

Figure 14.18 | Steady-state, photoinduced minority carrier concentrations and photocurrents in a "long" reverse-biased pn junction.

We can find the excess minority carrier hole concentration in the n region using the same type of analysis. Using the x' notation shown in Figure 14.17, we can write

$$\delta p_n = G_L \tau_{p0} - (G_L \tau_{p0} + p_{n0}) e^{-x'/L_p}$$
 (14.38)

Equations (14.37) and (14.38) are plotted in Figure 14.18. We may note that the steady-state values far from the space charge region are the same as were given previously.

The gradient in the minority carrier concentrations will produce diffusion currents in the pn junction. The diffusion current density at x = 0 due to minority carrier electrons is

$$J_{n1} = eD_n \frac{d(\delta n_p)}{dx} \Big|_{x=0} = eD_n \frac{d}{dx} [G_L \tau_{n0} - (G_L \tau_{n0} + n_{p0}) e^{-x/L_n}] \Big|_{x=0}$$

$$= \frac{eD_n}{L_n} (G_L \tau_{n0} + n_{p0})$$
(14.39)

Equation (14.39) can be written as

$$J_{n1} = eG_L L_n + \frac{eD_n n_{p0}}{L_m} (14.40)$$

The first term in Equation (14.40) is the steady-state photocurrent density while the second term is the ideal reverse saturation current density due to the minority carrier electrons.

The diffusion current density (in the x direction) at x' = 0 due to the minority carrier holes is

$$J_{p1} = eG_L L_p + \frac{eD_p p_{n0}}{L_n} (14.41)$$

Similarly, the first term is the steady-state photocurrent density and the second term is the ideal reverse saturation current density.

The total steady-state diode photocurrent density for the long diode is now

$$J_{L} = eG_{L}W + eG_{L}L_{n} + eG_{L}L_{n} = e(W + L_{n} + L_{n})G_{L}$$
 (14.42)

Again note that the photocurrent is in the reverse-biased direction through the diode. The photocurrent given by Equation (14.42) is the result of assuming uniform generation of excess carriers throughout the structure, a long diode, and steady state.

The time response of the diffusion components of the photocurrent is relatively slow, since these currents are the results of the diffusion of minority carriers toward the depletion region. The diffusion components of photocurrent are referred to as the delayed photocurrent.

Objective: Calculate the steady-state photocurrent density in a reverse-biased, long pn diode. Consider a silicon pn diode at T = 300 K with the following parameters:

EXAMPLE 14.5

$$N_a = 10^{16} \text{ cm}^{-3}$$
 $N_d = 10^{16} \text{ cm}^{-3}$
 $D_n = 25 \text{ cm}^2/\text{s}$ $D_p = 10 \text{ cm}^2/\text{s}$
 $\tau_{n0} = 5 \times 10^{-7} \text{ s}$ $\tau_{p0} = 10^{-7} \text{ s}$

Assume that a reverse-biased voltage of $V_R = 5$ volts is applied and let $G_L = 10^{21}$ cm⁻³-s⁻¹.

■ Solution

We may calculate various parameters as follows:

$$L_n = \sqrt{D_n \tau_{n0}} = \sqrt{(25)(5 \times 10^{-7})} = 35.4 \,\mu\text{m}$$

$$L_p = \sqrt{D_p \tau_{p0}} = \sqrt{(10)(10^{-7})} = 10.0 \,\mu\text{m}$$

$$V_{bi} = V_t \ln \left(\frac{N_a N_d}{n_i^2}\right) = (0.0259) \ln \left[\frac{(10^{16})(10^{16})}{(1.5 \times 10^{10})^2}\right] = 0.695 \,\text{V}$$

$$W = \left\{\frac{2\epsilon_s}{e} \left(\frac{N_a + N_d}{N_a N_d}\right) (V_{bi} + V_R)\right\}^{1/2}$$

$$= \left\{\frac{2(11.7)(8.85 \times 10^{-14})}{1.6 \times 10^{-19}} \cdot \frac{(2 \times 10^{16})}{(10^{16})(10^{16})} \cdot (0.695 + 5)\right\}^{1/2} = 1.21 \,\mu\text{m}$$

Finally, the steady-state photocurrent density is

$$J_L = e(W + L_n + L_p)G_L$$

= $(1.6 \times 10^{-19})(1.21 + 35.4 + 10.0) \times 10^{-4}(10^{21}) = 0.75 \text{ A/cm}^2$

Comment

Again, keep in mind that this photocurrent is in the reverse-biased direction through the diode and is many orders of magnitude larger than the reverse-biased saturation current density in the pn junction diode.

EXERCISE PROBLEM

Ex 14.5 The doping concentrations of the photodiode described in Example 14.5 are changed to $N_a = N_d = 10^{15} \text{ cm}^{-3}$. (a) Determine the steady-state photocurrent density. (b) Calculate the ratio of prompt photocurrent to steady-state photocurrent. $[\mathcal{E}LL0.0] = {}^{7}f/{}^{17}f$ (g) ${}^{1}_{7}\text{UD}/\text{V}/\text{L8}L.0 = {}^{7}f$ (p) ${}^{1}\text{SuV}$

In this example calculation, $L_n \gg W$ and $L_p \gg W$. In many pn junction structures, the assumption of a long diode will not be valid, so the photocurrent expression will have to be modified. In addition, the photon energy absorption may not be uniform throughout the pn structure. The effect of nonuniform absorption will be considered in the next section.

14.3.3 PIN Photodiode

In many photodetector applications, the speed of response is important; therefore, the prompt photocurrent generated in the space charge region is the only photocurrent of interest. To increase the photodetector sensitivity, the depletion region width should be made as large as possible. This can be achieved in a PIN photodiode.

The PIN diode consists of a p region and an n region separated by an intrinsic region. A sketch of a PIN diode is shown in Figure 14.19a. The intrinsic region width W is much larger than the space charge width of a normal pn junction. If a reverse bias is applied to the PIN diode, the space charge region extends completely through the intrinsic region.

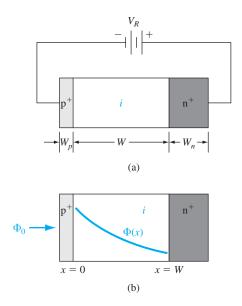


Figure 14.19 | (a) A reverse-biased PIN photodiode. (b) Geometry showing nonuniform photon absorption.

Assume that a photon flux Φ_0 is incident on the p⁺ region. If we assume that the p⁺ region width W_p is very thin, then the photon flux, as a function of distance, in the intrinsic region is $\Phi(x) = \Phi_0 e^{-\alpha x}$, where α is the photon absorption coefficient. This nonlinear photon absorption is shown in Figure 14.19b. The photocurrent density generated in the intrinsic region can be found as

$$J_L = e \int_0^W G_L dx = e \int_0^W \Phi_0 \alpha e^{-\alpha x} dx = e \Phi_0 (1 - e^{-\alpha W})$$
 (14.43)

This equation assumes that there is no electron–hole recombination within the space charge region and also that each photon absorbed creates one electron–hole pair.

