# 29
# *Inhomogeneous Semiconductors*

The Semiclassical Treatment of Inhomogeneous Solids

Fields and Carrier Densities in the Equilibrium *p-n* Junction

Elementary Picture of Rectification by a *p-n* Junction

Drift and Diffusion Currents

Collision and Recombination Times

Fields, Carrier Densities, and Currents in the Nonequilibrium *p-n* Junction

As used by music lovers, the term "solid state physics" refers only to the subject of inhomogeneous semiconductors, and it would be more accurate if it were this latter phrase that festooned the brows of countless tuners and amplifiers. The prevailing usage reflects the fact that modern solid state physics has had its most dramatic and extensive technological consequences through the electronic properties of semiconducting devices. These devices use semiconducting crystals in which the concentrations of donor and acceptor impurities have been made nonuniform in a carefully controlled manner. We shall not attempt to survey here the great variety of semiconducting devices, but will only describe the broad physical principles that underlie their operation. These principles come into play in determining how the densities and currents of electrons and holes are distributed in an inhomogeneous semiconductor, both in the absence and in the presence of an applied electrostatic potential.

The inhomogeneous semiconductors of interest are, ideally, single crystals in which the local concentration of donor and acceptor impurities varies with position. One way to make such crystals is to vary the concentration of impurities in the "melt" as the growing crystal is slowly extracted, thus producing a variation in impurity concentration along one spatial direction. Delicate methods of fabrication are needed because it is generally important, for efficient operation, that there be no great increase in electronic scattering associated with the variation in impurity concentration.

We shall illustrate the physics of inhomogeneous semiconductors by considering the simplest example, the *p-n* junction. This is a semiconducting crystal in which the impurity concentration varies only along a given direction (taken to be the $x$-axis) and only in a small region (taken to be around $x = 0$). For negative $x$ the crystal has a preponderance of acceptor impurities (i.e., it is *p*-type) while for positive $x$ it has a preponderance of donor impurities (i.e., it is *n*-type) (Figure 29.1). The manner in which the densities of donors and acceptors $N_d(x)$ and $N_a(x)$ vary with position is

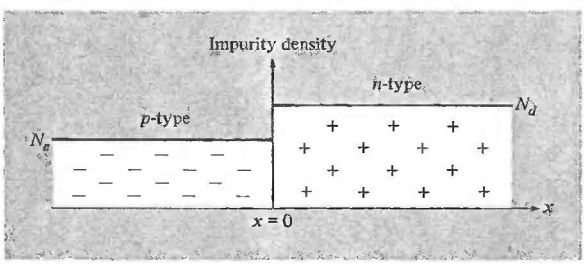


**Figure 29.1**
The impurity densities along a *p-n* junction in the case of an "abrupt junction," for which donor impurities dominate at positive $x$, and acceptor impurities at negative $x$. The donors are represented by (+) to indicate their charge when ionized, and the acceptors by (−). For a junction to be abrupt, the region about $x = 0$ where the impurity concentrations change must be narrow compared with the "depletion layer" in which the carrier densities are nonuniform. (Typical plots of the carrier densities are superimposed on this figure in Figure 29.3.)

called the "doping profile." The term "junction" is used to refer both to the device as a whole and, more specifically, to the transition region about $x = 0$ in which the doping profile is nonuniform.

As we shall see below, the nonuniformity in impurity concentrations induces a nonuniformity in the densities $n_c(x)$ and $p_v(x)$ of conduction band electrons and valence band holes, which in turn gives rise to a potential $\phi(x)$. The region in which these carrier densities are nonuniform is known as the "depletion layer" (or "space-charge region"). The depletion layer can extend for a range of about $10^2$ to $10^4$ Å around the (generally more narrow) transition region in which the doping profile varies, as we shall see below. Within the depletion layer, except near its boundaries, the total density of carriers is very much less than it is in the homogeneous regions farther away from the transition region. The existence of a depletion layer is one of the crucial properties of the $p$-$n$ junction. One of our main concerns will be to explain why such a layer is induced by the variation in impurity concentrations, and how its structure changes with the application of an external potential $V$.

