

# 3 Carrier Action

Carrier modeling under "rest" or equilibrium conditions, considered in Chapter 2, is important because it establishes the proper frame of reference. From a device standpoint, however, the zero current observed under equilibrium conditions is rather uninteresting. Only when a semiconductor system is perturbed, giving rise to carrier action or a net carrier response, can currents flow within and external to the semiconductor system. Action, carrier action, is the general concern of this chapter.

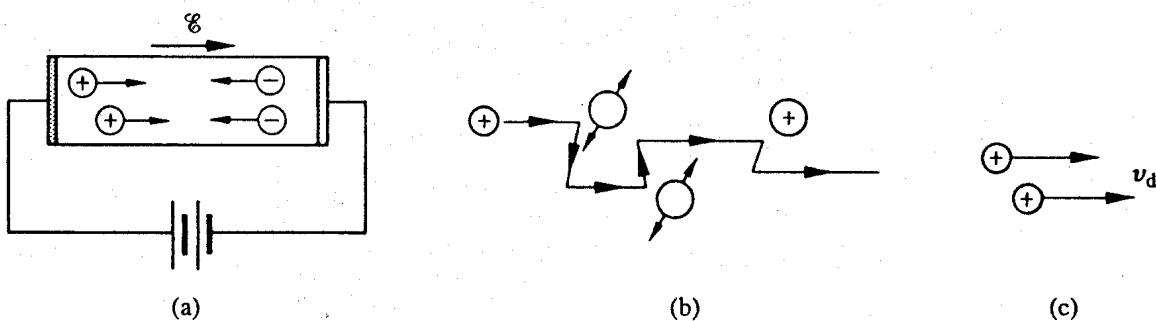
Under normal operating conditions the three primary types of carrier action occurring inside semiconductors are *drift*, *diffusion*, and *recombination-generation*. In this chapter we first describe each primary type of carrier action qualitatively and then quantitatively relate the action to the current flowing within the semiconductor. Special emphasis is placed on characterizing the "constants of the motion" associated with each type of action, and wherever appropriate the discussion is extended to subsidiary topics of a relevant nature. Although introduced individually, the various types of carrier action are understood to occur simultaneously inside any given semiconductor. Mathematically combining the various carrier activities next leads to the culmination of our carrier-action efforts; we obtain the basic set of starting equations employed in solving device problems of an electrical nature. Finally, simple example problems are considered to illustrate solution approaches and to introduce key supplemental concepts.

## 3.1 DRIFT

### 3.1.1 Definition–Visualization

*Drift*, by definition, is charged-particle motion in response to an applied electric field. Within semiconductors the drifting motion of the carriers on a microscopic scale can be described as follows: When an electric field ( $\mathcal{E}$ ) is applied across a semiconductor as visualized in Fig. 3.1(a), the resulting force on the carriers tends to accelerate the  $+q$  charged holes in the direction of the electric field and the  $-q$  charged electrons in the direction opposite to the electric field. Because of collisions with ionized impurity atoms and thermally agitated lattice atoms, however, the carrier acceleration is frequently interrupted (the carriers are said to be scattered). The net result, pictured in Fig. 3.1(b), is carrier motion generally along the direction of the electric field, but in a disjointed fashion involving repeated periods of acceleration and subsequent decelerating collisions.

The microscopic drifting motion of a single carrier is obviously complex and quite tedious to analyze in any detail. Fortunately, measurable quantities are *macroscopic*



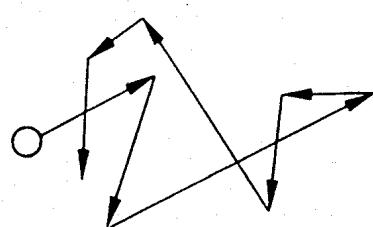
**Figure 3.1** Visualization of carrier drift: (a) motion of carriers within a biased semiconductor bar; (b) drifting hole on a microscopic or atomic scale; (c) carrier drift on a macroscopic scale.

observables that reflect the average or overall motion of the carriers. Averaging over all electrons or holes at any given time, we find that the resultant motion of each carrier type can be described in terms of a constant drift velocity,  $v_d$ . In other words, on a macroscopic scale, drift may be visualized (see Fig. 3.1c) as nothing more than all carriers of a given type moving along at a constant velocity in a direction parallel or antiparallel to the applied electric field.

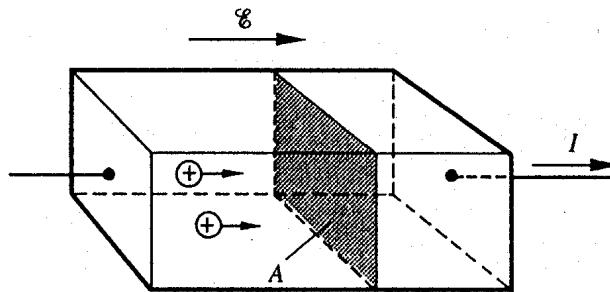
By way of clarification, it is important to point out that the drifting motion of the carriers arising in response to an applied electric field is actually superimposed upon the always-present thermal motion of the carriers. Electrons in the conduction band and holes in the valence band gain and lose energy via collisions with the semiconductor lattice and are nowhere near stationary even under equilibrium conditions. In fact, under equilibrium conditions the thermally related carrier velocities average  $\sim 1/1000$  the speed of light at room temperature! As pictured in Fig. 3.2, however, the thermal motion of the carriers is completely random. Thermal motion therefore averages out to zero on a macroscopic scale, does not contribute to current transport, and can be conceptually neglected.

### 3.1.2 Drift Current

Let us next turn to the task of developing an analytical expression for the current flowing within a semiconductor as a result of carrier drift. By definition



**Figure 3.2** Thermal motion of a carrier.



**Figure 3.3** Expanded view of a biased *p*-type semiconductor bar of cross-sectional area *A*.

*I* (current) = the charge per unit time crossing an arbitrarily chosen plane of observation oriented normal to the direction of current flow.

Considering the *p*-type semiconductor bar of cross-sectional area *A* shown in Fig. 3.3, and specifically noting the arbitrarily chosen  $v_d$ -normal plane lying within the bar, we can argue:

- $v_d t$  ... All holes this distance back from the  $v_d$ -normal plane will cross the plane in a time *t*.
- $v_d t A$  ... All holes in this volume will cross the plane in a time *t*.
- $p v_d t A$  ... Holes crossing the plane in a time *t*.
- $q p v_d t A$  ... Charge crossing the plane in a time *t*.
- $q p v_d A$  ... Charge crossing the plane per unit time.

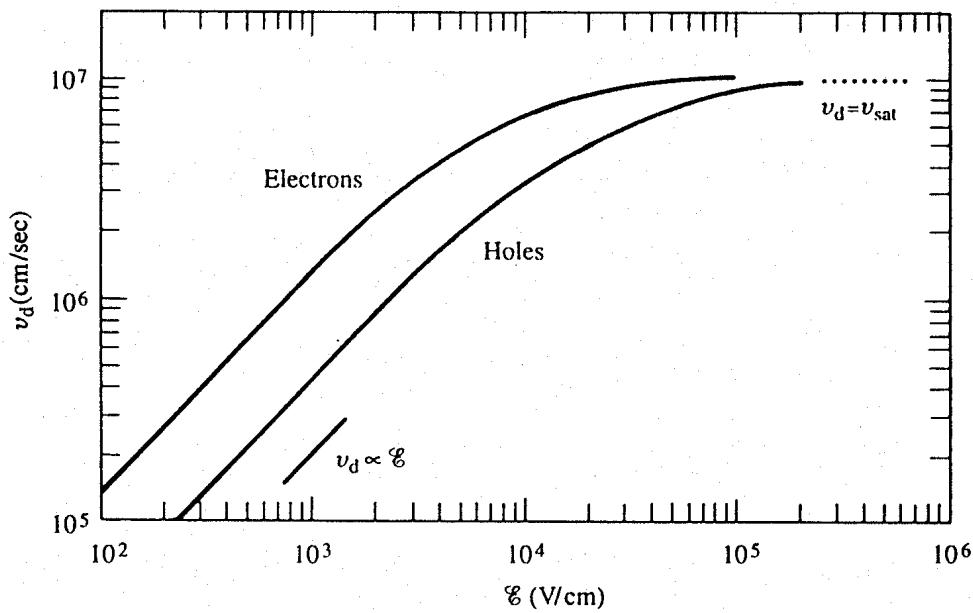
The word definition of the last quantity is clearly identical to the formal definition of current. Thus

$$I_{\text{Pldrift}} = q p v_d A \quad \text{hole drift current} \quad (3.1)$$

As a practical matter, the cross-sectional area *A* appearing in Eq. (3.1) and other current formulas is often excess baggage. Current, moreover, is generally thought of as a scalar quantity, while in reality it is obviously a vector. These deficiencies are overcome by introducing a related parameter known as the current density,  $\mathbf{J}$ .  $\mathbf{J}$  has the same orientation as the direction of current flow and is equal in magnitude to the current per unit area (or  $J = I/A$ ). By inspection, the current density associated with hole drift is simply

$$\mathbf{J}_{\text{Pldrift}} = q p v_d \quad (3.2)$$

Since the drift current arises in response to an applied electric field, it is reasonable to proceed one step further and seek a form of the current relationship that explicitly relates  $\mathbf{J}_{\text{Pldrift}}$  to the applied electric field. To this end, we make reference to the representative drift velocity versus electric field data presented in Fig. 3.4. Note that  $v_d$  is proportional to  $E$  at



**Figure 3.4** Measured drift velocity of the carriers in ultrapure silicon maintained at room temperature as a function of the applied electric field. Constructed from the data fits and the data respectively in Jacoboni et al.<sup>[4]</sup> and Smith et al.<sup>[5]</sup>

low electric fields, while at high electric fields  $v_d$  saturates and becomes independent of  $E$ . To be more precise,

$$v_d = \frac{\mu_0 E}{\left[ 1 + \left( \frac{\mu_0 E}{v_{sat}} \right)^\beta \right]^{1/\beta}} = \begin{cases} \mu_0 E & \dots E \rightarrow 0 \\ v_{sat} & \dots E \rightarrow \infty \end{cases} \quad (3.3)$$

where  $\beta \approx 1$  for holes and  $\beta \approx 2$  for electrons in silicon,  $\mu_0$  is the constant of proportionality between  $v_d$  and  $E$  at low to moderate electric fields, and  $v_{sat}$  is the limiting or *saturation velocity* approached at very high electric fields. Obviously, in the high-field limit  $v_d$  in Eq. (3.2) is simply replaced by  $v_{sat}$  and  $J_{P|drift}$  does not exhibit an  $E$ -field dependence. In the low-field limit, which is of greatest practical interest and assumed herein unless specified otherwise,  $v_d = \mu_p E$  ( $\mu_0 \rightarrow \mu_p$  for holes) and substitution into Eq. (3.2) yields

$$J_{P|drift} = q\mu_p p E \quad (3.4a)$$

Similarly, for electrons one obtains

$$J_{N|drift} = q\mu_n n E \quad (3.4b)$$

Respectively known as the *electron mobility* and *hole mobility*,  $\mu_n$  and  $\mu_p$  are always taken to be positive quantities. Note that, although electrons drift in the direction opposite to the applied electric field ( $v_d = -\mu_n E$ ), the current transported by negatively charged particles is in turn counter to the direction of drift ( $J_{Nl\text{drift}} = -qnv_d$ ). The net result, as indicated in Eq. (3.4b), is an electron current in the direction of the applied electric field.

### 3.1.3 Mobility

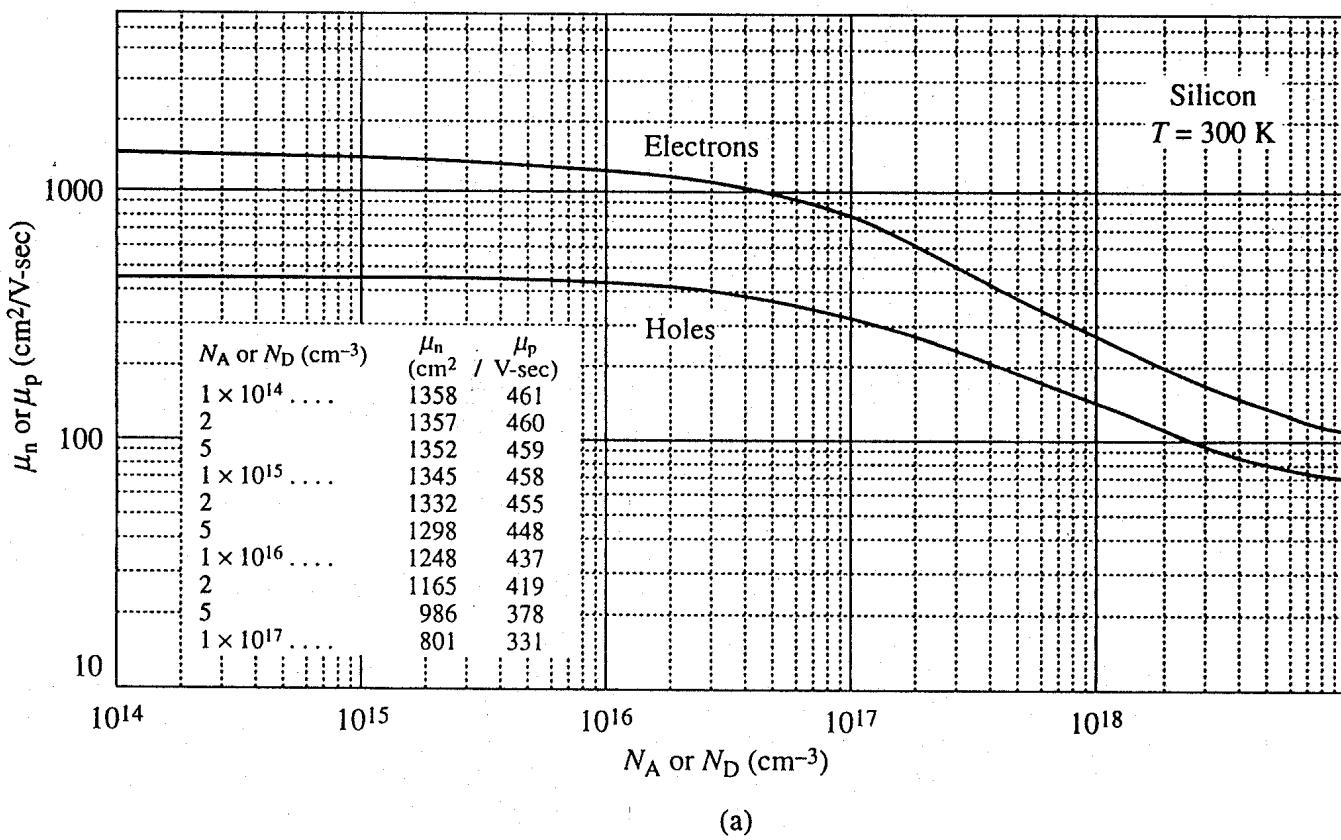
Mobility is obviously a central parameter in characterizing electron and hole transport due to drift. As further readings will reveal, the carrier mobilities also play a key role in characterizing the performance of many devices. It is reasonable therefore to examine  $\mu_n$  and  $\mu_p$  in some detail to enhance our general familiarity with the parameters and to establish a core of useful information for future reference.

**Standard Units:**  $\text{cm}^2/\text{V}\cdot\text{sec}$ .

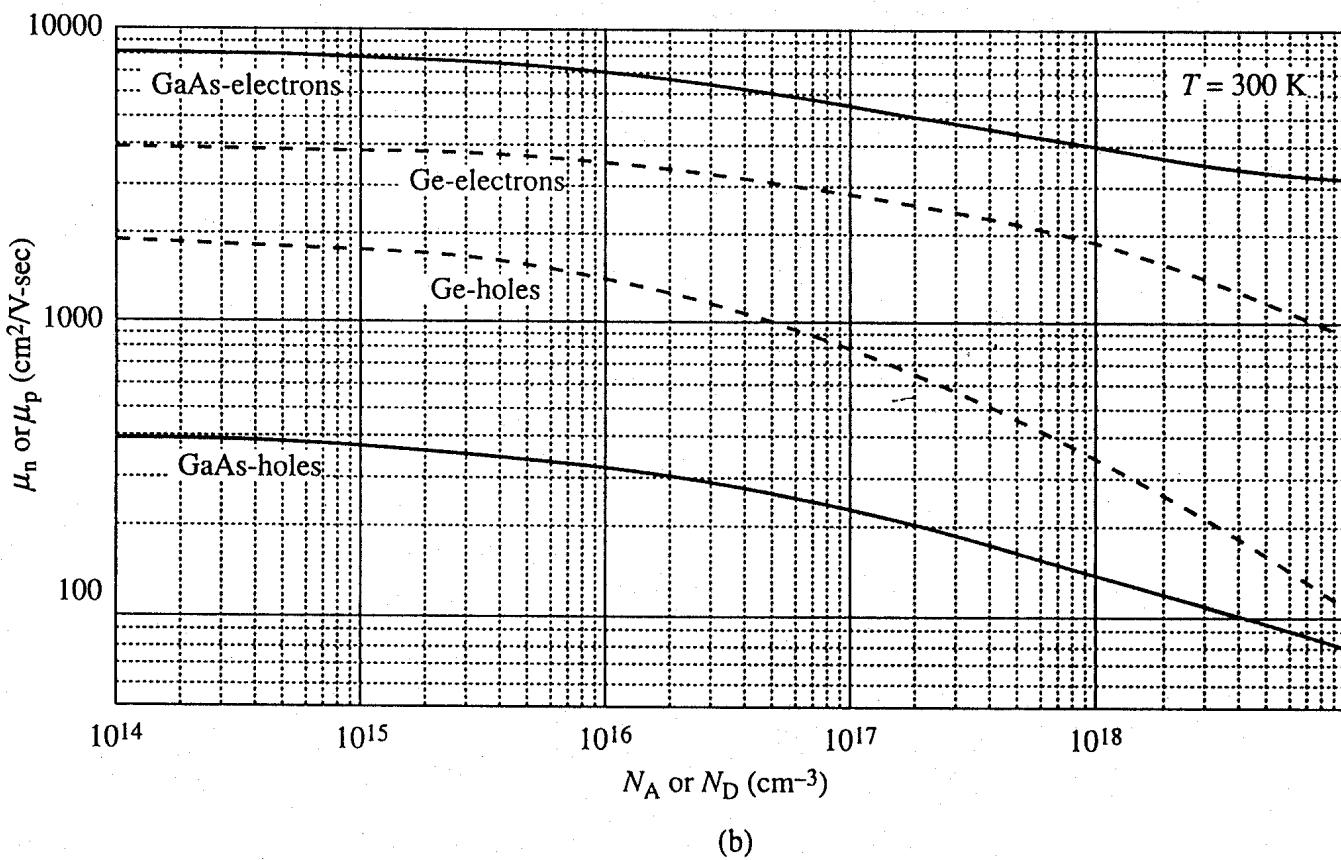
**Sample Numerical Values:**  $\mu_n \approx 1360 \text{ cm}^2/\text{V}\cdot\text{sec}$  and  $\mu_p \approx 460 \text{ cm}^2/\text{V}\cdot\text{sec}$  at 300 K in  $N_D = 10^{14}/\text{cm}^3$  and  $N_A = 10^{14}/\text{cm}^3$  doped Si, respectively. In uncompensated high-purity ( $N_D$  or  $N_A \leq 10^{15}/\text{cm}^3$ ) GaAs, the room-temperature mobilities are  $\mu_n \approx 8000 \text{ cm}^2/\text{V}\cdot\text{sec}$  and  $\mu_p \approx 400 \text{ cm}^2/\text{V}\cdot\text{sec}$ . The quoted values are useful for comparison purposes and when performing order-of-magnitude computations. Also note that  $\mu_n > \mu_p$  for both Si and GaAs. In the major semiconductors,  $\mu_n$  is consistently greater than  $\mu_p$  for a given doping and system temperature.

**Relationship to Scattering:** The word *mobility* in everyday usage refers to a general freedom of movement. Analogously, in semiconductor work the mobility parameter is a measure of the ease of carrier motion in a crystal. Increasing the motion-impeding collisions within a crystal decreases the mobility of the carriers. In other words, the carrier mobility varies inversely with the amount of scattering taking place within the semiconductor. As visualized in Fig. 3.1(b), the dominant scattering mechanisms in nondegenerately doped materials of device quality are typically (i) lattice scattering involving collisions with thermally agitated lattice atoms, and (ii) ionized impurity (i.e., donor-site and/or acceptor-site) scattering. Relative to lattice scattering, it should be emphasized that it is the thermal vibration, the *displacement* of lattice atoms from their lattice positions, that leads to carrier scattering. The internal field associated with the stationary array of atoms in a crystal is already taken into account in the effective mass formulation.

Although quantitative relationships connecting the mobility and scattering can become quite involved, it is readily established that  $\mu = q\langle\tau\rangle/m^*$ , where  $\langle\tau\rangle$  is the mean free time between collisions and  $m^*$  is the conductivity effective mass. Since increasing the number of motion-impeding collisions decreases the mean free time between collisions, we again conclude  $\mu$  varies inversely with the amount of scattering. However,  $\mu$  is also noted to vary inversely with the carrier effective mass—lighter carriers move more readily. The  $m_n^*$  in GaAs is significantly smaller than the  $m_n^*$  in Si, thereby explaining the higher mobility of the GaAs electrons.



(a)



(b)

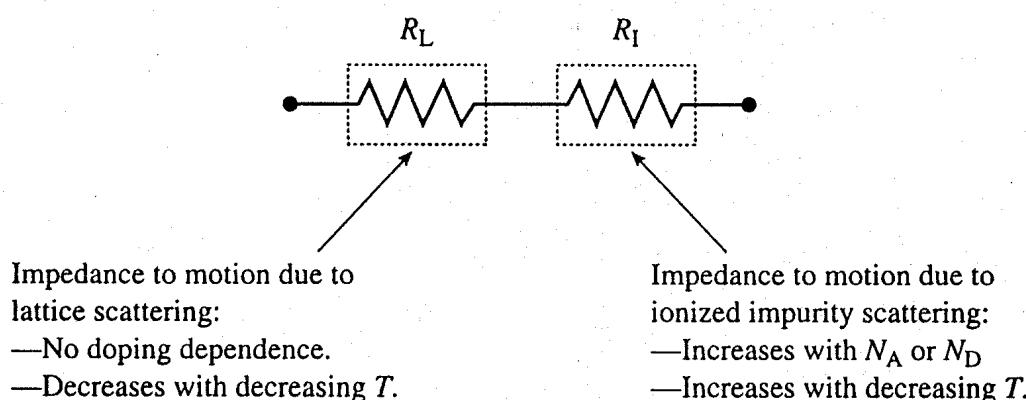
**Figure 3.5** Room temperature carrier mobilities as a function of the dopant concentration in (a) Si and (b) Ge and GaAs.  $\mu_n$  is the electron mobility;  $\mu_p$  is the hole mobility.

**Doping Dependence:** Figure 3.5 exhibits the observed doping dependence of the electron and hole mobilities in Si, Ge, and GaAs. All semiconductors exhibit the same general dependence. At low doping concentrations, below approximately  $10^{15}/\text{cm}^3$  in Si, the carrier mobilities are essentially independent of the doping concentration. For dopings in excess of  $\sim 10^{15}/\text{cm}^3$ , the mobilities monotonically decrease with increasing  $N_A$  or  $N_D$ .

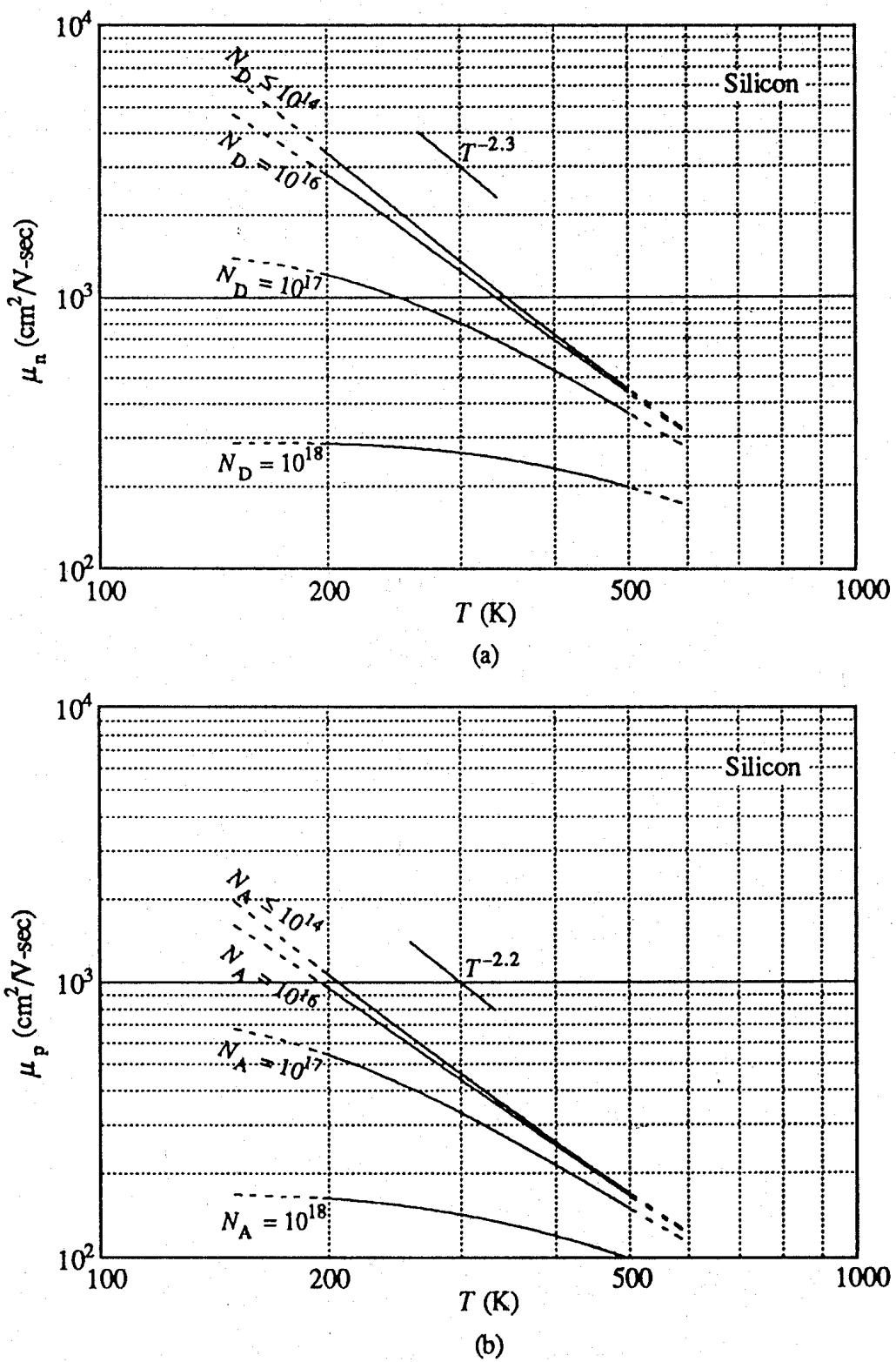
With the aid of Fig. 3.6 the explanation of the observed doping dependence is relatively straightforward. Invoking an electrical analogy, one can associate a resistance to motion with each of the scattering mechanisms. These resistances are in series. At sufficiently low doping levels, ionized impurity scattering can be neglected compared to lattice scattering, or in the analogy,  $R_{\text{TOTAL}} = R_L + R_I \approx R_L$ . When lattice scattering, which is not a function of  $N_A$  or  $N_D$ , becomes the dominant scattering mechanism, it automatically follows that the carrier mobilities will be likewise independent of  $N_A$  or  $N_D$ . For dopings in excess of  $\sim 10^{15}/\text{cm}^3$  in Si, ionized impurity scattering and the associated resistance to motion can no longer be neglected. Increasing the number of scattering centers by adding more and more acceptors or donors progressively increases the amount of ionized impurity scattering and systematically decreases the carrier mobilities.

**Temperature Dependence:** The temperature dependence of the electron and hole mobilities in Si with doping as a parameter is displayed in Fig. 3.7. For dopings of  $N_A$  or  $N_D \leq 10^{14}/\text{cm}^3$ , the data merge into a single curve and there is a near power-law increase in the carrier mobilities as the temperature is decreased. Roughly,  $\mu_n \propto T^{-2.3 \pm 0.1}$  and  $\mu_p \propto T^{-2.2 \pm 0.1}$ . For progressively higher dopings, the carrier mobilities still increase with decreasing temperature, but at a systematically decreasing rate. In fact, some  $N_D \geq 10^{18}/\text{cm}^3$  experimental data (not shown) exhibit a *decrease* in  $\mu_n$  as  $T$  is decreased below 200 K.

