
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

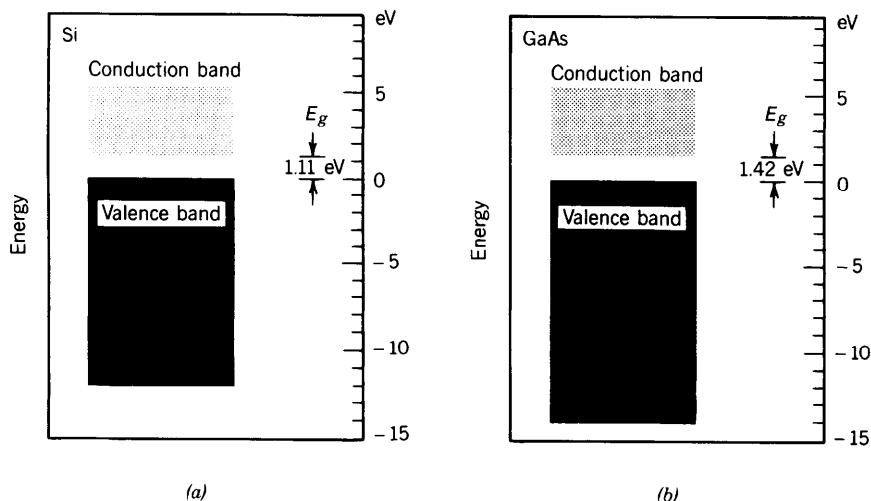


Figure 15.1-1 Energy bands: (a) in Si, and (b) in GaAs.

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 $E = p^2/2m_0 = \hbar^2k^2/2m_0$, 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

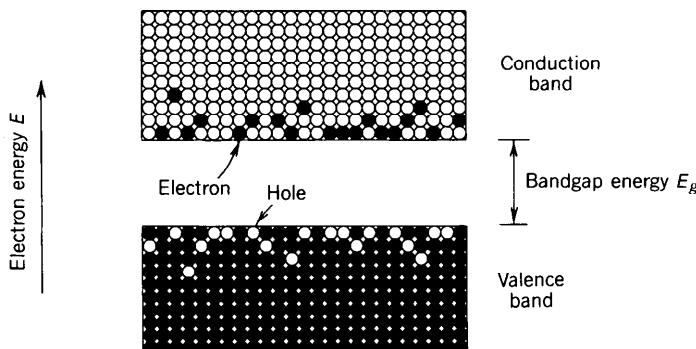


Figure 15.1-2 Electrons in the conduction band and holes in the valence band at $T > 0$ K.

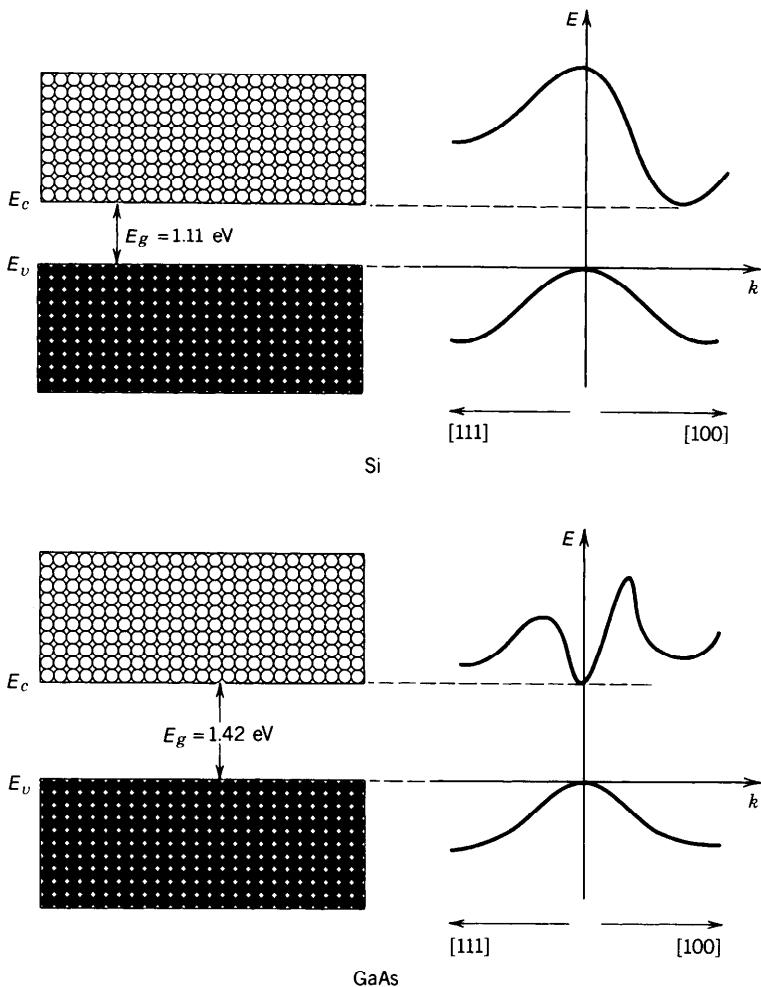


Figure 15.1-3 Cross section of the E - k function for Si and GaAs along the crystal directions [111] and [100].

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 \mathbf{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 \mathbf{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

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

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).

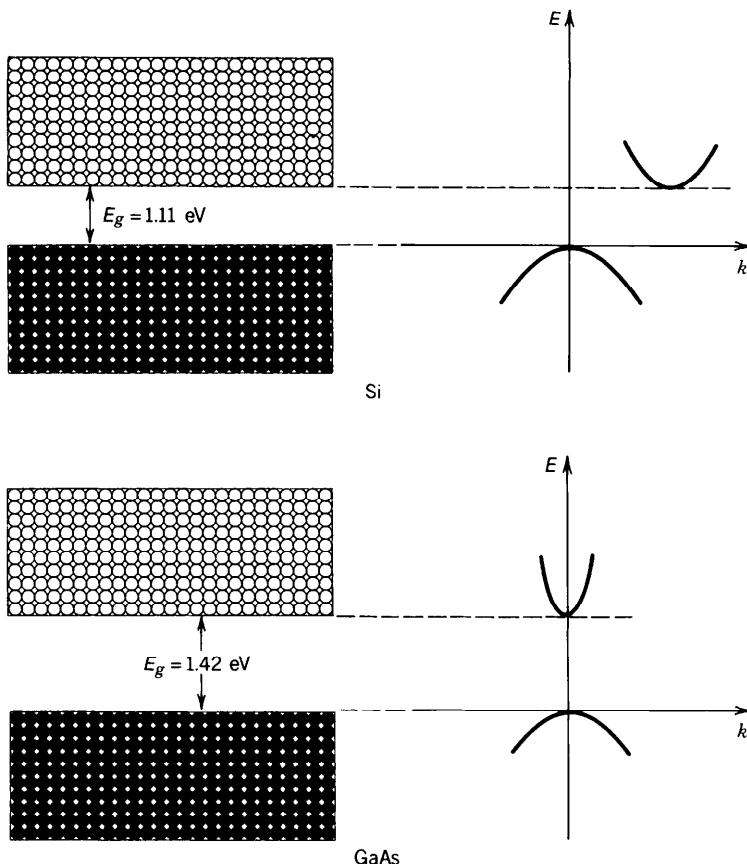


Figure 15.1-4 Approximating the E - k diagram at the bottom of the conduction band and at the top of the valence band of Si and GaAs by parabolas.

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) Gallium (Ga) Indium (In)	Silicon (Si) Germanium (Ge)	Phosphorus (P) Arsenic (As) Antimony (Sb)	Sulfur (S) Selenium (Se) Tellurium (Te)
Cadmium (Cd)				
Mercury (Hg)				

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.

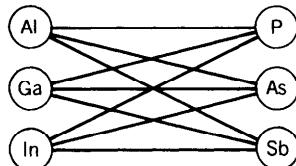
B. Semiconducting Materials

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

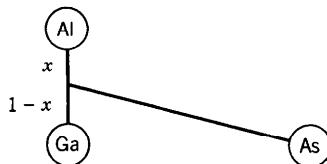
Several elements in group IV of the periodic table are semiconductors. Most important are **silicon** (Si) and **germanium** (Ge). At present most commercial electronic integrated circuits and devices are fabricated from Si. However, these materials are not useful for fabricating photon emitters because 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 $\lambda_g = hc_o/E_g$ (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,

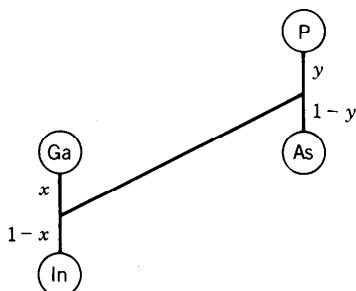
Ternary Semiconductors



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. $(Al_xGa_{1-x})As$, 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, $Al_xGa_{1-x}As$ 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 $Al_xGa_{1-x}As/GaAs$ 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 bandgap-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 $(In_{1-x}Ga_x)(As_{1-y}P_y)$, 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 $y = 2.16(1 - x)$, $(In_{1-x}Ga_x)(As_{1-y}P_y)$ 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 $\lambda_g = hc_o / E_g$, 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

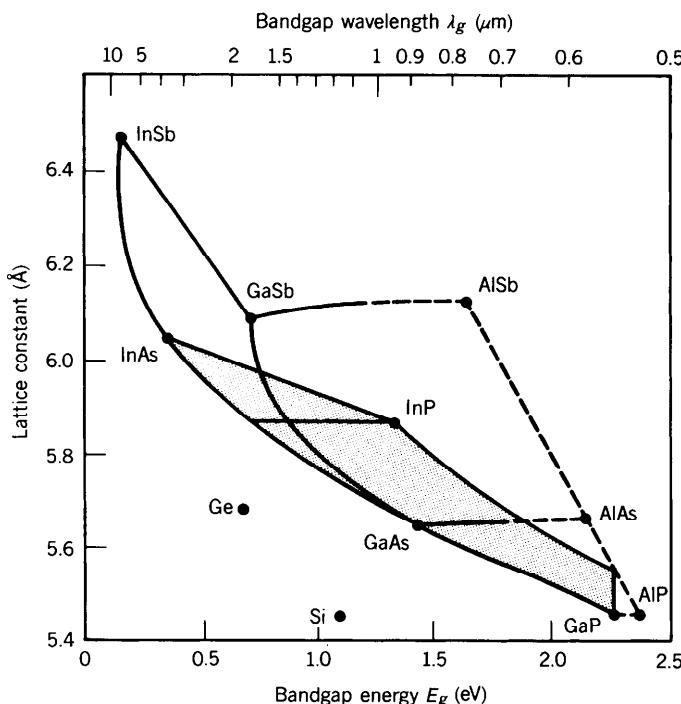


Figure 15.1-5 Lattice constants, bandgap energies, and bandgap wavelengths for Si, Ge, and nine III-V binary compounds. Ternary compounds can be formed from binary materials by motion along the line joining the two points that represent the binary materials. For example, $\text{Al}_x\text{Ga}_{1-x}\text{As}$ 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.

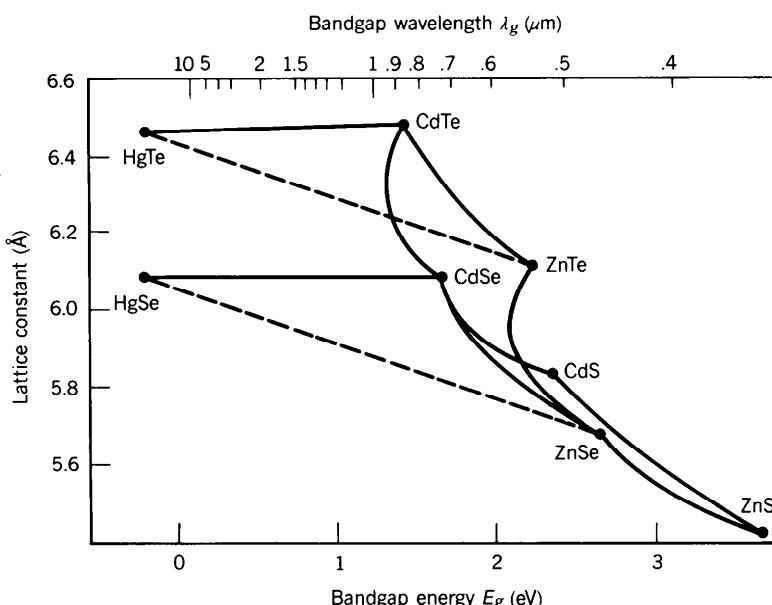


Figure 15.1-6 Lattice constants, bandgap energies, and bandgap wavelengths for some important II-VI binary compounds.

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 $\varrho(k)\Delta k = [(d/dk)(k^3/3\pi^2)]\Delta k = (k^2/\pi^2)\Delta k$, so that the density of states is

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

(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 $\varrho_c(E)\Delta E$ 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 $\varrho_c(E)$ and $\varrho(k)$ must be related by $\varrho_c(E)dE = \varrho(k)dk$. Thus the density of allowed energies in the conduction band is $\varrho_c(E) = \varrho(k)/(dE/dk)$. Similarly, the density of allowed energies in the valence band is $\varrho_v(E) = \varrho(k)/(dE/dk)$, 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

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

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

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

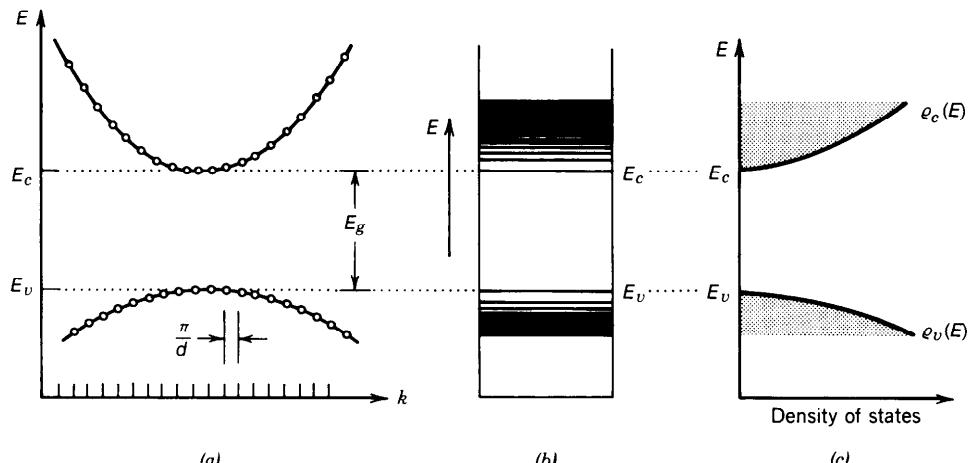


Figure 15.1-7 (a) Cross section of the $E-k$ diagram (e.g., in the direction of the k_1 component with k_2 and k_3 fixed). (b) Allowed energy levels (at all k). (c) Density of states near the edges of the conduction and valence bands. $\varrho_c(E)dE$ is the number of quantum states of energy between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

