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

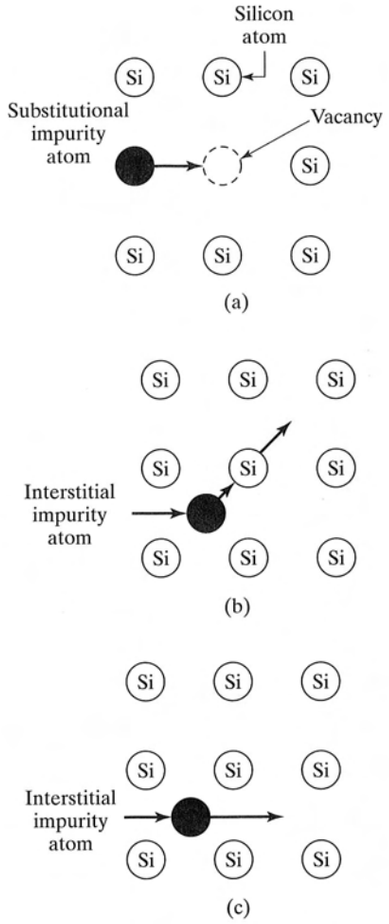


FIGURE 4.1

Atomic diffusion in a two-dimensional lattice. (a) Substitutional diffusion, in which the impurity moves among vacancies in the lattice; (b) interstitialcy mechanism, in which the impurity atom replaces a silicon atom in the lattice, and the silicon atom is displaced to a interstitial site; (c) interstitial diffusion, in which impurity atoms do not replace atoms in the crystal lattice.

proceeds much more rapidly than substitutional diffusion. The rapid diffusion rate makes interstitial diffusion difficult to control.

Impurity atoms need to occupy substitutional sites in the lattice in order to provide electrons or holes for conduction, as described in Volume I of this series [1]. Substitutional diffusion proceeds at a relatively low rate, because the supply of vacancies is limited, but this slow diffusion rate is actually an advantage, because it permits good control of the diffusion process.

4.2 MATHEMATICAL MODEL FOR DIFFUSION

The basic one-dimensional diffusion process follows *Fick's first law* of diffusion, presented in Chapter 3,

$$J = -D\,\partial N/\partial x \tag{4.1}$$

where J is the particle flux of the donor or acceptor impurity species, N is the concentration of the impurity, and D is the diffusion coefficient.

Fick's second law of diffusion may be derived using the continuity equation for the particle flux:

$$\partial N/\partial t = -\partial J/\partial x \tag{4.2}$$

Equation (4.2) states that the rate of increase of concentration with time is equal to the negative of the divergence of the particle flux. For the one-dimensional case, the divergence is equal to the gradient. Combining Eqs. (4.1) and (4.2) yields Fick's second law of diffusion:

$$\partial N/\partial t = D\,\partial^2 N/\partial x^2 \tag{4.3}$$

Here the diffusion coefficient D is assumed to be independent of position. This assumption is violated at high impurity concentrations. (See Section 4.6.3)

The partial differential equation in Eq. (4.3) can be solved by variable separation or Laplace transform techniques. Two specific types of boundary conditions are important in modeling impurity diffusion in silicon. The first is the *constant-source diffusion*, in which the surface concentration is held constant throughout the diffusion. The second is called a *limited-source diffusion*, in which a fixed quantity of the impurity species is deposited in a thin layer in the surface of the silicon.

4.2.1 Constant-Source Diffusion

During a constant-source diffusion, the impurity concentration is held constant at the surface of the wafer. Under this boundary condition, the solution to Eq. (4.3) is given by

$$N(x,t) = N_0 \operatorname{erfc}(x/2\sqrt{Dt}) \tag{4.4}$$

for a semiinfinite wafer in which N_0 is the impurity concentration at the wafer surface ($x = 0$). Such a diffusion is called a *complementary error function* (erfc) diffusion and is shown graphically in Fig. 4.2 on page 70. As time progresses, the diffusion front proceeds further and further into the wafer with the surface concentration remaining constant. The total number of impurity atoms per unit area in the silicon is called the *dose*, Q, with units of atoms/cm^2. Q increases with time, and an external impurity source must supply a continual flow of impurity atoms to the surface of the wafer. The dose is found by integrating the diffused impurity concentration throughout the silicon wafer.

$$Q = \int_0^\infty N(x,t)dx = 2N_0\sqrt{Dt/\pi} \tag{4.5}$$

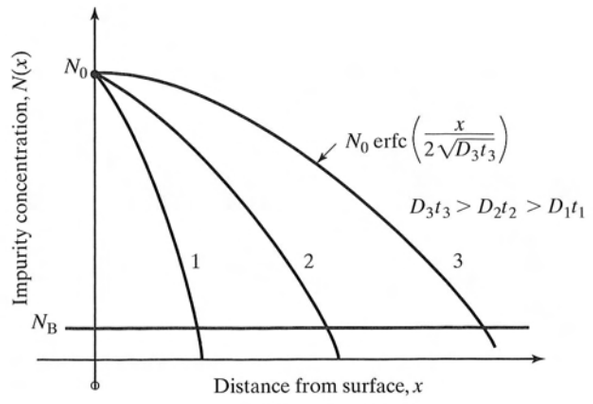


FIGURE 4.2

A constant-source diffusion results in a complementary error function impurity distribution. The surface concentration N_0 remains constant, and the diffusion moves deeper into the silicon wafer as the Dt product increases. Dt can change as a result of increasing diffusion time, increasing diffusion temperature, or a combination of both.

4.2.2 Limited-Source Diffusion

A limited-source diffusion is modeled mathematically using an impulse function at the silicon surface as the initial boundary condition. The magnitude of the impulse is equal to the dose Q. For this boundary condition in a semi-infinite wafer, the solution to Eq. (4.3) is given by the Gaussian distribution,

$$N(x,t) = (Q/\sqrt{\pi Dt})\exp-(x/2\sqrt{Dt})^2, \tag{4.6}$$

which is displayed graphically in Fig. 4.3. The dose remains constant throughout the limited-source diffusion process. As the diffusion front moves into the wafer, the surface concentration must decrease, so that the area under the curve can remain constant with time.

On a normalized logarithmic plot, the shapes of the Gaussian and complementary error function curves appear similar, as illustrated in Fig. 4.4. The erfc curve, however, falls off more rapidly than the Gaussian curve.

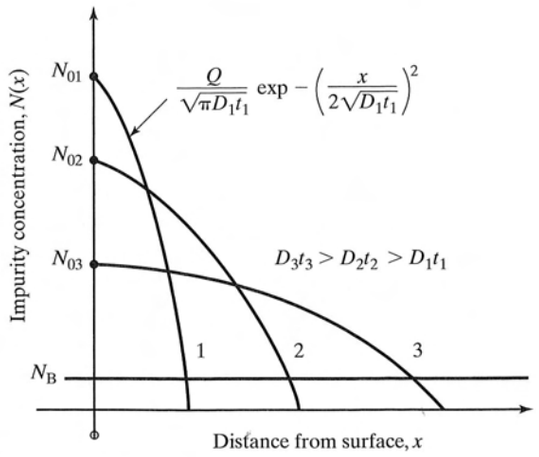


FIGURE 4.3

A Gaussian distribution results from a limited-source diffusion. As the Dt product increases, the diffusion front moves more deeply into the wafer, and the surface concentration decreases. The area (impurity dose) under each of the three curves is the same.

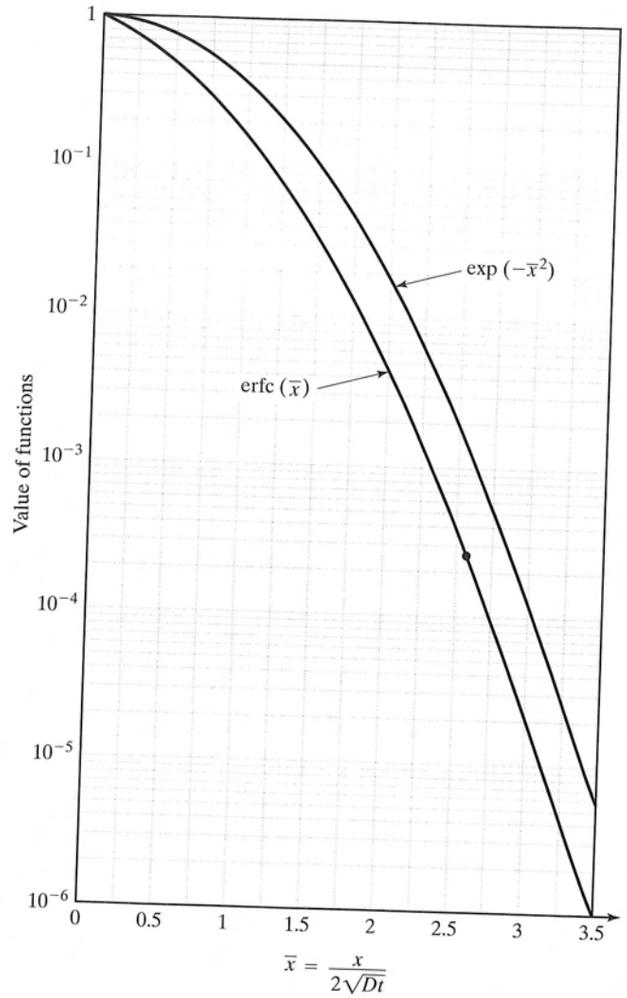


FIGURE 4.4

A graph comparing the Gaussian and complementary error function (erfc) profiles. We use this curve to evaluate the erfc and its inverse.

4.2.3 Two-Step Diffusion

A short constant-source diffusion is often followed by a limited-source diffusion, resulting in a two-step diffusion process. The constant-source diffusion step is used to establish a known dose in a shallow layer on the surface of the silicon and is called the *predeposition* step. The fixed dose approximates an impulse and serves as the impurity source for the second diffusion step.

The second diffusion is called the *drive-in* step and is used to move the diffusion front to the desired depth. If the Dt product for the drive-in step is much greater than the Dt product for the predeposition step, the resulting impurity profile is closely approximated by a Gaussian distribution, Eq. (4.6). If the Dt product for the drive-in step is much less than the Dt product for the predeposition step, the resulting impurity

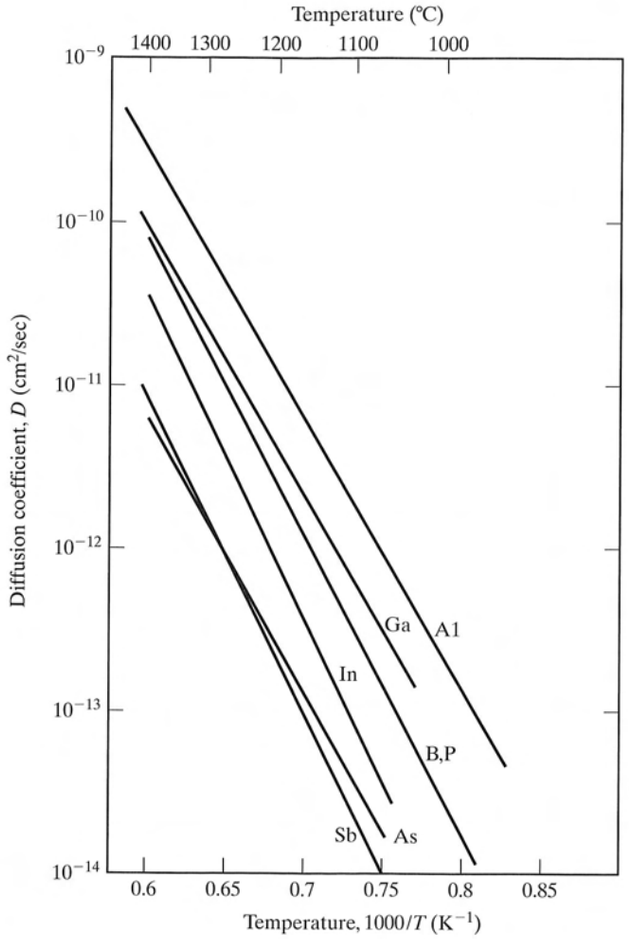


FIGURE 4.5

Diffusion constants in silicon for (a) substitutional diffusers (above) and (b) interstitial diffusers (next page). Copyright John Wiley & Sons, Inc.; reprinted with permission from Ref. [28].

profile is closely approximated by a complementary error function distribution, Eq. (4.4). An integral equation solution to the diffusion equation also exists for diffusion conditions that do not satisfy either inequality. (See Ref. [25].)

4.3 THE DIFFUSION COEFFICIENT

Figure 4.5 shows the temperature dependence of the diffusion coefficient D for (a) substitutional and (b) interstitial diffusers in silicon. The large difference between these coefficients is readily apparent. To achieve reasonable diffusion times with substitutional diffusers, temperatures in the range of 900 to 1200 °C are typically used. Interstitial diffusers are difficult to control, because of their large diffusion coefficients. (See Problem 4.19.)

Diffusion coefficients depend exponentially on temperature and follow the Arrhenius behavior introduced in Chapter 3:

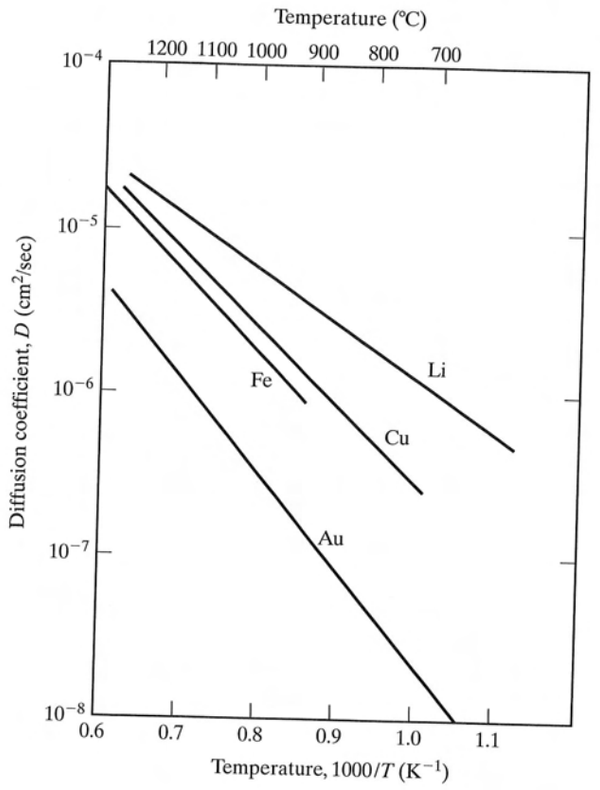


FIGURE 4.5

Continued.

$$D = D_0 \exp(-E_A/kT). \tag{4.7}$$

Values for D_0 and E_A can be determined from Fig. 4.5. Typical values for a number of impurities are given in Table 4.1.

