# *pn* JUNCTION DEVICES AND LIGHT EMITTING DIODES

Safa Kasap
Electrical Engineering Department
University of Saskatchewan
Canada

> "We consider alloyed or point contact junctions on *n*-type GaP. Then the light emitted with forward bias has a spectrum which is a comparatively narrow band, the position of the band depending on the impurities present in the GaP."
>
> J.W. Allen and P.E. Gibbons (1959)[1]
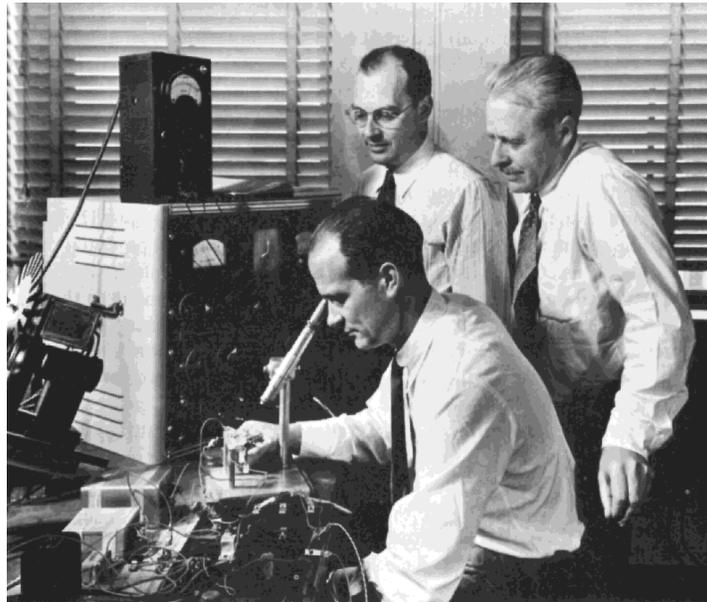
William Shockley, (seated), John Bardeen (left) and Walter Brattain (right) invented the transistor at Bell Labs and thereby ushered in a new era of semiconductor devices. The three inventors shared the Nobel prize in 1956. (Courtesy of Bell Laboratories.)

1. *pn* Junction Principles

1.1. Open Circuit

Consider what happens when one side of a sample of Si is doped *n*-type and the other *p*-type as shown in Figure 1 (a). Assume that there is an *abrupt* discontinuity between the *p* and *n* regions which we call the **metallurgical junction**, *M* in Figure 1 (a), where the fixed (immobile) ionized donors and the free electrons (in the conduction band, CB) in the *n*-region and fixed ionized acceptors and holes (in the valence band, VB) in the *p*-region are also shown.

---

[1] Allen and Gibbons (talking about their invention of the LED) in "Breakdown and light emission in gallium phosphide diodes", *Journal of Electronics*, Vol. 7, No. 6, p. 518, December 1959.

Due to the hole concentration gradient from the *p*-side, where $p = p_{po}$, to the *n*-side where $p = p_{no}$, holes *diffuse* towards the right and enter the *n*-region and recombine with the electrons (majority carriers) in this region. The *n*-side near the junction therefore becomes depleted of majority carriers and therefore has exposed positive donor ions (As$^+$) of concentration N_d. Similarly the electron concentration gradient drives the electrons by diffusion towards the left. Electrons diffusing into the *p*-side recombine with the holes (majority carriers) which exposes negative acceptor ions (B$^-$) of concentration N_a in this region. The regions on both sides of the junction *M* consequently become *depleted* of free carriers in comparison with the bulk *p* and *n* regions far away from the junction. There is therefore a **space charge layer** (SCL) around *M*. Figure 1 (b) shows the SCL, also known as the **depletion region** around *M*. Figure 1 (c) illustrates the hole and electron concentration profiles in which the vertical concentration scale is logarithmic. Note that we must, under equilibrium conditions (*e.g.*, no applied bias, or photoexcitation), have $pn = n_i^2$ everywhere.
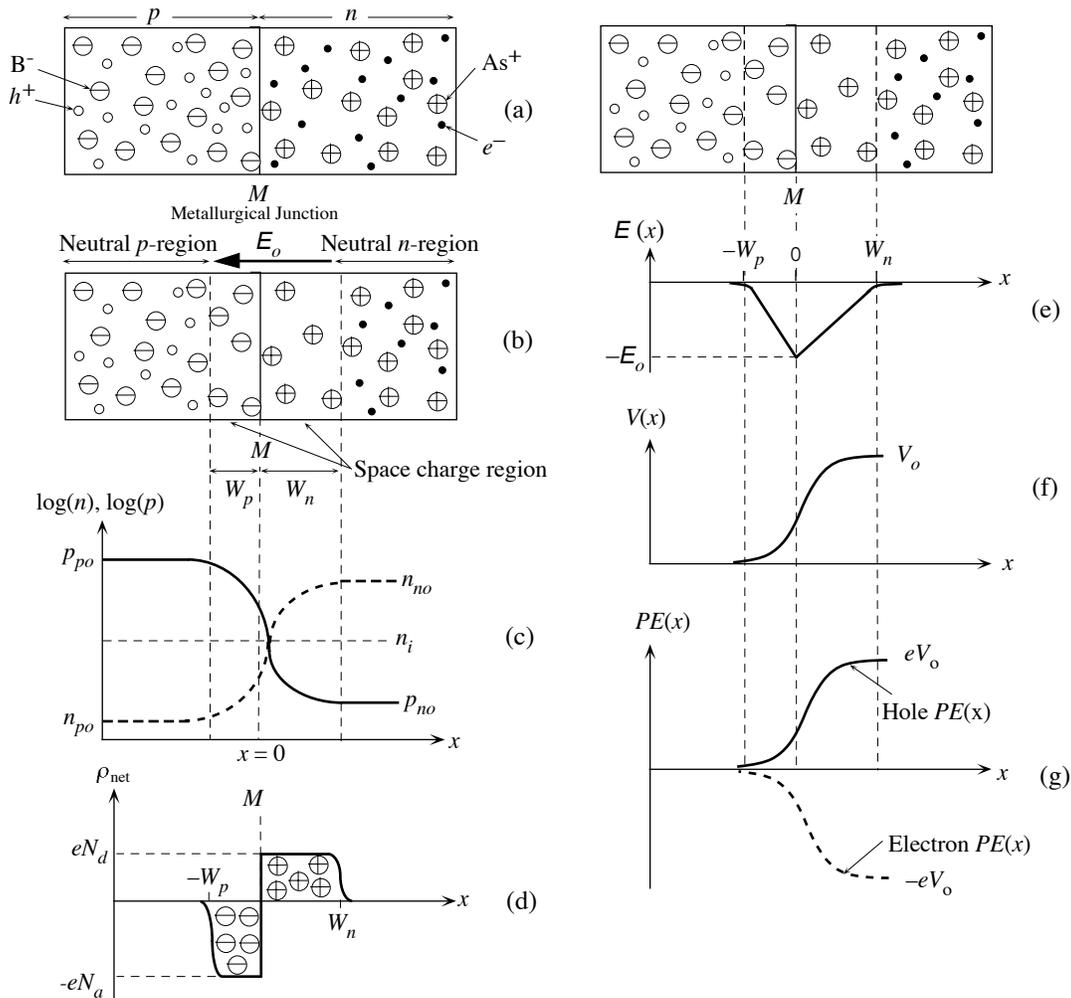
Properties of the *pn* junction.

