Micromanufacturing and Fabrication of Microelectronic Devices

PART V

The importance of the topics covered in the following two chapters can best be appreciated by considering the manufacture of a simple metal spur gear. It is important to recall that some gears are designed to transmit *power*, such as those in gear boxes, yet others transmit *motion*, such as in rack and pinion mechanisms in automobile steering systems. If the gear is, say, 100 mm (4 in.) in diameter, it can be produced by traditional methods, such as starting with a cast or forged blank, and machining and grinding it to its final shape and dimensions. A gear that is only 2 mm (0.080 in.) in diameter, on the other hand, can be difficult to produce by these methods. If sufficiently thin, the gear could, for example, be made from sheet metal, by fine blanking, chemical etching, or electroforming.

If the gear is only a *few micrometers* in size, it can be produced by such techniques as optical lithography, wet and dry chemical etching, and related processes. A gear that is only a *nanometer* in diameter would, however, be extremely difficult to produce; indeed, such a gear would have only a few tens of atoms across its surface.

The challenges faced in producing gears of increasingly smaller sizes is highly informative, and can be put into proper perspective by referring to the illustration of length scales shown in Fig. V.1. Conventional manufacturing processes, described in Chapters 11 through 27, typically produce parts that are larger than a millimeter or so, and can be described as visible to the naked eye. The sizes of such parts generally are referred to as **macroscale**, the word "macro" being derived from the Greek *makros*, meaning "long," and the processing of such parts is known as **macromanufacturing**. Macroscale is the most developed and best understood size range from a design and manufacturing standpoint, with a very wide variety of processes available for producing components of that size. All of the examples and case studies given thus far have been examples of macromanufacturing.

Note that the gear shown in Fig. V.1 is a few tens of micrometers across, and thus fits into the category of **micromanufacturing**, developed mostly for electronic devices of all types, including computer processors and memory chips, sensors, and magnetic storage devices. For the most part, this type of manufacturing relies heavily on lithography, wet and dry etching, and coating techniques. Examples of products that rely on micromanufacturing techniques are a wide variety of sensors and probes (see Fig. V.2), ink-jet printing heads, microactuators, magnetic hard-drive heads, and such devices as computer processors and memory chips. Microscale mechanical devices are still a relatively new technology, but one that has been developing rapidly.

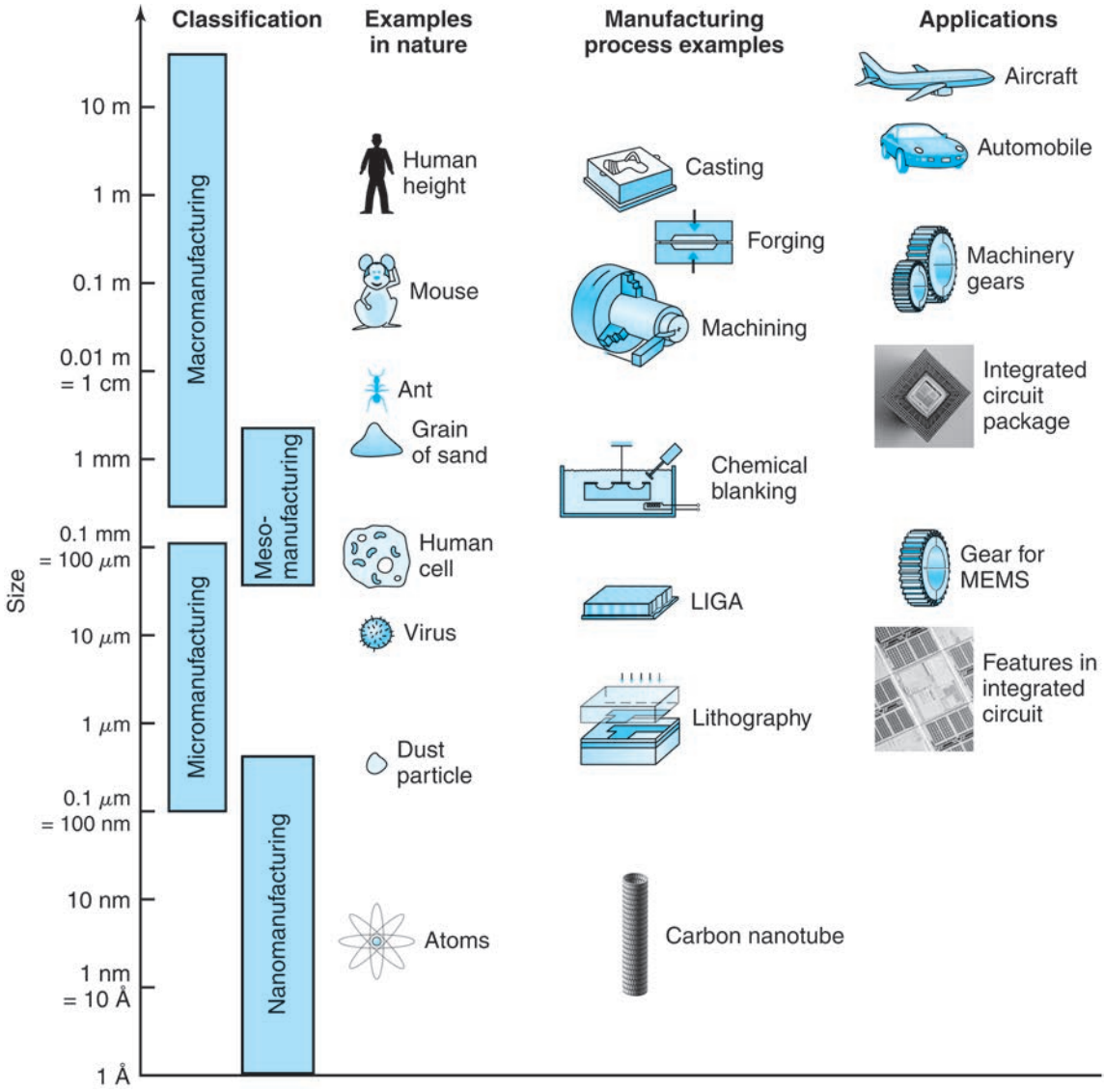


FIGURE V.1 Illustration of the regimes of macro-, meso-, micro-, and nanomanufacturing, the range of common sizes of parts, and the capabilities of manufacturing processes in producing these parts.

Mesomanufacturing overlaps macro- and micromanufacturing, as seen by the illustrations given in Fig. V.1. Examples of mesomanufacturing are extremely small motors, bearings, and components for miniature devices, such as hearing aids, stents, heart valves, and electronic toys, with components the same as the gear shown in Fig. V.1.

In **nanomanufacturing**, parts are produced at scales of one billionth of a meter, typically between 10^{-6} and 10^{-9} m in length; many of the features in integrated circuits are at this length scale. *Biomanufacturing*, covering such areas as molecularly engineered pharmaceutical products, genetic testing, gene therapy, and agricultural products, are at nanoscale level, and it is now recognized that many physical and

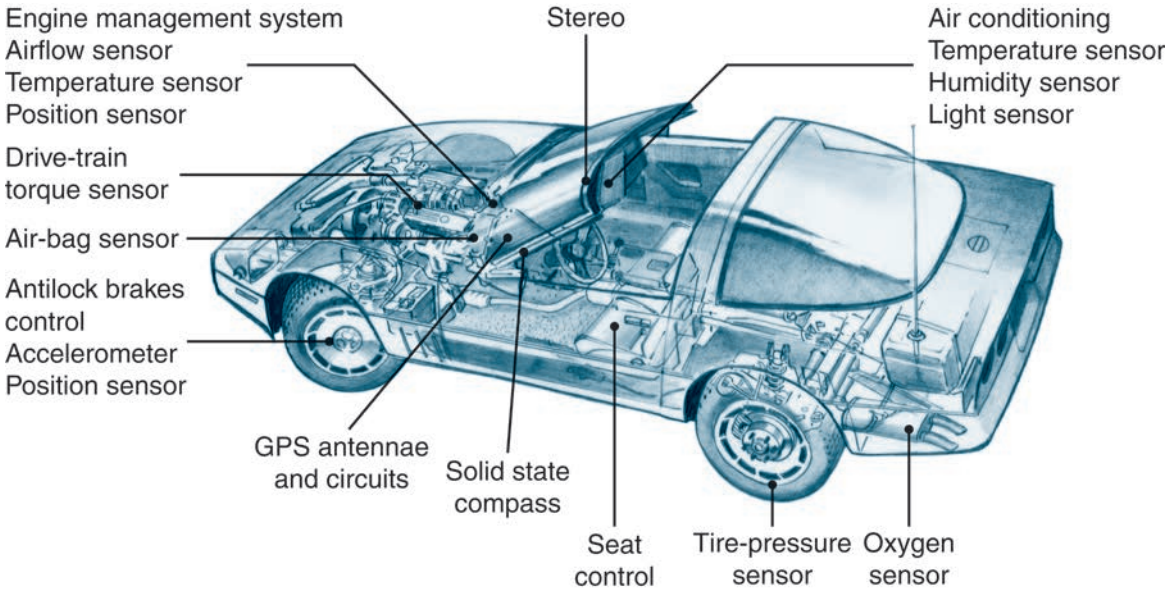


FIGURE V.2 Microelectronic and microelectromechanical devices and parts used in a typical automobile.

biological processes act at this scale and that nanomanufacturing holds much promise for future innovations.

In Chapter 28, the manufacture of silicon wafers and microelectronic devices is described, which include a wide variety of computer processors, memory devices, and integrated circuits. Controls, transportation, communications, engineering design and manufacturing, medicine, and entertainment all have been changed greatly by the wide availability of *metal-oxide-semiconductor* (MOS) *devices*, generally based on single-crystal silicon. Microelectronics is the best known and commercially important example of micromanufacturing, with some aspects of the applications exemplifying nanomanufacturing. The chapter also covers the techniques used in packaging and assembling integrated circuits onto printed circuit boards.

The production of microscale devices, that are mechanical and electrical in nature, is described in Chapter 29. Depending on their level of integration, these devices are called **micromechanical devices** or **microelectromechanical systems** (MEMS). While the historical origins of MEMS manufacture stem from the same processes used for microelectronic systems, and from identical processes and production sequences, still in use, several unique approaches also have been developed.

CHAPTER

28 Fabrication of Microelectronic Devices

- This chapter presents the science and technologies involved in the production of integrated circuits, a product that has fundamentally changed our society.
- The chapter begins by describing silicon, the preferred material for most integrated circuits, and its unique properties that make it attractive. Beginning with a cast ingot, the operations required to produce a wafer are described.
- The production of patterns on wafers is next discussed, including the processes of lithography, wet and dry etching, and doping.
- Metallization and testing are then described, as are the approaches for obtaining electrical connections from integrated circuits to circuit boards.
- The chapter concludes with a description of the different packages used for integrated circuits.

Typical parts produced: Computer processors, memory chips, printed circuit boards, and integrated circuits of all types.

28.1 Introduction

Although semiconducting materials have been used in electronics for a long time (the word semiconductor first appeared in 1838), it was the invention of the *transistor* in 1947 that set the stage for what would become one of the greatest technological achievements in all of history. **Microelectronics** has played an increasing role ever since the *integrated circuit* (IC) technology became the foundation for calculators, wrist watches, controls for home appliances and automobiles, information systems, telecommunications, robotics, space travel, weaponry, and personal computers.

The major advantages of today's ICs are their very small size and low cost. As their fabrication technology has become more advanced, the size and cost of such devices as transistors, diodes, resistors, and capacitors continue to decrease, and the global market has become highly competitive. More and more components can now be put onto a **chip**, a very small piece of semiconducting material on which the circuit is fabricated.

Typical chips produced today have sizes that are as small as 0.5 mm × 0.5 mm and, in rare cases, can be more than 50 mm × 50 mm, if not an entire wafer. New technologies now allow densities in the range of 10 million devices per chip (Fig. 28.1), a magnitude that has been termed **very large scale integration** (VLSI). Some of the advanced ICs may contain more than 100 million devices, termed **ultralarge-scale**

QR Code 28.1 Tour of an integrated circuit fabrication facility. (*Source:* Courtesy of Intel)

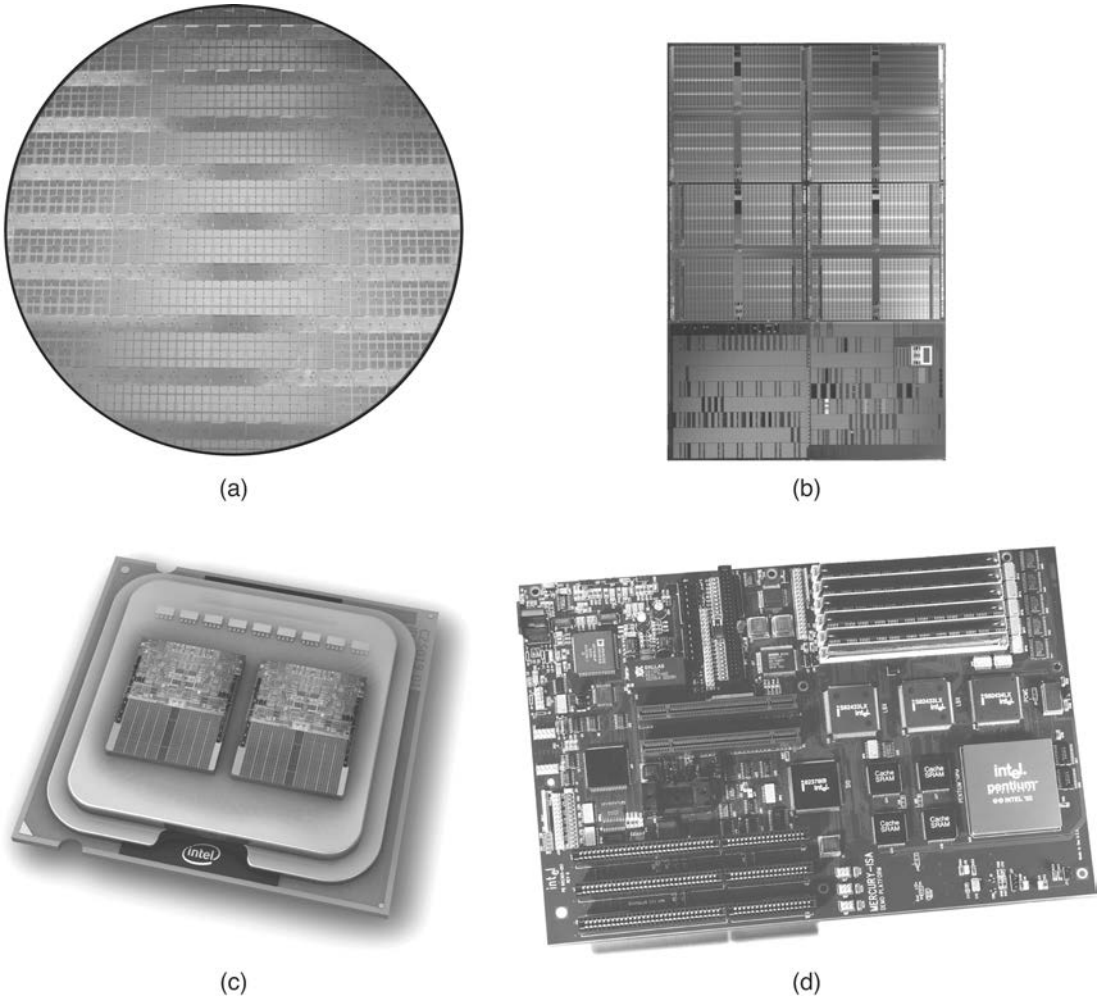


FIGURE 28.1 (a) A 300-mm (11.8-in.) wafer with a large number of dies fabricated onto its surface; (b) detail view of an Intel 45-nm chip, including a 153-Mbit SRAM (static random access memory) and logic test circuits; (c) image of the Intel® Itanium® 2 processor; and (d) Pentium® processor motherboard. *Source:* Courtesy of Intel Corporation.

integration (ULSI). The Intel® Itanium® processors, for example, have surpassed 2 billion transistors, and the Advanced Micro Devices Tahiti® graphic processing unit has surpassed 4.3 billion transistors.

Among more recent advances is **wafer-scale integration** (WSI), in which an entire silicon wafer is used to build a single device. This approach has been of greatest interest in the design of massively parallel supercomputers, including **three-dimensional integrated circuits** (3DICs), which use multiple layers of active circuits, maintaining connections both horizontally and vertically.

This chapter describes the processes that are currently in use in the fabrication of microelectronic devices and integrated circuits, following the basic sequence shown in Fig. 28.2. The major steps in fabricating a **metal-oxide-semiconductor field-effect transistor** (MOSFET), one of the dominant devices used in modern IC technology, are shown in Fig. 28.3.

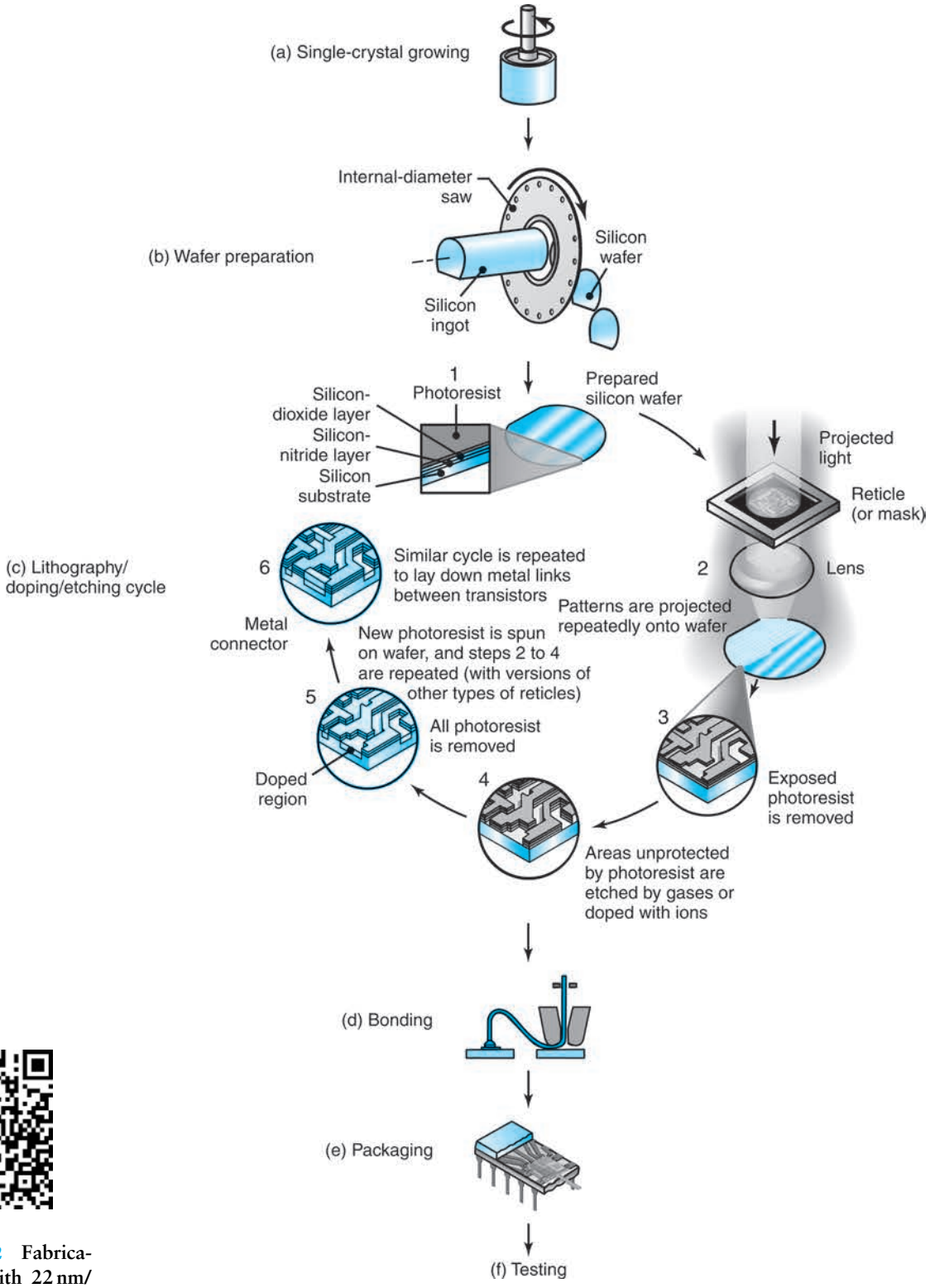


FIGURE 28.2 Outline of the general fabrication sequence for integrated circuits.