The general temperature dependence of the carrier mobilities is relatively easy to explain in the low doping limit. As noted in the doping dependence discussion, lattice scattering is the dominant scattering mechanism ( $R_{\text{TOTAL}} \approx R_L$ ) in lightly doped samples. Decreasing the system temperature causes an ever-decreasing thermal agitation of the semiconductor atoms, which in turn decreases the lattice scattering. The decreased scattering



**Figure 3.6** Electrical analogy for scattering in a semiconductor.  $R_L$  and  $R_I$  represent the impedance to motion due to lattice scattering and ionized impurity scattering, respectively.



**Figure 3.7** Temperature dependence of (a) electron and (b) hole mobilities in silicon for dopings ranging from  $\leq 10^{14}/\text{cm}^3$  to  $10^{18}/\text{cm}^3$ . The curves were constructed using the empirical fit relationships and parameters presented in Exercise 3.1. The dashed line portion of the curves correspond to a slight extension of the fit beyond the verified  $200 \text{ K} \leq T \leq 500 \text{ K}$  range of validity.

roughly follows a simple power-law dependence. Being inversely proportional to the amount of lattice scattering, the mobility of lightly doped samples is therefore expected to increase with decreasing temperature, varying roughly as the temperature to a negative power.

The more complex dependence of the higher doped samples reflects the added effect of ionized impurity scattering. In the electrical analogy,  $R_I$  can no longer be neglected. Moreover, whereas lattice scattering ( $R_L$ ) decreases with decreasing  $T$ , ionized impurity scattering ( $R_I$ ) increases with decreasing  $T$ . Ionized impurities become more and more effective in deflecting the charged carriers as the temperature and hence the speed of the carriers decreases. Thus ionized impurity scattering becomes a larger and larger percentage of the overall scattering as the temperature is decreased ( $R_{\text{TOTAL}} \rightarrow R_I$ ). Clearly, this explains the decreased slope of the mobility versus temperature dependence exhibited by the higher doped samples.

### (C) Exercise 3.1

There exist surprisingly accurate "empirical-fit" relationships that are widely employed to compute the carrier mobilities at a given doping and temperature. Figures 3.5 and 3.7 were constructed using such relationships. The form of a computational relationship is typically established on an empirical basis by noting the general functional dependencies predicted theoretically and observed experimentally. Parameters in the relationship are then adjusted until an acceptable match is obtained to the best available experimental data.

The majority carrier mobility versus doping at room temperature is popularly computed from

$$\mu = \mu_{\min} + \frac{\mu_0}{1 + (N/N_{\text{ref}})^{\alpha}}$$

where  $\mu$  is the carrier mobility ( $\mu_n$  or  $\mu_p$ ),  $N$  is the doping concentration ( $N_A$  or  $N_D$ ), and all other quantities are fit parameters. To model the temperature dependence, one additionally employs

$$A = A_{300} \left( \frac{T}{300} \right)^{\eta}$$

$A$  in the above equation represents  $\mu_{\min}$ ,  $\mu_0$ ,  $N_{\text{ref}}$ , or  $\alpha$ ;  $A_{300}$  is the 300 K value of the parameter,  $T$  is temperature in Kelvin, and  $\eta$  is the temperature exponent. The fit parameters appropriate for Si are listed in the following table:

Parameter	Value at 300 K		Temperature Exponent ( $\eta$ )
	Electrons	Holes	
$N_{\text{ref}} (\text{cm}^{-3})$	$1.3 \times 10^{17}$	$2.35 \times 10^{17}$	2.4
$\mu_{\text{min}} (\text{cm}^2/\text{V}\cdot\text{sec})$	92	54.3	-0.57
$\mu_0 (\text{cm}^2/\text{V}\cdot\text{sec})$	1268	406.9	-2.33 electrons -2.23 holes
$\alpha$	0.91	0.88	-0.146

**P:** (a) Construct a log-log plot of  $\mu_n$  and  $\mu_p$  versus  $N_A$  or  $N_D$  for  $10^{14}/\text{cm}^3 \leq N_A$  or  $N_D \leq 10^{19}/\text{cm}^3$  using the quoted fit relationship and the listed Si fit parameters. Compare your result with Fig. 3.5(a).

(b) Construct log-log plots of  $\mu_n$  versus  $T$  and  $\mu_p$  versus  $T$  for  $200 \text{ K} \leq T \leq 500 \text{ K}$  and  $N_D$  or  $N_A$  stepped in decade values from  $10^{14}/\text{cm}^3$  to  $10^{18}/\text{cm}^3$ . Compare your results with Figs. 3.7(a) and 3.7(b), respectively.

**S:** (a) MATLAB program script . . .

```
%Mobility versus Dopant Concentration (Si,300K)
```

```
%Fit Parameters
```

```
NDref=1.3e17; NAref=2.35e17;
μnmin=92; μpmin=54.3;
μn0=1268; μp0=406.9;
an=0.91; ap=0.88
```

```
%Mobility Calculation
```

```
N=logspace(14,19);
μn=μnmin+μn0./(1+(N/NDref).^an);
μp=μpmin+μp0./(1+(N/NAref).^ap);
```

```
%Plotting results
```

```
close
loglog(N,μn,N,μp); grid;
axis([1.0e14 1.0e19 1.0e1 1.0e4]);
xlabel('NA or ND (cm-3)');
ylabel('Mobility (cm2/V·sec)');
text(1.0e15,1500,'Electrons');
text(1.0e15,500,'Holes');
text(1.0e18,2000,'Si,300K');
```

The results obtained by running the preceding program should be numerically identical to Fig. 3.5(a).

(b) Part (b) is left for the reader to complete.

### 3.1.4 Resistivity

Resistivity is an important material parameter that is closely related to carrier drift. Qualitatively, *resistivity* is a measure of a material's inherent resistance to current flow—a “normalized” resistance that does not depend on the physical dimensions of the material. Quantitatively, resistivity ( $\rho$ ) is defined as the proportionality constant between the electric field impressed across a homogeneous material and the total particle current per unit area flowing in the material; that is,

$$\mathcal{E} = \rho \mathbf{J} \quad (3.5a)$$

or

$$\mathbf{J} = \sigma \mathcal{E} = \frac{1}{\rho} \mathcal{E} \quad (3.5b)$$

where  $\sigma = 1/\rho$  is the material *conductivity*. In a homogeneous material,  $\mathbf{J} = \mathbf{J}_{\text{drift}}$  and, as established with the aid of Eqs. (3.4),

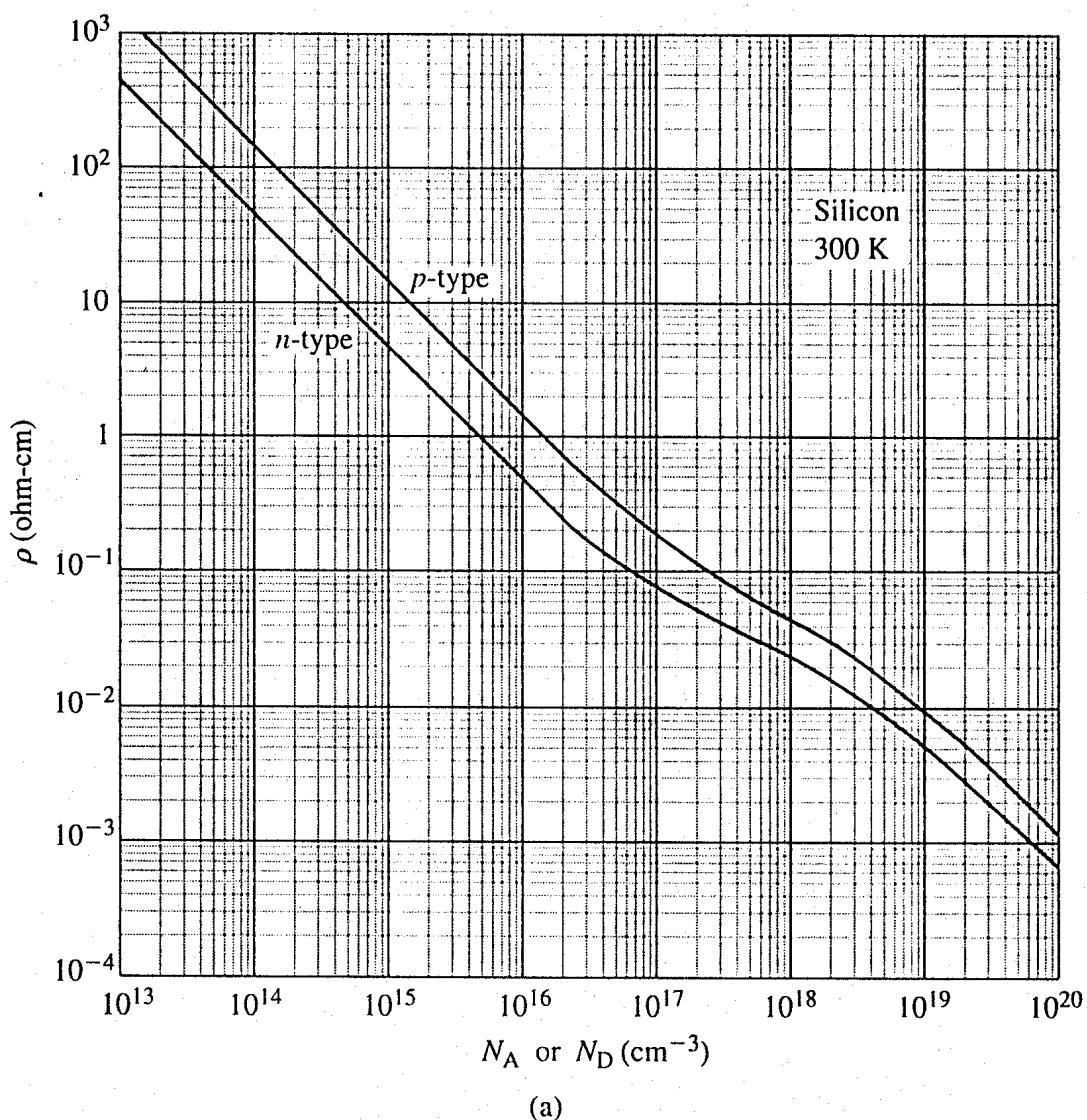
$$\mathbf{J}_{\text{drift}} = \mathbf{J}_{N_{\text{d}} \text{drift}} + \mathbf{J}_{P \text{drift}} = q(\mu_n n + \mu_p p) \mathcal{E} \quad (3.6)$$

It therefore follows that

$$\boxed{\rho = \frac{1}{q(\mu_n n + \mu_p p)}} \quad (3.7)$$

In a nondegenerate donor-doped semiconductor maintained in the extrinsic temperature region where  $N_D \gg n_i$ ,  $n \approx N_D$  and  $p \approx n_i^2/N_D \ll n$ . This result was established in Subsection 2.5.5. Thus, for typical dopings and mobilities,  $\mu_n n + \mu_p p \approx \mu_n N_D$  in an *n*-type semiconductor. Similar arguments yield  $\mu_n n + \mu_p p \approx \mu_p N_A$  in a *p*-type semiconductor. Consequently, under conditions normally encountered in Si samples maintained at or near room temperature, Eq. (3.7) simplifies to

$$\boxed{\rho = \frac{1}{q\mu_n N_D}} \quad \dots n\text{-type semiconductor} \quad (3.8a)$$



**Figure 3.8** Resistivity versus impurity concentration at 300 K in (a) Si and (b) other semiconductors. [(b) From Sze<sup>[2]</sup>, © 1981 by John Wiley & Sons, Inc. Reprinted with permission.]

and

$$\boxed{\rho = \frac{1}{q\mu_p N_A}} \quad \dots \text{p-type semiconductor} \quad (3.8b)$$

When combined with mobility-versus-doping data, Eqs. (3.8) provide a one-to-one correspondence between the resistivity, a directly measurable quantity, and the doping inside a semiconductor. In conjunction with plots of  $\rho$  versus doping (see Fig. 3.8), the measured resistivity is in fact routinely used to determine  $N_A$  or  $N_D$ .

The measured resistivity required in determining the doping can be obtained in a number of different ways. A seemingly straightforward approach would be to form the semiconductor into a bar, apply a bias  $V$  across contacts attached to the ends of the bar as in

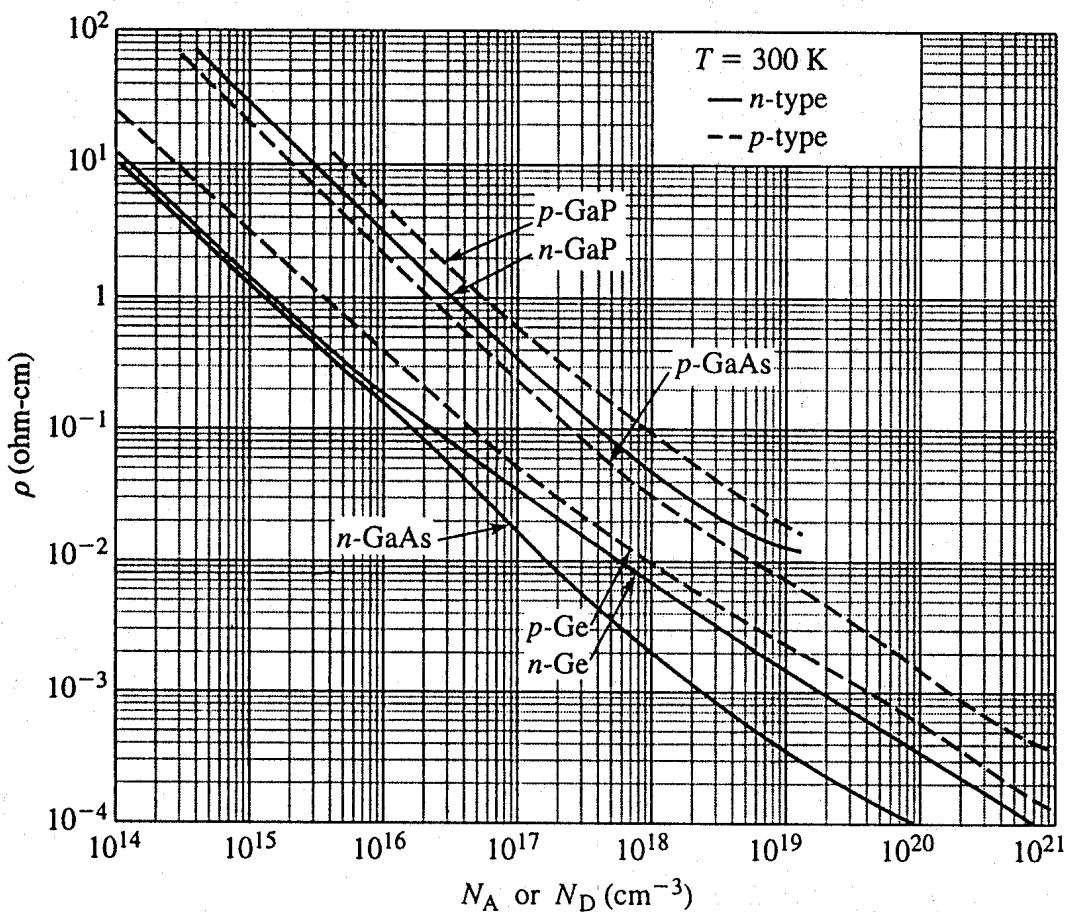


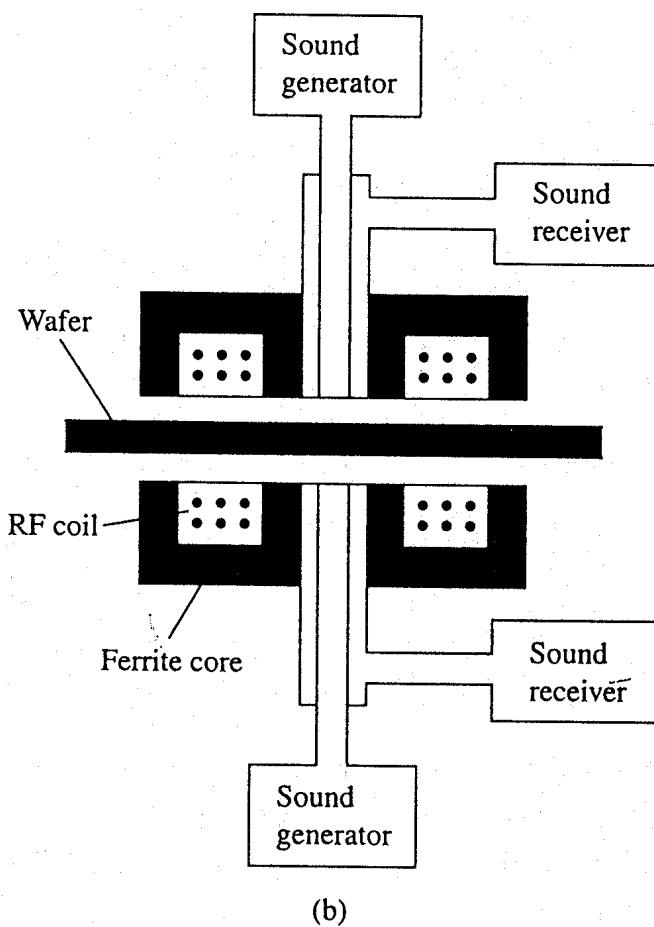
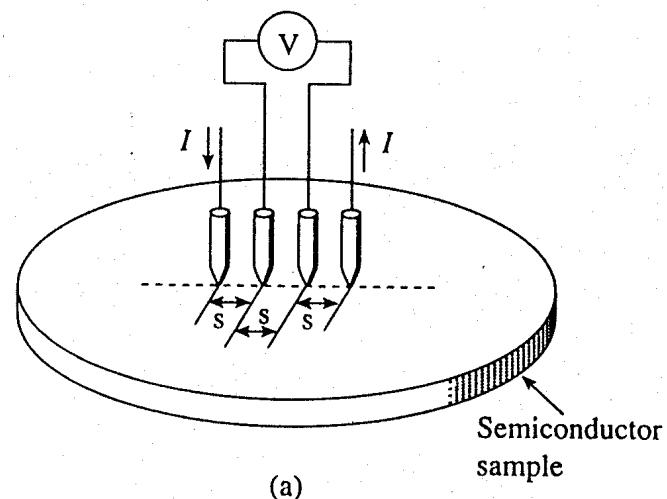
Figure 3.8 (b)

Fig. 3.1(a), measure the current  $I$  flowing in the circuit, and then deduce  $\rho$  from the measured resistance. [ $R(\text{resistance}) = V/I = \rho l/A$ , where  $l$  is the bar length and  $A$  is the cross-sectional area.] Unfortunately, the straightforward approach is deceptively difficult, is destructive (wastes semiconductor material), and is not readily adaptable to the wafers used in device processing.

A measurement method finding widespread usage in practice is the four-point probe technique. In the standard four-point probe technique, four collinear, evenly spaced probes, as shown in Fig. 3.9(a), are brought into contact with the surface of the semiconductor. A known current  $I$  is passed through the outer two probes, and the potential  $V$  thereby developed is measured across the inner two probes. The semiconductor resistivity is then computed from

$$\rho = 2\pi s \frac{V}{I} \Gamma \quad (3.9)$$

where  $s$  is the probe-to-probe spacing and  $\Gamma$  is a well-documented “correction” factor. The correction factor typically depends on the thickness of the sample and on whether the bottom of the semiconductor is touching an insulator or a metal. Commercial instruments are



**Figure 3.9** Resistivity measurement techniques. (a) Schematic drawing of the probe arrangement, placement, and biasing in the four-point probe measurement. (b) Schematic of a commercial eddy-current apparatus showing the RF coils and sonic components. [(b) From Schroder<sup>[6]</sup>, © 1990 by John Wiley & Sons, Inc. Reprinted with permission.]

available that automatically compute the appropriate correction factor based on the sample thickness entered by the operator. Unlike the semiconductor-bar measurement, the four-point probe technique is obviously easy to implement, causes only slight surface damage in the vicinity of the probe contacts, and is ideally suited for working with wafers. The surface damage, although slight, does exclude the technique from being used on wafers intended for producing high-yield devices and ICs.

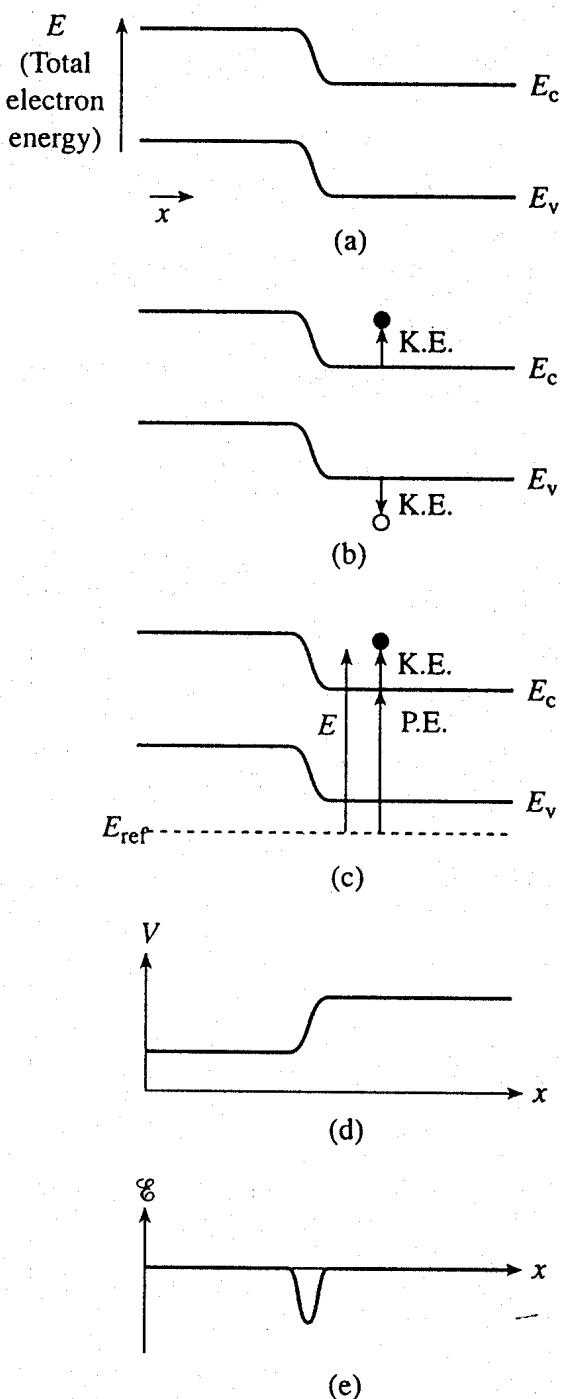
A second technique worthy of note utilizes a noncontacting eddy-current approach. The schematic of a commercial eddy-current apparatus is pictured in Fig. 3.9(b). There are fringing fields near a ferrite core excited with an RF coil. If a conducting material such as a semiconductor wafer is placed near the ferrite core, the fringing fields cause localized (eddy) currents to flow in the conducting material. The current flow in turn absorbs some of the RF power. The *sheet* resistivity of the conducting material is deduced by monitoring the power consumption in a calibrated system. The ultrasound generators and receivers built into the commercial apparatus are for the *in-situ* determination of the wafer thickness. The change in phase of ultrasound bounced off the top and bottom wafer surfaces permits the instrument to compute the intercore spacing occupied by the wafer. The sheet resistivity multiplied by the wafer thickness yields the wafer resistivity.

### 3.1.5 Band Bending

In previous encounters with the energy band diagram, we have consistently drawn  $E_c$  and  $E_v$  to be energies independent of the position coordinate  $x$ . When an electric field ( $\mathcal{E}$ ) exists inside a material, the band energies become a function of position. The resulting variation of  $E_c$  and  $E_v$  with position on the energy band diagram, exemplified by Fig. 3.10(a), is popularly referred to as "band bending."

Seeking to establish the precise relationship between the electric field within a semiconductor and the induced band bending, let us carefully re-examine the energy band diagram. The diagram itself, as emphasized in Fig. 3.10(a), is a plot of the allowed electron energies within the semiconductor as a function of position, where  $E$  increasing upward is understood to be the *total* energy of the electrons. Furthermore, we know from previous discussions that if an energy of precisely  $E_G$  is added to break an atom-atom bond, the created electron and hole energies would be  $E_c$  and  $E_v$ , respectively, and the created carriers would be effectively motionless. Absorbing an energy in excess of  $E_G$ , on the other hand, would in all probability give rise to an electron energy greater than  $E_c$  and a hole energy less than  $E_v$ , with both carriers moving around rapidly within the lattice. We are led, therefore, to interpret  $E - E_c$  to be the kinetic energy (K.E.) of the electrons and  $E_v - E$  to be the kinetic energy of the holes [see Fig. 3.10(b)]. Moreover, since the total energy equals the sum of the kinetic energy and the potential energy (P. E.),  $E_c$  minus the energy reference level ( $E_{ref}$ ) must equal the electron potential energy, as illustrated in Fig. 3.10(c). (Potential energy, it should be noted, is arbitrary to within a constant, and the position-independent reference energy,  $E_{ref}$ , may be chosen to be any convenient value.)

The potential energy is the key in relating the electric field within a semiconductor to positional variations in the energy bands. Specifically, assuming normal operational conditions, where magnetic field, temperature gradient, and stress-induced effects are negligible, only the force associated with an existing electric field can give rise to changes in the



**Figure 3.10** Relationship between band bending and the electrostatic variables inside a semiconductor: (a) sample energy band diagram exhibiting band bending; (b) identification of the carrier kinetic energies; (c) specification of the electron potential energy; (d) electrostatic potential and (e) electric field versus position dependence deduced from and associated with the part (a) energy band diagram.

potential energy of the carriers. Elementary physics, in fact, tells us that the potential energy of a  $-q$  charged particle under such conditions is simply related to the electrostatic potential  $V$  at a given point by

$$\text{P.E.} = -qV \quad (3.10)$$

Having previously concluded that

$$\text{P.E.} = E_c - E_{\text{ref}} \quad (3.11)$$

we can state

$$V = -\frac{1}{q}(E_c - E_{\text{ref}}) \quad (3.12)$$

By definition, moreover,

$$\mathcal{E} = -\nabla V \quad (3.13)$$

or, in one dimension,

$$\mathcal{E} = -\frac{dV}{dx} \quad (3.14)$$

Consequently,

$$\mathcal{E} = \frac{1}{q} \frac{dE_c}{dx} = \frac{1}{q} \frac{dE_v}{dx} = \frac{1}{q} \frac{dE_i}{dx} \quad (3.15)$$

The latter forms of Eq. (3.15) follow from the fact that  $E_c$ ,  $E_v$ , and  $E_i$  differ by only an additive constant.

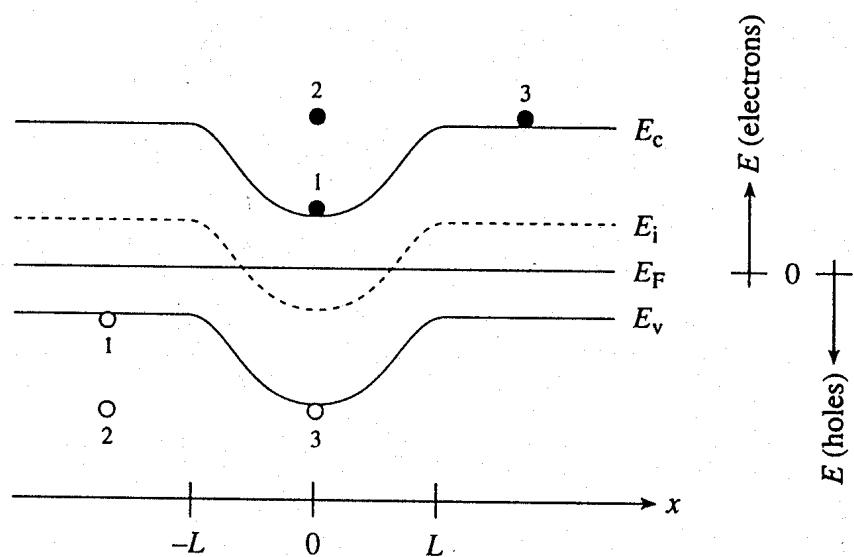
The preceding formulation provides a means of readily deducing the general form of the electrostatic variables associated with the "band bending" in Fig. 3.10(a) and other energy band diagrams. Making use of Eq. (3.12), or simply inverting  $E_c(x)$  in Fig. 3.10(a), one obtains the  $V$  versus  $x$  dependence presented in Fig. 3.10(d). [Please note that the electrostatic potential, like the potential energy, is arbitrary to within a constant—the Fig. 3.10(d) plot can be translated upward or downward along the voltage axis without modifying the physical situation inside the semiconductor.] Finally, recording the slope of  $E_c$  versus position, as dictated by Eq. (3.15), produces the  $\mathcal{E}$  versus  $x$  plot shown in Fig. 3.10(e).