Objective: Calculate the photocurrent density in a PIN photodiode.

EXAMPLE 14.6

Consider a silicon PIN diode with an intrinsic region width of $W = 20 \mu m$. Assume that the photon flux is 10^{17} cm⁻²-s⁻¹ and the absorption coefficient is $\alpha = 10^3$ cm⁻¹.

■ Solution

The generation rate of electron-hole pairs at the front edge of the intrinsic region is

$$G_{L1} = \alpha \Phi_0 = (10^3)(10^{17}) = 10^{20} \,\mathrm{cm}^{-3} - \mathrm{s}^{-1}$$

and the generation rate at the back edge of the intrinsic region is

$$G_{L2} = \alpha \Phi_0 e^{-\alpha W} = (10^3)(10^{17}) \exp\left[-(10^3)(20 \times 10^{-4})\right]$$

= 0.135 × 10²⁰ cm⁻³-s⁻¹

The generation rate is obviously not uniform throughout the intrinsic region. The photocurrent density is then

$$J_L = e\Phi_0(1 - e^{-\alpha W})$$
= $(1.6 \times 10^{-19})(10^{17})\{1 - \exp[-(10^3)(20 \times 10^{-4})]\}$
= 13.8 mA/cm^2

Comment

The prompt photocurrent density of a PIN photodiode will be larger than that of a regular photodiode since the space charge region is larger in a PIN photodiode.

EXERCISE PROBLEM

Ex 14.6 Repeat Example 14.6 for photon absorption coefficients of (a) $\alpha = 10^2$ cm⁻¹ and (b) $\alpha = 10^4$ cm⁻¹. [$_7$ tuɔ/ $_7$ tuo/ $_7$ tuo

In most situations, we will not have a long diode; thus, the steady-state photocurrent described by Equation (14.42) will not apply for most photodiodes.

14.3.4 Avalanche Photodiode

The avalanche photodiode is similar to the pn or PIN photodiode except that the bias applied to the avalanche photodiode is sufficiently large to cause impact ionization.

Electron-hole pairs are generated in the space charge region by photon absorption, as we have discussed previously. The photon-generated electrons and holes now generate additional electron-hole pairs through impact ionization. The avalanche photodiode now has a current gain introduced by the avalanche multiplication factor.

The electron–hole pairs generated by photon absorption and by impact ionization are swept out of the space charge region very quickly. If the saturation velocity is 10^7 cm/s in a depletion region that is 10μ m wide, then the transit time is

$$\tau_t = \frac{10^7}{10 \times 10^{-4}} = 100 \text{ ps}$$

The period of a modulation signal would be $2\tau_t$, so that the frequency would be

$$f = \frac{1}{2\tau_t} = \frac{1}{200 \times 10^{-12}} = 5 \text{ GHz}$$

If the avalanche photodiode current gain is 20, then the gain-bandwidth product is 100 GHz. The avalanche photodiode could respond to light waves modulated at microwave frequencies.

14.3.5 Phototransistor

A bipolar transistor can also be used as a photodetector. The phototransistor can have high gain through the transistor action. An npn bipolar phototransistor is shown in Figure 14.20a. This device has a large base–collector junction area and is usually operated with the base open circuited. Figure 14.20b shows the block diagram of the phototransistor. Electrons and holes generated in the reverse-biased B–C junction are swept out of the space charge region, producing a photocurrent I_L . Holes are swept

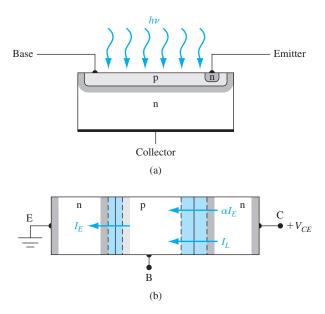


Figure 14.20 | (a) A bipolar phototransistor. (b) Block diagram of the open-base phototransistor.

into the p-type base, making the base positive with respect to the emitter. Since the B–E becomes forward-biased, electrons will be injected from the emitter back into the base, leading to the normal transistor action.

From Figure 14.20b, we see that

$$I_E = \alpha I_E + I_L \tag{14.44}$$

where I_L is the photon-generated current and α is the common base current gain. Since the base is an open circuit, we have $I_C = I_E$, so Equation (14.44) can be written as

$$I_C = \alpha I_C + I_L \tag{14.45}$$

Solving for I_C , we find

$$I_C = \frac{I_L}{1 - \alpha} \tag{14.46}$$

Relating α to β , the dc common emitter current gain, Equation (14.46) becomes

$$I_C = (1 + \beta)I_L \tag{14.47}$$

Equation (14.47) shows that the basic B–C photocurrent is multiplied by the factor $(1 + \beta)$. The phototransistor, then, amplifies the basic photocurrent.

With the relatively large B–C junction area, the frequency response of the phototransistor is limited by the B–C junction capacitance. Since the base is essentially the input to the device, the large B–C capacitance is multiplied by the Miller effect, so the frequency response of the phototransistor is further reduced. The phototransistor, however, is a lower-noise device than the avalanche photodiode.

Phototransistors can also be fabricated in heterostructures. The injection efficiency is increased as a result of the bandgap differences, as we discussed in Chapter 12. With the bandgap difference, the lightly doped base restriction no longer applies. A fairly heavily doped, narrow-base device can be fabricated with a high blocking voltage and a high gain.

TEST YOUR UNDERSTANDING

TYU 14.5 Consider a long silicon pn junction photodiode with the parameters given in Example 14.5. The cross-sectional area is $A = 10^{-3}$ cm². Assume the photodiode is reverse biased by a 5-volt battery in series with a 5 k Ω load resistor. An optical signal at a wavelength of $\lambda = 1$ μ m is incident on the photodiode producing a uniform generation rate of excess carriers throughout the entire device. Determine the incident intensity such that the voltage across the load resistor is 0.5 V. ($_{7}$ UD/M 997:0 = ^{a}I ·Su Ψ)

14.4 | PHOTOLUMINESCENCE AND ELECTROLUMINESCENCE

In the first section of this chapter, we have discussed the creation of excess electron—hole pairs by photon absorption. Eventually, excess electrons and holes recombine, and in direct bandgap materials the recombination process may result in the emission of a photon. The general property of light emission is referred to as luminescence.

When excess electrons and holes are created by photon absorption, photon emission from the recombination process is called photoluminescence.

Electroluminescence is the process of generating photon emission when the excitation of excess carriers is a result of an electric current caused by an applied electric field. We are mainly concerned here with injection electroluminescence, the result of injecting carriers across a pn junction. The light emitting diode and the pn junction laser diode are examples of this phenomenon. In these devices, electric energy, in the form of a current, is converted directly into photon energy.

14.4.1 Basic Transitions

Once electron-hole pairs are formed, there are several possible processes by which the electrons and holes can recombine. Some recombination processes may result in photon emission from direct bandgap materials, whereas other recombination processes in the same material may not.

