting to the reverse-blocking mode, a feature that is important in some applications. This type of signal diode will probably increase its utility with time over the next few years, because of the value of the fast recovery characteristic in microwave applications. Schottky barrier devices are made in larger sizes as well, and a small part of the rectifier market described below is supplied this way. Because of the limited reverse-blocking voltage of Schottky devices, however, the impact of this technique in the rectifier business is substantially less than it is for signal diodes.

Rectifiers are defined as simple *p-n* junction devices having two leads and carrying currents greater than 100 millamperes. They can be very large indeed; rectifiers rated in the thousands of amperes are made and are packaged in enclosures 8 cm (3 inches) or more in diameter. Devices for these applications may also have very high reverse blocking voltages, ranging up to 3,500 volts.

Still another two-terminal single-junction device is the Zener, or voltage reference, diode. These devices are designed and made to be used in the reverse-biased direction, with the voltage breakdown characteristic of the blocking junction carefully controlled. By passing a small current through the device in the reverse direction, beyond the normal reverse voltage rating of an ordinary device, a very stable reference voltage drop occurs. These devices are used widely to establish fixed values of voltage in power supplies and other specialized applications.

Another specialized rectifier is the variable capacitance or varactor diode. When a *p-n* junction is biased in the reverse direction, the free charge carriers are swept away from a region in the neighborhood of the *p-n* junction. This produces a layer free of any mobile charge, and in every respect this layer functions as a dielectric. A dielectric bounded on either side by a conducting medium forms a capacitor. Capacitors made in this fashion, in silicon, have the valuable feature that their capacitance can be controlled electrically. By changing the magnitude of the applied reverse voltage, the width of the blocking layer is changed and the capacitance also changes. By controlling the distribution of the dopant atoms on either side of the *p-n* junction, a variety of voltage-capacitance characteristics can be designed. These devices have increasing applications in electronically tuned radio and television receiving equipment.

Yet another specialized class of *p-n* junction diode is used for detection of light. If light is allowed to be absorbed in silicon, pairs of holes and electrons are created. If this occurs in the neighborhood of a reverse biased *p-n* junction, these holes and electrons form a current much larger than the small current that would pass through that junction in the dark. In connection with yet another specialized *p-n* junction device, the light-emitting diode, devices can be made for sensing holes in computer cards and for communicating signals between circuits at different potentials, for example.

Light-emitting diodes are not made of silicon, but of gallium-arsenic compounds. These compounds have the property that when holes and electrons recombine in a forward-biased junction, the energy that they represent appears in the form of light with a modest but useful efficiency. The wavelength of this light from pure gallium arsenide is in the near infrared and therefore invisible to the human eye. If part of the arsenic is replaced by phosphorus, however, diodes made of gallium arsenide-phosphide emit light which is bright red, the familiar color of the digits on many pocket calculators and electronic watches. This same red light is readily absorbed by silicon. If, therefore, a light-emitting diode and a photodiode are arranged in proximity, current passing through the light-emitting diode in the forward direction will be detected by the photodiode as a reverse current increase. Such devices are called opto-electronic couplers. Not all opto-electronic couplers use photodiodes. Many instead use a photosensitive transistor that provides greater sensitivity to the light signals.

Other specialized diodes are used as oscillators at ultrahigh frequencies. This class of very high frequency signal sources includes more than one type of *p-n* junction diode; tunnel diodes, IMPATT diodes, and Gunn effect diodes are all in this category. They all work in different ways, but the end result is similar. All generate ac signals of extremely high frequency that are useful in communications equipment in the microwave regions.

6.2 Transistors

All the devices heretofore discussed have two external leads. We will now consider three-leaded devices--transistors, of which there are many specialized types. Signal transistors, in which the power dissipation is 1 watt or less, are used in great variety but in decreasing numbers. The signal transistor is being displaced slowly by integrated circuits.

Larger transistors, capable of dissipating more than 1 watt of power, are classified as power transistors. They range to over 100 amperes in current-carrying capability, or up to nearly 2000 volts in voltage breakdown rating. All transistors share the common characteristic that the output current is a predictable and generally proportional function of the input current.