For simplicity we shall consider here only "abrupt junctions," in which the transition region is so sharp that variation in impurity concentrations[1] can be represented by a single discontinuous change at $x = 0$:

$$N_d(x) = \begin{cases} N_{d^+}, & x > 0 \\ 0, & x < 0 \end{cases},$$
$$N_a(x) = \begin{cases} 0, & x > 0 \\ N_{a^-}, & x < 0 \end{cases}. \tag{29.1}$$

Abrupt junctions are not only conceptually the simplest, but also the type of greatest practical interest. How sharp the actual transition region must be made for (29.1) to give a reasonable model of a physical junction will emerge in the analysis below. We shall find that a junction may be regarded as abrupt if the transition region in the actual doping profile is small in extent compared with the depletion layer. In most cases this permits the transition region to extend for 100 Å or more. A junction that cannot be treated as abrupt is called a "graded junction."

## THE SEMICLASSICAL MODEL

To calculate the response of an inhomogeneous semiconductor to an applied electrostatic potential, or even to compute the distribution of electric charge in the absence of an applied potential, one almost always uses the semiclassical model of Chapter 12. When a potential $\phi(x)$ is superimposed on the periodic potential of the crystal, the semiclassical model treats the electrons in the $n$th band as classical particles (i.e., as wave packets) governed by the Hamiltonian

$$H_n = \varepsilon_n\left(\frac{\mathbf{p}}{\hbar}\right) - e\phi(x). \tag{29.2}$$

[1] It is not essential that there be only donor impurities in the $n$-type region and only acceptor impurities in the $p$-type region. It suffices for each impurity type to be the dominant one in its own region. In what follows $N_d$ may be viewed as the excess density of donors over acceptors and $N_a$ as the excess density of acceptors over donors.

Such a treatment is valid provided that the potential $\phi(x)$ varies sufficiently slowly. How slow this variation must be is, in general, a very difficult question to answer. At the very least, one requires that the change in electrostatic energy $e\Delta\phi$ over a distance of the order of the lattice constant be small compared with the band gap $E_g$, but the condition may well be even more stringent than this.[2] In the case of the *p-n* junction the potential $\phi$ has almost all of its spatial variation within the depletion layer. There, as we shall see, the energy $e\phi$ changes by about $E_g$, over a distance that is typically a few hundred angstroms or more (so that the field in the depletion layer can be as large as $10^6$ volts per meter). Although this satisfies the minimum necessary condition for the validity of the semiclassical model (the change in $e\phi$ over a lattice constant is no more than a fraction of a percent of $E_g$), the variation is strong enough that one cannot exclude the possibility that the semiclassical description may break down in the depletion layer. Thus one should bear in mind the possibility that the field in the depletion layer may be strong enough to induce tunneling of electrons from valence band to conduction band levels, leading to a conductivity considerably in excess of the semiclassical prediction.

Having issued this warning, however, we shall follow the general practice of assuming the validity of the semiclassical description, so that we may explore its consequences. Before describing the semiclassical theory of the currents that flow in a *p-n* junction in the presence of an applied potential, we first examine the case of the *p-n* junction in thermal equilibrium, in the absence of applied potentials and current flow.

## THE *p-n* JUNCTION IN EQUILIBRIUM

We wish to determine the carrier densities and the electrostatic potential $\phi(x)$ induced by the nonuniform doping. We assume that nondegenerate conditions hold throughout the material, so that the carrier densities at each position $x$ have the "Maxwellian" forms analogous to the densities (28.12) we found in the uniform case. In the nonuniform case, to derive the carrier densities at position $x$ along the junction in the presence of a potential $\phi(x)$, the semiclassical procedure is simply to repeat the analysis for the uniform case, but using the semiclassical one-electron energy (29.2), in which each level is shifted by $-e\phi(x)$. Using the forms (28.3) of the $\varepsilon(\mathbf{k})$ appropriate to levels near the conduction band minimum or valence band maximum, we see that the effect of this is simply to shift the constants $\varepsilon_c$ and $\varepsilon_v$ by $-e\phi(x)$. Thus Eq. (28.12) for the equilibrium carrier densities is generalized to

$$n_c(x) = N_c(T) \exp\left\{-\frac{[\varepsilon_c - e\phi(x) - \mu]}{k_B T}\right\},$$
$$p_v(x) = P_v(T) \exp\left\{-\frac{[\mu - \varepsilon_v + e\phi(x)]}{k_B T}\right\}. \tag{29.3}$$