Figure 14.21a shows the basic interband transitions. Curve (i) corresponds to an intrinsic emission very close to the bandgap energy of the material. Curves (ii) and (iii) correspond to energetic electrons or holes. If either of these recombinations result in the emission of a photon, the energy of the emitted photon will be slightly larger than the bandgap energy. There will then be an emission spectrum and a bandwidth associated with the emission.

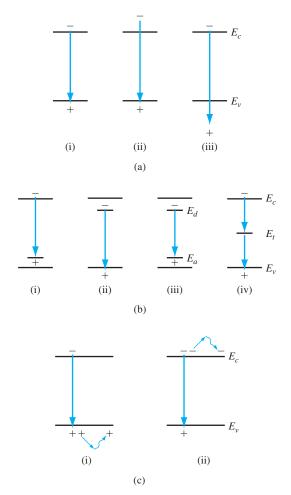
The possible recombination processes involving impurity or defect states are shown in Figure 14.21b. Curve (i) is the conduction band to acceptor transition, curve (ii) is the donor to valence-band transition, curve (iii) is the donor to acceptor transition, and curve (iv) is the recombination due to a deep trap. Curve (iv) is a nonradiative process corresponding to the Shockley–Read–Hall recombination process discussed in Chapter 6. The other recombination processes may or may not result in the emission of a photon.

Figure 14.21c shows the Auger recombination process, which can become important in direct bandgap materials with high doping concentrations. The Auger recombination process is a nonradiative process. The Auger recombination, in one case, shown in curve (i), is a recombination between an electron and hole, accompanied by the transfer of energy to another free hole. Similarly, in the second case, the recombination between an electron and hole can result in the transfer of energy to a free electron as shown in curve (ii). The third particle involved in this process will eventually lose its energy to the lattice in the form of heat. The process involving two holes and an electron would occur predominantly in heavily doped p-type materials, and the process involving two electrons and a hole would occur primarily in a heavily doped n-type material.

The recombination processes shown in Figure 14.21a indicate that the emission of a photon is not necessarily at a single, discrete energy, but can occur over a range of energies. The spontaneous emission rate generally has the form

$$I(\nu) \propto \nu^2 (h\nu - E_g)^{1/2} \exp\left[\frac{-(h\nu - E_g)}{kT}\right]$$
 (14.48)

where E_g is the bandgap energy. Figure 14.22 shows the emission spectra from gallium arsenide. The peak photon energy decreases with temperature because the



Wavelength (μ m)

1.0 0.95 0.9 0.85

100

0.036 eV

0.018 eV

Resolution \rightarrow ||

10-3

1.24 1.28 1.32 1.36 1.40 1.44 1.48 1.52

Photon energy (eV)

Figure 14.21 | Basic transitions in a semiconductor.

Figure 14.22 | GaAs diode emission spectra at T = 300 K and T = 77 K. (From Sze and Ng [17].)

bandgap energy decreases with temperature. We will show that the bandwidth of the emission spectra can be greatly reduced in a laser diode by using an optical resonator.

14.4.2 Luminescent Efficiency

We have shown that not all recombination processes are radiative. An efficient luminescent material is one in which radiative transitions predominate. The quantum efficiency is defined as the ratio of the radiative recombination rate to the total recombination rate for all processes. We can write

$$\eta_q = \frac{R_r}{R} \tag{14.49}$$

where η_q is the quantum efficiency, R_r is the radiative recombination rate, and R is the total recombination rate of the excess carriers. Since the recombination rate is

inversely proportional to lifetime, we can write the quantum efficiency in terms of lifetimes as

$$\eta_q = \frac{\tau_{nr}}{\tau_{nr} + \tau_r} \tag{14.50}$$

where τ_{nr} is the nonradiative lifetime and τ_r is the radiative lifetime. For a high luminescent efficiency, the nonradiative lifetimes must be large; thus, the probability of a nonradiative recombination is small compared to the radiative recombination.

The interband recombination rate of electrons and holes will be directly proportional to the number of electrons available and directly proportional to the number of available empty states (holes). We can write

$$R_r = Bnp \tag{14.51}$$

where R_r is the band-to-band radiative recombination rate and B is the constant of proportionality. The values of B for direct-bandgap materials are on the order of 10^6 larger than for indirect bandgap materials. The probability of a direct band-to-band radiative recombination transition in an indirect bandgap material is very unlikely.

One problem encountered with the emission of photons from a direct bandgap material is the reabsorption of the emitted photons. In general, the emitted photons will have energies $h\nu > E_g$, which means that the absorption coefficient is not zero for this energy. In order to generate a light output from a light emitting device, the process must take place near the surface. One possible solution to the reabsorption problem is to use heterojunction devices. These are discussed in later sections.

14.4.3 Materials

An important direct bandgap semiconductor material for optical devices is gallium arsenide. Another compound material that is of great interest is $Al_xGa_{1-x}As$. This material is a compound semiconductor in which the ratio of aluminum atoms to gallium atoms can be varied to achieve specific characteristics. Figure 14.23 shows the bandgap energy as a function of the mole fraction between aluminum and gallium. We can note from the figure that for 0 < x < 0.45, the alloy material is a direct bandgap material. For x > 0.45, the material becomes an indirect bandgap material, not suitable for optical devices. For 0 < x < 0.35, the bandgap energy can be expressed as

$$E_g = 1.424 + 1.247x \,\text{eV}$$
 (14.52)

Another compound semiconductor used for optical devices is the $GaAs_{1-x}P_x$ system. Figure 14.24a shows the bandgap energy as a function of the mole fraction x. For $0 \le x \le 0.45$, this material is also a direct bandgap material, and for x > 0.45, the bandgap becomes indirect. Figure 14.24b is the E versus k diagram, showing how the bandgap changes from direct to indirect as the mole fraction changes.

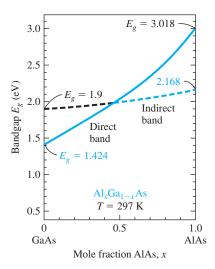


Figure 14.23 | Bandgap energy of $Al_xGa_{1-x}As$ as a function of the mole fraction x. (*From Sze [18]*.)

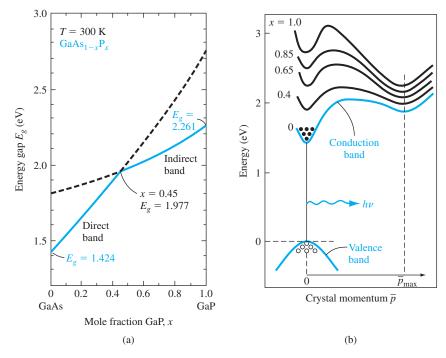


Figure 14.24 | (a) Bandgap energy of $GaAs_{1-x}P_x$ as a function of mole fraction x. (b) E versus k diagram of $GaAs_{1-x}P_x$ for various values of x. (*From Sze [18].*)

EXAMPLE 14.7

Objective: Determine the output wavelength of a $GaAs_{1-x}P_x$ material for two different mole fractions.

