3

The Physics of the Solar Cell

Jeffery L. Gray

Purdue University, West Lafayette, Indiana, USA

3.1 INTRODUCTION

Semiconductor solar cells are fundamentally quite simple devices. Semiconductors have the capacity to absorb light and to deliver a portion of the energy of the absorbed photons to carriers of electrical current – electrons and holes. A semiconductor diode separates and collects the carriers and conducts the generated electrical current preferentially in a specific direction. Thus, a solar cell is simply a semiconductor diode that has been carefully designed and constructed to efficiently absorb and convert light energy from the sun into electrical energy.

A simple conventional solar cell structure is depicted in Figure 3.1. Sunlight is incident from the top, on the front of the solar cell. A metallic grid forms one of the electrical contacts of the diode and allows light to fall on the semiconductor between the grid lines and thus be absorbed and converted into electrical energy. An antireflective layer between the grid lines increases the amount of light transmitted to the semiconductor. The semiconductor diode is fashioned when an *n*-type semiconductor and a *p*-type semiconductor are brought together to form a metallurgical junction. This is typically achieved through diffusion or implantation of specific impurities (dopants) or via a deposition process. The diode's other electrical contact is formed by a metallic layer on the back of the solar cell.

All electromagnetic radiation, including sunlight, can be viewed as being composed of particles called photons which carry specific amounts of energy determined by the spectral properties of their source. Photons also exhibit a wavelike character with the wavelength, λ, being related to the photon energy E_λ by

$$E_\lambda = \frac{hc}{\lambda} \qquad (3.1)$$

where h is Plank's constant and c is the speed of light. Only photons with sufficient energy to create an electron–hole pair, that is, those with energy greater than the semiconductor bandgap
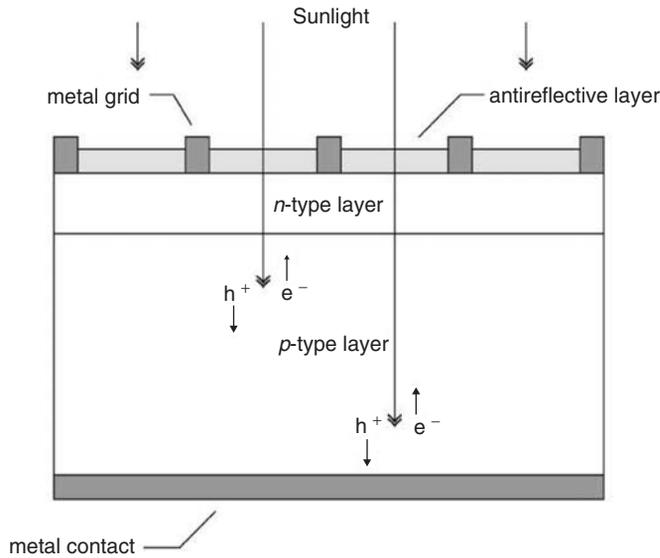
Figure 3.1 A schematic of a simple conventional solar cell. Creation of electron–hole pairs, e$^-$ and h$^+$, respectively, is depicted

(E_G), will contribute to the energy conversion process. Thus, the spectral composition of sunlight is an important consideration in the design of efficient solar cells.

The sun has a surface temperature of approximately 5762 K and its radiation spectrum can be approximated by a black body radiator at that temperature. Emission of radiation from the sun, as with all black body radiators, is isotropic. However, the Earth's great distance from the sun (approximately 93 million miles or 150 million kilometers) means that only those photons emitted directly at the Earth contribute to the solar spectrum as observed from the Earth. Therefore, for most practical purposes, the light falling on the Earth can be thought of as parallel streams of photons. Just above the Earth's atmosphere, the radiation intensity, or solar constant, is about 1.353 kW/m^2 [1] and the spectral distribution is referred to as an *air mass zero* (AM0) radiation spectrum. The air mass is a measure of how absorption in the atmosphere affects the spectral content and intensity of the solar radiation reaching the Earth's surface. The air mass number is given by [1]

$$\text{Air mass} = \frac{1}{\cos \theta} \tag{3.2}$$

where θ is the angle of incidence ($\theta = 0$ when the sun is directly overhead). The air mass number is always greater than or equal to one at the Earth's surface.

A widely used standard for comparing solar cell performance is the AM1.5 ($\theta = 48.2°$) spectrum normalized to a total power density of 1 kW/m^2. The spectral content of sunlight at the Earth's surface also has a diffuse (indirect) component due to scattering and reflection in the atmosphere and surrounding landscape, and can account for up to 20% of the light incident on a solar cell. The air mass number is therefore further defined by whether or not the measured spectrum includes the diffuse component. An AM1.5g (global) spectrum includes the diffuse component, while an AM1.5d (direct) does not. Black body ($T = 5762$ K), AM0, and AM1.5g radiation spectrums are shown in Figure 3.2. The air mass and solar radiation are described in more detail in Chapters 18 and 22.
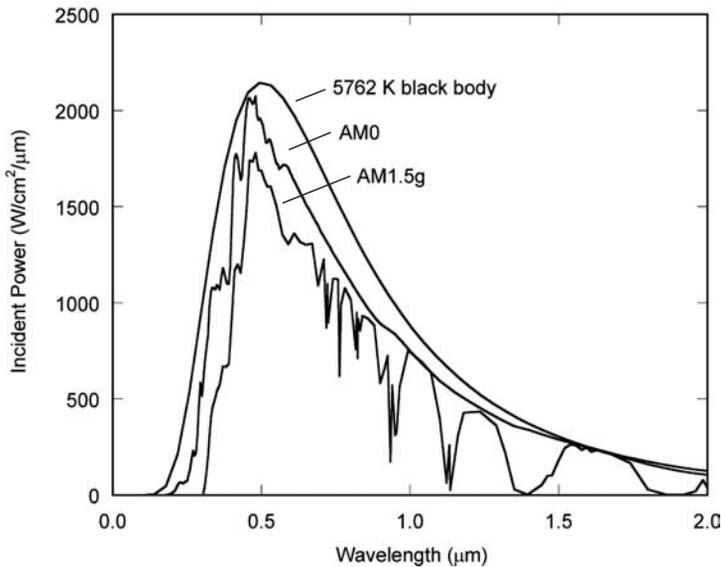
Figure 3.2   The radiation spectrum for a black body at 5780 K, an AM0 spectrum, and an AM1.5 global spectrum

The basic physical principles underlying the operation of solar cells are the subject of this chapter. First, a brief review of the fundamental properties of semiconductors is given that includes an overview of semiconductor band structure and carrier generation, recombination, and transport. Next, the electrostatic properties of the *pn*-junction diode are reviewed, followed by a description of the basic operating characteristics of the solar cell, including the derivation (based on the solution of the minority-carrier diffusion equation) of an expression for the current–voltage characteristic of an idealized solar cell. This is used to define the basic solar cell figures of merit, namely, the open-circuit voltage V_{OC}; the short-circuit current I_{SC}; the fill factor FF; the conversion efficiency η, and the collection efficiency η_C. Much of the discussion here will focus on how carrier recombination is the primary factor controlling solar cell performance. Finally, some additional topics relevant to solar cell operation, design and analysis are presented. These include the relationship between bandgap and efficiency, the solar cell spectral response, parasitic resistive effects, temperature effects, voltage-dependent collection, a brief introduction to some modern cell design concepts, and a brief overview of detailed numerical modeling of solar cells.

3.2 FUNDAMENTAL PROPERTIES OF SEMICONDUCTORS

An understanding of the operation of semiconductor solar cells requires familiarity with some basic concepts of solid-state physics. Here, an introduction is provided to the essential concepts needed to examine the physics of solar cells. More complete and rigorous treatments are available from a number of sources [2–6].

Solar cells can be fabricated from a number of semiconductor materials, most commonly silicon (Si) – crystalline, polycrystalline, and amorphous. Solar cells are also fabricated from other semiconductor materials such as GaAs, GaInP, Cu(InGa)Se$_2$, and CdTe, to name but a few. Solar cell materials are chosen largely on the basis of how well their absorption characteristics match the solar spectrum and upon their cost of fabrication. Silicon has been a common choice due to

Table 3.1 Abbreviated periodic
table of the elements

I	II	III	IV	V	VI
		B	C	N	O
		Al	Si	P	S
Cu	Zn	Ga	Ge	As	Se
Ag	Cd	In	Sn	Sb	Te

the fact that its absorption characteristics are a fairly good match to the solar spectrum, and silicon fabrication technology is well developed as a result of its pervasiveness in the semiconductor electronics industry.

3.2.1 Crystal Structure

Electronic grade semiconductors are very pure crystalline materials. Their crystalline nature means that their atoms are aligned in a regular periodic array. This periodicity, coupled with the atomic properties of the component elements, is what gives semiconductors their very useful electronic properties. An abbreviated periodic table of the elements is given in Table 3.1.

Note that silicon is in column IV, meaning that it has four valence electrons – that is, four electrons that can be shared with neighboring atoms to form covalent bonds with those neighbors. In crystalline silicon, the atoms are arranged in a *diamond lattice* (carbon is also a column IV element) with tetrahedral bonding – four bonds from each atom where the angle between any two bonds is 109.5°. Perhaps surprisingly, this arrangement can be represented by two interpenetrating face-centered cubic (*fcc*) unit cells where the second *fcc* unit cell is shifted one-fourth of the distance along the body diagonal of the first *fcc* unit cell. The lattice constant, ℓ, is the length of the edges of the cubic unit cell. The entire lattice can be constructed by stacking these unit cells. A similar arrangement, the *zincblende* lattice, occurs in many binary III–V and II–VI semiconductors such as GaAs (a III–V compound) and CdTe (a II–VI compound). For example, in GaAs, one interpenetrating *fcc* unit cell is composed entirely of gallium atoms and the other entirely of arsenic atoms. Note that the average valency is four for each compound, so that there are four bonds to and from each atom with each covalent bond involving two valence electrons. Some properties of semiconductors are dependent on the orientation of the crystal lattice, and casting the crystal structure in terms of a cubic unit cell makes identifying the orientation easier by means of Miller indices.

3.2.2 Energy Band Structure

Of more consequence to the physics of solar cells, however, is how the periodic crystalline structure of the semiconductor establishes its electronic properties. An electron moving in a semiconductor material is analogous to a particle confined to a three-dimensional box that has a complex interior structure, due primarily to the potential fields surrounding the component atom's nucleus and tightly bound core electrons. The dynamic behavior of the electron can be established from the electron wavefunction, ψ, which is obtained by solving the time-independent Schrödinger equation

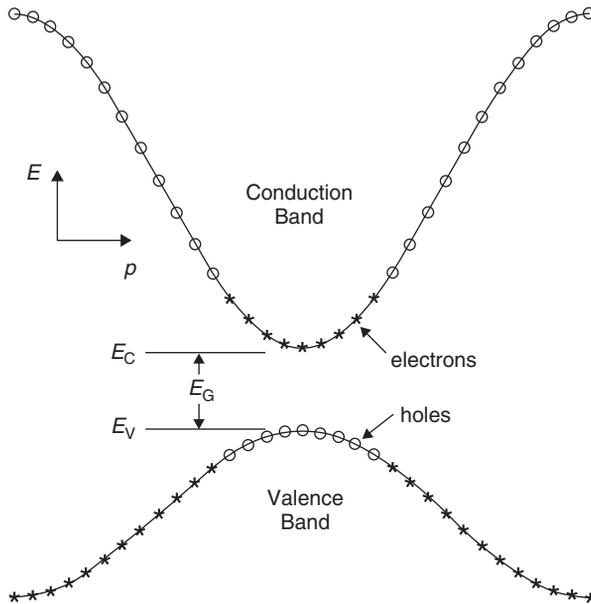$$\nabla^2 \psi + \frac{2m}{\hbar^2}[E - U(\vec{r})]\psi = 0 \qquad (3.3)$$

Figure 3.3   A simplified energy band diagram at $T > 0$ K for a direct bandgap (E_G) semiconductor. Electrons near the maxima in valence band have been thermally excited to the empty states near the conduction-band minima, leaving behind holes. The excited electrons and remaining holes are the negative and positive mobile charges that give semiconductors their unique transport properties

where m is electron mass, $\hbar$ is the reduced Planck constant, E is the energy of the electron, and $U(\vec{r})$ is the periodic potential energy inside the semiconductor. Solving this quantum mechanical equation is beyond the scope of this work, but suffice it to say that the solution defines the band structure (the allowed electron energies and the relationship between the electron's energy and momentum) of the semiconductor and, amazingly, tells us that the quantum mechanically computed motion of the electron in the crystal is, to a good approximation, like that of an electron in free space if its mass, m, is replaced by an effective mass m^* in Newton's second law of motion. Newton's second law of motion, from classical mechanics, is

$$F = m^* a \tag{3.4}$$

where F is the applied force and a is the acceleration of the electron.

A simplified energy band structure is illustrated in Figure 3.3. The allowed electron energies are plotted against the crystal momentum, $p = \hbar k$, where k is the wave vector (represented here as a scalar for simplicity) corresponding to the wavefunction solutions of the Schrödinger equation. Only the energy bands of immediate interest are shown – energy bands below the valence band are presumed to be fully occupied by electrons and those above the conduction band are presumed to be empty. The electron effective mass is defined by the curvature of the band as

$$m^* \equiv \left[\frac{d^2 E}{dp^2} \right]^{-1} = \left[\frac{1}{\hbar^2} \frac{d^2 E}{dk^2} \right]^{-1}. \tag{3.5}$$

Near the top of the valence band, the effective mass is actually negative. Electrons ($*$) fill the states from bottom to top and the states near the top of the valence band are empty ($\circ$) due to some

electrons being thermally excited into the conduction band. These empty states can conveniently be regarded as positively charged carriers of current called *holes* with a positive effective mass. It is conceptually much easier to deal with a relatively few number of holes that have a positive effective mass since they will behave like classical positively charged particles.

Notice that the effective mass is not constant within each band. The top of the valence band and the bottom of the conduction band are approximately parabolic in shape and therefore the electron effective mass (m_n^*) near the bottom of the conduction band is a constant, as is the hole effective mass (m_p^*) near the top of the valence band. This is a very practical assumption that greatly simplifies the modeling of semiconductor devices such as solar cells.

When the minimum of the conduction band occurs at the same value of the crystal momentum as the maximum of the valence band, as it does in Figure 3.3, the semiconductor is a *direct bandgap* semiconductor. When they do not align, the semiconductor is said to be an *indirect bandgap* semiconductor. This is especially important when the absorption of light by a semiconductor is considered later in this chapter.

Even amorphous materials exhibit a similar band structure. Over short distances, the atoms are arranged in a periodic manner and an electron wavefunction can be defined. The wavefunctions from these small regions overlap in such a way that a *mobility gap* can be defined, with electrons above the mobility gap defining the conduction band and holes below the gap defining the valence band. Unlike crystalline materials, however, there are a large number of localized energy states within the mobility gap (band tails and dangling bonds) that complicate the analysis of devices fabricated from these materials. Amorphous silicon (a-Si) solar cells are discussed in Chapter 12.

3.2.3 Conduction-band and Valence-band Densities of State

Now that the dynamics of the electron motion in a semiconductor has been approximated by a negatively charged particle with mass m_n^* in the conduction band and by a positively charged particle with mass m_p^* in the valence band, it is possible to calculate the density of states in each band. This again involves solving the time-independent Schrödinger equation for the wavefunction of a particle in a box, but in this case the box is empty. All the complexities of the periodic potentials of the component atoms have been incorporated into the effective mass. The density of states in the conduction band is given by [3]

$$g_C(E) = \frac{m_n^* \sqrt{2m_n^*(E - E_C)}}{\pi^2 \hbar^3} \text{ cm}^{-3} \text{eV}^{-1} \tag{3.6}$$

while the density of states in the valence band is given by

$$g_V(E) = \frac{m_p^* \sqrt{2m_p^*(E_V - E)}}{\pi^2 \hbar^3} \text{ cm}^{-3} \text{eV}^{-1}. \tag{3.7}$$

3.2.4 Equilibrium Carrier Concentrations

When the semiconductor is in thermal equilibrium (i.e. at a uniform temperature with no external injection or generation of carriers), the Fermi function determines the ratio of filled states to available states at each energy and is given by
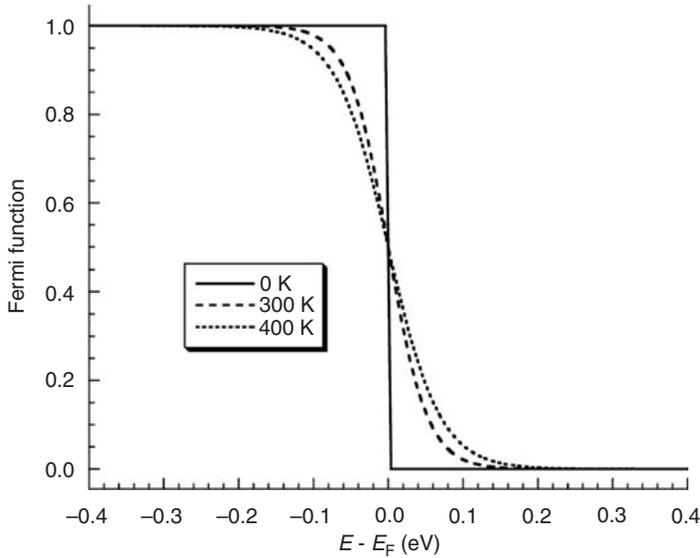
$$f(E) = \frac{1}{1 + e^{(E - E_F)/kT}} \tag{3.8}$$

Figure 3.4   The Fermi function at various temperatures

where E_F is the Fermi energy, k is Boltzmann's constant, and T is the Kelvin temperature. As seen in Figure 3.4, the Fermi function is a strong function of temperature. At absolute zero, it is a step function and all the states below E_F are filled with electrons and all those above E_F are completely empty. As the temperature increases, thermal excitation will leave some states below E_F empty, and the corresponding number of states above E_F will be filled with the excited electrons.