The potential $\phi(x)$ must be determined self-consistently, as that potential arising (via Poisson's equation) when the carrier densities have the forms (29.3). We examine

[2] A crude argument appropriate to metals is given in Appendix J. Analogous arguments (of comparable crudity) can be developed for semiconductors.

this problem in the special case (again, the case of major practical interest) in which far from the transition region on either side extrinsic conditions prevail, in which the impurities are fully "ionized" (see pages 583–584). Thus far away on the $n$-side the density of conduction band electrons is very nearly equal to the density $N_d$ of donors, while far away on the $p$-side the density of valence band holes is very nearly equal to the density $N_a$ of acceptors:

$$N_d = n_e(\infty) = N_c(T) \exp\left\{-\frac{[\mathcal{E}_c - e\phi(\infty) - \mu]}{k_B T}\right\},$$
$$N_a = p_v(-\infty) = P_v(T) \exp\left\{-\frac{[\mu - \mathcal{E}_v + e\phi(-\infty)]}{k_B T}\right\}. \tag{29.4}$$

Since the entire crystal is in thermal equilibrium, the chemical potential does not vary with position. In particular, the same value of $\mu$ appears in either of Eqs. (29.4); this immediately requires that the total potential drop across the junction be given by[3]

$$e\phi(\infty) - e\phi(-\infty) = \mathcal{E}_c - \mathcal{E}_v + k_B T \ln\left[\frac{N_d N_a}{N_c P_v}\right], \tag{29.5}$$

or

$$\boxed{e\,\Delta\phi = E_g + k_B T \ln\left[\frac{N_d N_a}{N_c P_v}\right].} \tag{29.6}$$

An alternative way of representing the information in (29.3) and (29.6) is sometimes helpful. If we define a position-dependent "electrochemical potential" $\mu_e(x)$ by

$$\mu_e(x) = \mu + e\phi(x), \tag{29.7}$$

then we may write the carrier densities (29.3) as

$$n_e(x) = N_c(T) \exp\left\{-\frac{[\mathcal{E}_c - \mu_e(x)]}{k_B T}\right\},$$
$$p_v(x) = P_v(T) \exp\left\{-\frac{[\mu_e(x) - \mathcal{E}_v]}{k_B T}\right\}. \tag{29.8}$$

These have precisely the form of the relations (28.12) for a homogeneous semiconductor, except that the constant chemical potential $\mu$ is replaced by the electrochemical potential $\mu_e(x)$. Thus $\mu_e(\infty)$ is the chemical potential of a homogeneous $n$-type crystal whose properties are identical to those of the inhomogeneous crystal far on the $n$-side of the transition region, while $\mu_e(-\infty)$ is the chemical potential of a homogeneous $p$-type crystal identical to the inhomogeneous crystal far on the $p$-side. The relation (29.6) can equally well be written as[4]

$$e\,\Delta\phi = \mu_e(\infty) - \mu_e(-\infty). \tag{29.9}$$

---

[3] The derivation of (29.5) requires the validity of (29.3) only far from the depletion layer, where $\phi$ is indeed slowly varying. It therefore holds even when the semiclassical model fails in the transition region.

[4] This follows directly from (29.7). Equation (29.9) is sometimes summarized in the rule that the total potential drop is such as to bring the "Fermi levels at the two ends of the junction" into coincidence. This point of view is evidently inspired by the representation of Figure 29.2b.

Figure 29.2a shows the electrochemical potential plotted as a function of position along the $p$-$n$ junction. We have assumed (as will be demonstrated below) that $\phi$ varies monotonically from one end to the other. Figure 29.2b shows an alternative representation of the same information in which the potential $\phi$ giving the position dependence in (29.3) is regarded as shifting $\mathcal{E}_c$ (or $\mathcal{E}_v$) rather than $\mu$. In either case, the significance of the diagrams is that at any particular position $x$ along the junction, the carrier densities are those that would be found in a piece of homogeneous material with the impurity concentrations prevailing at $x$, and with a chemical potential that is positioned with respect to the band edges as shown in the vertical section of the diagrams at $x$.

