
PHOTONS IN SEMICONDUCTORS

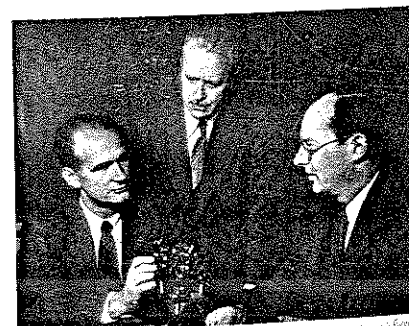
15.1 SEMICONDUCTORS

- A. Energy Bands and Charge Carriers
- B. Semiconducting Materials
- C. Electron and Hole Concentrations
- D. Generation, Recombination, and Injection
- E. Junctions
- F. Heterojunctions
- *G. Quantum Wells and Superlattices

15.2 INTERACTIONS OF PHOTONS WITH ELECTRONS AND HOLES

- A. Band-to-Band Absorption and Emission
- B. Rates of Absorption and Emission
- C. Refractive Index

William P. Shockley (1910–1989), left, Walter H. Brattain (1902–1987), center, and John Bardeen (1908–1991), right, shared the Nobel Prize in 1956 for showing that semiconductor devices could be used to achieve amplification.



Electronics is the technology of controlling the flow of electrons whereas photonics is the technology of controlling the flow of photons. Electronics and photonics have been joined together in semiconductor optoelectronic devices where photons generate mobile electrons, and electrons generate and control the flow of photons. The compatibility of semiconductor optoelectronic devices and electronic devices has, in recent years, led to substantive advances in both technologies. Semiconductors are used as optical detectors, sources (light-emitting diodes and lasers), amplifiers, waveguides, modulators, sensors, and nonlinear optical elements.

Semiconductors absorb and emit photons by undergoing transitions between different allowed energy levels, in accordance with the general theory of photon-atom interactions described in Chap. 12. However, as we indicated briefly there, semiconductors have properties that are unique in certain respects:

- A semiconductor material cannot be viewed as a collection of noninteracting atoms, each with its own individual energy levels. The proximity of the atoms in a solid results in one set of energy levels representing the entire system.
- The energy levels of semiconductors take the form of groups of closely spaced levels that form bands. In the absence of thermal excitations (at $T = 0$ K), these are either completely occupied by electrons or completely empty. The highest filled band is called the valence band, and the empty band above it is called the conduction band. The two bands are separated by an energy gap.
- Thermal and optical interactions can impart energy to an electron, causing it to jump across the gap from the valence band into the conduction band (leaving behind an empty state called a hole). The inverse process can also occur. An electron can decay from the conduction band into the valence band to fill an empty state (provided that one is accessible) by means of a process called electron-hole recombination. We therefore have two types of particles that carry electric current and can interact with photons: electrons and holes.

Two processes are fundamental to the operation of almost all semiconductor optoelectronic devices:

- *The absorption of a photon can create an electron-hole pair.* The mobile charge carriers resulting from absorption can alter the electrical properties of the material. One such effect, photoconductivity, is responsible for the operation of certain semiconductor photodetectors.
- *The recombination of an electron and a hole can result in the emission of a photon.* This process is responsible for the operation of semiconductor light sources. Spontaneous radiative electron-hole recombination is the underlying process of light generation in the light-emitting diode. Stimulated electron-hole recombination is the source of photons in the semiconductor laser.

In Sec. 15.1 we begin with a review of the properties of semiconductors that are important in semiconductor photonics; the reader is expected to be familiar with the basic principles of semiconductor physics. Section 15.2 provides an introduction to the optical properties of semiconductors. A simplified theory of absorption, spontaneous emission, and stimulated emission is developed using the theory of radiative atomic transitions developed in Chap. 12.

This, and the following two chapters, are to be regarded as a single unit. Chapter 16 deals with semiconductor optical sources such as the light-emitting diode and the injection laser diode. Chapter 17 is devoted to semiconductor photon detectors.

15.1 SEMICONDUCTORS

A semiconductor is a crystalline or amorphous solid whose electrical conductivity is typically intermediate between that of a metal and an insulator and can be changed significantly by altering the temperature or the impurity content of the material, or by illumination with light. The unique energy-level structure of semiconductor materials leads to special electrical and optical properties, as described later in this chapter. Electronic devices principally make use of silicon (Si) as a semiconductor material, but compounds such as gallium arsenide (GaAs) are of utmost importance to photonics (see Sec. 15.1B for a selected tabulation of other semiconductor materials).

A. Energy Bands and Charge Carriers

Energy Bands in Semiconductors

Atoms of solid-state materials have a sufficiently strong interaction that they cannot be treated as individual entities. Valence electrons are not attached (bound) to individual atoms; rather, they belong to the system of atoms as a whole. The solution of the Schrödinger equation for the electron energy, in the periodic potential created by the collection of atoms in a crystal lattice, results in a splitting of the atomic energy levels and the formation of energy bands (see Sec. 12.1). Each band contains a large number of finely separated discrete energy levels that can be approximated as a continuum. The valence and conduction bands are separated by a "forbidden" energy gap of width E_g (see Fig. 15.1-1), called the **bandgap energy**, which plays an important role in determining the electrical and optical properties of the material. Materials with a filled valence band and a large energy gap (> 3 eV) are electrical insulators; those for which the gap is small or nonexistent are conductors (see Fig. 12.1-5). Semiconductors have energy gaps that lie roughly in the range 0.1 to 3 eV.

Electrons and Holes

In accordance with the **Pauli exclusion principle**, no two electrons can occupy the same quantum state. Lower energy levels are filled first. In elemental semiconductors, such as Si and Ge, there are four valence electrons per atom; the valence band has a number of quantum states such that in the absence of thermal excitations the valence band is completely filled and the conduction band is completely empty. Consequently, the material cannot conduct electricity.