The equilibrium electron and hole concentrations (number per cm^3) are therefore

$$n_o = \int_{E_C}^{\infty} g_C(E) f(E) \mathrm{d}E = \frac{2N_C}{\sqrt{\pi}} F_{1/2}((E_F - E_C)/kT) \tag{3.9}$$

$$p_o = \int_{-\infty}^{E_V} g_V(E)[1 - f(E)] \mathrm{d}E = \frac{2N_V}{\sqrt{\pi}} F_{1/2}((E_V - E_F)/kT) \tag{3.10}$$

where $F_{1/2}(\xi)$ is the Fermi–Dirac integral of order 1/2,

$$F_{1/2}(\xi) = \int_0^{\infty} \frac{\sqrt{\xi'} \mathrm{d}\xi'}{1 + e^{\xi' - \xi}} \tag{3.11}$$

The conduction-band and valence-band effective densities of state ($\#/cm^3$), N_C and N_V, respectively, are given by

$$N_C = 2 \left(\frac{2\pi m_n^* kT}{h^2} \right)^{3/2} \tag{3.12}$$

and

$$N_V = 2 \left(\frac{2\pi m_p^* kT}{h^2} \right)^{3/2}. \tag{3.13}$$

When the Fermi energy, E_F, is sufficiently far ($>3\,kT$) from either band edge, the carrier concentrations can be well approximated (to within 2%) as [7]

$$n_o = N_C e^{(E_F - E_C)/kT} \tag{3.14}$$

and

$$p_o = N_V e^{(E_V - E_F)/kT}, \tag{3.15}$$

and the semiconductor is said to be *nondegenerate*. In nondegenerate semiconductors, the product of the equilibrium electron and hole concentrations is independent of the location of the Fermi energy and is just

$$p_o n_o = n_i^2 = N_C N_V e^{(E_V - E_C)/kT} = N_C N_V e^{-E_G/kT}. \tag{3.16}$$

In an undoped (intrinsic) semiconductor in thermal equilibrium, the number of electrons in the conduction band and the number of holes in the valence band are equal; $n_o = p_o = n_i$, where n_i is the intrinsic carrier concentration. The intrinsic carrier concentration can be computed from Equation (3.17), giving

$$n_i = \sqrt{N_C N_V} e^{(E_V - E_C)/2kT} = \sqrt{N_C N_V} e^{-E_G/2kT}. \tag{3.17}$$

The Fermi energy in an intrinsic semiconductor, $E_i = E_F$, is given by

$$E_i = \frac{E_V + E_C}{2} + \frac{kT}{2} \ln\left(\frac{N_V}{N_C}\right) \tag{3.18}$$

which is typically very close to the middle of the bandgap. The intrinsic carrier concentration is typically very small compared with the densities of states and typical doping densities ($n_i \approx 10^{10}\,\text{cm}^{-3}$ in Si) and intrinsic semiconductors behave very much like insulators; that is, they are not good conductors of electricity.

The number of electrons and holes in their respective bands, and hence the conductivity of the semiconductor, can be controlled through the introduction of specific impurities, or dopants, called *donors* and *acceptors*. For example, when semiconductor silicon is doped with phosphorus, one electron is donated to the conduction band for each atom of phosphorus introduced. From Table 3.1, it can be seen that phosphorous is in column V of the periodic table of elements and thus has five valence electrons. Four of these are used to satisfy the four covalent bonds of the silicon lattice and the fifth is available to fill an empty state in the conduction band. If silicon is doped with boron (valency of three, since it is in column III), each boron atom accepts an electron from the valence band, leaving behind a hole. All impurities introduce additional localized electronic states into the band structure, often within the forbidden band between E_C and E_V, as illustrated in Figure 3.5. If the energy of the state E_D introduced by a donor atom is sufficiently close to the conduction bandedge (within a few kT), there will be sufficient thermal energy to allow the extra electron to occupy a state in the conduction band. The donor state will then be positively charged (ionized) and must be considered when analyzing the electrostatics of the situation. Similarly, an acceptor atom will introduce a negatively charged (ionized) state at energy E_A. The controlled introduction of donor and acceptor impurities into a semiconductor allows the creation of the *n*-type (electrons are the primary carriers of electrical current) and *p*-type (holes are the primary carriers of electrical current) semiconductors, respectively. This is the basis for the construction
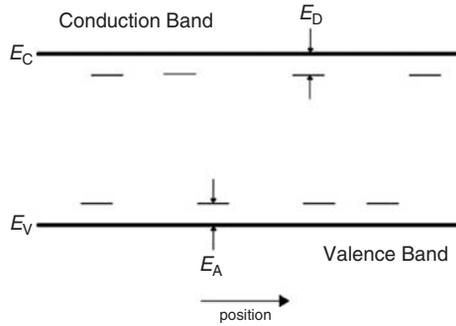
Figure 3.5 Donor and acceptor levels in a semiconductor. The nonuniform spatial distribution of these states reinforces the concept that these are localized states

of all semiconductor devices, including solar cells. The number of ionized donors and acceptors are given by [7]

$$N_{\mathrm{D}}^{+} = \frac{N_{\mathrm{D}}}{1 + g_{\mathrm{D}} \mathrm{e}^{(E_{\mathrm{F}} - E_{\mathrm{D}})/kT}} = \frac{N_{\mathrm{D}}}{1 + \mathrm{e}^{(E_{\mathrm{F}} - E_{\mathrm{D}}')/kT}} \tag{3.19}$$

and

$$N_{\mathrm{A}}^{-} = \frac{N_{\mathrm{A}}}{1 + g_{\mathrm{A}} \mathrm{e}^{(E_{\mathrm{A}} - E_{\mathrm{F}})/kT}} = \frac{N_{\mathrm{A}}}{1 + \mathrm{e}^{(E_{\mathrm{A}}' - E_{\mathrm{F}})/kT}} \tag{3.20}$$

where g_{D} and g_{A} are the donor and acceptor site degeneracy factors. Typically, $g_{\mathrm{D}} = 2$ and $g_{\mathrm{A}} = 4$. These factors are normally combined into the donor and the acceptor energies so that $E_D' = E_{\mathrm{D}} - kT \ln g_{\mathrm{D}}$ and $E_A' = E_{\mathrm{A}} + kT \ln g_{\mathrm{A}}$. Often, the donors and acceptors are assumed to be completely ionized so that $n_o \simeq N_D$ $n_o \simeq N_{\mathrm{D}}$ in n-type material and $p_o \simeq N_A$ in p-type material. The Fermi energy can then be written as

$$E_{\mathrm{F}} = E_{\mathrm{i}} + kT \ln \frac{N_{\mathrm{D}}}{n_{\mathrm{i}}} \tag{3.21}$$

in n-type material and as

$$E_{\mathrm{F}} = E_{\mathrm{i}} - kT \ln \frac{N_{\mathrm{A}}}{n_{\mathrm{i}}} \tag{3.22}$$

in p-type material.

When a very large concentration of dopants is introduced into the semiconductor, the dopants can no longer be thought of as a minor perturbation to the system. Their effect on the band structure must be considered. Typically, this so-called heavy doping effect manifests itself as a reduction in the bandgap, E_{G}, and thus an increase in the intrinsic carrier concentration, as can be seen from Equation (3.17). This bandgap narrowing (BGN) [8] is detrimental to solar cell performance and solar cells are typically designed to avoid this effect, though it may be a factor in the heavily doped regions near the solar cell contacts.

3.2.5 Light Absorption

The creation of electron–hole pairs via the absorption of sunlight is essential to the operation of solar cells. The excitation of an electron directly from the valence band (which leaves a hole behind)

to the conduction band is called *fundamental absorption*. Both the total energy and momentum of all particles involved in the absorption process must be conserved. Since the photon momentum, $p_\lambda = h/\lambda$, is very small compared with the range of the crystal momentum, $p = h/\ell$, the photon absorption process effectively conserves the momentum of the electron.[1] The absorption coefficient for a given photon energy, $h\nu$, is proportional to the probability, P_{12}, of the transition of an electron from the initial state E_1 to the final state E_2, the density of electrons in the initial state $g_V(E_1)$ and the density of available final states, and is then summed over all possible transitions between states where $E_2 - E_1 = h\nu$ [9],

$$\alpha(h\nu) \propto \sum P_{12} g_V(E_1) g_C(E_2), \tag{3.23}$$

assuming that all the valence-band states are full and all the conduction-band states are empty. Absorption results in creation of an electron–hole pair since a free electron excited into the conduction band leaves a free hole in the valence band.

In direct bandgap semiconductors, such as GaAs, GaInP, CdTe, and Cu(InGa)Se$_2$, the basic photon absorption process is illustrated in Figure 3.6. Both energy and momentum must be conserved in the transition. Every initial electron state with energy E_1 and crystal momentum p_1 in the valence band is associated with a final state in the conduction band at energy E_2 and crystal momentum p_2. Since the electron momentum is conserved, the crystal momentum of the final state is the same as the initial state, $p_1 \approx p_2 = p$.

Conservation of energy dictates that the energy of the absorbed photon is

$$h\nu = E_2 - E_1 \tag{3.24}$$

Since we have assumed parabolic bands,

$$E_V - E_1 = \frac{p^2}{2m_p^*} \tag{3.25}$$

and

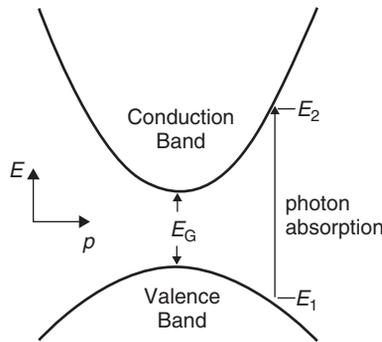$$E_2 - E_C = \frac{p^2}{2m_n^*} \tag{3.26}$$

Figure 3.6 Photon absorption in a direct bandgap semiconductor for an incident photon with energy $h\nu = E_2 - E_1 > E_G$

---

[1] The wavelength of sunlight, λ, is of the order of a micrometer (10^{-4} cm), while the lattice constant is a few angstroms (10^{-8} cm). Thus, the crystal momentum is several orders of magnitude larger than the photon momentum.

Combining Equations (3.25), (3.26), and (3.27) yields

$$h\upsilon - E_G = \frac{p^2}{2}\left(\frac{1}{m_n^*} + \frac{1}{m_p^*}\right) \tag{3.27}$$

and the absorption coefficient for direct transitions is [9]

$$\alpha(h\upsilon) \approx A^*(h\upsilon - E_G)^{1/2}, \tag{3.28}$$

where A^* is a constant. In some semiconductor materials, quantum selection rules do not allow transitions at $p = 0$, but allow them for $p \neq 0$. In such cases [9]

$$\alpha(h\upsilon) \approx \frac{B^*}{h\upsilon}(h\upsilon - E_G)^{3/2}, \tag{3.29}$$

where B^* is a constant.

In indirect band gap semiconductors such as Si and Ge, where the valence-band maximum occurs at a different crystal momentum from that of the conduction-band minimum, conservation of electron momentum necessitates that the photon absorption process involve an additional particle. Phonons, the particle representation of lattice vibrations in the semiconductor, are suited to this process because they are low-energy particles with relatively high momentum. This is illustrated in Figure 3.7. Notice that light absorption is facilitated by either phonon absorption or phonon emission. The absorption coefficient, when there is phonon absorption, is given by

$$\alpha_a(h\upsilon) = \frac{A(h\upsilon - E_G + E_{ph})^2}{e^{E_{ph}/kT} - 1} \tag{3.30}$$

and by

$$\alpha_e(h\upsilon) = \frac{A(h\upsilon - E_G - E_{ph})^2}{1 - e^{-E_{ph}/kT}} \tag{3.31}$$

when a phonon is emitted [9]. Because both processes are possible,

$$\alpha(h\upsilon) = \alpha_a(h\upsilon) + \alpha_e(h\upsilon). \tag{3.32}$$

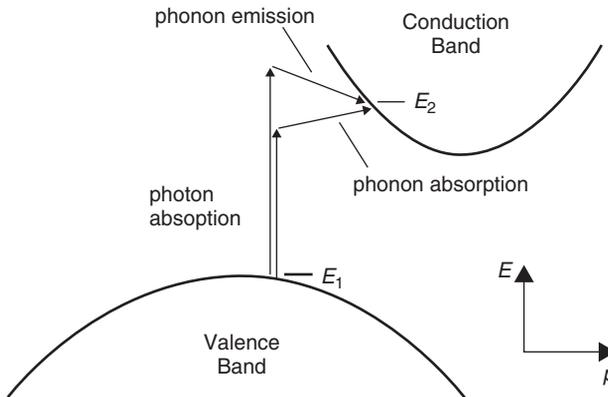**Figure 3.7** Photon absorption in an indirect bandgap semiconductor for a photon with energy $h\upsilon < E_2 - E_1$ and a photon with energy $h\upsilon > E_2 - E_1$. Energy and momentum in each case are conserved by the absorption and emission of a phonon, respectively
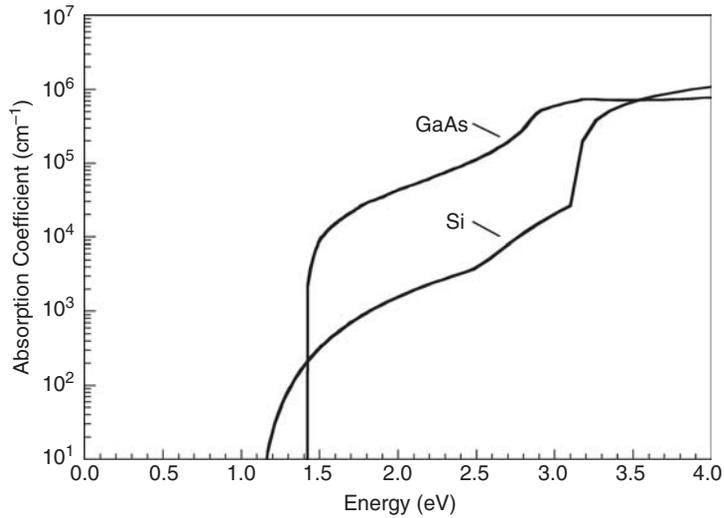
Figure 3.8 Absorption coefficient as a function of photon energy for Si (indirect bandgap) and GaAs (direct bandgap) at 300 K. Their bandgaps are 1.12 and 1.42 eV, respectively

Since both a phonon and an electron are needed to make the indirect gap absorption process possible, the absorption coefficient depends not only on the density of full initial electron states and empty final electron states but also on the availability of phonons (both emitted and absorbed) with the required momentum. Thus, compared with direct transitions, the absorption coefficient for indirect transitions is relatively small. As a result, light penetrates more deeply into indirect bandgap semiconductors than direct bandgap semiconductors. This is illustrated in Figure 3.8 for Si, an indirect bandgap semiconductor, and GaAs, a direct bandgap semiconductor. Similar spectra are shown for other semiconductors elsewhere in this handbook.

In both direct bandgap and indirect bandgap materials, a number of photon absorption processes are involved, though the mechanisms described above are the dominant ones. A direct transition, without phonon assistance, is possible in indirect bandgap materials if the photon energy is high enough (as seen in Figure 3.8 for Si at about 3.3 eV). Conversely, in direct bandgap materials, phonon-assisted absorption is also a possibility. Other mechanisms may also play a role in determining the optical absorption in semiconductors. These include absorption in the presence of an electric field (the Franz–Keldysh effect), absorption aided by localized states in the forbidden gap, and degeneracy effects when a significant number of states in the conduction band are not empty and/or when a significant number of state in the valence band are not full, as can happen in heavily doped materials (BGN) and under high-level injection (the Burstein–Moss shift). The net absorption coefficient is then the sum of the absorption coefficients due to all absorption processes or

$$\alpha(h\nu) = \sum_i \alpha_i(h\nu). \tag{3.33}$$

In practice, measured absorption coefficients or empirical expressions for the absorption coefficient are used in analysis and modeling. Chapter 17 has more details on extracting optical parameters from measurements and on the relation between optical and electric constants especially for thin film and conductive oxides, including heavily doped materials.

The rate of creation of electron–hole pairs (number of electron–hole pairs per cm^3 per second) as a function of position within a solar cell is

$$G(x) = (1 - s) \int_\lambda (1 - r(\lambda)) f(\lambda) \alpha(\lambda) e^{-\alpha x} \, d\lambda \tag{3.34}$$

where s is the grid-shadowing factor, $r(\lambda)$ is the reflectance, $\alpha(\lambda)$ is the absorption coefficient, and $f(\lambda)$ is the incident photon flux (number of photons incident per unit area per second per wavelength). The sunlight is assumed to be incident at $x = 0$. Here, the absorption coefficient has been cast in terms of the light's wavelength through the relationship $h\nu = hc/\lambda$. The photon flux, $f(\lambda)$, is obtained by dividing the incident power density at each wavelength by the photon energy.

Free-carrier absorption, in which electrons in the conduction band absorb the energy of a photon and move to an empty state higher in the conduction band (correspondingly for holes in the valence band), is typically only significant for photons with $E < E_G$ since the free-carrier absorption coefficient increases with increasing wavelength,

$$\alpha_{fc} \propto \lambda^\gamma \tag{3.35}$$

where $1.5 < \gamma < 3.5$ [9]. Thus, in single-junction solar cells, it does not affect the creation of electron–hole pairs and can be ignored (although free-carrier absorption can be exploited to probe the excess carrier concentrations in solar cells for the purpose of determining recombination parameters [10]). However, free-carrier absorption is a consideration in tandem solar cell systems in which a wide bandgap (E_{G1}) solar cell is stacked on top of a solar cell of smaller bandgap ($E_{G2} < E_{G1}$). Photons with energy too low to be absorbed in the top cell ($h\nu < E_{G1}$) will be transmitted to the bottom cell and be absorbed there (if $h\nu > E_{G2}$). Of course, more solar cells can be stacked as long as $E_{G1} > E_{G2} > E_{G3} \ldots$, and so on. The number of photons transmitted to the next cell in the stack will be reduced by whatever amount of free-carrier absorption occurs. This loss can be avoided by splitting the incident spectrum and directing the matched portion of the spectrum to each component solar cell of a multijuction system [11]. Multijunction solar cells are discussed more completely in Chapters 8 and 12.