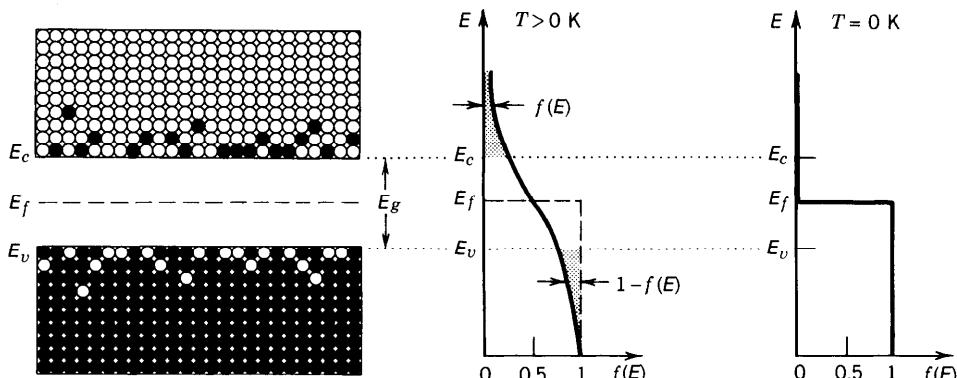


Figure 15.1-8 The Fermi function $f(E)$ is the probability that an energy level E is filled with an electron; $1 - f(E)$ is the probability that it is empty. In the valence band, $1 - f(E)$ is the probability that energy level E is occupied by a hole. At $T = 0$ K, $f(E) = 1$ for $E < E_f$, and $f(E) = 0$ 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

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

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-

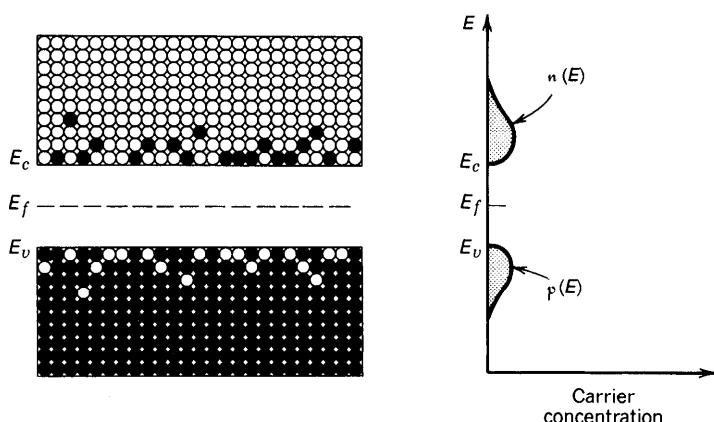


Figure 15.1-9 The concentrations of electrons and holes, $n(E)$ and $p(E)$, as a function of energy E in an intrinsic semiconductor. The total concentrations of electrons and holes are n and p , respectively.

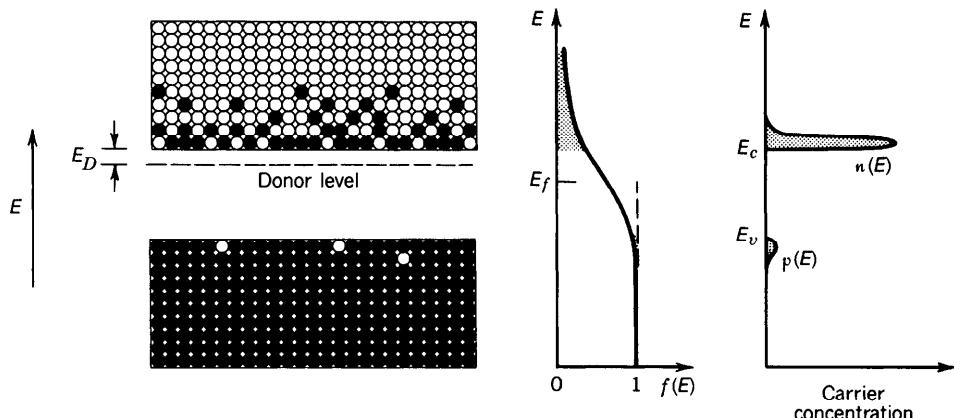


Figure 15.1-10 Energy-band diagram, Fermi function $f(E)$, and concentrations of mobile electrons and holes $n(E)$ and $p(E)$ in an *n*-type semiconductor.

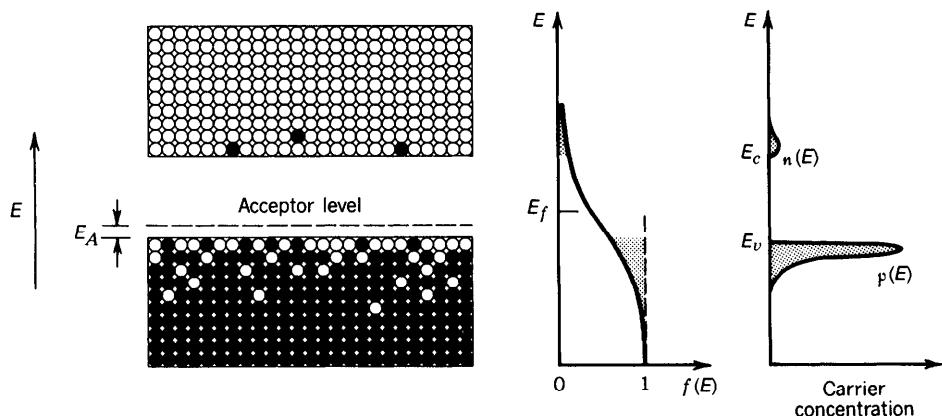


Figure 15.1-11 Energy-band diagram, Fermi function $f(E)$, and concentrations of mobile electrons $n(E)$ and holes $p(E)$ in a *p*-type semiconductor.

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 $T = 300$ K^a

	n_i (cm ⁻³)
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 $p = n^2/N_D$. 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 $n_p \neq n_i^2$. 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.

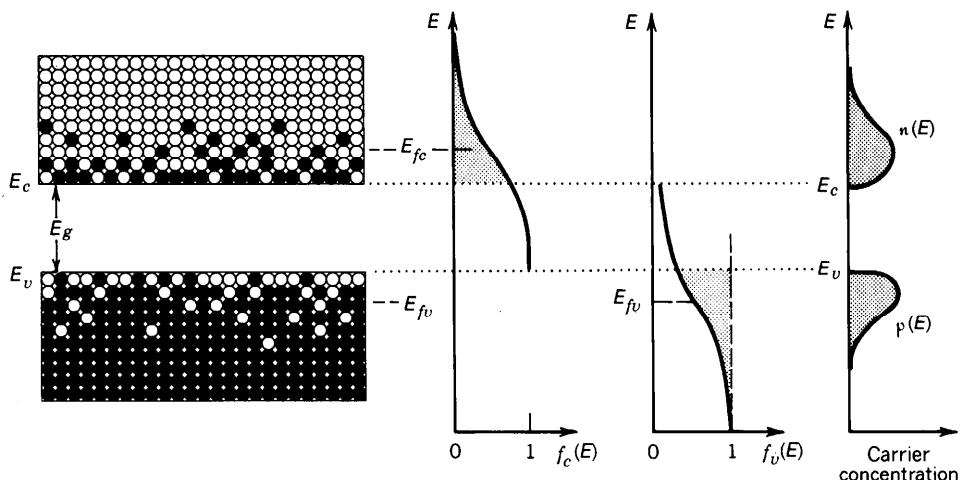


Figure 15.1-12 A semiconductor in quasi-equilibrium. The probability that a particular conduction-band energy level E is occupied by an electron is $f_c(E)$, the Fermi function with Fermi level E_{fc} . The probability that a valence-band energy level E is occupied by a hole is $1 - f_v(E)$, where $f_v(E)$ 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.

EXERCISE 15.1-2**Determination of the Quasi-Fermi Levels Given the Electron and Hole Concentrations**

- (a) Given the concentrations of electrons n and holes p in a semiconductor at $T = 0$ K, use (15.1-7) and (15.1-8) to show that the quasi-Fermi levels are

$$E_{fc} = E_c + (3\pi^2)^{2/3} \frac{\hbar^2}{2m_c} n^{2/3} \quad (15.1-13a)$$

$$E_{fv} = E_v - (3\pi^2)^{2/3} \frac{\hbar^2}{2m_v} p^{2/3} \quad (15.1-13b)$$

- (b) Show that these equations are approximately applicable at an arbitrary temperature T if n and p are sufficiently large so that $E_{fc} - E_c \gg k_B T$ and $E_v - E_{fv} \gg k_B T$, i.e., if the quasi-Fermi levels lie deeply within the conduction and valence bands.

D. Generation, Recombination, and Injection**Generation and Recombination in Thermal Equilibrium**

The thermal excitation of electrons from the valence band into the conduction band results in the *generation* of electron–hole pairs (Fig. 15.1-13). Thermal equilibrium requires that this generation process be accompanied by a simultaneous reverse process of deexcitation. This process, called **electron–hole recombination**, occurs when an electron decays from the conduction band to fill a hole in the valence band (Fig. 15.1-13). The energy released by the electron may take the form of an emitted photon, in which case the process is called **radiative recombination**. **Nonradiative recombination** can occur via a number of independent competing processes, including the transfer of energy to lattice vibrations (creating one or more phonons) or to another free electron (Auger process).

Recombination may also occur indirectly via traps or defect centers. These are energy levels associated with impurities or defects due to grain boundaries, dislocations, or other lattice imperfections, that lie within the energy bandgap. An impurity or defect state can act as a recombination center if it is capable of trapping both the

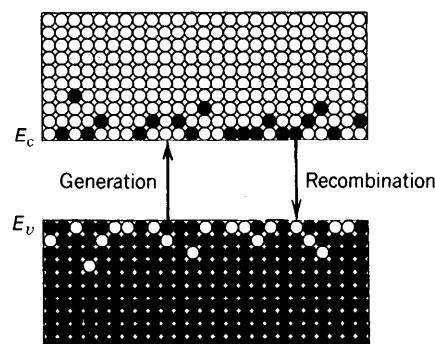


Figure 15.1-13 Electron–hole generation and recombination.

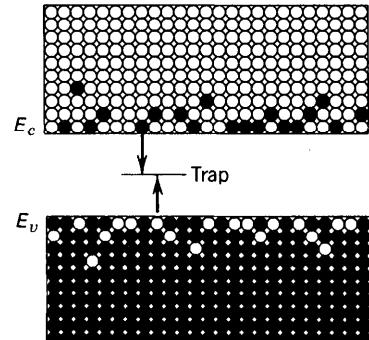


Figure 15.1-14 Electron–hole recombination via a trap.

electron and the hole, thereby increasing their probability of recombining (Fig. 15.1-14). Impurity-assisted recombination may be radiative or nonradiative.

Because it takes both an electron *and* a hole for a recombination to occur, the rate of recombination is proportional to the product of the concentrations of electrons and holes, i.e.,

$$\text{rate of recombination} = \tau n p, \quad (15.1-14)$$

where τ (cm^3/s) is a parameter that depends on the characteristics of the material, including its composition and defects, and on temperature; it also depends relatively weakly on the doping.

The equilibrium concentrations of electrons and holes n_0 and p_0 are established when the generation and recombination rates are in balance. In the steady state, the rate of recombination must equal the rate of generation. If G_0 is the rate of thermal electron–hole generation at a given temperature, then, in thermal equilibrium,

$$G_0 = \tau n_0 p_0.$$

The product of the electron and hole concentrations $n_0 p_0 = G_0 / \tau$ is approximately the same whether the material is *n*-type, *p*-type, or intrinsic. Thus $n_i^2 = G_0 / \tau$, which leads directly to the law of mass action $n_0 p_0 = n_i^2$. This law is therefore seen to be a consequence of the balance between generation and recombination in thermal equilibrium.

Electron–Hole Injection

A semiconductor in thermal equilibrium with carrier concentrations n_0 and p_0 has equal rates of generation and recombination, $G_0 = \tau n_0 p_0$. Now let additional electron–hole pairs be generated at a steady rate R (pairs per unit volume per unit time) by means of an external (nonthermal) injection mechanism. A new steady state will be reached in which the concentrations are $n = n_0 + \Delta n$ and $p = p_0 + \Delta p$. It is clear, however, that $\Delta n = \Delta p$ since the electrons and holes are created in pairs. Equating the new rates of generation and recombination, we obtain

$$G_0 + R = \tau n p. \quad (15.1-15)$$

Substituting $G_0 = \tau n_0 p_0$ into (15.1-15) leads to

$$R = \tau(n p - n_0 p_0) = \tau(n_0 \Delta n + p_0 \Delta n + \Delta n^2) = \tau \Delta n(n_0 + p_0 + \Delta n),$$

which we write in the form

$$R = \frac{\Delta n}{\tau}, \quad (15.1-16)$$

with

$$\tau = \frac{1}{\epsilon[(n_0 + p_0) + \Delta n]}. \quad (15.1-17)$$

For an injection rate such that $\Delta n \ll n_0 + p_0$,

$$\boxed{\tau \approx \frac{1}{\epsilon(n_0 + p_0)}}. \quad (15.1-18)$$

Excess-Carrier
Recombination Lifetime

In an *n*-type material, where $n_0 \gg p_0$, the recombination lifetime $\tau \approx 1/\epsilon n_0$ is inversely proportional to the electron concentration. Similarly, for a *p*-type material where $p_0 \gg n_0$, we obtain $\tau \approx 1/\epsilon p_0$. This simple formulation is not applicable when traps play an important role in the process.

The parameter τ may be regarded as the **electron–hole recombination lifetime** of the injected excess electron–hole pairs. This is readily understood by noting that the injected carrier concentration is governed by the rate equation

$$\frac{d(\Delta n)}{dt} = R - \frac{\Delta n}{\tau},$$

which is similar to (13.2-2). In the steady state $d(\Delta n)/dt = 0$ whereupon (15.1-16), which is like (13.2-10), is recovered. If the source of injection is suddenly removed (R becomes 0) at the time t_0 , then Δn decays exponentially with time constant τ , i.e., $\Delta n(t) = \Delta n(t_0) \exp[-(t - t_0)/\tau]$. In the presence of strong injection, on the other hand, τ is itself a function of Δn , as evident from (15.1-17), so that the rate equation is nonlinear and the decay is no longer exponential.

If the injection rate R is known, the steady-state injected concentration may be determined from

$$\boxed{\Delta n = R\tau}, \quad (15.1-19)$$

permitting the total concentrations $n = n_0 + \Delta n$ and $p = p_0 + \Delta n$ to be determined. Furthermore, if quasi-equilibrium is assumed, (15.1-8) may be used to determine the quasi-Fermi levels. Quasi-equilibrium is not inconsistent with the balance of generation and recombination assumed in the analysis above; it simply requires that the intraband equilibrium time be short in comparison with the recombination time τ .

This type of analysis will prove useful in developing theories of the semiconductor light-emitting diode and the semiconductor diode laser, which are based on enhancing light emission by means of carrier injection (see Chap. 16).

EXERCISE 15.1-3

Electron–Hole Pair Injection in GaAs. Assume that electron–hole pairs are injected into *n*-type GaAs ($E_g = 1.42$ eV, $m_c \approx 0.07m_0$, $m_v \approx 0.5m_0$) at a rate $R = 10^{23}$ per cm³ per second. The thermal equilibrium concentration of electrons is $n_0 = 10^{16}$ cm⁻³. If the recombination parameter $\tau = 10^{-11}$ cm³/s and $T = 300$ K, determine:

- (a) The equilibrium concentration of holes p_0 .
- (b) The recombination lifetime τ .
- (c) The steady-state excess concentration Δn .
- (d) The separation between the quasi-Fermi levels $E_{fc} - E_{fv}$, assuming that $T = 0$ K.

