Chapter 4

Excess Carriers in Semiconductors

OBJECTIVES

1. Understand how photons interact with direct and indirect band gap semiconductors
2. Understand generation–recombination of excess carriers, possibly through trap sites
3. Introduce quasi-Fermi levels in non-equilibrium
4. Calculate diffusion currents from carrier concentration gradients and diffusivity
5. Use the continuity equation to study time dependence of carrier concentrations

Most semiconductor devices operate by the creation of charge carriers in excess of the thermal equilibrium values. These excess carriers can be created by optical excitation or electron bombardment, or as we shall see in Chapter 5, they can be injected across a forward-biased p-n junction. However the excess carriers arise, they can dominate the conduction processes in the semiconductor material. In this chapter we shall investigate the creation of excess carriers by optical absorption and the resulting properties of photoluminescence and photoconductivity. We shall study more closely the mechanism of electron–hole pair recombination and the effects of carrier trapping. Finally, we shall discuss the diffusion of excess carriers due to a carrier gradient, which serves as a basic mechanism of current conduction along with the mechanism of drift in an electric field.

4.1 OPTICAL ABSORPTION[1]

An important technique for measuring the band gap energy of a semiconductor is the absorption of incident photons by the material. In this experiment photons of selected wavelengths are directed at the sample, and relative transmission of the various photons is observed. Since photons with energies greater than the band gap energy are absorbed while photons with energies less than the band gap are transmitted, this experiment gives an accurate measure of the band gap energy.

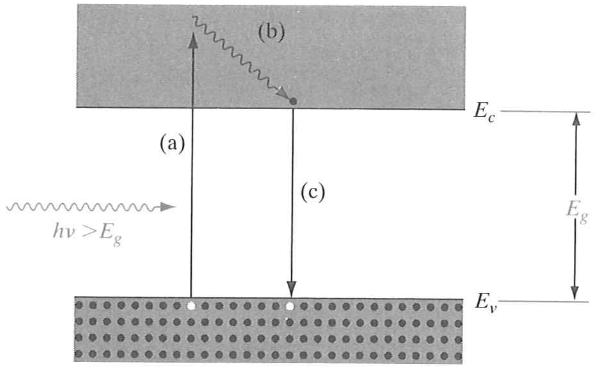


Figure 4–1 Optical absorption of a photon with $h\nu > E_g$: (a) An EHP is created during photon absorption; (b) the excited electron gives up energy to the lattice by scattering events; (c) the electron recombines with a hole in the valence band.

It is apparent that a photon with energy $h\nu \geq E_g$ can be absorbed in a semiconductor (Fig. 4-1). Since the valence band contains many electrons and the conduction band has many empty states into which the electrons may be excited, the probability of photon absorption is high. As Fig. 4-1 indicates, an electron excited to the conduction band by optical absorption may initially have more energy than is common for conduction band electrons (almost all electrons are near E_c unless the sample is very heavily doped). Thus the excited electron loses energy to the lattice in scattering events until its velocity reaches the thermal equilibrium velocity of other conduction band electrons. The electron and hole created by this absorption process are *excess carriers;* since they are out of balance with their environment, they must eventually recombine. While the excess carriers exist in their respective bands, however, they are free to contribute to the conductivity of the material.

A photon with energy less than E_g is unable to excite an electron from the valence band to the conduction band. Thus in a pure semiconductor, there is negligible absorption of photons with $h\nu < E_g$. This explains why some materials are transparent in certain wavelength ranges. We are able to "see through" certain insulators, such as a good NaCl crystal, because a large energy gap containing no electron states exists in the material. If the band gap is about 2 eV wide, only long wavelengths (infrared) and the red part of the visible spectrum are transmitted; on the other hand, a band gap of about 3 eV allows infrared and the entire visible spectrum to be transmitted.

If a beam of photons with $h\nu > E_g$ falls on a semiconductor, there will be some predictable amount of absorption, determined by the properties of the material. We would expect the ratio of transmitted to incident light intensity to depend on the photon wavelength and the thickness of the sample. To calculate this dependence, let us assume that a photon beam of intensity $\mathbf{I}_0$ (photons/cm^2-s) is directed at a sample of thickness l (Fig. 4-2). The beam contains only photons of wavelength λ, selected by a monochromator. As the beam passes through the sample, its intensity at a distance x from the surface can be calculated by considering the probability of absorption within

---

[1] In this context the word "optical" does not necessarily imply that the photons absorbed are in the visible part of the spectrum. Many semiconductors absorb photons in the infrared region, but this is included in the term "optical absorption."

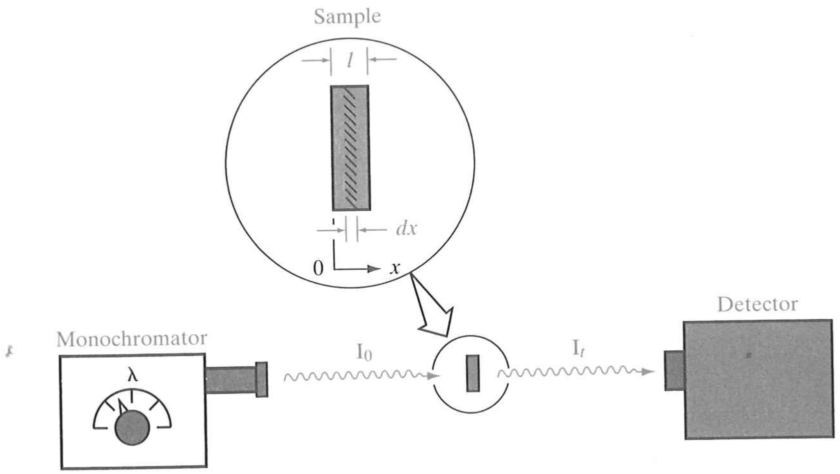


Figure 4–2 Optical absorption experiment.

any increment dx. Since a photon which has survived to x without absorption has no memory of how far it has traveled, its probability of absorption in any dx is constant. Thus the degradation of the intensity $-d\mathbf{I}(x)/dx$ is proportional to the intensity remaining at x:

$$-\frac{d\mathbf{I}(x)}{dx} = \alpha\mathbf{I}(x) \qquad (4\text{-}1)$$

The solution to this equation is

$$\mathbf{I}(x) = \mathbf{I}_0 e^{-\alpha x} \qquad (4\text{-}2)$$

and the intensity of light transmitted through the sample thickness l is

$$\mathbf{I}_t = \mathbf{I}_0 e^{-\alpha l} \qquad (4\text{-}3)$$

The coefficient α is called the *absorption coefficient* and has units of cm^{-1}. This coefficient will of course vary with the photon wavelength and with the material. In a typical plot of α vs. wavelength (Fig. 4-3), there is negligible absorption at long wavelengths ($h\nu$ small) and considerable absorption of photons with energies larger than E_g. According to Eq. (2-2), the relation between photon energy and wavelength is $E = hc/\lambda$. If E is given in electron volts and λ in micrometers, this becomes $E = 1.24/\lambda$.