As the temperature increases, however, some electrons will be thermally excited into the empty conduction band where there is an abundance of unoccupied states (see Fig. 15.1-2). There, the electrons can act as mobile carriers; they can drift in the crystal lattice under the effect of an applied electric field and thereby contribute to the electric current. Furthermore, the departure of an electron from the valence band provides an empty quantum state, allowing the remaining electrons in the valence band to exchange

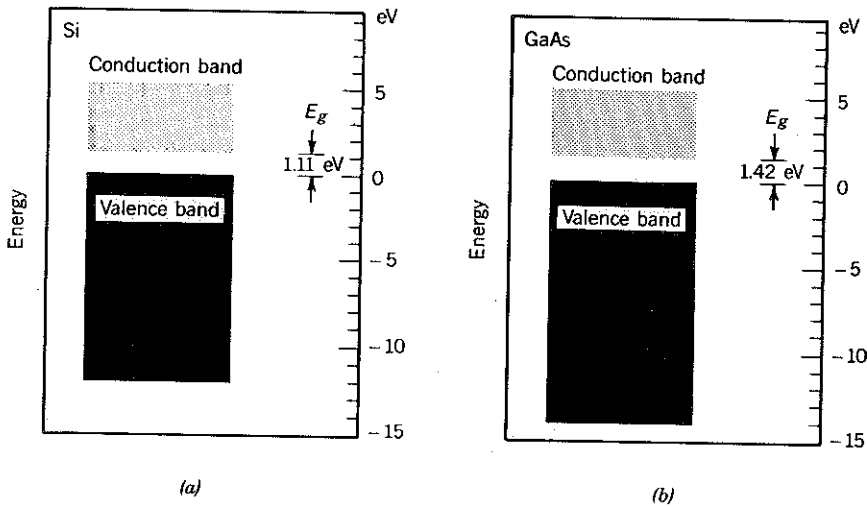


Figure 15.1-1 Energy bands: (a) in Si, and (b) in GaAs.

places with each other under the influence of an electric field. A motion of the "collection" of remaining electrons in the valence band occurs. This can equivalently be regarded as the motion, in the opposite direction, of the hole left behind by the departed electron. The hole therefore behaves as if it has a positive charge $+e$. The result of each electron excitation is, then, the creation of a free electron in the conduction band and a free hole in the valence band. The two charge carriers are free to drift under the effect of the applied electric field and thereby to generate an electric current. The material behaves as a *semiconductor* whose conductivity increases sharply with temperature as an increasing number of mobile carriers are thermally generated.

Energy-Momentum Relations

The energy E and momentum \mathbf{p} of an electron in free space are related by $E = p^2/2m_0 = \hbar^2 k^2/2m_0$, where p is the magnitude of the momentum and k is the magnitude of the wavevector $\mathbf{k} = \mathbf{p}/\hbar$ associated with the electron's wavefunction, and m_0 is the electron mass (9.1×10^{-31} kg). The E - k relation is a simple parabola.

The motion of electrons in the conduction band, and holes in the valence band, of a semiconductor are subject to different dynamics. They are governed by the Schrödinger

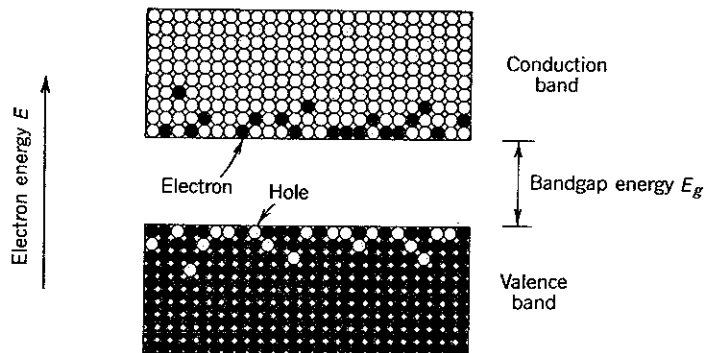


Figure 15.1-2 Electrons in the conduction band and holes in the valence band at $T > 0$ K.

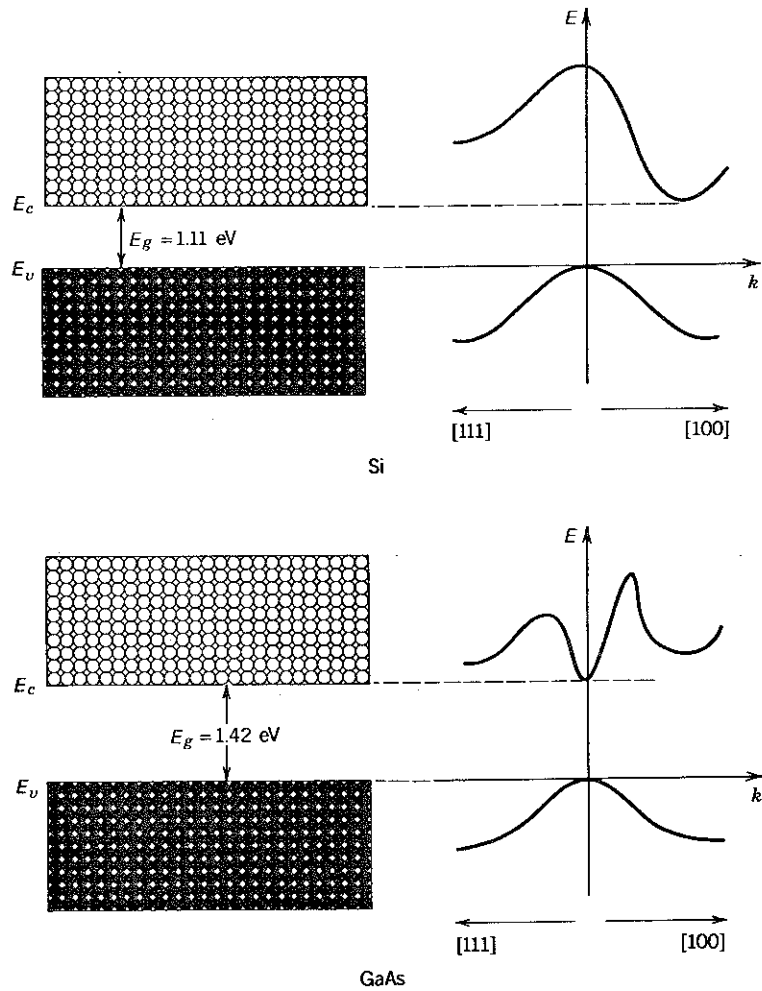


Figure 15.1-3 Cross section of the $E-k$ function for Si and GaAs along the crystal directions $[111]$ and $[100]$.

equation and the periodic lattice of the material. The $E-k$ relations are illustrated in Fig. 15.1-3 for Si and GaAs. The energy E is a periodic function of the components (k_1, k_2, k_3) of the vector k , with periodicities $(\pi/a_1, \pi/a_2, \pi/a_3)$, where a_1, a_2, a_3 are the crystal lattice constants. Figure 15.1-3 shows cross sections of this relation along two different directions of k . The energy of an electron in the conduction band depends not only on the magnitude of its momentum, but also on the direction in which it is traveling in the crystal.

Effective Mass

Near the bottom of the conduction band, the $E-k$ relation may be approximated by the parabola

$$E = E_c + \frac{\hbar^2 k^2}{2m_c}, \tag{15.1-1}$$

where E_c is the energy at the bottom of the conduction band and m_c is a constant representing the **effective mass** of the electron in the conduction band (see Fig. 15.1-4).