Figure 1

It is clear that there is an internal electric field E_o from positive ions to negative ions, *i.e.* $-x$ direction, which tries to drift the holes back into the *p*-region and electrons back into the *n*-region. This field drives the holes in the opposite direction to their diffusion. As shown in Figure 1 (b), E_o imposes a drift force on holes in the $-x$ direction, whereas the hole diffusion flux is in the $+x$ direction. A similar situation also applies for electrons with the electric field attempting to drift the electrons against diffusion

from *n* to the *p*-region. It is apparent that as more and more holes diffuse toward the right, and electrons toward the left, the internal field around *M* will increase until eventually an "equilibrium" is reached when the rate of holes diffusing towards the right is just balanced by holes drifting back to the left, driven by the field E_o. The electron diffusion and drift fluxes will also be balanced in equilibrium. For uniformly doped *p* and *n* regions, the net space charge density $\rho_{net}(x)$ across the semiconductor will be as shown in Figure 1 (d). The net space charge density ρ_{net} is negative and equal to $-eN_a$ in the SCL from $x = -W_p$ to $x = 0$ (*M* is at $x = 0$) and then positive and equal to $+ eN_d$ from $x = 0$ to W_n. The total charge on the left hand side must equal that on the right hand side for overall charge neutrality, so that

$$N_a W_p = N_d W_n \tag{1}$$

Figure 1 arbitrarily assumes that the donor concentration is less than the acceptor concentration, $N_d < N_a$. From Equation (1) this implies that $W_n > W_p$, i.e., the depletion region penetrates the *n*-side, lightly doped side, more than the *p*-side, heavily doped side. Indeed, if $N_a \gg N_d$, then the depletion region is almost entirely on the *n*-side. We generally indicate heavily doped regions with the superscript plus sign as p^+. The electric field, $E(x)$ and the net space charge density $\rho_{net}(x)$ at a point are related in electrostatics[2] by, $dE/dx = \rho_{net}(x)/\varepsilon$, where $\varepsilon = \varepsilon_o\varepsilon_r$ is the permittivity of the medium and ε_o and ε_r are the absolute and relative permittivity of the semiconductor material. We can thus integrate $\rho_{net}(x)$ across the diode and thus determine the electric field. The variation of the electric field across the *pn* junction is shown in Figure 1 (e). The negative field means that it is in the $-x$ direction. Note that $E(x)$ reaches a maximum value E_o at *M*.

The potential $V(x)$ at any point x can be found by integrating the electric field since by definition $E = -dV/dx$. Taking the potential on the *p*-side far away from *M* as zero (we have no applied voltage), which is an arbitrary reference level, then $V(x)$ increases in the depletion region towards the *n*-side as indicated in Figure 1 (f). Notice that on the *n*-side the potential reaches V_o which is called the **built-in potential**.

In an abrupt *pn* junction $\rho_{net}(x)$ can simply and approximately be described by step functions as displayed in Figure 1 (d). Using the step form of $\rho_{net}(x)$ in Figure 1 (d) and integrating it gives the electric field and the built-in potential,

$$E_o = -\frac{eN_d W_n}{\varepsilon} = -\frac{eN_a W_p}{\varepsilon} \qquad \text{\textit{Built-in field}} \tag{2}$$

and

$$V_o = -\frac{1}{2}E_o W_o = -\frac{eN_a N_d W_o^2}{2\varepsilon(N_a + N_a)} \qquad \text{\textit{Built-in potential}} \tag{3}$$

where $\varepsilon = \varepsilon_o\varepsilon_r$ and $W_o = W_n + W_p$ is the total width of the depletion region under zero applied voltage. If we know W_o, then W_n or W_p follow readily from Equation (1). Equation (3) is a relationship between the built-in voltage V_o and the depletion region width W_o. If we know V_o we can calculate W_o.

The simplest way to relate V_o to the doping parameters is to make use of Boltzmann statistics. For the system consisting of *p*- and *n*-type semiconductors together, in equilibrium, the Boltzmann statistics[3] demands that the concentrations n_1 and n_2 of carriers at potential energies E_1 and E_2 are related by

$$\frac{n_2}{n_1} = \exp\left[\frac{-(E_2 - E_1)}{k_B T}\right]$$

---

[2]This is called *Gauss's law in point* form and comes from Gauss's law in electrostatics. The integration of the electric field E over a closed surface S is related to the total charge $Q_{enclosed}$ enclosed, $\int E dS = Q_{enclosed}/\varepsilon$., where ε is the permittivity of the medium.

[3]We use Boltzmann statistics, *i.e.* $n(E) \propto \exp(-E/k_B T)$, because the concentration of electrons in the conduction band whether on the *n*-side or *p*-side is never so large that the Pauli exclusion principle becomes important. As long as the carrier concentration in the conduction band is much smaller than N_c we can use Boltzmann statistics.

where E is the potential energy, that is qV, where q is charge and V is voltage. Considering electrons ($q = -e$), we see from Figure 1 (g) that $E = 0$ on the p-side far away from M, where $n = n_{po}$, and $E = -eV_o$ on the n-side away from M, where $n = n_{no}$. Thus

$$n_{po}/n_{no} = \exp(-eV_o/k_BT) \tag{4}$$

This shows that V_o depends on n_{no} and n_{po} and hence on N_d and N_a. The corresponding equation for hole concentrations is similarly,

$$p_{no}/p_{po} = \exp(-eV_o/k_BT) \tag{5}$$

Thus, rearranging Equations (4) and (5) we obtain,

$$V_o = \frac{k_BT}{e}\ln\left(\frac{n_{no}}{n_{po}}\right) \qquad \text{and} \qquad V_o = \frac{k_BT}{e}\ln\left(\frac{p_{po}}{p_{no}}\right)$$

We can now write p_{po} and p_{no} in terms of the dopant concentrations inasmuch as $p_{po} = N_a$, $p_{no} = n_i^2/n_{no} = n_i^2/N_d$, so that V_o becomes

$$V_o = \frac{k_BT}{e}\ln\left(\frac{N_aN_d}{n_i^2}\right) \qquad\qquad \textit{Built-in potential} \tag{6}$$

Clearly V_o has been conveniently related to the dopant and parent material properties via N_a, N_d and n_i^2, which is given by $(N_cN_v)\exp(-E_g/k_BT)$. The built-in voltage (V_o) is the potential across a *pn* junction, going from p to n-type semiconductor, in an open circuit. It is *not* the voltage across the diode which is made up of V_o as well as the contact potentials at the metal to semiconductor junctions at the electrodes. If we add V_o and the contact potentials at the electroded ends we will find zero. Once we know the built-in potential V_o from Equation (6), we can then calculate the width W_o of the depletion region from Equation (3).

1.2. Forward Bias

(a)   (b)
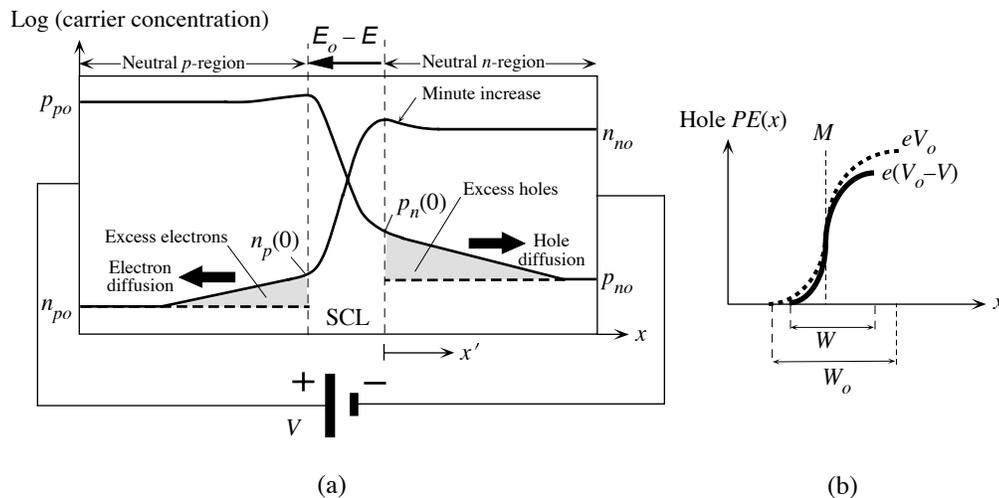
Forward biased *pn* junction and the injection of minority carriers (a) Carrier concentration profiles across the device under forward bias. (b). The hole potential energy with and without an applied bias. W is the width of the SCL with forward bias.