Figure 4-4 indicates the band gap energies of some of the common semiconductors, relative to the visible, infrared, and ultraviolet portions of the spectrum. We observe that GaAs, Si, Ge, and InSb lie outside the visible region, in the infrared. Other semiconductors, such as GaP and CdS, have band gaps wide enough to pass photons in the visible range. It is important to note here that a semiconductor absorbs photons with energies equal to the band gap, *or larger*. Thus Si absorbs not only band gap light (~1 μm) but also shorter wavelengths, including those in the visible part of the spectrum.

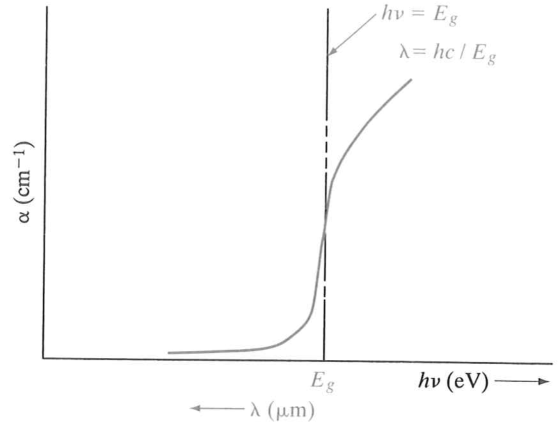


Figure 4–3 Dependence of optical absorption coefficient α for a semiconductor on the wavelength of incident light.

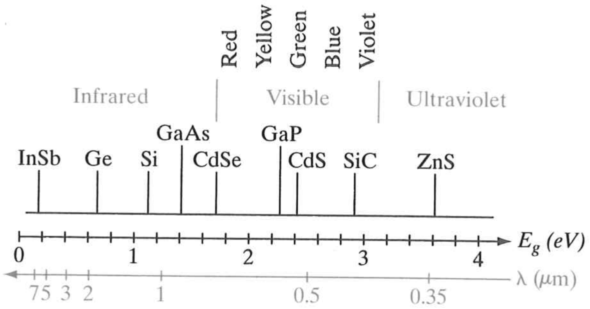


Figure 4–4 Band gaps of some common semiconductors relative to the optical spectrum.

4.2 LUMINESCENCE

When electron–hole pairs are generated in a semiconductor, or when carriers are excited into higher impurity levels from which they fall to their equilibrium states, light can be given off by the material. Many of the semiconductors are well suited for light emission, particularly the compound semiconductors with direct band gaps. The general property of light emission is called *luminescence.*[2] This overall category can be subdivided according to the excitation mechanism: If carriers are excited by photon absorption, the radiation resulting from the recombination of the excited carriers is called *photoluminescence;* if the excited carriers are created by high-energy electron bombardment of the material, the mechanism is called *cathodoluminescence;* if the excitation occurs by the introduction of current into the sample, the resulting luminescence is called *electroluminescence.* Other types of excitation are possible, but these three are the most important for device applications.

---

[2] The emission processes considered here should not be confused with radiation due to incandescence which occurs in heated materials. The various luminescent mechanisms can be considered "cold" processes as compared to the "hot" process of incandescence, which increases with temperature. In fact, most luminescent processes become more efficient as the temperature is lowered.

4.2.1 Photoluminescence

The simplest example of light emission from a semiconductor occurs for direct excitation and recombination of an EHP, as depicted in Fig. 3-5a. If the recombination occurs directly rather than via a defect level, band gap light is given off in the process. For steady state excitation, the recombination of EHPs occurs at the same rate as the generation, and one photon is emitted for each photon absorbed. Direct recombination is a fast process; the mean lifetime of the EHP is usually on the order of 10^{-8} s or less. Thus the emission of photons stops within approximately 10^{-8} s after the excitation is turned off. Such fast luminescent processes are often referred to as *fluorescence*. In some materials, however, emission continues for periods up to seconds or minutes after the excitation is removed. These slow processes are called *phosphorescence*, and the materials are called *phosphors*. An example of a slow process is shown in Fig. 4-5. This material contains a defect level (perhaps due to an impurity) in the band gap which has a strong tendency to temporarily capture (*trap*) electrons from the conduction band. The events depicted in the figure are as follows; (a) An incoming photon with $h\nu_1 > E_g$ is absorbed, creating an EHP; (b) the excited electron gives up energy to the lattice by scattering until it nears the bottom of the conduction band; (c) the electron is *trapped* by the impurity level E_t and remains trapped until it can be thermally reexcited to the conduction band (d); (e) finally direct recombination occurs as the electron falls to an empty state in the valence band, giving off a photon ($h\nu_2$) of approximately the band gap energy. The delay time between excitation and recombination can be relatively long if the probability of thermal reexcitation from the trap (d) is small. Even longer delay times result if the electron is retrapped several times before recombination. If the trapping probability is greater than the probability of recombination, an electron may make several trips between the trap and the conduction band before recombination finally occurs. In such material the emission of phosphorescent light persists for a relatively long time after the excitation is removed.

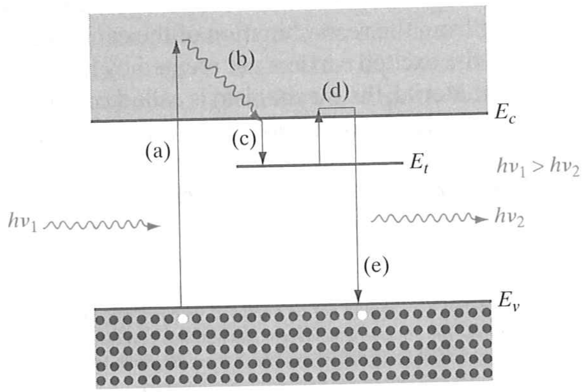


Figure 4–5 Excitation and recombination mechanisms in photoluminescence with a trapping level for electrons.

The color of light emitted by a phosphor such as ZnS depends primarily on the impurities present, since many radiative transitions involve impurity levels within the band gap. This selection of colors is particularly useful in the fabrication of a color television screen.