QR Code 28.2 Fabrication of a chip with 22 nm/3D transistors. (*Source:* Courtesy of Intel)

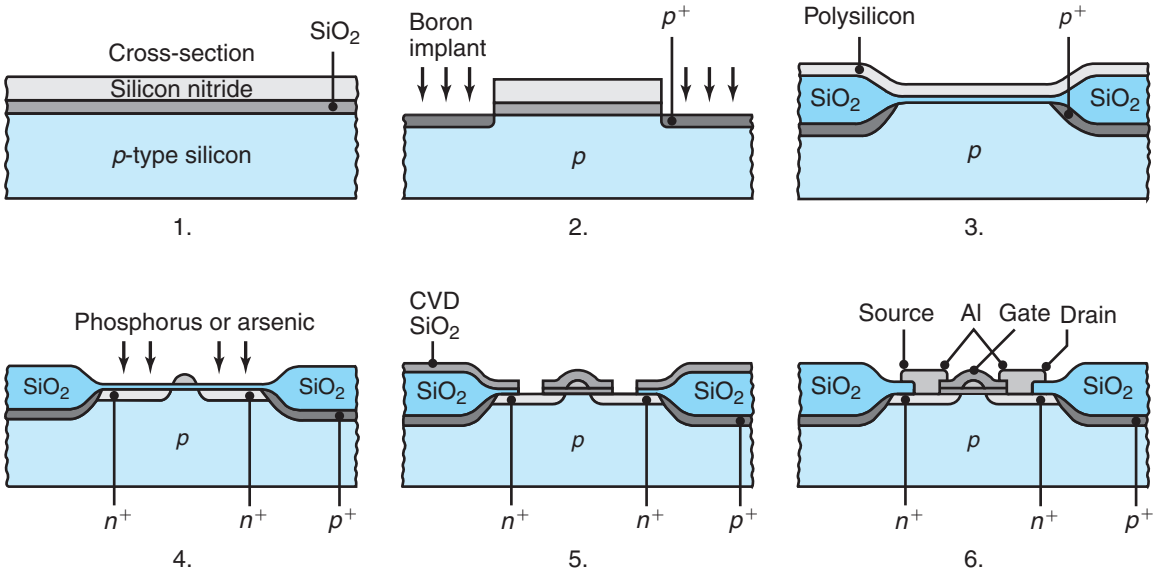


FIGURE 28.3 Cross-sectional views of the fabrication of a MOSFET transistor. *Source:* After R.C. Jaeger.

28.2 Clean Rooms

Clean rooms are essential for the production of integrated circuits, a fact that can be appreciated by noting the scale of manufacturing to be performed. Integrated circuits are typically a few millimeters in length, and the smallest features in a transistor on the circuit may be as small as a few tens of nanometers. This size range is smaller than particles that generally are not considered harmful, such as dust, smoke, and perfume; however, if these contaminants are present on the surface of a silicon wafer during its processing, they can seriously compromise the performance of the entire device.

There are various levels of clean rooms, defined by the **class** of the room. The size and the number of particles are significant in defining the class of a clean room, as shown in Fig. 28.4. The traditional classification system refers to the number of 0.5-μm or larger particles within a cubic foot of air. Thus, a Class-10 clean room has 10 or fewer such particles per cubic foot. This standard has been superseded by an ISO standard, but the traditional classification scheme is still widely used. Most clean rooms for microelectronics manufacturing range from Class 1 to Class 10; in comparison, the contamination level in modern hospitals is on the order of 10,000 particles per cubic foot.

To obtain controlled atmospheres that are free from particulate contamination, all ventilating air is passed through a *high-efficiency particulate air* (HEPA) filter. In addition, the air usually is

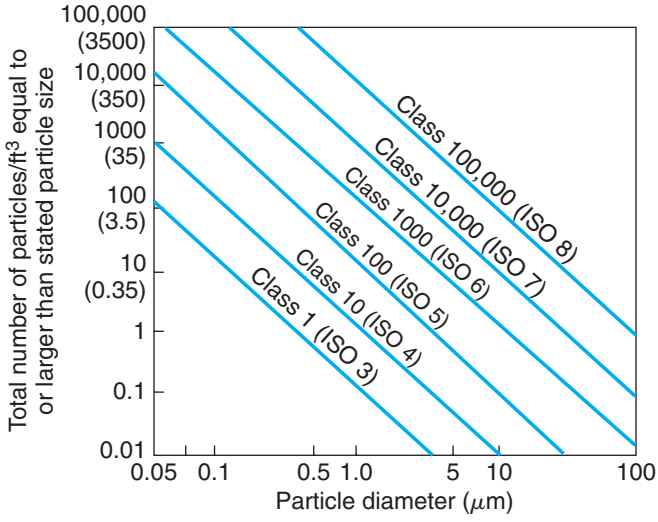


FIGURE 28.4 Allowable particle size counts for various clean-room classes; the numbers in parentheses on the ordinate are particle counts per cubic meter.

conditioned so that it is at 21°C (70°F) and 45% relative humidity. The largest source of contaminants in clean rooms are the workers themselves. Skin particles, hair, perfume, makeup, clothing, bacteria, and viruses are given off naturally by people, and in sufficiently large numbers to quickly compromise a Class-100 clean room. For these reasons, most clean rooms require special coverings, such as white laboratory coats, gloves, and hairnets, as well as the avoidance of perfumes and makeup. The most stringent clean rooms require full-body coverings, called *clean room smocks*. Other precautions include using (a) a ballpoint pen, instead of a pencil, to avoid objectionable graphite particles from pencils and (b) special clean-room paper, to prevent the accumulation of paper particles in the air.

Clean rooms are designed such that the cleanliness at critical *processing areas* is better than in the clean room in general. This is accomplished by always directing the filtered air *from top down* in the clean rooms, and from *floor to ceiling* in the service aisle, a goal that can be facilitated by laminar-flow hooded work areas. To minimize defects, no product is allowed in the service aisle.

28.3 Semiconductors and Silicon

As the name suggests, **semiconductor materials** have electrical properties that lie between those of conductors and insulators; they exhibit resistivities between 10^{-3} and 10^{8} Ω-cm. Semiconductors have become the foundation for electronic devices, because their electrical properties can be altered when controlled amounts of selected impurity atoms are added to their crystal structures. These impurity atoms, also known as **dopants**, have either one more valence electron (*n*-type, or negative, dopant) or one less valence electron (*p*-type, or positive, dopant) than the atoms in the semiconductor lattice.

For silicon, which is a Group IV element in the Periodic Table, typical *n*-type and *p*-type dopants include, respectively, phosphorus (Group V) and boron (Group III). The electrical operation of semiconductor devices can thus be controlled through the creation of regions with different doping types and concentrations.

Although the earliest electronic devices were fabricated on *germanium*, **silicon** has become the industry standard. The abundance of alternative forms of silicon in the crust of the Earth is second only to that of oxygen, making it economically attractive. Silicon's main advantage over germanium is its large energy gap (1.1 eV), as compared with that of germanium (0.66 eV). This energy gap allows silicon-based devices to operate at temperatures of about 150°C (270°F), higher than devices fabricated on germanium, which operate at about 100°C (180°F).

Another important processing advantage of silicon is that its oxide (*silicon dioxide*, SiO_2) is an excellent electrical *insulator*, and can be used for both isolation and passivation (see Section 3.8) purposes. By contrast, germanium oxide is water soluble and thus unsuitable for electronic devices. Moreover, the oxidized form of silicon allows the production of **metal-oxide-semiconductor** (MOS) devices, which are the basis for MOS transistors. These materials are used in memory devices, processors, and other devices, and are by far the largest volume of semiconductor material produced worldwide.

Structure of Silicon. The crystallographic structure of silicon is a diamond-type fcc structure, as shown in Fig. 28.5, along with the *Miller indices* of an fcc material. (Miller indices are a useful notation for identifying planes and directions within a unit cell; see also Section 1.3.) A crystallographic plane is defined by the reciprocal of its intercepts with the three axes. Because anisotropic etchants (see Section 28.8.1)

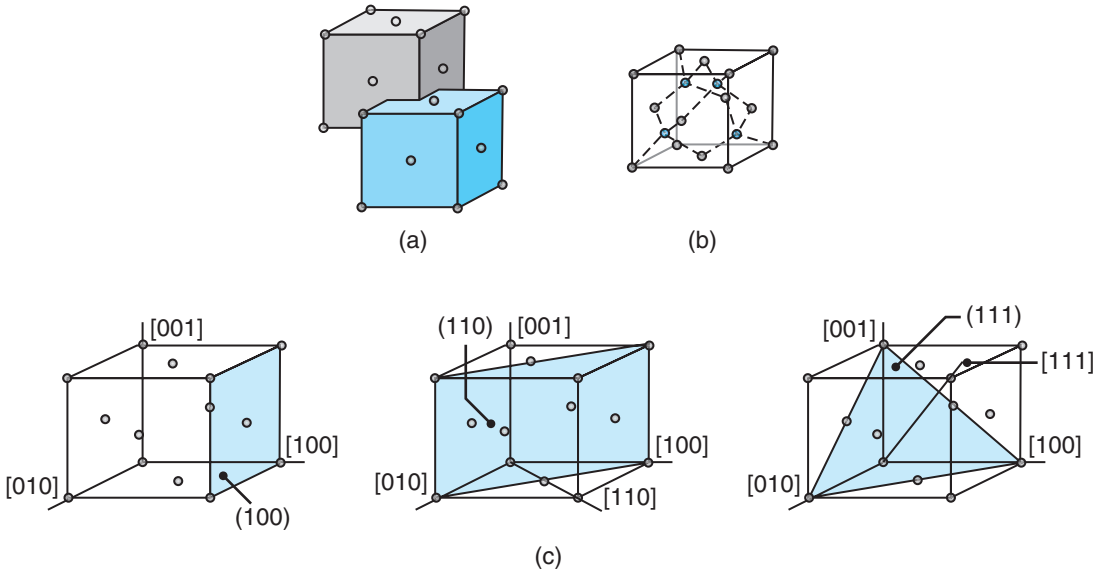


FIGURE 28.5 Crystallographic structure and Miller indices for silicon. (a) Construction of a diamond-type lattice from interpenetrating face-centered cubic cells; one of eight penetrating cells is shown. (b) Diamond-type lattice of silicon; the interiors have been shaded in color. (c) Miller indices for a cubic lattice.

preferentially remove material in certain crystallographic planes, the orientation of the silicon crystal in a wafer is an important consideration.

In spite of its advantages, however, silicon has a larger energy gap (1.1 eV) than germanium oxide and, therefore, has a higher maximum operating temperature (about 200°C; 400°F). This limitation has encouraged the development of *compound semiconductors*, specifically **gallium arsenide**. Its major advantage over silicon is its ability to emit light, thus allowing the fabrication of devices such as lasers and light-emitting diodes (LEDs).

Devices fabricated on gallium arsenide also have much higher operating speeds than those fabricated on silicon. Some of gallium arsenide's disadvantages, on the other hand, are its considerably higher cost, greater processing complications, and, most critically, the difficulty of growing high-quality oxide layers, the need for which is emphasized throughout the rest of this chapter.

28.4 Crystal Growing and Wafer Preparation

Silicon occurs naturally in the forms of silicon dioxide and various silicates. It must, however, undergo a series of purification steps in order to become the high-quality, defect-free, single-crystal material that is necessary for semiconductor device fabrication. The purification process begins by heating silica and carbon together in an electric furnace, which results in a 95–98% pure polycrystalline silicon. This material is converted to an alternative form, commonly trichlorosilane (a compound of silicon, hydrogen, and chlorine), which, in turn, is purified and decomposed in a high-temperature hydrogen atmosphere. The resulting product is extremely high quality *electronic-grade silicon* (EGS).

Single-crystal silicon usually is obtained through the **Czochralski**, or **CZ, process**, described in Section 11.5. The process utilizes a seed crystal that is dipped into

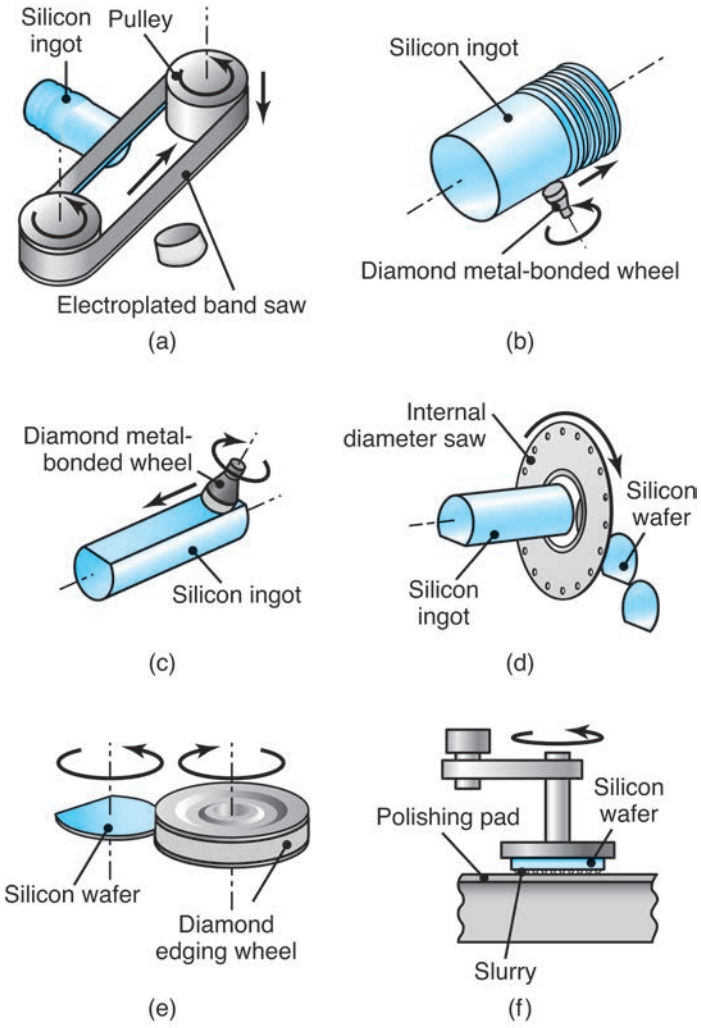


FIGURE 28.6 Finishing operations on a silicon ingot to produce wafers: (a) sawing the ends off the ingot; (b) grinding of the end and cylindrical surfaces of a silicon ingot; (c) machining of a notch or flat; (d) slicing of wafers; (e) end grinding of wafers; and (f) chemical–mechanical polishing of wafers.

a silicon melt, and is then pulled out slowly while being rotated. At this point, controlled amounts of impurities can be added to obtain a uniformly doped crystal. The result of the CZ process is a cylindrical single-crystal ingot, typically 100–300 mm (4–12 in.) in diameter and over 1 m (40 in.) in length. Because this technique does not allow for exact control of the ingot diameter, ingots are grown a few millimeters larger than the required size, and are ground to a desired diameter. Silicon **wafers** then are produced from silicon ingots, by a sequence of machining and finishing operations, illustrated in Fig. 28.6.

Next, the crystal is sliced into individual wafers, by using an inner-diameter diamond-encrusted blade (Fig. 24.28f), whereby a rotating, ring-shaped blade with its cutting edge on the inner diameter of the ring is utilized. While the substrate depth required for most electronic devices is no more than several microns, wafers typically are cut to a thickness of about 0.5 mm (0.02 in.). This thickness provides the physical support necessary for the absorption of temperature variations and the mechanical support needed during subsequent fabrication of the wafer.

The wafer is then ground along its edges using a diamond wheel. This operation gives the wafer a rounded profile, which is more resistant to chipping. Finally, the wafers must be polished and cleaned, to remove surface damage caused by the sawing process. This operation is commonly performed by *chemical–mechanical polishing*, also referred to as chemical–mechanical planarization, described in Section 26.7.

In order to properly control the manufacturing process, it is important to determine the *orientation* of the crystal in a wafer, which is done by notches or flats machined into them for identification, as shown in Fig. 28.7. Most commonly, the (100) or (111) plane of the crystal defines the wafer surface, although (110) surfaces also can be used for micromachining applications (Section 29.2). Wafers are also identified by a laser *scribe* mark, produced by the manufacturer. Laser scribing of information may take place on the front or on the back side of the wafer. The front side of some wafers has an exclusion edge area, 3–10 mm in size and reserved for the scribe information, such as lot numbers, orientation, and a wafer identification code unique to the particular manufacturer.

Wafers are typically processed in lots of 25 or 50, with 150–200 mm (6–8 in.) diameters each, or lots of 12–25 with 300-mm (12-in.) diameters each. In this way, they can be easily handled and transferred during subsequent processing steps. Because of the small device size and large wafer diameter, thousands of individual circuits can be placed on one wafer. Once processing is completed, the wafer is then sliced into individual **chips**, each containing one complete integrated circuit.

At this point, the single-crystal silicon wafer is ready for the fabrication of the integrated circuit or device. Fabrication takes place over the entire wafer surface, and thus many chips are produced at the same time, as shown in Fig. 28.1a. Because the number of chips that can be produced is dependent on the cross-sectional area of the wafer, advanced-circuit manufacturers have moved toward using larger single-crystal solid cylinders; 300-mm (12-in.) diameter wafers now are common, with 450-mm (18-in.) diameter wafers under development.

28.5 Film Deposition

Films are used extensively in microelectronic-device processing, particularly insulating and conducting types. Commonly deposited films include polysilicon, silicon nitride, silicon dioxide, and tungsten, titanium, and aluminum. In some cases, the wafers merely serve as a mechanical support, on which custom *epitaxial layers* are grown.

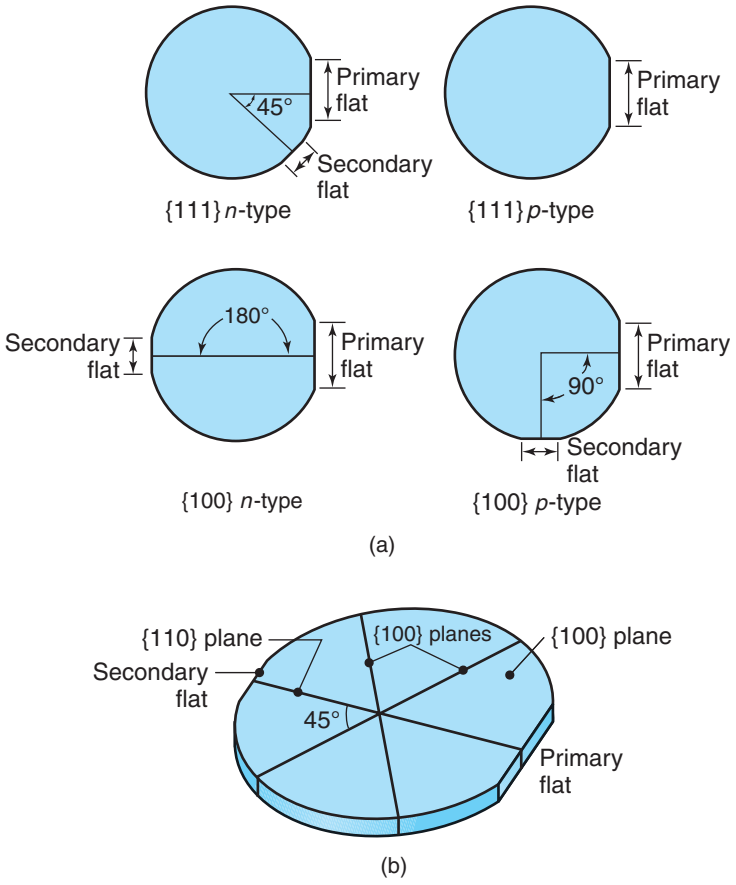


FIGURE 28.7 Identification of single-crystal wafers of silicon; this identification scheme is common for 150-mm (6-in.) diameter wafers, but notches are more common for larger wafers.

