(a) Isotropic wet chemical etching.

(b) Anisotropic dry etching with no undercutting.

2.4 **(a)** What type of photoresist must be used with each of the three mask levels (n-diffusion window, contact windows, and metal etch) used to fabricate the diode of Problem 2.2? Assume that the areas shown are dark on the mask (a "light-field mask").

(b) Draw a set of alignment marks suitable for use with the alignment sequence of Problem 2.2(b).

2.5 In Table 2.5, 193-nm DUV lithography is shown as an alternative for 120-nm technology generation. (a) What value of NA would be required based upon Eq. 2.1? (b) What is the depth of field for this system?

2.6 **(a)** What wavelength illumination is required to achieve $F = 0.25$ μm with $NA = 1$ and without the use of phase-shifting masks? What is the value of DF corresponding to your value of NA?

(b) Repeat for $NA = 0.5$.

2.7 Based upon the discussion in Section 2.4, what is the smallest feature size F that can be reproduced with a 193-nm optical source?

2.8 An extreme ultra violet (EUV) lithography source uses a 13-nm exposure wavelength. Based upon the discussion in Section 2.4, what is the smallest feature size F that can be reproduced with this source?

CHAPTER 3

Thermal Oxidation of Silicon

Upon exposure to oxygen, the surface of a silicon wafer oxidizes to form silicon dioxide. This native silicon dioxide film is a high-quality electrical insulator and can be used as a barrier material during impurity deposition. These two properties of silicon dioxide were the primary process factors leading to silicon becoming the dominant material in use today for the fabrication of integrated circuits. This chapter discusses the theory of oxide growth, the oxide growth processes, factors affecting oxide growth rate, impurity redistribution during oxidation, and techniques for selective oxidation of silicon. Methods for determining the thickness of the oxide film are also presented, and the SUPREM process simulation software is introduced.

3.1 THE OXIDATION PROCESS

Thermal oxidation of silicon is easily achieved by heating the wafer to a high temperature, typically 900 to 1200 °C, in an atmosphere containing either pure oxygen or water vapor. Both water vapor and oxygen move (diffuse) easily through silicon dioxide at these high temperatures. (See Fig. 3.1.) Oxygen arriving at the silicon surface can then combine with silicon to form silicon dioxide. The chemical reaction occurring at the silicon surface is

$$Si + O_2 \rightarrow SiO_2 \tag{3.1}$$

for dry oxygen and

$$Si + 2H_2O \rightarrow SiO_2 + 2H_2 \tag{3.2}$$

for water vapor. Silicon is consumed as the oxide grows, and the resulting oxide expands during growth, as shown in Fig. 3.2. The final oxide layer is approximately 54% above the original surface of the silicon and 46% below the original surface.

FIGURE 3.1

Diffusivities of hydrogen, oxygen, sodium, and water vapor in silicon glass. Copyright John Wiley & Sons, Inc. Reprinted with permission from Ref. [4].

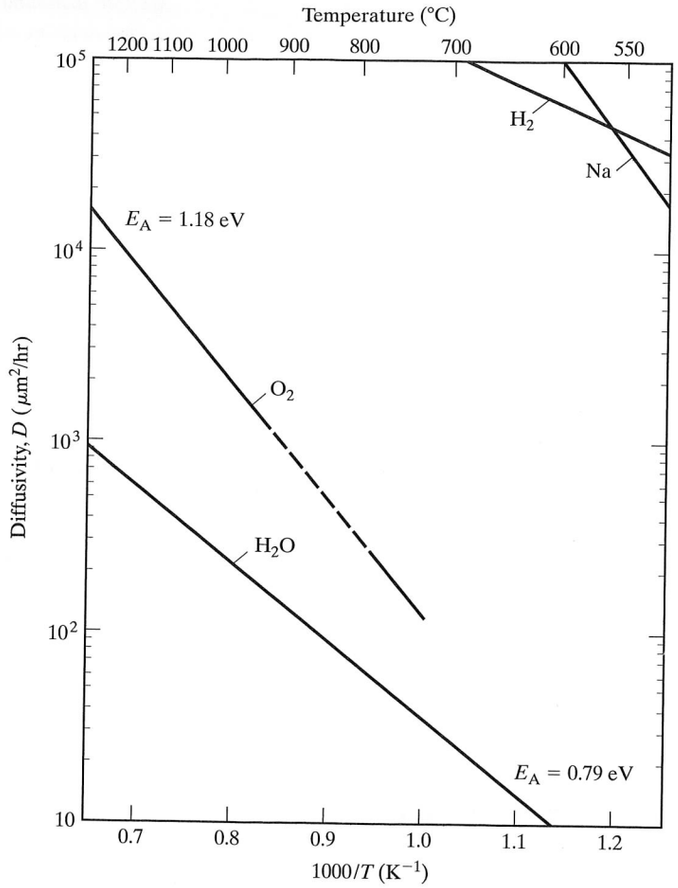
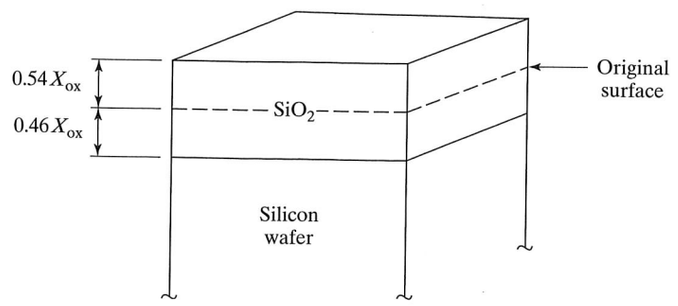


FIGURE 3.2

Formation of a silicon dioxide layer on the surface of a silicon wafer consumes silicon during growth of the layer. The oxide expands to fill a region approximately 54% above and 46% below the original surface of the wafer. The exact percentages depend on the density of the oxide.



3.2 MODELING OXIDATION

In order for oxidation to occur, oxygen must reach the silicon interface. As the oxide grows, oxygen must pass through more and more oxide, and the growth rate decreases as time goes on. A simple model for oxidation can be developed by assuming that

oxygen diffuses through the existing oxide layer. Fick's first law of diffusion states that the particle flow per unit area, J (called particle flux), is directly proportional to the concentration gradient of the particle:

$$J = -D\partial N(x,t)/\partial x, \tag{3.3}$$

where D is the diffusion coefficient and N is the particle concentration. The negative sign indicates that particles tend to move from a region of high concentration to a region of low concentration.

For our case of silicon oxidation, we will make the approximation that the oxygen flux passing through the oxide in Fig. 3.3 is constant everywhere in the oxide. (Oxygen does not accumulate in the oxide.) The oxygen flux J is then given by

$$J = -D(N_i - N_0)/X_o \text{ (number of particles/cm}^2 - \text{sec)}, \tag{3.4}$$

where X_o is the thickness of the oxide at a given time, and N_0 and N_i are the concentrations of the oxidizing species in the oxide at the oxide surface and silicon dioxide-silicon interface, respectively. At the silicon dioxide-silicon interface, we assume that the oxidation rate is proportional to the concentration of the oxidizing species so that the flux at the interface is

$$J = k_s N_i, \tag{3.5}$$

where k_s is called the rate constant for the reaction at the Si–SiO_2 interface. Eliminating N_i using Eqs. (3.4) and (3.5), we find that the flux J becomes

$$J = DN_0/(X_o + D/k_s). \tag{3.6}$$

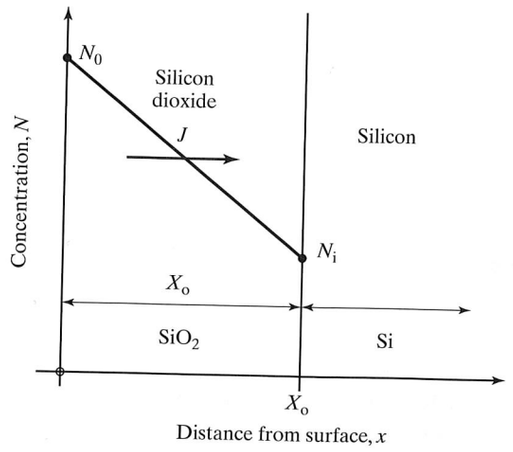


FIGURE 3.3

Model for thermal oxidation of silicon. X_o is the thickness of the silicon dioxide layer at any time t. J is the constant flux of oxygen diffusing through the layer, and N_0 and N_i represent the oxygen concentration at the oxide surface and silicon dioxide–silicon interface, respectively. Note that the oxide growth occurs at the silicon interface.

The rate of change of thickness of the oxide layer with time is then given by the oxidizing flux divided by the number of molecules M of the oxidizing species that are incorporated into a unit volume of the resulting oxide:

$$dX_o/dt = J/M = (DN_0/M)/(X_o + D/k_s). \tag{3.7}$$

This differential equation is easily solved using the boundary condition $X_o(t=0) = X_i$, which yields

$$t = X_o^2/B + X_o/(B/A) - \tau, \tag{3.8}$$

where $A = 2D/k_s$, $B = 2DN_0 / M$, and $\tau = X_i^2/B + X_i/(B/A)$. X_i is the initial thickness of oxide on the wafer, and τ represents the time which would have been required to grow the initial oxide. A thin native oxide layer (10 to 20 Å) is always present on silicon due to atmospheric oxidation, or X_i may represent a thicker oxide grown during previous oxidation steps. Solving Eq. (3.8) for $X_o(t)$ yields

$$X_o(t) = 0.5A\left[\left\{1 + \frac{4B}{A^2}(t + \tau)\right\}^{1/2} - 1\right]. \tag{3.9}$$

For short times with $(t + \tau) \ll A^2/4B$,

$$X_o(t) = (B/A)(t + \tau). \tag{3.10}$$

Oxide growth is proportional to time, and the ratio B/A is called the linear (growth) rate constant. In this region, growth rate is limited by the reaction at the silicon interface.