Included in signal and power transistor categories are two specialized devices used at frequencies above 300 megahertz. They are specialized in the sense that they are made with extremely small elements. For example, a microwave signal transistor to operate at 3 gigahertz has lines in its metallization pattern of 1 micrometer wide, spaced 1 micrometer apart. This very fine geometry puts extreme requirements on the photoresist processing techniques used to make these devices. Similarly, microwave power transistors also use a very large number of tiny elements in their construction to permit them to operate efficiently at very high frequencies. These devices comprise a rapidly developing segment of the semiconductor business; they are being used in high-frequency radio transmission and in cable television systems.

6.3 Thyristors

A device that has extremely important applications in industrial and power control equipment is the thyristor. It is sometimes referred to as the silicon-controlled rectifier, or SCR, and is used as a switching device. Its useful properties stem from the fact that when it is in its nonconducting state only a very small current exists, and when it is in the conducting state large currents can be passed with a small voltage drop. Thus a relatively small device can switch quite large amounts of power. Switching is initiated by means of current into a control terminal, called the gate. The thyristor is normally in the off-condition, and when a certain amount of current is conducted into the gate, it switches into the on-condition. The gate signal needed is only a small impulse, because once switching has occurred the gate loses control. The thyristor then remains in the on-state until the current through it is reduced to nearly zero by the external circuit. At this time the blocking state is recovered, and the device remains in the blocking state until another gate signal is received. Thyristors are available with current ratings as high as 2500 amperes and with voltage ratings as high as 5000 volts.

6.4 Integrated Circuits

An integrated circuit is nothing more than a collection of individual transistors, diodes, and resistors, all made extremely small, arrayed in a single piece of silicon and interconnected by a pattern of metallization on the surface of the die. The arrangement is made to perform a specific circuit function that may be quite simple, as in the case of a gate circuit for a computer, or that may be extremely complex and embody a large fraction of the circuitry for an entire television receiver. From the smallest to the largest integrated circuit, however, the building blocks are still the same.

The earliest types of integrated circuits were used in computers. Integrated circuits are particularly well suited to computer applications because computers are made up of large numbers of a few kinds of circuits; there may be tens of thousands of gate circuits in a computer. The earliest application of integrated circuits was simply to replace a gate circuit made of discrete parts with a gate circuit made in a single piece of silicon. Immediate economies of scale existed in this application, and integrated circuits have invaded the computer domain to the greatest degree.

Since computer circuits are generally arranged to operate in either an on-condition or an off-condition (binary circuitry), ICs used in this manner are referred to as digital circuits. Their behavior is defined only in the on or off state. Digital ICs are divided into families having certain characteristics in common. First, the input and output voltage swings of all devices of a family fall within the same ranges. Second, the power supply requirements are similar. Third, the relative speeds of all devices in the family are similar. These characteristics are achieved by using similar circuitry for all members of the family. Families are named according to the type of basic circuit element that is used. There are, among the obsolete families of digital circuits, such types as diode-transistor logic (DTL), resistor-transistor logic (RTL), and so on. These families have been replaced by families that are either faster, use less power, or are less expensive. Now the important bipolar digital families of ICs are transistor-transistor logic (TTL), emitter-coupled logic (ECL), and integrated-injection logic (I^2L). Within families there are varieties that differ in speed or power consumption or both. But, for example, all TTL circuits can work together, sharing common power supply voltages and common input-output voltage levels. Interconnection between families, however, is often more difficult, although it is done when necessary. The present bipolar digital families of interest mentioned above are expected to remain the significant members of this class of device for several years to come.

Digital circuits are further subdivided into bipolar integrated circuits, using transistors of the types discussed previously, or MOS (metal-oxide-silicon) circuits. MOS transistors are a special class of transistor in which an electric field outside the silicon controls the current parallel to and beneath the surface of a semiconductor chip. The electric field is created between a metal layer and the semiconductor, with a thin silicon dioxide dielectric interposed between (the metal-oxide-silicon structure). MOS digital circuits are distinguished by the fact that their power requirements are significantly lower than bipolar digital circuits, although in general their speed is also less. In many applications this is an acceptable compromise and has led to development of a variety of MOS device families. There are *p*-channel MOS (*p*MOS), *n*-channel MOS (*n*MOS), complementary MOS (CMOS), silicon-on-sapphire (SOS), and charge-coupled devices (CCD).