Epitaxy. Defined as the growth of a vapor deposit, *epitaxy*, or *electrodeposit*, occurs when the crystal orientation of the deposit is related directly to the crystal orientation in the underlying crystalline substrate. The advantages of processing on these deposited films, instead of on the actual wafer surface, include fewer impurities, especially carbon and oxygen, improved device performance, and the tailoring of material properties (which cannot be done on the wafers themselves).

Some of the major functions of deposited films are **masking** and protecting the semiconductor surface. In masking applications, the film must both inhibit the passage of dopants and concurrently display an ability to be etched into patterns of high resolution. Upon completion of device fabrication, films are applied to protect the underlying circuitry. Films used for masking and protecting include silicon dioxide, phosphosilicate glass (PSG), and silicon nitride. Each of these materials has distinct advantages, and they often are used in combination.

Conductive films are used primarily for device interconnection. These films must have a low electrical resistivity, be capable of carrying large currents, and be suitable for connection to terminal packaging leads with wire bonds. Generally, aluminum and copper are used for this purpose. Increasing circuit complexity has required up to six levels of conductive layers, all of which must be separated by insulating films.

Film Deposition. Films can be *deposited* by several techniques, involving a variety of pressures, temperatures, and vacuum systems (see also Chapter 34):

- One of the oldest and simplest methods is **vacuum deposition**, used primarily for depositing metal films. The metal is first heated in a vacuum to its point of

vaporization; upon evaporation, it forms a thin layer on the substrate surface. The heat of evaporation usually is generated by a heating filament or electron beam.

- **Sputtering** involves bombarding a target in a vacuum with high-energy ions, usually argon, aluminum, or copper. Sputtering systems generally include a DC power supply to produce the energized ions. As the ions impinge on the target, atoms are knocked off and are subsequently deposited on wafers mounted within the system. Although some argon may be trapped within the film, sputtering results in highly uniform coverage. Recent advances include using a radio-frequency power source (**RF sputtering**) and introducing magnetic fields (**magnetron sputtering**).
- In one of the most common techniques, **chemical-vapor deposition** (CVD), film is deposited by way of the reaction and/or decomposition of gaseous compounds. Using this technique, silicon dioxide is deposited routinely by the oxidation of silane or a chlorosilane. Figure 28.8a shows a continuous CVD reactor that operates at atmospheric pressure.

  A similar method that operates at lower pressures is referred to as **low-pressure chemical-vapor deposition** (LPCVD), as shown in Fig. 28.8b. Capable of coating hundreds of wafers at a time, this method results in a much higher production rate than that of atmospheric-pressure CVD, and provides superior film uniformity and with less consumption of carrier gases. The technique is commonly used for depositing polysilicon, silicon nitride, and silicon dioxide.
- **Plasma-enhanced chemical-vapor deposition** (PECVD) involves the processing of wafers in an RF plasma containing the source gases. This method has the advantage of maintaining a low wafer temperature during deposition.

Silicon **epitaxy** layers, in which the crystalline layer is formed, using the substrate as a seed crystal, can be grown by a variety of methods. If the silicon is deposited from the gaseous phase, the process is known as **vapor-phase epitaxy** (VPE). In another variation, the heated substrate is brought into contact with a liquid solution containing the material to be deposited, called **liquid-phase epitaxy** (LPE).

Another high-vacuum process, called **molecular-beam epitaxy** (MBE), utilizes evaporation to produce a thermal beam of molecules that are deposited on the heated substrate. This process results in a very high degree of purity. In addition, since the films are grown one atomic layer at a time, it is possible to have excellent control over doping profiles. This level of control is important especially in gallium-arsenide

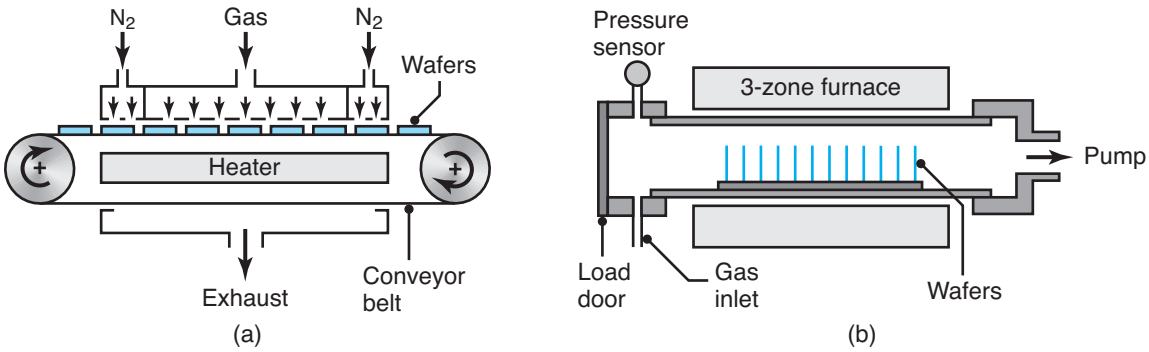


FIGURE 28.8 Schematic diagrams of a (a) continuous atmospheric-pressure CVD reactor and (b) low-pressure CVD. *Source:* After S.M. Sze.

technology; however, the MBE process has relatively low growth rates as compared to other conventional film-deposition techniques.

28.6 Oxidation

The term *oxidation* refers to the growth of an oxide layer as a result of the reaction of oxygen with the substrate material. Oxide films also can be formed by the previously described deposition techniques. Thermally grown oxides, described in this section, display a higher level of purity than deposited oxides, because they are grown directly from the high-quality substrate. However, methods of deposition must be used if the composition of the desired film is different from that of the substrate material.

Silicon dioxide is the most widely used oxide in IC technology today, and its excellent characteristics are one of the major reasons for the widespread use of silicon. Aside from its effectiveness in dopant masking and device isolation, silicon dioxide's most critical role is that of the *gate oxide* material. Silicon surfaces have an extremely high affinity for oxygen, and thus a freshly sawed slice of silicon will quickly grow a native oxide of 30–40 Å thickness.

- **Dry oxidation** is a relatively simple process, and is accomplished by elevating the substrate temperature, typically to about 750°–1100°C (1380°–2020°F), in an oxygen-rich environment. As a layer of oxide forms, the oxidizing agents must be able to pass through the oxide and reach the silicon surface, where the actual reaction takes place. Thus, an oxide layer does not continue to grow on top of itself, but rather, it grows from the silicon surface outward. Some of the silicon substrate is consumed in the oxidation process (Fig. 28.9).

  The ratio of oxide thickness to the amount of silicon consumed is found to be 1:0.44; thus, to obtain an oxide layer 1000 Å thick, approximately 440 Å of silicon will be consumed. This requirement does not present a problem, as substrates always are grown sufficiently thick. One important effect of the consumption of silicon is the rearrangement of dopants in the substrate near the interface. Because different impurities have different segregation coefficients or mobilities in silicon dioxide, some dopants become depleted away from the oxide interface while others pile up there. Consequently, processing parameters must be properly adjusted to compensate for this effect.
- **Wet oxidation** utilizes a water-vapor atmosphere as the agent. This method results in a considerably higher growth rate than that of dry oxidation, but it suffers from a lower oxide density and, therefore, a lower dielectric strength. The common practice in industry is to combine both dry and wet oxidation methods, by growing an oxide in a three-part layer: dry–wet–dry. This approach combines the advantages of wet oxidation's much higher growth rate and dry oxidation's high quality.

The two oxidation methods described are useful primarily for coating the entire silicon surface with oxide; however, it also may be necessary to oxidize only certain portions of the surface. The procedure of oxidizing only certain areas is called **selective oxidation**, and uses silicon nitride, which inhibits the passage of oxygen and water vapor. Thus, by covering certain areas with silicon nitride, the silicon under these areas remains unaffected while the uncovered areas are oxidized.

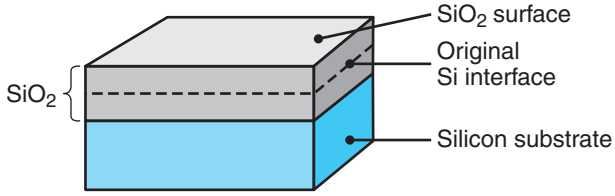


FIGURE 28.9 Growth of silicon dioxide, showing consumption of silicon. *Source:* After S.M. Sze.

TABLE 28.1

General Characteristics of Lithography Techniques

Method	Wavelength (nm)	Finest feature size (nm)
Ultraviolet (photolithography)	365	350
Deep UV	248	250
Extreme UV	10–20	30–100
X-ray	0.01–1	20–100
Electron beam	—	80
Immersion	193	11

Source: After P.K. Wright.

28.7 Lithography

Lithography is the process by which the geometric patterns that define devices are transferred to the substrate surface. A summary of lithographic techniques is given in Table 28.1, and a comparison of the basic lithography methods is shown in Fig. 28.10. The most common technique used today is **photolithography**, and most IC applications can be manufactured successfully with photolithography. *Electron-beam* and *X-ray lithography* are of great interest, because of their ability to transfer patterns with higher resolution, a necessary feature for the increased miniaturization of integrated circuits.

Photolithography. Photolithography uses a **reticle**, also called a **mask** or **photomask**, which is a glass or quartz plate with a pattern of the chip deposited onto it with a chromium film. The reticle image can be the same size as the desired structure on the chip, but it is often an enlarged image, usually 4× to 20× larger, although 10× magnification is the most common. The enlarged images

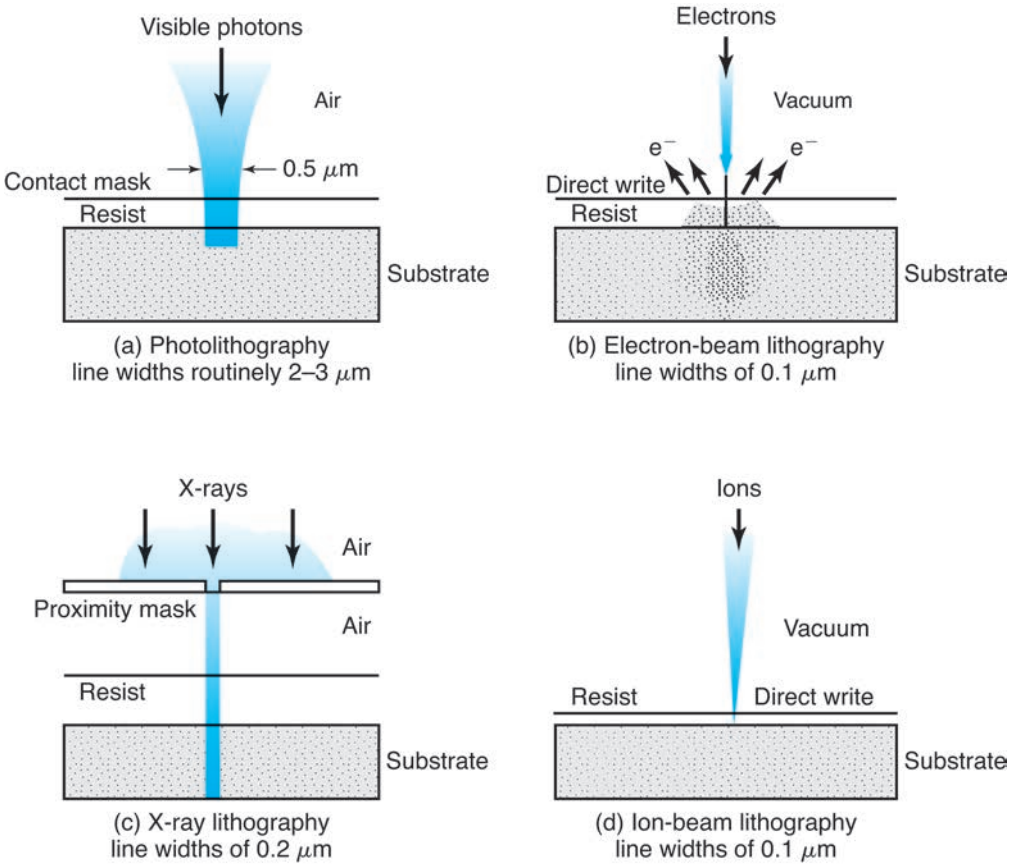


FIGURE 28.10 Comparison of lithography techniques; note that (a) and (c) involve masking to achieve pattern transfer, while (b) and (d) scribe the pattern without a mask, known as direct writing.

are then focused onto a wafer through a lens system, an operation that is referred to as *reduction lithography*.

In current practice, the lithographic process is applied to each microelectronic circuit as many as 25 times, each time using a different reticle to define the different areas of the working devices. Typically designed at several thousand times their final size, reticle patterns undergo a series of reductions before being applied permanently to a defect-free quartz plate. Computer-aided design (CAD; see Section 38.4) has had a major impact on reticle design and generation.

Cleanliness is especially important in lithography, and manufacturers now use robotics and specialized wafer-handling equipment in order to minimize contamination from dust and dirt. Once the film deposition process is completed and the desired reticle patterns have been generated, the wafer is cleaned and coated with a **photoresist** (PR), an organic polymer.

A photoresist consists of three principal components:

1. A polymer that changes its structure when exposed to radiation
2. A sensitizer that controls the reactions in the polymer
3. A solvent, to deliver the polymer in liquid form

Photoresist layers 0.5–2.5 μm (20–100 μin.) thick are produced by applying the photoresist to the substrate and then spinning it, at several thousand rpm, for 30 or 60 s to give uniform coverage (Fig. 28.11).

The next step in photolithography is **prebaking** the wafer, to remove the solvent from the photoresist and harden it. This step is carried out on a hot plate, heated to around 100°C. The pattern is transferred to the wafer through *stepper* or *step-and-scan* systems. With wafer steppers (Fig. 28.12a), the full image is exposed in one flash, and the reticle pattern is then refocused onto another adjacent section of the wafer. With step-and-scan systems (Fig. 28.12b), the exposing light source is focused into a line, and the reticle and wafer are translated simultaneously in opposite directions to transfer the pattern.

The wafer must be aligned carefully under the desired reticle. In this crucial step, called **registration**, the reticle must be aligned correctly with the previous layer on the wafer. Once the reticle is aligned, it is subjected to ultraviolet (UV) radiation. Upon development and removal of the exposed photoresist, a duplicate of the reticle

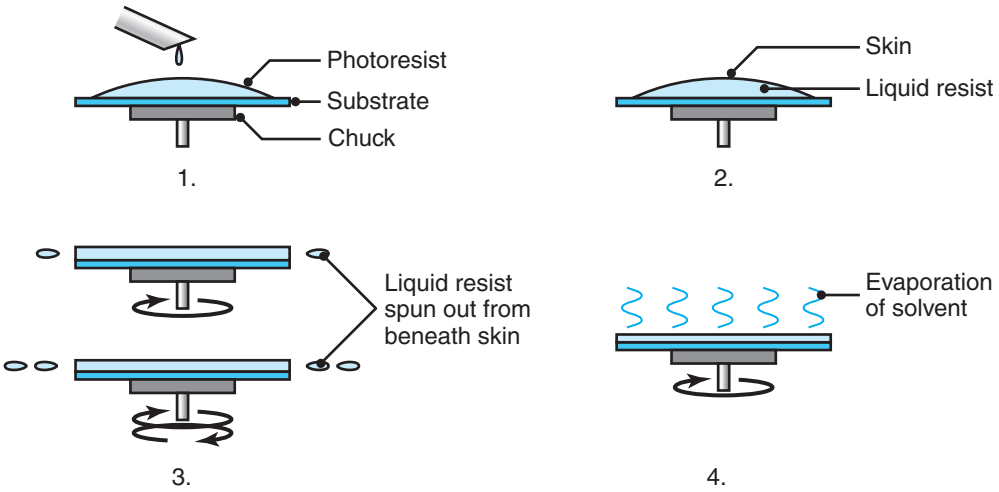


FIGURE 28.11 Spinning of an organic coating on a wafer.

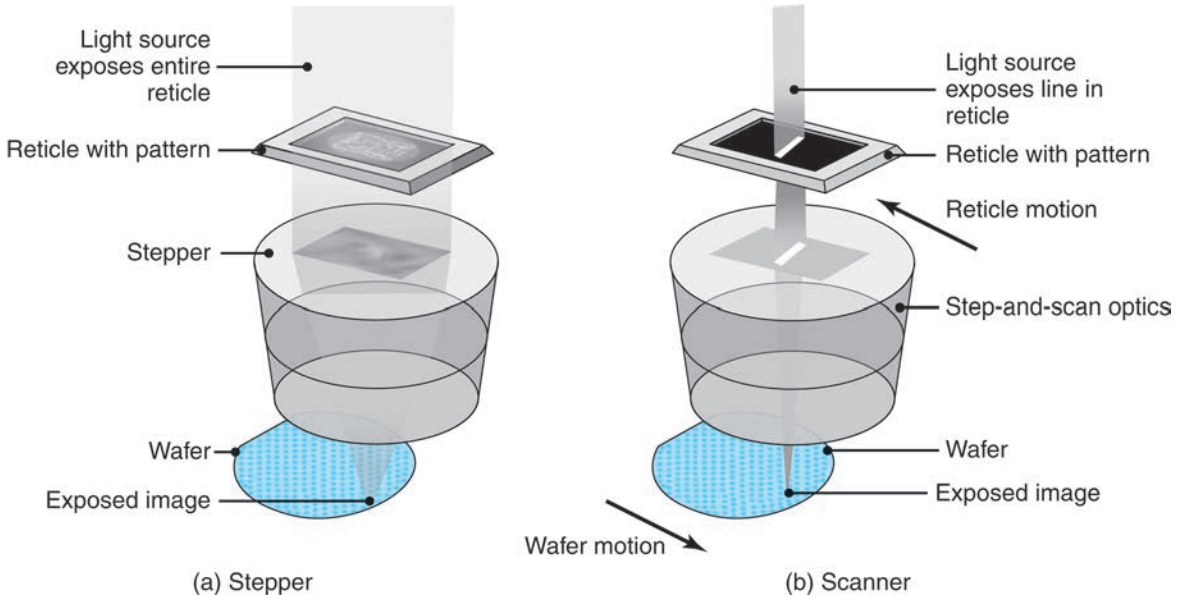


FIGURE 28.12 Schematic illustration of (a) wafer stepper technique for pattern transfer and (b) wafer step-and-scan technique.

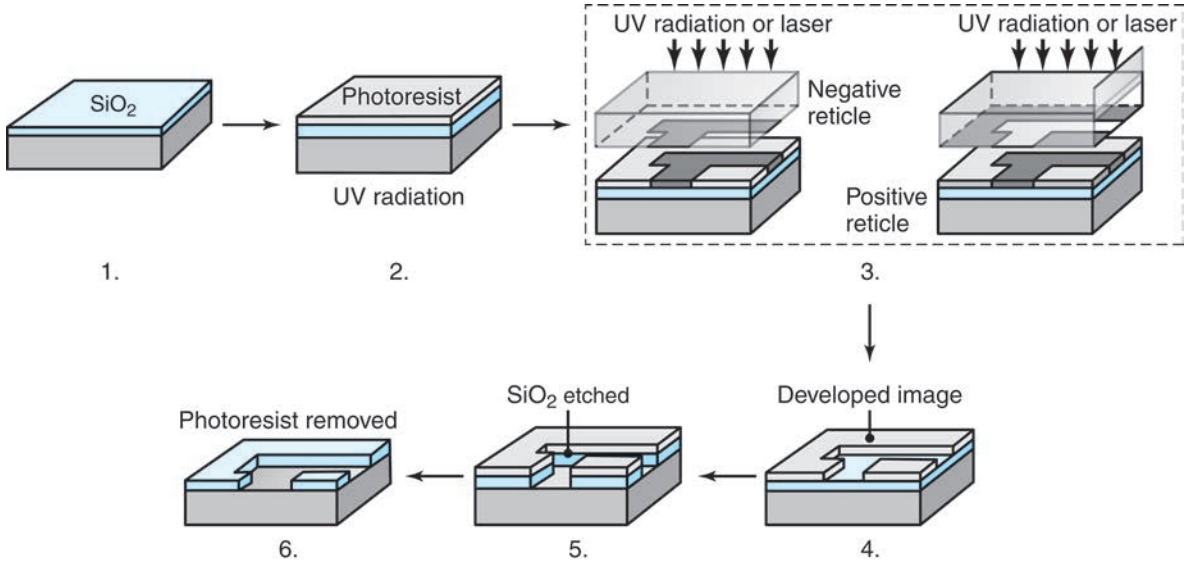


FIGURE 28.13 Pattern transfer by photolithography; note that the reticle in Step 3 can be a positive or negative image of the pattern.

pattern will appear in the photoresist layer. As seen in Fig. 28.13, the reticle can be a negative image or a positive image of the desired pattern. A positive reticle uses the UV radiation to break down the chains in the organic film, so that these films are removed preferentially by the developer. Positive masking is more common than negative masking because, with negative masking, the photoresist can swell and distort, thus making it unsuitable for small features. Newer negative photoresist materials do not have this problem.