For long times with $(t + \tau) \gg A^2/4B, t \gg \tau$

$$X_o = \sqrt{Bt}. \tag{3.11}$$

Oxide growth is proportional to the square root of time, and B is called the *parabolic rate constant*. The oxidation rate is diffusion limited in this region.

3.3 FACTORS INFLUENCING OXIDATION RATE

There is good experimental agreement with this simple theory. Figures 3.4 (page 47) and 3.5 (page 48) show experimental data for the parabolic and linear rate constants. The rate-constant data follow straight lines when plotted on a semilogarithmic scale versus reciprocal temperature. This type of behavior occurs in many natural systems and is referred to as an Arrhenius relationship. A mathematical model for this behavior is as follows:

$$D = D_0 \exp(-E_A/kT). \tag{3.12}$$

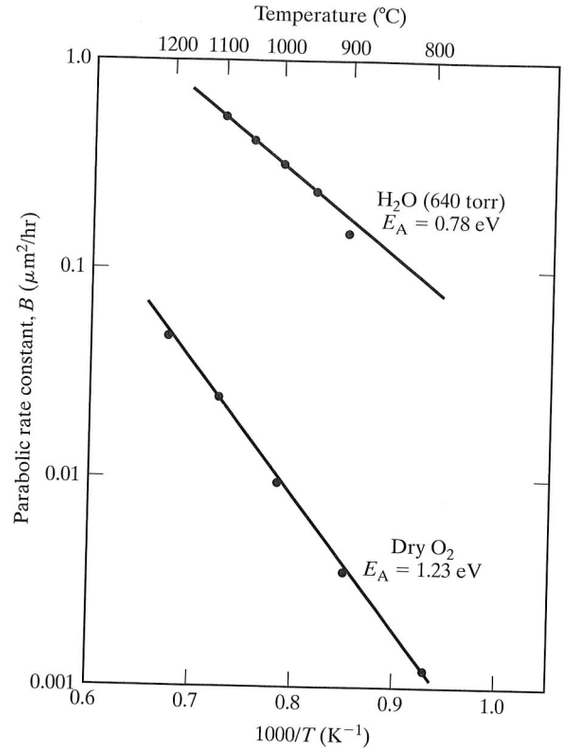


FIGURE 3.4

Dependence of the parabolic rate constant B on temperature for the thermal oxidation of silicon in pyrogenic H_2O (640 torr) or dry O_2. Reprinted by permission of the publisher, The Electrochemical Society, Inc., from Ref. [10].

TABLE 3.1 Values for Coefficient D_0 and Activation Energy E_A for Wet and Dry Oxygen*

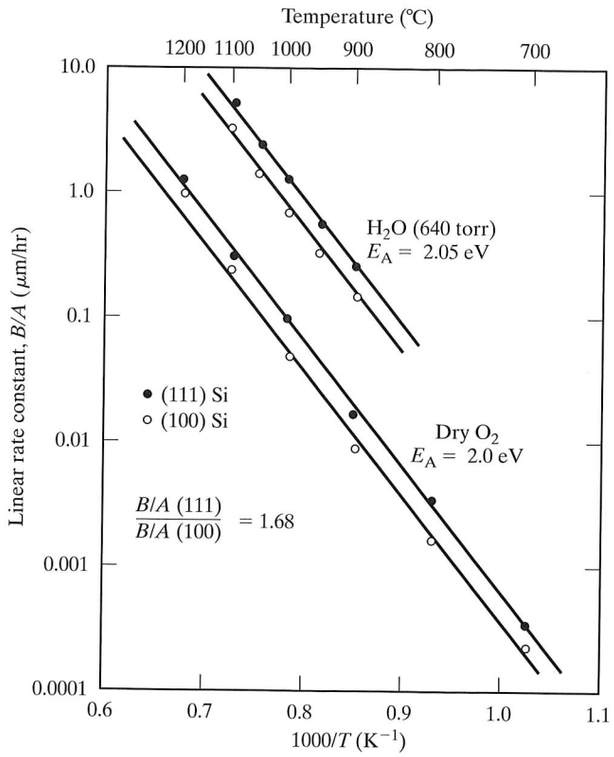
	Wet O_2(X_i = 0 nm)		Dry O_2(X_i = 25 nm)	
	D_0	E_A	D_0	E_A
<100> Silicon				
Linear (B/A)	9.70×10^7 μm/hr	2.05 eV	3.71×10^6 μm/hr	2.00 eV
Parabolic (B)	386 μm^2/hr	0.78 eV	772 μm^2/hr	1.23 eV
<111> Silicon				
Linear (B/A)	1.63×10^8 μm/hr	2.05 eV	6.23×10^6 μm/hr	2.00 eV
Parabolic (B)	386 μm^2/hr	0.78 eV	772 μm^2/hr	1.23 eV

*Data from Ref.[9]

Values for the coefficient D_0 and activation energy E_A for wet and dry oxygen are given in Table 3.1. For wet oxidation, a plot of the experimental data of oxide thickness versus oxidation time is consistent with an initial oxide thickness of approximately zero at $t = 0$. However, a similar plot for dry oxidation yields an initial oxide thickness of 250 Å for temperatures ranging from 800 to 1200 °C. Thus, a nonzero value for τ must be used in Eq. (3.8) for dry oxidation calculations. This large value of X_i indicates that our simple oxidation theory is not quite correct, and the reason for this value of X_i

FIGURE 3.5

Dependence of the linear rate constant B/A on temperature for the thermal oxidation of silicon in pyrogenic H_2O (640 torr) or dry O_2. Reprinted by permission of the publisher, The Electrochemical Society, Inc., from Ref. [10].



is not well understood. Graphs of oxide growth versus time, calculated using the values from Table 3.1, are given in Figs. 3.6 and 3.7 on page 49.

Equation (3.12) indicates the strong dependence of oxide growth on temperature. A number of other factors affect the oxidation rate, including wet and dry oxidation, pressure, crystal orientation, and impurity doping. Water vapor has a much higher solubility than oxygen in silicon dioxide, which accounts for the much higher oxide growth rate in a wet atmosphere. Slower growth results in a denser, higher quality oxide and is usually used for MOS gate oxides. More rapid growth in wet oxygen is used for thicker masking layers.

Equation (3.8) shows that both the linear and parabolic rate constants are proportional to N_0. N_0 is proportional to the partial pressure of the oxidizing species, so pressure can be used to control oxide growth rate. There is great interest in developing low-temperature processes for VLSI fabrication. High pressure is being used to increase oxidation rates at low temperatures. (See Fig. 3.8 on page 51.) In addition, very thin oxides (50 to 200 Å) are required for VLSI, and low-pressure oxidation is being investigated as a means of achieving controlled growth of very thin oxides.

Figures 3.4 through 3.7 also show the dependence of oxidation rate on substrate crystal orientation for the <111> and <100> materials most commonly used in bipolar and MOS processes, respectively. The crystal orientation changes the number of silicon bonds available at the silicon surface, which influences the oxide growth rate and quality of the silicon–silicon dioxide interface.

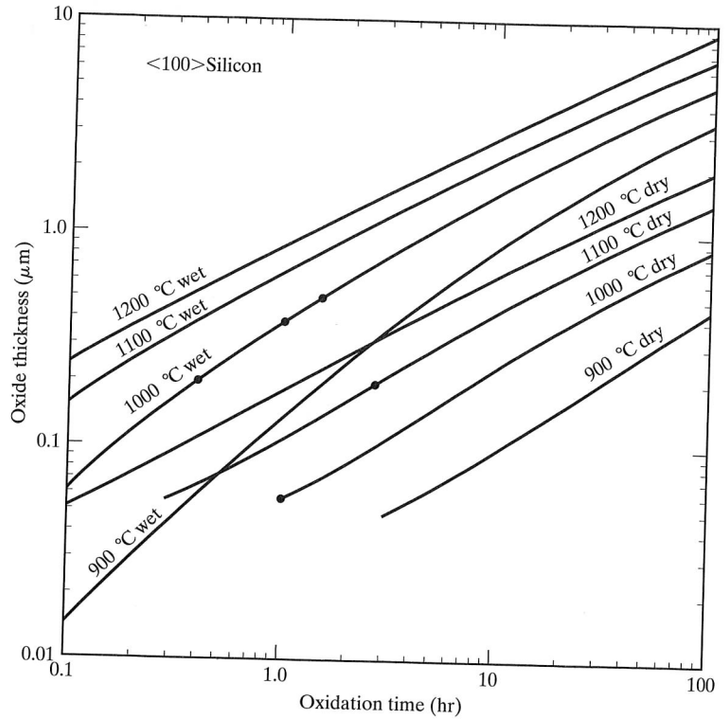


FIGURE 3.6

Wet and dry silicon dioxide growth for <100> silicon calculated using the data from Table 3.1. (The dots represent data used in examples.)