One of the most common examples of photoluminescence is the fluorescent lamp. Typically such a lamp is composed of a glass tube filled with gas (e.g., a mixture of argon and mercury), with a fluorescent coating on the inside of the tube. When an electric discharge is induced between electrodes in the tube, the excited atoms of the gas emit photons, largely in the visible and ultraviolet regions of the spectrum. This light is absorbed by the luminescent coating, and the visible photons are emitted. The efficiency of such a lamp is considerably better than that of an incandescent bulb, and the wavelength mixture in light given off can be adjusted by proper selection of the fluorescent material.

EXAMPLE 4–1

A 0.46-μm-thick sample of GaAs is illuminated with monochromatic light of $h\nu = 2$ eV. The absorption coefficient α is 5×10^4 cm^{-1}. The power incident on the sample is 10 mW.

(a) Find the total energy absorbed by the sample per second (J/s).
(b) Find the rate of excess thermal energy given up by the electrons to the lattice before recombination (J/s).
(c) Find the number of photons per second given off from recombination events, assuming perfect quantum efficiency.

SOLUTION

(a) From Eq. (4-3),

$$I_t = I_0 e^{-\alpha l} = 10^{-2} \exp(-5 \times 10^4 \times 0.46 \times 10^{-4})$$
$$= 10^{-2} e^{-2.3} = 10^{-3} \text{ W}$$

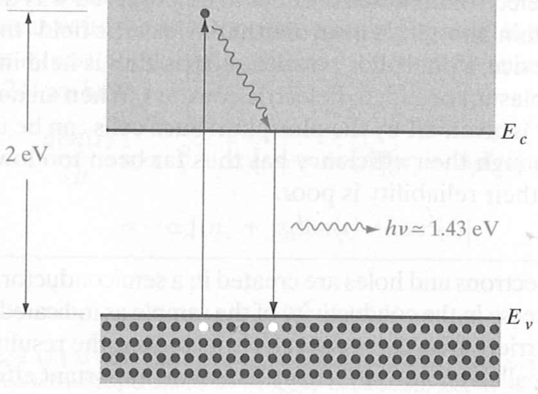


Figure 4–6 Excitation and band-to-band recombination leading to photoluminescence.

Thus the absorbed power is

$$10 - 1 = 9 \text{ mW} = 9 \times 10^{-3} \text{ J/s}$$

(b) The fraction of each photon energy unit which is converted to heat is

$$\frac{2 - 1.43}{2} = 0.285$$

Thus the amount of energy converted to heat per second is

$$0.285 \times 9 \times 10^{-3} = 2.57 \times 10^{-3} \text{ J/s}$$

(c) Assuming one emitted photon for each photon absorbed (perfect quantum efficiency), we have

$$\frac{9 \times 10^{-3} \text{ J/s}}{1.6 \times 10^{-19} \text{ J/eV} \times 2 \text{ eV/photon}} = 2.81 \times 10^{16} \text{ photons/s}$$

Alternative solution: Recombination radiation accounts for 9 - 2.57 = 6.43 mW at 1.43 eV/photon.

$$\frac{6.43 \times 10^{-3}}{1.6 \times 10^{-19} \times 1.43} = 2.81 \times 10^{16} \text{ photons/s}$$

4.2.2 Electroluminescence

There are many ways by which electrical energy can be used to generate photon emission in a solid. In LEDs an electric current causes the injection of minority carriers into regions of the crystal where they can recombine with majority carriers, resulting in the emission of recombination radiation. This important effect (*injection electroluminescence*) will be discussed in Chapter 8 in terms of p-n junction theory.

The first electroluminescent effect to be observed was the emission of photons by certain phosphors in an alternating electric field (the Destriau effect). In this device, a phosphor powder such as ZnS is held in a binder material (often a plastic) of a high dielectric constant. When an a-c electric field is applied, light is given off by the phosphor. Such cells can be useful as lighting panels, although their efficiency has thus far been too low for most applications and their reliability is poor.

4.3 CARRIER LIFETIME AND PHOTOCONDUCTIVITY

When excess electrons and holes are created in a semiconductor, there is a corresponding increase in the conductivity of the sample as indicated by Eq. (3-43). If the excess carriers arise from optical luminescence, the resulting increase in conductivity is called *photoconductivity.* This is an important effect, with useful applications in the analysis of semiconductor materials and in the operation of several types of devices. In this section we shall examine the mechanisms by which excess electrons and holes recombine and apply the recombination kinetics to the analysis of photoconductive devices. However, the importance of recombination is not limited to cases in which the excess carriers are created optically. In fact, virtually every semiconductor device depends in some way on the recombination of excess electrons and holes. Therefore, the concepts developed in this section will be used extensively in the analyses of diodes, transistors, lasers, and other devices in later chapters.

4.3.1 Direct Recombination of Electrons and Holes

It was pointed out in Section 3.1.4 that electrons in the conduction band of a semiconductor may make transitions to the valence band (i.e., recombine with holes in the valence band) either directly or indirectly. In direct recombination, an excess population of electrons and holes decays by electrons falling from the conduction band to empty states (holes) in the valence band. Energy lost by an electron in making the transition is given up as a photon. Direct recombination occurs *spontaneously;* that is, the probability that an electron and a hole will recombine is constant in time. As in the case of carrier scattering, this constant probability leads us to expect an exponential solution for the decay of the excess carriers. In this case the rate of decay of electrons at any time t is proportional to the number of electrons remaining at t and the number of holes, with some constant of proportionality for recombination, α_r. The *net* rate of change in the conduction band electron concentration is the thermal generation rate $\alpha_r n_i^2$ from Eq. (3-7) minus the recombination rate

$$\frac{dn(t)}{dt} = \alpha_r n_i^2 - \alpha_r n(t)p(t) \qquad (4\text{-}4)$$

Let us assume the excess electron-hole population is created at $t = 0$, for example by a short flash of light, and the initial excess electron and hole concentrations Δn and Δp are equal.[3] Then as the electrons and holes recombine in pairs, the instantaneous concentrations of excess carriers $\delta n(t)$ and $\delta p(t)$ are also equal. Thus we can write the total concentrations of Eq. (4-4) in terms of the equilibrium values n_0 and p_0 and the excess carrier concentrations $\delta n(t) = \delta p(t)$. Using Eq. (3-24) we have

$$\begin{aligned} \frac{d\delta n(t)}{dt} &= \alpha_r n_i^2 - \alpha_r[n_0 + \delta n(t)][p_0 + \delta p(t)] \\ &= -\alpha_r[(n_0 + p_0)\delta n(t) + \delta n^2(t)] \qquad (4\text{-}5) \end{aligned}$$

---

[3] We will use $\delta n(t)$ and $\delta p(t)$ to mean instantaneous excess carrier concentrations, and Δn, Δp for their values at $t = 0$. Later we will use similar symbolism for spatial distributions, such as $\delta n(x)$ and $\Delta n(x = 0)$.