In summary, the reader should be aware of the fact that the energy band diagram contains information relating to the electrostatic potential and electric field within the semiconductor. Moreover, the general form of the  $V$  and  $\mathcal{E}$  dependencies within the semiconductor can be obtained almost by inspection. To deduce the form of the  $V$  versus  $x$

relationship, merely sketch the "upside down" of  $E_c$  (or  $E_v$  or  $E_i$ ) versus  $x$ ; to determine the general  $\mathcal{E}$  versus  $x$  dependence, simply note the slope of  $E_c$  (or  $E_v$  or  $E_i$ ) as a function of position.

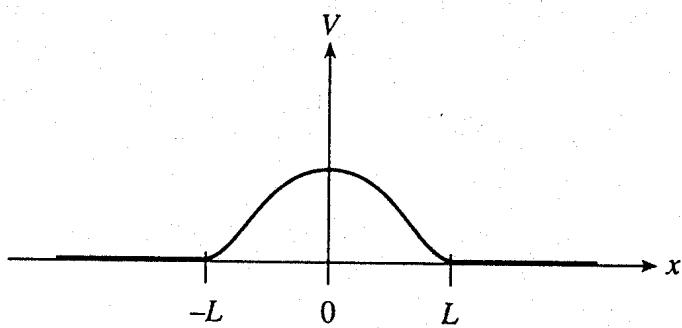
### Exercise 3.2

P: Consider the following energy band diagram. Take the semiconductor represented to be Si maintained at 300 K with  $E_i - E_F = E_G/4$  at  $x = \pm L$  and  $E_F - E_i = E_G/4$  at  $x = 0$ . Note the choice of  $E_F$  as the energy reference level and the identification of carriers at various points on the diagram.

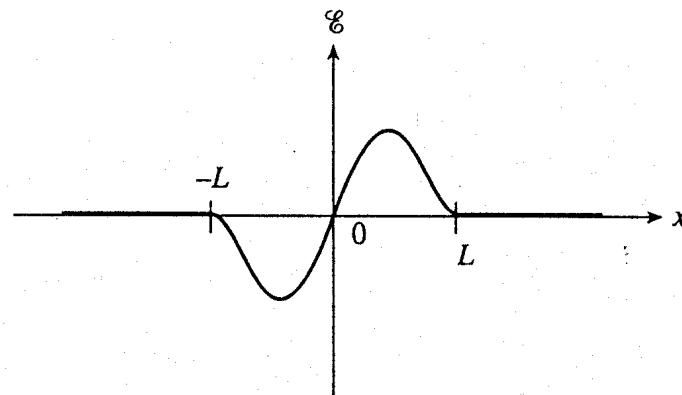


- Sketch the electrostatic potential ( $V$ ) inside the semiconductor as a function of  $x$ .
- Sketch the electric field ( $\mathcal{E}$ ) inside the semiconductor as a function of  $x$ .
- Ascertain the K.E. and P.E. of the electrons and holes pictured on the diagram.
- Determine the resistivity of the  $x > L$  portion of the semiconductor.

S: (a) The  $V$  versus  $x$  relationship must have the same functional form as the "upside down" of  $E_c$  (or  $E_i$  or  $E_v$ ). If the arbitrary voltage reference point is taken to be  $V = 0$  at  $x = L$ , then one concludes the  $V$  versus  $x$  dependence is as sketched here:



(b) The electric field,  $\mathcal{E}$ , is proportional to the slope of the bands:



(c) For the electrons, K.E. =  $E - E_c$  and P.E. =  $E_c - E_{ref} = E_c - E_F$ . Since the total energy of the holes increases *downward* on the diagram, K.E. =  $E_v - E$  and P.E. =  $E_{ref} - E_v = E_F - E_v$  for the holes. The energy differences are of course evaluated at the same  $x$ . The deduced K.E. and P.E. values are summarized in the following table:

Carrier	K.E. (eV)	P.E. (eV)
Electron 1	0	0.28
Electron 2	0.56	0.28
Electron 3	0	0.84
Hole 1	0	0.28
Hole 2	0.56	0.28
Hole 3	0	0.84

(d) In the  $x > L$  region,  $E_i - E_F = E_G/4 = 0.28$  eV and

$$N_A = p = n_i e^{(E_i - E_F)/kT} = 10^{10} e^{0.28/0.0259} = 4.96 \times 10^{14}/\text{cm}^3$$

From Fig. 3.5(a) one deduces  $\mu_p = 459 \text{ cm}^2/\text{V-sec}$ , and therefore,

$$\rho = \frac{1}{q\mu_p N_A} = \frac{1}{(1.6 \times 10^{-19})(459)(4.96 \times 10^{14})} = 27.5 \text{ ohm-cm}$$

Alternatively making use of the Fig. 3.8(a) Si resistivity plot, one similarly concludes  $\rho \approx 25 \text{ ohm-cm}$ .

## 3.2 DIFFUSION

### 3.2.1 Definition–Visualization

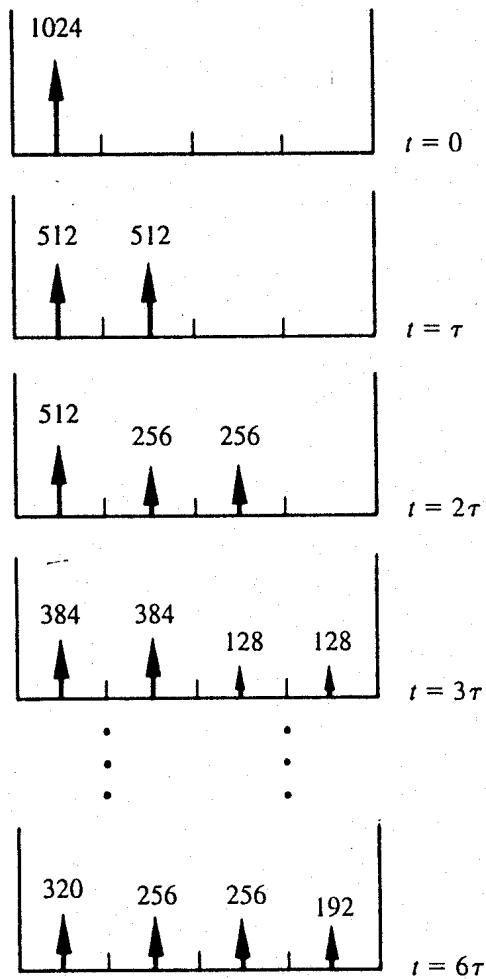
*Diffusion* is a process whereby particles tend to spread out or redistribute as a result of their random thermal motion, migrating on a macroscopic scale from regions of high particle concentration into regions of low particle concentration. If allowed to progress unabated, the diffusion process operates so as to produce a uniform distribution of particles. The diffusing entity, it should be noted, need not be charged; thermal motion, not interparticle repulsion, is the enabling action behind the diffusion process.

To cite an everyday example of diffusion, suppose an open bottle of perfume is placed in one corner of a room. Even in the absence of air currents, random thermal motion will spread the perfume molecules throughout the room in a relatively short period of time, with intermolecular collisions helping to uniformly redistribute the perfume to every nook and cranny within the room.

Seeking to obtain a more detailed understanding of the diffusion process, let us next conceptually “monitor” the process on a microscopic scale employing a simple hypothetical system. The system we propose to monitor is a one-dimensional box containing four compartments and 1024 mobile particles (see Fig. 3.11). The particles within the box obey certain stringent rules. Specifically, thermal motion causes all particles in a given compartment to “jump” into an adjacent compartment every  $\tau$  seconds. In keeping with the random nature of the motion, each and every particle has an equal probability of jumping to the left and to the right. Hitting an “external wall” while attempting to jump to the left or right reflects the particle back to its pre-jump position. Finally, at time  $t = 0$  it is assumed that all of the particles are confined in the left-most compartment.

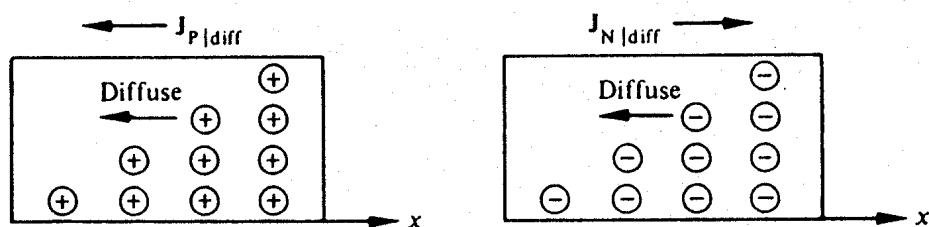
Figure 3.11 records the evolution of our 1024 particle system as a function of time. At time  $t = \tau$ , 512 of the 1024 particles originally in compartment 1 jump to the right and come to rest in compartment 2. The remaining 512 particles jump to the left and are reflected back into the left-most compartment. The end result is 512 particles each in compartments 1 and 2 after  $\tau$  seconds. At time  $t = 2\tau$ , 256 of the particles in compartment 2 jump into compartment 3 and the remainder jump back into compartment 1. In the meantime, 256 of the particles from compartment 1 jump into compartment 2 and 256 undergo a reflection at the left-hand wall. The net result after  $2\tau$  seconds is 512 particles in compartment 1 and 256 particles each in compartments 2 and 3. The state of the system after  $3\tau$  and  $6\tau$  seconds, also shown in Fig. 3.11, can be deduced in a similar manner. By  $t = 6\tau$ , the particles, once confined to the left-most compartment, have become almost uniformly distributed throughout the box, and it is unnecessary to consider later states. Indeed, the fundamental nature of the diffusion process is clearly self-evident from an examination of the existing states.

In semiconductors the diffusion process on a microscopic scale is similar to that occurring in the hypothetical system except, of course, the random motion of the diffusing particles is three-dimensional and not “compartmentalized.” On a *macroscopic* scale the net effect of diffusion is precisely the same within both the hypothetical system and semi-



**Figure 3.11** Diffusion on a microscopic scale in a hypothetical one-dimensional system. The numbers over the arrows indicate the number of particles in a given compartment; observation times are listed to the extreme right.

conductors; there is an overall migration of particles from regions of high particle concentration to regions of low particle concentration. Within semiconductors the mobile particles—the electrons and holes—are charged, and diffusion-related carrier transport therefore gives rise to particle currents as pictured in Fig. 3.12.



**Figure 3.12** Visualization of electron and hole diffusion on a macroscopic scale.

**(C) Exercise 3.3**

The “DiffDemo” MATLAB file listed below was written to help the user visualize the diffusion process. The program provides a pseudo-animation of a one-dimensional particle system similar to that described in the text. The  $y = [ ]$  statement in the program controls the initial condition, N specifies the maximum number of monitored “jumps,” and the number in the pause statements controls the time between jumps.

```
%Simulation of Diffusion (DiffDemo)
%One-dimensional system, right and left jumps equally probable

%Initialization
close
x=[0.5 1.5 2.5 3.5 4.5];
y=[1.0e6 0 0 0 0]; %NOTE: initial position can be changed
[xp,yp]=bar(x,y);
plot(xp,yp); text(0.5,1.1e6,'t = 0');
axis([0.5,0,1.2e6]);
pause(0.5)
N=15; %NOTE: increase N for extended run

%Computations and Plotting
for ii=1:N,
    %Diffusion step calculation
    bin(1)=round(y(1)/2 + y(2)/2);
    bin(2)=round(y(1)/2 + y(3)/2);
    bin(3)=round(y(2)/2 + y(4)/2);
    bin(4)=round(y(3)/2 + y(5)/2);
    bin(5)=round(y(4)/2 + y(5)/2);
    y=bin;
    %Plotting the result
    [xp,yp]=bar(x,y);
    axis(axis);
    plot(xp,yp); text(0.5,1.1e6,['t = ',num2str(ii)]);
    axis([0.5,0,1.2e6]);
    pause(0.5)
end
```

- P:** (a) Enter the program into your computer or locate the copy supplied on disk. Run the program. (The *command-period* on a Macintosh computer or *break* on an IBM-compatible computer can be used to prematurely terminate the program.)
- (b) Rerun the program after changing the initial conditions so that all of the carriers are in the middle box at  $t = 0$ .

(c) Experiment with initially placing particles in more than one box and changing the number in the pause statements. You might also try increasing the number of particle boxes.

S: Sample output is reproduced in Fig. E3.3.

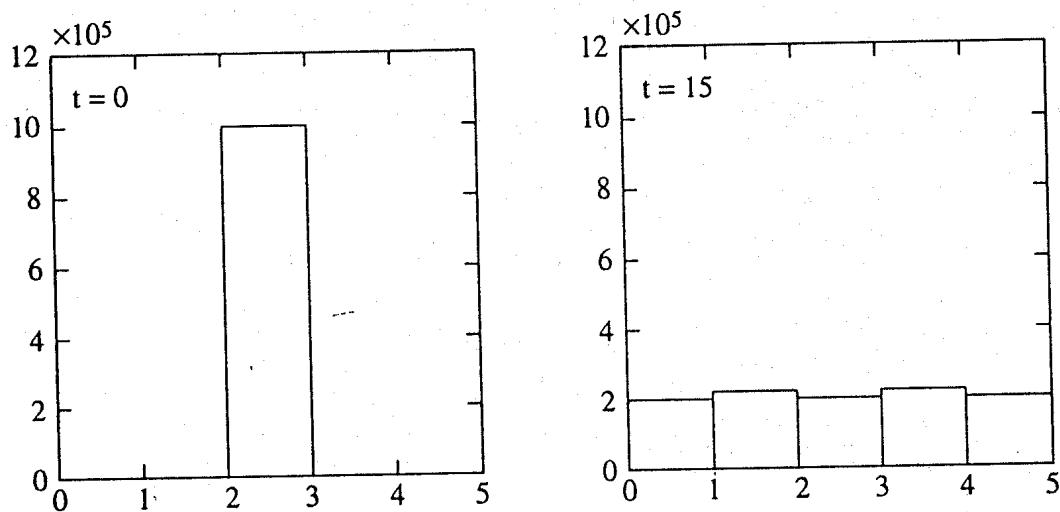


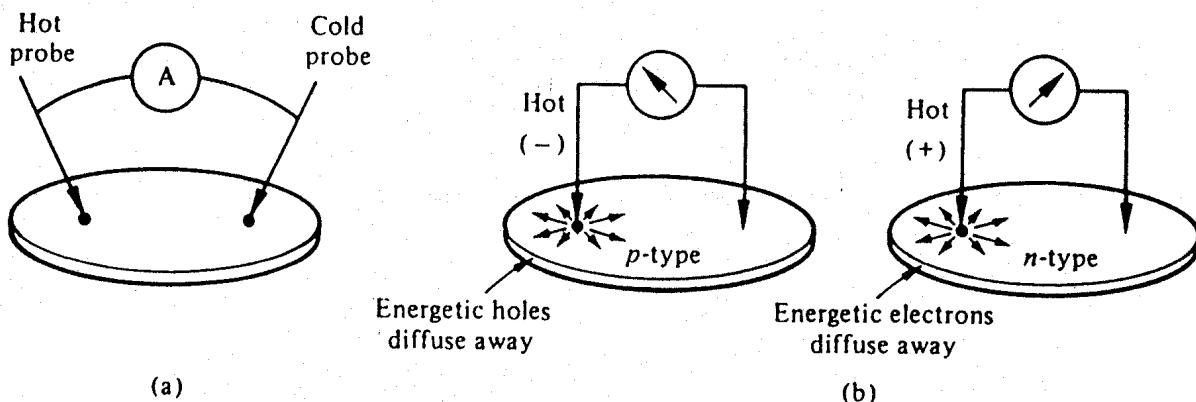
Figure E3.3

### 3.2.2 Hot-Point Probe Measurement

Before establishing the current associated with diffusion, let us digress somewhat and briefly consider the hot-point probe measurement. The *hot-point probe measurement* is a common technique for rapidly determining whether a semiconductor is *n*- or *p*-type. From a practical standpoint, knowledge of the semiconductor type is essential in device processing and must be known even before determining the doping concentration from resistivity measurements (refer back to Fig. 3.8). The hot-point probe “typing” experiment is considered here because it simultaneously provides an informative example of the diffusion process.

Examining Fig. 3.13(a), we find the only equipment required for performing the hot-point probe measurement is a hot-point probe, a cold-point probe, and a center-zero milliammeter. The hot-point probe is sometimes a refugee from a wood burning set or simply a soldering iron; the cold-point probe typically assumes the form of an electrical probe like that used with hand-held multimeters. No special requirements are imposed on the center-zero milliammeter connected between the probes. The measurement procedure itself is also extremely simple: After allowing the hot probe to heat up, one brings the two probes into contact with the semiconductor sample, and the ammeter deflects to the right or left, thereby indicating the semiconductor type. Observe that the spacing between the probes is set arbitrarily and can be reduced to enhance the ammeter deflection.

A simplified explanation of how the hot-point probe measurement works is presented in Fig. 3.13(b). In the vicinity of the probe contact the heated probe creates an increased



**Figure 3.13** The hot-point probe measurement: (a) required equipment; (b) simplified explanation of how the measurement works.

number of higher-energy carriers. These energetic carriers will be predominantly holes in the case of a *p*-type material and electrons in an *n*-type material. With more energetic carriers near the heated probe than elsewhere, diffusion acts so as to spread the higher-energy carriers throughout the semiconductor wafer. The net effect is a deficit of holes or a net negative charge in the vicinity of the hot-point probe for a *p*-type material, and a positive charge buildup near the heated probe within an *n*-type material. Accordingly, the center-zero meter deflects in a different direction for *p*- and *n*-type materials.

### 3.2.3 Diffusion and Total Currents

#### Diffusion Currents

In defining the diffusion process and in citing diffusion examples, we have sought to emphasize the direct correlation between diffusion and a spatial variation in particle numbers. For diffusion to occur, more of the diffusing particles must exist at one point than at other points or, in mathematical terms, there must be a nonzero concentration gradient ( $\nabla p \neq 0$  for holes,  $\nabla n \neq 0$  for electrons). Logically, moreover, the greater the concentration gradient, the larger the expected flux of particles. Quantitative analysis of the diffusion process indeed confirms the foregoing and leads to what is known as Fick's law:

$$\mathcal{F} = -D\nabla\eta \quad (3.16)$$

where  $\mathcal{F}$  is the flux or particles/cm<sup>2</sup>-sec crossing a plane perpendicular to the particle flow,  $\eta$  is the particle concentration, and  $D$  is the *diffusion coefficient*, a positive proportionality factor. The diffusion current density due to electrons and holes is obtained by simply multiplying the carrier flux by the carrier charge:

$$J_{P\text{ldiff}} = -qD_p\nabla p \quad (3.17a)$$

$$J_{N\text{ldiff}} = qD_n\nabla n \quad (3.17b)$$

The constants of proportionality,  $D_p$  and  $D_N$ , have units of  $\text{cm}^2/\text{sec}$  and are referred to as the hole and electron diffusion coefficients, respectively.

Upon examining Eqs. (3.17), note that the current directions deduced from the equations are consistent with the Fig. 3.12 visualization of the macroscopic diffusion currents. For the positive concentration gradient shown in Fig. 3.12 ( $dp/dx > 0$  and  $dn/dx > 0$  for the pictured one-dimensional situation), both holes and electrons will diffuse in the  $-x$  direction.  $J_{p\text{diff}}$  will therefore be negative or directed in the  $-x$  direction, while  $J_{N\text{diff}}$  will be oriented in the  $+x$  direction.

### Total Currents

The total or net carrier currents in a semiconductor arise as the combined result of drift and diffusion. Summing the respective  $n$  and  $p$  segments of Eqs. (3.4) and Eqs. (3.17), one obtains

$$\mathbf{J}_p = \mathbf{J}_{p\text{drift}} + \mathbf{J}_{p\text{diff}} = q\mu_p p \mathbf{E} - qD_p \nabla p \quad (3.18a)$$

$\downarrow$  drift       $\downarrow$  diffusion

$$\mathbf{J}_N = \mathbf{J}_{N\text{drift}} + \mathbf{J}_{N\text{diff}} = q\mu_n n \mathbf{E} + qD_N \nabla n \quad (3.18b)$$

The total particle current flowing in a semiconductor is in turn computed from

$$\mathbf{J} = \mathbf{J}_N + \mathbf{J}_p \quad (3.19)$$

The double boxes emphasize the importance of the total-current relationships; they are used directly or indirectly in the vast majority of device analyses.

### 3.2.4 Relating Diffusion Coefficients/Mobilities

The diffusion coefficients, the constants of the motion associated with diffusion, are obviously central parameters in characterizing carrier transport due to diffusion. Given the importance of the diffusion coefficients, one might anticipate an extended examination of relevant properties paralleling the mobility presentation in Subsection 3.1.3. Fortunately, an extended examination is unnecessary because the  $D$ 's and the  $\mu$ 's are all interrelated. It is only necessary to establish the connecting formula known as the Einstein relationship.

In deriving the Einstein relationship, we consider a *nonuniformly* doped semiconductor maintained under equilibrium conditions. Special facts related to the nonuniform doping and the equilibrium state are invoked in the derivation. These facts, important in themselves, are reviewed prior to presenting the derivation proper.

### Constancy of the Fermi Level

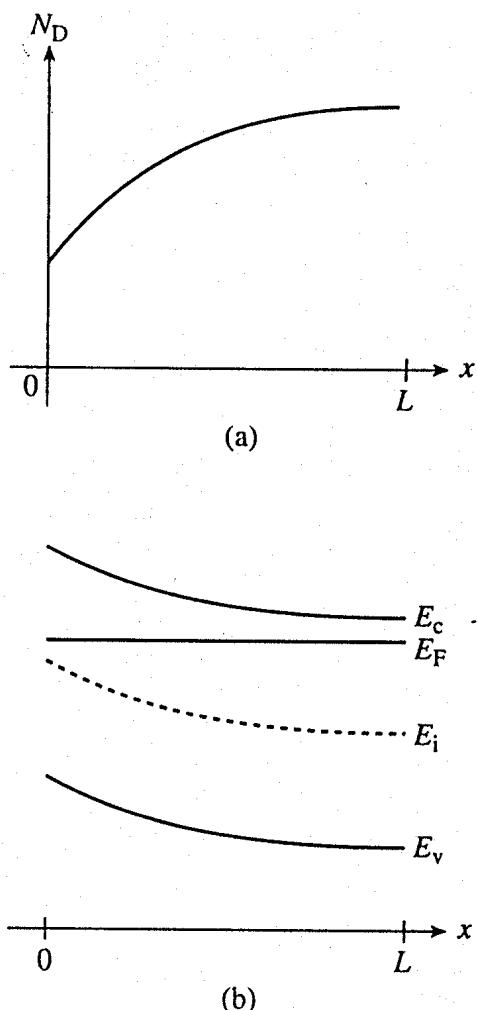
Consider a nondegenerate, nonuniformly doped  $n$ -type semiconductor sample, a sample where the doping concentration varies as a function of position. A concrete example of what we have in mind is shown in Fig. 3.14(a). Assuming equilibrium conditions prevail,

the energy band diagram characterizing the sample will have the form sketched in Fig. 3.14(b). In Chapter 2 the Fermi level in *uniformly* doped *n*-type semiconductors was found to move closer and closer to  $E_c$  when the donor doping was systematically increased (see Fig. 2.21). Consistent with this fact,  $E_c$  was drawn closer to  $E_F$  in going from  $x = 0$  to  $x = L$  on the Fig. 3.14(b) diagram. Construction of the diagram, however, could not have been completed without making use of a critical new piece of information. Specifically,

*under equilibrium conditions,  $dE_F/dx = dE_F/dy = dE_F/dz = 0$ ; i.e., the Fermi level inside a material or a group of materials in intimate contact is invariant as a function of position.*

The constancy of the Fermi level means that  $E_F$  is to appear as a horizontal line on equilibrium energy band diagrams.

The position independence of the Fermi energy is established by examining the trans-



**Figure 3.14** Nonuniformly doped semiconductor: (a) assumed doping variation with position; (b) corresponding equilibrium energy band diagram.

fer of carriers between allowed states with the same energy but at adjacent positions in an energy band. It is concluded the probability of filling the states at a given energy,  $f(E)$ , must be the same everywhere in the sample under equilibrium conditions. If this were not the case, carriers would preferentially transfer between states and thereby give rise to a net current. The existence of a net current is inconsistent with the specified equilibrium conditions. Referring to the Eq. (2.7) expression for  $f(E)$ , we find that the constancy of the Fermi function in turn requires the Fermi level to be independent of position.

### Current Flow Under Equilibrium Conditions

Under equilibrium conditions the total current is of course identically zero. Because electron and hole activity is totally decoupled under equilibrium conditions, the electron and hole current densities,  $J_N$  and  $J_P$ , must also independently vanish. Setting  $J_N = J_P = 0$  in Eq. (3.18), however, leads to the conclusion that the drift and diffusion currents associated with a given carrier are merely required to be of equal magnitude and opposite polarity. In fact, the drift and diffusion components will vanish under equilibrium conditions only if  $\mathcal{E} = 0$  and  $\nabla n = \nabla p = 0$ .

To provide a concrete illustration of non-vanishing equilibrium current components, let us return to an examination of the nonuniformly doped semiconductor sample characterized by Fig. 3.14. The dopant variation pictured in Fig. 3.14(a) clearly gives rise to a significant electron concentration gradient. An electron diffusion current flowing in the  $+x$  direction must exist inside the sample. The band bending shown in the band diagram of Fig. 3.14(b), moreover, implies the existence of a "built-in" electric field oriented in the  $-x$  direction. This gives rise to a drift current also in the  $-x$  direction. Note that the drift and diffusion components are of opposite polarity. Since equilibrium conditions are assumed to prevail, the components must also be of equal magnitude. In point of fact, the electric field arises inside the semiconductor to counteract the diffusive tendencies of the carriers caused by the nonuniform doping. The major point here is that, even under equilibrium condition, nonuniform doping will give rise to carrier concentration gradients, a built-in electric field, and non-zero current components.

### Einstein Relationship

Having laid the proper foundation, we can now proceed to derive the connecting formula between the  $D$ 's and the  $\mu$ 's known as the Einstein relationship. To simplify the development, we work in only one dimension. The sample under analysis is taken to be a nondegenerate, nonuniformly doped semiconductor maintained under equilibrium conditions. Citing the fact that the net carrier currents must be identically zero under equilibrium conditions, and focusing on the electrons, we can state

$$J_{N|drift} + J_{N|diff} = q\mu_n n \mathcal{E} + qD_N \frac{dn}{dx} = 0 \quad (3.20)$$

However,

$$\mathcal{E} = \frac{1}{q} \frac{dE_i}{dx} \quad (3.21)$$

(same as 3.15)

and

$$n = n_i e^{(E_F - E_i)/kT} \quad (3.22)$$

Moreover, with  $dE_F/dx = 0$  (due to the positional invariance of the Fermi level under equilibrium conditions),

$$\frac{dn}{dx} = -\frac{n_i}{kT} e^{(E_F - E_i)/kT} \frac{dE_i}{dx} = -\frac{q}{kT} n \mathcal{E} \quad (3.23)$$

Substituting  $dn/dx$  from Eq. (3.23) into Eq. (3.20), and rearranging the result slightly, one obtains

$$(qn\mathcal{E})\mu_n - (qn\mathcal{E})\frac{q}{kT} D_N = 0 \quad (3.24)$$

Since  $\mathcal{E} \neq 0$  (a consequence of the nonuniform doping), it follows from Eq. (3.24) that

$$\frac{D_N}{\mu_n} = \frac{kT}{q}$$

Einstein relationship for electrons (3.25a)

A similar argument for holes yields

$$\frac{D_P}{\mu_p} = \frac{kT}{q}$$

Einstein relationship for holes (3.25b)

Although it was established while assuming equilibrium conditions, we can present more elaborate arguments that show the Einstein relationship to be valid even under non-equilibrium conditions. The nondegenerate restriction, however, still applies; slightly modified forms of Eqs. (3.25) result when the argument is extended to degenerate mate-

rials. We should emphasize again that the preliminary results, such as the positional invariance of the equilibrium Fermi level and the situation inside a nonuniformly doped semiconductor under equilibrium conditions, are important in themselves. Relative to numerical values, note that  $kT/q$  is a voltage, and at room temperature (300 K) is equal to 0.0259 V.<sup>†</sup> Hence, for an  $N_D = 10^{14}/\text{cm}^3$  Si sample maintained at room temperature,  $D_N = (kT/q)\mu_n = (0.0259)(1358 \text{ cm}^2/\text{V}\cdot\text{sec}) = 35.2 \text{ cm}^2/\text{sec}$ . Finally, the Einstein relationship is one of the easiest equations to remember because it rhymes internally— $D$  over  $\mu$  equals  $kT$  over  $q$ . The rhyme holds even if the equation is inverted— $\mu$  over  $D$  equals  $q$  over  $kT$ .