Consider first GaAs and then GaAs_{1-x}P_x.

■ Solution

GaAs has a bandgap energy of $E_g = 1.42$ eV. This material would produce a photon output at a wavelength of

$$\lambda = \frac{1.24}{E} = \frac{1.24}{1.42} = 0.873 \ \mu \text{m}$$

This wavelength is in the infrared range and not in the visible range. If we desire a visible output with a wavelength of $\lambda = 0.653 \,\mu\text{m}$, for example, the bandgap energy would have to be

$$E = \frac{1.24}{\lambda} = \frac{1.24}{0.653} = 1.90 \text{ eV}$$

This bandgap energy would correspond to a mole fraction of approximately x = 0.4.

Comment

By changing the mole fraction in the $GaAs_{1-x}P_x$ system, the output can change from the infrared to the red spectrum.

EXERCISE PROBLEM

Ex 14.7 Determine the output wavelength of a $GaAs_{1-x}P_x$ material for mole fractions of (a) x = 0.15 and (b) x = 0.30.

 $[m_{\lambda} \delta 0 7.0 = \lambda(d); m_{\lambda} \delta 77.0 = \lambda(b).$

14.5 | LIGHT EMITTING DIODES

Photodetectors and solar cells convert optical energy into electrical energy—the photons generate excess electrons and holes, which produce an electric current. We might also apply a voltage across a pn junction resulting in a diode current, which in turn can produce photons and a light output. This inverse mechanism is called injection electroluminescence. This device is known as a Light Emitting Diode (LED). The spectral output of an LED may have a relatively wide wavelength bandwidth of between 30 and 40 nm. However, this emission spectrum is narrow enough so that a particular color is observed, provided the output is in the visible range.

14.5.1 Generation of Light

As we have discussed previously, photons may be emitted if an electron and hole recombine by a direct band-to-band recombination process in a direct bandgap material. The emission wavelength, from Equation (14.1), is

$$\lambda = \frac{hc}{E_g} = \frac{1.24}{E_g} \,\mu\text{m} \tag{14.53}$$

where E_g is the bandgap energy measured in electron-volts.

When a voltage is applied across a pn junction, electrons and holes are injected across the space charge region where they become excess minority carriers. These excess minority carriers diffuse into the neutral semiconductor regions where they recombine with majority carriers. If this recombination process is a direct band-to-band process, photons are emitted. The diode diffusion current is directly proportional to the recombination rate, so the output photon intensity will also be proportional to the ideal diode diffusion current. In gallium arsenide, electroluminescence originates primarily on the p side of the junction because the efficiency for electron injection is higher than that for hole injection.

14.5.2 Internal Quantum Efficiency

The *internal quantum efficiency* of an LED is the fraction of diode current that produces luminescence. The internal quantum efficiency is a function of the injection efficiency and a function of the percentage of radiative recombination events compared with the total number of recombination events.

The three current components in a forward-biased diode are the minority carrier electron diffusion current, the minority carrier hole diffusion current, and the space charge recombination current. These current densities can be written, respectively, as

$$J_n = \frac{eD_n n_{p0}}{L_n} \left[\exp\left(\frac{eV}{kT}\right) - 1 \right]$$
 (14.54a)

$$J_p = \frac{eD_p p_{n0}}{L_p} \left[\exp\left(\frac{eV}{kT}\right) - 1 \right]$$
 (14.54b)

and

$$J_R = \frac{en_i W}{2\tau_0} \left[\exp\left(\frac{eV}{2kT}\right) - 1 \right]$$
 (14.54c)

The recombination of electrons and holes within the space charge region is, in general, through traps near midgap and is a nonradiative process. Since luminescence is due primarily to the recombination of minority carrier electrons in GaAs, we can define an injection efficiency as the fraction of electron current to total current. Then

$$\gamma = \frac{J_n}{J_n + J_p + J_R} \tag{14.55}$$

where γ is the injection efficiency. We can make γ approach unity by using an n⁺p diode so that J_p is a small fraction of the diode current and by forward biasing the diode sufficiently so that J_R is a small fraction of the total diode current.

Once the electrons are injected into the p region, not all electrons will recombine radiatively. We can define the radiative and nonradiative recombination rates as

$$R_r = \frac{\delta n}{\tau_r} \tag{14.56a}$$

and

$$R_{nr} = \frac{\delta n}{\tau_{nr}} \tag{14.56b}$$

where τ_r and τ_{nr} are the radiative and nonradiative recombination lifetimes, respectively, and δn is the excess carrier concentration. The total recombination rate is

$$R = R_r + R_{nr} = \frac{\delta n}{\tau} = \frac{\delta n}{\tau_r} + \frac{\delta n}{\tau_{nr}}$$
 (14.57)

where τ is the net excess carrier lifetime.

The radiative efficiency is defined as the fraction of recombinations that are radiative. We can write

$$\eta = \frac{R_r}{R_r + R_{nr}} = \frac{\frac{1}{\tau_r}}{\frac{1}{\tau_r} + \frac{1}{\tau_{nr}}} = \frac{\tau}{\tau_r}$$
(14.58)

where η is the radiative efficiency. The nonradiative recombination rate is proportional to N_t , which is the density of nonradiative trapping sites within the forbidden bandgap. Obviously, the radiative efficiency increases as N_t is reduced.

The internal quantum efficiency is now written as

$$\eta_i = \gamma \eta \tag{14.59}$$

The radiative recombination rate is proportional to the p-type doping. As the p-type doping increases, the radiative recombination rate increases. However, the injection efficiency decreases as the p-type doping increases; therefore, there is an optimum doping that maximizes the internal quantum efficiency.

14.5.3 External Quantum Efficiency

One very important parameter of the LED is the *external quantum efficiency*: the fraction of generated photons that are actually emitted from the semiconductor. The external quantum efficiency is normally a much smaller number than the internal quantum efficiency. Once a photon has been produced in the semiconductor, there are three loss mechanisms the photon may encounter: photon absorption within the semiconductor, Fresnel loss, and critical angle loss.

Figure 14.25 shows a pn junction LED. Photons can be emitted in any direction. Since the emitted photon energy must be $h\nu \ge E_g$, these emitted photons can be reabsorbed within the semiconductor material. The majority of photons will actually be emitted away from the surface and reabsorbed in the semiconductor.

Photons must be emitted from the semiconductor into air; thus, the photons must be transmitted across a dielectric interface. Figure 14.26 shows the incident, reflected, and transmitted waves. The parameter \bar{n}_2 is the index of refraction for the semiconductor and \bar{n}_1 is the index of refraction for air. The reflection coefficient is

$$\Gamma = \left(\frac{\bar{n}_2 - \bar{n}_1}{\bar{n}_2 + \bar{n}_1}\right)^2 \tag{14.60}$$

This effect is called Fresnel loss. The reflection coefficient Γ is the fraction of incident photons that are reflected back into the semiconductor.

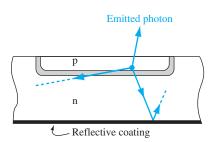


Figure 14.25 | Schematic of photon emission at the pn junction of an LED.