This nonlinear equation would be difficult to solve in its present form. Fortunately, it can be simplified for the case of low-level injection. If the excess carrier concentrations are small, we can neglect the δn^2 term. Furthermore, if the material is extrinsic, we can usually neglect the term representing the equilibrium minority carriers. For example, if the material is p-type ($p_0 \gg n_0$), Eq. (4-5) becomes

$$\frac{d\delta n(t)}{dt} = -\alpha_r p_0 \delta n(t) \tag{4-6}$$

The solution to this equation is an exponential decay from the original excess carrier concentration Δn:

$$\delta n(t) = \Delta n e^{-\alpha_r p_0 t} = \Delta n e^{-t/\tau_n} \tag{4-7}$$

Excess electrons in a p-type semiconductor recombine with a decay constant $\tau_n = (\alpha_r p_0)^{-1}$, called the *recombination lifetime.* Since the calculation is made in terms of the minority carriers, τ_n is often called the *minority carrier lifetime.* The decay of excess holes in n-type material occurs with $\tau_p = (\alpha_r n_0)^{-1}$. In the case of direct recombination, the excess majority carriers decay at exactly the same rate as the minority carriers.

There is a large percentage change in the minority carrier electron concentration in Example 4-2 and a small percentage change in the majority hole concentration. Basically, the approximations of extrinsic material and low-level injection allow us to represent $n(t)$ in Eq. (4-4) by the excess concentration $\delta n(t)$ and $p(t)$ by the equilibrium value p_0. Figure 4-7 indicates that this is a good approximation for the example. A more general expression for the carrier lifetime is

$$\tau_n = \frac{1}{\alpha_r(n_0 + p_0)} \tag{4-8}$$

This expression is valid for n- or p-type material if the injection level is low.

EXAMPLE 4-2 A numerical example may be helpful in visualizing the approximations made in the analysis of direct recombination. Let us assume a sample of GaAs is doped with 10^{15} acceptors/cm^3. The intrinsic carrier concentration of GaAs is approximately 10^6 cm^{-3}; thus the minority electron concentration is $n_0 = n_i^2/p_0 = 10^{-3}$ cm^{-3}. Certainly the approximation of $p_0 \gg n_0$ is valid in this case. Now if 10^{14} EHP/cm^3 are created at $t = 0$, we can calculate the decay of these carriers in time. The approximation of $\delta n \ll p_0$ is reasonable, as Fig. 4-7 indicates. This figure shows the decay in time of the excess populations for a carrier recombination lifetime of $\tau_n = \tau_p = 10^{-8}$ s.

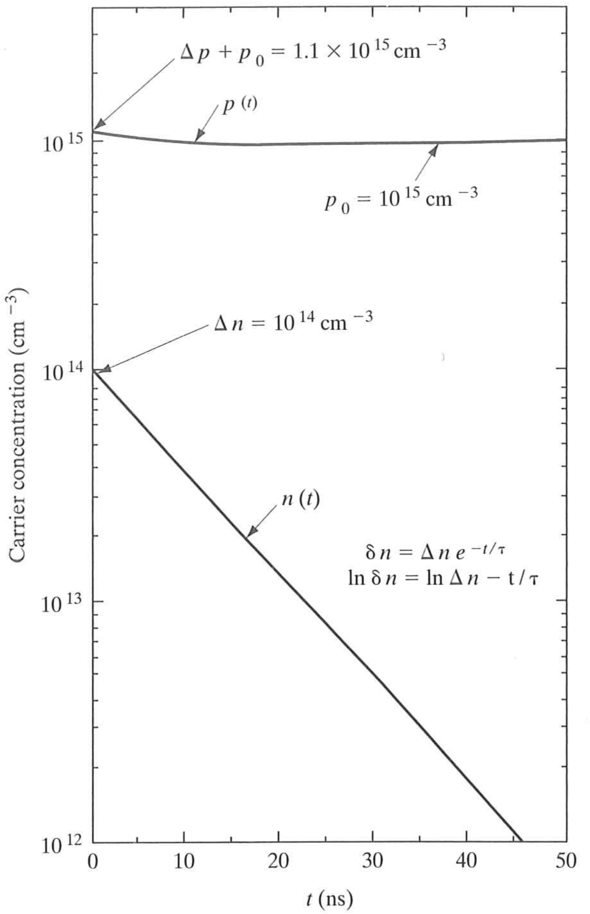


Figure 4–7 Decay of excess electrons and holes by recombination, for $\Delta n = \Delta p = 0.1 p_0$, with n_0 negligible, and $\tau = 10$ ns (Example 4–2). The exponential decay of $\delta n(t)$ is linear on this semilogarithmic graph.

4.3.2 Indirect Recombination; Trapping

In column IV semiconductors and in certain compounds, the probability of direct electron–hole recombination is very small (Appendix III). There is some band gap light given off by materials such as Si and Ge during recombination, but this radiation is very weak and may be detected only by sensitive equipment. The vast majority of the recombination events in indirect materials occur via *recombination levels* within the band gap, and the resulting energy loss by recombining electrons is usually given up to the lattice as heat rather than by the emission of photons. Any impurity or lattice defect can serve as a recombination center if it is capable of receiving a carrier of one type and subsequently capturing the opposite type of carrier, thereby annihilating the pair. For example, Fig. 4-8 illustrates a recombination level E_r which is below E_F at equilibrium and therefore is substantially filled with electrons. When excess electrons and holes are created in this material, each EHP recombines at E_r in two steps: (a) hole capture and (b) electron capture.

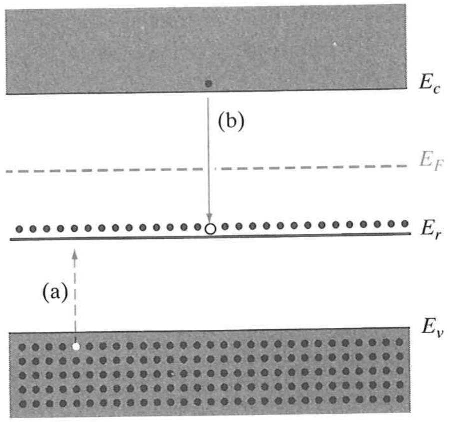


Figure 4–8
Capture processes at a recombination level: (a) hole capture at a filled recombination center; (b) electron capture at an empty center.