Internal Quantum Efficiency

The **internal quantum efficiency** η_i of a semiconductor material is defined as the ratio of the radiative electron–hole recombination rate to the total (radiative and nonradiative) recombination rate. This parameter is important because it determines the efficiency of light generation in a semiconductor material. The total rate of recombination is given by (15.1-14). If the parameter τ is split into a sum of radiative and nonradiative parts, $\tau = \tau_r + \tau_{nr}$, the internal quantum efficiency is

$$\eta_i = \frac{\tau_r}{\tau} = \frac{\tau_r}{\tau_r + \tau_{nr}}. \quad (15.1-20)$$

The internal quantum efficiency may also be written in terms of the recombination lifetimes since τ is inversely proportional to τ [see (15.1-18)]. Defining the radiative and nonradiative lifetimes τ_r and τ_{nr} , respectively, leads to

$$\frac{1}{\tau} = \frac{1}{\tau_r} + \frac{1}{\tau_{nr}}. \quad (15.1-21)$$

The internal quantum efficiency is then $\tau_r/\tau = (1/\tau_r)/(1/\tau)$, or

$$\eta_i = \frac{\tau}{\tau_r} = \frac{\tau_{nr}}{\tau_r + \tau_{nr}}.$$

(15.1-22)

Internal Quantum Efficiency

The radiative recombination lifetime τ_r governs the rates of photon absorption and emission, as explained in Sec. 15.2B. Its value depends on the carrier concentrations and the material parameter τ_r . For low to moderate injection rates,

$$\tau_r \approx \frac{1}{\tau_r(n_0 + p_0)}, \quad (15.1-23)$$

in accordance with (15.1-18). The nonradiative recombination lifetime is governed by a similar equation. However, if nonradiative recombination takes place via defect centers in the bandgap, τ_{nr} is more sensitive to the concentration of these centers than to the electron and hole concentrations.

TABLE 15.1-5 Approximate Values for Radiative Recombination Rates τ_r , Recombination Lifetimes, and Internal Quantum Efficiency η_i in Si and GaAs^a

	τ_r (cm ³ /s)	τ_r	τ_{nr}	τ	η_i
Si	10^{-15}	10 ms	100 ns	100 ns	10^{-5}
GaAs	10^{-10}	100 ns	100 ns	50 ns	0.5

^aUnder conditions of doping, temperature, and defect concentration specified in the text.

Approximate values for recombination rates and lifetimes in Si and GaAs are provided in Table 15.1-5. Order-of-magnitude values are given for τ_r and τ_r (assuming *n*-type material with a carrier concentration $n_0 = 10^{17}$ cm⁻³ at $T = 300$ K), τ_{nr} (assuming defect centers with a concentration of 10^{15} cm⁻³), τ , and the internal quantum efficiency η_i .

The radiative lifetime for Si is orders of magnitude larger than its overall lifetime, principally because it has an indirect bandgap. This results in a small internal quantum efficiency. For GaAs, on the other hand, the decay is largely via radiative transitions (it has a direct bandgap), and consequently the internal quantum efficiency is large. GaAs and other direct-gap materials are therefore useful for fabricating light-emitting structures, whereas Si and other indirect-gap materials are not.

E. Junctions

Junctions between differently doped regions of a semiconductor material are called **homojunctions**. An important example is the *p-n* junction, which is discussed in this subsection. Junctions between different semiconductor materials are called **heterojunctions**. These are discussed subsequently.

The p-n Junction

The *p-n* junction is a homojunction between a *p*-type and an *n*-type semiconductor. It acts as a diode which can serve in electronics as a rectifier, logic gate, voltage regulator (Zener diode), and tuner (varactor diode); and in optoelectronics as a light-emitting diode (LED), laser diode, photodetector, and solar cell.

A *p-n* junction consists of a *p*-type and an *n*-type section of the same semiconductor materials in metallurgical contact with each other. The *p*-type region has an abundance of holes (majority carriers) and few mobile electrons (minority carriers); the *n*-type region has an abundance of mobile electrons and few holes (Fig. 15.1-15). Both charge carriers are in continuous random thermal motion in all directions.

When the two regions are brought into contact (Fig. 15.1-16), the following sequence of events takes place:

- Electrons and holes diffuse from areas of high concentration toward areas of low concentration. Thus electrons diffuse away from the *n*-region into the *p*-region, leaving behind positively charged ionized donor atoms. In the *p*-region the electrons recombine with the abundant holes. Similarly, holes diffuse away from the *p*-region, leaving behind negatively charged ionized acceptor atoms. In the *n*-region the holes recombine with the abundant mobile electrons. This diffusion process cannot continue indefinitely, however, because it causes a disruption of the charge balance in the two regions.
- As a result, a narrow region on both sides of the junction becomes almost totally depleted of *mobile* charge carriers. This region is called the **depletion layer**. It

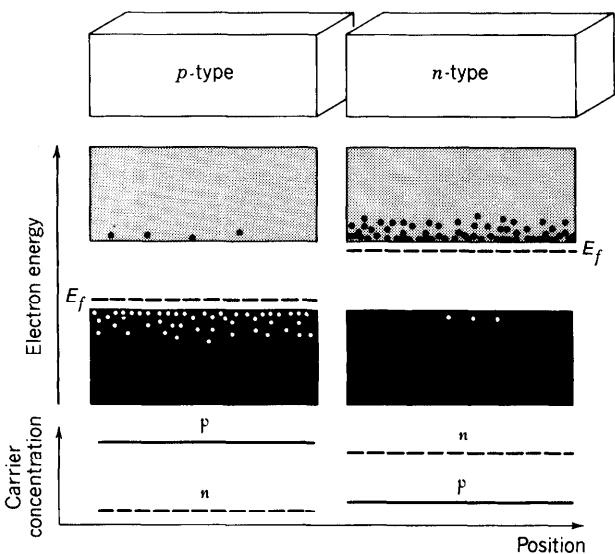


Figure 15.1-15 Energy levels and carrier concentrations of a *p*-type and an *n*-type semiconductor before contact.

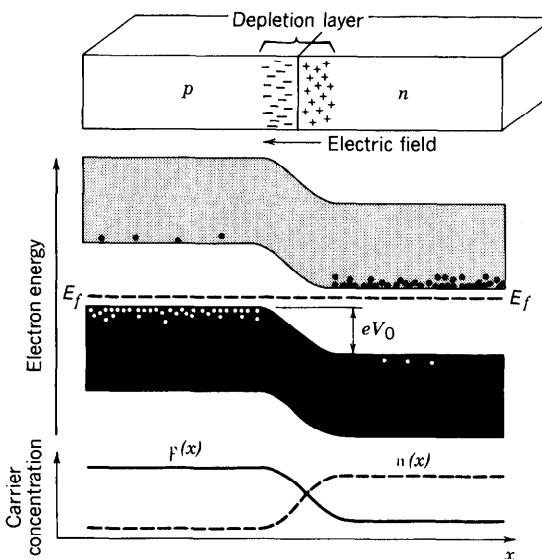


Figure 15.1-16 A *p-n* junction in thermal equilibrium at $T > 0$ K. The depletion-layer, energy-band diagram, and concentrations (on a logarithmic scale) of mobile electrons $n(x)$ and holes $p(x)$ are shown as functions of position x . The built-in potential difference V_0 corresponds to an energy eV_0 , where e is the magnitude of the electron charge.

contains only the *fixed* charges (positive ions on the *n*-side and negative ions on the *p*-side). The thickness of the depletion layer in each region is inversely proportional to the concentration of dopants in the region.

- The fixed charges create an electric field in the depletion layer which points from the *n*-side toward the *p*-side of the junction. This **built-in field** obstructs the diffusion of further mobile carriers through the junction region.
- An equilibrium condition is established that results in a net built-in potential difference V_0 between the two sides of the depletion layer, with the *n*-side exhibiting a higher potential than the *p*-side.
- The built-in potential provides a lower potential energy for an electron on the *n*-side relative to the *p*-side. As a result, the energy bands bend as shown in Fig. 15.1-16. In thermal equilibrium there is only a single Fermi function for the entire structure so that the Fermi levels in the *p*- and *n*-regions must align.
- No *net* current flows across the junction. The diffusion and drift currents cancel for the electrons and holes independently.

The Biased Junction

An externally applied potential will alter the potential difference between the *p*- and *n*-regions. This, in turn, will modify the flow of majority carriers, so that the junction can be used as a “gate.” If the junction is **forward biased** by applying a positive voltage V to the *p*-region (Fig. 15.1-17), its potential is increased with respect to the *n*-region, so that an electric field is produced in a direction opposite to that of the built-in field. The presence of the external bias voltage causes a departure from equilibrium and a misalignment of the Fermi levels in the *p*- and *n*-regions, as well as in the depletion layer. The presence of two Fermi levels in the depletion layer, E_{fc} and E_{fv} , represents a state of quasi-equilibrium.

The net effect of the forward bias is a reduction in the height of the potential-energy hill by an amount eV . The majority carrier current turns out to increase by an exponential factor $\exp(eV/k_B T)$ so that the net current becomes $i = i_s \exp(eV/k_B T) - i_s$, where i_s is a constant. The excess majority carrier holes and electrons that enter

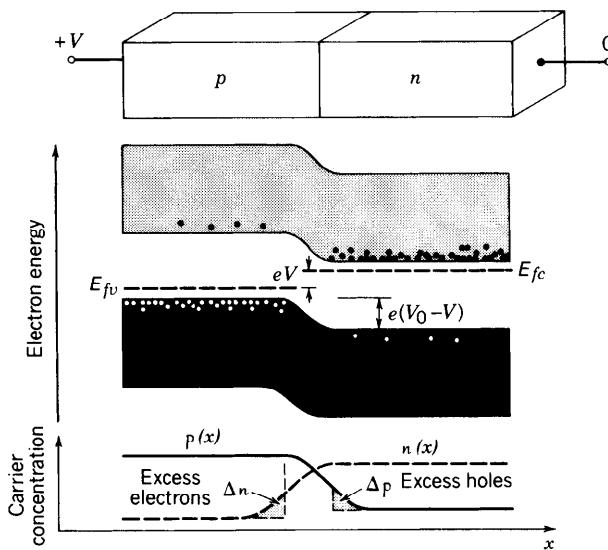


Figure 15.1-17 Energy-band diagram and carrier concentrations in a forward-biased *p*-*n* junction.

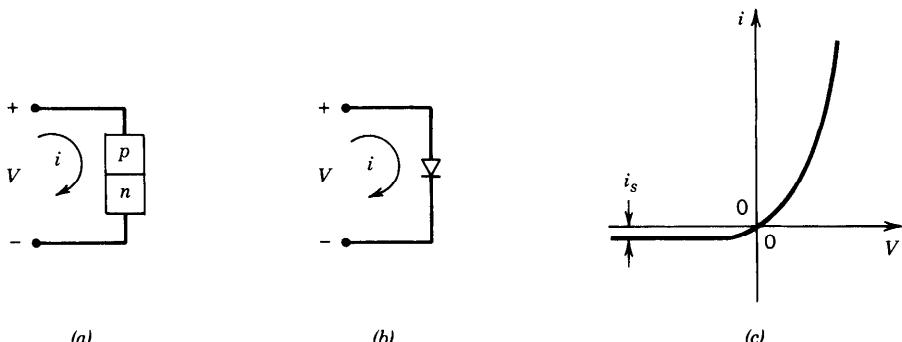


Figure 15.1-18 (a) Voltage and current in a *p*-*n* junction. (b) Circuit representation of the *p*-*n* junction diode. (c) Current–voltage characteristic of the ideal *p*-*n* junction diode.

the *n*- and *p*-regions, respectively, become minority carriers and recombine with the local majority carriers. Their concentration therefore decreases with distance from the junction as shown in Fig. 15.1-17. This process is known as **minority carrier injection**.

If the junction is **reverse biased** by applying a negative voltage *V* to the *p*-region, the height of the potential-energy hill is augmented by *eV*. This impedes the flow of majority carriers. The corresponding current is multiplied by the exponential factor $\exp(eV/k_B T)$, where *V* is negative; i.e., it is reduced. The net result for the current is $i = i_s \exp(eV/k_B T) - i_s$, so that a small current of magnitude $\approx i_s$ flows in the reverse direction when $|V| \gg k_B T/e$.

A *p*-*n* junction therefore acts as a diode with a current–voltage (*i*–*V*) characteristic

$$i = i_s \left[\exp\left(\frac{eV}{k_B T}\right) - 1 \right],$$

(15.1-24)
Ideal Diode
Characteristic

as illustrated in Fig. 15.1-18.

The response of a *p*-*n* junction to a dynamic (ac) applied voltage is determined by solving the set of differential equations governing the processes of electron and hole diffusion, drift (under the influence of the built-in and external electric fields), and recombination. These effects are important for determining the speed at which the diode can be operated. They may be conveniently modeled by two capacitances, a junction capacitance and a diffusion capacitance, in parallel with an ideal diode. The **junction capacitance** accounts for the time necessary to change the fixed positive and negative charges stored in the depletion layer when the applied voltage changes. The thickness *l* of the depletion layer turns out to be proportional to $(V_0 - V)^{1/2}$; it therefore increases under reverse-bias conditions (negative *V*) and decreases under forward-bias conditions (positive *V*). The junction capacitance $C = \epsilon A/l$ (where *A* is the area of the junction) is therefore inversely proportional to $(V_0 - V)^{1/2}$. The junction capacitance of a reverse-biased diode is smaller (and the *RC* response time is therefore shorter) than that of a forward-biased diode. The dependence of *C* on *V* is used to make voltage-variable capacitors (varactors).

Minority carrier injection in a forward-biased diode is described by the **diffusion capacitance**, which depends on the minority carrier lifetime and the operating current.

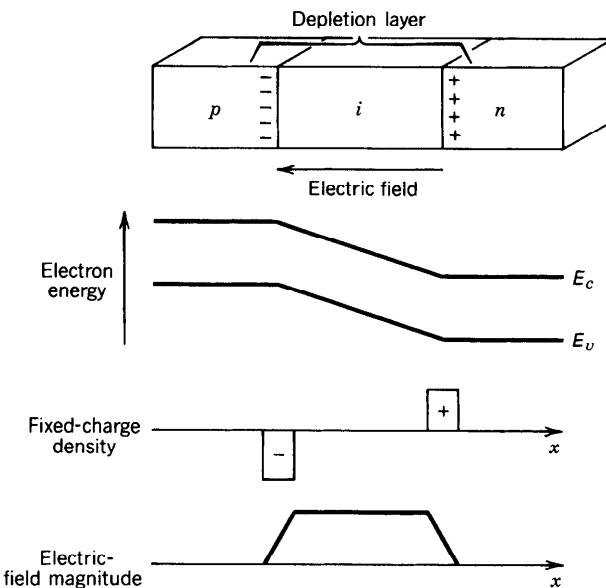


Figure 15.1-19 Electron energy, fixed-charge density, and electric field magnitude for a $p-i-n$ diode in thermal equilibrium.