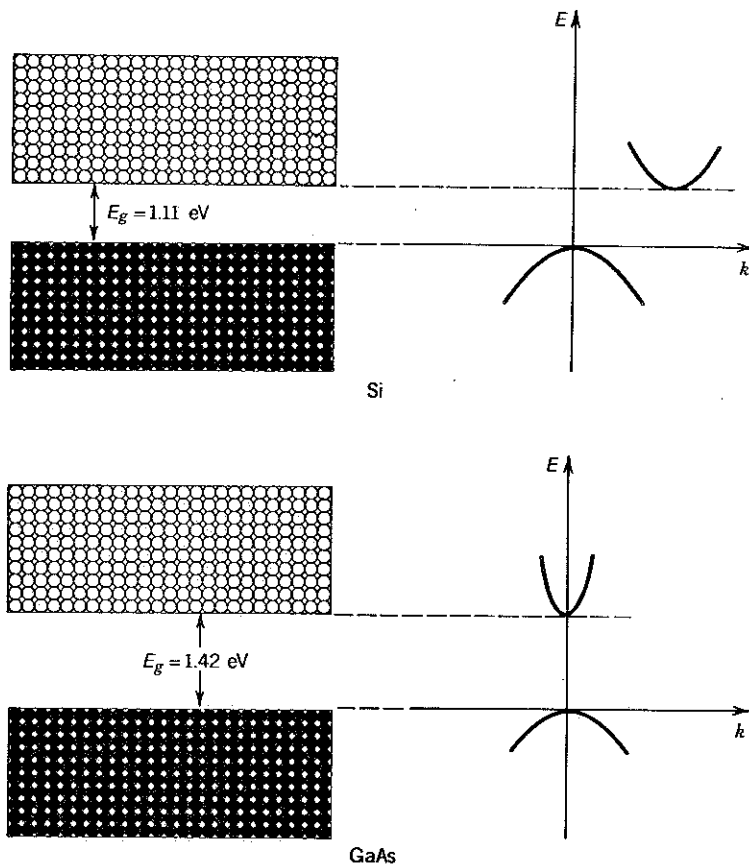


Figure 15.1-4 Approximating the E - k diagram at the bottom of the conduction band and at the top of the valence band of Si and GaAs by parabolas.

Similarly, near the top of the valence band,

$$E = E_v - \frac{\hbar^2 k^2}{2m_v}, \quad (15.1-2)$$

where $E_v = E_c - E_g$ is the energy at the top of the valence band and m_v is the effective mass of a hole in the valence band. In general, the effective mass depends on the crystal orientation and the particular band under consideration. Typical ratios of the averaged effective masses to the mass of the free electron m_0 are provided in Table 15.1-1 for Si and GaAs.

Direct- and Indirect-Gap Semiconductors

Semiconductors for which the valence-band maximum and the conduction-band minimum correspond to the same momentum (same k) are called **direct-gap** materials.

TABLE 15.1-1 Average Effective Masses of Electrons and Holes in Si and GaAs

	m_c/m_0	m_v/m_0
Si	0.33	0.5
GaAs	0.07	0.5

TABLE 15.1-2 A Section of the Periodic Table

II		III		IV		V		VI	
Zinc (Zn)		Aluminum (Al)		Silicon (Si)		Phosphorus (P)		Sulfur (S)	
Cadmium (Cd)		Gallium (Ga)		Germanium (Ge)		Arsenic (As)		Selenium (Se)	
Mercury (Hg)		Indium (In)				Antimony (Sb)		Tellurium (Te)	

Semiconductors for which this is not the case are known as **indirect-gap** materials. The distinction is important; a transition between the top of the valence band and the bottom of the conduction band in an indirect-gap semiconductor requires a substantial change in the electron's momentum. As is evident in Fig. 15.1-4, Si is an indirect-gap semiconductor, whereas GaAs is a direct-gap semiconductor. It will be shown subsequently that direct-gap semiconductors such as GaAs are efficient photon emitters, whereas indirect-gap semiconductors such as Si cannot be efficiently used as light emitters.

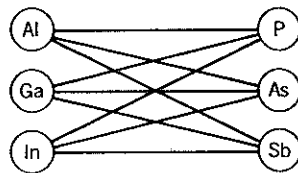
B. Semiconducting Materials

Table 15.1-2 reproduces a section of the periodic table of the elements, containing some of the important elements involved in semiconductor electronics and optoelectronics technology. Both elemental and compound semiconductors are of importance.

Elemental Semiconductors

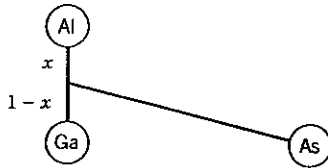
Several elements in group IV of the periodic table are semiconductors. Most important are **silicon (Si)** and **germanium (Ge)**. At present most commercial electronic integrated circuits and devices are fabricated from Si. However, these materials are not useful for fabricating photon emitters because of their indirect bandgap. Nevertheless, both are widely used for making photon detectors.

Binary Semiconductors



Compounds formed by combining an element in group III, such as aluminum (Al), gallium (Ga), or indium (In), with an element in group V, such as phosphorus (P), arsenic (As), or antimony (Sb), are important semiconductors. There are nine such III-V compounds. These are listed in Table 15.1-3, along with their bandgap energy E_g , bandgap wavelength $\lambda_g = hc_0/E_g$ (which is the free-space wavelength of a photon of energy E_g), and gap type (direct or indirect). The bandgap energies and the lattice constants of these compounds are also provided in Fig. 15.1-5. Various of these compounds are used for making photon detectors and sources (light-emitting diodes and lasers). The most important binary semiconductor for optoelectronic devices is gallium arsenide (GaAs). Furthermore,

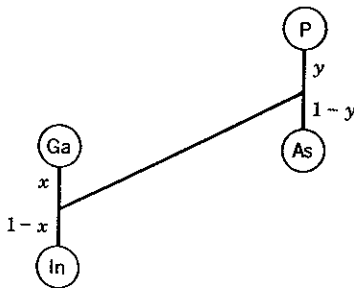
Ternary Semiconductors



GaAs is becoming increasingly important (relative to Si) as the basis of fast electronic devices and circuits.