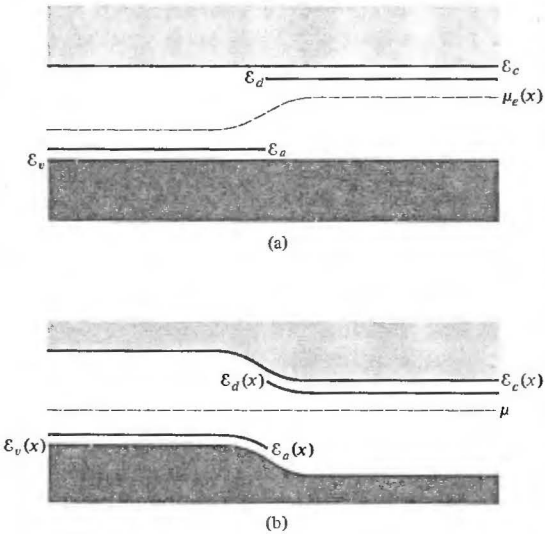


**Figure 29.2**
Two equivalent ways of representing the effect of the internal potential $\phi(x)$ on the electron and hole densities of a $p$-$n$ junction. (a) The electrochemical potential $\mu_e(x) = \mu + e\phi(x)$ is plotted along the $p$-$n$ junction. The carrier densities at any point $x$ are those that would be found in a uniform semiconductor characterized by the fixed band and impurity energies $\mathcal{E}_c$, $\mathcal{E}_v$, $\mathcal{E}_d$, and $\mathcal{E}_a$, at a chemical potential equal to $\mu_e(x)$. (b) Here $\mathcal{E}_c(x) = \mathcal{E}_c - e\phi(x)$ is the energy of an electron wave packet localized about $x$ formed from levels very near the conduction band minimum, and similarly for $\mathcal{E}_v(x)$. The energies of the local impurity levels are $\mathcal{E}_d(x) = \mathcal{E}_d - e\phi(x)$ and $\mathcal{E}_a(x) = \mathcal{E}_a - e\phi(x)$. The (constant) chemical potential is also shown. The carrier densities at any point $x$ are those that would be found in a uniform semiconductor characterized by band and impurity energies equal to $\mathcal{E}_c(x)$, $\mathcal{E}_d(x)$, $\mathcal{E}_a(x)$, and $\mathcal{E}_v(x)$ at the fixed chemical potential $\mu$.

Equation (29.6) (or its equivalent form, (29.9)) serves as the boundary condition in a differential equation determining the potential $\phi(x)$. The differential equation is simply Poisson's equation,[5]

$$-\nabla^2\phi = -\frac{d^2\phi}{dx^2} = \frac{4\pi\rho(x)}{\epsilon}, \tag{29.10}$$

[5] Here $\epsilon$ is the static dielectric constant of the semiconductor. The use of the macroscopic equation is possible because $\phi$ varies over the depletion layer, which is large on the interatomic scale.

relating the potential $\phi(x)$ to the charge distribution $\rho(x)$ giving rise to it. To express $\rho(x)$ in terms of $\phi$ and get a closed equation, we first note that if (as we have assumed) the impurities are fully ionized far from the junction, then they will remain fully ionized[6] at all $x$. Consequently the charge density due to the impurities and the carriers is[7]

$$\rho(x) = e[N_d(x) - N_a(x) - n_c(x) + p_v(x)]. \tag{29.11}$$

When the carrier and impurity densities (29.3) and (29.1) are substituted into the form (29.11) for the charge density, and the result is substituted into Poisson's equation (29.10), one finds a nonlinear differential equation for $\phi(x)$ whose exact solution usually requires numerical techniques.[8] However, a quite reasonable description of $\phi(x)$ may be had by exploiting the fact that the total change in $e\phi$ is of order $E_g \gg k_BT$. The relevance of this fact emerges when we combine (29.3) and (29.4) to write

$$\begin{aligned} n_c(x) &= N_d e^{-e[\phi(\infty)-\phi(x)]/k_BT}, \\ p_v(x) &= N_a e^{-e[\phi(x)-\phi(-\infty)]/k_BT}. \end{aligned} \tag{29.12}$$