Since the recombination centers in Fig. 4–8 are filled at equilibrium, the first event in the recombination process is hole capture. It is important to note that this event is equivalent to an electron at E_r falling to the valence band, leaving behind an empty state in the recombination level. Thus in hole capture, energy is *given up* as heat to the lattice. Similarly, energy is given up when a conduction band electron subsequently falls to the empty state in E_r. When both of these events have occurred, the recombination center is back to its original state (filled with an electron), but an EHP is missing. Thus one EHP recombination has taken place, and the center is ready to participate in another recombination event by capturing a hole.

The carrier lifetime resulting from indirect recombination is somewhat more complicated than is the case for direct recombination, since it is necessary to account for unequal times required for capturing each type of carrier. In particular, recombination is often delayed by the tendency for a captured carrier to be thermally reexcited to its original band before capture of the opposite type of carrier can occur (Section 4.2.1). For example, if electron capture (b) does not follow immediately after hole capture (a) in Fig. 4–8, the hole may be thermally reexcited to the valence band. Energy is required for this process, which is equivalent to a valence band electron being raised to the empty state in the recombination level. This process delays the recombination, since the hole must be captured again before recombination can be completed.

When a carrier is trapped temporarily at a center and then is reexcited without recombination taking place, the process is often called *temporary trapping.* Although the nomenclature varies somewhat, it is common to refer to an impurity or defect center as a *trapping center* (or simply *trap*) if, after capture of one type of carrier, the most probable next event is reexcitation. If the most probable next event is capture of the opposite type of carrier, the center is predominately a recombination center. The recombination can be slow or fast, depending on the average time the first carrier is held before the second carrier is captured. In general, trapping levels located deep in the band gap are slower in releasing trapped carriers than are the levels located near one of the bands. This results from the fact that more energy is required, for example, to reexcite a trapped electron from a center near the middle of the gap to the conduction band than is required to reexcite an electron from a level closer to the conduction band.

As an example of impurity levels in semiconductors, Fig. 4–9[4] shows the energy level positions of various impurities in Si. In this diagram a superscript indicates whether the impurity is positive (donor) or negative (acceptor) when ionized. Some impurities introduce multiple levels in the band gap; for example, Zn introduces a level (Zn^-) located 0.31 eV above the valence band and a second level ($\mathrm{Zn}^=$) near the middle of the gap. Each Zn impurity atom is capable of accepting two electrons from the semiconductor, one in the lower level and then one in the upper level.

The effects of recombination and trapping can be measured by a *photoconductive decay* experiment. As Fig. 4–7 shows, a population of excess electrons and holes disappears with a decay constant characteristic of the particular recombination process. The conductivity of the sample during the decay is

$$\sigma(t) = q[n(t)\mu_n + p(t)\mu_p] \tag{4-9}$$

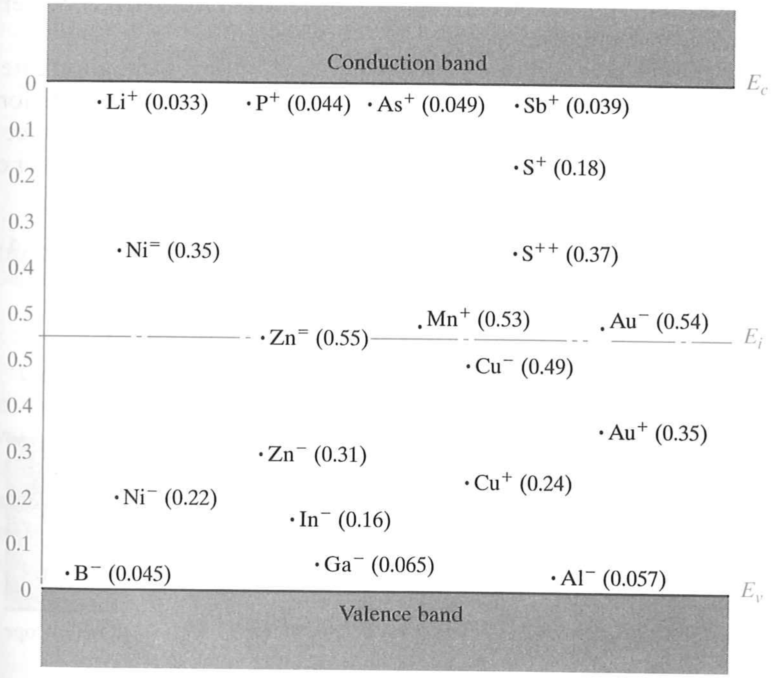


Figure 4–9
Energy levels of impurities in Si. The energies are measured from the nearest band edge (E_v or E_c); donor levels are designated by a plus sign and acceptors by a minus sign.

---

[4] References: S. M. Sze and J. C. Irvin, "Resistivity, Mobility, and Impurity Levels in GaAs, Ge and Si at 300 K," *Solid State Electronics,* vol. 11, pp. 599–602 (June 1968); E. Schibli and A. G. Milnes, "Deep Impurities in Silicon," *Materials Science and Engineering,* vol. 2, pp. 173–180 (1967).

Therefore, the time dependence of the carrier concentrations can be monitored by recording the sample resistance as a function of time. A typical experimental arrangement is shown schematically in Fig. 4–10. A source of short pulses of light is required, along with an oscilloscope for displaying the sample voltage as the resistance varies. Microsecond light pulses can be obtained by periodically discharging a capacitor through a flash tube containing a gas such as xenon. For shorter pulses, special techniques such as the use of a pulsed laser must be used.

4.3.3 Steady State Carrier Generation; Quasi-Fermi Levels

In the previous discussion we emphasized the transient decay of an excess EHP population. However, the various recombination mechanisms are also important in a sample at thermal equilibrium or with a steady state EHP generation–recombination balance.[5] For example, a semiconductor at equilibrium experiences thermal generation of EHPs at a rate $g(T) = g_i$ described by Eq. (3–7). This generation is balanced by the recombination rate so that the equilibrium concentrations of carriers n_0 and p_0 are maintained:

$$g(T) = \alpha_r n_i^2 = \alpha_r n_0 p_0 \tag{4–10}$$

This equilibrium rate balance can include generation from defect centers as well as band-to-band generation.