Compounds formed from two elements of group III with one element of group V (or one from group III with two from Group V) are important ternary semiconductors. $(\text{Al}_x\text{Ga}_{1-x})\text{As}$, for example, is a ternary compound with properties intermediate between those of AlAs and GaAs, depending on the compositional mixing ratio x (where x denotes the fraction of Ga atoms in GaAs replaced by Al atoms). The bandgap energy E_g for this material varies between 1.42 eV for GaAs and 2.16 eV for AlAs, as x is varied between 0 and 1. The material is represented by the line connecting GaAs and AlAs in Fig. 15.1-5. Because this line is nearly horizontal, $\text{Al}_x\text{Ga}_{1-x}\text{As}$ is lattice matched to GaAs (i.e., they have the same lattice constant). This means that a layer of a given composition can be grown on a layer of different composition without introducing strain in the material. The combination $\text{Al}_x\text{Ga}_{1-x}\text{As}/\text{GaAs}$ is highly important in current LED and semiconductor laser technology. Other III-V compound semiconductors of various compositions and bandgap types (direct/indirect) are indicated in the lattice-constant versus band-gap-energy diagram in Fig. 15.1-5.

Quaternary Semiconductors



These compounds are formed from a mixture of two elements from Group III with two elements from group V. Quaternary semiconductors offer more flexibility for the synthesis of materials with desired properties than do ternary semiconductors, since they provide an extra degree of freedom. An example is provided by the quaternary $(\text{In}_{1-x}\text{Ga}_x)(\text{As}_{1-y}\text{P}_y)$, whose bandgap energy E_g varies between 0.36 eV (InAs) and 2.26 eV (GaP) as the compositional mixing ratios x and y vary between 0 and 1. The shaded area in Fig. 15.1-5 indicates the range of energy gaps and lattice constants spanned by this compound. For mixing ratios x and y that satisfy $y = 2.16(1-x)$, $(\text{In}_{1-x}\text{Ga}_x)(\text{As}_{1-y}\text{P}_y)$ can be very well lattice matched to InP and therefore conveniently grown on it. These compounds are used in making semiconductor lasers and detectors.

TABLE 15.1-3 Selected Elemental and III-V Binary Semiconductors and Their Bandgap Energies E_g at $T = 300$ K, Bandgap Wavelengths $\lambda_g = hc_0 / E_g$, and Type of Gap (I = Indirect, D = Direct)

Material	Bandgap Energy E_g (eV)	Bandgap Wavelength λ_g (μm)	Type
Ge	0.66	1.88	I
Si	1.11	1.15	I
AlP	2.45	0.52	I
AlAs	2.16	0.57	I
AlSb	1.58	0.75	I
GaP	2.26	0.55	I
GaAs	1.42	0.87	D
GaSb	0.73	1.70	D
InP	1.35	0.92	D
InAs	0.36	3.5	D
InSb	0.17	7.3	D

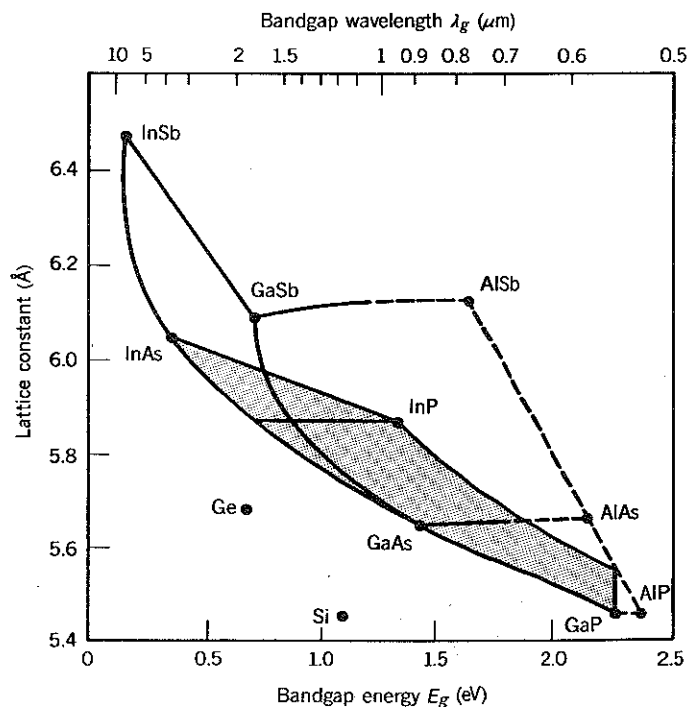


Figure 15.1-5 Lattice constants, bandgap energies, and bandgap wavelengths for Si, Ge, and nine III-V binary compounds. Ternary compounds can be formed from binary materials by motion along the line joining the two points that represent the binary materials. For example, $\text{Al}_x\text{Ga}_{1-x}\text{As}$ is represented by points on the line connecting GaAs and AlAs. As x varies from 0 to 1, the point moves along the line from GaAs to AlAs. Since this line is nearly horizontal, $\text{Al}_x\text{Ga}_{1-x}\text{As}$ is lattice matched to GaAs. Solid and dashed curves represent direct-gap and indirect-gap compositions, respectively. A material may have direct bandgap for one mixing ratio x and an indirect bandgap for a different x . A quaternary compound is represented by a point in the area formed by its four binary components. For example, $(\text{In}_{1-x}\text{Ga}_x)(\text{As}_{1-y}\text{P}_y)$ is represented by the shaded area with vertices at InAs, InP, GaP, and GaAs; the upper horizontal line represents compounds that are lattice matched to InP.

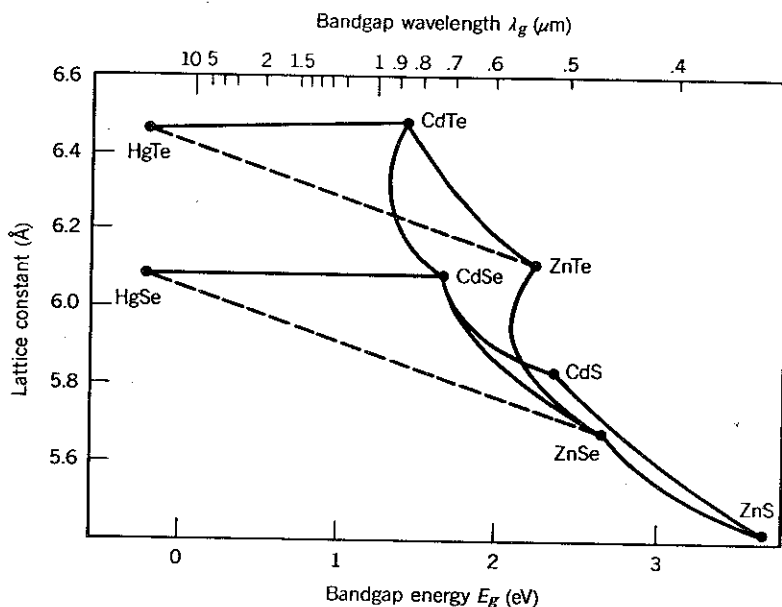


Figure 15.1-6 Lattice constants, bandgap energies, and bandgap wavelengths for some important II-VI binary compounds.