Suppose that the change in $\phi$ occurs within a region $-d_p \leqslant x \leqslant d_n$. Outside of this region, $\phi$ has its asymptotic value, and therefore $n_c = N_d$ on the $n$ side, $p_v = N_a$ on the $p$ side, and $\rho = 0$. Within the region, except quite near the boundaries, $e\phi$ differs by many $k_BT$ from its asymptotic value, so $n_c \ll N_d$, $p_v \ll N_a$. Thus, except in the vicinity of $x = -d_p$ and $x = d_n$, the charge density (29.11) between $-d_p$ and $d_n$ is quite accurately given by $\rho(x) = e[N_d(x) - N_a(x)]$, there being no appreciable carrier charge to cancel the charges of the "ionized" impurities. The points $x = -d_p$ and $x = d_n$ therefore mark the boundaries of the depletion layer.

Combining these observations, and using the form (29.1) for the impurity densities, we find that except for $x$ just greater than $-d_p$ or just less than $d_n$, Poisson's equation is well approximated by

$$\phi''(x) = \left\{ \begin{array}{ll} 0, & x > d_n, \\ \dfrac{-4\pi e N_d}{\epsilon}, & d_n > x > 0, \\ \dfrac{4\pi e N_a}{\epsilon}, & 0 > x > -d_p, \\ 0, & -d_p > x. \end{array} \right\} \tag{29.13}$$

---

[6] If $\phi$ is monotonic (as we shall find below) this follows from the fact that the degree of ionization of an impurity increases, the farther the chemical potential is from the impurity level. See Figure 29.2 and Eqs. (28.32) and (28.34).

[7] The density of holes on the far $n$-side has the very small value $p_v(\infty) = n_i^2/N_d$ required by the law of mass action. However, the density of electrons on the far $n$-side actually exceeds $N_d$ by this same small amount so as to insure that $n_c(\infty) - p_v(\infty) = N_d$. In computing the total charge density, if we ignore this small correction in $n_c$ (as we have done in writing (29.4)), then we should also ignore the small compensating density of holes on the far $n$-side. Similar remarks apply to the small concentration of electrons on the far $p$-side. These "minority carrier densities" have negligible effect on the total balance of charge. We shall see below, however, that they play an important role in determining the flow of currents in the presence of an applied potential.

[8] Some aspects of that equation are investigated in Problem 1.

This immediately integrates to give

$$\phi(x) = \begin{cases} \phi(\infty), & x > d_n, \\ \phi(\infty) - \left(\dfrac{2\pi e N_d}{\epsilon}\right)(x - d_n)^2, & d_n > x > 0, \\ \phi(-\infty) + \left(\dfrac{2\pi e N_a}{\epsilon}\right)(x + d_p)^2, & 0 > x > -d_p, \\ \phi(-\infty), & x < -d_p. \end{cases} \tag{29.14}$$

The boundary conditions (continuity of $\phi$ and its first derivative) are explicitly obeyed by the solution (29.14) at $x = -d_p$ and $x = d_n$. Requiring them to hold at $x = 0$ gives two additional equations that determine the lengths $d_n$ and $d_p$. Continuity of $\phi'$ at $x = 0$ implies that

$$N_d d_n = N_a d_p, \tag{29.15}$$

which is just the condition that the excess of positive charge on the $n$-side of the junction be equal to the excess of negative charge on the $p$-side. Continuity of $\phi$ at $x = 0$ requires that

$$\left(\frac{2\pi e}{\epsilon}\right)(N_d d_n^2 + N_a d_p^2) = \phi(\infty) - \phi(-\infty) = \Delta\phi. \tag{29.16}$$

Together with (29.15) this determines the lengths $d_n$ and $d_p$:

$$d_{n,p} = \left\{\frac{(N_a/N_d)^{\pm 1}}{(N_d + N_a)} \frac{\epsilon \Delta\phi}{2\pi e}\right\}^{1/2}. \tag{29.17}$$

To estimate the sizes of these lengths we may write Eq. (29.17) in the numerically more convenient form

$$d_{n,p} = 105\left\{\frac{(N_a/N_d)^{\pm 1}}{10^{-18}(N_d + N_a)}[\epsilon e\, \Delta\phi]_{\text{eV}}\right\}^{1/2} \text{Å}. \tag{29.18}$$