The $p-i-n$ Junction Diode

A $p-i-n$ diode is made by inserting a layer of intrinsic (or lightly doped) semiconductor material between a p -type region and an n -type region (Fig. 15.1-19). Because the depletion layer extends into each side of a junction by a distance inversely proportional to the doping concentration, the depletion layer of the $p-i$ junction penetrates deeply into the i -region. Similarly, the depletion layer of the $i-n$ junction extends well into the i -region. As a result, the $p-i-n$ diode can behave like a $p-n$ junction with a depletion layer that encompasses the entire intrinsic region. The electron energy, density of fixed charges, and the electric field in a $p-i-n$ diode in thermal equilibrium are illustrated in Fig. 15.1-19. One advantage of using a diode with a large depletion layer is its small junction capacitance and its consequent fast response. For this reason, $p-i-n$ diodes are favored over $p-n$ diodes for use as semiconductor photodiodes. The large depletion layer also permits an increased fraction of the incident light to be captured, thereby increasing the photodetection efficiency (see Sec. 17.3B).

F. Heterojunctions

Junctions between different semiconductor materials are called heterojunctions. Their development has been made possible by modern material growth techniques. Heterojunctions are used in novel bipolar and field-effect transistors, and in optical sources and detectors. They can provide substantial improvement in the performance of electronic and optoelectronic devices. In particular, in photonics the juxtaposition of different semiconductors can be advantageous in several respects:

- Junctions between materials of different bandgap create localized jumps in the energy-band diagram. A potential energy discontinuity provides a barrier that can be useful in preventing selected charge carriers from entering regions where they are undesired. This property may be used in a $p-n$ junction, for example, to

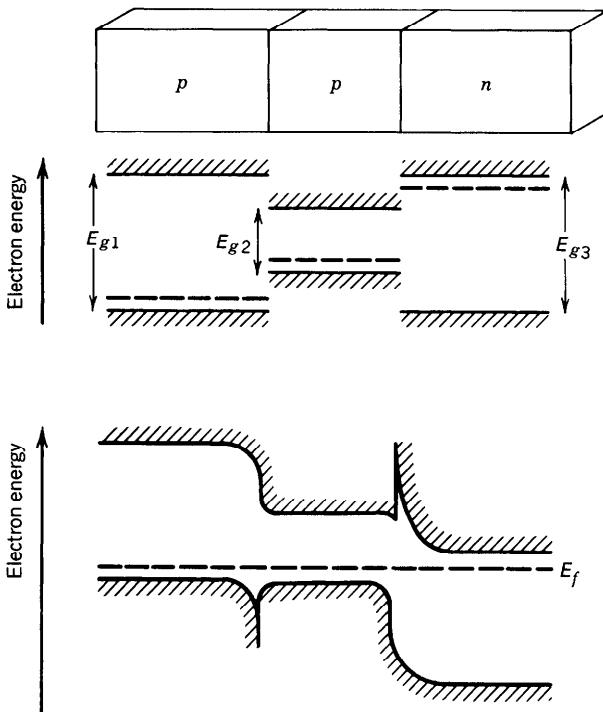


Figure 15.1-20 The $p-p-n$ double heterojunction structure. The middle layer is of narrower bandgap than the outer layers. In equilibrium, the Fermi levels align so that the edge of the conduction band drops sharply at the $p-p$ junction and the edge of the valence band drops sharply at the $p-n$ junction. The ratio of the difference in conduction-band energies to the difference in valence-band energies is known as the **band offset**. When the device is forward biased, these jumps act as barriers that confine the injected minority carriers. Electrons injected from the n -region, for example, are prevented from diffusing beyond the barrier at the $p-p$ junction. Similarly, holes injected from the p -region are not permitted to diffuse beyond the energy barrier at the $p-n$ junction. This double heterostructure therefore forces electrons and holes to occupy a narrow common region. This is essential for the efficient operation of an injection laser diode (see Secs. 16.2 and 16.3).

reduce the proportion of current carried by minority carriers, and thus to increase injection efficiency (see Fig. 15.1-20).

- Discontinuities in the energy-band diagram created by two heterojunctions can be useful for confining charge carriers to a desired region of space. For example, a layer of narrow bandgap material can be sandwiched between two layers of a wider bandgap material, as shown in the $p-p-n$ structure illustrated in Fig. 15.1-20 (which consists of a $p-p$ heterojunction and a $p-n$ heterojunction). This double heterostructure is effectively used in the fabrication of diode lasers, as explained in Sec. 16.3.
- Heterojunctions are useful for creating energy-band discontinuities that accelerate carriers at specific locations. The additional kinetic energy suddenly imparted to a carrier can be useful for selectively enhancing the probability of impact ionization in a multilayer avalanche photodiode (see Sec. 17.4A).
- Semiconductors of different bandgap type (direct and indirect) can be used in the same device to select regions of the structure where light is emitted. Only semiconductors of the direct-gap type can efficiently emit light (see Sec. 15.2).

- Semiconductors of different bandgap can be used in the same device to select regions of the structure where light is absorbed. Semiconductor materials whose bandgap energy is larger than the incident photon energy will be transparent, acting as a “window layer.”
- Heterojunctions of materials with different refractive indices can be used to create optical waveguides that confine and direct photons.

*G. Quantum Wells and Superlattices

Heterostructures of thin layers of semiconductor materials can be grown epitaxially, i.e., as lattice-matched layers of one semiconductor material over another, by using techniques such as molecular-beam epitaxy (MBE), liquid-phase epitaxy (LPE), and vapor-phase epitaxy (VPE), of which a common variant is metal-organic chemical vapor deposition (MOCVD). MBE makes use of molecular beams of the constituent elements that are caused to impinge on an appropriately prepared substrate in a high-vacuum environment, LPE uses the cooling of a saturated solution containing the constituents in contact with the substrate, and MOCVD uses gases in a reactor. The compositions and dopings of the individual layers are determined by manipulating the arrival rates of the molecules and the temperature of the substrate surface and can be made as thin as monolayers.

When the layer thickness is comparable to, or smaller than, the de Broglie wavelength of thermalized electrons (e.g., in GaAs the de Broglie wavelength ≈ 50 nm), the energy–momentum relation for a bulk semiconductor material no longer applies. Three structures offer substantial advantages for use in photonics: quantum wells, quantum wires, and quantum dots. The appropriate energy–momentum relations for these structures are derived below. Applications are deferred to subsequent chapters (see Secs. 16.3B and 17.4A).

Quantum Wells

A quantum well is a double heterojunction structure consisting of an ultrathin ($\lesssim 50$ nm) layer of semiconductor material whose bandgap is smaller than that of the surrounding material (Fig. 15.1-21). An example is provided by a thin layer of GaAs surrounded by AlGaAs (see Fig. 12.1-8). The sandwich forms conduction- and valence-band rectangular potential wells within which electrons and holes are confined: electrons in the conduction-band well and holes in the valence-band well. A sufficiently deep potential well can be approximated as an infinite potential well (see Fig. 12.1-9).

The energy levels E_q of a particle of mass m (m_c for electrons and m_v for holes) confined to a *one-dimensional* infinite rectangular well of full width d are determined by solving the time-independent Schrödinger equation. From Exercise 12.1-1,

$$E_q = \frac{\hbar^2 (q\pi/d)^2}{2m}, \quad q = 1, 2, \dots \quad (15.1-25)$$

As an example, the allowed energy levels of electrons in an infinitely deep GaAs well ($m_c = 0.07m_0$) of width $d = 10$ nm are $E_q = 54, 216, 486, \dots$ meV (recall that at $T = 300$ K, $k_B T = 26$ meV). The smaller the width of the well, the larger the separation between adjacent energy levels.

In the quantum-well structure shown in Fig. 15.1-21, electrons (and holes) are confined in the x direction to within a distance d_1 (the well thickness). However, they extend over much larger dimensions ($d_2, d_3 \gg d_1$) in the plane of the confining layer. Thus in the $y-z$ plane, they behave as if they were in bulk semiconductor. The

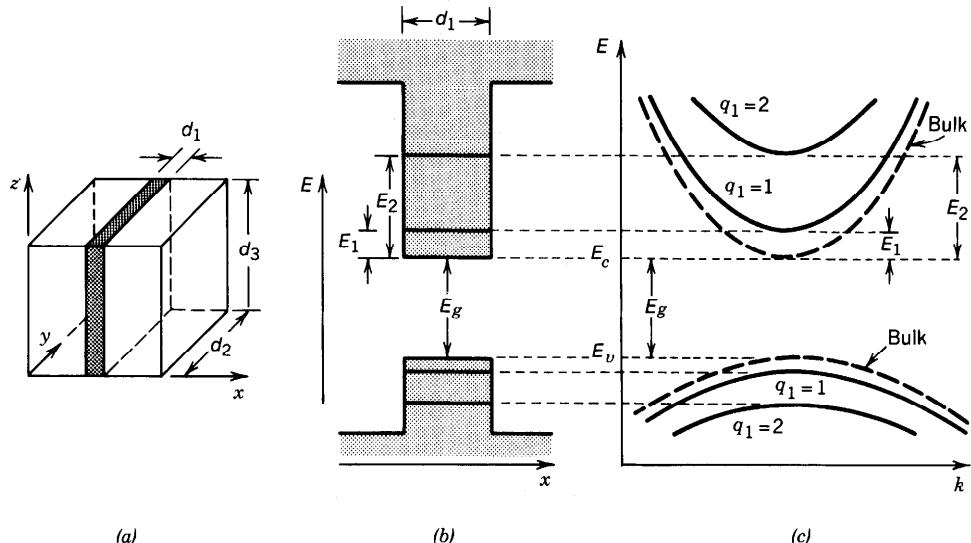


Figure 15.1-21 (a) Geometry of the quantum-well structure. (b) Energy-level diagram for electrons and holes in a quantum well. (c) Cross section of the E - k relation in the direction of k_2 or k_3 . The energy subbands are labeled by their quantum number $q_1 = 1, 2, \dots$. The E - k relation for bulk semiconductor is indicated by the dashed curves.

energy-momentum relation is

$$E = E_c + \frac{\hbar^2 k_1^2}{2m_c} + \frac{\hbar^2 k_2^2}{2m_c} + \frac{\hbar^2 k_3^2}{2m_c},$$

where $k_1 = q_1\pi/d_1$, $k_2 = q_2\pi/d_2$, $k_3 = q_3\pi/d_3$, and $q_1, q_2, q_3 = 1, 2, \dots$. Since $d_1 \ll d_2, d_3$, k_1 takes on well-separated discrete values, whereas k_2 and k_3 have finely spaced discrete values which may be approximated as a continuum. It follows that the energy-momentum relation for electrons in the conduction band of a quantum well is given by

$$E = E_c + E_{q1} + \frac{\hbar^2 k^2}{2m_c}, \quad q_1 = 1, 2, 3, \dots, \quad (15.1-26)$$

where k is the magnitude of a two-dimensional $\mathbf{k} = (k_2, k_3)$ vector in the y - z plane. Each quantum number q_1 corresponds to a subband whose lowest energy is $E_c + E_{q1}$. Similar relations apply for the valence band.

The energy-momentum relation for a bulk semiconductor is given by (15.1-1), where k is the magnitude of a three-dimensional wavevector $\mathbf{k} = (k_1, k_2, k_3)$. The sole distinction is that for the quantum well, k_1 takes on well-separated discrete values. As a result, the density of states associated with a quantum-well structure differs from that associated with bulk material, where the density of states is determined from the magnitude of the three-dimensional wavevector with components $k_1 = q_1\pi/d$, $k_2 = q_2\pi/d$, and $k_3 = q_3\pi/d$ for $d_1 = d_2 = d_3 = d$. The result is [see (15.1-3)] $\varrho(k) = k^2/\pi^2$ per unit volume, which yields the density of conduction-band states [see (15.1-4)]

and Fig. 15.1-7]

$$\varrho_c(E) = \frac{\sqrt{2} m_c^{3/2}}{\pi^2 \hbar^3} (E - E_c)^{1/2}, \quad E > 0. \quad (15.1-27)$$

In a quantum-well structure the density of states is obtained from the magnitude of the two-dimensional wavevector (k_2, k_3). For each quantum number q_1 the density of states is therefore $\varrho(k) = k/\pi$ states per unit area in the $y-z$ plane, and therefore $k/\pi d_1$ per unit volume. The densities $\varrho_c(E)$ and $\varrho(k)$ are related by $\varrho_c(E) dE = \varrho(k) dk = (k/\pi d_1) dk$. Finally, using the $E-k$ relation (15.1-26) we obtain $dE/dk = \hbar^2 k/m_c$, from which

$$\varrho_c(E) = \begin{cases} \frac{m_c}{\pi \hbar^2 d_1}, & E > E_c + E_{q1} \\ 0, & E < E_c + E_{q1}, \end{cases}$$

$$q_1 = 1, 2, \dots \quad (15.1-28)$$

Thus for each quantum number q_1 , the density of states per unit volume is constant when $E > E_c + E_{q1}$. The overall density of states is the sum of the densities for all values of q_1 , so that it exhibits the staircase distribution shown in Fig. 15.1-22. Each step of the staircase corresponds to a different quantum number q_1 and may be regarded as a subband within the conduction band (Fig. 15.1-21). The bottoms of these subbands move progressively higher for higher quantum numbers. It can be shown by substituting $E = E_c + E_{q1}$ in (15.1-27), and by using (15.1-25), that at $E = E_c + E_{q1}$ the quantum-well density of states is the same as that for the bulk. The density of states in the valence band has a similar staircase distribution.

In contrast with bulk semiconductor, the quantum-well structure exhibits a substantial density of states at its lowest allowed conduction-band energy level and at its highest allowed valence-band energy level. This property has a dramatic effect on the optical properties of the material, as discussed in Sec. 16.3G.

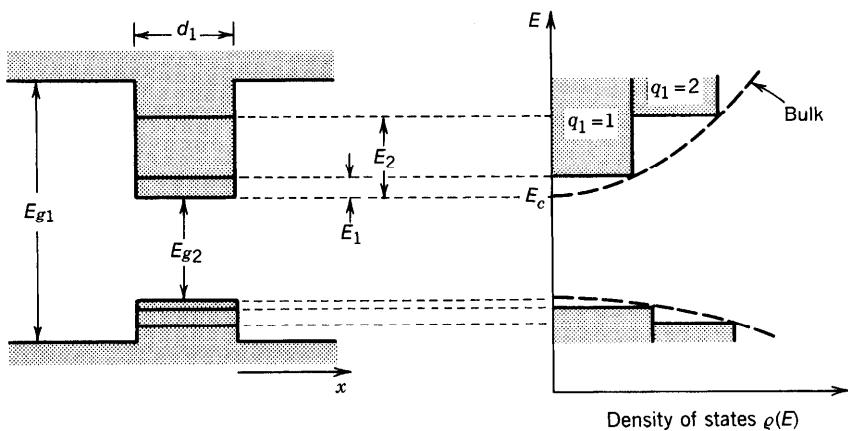


Figure 15.1-22 Density of states for a quantum-well structure (solid) and for a bulk semiconductor (dashed).

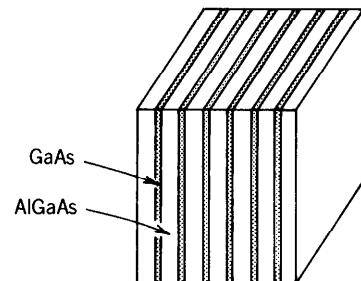


Figure 15.1-23 A multiquantum-well structure fabricated from alternating layers of AlGaAs and GaAs.