3.2.6 Recombination

When a semiconductor is taken out of thermal equilibrium, for instance by illumination and/or the injection of current, the concentrations of electrons (n) and holes (p) tend to relax back toward their equilibrium values through a process called *recombination* in which an electron falls from the conduction band to the valence band, thereby eliminating a valence-band hole. There are several recombination mechanisms important to the operation of solar cells – recombination through traps (defects) in the forbidden gap, radiative (band-to-band) recombination, and Auger recombination – that will be discussed here. These three processes are illustrated in Figure 3.9.

The net recombination rate per unit volume per second through a single level trap (SLT) located at energy $E = E_T$ within the forbidden gap, also commonly referred to as *Shockley–Read–Hall recombination*, is given by [12]

$$R_{SLT} = \frac{pn - n_i^2}{\tau_{SLT,n}(p + n_i e^{(E_i - E_T)/kT}) + \tau_{SLT,p}(n + n_i e^{(E_T - E_i)/kT})}. \tag{3.36}$$

The carrier lifetimes are given by
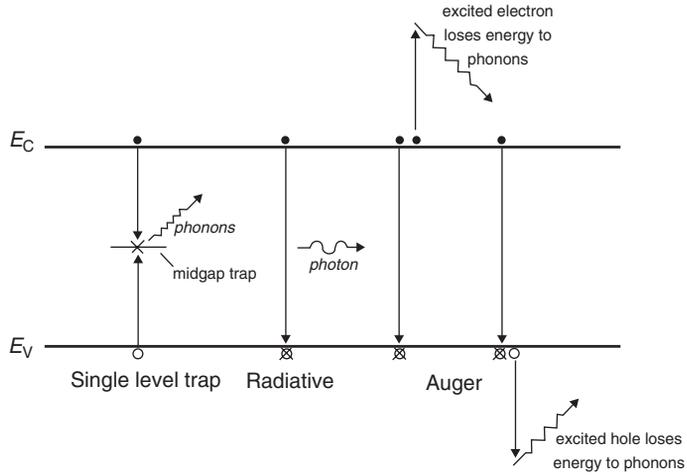
$$\tau_{SLT} = \frac{1}{\sigma v_{th} N_T} \tag{3.37}$$

Figure 3.9   Recombination processes in semiconductors

where σ is the capture cross-section (σ_n for electrons and σ_p for holes), v_{th} is the thermal velocity of the carriers, and N_T is the concentration of traps. The capture cross-section can be thought of as the size of the target presented to a carrier traveling through the semiconductor at velocity v_{th}. Small lifetimes correspond to high rates of recombination. If a trap presents a large target to the carrier, the recombination rate will be high (low carrier lifetime). When the velocity of the carrier is high, it has more opportunity within a given time period to encounter a trap and the carrier lifetime is low. Finally, the probability of interaction with a trap increases as the concentration of traps increases and the carrier lifetime is therefore inversely proportional to the trap concentration.

Some reasonable assumptions allow Equation (3.36) to be simplified. If the material is p-type ($p \approx p_o \gg n_o$), in low injection ($n_o \leq n \ll p_o$), and the trap energy is near the middle of the forbidden gap ($E_T \approx E_i$), the recombination rate can be written as

$$R_{\text{SLT}} \approx \frac{n - n_o}{\tau_{\text{SLT},n}}. \tag{3.38}$$

Notice that the recombination rate is solely dependent on the minority carrier. This is reasonable since there are far fewer minority carriers than majority carriers and one of each is necessary for there to be recombination.

If high-injection conditions prevail ($p \approx n \gg p_o, n_o$),

$$R_{\text{SLT}} \approx \frac{n}{\tau_{\text{SLT},p} + \tau_{\text{SLT},n}} \approx \frac{p}{\tau_{\text{SLT},p} + \tau_{\text{SLT},n}}. \tag{3.39}$$

In this case, the effective recombination lifetime is the sum of the two carrier lifetimes. While the recombination rate is high due to the large number of excess holes and electrons, the carrier lifetime is actually longer than in the case of low injection. This can be of significance in the base region of solar cells, especially concentrator cells (solar cells illuminated with concentrated sunlight), since the base is the least doped layer.

Radiative (band-to-band) recombination is simply the inverse of the optical generation process and is much more efficient in direct bandgap semiconductors than in indirect bandgap

semiconductors. When radiative recombination occurs, the energy of the electron is given to an emitted photon – this is how semiconductor lasers and light emitting diodes (LEDs) operate. In an indirect bandgap material, some of that energy is shared with a phonon. The net recombination rate due to radiative processes is given as

$$R_\lambda = B(pn - n_i^2) \tag{3.40}$$

If we have an n-type ($n \approx n_o \gg p_o$) semiconductor in low injection ($p_o \le p \ll n_o$), the net radiative recombination rate can be written in terms of an effective lifetime, $\tau_{\lambda,p}$,

$$R_\lambda \approx \frac{p - p_o}{\tau_{\lambda,p}} \tag{3.41}$$

where

$$\tau_{\lambda,p} = \frac{1}{n_o B}. \tag{3.42}$$

A similar expression can be written for p-type semiconductors. If high-injection conditions prevail ($p \approx n \gg p_o, n_o$), then

$$R_\lambda \approx Bp^2 \approx Bn^2. \tag{3.43}$$

Since photons with energies near that of the bandgap are emitted during this recombination process, it is possible for these photons to be reabsorbed before exiting the semiconductor. A well-designed direct bandgap solar cell can take advantage of this *photon recycling* and increase the effective lifetime [13].

Auger recombination is somewhat similar to radiative recombination, except that the energy of transition is given to another carrier (in either the conduction band or the valence band), as shown in Figure 3.9. This electron (or hole) then relaxes thermally (releasing its excess energy and momentum to phonons). Just as radiative recombination is the inverse process to optical absorption, Auger recombination is the inverse process to *impact ionization*, where an energetic electron collides with a crystal atom, breaking the bond and creating an electron–hole pair. The net recombination rate due to Auger processes is

$$R_{\text{Auger}} = (C_n n + C_p p)(pn - n_i^2) \tag{3.44}$$

In an n-type material in low injection (and assuming C_n and C_p are of comparable magnitudes), the net Auger recombination rate becomes

$$R_{\text{Auger}} \approx \frac{p - p_o}{\tau_{\text{Auger},p}} \tag{3.45}$$

with

$$\tau_{\text{Auger},p} = \frac{1}{C_n n_o^2}. \tag{3.46}$$

A similar expression can be derived for minority electron lifetime in p-type material. If high-injection conditions prevail ($p \approx n \gg p_o, n_o$), then

$$R_{\text{Auger}} \approx (C_n + C_p)p^3 \approx (C_n + C_p)n^3 \tag{3.47}$$

While the SLT recombination rate can be minimized by reducing the density of single-level traps and the radiative recombination rate can be minimized via photon recycling, the Auger recombination rate is a fundamental property of the semiconductor.

Each of these recombination processes occurs in parallel. And, there can be multiple and/or distributed traps[2] in the forbidden gap – in which case the net recombination is a sum of the contributions of each trap ($\sum_{\text{traps } i} R_{\text{SLT},i}$). Thus, the total recombination rate is the sum of rates due to each process

$$R = \left[\sum_{\text{traps } i} R_{\text{SLT},i} \right] + R_\lambda + R_{\text{Auger}}. \tag{3.48}$$

An effective minority-carrier lifetime for a doped material in low-level injection is given as

$$\frac{1}{\tau} = \left[\sum_{\text{traps } i} \frac{1}{\tau_{\text{SLT},i}} \right] + \frac{1}{\tau_\lambda} + \frac{1}{\tau_{\text{Auger}}}. \tag{3.49}$$

The distribution of traps in the energy gap for semiconductor materials can be influenced by the specific growth or processing conditions, impurities, and crystallographic defects.

Interfaces between two dissimilar materials, such as those that occur at the front surface of a solar cell, have a high concentration of defects due to the abrupt termination of the crystal lattice. These manifest themselves as a continuum of traps (*surface states*) within the forbidden gap at the surface and electrons and holes can recombine through them just as with bulk traps. These surface states are illustrated in Figure 3.10. Rather than giving a recombination rate per unit volume per second, surface states give a recombination rate per unit area per second. A general expression for surface recombination is [12]

$$R_{\text{S}} = \int_{E_{\text{V}}}^{E_{\text{C}}} \frac{pn - n_{\text{i}}^2}{(p + n_{\text{i}}e^{(E_{\text{i}} - E_{\text{t}})/kT})/s_n(E_{\text{t}}) + (n + n_{\text{i}}e^{(E_{\text{t}} - E_{\text{i}})/kT})/s_p(E_{\text{t}})} D_\Pi(E_{\text{t}}) \, dE_{\text{t}} \tag{3.50}$$

where E_{t} is the trap energy, $D_\Pi(E_{\text{t}})$ is the surface state (the concentration of traps is probably varies with trap energy), and $s_n(E_{\text{t}})$ and $s_p(E_{\text{t}})$ are surface recombination velocities, analogous to the carrier lifetimes for bulk traps. The surface recombination rate is generally written, for simplicity, as [12]

$$R_{\text{S}} = S_p(p - p_o) \tag{3.51}$$

in *n*-type material and as

$$R_{\text{S}} = S_n(n - n_o) \tag{3.52}$$

in *p*-type material. S_p and S_n are effective surface recombination velocities. It should be mentioned that these effective recombination velocities are not necessarily constants independent of carrier concentration, though they are commonly treated as such.

---

[2] It is unlikely that more than one trap will be involved in a single recombination event since the traps are spatially separated.
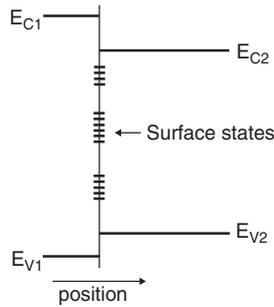
Figure 3.10 Illustration of surface states at a semiconductor surface or interface between dissimilar materials such as a semiconductor and an insulator (i.e., antireflective coating), two different semiconductors (heterojunction) or a metal and a semiconductor (Schottky contact)

3.2.7 Carrier Transport

As has already been established, electrons and holes in a semiconductor behave much like a free particle of the same electronic charge with effective masses of m_n^* and m_p^*, respectively. Thus, they are subject to the classical processes of drift and diffusion. Drift is a charged particle's response to an applied electric field. When an electric field is applied across a uniformly doped semiconductor, the bands bend upward in the direction of the applied electric field. Electrons in the conduction band, being negatively charged, move in the opposite direction to the applied field and holes in the valence band, being positively charged, move in the same direction as the applied field (Figure 3.11) – in other words, electrons *sink* and holes *float*. This is a useful conceptual tool for analyzing the motion of holes and electrons in semiconductor devices.

With nothing to impede their motion, the holes and electrons would continue to accelerate without bound. However, the semiconductor crystal is full of objects with which the carriers collide and are scattered. These objects include the component atoms of the crystal, dopant ions, crystal defects, and even other electrons and holes. On a microscopic scale, their motion is much like that of a ball in pinball machine, the carriers are constantly bouncing (scattering) off objects in the crystal,
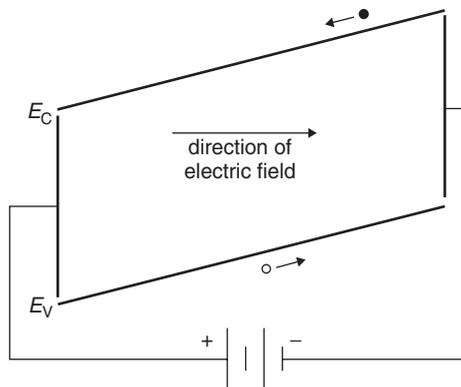

Figure 3.11 Illustration of the concept of drift in a semiconductor. Note that electrons and holes move in opposite directions. The electric field can be created by the internal built-in potential of the junction or by an externally applied bias voltage

but generally moving in the direction prescribed by the applied electric field, $\vec{E} = -\nabla\phi$, where ϕ is the electrostatic potential. The net effect is that the carriers appear to move, on a macroscopic scale, at a constant velocity, v_d, the drift velocity. The drift velocity is directly proportional to the electric field

$$|\vec{v}_d| = \left|\mu\vec{E}\right| = |\mu\nabla\phi| \tag{3.53}$$

where μ is the carrier mobility. The carrier mobility is generally independent of the electric field strength unless the field is very strong, a situation not typically encountered in solar cells. The drift current densities for holes and electrons can be written as

$$\vec{J}_p^{\text{drift}} = qp\vec{v}_{\text{d},p} = q\mu_p p\vec{E} = -q\mu_p p\nabla\phi \tag{3.54}$$

and

$$\vec{J}_n^{\text{drift}} = -qn\vec{v}_{\text{d},n} = q\mu_n n\vec{E} = -q\mu_n n\nabla\phi. \tag{3.55}$$

The most significant scattering mechanisms in solar cells are lattice (phonon) and ionized impurity scattering. These component mobilities can be written as

$$\mu_{\text{L}} = C_{\text{L}}T^{-3/2} \tag{3.56}$$

for lattice scattering and as

$$\mu_{\text{I}} = \frac{C_{\text{I}}T^{3/2}}{N_{\text{D}}^+ + N_{\text{A}}^-} \tag{3.57}$$

for ionized impurity scattering. These can then be combined using Mathiessen's rule to give the carrier mobility [14]

$$\frac{1}{\mu} = \frac{1}{\mu_{\text{L}}} + \frac{1}{\mu_{\text{I}}}. \tag{3.58}$$

This is a first-order approximation that neglects the velocity dependencies of the scattering mechanisms. These two types of mobility can be distinguished experimentally by their different dependencies on temperature and doping. A better approximation is [14]

$$\mu = \mu_{\text{L}}\left[1 + \left(\frac{6\mu_{\text{L}}}{\mu_{\text{I}}}\right)\left(\text{Ci}\left(\frac{6\mu_{\text{L}}}{\mu_{\text{I}}}\right)\cos\left(\frac{6\mu_{\text{L}}}{\mu_{\text{I}}}\right) + \left[\text{Si}\left(\frac{6\mu_{\text{L}}}{\mu_{\text{I}}}\right) - \frac{\pi}{2}\right]\sin\left(\frac{6\mu_{\text{L}}}{\mu_{\text{I}}}\right)\right)\right], \tag{3.59}$$

where Ci and Si (not to be confused with the abbreviation for silicon) are the cosine and sine integrals, respectively.

When modeling solar cells, it is more convenient to use measured data or empirical formulas. Carrier mobilities in Si at 300 K are well approximated by [14]

$$\mu_n = 92 + \frac{1268}{1 + \left(\dfrac{N_{\text{D}}^+ + N_{\text{A}}^-}{1.3 \times 10^{17}}\right)^{0.91}}\text{cm}^2/\text{V} - \text{s} \tag{3.60}$$

$$\mu_p = 54.3 + \frac{406.9}{1 + \left(\dfrac{N_{\text{D}}^+ + N_{\text{A}}^-}{2.35 \times 10^{17}}\right)^{0.88}}\text{cm}^2/\text{V} - \text{s} \tag{3.61}$$

and are plotted in Figure 3.12. At low impurity levels, the mobility is governed by intrinsic lattice scattering, while at high levels the mobility is governed by ionized impurity scattering.
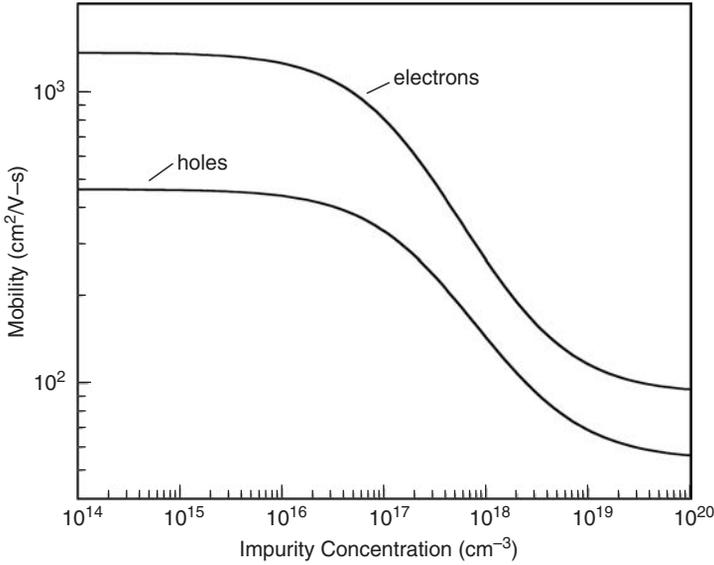
Figure 3.12   Electron and hole mobilities in silicon for $T = 300\,\text{K}$

Electrons and holes in semiconductors tend, as a result of their random thermal motion, to move (diffuse) from regions of high concentration to regions of low concentration. Much like the way the air in a balloon is distributed evenly within the volume of the balloon, carriers, in the absence of any external forces, will also tend to distribute themselves evenly within the semiconductor. This process is called *diffusion* and the diffusion current densities are given by

$$\vec{J}_p^{\text{diff}} = -qD_p\nabla p \tag{3.62}$$

$$\vec{J}_n^{\text{diff}} = qD_n\nabla n \tag{3.63}$$

where D_p and D_n are the hole and electron diffusion coefficients, respectively. Note that the currents are driven by the gradient of the carrier densities.