Compounds using elements from group II (e.g., Zn, Cd, Hg) and group VI (e.g., S, Se, Te) of the periodic table also form useful semiconductors, particularly at wavelengths shorter than $0.5 \mu\text{m}$ and longer than $5.0 \mu\text{m}$, as shown in Fig. 15.1-6. HgTe and CdTe, for example, are nearly lattice matched, so that the ternary semiconductor $\text{Hg}_x\text{Cd}_{1-x}\text{Te}$ is a useful material for fabricating photon detectors in the middle-infrared region of the spectrum. Also used in this range are IV-VI compounds such as $\text{Pb}_x\text{Sn}_{1-x}\text{Te}$ and $\text{Pb}_x\text{Sn}_{1-x}\text{Se}$. Applications include night vision, thermal imaging, and long-wavelength lightwave communications.

Doped Semiconductors

The electrical and optical properties of semiconductors can be substantially altered by adding small controlled amounts of specially chosen impurities, or **dopants**, which alter the concentration of mobile charge carriers by many orders of magnitude. Dopants with excess valence electrons (called **donors**) can be used to replace a small proportion of the normal atoms in the crystal lattice and thereby to create a predominance of mobile electrons; the material is then said to be an ***n*-type** semiconductor. Thus atoms from group V (e.g., P or As) replacing some of the group IV atoms in an elemental semiconductor, or atoms from group VI (e.g., Se or Te) replacing some of the group V atoms in a III-V binary semiconductor, produce an *n*-type material. Similarly, a ***p*-type** material can be made by using dopants with a deficiency of valence electrons, called **acceptors**. The result is a predominance of holes. Group-IV atoms in an elemental semiconductor replaced with some group-III atoms (e.g., B or In), or group-III atoms in a III-V binary semiconductor replaced with some group-II atoms (e.g., Zn or Cd), produce a *p*-type material. Group IV atoms act as donors in group III and as acceptors in group V, and therefore can be used to produce an excess of both electrons and holes in III-V materials.

Undoped semiconductors (i.e., semiconductors with no intentional doping) are referred to as **intrinsic** materials, whereas doped semiconductors are called **extrinsic**

materials. The concentrations of mobile electrons and holes are equal in an intrinsic semiconductor, $n = p \equiv n_i$, where n_i increases with temperature at an exponential rate. The concentration of mobile electrons in an n -type semiconductor (called **majority carriers**) is far greater than the concentration of holes (called **minority carriers**), i.e., $n \gg p$. The opposite is true in p -type semiconductors, for which holes are majority carriers and $p \gg n$. Doped semiconductors at room temperature typically have a majority carrier concentration that is approximately equal to the impurity concentration.

C. Electron and Hole Concentrations

Determining the concentration of carriers (electrons and holes) as a function of energy requires knowledge of:

- The density of allowed energy levels (density of states).
- The probability that each of these levels is occupied.

Density of States

The quantum state of an electron in a semiconductor material is characterized by its energy E , its wavevector \mathbf{k} [the magnitude of which is approximately related to E by (15.1-1) or (15.1-2)], and its spin. The state is described by a wavefunction satisfying certain boundary conditions.

An electron near the conduction band edge may be approximately described as a particle of mass m_c confined to a three-dimensional cubic box (of dimension d) with perfectly reflecting walls, i.e., a three-dimensional infinite rectangular potential well. The standing-wave solutions require that the components of the wavevector $\mathbf{k} = (k_x, k_y, k_z)$ assume the discrete values $\mathbf{k} = (q_1\pi/d, q_2\pi/d, q_3\pi/d)$, where the respective mode numbers, q_1, q_2, q_3 , are positive integers. This result is a three-dimensional generalization of the one-dimensional case discussed in Exercise 12.1-1. The tip of the vector \mathbf{k} must lie on the points of a lattice whose cubic unit cell has dimension π/d . There are therefore $(d/\pi)^3$ points per unit volume in \mathbf{k} -space. The number of states whose wavevectors \mathbf{k} have magnitudes between 0 and k is determined by counting the number of points lying within the positive octant of a sphere of radius k [with volume $\approx (\frac{1}{8})4\pi k^3/3 = \pi k^3/6$]. Because of the two possible values of the electron spin, each point in \mathbf{k} -space corresponds to two states. There are therefore approximately $2(\pi k^3/6)/(\pi/d)^3 = (k^3/3\pi^2)d^3$ such points in the volume d^3 and $(k^3/3\pi^2)$ points per unit volume. It follows that the number of states with electron wavenumbers between k and $k + \Delta k$, per unit volume, is $\rho(k)\Delta k = [(d/dk)(k^3/3\pi^2)]\Delta k = (k^2/\pi^2)\Delta k$, so that the density of states is

$$\rho(k) = \frac{k^2}{\pi^2}.$$

(15.1-3)

Density of States

This derivation is identical to that used for counting the number of modes that can be supported in a three-dimensional electromagnetic resonator (see Sec. 9.1C). In the case of electromagnetic modes there are two degrees of freedom associated with the field polarization (i.e., two photon spin values), whereas in the semiconductor case there are two spin values associated with the electron state. In resonator optics the allowed electromagnetic solutions for \mathbf{k} were converted into allowed frequencies through the linear frequency-wavenumber relation $\nu = ck/2\pi$. In semiconductor physics, on the other hand, the allowed solutions for \mathbf{k} are converted into allowed

energies through the quadratic energy-wavenumber relations given in (15.1-1) and (15.1-2).

If $\rho_c(E)\Delta E$ represents the number of conduction-band energy levels (per unit volume) lying between E and $E + \Delta E$, then, because of the one-to-one correspondence between E and k governed by (15.1-1), the densities $\rho_c(E)$ and $\rho(k)$ must be related by $\rho_c(E)dE = \rho(k)dk$. Thus the density of allowed energies in the conduction band is $\rho_c(E) = \rho(k)/(dE/dk)$. Similarly, the density of allowed energies in the valence band is $\rho_v(E) = \rho(k)/(dE/dk)$, where E is given by (15.1-2). The approximate quadratic $E-k$ relations (15.1-1) and (15.1-2), which are valid near the edges of the conduction band and valence band, respectively, are used to evaluate the derivative dE/dk for each band. The result that obtains is

$$\rho_c(E) = \frac{(2m_c)^{3/2}}{2\pi^2\hbar^3} (E - E_c)^{1/2}, \quad E \geq E_c \tag{15.1-4}$$

$$\rho_v(E) = \frac{(2m_v)^{3/2}}{2\pi^2\hbar^3} (E_v - E)^{1/2}, \quad E \leq E_v. \tag{15.1-5}$$

Density of States
Near Band Edges

The square-root relation is a result of the quadratic energy-wavenumber formulas for electrons and holes near the band edges. The dependence of the density of states on energy is illustrated in Fig. 15.1-7. It is zero at the band edge, increasing away from it at a rate that depends on the effective masses of the electrons and holes. The values of m_c and m_v for Si and GaAs that were provided in Table 15.1-1 are actually averaged values suitable for calculating the density of states.