### Exercise 3.4

**P:** In this exercise we continue the examination of the semiconductor sample represented by the energy band diagram presented in Exercise 3.2.

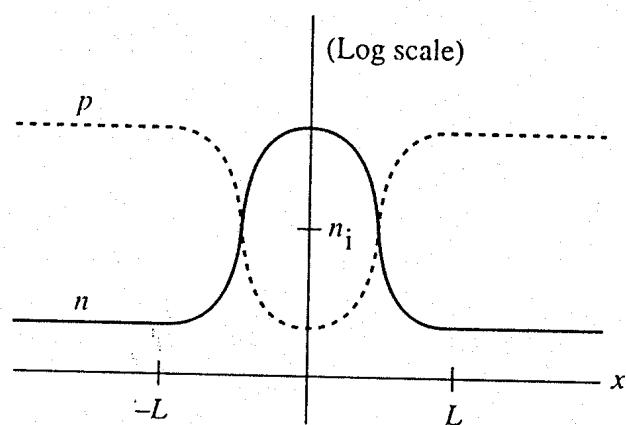
- (a) The semiconductor is in equilibrium. How does one deduce this fact from the given energy band diagram?
- (b) What is the electron current density ( $J_N$ ) and hole current density ( $J_P$ ) at  $x = \pm L/2$ ?
- (c) Roughly sketch  $n$  and  $p$  versus  $x$  inside the sample.
- (d) Is there an electron diffusion current at  $x = \pm L/2$ ? If there is a diffusion current at a given point, indicate the direction of *current* flow.
- (e) Noting the solution for the electric field presented in Exercise 3.2, is there an electron drift current at  $x = \pm L/2$ ? If there is a drift current at a given point, indicate the direction of current flow.
- (f) What is the hole diffusion coefficient ( $D_P$ ) in the  $x > L$  region of the semiconductor?

**S:** (a) The semiconductor is concluded to be in equilibrium because the Fermi level is invariant as a function of position.

- (b)  $J_N = 0$  and  $J_P = 0$  at both  $x = -L/2$  and  $x = +L/2$ . The net electron and hole currents are always identically zero everywhere under equilibrium conditions.

<sup>†</sup> At room temperature  $kT = 0.0259 \text{ eV} = (1.6 \times 10^{-19})(0.0259) \text{ joule}$ . Thus  $kT/q = (1.6 \times 10^{-19})(0.0259)/(1.6 \times 10^{-19} \text{ joule/coul} = 0.0259 \text{ V}$ .

(c) Since  $n = n_i \exp[(E_F - E_i)/kT]$  and  $p = n_i \exp[(E_i - E_F)/kT]$ , we conclude



(d) There is an electron diffusion current at both  $x = -L/2$  and  $x = +L/2$ . From the answer to the preceding question we note that  $dn/dx \neq 0$  at the cited points. With  $dn/dx > 0$ ,  $J_{N\text{ldiff}}$  at  $x = -L/2$  flows in the  $+x$  direction. Conversely,  $dn/dx < 0$  at  $x = L/2$  and  $J_{N\text{ldiff}}$  at this point flows in the  $-x$  direction.

(e) There is an electron drift current at both  $x = -L/2$  and  $x = +L/2$ . Since both  $n$  and  $\mathcal{E}$  are non-zero at  $x = \pm L/2$ , it follows that  $J_{N\text{ldrift}} = q\mu_n n \mathcal{E} \neq 0$ . The drift component of the current always has the same direction as the electric field;  $J_{N\text{ldrift}}$  is in the  $-x$  direction at  $x = -L/2$  and in the  $+x$  direction at  $x = +L/2$ . The drift component of the current must of course cancel the diffusion component of the current so that  $J_N = J_{N\text{ldrift}} + J_{N\text{ldiff}} = 0$  under equilibrium conditions. The direction-related answers to parts (d) and (e), summarized next, are consistent with this requirement:

	$x = -L/2$	$x = +L/2$
$J_{N\text{ldiff}}$	→	←
$J_{N\text{ldrift}}$	←	→

(f) In working part (d) of Exercise 3.2, we concluded the  $x > L$  region of the semiconductor (Si, 300 K) had a doping of  $N_A \cong 5 \times 10^{14}/\text{cm}^3$  and a corresponding hole mobility of  $\mu_p = 459 \text{ cm}^2/\text{V}\cdot\text{sec}$ . Thus, employing the Einstein relationship,  $D_p = (kT/q)\mu_p = (0.0259)(459) = [11.9 \text{ cm}^2/\text{sec}]$ .

### 3.3 RECOMBINATION-GENERATION

#### 3.3.1 Definition-Visualization

When a semiconductor is perturbed from the equilibrium state, an excess or deficit in the carrier concentrations relative to their equilibrium values is invariably created inside the semiconductor. Recombination-generation (R-G) is nature's order-restoring mechanism, the means whereby the carrier excess or deficit inside the semiconductor is stabilized (if the perturbation is maintained) or eliminated (if the perturbation is removed). Since non-equilibrium conditions prevail during device operation, recombination-generation more often than not plays a major role in shaping the characteristics exhibited by a device. Formally, recombination and generation can be defined as follows:

*Recombination*—a process whereby electrons and holes (carriers) are annihilated or destroyed.

*Generation*—a process whereby electrons and holes are created.

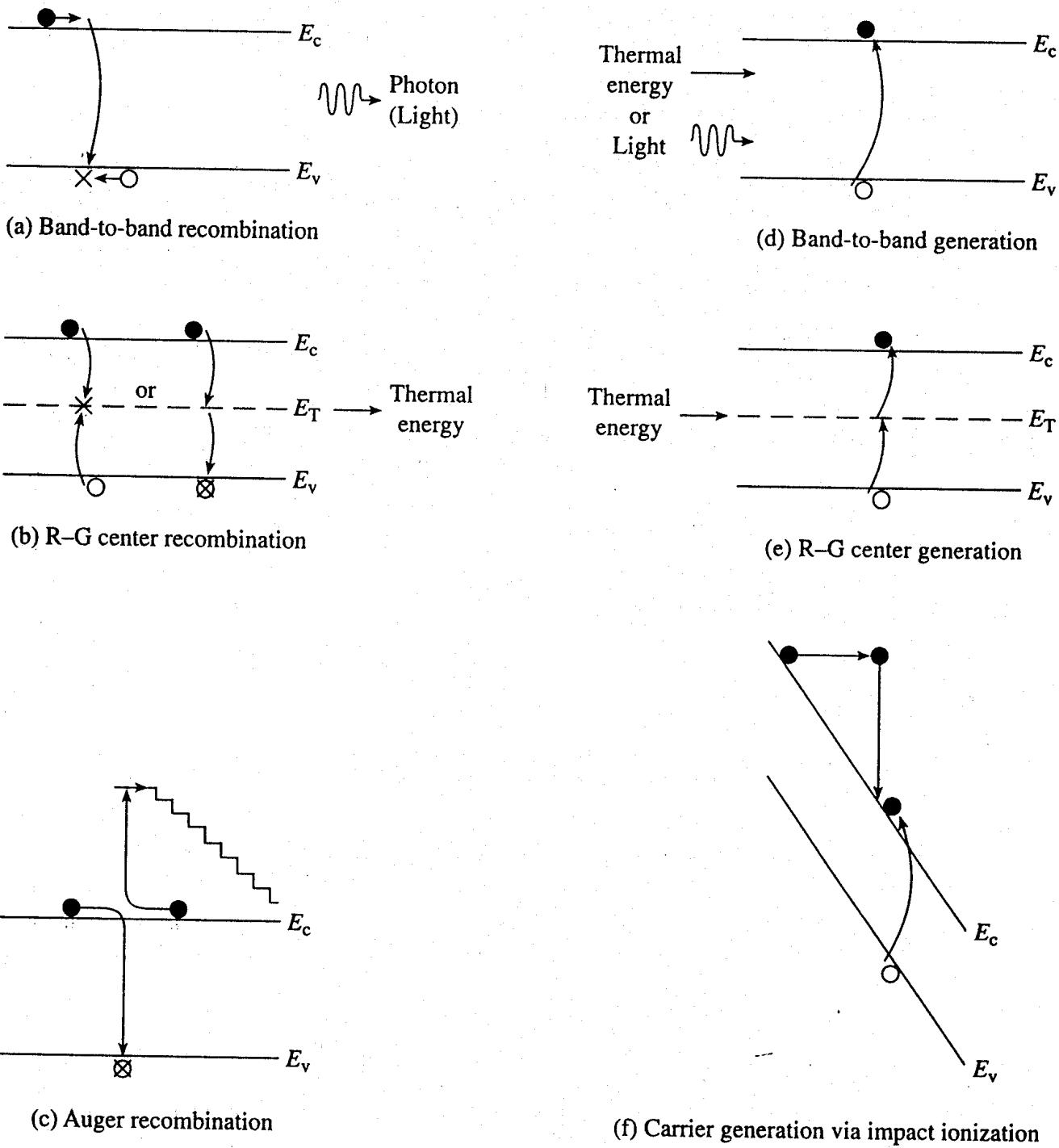
Unlike drift and diffusion, the terms *recombination* and *generation* do not refer to a single process. Rather, they are collective names for a group of similar processes; carriers can be created and destroyed within a semiconductor in a number of ways. The most common R-G processes are visualized in Fig. 3.15. The individual processes are described and discussed next.

#### Band-to-Band Recombination

Band-to-band recombination is conceptually the simplest of all recombination processes. As pictured in Fig. 3.15(a), it merely involves the direct annihilation of a conduction band electron and a valence band hole. An electron and hole moving in the semiconductor lattice stray into the same spatial vicinity and zap!—the electron and hole annihilate each other. The excess energy released during the process typically goes into the production of a photon (light).

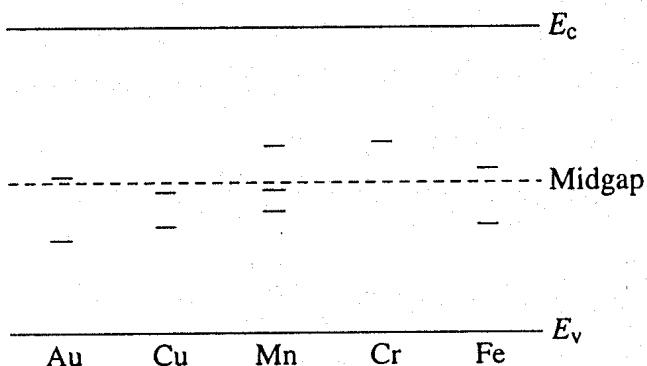
#### R-G Center Recombination

The process pictured in Fig. 3.15(b) involves a “third party,” or intermediary, and takes place only at special locations within the semiconductor known as R-G centers. Physically, R-G centers are lattice defects or special impurity atoms such as gold in Si. Lattice defects and the special atoms in the form of unintentional impurities are present even in semiconductors of the highest available purity. The R-G center concentration, however, is normally very low compared to the acceptor and donor concentrations in device-quality materials. The most important property of the R-G centers is the introduction of allowed electronic levels near the center of the band gap. These levels are identified by the  $E_T$  energy label in Fig. 3.15(b). The near midgap positioning of the levels is all-important because it distinguishes R-G centers from donors and acceptors. The midgap levels introduced by select R-G center impurities in Si are summarized in Fig. 3.16.



**Figure 3.15** Energy band visualization of recombination and generation processes.

As noted in Fig. 3.15(b), recombination at an R-G center is a two-step process. First, one type of carrier, say an electron, strays into the vicinity of an R-G center, is caught by the potential well associated with the center, loses energy, and is trapped. Subsequently, a hole comes along, is attracted to the trapped electron, loses energy, and annihilates the electron within the center. Alternatively, as also shown in Fig. 3.15(b), one can think of the electron as losing energy a second time and annihilating the hole in the valence band. R-G center recombination, also called *indirect recombination*, typically releases thermal energy (heat) during the process or, equivalently, produces lattice vibrations.



**Figure 3.16** Near-midgap energy levels introduced by some common impurities in Si. When an impurity introduces multiple levels, one of the levels tends to dominate in a given semiconductor sample.

### Auger Recombination

In the Auger (pronounced Oh-jay) process pictured in Fig. 3.15(c), band-to-band recombination occurs simultaneously with the collision between two like carriers. The energy released by the recombination is transferred during the collision to the surviving carrier. This highly energetic carrier subsequently “thermalizes”—loses energy in small steps through heat-producing collisions with the semiconductor lattice. The “staircase” in Fig. 3.15(c) represents the envisioned stepwise loss of energy.

### Generation Processes

Any of the foregoing recombination processes can be reversed to generate carriers. Band-to-band generation, where an electron is excited directly from the valence band into the conduction band, is pictured in Fig. 3.15(d). Note that either thermal energy or light can provide the energy required for the band-to-band transition. If thermal energy is absorbed, the process is alternatively referred to as *direct thermal generation*; if externally introduced light is absorbed, the process is called *photogeneration*. The thermally assisted generation of carriers with R–G centers acting as intermediaries is envisioned in Fig. 3.15(e). Finally, impact ionization, the inverse of Auger recombination, is visualized in Fig. 3.15(f). In this process an electron–hole pair is produced as a result of energy released when a highly energetic carrier collides with the crystal lattice. The generation of carriers through impact ionization routinely occurs in the high  $E$ -field regions of devices. We will have more to say about this process when treating breakdown in *pn* junctions.

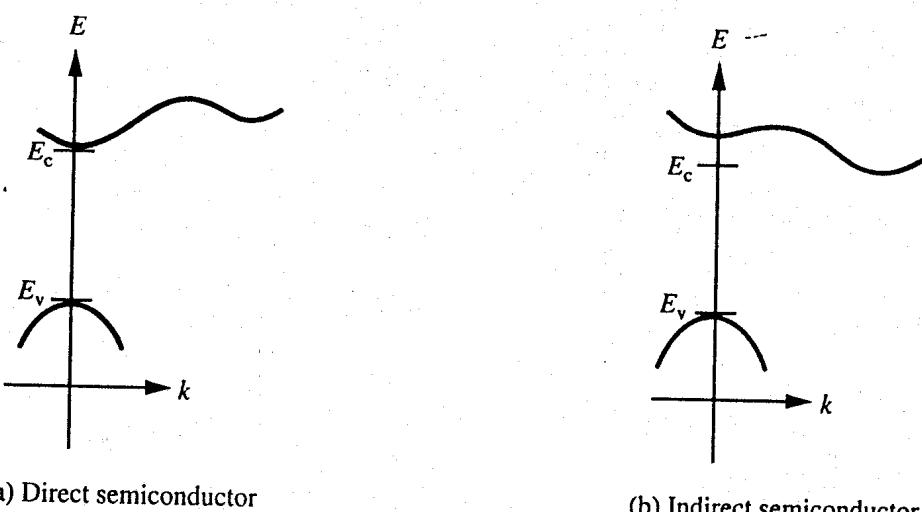
#### 3.3.2 Momentum Considerations

All of the various recombination–generation processes occur at all times in all semiconductors—even under equilibrium conditions. The critical issue is not whether the processes are occurring, but rather, the rates at which the various processes are occurring. Typically, one need be concerned only with the dominant process, the process proceeding at the fastest rate. A number of the processes are important only under special conditions or are highly improbable, occurring at a much slower rate than competing processes. Auger recombi-

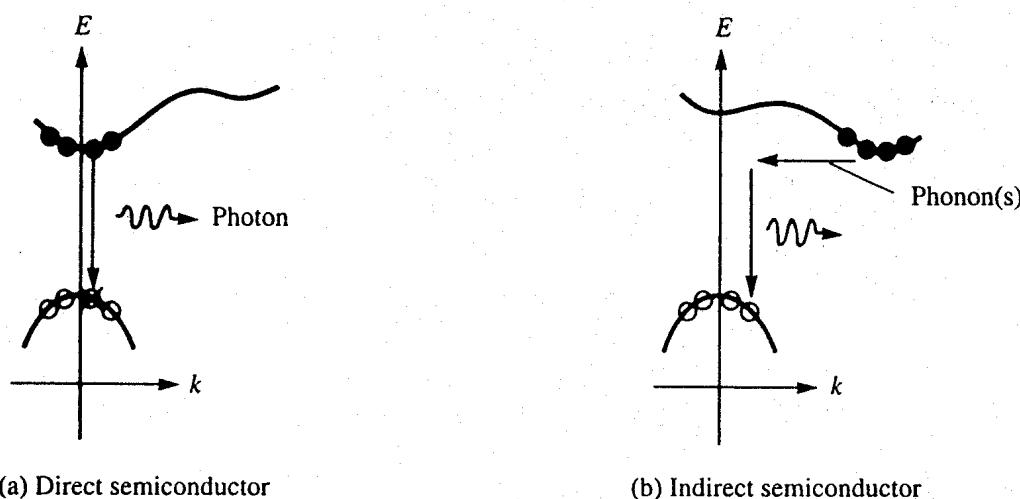
tion provides a case in point. Because the number of carrier-carrier collisions increases with increased carrier concentration, the frequency of Auger recombination likewise increases with carrier concentration. Auger recombination must be considered, for example, in treating highly doped regions of a device structure, but it is usually negligible otherwise. Similarly, significant impact ionization occurs only in the high  $E$ -field regions of a device. In a somewhat different vein, photogeneration at room temperature becomes a significant process only when the semiconductor is exposed to external illumination. What then is the R-G process pair deserving of special attention, the process pair that typically dominates in the low  $E$ -field regions of a nondegenerately doped semiconductor maintained at room temperature? From the information provided, the choice is obviously between thermal band-to-band R-G and recombination-generation via R-G centers.

From the energy band explanation alone one might speculate that band-to-band recombination-generation would dominate under the cited "standard" conditions. However, visualization of R-G processes using the energy band diagram can be misleading. The  $E$ - $x$  plot examines only changes in energy, whereas crystal momentum in addition to energy must be conserved in any R-G process. Momentum conservation requirements, as it turns out, play an important role in setting the process rate.

As first noted in Subsection 2.2.3, the momentum of an electron in an energy band, like the electron energy, can assume only certain quantized values. Momentum-related aspects of R-G processes are conveniently discussed with the aid of plots where the allowed electron energies in the conduction and valence bands are plotted versus the allowed momentum. These are referred to as  $E$ - $k$  plots, where  $k$  is a parameter proportional to the electron crystal momentum. The  $E$ - $k$  plots associated with semiconductors fall into two general categories. In the first, sketched in Fig. 3.17(a), both the minimum conduction band energy and maximum valence band energy occur at  $k = 0$ . In the second, pictured in Fig. 3.17(b), the conduction band minimum is displaced to a  $k \neq 0$ . Semiconductors exhibiting the former type of plot are referred to as *direct semiconductors*, and semiconductors exhibiting the latter type of plot, as *indirect semiconductors*. GaAs is a notable member of the direct semiconductor family; Ge and Si are both indirect semiconductors.



**Figure 3.17** General forms of  $E$ - $k$  plots for direct and indirect semiconductors.



**Figure 3.18**  $E$ - $k$  plot visualizations of recombination in direct and indirect semiconductors.

To make use of the  $E$ - $k$  plots in visualizing an R-G process, one needs to know the nature of transitions associated with the absorption or emission of photons (light) and lattice vibration quanta called *phonons*. Photons, being massless entities, carry very little momentum, and a photon-assisted transition is essentially vertical on the  $E$ - $k$  plot. Conversely, the thermal energy associated with lattice vibrations (phonons) is very small (in the 10–50 meV range), whereas the phonon momentum is comparatively large. Thus on an  $E$ - $k$  plot a phonon-assisted transition is essentially horizontal. It is also important to point out that, as pictured in Fig. 3.18, electrons and holes normally only occupy states very close to the  $E_c$  minimum and  $E_v$  maximum, respectively. This is of course consistent with our earlier discussion of the distribution of carriers in the bands.

Let us now specifically re-examine the band-to-band recombination process. In a direct semiconductor where the  $k$ -values of electrons and holes are all bunched near  $k = 0$ , little change in momentum is required for the recombination process to proceed. The conservation of both energy and momentum is readily met simply by the emission of a photon (see Fig. 3.18a). In an indirect semiconductor, on the other hand, there is a large change in momentum associated with the recombination process. The emission of a photon will conserve energy but cannot simultaneously conserve momentum. Thus for band-to-band recombination to proceed in an indirect semiconductor, the emission of a photon must be accompanied by the emission or absorption of a phonon (see Fig. 3.18b).

The rather involved nature of the band-to-band process in indirect semiconductors understandably leads to a vastly diminished recombination rate. Band-to-band recombination is in fact totally negligible compared to R-G center recombination in Si and other indirect semiconductors. Band-to-band recombination does proceed at a much faster rate in direct semiconductors and is the process producing the light observed from LEDs and junction lasers. Even in direct materials, however, the R-G center mechanism is often the dominant process. Given the universal importance of the R-G center process, and the fact that most devices are made from Si, which is an indirect material, the quantitative considerations in the next subsection are devoted primarily to the R-G center mechanism. Other processes are considered as the need arises.

### 3.3.3 R-G Statistics

R-G statistics is the technical name given to the mathematical characterization of recombination-generation processes. "Mathematical characterization" in the case of recombination-generation does not mean the development of a current-density relationship. An R-G event occurs at localized positions in the crystal, at a given  $x$ -value, and therefore the R-G action does not lead *per se* to charge transport. Rather, recombination-generation acts to change the carrier concentrations, thereby indirectly affecting the current flow. It is the time rate of change in the carrier concentrations ( $\partial n/\partial t$ ,  $\partial p/\partial t$ ) that must be specified to achieve the desired mathematical characterization. Because its characterization is relatively simple and its involvement in device operation quite common, we begin by considering photogeneration. The major portion of the subsection is of course devoted to recombination-generation via R-G centers, alternatively referred to as indirect thermal recombination-generation.

#### Photogeneration

Light striking the surface of a semiconductor as pictured in Fig. 3.19 will be partially reflected and partially transmitted into the material. Assume the light to be monochromatic with wavelength  $\lambda$  and frequency  $\nu$ . If the photon energy ( $h\nu$ ) is greater than the band gap energy, then the light will be absorbed and electron-hole pairs will be created as the light passes through the semiconductor. The decrease in intensity as monochromatic light passes through a material is given by

$$I = I_0 e^{-\alpha x} \quad (3.26)$$

where  $I_0$  is the light intensity just inside the material at  $x = 0^+$  and  $\alpha$  is the *absorption coefficient*. Note from Fig. 3.20 that  $\alpha$  is material dependent and a strong function of  $\lambda$ . Since there is a one-to-one correspondence between the absorption of photons and the creation of electron-hole pairs, the carrier creation rate should likewise exhibit an  $\exp(-\alpha x)$  dependence. Moreover, referring to Fig. 3.15(d), the photogeneration process always acts so as to create an *equal* number of electrons and holes. We therefore conclude

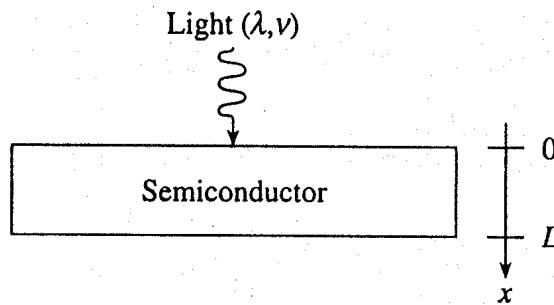
$$\left. \frac{\partial n}{\partial t} \right|_{\text{light}} = \left. \frac{\partial p}{\partial t} \right|_{\text{light}} = G_L(x, \lambda) \quad (3.27)$$

where, for the situation analyzed,

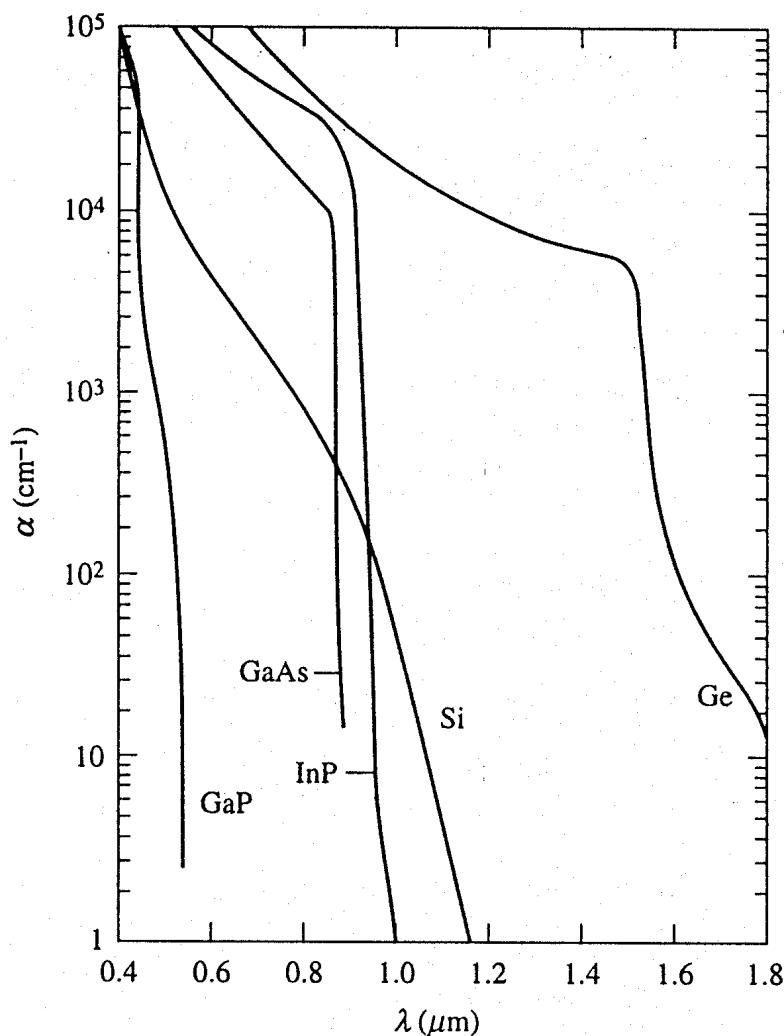
$$G_L(x, \lambda) = G_{L0} e^{-\alpha x} \quad (3.28)$$

$G_L$  is the simplified symbol for the photogeneration rate (a number/cm<sup>3</sup>-sec), while  $G_{L0}$  is the photogeneration rate at  $x = 0$ . The  $|_{\text{light}}$  designation means "due to light" and is necessary because, as we have seen,  $n$  and  $p$  can be affected by a number of processes.

In working problems involving the illumination of a sample or device, we will often invoke one of two simplifying assumptions: (i) The illumination is assumed to be uniform everywhere inside the sample and (ii) the light is assumed to be absorbed in an



**Figure 3.19** Semiconductor, light propagation, and coordinate orientation in the photogeneration analysis. The semiconductor is assumed to be sufficiently thick ( $L \gg 1/\alpha$ ) so that bottom surface reflections may be neglected.



**Figure 3.20** The absorption coefficient as a function of wavelength in Si and other select semiconductors. (From Schroder<sup>[6]</sup>, © 1990 by John Wiley & Sons, Inc. Reprinted with permission.)

infinitesimally thin layer near the semiconductor surface. Seeking to identify when the assumed conditions are approached in practice, we note first that  $1/\alpha$ , which has the dimensions of a length, represents the average depth of penetration of the light into a material. At wavelengths where  $h\nu \approx E_G$ , the absorption coefficient is very small and  $1/\alpha$  is very large, about 1 cm in Si. This makes  $\exp(-\alpha x) \approx 1$  for all practical depths. To a very good approximation,  $G_L \cong \text{constant}$  and the light is being uniformly absorbed. On the other hand, if  $\lambda = 0.4 \mu\text{m}$ , an  $\alpha(\text{Si}) \cong 10^5 \text{ cm}^{-1}$  is deduced from Fig. 3.20. At this wavelength the light penetration depth is only a few thousand angstroms and the assumption of surface absorption is closely approximated.