Figure 2

Consider what happens when a battery with a voltage V is connected across a *pn* junction so that the positive terminal of the battery is attached to the *p*-side and the negative terminal to the *n*-side (forward bias). The negative polarity of the supply will reduce the potential barrier V_o by V as shown in Figure 2 (a). The reason is that the bulk regions outside the SCL have high conductivities, due to plenty of majority carriers in the bulk, in comparison with the depletion region in which there are mainly immobile ions. Thus, the applied voltage drops mostly across the depletion width W. Consequently, V directly opposes V_o, and the potential barrier against diffusion is reduced to $(V_o - V)$ as depicted in Figure 2 (b). This has drastic consequences because the probability that a hole in the *p*-side will surmount this potential barrier and diffuse to the *n*-side now becomes proportional to $\exp[-e(V_o - V)/k_BT]$. In other words, the applied voltage effectively reduces the built-in potential and hence the built-in field which acts against diffusion. Consequently, many holes can now diffuse across the depletion region and enter the *n*-side. This results in the **injection of excess minority carriers**, holes into *n*-region. Similarly, excess electrons can now diffuse towards the *p*-side and enter this region and thereby become injected minority carriers.

When holes are injected into the neutral *n*-side, they draw some electrons from the bulk of *n*-side (and hence from the battery) so that there is a small increase in the electron concentration. This small increase in the majority carriers is necessary to balance the hole charges and maintain neutrality in the *n*-side.

The hole concentration, $p_n(0) = p_n(x' = 0)$, just outside the depletion region at $x' = 0$ (x' is measured from W_n) is due to the excess of holes diffusing as a result of the reduction in the built-in potential barrier. This concentration, $p_n(0)$ is determined by the probability of surmounting the new potential energy barrier $e(V_o - V)$,

$$p_n(0) = p_{po} \exp\left[\frac{-e(V_o - V)}{k_BT}\right] \tag{7}$$

This follows directly from the Boltzmann equation, by virtue of the hole potential energy rising by $e(V_o - V)$ from $x = -W_p$ to $x = W_n$, as indicated in Figure 2 (b), and at the same time the hole concentration falling from p_{po} to $p_n(0)$. By dividing Equation (7) by Equation (5) we get the effect of the applied voltage out directly which shows how the voltage V determines the amount of *excess* holes diffusing and arriving at the *n*-region,

$$p_n(0) = p_{no} \exp\left(\frac{eV}{k_BT}\right) \qquad \text{\textit{Law of the junction}} \tag{8}$$

which is called the **law of the junction**. Equation (8) describes the effect of the applied voltage V on the injected minority carrier concentration just outside the depletion region, $p_n(0)$. Obviously, with no applied voltage, $V = 0$ and $p_n(0) = p_{no}$ which is exactly what we expect.

Injected holes diffuse in the *n*-region and eventually recombine with electrons in this region; there are many electrons in the *n*-side. Those electrons lost by recombination are readily replenished by the negative terminal of the battery connected to this side. The current due to holes diffusing in the *n*-region can be sustained because more holes can be supplied by the *p*-region which itself can be replenished by the positive terminal of the battery.

Electrons are similarly injected from the *n*-side to the *p*-side. The electron concentration $n_p(0)$ just outside the depletion region at $x = -W_p$ is given by the equivalent of Equation (8) for electrons, *i.e.*

$$n_p(0) = n_{po} \exp\left(\frac{eV}{k_BT}\right) \qquad \text{\textit{Law of the junction}} \tag{9}$$

In the *p*-region, the injected electrons diffuse toward the positive terminal looking to be collected. As they diffuse they recombine with some of the many holes in this region. Those holes lost by recombination can be readily replenished by the positive terminal of the battery connected to this side. The

current due to the diffusion of electrons in the *p*-side can be maintained by the supply of electrons from the *n*-side which itself can be replenished by the negative terminal of the battery. It is apparent that an electric current can be maintained through a *pn* junction under forward bias, and that the current flow seems to be due to the *diffusion of minority carriers*. There is, however, some drift of majority carriers as well.

If the lengths of the *p* and *n*-regions are longer than the minority carrier diffusion lengths then we will be justified to expect the hole concentration $p_n(x')$ profile on the *n*-side to fall exponentially towards the thermal equilibrium value, p_{no}, as depicted in Figure 2 (a). If $\Delta p_n(x') = p_n(x') - p_{no}$ is the **excess carrier distribution**, then

$$\Delta p_n(x') = \Delta p_n(0)\exp(-x'/L_h) \tag{10}$$

where L_h is the **hole diffusion length** defined by $L_h = \sqrt{D_h \tau_h}$, in which D_h is the diffusion coefficient of holes and τ_h is the mean hole recombination lifetime (minority carrier lifetime) in the *n*-region. The diffusion length is the average distance diffused by a minority carrier before it disappears by recombination. Equation (10) can be derived by rigorous means but it is stated here as a reasonable result. The rate of recombination of injected holes at any point x' in the neutral *n*-region is proportional to the excess hole concentration at that point x'. In the steady state, this recombination rate at x' is just balanced by the rate of holes brought to x' by diffusion. This is the physical argument that leads to Equation (10).
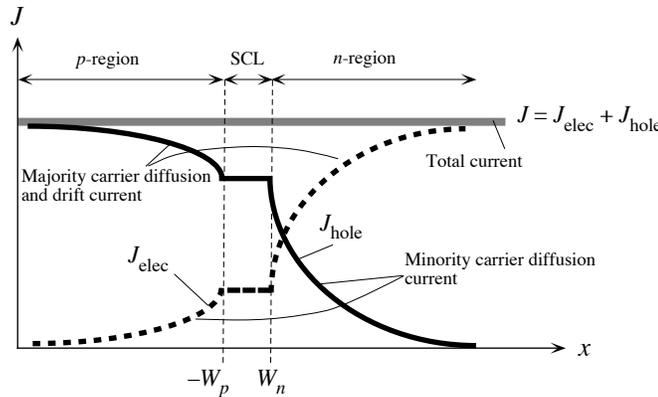
The total current anywhere in the device is constant. Just outside the depletion region it is primarily due to the diffusion of minority carriers and near the contacts it is primarily due to majority carrier drift.

Figure 3

The hole diffusion current density $J_{D,hole}$ is the *hole diffusion flux* multiplied by the hole charge[4],

$$J_{D,hole} = -eD_h \frac{dp_n(x')}{dx'} = -eD_h \frac{d\Delta p_n(x')}{dx'}$$

i.e
$$J_{D,\,hole} = \left(\frac{eD_h}{L_h}\right)\Delta p_n(0)\exp\left(-\frac{x'}{L_h}\right) \tag{11}$$

Although the above equation shows that hole diffusion current depends on location, the total current at any location is the sum of hole and electron contributions, which is independent of x as indicated in Figure 3. The decrease in the minority carrier diffusion current with x' is made up by the increase in the current due to the drift of the majority carriers as schematically shown in Figure 3. The field in the neutral region is not totally zero but a small value, just sufficient to drift the huge number of majority carriers there and maintain a constant current.

---

[4]The hole diffusion flux is $-D_h(dp/dx)$ and the diffusing charge is $+ e$.