Following the exposure and development sequence, **postbaking** the wafer drives off the solvent, and toughens and improves the adhesion of the remaining resist. In addition, a deep UV treatment (thereby baking the wafer to about 150°–200°C in ultraviolet light) can be used to further strengthen the resist against high-energy implants and dry etches. The underlying film not covered by the photoresist is then etched away (Section 28.8) or implanted (Section 28.9).

Following lithography, the developed photoresist must be removed, in a process called **stripping**. In *wet stripping*, the photoresist is dissolved by such solutions as acetone or strong acids. In this method, the solutions tend to lose potency in use. *Dry stripping* involves exposing the photoresist to an oxygen plasma, referred to as **ashing**. Dry stripping has become more common, because it (a) does not involve the disposal of consumed hazardous chemicals and (b) is easier to control and can result in exceptional surfaces.

One of the major issues in lithography is **line width**, which is the width of the smallest feature imprintable on the silicon surface, and is called **critical dimension** (CD). As circuit densities have escalated over the years, device sizes and features have become smaller and smaller; today, commercially feasible minimum critical dimension is 32 nm, with a major trend to obtain 16 nm or even smaller.

Because pattern resolution and device miniaturization have been limited by the wavelength of the radiation source used, the need has arisen to move to wavelengths shorter than those in the UV range, such as deep UV wavelengths, extreme UV wavelengths, electron beams, and X-rays. In these technologies, the photoresist is replaced by a similar resist that is sensitive to a specific range of shorter wavelengths.

Pitch Splitting. Multiexposure techniques have been developed to obtain higher resolution images than can be attained through conventional single-exposure lithography, and have been applied for sub-32 nm feature development. *Pitch splitting* is shown in Fig. 28.14, using conventional lithography in multiple stages. Recognizing that the spacing between features is the limiting dimension, pitch splitting involves breaking up the desired pattern into two complimentary portions and creating corresponding masks. By using two imaging steps, features can be developed in the substrate with twice the resolution of a single imaging step.

There are two forms of pitch splitting. In **double exposure** (DE), a mask exposes some of the desired trenches or regions in the photoresist, then a second mask is used to expose the remaining features (Fig. 28.14a). The photoresist is then exposed and the substrate is etched. **Double patterning** (DP) involves two sequential lithography and etch steps, so that it is sometimes referred to as the *LELE* (lithography-etch-lithography-etch) *process*.

Immersion Lithography. The resolution of lithography systems can be increased by inserting a fluid with a high refractive index between the final lens and the wafer, an approach called *immersion lithography*. Water has been mainly used to date, and has been the main approach used to attain feature sizes below 45 nm. Fluids with a refractive index higher than that of water also are being investigated to increase the resolutions of immersion lithography.

Immersion lithography requires careful process controls, especially thermal controls, since any bubbles that develop in the water will result in defects due to distortion of the light source.

Extreme Ultraviolet Lithography. The pattern resolution in photolithography and immersion lithography is ultimately limited by light diffraction. One of the means of reducing the effects of diffraction is to use ever shorter wavelengths. *Extreme ultraviolet lithography* (EUV) uses light at a wavelength of 13 nm, in order to obtain

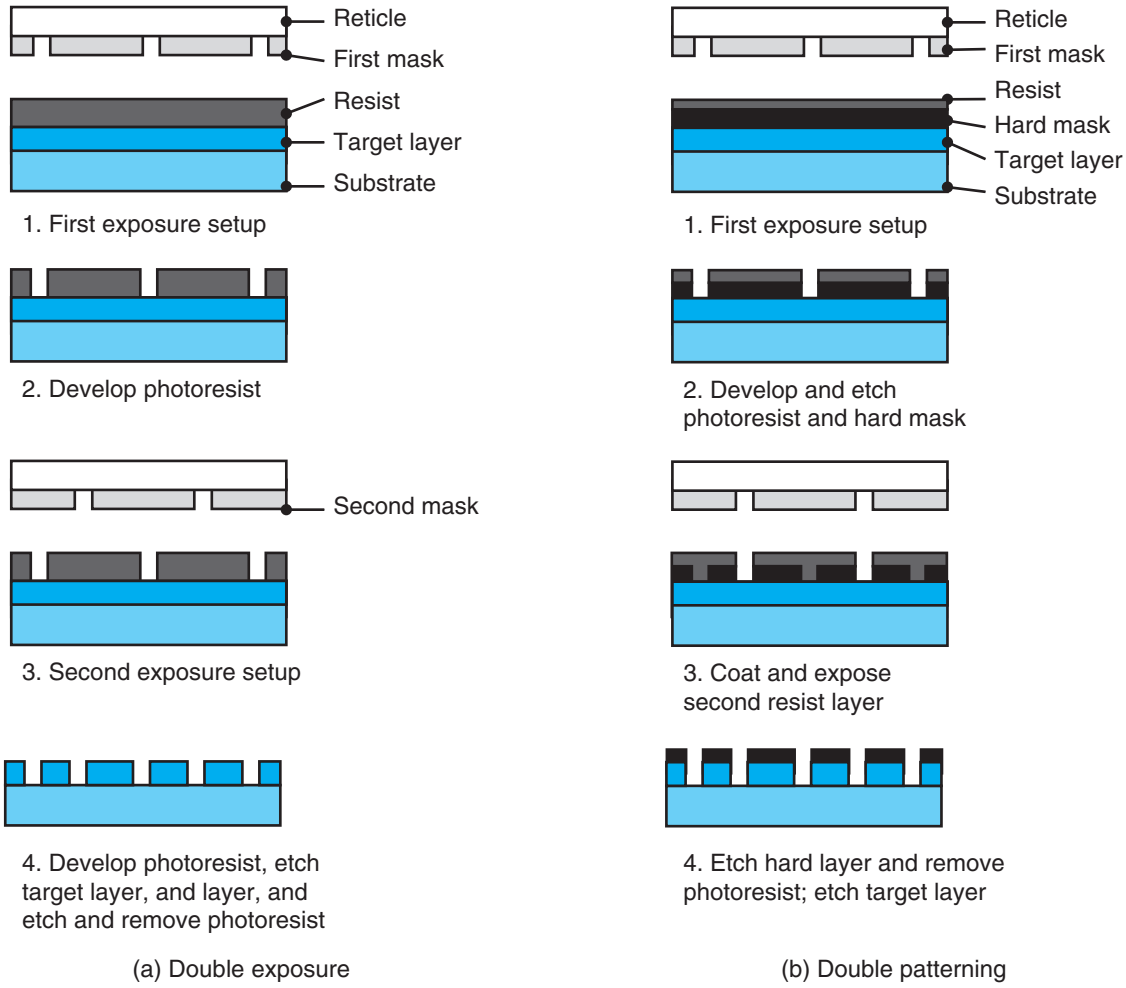


FIGURE 28.14 Pitch splitting lithography. (a) Double exposure (DE) process and (b) double patterning (DP) process, also known as the LELE (lithography-etch-lithography-etch) process.

features commonly in the range from 30 to 100 nm, but is expected to be useful for sub-25 nm features. Because glass lenses absorb some EUV light, the waves are focused through highly reflective molybdenum–silicon mirrors through the mask to the wafer surface.

X-Ray Lithography. Although photolithography is the most widely used lithography technique, it has fundamental resolution limitations associated with light diffraction. *X-ray lithography* is superior to photolithography, because of the shorter wavelength of the radiation and its very large depth of focus. These characteristics allow much finer patterns to be resolved, and make X-ray lithography far less susceptible to dust than photolithography. Moreover, the *aspect ratio* (defined as the ratio of depth to lateral dimension) can be higher than 100, whereas it is limited to around 10 with photolithography. However, to achieve this benefit, synchrotron radiation is required, which is expensive and available at only a few research laboratories.

Given the large capital investment required for a manufacturing facility, industry has preferred to refine and improve optical lithography, instead of investing new

capital into X-ray-based production. Currently, X-ray lithography is not widespread, although the LIGA process, described in Section 29.3, fully exploits the benefits of this technique.

Electron-beam and Ion-beam Lithography. Like X-ray lithography, *electron-beam* (e-beam) and *ion-beam* lithography are superior to photolithography, in terms of attainable resolutions. These two methods involve high current density in narrow electron or ion beams (known as *pencil sources*), which scan a pattern one pixel at a time onto a wafer. The masking is done by controlling the point-by-point transfer of the stored pattern, called *direct writing*, using software. These techniques have the advantages of accurate control of exposure over small areas of the wafer, large depth of focus, and low defect densities. Resolutions are limited to about 10 nm, because of electron scatter, although 2-nm resolutions have been reported for some materials.

It should be noted that the scan time significantly increases as the resolution improves, because more highly focused beams are required. The main drawback of these two techniques is that electron and ion beams have to be maintained in a vacuum, thus significantly increasing equipment complexity and production time. Moreover, the scan time for a wafer is much longer than that for other lithographic methods.

SCALPEL. In the SCALPEL (*sc*attering with *a*ngular *l*imitation *p*rojection *e*lectron-beam *l*ithography) process (Fig. 28.15), a mask is first produced from about a 0.1-μm-thick membrane of silicon nitride, and then patterned with an approximately 50-nm-thick coating of tungsten. High-energy electrons pass through both the silicon nitride and the tungsten, but the tungsten scatters the electrons widely, whereas

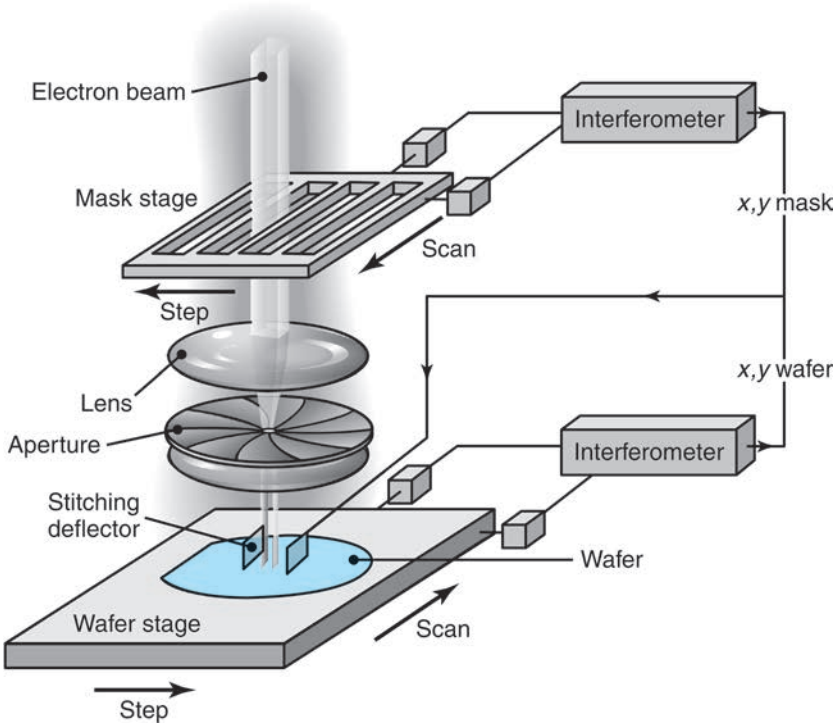


FIGURE 28.15 Schematic illustration of the SCALPEL process.

the silicon nitride results in very little scattering. An aperture blocks the scattered electrons, resulting in a high-quality image at the wafer.

The limitation to the SCALPEL process is the small-sized masks that are currently in use, but the process has high potential. Perhaps its most significant advantage is that energy does not need to be absorbed by the reticle; instead, it is blocked by the aperture, which is not as fragile or expensive as the reticle.

EXAMPLE 28.1 Moore's Law

G. Moore, an inventor of the integrated circuit and past chairman of Intel Corporation, observed in 1965 that the surface area of a single transistor is reduced by 50% every 12 months. In 1975, he revised this estimate to every two years; the resulting estimate is widely known as Moore's law, and it has been remarkably accurate. Figure 28.16 shows the historical progression of feature size in *dynamic random access memory* (DRAM) bits, as well as projected future developments. Looking ahead, however, one can note that there are some major impediments to the continued reliability of Moore's law. Among the more important ones are:

- To produce ever smaller features in a transistor requires that even more stringent manufacturing tolerances be achieved. For example, 180-nm line widths require ±14-nm dimensional tolerances, whereas 50-nm line widths require ±4 nm. Either requirement is especially problematic for the metal connection lines within the transistor.
- Smaller transistors can operate only if the dopant concentration is increased. Above a certain limit, however, the doping atoms cluster together. The result is that *p*-type and *n*-type silicon cannot be produced reliably at small length scales.
- The gate-switching energy of transistors has not been reduced at the same rate as their size; the result is increased power consumption in integrated circuits. This effect has a serious consequence, in that it is very difficult to dissipate the heat produced.
- At smaller length scales, microprocessors require lower voltages for proper operation. However, since the power consumption is still relatively high, very large currents are needed between the power-conversion devices and the central-processing units of modern microprocessors. These large currents result in resistive heating, thus compounding the heat-extraction problems.

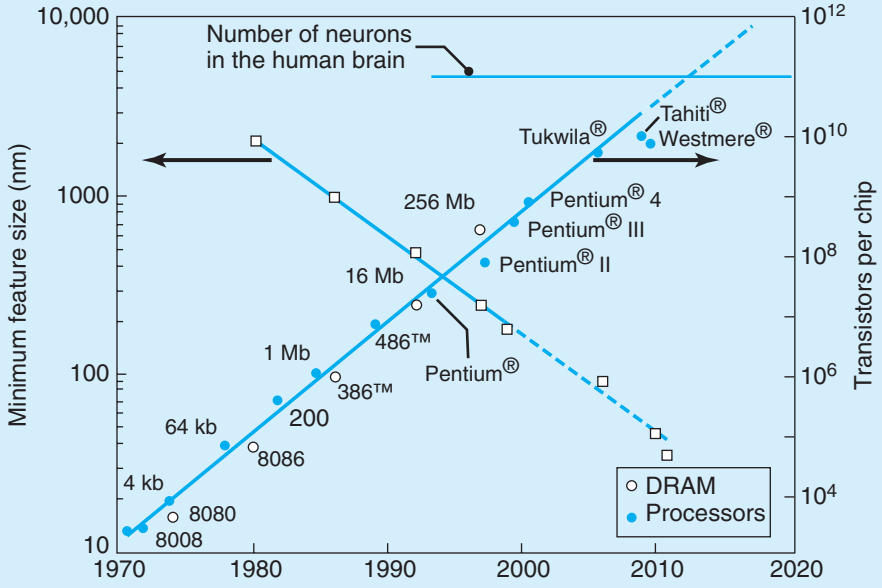


FIGURE 28.16 Illustration of Moore's law.

Much research continues to be directed toward overcoming these limitations. Moore's law was intended as a prediction of the short-term future of the semiconductor industry, and was put forward at a time when photolithography was the only option. During the four decades since the law was first stated, researchers often have identified seemingly insurmountable problems which, in turn, have been overcome.

Soft Lithography. *Soft lithography* refers to several processes for pattern transfer, all of which require that a master mold be created by one of the standard lithography techniques described above. The master mold is then used to produce an elastomeric pattern, or stamp, as shown in Fig. 28.17. An elastomer that has commonly been used for the stamp is silicone rubber (polydimethylsiloxane, PDMS), because it is chemically inert, is not hygroscopic (it does not swell due to humidity), and has good thermal stability, strength, durability, and surface properties.

Several PDMS stamps can be produced using the same pattern, and each stamp can be used several times. Some of the common soft lithography processes are:

1. **Microcontact printing** (μCP). In *microcontact printing*, the PDMS stamp is coated with an "ink" and then pressed against a surface. The peaks of the pattern are in contact with the opposing surface, and a thin layer of the ink is transferred, often only one molecule thick (called a *self-assembled monolayer* or *boundary film*; see Section 33.6). This film can serve as a mask for selective wet etching, described below, or it can be used to impart a desired chemistry onto the surface.
2. **Microtransfer molding** (μTM). In this process, shown in Fig. 28.18a, the recesses in the PDMS mold are filled with a liquid polymer precursor, and then pressed against a surface. After the polymer has cured, the mold is peeled off, leaving behind a pattern suitable for further processing.
3. **Micromolding in capillaries.** Called MIMIC, in this technique (Fig. 28.18b) the PDMS stamp pattern consists of channels that use capillary action to wick a liquid into the stamp, either from the side of the stamp or from reservoirs within the stamp itself. The liquid can be a thermosetting polymer, a ceramic sol gel, or suspensions of solids within liquid solvents. Good pattern replication can be

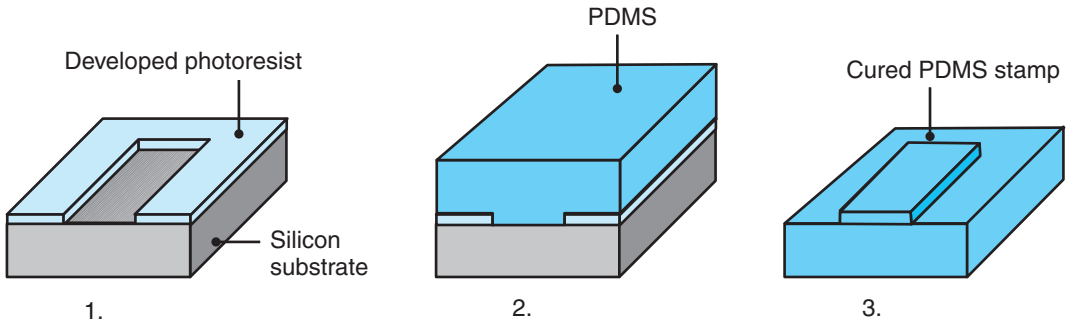


FIGURE 28.17 Production of a polydimethylsiloxane (PDMS) mold for soft lithography. 1. A developed photoresist is produced through standard lithography (see Fig. 28.13). 2. A PDMS stamp is cast over the photoresist. 3. The PDMS stamp is peeled off the substrate to produce a stamp. The stamp shown has been rotated to emphasize the replication of surface features; the master pattern can be used several times. *Source:* Based on Y. Xia and G.M. Whitesides.

