

2

Motion and Recombination of Electrons and Holes

CHAPTER OBJECTIVES

The first chapter builds the necessary model for understanding semiconductors at equilibrium. This chapter will consider how the electrons and holes respond to an electric field and to a gradient in the carrier concentration. It is the response of charge carriers to these disturbances that gives life to the myriad of semiconductor devices. This chapter also introduces recombination and its opposite, generation. They are nature's ways of restoring the carrier concentrations to the equilibrium value by annihilating and creating electron-hole pairs.

2.1 • THERMAL MOTION •

Even without an applied electric field, carriers are not at rest but possess finite kinetic energies. The average **kinetic energy** of electrons, $E - E_c$, can be calculated in the following manner:

$$\begin{aligned} \text{Average electron kinetic energy} &= \frac{\text{total kinetic energy}}{\text{number of electrons}} \\ &= \frac{\int f(E)D(E)(E - E_c)dE}{\int f(E)D(E)dE} \end{aligned} \quad (2.1.1)$$

The integration in Eq. (2.1.1) is to be carried out over the conduction band, and the same approximations used in the derivation of Eq. (1.8.5) can be used here. The result is

$$\text{Average kinetic energy} = \frac{3}{2}kT \quad (2.1.2)$$

It can be shown that Eq. (2.1.2) is true for both electrons and holes.¹ The kinetic energy in Eq. (2.1.2) may be used to estimate the **thermal velocity**, v_{th} , of electrons and holes by equating the energy with $m_n v_{th}^2/2$ or $m_p v_{th}^2/2$.

$$v_{th} = \sqrt{\frac{3kT}{m}} \quad (2.1.3)$$

EXAMPLE 2-1 Thermal Velocity

What are the approximate thermal velocities of electrons and holes in silicon at room temperature?

SOLUTION: Assume $T = 300$ K and recall $m_n = 0.26 m_0$.

$$\text{Kinetic energy} = \frac{1}{2} m_n v_{th}^2 = \frac{3}{2} kT$$

$$\begin{aligned} v_{th} &= \sqrt{\frac{3kT}{m}} \\ &= (3 \times 1.38 \times 10^{-23} \text{ J/K} \times (300 \text{ K} / 0.26 \times 9.1 \times 10^{-31} \text{ kg}))^{1/2} \\ &= 2.3 \times 10^5 \text{ m/s} = 2.3 \times 10^7 \text{ cm/s} \end{aligned} \quad (2.1.3)$$

Note that $1 \text{ J} = 1 \text{ kg} \cdot \text{m}^2/\text{s}^2$. Using $m_p = 0.39 m_0$ instead of m_n , one would find the hole thermal velocity to be $2.2 \times 10^7 \text{ cm/s}$. So, the typical thermal velocity of electrons and holes is $2.5 \times 10^7 \text{ cm/s}$, which is about 1000 times slower than the speed of light and 100 times faster than the sonic speed.

Electrons and holes move at the thermal velocity but not in a simple straight-line fashion. Their directions of motion change frequently due to **collisions** or **scattering** with imperfections in the crystal, more about which will be said in Section 2.2. The carriers move in a zigzag fashion as shown in Fig. 2-1. The **mean free time** between collisions is typically 10^{-13} s or 0.1 ps (picosecond), and the distance between collisions is a few tens of nanometers or a few hundred angstroms. The *net* thermal velocity (averaged over time or over a large number of carriers at any given time) is zero. Thus, thermal motion does not create a steady electric current, but it does introduce a thermal noise.

¹ In fact, Eq. (2.1.2) is applicable to many kinds of particles and is known as the **equal-partition principle** because the kinetic energy of motion is equally partitioned among the three dimensions (x, y, z) – $kT/2$ for each direction.

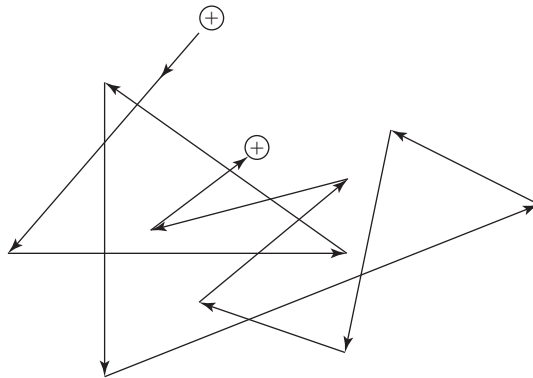


FIGURE 2-1 The thermal motion of an electron or a hole changes direction frequently by scattering off imperfections in the semiconductor crystal.

● Hot-Point Probe, Thermoelectric Generator and Cooler ●

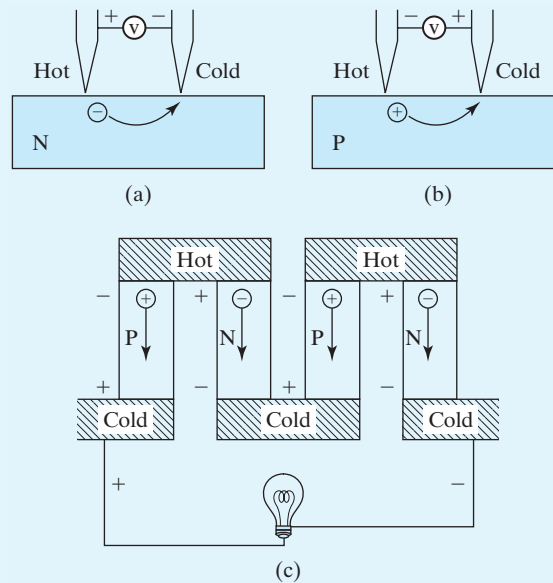


FIGURE 2-2 (a, b) Hot-point probe test can determine the doping type of a sample. (c) A thermoelectric generator converts heat into electric power.

Thermal motion can create a net current flow when there is a temperature difference. Figure 2-2a shows an N-type semiconductor sample. A cold (room-temperature) metal probe is placed on the sample close to a hot probe, perhaps a soldering iron. The electrons around the hot probe have higher thermal velocity and therefore on average move toward the cold side at a higher rate than the electrons on the cold side move to the hot side. The imbalance causes the electrons to accumulate on the cold side and build up a negative voltage, which is detected with a voltmeter.

Figure 2–2b shows that a positive voltage would be registered on the cold side if the sample is P-type. This is a practical and simple technique of testing the doping type of a semiconductor sample. It is called the **hot-point probe test**.

With engineering optimization, not only voltage but also significant electric power can be created with a temperature difference across P- and N-type semiconductor elements. Figure 2–2c is a schematic drawing of a **thermoelectric generator**. It powered some early space satellites using the radioactive decay of radioactive materials as the heat source. If, instead of extracting power from the device, current is fed into it, one set of the junctions become cooler than the other. The device is then a **thermoelectric cooler**. It can be used to cool a hot IC in a circuit board or beverages in a battery-powered portable cooler.

2.2 • DRIFT •

Drift is the motion of charge carriers caused by an electric field. Clearly, drift is usually at play when voltages are applied to a semiconductor device.