We can now use the Law of the junction to substitute for $\Delta p_n(0)$ in Equation (11) in terms of the applied voltage V in Equation (8). Further, we can eliminate p_{no} by $p_{no} = n_i^2/n_{no} = n_i^2/N_d$. Thus, at $x' = 0$, just outside the depletion region, from Equation (11) the hole diffusion current is

$$J_{D,\,\text{hole}} = \left(\frac{eD_h n_i^2}{L_h N_d}\right)\left[\exp\left(\frac{eV}{k_B T}\right) - 1\right] \qquad \text{\textit{Hole diffusion current}}$$

There is a similar expression for the electron diffusion current density $J_{D,\text{elec}}$ in the p-region. We will assume that the electron and hole currents do not change across the depletion region because, in general, the width of this region is narrow (and, for the time being, we neglect the recombination in the SCL). The electron current at $x = -W_p$ is the same as that at $x = W_n$. The total current density is then simply given by $J_{D,\text{hole}} + J_{D,\text{elec}}$, *i.e.*

$$J = \left(\frac{eD_h}{L_h N_d} + \frac{eD_e}{L_e N_a}\right)n_i^2\left[\exp\left(\frac{eV}{k_B T}\right) - 1\right]$$

or
$$J = J_{so}\left[\exp\left(\frac{eV}{k_B T}\right) - 1\right] \qquad \text{\textit{Shockley diode equation}} \qquad (12)$$

This is the familiar diode equation with $J_{so} = [(eD_h/L_h N_d) + (eD_e/L_e N_a)]n_i^2$. It is frequently called the **Shockley equation.** It represents the *diffusion of minority carriers* in the neutral regions. The constant J_{so} depends not only on the doping, N_d and N_a, but also on the material via n_i, D_h, D_e, L_h and L_e. It is known as the **reverse saturation current density**, because if we apply a reverse bias $V = -V_r$ greater than the thermal voltage $k_B T/e$ (25 mV), Equation (12) becomes $J = -J_{so}$.

So far we have assumed that, under a forward bias, the minority carriers diffusing and recombining in the neutral regions are supplied by the external current. However, some of the minority carriers will recombine in the depletion region. The external current must therefore also supply the carriers lost in the recombination process in the SCL. Consider for simplicity a symmetrical *pn* junction as in Figure 4 under forward bias. At the metallurgical junction at the center *C*, the hole and electron concentrations are p_M and n_M and are equal. We can find the SCL recombination current by considering electrons recombining in the *p*-side in W_p and holes recombining in the *n*-side in W_n as shown by the shaded areas *ABC* and *BCD* respectively in Figure 4. Suppose that the **mean hole recombination time** in W_n is τ_h and **mean electron recombination time** in W_p is τ_e. The rate at which the electrons in *ABC* are recombining is the area *ABC* (nearly all injected electrons) divided by τ_e. The electrons are replenished by the diode current. Similarly, the rate at which holes in *BCD* are recombining is the area *BCD* divided by τ_h. Thus, the recombination current density is

$$J_{\text{recom}} = \frac{eABC}{\tau_e} + \frac{eBCD}{\tau_h}$$

We can evaluate the areas *ABC* and *BCD* by taking them as triangles, $ABC \approx (\frac{1}{2})W_p n_M$ *etc.* so that

$$J_{\text{recom}} \approx \frac{e\frac{1}{2}W_p n_M}{\tau_e} + \frac{e\frac{1}{2}W_n p_M}{\tau_h}$$
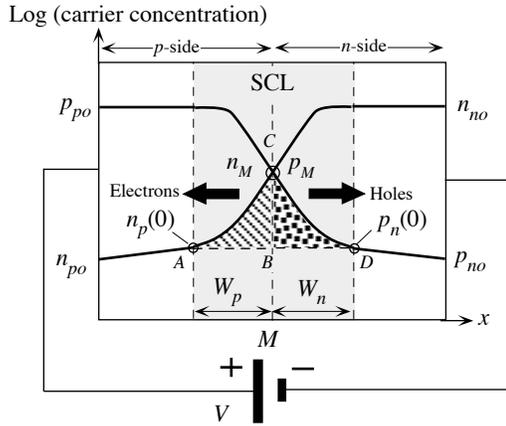
Forward biased *pn* junction and the injection of carriers and their recombination in the SCL.

Figure 4

Under steady state and equilibrium conditions, assuming a nondegenerate semiconductor, we can use Boltzmann statistics to relate these concentrations to the potential energy. At *A,* the potential is zero and at *M* it is $\frac{1}{2}e(V_o - V)$ so that

$$\frac{p_M}{p_{po}} = \exp\left[-\frac{e(V_o - V)}{2k_BT}\right]$$

Since V_o depends on dopant concentrations and n_i as in Equation (6) and further $p_{po} = N_a$, we can simplify the above to

$$p_M = n_i \exp\left(\frac{eV}{2k_BT}\right)$$

This means that the recombination current for $V > k_BT/e$ is given by,

$$J_{\text{recom}} = \frac{en_i}{2}\left[\frac{W_p}{\tau_e} + \frac{W_n}{\tau_h}\right]\exp\left(\frac{eV}{2k_BT}\right) \qquad \text{\textit{Recombination current}} \qquad (14)$$

From a better quantitative analysis, the expression for the recombination current can be shown to be[5],

$$J_{recom} = J_{ro}\left[\exp(eV/2k_BT) - 1\right]$$

where J_{ro} is the pre-exponential constant in Equation (14).

Equation (14) is the current that supplies the carriers that recombine in the depletion region. The total current into the diode will supply carriers for minority carrier diffusion in the neutral regions and recombination in the space charge layer so that it will be the sum of Equations (13) and (14). Generally the diode current is written as

$$I = I_o\left[\exp\left(\frac{eV}{\eta k_BT}\right) - 1\right] \qquad \text{\textit{Diode equation}} \qquad (15)$$

where I_o is a constant and η, called the **diode ideality factor**, is 1 for diffusion controlled and 2 for SCL recombination controlled characteristics. Figure 5 shows both the forward and the reverse *I–V* characteristics of a typical *pn* junction.

---

[5]This is generally proved in advanced texts.

$$I = I_o[\exp(eV/\eta k_B T) - 1]$$

Forward and reverse *I-V* characteristics of a *pn* junction (the positive and negative current axes have different scales and hence the discontinuity at the origin).
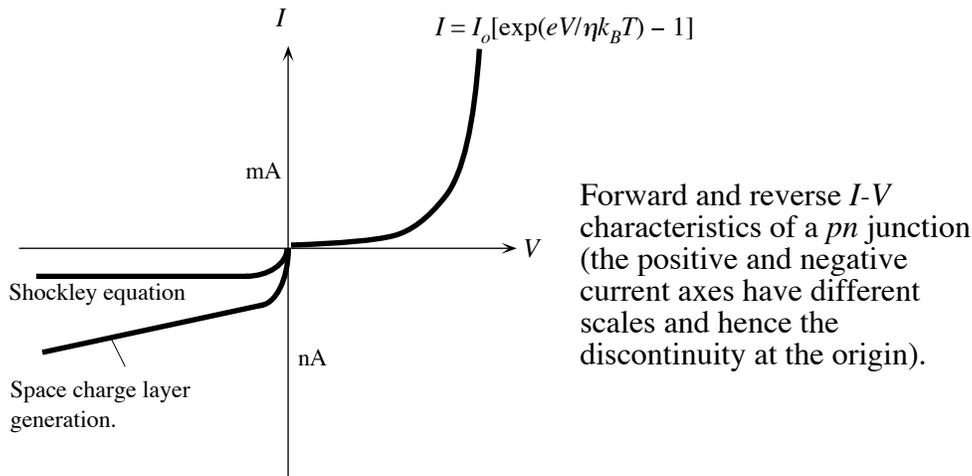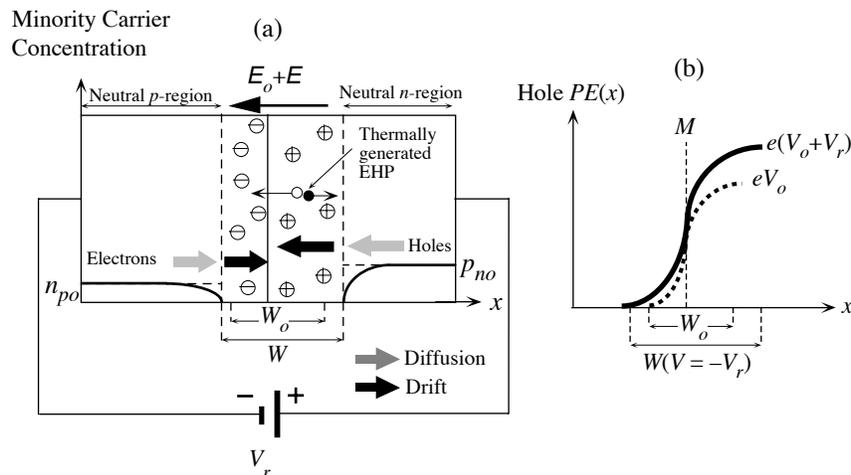
Figure 5