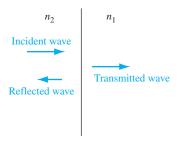


Figure 14.26 | Schematic of incident, reflected, and transmitted photons at a dielectric interface.

Objective: Calculate the reflection coefficient at a semiconductor–air interface.

Consider the interface between a GaAs semiconductor and air.

EXAMPLE 14.8

■ Solution

The index of refraction for GaAs is $\bar{n}_2 = 3.8$ at a wavelength of $\lambda = 0.70 \ \mu m$ and the index of refraction for air is $\bar{n}_1 = 1.0$. The reflection coefficient is

$$\Gamma = \left(\frac{\overline{n}_2 - \overline{n}_1}{\overline{n}_2 + \overline{n}_1}\right)^2 = \left(\frac{3.8 - 1.0}{3.8 + 1.0}\right)^2 = 0.34$$

Comment

A reflection coefficient of $\Gamma=0.34$ means that 34 percent of the photons incident from the gallium arsenide onto the semiconductor-air interface are reflected back into the semiconductor.

EXERCISE PROBLEM

Ex 14.8 At a wavelength of $\lambda = 0.70 \ \mu m$, the index of refraction for GaAs is $\overline{n}_2 = 3.8$ and that for GaP is $\overline{n}_2 = 3.2$. Consider a GaAs_{1-x}P_x material with a mole fraction x = 0.40. Assuming the index of refraction is a linear function of the mole fraction, determine the reflection coefficient, Γ , at the GaAs_{0.6}P_{0.4}—air interface. (SIE'0 = J'suV)

Photons incident on the semiconductor–air interface at an angle are refracted as shown in Figure 14.27. If the photons are incident on the interface at an angle greater than the critical angle θ_c , the photons experience total internal reflection. The critical angle is determined from Snell's law and is given by

$$\theta_c = \sin^{-1}\left(\frac{\overline{n}_1}{\overline{n}_2}\right) \tag{14.61}$$

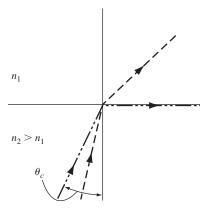


Figure 14.27 | Schematic showing refraction and total internal reflection at the critical angle at a dielectric interface.

EXAMPLE 14.9

Objective: Calculate the critical angle at a semiconductor–air interface.

Consider the interface between GaAs and air.

■ Solution

For GaAs, $\bar{n}_2 = 3.8$ at a wavelength of $\lambda = 0.70 \,\mu\text{m}$ and for air, $\bar{n}_1 = 1.0$. The critical angle is

$$\theta_c = \sin^{-1}\left(\frac{\overline{n}_1}{\overline{n}_2}\right) = \sin^{-1}\left(\frac{1.0}{3.8}\right) = 15.3^{\circ}$$

Comment

Any photon that is incident at an angle greater than 15.3° will be reflected back into the semiconductor.

EXERCISE PROBLEM

Ex 14.9 Repeat Example 14.9 for GaAs_{0.6}P_{0.4}. See Exercise Problem Ex 14.8 for a discussion of the dielectric constant.

$$(Ans. \theta_c = 16.3^\circ)$$

Figure 14.28a shows the external quantum efficiency plotted as a function of the p-type doping concentration and Figure 14.28b is a plot of the external efficiency as a function of junction depth below the surface. Both figures show that the external quantum efficiency is in the range of 1 to 3 percent.

14.5.4 LED Devices

The wavelength of the output signal of an LED is determined by the bandgap energy of the semiconductor. Gallium arsenide, a direct bandgap material, has a bandgap energy of $E_g = 1.42$ eV, which yields a wavelength of $\lambda = 0.873$ μ m. Comparing this wavelength to the visible spectrum, which is shown in Figure 14.5, the output

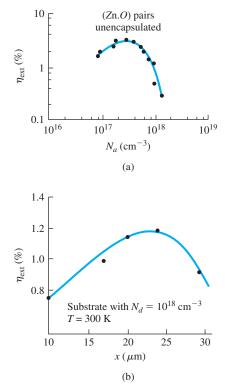


Figure 14.28 | (a) External quantum efficiency of a GaP LED versus acceptor doping. (b) External quantum efficiency of a GaAs LED versus junction depth. (From Yang [22].)

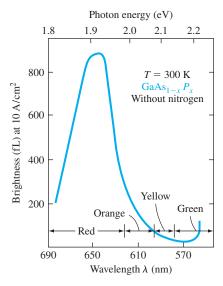


Figure 14.29 | Brightness of GaAsP diodes versus wavelength (or versus bandgap energy). (From Yang [22].)

of a GaAs LED is not in the visible range. For a visible output, the wavelength of the signal should be in the range of 0.4 to 0.72 μ m. This range of wavelengths corresponds to bandgap energies between approximately 1.7 and 3.1 eV.

GaAs_{1-x}P_x is a direct bandgap material for $0 \le x \le 0.45$, as shown in Figure 14.24. At x = 0.40, the bandgap energy is approximately $E_g = 1.9$ eV, which would produce an optical output in the red range. Figure 14.29 shows the brightness of GaAs_{1-x}P_x diodes for different values of x. The peak also occurs in the red range. By using planar technology, GaAs_{0.6}P_{0.4} monolithic arrays have been fabricated for numeric and alphanumeric displays. When the mole fraction x is greater than 0.45, the material changes to an indirect bandgap semiconductor so that the quantum efficiency is greatly reduced.

 $GaAl_xAs_{1-x}$ can be used in a heterojunction structure to form an LED. A device structure is shown in Figure 14.30. Electrons are injected from the wide-bandgap N-GaAl_{0.7}As_{0.3} into the narrow-bandgap p-GaAl_{0.6}As_{0.4}. The minority carrier electrons

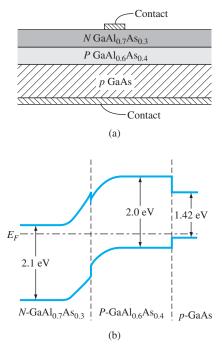


Figure 14.30 | The (a) cross section and (b) thermal equilibrium energy-band diagram of a GaAlAs heterojunction LED. (*From Yang [22]*.)

in the p material can recombine radiatively. Since $E_{gp} < E_{gN}$, the photons are emitted through the wide-bandgap N material with essentially no absorption. The wide bandgap N material acts as an optical window and the external quantum efficiency increases.

14.6 | LASER DIODES

The photon output of the LED is due to an electron giving up energy as it makes a transition from the conduction band to the valence band. The LED photon emission is spontaneous in that each band-to-band transition is an independent event. The spontaneous emission process yields a spectral output of the LED with a fairly wide bandwidth. If the structure and operating condition of the LED are modified, the device can operate in a new mode, producing a coherent spectral output with a bandwidth of wavelengths less than 0.1 nm. This new device is a laser diode, where laser stands for Light Amplification by Stimulated Emission of Radiation. Although there are many different types of lasers, we are here concerned only with the pn junction laser diode.