Multiquantum Wells and Superlattices

Multiple-layered structures of different semiconductor materials that alternate with each other are called **multiquantum-well** (MQW) structures (see Fig. 15.1-23). They can be fabricated such that the energy bandgap varies with position in any number of ways (see, e.g., Fig. 12.1-8). If the energy barriers between the adjacent wells are sufficiently thin so that electrons can readily tunnel through (quantum mechanically penetrate) the barriers between them, the discrete energy levels broaden into miniature bands in which case the multiquantum-well structure is also referred to as a **superlattice structure**. Multiquantum-well structures are used in lasers and photodetectors, and as nonlinear optical elements. A typical MQW structure might consist of 100 layers, each of which has thickness $\approx 10 \text{ nm}$ and contains some 40 atomic planes, so that the total thickness of the structure is $\approx 1 \mu\text{m}$. Such a structure would take about 1 hour to grow in an MBE machine.

Quantum Wires and Quantum Dots

A semiconductor material that takes the form of a thin wire of rectangular cross section, surrounded by a material of wider bandgap, is called a **quantum-wire** structure (Fig. 15.1-24). The wire acts as a potential well that narrowly confines electrons (and holes) in two directions (x, y). Assuming that the cross-sectional area is $d_1 d_2$, the energy-momentum relation in the conduction band is

$$E = E_c + E_{q1} + E_{q2} + \frac{\hbar^2 k^2}{2m_c}, \quad (15.1-29)$$

where

$$E_{q1} = \frac{\hbar^2 (q_1 \pi / d_1)^2}{2m_c}, \quad E_{q2} = \frac{\hbar^2 (q_2 \pi / d_2)^2}{2m_c}, \quad q_1, q_2 = 1, 2, \dots \quad (15.1-30)$$

and k is the wavevector component in the z direction (along the axis of the wire).

Each pair of quantum numbers (q_1, q_2) is associated with an energy subband with a density of states $\varrho(k) = 1/\pi$ per unit length of the wire and therefore $1/\pi d_1 d_2$ per unit volume. The corresponding density of states (per unit volume), as a function of energy, is

$$\varrho_c(E) = \begin{cases} \frac{(1/d_1 d_2)(m_c^{1/2}/\sqrt{2}\pi\hbar)}{(E - E_c - E_{q1} - E_{q2})^{1/2}}, & E > E_c + E_{q1} + E_{q2} \\ 0, & \text{otherwise,} \end{cases} \quad q_1, q_2 = 1, 2, \dots \quad (15.1-31)$$

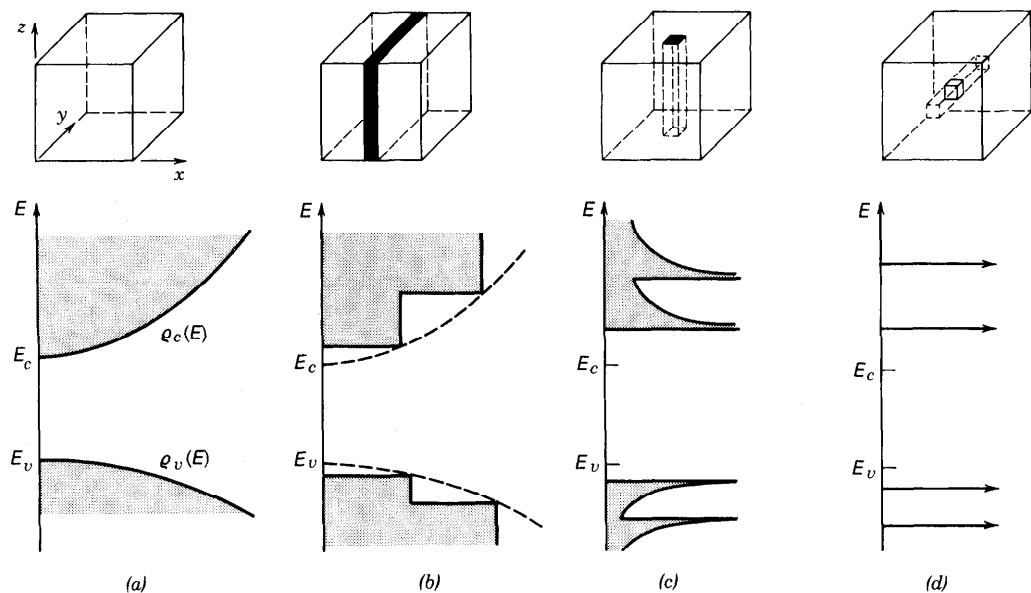


Figure 15.1-24 The density of states in different confinement configurations: (a) bulk; (b) quantum well; (c) quantum wire; (d) quantum dot. The conduction and valence bands split into overlapping subbands that become successively narrower as the electron motion is restricted in more dimensions.

These are decreasing functions of energy, as illustrated in Fig. 15.1-24(c). The energy subbands in a quantum wire are narrower than those in a quantum well.

In a **quantum-dot** structure, the electrons are narrowly confined in all three directions within a box of volume $d_1 d_2 d_3$. The energy is therefore quantized to

$$E = E_c + E_{q1} + E_{q2} + E_{q3},$$

where

$$E_{q1} = \frac{\hbar^2 (q_1 \pi / d_1)^2}{2m_c}, \quad E_{q2} = \frac{\hbar^2 (q_2 \pi / d_2)^2}{2m_c}, \quad E_{q3} = \frac{\hbar^2 (q_3 \pi / d_3)^2}{2m_c},$$

$$q_1, q_2, q_3 = 1, 2, \dots \quad (15.1-32)$$

The allowed energy levels are discrete and well separated so that the density of states is represented by a sequence of impulse functions (delta functions) at the allowed energies, as illustrated in Fig. 15.1-24(d). Quantum dots are often called artificial atoms. Even though they consist of perhaps tens of thousands of strongly interacting natural atoms, the discrete energy levels of the quantum dot can, in principle, be chosen at will by selecting a proper design.

15.2 INTERACTIONS OF PHOTONS WITH ELECTRONS AND HOLES

We now consider the basic optical properties of semiconductors, with an emphasis on the processes of absorption and emission that are important in the operation of photon sources and detectors.

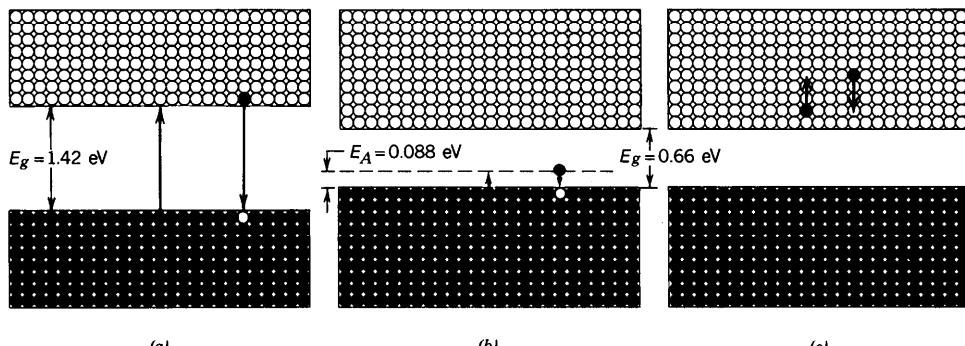


Figure 15.2-1 Examples of absorption and emission of photons in a semiconductor. (a) Band-to-band transitions in GaAs can result in the absorption or emission of photons of wavelength $\lambda_g = hc_o/E_g = 0.87 \mu\text{m}$. (b) The absorption of a photon of wavelength $\lambda_A = hc_o/E_A = 14 \mu\text{m}$ results in a valence-band to acceptor-level transition in Hg-doped Ge (Ge:Hg). (c) A free-carrier transition within the conduction band.

Several mechanisms can lead to the absorption and emission of photons in a semiconductor. The most important of these are:

- **Band-to-Band (Interband) Transitions.** An absorbed photon can result in an electron in the valence band making an upward transition to the conduction band, thereby creating an electron–hole pair [Fig. 15.2-1(a)]. Electron–hole recombination can result in the emission of a photon. Band-to-band transitions may be assisted by one or more phonons. A phonon is a quantum of the lattice vibrations that results from the thermal vibrations of the atoms in the material.
- **Impurity-to-Band Transitions.** An absorbed photon can result in a transition between a donor (or acceptor) level and a band in a doped semiconductor. In a *p*-type material, for example, a low-energy photon can lift an electron from the valence band to the acceptor level, where it becomes trapped by an acceptor atom [Fig. 15.2-1(b)]. A hole is created in the valence band and the acceptor atom is ionized. Or a hole may be trapped by an ionized acceptor atom; the result is that the electron decays from its acceptor level to recombine with the hole. The energy may be released radiatively (in the form of an emitted photon) or nonradiatively (in the form of phonons). The transition may also be assisted by traps in defect states, as illustrated in Fig. 15.1-14.
- **Free-Carrier (Intraband) Transitions.** An absorbed photon can impart its energy to an electron in a given band, causing it to move higher within that band. An electron in the conduction band, for example, can absorb a photon and move to a higher energy level within the conduction band [Fig. 15.2-1(c)]. This is followed by thermalization, a process whereby the electron relaxes down to the bottom of the conduction band while releasing its energy in the form of lattice vibrations.
- **Phonon Transitions.** Long-wavelength photons can release their energy by directly exciting lattice vibrations, i.e., by creating phonons.
- **Excitonic Transitions.** The absorption of a photon can result in the formation of an electron and a hole at some distance from each other but which are nevertheless bound together by their mutual Coulomb interaction. This entity, which is much like a hydrogen atom but with a hole rather than a proton, is called an **exciton**. A photon may be emitted as a result of the electron and hole recombining, thereby annihilating the exciton.

These transitions all contribute to the overall absorption coefficient, which is shown in Fig. 15.2-2 for Si and GaAs, and at greater magnification in Fig. 15.2-3 for a number

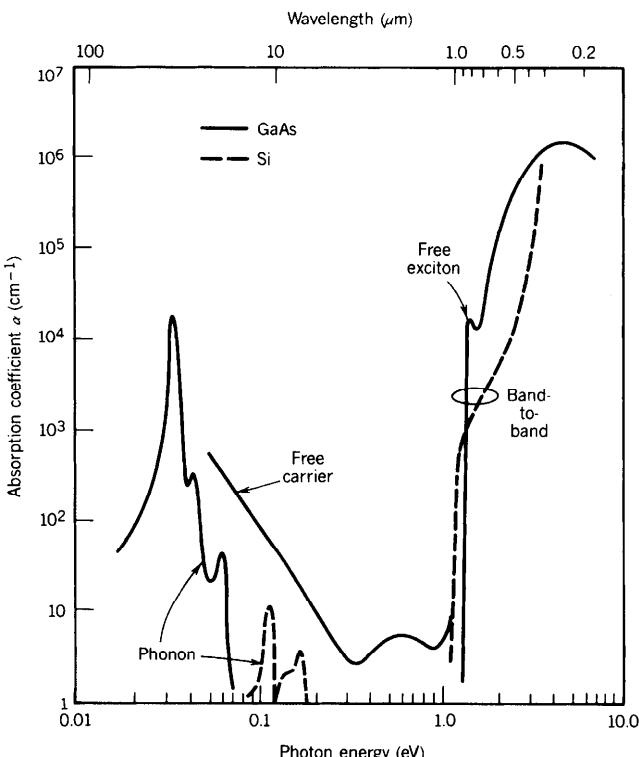


Figure 15.2-2 Observed optical absorption coefficient α versus photon energy for Si and GaAs in thermal equilibrium at $T = 300$ K. The bandgap energy E_g is 1.11 eV for Si and 1.42 eV for GaAs. Si is relatively transparent in the band $\lambda_o \approx 1.1$ to 12 μm , whereas intrinsic GaAs is relatively transparent in the band $\lambda_o \approx 0.87$ to 12 μm (see Fig. 5.5-1).

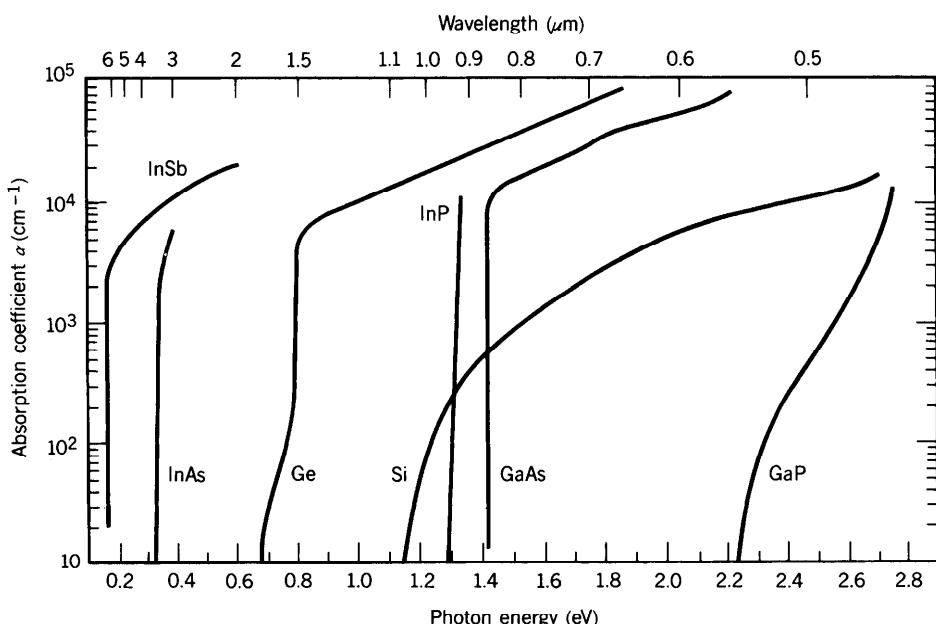


Figure 15.2-3 Absorption coefficient versus photon energy for Ge, Si, GaAs, and selected other III-V binary semiconductors at $T = 300$ K, on an expanded scale (Adapted from G. E. Stillman, V. M. Robbins, and N. Tabatabaie, III-V Compound Semiconductor Devices: Optical Detectors, *IEEE Transactions on Electron Devices*, vol. ED-31, pp. 1643–1655, © 1984 IEEE.)

of semiconductor materials. For photon energies greater than the bandgap energy E_g , the absorption is dominated by band-to-band transitions which form the basis of most photonic devices. The spectral region where the material changes from being relatively transparent ($h\nu < E_g$) to strongly absorbing ($h\nu > E_g$) is known as the **absorption edge**. Direct-gap semiconductors have a more abrupt absorption edge than indirect-gap materials, as is apparent from Figs. 15.2-2 and 15.2-3.

A. Band-to-Band Absorption and Emission

We now proceed to develop a simple theory of direct band-to-band photon absorption and emission, ignoring the other types of transitions.

Bandgap Wavelength

Direct band-to-band absorption and emission can take place only at frequencies for which the photon energy $h\nu > E_g$. The minimum frequency ν necessary for this to occur is $\nu_g = E_g/h$, so that the corresponding maximum wavelength is $\lambda_g = c_o/\nu_g = hc_o/E_g$. If the bandgap energy is given in eV (rather than joules), the bandgap wavelength $\lambda_g = hc_o/eE_g$ in μm is given by

$$\lambda_g = \frac{1.24}{E_g} \quad (15.2-1)$$

Bandgap Wavelength
 λ_g (μm) and E_g (eV)

The quantity λ_g is called the **bandgap wavelength** (or the **cutoff wavelength**); it is provided in Table 15.1-3 and in Figs. 15.1-5 and 15.1-6 for a number of semiconductor materials. The bandgap wavelength λ_g can be adjusted over a substantial range (from the infrared to the visible) by using III-V ternary and quaternary semiconductors of different composition, as is evident in Fig. 15.2-4.