If a steady light is shone on the sample, an optical generation rate g_{op} will be added to the thermal generation, and the carrier concentrations n and p will increase to new steady state values. We can write the balance between generation and recombination in terms of the equilibrium carrier concentrations and the departures from equilibrium δn and δp:

$$g(T) + g_{\text{op}} = \alpha_r np = \alpha_r(n_0 + \delta n)(p_0 + \delta p) \tag{4–11}$$

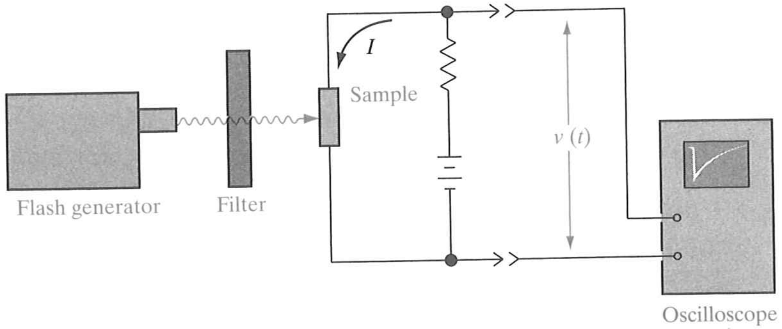


Figure 4–10 Experimental arrangement for photoconductive decay measurements, and a typical oscilloscope trace.

For steady state recombination and no trapping, $\delta n = \delta p$; thus Eq. (4–11) becomes

$$g(T) + g_{\text{op}} = \alpha_r n_0 p_0 + \alpha_r[(n_0 + p_0)\delta n + \delta n^2] \tag{4–12}$$

The term $\alpha_r n_0 p_0$ is just equal to the thermal generation rate $g(T)$. Thus, neglecting the δn^2 term for low-level excitation, we can rewrite Eq. (4–12) as

$$g_{\text{op}} = \alpha_r(n_0 + p_0)\delta n = \frac{\delta n}{\tau_n} \tag{4–13}$$

The excess carrier concentration can be written as

$$\delta n = \delta p = g_{\text{op}}\tau_n \tag{4–14}$$

More general expressions are given in Eq. (4–16), which allow for the case $\tau_p \neq \tau_n$, when trapping is present.

EXAMPLE 4–3

As a numerical example, let us assume that 10^{13} EHP/cm^3 are created optically every microsecond in a Si sample with $n_0 = 10^{14}$ cm^{-3} and $\tau_n = \tau_p = 2$ μsec. The steady state excess electron (or hole) concentration is then 2×10^{13} cm^{-3} from Eq. (4–14). While the percentage change in the majority electron concentration is small, the minority carrier concentration changes from

$$p_0 = n_i^2/n_0 = (2.25 \times 10^{20})/10^{14} = 2.25 \times 10^6 \text{ cm}^{-3} \qquad (\textit{equilibrium})$$

to

$$p = 2 \times 10^{13} \text{ cm}^{-3} \qquad (\textit{steady state})$$

Note that the equilibrium equation $n_0 p_0 = n_i^2$ cannot be used with the subscripts removed; that is, $np \neq n_i^2$ when excess carriers are present.

It is often desirable to refer to the steady state electron and hole concentrations in terms of Fermi levels, which can be included in band diagrams for various devices. The Fermi level E_F used in Eq. (3–25) is meaningful only when no excess carriers are present. However, we can write expressions for the steady state concentrations in the same *form* as the equilibrium expressions by defining separate *quasi-Fermi levels* F_n and F_p for electrons and holes. The resulting carrier concentration equations

$$\boxed{\begin{aligned} n &= n_i e^{(F_n - E_i)/kT} \\ p &= n_i e^{(E_i - F_p)/kT} \end{aligned}} \tag{4–15}$$

can be considered as defining relations for the quasi-Fermi levels.[6]

---

[5] The term *equilibrium* refers to a condition of no external excitation except for temperature, and no net motion of charge (e.g., a sample at a constant temperature, in the dark, with no fields applied). *Steady state* refers to a nonequilibrium condition in which all processes are constant and are balanced by opposing processes (e.g., a sample with a constant current or a constant optical generation of EHPs just balanced

[6] In some texts the quasi-Fermi level is called *IMREF*, which is Fermi spelled backward.

EXAMPLE 4–4 In Example 4–3, the steady state electron concentration is

$$n = n_0 + \delta n = 1.2 \times 10^{14} = (1.5 \times 10^{10})e^{(F_n - E_i)/0.0259}$$

where $kT \simeq 0.0259$ eV at room temperature. Thus the electron quasi-Fermi level position $F_n - E_i$ is found from

$$F_n - E_i = 0.0259 \ln(8 \times 10^3) = 0.233 \text{ eV}$$

and F_n lies 0.233 eV above the intrinsic level. By a similar calculation, the hole quasi-Fermi level lies 0.186 eV below E_i (Fig. 4–11). In this example, the equilibrium Fermi level is $0.0259 \ln(6.67 \times 10^3) = 0.228$ eV above the intrinsic level.

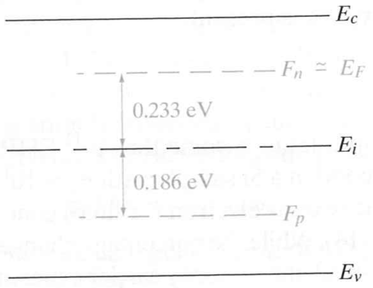


Figure 4–11 Quasi-Fermi levels F_n and F_p for a Si sample with $n_0 = 10^{14}$ cm^{-3}, $\tau_p = 2$ μs, and $g_{op} = 10^{19}$ EHP/cm^3-s (Example 4–4).

The quasi-Fermi levels of Fig. 4–11 illustrate dramatically the deviation from equilibrium caused by the optical excitation; the steady state F_n is only slightly above the equilibrium E_F, whereas F_p is greatly displaced below E_F. From the figure it is obvious that the excitation causes a large percentage change in minority carrier hole concentration and a relatively small change in the electron concentration.

In summary, the quasi-Fermi levels F_n and F_p are the steady state analogues of the equilibrium Fermi level E_F. When excess carriers are present, the deviations of F_n and F_p from E_F indicate how far the electron and hole populations are from the equilibrium values n_0 and p_0. A given concentration of excess EHPs causes a large shift in the minority carrier quasi-Fermi level compared with that for the majority carriers. The separation of the quasi-Fermi levels $F_n - F_p$ is a direct measure of the deviation from equilibrium (at equilibrium $F_n = F_p = E_F$). The concept of quasi-Fermi levels is very useful in visualizing minority and majority carrier concentrations in devices where these quantities vary with position.

4.3.4 Photoconductive Devices

There are a number of applications for devices which change their resistance when exposed to light. For example, such light detectors can be used in the home to control automatic night lights which turn on at dusk and turn off at dawn. They can also be used to measure illumination levels, as in exposure meters for cameras. Many systems include a light beam aimed at the photoconductor, which signals the presence of an object between the source and detector. Such systems are useful in moving-object counters, burglar alarms, and many other applications. Detectors are used in optical signaling systems in which information is transmitted by a light beam and is received at a photoconductive cell.