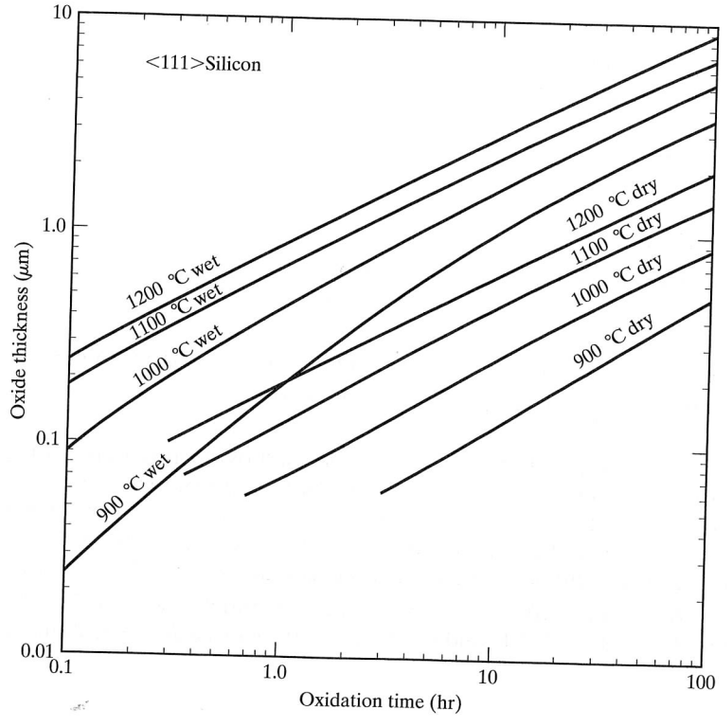
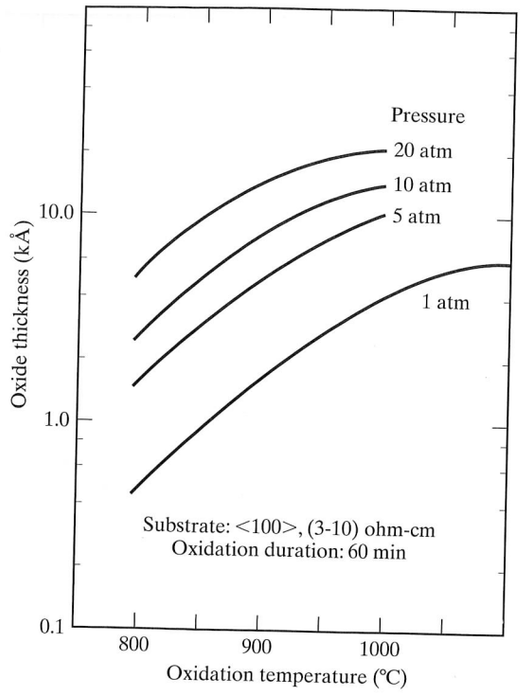


FIGURE 3.7

Wet and dry silicon dioxide growth for <111> silicon calculated using the data from Table 3.1.

FIGURE 3.8

Wet oxide growth at increased pressures. Reprinted with permission of Solid State Technology, published by Technical Publishing, a company of Dun and Bradstreet, from Ref. [12].



Example 3.1

According to Fig. 3.6, a 1-hr oxidation of <100> silicon at 1000 °C in dry oxygen will produce a silicon dioxide film approximately 580 Å (0.058 μm) thick. The same oxidation in wet oxygen will yield a film 3900 Å (0.39 μm) thick.

Example 3.2

A <100> wafer has a 2000-Å oxide on its surface.

(a) How long did it take to grow this oxide at 1100 °C in dry oxygen?

(b) The wafer is put back in the furnace in wet oxygen at 1000 °C. How long will it take to grow an additional 3000 Å of oxide? Solve this problem graphically using Figs. 3.6 and 3.7 as appropriate.

(c) Repeat part (b) using the oxidation theory presented in Eqs. (3.3) through (3.12).

Solution: **(a)** According to Fig. 3.6, it would take 2.8 hr to grow a 0.2-μm oxide in dry oxygen at 1100 °C.

(b) We can solve part (b) graphically using Fig. 3.6. The total oxide at the end of the oxidation would be 0.5 μm. If there were no oxide on the surface, it would take 1.5 hr to grow 0.5 μm. However, there is already a 0.2μm oxide on the surface, and the wafer "thinks" that it has already been in the furnace for 0.4 hr. The time required to grow the additional 0.3 μm of oxide is the difference in these two times: $\Delta t = (1.5 - 0.4)$ hr $= 1.1$ hr.

(c) From Table 3.1, $B = 3.86 \times 10^2 \exp(-0.78/kT)$ μm²/hr and $(B/A) = 0.97 \times 10^8 \exp(-2.05/kT)$ μm/hr. Using T = 1273 K and $k = 8.617 \times 10^{-5}$ eV/Kg, $B = 0.314$ μm²/hr and $(B/A) = 0.738$ μm/hr. Using these values and an initial oxide thickness of 0.2 μm yields a value of 0.398 hr for the effective initial oxidation time τ. Using τ and a final oxide thickness of 0.5 μm yields an oxidation time of 1.08 hr. Note that both the values of t and τ are close to those found in part (b). Of course, the graphical results depend on our ability to interpolate logarithmic scales!

Heavy doping of silicon also changes its oxidation characteristics. Phosphorus doping increases the linear rate constant without altering the parabolic rate constant. Boron doping, on the other hand, increases the parabolic rate constant but has little effect on the linear rate constant. These effects are related to impurity redistribution during oxidation, which is discussed in the next section.

3.4 DOPANT REDISTRIBUTION DURING OXIDATION

During oxidation, the impurity concentration changes in the silicon near the silicon–silicon dioxide interface. Boron and gallium tend to be depleted from the surface, whereas phosphorus, arsenic, and antimony pile up at the surface.

Impurity depletion and pileup depend on both the diffusion coefficient and the segregation coefficient of the impurity in the oxide. The segregation coefficient m is equal to the ratio of the equilibrium concentration of the impurity in silicon to that of the impurity in the oxide. Various possibilities are depicted in Fig. 3.9 on page 52. The value of m for boron is temperature dependent and is less than 0.3 at normal diffusion temperatures. Boron also diffuses slowly through SiO_2. Thus, boron is depleted from the silicon surface and remains in the oxide. (See Fig. 3.9(a).) The presence of hydrogen during oxide growth or impurity diffusion greatly enhances the diffusion of boron through oxide, resulting in enhanced depletion of boron at the silicon surface. (See Fig. 3.9(b).)

The value of m is approximately 10 for phosphorus, antimony, and arsenic. These elements are rejected by the oxide, and they diffuse slowly in the oxide, resulting in pileup at the silicon surface. (See Fig. 3.9(c).) In contrast, gallium has a segregation coefficient of 20, but it diffuses very rapidly through silicon dioxide. This combination causes gallium to deplete at the surface, as shown in Fig. 3.9(d).

The effects of boron depletion and phosphorus and arsenic pileup are particularly important in both bipolar and MOS processing. Process design must take both problems into account, and it is often necessary to add or change processing steps to overcome the effects of these phenomena.

3.5 MASKING PROPERTIES OF SILICON DIOXIDE

One of the most important properties of silicon dioxide is its ability to mask impurities during high-temperature diffusion. The diffusivities of antimony, arsenic, boron, and phosphorus in silicon dioxide are all orders of magnitude smaller than their corresponding values in silicon. Thus, SiO_2 films can be used effectively as a barrier layer to these elements. Relatively deep diffusion can take place in unprotected regions of silicon, whereas no significant impurity penetration will occur in regions covered by silicon dioxide.

FIGURE 3.9

The effects of oxidation on impurity profiles. (a) Slow diffusion in oxide (boron); (b) fast diffusion in oxide (boron with hydrogen ambient); (c) slow diffusion in oxide (phosphorus); (d) fast diffusion in oxide (gallium). C_B is the bulk concentration in the silicon. Copyright John Wiley & Sons, Inc. Reprinted with permission from Ref. [5].

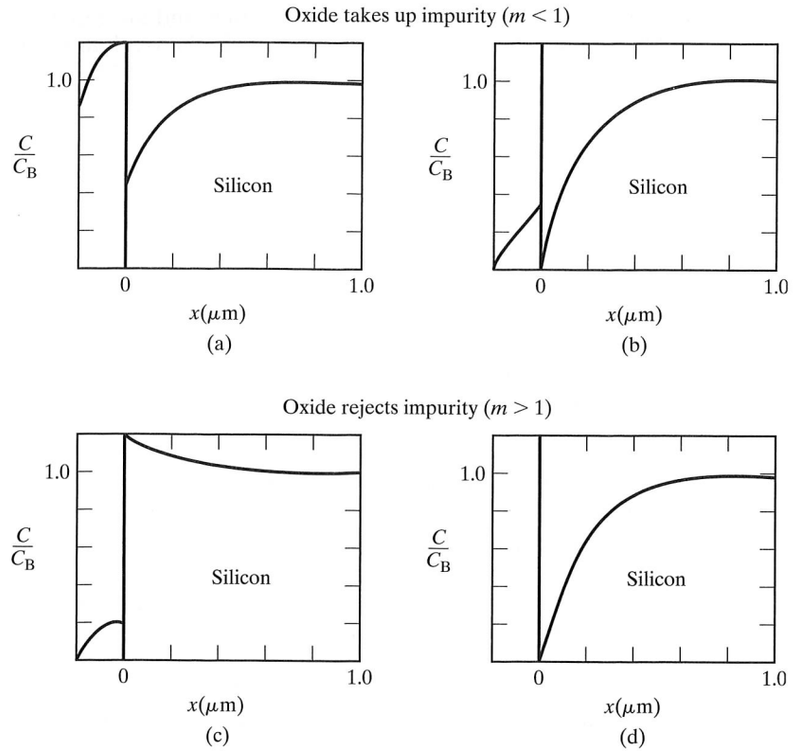


Figure 3.10 gives the SiO_2 thickness required to mask boron and phosphorus diffusions as a function of diffusion time and temperature. Note that silicon dioxide is much more effective in masking boron than in masking phosphorus. Arsenic and antimony diffuse more slowly than phosphorus, so an oxide thick enough to mask phosphorus is also sufficient to mask arsenic and antimony. Masking oxide thicknesses of 0.5 to 1.0 μm are typical in IC processes. The masking oxide would be considered to have failed if the impurity level under the mask were to reach a significant fraction (10%) of the background concentration in the silicon.