1.3. Reverse Bias

When a *pn* junction is reverse biased, the reverse current is typically very small, as shown in Figure 5. The reverse bias across a *pn* junction is illustrated in Figure 6 (a). The applied voltage drops mainly across the resistive depletion region, which becomes wider. The negative terminal will cause holes in the *p*-side to move away from the SCL, which results in more exposed negative acceptor ions and thus a wider SCL. Similarly, the positive terminal will attract electrons away from the SCL which exposes more positively charged donors. The depletion width on the *n*-side therefore also widens. The movement of electrons in the *n*-region towards the positive battery terminal cannot be sustained because there is no electron supply to this *n*-side. The *p*-side cannot supply electrons to the *n*-side because it has almost none. However, there is a small reverse current due to two causes.

Reverse biased *pn* junction. (a) Minority carrier profiles and the origin of the reverse current. (b) Hole *PE* across the junction under reverse bias.

Figure 6

The applied voltage increases the built-in potential barrier, as depicted in Figure 6 (b). The electric field in the SCL is larger than the built-in internal field E_o. The small number of holes on the *n*-side near

the depletion region become extracted and swept by the field across the SCL over to the *p*-side. This small current can be maintained by the diffusion of holes from the *n*-side bulk to the SCL boundary.

Assume that the reverse bias $V_r > 25$ mV $= k_BT/e$. The hole concentration $p_n(0)$ just outside the SCL is nearly zero by the law of the junction, Equation (8), whereas the hole concentration in the bulk (or near the negative terminal) is the equilibrium concentration p_{no}, which is small. There is therefore a small concentration gradient and hence a small hole diffusion current toward the SCL as shown in Figure 6 (a). Similarly, there is a small electron diffusion current from bulk *p*-side to the SCL. Within the SCL, these carriers are drifted by the field. This minority carrier diffusion current is the Shockley model. The reverse current is given by Equation (13) with a negative voltage which leads to a diode current density of $-J_{so}$ called the **reverse saturation current density**. The value of J_{so} depends only on the material via n_i, μ_h, μ_e, the dopant concentrations, *etc.* but not on the voltage ($V_r > k_BT/e$). Furthermore, as J_{so} depends on n_i^2, it is strongly temperature dependent. In some books, it is stated that thermal generation of minority carriers in the neutral region within a diffusion length to the SCL, the diffusion of these carriers to the SCL, and their subsequent drift through the SCL is the cause of the reverse current. This description, in essence, is identical to the Shockley model.

The thermal generation of electron hole pairs (EHPs) in the space charge region, as shown in Figure 6 (a), can also contribute to the observed reverse current, since the internal field in this layer will separate the electron and hole and drift them towards the neutral regions. This drift will result in an external current in addition to the reverse current due to the diffusion of minority carriers. The theoretical evaluation of SCL generation current involves an in-depth knowledge of the charge carrier generation processes via recombination centers which is discussed in advanced texts. Suppose that τ_g is the **mean time to generate an electron-hole pair** by virtue of the thermal vibrations of the lattice; τ_g is also called the **mean thermal generation time**. Given τ_g, the rate of thermal generation per unit volume must be n_i/τ_g because it takes on average τ_g seconds to create n_i number of EHPs per unit volume. Furthermore, since *WA*, where *A* is the cross sectional area, is the volume of the depletion region, the rate of EHP, or charge carrier, generation is $(AWn_i)/\tau_g$. Both holes and electrons drift in the depletion region and each contributes equally to the current. The observed current density must be $e(Wn_i)/\tau_g$. Therefore the reverse current density component due to thermal generation of electron-hole pairs within the SCL should be given by

$$J_{\text{gen}} = \frac{eWn_i}{\tau_g} \qquad\qquad \textit{EHP thermal generation in SCL} \qquad (16)$$

The reverse bias widens the width *W* of the depletion layer and hence increases J_{gen}. The total reverse current density J_{rev} is the sum of the diffusion and generation components, *i.e.*

$$J_{\text{rev}} = \left(\frac{eD_h}{L_hN_d} + \frac{eD_e}{L_eN_a}\right)n_i^2 + \frac{eWn_i}{\tau_g} \qquad\qquad \textit{Total reverse current} \qquad (17)$$

which is shown schematically in Figure 5. The thermal generation component J_{gen} in Equation (16) increases with reverse bias V_r because the SCL width *W* increases with V_r.

The terms in the reverse current in Equation (17) are predominantly controlled by n_i^2 and n_i. Their relative importance depends not only on the semiconductor properties but also on the temperature since $n_i \sim \exp(-E_g/2k_BT)$. Figure 7 shows the reverse current I_{rev} in dark in a Ge *pn* junction (a photodiode) plotted as $\ln(I_{\text{rev}})$ vs. $1/T$ to highlight the two different processes in Equation (17). The measurements in Figure 7 show that above 238 K, I_{rev} is controlled by n_i^2 because the slope of $\ln(I_{\text{rev}})$ vs. $1/T$ yields an E_g of approximately 0.63 eV, close to the expected E_g of about 0.66 eV in Ge. Below 238 K, I_{rev} is controlled by n_i because the slope of $\ln(I_{\text{rev}})$ vs. $1/T$ is equivalent to an $E_g/2$ of approximately 0.33 eV. In this range, the reverse current is due to EHP generation in the SCL via defects and impurities (recombination centers).
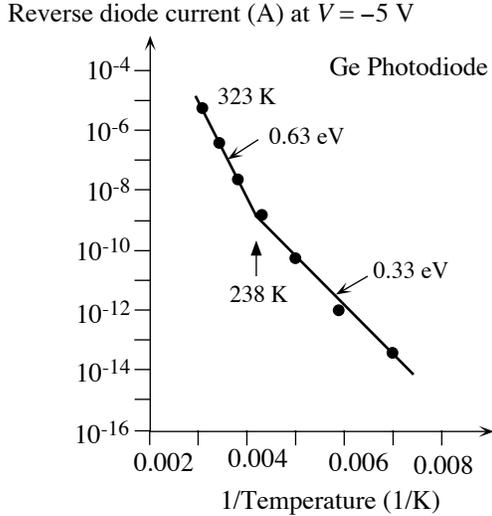
Reverse diode current (A) at $V = -5$ V

Reverse diode current in a Ge *pn* junction as a function of temperature in a $\ln(I_{rev})$ vs $1/T$ plot. Above 238 K, I_{rev} is controlled by n_i^2 and below 238 K it is controlled by n_i. The vertical axis is a logarithmic scale with actual current values. (From D. Scansen and S.O. Kasap, *Cnd. J. Physics.* **70**, 1070-1075, 1992.)