### Indirect Thermal Recombination-Generation

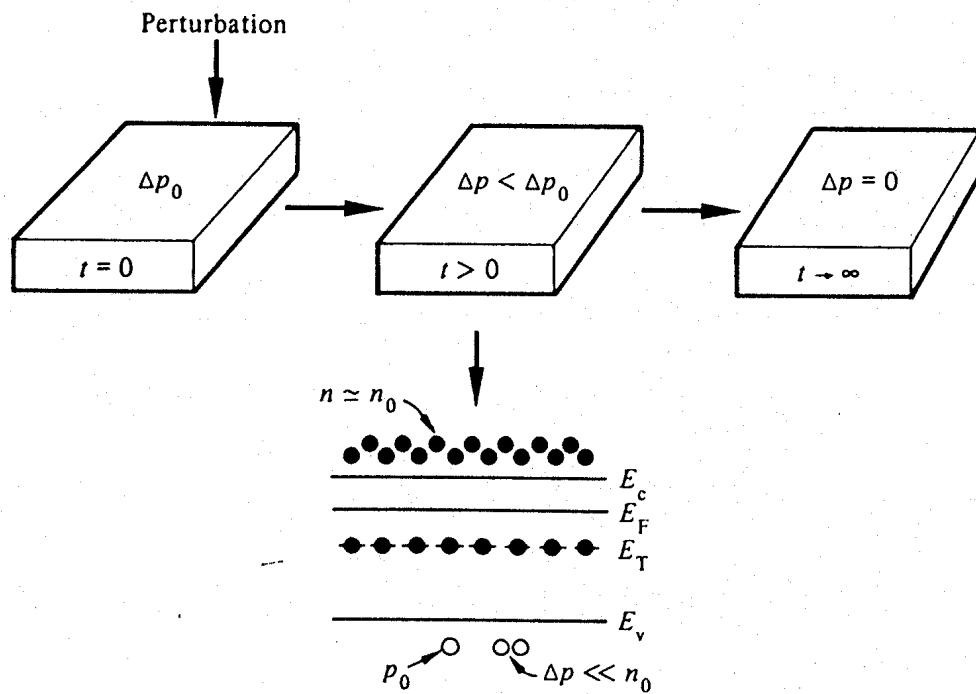
Although the general case development of R-G center statistics is beyond the scope of this text, it is possible to present arguments leading to the special-case expression employed in many practical problems. Let us begin by carefully defining the various carrier and center concentrations involved in the analysis.

$n_0, p_0$	... carrier concentrations in the material under analysis when equilibrium conditions prevail.
$n, p$	... carrier concentrations in the material under arbitrary conditions.
$\Delta n \equiv n - n_0$	deviations in the carrier concentrations from their equilibrium values.
$\Delta p \equiv p - p_0$	$\Delta n$ and $\Delta p$ can be both positive and negative, where a positive deviation corresponds to a carrier excess and a negative deviation corresponds to a carrier deficit.
$N_T$	... number of R-G centers/cm <sup>3</sup> .

In the special-case development, perturbation of the carrier concentrations from their equilibrium values must correspond to *low-level injection*. This requirement is just an alternative way of saying the perturbation must be relatively small. To be more precise,

$\text{low-level injection implies}$ $\Delta p \ll n_0, \quad n \approx n_0 \quad \text{in an } n\text{-type material}$ $\Delta n \ll p_0, \quad p \approx p_0 \quad \text{in a } p\text{-type material}$
-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Consider a specific example of  $N_D = 10^{14}/\text{cm}^3$  doped Si at room temperature subject to a perturbation where  $\Delta p = \Delta n = 10^9/\text{cm}^3$ . For the given material,  $n_0 \approx N_D = 10^{14}/\text{cm}^3$  and  $p_0 \approx n_i^2/N_D \approx 10^6/\text{cm}^3$ . Consequently,  $n = n_0 + \Delta n \approx n_0$  and  $\Delta p = 10^9/\text{cm}^3 \ll n_0 = 10^{14}/\text{cm}^3$ . The situation is clearly one of low-level injection. Observe, however, that  $\Delta p \gg p_0$ . Although the majority carrier concentration remains essentially unperturbed under low-level injection, the minority carrier concentration can, and routinely does, increase by many orders of magnitude.



**Figure 3.21** Situation inside an  $n$ -type semiconductor after a perturbation causing the low-level injection of excess holes.

We are now ready to analyze the specific situation envisioned in Fig. 3.21. A perturbation has caused a  $\Delta p \ll n_0$  excess of holes in an  $n$ -type semiconductor sample. Assuming the system is in the midst of relaxing back to the equilibrium state from the pictured perturbed state via the R-G center interaction, what factors would you expect to have the greatest effect on  $\partial p / \partial t|_R$ , the rate of hole recombination? Well, to eliminate a hole, the hole must make the transition from the valence band to an electron-filled R-G center. Logically, the greater the number of filled R-G centers, the greater the probability of a hole annihilating transition and the faster the rate of recombination. Under equilibrium conditions, essentially all of the R-G centers are filled with electrons because the equilibrium Fermi level is comfortably above  $E_T$ . With  $\Delta p \ll n_0$ , electrons always vastly outnumber holes and rapidly fill R-G levels that become vacant. This keeps the number of filled R-G centers during the relaxation process  $\approx N_T$ . Thus, we expect  $\partial p / \partial t|_R$  to be approximately proportional to  $N_T$ . Moreover, it also seems logical that the number of hole-annihilating transitions should increase almost linearly with the number of holes. All other factors being equal, the more holes available for annihilation, the greater the number of holes removed from the valence band per second. Consequently, we also expect  $\partial p / \partial t|_R$  to be approximately proportional to  $p$ . Considering additional factors leads to no additional dependencies. Hence, introducing a positive proportionality constant,  $c_p$ , and realizing  $\partial p / \partial t|_R$  is negative because  $p$  is decreasing, we conclude

$$\left. \frac{\partial p}{\partial t} \right|_R = -c_p N_T p \quad (3.29)$$

If the hole generation process is analyzed in a similar manner,  $\partial p/\partial t|_G$  is deduced to depend only on the number of empty R-G centers. This number, which is small for the situation under consideration, is held approximately constant at its equilibrium value over the range of perturbations where hole generation is significant. This means that  $\partial p/\partial t|_G$  can be replaced by  $\partial p/\partial t|_{G\text{-equilibrium}}$ . Moreover, the recombination and generation rates must precisely balance under equilibrium conditions,<sup>†</sup> or  $\partial p/\partial t|_G = \partial p/\partial t|_{G\text{-equilibrium}} = -\partial p/\partial t|_{R\text{-equilibrium}}$ . Thus, making use of Eq. (3.29), we obtain

$$\frac{\partial p}{\partial t} \Big|_G = c_p N_T p_0 \quad (3.30)$$

Finally, the net rate of change in the hole concentration due to the R-G center interaction is given by

$$\frac{\partial p}{\partial t} \Big|_{\substack{i\text{-thermal} \\ R-G}} = \frac{\partial p}{\partial t} \Big|_R + \frac{\partial p}{\partial t} \Big|_G = -c_p N_T (p - p_0) \quad (3.31)$$

or

$$\frac{\partial p}{\partial t} \Big|_{\substack{i\text{-thermal} \\ R-G}} = -c_p N_T \Delta p \quad \text{for holes in an } n\text{-type material} \quad (3.32a)$$

An analogous set of arguments yields

$$\frac{\partial n}{\partial t} \Big|_{\substack{i\text{-thermal} \\ R-G}} = -c_n N_T \Delta n \quad \text{for electrons in a } p\text{-type material} \quad (3.32b)$$

$c_n$  like  $c_p$  is a positive proportionality constant.  $c_n$  and  $c_p$  are referred to as the electron and hole *capture coefficients*, respectively.

Although completely functional as is, it is desirable to manipulate Eqs. (3.32) into a slightly more compact and meaningful form. An examination of the left-hand sides of Eqs. (3.32) reveals the dimensional units are those of a concentration divided by time. Since  $\Delta p$  and  $\Delta n$  on the right-hand sides of the same equations are also concentrations, the constants  $c_p N_T$  and  $c_n N_T$  must have units of 1/time. It is therefore reasonable to introduce the time constants

---

<sup>†</sup> Under equilibrium conditions each fundamental process and its inverse must self-balance independent of any other process that may be occurring inside the material. This is known as the *Principle of Detailed Balance*.

$$\tau_p = \frac{1}{c_p N_T} \quad (3.33a)$$

$$\tau_n = \frac{1}{c_n N_T} \quad (3.33b)$$

which, when substituted into Eqs. (3.32), yield

$$\left. \frac{\partial p}{\partial t} \right|_{\substack{i\text{-thermal} \\ R-G}} = - \frac{\Delta p}{\tau_p} \quad \text{for holes in an } n\text{-type material} \quad (3.34a)$$

$$\left. \frac{\partial n}{\partial t} \right|_{\substack{i\text{-thermal} \\ R-G}} = - \frac{\Delta n}{\tau_n} \quad \text{for electrons in a } p\text{-type material} \quad (3.34b)$$

Equations (3.34) are the desired end result, the special-case characterization of R-G center (indirect thermal) recombination-generation. Although steady state or slowly varying conditions are implicitly assumed in the development of Eqs. (3.34), the relationships can be applied with little error to most transient problems of interest. It should be restated that a  $\Delta p < 0$  is possible and will give rise to a  $\partial p / \partial t|_{i\text{-thermal } R-G} > 0$ . A positive  $\partial p / \partial t|_{i\text{-thermal } R-G}$  simply indicates that a carrier deficit exists inside the semiconductor and generation is occurring at a more rapid rate than recombination.  $\partial p / \partial t|_{i\text{-thermal } R-G}$  and  $\partial n / \partial t|_{i\text{-thermal } R-G}$ , it must be remembered, characterize the *net* effect of the thermal recombination and thermal generation processes.

As duly emphasized, the Eq. (3.34) relationships apply only to minority carriers and to situations meeting the low-level injection requirement. The more general steady state result<sup>[7]</sup>, valid for arbitrary injection levels and both carrier types in a nondegenerate semiconductor, is noted for completeness and future reference to be

$$\left. \frac{\partial p}{\partial t} \right|_{\substack{i\text{-thermal} \\ R-G}} = \left. \frac{\partial n}{\partial t} \right|_{\substack{i\text{-thermal} \\ R-G}} = \frac{n_i^2 - np}{\tau_p(n + n_1) + \tau_n(p + p_1)} \quad (3.35)$$

where

$$n_1 \equiv n_i e^{(E_T - E_i)/kT} \quad (3.36a)$$

$$p_1 \equiv n_i e^{(E_i - E_T)/kT} \quad (3.36b)$$

It is left as an exercise to show that the general-case relationship reduces to Eqs. (3.34) under the assumed special-case conditions.

### 3.3.4 Minority Carrier Lifetimes

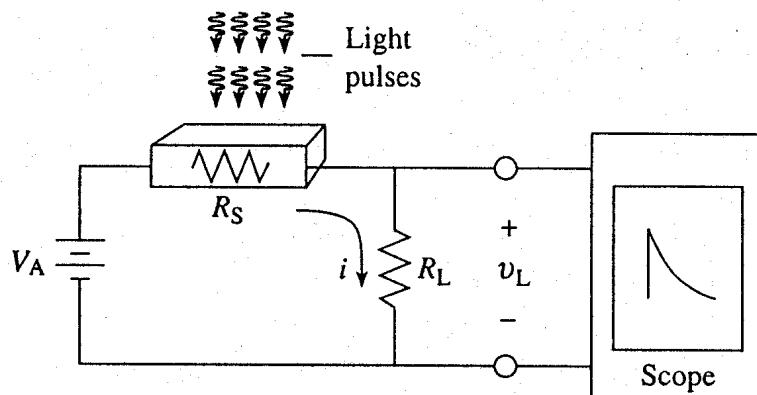
#### General Information

The time constants  $\tau_n$  and  $\tau_p$ , which were introduced without comment in writing down Eqs. (3.34), are obviously the “constants of the action” associated with recombination-generation. Seeking to provide insight relative to the standard interpretation and naming of the  $\tau$ 's, let us consider once again the situation pictured in Fig. 3.21. Examining the change in the hole concentration with time, it goes almost without saying that the excess holes do not all disappear at the same time. Rather, the hole excess present at  $t = 0$  is systematically eliminated, with some of the holes existing for only a short period and others “living” for comparatively long periods of time. If thermal recombination-generation is the sole process acting to relax the semiconductor, the average excess hole lifetime,  $\langle t \rangle$ , can be computed in a relatively straightforward manner. Without going into details, the computation yields  $\langle t \rangle = \tau_n$  (or  $\tau_p$ ). Physically, therefore,  $\tau_n$  and  $\tau_p$  have come to be interpreted as *the average time an excess minority carrier will live in a sea of majority carriers*. For identification purposes,  $\tau_n$  and  $\tau_p$  are simply referred to as the *minority carrier lifetimes*.

Like the  $\mu$ 's and the  $D$ 's, the  $\tau$ 's are important material parameters invariably required in the modeling of devices. Unlike the  $\mu$ 's and the  $D$ 's, there are few plots indicating the  $\tau$ -values to be employed in a given device structure. In fact, subsidiary experimental measurements (see the following subsection) must be performed to determine the minority carrier lifetime in a given semiconductor sample. The reason for the lack of catalogued information can be traced to the extreme variability of the  $\tau_n$  and  $\tau_p$  parameters. Referring to Eqs. (3.33), we see that the carrier lifetimes depend on the often poorly controlled R-G center concentration ( $N_T$ ), *not* on the carefully controlled doping parameters ( $N_A$  and  $N_D$ ). Moreover, the physical nature of the dominant R-G center concentration changes even within a given sample during device fabrication. A fabrication procedure called *gettering* can reduce the R-G center concentration to a very low level and give rise to a  $\tau_n(\tau_p) \sim 1$  msec in Si. The intentional introduction of gold into Si, on the other hand, can controllably increase the R-G center concentration and give rise to a  $\tau_n(\tau_p) \sim 1$  nsec. Minority carrier lifetimes in completed Si devices tend to lie about midway between the cited extremes.

#### A Lifetime Measurement

The carrier lifetimes are measured using a variety of different methods. The majority of measurements employ a device structure of some sort and deduce  $\tau_n$  and/or  $\tau_p$  by matching the theoretical and experimentally observed device characteristics. The measurement to be described employs a bar-like piece of semiconductor with ohmic contacts at its two ends. In a commercial implementation this structure is called a *photoconductor*, a type of photodetector. A situation very similar to that envisioned in Fig. 3.21 is created inside the semi-



**Figure 3.22** Schematic illustration of the photoconductivity decay measurement.

conductor during the measurement. The perturbation is in the form of a light pulse that gives rise to a carrier excess inside the semiconductor and a detectable change in conductivity. Once the light is removed the excess carriers are eliminated via recombination and the conductivity decays back to its value in the dark. The carrier lifetime is determined by monitoring the conductivity decay.

The experiment is schematically illustrated in Fig. 3.22.  $R_S$  is the sample resistance,  $R_L$  a load resistor,  $V_A$  the applied d.c. voltage, and  $v_L$  the measured load or output voltage. Obviously,  $R_S$  and therefore  $v_L$  are functions of time, with  $v_L$  reflecting the changes in conductivity inside the semiconductor sample.

We are specifically interested in how  $v_L$  changes with time after the light is turned off.  $v_L$  is related to the conductivity, which in turn is related to the excess carrier concentrations. Given an  $n$ -type sample and assuming uniform photogeneration throughout the semiconductor, we will verify later in the chapter (Section 3.5) that the decay of the excess minority carrier hole concentration ( $\Delta p$ ) is described by

$$\Delta p(t) = \Delta p_0 e^{-t/\tau_p} \quad (3.37)$$

$\Delta p_0$  is the excess hole concentration at  $t = 0$  when the light is turned off. Next considering the conductivity, we note

$$\sigma = 1/\rho = q(\mu_n n + \mu_p p) \quad (3.38a)$$

$$= q[\mu_n(n_0 + \Delta n) + \mu_p(p_0 + \Delta p)] \quad (3.38b)$$

$$= q\mu_n N_D + q(\mu_n + \mu_p)\Delta p \quad (3.39)$$

$$\begin{array}{c} \boxed{\phantom{0}} \\ \sigma_{\text{dark}} \end{array} \quad \begin{array}{c} \boxed{\phantom{0}} \\ \Delta\sigma(t) \end{array}$$

In establishing Eq. (3.39), we made use of the fact that  $n_0 \equiv N_D \gg p_0$  in an  $n$ -type semiconductor. Also  $\Delta n$  was assumed equal to  $\Delta p$ .  $\sigma_{\text{dark}}$  is the unperturbed sample conductivity in the dark and  $\Delta\sigma(t) \propto \Delta p$  represents the time decaying component of the conductivity.

The last step is to relate  $v_L$  and  $\Delta\sigma$ . Inspecting Fig. 3.22, one concludes

$$v_L = iR_L = V_A \frac{R_L}{R_L + R_S} \quad (3.40)$$

but

$$R_S = \frac{\rho l}{A} = \frac{l}{\sigma A} = \frac{l}{(\sigma_{\text{dark}} + \Delta\sigma)A} = R_{Sd} \left( \frac{\sigma_{\text{dark}}}{\sigma_{\text{dark}} + \Delta\sigma} \right) \quad (3.41)$$

where  $l$  is the length,  $A$  the cross-sectional area, and  $R_{Sd} = l/\sigma_{\text{dark}}A$  the dark resistance of the sample, respectively. Substituting Eq. (3.41) into Eq. (3.40) and rearranging yields

$$v_L = \frac{V_A}{1 + \frac{R_{Sd}/R_L}{1 + \Delta\sigma/\sigma_{\text{dark}}}} \quad (3.42)$$

or, if  $R_L$  is matched to  $R_{Sd}$  and the light intensity restricted so  $\Delta\sigma/\sigma_{\text{dark}} \ll 1$  (equivalent to specifying low-level injection conditions), we obtain

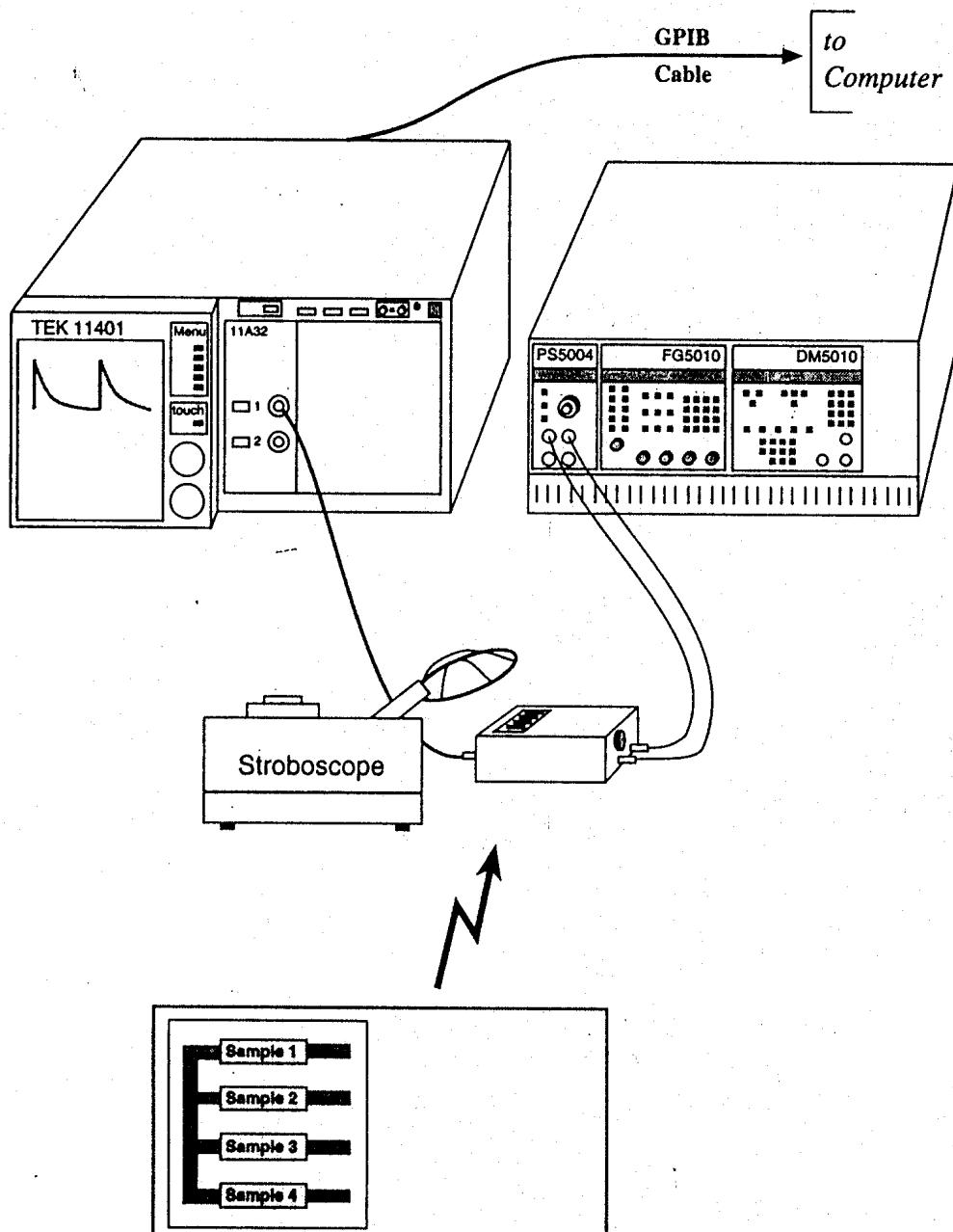
$$v_L \cong \frac{V_A}{2} \left( 1 + \frac{\Delta\sigma}{\sigma_{\text{dark}}} \right) = V_L + v_0 e^{-t/\tau_p} \quad (3.43)$$

□      □  
d.c.      decay transient

An exponential decay is predicted with  $\tau_p$  being the decay constant.

An actual implementation of the measurement is shown in Fig. 3.23.<sup>†</sup> The required  $V_A$  is supplied by the Tektronix PS5004 power supply (just about any power supply would do), while the Tektronix 11401 Digitizing Oscilloscope is used to capture the photoconductive decay. The bar-shaped Si test structures are housed in a special measurements box along with the  $R_L \cong R_{Sd}$  load resistor. The box permits convenient handling of the samples and facilitates switching between samples. The light pulses required in the measurement are derived from the General Radio 1531-AB Stroboscope. The GR Stroboscope puts out high-intensity light flashes of  $\sim 1 \mu\text{sec}$  duration spaced at controllable time intervals ranging from 2.5 msec to 0.5 sec. Fairly uniform photogeneration can be achieved in a sample

<sup>†</sup> This being the first text reference, it is perhaps worthwhile reiterating the General Introduction statement that measurement particulars, included at various points in the text, are derived from experiments performed in an undergraduate EE laboratory administered by the author. A somewhat expanded discussion of background theory, the measurement system, and measurement procedures can be found in R. F. Pierret, *Semiconductor Measurements Laboratory Operations Manual*, Supplement A, in the Modular Series on Solid State Devices, Addison-Wesley Publishing Co., Reading MA, © 1991.



**Figure 3.23** Photoconductive decay measurement system.

by first passing the stroboscope light through a small-diameter Si wafer. The Si wafer acts as a filter allowing only near-bandgap radiation to reach the sample under test. Although the 11401 Oscilloscope is capable of *in situ* data analysis, a computer is connected to the oscilloscope to add data manipulation and display flexibility.

An example of the photoconductive decay as displayed on the computer screen is reproduced in Fig. 3.24. The transient component of  $v_L$  in millivolts is plotted on the y-axis versus time in milliseconds on the x-axis. The part (a) linear-scale plot nicely exhibits the generally expected dependence. The part (b) semilog plot verifies the simple exponential nature of the decay (at least for times  $0.1 \text{ msec} \leq t \leq 0.6 \text{ msec}$ ) and is more useful for

The last step is to relate  $v_L$  and  $\Delta\sigma$ . Inspecting Fig. 3.22, one concludes

$$v_L = iR_L = V_A \frac{R_L}{R_L + R_S} \quad (3.40)$$

but

$$R_S = \frac{\rho l}{A} = \frac{l}{\sigma A} = \frac{l}{(\sigma_{\text{dark}} + \Delta\sigma)A} = R_{Sd} \left( \frac{\sigma_{\text{dark}}}{\sigma_{\text{dark}} + \Delta\sigma} \right) \quad (3.41)$$

where  $l$  is the length,  $A$  the cross-sectional area, and  $R_{Sd} = l/\sigma_{\text{dark}}A$  the dark resistance of the sample, respectively. Substituting Eq. (3.41) into Eq. (3.40) and rearranging yields

$$v_L = \frac{V_A}{1 + \frac{R_{Sd}/R_L}{1 + \Delta\sigma/\sigma_{\text{dark}}}} \quad (3.42)$$

or, if  $R_L$  is matched to  $R_{Sd}$  and the light intensity restricted so  $\Delta\sigma/\sigma_{\text{dark}} \ll 1$  (equivalent to specifying low-level injection conditions), we obtain

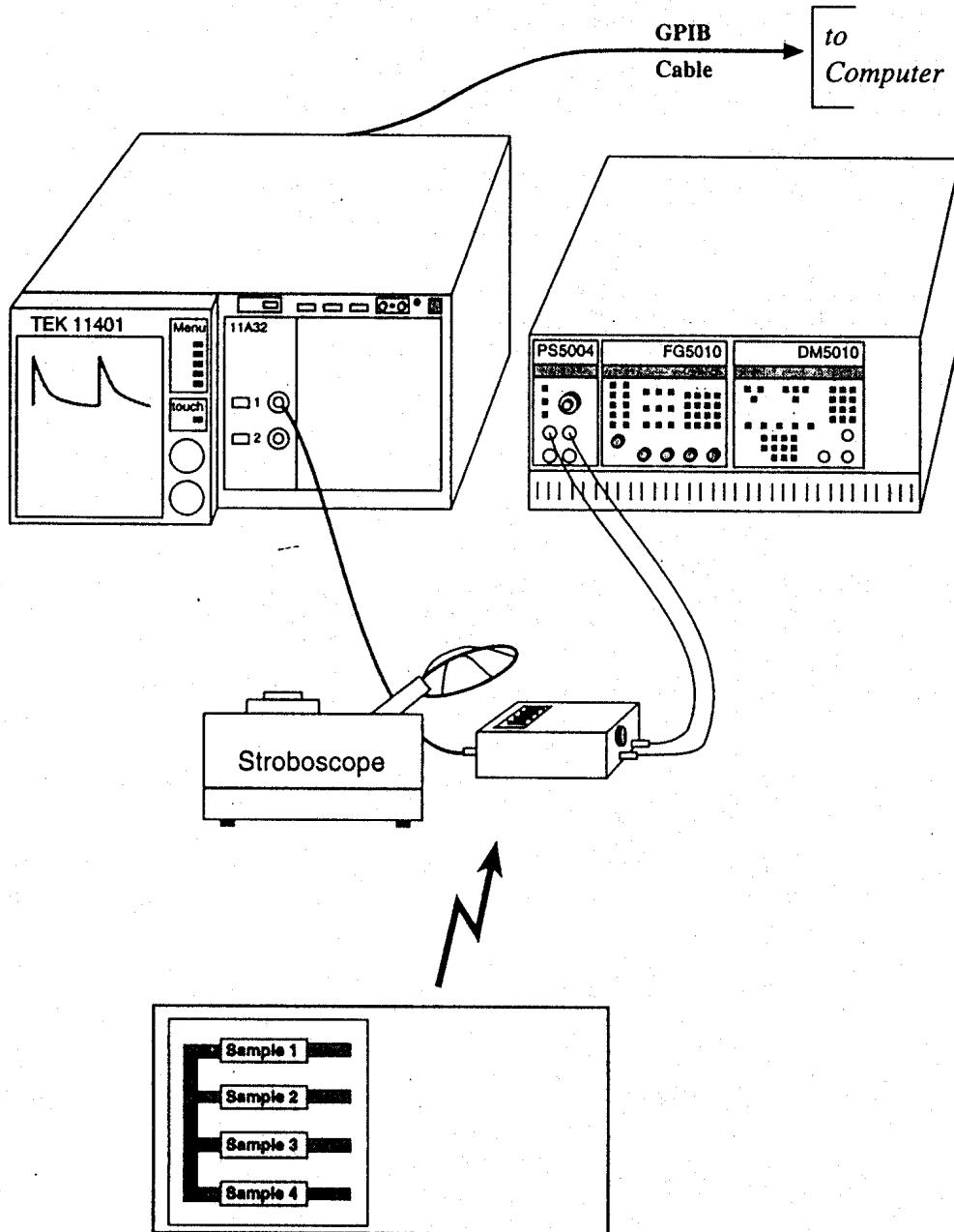
$$v_L \cong \frac{V_A}{2} \left( 1 + \frac{\Delta\sigma}{\sigma_{\text{dark}}} \right) = V_L + v_0 e^{-t/\tau_p} \quad (3.43)$$

d.c.
decay transient

An exponential decay is predicted with  $\tau_p$  being the decay constant.