2.2.1 Electron and Hole Mobilities

The average velocity of the carriers is no longer zero when an electric field \mathcal{E} is applied to the semiconductor. This nonzero velocity is called the **drift velocity**. The drift velocity is superimposed on the thermal motion as illustrated in Fig. 2–3. The drift velocity is so much more important than the thermal velocity in semiconductor devices that the term velocity usually means the *drift* velocity. A faster carrier velocity is desirable, for it allows a semiconductor device or circuit to operate at a higher speed. We can develop a model for the drift velocity using Fig. 2–3. Consider the case for holes. Assume that the **mean free time between collisions is τ_{mp}** and that the carrier loses its entire drift momentum, $m_p v$, after each collision.² The drift momentum gained between collisions is equal to the force, $q\mathcal{E}$, times the mean free time. Equating the loss to the gain, we can find the steady state drift velocity, v .

$$m_p v = q\mathcal{E} \tau_{mp} \quad (2.2.1)$$

$$v = \frac{q\mathcal{E} \tau_{mp}}{m_p} \quad (2.2.2)$$

Equation (2.2.2) is usually written as

$$v = \mu_p \mathcal{E} \quad (2.2.3a)$$

$$\mu_p = \frac{q \tau_{mp}}{m_p} \quad (2.2.4a)$$

² Actually, it may take more than one collision for the carrier to lose its drift momentum. Another name for τ_{mp} , the *mean time for momentum relaxation*, is therefore more accurate. Although we will study a simplified analysis, be assured that a full analysis does lead to the same results as presented here.

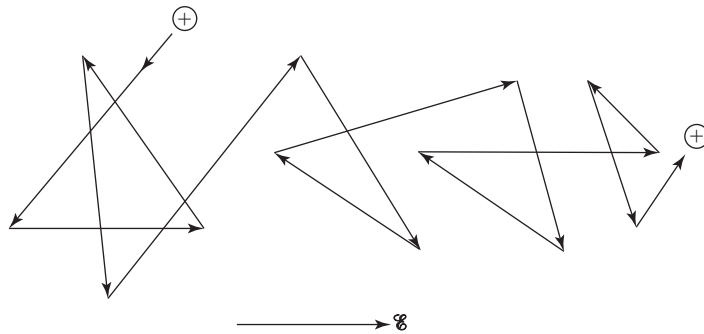


FIGURE 2-3 An electric field creates a drift velocity that is superimposed on the thermal velocity.

Equation (2.2.3a) simply says that the drift velocity is proportional to \mathcal{E} . The proportionality constant μ_p is the **hole mobility**, a metric of how mobile the holes are. Similarly, electron drift velocity and **electron mobility** are

$$v = -\mu_n \mathcal{E} \quad (2.2.3b)$$

$$\mu_n = \frac{q \tau_{mn}}{m_n} \quad (2.2.4b)$$

The negative sign in Eq. (2.2.3b) means that the electrons drift in a direction opposite to the field \mathcal{E} . They do so because the electron is negatively charged. We should memorize these statements rather than the negative sign.

Carrier mobility has the same dimension as v/\mathcal{E} , i.e., $\text{cm}^2/\text{V}\cdot\text{s}$. Table 2-1 shows some mobility values. Notice that GaAs has a much higher μ_n than Si (due to a smaller m_n). Thus, higher-speed transistors can be made with GaAs, which are typically used in communications equipment. InAs has an even higher μ_n , but the technology of fabricating InAs devices has not yet been fully developed.

TABLE 2-1 • Electron and hole mobilities at room temperature of selected lightly doped semiconductors.

	Si	Ge	GaAs	InAs
μ_n ($\text{cm}^2/\text{V}\cdot\text{s}$)	1400	3900	8500	30,000
μ_p ($\text{cm}^2/\text{V}\cdot\text{s}$)	470	1900	400	500

EXAMPLE 2-2 Drift Velocity, Mean Free Time, and Mean Free Path

Given $\mu_p = 470 \text{ cm}^2/\text{V}\cdot\text{s}$ for Si, what is the hole drift velocity at $\mathcal{E} = 10^3 \text{ V/cm}$? What is τ_{mp} and what is the average distance traveled between collisions, i.e., the **mean free path**?

SOLUTION: $v = \mu_p \mathcal{E} = 470 \text{ cm}^2/\text{V}\cdot\text{s} \times 10^3 \text{ V/cm} = 4.7 \times 10^5 \text{ cm/s}$

This is much lower than the thermal velocity, $\sim 2.1 \times 10^7 \text{ cm/s}$.

From Eq. (2.2.4b),

$$\begin{aligned}\tau_{mp} &= \mu_p m_p / q = 470 \text{ cm}^2 \times 0.39 \times 9.1 \times 10^{-31} \text{ kg} / 1.6 \times 10^{-19} \text{ C} \\ &= 0.047 \text{ m}^2 \times 2.2 \times 10^{-12} \text{ kg/C} = 1 \times 10^{-13} \text{ s} = 0.1 \text{ ps} \\ \text{Mean free path} &= \tau_{mp} v_{th} \sim 1 \times 10^{-13} \text{ s} \times 2.2 \times 10^7 \text{ cm/s} \\ &= 2.2 \times 10^{-6} \text{ cm} = 220 \text{ \AA} = 22 \text{ nm}\end{aligned}$$

2.2.2 Mechanisms of Carrier Scattering

We will now present a more detailed description of carrier collisions and show that τ_{mn} and τ_{mp} in Eq. (2.2.4) can vary significantly with temperature and the doping concentration.

What are the imperfections in the crystal that cause carrier collisions or scattering? There are two main causes: **phonon scattering** and **ionized impurity scattering**. Phonons are the particle representation of the vibration of the atoms in the crystal—the same sort of vibration that carries sound, hence the term phonons. Crystal vibration distorts the periodic crystal structure and thus scatters the electron waves. Instead of electron waves and vibration waves, it is more intuitive to think of electron particles scattering off phonon particles. The mobility due to phonon scattering alone, $\mu_{\text{phonon}} = q\tau_{\text{ph}}/m$, is proportional to τ_{ph} , the mean free time of phonon scattering. But, what determines the phonon scattering mean free time? Let us use the pinball machine for analogy. In a pinball machine, the mean time of collisions between the ball and the pins is inversely proportional to the density of the pins and the speed of the ball. Similarly, the mean free time of phonon scattering is inversely proportional to the phonon density and the electron speed, which is basically the thermal velocity. In addition, the phonon density is known to be proportional to the absolute temperature, T .

$$\mu_{\text{phonon}} \propto \tau_{\text{ph}}$$

$$\propto \frac{1}{\text{phonon density} \times \text{carrier thermal velocity}} \propto \frac{1}{T \cdot T^{1/2}} \propto T^{-3/2} \quad (2.2.5)$$

So, the **phonon scattering mobility decreases when temperature rises**. What about the impurity scattering mobility? The dopant ions are fixed charge in the semiconductor crystal. They can make electrons and holes change the direction of motion through the coulombic force. An electron can be scattered by either a donor (positive) ion or an acceptor (negative) ion as shown in Fig. 2-4. The same is true for a hole.

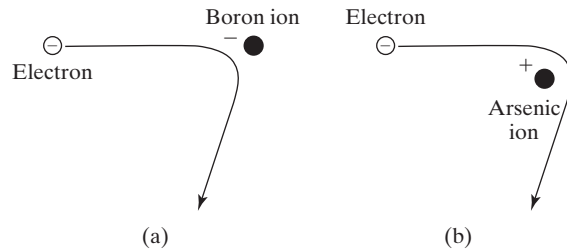


FIGURE 2-4 An electron can be scattered by an acceptor ion (a) and a donor ion (b) in a strikingly similar manner, even though the ions carry opposite types of charge. The same is true for a hole (not shown).