The MOS digital circuits were developed in the order given above, *p*MOS being oldest and CCD being newest. With the exception of CCDs, the speed also has increased with time through the various families from *p*MOS to SOS. The speed increase has not always been at the price of increased power consumption, however; CMOS has extremely low power consumption. As an additional feature, CCD devices are also photosensitive and can be used to make useful image-forming devices for use in TV cameras.

The low power consumption of MOS ICs has made possible the pocket calculator and the electronic watch, both of which must be sparing of power to give satisfactory battery life. More subtly, low power is also important in many other applications. Since the use of ICs permits great circuit complexity to be put in a very small volume, even the modest power dissipation of ICs becomes significant when many such devices are put together into one equipment enclosure.

In contrast with digital ICs in which the transistors and circuits are designed to operate in either one or the other of two states with rapid transitions between, linear integrated circuits are designed to have a predictable functional relationship between input and output. For instance, in an IC amplifier, the output is a magnified version of the input. The total gain through such a circuit may be 1 million or more. Design and manufacture are very much more difficult for linear integrated circuits than for digital circuits, a fact which has somewhat delayed their growth. Furthermore, there is a much larger variety of linear circuits than there is of digital circuits, and each variety of circuit requires a separate design. The cost of designing a complex integrated circuit may be as much as $500,000 so the task is not to be undertaken lightly. A substantial number of applications must exist beforehand. It should be noted that this cost covers only the design; if a new process sequence, not already in use in the factory, must be developed, the cost can easily be five times this amount. If the circuit is a microprocessor or microcomputer, the cost of developing the associated software will also run into the millions of dollars.

Six to eight linear integrated circuits can perform all of the functions required inside a television set with few exceptions, and such circuits now exist. Linear integrated circuits are also available to make AM radios, FM radios, low-power stereo amplifiers, etc. These applications share the common factors of a large market and relatively standardized circuitry.

In addition, one of the oldest types of linear integrated circuits, the operational amplifier (op amp), was developed for use in analog computers. Uses of op amps in other applications have now far overshadowed the applications in analog computers, a tribute to the ingenuity of users who, once a microelectronic device is available to do a particular job, are amazingly clever in finding other uses for the function. Many of these uses could not have been met at reasonable cost before.

6.5 Application Trends and Their Impacts

Cost and complexity trends in the industry are interesting to consider. In ICs, the average cost per component in 1961 was $1. By 1972, this figure was $0.01, and in 1975 it was $0.001. The trend for discrete devices is by no means as steep, but it is strongly downward nonetheless. This rapid decrease in cost with the passage of time has made today's minicomputers and pocket calculators economically possible.

On the other hand, complexity has grown. Where in 1959 one device per chip was the norm, by 1973 over 10,000 components per chip were possible. This increase rate corresponds closely to doubling each year. It continued at this rate for 16 years (a 65,000-fold increase by August 1975) but has slowed somewhat since that time. In 1980, development chips having about a half million components have been described. This growth has been made possible by reduction in size per component and increased chip size, together accounting for a factor of about 2500, and by ingenuity of device and circuit design that has contributed the remaining factor of 200.

The net effect of these trends has been to reduce substantially the cost of performing electronic functions. This change appears in products in a variety of ways:

1. Transistor radios, pocket calculators, and electronic watches have each, on their own time scale, been available at rapidly falling prices limited ultimately by the material and labor costs of the whole product. The electronics part of this total cost is relatively modest.

2. Television sets, major appliances, and automobiles tend to incorporate more and more sophistication at a relatively constant price. Communication and industrial control equipment also shares this trend. The electronics content of these tends to increase with time, to provide the added sophistication.

3. The size of the largest practical electronics system keeps increasing. This is not only for the reason that decreasing costs per component permit more components in an affordable system, but also because the failure rate per component in an IC has dropped at a rate almost identical to the rate of increase in the number of components per IC. The failure rate per IC package has not significantly changed. Practical system size tends to be limited more by the tolerable failure rate of the system.