Wide variability exists in diffusion coefficient data reported in the literature. We will use Eq. (4.7) and Table 4.1 in the examples and problems throughout the rest of this book. In general, calculations based on Eq. (4.7) and Table 4.1 can be used as guides. Most processes are then experimentally calibrated under the specific diffusion conditions in each laboratory.

Example 4.1

Calculate the diffusion coefficient for boron at 1100 °C.

Solution: From Table 4.1, $D_o = 10.5$ cm^2/sec and $E_A = 3.69$ eV. $T = 1373$ K.

$$D = 10.5 \exp - \frac{3.69}{(8.614 \times 10^{-5})(1373)} = 2.96 \times 10^{-13} \text{cm}^2/\text{sec}.$$

TABLE 4.1 Typical Diffusion Coefficient Values for a Number of Impurities.

Element	D_o(cm²/sec)	E_A(eV)
B	10.5	3.69
Al	8.00	3.47
Ga	3.60	3.51
In	16.5	3.90
P	10.5	3.69
As	0.32	3.56
Sb	5.60	3.95

4.4 SUCCESSIVE DIFFUSIONS

We are ultimately interested in the final impurity distribution after all processing is complete. A wafer typically goes through many time–temperature cycles during predeposition, drive-in, oxide growth, CVD, etc. For example, the base diffusion in a bipolar transistor will be followed by several high-temperature oxidations, as well as the emitter predeposition and drive-in cycles. These steps take place at different temperatures for different lengths of time. The effect of these steps is determined by calculating the total Dt product, $(Dt)_{tot}$, for the diffusion. $(Dt)_{tot}$ is equal to the sum of the Dt products for all high-temperature cycles affecting the diffusion:

$$(Dt)_{tot} = \Sigma_i D_i t_i \tag{4.8}$$

D_i and t_i are the diffusion coefficient and time associated with the ith processing step. $(Dt)_{tot}$ is then used in Eq. (4.4) or Eq. (4.6) to determine the final impurity distribution.

Example 4.2

Calculate $(Dt)_{tot}$ for a boron diffusion of 2 hours at 1100 °C followed by 5 hours at 1150 °C.

Solution: From Ex. 4.1 at $T = 1100$ °C, $D = 2.96 \times 10^{-13}$ cm²/sec. For $T = 1150$ °C,

$$D = 10.5 \exp - \frac{3.69}{(8.614 \times 10^{-5})(1423)} = 8.86 \times 10^{-13}\text{cm}^2/\text{sec},$$

$$(Dt)_{tot} = (2.96 \times 10^{-13}\ \text{cm}^2/\text{sec})(7200\ \text{sec}) + (8.86 \times 10^{-13}\ \text{cm}^2/\text{sec})(18000\ \text{sec})$$

$$(Dt)_{tot} = 1.81 \times 10^{-8}\text{cm}^2.$$

4.5 SOLID-SOLUBILITY LIMITS

At a given temperature, there is an upper limit to the amount of an impurity that can be absorbed by silicon. This quantity is called the *solid-solubility limit* for the impurity and is indicated by the solid lines in Fig. 4.6 for boron, phosphorus, antimony, and arsenic at normal diffusion temperatures. As can be seen in the figure, surface concentrations achieved through solid-solubility-limited diffusions will be quite high. For example, the solid-solubility limit of boron is approximately 3.3×10^{20}/cm³ at 1100° C, and 1.2×10^{21}/cm³ for phosphorus at the same temperature. High concentrations are desired for the emitter and subcollector diffusions in bipolar transistors and the source and drain diffusions in MOSFETs. However, solid-solubility-limited concentrations are too heavy for the base regions of bipolar transistors and for many resistors. The two-step diffusion process described in Section 4.2.3 overcomes this problem.

At high concentrations, only a fraction of the impurities actually contribute holes or electrons for conduction. The dotted lines in Fig. 4.6 show the "electrically active" portion of the impurity concentration. These curves will be referred to again in Section 4.7.2.

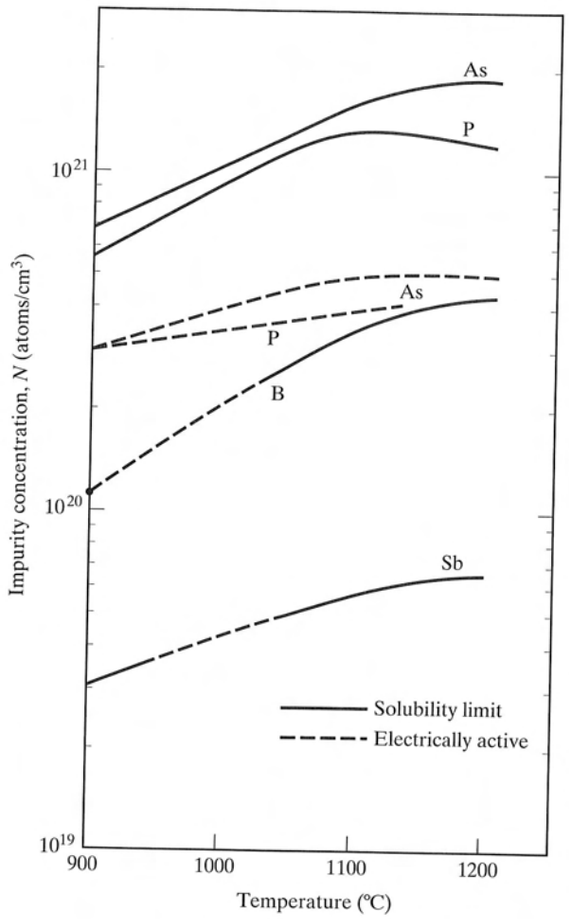


FIGURE 4.6

The solid-solubility and electrically active impurity-concentration limits in silicon for antimony, arsenic, boron, and phosphorus. Reprinted with permission from Ref. [29]. This paper was originally presented at the 1977 Spring Meeting of The Electrochemical Society, Inc., held in Philadelphia, Pennsylvania.

4.6 JUNCTION FORMATION AND CHARACTERIZATION

4.6.1 Vertical Diffusion and Junction Formation

The goal of most diffusions is to form *pn* junctions by converting *p*-type material to *n*-type material or vice versa. In Fig. 4.7, for example, the wafer is uniformly doped *n*-type material with a concentration indicated by N_B, and the diffusing impurity is boron. The point at which the diffused impurity profile intersects the background concentration is the *metallurgical junction depth*, x_j. The net impurity concentration at x_j is zero. Setting $N(x)$ equal to the background concentration N_B at $x = x_j$ yields

$$x_j = 2\sqrt{Dt \ln(N_0/N_B)} \tag{4.9}$$

and

$$x_j = 2\sqrt{Dt}\ \text{erfc}^{-1}(N_B/N_0) \tag{4.10}$$

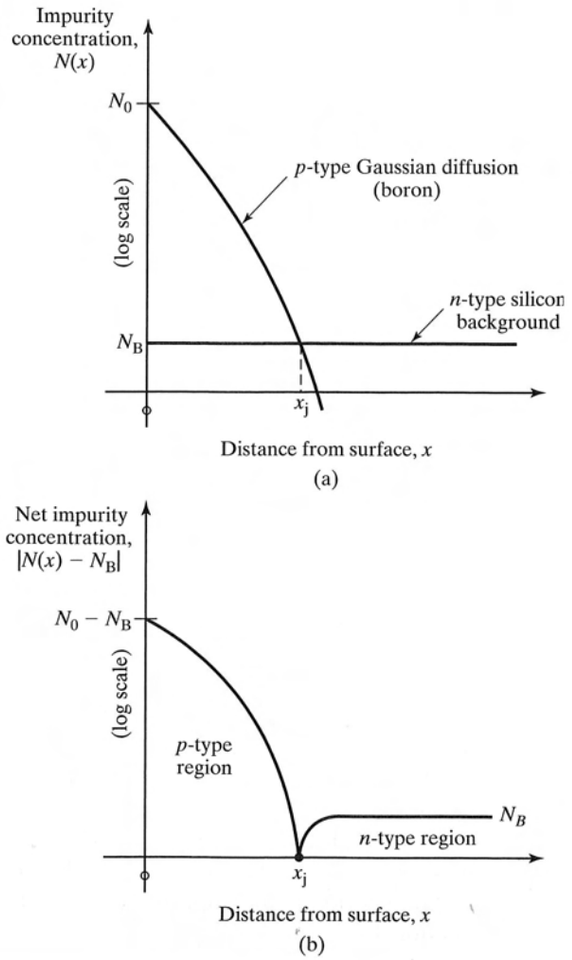


FIGURE 4.7

Formation of a *pn* junction by diffusion. (a) An example of a *p*-type Gaussian diffusion into a uniformly doped *n*-type wafer; (b) net impurity concentration in the wafer. The metallurgical junction occurs at the point $x = x_j$ where the net concentration is zero. The material is converted to *p*-type to the left of x_j and remains *n*-type to the right of x_j.

for the Gaussian and complementary error function distributions, respectively. In Fig. 4.7, the boron concentration N exceeds N_B to the left of the junction, and this region is *p*-type. To the right of x_j, N is less than N_B, and this region remains *n*-type.

We can use our scientific calculators to evaluate Eq. (4.9), and we will learn to evaluate the complementary error function expression using Fig. 4.4. To calculate the junction depth, we must know the background concentration N_B of the original wafer. Figure 4.8 gives the resistivity of *n*- and *p*-type silicon as a function of doping concentration. The background concentration can be determined using this figure when uniform concentrations of either donor or acceptor impurities are present in the silicon wafer.

Example 4.3

A boron diffusion is used to form the base region of an *npn* transistor in a 0.18-ohm-cm *n*-type silicon wafer. A solid-solubility-limited boron predeposition is performed at 900 °C for 15 min followed by a 5-hr drive-in at 1100 °C. Find the surface concentration and junction depth (a) following the predeposition step and (b) following the drive-in step.

Solution: The predeposition step is a solid-solubility-limited constant-source diffusion. Using Fig. 4.6, we see that the boron surface concentration is approximately $1.1 \times 10^{20}/\text{cm}^3$. The temperature of 900 °C equals 1173 K, which yields a diffusion coefficient $D_1 = 1.45 \times 10^{-15}$ cm²/sec, and $t_1 = 900$ sec (15 min). The constant-source diffusion results in an erfc profile, and the impurity profile following predeposition is given by

$$N(x) = 1.1 \times 10^{20}\ \text{erfc}(x/2\sqrt{D_1 t_1})\ \text{boron atoms/cm}^3.$$

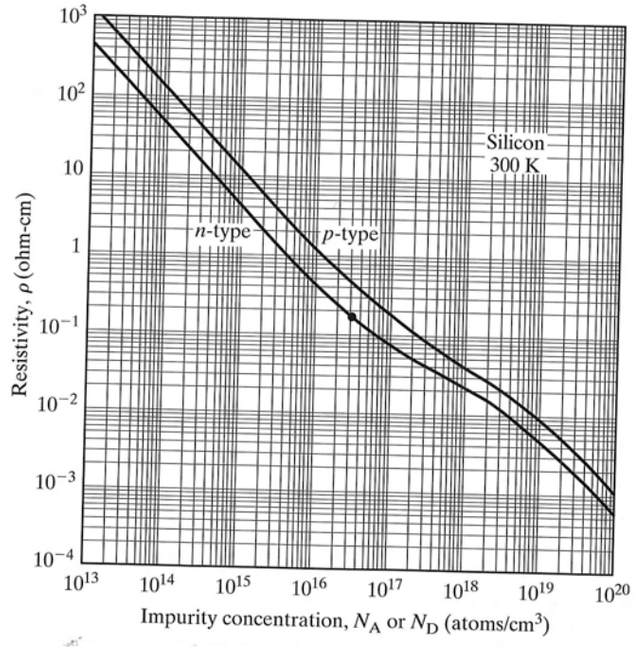


FIGURE 4.8

Room-temperature resistivity in *n*- and *p*-type silicon as a function of impurity concentration. (Note that these curves are valid for either donor or acceptor impurities but not for compensated material containing both types of impurities.) Copyright 1987 Addison-Wesley Publishing Company. Reprinted with permission from Ref. [3].

To find the junction depth x_j, we must find the point at which the concentration $N(x)$ is equal to the background concentration N_B. Using Fig. 4.8, we find that a 0.18-ohm-cm n-type wafer corresponds to a doping concentration of $3 \times 10^{16}/\text{cm}^3$. Thus,

$$1.1 \times 10^{20} \operatorname{erfc}(x_j/2\sqrt{D_1 t_1}) = 3 \times 10^{16}.$$

Solving for x_j yields

$$x_j = 2\sqrt{D_1 t_1}\, \operatorname{erfc}^{-1}(0.000273) = 2(\sqrt{1.31 \times 10^{-12}})(2.57)\text{cm} = 0.0587\ \mu\text{m}.$$

The value of $\operatorname{erfc}^{-1}(0.000273)$ is found with the aid of Fig. 4.4. The value 2.73×10^{-4} corresponds to the y-axis value of the complementary error function, and the corresponding value for the normalized distance is $x = 2.57$.

The dose in silicon is needed for the drive-in step and is equal to

$$Q = 2N_0\sqrt{D_1 t_1/\pi} = 2(1.1 \times 10^{20})\sqrt{(1.45 \times 10^{-15})(900)/\pi}\ \text{boron atoms/cm}^2$$

$$Q = 1.42 \times 10^{14}\ \text{boron atoms/cm}^2$$

At the drive-in temperature of 1100 °C (1373 K), $D_2 = 2.96 \times 10^{-13}$ cm^2/sec, and the drive-in time of 5 hr = 18000 sec. Assuming that a Gaussian profile results from the drive-in step, the final profile is given by

$$N(x) = 1.1 \times 10^{18} \exp - (x/2\sqrt{D_2 t_2})^2\ \text{boron atoms/cm}^3. \quad (4.3.1)$$

Setting Eq. (4.3.1) equal to the background concentration yields the final junction depth of 2.77 μm. Figure 4.9 on page 79 shows the concentrations at various points in the diffusion process.

We must check our assumption that the drive-in step results in a Gaussian profile. The Dt product for the predeposition step is 1.31×10^{-12} cm^2, and the Dt product for the drive-in step is 5.33×10^{-9} cm^2. Thus, $D_2 t_2 \gg D_1 t_1$, and our assumption is justified.