14.6.1 Stimulated Emission and Population Inversion

Figure 14.31a shows the case when an incident photon is absorbed and an electron is elevated from an energy state E_1 to an energy state E_2 . This process is known as induced absorption. If the electron spontaneously makes the transition back to the lower energy level with a photon being emitted, we have a spontaneous emission process as indicated in Figure 14.31b. On the other hand, if there is an incident photon at a time when an electron is in the higher energy state as shown in Figure 14.31c, the incident photon can interact with the electron, causing the electron to make a transition downward. The downward transition produces a photon. Since this process was initiated by the incident photon, the process is called *stimulated* or *induced emission*. Note that this stimulated emission process has produced two photons; thus, we can have optical gain or amplification. The two emitted photons are in phase so that the spectral output will be coherent.

In thermal equilibrium, the electron distribution in a semiconductor is determined by the Fermi–Dirac statistics. If the Boltzmann approximation applies, then we can write

$$\frac{N_2}{N_1} = \exp\left[\frac{-(E_2 - E_1)}{kT}\right] \tag{14.62}$$

where N_1 and N_2 are the electron concentrations in the energy levels E_1 and E_2 , respectively, and where $E_2 > E_1$. In thermal equilibrium, $N_2 < N_1$. The probability of an induced absorption event is exactly the same as that of an induced emission event. The number of photons absorbed is proportional to N_1 and the number of additional photons emitted is proportional to N_2 . In order to achieve optical amplification or for lasing action to occur, we must have $N_2 > N_1$; this is called population inversion. We cannot achieve lasing action at thermal equilibrium.

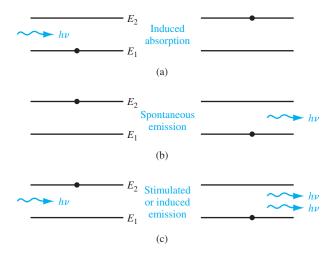


Figure 14.31 | Schematic diagram showing (a) induced absorption, (b) spontaneous emission, and (c) stimulated emission processes.

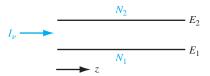


Figure 14.32 | Light propagating in z direction through a material with two energy levels.

Figure 14.32 shows the two energy levels with a light wave at an intensity I_{ν} propagating in the z direction. The change in intensity as a function of z can be written as

$$\frac{dI_{\nu}}{dz} \propto \frac{\text{\# photons emitted}}{\text{cm}^3} - \frac{\text{\# photons absorbed}}{\text{cm}^3}$$

or

$$\frac{dI_{\nu}}{dz} = N_2 W_i \cdot h\nu - N_1 W_i \cdot h\nu \tag{14.63}$$

where W_i is the induced transition probability. Equation (14.63) assumes no loss mechanisms and neglects the spontaneous transitions.

Equation (14.63) can be written as

$$\frac{dI_{\nu}}{dz} = \gamma(\nu)I_{\nu} \tag{14.64}$$

where $\gamma(\nu) \propto (N_2 - N_1)$ and is the amplification factor. From Equation (14.64), the intensity is

$$I_{\nu} = I_{\nu}(0)e^{\gamma(\nu)z} \tag{14.65}$$

Amplification occurs when $\gamma(\nu) > 0$ and absorption occurs when $\gamma(\nu) < 0$.

We can achieve population inversion and lasing in a forward-biased pn homojunction diode, if both sides of the junction are degenerately doped. Figure 14.33a shows the energy-band diagram of a degenerately doped pn junction in thermal equilibrium. The Fermi level is in the conduction band in the n-region and the Fermi level is in the valence band in the p region. Figure 14.33b shows the energy bands of the pn junction when a forward bias is applied. The gain factor in a pn homojunction diode is given by

$$\gamma(\nu) \propto \left\{ 1 - \exp\left[\frac{h\nu - (E_{Fn} - E_{Fp})}{kT}\right] \right\}$$
 (14.66)

In order for $\gamma(\nu) > 1$, we must have $h\nu < (E_{Fn} - E_{Fp})$, which implies that the junction must be degenerately doped since we also have the requirement that $h\nu \ge E_g$. In the vicinity of the junction, there is a region in which population inversion occurs. There are large numbers of electrons in the conduction band directly above a large number of empty states. If band-to-band recombination occurs, photons will be emitted with energies in the range $E_g < h\nu < (E_{Fn} - E_{Fp})$.

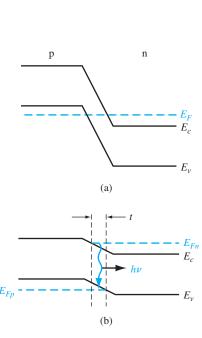


Figure 14.33 | (a) Degenerately doped pn junction at zero bias. (b) Degenerately doped pn junction under forward bias with photon emission.

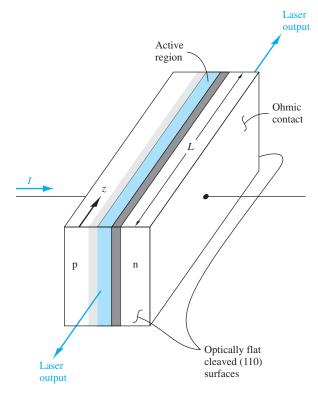


Figure 14.34 | A pn junction laser diode with cleaved (110) planes forming the Fabry-Perot cavity. (After Yang [22].)

14.6.2 Optical Cavity

Population inversion is one requirement for lasing action to occur. Coherent emission output is achieved by using an optical cavity. The cavity will cause a buildup of the optical intensity from positive feedback. A resonant cavity consisting of two parallel mirrors is known as a Fabry–Perot resonator. The resonant cavity can be fabricated, for example, by cleaving a gallium arsenide crystal along the (110) planes as shown in Figure 14.34. The optical wave propagates through the junction in the *z* direction, bouncing back and forth between the end mirrors. The mirrors are actually only partially reflecting so that a portion of the optical wave will be transmitted out of the junction.

For resonance, the length of the cavity L must be an integral number of half wavelengths, or

$$N\left(\frac{\lambda}{2}\right) = L \tag{14.67}$$

where N is an integer. Since λ is small and L is relatively large, there can be many resonant modes in the cavity. Figure 14.35a shows the resonant modes as a function of wavelength.

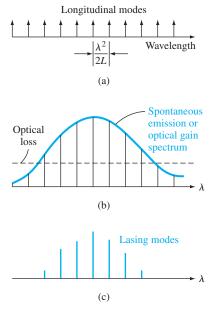


Figure 14.35 | Schematic diagram showing (a) resonant modes of a cavity with length *L*, (b) spontaneous emission curve, and (c) actual emission modes of a laser diode. (*After Yang [22]*.)

When a forward-bias current is applied to the pn junction, spontaneous emission will initially occur. The spontaneous emission spectrum is relatively broadband and is superimposed on the possible lasing modes as shown in Figure 14.35b. In order for lasing to be initiated, the spontaneous emission gain must be larger than the optical losses. By positive feedback in the cavity, lasing can occur at several specific wavelengths as indicated in Figure 14.35c.