In Fig. 2-4a, the repulsive coulombic force between the electron and the negative ion deflects the motion of the electron. Figure 2-4b shows that an attractive coulombic force can induce the same effect on the electron trajectory.³ The mobility due to impurity scattering is therefore inversely proportional to the sum of the donor and acceptor ion concentrations. It is also proportional to $T^{3/2}$. Why is the mobility higher, i.e., the scattering weaker, at a higher T ? At a higher temperature, the electron in Fig. 2-4 has a higher thermal velocity and flies by the ion in a shorter time, and its direction of motion is thus less affected by the ion. A sports analogy: a ball carrier that charges by a blocker at high speed gives the blocker less of a chance to stop him.

$$\mu_{\text{impurity}} \propto \frac{T^{3/2}}{N_a + N_d} \quad (2.2.6)$$

When there is more than one scattering mechanism, the total scattering rate ($1/\tau$) and therefore the total mobility are determined by the sum of the inverses.

$$\frac{1}{\tau} = \frac{1}{\tau_{\text{phonon}}} + \frac{1}{\tau_{\text{impurity}}} \quad (2.2.7a)$$

$$\frac{1}{\mu} = \frac{1}{\mu_{\text{phonon}}} + \frac{1}{\mu_{\text{impurity}}} \quad (2.2.7b)$$

Figure 2-5 shows that the silicon hole mobility is about one-third of the electron mobility. Part of this difference in mobility can be explained by the difference in the effective mass (see Eq. (2.2.4)). The rest is attributable to the difference in the scattering mean free time. The mobilities may be expressed as (with N_a and N_d in per cubic centimeter) [1, 2].

$$\mu_p (\text{cm}^2/\text{V} \cdot \text{s}) = \frac{420}{1 + [(N_a + N_d)/1.6 \times 10^{17}]^{0.7}} + 50 \quad (2.2.8)$$

³ This is how a space probe uses the attractive gravitational force of a planet to change its course in a “slingshot” manner.

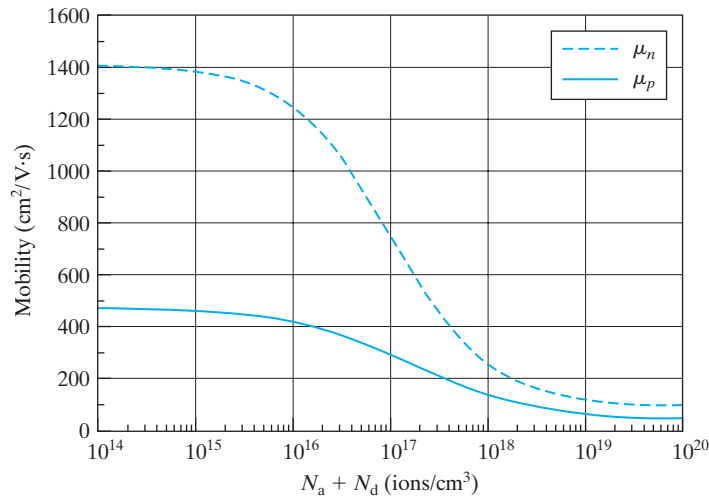


FIGURE 2-5 The electron and hole mobilities of silicon at 300 K. At low dopant concentration, the electron mobility is dominated by phonon scattering; at high dopant concentration, it is dominated by impurity ion scattering. (After [3].)

$$\mu_n(\text{cm}^2/\text{V} \cdot \text{s}) = \frac{1318}{1 + [(N_a + N_d)/1 \times 10^{17}]^{0.85}} + 92 \quad (2.2.9)$$

You may notice that the inverse proportionality to dopant density (Eq. (2.2.6)) is not followed in Eqs. (2.2.8) and (2.2.9) at the limit of very large N_a or N_d . The reason is **free-carrier screening**. When the carrier concentration is large, the carriers can distribute themselves to partially screen out the coulombic field of the dopant ions.

QUESTION • The electron mobility of an N-type silicon sample at room temperature is measured to be $600 \text{ cm}^2/\text{V} \cdot \text{s}$. Independent measurement shows that the electron concentration is $n = 5 \times 10^{16} \text{ cm}^{-3}$. According to Fig. 2-5, μ_n should be significantly larger than $600 \text{ cm}^2/\text{V} \cdot \text{s}$ if $N_d = 5 \times 10^{16} \text{ cm}^{-3}$. What do you think may be responsible for the discrepancy? Be as quantitative as you can. Hint: Consider possible dopant compensation (Section 1.9).

Figure 2-6 shows a schematic plot of the temperature dependence of μ_n . At small dopant concentrations, μ decreases with increasing T , indicative of the dominance of phonon scattering (Eq. (2.2.5)). At very high dopant concentration and low temperature, where impurity scattering is expected to dominate, μ indeed increases with increasing T (Eq. (2.2.6)). The mobility data in Figs. 2-5 and 2-6 (and Fig. 2-8) do not agree perfectly. This goes to show that it is not easy to measure mobility accurately and that we should presume the existence of uncertainties in experimental data in general.

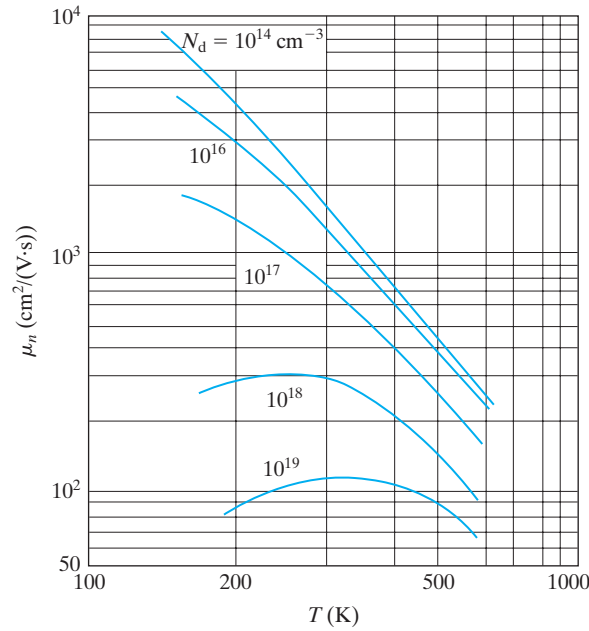


FIGURE 2-6 Temperature dependence of the electron mobility in Si. (After [4], reprinted by permission of John Wiley & Sons, Inc.)

• Velocity Saturation and Ballistic Transport •

In small devices, the electric field can easily reach 10^5 V/cm . If the electron mobility is $10^3 \text{ cm}^2/\text{V}\cdot\text{s}$, the drift velocity, according to $v = \mathcal{E}\mu$, should be 10^8 cm/s . In reality, electron and hole velocities saturate at around 10^7 cm/s and do not increase beyond that, no matter how large \mathcal{E} is (as shown in Figs. 6-20 and 6-21). The culprit is **optical phonon scattering**. Optical phonons are high-energy phonons that interact strongly with the electrons and holes. When the kinetic energy of a carrier exceeds the optical phonon energy, E_{opt} , it generates an optical phonon and loses the kinetic energy. Therefore, the velocity does not rise above **saturation velocity**, v_{sat} .

Mobility and even velocity saturation (see Section 2.2.2) are concepts that describe the carrier motion averaged over many scattering events. These concepts become fuzzy when we deal with devices whose sizes are smaller than the mean free path. The motion of carriers in a nearly scattering-free environment is called **ballistic transport**. Section 6.12 presents an example of this situation.

$$\frac{1}{2}mv_{\text{sat}}^2 \approx E_{\text{opt}}$$

$$v_{\text{sat}} \approx \sqrt{2E_{\text{opt}}/m}$$

E_{opt} is about 40 meV, which puts v_{sat} at around 10^7 cm/s . **Velocity saturation** has a deleterious effect on device speed as shown in Chapter 6.

2.2.3 Drift Current and Conductivity

Let us turn our attention to the current that flows in a semiconductor as a result of carrier drift. The current density, J , is the charge per second crossing a unit area plane normal to the direction of current flow. In the P-type semiconductor bar of unit area shown in Fig. 2–7, the hole current density is

$$J_{p,\text{drift}} = qp v \quad (2.2.10)$$

For example, if $p = 10^{15} \text{ cm}^{-3}$ and $v = 10^4 \text{ cm/s}$, then $J_{p,\text{drift}} = 1.6 \times 10^{-19} \text{ C} \times 10^{15} \text{ cm}^{-3} \times 10^4 \text{ cm/s} = 1.6 \text{ C/s} \cdot \text{cm}^2 = 1.6 \text{ A/cm}^2$.