4.6.2 Lateral Diffusion

During diffusion, impurities not only diffuse vertically, but also move laterally under the edge of any diffusion barrier. Figure 4.10 on page 80 presents the results of computer simulation of the two-dimensional diffusion process. The normalized impurity concentrations can be used to find the ratio of lateral to vertical diffusion. Lateral diffusion is an important effect-coupling device and process design and was an important factor driving the development of self-aligned polysilicon-gate MOS processes. The interaction of lateral diffusion and device layout will be discussed in greater detail in Chapters 9 and 10.

Example 4.4

An erfc diffusion results in a junction depth of 2 μm and a surface concentration of $1 \times 10^{20}/\text{cm}^3$. The background concentration of the wafer is $1 \times 10^{16}/\text{cm}^3$. What is the lateral diffusion underneath the edge of the mask?

Solution: The junction occurs at $N = N_B$, and $N_0/N_B = 10^4$. Using Fig. 4.10(a), we find that the ratio of lateral diffusion to vertical diffusion is 2.4/2.75, or 0.87. The lateral junction depth is therefore 1.74 μm.

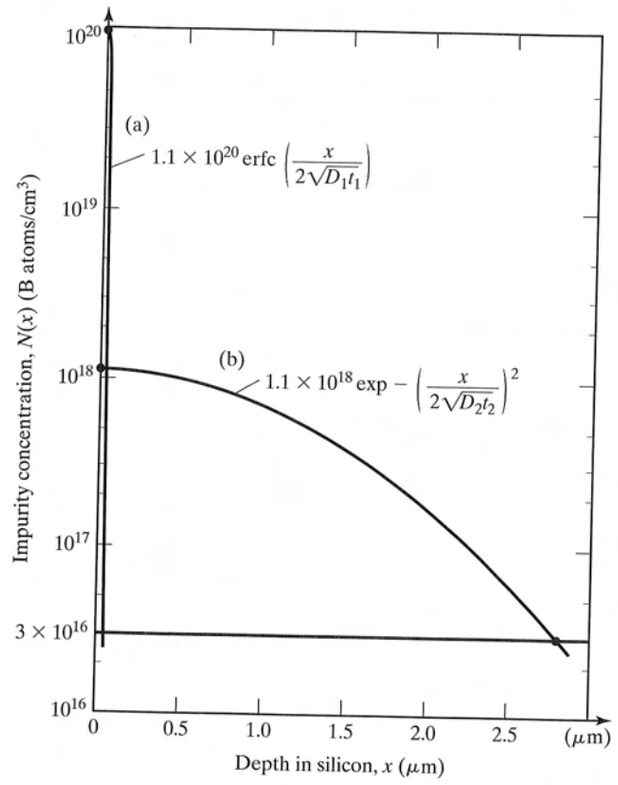


FIGURE 4.9

Calculated boron impurity profiles for Example 4.3. (a) Following the predeposition step at 900°C for 15 min; (b) following a subsequent 5-hr drive-in step at 1100°C. The final junction depth is 2.77 μm with a surface concentration of $1.1 \times 10^{18}/\text{cm}^3$. The initial profile approximates an impulse.

4.6.3 Concentration-Dependent Diffusion

Diffusion follows the theory of Section 4.3 as long as the impurity concentration remains below the value of the intrinsic-carrier concentration n_i at the diffusion temperature. Above this concentration, the diffusion coefficient becomes concentration dependent, and each of the common impurities exhibits a different behavior.

The diffusion equation can be solved analytically for linear, parabolic, and cubic dependencies of the diffusion coefficient on concentration. The results are presented in Fig. 4.11 on page 81, in which D_{sur} represents the diffusion coefficient at the surface. In general, concentration-dependent diffusion results in a much more abrupt profile than for the case of a constant-diffusion coefficient. These highly abrupt profiles are actually of value in formation of the shallow junctions desired in scaled VLSI devices.

Boron and arsenic can be modeled by the first-order dependence in Fig. 4.11, resulting in the analytical relations between junction depth, sheet resistance, total dose, and surface concentration given in Table 4.2 [7–9] on page 81.

High-concentration phosphorus diffusion results in a more complicated profile than that of boron or arsenic. Figure 4.12 on page 82 depicts typical shallow phosphorus diffusion profiles. As phosphorus diffuses into the wafer, the diffusion coefficient becomes enhanced at concentrations below approximately $10^{19}/\text{cm}^3$, resulting in a distinct "kink" in the profile. The kink effect represents a practical limitation to the use of phosphorus for the source–drain diffusions and emitter diffusions of shallow MOS and

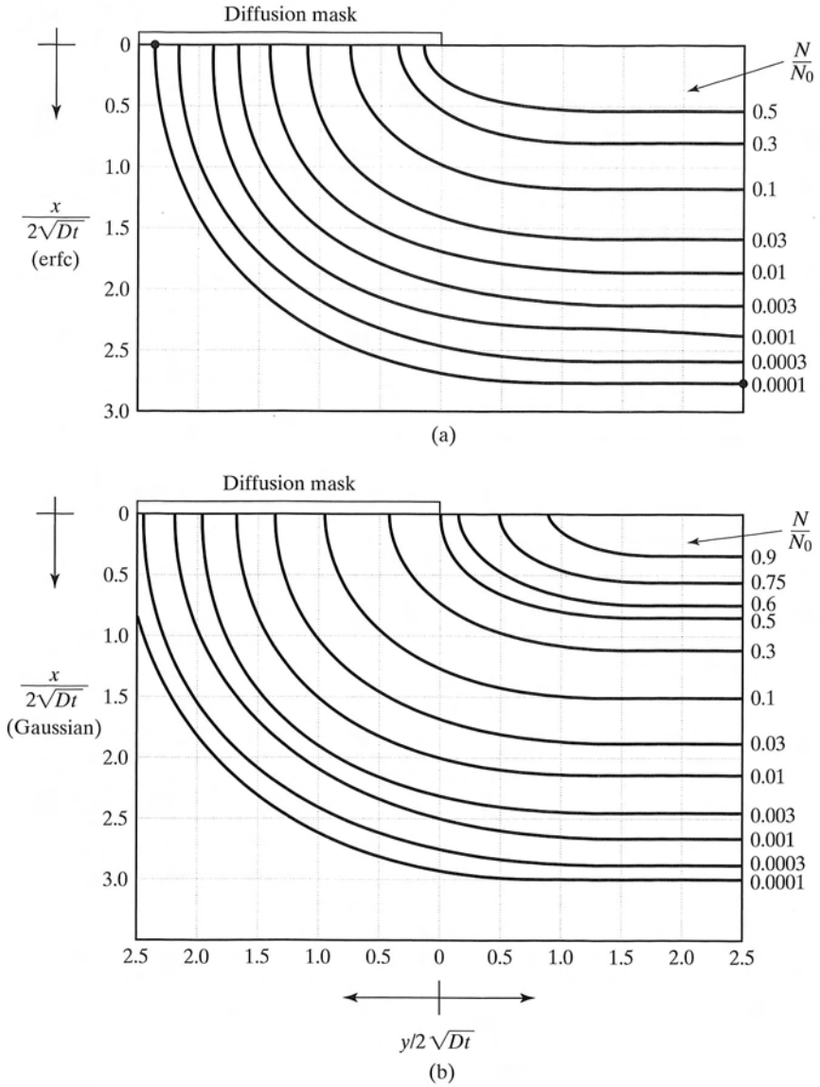


FIGURE 4.10

Normalized two-dimensional complementary error function and Gaussian diffusions near the edge of a window in the barrier layer. Copyright 1965 by International Business Machines Corporation; reprinted with permission from Ref. [4].

bipolar devices. Most MOS and bipolar VLSI processes now use arsenic to avoid this problem. Complex mathematical models describing the diffusion of phosphorus may be found in Refs. [10] and [11].

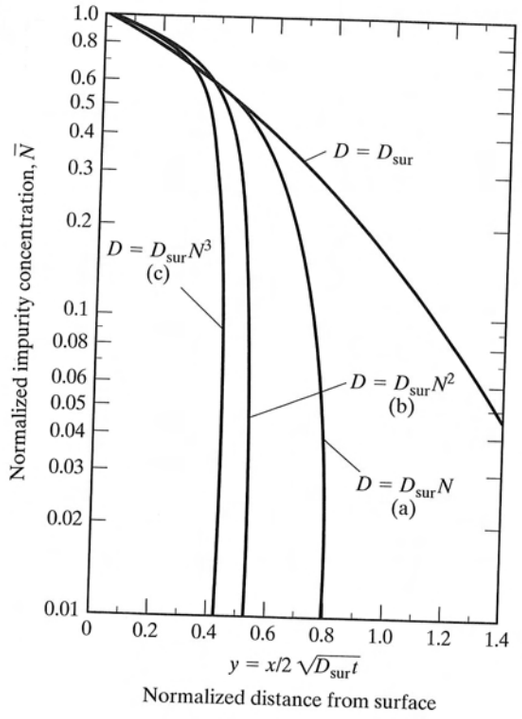


FIGURE 4.11

Diffusion profiles for concentration-dependent diffusion. Copyright 1963 by the American Physical Society. Reprinted with permission from Ref. [6].

TABLE 4.2 Properties of High-Concentration Arsenic and Boron Diffusions

Element	x_j(cm)	D(cm^2/sec)	N_0(cm^{-3})	Q(cm^{-2})
Arsenic	$2.29\sqrt{N_0Dt/n_i}$*	22.9 exp (− 4.1/kT)	$1.56\times10^{17}(R_sx_j)^{-1}$	$0.55N_0x_j$
Boron	$2.45\sqrt{N_0Dt/n_i}$*	3.17 exp (− 3.59/kT)	$2.78\times10^{17}(R_sx_j)^{-1}$	$0.67N_0x_j$

*The value of n_i must be calculated at the diffusion temperature.

4.7 SHEET RESISTANCE

In diffused layers, resistivity is a strong function of depth. For circuit and device design, it is convenient to work with a new parameter, R_s, called *sheet resistance*, which eliminates the need to know the details of the diffused-layer profile.

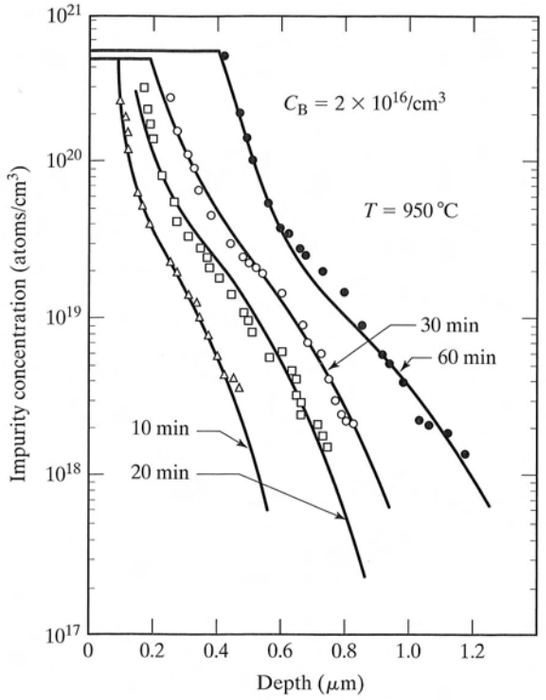


FIGURE 4.12

Shallow phosphorus diffusion profiles for constant-source diffusions at 950 °C. Copyright 1969 IEEE. Reprinted with permission from Ref. [10].

4.7.1 Sheet-Resistance Definition

Let us first consider the resistance R of the rectangular block of uniformly doped material in Fig. 4.13. R is given by

$$R = \rho L/A, \tag{4.11}$$

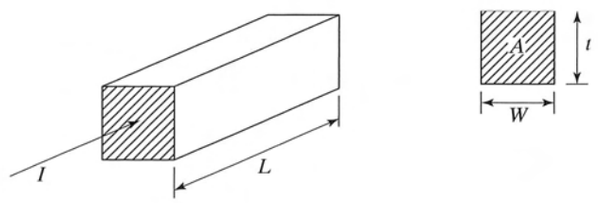


FIGURE 4.13

Resistance of a block of material having uniform resistivity. A uniform current distribution is entering the material perpendicular to the end of the block. The ratio of resistivity to thickness is called the *sheet resistance* of the material.

$$R = \rho\frac{L}{A} \qquad \rho = \frac{1}{\sigma} \qquad \sigma = q\,(\mu_n n + \mu_p p)$$

where ρ is the material's resistivity, and L and A represent the length and cross-sectional area of the block, respectively. Resistance is proportional to the material resistivity. If the length of the block is made longer, the resistance increases, and the resistance is inversely proportional to the cross-sectional area.

Using W as the width of the sample and t as the thickness of the sample, the resistance may be rewritten as

$$R = (\rho/t)(L/W) = R_s(L/W), \tag{4.12}$$

where $R_s = (\rho/t)$ is called the *sheet resistance* of the layer of material. Given the sheet resistance R_s, a circuit designer need specify only the length and width of the resistor to define its value. Strictly speaking, the unit for sheet resistance is the ohm, since the ratio L/W is unitless. To avoid confusion between R and R_s, sheet resistance is given the special descriptive unit of ohms per square. The ratio L/W represents the number of unit squares of material in the resistor.

Figure 4.14 shows top and side views of two typical dumbbell-shaped resistors with top contacts at the ends. The body of each resistor is seven "squares" long. If the sheet resistance of the diffusion were 50 ohms per square, each resistor would have a resistance of 350 ohms. The portion of the resistor surrounding the contacts also contributes to the total resistance of the structure. Figure 4.15 on page 84 presents the effective number of squares contributed by various end and corner configurations. For the resistor of Fig. 4.14, each end adds approximately 0.65 squares to the resistor and the total resistance would be 415 Ω. Note that lateral diffusion under the edges of the mask may change both the geometry of the contacts, as well as the number of squares in the body of the resistor. (See Problems 4.8–4.10.)