Probability of Occupancy

In the absence of thermal excitation (at $T = 0$ K), all electrons occupy the lowest possible energy levels, subject to the Pauli exclusion principle. The valence band is then completely filled (there are no holes) and the conduction band is completely empty (it

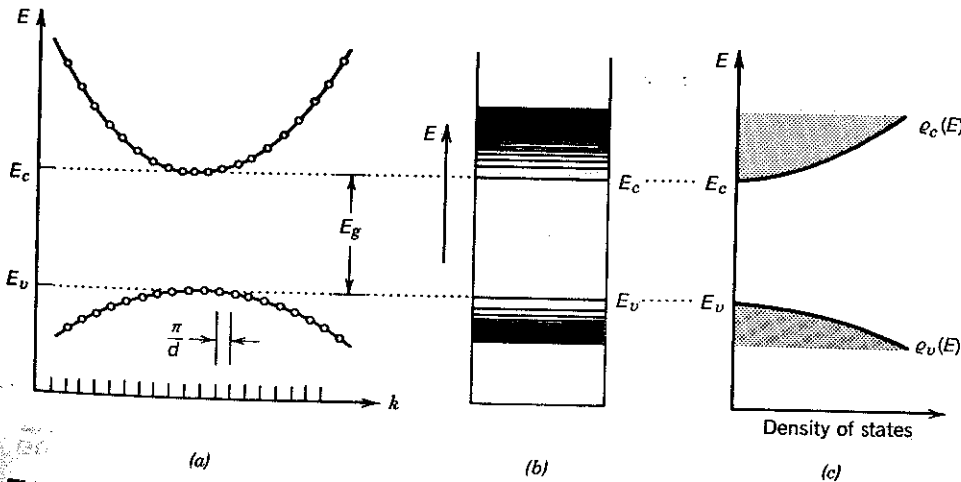


Figure 15.1-7 (a) Cross section of the $E-k$ diagram (e.g., in the direction of the k_1 component with k_2 and k_3 fixed). (b) Allowed energy levels (at all k). (c) Density of states near the edges of the conduction and valence bands. $\rho_c(E)dE$ is the number of quantum states of energy between E and $E + dE$, per unit volume, in the conduction band. $\rho_v(E)$ has an analogous interpretation for the valence band.

contains no electrons). When the temperature is raised, thermal excitations raise some electrons from the valence band to the conduction band, leaving behind empty states in the valence band (holes). The laws of statistical mechanics dictate that under conditions of thermal equilibrium at temperature T , the probability that a given state of energy E is occupied by an electron is determined by the Fermi function

$$f(E) = \frac{1}{\exp[(E - E_f)/k_B T] + 1}, \quad (15.1-6)$$

Fermi Function

where k_B is Boltzmann's constant (at $T = 300$ K, $k_B T = 0.026$ eV) and E_f is a constant known as the Fermi energy or Fermi level. This function is also known as the Fermi-Dirac distribution. The energy level E is either occupied [with probability $f(E)$], or it is empty [with probability $1 - f(E)$]. The probabilities $f(E)$ and $1 - f(E)$ depend on the energy E in accordance with (15.1-6). The function $f(E)$ is not itself a probability distribution, and it does not integrate to unity; rather, it is a sequence of occupation probabilities of successive energy levels.

Because $f(E_f) = \frac{1}{2}$ whatever the temperature T , the Fermi level is that energy level for which the probability of occupancy (if there were an allowed state there) would be $\frac{1}{2}$. The Fermi function is a monotonically decreasing function of E (Fig. 15.1-8). At $T = 0$ K, $f(E)$ is 0 for $E > E_f$ and 1 for $E \leq E_f$. This establishes the significance of E_f ; it is the division between the occupied and unoccupied energy levels at $T = 0$ K. Since $f(E)$ is the probability that the energy level E is occupied, $1 - f(E)$ is the probability that it is empty, i.e., that it is occupied by a hole (if E lies in the valence band). Thus for energy level E :

$f(E)$ = probability of occupancy by an electron

$1 - f(E)$ = probability of occupancy by a hole (valence band).

These functions are symmetric about the Fermi level.

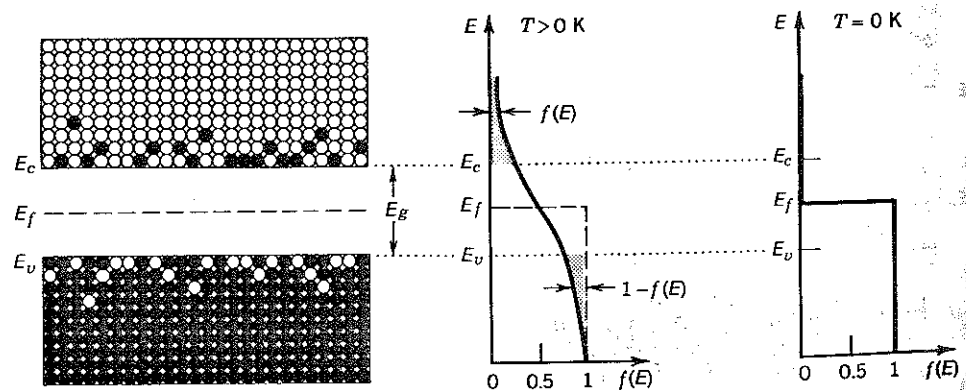


Figure 15.1-8 The Fermi function $f(E)$ is the probability that an energy level E is filled with an electron; $1 - f(E)$ is the probability that it is empty. In the valence band, $1 - f(E)$ is the probability that energy level E is occupied by a hole. At $T = 0$ K, $f(E) = 1$ for $E < E_f$, and $f(E) = 0$ for $E > E_f$; i.e., there are no electrons in the conduction band and no holes in the valence band.

When $E - E_f \gg k_B T$, $f(E) \approx \exp[-(E - E_f)/k_B T]$, so that the high-energy tail of the Fermi function in the conduction band decreases exponentially with increasing energy. The Fermi function is then proportional to the Boltzmann distribution, which describes the exponential energy dependence of the fraction of a population of atoms excited to a given energy level (see Sec. 12.1B). By symmetry, when $E < E_f$ and $E_f - E \gg k_B T$, $1 - f(E) \approx \exp[-(E_f - E)/k_B T]$; i.e., the probability of occupancy by holes in the valence band decreases exponentially as the energy decreases well below the Fermi level.