The graph for boron is valid for an environment that contains no hydrogen! As mentioned earlier, the presence of hydrogen greatly enhances the boron diffusivity. Wet oxidation releases hydrogen, and care must be taken to avoid boron diffusion in the presence of water vapor.

As mentioned in Section 3.4, gallium diffuses rapidly through SiO_2, as does aluminum, and silicon dioxide cannot be used as a barrier for these elements. However, silicon nitride can be used effectively to prevent diffusion of these impurities.

3.6 TECHNOLOGY OF OXIDATION

Thermal oxidation of silicon is typically carried out in a high-temperature furnace. The furnace walls may be made of quartz, polycrystalline silicon, or silicon carbide and are specially fabricated to prevent sodium contamination during oxidation. The furnaces

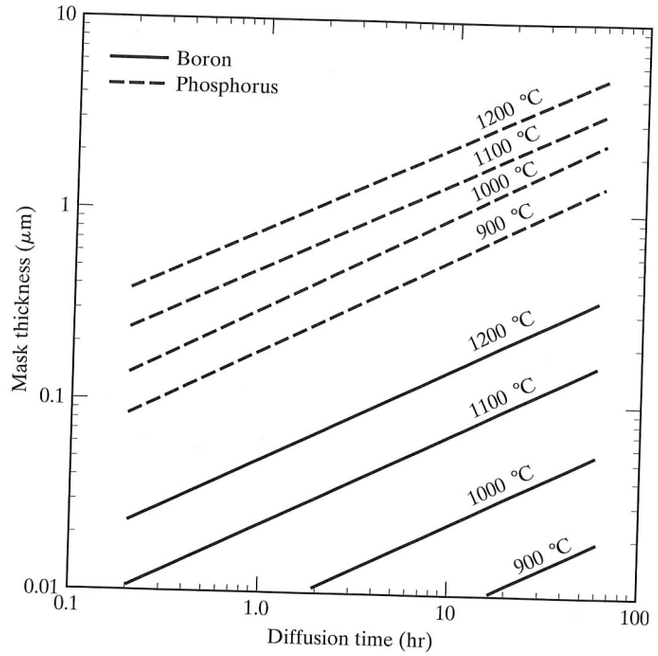


FIGURE 3.10

Thickness of silicon dioxide needed to mask boron and phosphorus diffusions as a function of diffusion time and temperature.

used for many years were manufactured with horizontal oxidation tubes. The wafers are placed upright on edge in a quartz boat and pushed slowly into the furnace. The furnace is maintained at a temperature between 800 and 1200 °C. Three-zone resistance-heated furnaces maintain the temperature within a fraction of a degree over a distance of 0.5 m in the center zone. Vertical furnaces are now commonly used for processing wafers with diameters of 150 mm and larger. A photograph of horizontal and vertical furnaces used for oxidation and diffusion appears in Fig. 3.11.

The furnace is continually purged with an inert gas such as nitrogen prior to oxidation. Oxidation begins by introducing the oxidizing species into the furnace in gaseous form. Extremely high-purity oxygen is available and is used for dry oxidation. Water vapor may be introduced by passing oxygen through a bubbler containing deionized water heated to 95 °C. The oxygen serves as a transport gas to carry the water vapor into the furnace. High-purity water vapor can also be obtained by burning hydrogen and oxygen in the furnace tube. Steam is not often used because it tends to pit the silicon surface.

3.7 OXIDE QUALITY

Wet oxidation is used to grow relatively thick oxides used for masking. An oxidation growth cycle usually consists of a sequence of dry-wet-dry oxidations. Most of the oxide is grown during the wet oxidation phase, since the growth rate is much higher in the presence of water. Dry oxidation results in a higher density oxide than that achieved with wet oxidation. Higher density in turn results in a higher breakdown voltage (5 to 10 MV/cm). To maintain good process control, the thin gate oxides (<1000 Å) of MOS devices are usually formed using dry oxidation.

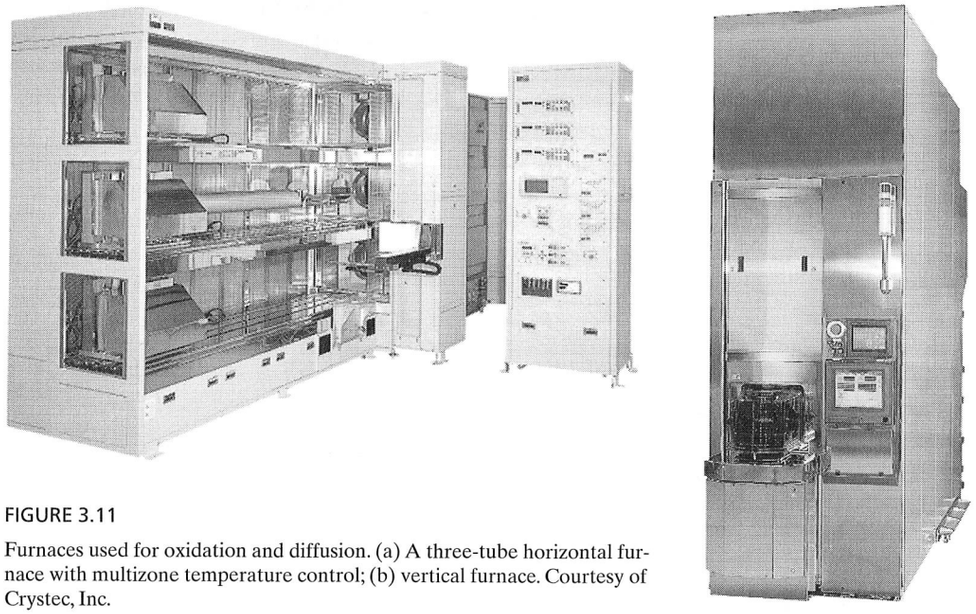


FIGURE 3.11

Furnaces used for oxidation and diffusion. (a) A three-tube horizontal furnace with multizone temperature control; (b) vertical furnace. Courtesy of Crystec, Inc.

MOS devices are usually fabricated on wafers having a <100> surface orientation. The <100> orientation results in the smallest number of unsatisfied silicon bonds at the Si–SiO_2 interface, and the choice of the <100> orientation yields the lowest number of interface traps.*

Sodium ions are highly mobile in SiO_2 films (see Fig. 3.1), and contamination of MOS gate oxides was a difficult problem to overcome in the early days of the integrated-circuit industry.* Bipolar devices are much more tolerant of oxide contamination than MOS devices, and this was a primary factor in the early dominance of bipolar integrated circuits.

Sodium-ion contamination results in mobile positive charge in the oxide. In addition, a substantial level of positive fixed oxide exists at the Si–SiO_2 interface.* These charge centers attract electrons to the surface of MOS transistors, resulting in a negative shift in the threshold voltage of the MOS devices. NMOS devices tend to become depletion-mode devices. PMOS devices remain enhancement-mode devices, but have more negative threshold voltages. The first successful MOS processes were therefore PMOS processes. As the industry was able to improve overall oxide quality, NMOS processes became dominant because of the mobility advantage of electrons over holes.

It was discovered that the effects of sodium contamination can be greatly reduced by adding chlorine during oxidation. Chlorine is incorporated into the oxide and immobilizes the sodium ions. A small amount (6% or less) of anhydrous HCl can be added to the oxidizing gas. Gaseous chlorine, oxygen, or nitrogen can also be bubbled through trichloroethylene (C_2HCl_3). It should also be noted that the presence of chlorine during dry oxidation results in an increase in both the linear and parabolic rate constants.

*See Volume IV in the Modular Series on Solid State Devices, *Field Effect Devices*, Chapter 4, for an excellent discussion of oxide quality.

3.8 SELECTIVE OXIDATION AND SHALLOW TRENCH FORMATION

The oxidation processes described above generally form an oxide film over the complete surface of the silicon wafer. The ability to selectively oxidize the silicon surface has become very important in high-density bipolar and MOS processes. Selective oxidation processes result in improved device packing density and more planar final structures. The techniques utilized for localized oxidation of silicon are generally referred to as LOCOS processes.