Employing Eq. (2.2.3a), Eq. (2.2.10) can be written as

$$J_{p,\text{drift}} = qp\mu_p \mathcal{E} \quad (2.2.11)$$

Similarly, the electron current density can be expressed as

$$J_{n,\text{drift}} = -qn v = qn\mu_n \mathcal{E} \quad (2.2.12)$$

The total drift current density is the sum of the electron and the hole components:

$$J_{\text{drift}} = J_{n,\text{drift}} + J_{p,\text{drift}} = (qn\mu_n + qp\mu_p) \mathcal{E} \quad (2.2.13)$$

The quantity in the parentheses is the **conductivity**, σ , of the semiconductor

$$\sigma = qn\mu_n + qp\mu_p \quad (2.2.14)$$

Usually only one of the components in Eq. (2.2.14) is significant because of the large ratio between the majority and minority carrier densities. The resistivity, ρ , is the reciprocal of the conductivity. The standard units of σ and ρ are $\text{A/V} \cdot \text{cm}$ (or S/cm , S being siemens) and $\Omega \cdot \text{cm}$, respectively. ρ is shown as a function of the dopant density in Fig. 2–8.

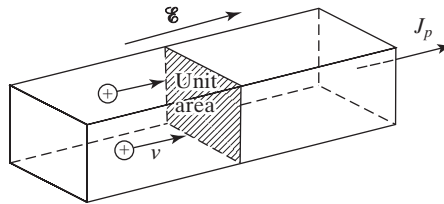


FIGURE 2–7 A P-type semiconductor bar of unit area is used to demonstrate the concept of current density.

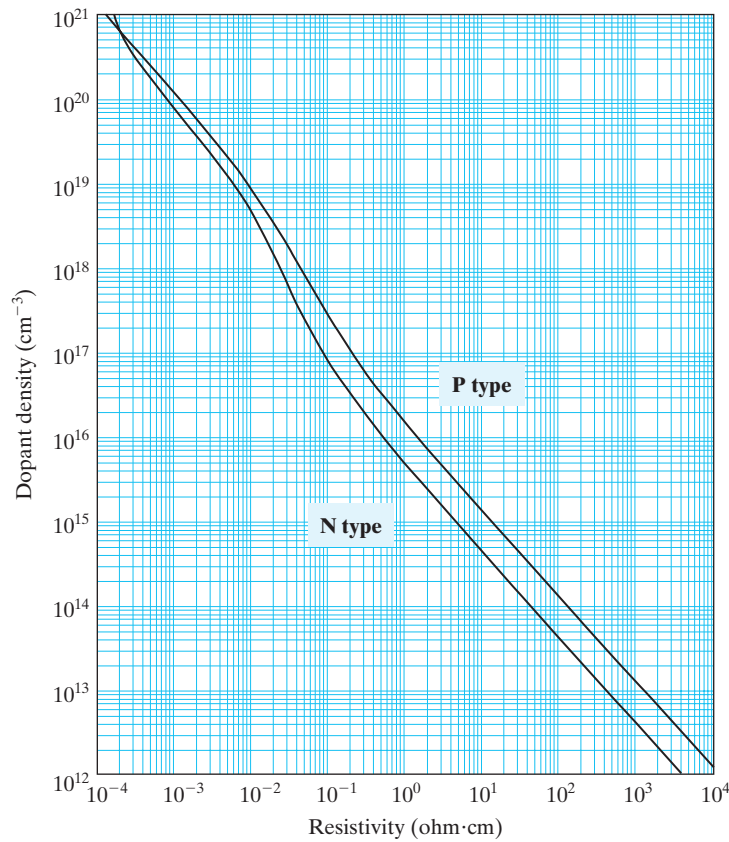


FIGURE 2-8 Conversion between resistivity and dopant density of silicon at room temperature. (After [3].)

EXAMPLE 2-3 Temperature Dependence of Resistance

- What is the resistivity, ρ , of silicon doped with 10^{17}cm^{-3} of arsenic?
- What is the resistance, R , of a piece of this silicon material $1\text{ }\mu\text{m}$ long and $0.1\text{ }\mu\text{m}^2$ in cross-sectional area?
- By what factor will R increase (or decrease) from $T = 300\text{ K}$ to $T = 400\text{ K}$?
- What As concentration should one choose if she wishes to minimize the change in (c)?

SOLUTION:

- Using the N-type curve in Fig. 2-8, we find that $\rho = 0.084\text{ }\Omega\cdot\text{cm}$. You can also answer this question using Fig. 2-5 and Eq. (2.2.14).
- $R = \rho \times \text{length} / \text{area} = 0.084\text{ }\Omega\cdot\text{cm} \times 1\text{ }\mu\text{m} / 0.1\text{ }\mu\text{m}^2$
 $= 0.084\text{ }\Omega\cdot\text{cm} \times 10^{-4}\text{cm} / 10^{-9}\text{cm}^2 = 8.4 \times 10^3\text{ }\Omega$.

- c. The temperature dependent factor in σ (Eq. (2.2.14)) (and therefore in ρ) is μ_n . Figure 2-6 (10^{17}cm^{-3} curve) shows μ_n to decrease from 770 at 300 K to 400 at 400K. We conclude that R **increases** by

$$\frac{770}{400} = 1.93$$

- d. The 10^{19}cm^{-3} curve indicates nearly equal mobilities at 300 and 400K. Therefore, that N_d would be a good choice. $1.1 \times 10^{19}\text{cm}^{-3}$ would be an even better choice.

2.3 • DIFFUSION CURRENT •

In addition to the drift current, there is a second component of current called the **diffusion current**. Diffusion current is generally not an important consideration in metals because of their high conductivities. The low conductivity and the ease of creating nonuniform carrier densities make diffusion an important process in semiconductors.

Diffusion is the result of particles undergoing thermal motion as depicted in Fig. 2-1. It is the familiar process by which particles move from a point of higher particle density toward a point of lower density, as shown in Fig. 2-9. The aroma of a cup of coffee travels across a room by the diffusion of flavor molecules through the air. A drop of soy sauce spreads in a bowl of clear soup by diffusion, too.

It is known that the rate of particle movement by diffusion is proportional to the concentration gradient. If the electron concentration is not uniform, there will be an electron diffusion current, which is proportional to the gradient of the electron concentration.

$$J_{n,\text{diffusion}} \propto \frac{dn}{dx} \quad (2.3.1)$$

$$J_{n,\text{diffusion}} = qD_n \frac{dn}{dx} \quad (2.3.2)$$

We have introduced the proportional constant qD_n . q is the elementary charge ($+1.6 \times 10^{-19}\text{C}$), and D_n is called the electron **diffusion constant**. The larger the D_n

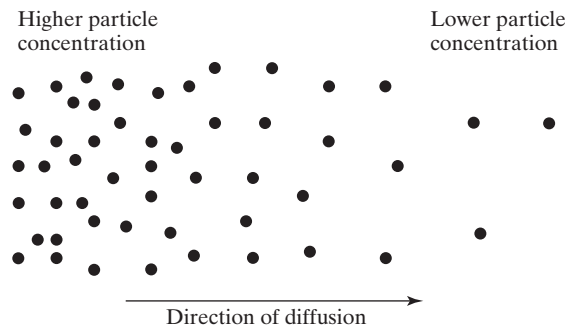


FIGURE 2-9 Particles diffuse from high-concentration locations toward low-concentration locations.