Thermal-Equilibrium Carrier Concentrations

Let $n(E)\Delta E$ and $p(E)\Delta E$ be the number of electrons and holes per unit volume, respectively, with energy lying between E and $E + \Delta E$. The densities $n(E)$ and $p(E)$ can be obtained by multiplying the densities of states at energy level E by the probabilities of occupancy of the level by electrons or holes, so that

$$n(E) = g_c(E)f(E), \quad p(E) = g_v(E)[1 - f(E)]. \quad (15.1-7)$$

The concentrations (populations per unit volume) of electrons and holes n and p are then obtained from the integrals

$$n = \int_{E_c}^{\infty} n(E) dE, \quad p = \int_{-\infty}^{E_v} p(E) dE. \quad (15.1-8)$$

In an intrinsic (pure) semiconductor at any temperature, $n = p$ because thermal excitations always create electrons and holes in pairs. The Fermi level must therefore be placed at an energy level such that $n = p$. If $m_v = m_c$, the functions $n(E)$ and $p(E)$ are symmetric, so that E_f must lie precisely in the middle of the bandgap (Fig. 15.1-9). In most intrinsic semiconductors the Fermi level does indeed lie near the middle of the bandgap.

The energy-band diagrams, Fermi functions, and equilibrium concentrations of electrons and holes for n -type and p -type doped semiconductors are illustrated in Figs. 15.1-10 and 15.1-11, respectively. Donor electrons occupy an energy E_D slightly below the conduction-band edge so that they are easily raised to it. If $E_D = 0.01$ eV, for example, at room temperature ($k_B T = 0.026$ eV) most donor electrons will be ther-

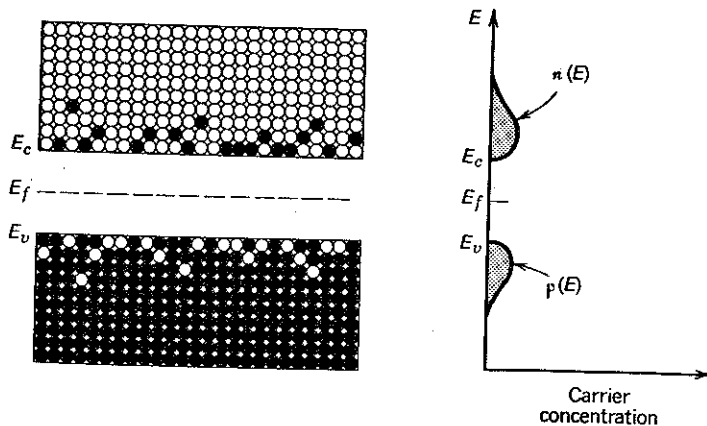


Figure 15.1-9 The concentrations of electrons and holes, $n(E)$ and $p(E)$, as a function of energy E in an intrinsic semiconductor. The total concentrations of electrons and holes are n and p , respectively.

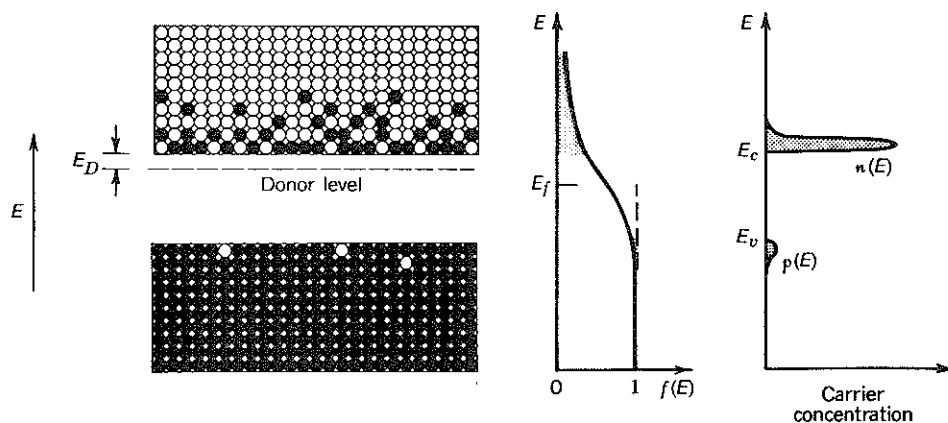


Figure 15.1-10 Energy-band diagram, Fermi function $f(E)$, and concentrations of mobile electrons and holes $n(E)$ and $p(E)$ in an n -type semiconductor.

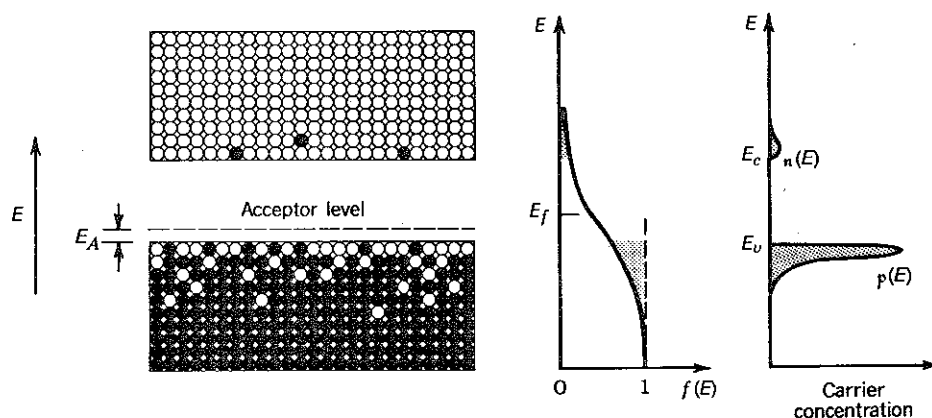


Figure 15.1-11 Energy-band diagram, Fermi function $f(E)$, and concentrations of mobile electrons and holes $n(E)$ and $p(E)$ in a p -type semiconductor.

mally excited into the conduction band. As a result, the Fermi level [where $f(E_f) = \frac{1}{2}$] lies above the middle of the bandgap. For a p -type semiconductor, the acceptor energy level lies at an energy E_A just above the valence-band edge so that the Fermi level is below the middle of the bandgap. Our attention has been directed to the mobile carriers in doped semiconductors. These materials are, of course, electrically neutral as assured by the fixed donor and acceptor ions, so that $n + N_A = p + N_D$ where N_A and N_D are, respectively, the number of ionized acceptors and donors per unit volume.