14.6.3 Threshold Current

The optical intensity in the device can be written from Equation (14.65) as $I_{\nu} \propto e^{\gamma(\nu)z}$, where $\gamma(\nu)$ is the amplification factor. We have two basic loss mechanisms. The first is the photon absorption in the semiconductor material. We can write

$$I_{\nu} \propto e^{-\alpha(\nu)z} \tag{14.68}$$

where $\alpha(\nu)$ is the absorption coefficient. The second loss mechanism is due to the partial transmission of the optical signal through the ends, or through the partially reflecting mirrors.

At the onset of lasing, which is known as threshold, the optical loss of one round trip through the cavity is just offset by the optical gain. The threshold condition is then expressed as

$$\Gamma_1 \Gamma_2 \exp\left[(2\gamma_t(\nu) - 2\alpha(\nu))L \right] = 1 \tag{14.69}$$

where Γ_1 and Γ_2 are the reflectivity coefficients of the two end mirrors. For the case when the optical mirrors are cleaved (110) surfaces of gallium arsenide, the reflectivity coefficients are given approximately by

$$\Gamma_1 = \Gamma_2 = \left(\frac{\overline{n}_2 - \overline{n}_1}{\overline{n}_2 + \overline{n}_1}\right)^2 \tag{14.70}$$

where \bar{n}_2 and \bar{n}_1 are the index of refraction parameters for the semiconductor and air, respectively. The parameter $\gamma_t(\nu)$ is the optical gain at threshold.

The optical gain at threshold, $\gamma_t(\nu)$, may be determined from Equation (14.69) as

$$\gamma_t(\nu) = \alpha + \frac{1}{2L} \ln\left(\frac{1}{\Gamma_1 \Gamma_2}\right) \tag{14.71}$$

Since the optical gain is a function of the pn junction current, we can define a threshold current density as

$$J_{th} = \frac{1}{\beta} \left[\alpha + \frac{1}{2L} \ln \left(\frac{1}{\Gamma_1 \Gamma_2} \right) \right]$$
 (14.72)

where β can be determined theoretically or experimentally. Figure 14.36 shows the threshold current density as a function of the mirror losses. We may note the relatively high threshold current density for a pn junction laser diode.

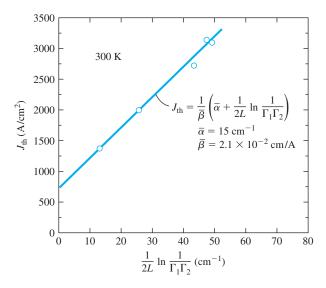


Figure 14.36 | Threshold current density of a laser diode as a function of Fabry-Perot cavity end losses. (*After Yang [22].*)

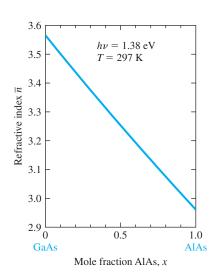


Figure 14.37 | Index of refraction of $Al_xGa_{1-x}As$ as a function of mole fraction x. (From Sze [18].)

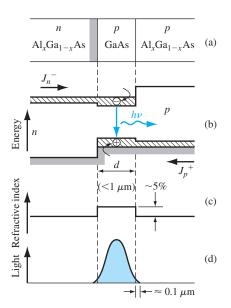


Figure 14.38 (a) Basic double heterojunction structure. (b) Energy-band diagram under forward bias. (c) Refractive index change through the structure. (d) Confinement of light in the dielectric waveguide. (From Yang [22].)

14.6.4 Device Structures and Characteristics

We have seen that in a homojunction LED, the photons may be emitted in any direction, which lowers the external quantum efficiency. Significant improvement in device characteristics can be made if the emitted photons are confined to a region near the junction. This confinement can be achieved by using an optical dielectric waveguide. The basic device is a three-layered, double heterojunction structure known as a double heterojunction laser. A requirement for a dielectric waveguide is that the index of refraction of the center material be larger than that of the other two dielectrics. Figure 14.37 shows the index of refraction for the AlGaAs system. We may note that GaAs has the highest index of refraction.

An example of a double heterojunction laser is shown in Figure 14.38a. A thin p-GaAs layer is between P-AlGaAs and N-AlGaAs layers. A simplified energy-band diagram is shown in Figure 14.38b for the forward-biased diode. Electrons are injected from the N-AlGaAs into the p-GaAs. Population inversion is easily obtained since the conduction band potential barrier prevents the electrons from diffusing into the P-AlGaAs region. Radiative recombination is then confined to the p-GaAs region. Since the index of refraction of GaAs is larger than that of AlGaAs, the light wave is also confined to the GaAs region. An optical cavity can be formed by cleaving the semiconductor perpendicular to the N-AlGaAs-p-GaAs junction.

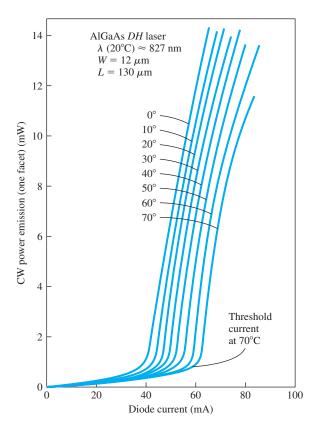


Figure 14.39 | Typical output power versus laser diode current at various temperatures. (From Yang [22].)

Typical optical output versus diode current characteristics are shown in Figure 14.39. The threshold current is defined to be the current at the breakpoint. At low currents, the output spectrum is very wide and is the result of the spontaneous transitions. When the diode current is slightly above the threshold value, the various resonant frequencies are observed. When the diode current becomes large, a single dominant mode with a narrow bandwidth is produced.

The performance of the laser diode can be further improved if a very narrow recombination region is used with a somewhat wider optical waveguide. Very complex structures using multilayers of compound semiconductor materials have been fabricated in a continuing effort to improve semiconductor laser performance.

14.7 | SUMMARY

The absorption or emission of light (photons) in semiconductors leads to the study of a general class of devices called optoelectronics. A few of these devices have been discussed and analyzed in this chapter.