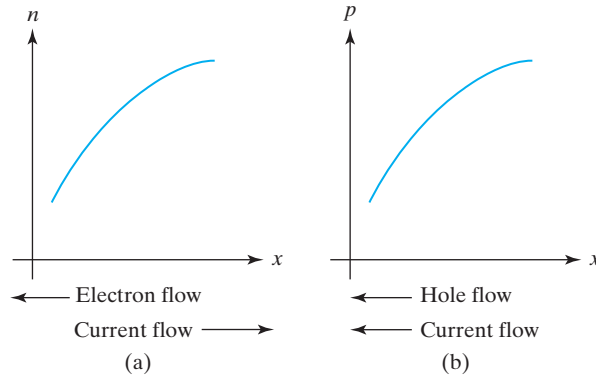
2.4 • Relation Between the Energy Diagram and V, \mathcal{E} 

FIGURE 2-10 A positive slope of carrier concentration produces a positive electron diffusion current (a), but a negative hole diffusion current (b).

is, the faster the electrons diffuse. In Section 2.4, we will see what determines D_n . For holes,

$$J_{p,\text{diffusion}} = -qD_p \frac{dp}{dx} \quad (2.3.3)$$

Equation (2.3.3) has a negative sign, while Eq. (2.3.2) has a positive sign. Instead of memorizing the signs, memorize Fig. 2-10. In Fig. 2-10, (a) shows a positive dn/dx (n increases as x increases) and (b) shows a positive dp/dx . In (a), electrons diffuse to the left (toward the lower concentration point). Because electrons carry negative charge, the diffusion current flows to the *right*. In (b), holes diffuse to the left, too. Because holes are positively charged, the hole current flows to the *left*, i.e., the current is negative.

In general, both drift and diffusion may contribute to the current. Therefore,

$$J_n = J_{n,\text{drift}} + J_{n,\text{diffusion}} = qn\mu_n \mathcal{E} + qD_n \frac{dn}{dx} \quad (2.3.4)$$

$$J_p = J_{p,\text{drift}} + J_{p,\text{diffusion}} = qp\mu_p \mathcal{E} - qD_p \frac{dp}{dx} \quad (2.3.5)$$

$$J = J_n + J_p \quad (2.3.6)$$

2.4 • RELATION BETWEEN THE ENERGY DIAGRAM AND V, \mathcal{E} •

When a voltage is applied across a piece of semiconductor as shown in Fig. 2-11a, it alters the band diagram. By definition, a positive voltage raises the potential energy of a positive charge and lowers the energy of a negative charge. It therefore lowers the energy diagrams since the energy diagram plots the energy of an electron (a negative charge). Figure 2-11c shows that the energy diagram is lower (at the left) where the voltage is higher. The band diagram is higher where the voltage is lower. (E_c and E_v are always separated by a constant, E_g .) *The point to remember is that E_c and E_v vary in the opposite direction from the voltage. E_c and E_v are higher where the voltage is lower.* That is to say

$$E_c(x) = \text{constant} - qV(x) \quad (2.4.1)$$

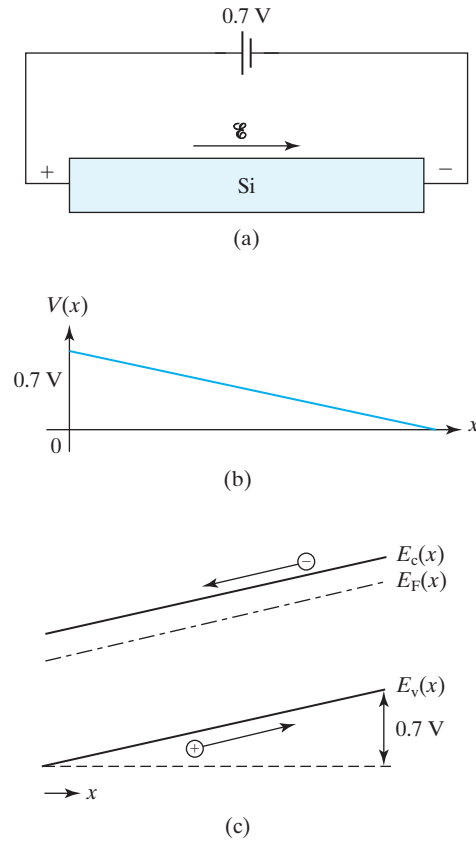


FIGURE 2-11 Energy band diagram of a semiconductor under an applied voltage. 0.7 eV is an arbitrary value.

The q takes care of the difference between the units, eV and V. The “constant” takes care of the unspecified and inconsequential zero references for E_c and V . The “constant” drops out when one considers the electric field.

$$\mathcal{E}(x) \equiv -\frac{dV}{dx} = \frac{1}{q} \frac{dE_c}{dx} = \frac{1}{q} \frac{dE_v}{dx} \quad (2.4.2)$$

In other words, the slope of E_c and E_v indicates the electric field. The direction of \mathcal{E} in Fig. 2-11c is consistent with Eq. (2.4.2). Figure 2-11c suggests the following analogies: *the electrons roll downhill like stones in the energy band diagram and the holes float up like bubbles.*

2.5 • EINSTEIN RELATIONSHIP BETWEEN D AND μ •

Consider a bar of semiconductor, whose band diagram is shown in Fig. 2-12. The semiconductor is at equilibrium, and therefore the Fermi level E_F is constant (Section 1.7). The left side is more heavily doped than the right side, and so E_c is

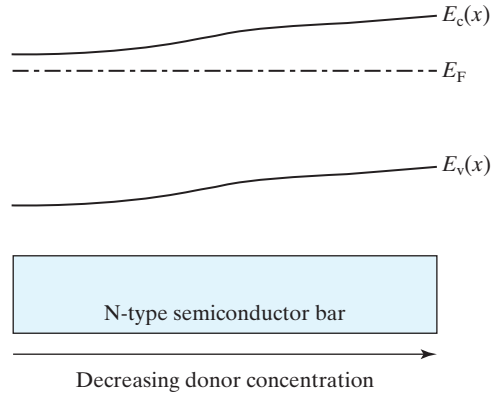


FIGURE 2-12 A piece of N-type semiconductor in which the dopant density decreases toward the right.

closer to E_F on the left side. Because E_c is not a constant, there is an electric field equal to $(1/q) dE_c/dx$, according to Eq. (2.4.2). This field is internally created and is as real as a field created by an external voltage. Because the semiconductor is at equilibrium, there cannot be any J_n (or J_p). From Eq. (2.3.4),

$$J_n = 0 = qn\mu_n \mathcal{E} + qD_n \frac{dn}{dx} \quad (2.5.1)$$

Recall Eq. (1.8.5)

$$n = N_c e^{-(E_c - E_F)/kT} \quad (1.8.5)$$

$$\frac{dn}{dx} = \frac{-N_c}{kT} e^{-(E_c - E_F)/kT} \frac{dE_c}{dx} \quad (2.5.2)$$

$$= \frac{-n}{kT} \frac{dE_c}{dx} \quad (\text{Eq. (1.8.5) is used}) \quad (2.5.3)$$

$$= \frac{-n}{kT} q \mathcal{E} \quad (\text{Eq. (2.4.2) is used}) \quad (2.5.4)$$

dn/dx in Eq. (2.5.1) will now be substituted with Eq. (2.5.4)

$$0 = qn\mu_n \mathcal{E} - qn \frac{qD_n}{kT} \mathcal{E} \quad (2.5.5)$$

$$\therefore \boxed{D_n = \frac{kT}{q} \mu_n} \quad (2.5.6a)$$

The electron drift and diffusion currents will perfectly cancel each other out for an arbitrary doping profile, if and only if Eq.(2.5.6a) is satisfied. This remarkably simple relationship between the diffusion constant, D_n , and mobility was derived by Albert Einstein. A close relationship between μ and D becomes plausible when one realizes that all scattering mechanisms (e.g. phonon and

impurity scatterings) that impede electron drift would also impede electron diffusion. Equation (2.5.6a) and its counterpart for holes

$$D_p = \frac{kT}{q} \mu_p \quad (2.5.6b)$$

are known as the **Einstein relationship**.