The quantity $\epsilon e\, \Delta\phi$ is typically of order 1 eV, and since typical impurity concentrations are in the range from $10^{14}$ to $10^{18}$ per cubic centimeter, the lengths $d_n$ and $d_p$, which give the extent of the depletion layer, will generally be from $10^4$ to $10^2$ Å. The field within the depletion layer is of order $\Delta\phi/(d_n + d_p)$, and for $d$'s of this size is therefore in the range from $10^5$ to $10^7$ volts per meter, for an energy gap of 0.1 eV.

The resulting picture of the depletion layer is shown in Figure 29.3. The potential $\phi$ varies monotonically through the layer, as asserted above. Except at the boundaries of the layer, the carrier concentrations are negligible compared with the impurity concentrations, so the charge density is that of the ionized impurities. Outside of the depletion layer the carrier concentrations balance the impurity concentrations, and the charge density is zero.

The mechanism establishing such a region of sharply reduced carrier densities is relatively simple. Suppose that one initially were able to impose carrier concentrations that gave charge neutrality at every point in the crystal. Such a configuration could not be maintained, for electrons would begin to diffuse from the $n$-side (where their concentration was high) to the $p$-side (where their concentration was very low), and

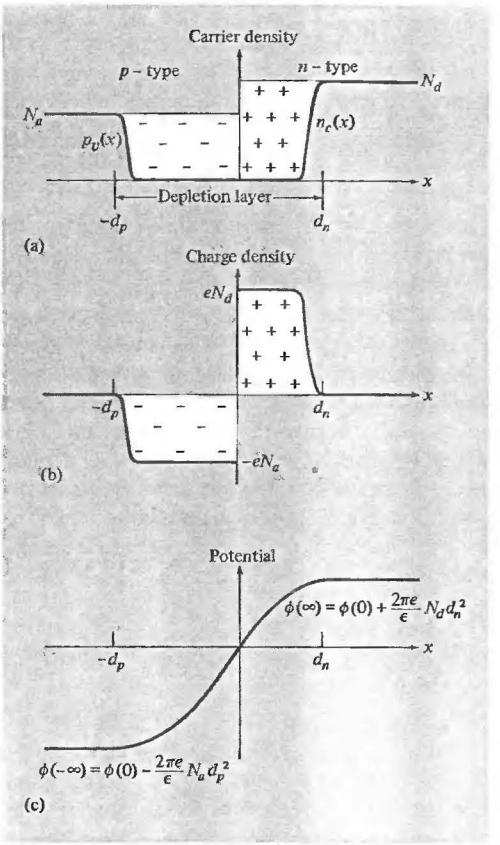


**Figure 29.3**
(a) Carrier densities, (b) charge density, and (c) potential $\phi(x)$ plotted vs. position across an abrupt *p-n* junction. In the analysis in the text the approximation was made that the carrier densities and charge density are constants except for discontinuous changes at $x = -d_p$ and $x = d_n$. More precisely (see Problem 1), these quantities undergo rapid change over regions just within the depletion layer whose extent is a fraction of order $(k_B T/E_g)^{1/2}$ of the total extent of the depletion layer. The extent of the depletion layer is typically from $10^2$ to $10^4$ Å.

holes would diffuse in the opposite direction. As this diffusion continued, the resulting transfer of charge would build up an electric field opposing further diffusive currents, until an equilibrium configuration was reached in which the effect of the field on the currents precisely canceled the effect of diffusion. Because the carriers are highly mobile, in this equilibrium configuration the carrier densities are very low wherever the field has an appreciable value. This is precisely the state of affairs depicted in Figure 29.3.