Absorption and Emission

Electron excitation from the valence to the conduction band may be induced by the absorption of a photon of appropriate energy ($h\nu > E_g$). An electron-hole pair is generated [Fig. 15.2-5(a)]. This adds to the concentration of mobile charge carriers and

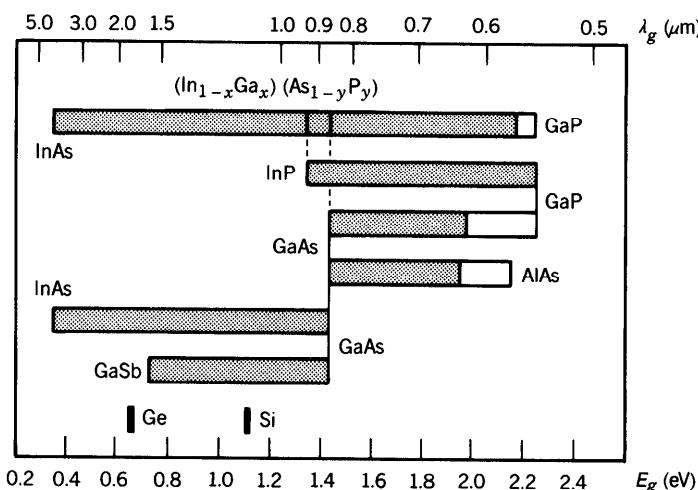


Figure 15.2-4 Bandgap energy E_g and corresponding bandgap wavelength λ_g for selected elemental and III-V binary, ternary, and quaternary semiconductor materials. The shaded regions represent compositions for which the materials are direct-gap semiconductors.

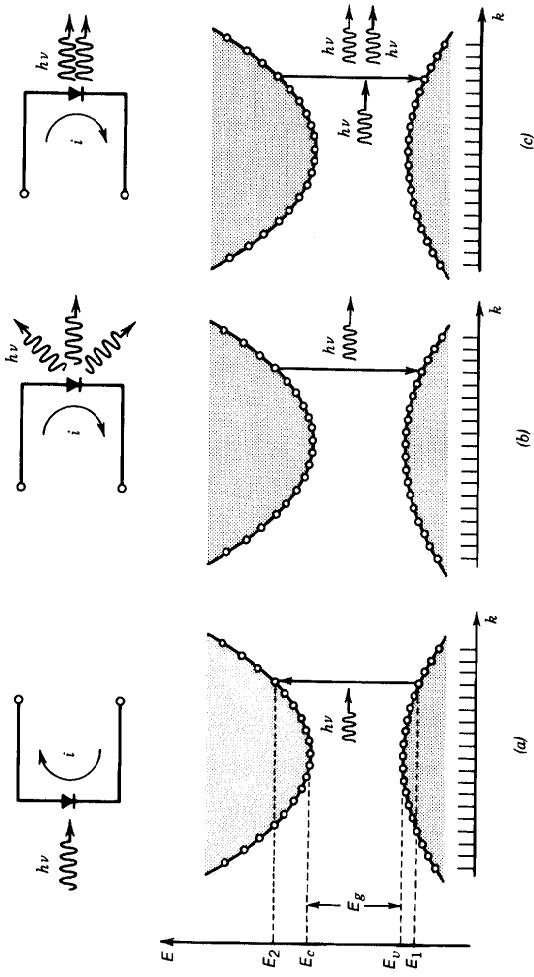


Figure 15.2-5 (a) The absorption of a photon results in the generation of an electron–hole pair. This process is used in the photodetection of light. (b) The recombination of an electron–hole pair results in the spontaneous emission of a photon. Light-emitting diodes (LEDs) operate on this basis. (c) Electron–hole recombination can be stimulated by a photon. The result is the induced emission of an identical photon. This is the underlying process responsible for the operation of semiconductor injection lasers.

increases the conductivity of the material. The material behaves as a photoconductor with a conductivity proportional to the photon flux. This effect is used to detect light, as discussed in Chap. 17.

Electron deexcitation from the conduction to the valence band (electron–hole recombination) may result in the spontaneous emission of a photon of energy $h\nu > E_g$ [Fig. 15.2-5(b)], or in the stimulated emission of a photon (see Sec. 12.2), provided that a photon of energy $h\nu > E_g$ is present [Fig. 15.2-5(c)]. Spontaneous emission is the underlying phenomenon on which the light-emitting diode is based, as will be seen in Sec. 16.1. Stimulated emission is responsible for the operation of semiconductor amplifiers and lasers, as will be seen in Secs. 16.2 and 16.3.

Conditions for Absorption and Emission

- *Conservation of Energy.* The absorption or emission of a photon of energy $h\nu$ requires that the energies of the two states involved in the interaction (E_1 and E_2 in the valence band and conduction band, respectively) be separated by $h\nu$. Thus, for photon emission to occur by electron–hole recombination, for example, an electron occupying an energy level E_2 must interact with a hole occupying an energy level E_1 , such that energy is conserved, i.e.,

$$E_2 - E_1 = h\nu. \quad (15.2-2)$$

- *Conservation of Momentum.* Momentum must also be conserved in the process of photon emission/absorption, so that $p_2 - p_1 = h\nu/c = h/\lambda$, or $k_2 - k_1 = 2\pi/\lambda$. The photon-momentum magnitude h/λ is, however, very small in comparison with the range of values that electrons and holes can assume. The semiconductor E - k diagram extends to values of k of the order $2\pi/a$, where the lattice constant a is much smaller than the wavelength λ , so that $2\pi/\lambda \ll 2\pi/a$. The momenta of the electron and the hole involved in interaction with the photon are therefore roughly equal. This condition, $k_2 \approx k_1$, is called the ***k-selection rule***. Transitions that obey this rule are represented in the E - k diagram (Fig. 15.2-5) by vertical lines, indicating that the change in k is negligible on the scale of the diagram.

- *Energies and Momenta of the Electron and Hole with Which a Photon Interacts.* As is apparent from Fig. 15.2-5, conservation of energy and momentum require that a photon of frequency ν interact with electrons and holes of specific energies and momentum determined by the semiconductor E - k relation. Using (15.1-1) and (15.1-2) to approximate this relation for a direct-gap semiconductor by two parabolas, and writing $E_c - E_v = E_g$, (15.2-2) may be written in the form

$$E_2 - E_1 = \frac{\hbar^2 k^2}{2m_v} + E_g + \frac{\hbar^2 k^2}{2m_c} = h\nu, \quad (15.2-3)$$

from which

$$k^2 = \frac{2m_r}{\hbar^2} (h\nu - E_g), \quad (15.2-4)$$

where

$$\frac{1}{m_r} = \frac{1}{m_v} + \frac{1}{m_c}. \quad (15.2-5)$$

Substituting (15.2-4) into (15.1-1), the energy levels E_1 and E_2 with which the photon interacts are therefore

$$E_2 = E_c + \frac{m_r}{m_c} (h\nu - E_g) \quad (15.2-6)$$

$$E_1 = E_v - \frac{m_r}{m_v} (h\nu - E_g) = E_2 - h\nu. \quad (15.2-7)$$

Energies of Electron
and Hole Interacting
with a Photon $h\nu$

In the special case where $m_c = m_v$, we obtain $E_2 = E_c + \frac{1}{2}(h\nu - E_g)$, as required by symmetry.

- *Optical Joint Density of States.* We now determine the density of states $\varrho(\nu)$ with which a photon of energy $h\nu$ interacts under conditions of energy and momentum conservation in a direct-gap semiconductor. This quantity incorporates the density of states in both the conduction and valence bands and is called the optical joint density of states. The one-to-one correspondence between E_2 and ν , embodied in (15.2-6), permits us to readily relate $\varrho(\nu)$ to the density of states $\varrho_c(E_2)$ in the conduction band by use of the incremental relation $\varrho_c(E_2) dE_2 = \varrho(\nu) d\nu$, from which $\varrho(\nu) = (dE_2/d\nu)\varrho_c(E_2)$, so that

$$\varrho(\nu) = \frac{hm_r}{m_c} \varrho_c(E_2). \quad (15.2-8)$$

Using (15.1-4) and (15.2-6), we finally obtain the number of states per unit volume per unit frequency:

$$\varrho(\nu) = \frac{(2m_r)^{3/2}}{\pi\hbar^2} (\hbar\nu - E_g)^{1/2}, \quad h\nu \geq E_g, \quad (15.2-9)$$

Optical Joint
Density of
States

which is illustrated in Fig. 15.2-6. The one-to-one correspondence between E_1 and ν in (15.2-7), together with $\varrho_v(E_1)$ from (15.1-5), results in an expression for $\varrho(\nu)$ identical to (15.2-9).

- *Photon Emission Is Unlikely in an Indirect-Gap Semiconductor.* Radiative electron-hole recombination is unlikely in an indirect-gap semiconductor. This is because transitions from near the bottom of the conduction band to near the top of the valence band (where electrons and holes, respectively, are most likely to reside) requires an exchange of momentum that cannot be accommodated by the emitted photon. Momentum may be conserved, however, by the participation of

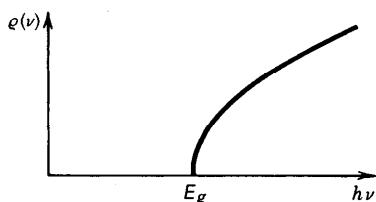


Figure 15.2-6 The density of states with which a photon of energy $h\nu$ interacts increases with $h\nu - E_g$ in accordance with a square-root law.

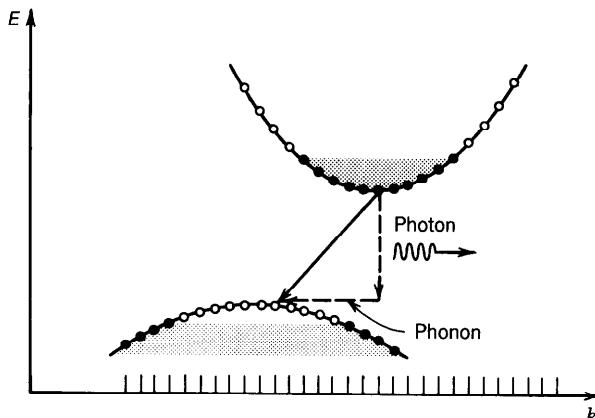


Figure 15.2-7 Photon emission in an indirect-gap semiconductor. The recombination of an electron near the bottom of the conduction band with a hole near the top of the valence band requires the exchange of energy *and* momentum. The energy may be carried off by a photon, but one or more phonons are required to conserve momentum. This type of multiparticle interaction is unlikely.

phonons in the interaction. Phonons can carry relatively large momenta but typically have small energies (≈ 0.01 to 0.1 eV; see Fig. 15.2-2), so their transitions appear horizontal on the E - k diagram (see Fig. 15.2-7). The net result is that momentum is conserved, but the k -selection rule is violated. Because phonon-assisted emission involves the participation of three bodies (electron, photon, and phonon), the probability of their occurrence is quite low. Thus Si, which is an indirect-gap semiconductor, has a substantially lower radiative recombination rate than does GaAs, which is a direct-gap semiconductor (see Table 15.1-5). Si is therefore not an efficient light emitter, whereas GaAs is.

- **Photon Absorption is Not Unlikely in an Indirect-Gap Semiconductor.** Although photon absorption also requires energy and momentum conservation in an indirect-gap semiconductor, this is readily achieved by means of a two-step process (Fig. 15.2-8). The electron is first excited to a high energy level within the

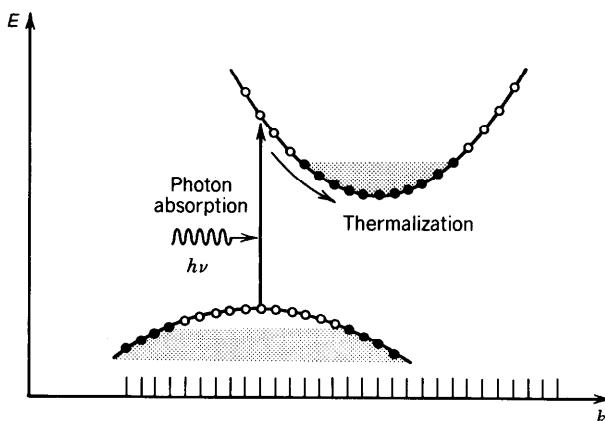


Figure 15.2-8 Photon absorption in an indirect-gap semiconductor. The photon generates an excited electron and a hole by a vertical transition; the carriers then undergo fast transitions to the bottom of the conduction band and top of the valence band, respectively, releasing their energy in the form of phonons. Since the process is sequential it is not unlikely.

conduction band by a vertical transition. It then quickly relaxes to the bottom of the conduction band by a process called thermalization in which its momentum is transferred to phonons. The generated hole behaves similarly. Since the process occurs sequentially, it does not require the simultaneous presence of three bodies and is thus not unlikely. Si is therefore an efficient photon detector, as is GaAs.

B. Rates of Absorption and Emission

We now proceed to determine the probability densities of a photon of energy $h\nu$ being emitted or absorbed by a semiconductor material in a direct band-to-band transition. Conservation of energy and momentum, in the form of (15.2-6), (15.2-7), and (15.2-4), determine the energies E_1 and E_2 , and the momentum $\hbar k$, of the electrons and holes with which the photon may interact.

Three factors determine these probability densities: the occupancy probabilities, the transition probabilities, and the density of states. We consider these in turn.

Occupancy Probabilities

The occupancy conditions for photon emission and absorption by means of transitions between the discrete energy levels E_1 and E_2 are the following:

Emission condition: A conduction-band state of energy E_2 is filled (with an electron) and a valence-band state of energy E_1 is empty (i.e., filled with a hole).

Absorption condition: A conduction-band state of energy E_2 is empty and a valence-band state of energy E_1 is filled.

The probabilities that these occupancy conditions are satisfied for various values of E_1 and E_2 are determined from the appropriate Fermi functions $f_c(E)$ and $f_v(E)$ associated with the conduction and valence bands of a semiconductor in quasi-equilibrium. Thus the probability $f_e(\nu)$ that the emission condition is satisfied for a photon of energy $h\nu$ is the product of the probabilities that the upper state is filled and that the lower state is empty (these are independent events), i.e.,

$$f_e(\nu) = f_c(E_2)[1 - f_v(E_1)]. \quad (15.2-10)$$

E_1 and E_2 are related to ν by (15.2-6) and (15.2-7). Similarly, the probability $f_a(\nu)$ that the absorption condition is satisfied is

$$f_a(\nu) = [1 - f_c(E_2)]f_v(E_1). \quad (15.2-11)$$

EXERCISE 15.2-1

Requirement for the Photon Emission Rate to Exceed the Absorption Rate

- (a) For a semiconductor in thermal equilibrium, show that $f_e(\nu)$ is always smaller than $f_a(\nu)$ so that the rate of photon emission cannot exceed the rate of photon absorption.
- (b) For a semiconductor in quasi-equilibrium ($E_{fc} \neq E_{fv}$), with radiative transitions occurring between a conduction-band state of energy E_2 and a valence-band state of energy

E_1 with the same k , show that emission is more likely than absorption if the separation between the quasi-Fermi levels is larger than the photon energy, i.e., if

$$E_{fc} - E_{fv} > h\nu. \quad (15.2-12)$$

Condition for
Net Emission

What does this condition imply about the locations of E_{fc} relative to E_c and E_{fv} relative to E_v ?