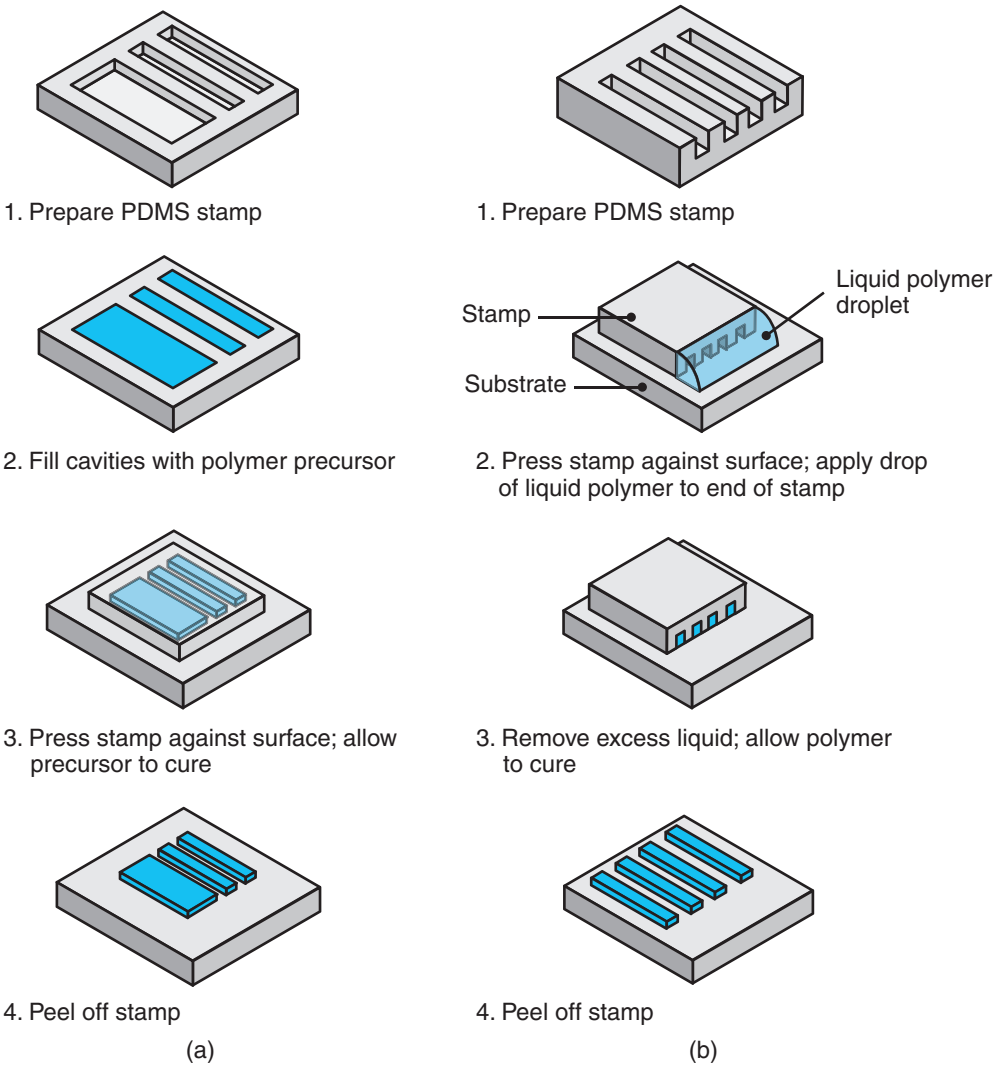


FIGURE 28.18 Soft lithography techniques. (a) Microtransfer molding (μTM) and (b) micromolding in capillaries (MIMIC). *Source:* After Y. Xia and G.M. Whitesides.

obtained, as long as the channel aspect ratio is moderate and the actual channel dimensions allow fluid flow. The dimensions required depend on the liquid used. The MIMIC process has been used to produce all-polymer field-effect transistors and diodes, and has various applications in sensors (Section 37.7).

28.8 Etching

Etching is the process by which entire films or particular sections of films are removed. One of the key criteria in this process is **selectivity**, that is, the ability to etch one material without etching another. In silicon technology, an etching process must etch the silicon-dioxide layer effectively, with minimal removal of either the underlying silicon or the resist material. In addition, polysilicon and metals must be etched into high-resolution lines with vertical wall profiles, and with minimal removal of either the underlying insulating film or the photoresist. Typical etch rates range from hundreds

TABLE 28.2

General Characteristics of Silicon Etching Operations

Wet etching	Temperature (°C)	Etch rate (μm/min)	{111}/{100} selectivity	Nitride etch rate (nm/min)	SiO_2 etch rate (nm/min)	p^{++} etch stop
Wet etching						
$HF:HNO_3:CH_3COOH$	25	1–20	—	Low	10–30	No
KOH	70–90	0.5–2	100:1	<1	10	Yes
Ethylene-diamine pyrocatechol (EDP)	115	0.75	35:1	0.1	0.2	Yes
$N(CH_3)_4OH$ (TMAH)	90	0.5–1.5	50:1	<0.1	<0.1	Yes
Dry (plasma) etching						
SF_6	0–100	0.1–0.5	—	200	10	No
SF_6/C_4F_8 (DRIE)	20–80	1–3	—	200	10	No

Source: Based on Kovacs, G.T.A., Maluf, N.I., and Peterson, K.E., "Bulk micromachining of silicon" in *Integrated Sensors, Microactuators and Microsystems (MEMS)*, pp. 1536–1551, K.D. Wise (ed), Proceedings of the IEEE, v. 86, No. 8, August 1998.

to several thousands of angstroms per minute, and selectivities can range from 1:1 to 100:1. A summary of etching processes and etchants is given in Tables 28.2 and 28.3.

28.8.1 Wet Etching

Video Solution 28.1 Wet Etching

Wet etching involves immersing the wafers in a liquid solution, usually acidic. A primary feature of most wet-etching operations is that they are *isotropic*; that is, they etch in all directions of the workpiece at the same rate. Isotropy results in *undercuts* beneath the mask material (see, for example, Figs. 27.3b and 28.19a), and thus limits the resolution of geometric features in the substrate.

Effective etching requires the following conditions:

1. Etchant transport to the surface
2. A chemical reaction
3. Transport of reaction products away from the surface
4. Ability to stop the etching process rapidly, known as *etch stop*, in order to obtain superior pattern transfer, usually by using an underlying layer with high selectivity

If the first or third condition listed above limits the speed of the process, agitation or stirring of the etchant can be employed to increase etching rates. If the second condition limits the speed of the process, the etching rate will strongly depend on temperature, etching material, and solution composition. Reliable etching thus requires both proper temperature control and repeatable stirring capability.

Isotropic Etchants. These etchants are widely used for:

- Removing damaged surfaces
- Rounding sharply etched corners, to avoid stress concentrations
- Reducing roughness developed after anisotropic etching
- Creating structures in single-crystal slices
- Evaluating defects

Fabrication of microelectronic devices as well as microelectromechanical systems, described in Chapter 29, requires the precise machining of structures, done through masking. With isotropic etchants, however, masking can be a challenge, because the strong acids (a) etch aggressively at a rate of up to 50 μm/min, with an etchant of 66% HNO_3 and 34% HF, although etch rates of 0.1–1 μm/s are more typical,

TABLE 28.3

Comparison of Etch Rates for Selected Etchants and Target Materials

Etchant	Target material	Etch rate (nm/min)[a] Polysilicon n^+	Polysilicon, undoped	Silicon dioxide	Silicon nitride	Phospho-silicate glass, annealed	Aluminum	Titanium	Photoresist (OCG-820PR)
Wet etchants									
Concentrated HF (49%)	Silicon oxides	0	—	2300	14	3600	4.2	>1000	0
25:1 HF:H_2O	Silicon oxides	0	0	9.7	0.6	150	—	—	0
5:1 BHF[b]	Silicon oxides	9	2	100	0.9	440	140	>1000	0
Silicon etchant (126 HNO_3:60 H_2O:5 NH_4F)	Silicon	310	100	9	0.2	170	400	300	0
Aluminum etchant (16 H_3PO_4:1 HNO_3:1 HAc:2 H_2)	Aluminum	<1	<1	0	0	<1	660	0	0
Titanium etchant (20 H_2O:1 H_2O_2:1 HF)	Titanium	1.2	—	12	0.8	210	>10	880	0
Piranha (50 H_2SO_4:1 H_2O_2)	Metals and organics (cleaning off)	0	0	0	0	0	180	240	>10
Acetone (CH_3COOH)	Photoresist	0	0	0	0	0	0	0	>4000
Dry Etchants									
$CF_4 + CHF_3 + He$, 450 W	Silicon oxides	190	210	470	180	620	—	>1000	220
$SF_6 + He$, 100 W	Silicon nitrides	73	67	31	82	61	—	>1000	69
SF_6, 12.5 W	Thin silicon nitrides	170	280	110	280	140	—	>1000	310
O_2, 400 W	Ashing photoresist	0	0	0	0	0	0	0	340

Notes:

[a]Results are for fresh solutions at room temperature, unless noted. Actual etch rates will vary with temperature and prior use of solution, area of exposure of film, other materials present, film impurities, and microstructure.

[b]Buffered hydrofluoric acid (33% NH4F and 8.3% HF by weight).

Source: After K. Williams and R. Muller.

(b) produce rounded cavities, and (c) are etch rate that is highly sensitive to agitation, thus lateral and vertical features are difficult to control. Because the size of the features in an integrated circuit determines its performance, there is a strong need to produce extremely small and well-defined structures. Such small features cannot be attained through isotropic etching, because of the poor definition resulting from undercutting of masks.

Anisotropic Etching. This situation occurs when etching is strongly dependent on compositional or structural variations in the material. There are two basic types of anisotropic etching: *orientation-dependent etching* (ODE) and *vertical etching*. Orientation-dependent etching commonly occurs in a single crystal, when etching takes place at different rates in different directions, as shown in Fig. 28.19b. Most vertical etching is done with dry plasmas, as described later.

Anisotropic etchants produce geometric shapes, with walls that are defined by the crystallographic planes that resist the etchants. For example, Fig. 28.20 shows the vertical etch rate for silicon as a function of temperature. Note that the etching rate is more than one order of magnitude lower in the [111] crystal direction than in the other directions; therefore, well-defined walls can be obtained along the [111] crystal direction.

The **anisotropy ratio** for etching is defined by

$$\text{AR} = \frac{E_1}{E_2}, \tag{28.1}$$

where E is the etch rate in the crystallographic direction of interest, and the two subscripts refer to two directions for the materials of interest. The anisotropy ratio is unity for isotropic etchants, but can be as high as 400/200/1 for (110)/(100)/(111) silicon. The ⟨111⟩ planes always etch at the lowest rate, but the ⟨100⟩ and ⟨110⟩ planes can be controlled through etchant chemistry.

Masking is also a concern in anisotropic etching, but for different reasons than those for isotropic etching. Anisotropic etching is slow, typically 3 μm/min, with anisotropic etching through a wafer taking as much as several hours. Silicon oxide may etch too rapidly to be used as a mask, hence a high-density silicon nitride mask may be needed.

Often, it is important to rapidly halt the etching process, especially when thin membranes are to be manufactured or features with very precise thickness control are required. Conceptually, rapid halting can be accomplished by removing the wafer from the etching solution. However, etching depends to a great extent on the ability to circulate fresh etchants to the desired locations. Since the circulation varies across a wafer's surface, this strategy for halting the etching process would lead to large variations in the etched depth.

The most common approach to obtain uniform feature sizes across a wafer is to use a boron etch stop (Fig. 28.21), whereby a boron layer is diffused or implanted into the silicon. Examples of common etch stops are the placement of a boron-doped layer beneath silicon or the placement of silicon dioxide (SiO_2) beneath silicon nitride (Si_3N_4). Because anisotropic etchants do not attack boron-doped silicon as aggressively as they do undoped silicon, surface features or membranes can be created by **back etching**.

Several etchant formulations have been developed, including hydrofluoric acid, phosphoric acid, mixtures of nitric acid and hydrofluoric acid, potassium hydrochloride, and mixtures of phosphoric acid, nitric acid, acetic acid, and water. Wafer *cleaning* is done with a solution consisting of sulfuric acid and peroxide, called *Piranha solution*, a trade name. Photoresist can be removed with these solutions, although acetone is commonly used for this purpose.

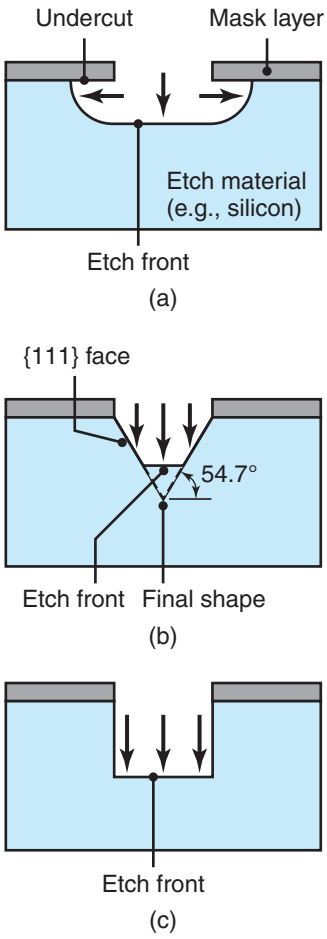


FIGURE 28.19 Etching directionality. (a) Isotropic etching; etch proceeds vertically and horizontally at approximately the same rate, with significant mask undercut. (b) Orientation-dependent etching (ODE); etch proceeds vertically, terminating on {111} crystal planes with little mask undercut. (c) Vertical etching; etch proceeds vertically with little mask undercut. *Source:* After K. Williams and R. Muller.

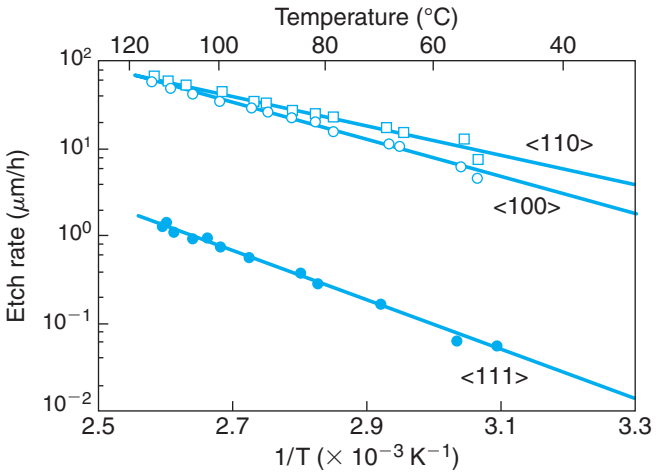


FIGURE 28.20 Etch rates of silicon in different crystallographic orientations, using ethylene-diamine/pyrocatechol-in-water as the solution. *Source:* After H. Seidel.

28.8.2 Dry Etching

Integrated circuits are now etched exclusively by *dry etching*, which involves the use of chemical reactants in a low-pressure system. In contrast to wet etching, dry etching can have a high degree of directionality, resulting in highly anisotropic etching profiles (Fig. 28.19c). Also, the dry-etching process requires only small amounts of the reactant gases, whereas the solutions used in wet etching have to be refreshed periodically.

Dry etching usually involves a plasma or discharge in areas of high electric and magnetic fields; any gases that are present are dissociated to form ions, photons, electrons, or highly reactive molecules. Table 28.2 lists some of the more common dry etchants, their target materials, and typical etch rates. There are several specialized dry-etching techniques.

Sputter Etching. This process removes material by bombarding the surface with noble gas ions, usually Ar^+. The gas is ionized in the presence of a cathode and an anode (Fig. 28.22). If a silicon wafer is the target, the momentum transfer associated with the bombardment of atoms causes bond breakage and material to be ejected or sputtered. If the silicon chip is the substrate, then the material in the target is deposited onto the silicon, after it has been sputtered by the ionized gas.

The major concerns in sputter etching are:

- The ejected material can be redeposited onto the target, especially with large aspect ratios
- Sputtering can cause damage or excessive erosion of the material
- Sputter etching is not material selective, and because most materials sputter at about the same rate, masking is difficult
- The sputter etching process is slow, with etch rates limited to tens of nm/min
- The photoresist is difficult to remove

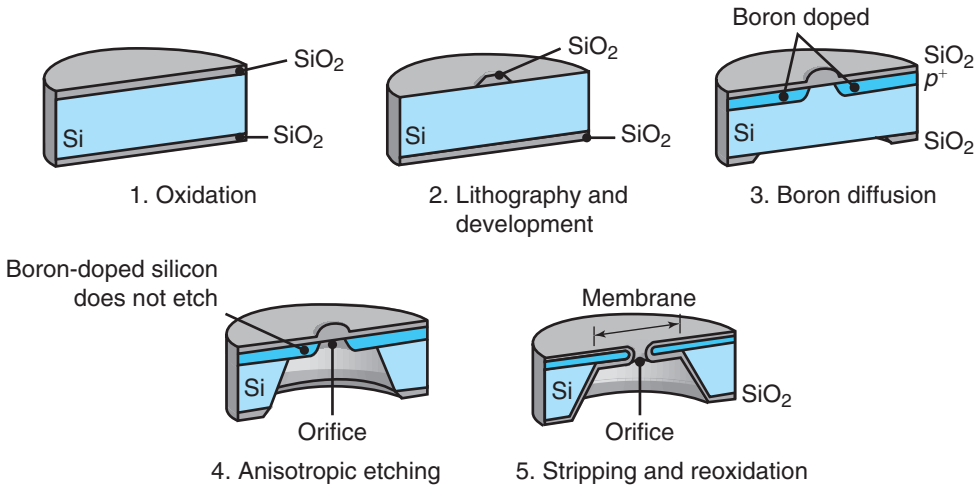


FIGURE 28.21 Application of a boron etch stop and back etching to form a membrane and orifice. *Source:* Based on I. Brodie and J.J. Murray.

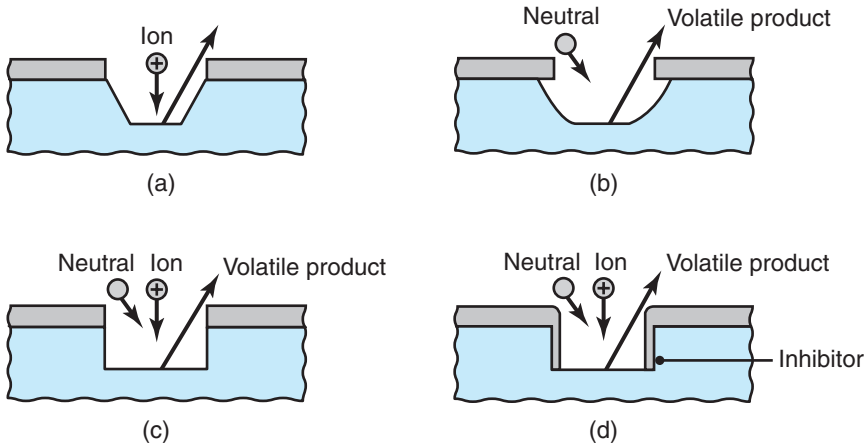


FIGURE 28.22 Machining profiles associated with different dry-etching techniques: (a) sputtering; (b) chemical; (c) ion-enhanced energetic; and (d) ion-enhanced inhibitor. *Source:* After M. Madou.

Reactive Plasma Etching. Also referred to as **dry chemical etching**, this process involves chlorine or fluorine ions (generated by RF excitation), and other molecular species, that diffuse into and chemically react with the substrate. As a result, a volatile compound is formed, which is then removed by a vacuum system. The mechanism of reactive plasma etching is illustrated in Fig. 28.23:

1. A reactive species, such as CF_4, is first produced, and it dissociates upon impact with energetic electrons, to produce fluorine atoms.
2. The reactive species then bombard and diffuse into the surface.
3. The reactive species chemically reacts to form a volatile compound.
4. The reactant desorbs from the surface.
5. It then diffuses into the bulk gas, where it is removed by a vacuum system.

Some reactants polymerize on the surface, thus requiring additional removal of material, either with oxygen in the plasma reactor or by an external *ashing* operation (see *dry stripping* in Section 28.7). The electrical charge of the reactive species is not high enough to cause damage through impact on the surface, hence no sputtering occurs; thus, the etching is isotropic and undercutting of the mask takes place (Fig. 28.19a).

Physical–chemical Etching. Processes such as *reactive ion-beam etching* (RIBE) and *chemically assisted ion-beam etching* (CAIBE) combine the advantages of physical and chemical etching. These processes use a chemically reactive species to drive material removal, but are assisted physically by the impact of ions onto the surface. In RIBE, also known as *deep reactive-ion etching* (DRIE), vertical trenches hundreds of nanometers deep can be produced by periodically interrupting the etching process and depositing a polymer layer, as shown in Fig. 28.23d.