In thermal equilibrium, there can be no net hole current and no net electron current – in other words, the drift and diffusion currents must exactly balance. In nondegenerate materials, this leads to the Einstein relationship

$$\frac{D}{\mu} = \frac{kT}{q} \tag{3.64}$$

and allows the diffusion coefficient to be directly computed from the mobility. Generalized forms of the Einstein relationship, valid for degenerate materials, are

$$\frac{D_n}{\mu_n} = \frac{1}{q}n\left[\frac{dn}{dE_{\text{F}}}\right]^{-1} \tag{3.65}$$

and

$$\frac{D_p}{\mu_p} = -\frac{1}{q}p\left[\frac{dp}{dE_{\text{F}}}\right]^{-1}. \tag{3.66}$$

The diffusion coefficient actually increases when degeneracy effects come into play.

The total hole and electron currents (vector quantities) are the sum of their drift and diffusion components

$$\vec{J}_p = \vec{J}_p^{\text{drift}} + \vec{J}_p^{\text{diff}} = q\mu_p p\vec{E} - qD_p\nabla p = -q\mu_p p\nabla\phi - qD_p\nabla p \tag{3.67}$$

$$\vec{J}_n = \vec{J}_n^{\text{drift}} + \vec{J}_n^{\text{diff}} = q\mu_n n\vec{E} + qD_n\nabla n = -q\mu_n n\nabla\phi + qD_n\nabla n \tag{3.68}$$

The total current is then

$$\vec{J} = \vec{J}_p + \vec{J}_n + \vec{J}_{\text{disp}} \tag{3.69}$$

where $\vec{J}_{\text{disp}}$ is the *displacement current* given by

$$\vec{J}_{\text{disp}} = \frac{\partial\vec{D}}{\partial t}. \tag{3.70}$$

$\vec{D} = \varepsilon\vec{E}$ is the dielectric displacement field, where ε is the electric permittivity of the semiconductor. The displacement current can be neglected in solar cells since they are static (dc) devices.

3.2.8 Semiconductor Equations

The operation of most semiconductor devices, including solar cells, can be described by the so-called semiconductor device equations, first described by Van Roosbroeck in 1950 [15]. A generalized form of these equations is given here.[3]

$$\nabla \cdot \varepsilon\vec{E} = q(p - n + N) \tag{3.71}$$

This is a form of Poisson's equation, where N is the net charge due to dopants and other trapped charges. The hole and electron continuity equations are

$$\nabla \cdot \vec{J}_p = q\left(G - R_p - \frac{\partial p}{\partial t}\right) \tag{3.72}$$

$$\nabla \cdot \vec{J}_n = q\left(R_n - G + \frac{\partial n}{\partial t}\right) \tag{3.73}$$

where G is the optical generation rate of electron–hole pairs. Thermal generation is included in R_p and R_n. The hole and electron current densities are given by (Equations 3.67 and 3.68)

$$\vec{J}_p = -q\mu_p p\nabla(\phi - \phi_p) - kT\mu_p\nabla p \tag{3.74}$$

and

$$\vec{J}_n = -q\mu_n n\nabla(\varphi + \varphi_n) + kT\mu_n\nabla n. \tag{3.75}$$

Two new terms, ϕ_p and ϕ_n, have been introduced here. These are the so-called band parameters that account for degeneracy and a spatially varying bandgap (heterostructure solar cells) and electron affinity [17]. These terms were ignored in the preceding discussion and can usually be ignored in nondegenerate homostructure solar cells.

---

[3] In some photovoltaic materials such as GaInN, polarization is important and Poisson's equation becomes $\nabla \cdot (\varepsilon\vec{E} + \vec{P}) = q(p - n + N)$, where $\vec{P}$ is the polarization [16].

The intent here is to derive a simple analytic expression for the current–voltage characteristic of a solar cell, and so some simplifications are in order. It should be noted, however, that a complete description of the operation of solar cells can be obtained by solving the full set of coupled partial differential equations, Equations (3.71–3.75). The numerical solution of these equations is briefly addressed later in this chapter.

3.2.9 Minority-carrier Diffusion Equation

In a uniformly doped semiconductor, the bandgap and electric permittivity are independent of position. Since the doping is uniform, the carrier mobilities and diffusion coefficients are also independent of position. As we are mainly interested in the steady-state operation of the solar cell, the semiconductor equations reduce to

$$\frac{d\vec{E}}{dx} = \frac{q}{\varepsilon}(p - n + N_{\mathrm{D}} - N_{\mathrm{A}}) \tag{3.76}$$

$$q\mu_p \frac{d}{dx}(p\vec{E}) - qD_p\frac{d^2 p}{dx^2} = q(G - R) \tag{3.77}$$

and

$$q\mu_n \frac{d}{dx}(n\vec{E}) + qD_n\frac{d^2 n}{dx^2} = q(R - G) \tag{3.78}$$

In regions sufficiently far from the pn-junction of the solar cell (quasi-neutral regions), the electric field is very small. When considering the minority carrier (holes in the n-type material and electrons in the p-type material) and low-level injection ($\Delta p = \Delta n \ll N_{\mathrm{D}}, N_{\mathrm{A}}$), the drift current can be neglected with respect to the diffusion current. Under low-level injection, R simplifies to

$$R = \frac{n_P - n_{Po}}{\tau_n} = \frac{\Delta n_P}{\tau_n} \tag{3.79}$$

in the p-type region and to

$$R = \frac{p_N - p_{No}}{\tau_p} = \frac{\Delta p_N}{\tau_p} \tag{3.80}$$

in the n-type region. Δp_N and Δn_P are the excess minority-carrier concentrations. The minority-carrier lifetimes, τ_n and τ_p, are given by Equation (3.49). For clarity, the capitalized subscripts, P and N, are used to indicate quantities in p-type and n-type regions, respectively, when it may not be otherwise apparent. Lowercase subscripts, p and n, refer to quantities associated with minority holes and electrons, respectively. For example, Δn_P is the minority electron concentration in the p-type material.

Thus, Equations (3.77) and (3.78) each reduce to what is commonly referred to as the *minority-carrier diffusion equation*. It can be written as

$$D_p\frac{d^2 \Delta p_N}{dx^2} - \frac{\Delta p_N}{\tau_p} = -G(x) \tag{3.81}$$

in n-type material and as

$$D_n\frac{d^2 \Delta n_P}{dx^2} - \frac{\Delta n_P}{\tau_n} = -G(x) \tag{3.82}$$

in *p*-type material. The minority-carrier diffusion equation is often used to analyze the operation of semiconductor devices, including solar cells, and will be used in this way later in this chapter.

3.2.10 *pn*-junction Diode Electrostatics

Where an *n*-type semiconductor comes into contact with a *p*-type semiconductor, a *pn*-junction is formed. In thermal equilibrium there is no net current flow and by definition the Fermi energy must be independent of position. Since there is a concentration difference of holes and electrons between the two types of semiconductors, holes diffuse from the *p*-type region into the *n*-type region and, similarly, electrons from the *n*-type material diffuse into the *p*-type region. As the carriers diffuse, the charged impurities (ionized acceptors in the *p*-type material and ionized donors in the *n*-type material) are uncovered – that is, they are no longer screened by the majority carrier. As these impurity charges are uncovered, an electric field (or electrostatic potential difference) is produced, which counteracts the diffusion of the holes and electrons. In thermal equilibrium, the diffusion and drift currents for each carrier type exactly balance, so there is no net current flow. The transition region between the *n*-type and the *p*-type semiconductors is called the *space-charge region*. It is also often called the *depletion region*, since it is effectively depleted of both holes and electrons. Assuming that the *p*-type and the *n*-type regions are sufficiently thick, the regions on either side of the depletion region are essentially charge-neutral (often termed *quasi-neutral*). The electrostatic potential difference resulting from the junction formation is called the *built-in voltage*, V_{bi}. It arises from the electric field created by the exposure of the positive and the negative space charge in the depletion region.

The electrostatics of this situation (assuming a single acceptor and a single donor level) are governed by Poisson's equation

$$\nabla^2 \phi = \frac{q}{\varepsilon}(n_o - p_o + N_A^- - N_D^+) \tag{3.83}$$

where ϕ is the electrostatic potential, q is magnitude of the electron charge, ε is the electric permittivity of the semiconductor, p_o is the equilibrium hole concentration, n_o is the equilibrium electron concentration, N_A^- is the ionized acceptor concentration, and N_D^+ is the ionized donor concentration. Equation (3.83) is a restatement of Equation (3.71) for the given conditions.

This equation is easily solved numerically; however, an approximate analytic solution for an abrupt *pn*-junction can be obtained that lends physical insight into the formation of the space-charge region. Figure 3.13 depicts a simple one-dimensional (1D) *pn*-junction solar cell (diode), with the metallurgical junction at $x = 0$, which is uniformly doped, with a doping density of N_D on the *n*-type side and of N_A on the *p*-type side. For simplicity, it is assumed that the each side is nondegenerately doped and that the dopants are fully ionized. In this example, the *n*-type side is assumed to be more heavily doped (n^+) than the *p*-type side.

Within the depletion region, defined by $-x_N < x < x_P$, it can be assumed that p_o and n_o are both negligible compared to $|N_A - N_D|$ so that Equation (3.83) can be simplified to

$$\nabla^2 \phi = -\frac{q}{\varepsilon} N_D, \ \text{for} -x_N < x < 0 \ \text{and}$$

$$\nabla^2 \phi = \frac{q}{\varepsilon} N_A, \quad \text{for } 0 < x < x_P \tag{3.84}$$

Outside the depletion region, charge neutrality is assumed and

$$\nabla^2 \phi = 0, \ \text{for } x \leq -x_N \ \text{and } x \geq x_P. \tag{3.85}$$
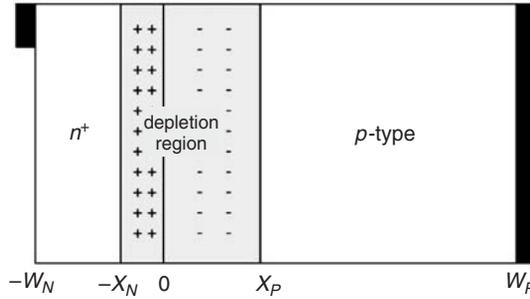
Figure 3.13 Simple solar cell structure used to analyze the operation of a solar cell. Free carriers have diffused across the junction ($x = 0$) leaving a space-charge or depletion region practically devoid of any free or mobile charges. The fixed charges in the depletion region are due to ionized donors on the n-side and ionized acceptors on the p-side

This is commonly referred to as the *depletion approximation*. The regions on either side of the depletion regions are the quasi-neutral regions.

The electrostatic potential difference across the junction is the built-in voltage, V_{bi}, and can be obtained by integrating the electric field, $\vec{E} = -\nabla \phi$.

$$\int_{-x_N}^{x_P} \vec{E}\,dx = -\int_{-x_N}^{x_P} \frac{d\phi}{dx}\,dx = -\int_{V(-x_N)}^{V(x_P)} d\phi = \phi(-x_N) - \phi(x_P) = V_{bi} \tag{3.86}$$

Solving Equations (3.84) and (3.85) and defining $\phi(x_P) = 0$, gives

$$\phi(x) = \begin{cases} V_{bi}, & x \le -x_N \\ V_{bi} - \dfrac{q N_D}{2\varepsilon}(x + x_N)^2, & -x_N < x \le 0 \\ \dfrac{q N_A}{2\varepsilon}(x - x_P)^2, & 0 \le x < x_P \\ 0, & x \ge x_P \end{cases} \tag{3.87}$$

The electrostatic potential must be continuous at $x = 0$. Therefore, from Equation (3.87),

$$V_{bi} - \frac{q N_D}{2\varepsilon}x_N^2 = \frac{q N_A}{2\varepsilon}x_P^2 \tag{3.88}$$

In the absence of any interface charge at the metallurgical junction, the electric field is also continuous at this point (really, it is the displacement field, $\vec{D} = \varepsilon\vec{E}$, that is continuous, but in this example, ε is independent of position), and

$$x_N N_D = x_P N_A \tag{3.89}$$

This is simply a statement that the total charge in either side of the depletion region exactly balance each other and therefore the depletion region extends furthest into the more lightly doped side.

Solving Equations (3.88) and (3.89) for the depletion width, W_D, gives[4]

$$W_D = x_N + x_P = \sqrt{\frac{2\varepsilon}{q} \left(\frac{N_A + N_D}{N_A N_D} \right) V_{bi}}. \tag{3.90}$$

Under nonequilibrium conditions, the electrostatic potential difference across the junction is modified by the applied voltage V which is zero in thermal equilibrium. As a consequence, the depletion width is dependent on the applied voltage,

$$W_D(V) = x_N + x_P = \sqrt{\frac{2\varepsilon}{q} \left(\frac{N_A + N_D}{N_A N_D} \right) (V_{bi} - V)}. \tag{3.91}$$

As previously stated, the built-in voltage, V_{bi}, can be calculated by noting that, under thermal equilibrium, the net hole and electron currents are zero. The hole current density is

$$\vec{J}_p = q \mu_p p_o \vec{E} - q D_p \nabla p = 0. \tag{3.92}$$

Thus, in 1D and utilizing the Einstein relationship, the electric field can be written as

$$\vec{E} = \frac{kT}{q} \frac{1}{p_o} \frac{\mathrm{d}p_o}{\mathrm{d}x} \tag{3.93}$$

Rewriting Equation (3.86) and substituting Equation (3.93) yields

$$V_{bi} = \int_{-x_N}^{x_P} E\vec{\mathrm{d}}x = \int_{-x_N}^{x_P} \frac{kT}{q} \frac{1}{p_o} \frac{\mathrm{d}p_o}{\mathrm{d}x} \mathrm{d}x = \frac{kT}{q} \int_{p_o(-x_N)}^{p_o(x_P)} \frac{\mathrm{d}p_o}{p_o} = \frac{kT}{q} \ln \left[\frac{p_o(x_P)}{p_o(-x_N)} \right] \tag{3.94}$$

Since we have assumed nondegeneracy, $p_o(x_P) = N_A$ and $p_o(-x_N) = n_i^2/N_D$. Therefore,

$$V_{bi} = \frac{kT}{q} \ln \left[\frac{N_D N_A}{n_i^2} \right]. \tag{3.95}$$

Figure 3.14 shows the equilibrium energy band diagram (a), electric field (b), and charge density (c) for a simple abrupt pn-junction silicon diode in the vicinity of the depletion region. The conduction band edge is given by $E_C(x) = E_0 - q\phi(x) - \chi$, the valence band edge by $E_V(x) = E_C(x) - E_G$, and the intrinsic energy by Equation (3.18). E_0, defined as the vacuum energy, serves as a convenient reference point and is universally constant with position. An electron at the vacuum energy is, by definition, completely free of influence from all external forces. The electron affinity χ is the minimum energy needed to free an electron from the bottom of the conduction band and take it to the vacuum level. The electric field is a result of the uncovered ionized donors and acceptors, and opposes the diffusion of electrons and holes in the quasi-neutral regions. The charge density plot illustrates the balance of charge between the two sides

---

[4] A somewhat more rigorous treatment of equation 3.89 would yield a factor of $2kT/q$ which is $\sim 50\,\mathrm{mV}$ at 300 K, or

$$W_D = \sqrt{\frac{2\varepsilon}{q} \left(\frac{N_A + N_D}{N_A N_D} \right) (V_{bi} - 2kT/q)} \quad [3].$$
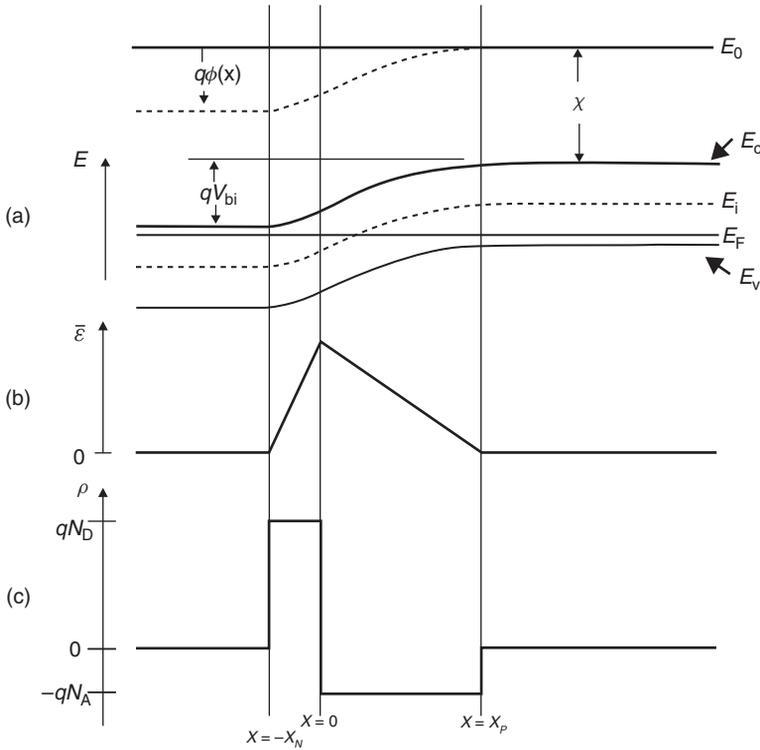
Figure 3.14 Equilibrium conditions in a solar cell: (a) energy bands; (b) electric field; (c) charge density

of the depletion region. In heterostructures, both the bandgap and the electron affinity are position-dependent – making the calculation of the junction electrostatics and energy band diagram more complex, as discussed in Section 3.4.8.

3.2.11 Summary

The fundamental physical principles relevant to solar cell operation have been reviewed and the basic solar cell structure has now been established (Figures 3.1 and 3.13). A solar cell is simply a *pn*-junction diode consisting of two quasi-neutral regions on either side of a depletion region with an electrical contact made to each quasi-neutral region. Typically, the more heavily doped quasi-neutral region is called the *emitter* (the *n*-type region in Figure 3.13) and the more lightly doped region is called the *base* (the *p*-type region in Figure 3.13). The base region is also often referred to as the *absorber region* since the emitter region is usually very thin and most of the light absorption occurs in the base. This basic structure will now serve as the basis for deriving the fundamental operating characteristics of the solar cell.