An actual implementation of the measurement is shown in Fig. 3.23.<sup>†</sup> The required  $V_A$  is supplied by the Tektronix PS5004 power supply (just about any power supply would do), while the Tektronix 11401 Digitizing Oscilloscope is used to capture the photoconductive decay. The bar-shaped Si test structures are housed in a special measurements box along with the  $R_L \cong R_{Sd}$  load resistor. The box permits convenient handling of the samples and facilitates switching between samples. The light pulses required in the measurement are derived from the General Radio 1531-AB Stroboscope. The GR Stroboscope puts out high-intensity light flashes of  $\sim 1 \mu\text{sec}$  duration spaced at controllable time intervals ranging from 2.5 msec to 0.5 sec. Fairly uniform photogeneration can be achieved in a sample

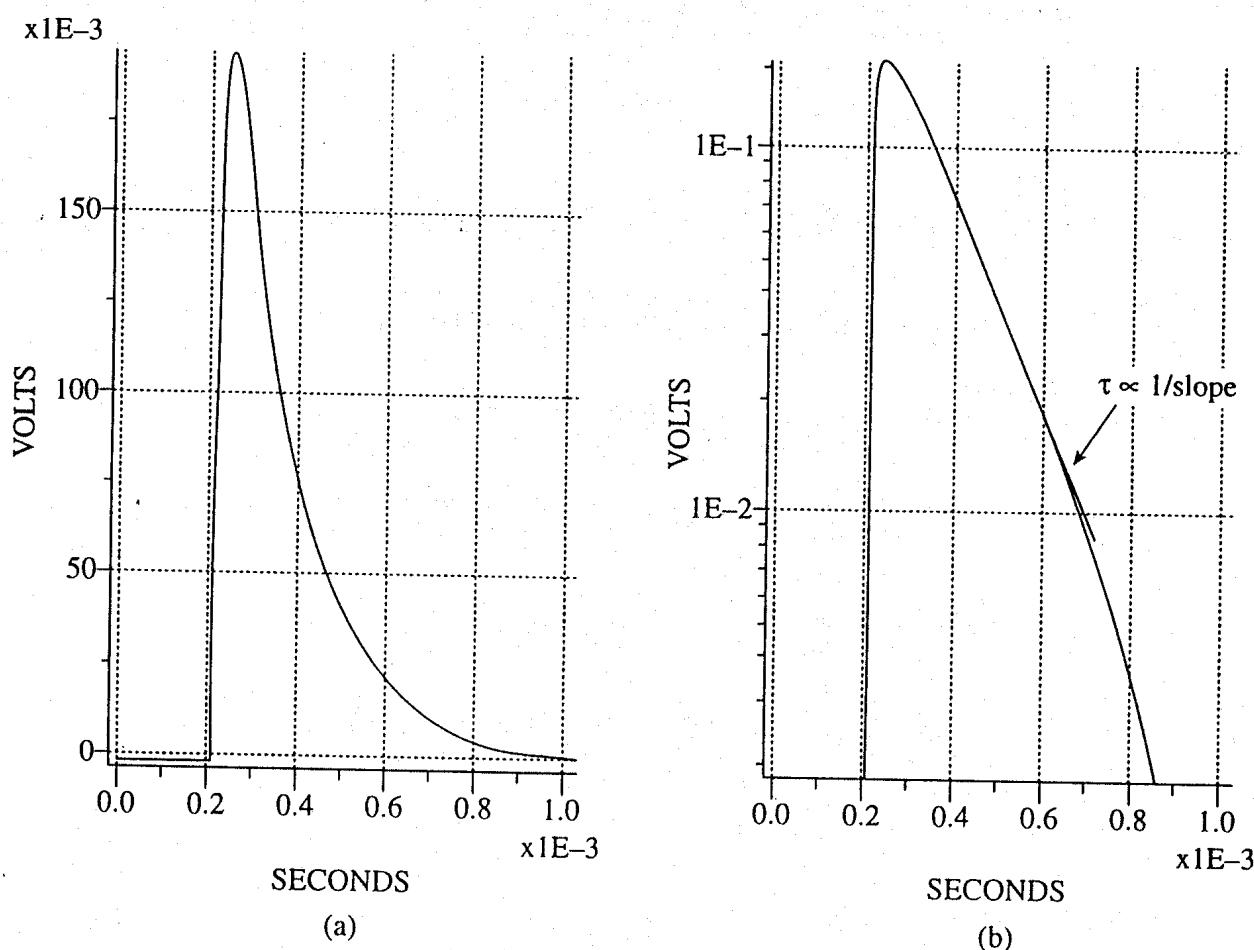
<sup>†</sup> This being the first text reference, it is perhaps worthwhile reiterating the General Introduction statement that measurement particulars, included at various points in the text, are derived from experiments performed in an undergraduate EE laboratory administered by the author. A somewhat expanded discussion of background theory, the measurement system, and measurement procedures can be found in R. F. Pierret, *Semiconductor Measurements Laboratory Operations Manual*, Supplement A, in the Modular Series on Solid State Devices, Addison-Wesley Publishing Co., Reading MA, © 1991.



**Figure 3.23** Photoconductive decay measurement system.

by first passing the stroboscope light through a small-diameter Si wafer. The Si wafer acts as a filter allowing only near-bandgap radiation to reach the sample under test. Although the 11401 Oscilloscope is capable of *in situ* data analysis, a computer is connected to the oscilloscope to add data manipulation and display flexibility.

An example of the photoconductive decay as displayed on the computer screen is reproduced in Fig. 3.24. The transient component of  $v_L$  in millivolts is plotted on the y-axis versus time in milliseconds on the x-axis. The part (a) linear-scale plot nicely exhibits the generally expected dependence. The part (b) semilog plot verifies the simple exponential nature of the decay (at least for times  $0.1 \text{ msec} \leq t \leq 0.6 \text{ msec}$ ) and is more useful for



**Figure 3.24** Photoconductivity transient response. The time-varying component of  $v_L$  in millivolts is plotted on the  $y$ -axis versus time in milliseconds on the  $x$ -axis. (a) is a linear and (b) is a semilog plot.

extracting the minority carrier lifetime. Specifically, from the slope of the straight-line region in the latter plot, the minority carrier lifetime in the probed sample is readily deduced to be  $\tau_p \equiv 150 \mu\text{sec}$ .

### 3.4 EQUATIONS OF STATE

In the first three sections of this chapter we examined and separately modeled the primary types of carrier action taking place inside semiconductors. Within actual semiconductors all the various types of carrier action occur at the same time, and the state of any semiconductor can be determined only by taking into account the combined effect of the individual types of carrier action. "Putting it all together," so to speak, leads to the basic set of starting equations employed in solving device problems, herein referred to as the *equations of state*. The first portion of this section is devoted to developing the equations of state; the latter portion includes a summary of common simplifications, a listing of special-case solutions, and problem-solving examples.

### 3.4.1 Continuity Equations

Each and every type of carrier action—whether it be drift, diffusion, indirect or direct thermal recombination, indirect or direct generation, or some other type of carrier action—gives rise to a change in the carrier concentrations with time. The combined effect of all types of carrier action can therefore be taken into account by equating the overall change in the carrier concentrations per unit time ( $\partial n/\partial t$  or  $\partial p/\partial t$ ) to the sum of the  $\partial n/\partial t$ 's or  $\partial p/\partial t$ 's due to the individual processes; that is,

$$\frac{\partial n}{\partial t} = \left. \frac{\partial n}{\partial t} \right|_{\text{drift}} + \left. \frac{\partial n}{\partial t} \right|_{\text{diff}} + \left. \frac{\partial n}{\partial t} \right|_{\substack{\text{thermal} \\ \text{R-G}}} + \left. \frac{\partial n}{\partial t} \right|_{\substack{\text{other processes} \\ (\text{light, etc.})}} \quad (3.44a)$$

$$\frac{\partial p}{\partial t} = \left. \frac{\partial p}{\partial t} \right|_{\text{drift}} + \left. \frac{\partial p}{\partial t} \right|_{\text{diff}} + \left. \frac{\partial p}{\partial t} \right|_{\substack{\text{thermal} \\ \text{R-G}}} + \left. \frac{\partial p}{\partial t} \right|_{\substack{\text{other processes} \\ (\text{light, etc.})}} \quad (3.44b)$$

The overall effect of the individual processes is established, in essence, by invoking the requirement of conservation of carriers. Electrons and holes cannot mysteriously appear and disappear at a given point, but must be transported to or created at the given point via some type of ongoing carrier action. There must be a spatial and time continuity in the carrier concentrations. For this reason, Eqs. (3.44) are known as the *continuity equations*.

The continuity equations can be written in a somewhat more compact form by noting

$$\left. \frac{\partial n}{\partial t} \right|_{\text{drift}} + \left. \frac{\partial n}{\partial t} \right|_{\text{diff}} = \frac{1}{q} \left( \frac{\partial J_{Nx}}{\partial x} + \frac{\partial J_{Ny}}{\partial y} + \frac{\partial J_{Nz}}{\partial z} \right) = \frac{1}{q} \nabla \cdot \mathbf{J}_N \quad (3.45a)$$

$$\left. \frac{\partial p}{\partial t} \right|_{\text{drift}} + \left. \frac{\partial p}{\partial t} \right|_{\text{diff}} = -\frac{1}{q} \left( \frac{\partial J_{Px}}{\partial x} + \frac{\partial J_{Py}}{\partial y} + \frac{\partial J_{Pz}}{\partial z} \right) = -\frac{1}{q} \nabla \cdot \mathbf{J}_P \quad (3.45b)$$

Equations (3.45), which can be established by a straightforward mathematical manipulation, merely state that there will be a change in the carrier concentrations within a given small region of the semiconductor if an imbalance exists between the total carrier currents into and out of the region. Utilizing Eqs. (3.45), we obtain

$$\frac{\partial n}{\partial t} = \frac{1}{q} \nabla \cdot \mathbf{J}_N + \left. \frac{\partial n}{\partial t} \right|_{\substack{\text{thermal} \\ \text{R-G}}} + \left. \frac{\partial n}{\partial t} \right|_{\text{other processes}} \quad (3.46a)$$

$$\frac{\partial p}{\partial t} = -\frac{1}{q} \nabla \cdot \mathbf{J}_P + \left. \frac{\partial p}{\partial t} \right|_{\substack{\text{thermal} \\ \text{R-G}}} + \left. \frac{\partial p}{\partial t} \right|_{\text{other processes}} \quad (3.46b)$$

The (3.46) continuity equations are completely general and directly or indirectly constitute the starting point in most device analyses. In computer simulations the continuity

equations are often employed directly. The appropriate relationships for  $\partial n / \partial t|_{\text{thermal R-G}}$ ,  $\partial p / \partial t|_{\text{thermal R-G}}$  [which may or may not be the special-case relationships given by Eqs. (3.34)], along with the concentration changes due to "other processes," are substituted into Eqs. (3.46), and numerical solutions are sought for  $n(x, y, z, t)$  and  $p(x, y, z, t)$ . In problems where a closed-form type of solution is desired, the continuity equations are typically utilized only in an indirect fashion. The actual starting point in such analyses is a simplified version of the continuity equations to be established in the next subsection.

### 3.4.2 Minority Carrier Diffusion Equations

The "workhorse" minority carrier diffusion equations are derived from the continuity equations by invoking the following set of simplifying assumptions:

- (1) The particular system under analysis is *one-dimensional*; i.e., all variables are at most a function of just one coordinate (say the  $x$ -coordinate).
- (2) The analysis is limited or restricted to *minority carriers*.
- (3)  $\mathcal{E} \approx 0$  in the semiconductor or regions of the semiconductor subject to analysis.
- (4) The equilibrium minority carrier concentrations are not a function of position. In other words,  $n_0 \neq n_0(x)$ ,  $p_0 \neq p_0(x)$ .
- (5) *Low-level injection* conditions prevail.
- (6) *Indirect* thermal recombination-generation is the dominant thermal R-G mechanism.
- (7) There are *no* "other processes," except possibly photogeneration, taking place within the system.

Working on the continuity equation for electrons, we note that

$$\frac{1}{q} \nabla \cdot \mathbf{J}_N \rightarrow \frac{1}{q} \frac{\partial J_N}{\partial x} \quad (3.47)$$

if the system is one-dimensional. Moreover,

$$J_N = q\mu_n n \mathcal{E} + qD_N \frac{\partial n}{\partial x} \approx qD_N \frac{\partial n}{\partial x} \quad (3.48)$$

when  $\mathcal{E} \approx 0$  and one is concerned only with minority carriers [simplifications (2) and (3)]. By way of explanation, the drift component can be neglected in the current density expression because  $\mathcal{E}$  is small by assumption and minority carrier concentrations are also small, making the  $n\mathcal{E}$  product extremely small. (Note that the same argument cannot be applied to majority carriers.) Since by assumption  $n_0 \neq n_0(x)$ , and by definition  $n = n_0 + \Delta n$ , we can also write

$$\frac{\partial n}{\partial x} = \frac{\partial n_0}{\partial x} + \frac{\partial \Delta n}{\partial x} = \frac{\partial \Delta n}{\partial x} \quad (3.49)$$

Combining Eqs. (3.47) through (3.49) yields

$$\frac{1}{q} \nabla \cdot \mathbf{J}_N \rightarrow D_N \frac{\partial^2 \Delta n}{\partial x^2} \quad (3.50)$$

Turning to the remaining terms in the continuity equation for electrons, the assumed dominance of recombination-generation via R-G centers, combined with the low-level injection and minority carrier restrictions, allows us to replace the thermal R-G term with the Eq. (3.34) special-case expression.

$$\left. \frac{\partial n}{\partial t} \right|_{R-G}^{\text{thermal}} = - \frac{\Delta n}{\tau_n} \quad (3.51)$$

In addition, applying simplification (7) yields

$$\left. \frac{\partial n}{\partial t} \right|_{\substack{\text{other} \\ \text{processes}}} = G_L \quad (3.52)$$

where it is understood that  $G_L = 0$  if the semiconductor is not subject to illumination. Finally, the equilibrium electron concentration is never a function of time,  $n_0 \neq n_0(t)$ , and we can therefore write

$$\frac{\partial n}{\partial t} = \frac{\partial n_0}{\partial t} + \frac{\partial \Delta n}{\partial t} = \frac{\partial \Delta n}{\partial t} \quad (3.53)$$

Substituting Eqs. (3.50) through (3.53) into the (3.46a) continuity equation, and simultaneously recording the analogous result for holes, one obtains

$$\frac{\partial \Delta n_p}{\partial t} = D_N \frac{\partial^2 \Delta n_p}{\partial x^2} - \frac{\Delta n_p}{\tau_n} + G_L \quad (3.54a)$$

Minority carrier diffusion equations (3.54b)

$$\frac{\partial \Delta p_n}{\partial t} = D_P \frac{\partial^2 \Delta p_n}{\partial x^2} - \frac{\Delta p_n}{\tau_p} + G_L \quad (3.54b)$$

Subscripts are added to the carrier concentrations in Eqs. (3.54) to remind the user that the equations are valid only for minority carriers, applying to electrons in *p*-type materials and to holes in *n*-type materials.

**Table 3.1** Common Diffusion Equation Simplifications.

<i>Simplification</i>	<i>Effect</i>
Steady state	$\frac{\partial \Delta n_p}{\partial t} \rightarrow 0 \quad \left( \frac{\partial \Delta p_n}{\partial t} \rightarrow 0 \right)$
No concentration gradient or no diffusion current	$D_N \frac{\partial^2 \Delta n_p}{\partial x^2} \rightarrow 0 \quad \left( D_p \frac{\partial^2 \Delta p_n}{\partial x^2} \rightarrow 0 \right)$
No drift current or $\mathcal{E} = 0$	No further simplification. ( $\mathcal{E} \approx 0$ is assumed in the derivation.)
No thermal R-G	$\frac{\Delta n_p}{\tau_n} \rightarrow 0 \quad \left( \frac{\Delta p_n}{\tau_p} \rightarrow 0 \right)$
No light	$G_L \rightarrow 0$

### 3.4.3 Simplifications and Solutions

In the course of performing device analyses, the conditions of a problem often permit additional simplifications that drastically reduce the complexity of the minority carrier diffusion equations. Common simplifications and their effect on the minority carrier diffusion equations are summarized in Table 3.1. Extensively utilized solutions to simplified forms of the minority carrier diffusion equations are collected for future reference in Table 3.2.

### 3.4.4 Problem Solving

Although we have presented all necessary information, it may not be completely obvious how one works with the diffusion equations to obtain the carrier concentration solutions. Device analyses of course provide examples of "problem" solutions, and a number of device analyses are presented later in the text. Nevertheless, it is worthwhile to examine a few well-defined simple problems to illustrate problem-solving procedures. The chosen sample problems also provide a basis for the two supplemental concepts introduced in the next section.

#### Sample Problem No. 1

**P:** A uniformly donor-doped silicon wafer maintained at room temperature is suddenly illuminated with light at time  $t = 0$ . Assuming  $N_D = 10^{15}/\text{cm}^3$ ,  $\tau_p = 10^{-6}$  sec, and a light-induced creation of  $10^{17}$  electrons and holes per  $\text{cm}^3\text{-sec}$  throughout the semiconductor, determine  $\Delta p_n(t)$  for  $t > 0$ .

**S: Step 1**—Review precisely what information is given or implied in the statement of the problem.

**Table 3.2** Common Special-Case Diffusion Equation Solutions.**Solution no. 1**

*GIVEN:* Steady state, no light.

**SIMPLIFIED**

$$\text{DIFF. EQN: } 0 = D_N \frac{d^2 \Delta n_p}{dx^2} - \frac{\Delta n_p}{\tau_n}$$

$$\text{SOLUTION: } \Delta n_p(x) = Ae^{-x/L_N} + Be^{x/L_N}$$

$$\text{where } L_N \equiv \sqrt{D_n \tau_n}$$

and  $A, B$  are solution constants.

**Solution no. 2**

*GIVEN:* No concentration gradient, no light.

**SIMPLIFIED**

$$\text{DIFF. EQN: } \frac{d\Delta n_p}{dt} = -\frac{\Delta n_p}{\tau_n}$$

$$\text{SOLUTION: } \Delta n_p(t) = \Delta n_p(0)e^{-t/\tau_n}$$

**Solution no. 3**

*GIVEN:* Steady state, no concentration gradient.

**SIMPLIFIED**

$$\text{DIFF. EQN: } 0 = -\frac{\Delta n_p}{\tau_n} + G_L$$

$$\text{SOLUTION: } \Delta n_p = G_L \tau_n$$

**Solution no. 4**

*GIVEN:* Steady state, no R-G, no light.

**SIMPLIFIED**

$$\text{DIFF. EQN: } 0 = D_N \frac{d^2 \Delta n_p}{dx^2} \quad \text{or} \quad 0 = \frac{d^2 \Delta n_p}{dx^2}$$

$$\text{SOLUTION: } \Delta n_p(x) = A + Bx$$

The semiconductor is silicon,  $T = 300$  K, the donor doping is the same everywhere with  $N_D = 10^{15}/\text{cm}^3$ , and  $G_L = 10^{17}/\text{cm}^3\text{-sec}$  at all points inside the semiconductor. Also, the statement of the problem *implies* equilibrium conditions exist for  $t < 0$ .

**Step 2**—Characterize the system under equilibrium conditions.

In Si at room temperature  $n_i = 10^{10}/\text{cm}^3$ . Since  $N_D \gg n_i$ ,  $n_0 = N_D = 10^{15}/\text{cm}^3$  and  $p_0 = n_i^2/N_D = 10^5/\text{cm}^3$ . With the doping being uniform, the equilibrium  $n_0$  and  $p_0$  values are the same everywhere throughout the semiconductor.

**Step 3**—Analyze the problem qualitatively.

Prior to  $t = 0$ , equilibrium conditions prevail and  $\Delta p_n = 0$ . Starting at  $t = 0$  the light creates added electrons and holes and  $\Delta p_n$  will begin to increase. The growing excess

carrier numbers, however, in turn lead to an increased indirect thermal recombination rate which is proportional to  $\Delta p_n$ . Consequently, as  $\Delta p_n$  grows as a result of photogeneration, more and more of the excess holes are eliminated per second by recombination through R-G centers. Eventually, a point is reached where the carriers annihilated per second by indirect thermal recombination balance the carriers created per second by the light, and a steady state condition is attained.

Summarizing, we expect  $\Delta p_n(t)$  to start from zero at  $t = 0$ , to build up at a decreasing rate, and to ultimately become constant. Since the light generation and thermal recombination rates must balance under steady state conditions, we can even state  $G_L = \Delta p_n(t \rightarrow \infty)/\tau_p$  or  $\Delta p_n(t \rightarrow \infty) = \Delta p_{n\max} = G_L \tau_p$  if low-level injection prevails.

*Step 4*—Perform a quantitative analysis.

The minority carrier diffusion equation is the starting point for most first-order quantitative analyses. After examining the problem for obvious conditions that would invalidate the use of the diffusion equation, the appropriate minority carrier diffusion equation is written down, the equation is simplified, and a solution is sought subject to boundary conditions stated or implied in the problem.

For the problem under consideration a cursory inspection reveals that all simplifying assumptions involved in deriving the diffusion equations are readily satisfied. Specifically, only the minority carrier concentration is of interest; the equilibrium carrier concentrations are not a function of position, indirect thermal R-G is dominant in Si, and there are no "other processes" except for photogeneration. Because the photogeneration is uniform throughout the semiconductor, the perturbed carrier concentrations are also position-independent and the electric field  $\mathcal{E}$  must clearly be zero in the perturbed system. Finally, a  $\Delta p_{n\max} = G_L \tau_p = 10^{11}/\text{cm}^3 \ll n_0 = 10^{15}/\text{cm}^3$  is consistent with low-level injection prevailing at all times.

With no obstacles to utilizing the diffusion equation, the desired quantitative solution can now be obtained by solving

$$\frac{\partial \Delta p_n}{\partial t} = D_p \frac{\partial^2 \Delta p_n}{\partial x^2} - \frac{\Delta p_n}{\tau_p} + G_L \quad (3.55)$$

subject to the boundary condition

$$\Delta p_n(t)|_{t=0} = 0 \quad (3.56)$$

Since  $\Delta p_n$  is not a function of position, the diffusion equation becomes an ordinary differential equation and simplifies to

$$\frac{d\Delta p_n}{dt} + \frac{\Delta p_n}{\tau_p} = G_L \quad (3.57)$$

The general solution of Eq. (3.57) is

$$\Delta p_n(t) = G_L \tau_p + A e^{-t/\tau_p} \quad (3.58)$$

Applying the boundary condition yields

$$A = -G_L \tau_p \quad (3.59)$$

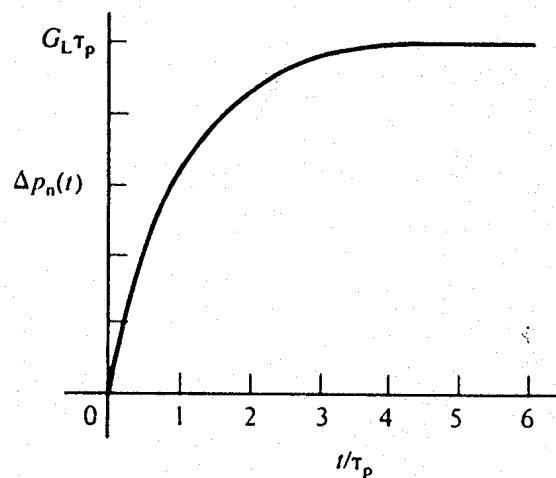
and

$$\Delta p_n(t) = G_L \tau_p (1 - e^{-t/\tau_p}) \quad \Leftarrow \text{solution} \quad (3.60)$$

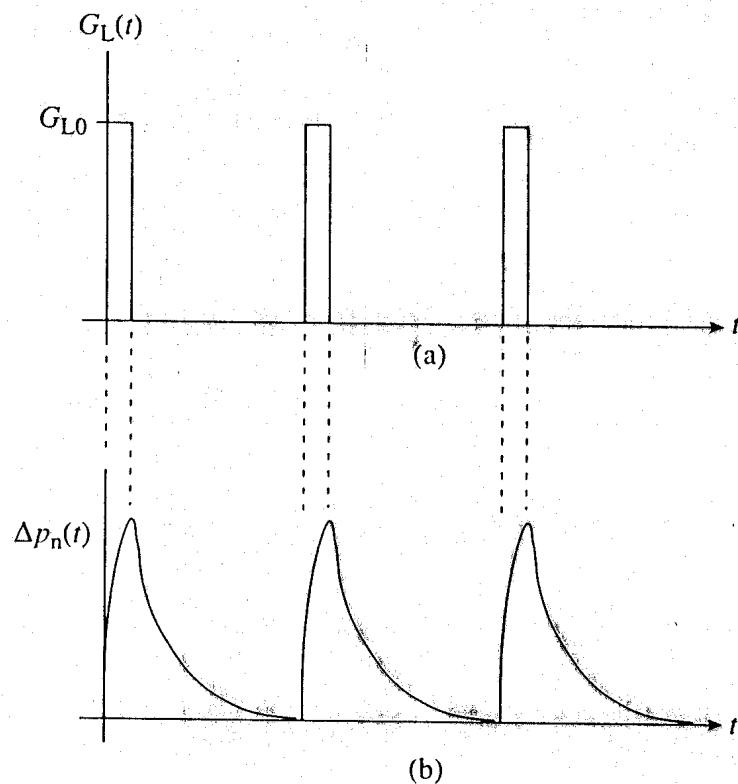
**Step 5—Examine the solution.**

Failing to examine the mathematical solution to a problem is like growing vegetables and then failing to eat the produce. Relative to the Eq. (3.60) result,  $G_L \tau_p$  has the dimensions of a concentration (number/cm<sup>3</sup>) and the solution is at least dimensionally correct. A plot of the Eq. (3.60) result is shown in Fig. 3.25. Note that, in agreement with qualitative predictions,  $\Delta p_n(t)$  starts from zero at  $t = 0$  and eventually saturates at  $G_L \tau_p$  after a few  $\tau_p$ .

**Epilogue**—We would be remiss if we did not point out the connection between the hypothetical problem just completed and the photoconductive decay measurement described in Subsection 3.3.4. Light output from the stroboscope used in the measurement can be modeled to first order by the pulse train pictured in Fig. 3.26(a). The Eq. (3.60) solution therefore approximately describes the carrier build-up during a light pulse. It should be noted, however, that the stroboscope light pulses have a duration of  $t_{\text{on}} \approx 1 \mu\text{sec}$  compared to a minority carrier lifetime of  $\tau_p \approx 150 \mu\text{sec}$ . With  $t_{\text{on}}/\tau_p \ll 1$ , one sees only the very



**Figure 3.25** Solution to Sample Problem No. 1. Photogeneration-induced increase in the excess hole concentration as a function of time.



**Figure 3.26** (a) Approximate model for the light output from the stroboscope in the photoconductivity decay measurement. (b) Sketch of the combined light-on/light-off solution for the excess minority-carrier concentration.

first portion of the Fig. 3.25 transient before the light is turned off and the semiconductor is allowed to decay back to equilibrium.

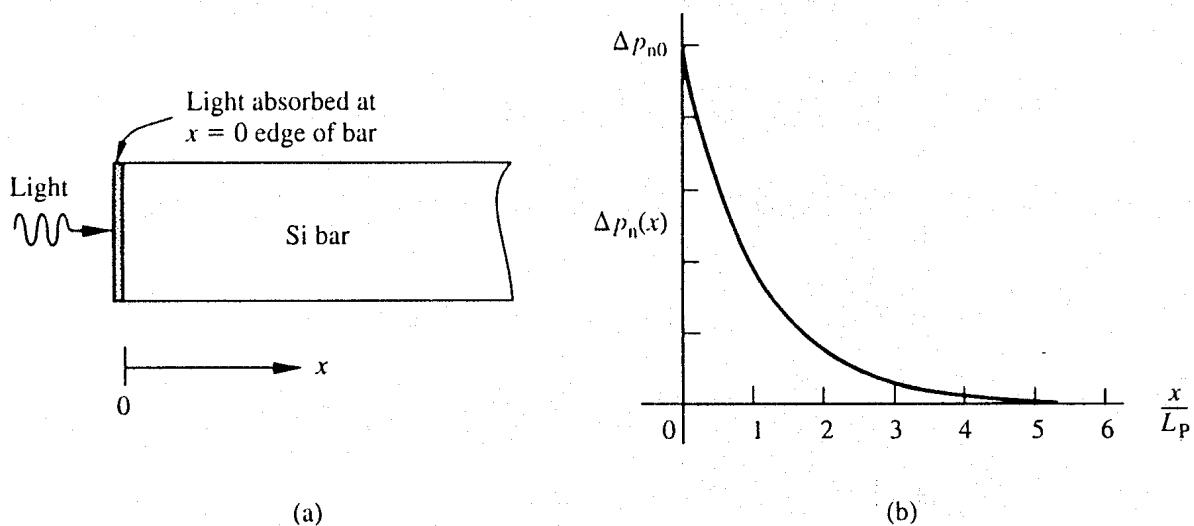
We are also interested in verifying the correctness of the light-off  $\Delta p(t)$  expression used in the photoconductivity decay analysis. Derivation of this expression closely parallels the light-on sample problem solution. Exceptions are  $\Delta p_n$  at the beginning of the light-off transient equals  $\Delta p_n$  at the end of the light-on transient and  $G_L$  is set equal to zero in Eqs. (3.55) and (3.57). The solution to Eq. (3.57) with  $G_L = 0$  is the same as Solution No. 2 in Table 3.2; i.e.,

$$\Delta p_n(t) = \Delta p_n(0)e^{-t/\tau_p} \quad (3.61)$$

where  $t = 0$  has been reset to correspond to the beginning of the light-off transient. Equation (3.37) in the measurement analysis and Eq. (3.61) are of course identical. For completeness, the combined light-on and light-off  $\Delta p_n(t)$  solution is sketched in Fig. 3.26(b).