Oxygen and water vapor do not diffuse well through silicon nitride. Figure 3.12 shows a MOS process using selective oxidation in which silicon nitride is used as an oxidation barrier. A thin layer (10 to 20 nm) of silicon dioxide is first grown on the wafer to protect the silicon surface. Next, a layer of silicon nitride is deposited over the surface and patterned using photolithography. The wafer then goes through a thermal oxidation step. Oxide grows wherever the wafer is not protected by silicon nitride. This process results in the so-called *semirecessed oxide structure.*

Some oxide growth occurs under the edges of the nitride and causes the nitride to bend up at the edges of the masked area. The penetration of the oxide underneath the nitride results in a "bird's beak" structure. Formation of the bird's beak in Fig. 3.12 leads to loss of geometry control in VLSI structures, so minimization of the bird's beak phenomenon is an important goal in VLSI process design.

A *fully recessed oxide* can be formed by etching the silicon prior to oxidation. This process can yield a very planar surface after the removal of the nitride mask. However, subsequent processing reduces the advantage of this process over the semirecessed version, and most processes today use some form of semirecessed oxidation.

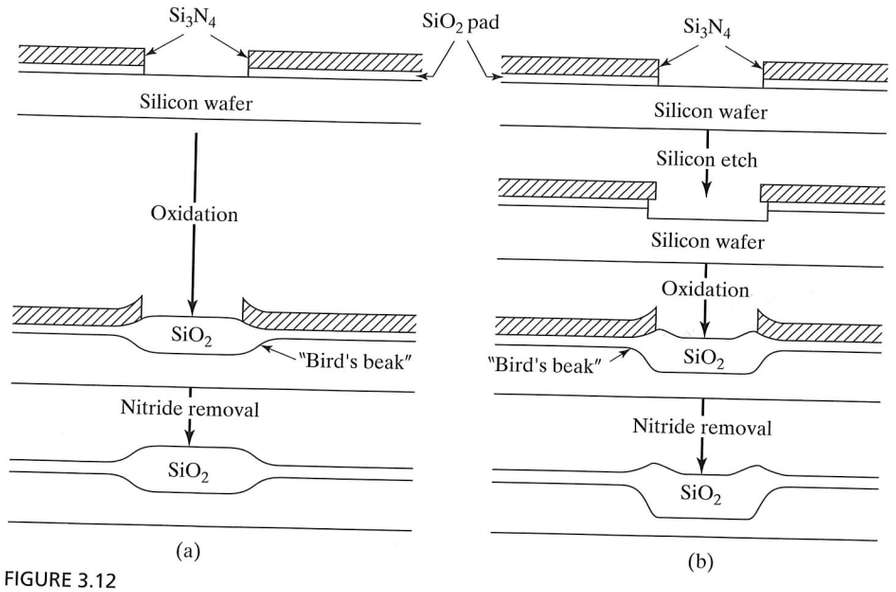


FIGURE 3.12

Cross section depicting process sequence for local oxidation of silicon (LOCOS): (a) semirecessed and (b) fully recessed structures.

3.8.1 Trench Isolation

Some form of shallow or deep refilled trench isolation is utilized in most of today's advanced MOS and bipolar processes. Figure 3.13(a) [18] on page 57 depicts formation of deep trenches filled with polysilicon in combination with a LOCOS field oxidation. A thin oxide pad is grown on silicon followed by deposition of a silicon nitride layer. Lithography is used to define openings in the nitride where trenches will be formed. The trenches are etched using reactive-ion etching and can be quite deep with very high aspect ratios. The surface of the trench is passivated with a thin layer of thermally grown oxide, and then the trench is refilled with deposited polysilicon. The final structure is produced by etching back any excess polysilicon, using a lithography step to remove the nitride layer where oxidation is desired, and growing the semirecessed oxide layer. The polysilicon may be doped, and similar structures are used to form trench capacitors for use as storage elements in some DRAM technologies [19]. The capacitor is formed between the polysilicon and the substrate, and the trenches may be as much as 5–10 μm in depth.*

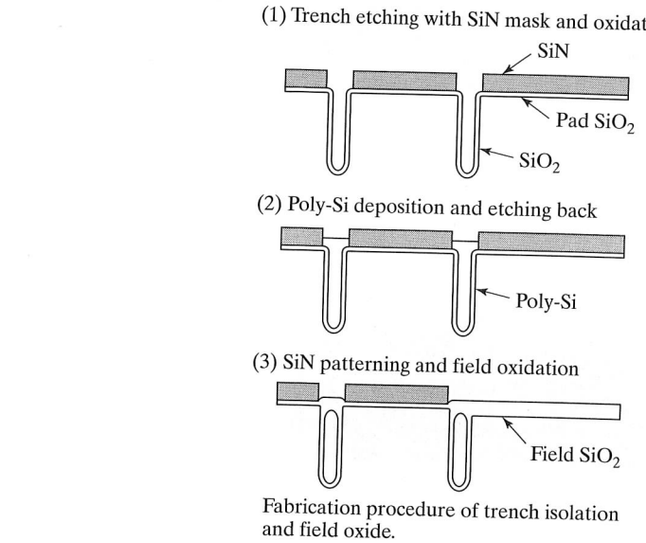
Shallow trench isolation is used to provide isolation between transistors in most MOS and bipolar technologies (see Chapters 9 and 10) with feature sizes below 0.5 μm. Figure 3.13(b) depicts one possible process flow for forming shallow trenches [20]. A shallow trench with tapered sidewalls is etched in the silicon following patterning of the nitride layer. The pad oxide may be etched away slightly to round the corners of the final structure. A thin oxide layer is grown as a liner on the trench walls, and the trench is then filled with an oxide deposited using decomposition of TEOS or via a high-density plasma deposition.* A process called chemical mechanical polishing (CMP), discussed in more detail in the next section, is used to remove the excess oxide and create a highly planar surface. In the final step, the nitride may be removed, leaving the shallow trench isolation between two silicon regions. Figure 3.14 on page 58 shows the use of both deep and shallow trench isolation in an advanced bipolar process.

3.8.2 Chemical Mechanical Polishing (CMP)

CMP [21–24] was introduced into fabrication processing during the early 1990s and is now widely used in both bipolar and MOS processes to achieve the highly planar topologies required in deep submicron lithography. A conceptual diagram of a CMP apparatus is given in Fig. 3.15 on page 58. The wafer is mounted on a carrier and is brought into contact with a polishing pad mounted on a rotating platten. A liquid slurry is continuously dispensed onto the surface of the polishing pad. A combination of the vertical force between the wafer and the abrasive pad as well as the chemical action of the slurry is used to polish the surface to a highly planar state. In the case of formation of the shallow trenches, the nitride layer serves as a polishing stop. Polishing terminates when the nitride layer is fully exposed.

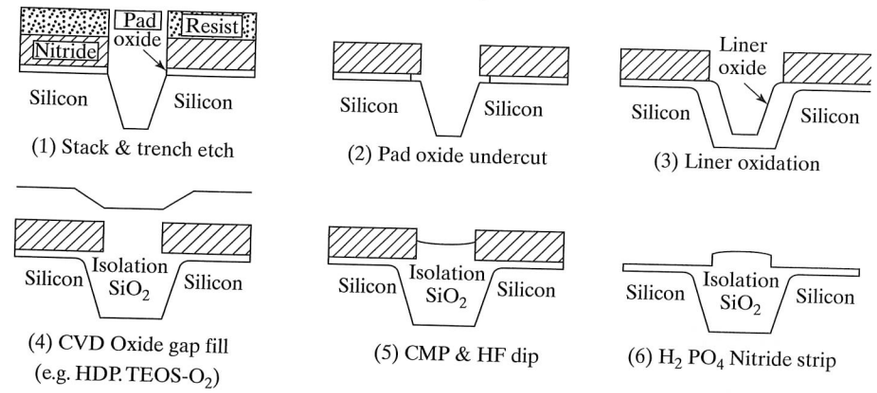
Actual structures differ slightly from the idealized planar surface, as depicted in Fig. 3.16 on page 59. A slight amount of "dishing" of the oxide and corner rounding can both occur due to polishing rate differences between the various materials. Some erosion of the nitride layer may also occur prior to process endpoint detection.

* See Chapter 6.



Fabrication procedure of trench isolation and field oxide.

(a) Deep-trench process



(b) Steps in a typical STI process flow

FIGURE 3.13

Trench isolation structures. (a) Deep trench isolation. Copyright 1996 IEEE. Reprinted with permission from Ref. [18]. (b) Shallow trench isolation. Copyright 1998 IEEE. Reprinted with permission from Ref. [20].

3.9 OXIDE THICKNESS CHARACTERIZATION

One of the simplest methods for determining the thickness of an oxide is to compare the color of the wafer with the reference color chart in Table 3.2 on page 60. When a wafer is illuminated with white light perpendicular to the surface, the light penetrates the oxide film and is reflected by the underlying silicon wafer. Constructive interference causes enhancement of a certain wavelength in the reflected light, and the color

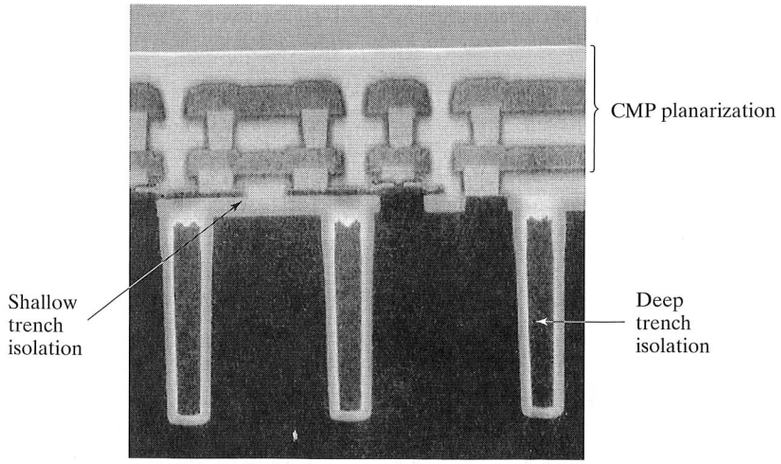


FIGURE 3.14

Microphotograph of actual deep and shallow trench applied to SiGe HBT technology. Copyright 1998 IEEE. Reprinted with permission from Ref. [31].