3.3 SOLAR CELL FUNDAMENTALS

The basic current–voltage characteristic of the solar cell can be derived by solving the minority-carrier diffusion equation with appropriate boundary conditions.

3.3.1 Solar Cell Boundary Conditions

In Figure 3.13, at $x = -W_N$, the usual assumption is that the front contact can be treated as an ideal ohmic contact, i.e.

$$\Delta p(-W_N) = 0. \tag{3.96}$$

However, since the front contact is usually a grid with metal contacting the semiconductor on only a small percentage of the front surface, modeling the front surface with an effective surface recombination velocity is more realistic. This effective recombination velocity models the combined effects of the ohmic contact and the antireflective passivation layer (SiO_2 in silicon solar cells). In this case, the boundary condition at $x = -W_N$ is

$$\frac{\mathrm{d}\Delta p}{\mathrm{d}x} = \frac{S_{\mathrm{F,eff}}}{D_p} \Delta p(-W_N) \tag{3.97}$$

where $S_{\mathrm{F,eff}}$ is the effective front surface recombination velocity. As $S_{\mathrm{F,eff}} \to \infty$, $\Delta p \to 0$, and the boundary condition given by Equation (3.97) reduces to that of an ideal ohmic contact (Equation 3.96). In reality, $S_{\mathrm{F,eff}}$ depends upon a number of parameters and is bias dependent. This will be discussed in more detail later.

The back contact can also be treated as an ideal ohmic contact, so that

$$\Delta n(W_P) = 0. \tag{3.98}$$

However, solar cells are often fabricated with a *back-surface field* (BSF), a thin, more heavily doped region at the back of the base region. An even more effective BSF can be created by inserting a wider bandgap semiconductor material at the back of the solar cell (a heterojunction). The BSF keeps minority carriers away from the back ohmic contact and increases their chances of being collected and it can be modeled by an effective, and relatively low, surface recombination velocity. This boundary condition is then

$$\left.\frac{\mathrm{d}\Delta n}{\mathrm{d}x}\right|_{x=W_P} = -\frac{S_{\mathrm{BSF}}}{D_n} \Delta n(W_P), \tag{3.99}$$

where S_{BSF} is the effective surface recombination velocity at the BSF.

All that remains now is to determine suitable boundary conditions at $x = -x_N$ and $x = x_P$. These boundary conditions are commonly referred to as the *law of the junction*.

Under equilibrium conditions, zero applied voltage and no illumination, the Fermi energy, E_F, is constant with position. When a bias voltage is applied, it is convenient to introduce the concept of quasi-Fermi energies. It was shown earlier that the equilibrium carrier concentrations could be related to the Fermi energy (Equations 3.14 and 3.15). Under nonequilibrium conditions, similar relationships hold. Assuming the semiconductor is nondegenerate,

$$p = n_i \mathrm{e}^{(E_i - F_P)/kT} \tag{3.100}$$

and

$$n = n_i \mathrm{e}^{(F_N - E_i)/kT} \tag{3.101}$$

It is evident that, under equilibrium conditions, $F_P = F_N = E_F$. Under nonequilibrium conditions, assuming that the majority carrier concentrations at the contacts retain their equilibrium

values, the applied voltage can be written as

$$qV = F_N(-W_N) - F_P(W_P) \tag{3.102}$$

Since, in low-level injection, the majority carrier concentrations are essentially constant throughout the quasi-neutral regions, that is, $p_P(x_P \le x \le W_P) = N_A$ and $n_N(-W_N \le x \le -x_N) = N_D$, $F_N(-W_N) = F_N(-x_N)$ and $F_P(W_P) = F_P(x_P)$. Then, assuming that both the quasi-Fermi energies remain constant inside the depletion region,

$$qV = F_N(x) - F_P(x) \tag{3.103}$$

for $-x_N \le x \le x_P$, that is, everywhere inside the depletion region. Using Equations (3.100) and (3.101), this leads directly to the *law of the junction*, the boundary conditions used at the edges of the depletion region,

$$p_N(-x_N) = \frac{n_i^2}{N_D} e^{qV/kT} \tag{3.104}$$

and

$$n_P(x_P) = \frac{n_i^2}{N_A} e^{qV/kT}. \tag{3.105}$$

3.3.2 Generation Rate

For light incident at the front of the solar cell, $x = -W_N$, the optical generation rate takes the form (see Equation 3.34)

$$G(x) = (1-s) \int_\lambda (1 - r(\lambda)) f(\lambda) \alpha(\lambda) e^{-\alpha(x+W_N)} \, d\lambda. \tag{3.106}$$

Essentially, only photons with $\lambda \le hc/E_G$ contribute to the generation rate.

3.3.3 Solution of the Minority-carrier Diffusion Equation

Using the boundary conditions defined by Equations (3.97), (3.99), (3.104), and (3.105) and the generation rate given by Equation (3.106), the solution to the minority-carrier diffusion equation, Equations (3.81) and (3.82), is easily shown to be

$$\Delta p_N(x) = A_N \sinh[(x + x_N)/L_p] + B_N \cosh[(x + x_N)/L_p] + \Delta p'_N(x) \tag{3.107}$$

in the n-type region and

$$\Delta n_P(x) = A_P \sinh[(x - x_P)/L_n] + B_P \cosh[(x - x_P)/L_n] + \Delta n'_P(x) \tag{3.108}$$

in the p-type region. The particular solutions, $\Delta p'_N(x)$ and $\Delta n'_P(x)$, due to $G(x)$ are given by

$$\Delta p'_N(x) = -(1-s) \int_\lambda \frac{\tau_p}{(L_p^2 \alpha^2 - 1)} [1 - r(\lambda)] f(\lambda) \alpha(\lambda) e^{-\alpha(x+W_N)} \, d\lambda \tag{3.109}$$

and

$$\Delta n'_P(x) = -(1-s) \int_\lambda \frac{\tau_n}{(L_n^2 \alpha^2 - 1)} [1 - r(\lambda)] f(\lambda) \alpha(\lambda) e^{-\alpha(x+W_N)} \, d\lambda. \tag{3.110}$$

Using the boundary conditions set above, A_N, B_N, A_P, and B_P in Equations (3.107) and (3.108) are readily solved for and are needed to obtain the diode current–voltage ($I-V$) characteristics.

3.3.4 Derivation of the Solar Cell $I-V$ Characteristic

The minority-carrier current densities in the quasi-neutral regions are just the diffusion currents, because the electric field is negligible. Using the active sign convention for the current (since a solar cell is typically thought of as a battery) gives

$$\vec{J}_{p,N}(x) = -qD_p \frac{d\Delta p_N}{dx} \tag{3.111}$$

and

$$\vec{J}_{n,P}(x) = qD_n \frac{d\Delta n_P}{dx} \tag{3.112}$$

The total current is given by

$$I = A[J_p(x) + J_n(x)] \tag{3.113}$$

and is true everywhere within the solar cell (A is the area of the solar cell). Equations (3.111) and (3.112) give only the hole current in the n-type region and the electron current in the p-type region, not both at the same point. However, integrating Equation (3.73), the electron continuity equation, over the depletion region, gives

$$\int_{-x_N}^{x_P} \frac{d\vec{J}_n dx}{dx} dx = \vec{J}_n(x_P) - \vec{J}_n(-x_N) = q \int_{-x_N}^{x_P} [R(x) - G(x)]dx \tag{3.114}$$

$G(x)$ is easily integrated and the integral of the recombination rate can be approximated by assuming that the recombination rate is constant within the depletion region and is $R(x_m)$ where x_m is the point at which $p_D(x_m) = n_D(x_m)$ and corresponds to the maximum recombination rate in the depletion region. If recombination via a midgap single level trap is assumed, then, from Equations (3.36), (3.100), (3.101), and (3.103), the recombination rate in the depletion region is

$$R_D = \frac{p_D n_D - n_i^2}{\tau_n(p_D + n_i) + \tau_p(n_D + n_i)} = \frac{n_D^2 - n_i^2}{(\tau_n + \tau_p)(n_D + n_i)} = \frac{n_D - n_i}{(\tau_n + \tau_p)} = \frac{n_i(e^{qV/2kT} - 1)}{\tau_D} \tag{3.115}$$

where τ_D is the effective lifetime in the depletion region. From Equation (3.114), $\vec{J}_n(-x_N)$, the majority carrier current at $x = -x_N$, can now be written as

$$\vec{J}_n(-x_N) = \vec{J}_n(x_P) + q \int_{-x_N}^{x_P} G(x)\, dx - q \int_{-x_N}^{x_P} R_D\, dx$$

$$= \vec{J}_n(x_P) + q(1-s) \int_{\lambda} [1 - r(\lambda)]f(\lambda)e[^{-\alpha(W_N-x_N)} - e^{-\alpha(W_N+x_P)}]\, d\lambda$$

$$-q \frac{W_D n_i}{\tau_D}(e^{qV/2kT} - 1) \tag{3.116}$$

where $W_D = x_P + x_N$. Substituting into Equation (3.113), the total current is now

$$I = A\left[J_p(-x_N) + J_n(x_P) + J_D - q\frac{W_D n_i}{\tau_D}(e^{qV/2kT} - 1) \right] \tag{3.117}$$

where

$$J_{\mathrm{D}} = q(1-s) \int_{\lambda} [1 - r(\lambda)] f(\lambda)(\mathrm{e}^{-\alpha(W_N - x_N)} - \mathrm{e}^{-\alpha(W_N + x_P)})\, \mathrm{d}\lambda \tag{3.118}$$

is the generation current density from the depletion region and A is the area of the solar cell. The last term of Equation (3.117) represents recombination in the space-charge region.

The solutions to the minority-carrier diffusion equation, Equations (3.107) and (3.108), can be used to evaluate the minority-carrier current densities, Equations (3.111) and (3.112). These can then be substituted into Equation (3.117), which, with some algebraic manipulation, yields the solar cell current–voltage characteristic

$$I = I_{\mathrm{SC}} - I_{o1}(\mathrm{e}^{qV/kT} - 1) - I_{o2}(\mathrm{e}^{qV/2kT} - 1). \tag{3.119}$$

where I_{SC} is the short-circuit current and is the sum of the contributions from each of the three regions: the n-type region (I_{SCN}), the depletion region ($I_{\mathrm{SCD}} = AJ_{\mathrm{D}}$), and the p-type region (I_{SCP})

$$I_{\mathrm{SC}} = I_{\mathrm{SCN}} + I_{\mathrm{SCD}} + I_{\mathrm{SCP}} \tag{3.120}$$

where

$$I_{\mathrm{SCN}} = qAD_p \left[\frac{\Delta p'(-x_N)T_{p1} - S_{\mathrm{F,eff}}\Delta p'(-W_N) + D_p \left.\dfrac{\mathrm{d}\Delta p'}{\mathrm{d}x}\right|_{x=-W_N}}{L_p T_{p2}} - \left.\frac{\mathrm{d}\Delta p'}{\mathrm{d}x}\right|_{x=-x_N} \right] \tag{3.121}$$

with

$$T_{p1} = D_p/L_p \sinh[(W_N - x_N/L_p] + S_{\mathrm{F,eff}} \cosh[(W_N - x_N/L_p] \tag{3.122}$$

$$T_{p2} = D_p/L_p \cosh[(W_N - x_N/L_p] + S_{\mathrm{F,eff}} \sinh[(W_N - x_N/L_p] \tag{3.123}$$

and

$$I_{\mathrm{SCP}} = qAD_n \left[\frac{\Delta n'(x_P)T_{n1} - S_{\mathrm{BSF}}\Delta n'(W_P) - D_n \left.\dfrac{\mathrm{d}\Delta n'}{\mathrm{d}x}\right|_{x=W_P}}{L_n T_{n2}} + \left.\frac{\mathrm{d}\Delta n'}{\mathrm{d}x}\right|_{x=x_P} \right] \tag{3.124}$$

with

$$T_{n1} = D_n/L_n \sinh[(W_P - x_P)/L_n] + S_{\mathrm{BSF}} \cosh[(W_P - x_P)/L_n] \tag{3.125}$$

$$T_{n2} = D_n/L_n \cosh[(W_P - x_P)/L_n] + S_{\mathrm{BSF}} \sinh[(W_P - x_P)/L_n] \tag{3.126}$$

I_{o1} is the dark saturation current due to recombination in the quasi-neutral regions,

$$I_{o1} = I_{o1,p} + I_{o1,n} \tag{3.127}$$

with

$$I_{o1,p} = qA \frac{n_{\mathrm{i}}^2}{N_{\mathrm{D}}} \frac{D_p}{L_p} \left\{ \frac{D_p/L_p \sinh[(W_N - x_N)/L_p] + S_{\mathrm{F,eff}} \cosh[(W_N - x_N/L_p]}{D_p/L_p \cosh[(W_N - x_N)/L_p] + S_{\mathrm{F,eff}} \sinh[(W_N - x_N/L_p]} \right\} \tag{3.128}$$

and

$$I_{o1,n} = qA\frac{n_i^2}{N_A}\frac{D_n}{L_n}\left\{\frac{D_n/L_n \sinh[(W_P - x_P)/L_n] + S_{BSF}\cosh[(W_P - x_P/L_n]}{D_n/L_n \cosh[(W_P - x_P)/L_n] + S_{BSF}\sinh[(W_P - x_P/L_n]}\right\} \quad (3.129)$$

These are very general expressions for the dark saturation current and reduce to more familiar forms when appropriate assumptions are made, as will be seen later.

I_{o2} is the dark saturation current due to recombination in the space-charge region,

$$I_{o2} = qA\frac{W_D n_i}{\tau_D} \quad (3.130)$$

and is bias-dependent since the depletion width, W_D, is a function of the applied voltage (Equation 3.91).

3.3.5 Interpreting the Solar Cell $I-V$ Characteristic

Equation (3.119), repeated here, is a general expression for the current produced by a solar cell.

$$I = I_{SC} - I_{o1}(e^{qV/kT} - 1) - I_{o2}(e^{qV/2kT} - 1) \quad (3.131)$$

The short-circuit current and dark saturation currents are given by rather complex expressions (Equations 3.120, 3.127, 3.128, 3.129, and 3.130) that depend on the solar cell structure, material properties, and the operating conditions. A full understanding of solar cell operation requires detailed examination of these terms. However, much can be learned about solar cell operation by examining the basic form of Equation (3.131). From a circuit perspective, it is apparent that a solar cell can be modeled by an ideal current source I_{SC} in parallel with two diodes – one with an ideality factor of 1 and the other with an ideality factor of 2, as shown in Figure 3.15. Note that the direction of the current source is such that it serves to forward-bias the diodes.

The current–voltage ($I-V$) characteristic of a typical silicon solar cell is plotted in Figure 3.16 for the parameter values given in Table 3.2. Note that it is the minority-carrier properties which determine the solar cell behavior, as indicated by Equations (3.119–3.129). For simplicity, the dark current due to the depletion region (diode 2) has been ignored (a reasonable and common assumption for a good solar cell, especially at larger forward biases). It illustrates several important figures of merit for solar cells – the short-circuit current, the open-circuit voltage,
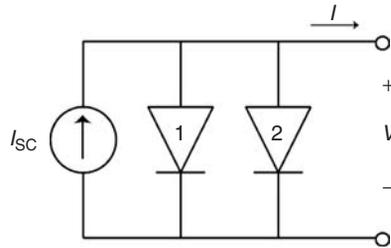

Figure 3.15 Simple solar cell circuit model. Diode 1 represents the recombination current in the quasi-neutral regions ($\propto e^{qV/kT}$), while diode 2 represents recombination in the depletion region ($\propto e^{qV/2kT}$)
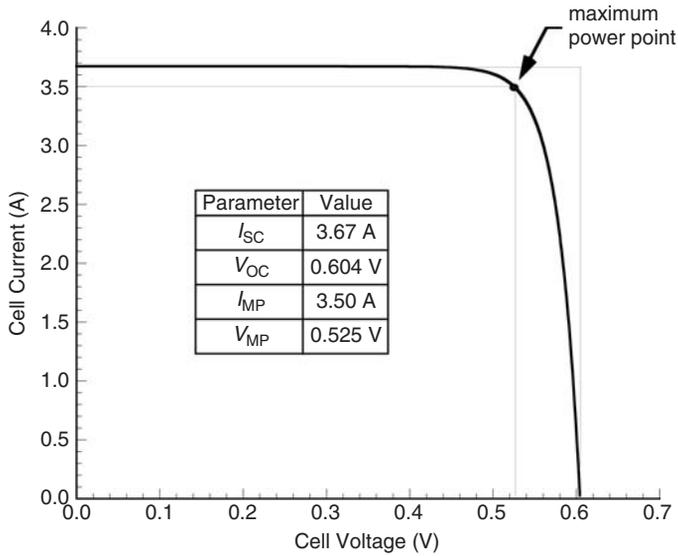
Parameter	Value
I_{SC}	3.67 A
V_{OC}	0.604 V
I_{MP}	3.50 A
V_{MP}	0.525 V

Figure 3.16 Current–voltage characteristic calculated for the silicon solar cell defined by Table 3.2 (area $A = 100 \text{ cm}^2$)

Table 3.2 Si solar cell model parameters

Parameter	n-type Si emitter	p-type Si base
Thickness	$W_N = 0.35\,\mu\text{m}$	$W_P = 300\,\mu\text{m}$
Doping density	$N_D = 1 \times 10^{20} \text{ cm}^{-3}$	$N_A = 1 \times 10^{15} \text{ cm}^{-3}$
Surface recombination	$D_p = 1.5 \text{ cm}^{-2}/\text{V s}$	$D_n = 35 \text{ cm}^{-2}/\text{V s}$
Minority-carrier diffusivity	$S_{F,\text{eff}} = 3 \times 10^4 \text{ cm/s}$	$S_{BSF} = 100 \text{ cm/s}$
Minority-carrier lifetime	$\tau_p = 1\,\mu\text{s}$	$\tau_n = 350\,\mu\text{s}$
Minority-carrier diffusion length	$L_p = 12\,\mu\text{m}$	$L_n = 1100\,\mu\text{m}$

and the fill factor. At small applied voltages, the diode current is negligible and the current is just the short-circuit current, I_{SC}, as can be seen when V is set to zero in Equation (3.131). When the applied voltage is high enough so that the diode current (recombination current) becomes significant, the solar cell current drops quickly.