### Sample Problem No. 2

**P:** As pictured in Fig. 3.27(a), the  $x = 0$  end of a uniformly doped semi-infinite bar of silicon with  $N_D = 10^{15}/\text{cm}^3$  is illuminated so as to create  $\Delta p_{n0} = 10^{10}/\text{cm}^3$  excess holes at



**Figure 3.27** (a) Pictorial definition of Sample Problem No. 2. (b) Solution to Sample Problem No. 2 showing the excess hole concentration inside the Si bar as a function of position.

$x = 0$ . The wavelength of the illumination is such that no light penetrates into the interior ( $x > 0$ ) of the bar. Determine  $\Delta p_n(x)$ .

**S:** The semiconductor is again silicon uniformly doped with an  $N_D = 10^{15}/\text{cm}^3$ . Steady state conditions are inferred from the statement of the problem, since we are asked for  $\Delta p_n(x)$  and not  $\Delta p_n(x, t)$ . Moreover, at  $x = 0$ ,  $\Delta p_n(0) = \Delta p_{n0} = 10^{10}/\text{cm}^3$ , and  $\Delta p_n \rightarrow 0$  as  $x \rightarrow \infty$ . The latter boundary condition follows from the semi-infinite nature of the bar. The perturbation in  $\Delta p_n$  due to the nonpenetrating light can't possibly extend out to  $x = \infty$ . The nonpenetrating nature of the light allows us to set  $G_L = 0$  for  $x > 0$ . Finally, note that the problem statement fails to mention the temperature of operation. When this happens, it is reasonable to assume an intended  $T = 300$  K.

If the light were removed, the silicon bar in Sample Problem 2 maintained at 300 K would revert to an equilibrium condition identical to that described in Sample Problem 1. Under equilibrium conditions, then,  $n_0 = 10^{15}/\text{cm}^3$ ,  $p_0 = 10^{10}/\text{cm}^3$ , and the carrier concentrations are uniform throughout the semiconductor bar.

Qualitatively it is a simple matter to predict the expected effect of the nonpenetrating light on the silicon bar. The light first creates excess carriers right at  $x = 0$ . With more carriers at one point than elsewhere in the bar, the diffusion process next comes into play and the carrier excess spreads into the semiconductor proper. At the same time, however, the appearance of an excess hole concentration inside the bar enhances the thermal recombination rate. Thus, as the diffusing holes move into the bar their numbers are reduced by recombination. In addition, since the minority carrier holes live for only a limited period, a time  $\tau_p$  on the average, fewer and fewer excess holes will survive as the depth of penetration into the bar becomes larger and larger. Under steady state conditions it is therefore reasonable to expect an excess distribution of holes near  $x = 0$ , with  $\Delta p_n(x)$  monotonically decreasing from  $\Delta p_{n0}$  at  $x = 0$  to  $\Delta p_n = 0$  as  $x \rightarrow \infty$ .

In preparation for obtaining a quantitative solution, we observe that the system under consideration is one-dimensional, the analysis is restricted to the minority carrier holes, the equilibrium carrier concentrations are position-independent, indirect thermal R-G dominates, there are no "other processes" for  $x > 0$ , and low-level injection conditions clearly prevail ( $\Delta p_{n\text{max}} = \Delta p_{n0} = 10^{10}/\text{cm}^3 \ll n_0 = 10^{15}/\text{cm}^3$ ). The only question that might be raised concerning the use of the diffusion equation as the starting point for the quantitative analysis is whether  $\mathcal{E} \approx 0$ . With the light on, a nonuniform distribution of holes and associated distribution of positive charge will appear near the  $x = 0$  surface. The excess hole pile-up, however, is very small ( $\Delta p_{n\text{max}} \approx n_i$ ) and the associated electric field is therefore expected to be correspondingly small. Moreover, in problems of this type it is found that the majority carriers, negatively charged electrons in the given problem, redistribute in such a way as to partly cancel the minority carrier charge. Thus experience indicates the  $\mathcal{E} \approx 0$  assumption to be reasonable, and use of the minority carrier diffusion equation to be justified.

Under steady state conditions with  $G_L = 0$  for  $x > 0$  the hole diffusion equation reduces to the form

$$D_p \frac{d^2 \Delta p_n}{dx^2} - \frac{\Delta p_n}{\tau_p} = 0 \quad \text{for } x > 0 \quad (3.62)$$

which is to be solved subject to the boundary conditions

$$\Delta p_{n|x=0^+} = \Delta p_{n|x=0} = \Delta p_{n0} \quad (3.63)$$

and

$$\Delta p_{n|x \rightarrow \infty} = 0 \quad (3.64)$$

Equation (3.62) should be recognized as one of the simplified diffusion equations cited in Table 3.2, with the general solution

$$\Delta p_n(x) = A e^{-x/L_p} + B e^{x/L_p} \quad (3.65)$$

where

$$L_p \equiv \sqrt{D_p \tau_p} \quad (3.66)$$

Because  $\exp(x/L_p) \rightarrow \infty$  as  $x \rightarrow \infty$ , the only way that the Eq. (3.64) boundary condition can be satisfied is for  $B$  to be identically zero. With  $B = 0$ , application of the Eq. (3.63) boundary condition yields

$$A = \Delta p_{n0} \quad (3.67)$$

and

$$\Delta p_n(x) = \Delta p_{n0} e^{-x/L_p} \quad \Leftarrow \text{solution} \quad (3.68)$$

The Eq. (3.68) result is plotted in Fig. 3.27(b). In agreement with qualitative arguments, the nonpenetrating light merely gives rise to a monotonically decreasing  $\Delta p_n(x)$  starting from  $\Delta p_{n0}$  at  $x = 0$  and decreasing to  $\Delta p_n = 0$  as  $x \rightarrow \infty$ . Note that the precise functional form of the falloff in the excess carrier concentration is exponential with a characteristic decay length equal to  $L_p$ .

## 3.5 SUPPLEMENTAL CONCEPTS

### 3.5.1 Diffusion Lengths

The situation just encountered in Sample Problem No. 2—the creation (or appearance) of an excess of minority carriers along a given plane in a semiconductor, the subsequent diffusion of the excess from the point of injection, and the exponential falloff in the excess carrier concentration characterized by a decay length ( $L_p$ )—occurs often enough in semiconductor analyses that the characteristic length has been given a special name. Specifically,

$$L_p \equiv \sqrt{D_p \tau_p} \quad \text{associated with the minority carrier holes in an } n\text{-type material} \quad (3.69a)$$

and

$$L_N \equiv \sqrt{D_N \tau_n} \quad \text{associated with the minority carrier electrons in a } p\text{-type material} \quad (3.69b)$$

are referred to as *minority carrier diffusion lengths*.

Physically,  $L_p$  and  $L_N$  represent the average distance minority carriers can diffuse into a sea of majority carriers before being annihilated. This interpretation is clearly consistent with Sample Problem No. 2, where the average position of the excess minority carriers inside the semiconductor bar is

$$\langle x \rangle = \int_0^\infty x \Delta p_n(x) dx / \int_0^\infty \Delta p_n(x) dx = L_p \quad (3.70)$$

For memory purposes, minority carrier diffusion into a sea of majority carriers might be likened to a small group of animals attempting to cross a piranha-infested stretch of the Amazon River. In the analogy,  $L_p$  and  $L_N$  correspond to the average distance the animals advance into the river before being eaten.

Seeking an idea as to the size of diffusion lengths, let us assume  $T = 300$  K, the semiconductor is  $N_D = 10^{15}/\text{cm}^3$  doped Si, and  $\tau_p = 10^{-6}$  sec. In this sample

$$\begin{aligned} L_p &= \sqrt{D_p \tau_p} = \sqrt{(kT/q)\mu_p \tau_p} = [(0.0259)(458)(10^{-6})]^{1/2} \\ &= 3.44 \times 10^{-3} \text{ cm} \end{aligned}$$

Although the computed value is fairly representative, it should be understood that diffusion lengths can vary over several orders of magnitude because of wide variations in the carrier lifetimes.

### 3.5.2 Quasi-Fermi Levels

*Quasi-Fermi* levels are energy levels used to specify the carrier concentrations inside a semiconductor under *nonequilibrium* conditions.

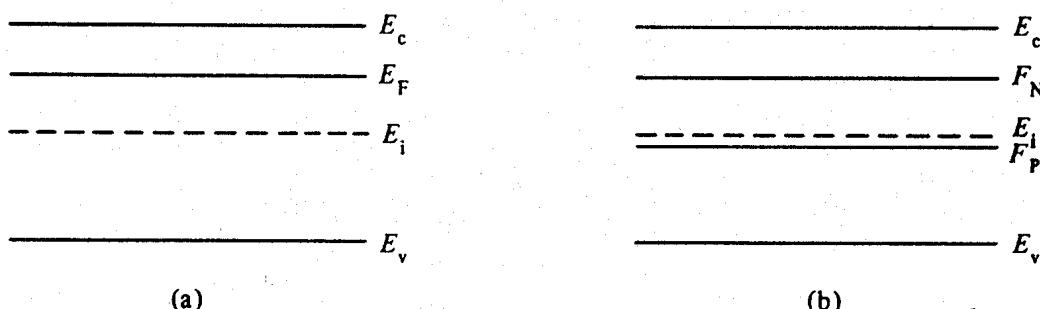
To understand the need for introducing quasi-Fermi levels, let us first refer to Sample Problem No. 1. In this problem, equilibrium conditions prevailed prior to  $t = 0$ , with  $n_0 = N_D = 10^{15}/\text{cm}^3$  and  $p_0 = 10^5/\text{cm}^3$ . The energy band diagram describing the equilibrium situation is shown in Fig. 3.28(a). A simple inspection of the energy band diagram and the Fermi level positioning also conveys the equilibrium carrier concentrations, since

$$n_0 = n_i e^{(E_F - E_i)/kT} \quad (3.71a)$$

$$p_0 = n_i e^{(E_i - E_F)/kT} \quad (3.71b)$$

The point we wish to emphasize is that under equilibrium conditions there is a one-to-one correspondence between the Fermi level and the carrier concentrations. Knowledge of  $E_F$  completely specifies  $n_0$  and  $p_0$  and vice versa.

Let us turn next to the nonequilibrium (steady state) situation inside the Problem 1 semiconductor at times  $t \gg \tau_p$ . For times  $t \gg \tau_p$ ,  $\Delta p_n = G_L \tau_p = 10^{11}/\text{cm}^3$ ,  $p = p_0 + \Delta p \cong 10^{11}/\text{cm}^3$ , and  $n \cong n_0 = 10^{15}/\text{cm}^3$ . Although  $n$  remains essentially unperturbed,  $p$  has



**Figure 3.28** Sample use of quasi-Fermi levels. Energy band description of the situation inside the semiconductor of Sample Problem No. 1 under (a) equilibrium conditions and (b) nonequilibrium conditions ( $t \gg \tau_p$ ).

increased by many orders of magnitude; and it is clear that the Fig. 3.28(a) diagram no longer describes the state of the system. In fact, the Fermi level is defined only for a system under equilibrium conditions and cannot be used to deduce the carrier concentrations inside a system in a nonequilibrium state.

The convenience of being able to deduce the carrier concentrations by inspection from the energy band diagram is extended to nonequilibrium conditions through the use of quasi-Fermi levels. This is accomplished by introducing two energies,  $F_N$  (the quasi-Fermi level for electrons) and  $F_P$  (the quasi-Fermi level for holes), which are *by definition* related to the nonequilibrium carrier concentrations in the same way  $E_F$  is related to the equilibrium carrier concentrations. To be specific, under nonequilibrium conditions and assuming the semiconductor to be nondegenerate,

$$n \equiv n_i e^{(F_N - E_i)/kT} \quad \text{or} \quad F_N \equiv E_i + kT \ln\left(\frac{n}{n_i}\right) \quad (3.72a)$$

and

$$p \equiv n_i e^{(E_i - F_P)/kT} \quad \text{or} \quad F_P \equiv E_i - kT \ln\left(\frac{p}{n_i}\right) \quad (3.72b)$$

Note that  $F_N$  and  $F_P$  are conceptual constructs, with the values of  $F_N$  and  $F_P$  being totally determined from a prior knowledge of  $n$  and  $p$ . Moreover, the quasi-Fermi level definitions have been carefully chosen so that when a perturbed system relaxes back toward equilibrium,  $F_N \rightarrow E_F$ ,  $F_P \rightarrow E_F$ , and Eqs. (3.72)  $\rightarrow$  Eqs. (3.71).

To provide a straightforward application of the quasi-Fermi level formalism, let us again consider the  $t \gg \tau_p$  state of the semiconductor in Sample Problem No. 1. First, since  $n \approx n_0$ ,  $F_N \approx E_F$ . Next, substitution of  $p = 10^{11}/\text{cm}^3$  ( $n_i = 10^{10}/\text{cm}^3$  and  $kT = 0.0259 \text{ eV}$ ) into Eq. (3.72b) yields  $F_P = E_i - 0.06 \text{ eV}$ . By eliminating  $E_F$  from the energy band diagram and drawing lines at the appropriate energies to represent  $F_N$  and  $F_P$ , one obtains the diagram displayed in Fig. 3.28(b). Figure 3.28(b) clearly conveys to any observer that the system under analysis is in a nonequilibrium state. When referenced to part (a), part (b) of Fig. 3.28 further indicates at a glance that low-level injection is taking place inside the semiconductor, creating a concentration of minority carrier holes in excess of  $n_i$ . A second example use of the quasi-Fermi level formalism, a use based on Sample Problem No. 2 and involving position dependent quasi-Fermi levels, is presented in Exercise 3.5 at the end of the section.

As a final point, it should be mentioned that the quasi-Fermi level formalism can be used to recast some of the carrier-action relationships in a more compact form. For example, the standard form of the equation for the total hole current reads

$$\mathbf{J}_P = q\mu_p p \mathbf{E} - qD_P \nabla p \quad (3.73)$$

(same as 3.18a)

Differentiating both sides of Eq. (3.72b) with respect to position, one obtains

$$\nabla p = \left( \frac{n_i}{kT} \right) e^{(E_i - F_p)/kT} (\nabla E_i - \nabla F_p) \quad (3.74a)$$

$$= \left( \frac{qp}{kT} \right) \mathcal{E} - \left( \frac{p}{kT} \right) \nabla F_p \quad (3.74b)$$

The identity  $\mathcal{E} = \nabla E_i/q$  (the three-dimensional version of Eq. 3.15) is employed in progressing from Eq. (3.74a) to Eq. (3.74b). Next, eliminating  $\nabla p$  in Eq. (3.73) using Eq. (3.74b) gives

$$\mathbf{J}_p = q \left( \mu_p - \frac{qD_p}{kT} \right) p \mathcal{E} + \left( \frac{qD_p}{kT} \right) p \nabla F_p \quad (3.75)$$

From the Einstein relationship, however,  $qD_p/kT = \mu_p$ . We therefore conclude

$$\mathbf{J}_p = \mu_p p \nabla F_p \quad (3.76a)$$

Similarly,

$$\mathbf{J}_n = \mu_n n \nabla F_n \quad (3.76b)$$

Since  $\mathbf{J}_p \propto \nabla F_p$  and  $\mathbf{J}_n \propto \nabla F_n$  in Eqs. (3.76), one is led to a very interesting general interpretation of energy band diagrams containing quasi-Fermi levels. Namely, a quasi-Fermi level that varies with position ( $dF_p/dx \neq 0$  or  $dF_n/dx \neq 0$ ) indicates at a glance that current is flowing inside the semiconductor.

### Exercise 3.5

**P:** In Sample Problem No. 2, nonpenetrating illumination of a semiconductor bar was found to cause a steady state, excess-hole concentration of  $\Delta p_n(x) = \Delta p_{n0} \exp(-x/L_p)$ . Given the prevailing low-level injection conditions, and noting that  $p = p_0 + \Delta p$ , we can therefore state

$$n \approx n_0$$

$$p = p_0 + \Delta p_{n0} e^{-x/L_p}$$

for the illuminated sample.

- (a) Making use of Eqs. (3.72), establish relationships for  $F_n$  and  $F_p$  in the illuminated bar.
- (b) Show that  $F_p$  is a linear function of  $x$  at points where  $\Delta p_n(x) \gg p_0$ .

- (c) Using the results of parts (a) and (b), sketch the energy band diagrams describing the semiconductor bar of Sample Problem No. 2 under equilibrium and illuminated steady state conditions. (Assume  $\mathcal{E} = 0$  in the illuminated bar.)
- (d) Is there a hole current in the illuminated bar under steady state conditions? Explain.
- (e) Is there an electron current in the illuminated bar under steady state conditions? Explain.

S: (a) Since  $n \approx n_0$ , it follows from Eq. (3.72a) that  $F_N \approx E_F$ . Likewise, substituting the preceding  $p$ -expression into Eq. (3.72b), we conclude

$$F_P = E_i - kT \ln(p/n_i) = E_i - kT \ln[p_0/n_i + (\Delta p_{n0}/n_i)e^{-x/L_P}]$$

(b) If  $\Delta p_n(x) \gg p_0$ , then  $(\Delta p_{n0}/n_i) \exp(-x/L_P) \gg p_0/n_i$  and

$$F_P \approx E_i - kT \ln[(\Delta p_{n0}/n_i)e^{-x/L_P}]$$

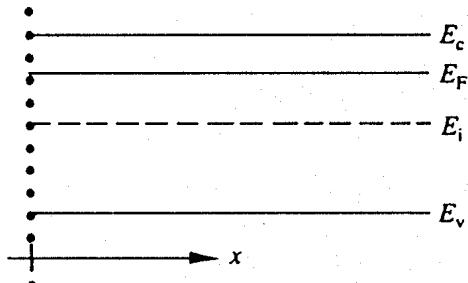
or

$$F_P = E_i - kT \ln(\Delta p_{n0}/n_i) + kT(x/L_P)$$

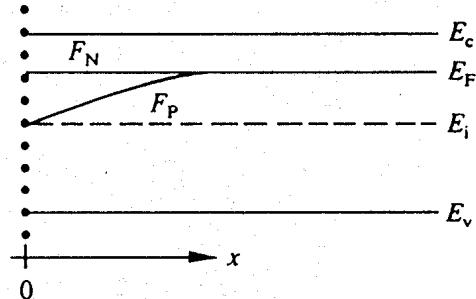
(c) We know from Sample Problem No. 2 that  $\Delta p_{n0} = 10^{10}/\text{cm}^3$ ,  $n_i = 10^{10}/\text{cm}^3$ , and  $p_0 = n_i^2/N_D = 10^5/\text{cm}^3$ . Thus

- (i) Near  $x = 0$ ,  $\Delta p_n(x) \gg p_0$  and  $F_P$  is a linear function of  $x$ .
- (ii) At  $x = 0$ ,  $\Delta p_{n0} = n_i$  and we deduce from the part (b) result that  $F_P = E_i$ .
- (iii) For large  $x$ ,  $F_P$  eventually approaches  $F_N = E_F$ .
- (iv)  $F_N - E_i \approx E_F - E_i = kT \ln(N_D/n_i) = 0.30 \text{ eV}$ .

Utilizing the preceding information, one concludes



Equilibrium



Steady state illuminated

- (d) Assuming  $p \neq 0$ , it follows from Eq. (3.76a) that there will be a hole current whenever  $dF_p/dx \neq 0$ . There is obviously a hole current in the illuminated bar near  $x = 0$ .

(e) Appearances can sometimes be deceiving;  $J_N \neq 0$  near  $x = 0$ ! One might conclude from the part (c) result that  $dF_N/dx = 0$  and therefore  $J_N = 0$ . Under steady state conditions, however, one must have  $J(\text{total current}) = J_N + J_P = \text{constant}$  at all points in the bar. Since no current leaves the bar at  $x = 0$ ,  $J$  moreover must vanish at  $x = 0$ . Thus  $J_N(x) = -J_P(x)$  and we know  $J_P \neq 0$  near  $x = 0$ . The apparent discrepancy here stems from the fact that  $J_N$  is proportional to both  $n$  and  $\nabla F_N$ . Because the majority carrier electron concentration is much larger than the minority carrier hole concentration,  $dF_N/dx$  must be correspondingly smaller than  $dF_P/dx$ . The slope in  $F_N$  simply cannot be detected by inspecting the energy band diagram.

### 3.6 SUMMARY AND CONCLUDING COMMENTS

Most of the chapter was devoted to examining the three primary types of carrier action occurring inside of semiconductors: drift, diffusion, and recombination-generation. In each case the carrier action was first defined and visualized. Drift is charged particle motion in response to an applied electric field. Diffusion is migration from regions of high particle concentration to regions of low particle concentration due to the random thermal motion of the carriers. Recombination and generation are respectively the annihilation and creation of carriers. Next the quantitative effect of each type of carrier action was analyzed. Drift and diffusion give rise to particle currents (Eqs. 3.4 and 3.17–3.19); recombination-generation acts to change the carrier concentrations as a function of time (Eqs. 3.34). Associated with the quantitative analysis of each carrier action there arises a “constant of the motion,” an important material-based parameter that specifies the magnitude of the carrier action in a given semiconductor sample. The carrier mobilities, the diffusion coefficients, and the minority carrier lifetimes are the material parameters associated with drift, diffusion, and recombination-generation, respectively. In the major semiconductors the carrier mobilities are very well characterized as a function of doping and temperature. Select mobility data are graphed in Figs. (3.5) and (3.7); an empirical fit relationship for the carrier mobilities in Si is presented in Exercise 3.1. The diffusion coefficients in nondegenerate semiconductors are computed from the carrier mobilities using the Einstein relationship (Eqs. 3.25). Conversely, an experimental measurement is normally performed to determine the carrier lifetimes in a given semiconductor sample.

Although introduced and examined individually, the various types of carrier action occur simultaneously inside semiconductors. Mathematically combining the overall effect of the various carrier activities leads to the continuity equations (Eqs. 3.46). The continuity equations in turn simplify to the minority carrier diffusion equations (Eqs. 3.54) under conditions encountered in many practical problems. Additional simplifications and their effect on the minority carrier diffusion equations are noted in Table 3.1; extensively utilized solutions to simplified forms of the equations are collected in Table 3.2. The continuity or minority carrier diffusion equations combined with the other relationships established in the chapter allow one to model the state of a semiconductor subject to an external perturbation. The more important working relationships are organized and repeated in Table 3.3.

**Table 3.3** Carrier Action Equation Summary.

<i>Equations of State</i>			
$\frac{\partial n}{\partial t} = \frac{1}{q} \nabla \cdot \mathbf{J}_N + \left. \frac{\partial n}{\partial t} \right _{R-G}^{\text{thermal}} + \left. \frac{\partial n}{\partial t} \right _{R-G}^{\text{other processes}}$		$\frac{\partial \Delta n_p}{\partial t} = D_N \frac{\partial^2 \Delta n_p}{\partial x^2} - \frac{\Delta n_p}{\tau_n} + G_L$	
$\frac{\partial p}{\partial t} = -\frac{1}{q} \nabla \cdot \mathbf{J}_P + \left. \frac{\partial p}{\partial t} \right _{R-G}^{\text{thermal}} + \left. \frac{\partial p}{\partial t} \right _{R-G}^{\text{other processes}}$		$\frac{\partial \Delta p_n}{\partial t} = D_P \frac{\partial^2 \Delta p_n}{\partial x^2} - \frac{\Delta p_n}{\tau_p} + G_L$	
<i>Current and R-G Relationships</i>			
$\mathbf{J}_N = \mathbf{J}_{N\text{drift}} + \mathbf{J}_{N\text{diff}} = q\mu_n n \mathcal{E} + qD_N \nabla n$		$\left. \frac{\partial n}{\partial t} \right _{i-\text{thermal}}^{\text{R-G}} = -\frac{\Delta n}{\tau_n}$	
$\Updownarrow \text{drift}$	$\Updownarrow \text{diffusion}$		
$\mathbf{J}_P = \mathbf{J}_{P\text{drift}} + \mathbf{J}_{P\text{diff}} = q\mu_p p \mathcal{E} - qD_P \nabla p$		$\left. \frac{\partial p}{\partial t} \right _{i-\text{thermal}}^{\text{R-G}} = -\frac{\Delta p}{\tau_p}$	
$\mathbf{J} = \mathbf{J}_N + \mathbf{J}_P$			
<i>Key Parametric Relationships</i>			
$L_N \equiv \sqrt{D_N \tau_n}$	$\frac{D_N}{\mu_n} = \frac{kT}{q}$	$\tau_n = \frac{1}{c_n N_T}$	
$L_P \equiv \sqrt{D_P \tau_p}$	$\frac{D_P}{\mu_p} = \frac{kT}{q}$	$\tau_p = \frac{1}{c_p N_T}$	
<i>Resistivity and Electrostatic Relationships</i>			
$\rho = \frac{1}{q(\mu_n n + \mu_p p)}$	$\rho = \frac{1}{q\mu_n N_D}$		... n-type semiconductor
	$\rho = \frac{1}{q\mu_p N_A}$		... p-type semiconductor
$\mathcal{E} = \frac{1}{q} \frac{dE_c}{dx} = \frac{1}{q} \frac{dE_v}{dx} = \frac{1}{q} \frac{dE_i}{dx}$		$V = -\frac{1}{q}(E_c - E_{\text{ref}})$	
<i>Quasi-Fermi Level Relationships</i>			
$F_N \equiv E_i + kT \ln\left(\frac{n}{n_i}\right)$		$\mathbf{J}_N = \mu_n n \nabla F_N$	
$F_P \equiv E_i - kT \ln\left(\frac{p}{n_i}\right)$		$\mathbf{J}_P = \mu_p p \nabla F_P$	

A number of topics related to carrier action were also addressed in the chapter, including resistivity and resistivity measurements, the hot-point probe measurement, constancy of the equilibrium Fermi level, nonuniform doping and the associated built-in electric field,  $E$ - $k$  diagrams, measurement of the minority carrier lifetimes, and diffusion lengths. Moreover, although not identified as such, the energy band diagram was subject to further development. Specifically, it was pointed out that the existence of an electric field inside the semiconductor causes band bending or a variation of the energy bands with position. By simply inspecting an energy band diagram, it is possible to ascertain the general functional dependence of the electrostatic potential and electric field present in the material. In the discussion of recombination-generation, another level was added near midgap. This level, arising from R-G centers, plays a dominant role in the thermal communication between the energy bands. Lastly, quasi-Fermi levels were introduced to describe nonequilibrium conditions.

## PROBLEMS

CHAPTER 3 PROBLEM INFORMATION TABLE

<i>Problem</i>	<i>Complete After</i>	<i>Difficulty Level</i>	<i>Suggested Point Weighting</i>	<i>Short Description</i>
3.1	3.3.1	1	16 (2 each part)	Energy band visualization
3.2	3.1.4 (a-d) 3.3.4 (e-h)	1-2	16 (2 each part)	Short answer
● 3.3	3.1.3	3	25	$\mu$ vs. $T$ plots
● 3.4	"	3	18 (b-15, c-3)	$\mu_n$ vs. $T$ student data
3.5	3.1.4	a-1, b-3	12 (a-6, b-6)	Intrinsic/maximum $\rho$
3.6	"	2	15 (3 each part)	Resistivity questions
● 3.7	"	2	15	$\rho$ vs. $N_A$ , $N_D$ plot
● 3.8	"	3-4	20 (10 each part)	$N_D(x)$ variation in resistor
● 3.9	"	3	18 (a-12, b-6)	$\rho$ vs. $T$ plots
3.10	"	4	20 (5 each part)	Temperature sensor
3.11	3.1.5	2	5	Compute thermal speed
3.12	3.1.5 (a-d) 3.2.4 (e, f)	2, e-3	16/diagram (a::c-2, d-4, e-2, f-4)	Interpret E-band diagrams
3.13	3.2.4	3	15 (a-10, b-3, c-2)	Built-in electric field
● 3.14	"	2	15 (a-3, b-10, c-2)	$D$ vs. $N_A$ , $N_D$ computation
3.15	3.3.3	3	10	Eq. (3.35) $\rightarrow$ Eqs. (3.34)
3.16	3.4.2	1	6 (2 each part)	Simple Diff. Eq. questions

3.17	3.4.4	1	6	Diff. Eq., no R-G region
3.18	"	2	12 (a::c-2, d-6)	Diff. Eq., mysterious ray
3.19	"	2	8	Diff. Eq., $G_L \rightarrow G_L/2$
3.20	"	2-3	10	Diff. Eq., light + edge R
3.21	"	3	15 (a::d-2, e-7)	Diff. Eq., double-ended
3.22	"	3-4	15 (a::c-2, d-9)	Diff. Eq., half-illuminated
3.23	"	2	8 (a-2, b-4, c-2)	CdS Photoconductor
3.24	3.5.2	2	12 (a-2, b-4, c::e-2)	Quasi-Fermi levels
3.25	"	3-4	15	Quasi-Fermi levels
• 3.26	"	3-4	15 (a-13, b-2)	Plot quasi-Fermi levels

### 3.1 Using the energy band diagram, indicate how one visualizes

- (a) The existence of an electric field inside a semiconductor.
- (b) An electron with a K.E. = 0.
- (c) A hole with a K.E. =  $E_G/4$ .
- (d) Photogeneration.
- (e) Direct thermal generation.
- (f) Band-to-band recombination.
- (g) Recombination via R-G centers.
- (h) Generation via R-G centers.