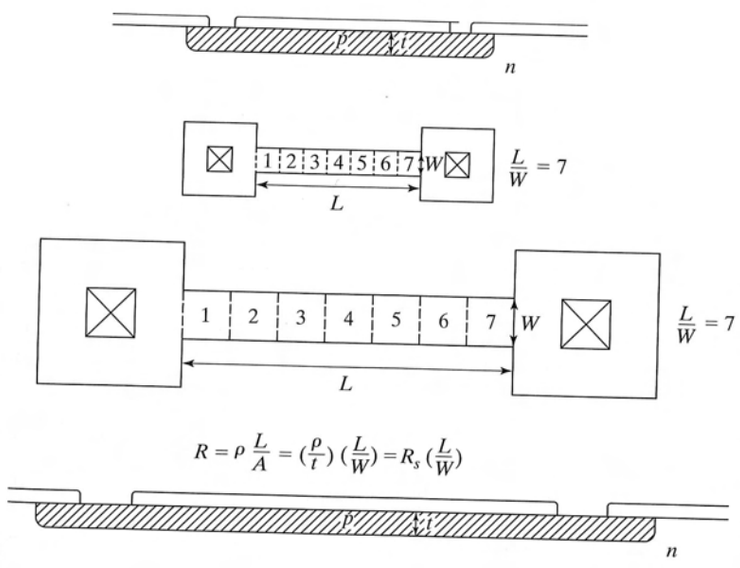


FIGURE 4.14

Top and side views of two diffused resistors of different physical size having equal values of resistance. Each resistor has a ratio L/W equal to 7 squares. Each end of the resistor contributes approximately 0.65 additional squares.

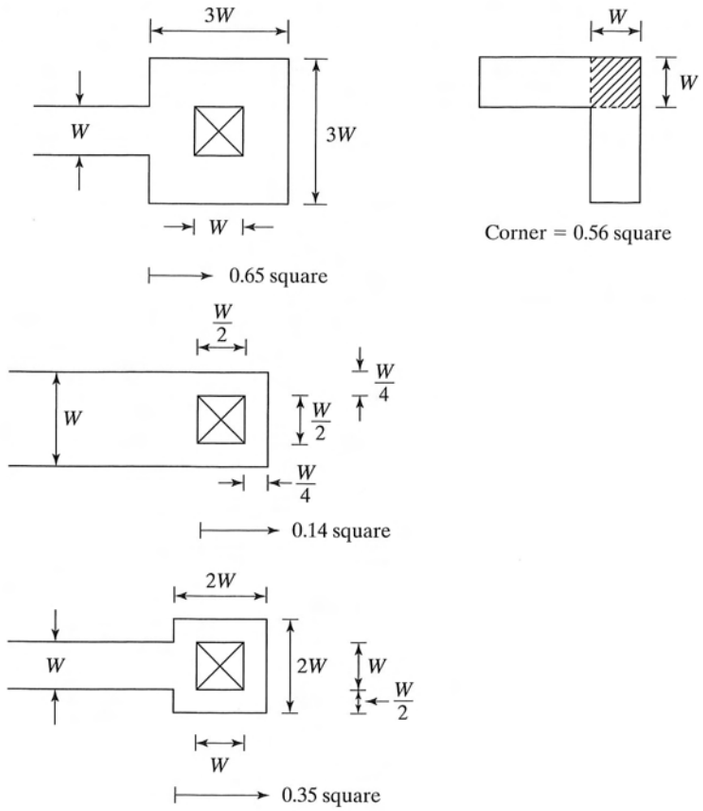


FIGURE 4.15

Effective square contributions of various resistor end and corner configurations.

Example 4.5

A diffusion with a sheet resistance of 200 Ω/□ is used to fabricate a 5 kΩ resistor with the dumbbell shape similar to that in Fig. 4.14. How many squares are required in the body of the resistor?

Solution: The total number of squares required is $N_{tot} = (5000\ \Omega) / (200\ \Omega/\square) = 25\ \square$. Subtracting the contribution of the two contacts gives $N = 25 - 2\ (0.65) = 23.7\ \square$ for the body of the resistor. Note that the resistor body need not be an interger number of squares.

4.7.2 Irvin's Curves

From Section 4.2, we know that the impurity concentration resulting from a diffusion varies rapidly between the surface and the junction. Thus, ρ is a function of depth for diffused resistors. For diffused layers, we define the sheet resistance R_s in terms of the average resistivity of the layer

$$\overline{\rho} = \frac{1}{\overline{\sigma}} = \frac{1}{\frac{1}{x_j}\int_0^{x_j} \sigma(x)dx},$$

$$R_s = \frac{\overline{\rho}}{x_j} = [\int_0^{x_j} \sigma(x)dx]^{-1}.$$

In extrinsic material, this expression can be approximated by

$$R_s = [\int_0^{x_j} q\mu N(x)dx]^{-1}, \tag{4.13}$$

where x_j is the junction depth, μ is the majority-carrier mobility, and $N(x)$ is the net impurity concentration. We neglect the depletion of charge carriers near the junction x_j.

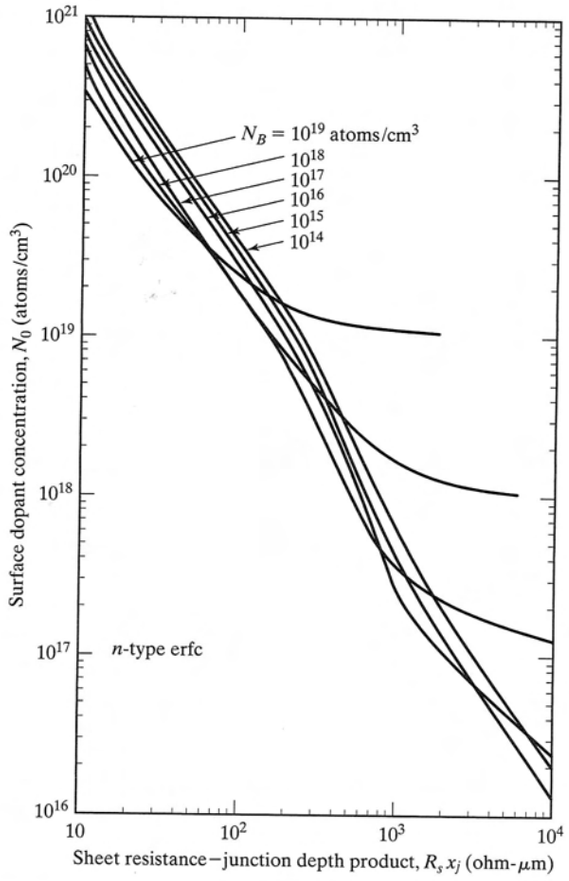
For a given diffusion profile, sheet resistance is uniquely related to the surface concentration of the diffused layer and the background concentration of the wafer. Equation (4.13) was evaluated numerically by Irvin [5], and a number of Irvin's results have been combined into Figs. 4.16(a)–(d) [2]on page 86-87. These figures plot surface concentration versus the $R_s x_j$ product and are used to find the sheet resistance and surface concentration of diffused layers.

Example 4.6

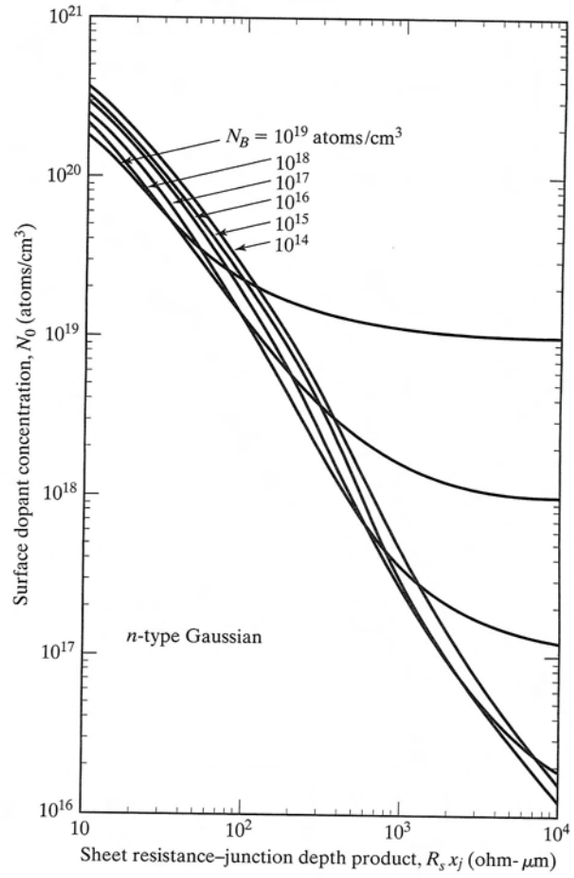
Find the sheet resistance of the boron diffusion from Example 4.2.

Solution: From Example 4.3, the background concentration of the wafer is $3 \times 10^{16}/\text{cm}^3$, the surface concentration is $1.1 \times 10^{18}/\text{cm}^3$, and the junction depth is 2.77 μm. The diffusion resulted in a *p*-type Gaussian layer. Using Fig. 4.16(d), we find that the $R_s x_j$ product is found to be approximately 800 ohm-μm. Dividing by a junction depth of 2.77 μm yields a sheet resistance of 289 ohms/square.

Sheet resistance is an electrical quantity that depends on the majority-carrier concentration. As shown in Fig. 4.6, the electrically active impurity concentration for phosphorus and arsenic is considerably less than the total impurity concentration at high doping levels. In order to use Irvin's curves at high doping levels, the vertical axis, which is labeled "surface dopant density," should be interpreted to be the electrically active dopant concentration at the surface.



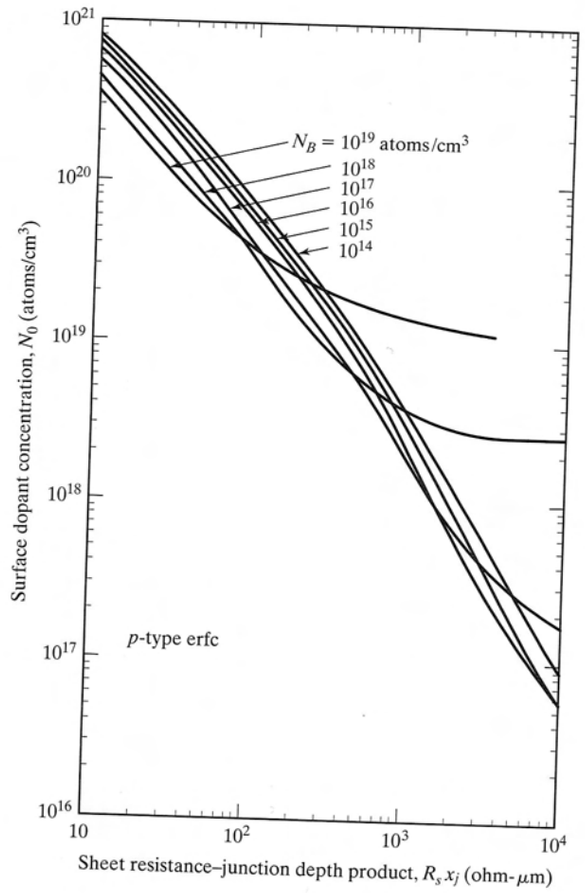
(a)



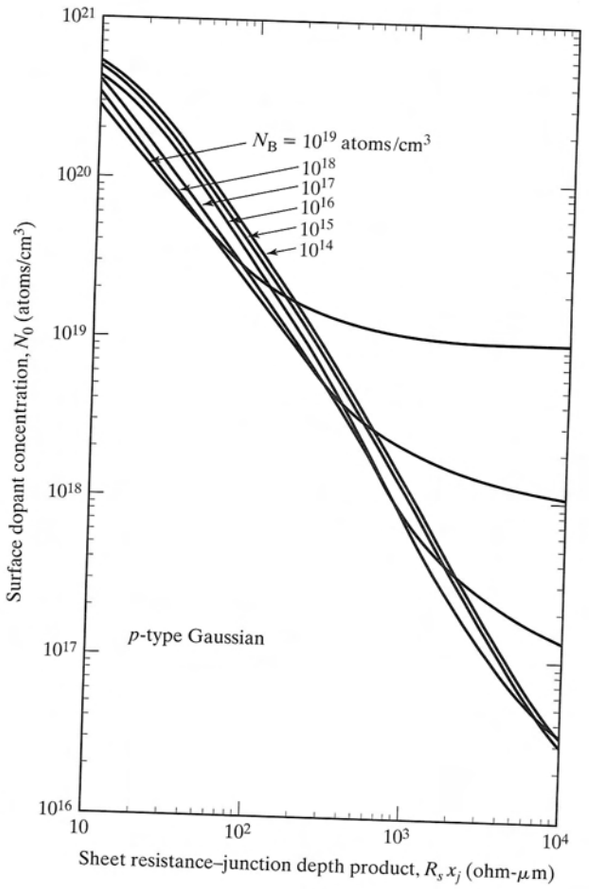
(b)

FIGURE 4.16

Surface impurity concentration versus the sheet resistance–junction depth product for different silicon background concentrations at 300 K. (a) *n*-type erfc distribution; (b) *n*-type Gaussian distribution; (c) *p*-type erfc distribution; (d) *p*-type Gaussian distribution. After Ref. [2]. Reprinted from Ref. [5] with permission from the *AT&T Technical Journal*. Copyright 1962 AT&T.



(c)



(d)

FIGURE 4.16

Continued.

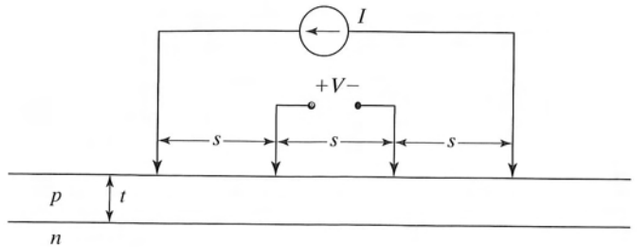


FIGURE 4.17

Four-point probe with probe spacing s used for direct measurement of bulk wafer resistivity and the sheet resistance of thin diffused layers. A known current is forced throught the outer probes, and the voltage developed is measured across the inner probes. (See Eqs. (4.14) through (4.16).)