Figure 7

2. *pn* Junction Depletion Layer Capacitance

It is apparent that the depletion region of a *pn* junction has positive and negative charges separated over a distance W similar to a parallel plate capacitor as indicated in Figure 1 (d). If A is the cross sectional area, the stored charge in the depletion region is $+Q = eN_dW_nA$ on the *n*-side and $-Q = -eN_aW_pA$ on the *p*-side. Unlike in the case of a parallel plate capacitor, Q does not depend linearly on the voltage V across the device. It is useful to define an incremental capacitance which relates the incremental charge stored to an incremental voltage change across the *pn* junction. When the voltage V across a *pn* junction changes by dV, to $V + dV$, then W also changes and, as a result, the amount of charge in the depletion region becomes $Q + dQ$. The **depletion layer capacitance** C_{dep} is defined by

$$C_{dep} = \left| \frac{dQ}{dV} \right| \qquad \qquad \textit{Depletion layer capacitance} \qquad (18)$$

If the applied voltage is V then the voltage across the depletion layer W is $V_o - V$ and Equation (3) in this case becomes,

$$W = \left[\frac{2\varepsilon(N_a + N_d)(V_o - V)}{eN_aN_d} \right]^{1/2} \qquad \textit{SCL width and voltage} \qquad (19)$$

The amount of charge (on any one side of the depletion layer) is, $|Q| = eN_dW_nA = eN_aW_pA$, and $W = W_n + W_p$. We can therefore substitute for W in Equation (19) in terms of Q and then differentiate it to obtain dQ/dV. The final result for the depletion capacitance is

$$C_{dep} = \frac{\varepsilon A}{W} = \frac{A}{(V_o - V)^{1/2}} \left[\frac{e\varepsilon(N_aN_d)}{2(N_a + N_d)} \right]^{1/2} \qquad \textit{Depletion layer capacitance} \qquad (20)$$

We should note that C_{dep} is given by the same expressions as that for the parallel plate capacitor, $\varepsilon A/W$, but with W being voltage dependent by virtue of Equation (19). Putting a reverse bias $V = -V_r$ in Equation (20) shows C_{dep} decreases with increasing V_r. Typically C_{dep} under reverse bias is of the order of a few picofarads.

3. Recombination Lifetime and Weak and Strong Injection

Consider recombination in a direct bandgap semiconductor, for example doped GaAs. Recombination involves a direct meeting of an electron-hole pair. Suppose that excess electrons and holes have been injected, as would be in a *pn* junction under forward bias, and that Δn_p is the excess electron concentration and Δp_p is the excess hole concentration in the *neutral p*-side of a GaAs *pn* junction. Injected electron and hole concentrations would be the same to maintain charge neutrality, that is, $\Delta n_p = \Delta p_p$. Thus, at any instant,

$$n_p = n_{po} + \Delta n_p = \text{instantaneous } minority \text{ carrier concentration}$$

and $$p_p = p_{po} + \Delta n_p = \text{instantaneous } majority \text{ carrier concentration}$$

The instantaneous recombination rate will be proportional to both the electron and hole concentrations at that instant , that is $n_p p_p$. Suppose that the thermal generation rate of EHPs is G_{thermal}. The net rate of *change* of Δn_p is

$$\partial \Delta n_p / \partial t = -B n_p p_p + G_{\text{thermal}} \tag{21}$$

where B is called the **direct recombination capture coefficient.** In equilibrium, $\partial \Delta n_p / \partial t = 0$ so that setting Equation (21) to zero and using $n_p = n_{po}$ and $p_p = p_{po}$, where the subscript o refers to thermal equilibrium concentrations, we find $G_{\text{thermal}} = B n_{po} p_{po}$. Thus, the rate of change in Δn_p is

$$\frac{\partial \Delta n_p}{\partial t} = -B\left(n_p p_p - n_{po} p_{po}\right) \qquad \textit{Rate of change due to recombination} \tag{22}$$

In many instances the rate of change $\partial \Delta n_p / \partial t$ is proportional to Δn_p and an **excess minority carrier recombination time (lifetime)** τ_e is defined by

$$\frac{\partial \Delta n_p}{\partial t} = -\frac{\Delta n_p}{\tau_e} \qquad \textit{Recombination time definition} \tag{23}$$

Consider practical cases where the injected excess minority carrier concentration Δn_p is much greater than the actual equilibrium minority carrier concentration n_{po}. There are two conditions on Δn_p corresponding to weak and strong injection based on Δn_p compared with the majority carrier concentration p_{po}.

In **weak injection**, $\Delta n_p << p_{po}$. Then $n_p \approx \Delta n_p$ and $p_p \approx p_{po} + \Delta p_p \approx p_{po} \approx N_a =$ acceptor concentration. Therefore, with these approximations in Equation (22) we obtain,

$$\partial \Delta n_p / \partial t = -B_o N_a \Delta n_p \tag{24}$$

Thus, comparing with Equation (23),

$$\tau_e = 1/BN_a \qquad \textit{Weak injection lifetime} \tag{25}$$

and is constant under weak injection conditions as here.

In **strong injection**, $\Delta n_p >> p_{po}$. Then it is easy to show that with this condition, Equation (22) becomes

$$\partial \Delta n_p / \partial t = -B_o \Delta p_p \Delta n_p = B(\Delta n_p)^2 \tag{26}$$

so that under high level injection conditions the lifetime τ_e is inversely proportional to the injected carrier concentration. When a light emitting diode (LED) is modulated under high injection levels for example, the

lifetime of the minority carriers is therefore not constant, which in turn leads to distortion of the modulated light output.

3.1. Example: A direct bandgap *pn* junction

Consider a symmetrical GaAs *pn* junction which has the following properties. N_a (p-side doping) = N_d (n-side doping) = 10^{17} cm^{-3} (or 10^{23} m^{-3}), $B = 7.21 \times 10^{-16}$ m^3 s^{-1}, cross-sectional area $A = 1$ mm^2. Suppose that the forward voltage across the diode is 1 V. What is the diode current due to minority carrier diffusion at 300 K, assuming direct recombination? If the mean minority carrier lifetime in the depletion region were to be the same as this lifetime (this will not generally be the case), what would be the recombination component of the diode current?

GaAs has an intrinsic concentration (n_i) of 1.8×10^6 cm^{-3} and a relative permittivity (ε_r) of 13.2. The hole drift mobility (μ_h) in the *n*-side is 250 cm^2 V^{-1} s^{-1} and electron drift mobility (μ_e) in the *p*-side is 5000 cm^2 V^{-1} s^{-1} (these are at the doping levels given).

Solution

Assuming weak injection, we can calculate the recombination times τ_e and τ_h for electrons and holes recombining in the neutral *p* and *n*-regions respectively. Using S.I. units throughout, and taking $k_B T/e = 0.02585$ V, and since this is a symmetric device,

$$\tau_h = \tau_e \approx \frac{1}{BN_a} = \frac{1}{(7.21 \times 10^{-16} \ \text{m}^3 \ \text{s}^{-1})(1 \times 10^{23} \ \text{m}^{-3})} = 1.39 \times 10^{-8} \ \text{s}$$

To find the Shockley current we need the diffusion coefficients and lengths. The Einstein relation[6] gives the diffusion coefficients as

$$D_h = \mu_h k_B T/e = (0.02585 \ \text{V})(250 \times 10^{-4} \ \text{m}^2 \ \text{V}^{-1} \ \text{s}^{-1}) = 6.46 \times 10^{-4} \ \text{m}^2 \ \text{s}^{-1}$$

and $$D_e = \mu_e k_B T/e = (0.02585 \ \text{V})(5000 \times 10^{-4} \ \text{m}^2 \ \text{V}^{-1} \ \text{s}^{-1}) = 1.29 \times 10^{-2} \ \text{m}^2 \ \text{s}^{-1}$$

The diffusion lengths are

$$L_h = (D_h \tau_h)^{1/2} = [(6.46 \times 10^{-4} \ \text{m}^2 \ \text{s}^{-1})(1.39 \times 10^{-8} \ \text{s})]^{1/2} = 3.00 \times 10^{-6} \ \text{m}$$

and $$L_e = (D_e \tau_e)^{1/2} = [(1.29 \times 10^{-2} \ \text{m}^2 \ \text{s}^{-1})(1.39 \times 10^{-8} \ \text{s})]^{1/2} = 1.34 \times 10^{-5} \ \text{m}$$

Notice that the electrons diffuse much further in the *p*-side.