In CAIBE, ion bombardment can assist dry chemical etching by:

- Making the surface more reactive
- Clearing the surface of reaction products and allowing the chemically reactive species access to the cleared areas

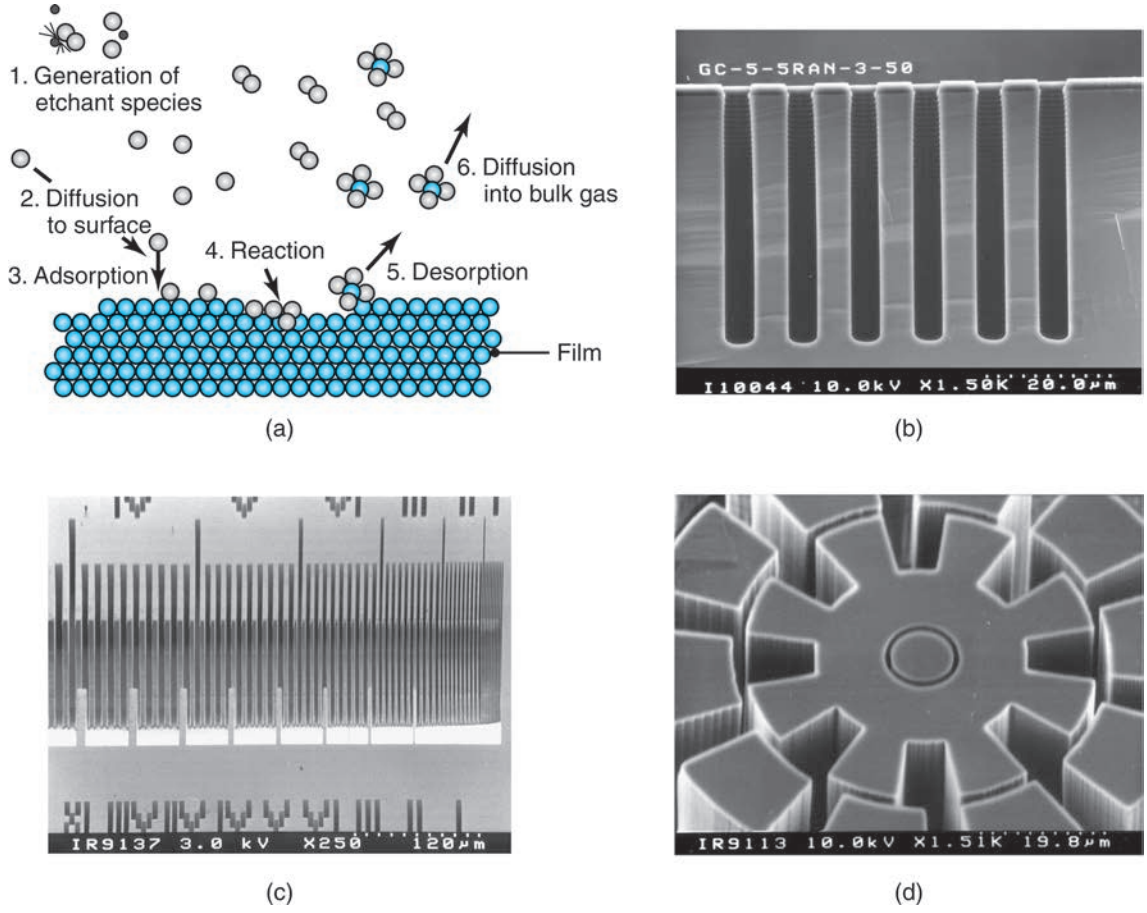


FIGURE 28.23 (a) Schematic illustration of reactive plasma etching. (b) Example of a deep reactive-ion etched trench; note the periodic undercuts, or scallops. (c) Near-vertical sidewalls produced through deep reactive-ion etching (DRIE), an anisotropic-etching process. (d) An example of cryogenic dry etching, showing a 145-μm deep structure etched into silicon with the use of a 2.0-μm-thick oxide masking layer; the substrate temperature was $-140°C$ during etching. *Source:* (a) Based on M. Madou. (b) through (d) After R. Kassing and I.W. Rangelow, University of Kassel, Germany.

- Providing the energy to drive surface chemical reactions; however, the neutral species do most of the etching

Physical–chemical etching is extremely useful because the ion bombardment is directional, so that etching is anisotropic. Also, the ion-bombardment energy is low and does not contribute much to mask removal, thus allowing the generation of near-vertical walls with very large aspect ratios. Since the ion bombardment does not remove material directly, masks can be used.

Cryogenic Dry Etching. This method is used to obtain very deep features with vertical walls. The workpiece is first lowered to cryogenic temperatures, then the CAIBE process takes place. The very low temperatures involved ensure that insufficient energy is available for a surface chemical reaction to take place, unless ion bombardment is normal to the surface. Oblique impacts, such as those occurring on sidewalls in deep crevices, cannot drive the chemical reactions.

Because dry etching is not selective, etch stops cannot be applied directly and dry-etching reactions must be terminated when the target film is removed. This can be done by measuring the wavelength of light being emitted during a reaction; when the target film is removed, the wavelength of light emitted will change and can be detected with proper sensors.

EXAMPLE 28.2 Comparison of Wet and Dry Etching

Consider the case where a ⟨100⟩ wafer (see Fig. 28.5) has an oxide mask placed on it, in order to produce square or rectangular holes. The sides of the square are oriented precisely within the ⟨100⟩ direction of the wafer surface, as shown in Fig. 28.24.

Isotropic etching results in the cavity shown in Fig. 28.24a, and since etching occurs at constant rates in all directions, a rounded cavity that undercuts the mask is produced. An orientation-dependent etchant produces the cavity shown in Fig. 28.24b. Because etching is much faster in the ⟨100⟩ and ⟨110⟩ directions than in the ⟨111⟩ direction, sidewalls defined by the ⟨111⟩ plane are generated. For silicon, these sidewalls are at an angle of 54.74° to the surface.

The effect of a larger mask or shorter etch time is shown in Fig. 28.24c. The resultant pit is defined by ⟨111⟩ sidewalls and by a bottom in the ⟨100⟩ direction parallel to the surface. A rectangular mask and the resulting pit are shown in Fig. 28.24d. Deep reactive-ion etching is depicted in Fig. 28.24e. Note that a polymer layer is deposited periodically onto the hole sidewalls, to allow for deep pockets, but scalloping (greatly exaggerated in the figure) is unavoidable. A hole resulting from CAIBE is shown in Fig. 28.24f.

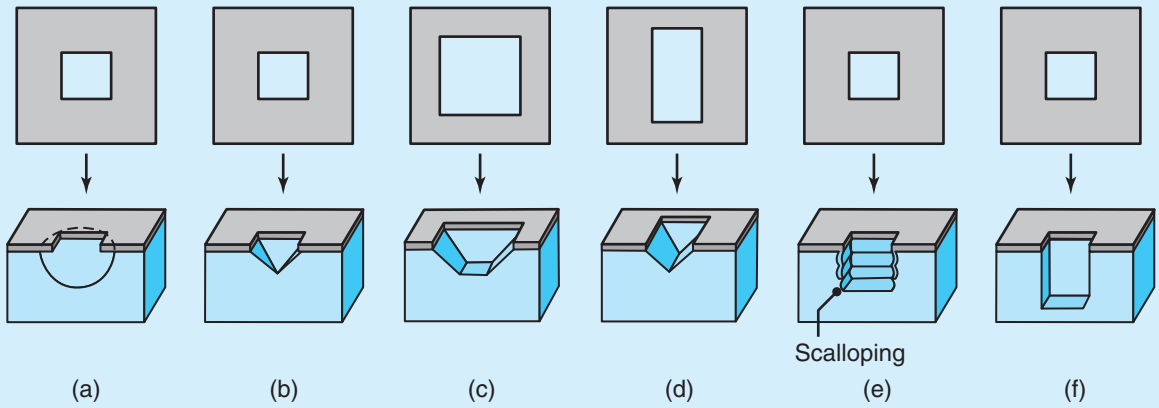


FIGURE 28.24 Holes generated from a square mask in (a) isotropic (wet) etching; (b) orientation-dependent etching (ODE); (c) ODE with a larger hole; (d) ODE of a rectangular hole; (e) deep reactive-ion etching; and (f) vertical etching. *Source:* After M. Madou.

28.9 Diffusion and Ion Implantation

Recall that the operation of microelectronic devices depends on regions that have different doping types and concentrations. The electrical characteristics of these regions are altered through the introduction of dopants into the substrate, by *diffusion* and *ion-implantation processes*. This step in the fabrication sequence is repeated several times, since many different regions of microelectronic devices must be defined.

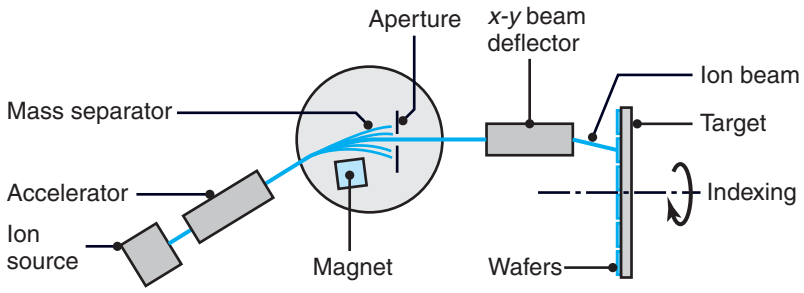


FIGURE 28.25 Schematic illustration of an apparatus for ion implantation.

In the diffusion process, the movement of atoms is a result of thermal excitation. Dopants can be introduced to the substrate surface in the form of a deposited film or the substrate can be placed in a vapor, containing the dopant source. The process takes place at temperatures usually 800°–1200°C (1500°–2200°F). Dopant movement within the substrate is strictly a function of temperature, time, and the diffusion coefficient (or *diffusivity*) of the dopant species, as well as the type and quality of the substrate material.

Because of the nature of diffusion, the dopant concentration is very high at the substrate surface, and drops off sharply away from the surface. To obtain a more uniform concentration within the substrate, the wafer is heated further to drive in the dopants, in a process called **drive-in diffusion**. Diffusion, whether desired or not, is highly isotropic and always occurs at high temperatures, a phenomenon that is taken into account during subsequent processing steps.

Ion implantation is a much more extensive process and requires specialized equipment (Fig. 28.25; see also Section 34.7). Implantation is accomplished by accelerating the ions through a high-voltage field of as much as 1 million electron volts, and then by choosing the desired dopant by means of a mass separator. In a manner similar to that of cathode-ray tubes, the beam is swept across the wafer by sets of deflection plates, thus ensuring uniformity of coverage of the substrate. The whole implantation operation must be performed in a vacuum.

The high-velocity impact of ions on the silicon surface damages the lattice structure and results in lower electron mobilities. Although this condition is undesirable, the damage can be repaired by an annealing step, which involves heating the substrate to relatively low temperatures, usually 400°–800°C (750°–1500°F), for a period of 15–30 min. Annealing provides the energy that the silicon lattice needs to rearrange and mend itself. Another important function of annealing is driving in the implanted dopants. Implantation alone imbeds the dopants less than half a micron below the silicon surface; annealing enables the dopants to diffuse to a more desirable depth of a few microns.

EXAMPLE 28.3 Processing of a *p*-type Region in *n*-type Silicon

Given: It is desired to create a *p*-type region within a sample of *n*-type silicon.

Find: Draw cross-sections of the sample at each processing step in order to accomplish this task.

Solution: Refer to Fig. 28.26. This simple device, known as a *pn-junction diode*, is the foundation for most semiconductor devices.

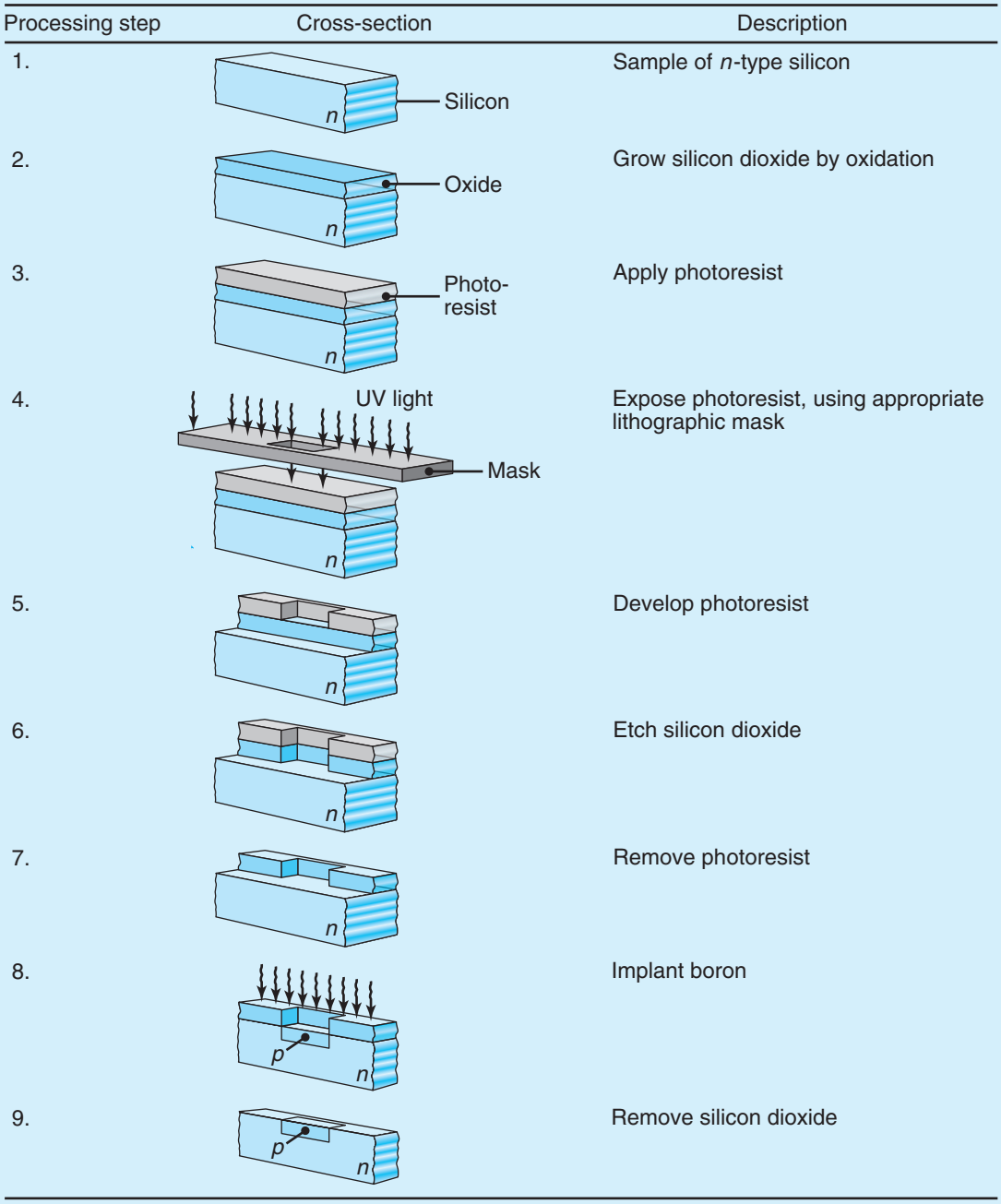
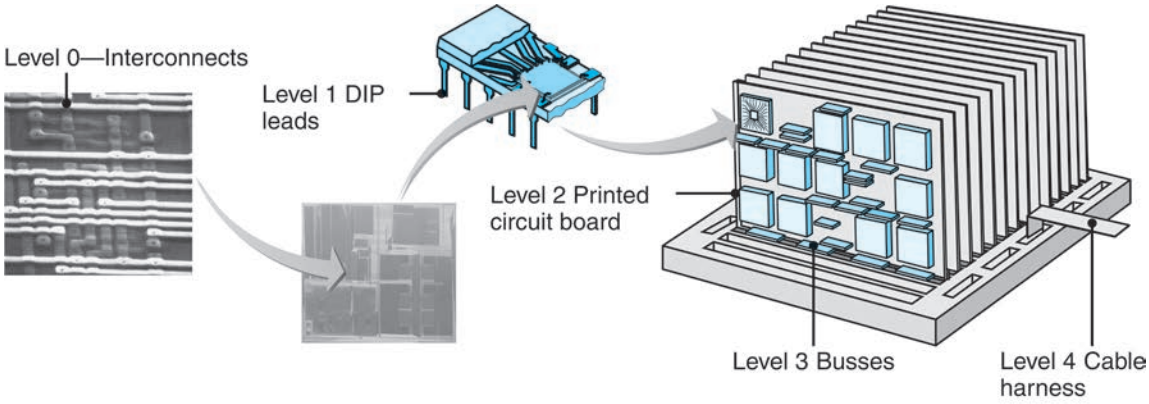


FIGURE 28.26 Fabrication sequence for a *pn*-junction diode.

28.10 Metallization and Testing

The preceding sections focused on device fabrication. Generating a complete and functional integrated circuit requires that these devices be interconnected, and this must take place on a number of levels (Fig. 28.27). **Interconnections** are made using



Level	Element example	Interconnection method
Level 0	Transistor within an IC	IC metallization
Level 1	ICs, other discrete components	Package leads or module interconnections
Level 2	IC packages	Printed circuit board
Level 3	Printed circuit boards	Connectors (busses)
Level 4	Chassis or box	Connectors/cable harnesses
Level 5	System, e.g., computer	

FIGURE 28.27 Connections between elements in the hierarchy for integrated circuits.

metals that exhibit low electrical resistance and good adhesion to dielectric insulator surfaces. Aluminum and aluminum–copper alloys remain the most commonly used materials for this purpose in VLSI technology today.

Because device dimensions continue to shrink, **electromigration** has become more of a concern with aluminum interconnects, a process by which aluminum atoms are moved physically by the impact of drifting electrons under high currents. In extreme cases, electromigration can lead to severed or shorted metal lines. Solutions to the problem include (a) addition of sandwiched metal layers, such as tungsten and titanium and (b) use of pure copper, which displays lower resistivity and has significantly better electromigration performance than aluminum. Metals are deposited by standard deposition techniques, an operation called **metallization**. Modern ICs typically have one to six layers of metallization, each layer of metal being insulated by a dielectric. Interconnection patterns are generated through lithographic and etching processes.

Planarization, that is, producing a planar surface of interlayer dielectrics, is critical to the reduction of metal shorts and the line width variation of the interconnect. A common method used to achieve a planar surface is a uniform oxide-etch process that smoothens out the peaks and valleys of the dielectric layer. Planarizing high-density interconnects has now become the process of **chemical–mechanical polishing** (CMP), described in Section 26.7. This process entails physically polishing the wafer surface, in a manner similar to that by which a disk or belt sander flattens the ridges

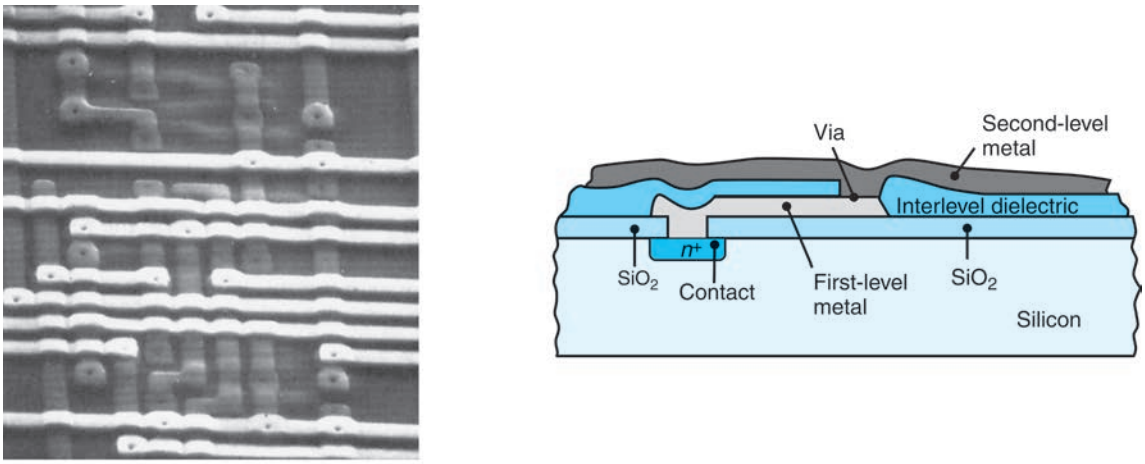


FIGURE 28.28 (a) Scanning-electron microscope (SEM) photograph of a two-level metal interconnect; note the varying surface topography. (b) Schematic illustration of a two-level metal interconnect structure. *Source:* (a) Courtesy of National Semiconductor Corporation. (b) After R.C. Jaeger.

in a piece of wood. A typical CMP process combines an abrasive medium with a polishing compound or slurry, and can polish a wafer to within 0.03 μm (1.2 μin.) of being perfectly flat, with an R_q roughness (see Section 4.4) on the order of 0.1 nm for a new silicon wafer.