4.7.3 The Four-Point Probe

A special instrument called a *four-point probe* may be used to measure the bulk resistivity of starting wafers and the sheet resistance of shallow diffused layers. As shown schematically in Fig. 4.17, a fixed current is injected into the wafer through the two outer probes, and the resulting voltage is measured between the two inner probes. If probes with a uniform spacing s are placed on an infinite slab of material, then the resistivity is given by

$$\rho = 2\pi sV/I \text{ ohm}-\text{meters for } t \gg s \tag{4.14}$$

and

$$\rho = (\pi t/\ln 2)V/I \text{ ohm}-\text{meters for } s \gg t. \tag{4.15}$$

For shallow layers, Eq. (4.15) gives the sheet resistance as

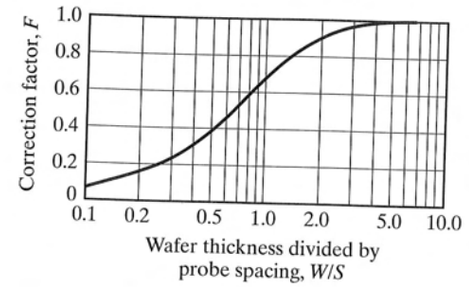
$$R_s = \rho/t = (\pi/\ln 2)V/I = 4.53\,V/I \text{ ohm}-\text{meters for } s \gg t. \tag{4.16}$$

The approximation used in Eqs. (4.15) and (4.16) is easily met for shallow diffused layers in silicon. Unfortunately, silicon wafers are often thinner than the probe spacing s, and the approximation in Eq. (4.14) is not valid. Correction factors are given in Fig. 4.18 for thin wafers and for small-diameter wafers [12].

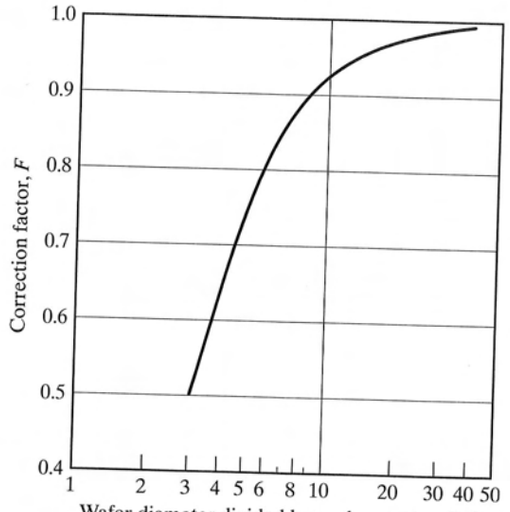
4.7.4 Van der Pauw's Method

The sheet resistance of an arbitrarily shaped sample of material may be measured by placing four contacts on the periphery of the sample. A current is injected through one pair of the contacts, and the voltage is measured across another pair of contacts. Van der Pauw [13,14] demonstrated that two of these measurements can be related by Eq. (4.17):

$$\exp(-\pi t R_{AB,CD}/\rho) + \exp(-\pi t R_{BC,DA}/\rho) = 1. \tag{4.17}$$



(a)



(b)

FIGURE 4.18

Four-point-probe correction factors, F, used to correct for (a) wafers which are relatively thick compared to the probe spacing s and (b) wafers of finite diameter. In each case $\rho = F\rho_{\text{measured}}$. (a) Copyright 1975 by McGraw-Hill Book Company. Reprinted with permission from Ref. [12]. (b) Reprinted from Ref. [30] with permission from the AT&T Technical Journal. Copyright 1958 AT&T.

Here $R_{AB,CD} = V_{CD}/I_{AB}$ and $R_{BC,DA} = V_{DA}/I_{BC}$. For a symmetrical structure such as a square or a circle,

$$R_{AB,CD} = R_{BC,DA}$$

and

$$R_s = \rho/t = (\pi/\ln 2)V_{CD}/I_{AB}. \tag{4.18}$$

Note the similarity to Eq. (4.16).

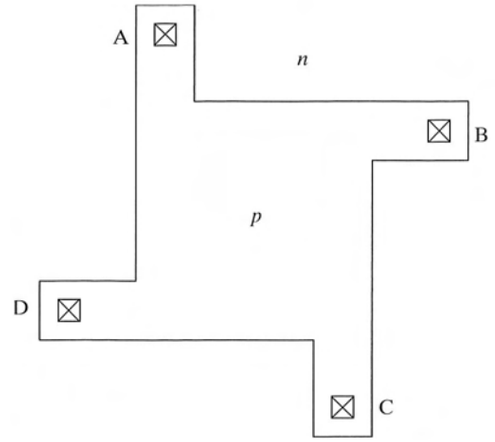


FIGURE 4.19

A simple van der Pauw test structure used to measure the sheet resistance of a diffused layer. Sheet resistance is calculated using Eq. (4.18).

Specially designed sheet-resistance test structures are often included on wafers so that the sheet resistances of *n*-type and *p*-type diffusions can be measured after final processing of the wafer. A sample structure is shown in Fig. 4.19.

4.8 JUNCTION-DEPTH AND IMPURITY PROFILE MEASUREMENT

Test wafers are normally processed in parallel with the actual IC wafers. No masking is done on the test wafer so that diffusion may take place across its full surface. The test wafer provides a large area for experimental characterization of junction depth. Alternatively, special test dice replace a few of the normal die sites on each wafer. These test dice provide an array of test structures for monitoring process and device characteristics during the various phases of the process.

4.8.1 Grove-and-Stain and Angle-Lap Methods

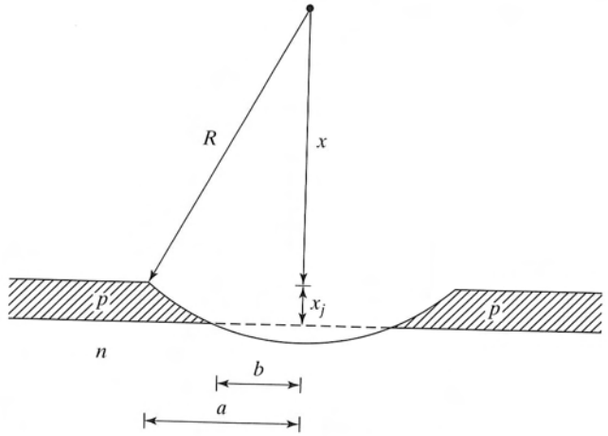
Two relatively simple methods can be used to measure the junction depth of diffused layers. In the first, known as the *groove-and-stain* method, a cylindrical groove is mechanically ground into the surface of the wafer, as shown in Fig. 4.20. If the radius R of the grinding tool is known, the junction depth x is easily found to be

$$x_j = \sqrt{(R^2 - b^2)} - \sqrt{(R^2 - a^2)}. \tag{4.19}$$

If the radius R is much larger than both distances a and b, then the junction depth is given approximately by

$$x_j = (a^2 - b^2)/2R = (a + b)(a - b)/2R. \tag{4.20}$$

After the grooving operation, the junction is delineated using a chemical etchant that stains the *pn* junction. Concentrated hydrofluoric acid with 0.1 to 0.5% nitric acid



$$(x + x_j) - x = \sqrt{R^2 - b^2} - \sqrt{R^2 - a^2} = R\left(\sqrt{1 - (\frac{b}{R})^2} - \sqrt{1 - (\frac{a}{R})^2}\right)$$

$$\doteq R\left\{(1 - \frac{1}{2}\frac{b^2}{R^2}) - (1 - \frac{1}{2}\frac{a^2}{R^2})\right\} \quad \begin{matrix} R >> a \\ R >> b \end{matrix}$$

$$x_j \doteq \frac{a^2 - b^2}{2R} = \frac{(a+b)(a-b)}{2R}$$

FIGURE 4.20

Junction-depth measurement by the groove-and-stain technique. The distances a and b are measured through a microscope, and the junction depth is calculated using Eq. (4.20).

can be used as a stain that is enhanced through exposure to high-intensity light [12]. The distances a and b are measured through a microscope, and the junction depth is calculated using Eq. (4.20).

The second technique is the *angle-lap* method. A piece of the wafer is mounted on a special fixture that permits the edge of the wafer to be lapped at an angle between 1 and 5°, as depicted in Fig. 4.21 on page 92. The junction depth is magnified so that the distance on the lapped surface is given by

$$x_j = d \tan\theta = N\lambda/2, \tag{4.21}$$

where θ is the angle of the fixture. An optically flat piece of glass is placed over the lapped region, and the test structure is illuminated with a collimated monochromatic beam of light with wavelength λ, typically from a sodium vapor lamp. The resulting interference pattern has fringe lines that are approximately 0.29 μm apart. The number N of fringes is counted through a microscope, and the junction depth may be found using Eq. (4.21). The usefulness of this method becomes limited for very shallow junctions. The analytic techniques discussed in the next section provide more general characterization capability for shallow structures.

4.8.2 Impurity-Profile Measurement

Spreading-resistance measurements and *Secondary Ion Mass Spectroscopy* (SIMS) are two techniques that are widely used for measurement of impurity profiles in semiconductors. Note that both methods are destructive; that is, they modify or destroy the region being characterized.

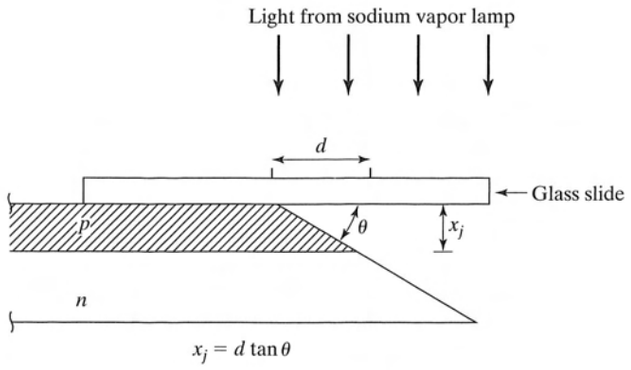


FIGURE 4.21

Junction depth measurement by the angle-lap and stain method. Interference fringe lines are used to measure the distance d, which is related to the junction depth using Eq. (4.21).

In the spreading-resistance method, a region of the semiconductor is angle lapped in a manner similar to that used for junction measurement discussed in Section 4.8.1. The resistivity of the layer is measured as a function of depth on the angle-lapped surface using a two-point probe. From this information, the impurity concentration and impurity type can be calculated. Figure 4.22 on page 93 presents an example of the impurity profile and junction depths determined from spreading resistance measurements.

In the SIMS method [15–16] depicted in Fig. 4.23(a) on page 94, a low-energy (1–20 keV) ion beam of say, cesium or oxygen, is used to remove (sputter) atoms from the surface, one or two atomic layers at a time. A small percentage of the atoms that are removed from the surface are ionized, and these ions are collected and analyzed by a mass spectrometer, which identifies the atomic species. The analysis is performed continuously during the sputtering process, and a profile of atomic distibution versus depth is produced as shown in Fig. 4.23(b). Mass removal proceeds at a rate of 2–5 Å/sec and can be used to a depth of a few microns. SIMS is the only surface-analysis tool with the sensitivity needed to characterize impurity profiles in silicon. Examples of typical sensitivites achievable with SIMS are provided in Table 4.3.

TABLE 4.3 SIMS Analysis in Silicon.

Element	Ion Beam	Sensitivity
Arsenic	Cesium	5×10^{14}/cm^3
Boron	Oxygen	1×10^{13}/cm^3
Phosphorus	Cesium	5×10^{15}/cm^3
Oxygen	Cesium	1×10^{17}/cm^3

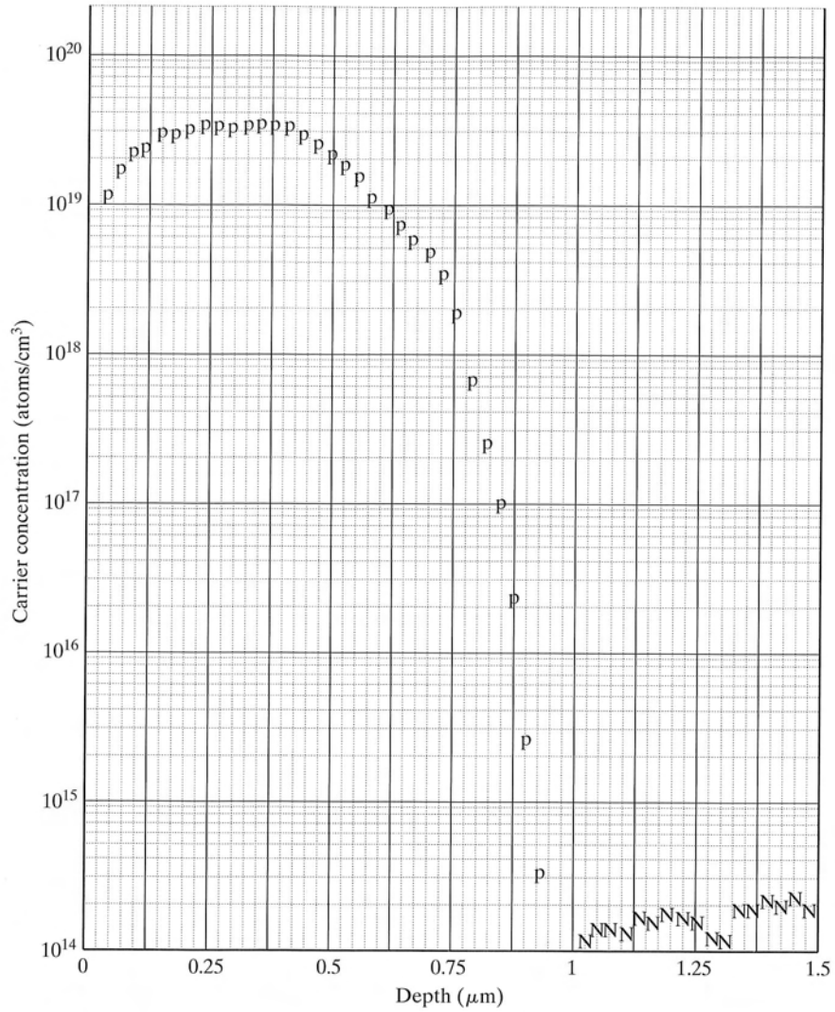


FIGURE 4.22

Example of an impurity profile measured using the spreading resistance method.

4.9 DIFFUSION SIMULATION

The SUPREM program introduced in Chapter 3 includes complete models for diffusion. SUPREM can simulate simple one-dimensional diffusion, as well as highly complex two-dimensional diffusion through a mask window, as depicted in Fig. 4.10. (See Problem 4.20.)

As a simple example, a portion of the input description of the two-step diffusion from Ex. 4.3 is given next, along with a plot of the corresponding output data in Fig.