Table 3.2 shows the huge asymmetry between the n-emitter and the p-base in a typical solar cell. The emitter is ~1000 times thinner, 10 000 times more heavily doped, and its diffusion length is ~100 times shorter than the corresponding quantities in the base.

At open-circuit ($I = 0$), all the light-generated current I_{SC} is flowing through diode 1 (diode ignored, as assumed above), so the open-circuit voltage can be written as

$$V_{OC} = \frac{kT}{q} \ln \frac{I_{SC} + I_{o1}}{I_{o1}} \approx \frac{kT}{q} \ln \frac{I_{SC}}{I_{o1}}, \tag{3.132}$$

where $I_{SC} \gg I_{o1}$.

Of particular interest is the point on the $I-V$ curve where the power produced is at a maximum. This is referred to as the *maximum power point* with $V = V_{\mathrm{MP}}$ and $I = I_{\mathrm{MP}}$. As seen in Figure 3.16, this point defines a rectangle whose area, given by $P_{\mathrm{MP}} = V_{\mathrm{MP}} I_{\mathrm{MP}}$, is the largest rectangle for any point on the $I-V$ curve. The maximum power point is found by solving

$$\left.\frac{\partial P}{\partial V}\right|_{V=V_{\mathrm{MP}}} = \left.\frac{\partial (IV)}{\partial V}\right|_{V=V_{\mathrm{MP}}} = \left.\left[I + V\frac{\partial I}{\partial V}\right]\right|_{V=V_{\mathrm{MP}}} = 0 \tag{3.133}$$

for $V = V_{\mathrm{MP}}$. The current at the maximum power point, I_{MP}, is then found by evaluating Equation (3.131) at $V = V_{\mathrm{MP}}$.

The rectangle-defined by V_{OC} and I_{SC} provides a convenient reference for describing the maximum power point. The fill factor, FF, is a measure of the *squareness* of the $I-V$ characteristic and is always less than one. It is the ratio of the areas of the two rectangles shown in Figure 3.16 or

$$FF = \frac{V_{\mathrm{MP}} I_{\mathrm{MP}}}{V_{\mathrm{OC}} I_{\mathrm{SC}}} = \frac{P_{\mathrm{MP}}}{V_{\mathrm{OC}} I_{\mathrm{SC}}}. \tag{3.134}$$

Arguably, the most important figure of merit for a solar cell is its power conversion efficiency, η, which is defined as

$$\eta = \frac{P_{\mathrm{MP}}}{P_{\mathrm{in}}} = \frac{FF V_{\mathrm{OC}} I_{\mathrm{SC}}}{P_{\mathrm{in}}} \tag{3.135}$$

The incident power P_{in} is determined by the properties of the light spectrum incident upon the solar cell. Further information regarding experimental determination of these parameters appears in Chapter 18.

Another important figure of merit is the collection efficiency, which can be defined relative to both optical and recombination losses as an *external* collection efficiency

$$\eta_{\mathrm{C}}^{\mathrm{ext}} = \frac{I_{\mathrm{SC}}}{I_{inc}} \tag{3.136}$$

where

$$I_{inc} = qA \int_{\lambda<\lambda_G} f(\lambda)\, d\lambda \tag{3.137}$$

is the maximum possible photocurrent that would result if all photons with $E > E_G$ ($\lambda < \lambda_G = hc/E_G$) created electron–hole pairs that were collected. The collection efficiency can also be defined with respect to recombination losses as the internal collection efficiency

$$\eta_{\mathrm{C}}^{\mathrm{int}} = \frac{I_{\mathrm{SC}}}{I_{\mathrm{gen}}} \tag{3.138}$$

where

$$I_{\mathrm{gen}} = qA(1-s) \int_{\lambda<\lambda_G} [1 - r(\lambda) f(\lambda)(1 - e^{-\alpha(W_N + W_P)})\, d\lambda \tag{3.139}$$

is the light-generated current. This represents what the short-circuit current would be if every photon that is absorbed is collected and contributes to the short-circuit current. $I_{\mathrm{gen}} = I_{\mathrm{inc}}$ when there is no grid shadowing, no reflective losses, and the solar cell has infinite optical thickness.

3.3.6 Properties of Efficient Solar Cells

Using these figures of merit, the properties of a good (efficient) solar cell can be ascertained. From Equation (3.135), it is clear that an efficient solar cell will have a high short-circuit current I_{SC}, a high open-circuit voltage, V_{OC} and a fill factor FF as close as possible to 1. A more detailed understanding of what influences the solar cell efficiency can be obtained by rewriting the efficiency as [18]

$$\eta = \frac{P_{max}}{P_{in}} = \eta_{ideal}\eta_{photon}FF\eta_V\eta_C^{int},$$
(3.140)

where FF and η_C^{int} have been previously defined (Equations 3.134 and 3.138, respectively) and η_{ideal}, η_{photon}, and η_V are defined below.

Assuming the maximum energy that can be extracted from an absorbed photon is E_G, the ideal efficiency can be expressed as

$$\eta_{ideal}(E_G) = \frac{\frac{1}{q}E_G I_{inc}}{P_{in}} = \frac{E_G}{(P_{in}/A)}\int_{\lambda<\lambda_G} f(\lambda)\,d\lambda.$$
(3.141)

Since only photons with $h\nu > E_G$ can create electron–hole pairs and contribute to the output power of the solar cell, it is clear that the bandgap determines how well the solar cell is coupled to the solar spectrum. A simple analysis can be performed to predict the ideal efficiency. This is plotted in Figure 3.17 for an AM1.5 global spectrum and shows a maximum efficiency of 48% at about $E_G = 1.1$ eV, close to the bandgap of silicon, although bandgaps between 1.0 and 1.6 eV have comparable ideal efficiencies.
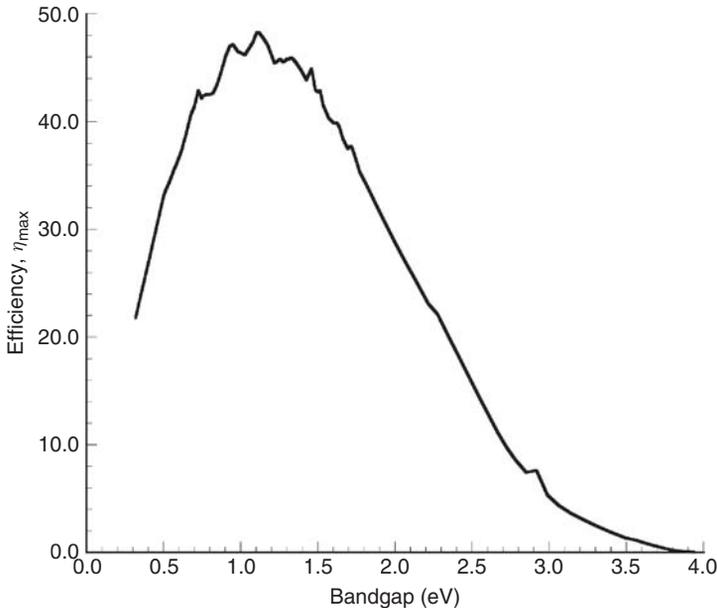

Figure 3.17 Ideal efficiency as a function of semiconductor band gap for an AM1.5 global spectrum

Of course, this assumes that $V_{OC} = \frac{1}{q} E_G$ and $FF = 1$, which are obvious exaggerations. Perfect light trapping is also assumed so that $I_{SC} = I_{inc}$, but that is a more realistic prospect. However, this quantity does serve to set an upper bound on the efficiency of a single-junction solar cell. Of course, multijuction photovoltaic systems will have a higher ideal efficiency. More complete analyses of the theoretical limits of solar cells are given elsewhere [19–21] and are also discussed in Chapter 4 of this handbook.

The photon efficiency η_{photon} accounts for photons that are reflected, transmitted through, or otherwise not absorbed in the solar cell and can be written as

$$\eta_{photon} = \frac{I_{gen}}{I_{inc}} = \frac{\eta_C^{ext}}{\eta_C^{int}}. \tag{3.142}$$

To maximize η_{photon} ($\eta_{photon} \to 1$ when $I_{gen} \to I_{inc}$ or, equivalently, $\eta_C^{ext} \to \eta_C^{int}$), the solar cell should be designed with a minimum amount of grid shadowing s, minimum reflectance $r(\lambda)$, and be optically thick enough such that nearly all the photons with $E > E_G$ are absorbed.

A transcendental relationship between V_{OC} and V_{MP} can be obtained from the solution of Equation (3.131) for the single-diode model, from which the following semi-empirical expression for the fill factor can be extracted [22]

$$FF = \frac{V_{OC} - \frac{kT}{q} \ln[q V_{OC}/kT + 0.72]}{V_{OC} + kT/q}. \tag{3.143}$$

It can be seen that FF is a weak function of the open-circuit voltage, increasing slowly as the open-circuit voltage increases. This expression neglects any series and shunt resistances which tend to degrade the fill factor, as will be discussed later in this chapter.

The voltage efficiency η_V is the ratio of the open-circuit voltage to the bandgap voltage

$$\eta_V = \frac{V_{OC}}{\frac{1}{q} E_G}. \tag{3.144}$$

Empirically, the best solar cells have an open-circuit voltage approximately $0.4\,\text{V}$ less than the bandgap voltage (no solar concentration). For silicon, this gives $\eta_V = 0.643$. It is clearly desirable to have the open-circuit voltage approach the bandgap voltage and this is one of the challenges in the development of next generation solar cells. At open-circuit, since there is no flow of carriers out of the devices, every electron–hole pair must recombine. It is the rate of this recombination, or the reverse saturation current, that constrains the open-circuit voltage. The open-circuit voltage (Equation 3.132)

$$V_{OC} \approx \frac{kT}{q} \ln \frac{I_{SC}}{I_{o1}} \tag{3.145}$$

is logarithmically proportional to the short-circuit current and to the reciprocal of the reverse saturation current I_{o1}. Therefore, reducing the saturation current will increase the open-circuit voltage. From Equations (3.128) and (3.129), it is obvious that $I_{o1} \to 0$ as $\tau \to \infty$ and $S \to 0$.

The final term in Equation (3.140) is for the internal collection efficiency, which was defined previously in Equation (3.138), $I_{SC} = \eta_C^{int} I_{gen}$ is dependent on the recombination in the solar cell and will approach 1 as $\tau \to \infty$ and $S \to 0$. Voltage-dependent collection can compensate for low effective carrier lifetimes [23] in achieving a higher short-circuit current, but the open-circuit voltage and fill factor do not benefit from this effect. This effect is briefly discussed later in this chapter.

From this discussion, it can be seen that the design of an efficient solar cell has several key goals:

1. Selection of a semiconductor material with a bandgap well matched to the solar spectrum, i.e. maximizing η_{ideal}.
2. Minimizing optical losses such as grid shadowing, reflectance, and absorption in the optical components, as well as maximizing the optical thickness of the solar cell, thereby maximizing η_{photon}.
3. Minimizing series and shunt resistances in the cell and its connections, thereby maximizing the fill factor, FF.
4. Minimization of the bulk and surface recombination rates, maximizing η_V and hence the open-circuit voltage.
5. Minimization of the bulk and surface recombination rates will also maximize the internal collection efficiency η_C^{int} and hence the short-circuit current.

Simultaneous achievement of all of these goals will result in a very efficient solar cell. For a silicon solar cell with $V_{OC} = 0.72$ V, the predicted efficiency from Equation (3.140) (assuming $\eta_{photon} = \eta_C^{int} = 1$ under AM1.5 global illumination is 26.2% – just slightly higher than the best reported silicon solar cells [24].

It is evident that, despite the apparent complexity of the expressions describing the fundamental operation of solar cells, the basic operating principles are easy to understand, as illustrated by the above discussion.

3.3.7 Lifetime and Surface Recombination Effects

The solar cell characteristics previously derived (Equations 3.119–3.132) allow examination of the dependence of the solar cell performance on specific sources of recombination. Figure 3.18 shows how the base minority-carrier lifetime affects V_{OC}, I_{SC}, and the FF. Unless otherwise stated, the parameters of Table 3.2 are used to compute the solar cell performance. Short lifetimes mean that the diffusion length in the base is much less than the base thickness and carriers created deeper than about one diffusion length in the base are unlikely to be collected. When this is true ($L_n \ll W_P$), the contribution to the dark saturation current in the base (Equation 3.129) becomes

$$I_{o1,n} = qA \frac{n_i^2}{N_A} \frac{D_n}{L_n} \tag{3.146}$$

and is commonly referred to as the *long-base approximation*. In this case, the BSF has no effect on the dark saturation current. On the other hand, when the base minority-carrier lifetime is long ($L_n \gg W_P$), the carriers readily come in contact with the BSF and the dark saturation current is a strong function of S_{BSF}

$$I_{o1,n} = qA \frac{n_i^2}{N_A} \frac{D_n}{(W_P - X_P)} \frac{S_{BSF}}{S_{BSF} + D_n/(W_P - x_P)} \tag{3.147}$$

When S_{BSF} is very large (i.e. no BSF), this reduces to the more familiar short-base approximation

$$I_{o1,n} = qA \frac{n_i^2}{N_A} \frac{D_n}{(W_P - x_P)}. \tag{3.148}$$

Figure 3.19 shows how S_{BSF} affects V_{OC}, I_{SC}, and FF. Notice that the break point in the curves occurs when $S_{BSF} \approx D_n/W_P = 1000$ cm/s, as can be inferred from Equation (3.147).
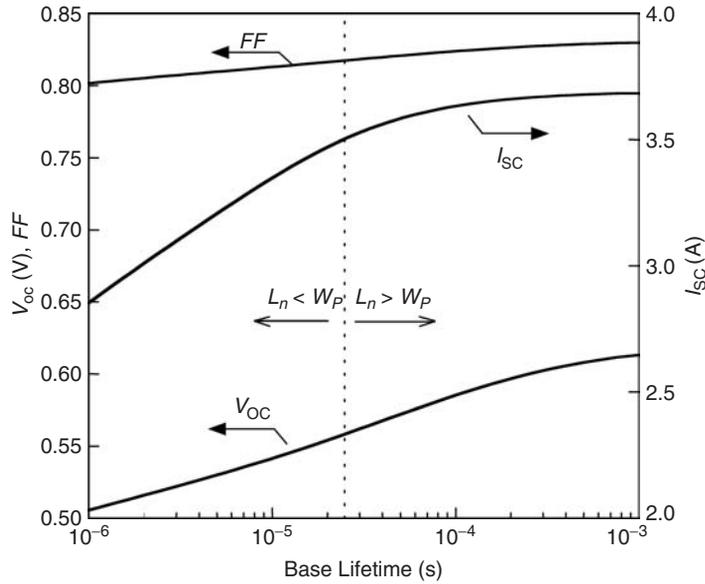
Figure 3.18 Effect of base lifetime on solar cell performance for the solar cell parameters in Table 3.2. The minority-carrier diffusion length ($L_n = \sqrt{D_n \tau_n}$) is equal to the base thickness (W_P) when $\tau_n = 25.7\,\mu s$

       Front surface recombination for solar cells with contact grids on the front of the device is really an average over the front surface area of the relatively low surface recombination velocity between the grid lines and the very high surface recombination velocity of the ohmic contact. An expression for the effective front surface recombination velocity is given by [25]

$$S_{F,eff} = \frac{(1-s)S_F \bar{G}_N \tau_p \left(\cosh \dfrac{W_N}{L_p} - 1 \right) + p_o(e^{qV/A_okT} - 1) \left[s\dfrac{D_p}{L_p} \dfrac{\cosh \dfrac{W_N}{L_p}}{\sinh \dfrac{W_N}{L_p}} + S_F \right]}{(1-s) \left[p_o(e^{qV/A_okT} - 1) + \bar{G}_N \tau_p \left(\cosh \dfrac{W_N}{L_p} - 1 \right) \right]}$$

(3.149)

where S_F is the surface recombination velocity between the grid lines and $\bar{G}$ is the average generation rate in the emitter region. It is obvious that $S_{F,eff}$ is dependent upon the solar cell operation point. This is better seen in Table 3.3 where some special cases are illustrated (assuming $L_p \gg W_N$).