EXAMPLE 2-4 Diffusion Constant

What is the hole diffusion constant in a piece of silicon doped with $3 \times 10^{15} \text{cm}^{-3}$ of donors and $7 \times 10^{15} \text{cm}^{-3}$ of acceptors at 300 K? at 400 K?

SOLUTION: Figure 2-5 shows that, for $N_d + N_a = 3 \times 10^{15} + 7 \times 10^{15} = 1 \times 10^{16} \text{cm}^{-3}$, μ_p is about $410 \text{ cm}^2/\text{V}\cdot\text{s}$ at 300 K. $D_p = (kT/q)\mu_p = 26 \text{ mV} \times 410 \text{ cm}^2/\text{V}\cdot\text{s} = 11 \text{ cm}^2/\text{s}$. Do you remember $kT/q = 26 \text{ mV}$ at room temperature?

For the 400 K case, we turn to Fig. 2-6, which shows for $N = 10^{16} \text{cm}^{-3}$ and $T = 127^\circ\text{C}$, $\mu_p = 220 \text{ cm}^2/\text{V}\cdot\text{s}$.

$$D_p = \mu_p(kT/q) = 220 \text{ cm}^2/\text{V}\cdot\text{s} \times 26 \text{ mV} \times (400 \text{ K}/300 \text{ K}) = 7.6 \text{ cm}^2/\text{V}\cdot\text{s}.$$

2.6 • ELECTRON-HOLE RECOMBINATION •

The electron and hole concentrations introduced in Chapter 1 (for example, $n = N_d$ and $p = n_i^2/N_d$ for an N-type sample) are the **equilibrium carrier concentrations** and they will be denoted with n_0 and p_0 from now on. The electron and hole concentrations can be different from n_0 and p_0 , for example, when light shines on the sample and generates electrons and holes (as shown in Fig. 1-12). The differences are known as the **excess carrier concentrations** denoted by n' and p' .

$$n \equiv n_0 + n' \quad (2.6.1a)$$

$$p \equiv p_0 + p' \quad (2.6.1b)$$

If n' and p' are created by light, n' and p' are equal because the electrons and holes are created in pairs. If n' and p' are introduced by other means, they will still be equal to each other because of **charge neutrality** (Eq. (1.9.1)). Since charge neutrality is satisfied at equilibrium when $n' = p' = 0$, any time a non-zero n' is present, an equal p' must be present to maintain the charge neutrality, and vice versa. Otherwise, the net charge will attract or repel the abundant majority carriers until neutrality is restored. This conclusion for a charge neutral sample can be written as

$$n' \equiv p' \quad (2.6.2)$$

If the light is suddenly turned off, n' and p' will decay with time until they become zero and n and p return to their equilibrium values, n_0 and p_0 . The process of decay is **recombination**, whereby an electron and a hole recombine and annihilate each other. The time constant of the decay is called the **recombination time** or **carrier lifetime**, τ .

$$\frac{dn'}{dt} = -\frac{n'}{\tau} = -\frac{p'}{\tau} \quad (2.6.3)$$

The recombination rate (per cubic centimeter per second) is proportional to n' and p' .

$$\text{Recombination rate} = \frac{n'}{\tau} = \frac{p'}{\tau} \quad (2.6.4)$$

τ has the dimension of time and is typically around $1\mu\text{s}$ in Si. It may range from 1 ns to 1 ms, depending on the density of trace metal impurities such as Au and Pt, which form traps in the band gap with several energy levels deep in the band gap. These **deep traps** can capture electrons or holes to facilitate recombination (as shown in Fig. 2–13) and thus shorten the recombination time (they are also called **recombination centers**). Too small a τ is bad for device leakage current, and hence extreme cleanliness is maintained in the semiconductor fabrication plants partly to avoid these metallic contaminants. The other recombination process shown in Fig. 2–13, **direct recombination**, or **radiative recombination**, is very inefficient and unimportant in silicon because the electrons and holes at the edges of the band gap do not have the same wave vectors (see Section 1.5.1 and Figure 4–27b). These types of semiconductors are called **indirect gap semiconductors**. The radiative recombination process is very efficient in **direct gap semiconductors** such as GaAs because the electrons and holes have the same wave vectors (see Figure 4–27a) and is responsible for light emission in light-emitting diodes and lasers, which will be presented in Chapter 4.

Echoes in a canyon eventually die out. So do ripples produced by a cast stone. Nature provides ways to restore equilibrium. Recombination is nature's way of restoring n' and p' to zero.

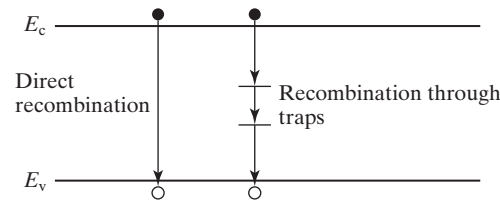


FIGURE 2–13 An electron–hole pair recombines when an electron drops from the conduction band into the valence band. In silicon, direct recombination is unimportant and the lifetime is highly variable and determined by the density of recombination centers.

EXAMPLE 2–5 Photoconductors

A bar of Si is doped with boron at 10^{15}cm^{-3} . It is exposed to light such that electron–hole pairs are generated throughout the volume of the bar at the rate of $10^{20}/\text{s}\cdot\text{cm}^3$. The recombination lifetime is $10\mu\text{s}$. What are (a) p_0 , (b) n_0 , (c) p' , (d) n' , (e) p , (f) n , and (g) the np product? (h) If the light is suddenly turned off at $t = 0$, find $n'(t)$ for $t > 0$.

SOLUTION:

a. $p_0 = N_a = 10^{15}\text{cm}^{-3}$ is the equilibrium hole concentration.

b. $n_0 = n_i^2/p_0 \approx 10^5\text{cm}^{-3}$ is the equilibrium electron concentration.

- c. In steady state, the rate of generation is equal to the rate of recombination (Eq. (2.6.4)).

$$10^{20}/\text{s}\cdot\text{cm}^3 = p'/\tau$$

$$\therefore p' = 10^{20}/\text{s}\cdot\text{cm}^3 \cdot t = 10^{20}/\text{s}\cdot\text{cm}^3 \cdot 10^{-5}\text{s} = 10^{15}\text{cm}^{-3}$$

- d. $n' = p' = 10^{15}\text{cm}^{-3}$ (Eq. (2.6.2))

e. $p = p_0 + p' = 10^{15}\text{cm}^{-3} + 10^{15}\text{cm}^{-3} = 2 \times 10^{15}\text{cm}^{-3}$

- f. $n = n_0 + n' = 10^5\text{cm}^{-3} + 10^{15}\text{cm}^{-3} \approx 10^{15}\text{cm}^{-3}$. **The non-equilibrium minority carrier concentration is often much much larger than the small equilibrium concentration.**

- g. $np \approx 2 \times 10^{15}\text{cm}^{-3} \times 10^{15}\text{cm}^{-3} = 2 \times 10^{30}\text{cm}^{-6} \gg n_i^2 = 10^{20}\text{cm}^{-6}$. **The np product can be very different from n_i^2 .**

- h. The solution to Eq. (2.6.4) is

$$n'(t) = n'(0)e^{-t/\tau} = 10^{15}\text{cm}^{-3}e^{-t/\tau}$$

Therefore, n' decays exponentially toward its equilibrium value of zero. The characteristic time of the exponential decay is the carrier lifetime, τ .

2.7 • THERMAL GENERATION •

The reverse process of recombination is called **thermal generation**. At any nonzero temperature, electron-hole pairs are constantly being generated *and* lost (by recombination). If $n' = p' = 0$, the rate of recombination equals the rate of generation and the net rate of change is zero. If n' is positive, there is a net recombination rate as shown in Eq. (2.6.4). If n' is negative, i.e., there are fewer electrons than the equilibrium concentration, nature sees to it that there is a net rate of thermal generation rather than recombination. Equation (2.6.3) confirms this fact by predicting a positive dn'/dt . Later we will see that thermal generation is responsible for the leakage current in diodes.