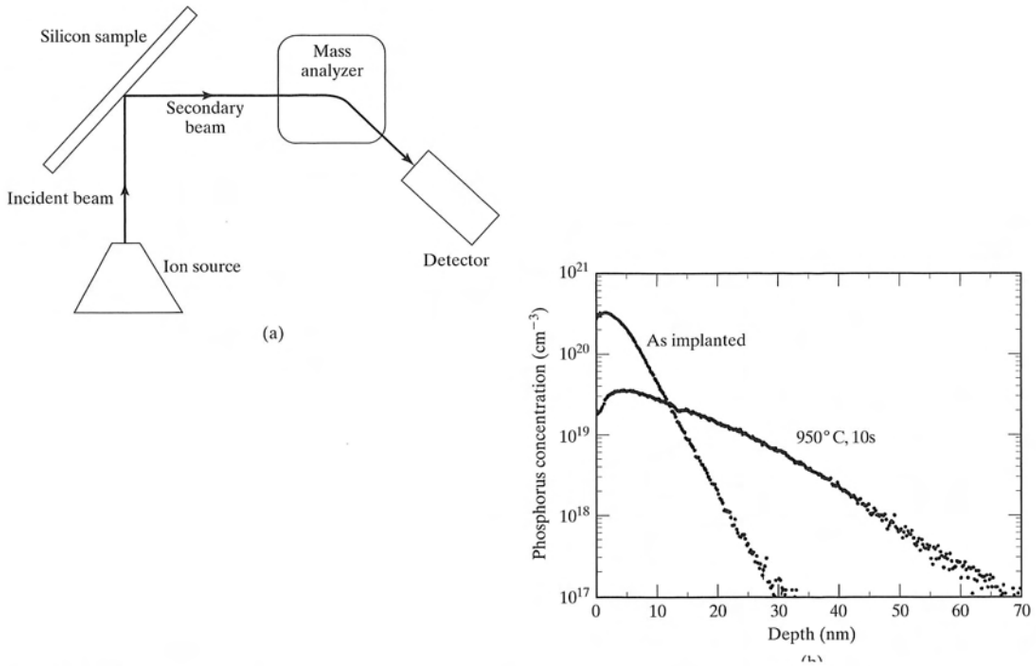


FIGURE 4.23

(a) Concept of a SIMS analysis system. (b) Example of an impurity profile measured using the SIMS analysis. Copyright 1997 IEEE. Reprinted with permission from Ref. [17].

4.24 on page 95. The input file defines the starting material to be a phosphorus doped <100> wafer with a resistivity of 0.18 Ω-cm. The predeposition takes place at 900 °C for 15 minutes and the drive-in occurs at 1100 °C for 300 minutes. (Output control statements are not included in the listings.)

Following the predeposition step, the SUPREM simulation results predict that the surface concentration will be $N_0 = 3 \times 10^{20}$/ cm^3 and that the junction depth will be $x_j = 0.1$ μm. After the drive-in step, the final values of N_0 and x_j are predicted to be 10^{17}/cm^3 and 2.0 μm, respectively. The sheet resistance of the diffused layer is estimated to be approximately 500 Ω/□. The simulation results show depletion of both boron and phosphorus near the wafer surface due to out-diffusion and indicate that the peak of the boron profile is actually below the silicon surface. These are features that we do not attempt to include in our basic hand analyses, and they show the power and importance of using the sophisticated computer simulation tools.

```
$TWO STEP DIFFUSION
INITIALIZE <100> PHOS=0.18 RESISTIVITY
DIFFUSE TEMP=900 TIME=15 BORON=1E21
...
...
DIFFUSION TEMP=1100 TIME=300
...
...
```

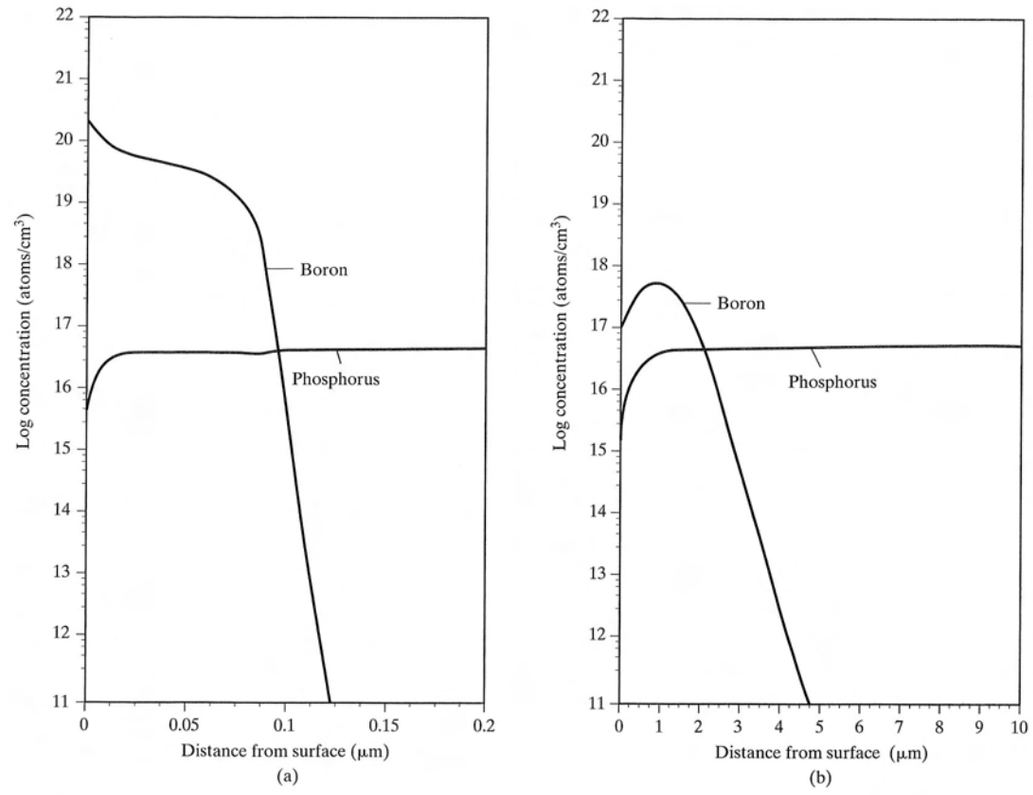



FIGURE 4.24
SUPREM simulation results for two-step boron diffusion into the phosphorus doped wafer from Ex. 4.3.

4.10 DIFFUSION SYSTEMS

The open-furnace-tube system using solid, liquid, or gaseous sources, as depicted in Fig. 4.25, yields good reproducibility and is a common diffusion technology used in IC fabrication. Three-zone horizontal furnaces can be used for diffusion. Wafers are placed in a quartz boat and positioned in the center zone of the furnace, where they are heated to a high temperature. Impurities are transported to the silicon surface, and then diffuse into the wafer.

Most common silicon dopants can be applied using liquid spin-on sources. These spin-on dopants are versatile, safe, and easy to apply, but the uniformity is often poorer than with other impurity sources. To achieve good quality control, most production systems use other solid, liquid, or gaseous impurity sources.

In one type of solid-source system, carrier gases (usually N_2 or O_2) flow at a controlled rate over a source boat placed in the furnace tube. The carrier gas picks up the vapor from the source and transports the dopant species to the wafer, where it is deposited on the surface of the wafer. The temperature of the source is controlled to maintain the desired vapor pressure. The source can be placed in a low-temperature section of the furnace or may be external to the furnace. Solid boron and phosphorus

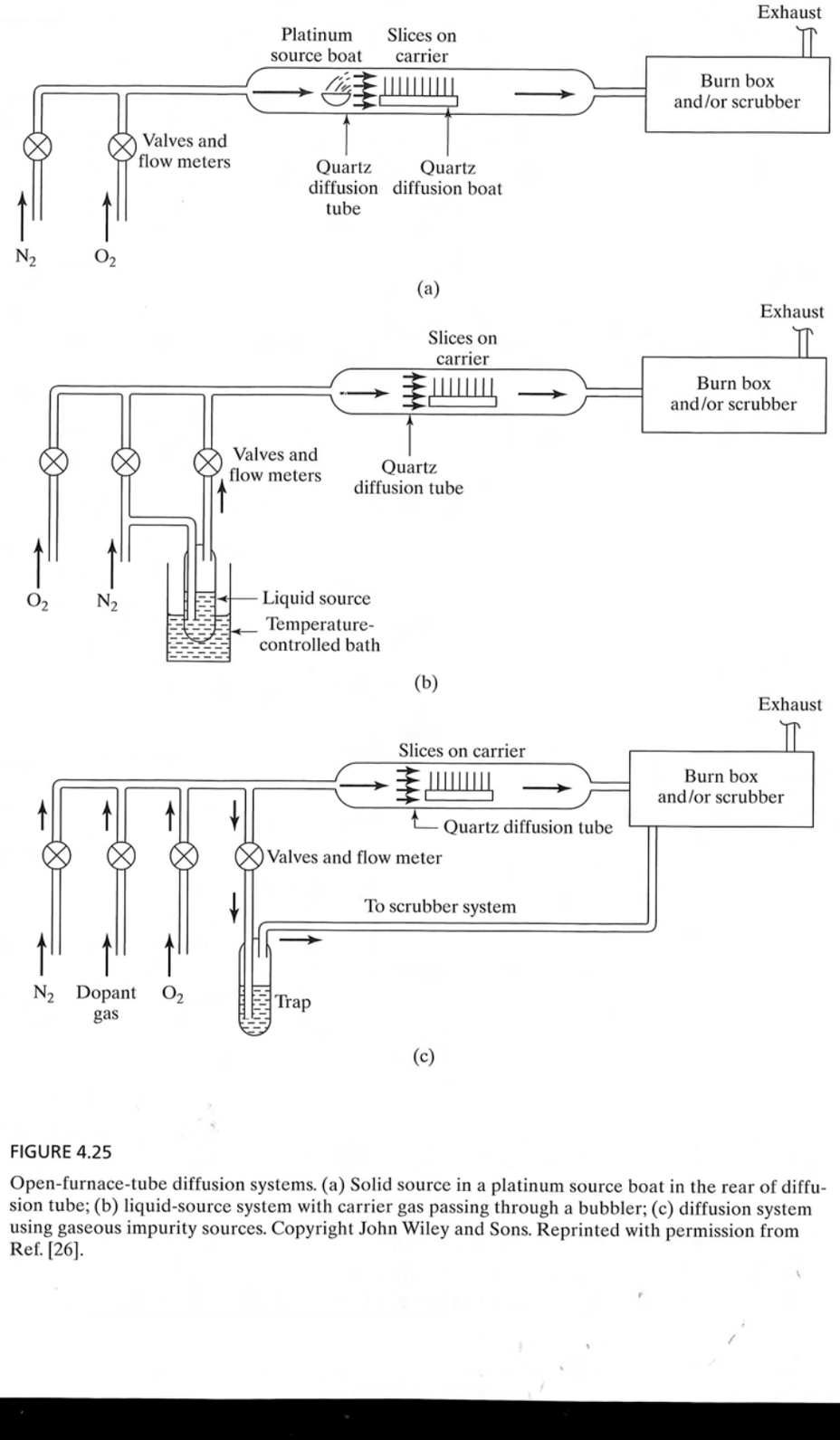


FIGURE 4.25

Open-furnace-tube diffusion systems. (a) Solid source in a platinum source boat in the rear of diffusion tube; (b) liquid-source system with carrier gas passing through a bubbler; (c) diffusion system using gaseous impurity sources. Copyright John Wiley and Sons. Reprinted with permission from Ref. [26].

impurity sources are also available in wafer form and are placed in the boat between adjacent pairs of silicon wafers.

In liquid-source systems, a carrier gas passes through a bubbler, where it picks up the vapor of the liquid source. The gas carries the vapor into the furnace tube, where it reacts with the surface of the silicon wafer.

Gas-source systems supply the dopant species directly to the furnace tube in the gaseous state. The common gas sources are extremely toxic, and additional input purging and trapping systems are required to ensure that all the source gas is removed from the system before wafer entry or removal. In addition, most diffusion processes either do not use all of the source gas or produce undesirable reaction by-products. Therefore, the output of diffusion systems must be processed by burning or by chemical or water scrubbing before being exhausted into the atmosphere.

Boron is the only commonly used *p*-type dopant. The diffusion coefficients of aluminum and gallium are quite high in silicon dioxide, and these elements cannot be masked effectively by SiO_2. Indium is not used, because it is a relatively deep-level acceptor ($E_A - E_V = 0.14$ eV).

In contrast, antimony, phosphorus, and arsenic can all be masked by silicon dioxide and are all routinely used as *n*-type dopants in silicon processing.

4.10.1 Boron Diffusion

Boron has a high solubility in silicon and can achieve active surface concentrations up to $4 \times 10^{20}/\text{cm}^3$. (See Fig. 4.6.) Elemental boron is inert up to temperatures exceeding the melting point of silicon. A surface reaction with boron trioxide (B_2O_3) is used to introduce boron to the silicon surface:

$$2B_2O_3 + 3Si \rightleftharpoons 4B + 3SiO_2. \tag{4.22}$$

An excess amount of boron trioxide can cause a brown boron skin to form that is very difficult to remove with most acids. Boron skin formation can be minimized by performing the diffusions in an oxidizing atmosphere containing 3 to 10% oxygen. In a two-step diffusion, the boron predeposition step is commonly followed by a short wet-oxidation step to assist in removal of the boron skin prior to drive-in.

Common solid sources of boron include trimethylborate (TMB) and boron nitride wafers. TMB is a solid with high vapor pressure at room temperature. The TMB source is normally placed outside the diffusion furnace and cooled below room temperature during use. TMB vapor reacts in the furnace tube with oxygen to form boron trioxide, water, and carbon dioxide:

$$2(CH_3O)_3B + 9O_2 \xrightarrow{900^\circ C} B_2O_3 + 6CO_2 + 9H_2O. \tag{4.23}$$

Unreacted TMB must be scrubbed from the exhaust stream.

Boron nitride is a solid source available in wafer form. Activated wafers are placed in every third slot in the same quartz boat used to hold the silicon wafers. A silicon wafer faces each side of the oxidized boron nitride wafer, and boron trioxide is transferred directly to the surface of the silicon wafer during high-temperature diffusion. A small flow of inert gas such as nitrogen is used to keep contaminants out of the tube during diffusion.

The most common liquid source for boron is boron tribromide (BBr_3). The reaction is

$$4BBr_3 + 3O_2 \rightarrow 2B_2O_3 + 6Br_2. \tag{4.24}$$

Free bromine combines easily with metallic impurities and is useful in removing (gettering) metallic impurities during diffusion. Bromine, as well as unused boron tribromide, is in the exhaust stream, so the outlet gases must be carefully cleaned.

The primary gaseous source of boron is diborane (B_2H_6). Diborane is a highly poisonous and explosive gas. Table 4.4 summarizes the ACGIH recommendations for the maximum permissible exposure to the common gases used as diffusion sources. Extreme care must be taken in using these gases. To reduce the risk of handling, diborane is usually diluted with 99.9% argon or nitrogen by volume.