3.4 ADDITIONAL TOPICS

3.4.1 Spectral Response

The spectral response, $SR(\lambda)$, of a solar cell permits an examination of how photons of different wavelengths (energy) contribute to the short-circuit current. Just as the collection efficiency can be
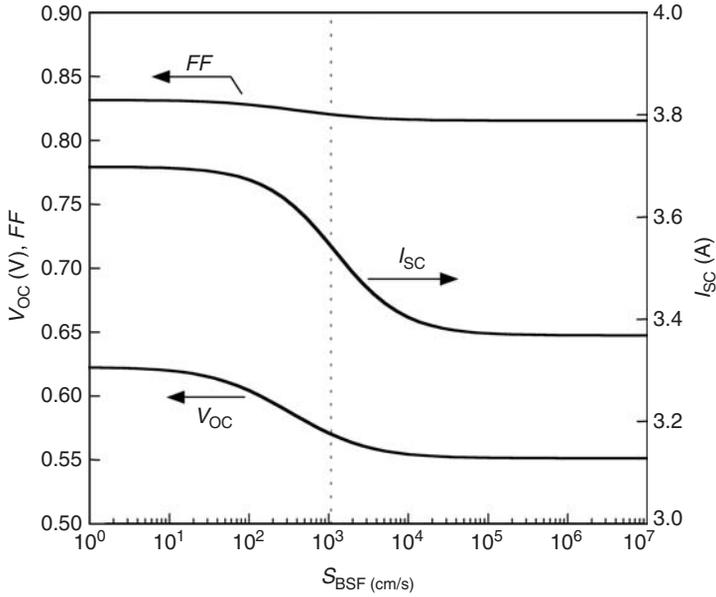
Figure 3.19 Effect of the back-surface field recombination velocity on solar cell performance. All other parameters are from Table 3.2

Table 3.3   Special cases of $S_{F,eff}$

No grid ($s = 0$)	$S_{F,eff} = S_F$
Full grid ($s = 1$)	$S_{F,eff} \to \infty$
Dark ($\overline{G} = 0$)	$S_{F,eff} = \dfrac{S_F + s D_p / W_N}{1 - s}$
Short-circuit (V = 0)	$S_{F,eff} = S_F$
V large ($\approx V_{OC}$)	$S_{F,eff} = \dfrac{S_F + s D_p / W_N}{1 - s}$

measured as either an external or internal collection efficiency, so can the spectral response. The spectral response is defined as the *short-circuit current* $I_{SC}(\lambda)$, resulting from a single wavelength of light normalized by the maximum possible current. The external spectral response is defined as

$$SR_{ext} = \frac{I_{SC}(\lambda)}{qAf(\lambda)} \tag{3.150}$$

and the internal spectral response as

$$SR_{int} = \frac{I_{SC}(\lambda)}{qA(1 - s)(1 - r(\lambda))f(\lambda)(e^{-\alpha(\lambda)W_{opt}} - 1)}, \tag{3.151}$$

where W_{opt} is the optical thickness of the solar cell (technically, also a function of wavelength). Experimentally, the external spectral response is measured. The internal spectral response is
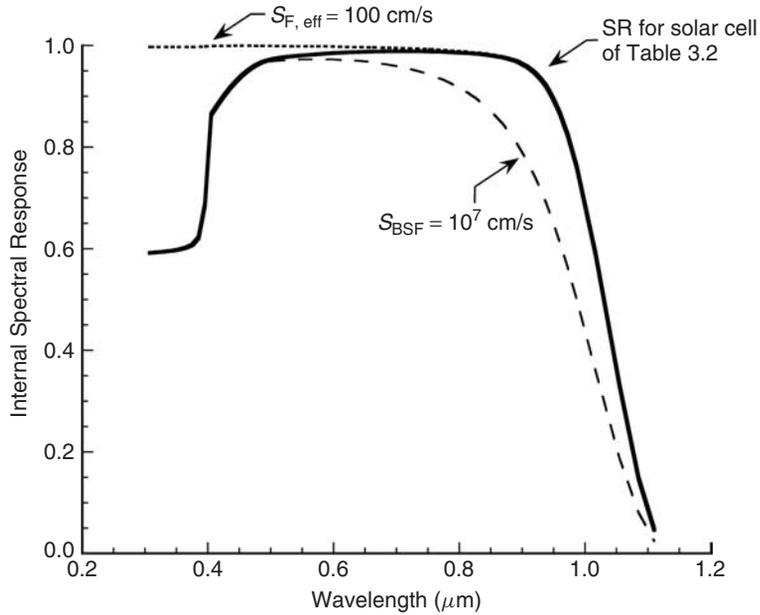
Figure 3.20   Internal spectral response of the silicon solar cell defined in Table 3.2

determined from it, along with the knowledge of the grid shadowing, reflectance, and optical thickness. W_{opt} can be greater than the cell thickness if light-trapping methods are used. Such methods include textured surfaces [26] and back-surface reflectors [27] and are discussed in Chapters 11 and 12. The short-circuit current can be written in terms of the external spectral response as

$$I_{SC} = \int_\lambda SR_{ext}(\lambda) f(\lambda) \, d\lambda. \tag{3.152}$$

The internal spectral response gives an indication of which sources of recombination are affecting the cell performance. This is demonstrated in Figure 3.20 where the internal spectral response of the silicon solar cell described by the parameters of Table 3.2 is shown. Also shown is the spectral response when $S_{F,eff} = 100$ cm/s (a well-passivated front surface) and the spectral response when $S_{BSF} = 1 \times 10^7$ cm/s (in effect, no BSF). The short-wavelength response improves dramatically when the front surface is passivated since the absorption coefficient is highest for short-wavelength (high-energy) photons. Conversely, removing the BSF makes it more likely that electrons created deep within the base region of the solar cell (those created by the long-wavelength, low-energy photons) will recombine at the back contact and therefore, the long-wavelength response is dramatically reduced.

3.4.2 Parasitic Resistance Effects

Equation (3.143) neglects the parasitic series and shunt resistances typically associated with real solar cells. Incorporating these resistances into the circuit model of Figure 3.15, as shown in

Figure 3.21, yields

$$I = I'_{SC} - I_{o1}(e^{q(V+IR_S)/kT} - 1) - I_{o2}(e^{q(V+IR_S)/2kT} - 1) - \frac{(V + IR_S)}{R_{Sh}} \tag{3.153}$$

where I'_{SC} is the short-circuit current when there are no parasitic resistances. The effect of these parasitic resistances on the $I-V$ characteristic is shown in Figures 3.22 and 3.23. As can also be seen in Equation (3.153), the shunt resistance R_{Sh} has no effect on the short-circuit current, but reduces the open-circuit voltage. Conversely, the series resistance R_S has no effect on the open-circuit voltage, but reduces the short-circuit current. Sources of series resistance include the metal contacts, particularly the front grid, and the transverse flow of current in the solar cell emitter to the front grid.
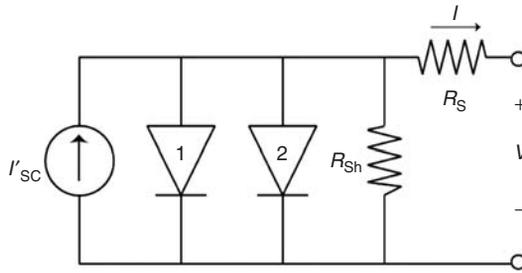

Figure 3.21   Solar cell circuit model including the parasitic series and shunt resistances
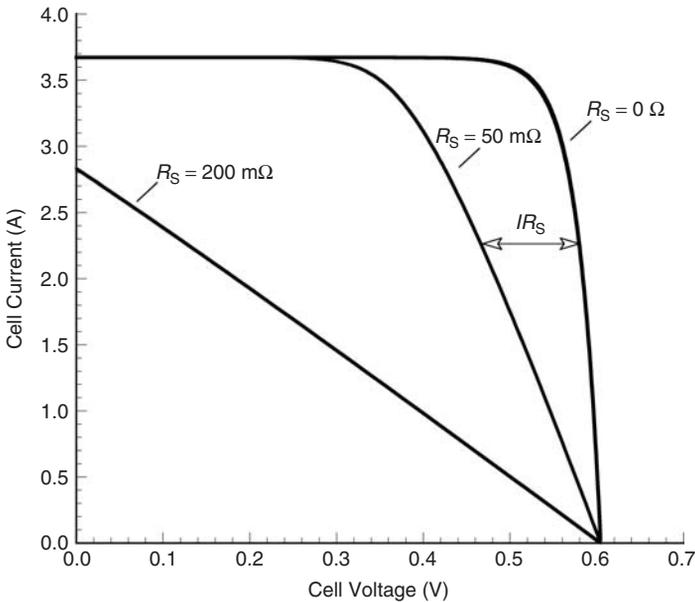

Figure 3.22   Effect of series resistance on the current–voltage characteristic of a solar cell ($R_{Sh} \to \infty$)
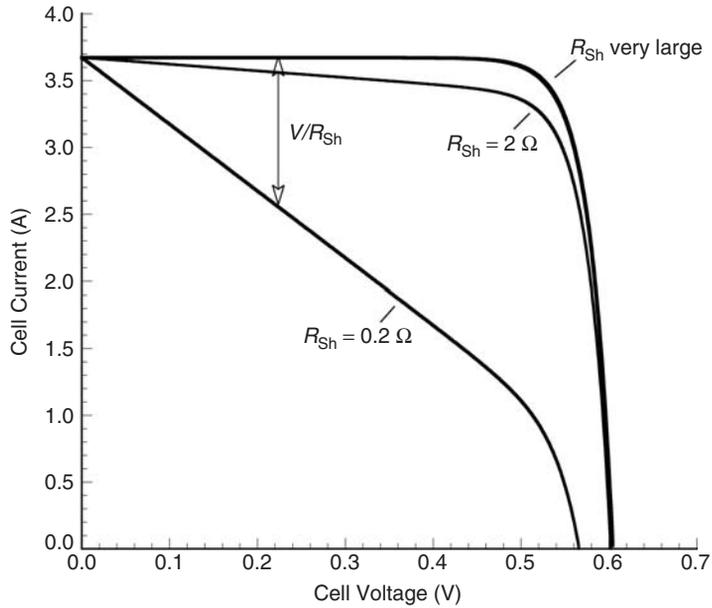
Figure 3.23   Effect of shunt resistance on the current−voltage characteristic of a solar cell ($R_S = 0$)

It is often more convenient to rewrite Equation (1.153) as

$$I = I'_{SC} - I_o(e^{q(V+IR_S)/A_okT} - 1) - \frac{(V + IR_S)}{R_{Sh}} \tag{3.154}$$

where A_o is the diode ideality (quality) factor and typically has a value between 1 and 2, with $A_o \approx 1$ for diode dominated by recombination in the quasi-neutral regions and $A_o \to 2$ when recombination in the depletion region dominates. In solar cells where the recombination in each region is comparable, A_o is somewhere in between. At short-circuit, Equation (3.154) becomes

$$I_{SC} = I'_{SC} - I_o(e^{qI_{SC}R_S/A_okT} - 1) - I_{SC}R_S/R_{Sh} \tag{3.155}$$

and at open-circuit, it becomes

$$0 = I'_{SC} - I_o(e^{V_{OC}/A_okT} - 1) - V_{OC}/R_{Sh}. \tag{3.156}$$

When $\log(I_{SC})$ is plotted versus V_{OC} (where I_{SC} and V_{OC} are obtained over a range of illumination intensities), there is typically a regime where neither the series nor shunt resistances are important, as illustrated in Figure 3.24. The slope of this line will yield the diode ideality factor A_o, while the y-intercept will give I_o. In the regime where only series resistance is important, Equations (3.155) and (3.156) can be combined to give

$$I_{SC}R_S = \frac{A_okT}{q} \ln\left[\frac{I_o e^{qV_{OC}/A_okT} - I_{SC}}{I_o}\right] \tag{3.157}$$

and a plot of I_{SC} versus $\log[I_o e^{qV_{OC}/A_okT} - I_{SC}]$ will then permit R_S to be extracted from the slope of this line. Similarly, in the regime where only R_{Sh} is important, Equations (3.155) and (3.156)
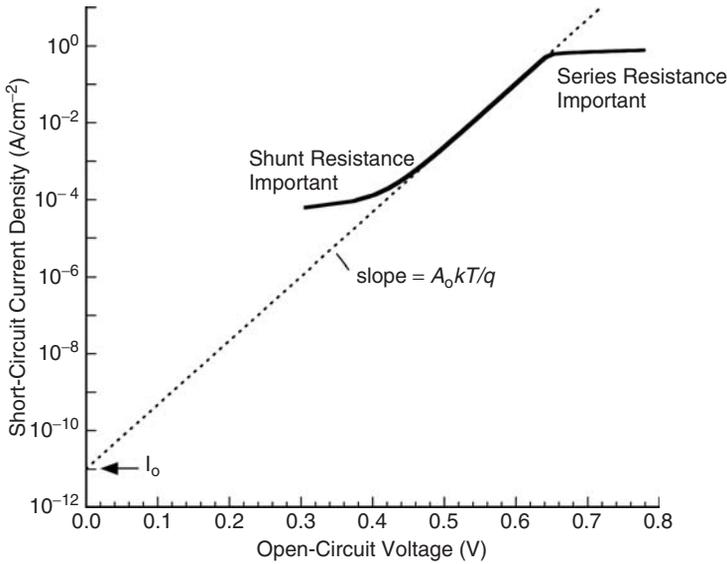
Figure 3.24   Short-circuit current versus open-circuit voltage plot illustrating the effects of series and shunt resistances

can be combined to give

$$\frac{V_{OC}}{R_{Sh}} = I_{SC} - I_o e^{qV_{OC}/A_o kT} \qquad (3.158)$$

and R_{Sh} can be determined from the slope of the line given by plotting V_{OC} versus $[I_{SC} - I_o e^{qV_{OC}/A_o kT}]$. If the series and shunt resistances are such that there is no regime where they can be neglected, the parameters can, with patience, be extracted through the process of trial and error.

3.4.3 Temperature Effects

From Equations (3.128) and (3.129), it is apparent that

$$I_{o1,n}, I_{o1,p} \propto n_i^2 \qquad (3.159)$$

and from Equation (3.130) that

$$I_{o2} \propto n_i. \qquad (3.160)$$

An increase in the intrinsic carrier concentration increases the dark saturation (recombination) current and results in a decrease in the open-circuit voltage, as can be seen from Equation (3.145). The dark saturation current contains other temperature-dependent terms (D, τ, and S), but the temperature dependence of the intrinsic carrier concentration dominates. The intrinsic carrier concentration is given by Equation (3.17), which when combined with Equations (3.12) and (3.13) yields

$$n_i = 2(m_n^* m_p^*)^{3/4} \left(\frac{2\pi kT}{h^2} \right)^{3/2} e^{-E_G/2kT}. \qquad (3.161)$$

The effective masses are generally taken to be weak functions of temperature. The bandgap decreases with temperature, and its temperature dependence is well modeled by

$$E_G(T) = E_G(0) - \frac{aT^2}{T + \beta}.$$ (3.162)

where α and β are constants specific to each semiconductor. It is clear that as the temperature increases, n_i increases, and thus recombination increases, and cell performance is impaired. Bandgap narrowing, referred to earlier, is a reduction in bandgap due to high doping and also serves to increase n_i and impair solar cell performance.

The open-circuit current expression, Equation (3.145), can be rearranged and the temperature dependence explicitly included to give

$$I_{SC} \approx I_{o1} e^{qV_{OC}/kT} \approx BT^\zeta e^{-E_G(0)/kT} e^{qV_{OC}/kT}$$ (3.163)

where B is a temperature-independent constant and $T^\zeta e^{-E_G(0)/kT}$ accounts for the temperature dependence of the saturation current. The short-circuit current is relatively unaffected by temperature under typical operating conditions, so by differentiating with respect to T, the temperature dependence of the open-circuit voltage can be expressed as [22]

$$\frac{dV_{OC}}{dT} = -\frac{\frac{1}{q}E_G(0) - V_{OC} + \zeta\frac{kT}{q}}{T}$$ (3.164)

which for silicon at 300 K corresponds to about $-2.3\,\text{mV}/°\text{C}$. Equation (3.163) can be rearranged as follows:

$$V_{OC}(T) = \frac{1}{q}E_G(0) - \frac{kT}{q}\ln\left(\frac{BT^\zeta}{I_{SC}}\right).$$ (3.165)

V_{OC} varies roughly linearly with temperature and an extrapolation of V_{OC} to $T = 0$ is the bandgap voltage since $\lim_{T \to 0}[T \ln T] = 0$.

The reason this is so important is that modules typically operate at $20–40°\text{C}$ above ambient, depending on the module design and sunlight intensity. Typical Si modules have a negative temperature coefficient of power output of -0.4 to -0.5% relative per $°\text{C}$, largely due to the temperature dependence of V_{OC} as indicated in Equation 3.165. The effect of temperature on module performance is discussed further in Chapters 18 and 19.

3.4.4 Concentrator Solar Cells

Operating solar cells under concentrated illumination offers two main advantages. The first is that since fewer solar cells are required to collect the sunlight falling on a given area, their cost of manufacture can be higher than that for cells designed for unconcentrated illumination, and they are therefore presumably of higher quality (efficiency). The second is that operation under concentrated illumination offers an advantage in the solar cell efficiency. If sunlight is concentrated by a factor of X (X suns illumination), the short-circuit at that concentration is

$$I_{SC}^{X\text{suns}} = XI_{SC}^{1\text{sun}}.$$ (3.166)

This is assuming that the semiconductor parameters are unaffected by the illumination level and that the cell temperature is the same at both levels of illumination – not necessarily valid assumptions,

especially at very large X. However, these assumptions will allow the demonstration of the potential efficiency of concentrator solar cells. Substituting Equation (3.166) into Equation (3.135) gives

$$\eta = \frac{FF^{\text{Xsuns}} V_{\text{OC}}^{\text{Xsuns}} I_{\text{SC}}^{\text{Xsuns}}}{P_{\text{in}}^{\text{Xsuns}}} = \frac{FF^{\text{Xsuns}} V_{\text{OC}}^{\text{Xsuns}} X I_{\text{SC}}^{\text{1sun}}}{X P_{\text{in}}^{\text{1sun}}} = \frac{FF^{\text{Xsuns}} V_{\text{OC}}^{\text{Xsuns}} I_{\text{SC}}^{\text{1sun}}}{P_{\text{in}}^{\text{1sun}}} \tag{3.167}$$

From Equation (3.132),

$$V_{\text{OC}}^{\text{Xsuns}} = V_{\text{OC}}^{\text{1sun}} + \frac{kT}{q} \ln X. \tag{3.168}$$

FF is a function of V_{OC} (Equation 3.143), so

$$\eta^{\text{Xsuns}} = \eta^{\text{1sun}} \left(\frac{FF^{\text{Xsuns}}}{FF^{\text{1sun}}} \right) \left(1 + \frac{\frac{kT}{q} \ln X}{V_{\text{OC}}^{\text{1sun}}} \right). \tag{3.169}$$

Both factors multiplying the 1 sun efficiency increase as the illumination concentration increases. Therefore, the efficiency of concentrator cells increases as the illumination concentration increases. For a silicon solar cell with $V_{\text{OC}}^{\text{1sun}} = 0.72\text{V}$, the efficiency at 1000 suns can potentially be more than 25% higher than its 1 sun value.