When the np product is equal to n_i^2 , the rate of thermal generation is equal to the rate of recombination. Under this condition, n and p are said to be at thermal equilibrium. When $np > n_i^2$, there is net recombination; when $np < n_i^2$, there is net generation. The terms recombination and generation rates generally refer to the *net* rates of recombination and generation.

2.8 • QUASI-EQUILIBRIUM AND QUASI-FERMI LEVELS •

Whenever $np \neq n_i^2$, the semiconductor is not at equilibrium. More precisely, the electrons and holes are not at equilibrium with each other. Nonetheless, we would like to preserve and use, as much as possible, the following equilibrium relationships, which we have found to be very useful.

$$n = N_c e^{-(E_c - E_F)/kT} \quad (1.8.5)$$

$$p = N_v e^{-(E_F - E_v)/kT} \quad (1.8.8)$$

The problem is that the above equations, when multiplied together, lead to $np = n_i^2$. We saw in Example 2-5(g) that the presence of excess carriers can easily make the np product much larger than n_i^2 . This problem can be addressed by introducing two **quasi-Fermi levels**, E_{Fn} and E_{Fp} , such that

$$n = N_c e^{-(E_c - E_{Fn})/kT} \quad (2.8.1)$$

$$p = N_v e^{-(E_{Fp} - E_v)/kT} \quad (2.8.2)$$

E_{Fn} and E_{Fp} are the electron and hole quasi-Fermi levels. When electrons and holes are at equilibrium, i.e., when $np = n_i^2$, E_{Fn} and E_{Fp} coincide and this is known as E_F . Otherwise, $E_{Fn} \neq E_{Fp}$. Equations (2.8.1) and (2.8.2) indicate that even when electrons and holes, as two groups, are not at equilibrium with each other, the electrons (and holes) can still be at equilibrium among themselves. Electrons and holes, as two groups of particles, can get out of equilibrium easily because they are only loosely coupled by the recombination/generation mechanism, which is a slow process (has a long time constant around 1 μ s). In contrast, the electrons (or holes) are strongly coupled among themselves by exchanging positions and energy through thermal motion at high speed and by scattering with 0.1 ps mean free time. The usefulness of this **quasi-equilibrium** concept will become clear in later applications.

EXAMPLE 2-6 Quasi-Fermi Levels and Low-Level Injection

Consider an Si sample with $N_d = 10^{17} \text{ cm}^{-3}$.

- Find the location of E_F .
- Find the location of E_{Fn} and E_{Fp} when excess carriers are introduced such that $n' = p' = 10^{15} \text{ cm}^{-3}$.

Notice that n' and p' are much less than the majority carrier concentration. This condition is commonly assumed and is called **low-level injection**. The opposite condition, **high-level injection**, is often encountered in bipolar transistors (Section 8.4.4).

SOLUTION:

- Using Eq. (1.8.5)

$$n = N_d = 10^{17} \text{ cm}^{-3} = N_c e^{-(E_c - E_F)/kT}$$

$$E_c - E_F = kT \ln \frac{N_c}{10^{17} \text{ cm}^{-3}} = 26 \text{ meV} \cdot \ln \frac{2.8 \times 10^{19} \text{ cm}^{-3}}{10^{17} \text{ cm}^{-3}} = 0.15 \text{ eV}$$

E_F is below E_c by 0.15 eV.

b. $n = n_0 + n' = N_d + n' = 1.01 \times 10^{17} \text{ cm}^{-3}$

Using Eq. (2.8.1),

$$1.01 \times 10^{17} \text{ cm}^{-3} = N_c e^{-(E_c - E_{Fn})/kT}$$

$$E_c - E_{Fn} = kT \ln \frac{N_c}{1.01 \times 10^{17} \text{ cm}^{-3}} = 26 \text{ meV} \cdot \ln \frac{2.8 \times 10^{19} \text{ cm}^{-3}}{1.01 \times 10^{17} \text{ cm}^{-3}} = 0.15 \text{ eV}$$

E_{Fn} is basically unchanged from the E_F in (a) as illustrated in Fig. 2–14.

$$p = p_0 + p' = \frac{n_i^2}{N_d} + p' = 10^3 \text{ cm}^{-3} + 10^{15} \text{ cm}^{-3} = 10^{15} \text{ cm}^{-3}$$

Using Eq. (2.8.2)

$$10^{15} \text{ cm}^{-3} = N_v e^{-(E_{Fp} - E_v)/kT}$$

$$E_{Fp} - E_v = kT \ln \frac{N_v}{10^{15} \text{ cm}^{-3}} = 26 \text{ meV} \cdot \ln \frac{1.04 \times 10^{19} \text{ cm}^{-3}}{10^{15} \text{ cm}^{-3}} = 0.24 \text{ eV}$$

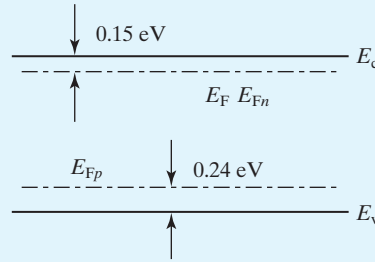


FIGURE 2–14 Location of E_F , E_{Fn} , and E_{Fp} .

2.9 • CHAPTER SUMMARY •

In the presence of an electric field, charge carriers gain a **drift velocity** and produce a **drift current density** in proportion to \mathcal{E} :

$$v_p = \mu_p \mathcal{E} \quad (2.2.3a)$$

$$v_n = -\mu_n \mathcal{E} \quad (2.2.3b)$$

$$J_{p,\text{drift}} = qp\mu_p \mathcal{E} \quad (2.2.11)$$

$$J_{n,\text{drift}} = qn\mu_n \mathcal{E} \quad (2.2.12)$$

μ_p and μ_n are called the **hole** and **electron mobility**. They are determined by how frequently the carriers collide with phonons or dopant ions and lose their drift

momentum. Mobilities are functions of temperature and the total dopant concentration. Measured mobility data are routinely presented in figures.

The second important transport mechanism is **diffusion**. **Diffusion current density** is proportional to the gradient of the carrier concentration.

$$J_{n,\text{diffusion}} = qD_n \frac{dn}{dx} \quad (2.3.2)$$

$$J_{p,\text{diffusion}} = -qD_p \frac{dp}{dx} \quad (2.3.3)$$

D_n and D_p are the electron and hole diffusion constants. Both drift and diffusion are perturbations to the same thermal motion, and both are slowed down by the same collisions that are responsible for the zigzag paths of the thermal motion. As a result, D and μ are related by the **Einstein relationship**:

$$D_n = \frac{kT}{q} \mu_n \quad (2.5.6a)$$

$$D_p = \frac{kT}{q} \mu_p \quad (2.5.6b)$$

The sum of J_{drift} and $J_{\text{diffusion}}$ is the total current density.

The minority carrier concentration, e.g., p in an N-type semiconductor, can easily be increased from its **equilibrium concentration** p_0 by orders of magnitude with light or by another means to be presented in Chapter 4, so that

$$p \equiv p_0 + p' \gg p_0 \quad (2.6.1b)$$

p' is the **excess hole concentration**. In a charge neutral region,

$$n' \equiv p' \quad (2.6.2)$$

Charge nonneutrality will generate an electric field that causes the majority carriers to redistribute until neutrality is achieved.

The electron–hole **recombination rate** is proportional to n' ($=p'$):

$$\text{Recombination rate} = \frac{n'}{\tau} = \frac{p'}{\tau}$$

τ is the recombination lifetime and ranges from nanoseconds to milliseconds for Si, depending on the density of trace metal impurities that form deep traps.