## ELEMENTARY PICTURE OF RECTIFICATION BY A *p-n* JUNCTION

We now consider the behavior of a *p-n* junction when an external voltage $V$ is applied. We shall take $V$ to be positive if its application raises the potential of the *p*-side with respect to the *n*-side. When $V = 0$ we found above that there is a depletion layer some $10^2$ to $10^4$ Å in extent about the transition point where the doping changes from *p*-type to *n*-type, in which the density of carriers is reduced greatly below its value in the homogeneous regions farther away. Because of its greatly reduced carrier

density, the depletion layer will have a much higher electrical resistance than the homogeneous regions, and the whole device can therefore be viewed as a series circuit in which a relatively high resistance is sandwiched between two relatively low resistances. When a potential $V$ is applied across such a circuit, almost all of the potential drop will occur across the region of high resistance. Thus even in the presence of an applied potential $V$, we expect that the potential $\phi(x)$ along the device will vary appreciably only within the depletion layer. When $V = 0$ we found that $\phi(x)$ rose from the $p$-side of the depletion layer to the $n$-side by the amount (which we now denote by $(\Delta\phi)_0$) given by Eq. (29.6), so we conclude that when $V \neq 0$ the change in potential across the depletion layer is modified to

$$\Delta\phi = (\Delta\phi)_0 - V. \tag{29.19}$$

Associated with this change in potential drop across the depletion layer, there is a change in the size of the layer. The lengths $d_n$ and $d_p$ giving the extent of the layer on the $n$- and $p$-sides of the junction are determined by Eqs. (29.15) and (29.16), which use only the value of the total potential drop across the layer, and the assumption that the carrier densities are greatly reduced throughout almost all of the layer. We shall find below that this assumption remains valid when $V \neq 0$, and therefore $d_n$ and $d_p$ continue to be given by Eq. (29.17) provided that we take the value of $\Delta\phi$ to be $(\Delta\phi)_0 - V$. Since $d_n$ and $d_p$ vary as $(\Delta\phi)^{1/2}$ according to Eq. (29.17), we conclude that when $V \neq 0$,

$$d_{n,p}(V) = d_{n,p}(0)\left[1 - \frac{V}{(\Delta\phi)_0}\right]^{1/2}. \tag{29.20}$$

This behavior of $\phi$ and the extent of the depletion layer are illustrated in Figure 29.4.

To deduce the dependence on $V$ of the current that flows when a $p$-$n$ junction is "biased" by the application of an external voltage, we must consider separately the currents of electrons and holes. Throughout the discussion that follows we shall use the symbol $J$ for number current densities and $j$ for electrical current densities, so that

$$j_e = -eJ_e, \quad j_h = eJ_h. \tag{29.21}$$

When $V = 0$, both $J_e$ and $J_h$ vanish. This does not, of course, mean that no individual carriers flow across the junction, but only that as many electrons (or holes) flow in one direction as in the other. When $V \neq 0$, this balance is disrupted. Consider, for example, the current of holes across the depletion layer. This has two components:

1. A hole current flows from the $n$- to the $p$-side of the junction, known as the hole *generation current*. As the name indicates, this current arises from holes that are generated just on the $n$-side of the depletion layer by the thermal excitation of electrons out of valence band levels. Although the density of such holes on the $n$-side ("minority carriers") is minute compared with the density of electrons ("majority carriers"), they play an important role in carrying current across the junction. This is because any such hole that wanders into the depletion layer is immediately swept over to the $p$-side of the junction by the strong electric field

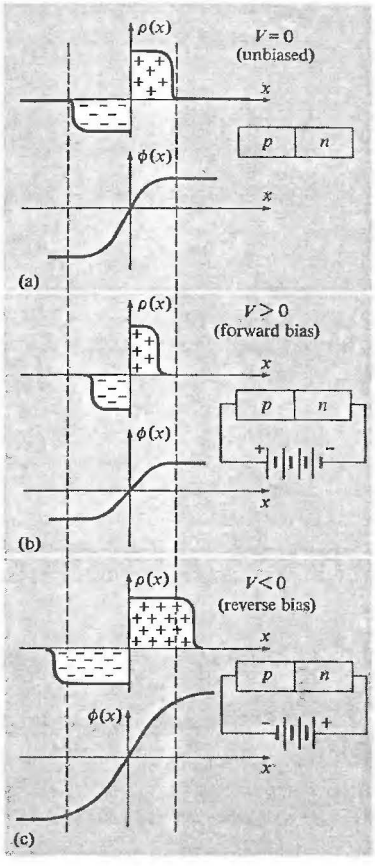


**Figure 29.4**
The charge density $\rho$ and potential $\phi$ in the depletion layer (a) for the unbiased junction, (b) for the junction with $V > 0$ (forward bias), and (c) for the junction with $V < 0$ (reverse bias). The positions $x = d_n$ and $x = -d_p$ that mark the boundaries of the depletion layer when $V = 0$ are given by the dashed lines. The depletion layer and change in $\phi$ are reduced by a forward bias and increased by a reverse bias.