Of course, there are many obstacles to achieving this. Concentrator cells must be cooled, since an increase in operating temperature reduces V_{OC}, and hence the cell efficiency. In real devices, the FF^{Xsuns} eventually decreases with increasing solar cell current due to parasitic series resistance. Concentrator solar cells are discussed in more detail in Chapter 10.

3.4.5 High-level Injection

In high-level injection, the excess carrier concentrations greatly exceed the doping in the base region, so $\Delta p \approx \Delta n \approx n \approx p$ if the carriers are moving generally in the same direction. This occurs with back-contact solar cells, such as the silicon point-contact solar cell [28], which is illustrated in Figure 3.25. Since both electrical contacts are on the back, there is no grid shadowing. These cells are typically used in concentrator application and high-level injection conditions pervade. Assuming high-level injection, a simple analysis is possible.

Returning to Equations (3.77) and (3.78), it can be seen that in high-level injection, the electric field can be eliminated (it is not necessarily zero), resulting in the ambipolar diffusion equation

$$D_{\text{a}} \frac{d^2 p}{dx^2} - \frac{p}{\tau_n + \tau_p} = -G(x), \tag{3.170}$$

where the ambipolar diffusion coefficient is given by

$$D_{\text{a}} = \frac{2 D_n D_p}{D_n + D_p}. \tag{3.171}$$

In silicon, where $D_n / D_p \approx 3$ over a wide range of doping, the ambipolar diffusion coefficient is $D_{\text{a}} \approx 3/2 D_p \approx 1/2 D_n$ and, if we also assume $\tau_p \approx \tau_n$, the ambipolar diffusion length is

$$L_{\text{a}} \approx \sqrt{3} L_p \approx L_n. \tag{3.172}$$

Thus, the increased high-injection lifetime (see Equation 3.40) offsets the reduced ambipolar diffusion coefficient.
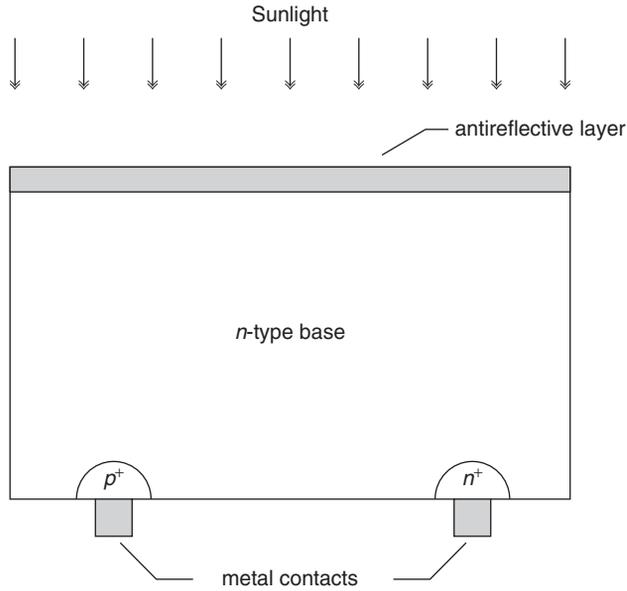
Figure 3.25   Schematic of a back-contact solar cell

It is crucial that the front surface of a back-contacted cell be well passivated, so we will assume that $S_F = 0$. We will further assume that optical generation is uniform throughout the base region. At open-circuit, with these assumptions, $\mathrm{d}^2 p/\mathrm{d}x^2 = 0$ and therefore

$$V_{\mathrm{OC}} = \frac{2kT}{q} \ln \left[\frac{G(\tau_n + \tau_p)}{n_{\mathrm{i}}} \right].$$

(3.173)

The short-circuit current (with $p \simeq 0$ at the back of the cell) is

$$I_{\mathrm{SC}} = q A L_{\mathrm{a}} G \sinh(W_{\mathrm{B}}/L_{\mathrm{a}})$$

(3.174)

which, when $L_{\mathrm{a}} \gg W_{\mathrm{B}}$, becomes

$$I_{\mathrm{SC}} = q A W_{\mathrm{B}} G.$$

(3.175)

3.4.6 *p-i-n* Solar Cells and Voltage-dependent Collection

The *p-i-n* solar cell takes advantage of the fact that in many semiconductor materials, especially direct bandgap semiconductors (i.e. large absorption coefficient), most of the electron–hole pairs are created very near the surface. If an intrinsic (undoped) layer is placed between the (very thin) *n* and *p* regions, the depletion region thickness is the most significant fraction of the total solar cell thickness, as illustrated in Figure 3.26. Carrier collection is now aided by the electric field in the depletion region, which helps offset the low lifetimes in some materials, such as amorphous silicon (see Chapter 12). The $I-V$ characteristic of a *p-i-n* solar cell can be described with minor modifications to the previously derived expressions. The most significant modification is to Equation (3.130) where the depletion width is now written as

$$W_{\mathrm{D}} = \chi_N + W_{\mathrm{I}} + \chi_P$$

(3.176)

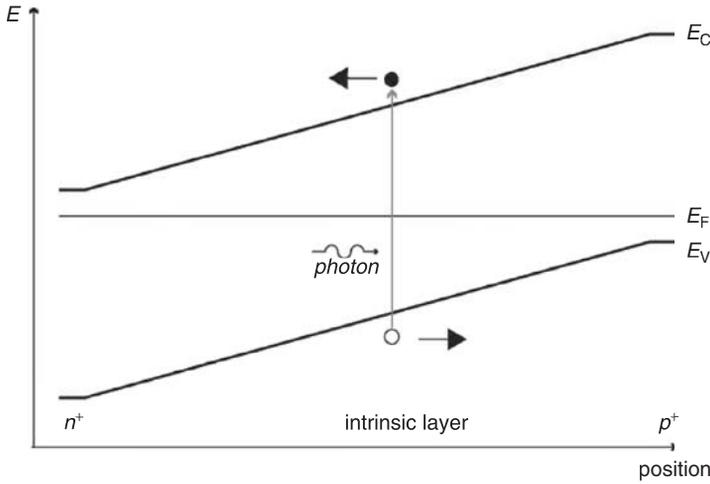**Figure 3.26**   Band diagram of a *p-i-n* solar cell illustrating field-enhanced collection

where W_I is the thickness of the intrinsic layer. Since χ_N and χ_P are very thin, short-base approximations are in order for Equations (3.128) and (3.129). Also, there is no BSF ($S_{BSF} \rightarrow \infty$).

As mentioned above, the electric field in the depletion region of *p-i-n* solar cells aids in the collection of carriers. This is referred to as voltage-dependent collection (VDC) and is an important effect in any solar cell in which there is significant photogeneration in the depletion region, as is the case in most thin film solar cells. At the maximum power point and at open-circuit, the electric field in the depletion region is lower than at short-circuit, which leads to a lower *FF* and V_{OC} than might otherwise be expected. An excellent analysis of VDC in CdTe/CdS solar cells is given in [23].

3.4.7 Heterojunction Solar Cells

Reducing the recombination losses in the emitter will improve the efficiency of the solar cell. This can be accomplished by reducing the junction area [19]. Another way is by using a wider bandgap material for the emitter of the solar cell, as shown in Figure 3.27 for an n^+p solar cell. Ideally, minority-carriers in each region are collected by the junction to become majority-carriers in the opposite region. Majority-carriers injected into the opposite region that become minority-carriers are a source of recombination and reduce the efficiency of the solar cell.

The larger barrier presented to holes by the wider bandgap emitter of a heterojunction solar cell substantially reduces the number of holes injected into the emitter, thus reducing recombination in the emitter and improving the efficiency of the solar cell. The arrows in Figure 3.27 illustrate this point. This can also be seen analytically from Equation (3.128), the emitter component of the dark saturation current, which is reproduced below with $n_i = n_{i,emitter}$

$$I_{o1,p} = qA\frac{n_{i,emitter}^2}{N_D}\frac{D_p}{L_p}\left\{\frac{D_p/L_p \sinh[(W_N - x_N)/L_p] + S_{F,eff}\cosh[(W_N - x_N/L_p]}{D_p/L_p \cosh[(W_N - x_N)/L_p] + S_{F,eff}\sinh[(W_N - x_N/L_p]}\right\} \quad (3.177)$$

Recall from Equation (3.17) that

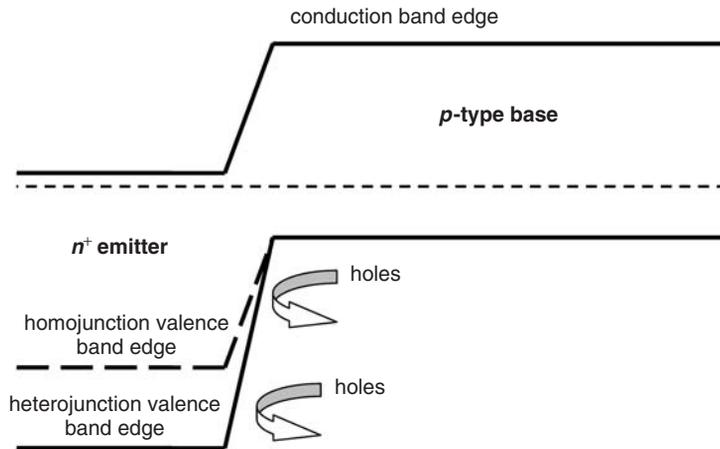$$n_{i,emitter} = \sqrt{N_C N_V}e^{-E_{G,emitter}/2kT}. \quad (3.178)$$

Figure 3.27   Band diagram of a heterojunction solar cell

The intrinsic carrier concentration $n_{i,emitter}$ is much smaller for a wider bandgap emitter, thus reducing the emitter component of the dark saturation current and therefore the net recombination in the emitter region.

An additional advantage of a heterojunction solar cell is that incident photons with energy less than the emitter bandgap energy (and higher than the base region bandgap) will be absorbed in the base region of the solar cell rather than near the front surface where recombination can be high (as would be the case in a homojunction). Heterojunction solar cells are discussed in more detail in Chapters 8, 13, and 14.

3.4.8 Detailed Numerical Modeling

While analytic solutions such as those discussed thus far in this chapter provide an intuitive under-standing of solar cell performance and are therefore very important, they are limited in their accuracy due to the many simplifying assumptions that must be made in order to obtain them. It is rather straightforward to solve the semiconductor equations numerically without the need to make so many simplifications. Several computer codes have been written that solve the semiconductor equations for the explicit purpose of modeling solar cells: PC-1D [29], AMPS [30], ADEPT [31], and its predecessors [32, 33], for example.

The basic design of these computer programs is very similar. The semiconductor equations (three coupled nonlinear partial differential equations) are cast in a normalized form [34] to simplify the calculations. Finite difference or finite element methods are then used to discretize the equations on a mesh (grid), resulting in a set of three coupled nonlinear difference equations. Using appro-priately discretized boundary conditions, these equations are solved iteratively using a generalized Newton method to obtain the carrier concentrations and electric potential at each mesh point. Each Newton iteration involves the solution of a very large matrix equation of order $3N$, where N is the number of mesh points. One-dimensional simulations typically utilize on the order of 1000 mesh points, so the matrix is 3000×3000. In 2D, the minimum mesh is typically at least 100×100, so $N = 10^4$ and the matrix is of order 3×10^4 and contains 9×10^8 elements. Fortunately, the matrices are sparse and can be solved using considerably less computer memory than one would expect at first glance.

Numerical simulation allows analysis of solar cell designs and operating conditions for which simple analytic expressions are inadequate. The necessity of ignoring the spatial variation of parameters is eliminated and more accurate representations of the solar cell are possible. In particular, nonuniform doping, heterojunction solar cells (the bandgap varies spatially), amorphous silicon solar cells (complex trapping/recombination mechanisms, field-assisted collection), and concentrator solar cells (high-level injection, 2D/3D effects) can all be modeled with more precision.

3.5 SUMMARY

It has been the objective of this chapter to give the reader a basic understanding of the physical principles that underlie the operation of solar cells. Toward that end, the fundamental physical characteristics of solar cell materials that permit the conversion of light into electricity have been reviewed. These characteristics include the ability of semiconductors to absorb photons by conferring that energy to carriers of electrical current and the ability of semiconductor materials to conduct electricity.

The basic operating principles of the solar cell (a carefully designed pn-junction diode) were derived from the (simplified) equations describing the dynamics of holes and electrons in semiconductors. This led to the definition of the solar cell figures of merit – the open-circuit voltage (V_{OC}), the short-circuit current (I_{SC}), the fill factor (FF), and the cell efficiency (η). The two key factors determining solar cell efficiency – electron–hole pair generation and recombination – were identified and discussed. In particular, the need to minimize all sources of recombination in the solar cells was demonstrated through examples.

The importance of matching the bandgap of the solar cell material to the solar spectrum was also discussed and it was shown that silicon, with a bandgap of $1.12\,\mathrm{eV}$, is a reasonably good match to the solar spectrum. The effects of parasitic resistances and temperature on solar cell performance were examined and, finally, some advanced cell concepts were briefly introduced. Many of these topics will be expanded upon in the following chapters of this handbook.

REFERENCES

1. Green M, *Solar Cells: Operating Principles, Technology, and System Applications*, Chap. 1, Prentice Hall, Englewood Cliffs, NJ, 1–12 (1982).
2. Pierret R, in Pierret R, Neudeck G (Eds), *Modular Series on Solid State Devices, Volume VI: Advanced Semiconductor Fundamentals*, Addison-Wesley, Reading, MA (1987).
3. Sze S, *Physics of Semiconductor Devices*, 2nd Edition, John Wiley & Sons, Inc., New York, NY (1981).
4. Böer K, *Survey of Semiconductor Physics: Electrons and Other Particles in Bulk Semiconductors*, Van Nostrand Reinhold, New York, NY (1990).
5. Shur M, *Physics of Semiconductor Devices*, Prentice Hall, Englewood Cliffs, NJ (1990).
6. Singh J, *Physics of Semiconductors and Their Heterostructures*, McGraw-Hill, New York, NY (1993).
7. Pierret R, *Semiconductor Device Fundamentals*, Chap. 2, Addison-Wesley, Reading, MA, 23–74 (1996).
8. Slotboom J, De Graff H, *Solid-State Electron.* **19**, 857–862 (1976).
9. Pankove J, *Optical Processes in Semiconductors*, Chap. 3, Dover Publications, New York, NY, 34–81 (1971).
10. Sanii F, Giles F, Schwartz R, Gray J, *Solid-State Electron.* **35**, 311–317 (1992).
11. Barnett A, Honsberg C, Kirkpatrick D, *et al.*, *Proc. 4th World Conference on Photovoltaic Energy Conversion*, 2560–2564 (2006).

12. Pierret R, in Pierret R, Neudeck G (Eds), *Modular Series on Solid State Devices, Volume VI: Advanced Semiconductor Fundamentals*, Chap. 5, Addison-Wesley, Reading, MA, 139–179 (1987).
13. Dubin S, Gray J, *IEEE Trans. Electron Devices* **41**, 239–245 (1994).
14. Pierret R, in Pierret R, Neudeck G (Eds), *Modular Series on Solid State Devices, Volume VI: Advanced Semiconductor Fundamentals*, Chap. 6, Addison-Wesley, Reading, MA (1987).
15. Van Roosbroeck W, *Bell Syst. Tech. J.* **29**, 560–607 (1950).
16. Sacconi F, Di Carlo A, Lugli P, Morkoc H, *IEEE Trans. Electron Devices* **48**, 450–457 (2001).
17. Lundstrom M, Schulke R, *IEEE Trans. Electron Devices* **30**, 1151–1159 (1983).
18. Gray J, Haas A, Wilcox J, Schwartz R, *Proc. 33rd IEEE Photovoltaic Specialist Conf.*, 1–6 (2008).
19. Gray J, Schwartz R, *Proc. 18th IEEE Photovoltaic Specialist Conf.*, 568–572 (1985).
20. Mathers C, *J. Appl. Phys.* **48**, 3181, 3182 (1977).
21. Shockley W, Queisser H, *J. Appl. Phys.* **32**, 510 (1961).
22. Green M, *Solar Cells: Operating Principles, Technology, and System Applications*, Chap. 5, Prentice Hall, Englewood Cliffs, NJ, 85–102 (1982).
23. Hegedus S, Desai D, Thompson C, *Prog Photovolt: Res. Applic.* **15**, 587–602 (2007).
24. Green M, Emery K, King D, Hisikawa Y, Warta W, *Prog Photovolt: Res. Applic.* **14**, 45–51 (2006).
25. Gray J, *Two-Dimensional Modeling of Silicon Solar Cells*, Ph.D. thesis, Purdue University, West Lafayette, IN (1982).
26. Baraona C, Brandhorst Jr. H, *Proc. 11th IEEE Photovoltaic Specialist Conf.*, 44–48 (1975).
27. Chai A, *Proc. 14th IEEE Photovoltaic Specialist Conf.*, 156–160 (1980).
28. Sinton R, Kwark Y, Swanson R, *Proc. 18th IEEE Photovoltaic Specialist Conf.*, 61–64 (1984).
29. Rover D, Basore P, Thorson G, *Proc. 18th IEEE Photovoltaic Specialist Conf.*, 703–709 (1984).
30. Rubinelli F *et al.*, *Proc. 22nd IEEE Photovoltaic Specialist Conf.*, 1405–1408 (1991).
31. Gray J, *Proc. 22nd IEEE Photovoltaic Specialist Conf.*, 436–438 (1991).
32. Lundstrom M, *Numerical Analysis of Silicon Solar Cells*, Ph.D. thesis, Purdue University, West Lafayette, IN (1980).
33. Gray J, *IEEE Trans. Electron Devices* **36**, 906–912 (1989).
34. Snowden C, *Introduction to Semiconductor Device Modeling*, Chap. 2, World Scientific, Singapore, 14–36 (1986).