EXERCISE 15.1-1

Exponential Approximation of the Fermi Function. When $E - E_f \gg k_B T$, the Fermi function $f(E)$ may be approximated by an exponential function. Similarly, when $E_f - E \gg k_B T$, $1 - f(E)$ may be approximated by an exponential function. These conditions apply when the Fermi level lies within the bandgap, but away from its edges by an energy of at least several times $k_B T$ (at room temperature $k_B T \approx 0.026$ eV whereas $E_g = 1.11$ eV in Si and 1.42 eV in GaAs). Using these approximations, which apply for both intrinsic and

doped semiconductors, show that (15.1-8) gives

$$n = N_c \exp\left(-\frac{E_c - E_f}{k_B T}\right) \quad (15.1-9a)$$

$$p = N_v \exp\left(-\frac{E_f - E_v}{k_B T}\right) \quad (15.1-9b)$$

$$np = N_c N_v \exp\left(-\frac{E_g}{k_B T}\right), \quad (15.1-10a)$$

where $N_c = 2(2\pi m_c k_B T/h^2)^{3/2}$ and $N_v = 2(2\pi m_v k_B T/h^2)^{3/2}$. Verify that if E_f is closer to the conduction band and $m_v = m_c$, then $n > p$ whereas if it is closer to the valence band, then $p > n$.

Law of Mass Action

Equation (15.1-10a) reveals that the product

$$np = 4\left(\frac{2\pi k_B T}{h^2}\right)^3 (m_c m_v)^{3/2} \exp\left(-\frac{E_g}{k_B T}\right) \quad (15.1-10b)$$

is independent of the location of the Fermi level E_f within the bandgap and the semiconductor doping level, provided that the exponential approximation to the Fermi function is valid. The constancy of the concentration product is called the **law of mass action**. For an intrinsic semiconductor, $n = p \equiv n_i$. Combining this relation with (15.1-10a) then leads to

$$n_i \approx (N_c N_v)^{1/2} \exp\left(-\frac{E_g}{2k_B T}\right), \quad (15.1-11)$$

Intrinsic Carrier
Concentration

revealing that the intrinsic concentration of electrons and holes increases with temperature T at an exponential rate. The law of mass action may therefore be written in the form

$$np = n_i^2. \quad (15.1-12)$$

Law of Mass Action

The values of n_i for different materials vary because of differences in the bandgap energies and effective masses. For Si and GaAs, the room temperature values of intrinsic carrier concentrations are provided in Table 15.1-4.

The law of mass action is useful for determining the concentrations of electrons and holes in doped semiconductors. A moderately doped n -type material, for example, has

TABLE 15.1-4 Intrinsic Concentrations in Si and GaAs at $T = 300 \text{ K}^a$

	$n_i \text{ (cm}^{-3}\text{)}$
Si	1.5×10^{10}
GaAs	1.8×10^6

^aSubstitution of the values of m_c and m_v given in Table 15.1-1, and E_g given in Table 15.1-3, into (15.1-11) will not yield the precise values of n_i given here because of the sensitivity of the formula to the precise values of the parameters.

a concentration of electrons n that is essentially equal to the donor concentration N_D . Using the law of mass action, the hole concentration can be determined from $p = n_i^2/N_D$. Knowledge of n and p allows the Fermi level to be determined by the use of (15.1-8). As long as the Fermi level lies within the bandgap, at an energy greater than several times $k_B T$ from its edges, the approximate relations in (15.1-9) can be used to determine it directly.

If the Fermi level lies inside the conduction (or valence) band, the material is referred to as a **degenerate semiconductor**. In that case, the exponential approximation to the Fermi function cannot be used, so that $np \neq n_i^2$. The carrier concentrations must then be obtained by numerical solution. Under conditions of very heavy doping, the donor (acceptor) impurity band actually merges with the conduction (valence) band to become what is called the **band tail**. This results in an effective decrease of the bandgap.

Quasi-Equilibrium Carrier Concentrations

The occupancy probabilities and carrier concentrations provided above are applicable only for a semiconductor in thermal equilibrium. They are not valid when thermal equilibrium is disturbed. There are, nevertheless, situations in which the conduction-band electrons are in thermal equilibrium among themselves, as are the valence-band holes, but the electrons and holes are not in mutual thermal equilibrium. This can occur, for example, when an external electric current or photon flux induces band-to-band transitions at too high a rate for interband equilibrium to be achieved. This situation, which is known as quasi-equilibrium, arises when the relaxation (decay) times for transitions within each of the bands are much shorter than the relaxation time between the two bands. Typically, the intraband relaxation time $< 10^{-12}$ s, whereas the radiative electron-hole recombination time $\approx 10^{-9}$ s.

Under these circumstances, it is appropriate to use a separate Fermi function for each band; the two Fermi levels are then denoted E_{fc} and E_{fv} and are known as **quasi-Fermi levels** (Fig. 15.1-12). When E_{fc} and E_{fv} lie well inside the conduction and valence bands, respectively, the concentrations of *both* electrons and holes can be quite large.

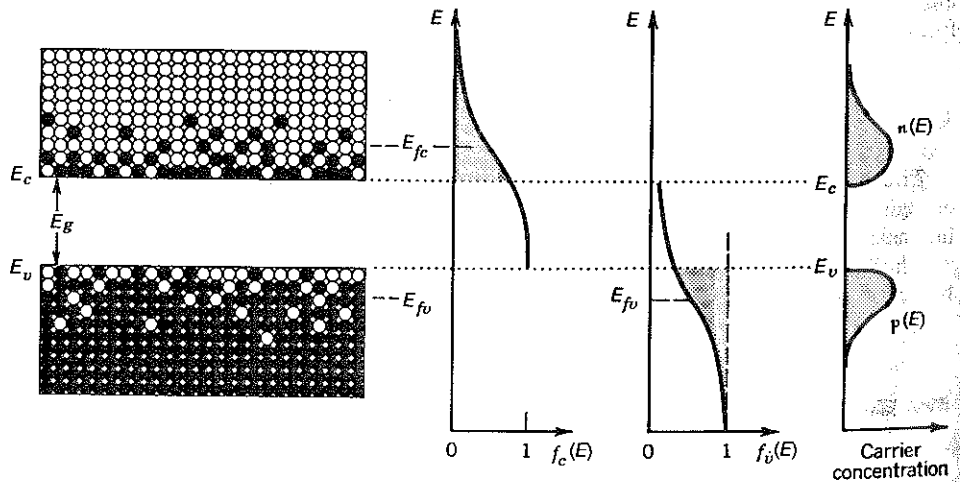


Figure 15.1-12 A semiconductor in quasi-equilibrium. The probability that a particular conduction-band energy level E is occupied by an electron is $f_c(E)$, the Fermi function with Fermi level E_{fc} . The probability that a valence-band energy level E is occupied by a hole is $1 - f_v(E)$, where $f_v(E)$ is the Fermi function with Fermi level E_{fv} . The concentrations of electrons and holes are $n(E)$ and $p(E)$, respectively. Both can be large.