Transition Probabilities

Satisfying the emission/absorption occupancy condition does not assure that the emission/absorption actually takes place. These processes are governed by the probabilistic laws of interaction between photons and atomic systems examined at length in Secs. 12.2A to C (see also Exercise 12.2-1). As they relate to semiconductors, these laws are generally expressed in terms of emission into (or absorption from) a narrow band of frequencies between ν and $\nu + d\nu$:

A radiative transition between two discrete energy levels E_1 and E_2 is characterized by a transition cross section $\sigma(\nu) = (\lambda^2/8\pi t_{sp})g(\nu)$, where ν is the frequency, t_{sp} is the spontaneous lifetime, and $g(\nu)$ is the lineshape function (which has linewidth $\Delta\nu$ centered about the transition frequency $\nu_0 = (E_2 - E_1)/h$ and has unity area). In semiconductors the radiative electron-hole recombination lifetime τ_r , which was discussed in Sec. 15.1D, plays the role of t_{sp} so that

$$\sigma(\nu) = \frac{\lambda^2}{8\pi\tau_r} g(\nu). \quad (15.2-13)$$

- If the occupancy condition for emission is satisfied, the probability density (per unit time) for the spontaneous emission of a photon into any of the available radiation modes in the narrow frequency band between ν and $\nu + d\nu$ is

$$P_{sp}(\nu) d\nu = \frac{1}{\tau_r} g(\nu) d\nu. \quad (15.2-14)$$

- If the occupancy condition for emission is satisfied *and* a mean photon-flux spectral density ϕ_ν (photons per unit time per unit area per unit frequency) at frequency ν is present, the probability density (per unit time) for the stimulated emission of one photon into the narrow frequency band between ν and $\nu + d\nu$ is

$$W_i(\nu) d\nu = \phi_\nu \sigma(\nu) d\nu = \phi_\nu \frac{\lambda^2}{8\pi\tau_r} g(\nu) d\nu. \quad (15.2-15)$$

- If the occupancy condition for absorption is satisfied *and* a mean photon-flux spectral density ϕ_ν at frequency ν is present, the probability density for the absorption of one photon from the narrow frequency band between ν and $\nu + d\nu$ is also given by (15.2-15).

Since each transition has a different central frequency ν_0 , and since we are considering a collection of such transitions, we explicitly label the central frequency of the transition by writing $g(\nu)$ as $g_{\nu_0}(\nu)$. In semiconductors the homogeneously broadened lineshape function $g_{\nu_0}(\nu)$ associated with a pair of energy levels generally has its origin in electron–phonon collision broadening. It therefore typically exhibits a Lorentzian lineshape [see (12.2-27) and (12.2-30)] with width $\Delta\nu \approx 1/\pi T_2$, where the electron–phonon collision time T_2 is of the order of picoseconds. If $T_2 = 1$ ps, for example, then $\Delta\nu = 318$ GHz, corresponding to an energy width $h\Delta\nu \approx 1.3$ meV. The radiative lifetime broadening of the levels is negligible in comparison with collisional broadening.

Overall Emission and Absorption Transition Rates

For a pair of energy levels separated by $E_2 - E_1 = h\nu_0$, the rates of spontaneous emission, stimulated emission, and absorption of photons of energy $h\nu$ (photons per second per hertz per cm^3 of the semiconductor) at the frequency ν are obtained as follows. The appropriate transition probability density $P_{\text{sp}}(\nu)$ or $W_i(\nu)$ [as given in (15.2-14) or (15.2-15)] is multiplied by the appropriate occupation probability $f_e(\nu_0)$ or $f_a(\nu_0)$ [as given in (15.2-10) or (15.2-11)], and by the density of states that can interact with the photon $\varrho(\nu_0)$ [as given in (15.2-9)]. The overall transition rate for all allowed frequencies ν_0 is then calculated by integrating over ν_0 .

The rate of spontaneous emission at frequency ν , for example, is therefore given by

$$r_{\text{sp}}(\nu) = \int [(1/\tau_r) g_{\nu_0}(\nu)] f_e(\nu_0) \varrho(\nu_0) d\nu_0.$$

When the collision-broadened width $\Delta\nu$ is substantially less than the width of the function $f_e(\nu_0)\varrho(\nu_0)$, which is the usual situation, $g_{\nu_0}(\nu)$ may be approximated by $\delta(\nu - \nu_0)$, whereupon the transition rate simplifies to $r_{\text{sp}}(\nu) = (1/\tau_r)\varrho(\nu)f_e(\nu)$. The rates of stimulated emission and absorption are obtained in similar fashion, so that the following formulas emerge:

$$r_{\text{sp}}(\nu) = \frac{1}{\tau_r} \varrho(\nu) f_e(\nu) \quad (15.2-16)$$

$$r_{\text{st}}(\nu) = \phi_\nu \frac{\lambda^2}{8\pi\tau_r} \varrho(\nu) f_e(\nu) \quad (15.2-17)$$

$$r_{\text{ab}}(\nu) = \phi_\nu \frac{\lambda^2}{8\pi\tau_r} \varrho(\nu) f_a(\nu). \quad (15.2-18)$$

Rates of
Spontaneous Emission
Stimulated Emission
and Absorption

These equations, together with (15.2-9) to (15.2-11), permit the rates of spontaneous emission, stimulated emission, and absorption arising from direct band-to-band transitions (photons per second per hertz per cm^3) to be calculated in the presence of a mean photon-flux spectral density ϕ_ν (photons per second per cm^2 per hertz). The products $\varrho(\nu)f_e(\nu)$ and $\varrho(\nu)f_a(\nu)$ are similar to the products of the lineshape function and the atomic number densities in the upper and lower levels, $g(\nu)N_2$ and $g(\nu)N_1$, respectively, used in Chaps. 12 to 14 to study emission and absorption in atomic systems.

The determination of the occupancy probabilities $f_e(\nu)$ and $f_a(\nu)$ requires knowledge of the quasi-Fermi levels E_{fc} and E_{fv} . It is through the control of these two parameters (by the application of an external bias to a $p-n$ junction, for example) that the emission and absorption rates are modified to produce semiconductor photonic devices that carry out different functions. Equation (15.2-16) is the basic result that describes the operation of the light-emitting diode (LED), a semiconductor photon source based on spontaneous emission (see Sec. 16.1). Equation (15.2-17) is applicable to semiconductor optical amplifiers and injection lasers, which operate on the basis of stimulated emission (see Secs. 16.2 and 16.3). Equation (15.2-18) is appropriate for semiconductor photon detectors which function by means of photon absorption (see Chap. 17).

Spontaneous Emission Spectral Density in Thermal Equilibrium

A semiconductor in thermal equilibrium has only a single Fermi function so that (15.2-10) becomes $f_e(\nu) = f(E_2)[1 - f(E_1)]$. If the Fermi level lies within the bandgap, away from the band edges by at least several times $k_B T$, use may be made of the exponential approximations to the Fermi functions, $f(E_2) \approx \exp[-(E_2 - E_f)/k_B T]$ and $1 - f(E_1) \approx \exp[-(E_f - E_1)/k_B T]$, whereupon $f_e(\nu) \approx \exp[-(E_2 - E_1)/k_B T]$, i.e.,

$$f_e(\nu) \approx \exp\left(-\frac{h\nu}{k_B T}\right). \quad (15.2-19)$$

Substituting (15.2-9) for $\varrho(\nu)$ and (15.2-19) for $f_e(\nu)$ into (15.2-16) therefore provides

$$r_{sp}(\nu) \approx D_0(h\nu - E_g)^{1/2} \exp\left(-\frac{h\nu - E_g}{k_B T}\right), \quad h\nu \geq E_g, \quad (15.2-20)$$

where

$$D_0 = \frac{(2m_r)^{3/2}}{\pi\hbar^2\tau_r} \exp\left(-\frac{E_g}{k_B T}\right) \quad (15.2-21)$$

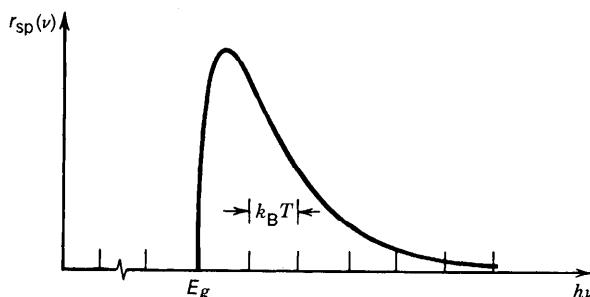


Figure 15.2-9 Spectral density of the direct band-to-band spontaneous emission rate $r_{sp}(\nu)$ (photons per second per hertz per cm^3) from a semiconductor in thermal equilibrium as a function of $h\nu$. The spectrum has a low-frequency cutoff at $\nu = E_g/h$ and extends over a width of approximately $2k_B T/h$.

is a parameter that increases with temperature at an exponential rate. The spontaneous emission rate, which is plotted versus $h\nu$ in Fig. 15.2-9, takes the form of two factors: a power-law increasing function of $h\nu - E_g$ arising from the density of states and an exponentially decreasing function of $h\nu - E_g$ arising from the Fermi function.

The spontaneous emission rate can be increased by increasing $f_e(\nu)$. In accordance with (15.2-10), this can be achieved by purposely causing the material to depart from thermal equilibrium in such a way that $f_c(E_2)$ is made large and $f_v(E_1)$ is made small. This assures an abundance of *both* electrons and holes, which is the desired condition for the operation of an LED, as discussed in Sec. 16.1.

Gain Coefficient in Quasi-Equilibrium

The net gain coefficient $\gamma_0(\nu)$ corresponding to the rates of stimulated emission and absorption in (15.2-17) and (15.2-18) is determined by taking a cylinder of unit area and incremental length dz and assuming that a mean photon-flux spectral density is directed along its axis (as shown in Fig. 13.1-1). If $\phi_v(z)$ and $\phi_v(z) + d\phi_v(z)$ are the mean photon-flux spectral densities entering and leaving the cylinder, respectively, $d\phi_v(z)$ must be the mean photon-flux spectral density emitted from within the cylinder. The incremental number of photons, per unit time per unit frequency per unit area, is simply the number of photons gained, per unit time per unit frequency per unit volume [$r_{st}(\nu) - r_{ab}(\nu)$] multiplied by the thickness of the cylinder dz , i.e., $d\phi_v(z) = [r_{st}(\nu) - r_{ab}(\nu)] dz$. Substituting from (15.2-17) and (15.2-18), we obtain

$$\frac{d\phi_v(z)}{dz} = \frac{\lambda^2}{8\pi\tau_r} \varrho(\nu) [f_e(\nu) - f_a(\nu)] \phi_v(z) = \gamma_0(\nu) \phi_v(z). \quad (15.2-22)$$

The net gain coefficient is therefore

$$\gamma_0(\nu) = \frac{\lambda^2}{8\pi\tau_r} \varrho(\nu) f_g(\nu), \quad (15.2-23)$$

Gain
Coefficient

where the Fermi inversion factor is given by

$$f_g(\nu) \equiv f_e(\nu) - f_a(\nu) = f_c(E_2) - f_v(E_1), \quad (15.2-24)$$

as may be seen from (15.2-10) and (15.2-11), with E_1 and E_2 related to ν by (15.2-6) and (15.2-7). Using (15.2-9), the gain coefficient may be cast in the form

$$\gamma_0(\nu) = D_1 (h\nu - E_g)^{1/2} f_g(\nu), \quad h\nu > E_g, \quad (15.2-25a)$$

with

$$D_1 = \frac{\sqrt{2} m_r^{3/2} \lambda^2}{h^2 \tau_r}. \quad (15.2-25b)$$

The sign and spectral form of the Fermi inversion factor $f_g(\nu)$ are governed by the quasi-Fermi levels E_{fc} and E_{fv} , which, in turn, depend on the state of excitation of the carriers in the semiconductor. As shown in Exercise 15.2-1, this factor is positive (corresponding to a population inversion and net gain) only when $E_{fc} - E_{fv} > h\nu$. When the semiconductor is pumped to a sufficiently high level by means of an external energy source, this condition may be satisfied and net gain achieved, as we shall see in

Sec. 16.2. This is the physics underlying the operation of semiconductor optical amplifiers and injection lasers.

Absorption Coefficient in Thermal Equilibrium

A semiconductor in thermal equilibrium has only a single Fermi level $E_f = E_{fc} = E_{fv}$, so that

$$f_c(E) = f_v(E) = f(E) = \frac{1}{\exp[(E - E_f)/k_B T] + 1}. \quad (15.2-26)$$

The factor $f_g(\nu) = f_c(E_2) - f_v(E_1) = f(E_2) - f(E_1) < 0$, and therefore the gain coefficient $\gamma_0(\nu)$ is always negative [since $E_2 > E_1$ and $f(E)$ decreases monotonically with E]. This is true whatever the location of the Fermi level E_f . Thus a semiconductor in thermal equilibrium, whether it be intrinsic or doped, always attenuates light. The attenuation (or absorption) coefficient, $\alpha(\nu) = -\gamma_0(\nu)$, is therefore

$$\alpha(\nu) = D_1(h\nu - E_g)^{1/2}[f(E_1) - f(E_2)], \quad (15.2-27)$$

Absorption
Coefficient

where E_1 and E_2 are given by (15.2-7) and (15.2-6), respectively, and D_1 is given by (15.2-25b).

If E_f lies within the bandgap but away from the band edges by an energy of at least several times $k_B T$, then $f(E_1) \approx 1$ and $f(E_2) \approx 0$ so that $[f(E_1) - f(E_2)] \approx 1$. In that case, the direct band-to-band contribution to the absorption coefficient is

$$\alpha(\nu) \approx \frac{\sqrt{2} c^2 m_r^{3/2}}{\tau_r} \frac{1}{(h\nu)^2} (h\nu - E_g)^{1/2}. \quad (15.2-28)$$

As the temperature increases, $f(E_1) - f(E_2)$ decreases below unity and the absorption coefficient is reduced. Equation (15.2-28) is plotted in Fig. 15.2-10 for GaAs, using the following parameters: $n = 3.6$, $m_c = 0.07m_0$, $m_v = 0.5m_0$, $m_0 = 9.1 \times 10^{-31}$ kg, a

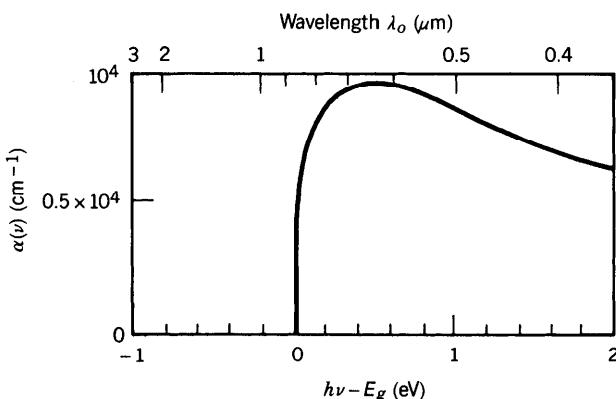


Figure 15.2-10 Calculated absorption coefficient $\alpha(\nu)$ (cm^{-1}) resulting from direct band-to-band transitions as a function of the photon energy $h\nu$ (eV) and wavelength λ_o (μm) for GaAs. This should be compared with the experimental result shown in Fig. 15.2-3, which includes all absorption mechanisms.

doping level such that $\tau_r = 0.4$ ns (this differs from that given in Table 15.1-5 because of the difference in doping level), $E_g = 1.42$ eV, and a temperature such that $[f(E_1) - f(E_2)] \approx 1$.