Layers of metal are connected together by **vias**, and access to the devices on the substrate is achieved through **contacts** (Fig. 28.28). As devices continue to become smaller and faster, the size and speed of some chips have become limited by the metallization process itself. Wafer processing is completed upon application of a *passivation layer*, usually silicon nitride (Si_3N_4). The silicon nitride acts as a barrier to sodium ions and also provides excellent scratch resistance.

The next step is to test each of the individual circuits on the wafer (Fig. 28.29). Each chip, also known as a **die**, is tested by a computer-controlled probe platform, containing needlelike probes that access the bonding pads on the die. The probes are of two forms:

1. **Test patterns or structures:** The probe measures test structures, often outside of the active dice, placed in the scribe line (the empty space between dies); these probes consist of transistors and interconnect structures that measure various processing parameters, such as resistivity, contact resistance, and electromigration.
2. **Direct probe:** This approach involves 100% testing on the bond pads of each die.

The platform scans across the wafer and uses computer-generated timing waveforms, to test whether each circuit is functioning properly. If a chip is defective, it is marked with a drop of ink. Up to one-third of the cost of a microelectronic circuit can be incurred during this testing.

After the wafer-level testing is completed, back grinding may be done to remove a large amount of the original substrate. The final die thickness depends on the packaging requirement, but anywhere from 25 to 75% of the wafer thickness may be removed. After back grinding, each die is separated from the wafer. *Diamond*

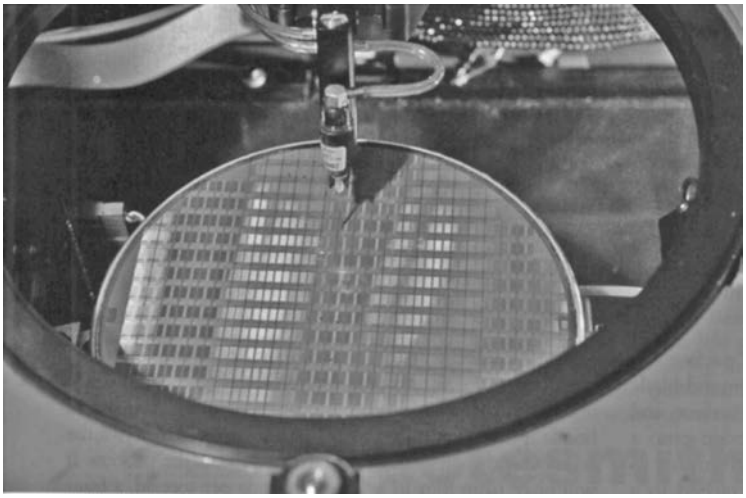


FIGURE 28.29 A probe (top center) checking for defects in a wafer; an ink mark is placed on each defective die. *Source:* Courtesy of Intel Corp.

sawing is a commonly used separation technique and results in very straight edges, with minimal chipping and cracking damage. The chips are then sorted, the functional dice are sent on for packaging, and the inked dice are discarded.

28.11 Wire Bonding and Packaging

The working dice must be attached to a more rugged foundation to ensure reliability. One simple method is to fasten a die to its packaging material with epoxy cement; another method makes use of a eutectic bond, made by heating metal-alloy systems (see Section 4.3). One widely used mixture is 96.4% Au and 3.6% Si, which has a eutectic point at 370°C (700°F).

Once the chip has been attached to its substrate, it must be connected electrically to the package leads. This is accomplished by *wire bonding* very thin (25 μm diameter; 0.001 μin.) gold wires from the package leads to bonding pads, located around the perimeter or down the center of the die (Fig. 28.30). The bonding pads on the die are typically drawn at 75–100 μm (0.003–0.004 in.) per side, and the bond wires are attached by means of thermocompression, ultrasonic, or thermosonic techniques (Fig. 28.31).

The connected circuit is now ready for final packaging. The *packaging* process largely determines the overall cost of each completed IC since the circuits are mass produced on the wafer, but are then packaged individually. Packages are available in a wide variety of styles; the appropriate one must reflect the operating requirements. Consideration of a circuit's package includes the chip size, number of external leads, operating environment, heat dissipation, and power requirements; for example, ICs used for military and industrial applications require packages of particularly high strength, toughness, and resistance to temperature.

Packages are produced from polymers, metals, or ceramics. Metal containers are made from alloys such as Kovar (an iron–cobalt–nickel alloy with a low coefficient of expansion; see Section 3.6), which provide a hermetic seal and good thermal conductivity, but are limited in the number of leads that can be used. Ceramic packages usually are produced from aluminum oxide (Al_2O_3), are hermetic, and

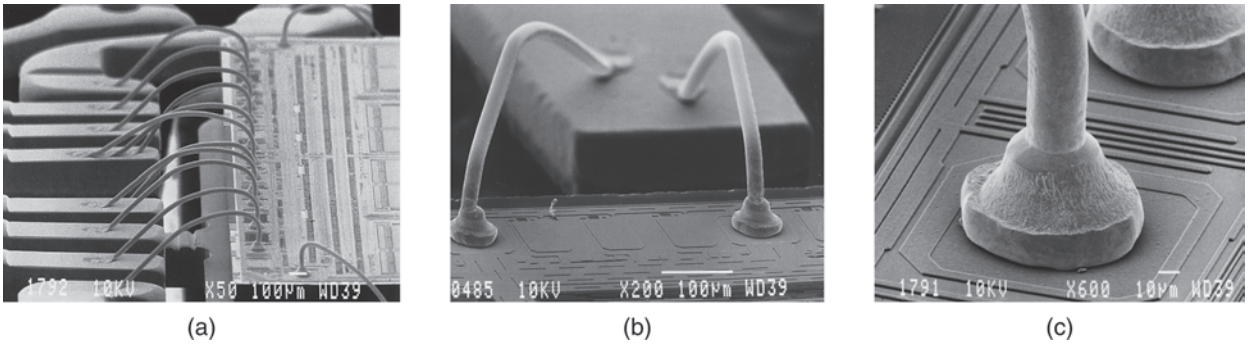


FIGURE 28.30 (a) SEM photograph of wire bonds connecting package leads (left-hand side) to die bonding pads. (b) and (c) Detailed views of (a). *Source:* Courtesy of Micron Technology, Inc.

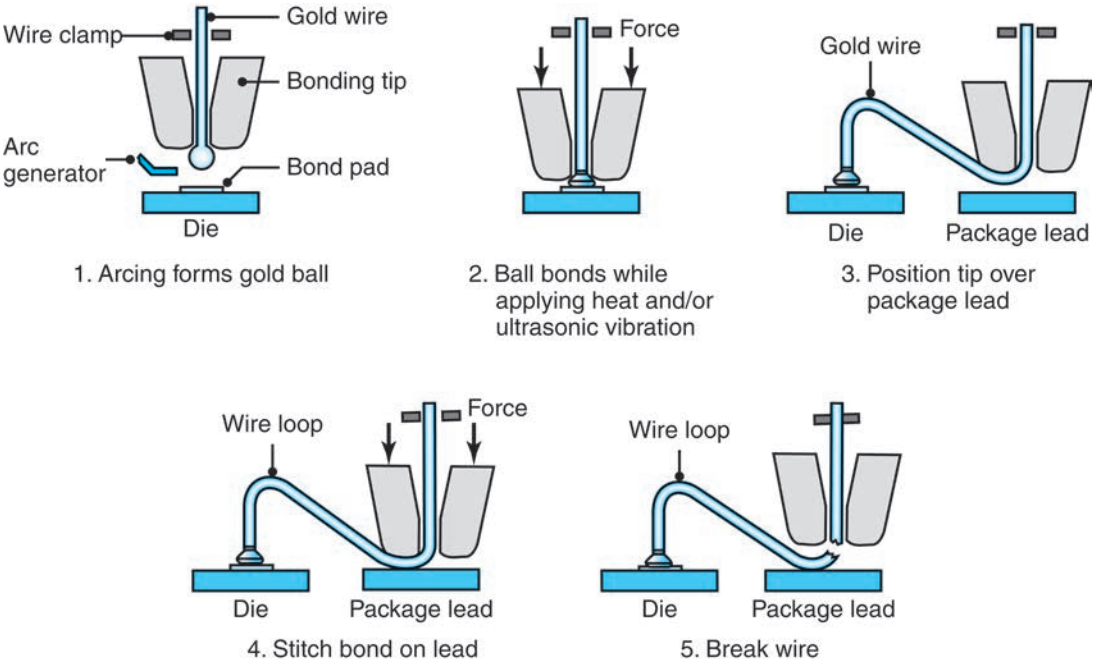


FIGURE 28.31 Schematic illustration of thermosonic welding of gold wires from package leads to bonding pads.

have good thermal conductivity, but have higher lead counts than metal packages; they are also more expensive. Plastic packages are inexpensive and have high lead counts, but they cannot withstand high temperatures and are not hermetic.

An older style of packaging is the **dual-in-line package** (DIP), shown schematically in Fig. 28.32a. Characterized by low cost and ease of handling, DIP packages are made of thermoplastics, epoxies, or ceramics, and can have from 2 to 500 external leads. Ceramic packages are designed for use over a broader temperature range and in high-reliability and military applications, but cost considerably more than plastic packages.

Figure 28.32b shows a *flat ceramic package* in which the package and all of the leads are in the same plane. This package style does not offer the ease of handling

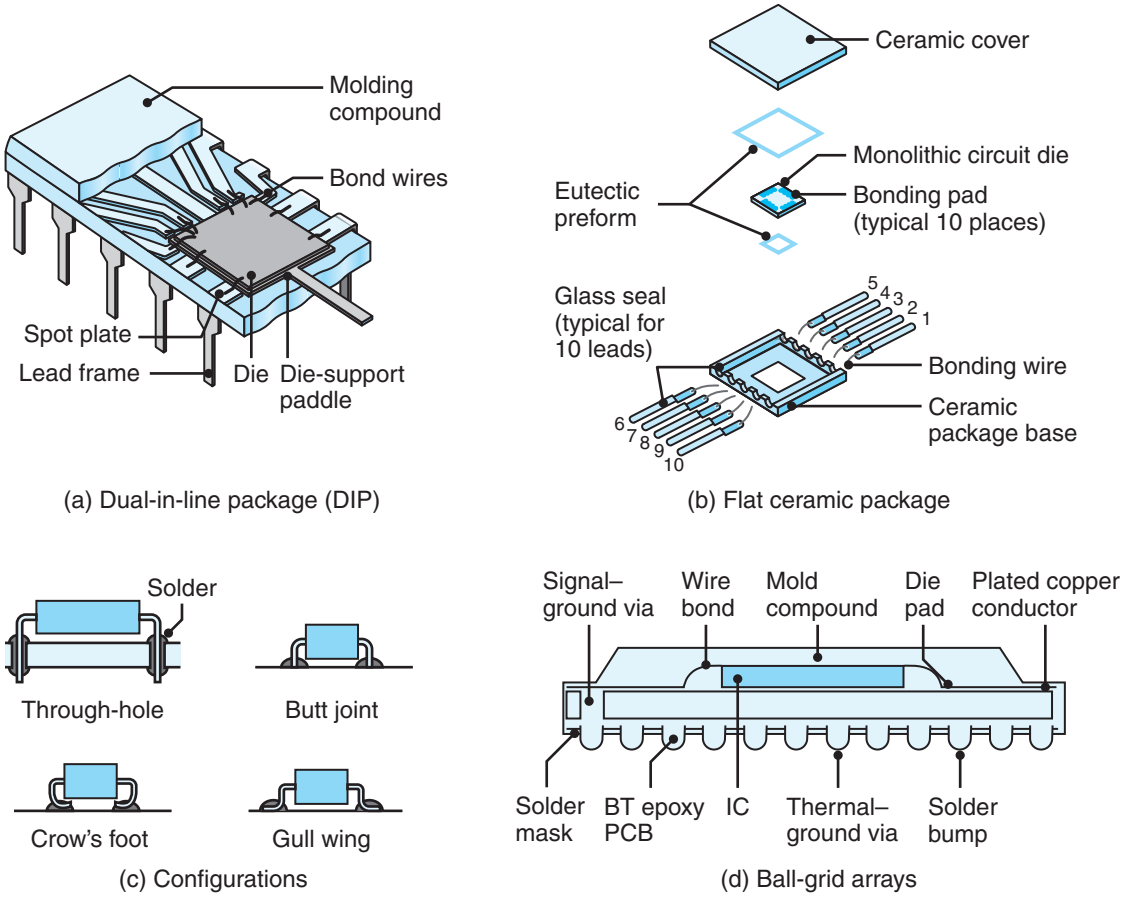


FIGURE 28.32 Schematic illustrations of various IC packages.

or the modular design of the DIP package. For that reason, it usually is affixed permanently to a multiple-level circuit board, in which the low profile of the flat package is necessary.

Surface-mount packages have become common for today's integrated circuits. Some examples are shown in Fig. 28.32c; note that the main difference among them is in the shape of the connectors. The DIP connection to the surface board is *via prongs*, which are inserted into corresponding holes, while a surface mount is soldered onto a specially fabricated pad or *land* (a raised solder platform for interconnections among components in a printed circuit board). Package size and layouts are selected from standard patterns, and usually require adhesive bonding of the package to the board, followed by **wave soldering** of the connections (see Section 32.3.3).

Faster and more versatile chips require increasingly tightly spaced connections. **Pin-grid arrays** (PGAs) use tightly packed pins that connect onto printed circuit boards by way of through-holes. PGAs and other in-line and surface-mount packages are, however, extremely susceptible to plastic deformation of the wires and legs, especially with small-diameter, closely spaced wires. One way of achieving tight packing of connections, and avoiding the difficulties of slender connections, is through **ball-grid arrays** (BGAs), as shown in Fig. 28.32d. This type of array has a solder-plated coating on a number of closely spaced metal balls on the underside of the package. The spacing between the balls can be as small as 50 μm (2000 μin.), but more

commonly it is standardized as 1.0 mm (0.040 in.), 1.27 mm (0.050 in.), or 1.5 mm (0.060 in.).

Although BGAs can be designed with over 1000 connections, this is extremely rare and usually 200–300 connections are sufficient for demanding applications. By using the **reflow soldering** technique (see Section 32.3.2), the solder serves to center the BGA by surface tension, thus resulting in well-defined electrical connections for each ball. After the chip has been sealed in the package, it undergoes final testing. Because one of the main purposes of packaging is isolation from the environment, testing at this stage usually involves such factors as heat, humidity, mechanical shock, corrosion, and vibration. Destructive tests also are performed to determine the effectiveness of sealing.

Chip on Board. *Chip on board* (COB) designs refer to the direct placement of chips onto an adhesive layer on a circuit board. Electrical connections are then made by wire bonding the chips directly to the pads on the circuit board. After wire bonding, final encapsulation with an epoxy is necessary, not only to attach the IC package more securely to the printed circuit board but also to transfer heat evenly during its operation.

Flip-chip on Board. The flip-chip on board (FOB) technology, illustrated in Fig. 28.33, involves the direct placement of a chip with solder bumps onto an array of pads on the circuit board. The main advantage to flip chips, and ball-grid array packages, is that the space around the package, normally reserved for bond pads, is saved; thus, a higher level of miniaturization can be achieved.

System in Package. A trend that allows for more compact devices involves incorporating more than one integrated circuit into a package. Figure 28.34 illustrates the major categories of *system-in-package* (SiP) designs. Although these packages can be integrated horizontally, vertical integration through stacked or embedded structures (Figs. 28.34b and c) has the advantage of achieving performance increases over conventional packages. These benefits have been described as "more than Moore" (see Example 28.1), although SiPs also have other advantages, such as (a) they present reduced size and less noise, (b) cross talk between chips can be better isolated, and (c) individual chips can be upgraded more easily. On the other hand, these packages are more complex, require higher power density and associated heat extraction, and are more expensive than conventional packages.

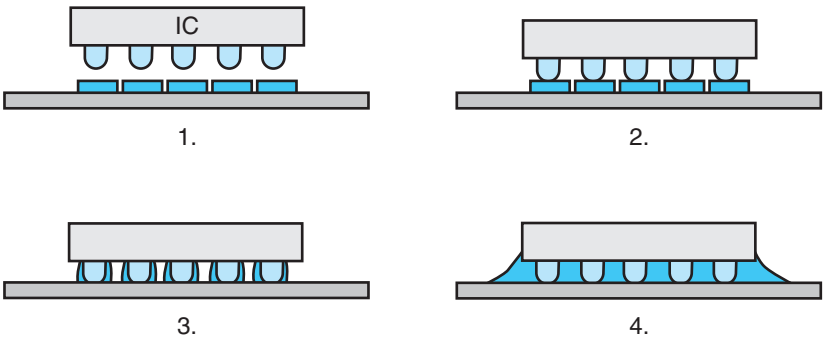


FIGURE 28.33 Illustration of flip-chip technology. Flip-chip package with 1. Solder-plated metal balls and pads on the printed circuit board. 2. Flux application and placement. 3. Reflow soldering. 4. Encapsulation.

SiP packages can, however, be made very simple by incorporating more than one chip inside a single package, as shown in Fig. 28.34a. To preserve area on a circuit board, chips and/or flip chips can be stacked and bonded to a circuit board, to produce three-dimensional integrated circuits, as illustrated in Fig. 28.34b. Here, an

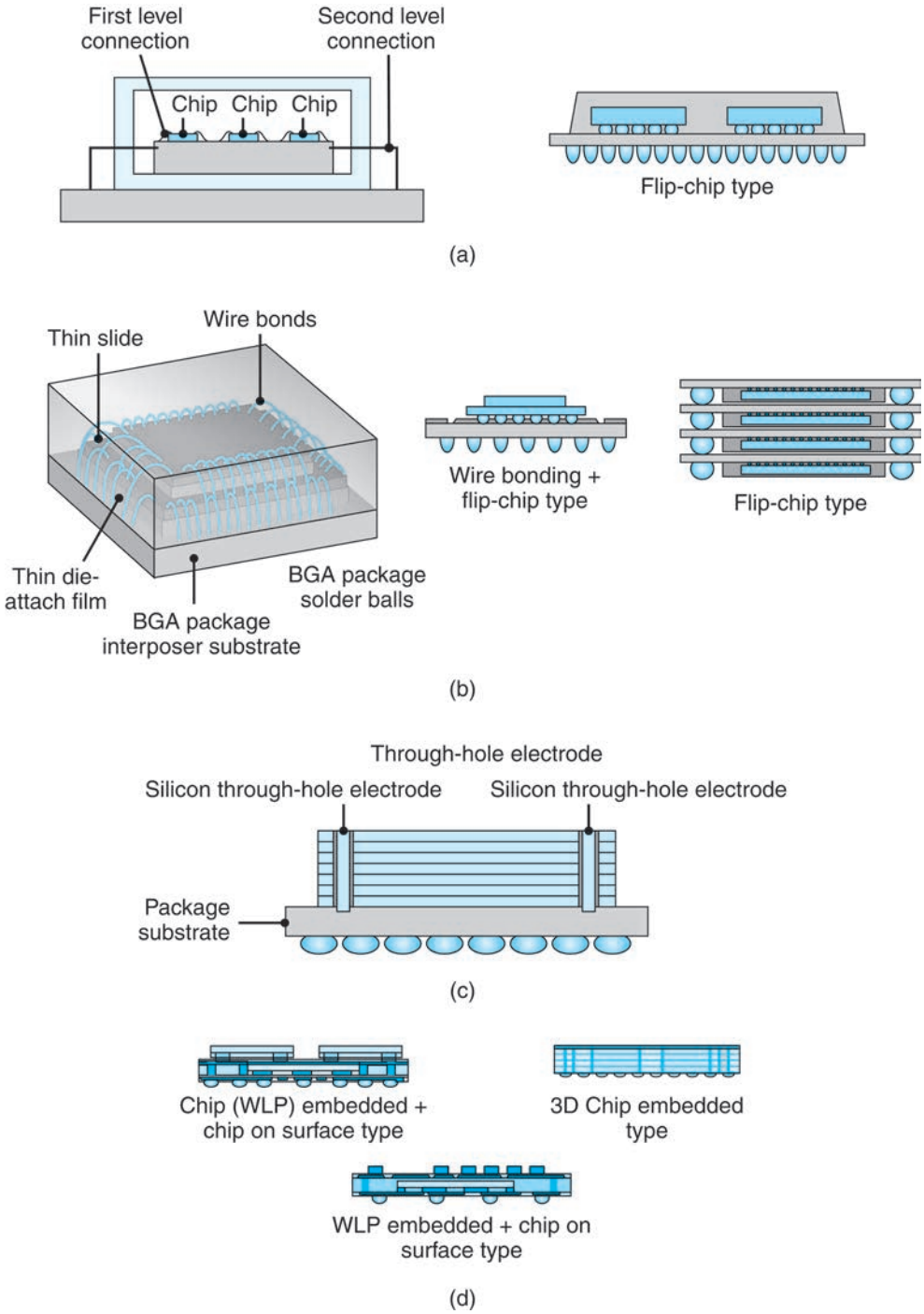


FIGURE 28.34 Major categories of system-in-package designs. (a) Horizontal placement, or multichip modules (MCMs); (b) interposer-type stacked structure; (c) interposerless stacked structure with through-silicon vias; and (d) embedded structure. WLP = wafer level package.

interposing layer, commonly an adhesive, separates the chips and electrically isolates adjacent layers. An alternative is to employ a so-called *interposerless* structure, using **through-silicon vias** (TSVs) instead of wire bonding, to provide electrical connections to all layers. TSVs are sometimes considered a packaging feature, but it has been noted that this is perhaps a case of 3D integration of a wafer, as shown in Fig. 28.34c.