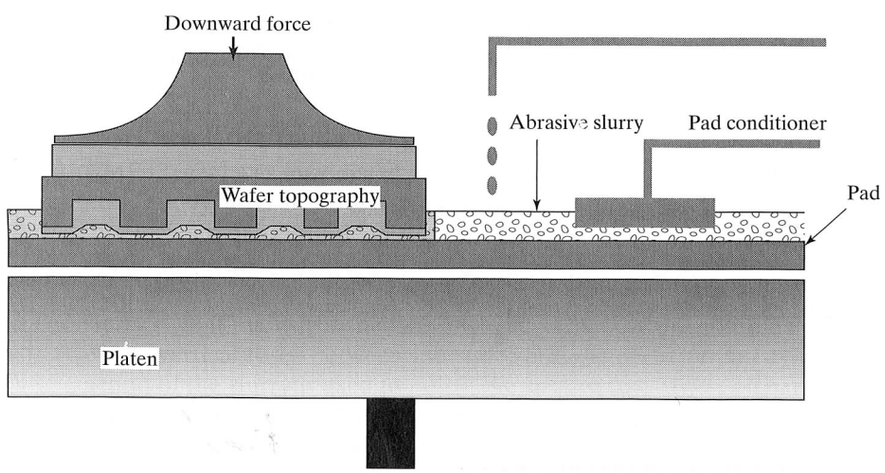


FIGURE 3.15

Chemical mechanical polishing technique.

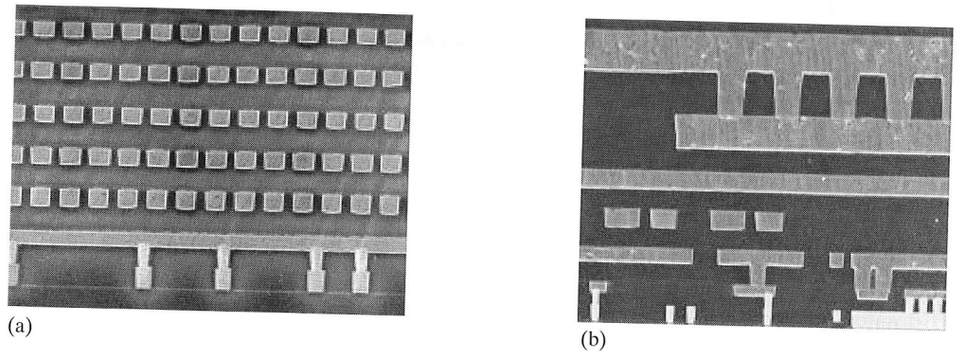


FIGURE 3.16

Multilevel metallization fabricated with chemical mechanical polishing.
(a) SEM of 6-level thin-wire copper. First-level copper is connected by Tungsten studs to Tungsten local interconnect. (b) SEM of 6-level copper with low RC metallization on levels 5 and 6. Copyright 1997 IEEE. Reprinted with permission from Ref. [24].

of the wafer corresponds to the enhanced wavelength. Constructive interference occurs when the path length in the oxide ($2X_o$) is equal to an integer multiple of one wavelength of light in the oxide.

$$2X_o = k\lambda/n, \tag{3.13}$$

where the number k is any integer greater than zero and n is the refractive index of the oxide ($n = 1.46$ for SiO_2).

As an example, a wafer with a 5000-Å silicon dioxide layer will appear blue green. Color-chart comparisons are quite subjective, and the colors vary periodically with thickness. In addition, care must be exercised to determine the color from a position perpendicular to the wafer. The color chart (Table 3.2) is only valid for vertical illumination with fluorescent light.

Accurate thickness measurement can be achieved with an instrument called an *ellipsometer*, and this instrument is often used to make an accurate reference color chart. Polarized monochromatic light is used to illuminate the wafer at an angle to the surface. Light is reflected from both the oxide and silicon surfaces. The differences in polarization are measured, and the oxide thickness can then be calculated [17].

A mechanical surface profiler can also be used to measure film thickness. The oxide is partially etched from the surface of a test wafer to expose a step between the wafer and oxide surfaces. A stylus is mechanically scanned over the surface of the wafer, and thickness variations are recorded by a computer. Films ranging from less than 0.01 μm to more than 5 μm can be measured with this instrument.

TABLE 3.2 Color Chart for Thermally Grown SiO_2 Films Observed Perpendicularly Under Daylight Fluorescent Lighting. Copyright 1964 by International Business Machines Corporation; reprinted with permission from Ref. [11].

Film Thickness (μm)	Color and Comments	Film Thickness (μm)	Color and Comments
0.05	Tan	0.63	Violet red
0.07	Brown	0.68	"Bluish" (not blue but borderline
0.10	Dark violet to red violet		between violet and blue green; appears
0.12	Royal blue		more like a mixture between violet
0.15	Light blue to metallic blue		red and blue green and looks grayish)
0.17	Metallic to very light	0.72	Blue green to green (quite broad)
	yellow green	0.77	"Yellowish"
0.20	Light gold or yellow;	0.80	Orange (rather broad for orange)
	slightly metallic	0.82	Salmon
0.22	Gold with slight	0.85	Dull, light red violet
	yellow orange	0.86	Violet
0.25	Orange to melon	0.87	Blue violet
0.27	Red violet	0.89	Blue
0.30	Blue to violet blue	0.92	Blue green
0.31	Blue	0.95	Dull yellow green
0.32	Blue to blue green	0.97	Yellow to "yellowish"
0.34	Light green	0.99	Orange
0.35	Green to yellow green	1.00	Carnation pink
0.36	Yellow green	1.02	Violet red
0.37	Green yellow	1.05	Red violet
0.39	Yellow	1.06	Violet
0.41	Light orange	1.07	Blue violet
0.42	Carnation pink	1.10	Green
0.44	Violet red	1.11	Yellow green
0.46	Red violet	1.12	Green
0.47	Violet	1.18	Violet
0.48	Blue violet	1.19	Red violet
0.49	Blue	1.21	Violet red
0.50	Blue green	1.24	Carnation pink to salmon
0.52	Green (broad)	1.25	Orange
0.54	Yellow green	1.28	"Yellowish"
0.56	Green yellow	1.32	Sky blue to green blue
0.57	Yellow to "yellowish" (not yellow but is in the position where yellow is to be expected; at times appears to be light creamy gray or metallic)	1.40	Orange
		1.45	Violet
		1.46	Blue violet
		1.50	Blue
0.58	Light orange or yellow to pink	1.54	Dull yellow green
0.60	Carnation pink		

Accurate film thickness measurements can also be achieved using light-interference effects in microscopy, and automated interference-based equipment is commercially available for thin-film characterization.

3.10 PROCESS SIMULATION

As the scale of integrated circuits is reduced, accurate knowledge of one-, two-, and three-dimensional structural details of the process becomes more and more important. At the same time, experimental examination and characterization of the structures is a difficult and time-consuming task in deep submicron fabrication. For these reasons, computer simulation continues to grow in importance throughout the VLSI fabrication process.

Sophisticated computer programs that can predict the results of various fabrication steps have been available for many years [25–30]. These programs solve the generalized differential equations that model various fabrication processes and include the ability to simulate multidimensional oxide growth with its attendant moving $Si–SiO_2$ boundary, impurity segregation during oxide growth, and dopant evaporation from the surface as described in this chapter, as well as the various deposition, diffusion, epitaxial growth, and ion implantation processes studied in upcoming chapters. The detailed structures resulting from etching and recessed oxidation can also be simulated. Other programs have been developed to model lithography processes such as photoresist exposure and development.

One of the most widely used of these programs is the Stanford University Process Engineering Modeling program (SUPREM) [25–27] and its commercial implementations. The use of SUPREM requires specification of the process steps including times, temperature profiles, and other ambient conditions for oxidation, diffusion, ion implantation, film deposition, and etching. The program can model the one- and two-dimensional structures resulting from oxidation, as well as predicting the impurity profiles in the substrate, oxide, and polysilicon layers.

An example of a one-dimensional oxidation simulation is given in the SUPREM IV listing in Table 3.3 on page 63. The input listing describes a complex dry-wet-dry oxidation cycle, including the ramp-up of the furnace from one temperature to another and various ambient gas conditions at different steps in the oxidation cycle. The output data includes the oxide thickness and impurity dose in the oxide. Graphical output of the data is shown in Fig. 3.17. Incorporation of boron in the oxide and its depletion from the substrate are both clearly evident in the plotted results. Note that the oxide extends both above and below the original silicon surface defined as $x = 0$ on the horizontal axis. Most of the oxide growth occurs during the steam oxidation. The width of the oxide agrees well with a simple estimate from Fig. 3.6, based upon a single 5-hr wet oxidation at 1100 C: $X_o = 1.5$ μm.

SUMMARY

Silicon dioxide provides a high-quality insulating barrier on the surface of the silicon wafer. In addition, this layer can serve as a barrier layer during subsequent impurity-diffusion process steps. These two factors have allowed silicon to become the dominant semiconductor material in use today.