Considerations in choosing a photoconductor for a given application include the sensitive wavelength range, time response, and optical sensitivity of the material. In general, semiconductors are most sensitive to photons with energies equal to the band gap or slightly more energetic than band gap. Less energetic photons are not absorbed, and photons with $h\nu \gg E_g$ are absorbed at the surface and contribute little to the bulk conductivity. Therefore, the table of band gaps (Appendix III) indicates the photon energies to which most semiconductor photodetectors respond. For example, CdS ($E_g = 2.42$ eV) is commonly used as a photoconductor in the visible range, and narrow-gap materials such as Ge (0.67 eV) and InSb (0.18 eV) are useful in the infrared portion of the spectrum. Some photoconductors respond to excitations of carriers from impurity levels within the band gap and therefore are sensitive to photons of less than band gap energy.

The optical sensitivity of a photoconductor can be evaluated by examining the steady state excess carrier concentrations generated by an optical generation rate g_{op}. If the mean time each carrier spends in its respective band before capture is τ_n and τ_p, we have

$$\delta n = \tau_n g_{op} \quad \text{and} \quad \delta p = \tau_p g_{op} \tag{4–16}$$

and the photoconductivity change is

$$\Delta\sigma = q g_{op}(\tau_n \mu_n + \tau_p \mu_p) \tag{4–17}$$

For simple recombination, τ_n and τ_p will be equal. If trapping is present, however, one of the carriers may spend little time in its band before being trapped. From Eq. (4–17) it is obvious that for maximum photoconductive response, we want high mobilities and long lifetimes. Some semiconductors are especially good candidates for photoconductive devices because of their high mobility; for example, InSb has an electron mobility of about 10^5 cm^2/V-s and therefore is used as a sensitive infrared detector in many applications.

The time response of a photoconductive cell is limited by the recombination times, the degree of carrier trapping, and the time required for carriers to drift through the device in an electric field. Often these properties can be adjusted by proper choice of material and device geometry, but in some cases improvements in response time are made at the expense of sensitivity. For example, the drift time can be reduced by making the device short, but this substantially reduces the responsive area of the device. In addition, it is often desirable that the device have a large dark resistance, and for this reason, shortening the length may not be practical. There is usually a compromise between sensitivity, response time, dark resistance, and other requirements in choosing a device for a particular application.

4.4 DIFFUSION OF CARRIERS

When excess carriers are created nonuniformly in a semiconductor, the electron and hole concentrations vary with position in the sample. Any such spatial variation (*gradient*) in n and p calls for a net motion of the carriers from regions of high carrier concentration to regions of low carrier concentration. This type of motion is called *diffusion* and represents an important charge transport process in semiconductors. The two basic processes of current conduction are diffusion due to a carrier gradient and drift in an electric field.

4.4.1 Diffusion Processes

When a bottle of perfume is opened in one corner of a closed room, the scent is soon detected throughout the room. If there is no convection or other net motion of air, the scent spreads by diffusion. The diffusion is the natural result of the *random motion* of the individual molecules. Consider, for example, a volume of arbitrary shape with scented air molecules inside and unscented molecules outside the volume. All the molecules undergo random thermal motion and collisions with other molecules. Thus each molecule moves in an arbitrary direction until it collides with another air molecule, after which it moves in a new direction. If the motion is truly random, a molecule at the edge of the volume has equal probabilities of moving into or out of the volume on its next step (assuming the curvature of the surface is negligible on the molecular scale). Therefore, after a mean free time $\bar{t}$, half the molecules at the edge will have moved into the volume and half will have moved out of the volume. The net effect is that the volume containing scented molecules has increased. This process will continue until the molecules are uniformly distributed in the room. Only then will a given volume gain as many molecules as it loses in a given time. In other words, net diffusion will continue as long as gradients exist in the distribution of scented molecules.

Carriers in a semiconductor diffuse in a carrier gradient by random thermal motion and scattering from the lattice and impurities. For example, a pulse of excess electrons injected at $x = 0$ at time $t = 0$ will spread out in time as shown in Fig. 4–12. Initially, the excess electrons are concentrated at $x = 0$; as time passes, however, electrons diffuse to regions of low electron concentration until finally $n(x)$ is constant.

We can calculate the rate at which the electrons diffuse in a one-dimensional problem by considering an arbitrary distribution $n(x)$ such as Fig. 4–13a. Since the mean free path $\bar{l}$ between collisions is a small incremental distance, we can divide x into segments $\bar{l}$ wide, with $n(x)$ evaluated at the center of each segment (Fig. 4–13b).

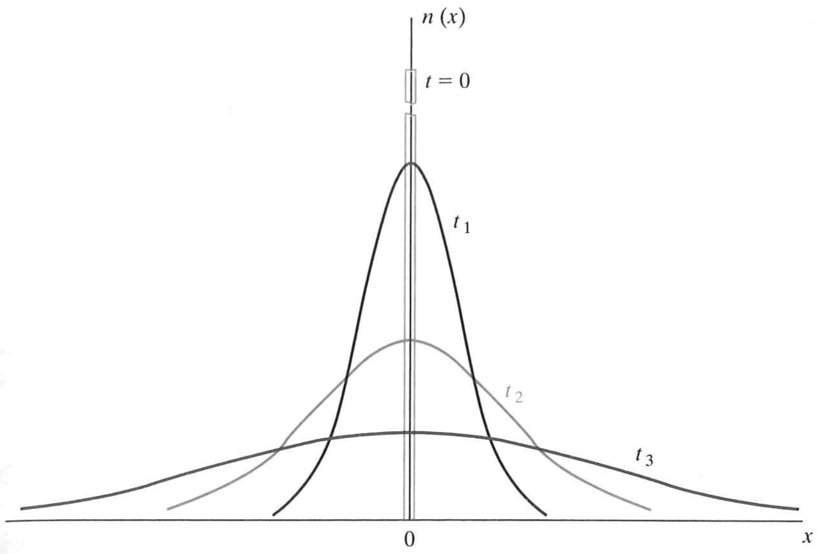
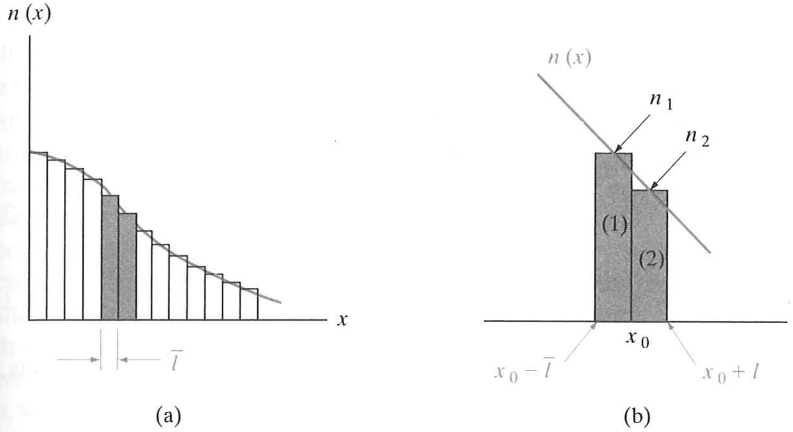


Figure 4–12
Spreading of a pulse of electrons by diffusion.