EXERCISE 15.2-2

Wavelength of Maximum Band-to-Band Absorption. Use (15.2-28) to determine the (free-space) wavelength λ_p at which the absorption coefficient of a semiconductor in thermal equilibrium is maximum. Calculate the value of λ_p for GaAs. Note that this result applies only to absorption by direct band-to-band transitions.

C. Refractive Index

The ability to control the refractive index of a semiconductor is important in the design of many photonic devices, particularly those that make use of optical waveguides, integrated optics, and injection laser diodes. Semiconductor materials are dispersive, so that the refractive index is dependent on the wavelength. Indeed, it is related to the absorption coefficient $\alpha(\nu)$ inasmuch as the real and imaginary parts of the susceptibility must satisfy the Kramers–Kronig relations (see Sec. 5.5B and Sec. B.1 of Appendix B). The refractive index also depends on temperature and on doping level, as is clear from the curves in Fig. 15.2-11 for GaAs.

The refractive indices of selected elemental and binary semiconductors, under specific conditions and near the bandgap wavelength, are provided in Table 15.2-1.

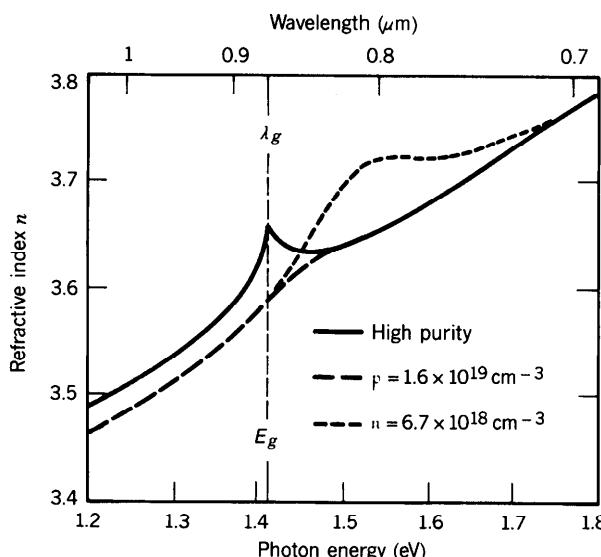


Figure 15.2-11 Refractive index for high-purity, p -type, and n -type GaAs at 300 K, as a function of photon energy (wavelength). The peak in the high-purity curve at the bandgap wavelength is associated with free excitons. (Adapted from H. C. Casey, Jr., and M. B. Panish, *Heterostructure Lasers*, part A, *Fundamental Principles*, Academic Press, New York, 1978.)

TABLE 15.2-1 Refractive Indices of Selected Semiconductor Materials at $T = 300$ K for Photon Energies Near the Bandgap Energy of the Material ($h\nu \approx E_g$)^a

Material	Refractive Index
Elemental semiconductors	
Ge	4.0
Si	3.5
III-V binary semiconductors	
AlP	3.0
AlAs	3.2
AlSb	3.8
GaP	3.3
GaAs	3.6
GaSb	4.0
InP	3.5
InAs	3.8
InSb	4.2

^aThe refractive indices of ternary and quaternary semiconductors can be approximated by linear interpolation between the refractive indices of their components.

READING LIST

Books on Semiconductor Physics and Devices

- B. G. Streetman, *Solid State Electronic Devices*, Prentice-Hall, Englewood Cliffs, NJ, 3rd ed. 1990.
 S. Wang, *Fundamentals of Semiconductor Theory and Device Physics*, Prentice-Hall, Englewood Cliffs, NJ, 1989.
 E. S. Yang, *Microelectronic Devices*, McGraw-Hill, New York, 1988.
 K. Hess, *Advanced Theory of Semiconductor Devices*, Prentice-Hall, Englewood Cliffs, NJ, 1988.
 C. Kittel, *Introduction to Solid State Physics*, Wiley, New York, 6th ed. 1986.
 D. A. Fraser, *The Physics of Semiconductor Devices*, Clarendon Press, Oxford, 4th ed. 1986.
 S. M. Sze, *Semiconductor Devices: Physics and Technology*, Wiley, New York, 1985.
 K. Seeger, *Semiconductor Physics*, Springer-Verlag, Berlin, 2nd ed. 1982.
 S. M. Sze, *Physics of Semiconductor Devices*, Wiley, New York, 2nd ed. 1981.
 O. Madelung, *Introduction to Solid State Theory*, Springer-Verlag, Berlin, 1978.
 R. A. Smith, *Semiconductors*, Cambridge University Press, New York, 2nd ed. 1978.
 N. W. Ashcroft and N. D. Mermin, *Solid State Physics*, Holt, Rinehart and Winston, New York, 1976.
 A. van der Ziel, *Solid State Physical Electronics*, Prentice-Hall, Englewood Cliffs, NJ, 3rd ed. 1976.
 D. H. Navon, *Electronic Materials and Devices*, Houghton Mifflin, Boston, 1975.
 W. A. Harrison, *Solid State Theory*, McGraw-Hill, New York, 1970.
 C. A. Wert and R. M. Thomson, *Physics of Solids*, McGraw-Hill, New York, 1970.
 J. M. Ziman, *Principles of the Theory of Solids*, Wiley, New York, 1968.
 A. S. Grove, *Physics and Technology of Semiconductor Devices*, Wiley, New York, 1967.

Books on Optoelectronics

- J. Wilson and J. F. B. Hawkes, *Optoelectronics*, Prentice-Hall, Englewood Cliffs, NJ, 2nd ed. 1989.

- M. L. Cohen and J. R. Chelikowsky, *Electronic Structure and Optical Properties of Semiconductors*, Springer-Verlag, New York, 2nd ed. 1989.
- J. Gowar, *Optical Communication Systems*, Prentice-Hall, Englewood Cliffs, NJ, 1984.
- H. Kressel, ed., *Semiconductor Devices for Optical Communications*, Springer-Verlag, New York, 2nd ed. 1982.
- T. S. Moss, G. J. Burrell, and B. Ellis, *Semiconductor Opto-electronics*, Wiley, New York, 1973.
- J. I. Pankove, *Optical Processes in Semiconductors*, Prentice-Hall, Englewood Cliffs, NJ, 1971; Dover, New York, 1975.

Books on Heterostructures and Quantum-Well Structures

- C. Weisbuch and B. Vinter, *Quantum Semiconductor Structures*, Academic Press, Orlando, FL, 1991.
- F. Capasso, ed., *Physics of Quantum Electron Devices*, Springer-Verlag, New York, 1990.
- R. Dingle, *Applications of Multiquantum Wells, Selective Doping, and Super-Lattices*, Academic Press, New York, 1987.
- F. Capasso and G. Margaritondo, eds., *Heterojunction Band Discontinuities*, North-Holland, Amsterdam, 1987.
- H. C. Casey, Jr., and M. B. Panish, *Heterostructure Lasers*, part A, *Fundamental Principles*, Academic Press, New York, 1978.
- H. C. Casey, Jr., and M. B. Panish, *Heterostructure Lasers*, part B, *Materials and Operating Characteristics*, Academic Press, New York, 1978.
- H. Kressel and J. K. Butler, *Semiconductor Lasers and Heterojunction LEDs*, Academic Press, New York, 1977.
- A. G. Milnes and D. L. Feucht, *Heterojunctions and Metal-Semiconductor Junctions*, Academic Press, New York, 1972.

Special Journal Issues

- Special issue on quantum-well heterostructures and superlattices, *IEEE Journal of Quantum Electronics*, vol. QE-24, no. 8, 1988.
- Special issue on semiconductor quantum wells and superlattices: physics and applications, *IEEE Journal of Quantum Electronics*, vol. QE-22, no. 9, 1986.

Articles

- E. Corcoran, Diminishing Dimensions, *Scientific American*, vol. 263, no. 5, pp. 122–131, 1990.
- D. A. B. Miller, Optoelectronic Applications of Quantum Wells, *Optics and Photonics News*, vol. 1, no. 2, pp. 7–15, 1990.
- S. Schmitt-Rink, D. S. Chemla, and D. A. B. Miller, Linear and Nonlinear Optical Properties of Semiconductor Quantum Wells, *Advances in Physics*, vol. 38, pp. 89–188, 1989.
- W. D. Goodhue, Using Molecular-Beam Epitaxy to Fabricate Quantum-Well Devices, *Lincoln Laboratory Journal*, vol. 2, no. 2, pp. 183–206, 1989.
- S. R. Forrest, Organic-on-Inorganic Semiconductor Heterojunctions: Building Block for the Next Generation of Optoelectronic Devices?, *IEEE Circuits and Devices Magazine*, vol. 5, no. 3, pp. 33–37, 41, 1989.
- A. M. Glass, Optical Materials, *Science*, vol. 235, pp. 1003–1009, 1987.
- L. Esaki, A Bird's-Eye View on the Evolution of Semiconductor Superlattices and Quantum Wells," *IEEE Journal of Quantum Electronics*, vol. QE-22, pp. 1611–1624, 1986.
- D. S. Chemla, Quantum Wells for Photonics, *Physics Today*, vol. 38, no. 5, pp. 56–64, 1985.

PROBLEMS

- 15.1-1 **Fermi Level of an Intrinsic Semiconductor.** Given the expressions for the thermal equilibrium carrier concentrations in the conduction and valence bands [(15.1-9a) and (15.1-9b)]:
 (a) Determine an expression for the Fermi level E_f of an intrinsic semiconductor and show that it falls exactly in the middle of the bandgap only when the effective mass of the electrons m_e is precisely equal to the effective mass of the holes m_h .
 (b) Determine an expression for the Fermi level of a doped semiconductor as a function of the doping level and the Fermi level determined in part (a).
- 15.1-2 **Electron–Hole Recombination Under Strong Injection.** Consider electron–hole recombination under conditions of strong carrier-pair injection such that the recombination lifetime can be approximated by $\tau = 1/\epsilon \Delta n$, where ϵ is the recombination parameter of the material and Δn is the injection-generated excess carrier concentration. Assuming that the source of injection R is set to zero at $t = t_0$, find an analytic expression for $\Delta n(t)$, demonstrating that it exhibits power-law rather than exponential behavior.
- *15.1-3 **Energy Levels in a GaAs/AlGaAs Quantum Well.** (a) Draw the energy-band diagram of a single-crystal multiquantum-well structure of GaAs/AlGaAs to scale on the energy axis when the AlGaAs has the composition $\text{Al}_{0.3}\text{Ga}_{0.7}\text{As}$. The bandgap of GaAs, $E_g(\text{GaAs})$, is 1.42 eV; the bandgap of AlGaAs increases above that of GaAs by ≈ 12.47 meV for each 1% Al increase in the composition. Because of the inherent characteristics of these two materials, the depth of the GaAs conduction-band quantum well is about 60% of the total conduction-plus-valence band quantum-well depths.
 (b) Assume that a GaAs conduction-band well has depth as determined in part (a) above and precisely the same energy levels as the finite square well shown in Fig. 12.1-9(b), for which $(mV_0d^2/2\hbar^2)^{1/2} = 4$, where V_0 is the depth of the well. Find the total width d of the GaAs conduction-band well. The effective mass of an electron in the conduction band of GaAs is $m_e \approx 0.07m_0 = 0.64 \times 10^{-31}$ kg.
- 15.2-1 **Validity of the Approximation for Absorption/Emission Rates.** The derivation of the rate of spontaneous emission made use of the approximation $g_{\nu_0}(\nu) \approx \delta(\nu - \nu_0)$ in the course of evaluating the integral
- $$r_{sp}(\nu) = \int \left[\frac{1}{\tau_r} g_{\nu_0}(\nu) \right] f_e(\nu_0) \varrho(\nu_0) d\nu_0.$$
- (a) Demonstrate that this approximation is satisfactory for GaAs by plotting the functions $g_{\nu_0}(\nu)$, $f_e(\nu_0)$, and $\varrho(\nu_0)$ at $T = 300$ K and comparing their widths. GaAs is collisionally lifetime broadened with $T_2 \approx 1$ ps.
 (b) Repeat part (a) for the rate of absorption in thermal equilibrium.
- 15.2-2 **Peak Spontaneous Emission Rate in Thermal Equilibrium.** (a) Determine the photon energy $h\nu_p$ at which the direct band-to-band spontaneous emission rate from a semiconductor material in thermal equilibrium achieves its maximum value when the Fermi level lies within the bandgap and away from the band edges by at least several times $k_B T$.
 (b) Show that this peak rate (photons per second per hertz per cm^3) is given by

$$r_{sp}(\nu_p) = \frac{D_0}{\sqrt{2e}} (k_B T)^{1/2} = \frac{2(m_r)^{3/2}}{\sqrt{e} \pi \hbar^2 \tau_r} (k_B T)^{1/2} \exp\left(-\frac{E_g}{k_B T}\right).$$

- (c) What is the effect of doping on this result?
 (d) Assuming that $\tau_r = 0.4$ ns, $m_c = 0.07m_0$, $m_v = 0.5m_0$, and $E_g = 1.42$ eV, find the peak rate in GaAs at $T = 300$ K.

- 15.2-3 **Radiative Recombination Rate in Thermal Equilibrium.** (a) Show that the direct band-to-band spontaneous emission rate integrated over all emission frequencies (photons per second per cm^3) is given by

$$\int_0^\infty r_{\text{sp}}(\nu) d\nu = D_0 \frac{\sqrt{\pi}}{2h} (k_B T)^{3/2} = \frac{(m_r)^{3/2}}{\sqrt{2} \pi^{3/2} \hbar^3 \tau_r} (k_B T)^{3/2} \exp\left(-\frac{E_g}{k_B T}\right),$$

provided that the Fermi level is within the semiconductor energy gap and away from the band edges. [Note: $\int_0^\infty x^{1/2} e^{-\mu x} dx = (\sqrt{\pi}/2)\mu^{-3/2}$.]

(b) Compare this with the approximate integrated rate obtained by multiplying the peak rate obtained in Problem 15.2-2 by the approximate frequency width $2k_B T/h$ shown in Fig. 15.2-9.

(c) Using (15.1-10b), set the phenomenological equilibrium radiative recombination rate $\tau_r n_p = \tau_r n_i^2$ (photons per second per cm^3) introduced in Sec. 15.1D equal to the direct band-to-band result derived in (a) to obtain the expression for the radiative recombination rate

$$\tau_r = \frac{\sqrt{2} \pi^{3/2} \hbar^3}{(m_c + m_v)^{3/2}} \frac{1}{(k_B T)^{3/2} \tau_r}.$$

(d) Use the result in (c) to find the value of τ_r for GaAs at $T = 300$ K using $m_c = 0.07m_0$, $m_v = 0.5m_0$, and $\tau_r = 0.4$ ns. Compare this with the value provided in Table 15.1-5 on page 563 ($\tau_r \approx 10^{-10} \text{ cm}^3/\text{s}$).