### 3.2 Short Answer

- (a) An average hole drift velocity of  $10^3$  cm/sec results when 2 V is applied across a 1-cm-long semiconductor bar. What is the hole mobility inside the bar?
- (b) Name the two dominant carrier *scattering* mechanisms in nondegenerately doped semiconductors of device quality.
- (c) For a given semiconductor the carrier mobilities in intrinsic material are (choose one: higher than, lower than, the same as) those in heavily doped material. Briefly explain why the mobilities in intrinsic material are (chosen answer) those in heavily doped material.
- (d) Two GaAs wafers, one *n*-type and one *p*-type, are uniformly doped such that  $N_D(\text{wafer 1}) = N_A(\text{wafer 2}) \gg n_i$ . Which wafer will exhibit the larger resistivity? Explain.

- (e) The electron mobility in a silicon sample is determined to be  $1300 \text{ cm}^2/\text{V}\cdot\text{sec}$  at room temperature. What is the electron diffusion coefficient?
- (f) What is the algebraic statement of low-level injection?
- (g) Light is used to create excess carriers in silicon. These excess carriers will predominantly recombine via (choose one: band-to-band, R-G center, or photo) recombination.
- (h) Prior to processing, a silicon sample contains  $N_D = 10^{14}/\text{cm}^3$  donors and  $N_T = 10^{11}/\text{cm}^3$  R-G centers. After processing (say in the fabrication of a device), the sample contains  $N_D = 10^{17}/\text{cm}^3$  donors and  $N_T = 10^{10}/\text{cm}^3$  R-G centers. Did processing increase or decrease the minority carrier lifetime? Explain.

• **3.3** Complete part (b) of Exercise 3.1.

- **3.4** (a) (Optional) Read the Experiment No. 7 Introduction and Measurement System description in R. F. Pierret, *Semiconductor Measurement Laboratory Operations Manual*, Supplement A, in the Modular Series on Solid State Devices, Addison-Wesley Publishing Co., Reading, MA, © 1991.
- (b) The measurement described in the part (a) reference is used to determine the mobility in an  $N_D < 10^{14}/\text{cm}^3$  Si sample as a function of temperature from roughly room temperature to  $T = 150 \text{ K}$ . Representative  $\mu_n$  versus  $T$  (K) data derived by undergraduate students is tabulated in the following table. Assuming  $\mu_n \propto T^{-b}$  [or  $\ln(\mu_n) = \text{constant} - b \ln(T)$ ] can be fitted to the data, use the MATLAB polyfit function or an equivalent least squares fit<sup>†</sup> to determine the best fit value of the power factor  $b$ . Draw the fit line and note the  $b$ -fit value on a log-log plot of the data.
  - (c) Compare your part (b) result with Fig. 3.7(a). Is it significant that the experimental data was derived from a lowly doped sample? Is the experimentally derived  $b$ -value consistent with the  $b$ -value noted on the text plot? Are the measured mobility values of the proper magnitude?

<sup>†</sup> Given  $N$  data points,  $(x_i, y_i)$  with  $i = 1$  to  $N$ , the straight line  $y = a + bx$ , which is the “best fit” to the data, will have a  $y$ -axis intercept and slope respectively equal to

$$a = \frac{\sum y_i \sum x_i^2 - \sum x_i \sum x_i y_i}{N \sum x_i^2 - (\sum x_i)^2} \quad \text{and} \quad b = \frac{N \sum x_i y_i - \sum x_i \sum y_i}{N \sum x_i^2 - (\sum x_i)^2}$$

All summations are from  $i = 1$  to  $i = N$ . The “best fit” criterion leading to the expressions for  $a$  and  $b$  is that the square of the difference between the data points and the fitted line is minimized; hence the name “least squares fit.”

$T$ (K)	$\mu_n$ (cm <sup>2</sup> /V-sec)
290	1501
280	1646
270	1805
260	1985
250	2185
240	2415
230	2675
220	2978
210	3306
200	3743
190	4209
180	4619
170	5216
160	5910
150	6757

### 3.5 Intrinsic and Maximum Resistivity

- (a) Determine the resistivity of intrinsic Ge, Si, and GaAs at 300 K.
- (b) Determine the maximum possible resistivity of Ge, Si, and GaAs at 300 K.

### 3.6 More Resistivity Questions

- (a) A silicon sample maintained at room temperature is uniformly doped with  $N_D = 10^{16}/\text{cm}^3$  donors. Calculate the resistivity of the sample using Eq. (3.8a). Compare your calculated result with the  $\rho$  deduced from Fig. 3.8(a).
- (b) 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.)
- (c) Compute the resistivity of intrinsic ( $N_A = 0$ ,  $N_D = 0$ ) silicon at room temperature. How does your result here compare with that for part (b)?

- (d) A 500-ohm resistor is to be made from a bar-shaped piece of *n*-type Si. The bar has a cross-sectional area of  $10^{-2} \text{ cm}^2$  and a current-carrying length of 1 cm. Determine the doping required.
- (e) A lightly doped ( $N_D < 10^{14}/\text{cm}^3$ ) Si sample is heated up from room temperature to 100°C.  $N_D \gg n_i$  at both room temperature and 100°C. Is the resistivity of the sample expected to increase or decrease? Explain.
- 3.7 Making use of the mobility fit relationships and parameters quoted in Exercise 3.1, construct a plot of the silicon resistivity versus impurity concentration at  $T = 300 \text{ K}$ . Include curves for both *n*- and *p*-type silicon over the range  $10^{13}/\text{cm}^3 \leq N_A \text{ or } N_D \leq 10^{20}/\text{cm}^3$ . Compare your result with Fig. 3.8(a).
- 3.8 Resistors in ICs are sometimes thin semiconductor layers near the surface of a wafer. However, formation of the layer by diffusion or ion implantation (discussed in Chapter 4) is likely to give rise to a doping concentration that varies with depth into the layer. Let us examine how the resistance is computed when the doping varies with depth.
- (a) Given a bar-shaped layer of width  $W$ , length  $L$ , and depth  $d$ , and assuming an arbitrary  $N_D(x)$  variation with the depth  $x$  from the wafer surface, show that the resistance of the layer is to be computed from
- $$R = \frac{L}{W} \left[ \frac{1}{q \int_0^d \mu_n(x) N_D(x) dx} \right]$$
- (b) Taking  $N_D(x) = N_{D0} \exp(-ax) + N_{DB}$ , compute and plot  $R$  versus  $N_{D0}$  for  $10^{14}/\text{cm}^3 \leq N_{D0} \leq 10^{18}/\text{cm}^3$  when  $L = W$ ,  $N_{DB} = 10^{14}/\text{cm}^3$ ,  $d = 5 \mu\text{m}$ , and  $1/a = 1 \mu\text{m}$ .
- 3.9 (a) Modify your  $\mu$  versus  $T$  program from Problem 3.3 to compute and construct *semi-log* plots of  $\rho$  (log scale) versus  $T$  for  $200 \text{ K} \leq T \leq 500 \text{ K}$  and  $N_D$  or  $N_A$  stepped in decade values from  $10^{14}/\text{cm}^3$  to  $10^{18}/\text{cm}^3$ . [NOTE: The lower doping curves are in error at the higher temperatures if Eqs. (3.8) are used to compute  $\rho$ . Can you explain why? Assuming Eqs. (3.8) were used in the computation, how must the  $\rho$  versus  $T$  program be modified to correct the error?]
- (b) Compare your *n*-type plot with Fig. 7 in Li and Thurber, Solid-State Electronics, 20, 609 (1977). Compare your *p*-type plot with Fig. 7 in Li, Solid-State Electronics, 21, 1109 (1978). Comment on the comparisons.
- 3.10 Your boss asks you to construct a temperature sensor for measuring the ambient outside temperature in W. Lafayette IN ( $-30^\circ\text{C} \leq T \leq 40^\circ\text{C}$ ). You decide to base operation of the temperature sensor on the change in resistance associated with a bar-shaped piece of Si.
- (a) Restricting yourself to nondegenerate dopings and reasonable device dimensions, and also requiring the resistance to be readily measured with a hand-held multimeter (say  $1 \Omega \leq R \leq 1000 \Omega$ ), indicate the doping and dimensions of your sensor.

- (b) Derive an expression for the sensitivity ( $dR/dT$  in  $\Omega/\text{ }^\circ\text{C}$ ) of your sensor over the operating temperature range. Relative to sensitivity, is it preferable to use low or high dopings? Explain.
- (c) What are the upper and lower temperature limits of operation of your sensor (approximately)? Explain.
- (d) Plot  $R$  versus  $T$  of your sensor over the required temperature range of operation.

**3.11** Thermal energy alone was noted to contribute to relatively high carrier velocities inside of a semiconductor. Considering a nondegenerate semiconductor maintained at room temperature, compute the thermal velocity of electrons having a kinetic energy corresponding to the peak of the electron distribution in the conduction band. Set  $m^* = m_0$  in performing your calculation. (This problem assumes the prior working of either Problem 2.7 or Problem 2.8.)

### 3.12 Interpretation of Energy Band Diagrams

Six different silicon samples maintained at 300 K are characterized by the energy band diagrams in Fig. P3.12. Answer the questions that follow after choosing a specific diagram for analysis. Possibly repeat using other energy band diagrams. (Excessive repetitions have been known to lead to the onset of insanity.)

- (a) Do equilibrium conditions prevail? How do you know?
- (b) Sketch the electrostatic potential ( $V$ ) inside the semiconductor as a function of  $x$ .
- (c) Sketch the electric field ( $E$ ) inside the semiconductor as a function of  $x$ .
- (d) The carrier pictured on the diagram moves back and forth between  $x = 0$  and  $x = L$  without changing its total energy. Sketch the K.E. and P.E. of the carrier as a function of position inside the semiconductor. Let  $E_F$  be the energy reference level.
- (e) Roughly sketch  $n$  and  $p$  versus  $x$ .
- (f) On the same set of coordinates, make a rough sketch of the electron drift-current density ( $J_{N\text{ldrift}}$ ) and the electron diffusion-current density ( $J_{N\text{ldiff}}$ ) inside the Si sample as a function of position. Be sure to graph the proper polarity of the current densities at all points and clearly identify your two current components. Also briefly explain how you arrived at your sketch.

**3.13** The nonuniform doping in the central region of bipolar junction transistors (BJTs) creates a built-in field that assists minority carriers across the region and increases the maximum operating speed of the device. Suppose the BJT is a Si device maintained under equilibrium conditions at room temperature with a central region of length  $L$ . Moreover, the nonuniform acceptor doping is such that

$$p(x) \cong N_A(x) = n_i e^{(a-x)/b} \quad \dots \quad 0 \leq x \leq L$$

where  $a = 1.8 \mu\text{m}$ ,  $b = 0.1 \mu\text{m}$ , and  $L = 0.8 \mu\text{m}$ .

- (a) Draw the energy band diagram for the  $0 \leq x \leq L$  region specifically showing  $E_c$ ,  $E_F$ ,  $E_i$ , and  $E_v$  on your diagram. Explain how you arrived at your diagram.

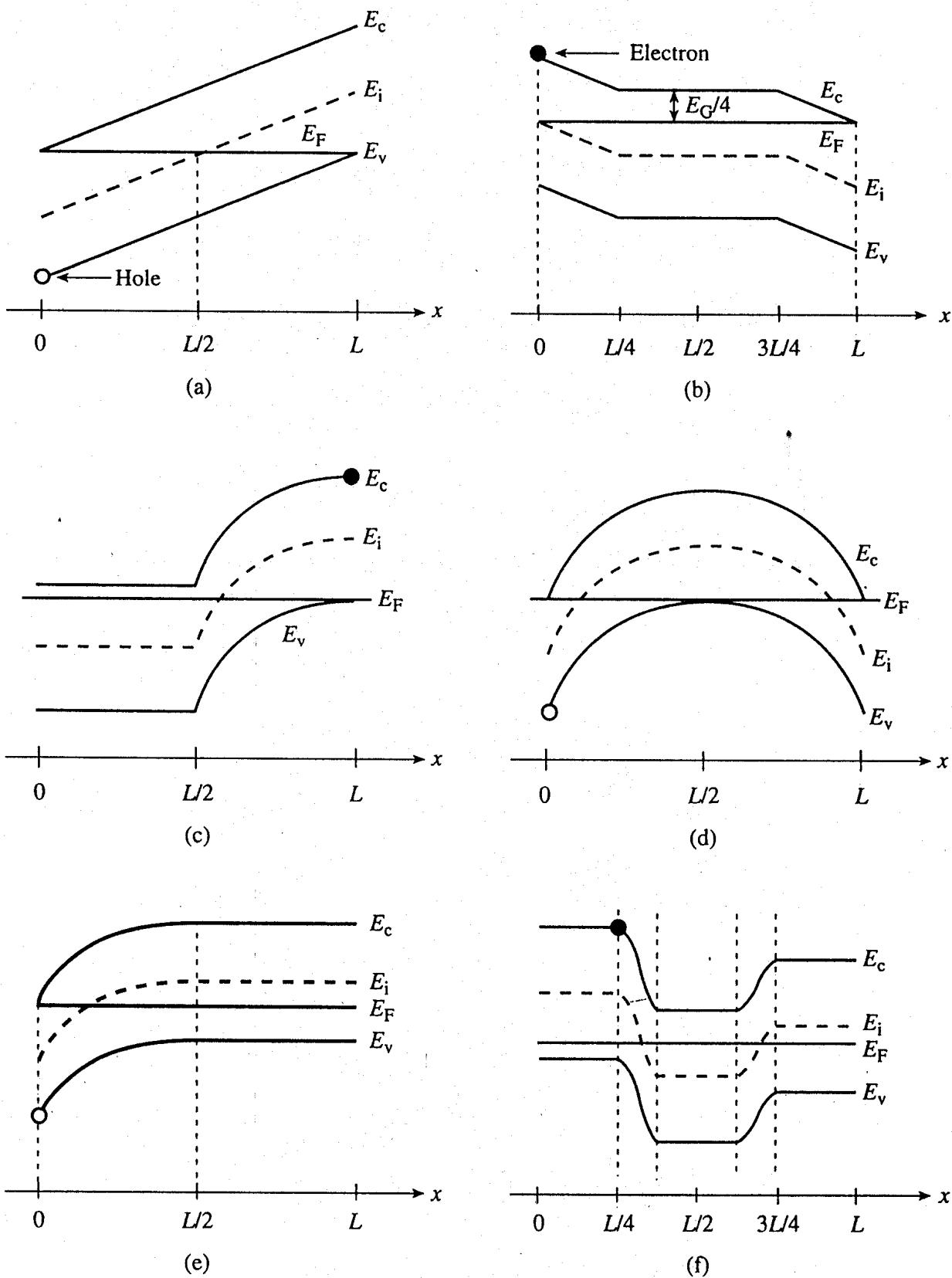


Figure P3.12

- (b) Make a sketch of the  $\mathcal{E}$ -field inside the region as a function of position, and compute the value of  $\mathcal{E}$  at  $x = L/2$ .
- (c) Is the built-in electric field such as to aid the motion of minority carrier electrons in going from  $x = 0$  to  $x = L$ ? Explain.

- 3.14** (a) Based on information found in the text proper, *roughly sketch* the expected variation of  $D_N$  and  $D_P$  versus doping appropriate for  $10^{14}/\text{cm}^3 \leq N_A$  or  $N_D \leq 10^{18}/\text{cm}^3$  doped Si maintained at  $T = 300$  K. Explain how you arrived at the form of your sketch.
- (b) Making use of the fit relationships and parameters quoted in Exercise 3.1, construct a plot of  $D_N$  and  $D_P$  versus  $N_A$  or  $N_D$  for  $10^{14}/\text{cm}^3 \leq N_A$  or  $N_D \leq 10^{18}/\text{cm}^3$  doped Si maintained at  $T = 300$  K.
- (c) Why was the upper doping limit taken to be  $10^{18}/\text{cm}^3$  in the part (b) computation?

- 3.15** Taking  $E_T \cong E_i$  so that  $n_1 \cong p_1 \cong n_i$ , setting  $\Delta n = \Delta p$ , and assuming  $\tau_n$  is comparable to  $\tau_p$ , show that the general-case R-G relationship of Eq. (3.35) reduces to the special-case relationships of Eqs. (3.34) when one has low-level injection in a specific material type.

- 3.16**  $\partial \Delta n_p / \partial t = D_N \partial^2 \Delta n_p / \partial x^2 - \Delta n_p / \tau_n + G_L$  is known as the minority carrier diffusion equation for electrons.

- (a) Why is it called a *diffusion equation*?
- (b) Why is it referred to as a *minority carrier equation*?
- (c) The equation is valid only under low-level injection conditions. Why?

- 3.17** Show that, under steady state conditions

$$\Delta p_n(x) = \Delta p_{n0}(1 - x/L) \quad \dots \quad 0 \leq x \leq L$$

is the special-case solution of the minority carrier diffusion equation that will result if (1) one assumes *all* recombination-generation processes are negligible within an *n*-type semiconductor of length  $L$ , and (2) one employs the boundary conditions  $\Delta p_n(0) = \Delta p_{n0}$  and  $\Delta p_n(L) = 0$ . (Neglecting recombination-generation is an excellent approximation when  $L$  is much less than a minority carrier diffusion length. A  $\Delta p(x)$  solution of the above type is frequently encountered in practical problems.)

- 3.18** The earth is hit by a mysterious ray that momentarily wipes out all minority carriers. Majority carriers are unaffected. Initially in equilibrium and not affected by room light, a uniformly doped silicon wafer sitting on your desk is struck by the ray at time  $t = 0$ . The wafer doping is  $N_A = 10^{16}/\text{cm}^3$ ,  $\tau_n = 10^{-6}$  sec, and  $T = 300$  K.

- (a) What is  $\Delta n$  at  $t = 0^+$ ? ( $t = 0^+$  is an imperceptible fraction of a second after  $t = 0$ .)
- (b) Does generation or recombination dominate at  $t = 0^+$ ? Explain.
- (c) Do low-level injection conditions exist inside the wafer at  $t = 0^+$ ? Explain.
- (d) Starting from the appropriate differential equation, derive  $\Delta n_p(t)$  for  $t > 0$ .

**3.19** A silicon wafer ( $N_A = 10^{14}/\text{cm}^3$ ,  $\tau_n = 1 \mu\text{sec}$ ,  $T = \text{room temperature}$ ) is first illuminated for a time  $t \gg \tau_n$  with light which generates  $G_{L0} = 10^{16}$  electron-hole pairs per  $\text{cm}^3\text{-sec}$  uniformly throughout the volume of the silicon. At time  $t = 0$  the light intensity is reduced, making  $G_L = G_{L0}/2$  for  $t \geq 0$ . Determine  $\Delta n_p(t)$  for  $t \geq 0$ .

**3.20** A semi-infinite *p*-type bar (see Fig. P3.20) is illuminated with light which generates  $G_L$  electron-hole pairs per  $\text{cm}^3\text{-sec}$  uniformly throughout the volume of the semiconductor. Simultaneously, carriers are extracted at  $x = 0$  making  $\Delta n_p = 0$  at  $x = 0$ . Assuming a steady state condition has been established and  $\Delta n_p(x) \ll p_0$  for all  $x$ , determine  $\Delta n_p(x)$ .

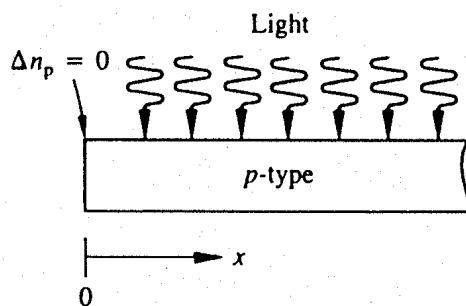


Figure P3.20

**3.21** The two ends of a uniformly doped *n*-type Si bar of length  $L$  are simultaneously illuminated so as to create  $\gamma N_D$  excess holes at both  $x = 0$  and  $x = L$  (see Fig. P3.21). The wavelength and intensity of the illumination are such that no light penetrates into the interior ( $0 < x < L$ ) of the bar and  $\gamma = 10^{-3}$ . Also, steady state conditions prevail,  $T = 300 \text{ K}$ , and  $N_D \gg n_i$ .

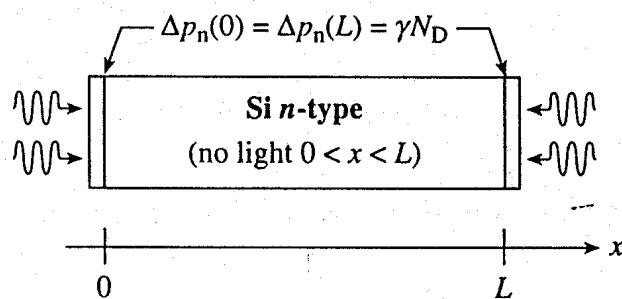


Figure P3.21

- Based on *qualitative reasoning*, sketch the expected general form of the  $\Delta p_n(x)$  solution.
- Do low-level injection conditions prevail inside the illuminated bar? Explain.
- Write down the differential equation (simplest form possible) you must solve to determine  $\Delta p_n(x)$  inside the bar.
- Write down the general form of the  $\Delta p_n(x)$  solution and the boundary condition(s) appropriate for this particular problem.

- (e) Establish an expression for the hole current ( $J_p$ ) flowing in the illuminated bar at  $x = 0$ . [Your answer may be left in terms of the arbitrary constant(s) appearing in the general form of the  $\Delta p_n(x)$  solution.]

**3.22** As pictured in Fig. P3.22, the  $x > 0$  portion of an infinite semiconductor is illuminated with light. The light generates  $G_L = 10^{15}$  electron-hole pairs/cm<sup>3</sup>-sec uniformly throughout the  $x > 0$  region of the bar.  $G_L = 0$  for  $x < 0$ , steady state conditions prevail, the semiconductor is silicon, the entire bar is uniformly doped with  $N_D = 10^{18}/\text{cm}^3$ ,  $\tau_p = 10^{-6}$  sec, and  $T = 300$  K.

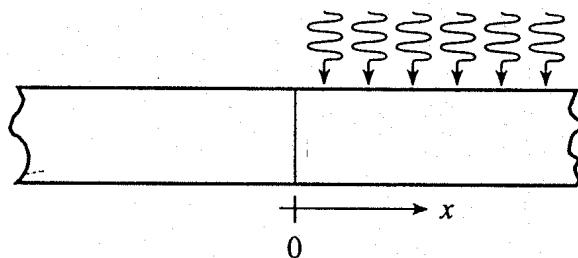


Figure P3.22

- (a) What is the hole concentration at  $x = -\infty$ ? Explain.
- (b) What is the hole concentration at  $x = +\infty$ ? Explain.
- (c) Do low-level injection conditions prevail? Explain.
- (d) Determine  $\Delta p_n(x)$  for all  $x$ .

NOTE: (1) Separate  $\Delta p_n(x)$  expressions apply for  $x > 0$  and  $x < 0$ .  
 (2) Both  $\Delta p_n$  and  $d\Delta p_n/dx$  must be continuous at  $x = 0$ .

**3.23** CdS is the most widely used material for constructing commercial photoconductors operating in the visible portion of the spectrum. The CdS photoconductor has a high sensitivity and its spectral response closely matches that of the human eye. A model VT333 CdS Photoconductor is pictured in Fig. P3.23.

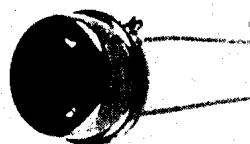


Figure P3.23

- (a) Speculate why the conducting film has a snake-like pattern.
- (b) The VT333 resistor pattern is approximately 0.3 mm wide and 3 cm long. Estimating the deposited CdS film to be  $5 \mu\text{m}$  thick,  $N_D = 10^{13}/\text{cm}^3 \gg n_i$ , and  $\mu_n = 100 \text{ cm}^2/\text{V}\cdot\text{sec}$ , compute the dark resistance of the device.
- (c) The VT333 exhibited a  $250 \Omega$  resistance when illuminated with a microscope light. Can the usual relationships be used to determine the  $G_L$  required to produce the cited resistance? Explain.

**3.24** The equilibrium and steady state conditions before and after illumination of a semiconductor are characterized by the energy band diagrams shown in Fig. P3.24.  $T = 300 \text{ K}$ ,  $n_i = 10^{10}/\text{cm}^3$ ,  $\mu_n = 1345 \text{ cm}^2/\text{V}\cdot\text{sec}$ , and  $\mu_p = 458 \text{ cm}^2/\text{V}\cdot\text{sec}$ . From the information provided, determine

- $n_0$  and  $p_0$ , the equilibrium carrier concentrations.
- $n$  and  $p$  under steady state conditions.
- $N_D$ .
- Do we have "low-level injection" when the semiconductor is illuminated? Explain.
- What is the resistivity of the semiconductor before and after illumination?

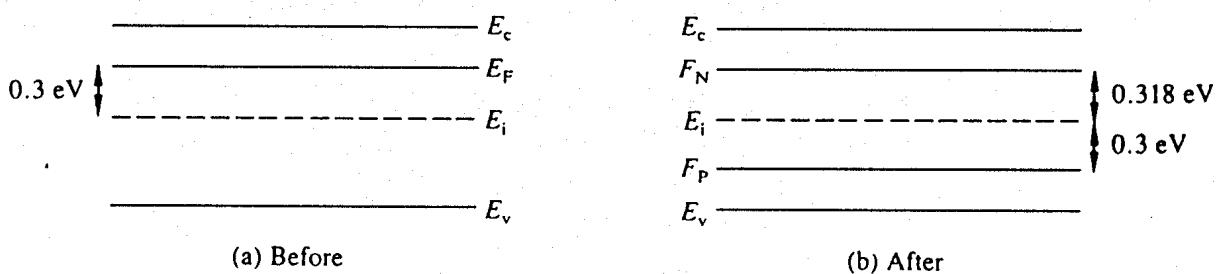


Figure P3.24

**3.25** A portion ( $0 \leq x \leq L$ ) of a Si sample, uniformly doped with  $N_D = 10^{15}/\text{cm}^3$  donors and maintained at room temperature, is subject to a *steady state* perturbation such that

$$\begin{aligned} n &\simeq N_D \\ p &= n_i(1 - x/L) + n_i^2/N_D \quad \dots 0 \leq x \leq L \end{aligned}$$

Since  $n \simeq N_D$ , it is reasonable to assume  $\mathcal{E} \simeq 0$  in the  $0 \leq x \leq L$  region. Given  $\mathcal{E} \simeq 0$ , sketch the energy band diagram for the perturbed region specifically including  $E_c$ ,  $E_i$ ,  $E_v$ ,  $F_N$ , and  $F_P$  on your diagram.

- **3.26 (a)** Write a MATLAB (computer) program that automatically plots  $F_N$  and  $F_P$  versus  $x$  for the experimental situation first considered in Sample Problem No. 2 and further analyzed in Exercise 3.5. Assume a Si sample maintained at 300 K with  $\tau_p = 10^{-6}$  sec. In writing the program, take  $N_D$  and  $\Delta p_{n0}$  to be input variables. Scale the  $x$ -axis in  $L_p$  units with  $(x/L_p)_{\max} = \ln[100\Delta p_{n0}/p_0]$ . Scale the  $y$ -axis in  $kT$  units plotting  $(F_N - E_v)/kT$  and  $(F_P - E_v)/kT$ . Set  $y_{\min} = -5$  and  $y_{\max} = 45$ . Identify the positions of  $E_v$  and  $E_c$  on your plot.
- (b) Run your program with  $N_D = 10^{15}/\text{cm}^3$  and  $\Delta p_{n0} = 10^{10}/\text{cm}^3$ . Compare the resultant plot with the corresponding diagram sketched in Exercise 3.5.
- (c) Remembering the assumed low-level injection requires  $\Delta p_{n0} \ll N_D$ , try running your program using different values of  $N_D$  and  $\Delta p_{n0}$ .