When excess minority carriers are present, the pn product can be orders of magnitude larger than n_i^2 . Clearly, electrons and holes as two groups of particles are *not* at equilibrium with each other. Within each group, however, the carriers are still at equilibrium among themselves and share one common (quasi) Fermi level at different locations. This situation is called **quasi-equilibrium** and the following relationships are useful:

$$n = N_c e^{-(E_c - E_{Fn})/kT} \quad (2.8.1)$$

$$p = N_v e^{-(E_{Fp} - E_v)/kT} \quad (2.8.2)$$

E_{Fn} and E_{Fp} are the quasi-Fermi levels of electrons and holes.

● PROBLEMS ●

● Mobility ●

- 2.1** (a) For an electron mobility of $500 \text{ cm}^2/\text{V}\cdot\text{s}$, calculate the time between collisions. (Take $m_n = m_0$ in these calculations.)
- (b) For an electric field 100 V/cm , calculate the distance an electron travels by drift between collisions.
- 2.2** An electron is moving in a piece of very lightly doped silicon under an applied field such that its drift velocity is one-tenth of its thermal velocity. Calculate the average number of collisions it will experience in traversing by drift a region $1 \mu\text{m}$ long. What is the voltage across this region?
- 2.3** The electron mobility is determined by collisions that come in two flavors:
- (1) scattering due to phonons (lattice vibrations) and
 - (2) scattering due to ionized impurities.

The mobilities from phonon interactions alone, μ_1 , and from ionized impurities alone μ_2 , depend on the electron effective mass m_n , ionized impurity density N_i , and temperature as follows:

$$\mu_1 = \frac{2.2 \times 10^{-70}}{m_n^{5/2} T^{3/2}} k_g^{5/2} K^{3/2} \text{ cm}^2 / (\text{V} \cdot \text{s})$$

$$\mu_2 = \frac{7.45 T^{3/2}}{m_n^{1/2} N_i} k_g^{1/2} / (\text{cm} \cdot K^{3/2} \cdot \text{V} \cdot \text{s})$$

Consider a uniformly doped N-type semiconductor with $N_d = 10^{17} \text{ cm}^{-3}$ and $m_n = 0.27m_0$.

- (a) Make a plot of $\log(\mu_1)$ and $\log(\mu_2)$ versus temperature from 100 to 700 K.
- (b) What is the total electron mobility at 300 K?
- (c) Calculate the electron drift current density if the sample is biased as shown in Fig. 2-15 ($T = 300 \text{ K}$):

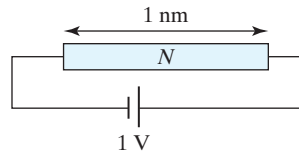


FIGURE 2-15

● Drift ●

- 2.4** Phosphorus donor atoms at a concentration of 10^{16} cm^{-3} are distributed uniformly throughout a silicon sample.
- (a) What is the sample resistivity at 300 K?
 - (b) If 10^{17} cm^{-3} of boron is included in addition to the phosphorus, what is the resulting resistivity and conductivity type (N-type or P-type material)?
 - (c) Sketch the energy band diagram for part (a) and for part (b) and show the position of the Fermi level.

- 2.5** An N-type silicon sample has a uniform density $N_d = 10^{17} \text{ cm}^{-3}$ of arsenic, and a P-type sample has $N_a = 10^{15} \text{ cm}^{-3}$. A third sample has both impurities present at the same time.
- Find the equilibrium minority carrier concentrations at 300 K in each sample.
 - Find the conductivity of each sample at 300 K.
 - Find the Fermi level in each material at 300 K with respect to either the conduction band edge (E_c) or the valence band edge (E_v).
- 2.6**
- A silicon sample maintained at $T = 300 \text{ K}$ is uniformly doped with $N_d = 10^{16} \text{ cm}^{-3}$ donors. Calculate the resistivity of the sample.
 - The silicon sample of part (a) is “compensated” by adding $N_a = 10^{16} \text{ cm}^{-3}$ acceptors. Calculate the resistivity of the compensated sample. (Exercise caution in choosing the mobility values to be employed in this part of the problem.)
 - Compute the resistivity of intrinsic ($N_a = 0, N_d = 0$) silicon at $T = 300 \text{ K}$. Compare it with the result of part (b) and comment.
- 2.7** A sample of N-type silicon is at the room temperature. When an electric field with a strength of 1000 V/cm is applied to the sample, the hole velocity is measured and found to be $2 \times 10^5 \text{ cm/sec}$.
- Estimate the thermal equilibrium electron and hole densities, indicating which is the minority carrier.
 - Find the position of E_F with respect to E_c and E_v .
 - The sample is used to make an integrated circuit resistor. The width and height of the sample are $10 \text{ } \mu\text{m}$ and $1.5 \text{ } \mu\text{m}$, respectively, and the length of the sample is $20 \text{ } \mu\text{m}$. Calculate the resistance of the sample.

● Diffusion ●

- 2.8** A general relationship for the current density carried by electrons of density n is $J = qnv$, where q is the electronic charge and v is the electron velocity.
- Find the velocity of electrons, $v(x)$, that are moving only by diffusion if they have a density distribution of $n(x) = n_0 \exp(-x/\lambda)$. The electric field is zero.
 - What would be the electric field, $\mathcal{E}(x)$, that would lead to an electron drift velocity equal to that of the diffusion velocity in part (a)?
 - At 300 K, what value of λ would make the field in part (b) to be 1000 V/cm ?
- 2.9** Figure 2–16 is a part of the energy band diagram of a P-type semiconductor bar under equilibrium conditions (i.e., E_F is constant). The valence band edge is sloped because doping is nonuniform along the bar. Assume that E_v rises with a slope of Δ/L .

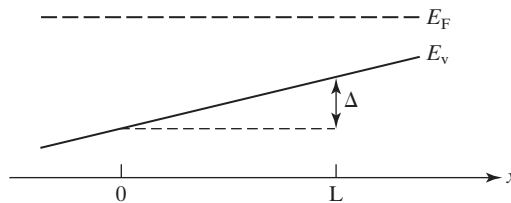


FIGURE 2-16

- Write an expression for the electric field inside this semiconductor bar.
- Within the Boltzmann approximation, what is the electron concentration $n(x)$ along the bar? Assume that $n(x = 0)$ is n_0 . Express your answer in terms of n_0 , Δ , and L .

- (c) Given that the semiconductor bar is under equilibrium, the total electron and hole currents are individually zero. Use this fact and your answers to parts (a) and (b) to derive the Einstein relation ($D_n/\mu_n = kT/q$) relating electron mobility and diffusion constant.

● REFERENCES ●

1. Baccarani, G., and P. Ostojia. "Electron Mobility Empirically Related to the Phosphorus Concentration in Silicon." *Solid State Electronics*, 18, 1975, 579.
2. Antoniadis, D. A., A. G. Gonzalez, and R. W. Dutton. "Boron in Near-Intrinsic <100> and <111> Silicon Under Inert and Oxidizing Ambients—Diffusion and Segregation," *J. Electrochem. Soc.: Solid-State Science and Technology*, 5, 1978, 813.
3. Beadle, W. E., J. C. C. Tsai, and R. D. Plummer. *Quick Reference Manual for Silicon Integrated Circuit Technology*. New York: John Wiley & Sons, 1985. 2–22, 2–23, 2–27.
4. Sze, S. M. *Semiconductor Devices*. New York: Wiley, 1985, 33.

● GENERAL REFERENCES ●

1. Neamen, D. A. *Semiconductor Physics and Devices*, 3rd ed. New York: McGraw-Hill, 2003.
2. Pierret, R. F. *Semiconductor Device Fundamentals*. Reading, MA: Addison-Wesley, 1996.
3. Sze, S. M. *Semiconductor Devices*. New York: Wiley, 1985.