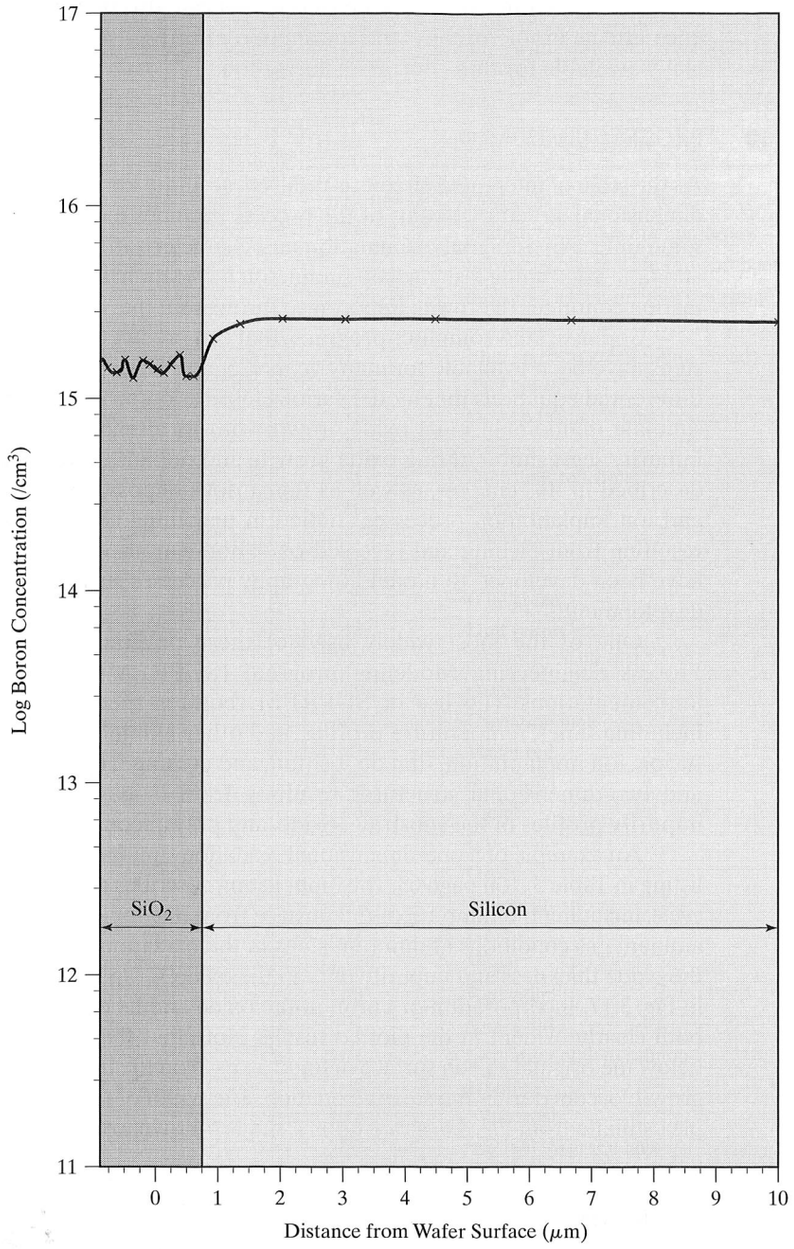


FIGURE 3.17

Results of SUPREM simulation of oxide growth on boron doped silicon wafers.

TABLE 3.3 SUPREM-IV Simulation Example

```
$ Multistep Oxidation

$ Use Automatic Grid Generation and Adaptive Grid

INITIALIZE <100> BORON=5 RESISTIV
DIFFUSION TEMP=950 TIME=30 F.N2=5
DIFFUSION TEMP=950 TIME=30 T.FINAL=1100 F.O2=5
DIFFUSION TEMP=1100 TIME=300 STEAM
DIFFUSION TEMP=1100 TIME=60 F.02=5
DIFFUSION TEMP=1100 TIME=60 T.FINAL=800 F.N2=5
$ Print layer information
....
....
$ Plot results
....
....
```

A native oxide layer several tens of angstroms thick forms on the surface of silicon immediately upon exposure to oxygen even at room temperature. The thickness of this oxide layer may be readily measured from the accumulation-region capacitance of a MOS test capacitor. Thicker layers of silicon dioxide are conveniently grown in high-temperature oxidation furnaces using both wet and dry oxygen. Oxidation occurs much more rapidly in wet oxygen than in dry oxygen. However, the dry-oxygen environment produces a higher quality oxide and is usually used for the growth of MOS gate oxides. Thin oxides grow in direct proportion to time. However, as the oxide becomes thicker, grow rate slows and becomes proportional to the square root of time. These two growth regions are characterized by the linear and parabolic growth-rate constants.

Oxide cleanness is extremely important for MOS processes, and great care is exercised to prevent sodium contamination of the oxide. The addition of chlorine during oxidation improves oxide quality. Finally, oxidation alters the impurity distribution at the surface of the silicon wafer. Boron tends to be depleted from the silicon surface, whereas phosphorus tends to pile up at the silicon surface.

Oxidation thickness can be accurately measured using ellipsometers, interference microscopes, and mechanical surface profilers or can be estimated from the apparent color of the oxide under vertical illumination with white light.

REFERENCES

[1] W. R. Runyon and K. E. Bean, *Semiconductor Integrated Circuit Processing Technology*, Addison Wesley, Reading, MA, 1990.

[2] S. A. Campbell, *The Science and Engineering of Microelectronic Fabrication*, Oxford University Press, 1996.

[3] S. M. Sze, *VLSI Technology*, McGraw-Hill, New York, 1983.

[4] S. K. Ghandhi, *VLSI Fabrication Principles*, John Wiley & Sons, New York, 1983.

[5] R. A. Colclaser, *Microelectronics—Processing and Device Design*, John Wiley & Sons, New York, 1980.

[6] W. R. Runyan, *Semiconductor Measurements and Instrumentation*, McGraw-Hill, New York, 1975.

[7] S. K. Ghandhi, *The Theory and Practice of Microelectronics*, John Wiley & Sons, New York, 1968.

[8] B. E. Deal and A. S. Grove, "General Relationship for the Thermal Oxidation of Silicon," *Journal of Applied Physics*, *36*, 3770–3778 (December, 1965).

[9] B. E. Deal, "Thermal Oxidation Kinetics of Silicon in Pyrogenic H_2O and 5% HCL/H_2O Mixtures," *Journal of the Electrochemical Society*, *125*, 576–579 (April, 1978).

[10] B. E. Deal, "The Oxidation of Silicon in Dry Oxygen, Wet Oxygen and Steam," *Journal of the Electrochemical Society*, *110*, 527–533 (June, 1963).

[11] W. A. Pliskin and E. E. Conrad, "Nondestructive Determination of Thickness and Refractive Index of Transparent Films," *IBM Journal of Research and Development*, *8*, 43–51 (January, 1964).

[12] S. C. Su, "Low Temperature Silicon Processing Techniques for VLSIC Fabrication," *Solid-State Technology*, *24*, 72–82 (March, 1981).

[13] A. S. Grove, O. Leistiko, and C. T. Sah, "Redistribution of Acceptor and Donor Impurities During Thermal Oxidation of Silicon," *Journal of Applied Physics*, *35*, 2695–2701 (September, 1964).

[14] A. S. Grove, *Physics and Technology of Semiconductor Devices*, John Wiley & Sons, New York, 1967.

[15] E. H. Nicollian and J. R. Brews, *MOS Physics and Technology*, John Wiley & Sons, New York, 1982.

[16] M. Ghezzo and D. M. Brown, "Diffusivity Summary of B, Ga, P, As and Sb in SiO_2," *Journal of the Electrochemical Society*, *120*, 146–148 (January, 1973).

[17] E. Passaglia, R. R. Stromberg, and J. Kruger, Eds., *Ellipsometry in the Measurement of Surfaces and Thin Films*, National Bureau of Standards, Miscellaneous Publication #256, 1964.

[18] H. Kotaki et al., "Novel Bulk Dynamic Threshold Voltage MOSFET (B-DTMOS) with Advanced Isolation (SITOS) and Gate to Shallow-Well Contact (SSS-C) Processes for Ultra Low Power Dual Gate CMOS," *IEEE IEDM Technical Digest*, pp. 459–462, December 1996.

[19] K. P. Muller, B. Flietner, C. L. Hwang, R. L. Kleinhenz, T. Nakao, R. Ranade, Y. Tsunashima, and T. Mii, "Trench Storage Node Technology for Gigabit DRAM Generations," *IEEE IEDM Technical Digest*, pp. 507–510, December 1996.

[20] M. Nandakumar, A. Chatterjee, S. Sridhar, K. Joyner, M. Rodder, and I-C. Chen, "Shallow Trench Isolation for Advanced ULSI CMOS Technologies," *IEEE IEDM Technical Digest*, pp. 133–136, December 1998.

[21] J. M. Steigerwald et al., "Pattern Geometry in Chemical-Mechanical Polishing of Inlaid Copper Structures," *Journal of the Electrochemical Society*, *141*, no. 10: 2842–2848, Oct. 1994.

[22] B. Stine et al., "Rapid Characterization and Modeling of Pattern-Dependent Variation in Chemical-Mechanical Polishing," *IEEE Trans. Semiconductor Manufacturing*, *11*, no. 1, February 1998.