In power semiconductors, the trends are different in kind but similarly broadening in their impact on our daily life. Applications of power devices are most apparent in variable lighting controls, speed controls on power tools and appliances, and in other major household appliance control functions. Less visible to the consumer but of more economic importance are industrial controls that are at the heart of many machine tools and that operate many industrial processes. The economic effect stems from the improved productivity per man-hour that is provided by these sophisticated kinds of plants and machinery.

The growing concern for our energy sources has caused the development of many electric power-generating schemes, such as solar cells and windmills, whose output is either direct current or perhaps alternating current of varying frequency. The output from these sources requires modification to be compatible with the existing constant-frequency ac power distribution system, and solid-state power-converting equipment will be necessary.

A final example is the use of power semiconductor devices as the "muscle-power" to operate mechanical systems under the primary control of electronic logic circuits. Such systems are already in use in antiskid controls on large vehicles and are being considered for a variety of other vehicular subsystems. The need for power solid-state devices to do the work, under control of logic circuits (computers) that make decisions, is not unique to the automotive field, of course, but the example illustrates the concept.

7. SUMMARY

We have explored briefly the wide range of technologies used in the semiconductor industry to transform simple, pure crystals of a few kinds into a constantly growing array of solid-state devices and circuits. These have also been described. The manufacturing process is materials- and process-oriented, requiring a wide variety of technical professionals for its conception, development, and implementation. It is an industry of extremes: highest purity, smallest linewidths, most complex products, simple concepts, and bewildering processes. In less than 35 years since the transistor was invented, a nearly $15 billion annual flood of semiconductors worldwide has radically changed the electronics industry and our daily lives. Semiconductor devices are at the heart of nearly every electronic product and are the indispensible components in almost $80 billion worth of goods in 1980 in the United States alone.

Semiconductor manufacturing operations are limited at any particular time by materials, processing, and measurement problems. As today's problems are solved and new progress is made, new problems arise and impose new limits. The past performance of the industry in controlling processes, improving yields, and raising reliability levels has resulted in dramatic reductions in cost. In turn, lowered cost has allowed penetration of electronic products into new areas. Fundamental physical limits will ultimately determine how complex an integrated circuit can become, for example, but there is still a long way to go.

NBS-114A (REV. 2-80)

U.S. DEPT. OF COMM.
BIBLIOGRAPHIC DATA SHEET *(See instructions)*

1. PUBLICATION OR REPORT NO.
NBSIR 81-2197

2. Performing Organ. Report No.

3. Publication Date
January 1981

4. TITLE AND SUBTITLE
Semiconductor Technology for the Non-Technologist

5. AUTHOR(S)
Robert I. Scace

6. PERFORMING ORGANIZATION *(If joint or other than NBS, see instructions)*

NATIONAL BUREAU OF STANDARDS
DEPARTMENT OF COMMERCE
WASHINGTON, D.C. 20234

7. Contract/Grant No.
N/A

8. Type of Report & Period Covered
N/A

9. SPONSORING ORGANIZATION NAME AND COMPLETE ADDRESS *(Street, City, State, ZIP)*

10. SUPPLEMENTARY NOTES

☐ Document describes a computer program; SF-185, FIPS Software Summary, is attached.

11. ABSTRACT *(A 200-word or less factual summary of most significant information. If document includes a significant bibliography or literature survey, mention it here)*

The properties of semiconductor materials, the methods of processing them, and the solid-state products made from them are described in terms intended to be understandable by the lay person. The semiconductor industry has grown at a rate of 21 percent per year compounded for the last twenty years, and its products have declined in unit cost by a factor of five in current dollars (a factor of ten in constant dollars) in the same period. This very satisfactory but anomalous behavior has attracted the interest of many who are not familiar with the technology of the industry, yet who need to have some understanding of it. This report is intended to help meet that need.

12. KEY WORDS *(Six to twelve entries; alphabetical order; capitalize only proper names; and separate key words by semicolons)*

Integrated circuits; semiconductors; semiconductor devices; semiconductor processes; semiconductor technology; silicon.

13. AVAILABILITY

☒ Unlimited
☐ For Official Distribution. Do Not Release to NTIS
☐ Order From Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402.
☒ Order From National Technical Information Service (NTIS), Springfield, VA. 22161

14. NO. OF PRINTED PAGES
28

15. Price
$6.50