(a)

(b)

Figure 4–13
An arbitrary electron concentration gradient in one dimension: (a) division of $n(x)$ into segments of length equal to a mean free path for the electrons; (b) expanded view of two of the segments centered at x_0.

The electrons in segment (1) to the left of x_0 in Fig. 4–13b have equal chances of moving left or right, and in a mean free time $\bar{t}$ one-half of them will move into segment (2). The same is true of electrons within one mean free path of x_0 to the right; one-half of these electrons will move through x_0 from right to left in a mean free time. Therefore, the *net* number of electrons passing x_0 from left to right in one mean free time is $\frac{1}{2}(n_1\bar{l}A) - \frac{1}{2}(n_2\bar{l}A)$, where the area perpendicular to x is A. The rate of electron flow in the $+x$-direction per unit area (the electron flux density ϕ_n) is given by

$$\phi_n(x_0) = \frac{\bar{l}}{2\bar{t}}(n_1 - n_2) \tag{4-18}$$

Since the mean free path $\bar{l}$ is a small differential length, the difference in electron concentration $(n_1 - n_2)$ can be written as

$$n_1 - n_2 = \frac{n(x) - n(x + \Delta x)}{\Delta x}\bar{l} \tag{4-19}$$

where x is taken at the center of segment (1) and $\Delta x = \bar{l}$. In the limit of small Δx (i.e., small mean free path $\bar{l}$ between scattering collisions), Eq. (4–18) can be written in terms of the carrier gradient $dn(x)/dx$:

$$\phi_n(x) = \frac{\bar{l}^2}{2\bar{t}} \lim_{\Delta x \to 0} \frac{n(x) - n(x + \Delta x)}{\Delta x} = \frac{-\bar{l}^2}{2\bar{t}} \frac{dn(x)}{dx} \tag{4-20}$$

The quantity $\bar{l}^2/2\bar{t}$ is called the *electron diffusion coefficient*[7] D_n, with units cm^2/s. The minus sign in Eq. (4–20) arises from the definition of the derivative; it simply indicates that the net motion of electrons due to diffusion is in the direction of *decreasing* electron concentration. This is the result we expect, since net diffusion occurs from regions of high particle concentration to regions of low particle concentration. By identical arguments, we can show that holes in a hole concentration gradient move with a diffusion coefficient D_p. Thus

$$\phi_n(x) = -D_n \frac{dn(x)}{dx} \tag{4-21a}$$

$$\phi_p(x) = -D_p \frac{dp(x)}{dx} \tag{4-21b}$$

The diffusion current crossing a unit area (the current density) is the particle flux density multiplied by the charge of the carrier:

$$J_n(\text{diff.}) = -(-q)D_n \frac{dn(x)}{dx} = +qD_n \frac{dn(x)}{dx} \tag{4-22a}$$

$$J_p(\text{diff.}) = -(+q)D_p \frac{dp(x)}{dx} = -qD_p \frac{dp(x)}{dx} \tag{4-22b}$$

---

[7] If motion in three dimensions were included, the diffusion would be smaller in the x-direction. Actually, the diffusion coefficient should be calculated from the true energy distributions and scattering mechanisms. Diffusion coefficients are usually determined experimentally for a particular material, as described in Section 4.4.5.

It is important to note that electrons and holes move together in a carrier gradient [Eqs. (4–21)], but the resulting currents are in opposite directions [Eqs. (4–22)] because of the opposite charge of electrons and holes.

4.4.2 Diffusion and Drift of Carriers; Built-in Fields

If an electric field is present in addition to the carrier gradient, the current densities will each have a drift component and a diffusion component

$$\boxed{\begin{aligned} J_n(x) &= \underbrace{q\mu_n n(x)\mathscr{E}(x)}_{\text{drift}} + \underbrace{qD_n \frac{dn(x)}{dx}}_{\text{diffusion}} && \text{(4-23a)} \\ J_p(x) &= q\mu_p p(x)\mathscr{E}(x) - qD_p \frac{dp(x)}{dx} && \text{(4-23b)} \end{aligned}}$$

and the total current density is the sum of the contributions due to electrons and holes:

$$J(x) = J_n(x) + J_p(x) \tag{4-24}$$

We can best visualize the relation between the particle flow and the current of Eqs. (4–23) by considering a diagram such as shown in Fig. 4–14. In this figure an electric field is assumed to be in the x-direction, along with carrier distributions $n(x)$ and $p(x)$ which decrease with increasing x. Thus the derivatives in Eqs. (4–21) are negative, and diffusion takes place in the $+x$-direction. The resulting electron and hole diffusion currents [J_n (diff.) and J_p (diff.)] are in opposite directions, according to Eqs. (4–22). Holes drift in the direction of the electric field [ϕ_p (drift)], whereas electrons drift in the opposite direction because of their negative charge. The resulting drift current is in the $+x$-direction in each case. Note that the drift and diffusion components of the current are additive for holes when the field is in the direction of decreasing hole concentration, whereas the two components are subtractive for electrons under similar conditions. The total current may be due primarily to the flow of electrons or holes, depending on the relative concentrations and the relative magnitudes and directions of electric field and carrier gradients.

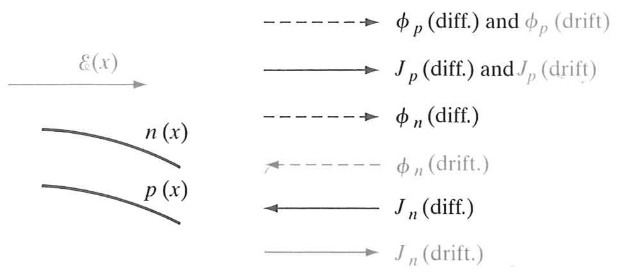


Figure 4–14
Drift and diffusion directions for electrons and holes in a carrier gradient and an electric field. Particle flow directions are indicated by dashed arrows, and the resulting currents are indicated by solid arrows.