The reverse saturation current due to diffusion in the neutral regions is

$$I_{so} = A\left(\frac{D_h}{L_h N_d} + \frac{D_e}{L_e N_a} \right) e n_i^2$$

$$I_{so} = (10^{-6} \ \text{m}^2)\left[\frac{(6.46 \times 10^{-4} \ \text{m}^2 \ \text{s}^{-1})}{(3.00 \times 10^{-6} \ \text{m})(10^{23} \ \text{m}^{-3})} + \frac{(1.29 \times 10^{-2} \ \text{m}^2 \ \text{s}^{-1})}{(1.34 \times 10^{-5} \ \text{m})(10^{23} \ \text{m}^{-3})} \right]$$

$$\times (1.6 \times 10^{-19} \ \text{C})(1.8 \times 10^{12} \ \text{m}^{-3})^2$$

$$I_{so} = 6.13 \times 10^{-21} \ \text{A}$$

Thus the forward diffusion current is

---

[6] *Principles of Electrical Engineering Materials and Devices,* S.O. Kasap (McGraw-Hill Chicago, 1997, 2000), Ch. 5 derives the Einstein relation.

$$I_{\text{diff}} = I_{so} \exp\left(\frac{eV}{k_B T}\right)$$

$$I_{\text{diff}} = (6.13 \times 10^{-21} \ \text{A}) \exp\left[\frac{1.0 \ \text{V}}{0.02585 \ \text{V}}\right] = 3.9 \times 10^{-4} \ \text{A}$$

Recombination component of the current is quite difficult to calculate for two reasons. First we have to determine whether the recombination in the SCL is direct or indirect and then we need to find a mean recombination time if we wish to use the recombination current in Equation (14). Suppose that as a first order we assume that these recombination times are as above.

The built-in voltage V_o is given by

$$V_o = \frac{k_B T}{e} \ln\left(\frac{N_a N_d}{n_i^2}\right) = (0.02585 \ \text{V}) \ln\left[\frac{(10^{23} \ \text{m}^{-3})(10^{23} \ \text{m}^{-3})}{(1.8 \times 10^{12} \ \text{m}^{-3})^2}\right] = 1.28 \ \text{V}$$

Depletion layer width W is

$$W = \left[\frac{2\varepsilon(N_a + N_d)(V_o - V)}{e N_a N_d}\right]^{1/2}$$

$$W = \left[\frac{2(13.2)(8.85 \times 10^{-12} \ \text{F m}^{-1})(10^{23} \ \text{m}^{-3} + 10^{23} \ \text{m}^{-3})(1.28 \ \text{V} - 1 \ \text{V})}{(1.6 \times 10^{-19} \ \text{C})(10^{23} \ \text{m}^{-3})(10^{23} \ \text{m}^{-3})}\right]^{1/2}$$

$W = 9.0 \times 10^{-8}$ m, or 0.090 μm

As this is a symmetric diode, $W_p = W_n = {}^1\!/_2 W$. The pre-exponential I_{ro} is

$$I_{ro} = \frac{A e n_i}{2}\left[\frac{W_p}{\tau_e} + \frac{W_n}{\tau_h}\right] = \frac{A e n_i}{2}\left(\frac{W}{\tau_e}\right)$$

$$I_{ro} \approx \frac{(10^{-6} \ \text{m}^2)(1.6 \times 10^{-19} \ \text{C})(1.8 \times 10^{12} \ \text{m}^{-3})}{2}\left(\frac{9.0 \times 10^{-8} \ \text{m}}{1.39 \times 10^{-8} \ \text{s}}\right)$$

$I_{ro} \approx 9.4 \times 10^{-13}$ A

so that

$$I_{\text{recom}} \approx I_{ro} \exp\left(\frac{eV}{2 k_B T}\right)$$

$$I_{\text{recom}} \approx (9.4 \times 10^{-13} \ \text{A}) \exp\left[\frac{1.0 \ \text{V}}{2(0.02585 \ \text{V})}\right] = 2.4 \times 10^{-4} \ \text{A}$$

This is only a rough estimate on the upper bound for the recombination current based on various assumptions and direct recombination. In the presence of crystal defects or impurities, which encourage indirect recombination, there will be an additional contribution to the recombination current.

4. The *pn* Junction Band Diagram

4.1. Open Circuit

Figure 8 (a) shows the energy band diagram of a *pn* junction under open circuit. If E_{Fp} and E_{Fn} are the Fermi levels in the *p* and *n*-sides, then in equilibrium and in the dark, the Fermi level must be uniform through the two materials as depicted in Figure 8 (a). Far away from the metallurgical junction M, in the

bulk of the *n*-type semiconductor we should still have an *n*-type semiconductor and $E_c - E_{Fn}$ should be the same as in the isolated *n*-type material. Similarly, $E_{Fp} - E_v$ far away from *M* inside the *p*-type material should also be the same as in the isolated *p*-type material. These features are sketched in Figure 8 (a) keeping E_{Fp} and E_{Fn} the same through the whole system and, of course, keeping the bandgap, $E_c - E_v$, the same. Clearly to draw the energy band diagram we have to bend the bands, E_c and E_v, near the junction at *M* because E_c on the *n*-side is close to E_{Fn}, whereas on the *p*-side it is far away from E_{Fp}.

The instant the two semiconductors are brought together to form the junction, electrons diffuse from the *n*-side to the *p*-side and as they do so they deplete the *n*-side near the junction. Thus E_c must move away from E_{Fn} as we move toward *M*, which is exactly what is sketched in Figure 8 (a). Holes diffuse from the *p*-side to the *n*-side and the loss of holes in the *p*-type material near the junction means that E_v moves away from E_{Fp} as we move towards *M* which is also in the figure. Furthermore, as electrons and holes diffuse towards each other, most of them recombine and disappear around *M*, which leads to the formation of a **space charge layer** (SCL) as we saw in Figure 1 (b). The SCL zone around the metallurgical junction has therefore been *depleted* of carriers compared with the bulk.

The electrostatic potential energy (*PE*) of the electron decreases from 0 inside the *p*-region to $-eV_o$ inside the *n*-region as shown in Figure 1 (g). The total energy of the electron must therefore decrease going from the *p* to the *n*-region by an amount eV_o. In other words, the electron in the *n*-side at E_c must overcome a *PE* barrier to go over to E_c in the *p*-side. This *PE* barrier is eV_o where V_o is the built-in potential which we evaluated previously. Band bending around *M* therefore accounts not only for the variation of electron and hole concentrations in this region but also for the effect of the built-in potential (and hence the built-in field as the two are related). The diffusion of electrons from *n*-side to *p*-side is prevented by the built-in *PE* barrier eV_o. This barrier also prevents holes from diffusing from the *p* to the *n*-side.

It should be noted that in the SCL region, the Fermi level is neither close to E_c nor E_v, compared with the bulk or neutral semiconductor regions. This means that both *n* and *p* in this zone are much less than their bulk values n_{no} and p_{po}. The metallurgical junction zone has been depleted of carriers compared with the bulk. Any applied voltage must therefore drop across the SCL.