- The photon absorption process has been discussed and the absorption coefficient data for semiconductors has been presented.
- Solar cells convert optical power into electrical power. The simple pn junction solar cell was initially considered. The short-circuit current, open-circuit voltage, and maximum power were considered.
- Heterojunction and amorphous silicon solar cells were also considered. Heterojunction cells can be fabricated that tend to increase the conversion efficiency and produce relatively large open-circuit voltages. Amorphous silicon offers the possibility of low-cost, large-area solar cell arrays.
- Photodetectors are semiconductor devices that convert optical signals into electrical signals. The photoconductor is perhaps the simplest type of photodetector. The change in conductivity of the semiconductor due to the creation of excess electrons and holes by the incident photons is the basis of this device.
- Photodiodes are diodes that have reverse-biased voltages applied. Excess carriers that are created by incident photons in the space-charge region are swept out by the electric field creating a photocurrent. The photocurrent is directly proportional to the incident photon intensity. PIN and avalanche photodiodes are variations of the basic photodiode.
- The photocurrent generated in a phototransistor is multiplied by the transistor gain. However, the time response of the phototransistor may be slower than that of a photodiode because of the Miller effect and Miller capacitance.
- The inverse mechanism of photon absorption in a pn junction is injection electroluminescence. The recombination of excess electrons and holes in a direct bandgap semiconductor can result in the emission of photons.
- The light emitting diodes (LEDs) are the class of pn junction diodes whose photon output is a result of spontaneous recombinations of excess electrons and holes. A fairly wide bandwidth in the output signal, on the order of 30 nm, is a result of the spontaneous process.
- The output of a laser diode is the result of stimulated emission. An optical cavity, or Fabry-Perot resonator, is used in conjunction with a diode so that the photon output is in phase, or coherent. Multilayered heterojunction structures can be fabricated to improve the laser diode characteristics.

GLOSSARY OF IMPORTANT TERMS

absorption coefficient The relative number of photons absorbed per unit distance in a semi-conductor and denoted by the parameter α .

conversion efficiency The ratio of output electrical power to incident optical power in a solar cell.

delayed photocurrent The component of photocurrent in a semiconductor device due to diffusion currents.

external quantum efficiency The ratio of emitted photons to generated photons in a semi-conductor device.

fill factor The ratio $I_m V_m$ to $I_{sc} V_{oc}$, which is a measure of the realizable power from a solar cell. The parameters I_m and V_m are the current and voltage at the maximum power point, respectively, and I_{sc} and V_{oc} are the short-circuit current and open-circuit voltage.

fresnel loss The ratio of reflected to incident photons at an interface due to a change in the index of refraction.

internal quantum efficiency The fraction of diode current that produces luminescence.

LASER diode An acronym for Light Amplification by Stimulated Emission of Radiation; the stimulated emission of photons produced in a forward-biased pn junction in conjunction with an optical cavity.

LED An acronym for Light Emitting Diode; the spontaneous photon emission due to electron–hole recombination in a forward-biased pn junction.

luminescence The general property of light emission.

open-circuit voltage The voltage generated across the open-circuited terminals of a solar cell.

photocurrent The current generated in a semiconductor device due to the flow of excess carriers generated by the absorption of photons.

population inversion The condition whereby the concentration of electrons in one energy state is greater than that in a lower energy state; a nonequilibrium condition.

prompt photocurrent The component of photocurrent generated within the space charge region of a semiconductor device.

radiative recombination The recombination process of electrons and holes that produces a photon, such as the direct band-to-band transition in gallium arsenide.

short-circuit current The current produced in a solar cell when the two terminals are shorted together.

stimulated emission The process whereby an electron is induced by an incident photon to make a transition to a lower energy state, emitting a second photon.

CHECKPOINT

After studying this chapter, the reader should have the ability to:

- Describe the optical absorption process in semiconductors. When is optical absorption essentially zero?
- Describe the basic operation and characteristics of a solar cell, including the shortcircuit current and open-circuit voltage.
- Discuss the factors that contribute to the solar cell conversion efficiency.
- Describe the advantages and disadvantages of an amorphous silicon solar cell.
- Describe the characteristics of a photoconductor, including the concept of the photoconductor gain.
- Discuss the operation and characteristics of a simple pn junction photodiode.
- Discuss the advantages of PIN and avalanche photodiodes compared to the simple pn junction photodiode.
- Discuss the operation and characteristics of a phototransistor.
- Describe the operation of an LED.
- Describe the operation of a laser diode.

REVIEW QUESTIONS

- 1. Sketch the general shape of the optical absorption coefficient in a semiconductor as a function of wavelength. When does the absorption coefficient become zero?
- 2. Sketch the *I–V* characteristic of a pn junction solar cell. Define short-circuit current and open-circuit voltage.

- 3. Discuss how a pn junction solar cell becomes forward biased.
- **4.** Write an expression for the steady-state photocurrent in a simple photoconductor.
- 5. What is the source of prompt photocurrent in a photodiode? Does the prompt photocurrent depend on the reverse-biased voltage? Why or why not.
- **6.** Sketch the cross section of a phototransistor and show the currents that are created by incident photons. Explain how current gain is achieved.
- Explain the basic operation of an LED. State two factors that affect the efficiency of the device.
- **8.** How can different colors be obtained in an LED?
- 9. Discuss the difference between an LED and a laser diode.
- **10.** Discuss the concept of population inversion in a laser diode.

PROBLEMS

Section 14.1 Optical Absorption

- 14.1 Determine the maximum wavelength λ of a light source that can generate electronhole pairs in (a) Si, (b) Ge, (c) GaAs, and (d) InP.
- 14.2 (a) Two sources generate light at wavelengths of $\lambda = 480$ nm and $\lambda = 725$ nm, respectively. What are the corresponding photon energies? (b) Three sources generate light with photon energies of E = 0.87 eV, E = 1.32 eV, and E = 1.90 eV, respectively. What are the corresponding wavelengths?
- 14.3 (a) A sample of GaAs is 1.2 μ m thick. The sample is illuminated with a light source that generates photons with energies of $h\nu = 1.65$ eV. Determine the (i) absorption coefficient and (ii) fraction of energy that is absorbed in the material. (b) Repeat part (a) for a sample of GaAs that is 0.80 μ m thick and is illuminated with photons with energies of $h\nu = 1.90$ eV.
- 14.4 A light source with $h\nu = 1.3$ eV and at a power density of 10^{-2} W/cm² is incident on a thin slab of silicon. The excess minority carrier lifetime is 10^{-6} s. Determine the electron–hole generation rate and the steady-state excess carrier concentration. Neglect surface effects.
- 14.5 An n-type GaAs sample has a minority carrier lifetime of $\tau_p = 2 \times 10^{-7}$ s. Incident photons with energies $h\nu = 1.65$ eV generate an excess carrier concentration of $\delta p = 5 \times 10^{15}$ cm⁻³ at the surface of the semiconductor. (a) Determine the incident power required. (b) At what distance in the semiconductor does the generation rate drop to 10 percent of that at the surface?
- Consider a silicon semiconductor that is illuminated with photons with energies $h\nu = 1.40$ eV. (a) Determine the thickness of the material such that 90 percent of the energy is absorbed. (b) Determine the thickness of the material such that 30 percent of the energy is transmitted through the material.
- 14.7 If the thickness of a GaAs semiconductor is 1 μ m and 50 percent of the incident monochromic photon energy is absorbed, determine the incident photon energy and wavelength.
- *14.8 Consider monochromatic light at an intensity $I_{\nu 0}$ incident on the surface at x=0 of an n-type semiconductor that extends to $x=\infty$. Assume the electric field is zero in the semiconductor and assume a surface recombination velocity, s. Taking into

^{*}Asterisks next to problems indicate problems that are more difficult.