Diborane oxidizes in either oxygen or carbon dioxide to form boron trioxide:

$$B_2H_6 + 3O_2 \xrightarrow{300^\circ C} B_2O_3 + 3H_2O \tag{4.25}$$

and

$$B_2H_6 + 6CO_2 \rightarrow B_2O_3 + 6CO + 3H_2O. \tag{4.26}$$

Both systems must provide a means for purging diborane from the input to the diffusion tube, and the output must be scrubbed to eliminate residual diborane and carbon monoxide.

4.10.2 Phosphorus Diffusion

Phosphorus has a higher solubility in silicon than does boron, and surface concentrations in the low $10^{21}/cm^3$ range can be achieved during high-temperature diffusion. Phosphorus is introduced into silicon through the reaction of phosphorus pentoxide at the wafer surface:

$$2P_2O_5 + 5Si \rightleftharpoons 4P + 5SiO_2. \tag{4.27}$$

Solid P_2O_5 wafers can be used as a solid source for phosphorus, as can ammonium monophosphate ($NH_4H_2PO_4$) and ammonium diphosphate [$(NH_4)_2H_2PO_4$] in wafer form. However, the most popular diffusion systems use either liquid or gaseous sources. Phosphorus oxychloride ($POCl_3$) is a liquid at room temperature. A carrier gas is passed through a bubbler and brings the vapor into the diffusion furnace. The gas stream also contains oxygen, and P_2O_5 is deposited on the surface of the wafers:

$$4POCl_3 + 3O_2 \rightarrow 2P_2O_5 + 6Cl_2. \tag{4.28}$$

Liberated chlorine gas serves as a gettering agent, and Cl_2 and $POCl_3$ must be removed from the exhaust stream.

TABLE 4.4 Threshold Limit Recommendations for Common Gaseous Sources [24] *

Source	8-h exposure level (ppm)	Life-threatening exposure	Comments
Diborane (B_2H_6)	0.10	160 ppm for 15 min	Colorless, sickly sweet, extremely toxic, flammable.
Phosphine (PH_3)	0.30	400 ppm for 30 min	Colorless, decaying fish odor, extremely toxic, flammable. A few minutes' exposure to 2000 ppm can be lethal.
Arsine (AsH_3)	0.05	6–15 ppm for 30 min	Colorless, garlic odor, extremely toxic. A few minutes' exposure to 500 ppm can be lethal.
Silane (SiH_4)	0.50	Unknown	Repulsive odor, burns in air, explosive, poorly understood.
Dichlorosilane (SiH_2Cl_2)	5.00	...	Colorless, flammable, toxic. Irritating odor provides adequate warning for voluntary withdrawal from contaminated areas.

*Data from the 1979 American Conference of Governmental Hygienists (ACGIH).

Phosphine, PH_3, is a highly toxic, explosive gas used as the gaseous source for phosphorus. It is also supplied in dilute form with 99.9% argon or nitrogen. Phosphine is oxidized with oxygen in the furnace:

$$2PH_3 + 4O_2 \rightarrow P_2O_5 + 3H_2O. \tag{4.29}$$

Unreacted phosphine must be cleaned from the exhaust gases, and the gas delivery system must be able to purge phosphine from the input to the tube.

4.10.3 Arsenic Diffusion

Arsenic has the highest solubility of any of the common dopants in silicon, with surface concentrations reaching $2 \times 10^{21}/cm^3$. The surface reaction involves arsenic trioxide:

$$2As_2O_3 + 3Si \rightleftharpoons 3SiO_2 + 4As. \tag{4.30}$$

Oxide vapors can be carried into the furnace tube from a solid diffusion source by a nitrogen carrier gas. However, evaporation of arsenic from the surface limits surface concentrations to below $3 \times 10^{19}/cm^3$. The exhaust must be carefully cleaned, because of the presence of arsenic.

Arsine gas may be used as a source, but it is extremely toxic and also produces relatively low surface concentrations. The problems with arsenic deposition and safety delayed its widespread use in silicon processing until ion implantation was developed in the early 1970s. Ion implantation is now the preferred technique for introducing arsenic into silicon. (Chapter 5 is devoted to the subject of ion implantation.)

4.10.4 Antimony Diffusion

Antimony, like arsenic, has a low diffusion coefficient and has been used for a long time for buried layers in bipolar processes. Antimony trioxide is a solid source that is placed in a two-zone furnace in which the source is maintained at a temperature of 600 to 650 °C. Antimony is introduced at the silicon surface as in the other cases:

$$2Sb_2O_3 + 3Si \rightleftharpoons 3SiO_2 + 4Sb. \tag{4.31}$$

A liquid source, antimony pentachloride (Sb_3Cl_5), has been successfully used with oxygen as a carrier gas passing through a bubbler. The gas stabine (SbH_3) is unstable and is not used for antimony diffusion.

4.11 GETTERING

A process called gettering is often used to improve the quality of the silicon wafer as one of the first steps in the fabrication process. Gettering is used to remove unwanted impurities, typically heavy metals such as copper, gold, iron, and nickel, from the surface where the devices will be fabricated. These unwanted impurities can reduce both lifetime and mobility in silicon. The heavy metals tend to be fast diffusers in silicon and have high solubility in heavily doped *n*-type silicon. In general, the gettering techniques attempt to provide locations away from the active device regions where the undesired impurities can precipitate and be immobilized.

A number of different backside gettering techniques have been used successfully. In one of the earliest techniques, a heavily doped phosphorus diffusion is applied to the back of the wafer. In older bipolar processes, this gettering step often occurred naturally during the emitter diffusion step. Damage to the backside of the wafer is also effective and can be introduced by sandblasting the back of the wafer. Internal wafer stress can assist the gettering process and can be introduced into the wafer by depositing thin layers (0.1–0.5 μm) of either silicon nitride or low-temperature polysilicon on the backside of the wafer. Implantation of argon atoms is another method used to introduce damage into the wafer.

Oxygen, at levels as high as $10^{18}/cm^3$, is incorporated into silicon wafers during the crystal growth process. This oxygen can combine with other undesired impurities to form precipitates, and this technique is commonly referred to as intrinsic gettering. To be effective, the oxygen level must be controlled reasonably well during crystal growth, and a specific heat treatment cycle must be utilized. Defects in the original crystal structure or those introduced by epitaxial growth can also provide gettering sites. An excellent introduction to gettering can be found in reference [31].

SUMMARY

In Chapter 4, we have discussed the formation of *pn* junctions using high-temperature diffusion. Mathematical models for diffusion have been presented, and the behavior of common *n*- and *p*-type dopants in silicon has been discussed. A key parameter governing the diffusion process is the diffusion coefficient, which is highly temperature dependent, following an Arrhenius relationship. Boron, phosphorus, antimony, and arsenic all have reasonable diffusion coefficients in silicon at temperatures between 900 and 1200° C, and they can be conveniently masked by a barrier layer of silicon dioxide. Gallium and aluminum are not easily masked by SiO_2 and are seldom used, and indium is not used, because of its large activation energy. At high concentrations, diffusion coefficients become concentration-dependent, causing diffused profiles to differ substantially from predictions of simple theories.

Two types of diffusions are most often used. If the surface concentration is maintained constant throughout the diffusion process, then a complementary error function (erfc) distribution is obtained. In the erfc case, the surface concentration is usually set by the solid-solubility limit of the impurity in silicon. If a fixed dose of impurity is diffused into silicon, a Gaussian diffusion profile is achieved. These two cases are often combined in a two-step process to obtain lower surface concentrations than those achievable with a solid-solubility-limited diffusion. As the complexity of fabrication processes grows, simulation with process modeling programs such as SUPREM is becoming ever more important. The SUPREM program includes comprehensive models for single- and multi-dimensional diffusion, and it represents an extremely useful tool for modeling advanced device structures.

The concept of sheet resistance has been introduced, and Irvin's curves have been used to relate the sheet resistance, junction depth, and surface concentration of diffused layers. Techniques for calculating and measuring junction depth have also been presented. Resistor fabrication has been discussed, including end and corner effects, as well as the effects of lateral diffusion under the edges of diffusion barriers.

High-temperature open-furnace diffusion systems are routinely used for diffusion with solid, liquid, and gaseous impurity sources. Boron, phosphorus, and antimony are all easily introduced into silicon using high-temperature diffusion. However, arsenic deposition by diffusion is much more difficult, and today it is usually accomplished using ion implantation. (See Chapter 5.) As with many chemicals used in IC fabrication, some of the sources used for diffusion are extremely toxic and must be handled with great care.

REFERENCES

[1] R. F. Pierret, *Semiconductor Fundamentals*, Volume I in the Modular Series on Solid State Devices, Addison-Wesley, Reading, MA, 1983.

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

[3] R. F. Pierret, *Advanced Semiconductor Fundamentals*, Volume VI in the Modular Series on Solid State Devices, Addison-Wesley, Reading, MA, 1987.

[4] D. P. Kennedy and R. R. O'Brien, "Analysis of the Impurity Atom Distribution Near the Diffusion Mask for a Planar *p–n* Junction," *IBM Journal of Research and Development*, *9*, 179–186 (May, 1965).

[5] J. C. Irvin, "Resistivity of Bulk Silicon and of Diffused Layers in Silicon," *Bell System Technical Journal*, *41*, 387–410 (March, 1962).

[6] L. R. Weisberg and J. Blanc, "Diffusion with Interstitial-Substitutional Equilibrium: Zinc in GaAs," *Physical Review*, *131*, 1548–1552 (August 15, 1963).

[7] R. B. Fair, "Boron Diffusion in Silicon—Concentration and Orientation Dependence, Background Effects, and Profile Estimation," *Journal of the Electrochemical Society*, *122*, 800–805 (June, 1975).

[8] R. B. Fair, "Profile Estimation of High-Concentration Arsenic Diffusions in Silicon," *Journal of Applied Physics*, *43*, 1278–1280 (March, 1972).

[9] R. B. Fair and J. C. C. Tsai, "Profile Parameters of Implanted-Diffused Arsenic Layers in Silicon," *Journal of the Electrochemical Society*, *123*, 583–586 (1976).

[10] J. C. C. Tsai, "Shallow Phosphorus Diffusion Profiles in Silicon," *Proceedings of the IEEE*, *57*, 1499–1506 (September, 1969).

[11] R. B. Fair and J. C. C. Tsai, "A Quantitative Model for the Diffusion of Phosphorus in Silicon and the Emitter Dip Effect," *Journal of the Electrochemical Society*, *124*, 1107–1118 (July, 1977).

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

[13] L. J. van der Pauw, "A Method of Measuring Specific Resistivity and Hall Effect of Discs of Arbitrary Shape," *Philips Research Reports*, *13*, 1–9 (February, 1958).

[14] R. Chwang, B. J. Smith, and C. R. Crowell, "Contact Size Effects on the van der Pauw Method for Resistivity and Hall Coefficient Measurements," *Solid-State Electronics*, *17*, 1217–1227 (December, 1974).

[15] P. F. Kane, and G. B. Larrabee, *Characterization of Semiconductor Materials*, McGraw-Hill Book Company, New York, 1970.

[16] (a) C. W. White and W. H. Cristie, "The Use of RBS and SIMS to Measure Dopant Profile Changes in Silicon by Pulsed Laser Annealing," *Solid-State Technology*, pp. 109–116, September 1980. (b) J. M. Anthony a et al., "Super SIMS for Ultra Sensitive Impurity Analysis," *Proceedings of the Materials Research Society Symposium 69*, pp. 311–316, 1986.

[17] A. Agarwal et al., "Boron-enhanced-Diffusion of Boron: The Limiting Factor for Ultra-shallow Junctions," *IEEE IEDM Technical Digest*, pp. 467–470, December 1997.

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

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

[20] D. Chin, M. Kump, H. G. Lee, and R. W. Dutton, "Process Design Using Coupled 2D Process and Device Simulators," *IEEE IEDM Digest*, pp. 223–226 (December, 1980).

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

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

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

[24] *Matheson Gas Data Book*, Matheson Gas Products, 1980.

[25] A. B. Glaser and G. E. Subak-Sharpe, *Integrated Circuit Engineering*, Addison-Wesley, Reading, MA, 1979.

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

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

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

[29] R. B. Fair, "Recent Advances in Implantation and Diffusion Modeling for the Design and Process Control of Bipolar ICs," *Semiconductor Silicon 1977*, *PV 77-2*, pp. 968–985.

[30] F. M. Smits, "Measurement of Sheet Resistivities with the Four Point Probe," *Bell System Technical Journal*, *37*, 711–718 (May, 1958).

[31] W. R. Runyan and K. E. Bean, *Semiconductor Integrated Circuit Processing Technology*, Addison Wesley Publishing Company, Reading, MA, 1990.

PROBLEMS

4.1 A phosphorus diffusion has a surface concentration of $5 \times 10^{18}/\text{cm}^3$, and the background concentration of the *p*-type wafer is $1 \times 10^{15}/\text{cm}^3$. The *Dt* product for the diffusion is 10^{-8} cm^2.

- **(a)** Find the junction depth for a Gaussian distribution.
- **(b)** Find the junction depth for an erfc profile.
- **(c)** What is the sheet resistance of the two diffusions?
- **(d)** Draw a graph of the two profiles.

4.2 A 5-hr boron diffusion is to be performed at 1100 °C.

- **(a)** What thickness of silicon dioxide is required to mask this diffusion?
- **(b)** Repeat part (a) for phosphorus.

4.3 A boron diffusion into a 1-ohm-cm *n*-type wafer results in a Gaussian profile with a surface concentration of $5 \times 10^{18}/\text{cm}^3$ and a junction depth of 4 μm.

- **(a)** How long did the diffusion take if the diffusion temperature was 1100 °C?
- **(b)** What was the sheet resistance of the layer?
- **(c)** What is the dose in the layer?
- **(d)** The boron dose was deposited by a solid-solubility-limited diffusion. Design a diffusion schedule (temperature and time) for this predeposition step.

4.4 The boron diffusion in Problem 4.3 is followed by a solid-solubility-limited phosphorus diffusion for 30 min at 950 °C. Assume that the boron profile does not change during the phosphorus diffusion.

(a) Find the junction depth of the new phosphorus layer. Assume an erfc profile.

(b) Find the junction depth based on the concentration-dependent diffusion data presented in Fig. 4.12.

(c) Calculate the total Dt product for Prob. 4.3 and compare the result to the Dt product for this problem. Is the assumption in the problem statement justified?