[23] H. Nojo, M. Kodera, and R. Nakata, "Slurry Engineering for Self-Stopping Dishing Free SiO_2-CMP," *IEEE IEDM Technical Digest*, pp. 349–352, December 1996.

[24] D. Edelstein et al., "Full Copper Wiring in a Sub-0.25μm CMOS ULSI Technology," *IEEE IEDM Technical Digest*, pp. 773–776, December 1997.

[25] D. A. Antoniadis and R. W. Dutton, "Models for Computer Simulation of Complete IC Fabrication Processes," *IEEE Journal of Solid-State Circuits*, SC-14, pp. 412–422, April 1979.

[26] C. P. Ho, J. D. Plummer, S. E. Hansen, and R. W. Dutton, "VLSI Process Modeling—SUPREM III," *IEEE Trans. Electron Devices*, ED-30, pp. 1438–1453, November 1983.

[27] D. Chin, M. Kump, H.G. Lee, and R. W. Dutton, "Process Design Using Coupled 2D Process and Device Simulators," *IEEE IEDM Technical Digest*, pp. 223–226, December 1986.

[28] C. D. Maldanado, F. Z. Custode, S. A. Louie, and R. K. Pancholy, "Two-Dimensional Simulation of a 2μm CMOS Process Using ROMANS II," *IEEE Trans. Electron Devices*, ED-30, pp. 1462–1469, November 1983.

[29] M. E. Law, C. S. Rafferty, and R. W. Dutton, "New n-well Fabrication Techniques Based upon 2D Process Simulation," *IEEE IEDM Technical Digest*, pp. 518–521, December 1986.

[30] R. W. Dutton, "Modeling and Simulation for VLSI," *IEEE IEDM Technical Digest*, pp. 2–7, December 1986.

[31] John D. Cressler, "Re-engineering Silicon: Si-Ge Heterojunction Bipoloar Technology," *IEEE Spectrum*, pp. 49-55, March 1995.

PROBLEMS

3.1 How long does it take to grow 100 nm of oxide in wet oxygen at 1000 °C (assume 100 silicon)? In dry oxygen? Which process would be preferred?

3.2 A 1.2-μm silicon dioxide film is grown on a <100> silicon wafer in wet oxygen at 1100 °C. How long does it take to grow the first 0.4 μm? The second 0.4μm? The final 0.4 μm?

3.3 Derive Eq. (3.8) by solving differential Eq. (3.7).

3.4 A 3-μm silicon dioxide film is grown on a <100> silicon wafer in wet oxygen at 1150 C. How long does it take to grow the initial 1 μm oxide? The second micron? The final micron?

3.5 The gate oxide for a CMOS process on <100> silicon is to have a thickness of 10 nm (100 Å). Calculate the time required to grow this oxide at 850° C in wet oxygen using Eq. 3.8. Repeat for 1000° C. (Be careful!) Do either of these possible processes seem to be controllable? If so, which one.

3.6 A 2-μm SiO_2 film is needed as the initial oxide on a <100> silicon wafer. (a) Find the growth time in wet oxygen at 1150° C using Fig. 3.6. (b) Use Eq. (3.8) to calculate the growth time.

3.7 A 1-μm SiO_2 film is needed as the initial oxide on a <100> silicon wafer. (a) Find the growth time in wet oxygen at 1050° C using Fig. 3.6. (b) Repeat the calculation for dry oxygen. (c) Use Eq. (3.8) to calculate the growth times.

3.8 A dry-wet-dry oxidation cycle of 30 min./120 min./30 min. is performed at 1100° C. (a) What is the final oxide thickness for a <100> silicon wafer? Use Eq. 3.8. (b) What is the final oxide thickness for a <111> silicon wafer?

3.9 An one-hour dry oxidation at 1000° C is followed by a 5-hour wet oxidation at 1100° C. (a) Calculate the oxide thickness after each step for a <100> wafer. (b) Find the final oxide thickness graphically.

3.10 An one-hour dry oxidation at 1100° C is followed by a 5-hour wet oxidation at 1100° C. (a) Calculate the oxide thickness after each step for a <111> wafer. (b) Find the final oxide thickness graphically.

3.11 A square window is etched through 200 nm of oxide prior to a second oxidation, as in Example 3.2. The second oxidation grows 300 nm of oxide in the thick oxide region at 100° C. Make a scale drawing of the cross section of the wafer after the second oxidation. What are the colors of the various regions under vertical illumination by white light? Assume a temperature of 1100° C.

3.12 A 10-μm square window is etched through a 1-μm thick oxide on a silicon wafer. The wafer is reoxidized to grow a new 1-μm thick oxide film in the window at 1100° C. (a) Draw a cross section of the wafer following the second oxidation. (b) What are the colors of the oxide under vertical illumination by white light? Assume a temperature of 1100° C.

3.13 A <100> silicon wafer has 400 nm of oxide on its surface. How long will it take to grow an additional 1μm of oxide in wet oxygen at 1100° C? Compare graphical and mathematical results. What is the color of the final oxide under vertical illumination by white light?

3.14 How much oxide is needed to mask a 4-hr boron diffusion at 1150° C? A 1-hr phosphorus diffusion at 1050° C?

3.15 The isolation diffusion in a bipolar process uses a 15-hour boron diffusion at 1150° C into a <111> silicon wafer. How much oxide is required as a barrier layer for this diffusion?

3.16 The *n*-well in a <100> CMOS process is formed with a 20-hour phosphorus diffusion at 1200° C. How much oxide is required as a barrier layer for this diffusion?

3.17 What is the color of the oxide in Problem 3.7? (b) How about in Problem 3.6?

3.18 Yellow light has a wavelength of approximately 0.57 μm. Calculate the thicknesses of silicon dioxide that will appear yellow under vertical illumination by white light. Consider oxide thicknesses less than 1.5 μm. Compare with the color chart (Table 3.2).

3.19 Write a computer program to calculate the linear and parabolic rate constants for wet and dry oxidation for temperatures of 950, 1000, 1050, 1100, 1150, and 1200 C. Assume <100> silicon.

3.20 Write a computer program to calculate the time required to grow a given thickness of oxide, based on the theory of Section 3.2. The user should be able to specify desired oxide thickness, wet or dry oxidation conditions, temperature, and orientation of the silicon wafer.

3.21 (a) Use SUPREM to simulate the oxide growth in Problem 3.8 on a <100> wafer doped with 10^{15} boron atoms/cm^3. Plot the final doping profiles in the silicon and oxide. (b) Repeat for a wafer doped with 10^{15} arsenic atoms/cm^3. (c) Compare the oxide thickness to hand calculations.

3.22 (a) Use SUPREM to simulate the oxide growth in Problem 3.8 on a <111> wafer doped with 3×10^{15} boron atoms/cm^3. (b) Compare the oxide thickness to hand calculations.

3.23 (a) Use SUPREM to simulate the oxide growth in Problem 3.6 on a 5 Ω-cm <100> boron-doped wafer. Plot the concentration of boron in the oxide and substrate. (b) Repeat for a 5 Ω-cm wafer doped with phosphorus atoms.

3.24 (a) Use SUPREM to simulate the oxide growth in Problem 3.7 on a on a <100> wafer doped with 5×10^{15} boron atoms/cm^3. Plot the concentration of boron in the oxide and substrate. (b) Repeat for a wafer with a doping of 5×10^{15} phosphorus atoms/cm^3.

3.25 Use SUPREM to calculate the thicknesses of the oxides in the two regions in Problem 3.12.

CHAPTER 4

Diffusion

High-temperature diffusion has historically been one of the most important processing steps used in the fabrication of monolithic integrated circuits. For many years, diffusion was the primary method of introducing impurities such as boron, phosphorus, and antimony into silicon to control the majority-carrier type and resistivity of layers formed in the wafer. Today, diffusion is used in the formation of "deep" layers exceeding a few tenths of a micron in depth. However, most deposition steps utilize the ion-implantation and rapid thermal annealing processes that will be explored in Chapter 5. We must still study the diffusion process in order to understand its limitations and the various problems associated with redistribution of impurities as they are added to silicon. In this chapter, we explore the theoretical and practical aspects of the diffusion process, the characterization of diffused layer sheet resistance, and the determination of junction depth. Physical diffusion systems and solid, liquid, and gaseous impurity sources are all discussed.

4.1 THE DIFFUSION PROCESS

The diffusion process begins with the deposition of a shallow high-concentration layer of the desired impurity in the silicon surface through windows etched in the protective barrier layer. At high temperatures (900 to 1200 °C), the impurity atoms move from the surface into the silicon crystal via the *substitutional* or *interstitial* diffusion mechanisms illustrated in Fig. 4.1 on page 68.

In the case of substitutional diffusion, the impurity atom hops from one crystal lattice site to another. The impurity atom thereby "substitutes" for a silicon atom in the lattice. Vacancies must be present in the silicon lattice in order for the substitutional process to occur. Statistically, a certain number of vacancies will always exist in the lattice. At high temperatures, vacancies may also be created by displacing silicon atoms from their normal lattice positions into the vacant *interstitial* space between lattice sites. The substitutional diffusion process in which silicon atoms are displaced into interstitial sites is called *interstitialcy* diffusion.

Considerable space exists between atoms in the silicon lattice, and certain impurity atoms diffuse through the crystal by jumping from one interstitial site to another. Since this mechanism does not require the presence of vacancies, interstitial diffusion