28.12 Yield and Reliability

Yield is defined as the ratio of functional chips to the total number of chips produced. The overall yield of the total IC manufacturing process is the product of the wafer yield, bonding yield, packaging yield, and test yield. This quantity can range from only a few percent for new processes to more than 90% for mature manufacturing lines. Most yield loss occurs during wafer processing, due to its more complex nature. Wafers are commonly separated into regions of good and bad chips. Failures at this stage can arise from point defects (such as oxide pinholes), film contamination, metal particles, and area defects (such as uneven film deposition or nonuniformity of the etch).

Video Solution 28.2 Yield and Reliability of Integrated Circuits

A major concern about completed ICs is their **reliability** and **failure rate**. Since no device has an infinite lifetime, statistical methods are used to characterize the expected lifetimes and failure rates of microelectronic devices. The unit of failure rate is the FIT (failure in time), defined as *one failure per 1 billion device-hours*. However, complete systems may have millions of devices, so the overall failure rate in entire systems is correspondingly higher.

Equally important in failure analysis is determining the *failure mechanism*, that is, the actual process that causes the device to fail. Common failures due to processing involve:

- Diffusion regions: nonuniform current flow and junction breakdown
- Oxide layers: dielectric breakdown and accumulation of surface charge
- Lithography: uneven definition of features and mask misalignment
- Metal layers: poor contact and electromigration, resulting from high current densities
- Other failures, originating in improper chip mounting, poorly formed wire bonds, or loss of the package's hermetic seal

Because device lifetimes are very long, it is impractical to study device failure under normal operating conditions. One method of studying failures efficiently is **accelerated life testing**, which involves accelerating the conditions whose effects cause device breakdown. Cyclic variations in temperature, humidity, voltage, and current are used to stress the components. Chip mounting and packaging are strained by cyclical temperature variations. The statistical data taken from these tests are then used to predict device-failure modes and device life under normal operating conditions.

28.13 Printed Circuit Boards

Packaged ICs seldom are used alone; rather, they usually are combined with other ICs to serve as building blocks of a yet larger system. A *printed circuit board* (PCB) is the substrate for the final interconnections among all of the completed chips, and serves as the communication link between the outside world and the microelectronic circuitry within each packaged IC. In addition to possessing ICs, circuit boards usually contain discrete circuit components (such as resistors and capacitors), which take up too much

"real estate" on the limited silicon surface, have special power-dissipation requirements, or cannot be implemented on a chip. Other common discrete components are inductors (that cannot be integrated onto the silicon surface), high-performance transistors, large capacitors, precision resistors, and crystals (for frequency control).

A PCB is basically a plastic (resin) material, containing several layers of copper foil (Fig. 28.35). *Single-sided PCBs* have copper tracks on only one side of an insulating substrate; *double-sided boards* have copper tracks on both sides. *Multilayered boards* also can be constructed from alternating layers of copper and insulator, but single-sided boards are the simplest form of circuit board.

Double-sided boards usually must have locations where electrical connectivity is established between the features on both sides of the board; this is accomplished with vias, as shown in Fig. 28.35. Multilayered boards can have partial, buried, or through-hole vias to allow for extremely flexible PCBs. Double-sided and multilayered boards are preferable, because IC packages can be bonded to both sides of the board, thus allowing for more compact designs.

The insulating material is usually an epoxy resin, 0.25–3 mm (0.01–0.12 in.) thick, reinforced with an epoxy-glass fiber, and is referred to as E-glass (see Section 9.2.1). The assembly is produced by impregnating sheets of glass fiber with epoxy, and pressing the layers together between hot plates or rolls. The heat and pressure cure the board, resulting in a stiff and strong basis for printed circuit boards.

Boards are first sheared to a desired size, and about 3-mm-diameter locating holes are then drilled or punched into the board's corners, to permit alignment and proper location of the board within the chip-insertion machines. Holes for vias and connections are punched or produced through CNC drilling (Section 37.3); stacks of boards can be drilled simultaneously to increase production rates.

The conductive patterns on circuit boards are defined by lithography, although originally they were produced through screen-printing technologies, hence the term *printed* circuit board or *printed wiring board* (PWB). In the *subtractive method*, a copper foil is bonded to the circuit board. The desired pattern on the board is defined

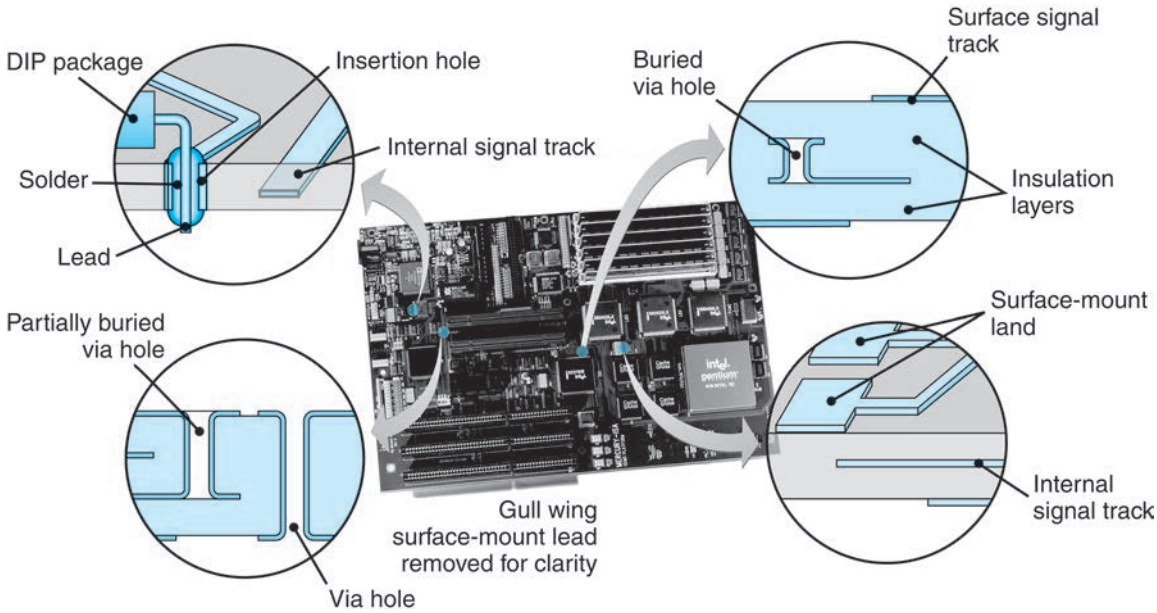


FIGURE 28.35 Printed circuit board structures and design features.

by a positive mask, developed through photolithography, and the remaining copper is removed through wet etching. In the *additive method*, a negative mask is placed directly onto an insulator substrate, to define the desired shape. Electroless plating and electroplating of copper serve to define the connections, tracks, and lands on the circuit board.

The ICs and other discrete components are then soldered to the board. This is the final step in making both the ICs and the microelectronic devices they contain into larger systems, through connections on PCBs. *Wave soldering* and *reflow paste soldering* (see Section 32.3.3 and Example 32.1) are the preferred methods of soldering ICs onto circuit boards.

Basic design considerations in laying out PCBs are:

1. Wave soldering should be used only on one side of the board; thus, all through-hole mounted components should be inserted from the same side of the board. Surface-mount devices placed on the insertion side of the board must be reflow soldered in place; surface-mount devices on the lead side can be wave soldered.
2. To allow good solder flow in wave soldering, IC packages should be laid out carefully on the PCB. Inserting the packages in the same direction is advantageous for automated placing, because random orientations can cause problems in the flow of solder across all of the connections.
3. The spacing of ICs is determined mainly by the need to remove heat during the operation. Sufficient clearance between packages and adjacent boards is thus required to allow forced airflow and heat convection.
4. There should be sufficient space also around each IC package to allow for reworking and repairing without disturbing adjacent devices.

SUMMARY

- The microelectronics industry continues to develop rapidly, and possibilities for new device concepts and circuit designs appear to be endless. The fabrication of microelectronic devices and integrated circuits involves several different types of processes, many of which have been adapted from those of other fields in manufacturing.
- A rough shape of single-crystal silicon is first obtained by the Czochralski process. This shape is ground to a solid cylinder of well-controlled dimensions, and a notch or flat is machined into the cylinder. The cylinder is then sliced into wafers, which are ground on their edges and subjected to chemical–mechanical polishing to complete the wafer.
- After bare wafers have been prepared, they undergo repeated oxidation or film deposition, and lithographic or etching steps to open windows in the oxide layer in order to access the silicon substrate.
- Wet etching is isotropic and relatively fast. Dry etching, using gas plasmas, is anisotropic and allows for more accurate lithography and large-scale integration of integrated circuits.
- After each of the processing cycles is completed, dopants are introduced into various regions of the silicon structure, through diffusion and ion implantation. The devices are then interconnected by multiple metal layers, and the completed circuit is packaged and made accessible through electrical connections.
- The packaged circuit and other discrete devices are then soldered to a printed circuit board for final installation.

KEY TERMS

Accelerated-life testing
Bonding
Chemical–mechanical polishing
Chemical-vapor deposition
Chip
Chip on board
Contacts
Critical dimension
Czochralski process
Die
Diffusion
Dopants
Dry etching
Dry oxidation
Dual-in-line package
Electromigration
Epitaxy
Etching
Evaporation
Failure rate
Film deposition
Flip-chip on board
Gallium arsenide
Immersion lithography
Integrated circuit
Ion implantation
LELE process
Line width
Lithography
Masking
Metal-oxide-semiconductor field-effect transistor
Metallization
Microcontact printing
Micromolding in capillaries
Microtransfer molding
Oxidation
Packaging
Photoresist
Pitch splitting lithography
Planarization
Postbaking
Prebaking
Printed circuit board
Registration
Reliability
Reticle
SCALPEL
Selective oxidation
Selectivity
Semiconductor
Silicon
Soft lithography
Sputtering
Surface-mount package
System in package
Three-dimensional integrated circuits
Very large scale integration
Vias
Wafer
Wafer-scale integration
Wet etching
Wet oxidation
Wire bonding
Yield

BIBLIOGRAPHY

Campbell, S.A., **The Science and Engineering of Microelectronic Fabrication**, 2nd ed., Oxford University Press, 2001.

Doering, R., and Nishi, Y., **Handbook of Semiconductor Manufacturing Technology**, 2nd ed., CRC Press, 2008.

Geng, H., **Semiconductor Manufacturing Handbook**, McGraw-Hill, 2005.

Harper, C.A. (ed.), **Electronic Packaging and Interconnection Handbook**, 4th ed., McGraw-Hill, 2004.

Jackson, K.A., and Schroter, W. (eds.), **Handbook of Semiconductor Technology: Processing of Semiconductors**, 2 vols., Wiley, 2000.

Jaeger, R.C., **Introduction to Microelectronic Fabrication**, 2nd ed., Prentice Hall, 2001.

Khandpur, R.S., **Printed Circuit Boards: Design, Fabrication, and Assembly**, McGraw-Hill, 2005.

May, G.S., and Spanos, C.J., **Fundamentals of Semiconductor Manufacturing and Process Control**, Wiley, 2006.

Nishi, Y., and Doering, R., **Handbook of Semiconductor Manufacturing Technology**, 2nd ed., CRC Press, 2007.

Rizvi, S., **Handbook of Photomask Manufacturing Technology**, CRC Press, 2005.

Schroder, D.K., **Semiconductor Material and Device Characterization**, 3rd ed., Wiley-Interscience, 2006.

Ulrich, R.K., and Brown, W.D. (eds.), **Advanced Electronic Packaging**, Wiley, 2006.

Van Zant, P., **Microchip Fabrication**, 5th ed., McGraw-Hill, 2004.

Yoo, C.S., **Semiconductor Manufacturing Technology**, World Scientific Publishing, 2008.

REVIEW QUESTIONS

28.1 Define the terms wafer, chip, die, device, integrated circuit, line width, registration, surface mount, accelerated-life testing, and yield.

28.2 Why is silicon the semiconductor most used in IC technology?

28.3 What do the abbreviations BJT, MOSFET, VLSI, IC, CVD, CMP, LELE, and DIP stand for?

28.4 Explain the differences between wet and dry oxidation.

28.5 Explain the differences between wet and dry etching.

28.6 What are the purposes of prebaking and postbaking in lithography?

28.7 Define selectivity and isotropy, and their importance in relation to etching.

28.8 Compare the diffusion and ion-implantation processes.

28.9 Explain the difference between evaporation and sputtering.

28.10 What are the levels of interconnection?

28.11 Which is cleaner, a Class-10 or a Class-1 clean room?

28.12 Review Fig. 28.2 and describe the fabrication sequence for integrated circuits.
28.13 What is a via? Why is it important?
28.14 Describe how electrical connections are established between a die and a package.
28.15 What is a flip chip?
28.16 Describe the procedures of image splitting lithography and immersion lithography.

QUALITATIVE PROBLEMS

28.17 Comment on your observations regarding the contents of Fig. V.1.
28.18 Describe how *n*-type and *p*-type dopants differ.
28.19 How is silicon nitride used in oxidation?
28.20 How is epitaxy different from other techniques used for deposition? Explain.
28.21 Note that, in a horizontal epitaxial reactor (see Fig. P28.21), the wafers are placed on a stage (susceptor) that is tilted by a small amount, usually 1°–3°. Explain why this is done.

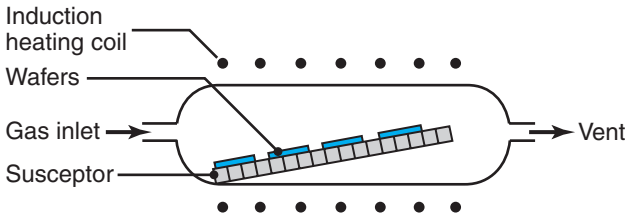


FIGURE P28.21

28.22 The table that follows describes three wafer-manufacturing changes: increasing the wafer diameter, reducing the chip size, and increasing process complexity. Complete the table by filling in "increase," "decrease," or "no change," and indicate the effect that each change would have on the wafer yield and on the overall number of functional chips.

Change	Wafer yield	Number of functional chips
Increase wafer diameter		
Reduce chip size		
Increase process complexity		

28.23 The speed of a transistor is directly proportional to the width of its polysilicon gate; thus, a narrower gate results in a faster transistor and a wider gate in a slower transistor. Knowing that the manufacturing process has a certain variation for the gate width (say, $\pm 0.1\ \mu$m), how would a designer modify the gate size of a critical circuit in order to minimize its variation in speed? Are there any negative effects of this change?
28.24 What is accelerated life testing? Why is it practiced?
28.25 Explain the difference between a die, a chip, and a wafer.
28.26 A common problem in ion implantation is channeling, in which the high-velocity ions travel deep into the material via channels along the crystallographic planes before finally being stopped. How could this effect be avoided? Explain.
28.27 Examine the hole profiles shown in Fig. P28.27 and explain how they might be produced.

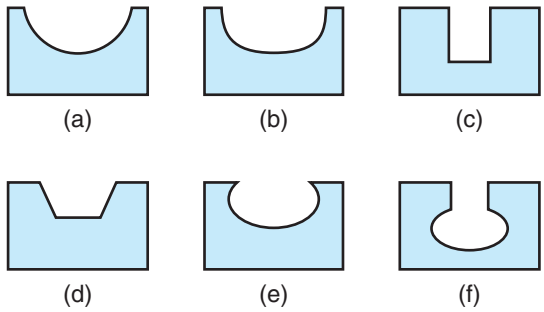


FIGURE P28.27

28.28 Referring to Fig. 28.24, sketch the shape of the holes generated from a circular mask.

QUANTITATIVE PROBLEMS

28.29 A certain wafer manufacturer produces two equal-sized wafers, one containing 500 chips and the other containing 200. After testing, it is observed that 50 chips on each wafer are defective. What are the yields of these two wafers? Can any relationship be drawn between chip size and yield?

28.30 A chlorine-based polysilicon etching process displays a polysilicon-to-resist selectivity of 5:1 and a polysilicon-to-oxide selectivity of 60:1. How much resist and exposed oxide will be consumed in etching 3500 Å of polysilicon? What would the polysilicon-to-oxide selectivity have to be in order to reduce the loss to only 40 Å of exposed oxide?

28.31 During a processing sequence, three silicon-dioxide layers are grown by oxidation to 2500 Å, 4000 Å, and 1500 Å, respectively. How much of the silicon substrate is consumed?

28.32 A certain design rule calls for metal lines to be no less than 2-μm wide. If a 1-μm thick metal layer is to be wet etched, what is the minimum photoresist width allowed (assuming that the wet etching is perfectly isotropic)? What would be the minimum photoresist width if a perfectly anisotropic dry-etching process is used?

28.33 Using Fig. 28.20, obtain mathematical expressions for the etch rate as a function of temperature.

28.34 If a square mask of side length 100 μm is placed on a {100} plane and oriented with a side in the ⟨100⟩ direction, how long will it take to etch a hole 5 μm deep at 80°C using ethylene–diamine/pyrocatechol? Sketch the resulting profile.

28.35 Obtain an expression for the width of the trench bottom as a function of time for the mask shown in Fig. 28.19.

SYNTHESIS, DESIGN, AND PROJECTS

28.36 Describe products that would not exist today without the knowledge and techniques described in this chapter. Explain.

28.37 Inspect various electronic and computer equipment, take them apart as much as you can, and identify components that may have been manufactured by the techniques described in this chapter.

28.38 Describe your understanding of the important features of clean rooms and how they are maintained.

28.39 Make a survey of the necessity for clean rooms in various industries, including the medical, pharmacological, and aerospace industries, and what their requirements are.

28.40 Review the technical literature, and give further details regarding the type and shape of the abrasive wheel used in the wafer-cutting process shown in Step 2 in Fig. 28.2. (See also Chapter 26.)

28.41 List and discuss the technologies that have enabled the manufacture of the products described in this chapter.

28.42 Estimate the time required to etch a spur gear blank from a 75-mm thick slug of silicon.

28.43 Microelectronic devices may be subjected to hostile environments, such as high temperature, humidity, and vibration, as well as physical abuse, such as being dropped onto a hard surface. Describe your thoughts on how you would go about testing these devices for their endurance under these conditions. Are there any industry standards regarding such tests? Explain.

28.44 Review the specific devices, shown in Fig. V.2. Choose any one of these devices, and investigate what they are, what their characteristics are, how they are manufactured, and what their costs are.