4.5 The p-well in a CMOS process is to be formed by a two-step boron diffusion into a 5-ohm-cm n-type substrate. The sheet resistance of the well is 1000 ohms per square and the junction depth is 7.5 μm.

(a) Design a reasonable diffusion schedule for the drive-in step that produces the p-well.

(b) What is the final surface concentration in the p-well?

(c) What is the dose required to form the well?

(d) Can this dose be achieved using a solid-solubility-limited diffusion with diffusion temperatures of 900 °C or above? Explain.

4.6 **(a)** Calculate the Dt product required to form a 0.2-μm-deep source–drain diffusion for an NMOS transistor using a solid-solubility-limited arsenic deposition at 1000° C into a wafer with a backgound concentration of $3 \times 10^{16}/\text{cm}^3$.

(b) What is the diffusion time? Does this time seem like a reasonable process?

(c) Recalculate Dt based upon the model in Fig. 4.20(e) and Table 4.2.

4.7 The channel length of a silicon-gate NMOS transistor is the spacing between the source and drain diffusions as shown in Fig. P4.7a. The spacing between the source and drain diffusion openings is 3 μm on the masking oxide used to make the transistor. The source/drain junctions are diffused to a depth of 0.5 μm using a constant-source diffusion. The surface concentration is $1 \times 10^{20}/\text{cm}^3$ and the wafer has a concentration of $1 \times 10^{16}/\text{cm}^3$. What is the channel length in the actual device after the diffusion is completed?

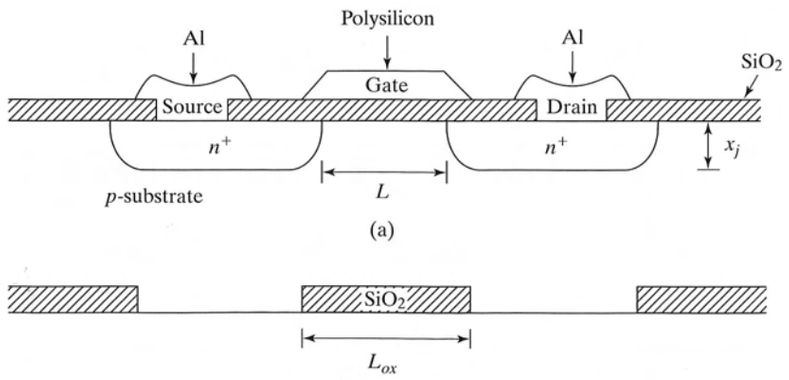


FIGURE P4.7

4.8 **(a)** What is the total number of squares in the resistor shown in Fig. P4.8, assuming that its geometry is specified precisely by the mask dimensions?

(b) The resistor is actually formed from a p-type base diffusion with a 6-μm junction depth. What is the actual number of squares in this resistor, assuming that the lateral diffusion under the edge of the mask is 5 μm.

(c) What would be the resistance of the resistors in parts (a) and (b) if the surface concentration of the base diffusion was 5×10^{18} boron atoms/cm^3, the bulk concentration $10^{15}/\text{cm}^3$, and the junction depth 6 μm.

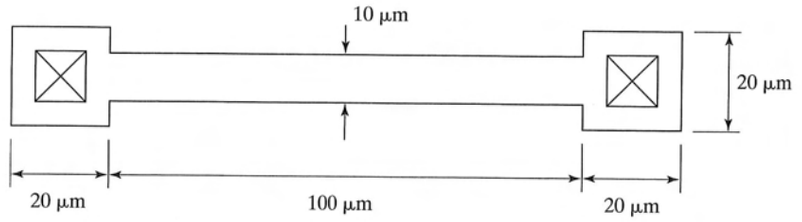


FIGURE P4.8

4.9 **(a)** What is the total number of squares in the resistor drawn in Fig. P4.9, assuming that its horizontal geometry is identical to that on the mask in the figure.

(b) The resistor is actually formed from a 3-μm-deep diffusion that has a surface concentration of $3 \times 10^{18}/\text{cm}^3$ in a wafer with background concentration of $10^{16}/\text{cm}^3$. What is actual number of squares in the resistor? (Use Fig. 4.10.)

(c) What is the resistance of the resistor if the diffusion is an n-type Guassian layer?

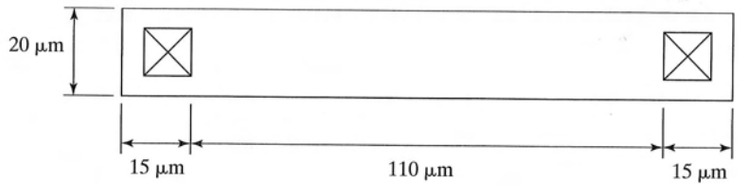


FIGURE P4.9

4.10 **(a)** Draw cross sections of the resistor in Fig. P4.10 through A′–A and B′–B if the lateral diffusion = 0.5 λ and vertical diffusion = λ.

(b) What is the total number of squares in the resistor as drawn on the masks?

(c) What is the actual number of squares in the resistor for the diffusions given in part (a)?

(d) Suppose the resistor is formed from a 2-μm-deep diffusion that has a surface concentration of $10^{19}/cm^3$ into a background concentration of $10^{16}/cm^3$. What is the actual number of squares in the resistor if λ = 2 μm? (Use Fig. 4.10.)

(e) A processing error occurred and the actual junction depth in part (d) is discovered to be 3 μm. Estimate the new number of squares in the resistor.

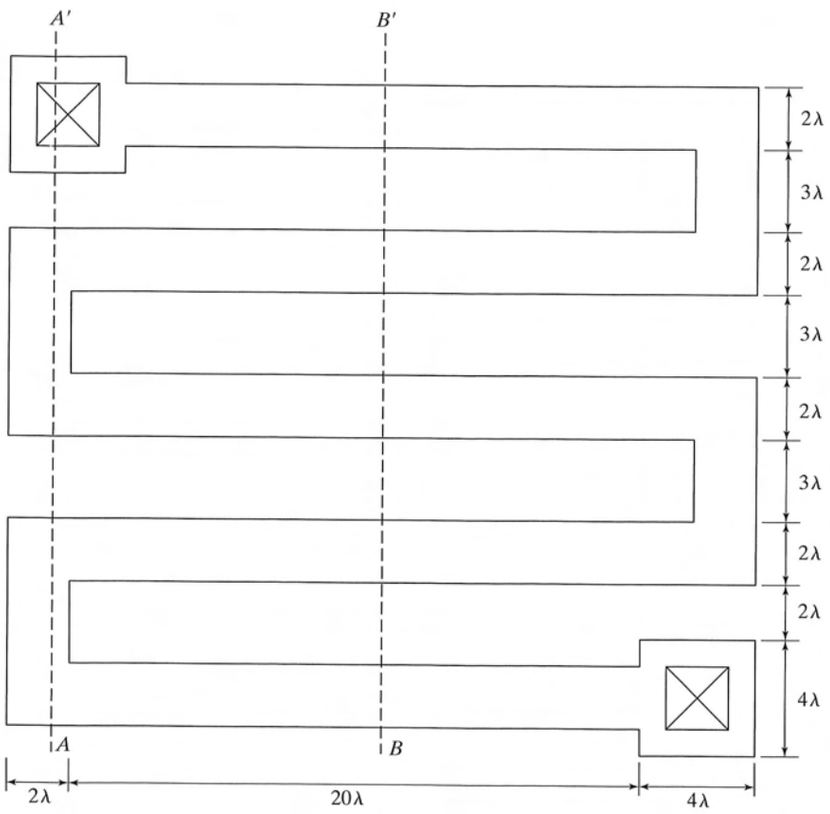


FIGURE P4.10

4.11 In practice, wafers are slowly pushed into and pulled out of the furnace, or the furnace temperature may be changed with time. Assume that the furnace temperature is being ramped down with time: $T = T_0 - Rt$, where T_0 is the initial temperature and R is the temperature change per second. Show that the effective Dt product defined by

$$(Dt)_{\text{eff}} = \int_0^{t_0} D(t)\, dt$$

where t_0 is the ramp-down time is given by

$$(Dt)_{\text{eff}} = D(T_0)(kT_0^2/RE_A)$$

where

$$D(T_0) = D_0 \exp(-E_A/kT_0)$$

4.12 Determine the sensitivity of junction depth to changes in furnace temperature by calculating $(dx_j/x_j)/(dT/T)$ for a Gaussian diffusion profile. What fractional change in junction depth will occur at 1100° C if the furnace temperature is in error by 10 °C?

4.13 Design (choose times and temperatures) a two-step diffusion to form a 5-μm-deep n-type layer with a surface concentration of $5 \times 10^{16}/cm^3$ in a 10 Ω-cm p-type substrate. (This layer could be the n-well for a CMOS process.)

4.14 Use Eq. (4.13) to calculate the sheet resistance of a p-type Gaussian diffusion having a surface concentration of $2 \times 10^{18}/cm^3$ and a junction depth of 2 μm. Assume that the hole mobility has a constant value of 300 cm^2/V-sec and $N_B = 10^{16}/cm^3$.

4.15 An n-type Gaussian diffusion has a surface conc. of $10^{20}/cm^3$ and a junction depth of 2 μm. $N_B = 5 \times 10^{16}/cm^3$. (a) Calculate the sheet resistance contribution of the first 0.5 μm of the layer, the second 0.5 μm, the third 0.5 μm, and the last 0.5 μm. (b) What is the sheet resistance of the total diffusion? Assume that the electron mobility has a constant value of 100 cm^2/V-sec. (c) Compare the calculation to the prediction from Irvin's curves.

4.16 Rework Example 4.3 using the concentration-dependent boron diffusion expressions for the predeposition calculations. Find the new surface concentration and junction depth following the drive-in step, and compare the results with those presented in the example. At 900 °C, $n_i = 4 \times 10^{18}/cm^3$.

4.17 Derive the expressions for the Gaussian and complementary-error-function solutions to the diffusion equation.

4.18 What is the minimum sheet resistance to be expected from shallow arsenic- and boron-doped regions if the regions are 1 μm deep? 0.25 μm deep? Make use of Figs. 4.6 and 4.16. Assume that the region is uniformly doped. Compare your results to the equations presented in Section 4.6.

4.19 Gold is diffused into a silicon wafer using a constant-source diffusion with a surface concentration of $10^{18}/cm^3$. How long does it take the gold to diffuse completely through a silicon wafer 400 μm thick with a background concentration of $10^{16}/cm^3$ at a temperature of 1000° C?

4.20 Use SUPREM to simulate the (two-dimensional) results of the two-step diffusion from Ex. 4.3 through a 5-μm-wide opening in a silicon dioxide barrier layer. Plot the results.

4.21 **(a)** Use SUPREM to simulate the diffusion profile of Example 4.3. Compare the simulation results with those given in the example.

(b) Follow the boron diffusion by the growth of a 500-nm layer of oxide in wet oxygen at 1100° C. Discuss what has happened to the boron concentration at the Si–SiO_2 interface.

(c) Add a 30-min solid-solubility-limited phosphorus diffusion at 950° C.

(d) The phosphorus diffusion created a new pn junction. Update the hand calculations for the boron impurity profile of Ex. 4.3 and estimate the location of both pn junctions with the aid of Fig. 4.12. Compare your results with those of SUPREM in part (c).

4.22 A current of 10 μA is injected into a van der Pauw structure having a sheet resistance of 300 Ω/□. What is the voltage that should be measured at the second set of terminals?

4.23 A gas cylinder contains 100 ft^3 of a mixture of diborane and argon. The diborane represents 0.1% by volume. An accident occurs and the complete cylinder is released into a room measuring $10 \times 12 \times 8$ ft.

(a) What will be the equilibrium concentration of diborane in the room in ppm?

(b) Compare this level with the toxic level based on Table 4.3.

(c) Would your answer to part (b) change if the gas cylinder contained arsine?

4.24 **(a)** Numerically calculate the sheet resistance of the diffusion in Problem 4.15 if the electron mobility can be described by

$$\mu_n = \left[92 + \frac{1270}{1 + \left(\frac{N}{1.3 \times 10^{17}} \right)^{.091}} \right] \frac{cm^2}{V-sec}$$

(b) Repeat for a similar p-type Gaussian layer in an n-type substrate if the hole moblility is given by

$$\mu_p = \left[48 + \frac{447}{1 + \left(\frac{N}{6.3 \times 10^{16}} \right)^{.076}} \right] \frac{cm^2}{V-sec}$$

CHAPTER 5

Ion Implantation

Ion implantation offers many advantages over diffusion for the introduction of impurity atoms into the silicon wafer and has become a workhorse technology in modern IC fabrication. In this chapter, we will first discuss ion implantation technology and mathematical modeling of the impurity distributions obtained with ion implantation. We will subsequently explore deviations from the model caused by nonideal behavior and will discuss annealing techniques used to remove crystal damage caused by the implantation process.

5.1 IMPLANTATION TECHNOLOGY

An ion implanter is a high-voltage particle accelerator producing a high-velocity beam of impurity ions that can penetrate the surface of silicon target wafers. The following list shows the basic parts of the system, shown schematically in Fig. 5.1, beginning with the impurity-source end of the system.

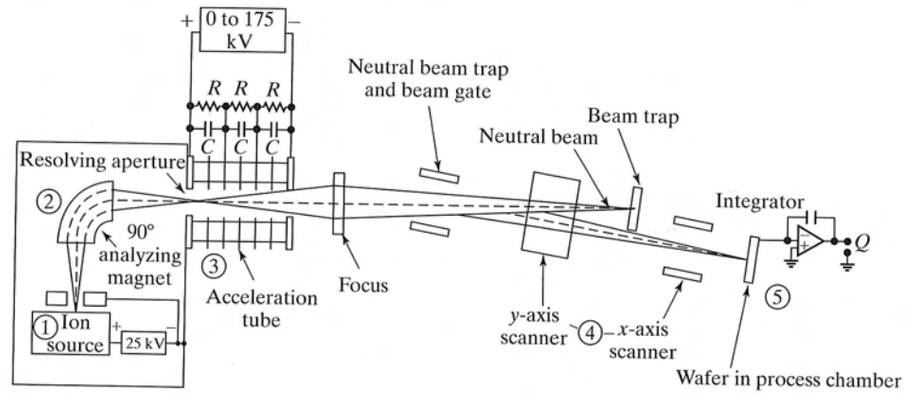


FIGURE 5.1

Schematic drawing of a typical ion implanter showing (1) ion source, (2) mass spectrometer, (3) high-voltage accelerator column, (4) *x*- and *y*-axis deflection system, and (5) target chamber.