that prevails within the layer. The resulting generation current is insensitive to the size of the potential drop across the depletion layer, since any hole, having entered the layer from the $n$-side, will be swept through to the $p$-side.[9]

2. A hole current flows from the $p$- to the $n$-side of the junction, known as the hole *recombination current*.[10] The electric field in the depletion layer acts to oppose such a current, and only holes that arrive at the edge of the depletion layer with a thermal energy sufficient to surmount the potential barrier will contribute to

[9] The density of holes giving rise to the hole generation current will also be insensitive to the size of $V$, provided that $eV$ is small compared with $E_g$, for this density is entirely determined by the law of mass action and the density of electrons. The latter density differs only slightly from the value $N_c$ outside of the depletion layer when $eV$ is small compared with $E_g$, as will emerge from the more detailed analysis below.

[10] So named because of the fate suffered by such holes upon arriving on the $n$-side of the junction, where one of the abundant electrons will eventually drop into the empty level that constitutes the hole.

the recombination current. The number of such holes is proportional to $e^{-e\Delta\phi/k_BT}$, and therefore[11]

$$J_h^{\text{rec}} \propto e^{-e[(\Delta\phi)_0 - V]/k_BT}. \tag{29.22}$$

In contrast to the generation current, the recombination current is highly sensitive to the applied voltage $V$. We can compare their magnitudes by noting that when $V = 0$ there can be no net hole current across the junction:

$$J_h^{\text{rec}}\big|_{V=0} = J_h^{\text{gen}}. \tag{29.23}$$

Taken together with Eq. (29.22), this requires that

$$J_h^{\text{rec}} = J_h^{\text{gen}} e^{eV/k_BT}. \tag{29.24}$$

The total current of holes flowing from the $p$- to the $n$-side of the junction is given by the recombination current minus the generation current:

$$J_h = J_h^{\text{rec}} - J_h^{\text{gen}} = J_h^{\text{gen}}(e^{eV/k_BT} - 1). \tag{29.25}$$

The same analysis applies to the components of the electron current, except that the generation and recombination currents of electrons flow oppositely to the corresponding currents of holes. Since, however, the electrons are oppositely charged, the electrical generation and recombination currents of electrons are parallel to the electrical generation and recombination currents of holes. The total electrical current density is thus:

$$\boxed{j = e(J_h^{\text{gen}} + J_e^{\text{gen}})(e^{eV/k_BT} - 1).} \tag{29.26}$$

This has the highly asymmetric form characteristic of rectifiers, as shown in Figure 29.5.

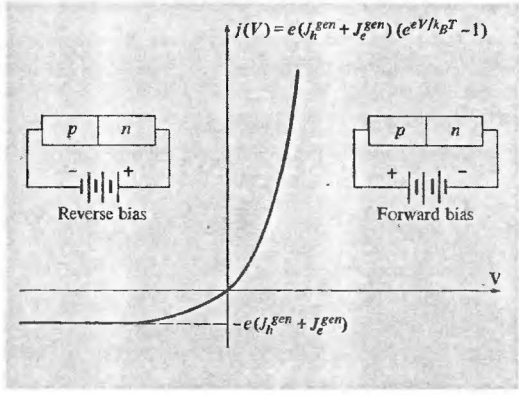


**Figure 29.5**
Current vs. applied voltage $V$ for a $p$-$n$ junction. The relation is valid for eV small compared with the energy gap, $E_g$. The saturation current $(eJ_h^{\text{gen}} + eJ_e^{\text{gen}})$ varies with temperature as $e^{-E_g/k_BT}$, as established below.

[11] In assuming that (29.22) gives the dominant dependence of the hole recombination current on $V$, we are assuming that the density of holes just on the $p$-side of the depletion layer differs only slightly from $N_a$. We shall find that this is also the case provided that $eV$ is small compared with the energy gap $E_g$.