4.2. Forward and Reverse Bias

When the *pn* junction is forward biased, majority of the applied voltage drops across the depletion region so that the applied voltage is in opposition to the built-in potential, V_o. Figure 8 (b) shows the effect of forward bias which is to reduce the *PE* barrier from eV_o to $e(V_o - V)$. The electrons at E_c in the *n*-side can now readily overcome the *PE* barrier and diffuse to the *p*-side. The diffusing electrons from the *n*-side can be easily replenished by the negative terminal of the battery connected to this side. Similarly holes can now diffuse from *p* to *n*-side. The positive terminal of the battery can replenish those holes diffusing away from the *p*-side. There is therefore a current flow through the junction and around the circuit.
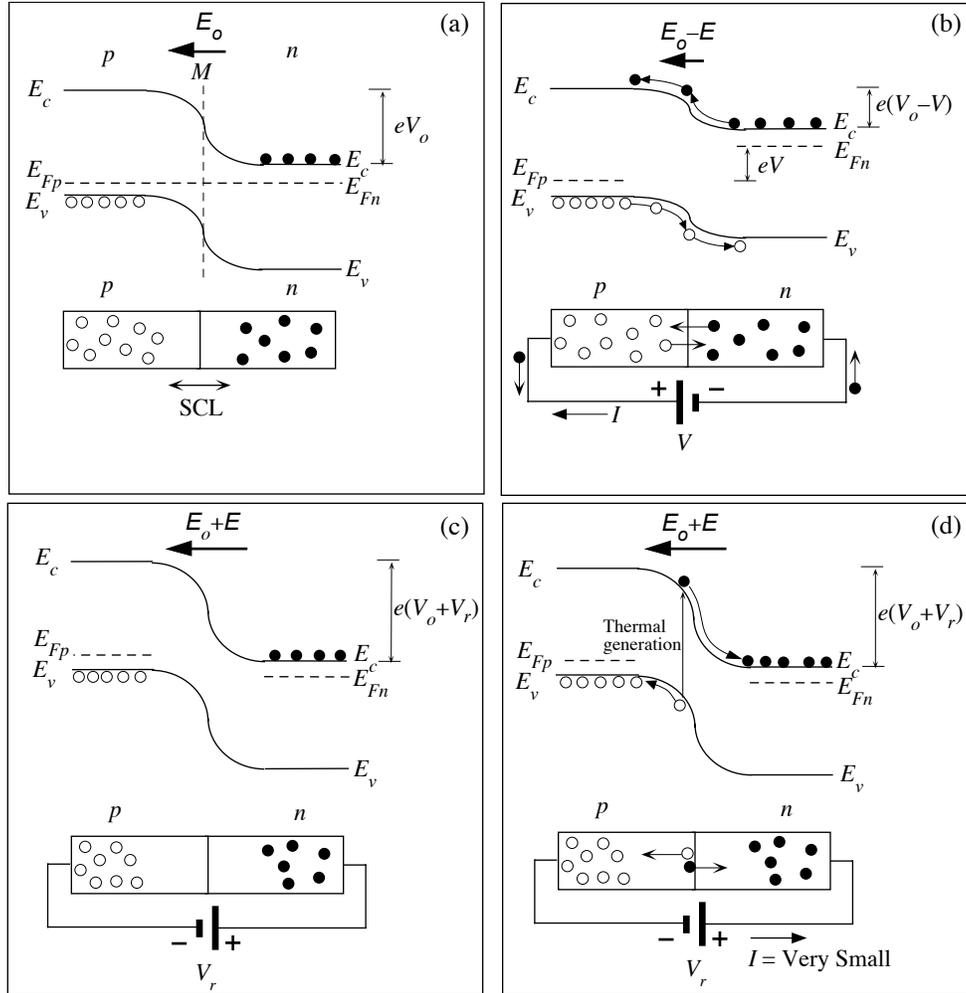
Energy band diagrams for a *pn* junction under (a) open circuit, (b) forward bias and (c) reverse bias conditions. (d) Thermal generation of electron hole pairs in the depletion region results in a small reverse current.

Figure 8

The probability that an electron at E_c in the *n*-side overcomes the new *PE* barrier and diffuses to E_c in the *p*-side is now proportional to the Boltzmann factor $\exp[-e(V_o - V)/k_BT]$. The latter increases enormously even for small forward voltages. Thus, a major diffusion of electrons from *n* to *p*-side takes place. Similar ideas also apply to holes at E_v in the *p*-side which also overcome the barrier $e(V_o - V)$ to diffuse into the *n*-side. Since the forward current is due to the number of electrons and holes overcoming the barrier it is also proportional to $\exp[-e(V_o - V)/k_BT]$ or $\exp(eV/k_BT)$.

When a reverse bias, $V = -V_r$, is applied to the *pn* junction the voltage again drops across the SCL. In this case, however, V_r adds to the built-in potential V_o so that the *PE* barrier becomes $e(V_o + V_r)$, as shown in Figure 8 (c). The field in the SCL at *M* increases to $E_o + E$ where E is the applied field (it is not simply V/W). There is hardly any reverse current because if an electron were to leave the *n*-side to travel to the positive terminal, it cannot be replenished from the *p*-side (virtually no electrons on the *p*-side). There is, however, a small reverse current arising from thermal generation of electron hole pairs (EHPs) in the SCL and thermal generation of minority carriers within a diffusion length to the SCL. When an EHP is thermally generated in the SCL, as shown in Figure 8 (d), the field here separates the pair. The

electron falls the *PE* hill, down E_c, to the *n*-side to be collected by the battery. Similarly the hole falls down its own *PE* hill (energy increases downward for holes) to make it to the *p*-side. The process of falling down a *PE* hill is the same process as being driven by a field, in this case by $E_o + E$. A thermally generated hole in the *n*-side within a diffusion length can diffuse to the SCL and then drift across the SCL, which would result in a reverse current. Similarly a thermally generated electron in the *p*-side within a diffusion length to the SCL can also lead to a reverse current. All these reverse current components are very small compared with the forward current as they depend on the rate of thermal generation.

5. Light Emitting Diodes

5.1. Principles

A light emitting diode (LED) is essentially a *pn* junction diode typically made from a direct bandgap semiconductor, for example GaAs, in which the electron-hole pair (EHP) recombination results in the emission of a photon. The emitted photon energy is therefore approximately equal to the bandgap energy, $hv \approx E_g$. Figure 9 (a) shows the energy band diagram of an unbiased pn^+ junction device in which the *n*-side is more heavily doped than the *p*-side. The band diagram is drawn to keep the Fermi level uniform through the device which is a requirement of equilibrium with no applied bias. The depletion region in a pn^+ device extends mainly into the *p*-side. There is a potential energy (*PE*) barrier eV_o from E_c on the *n*-side to E_c on the *p*-side, that is $\Delta E_c = eV_o$, where V_o is the *built-in voltage*. The higher concentration of conduction (free) electrons in the *n*-side encourages the diffusion of conduction electrons from the *n* to the *p*-side. This net electron diffusion, however, is prevented by the electron *PE* barrier eV_o.
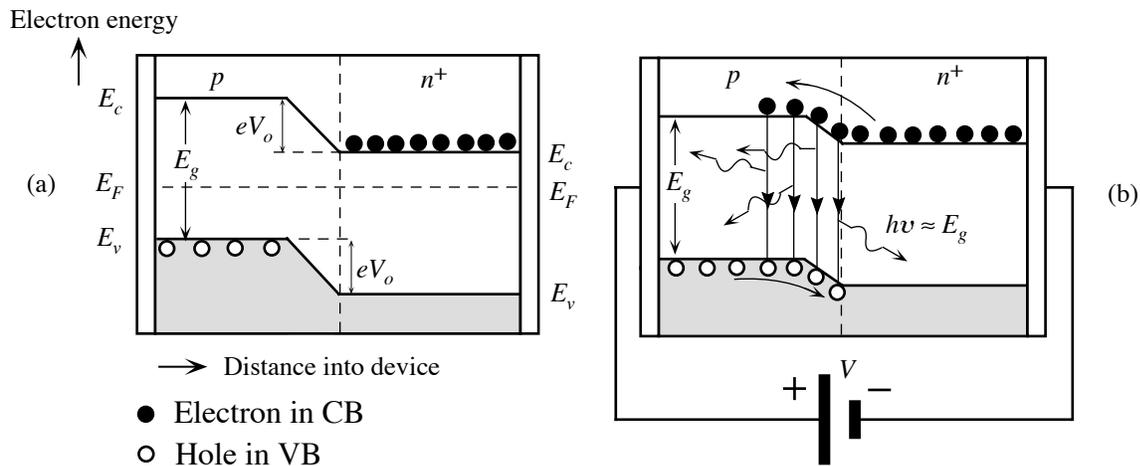
(a) The energy band diagram of a pn^+ (heavily *n*-type doped) junction without any bias. Built-in potential V_O prevents electrons from diffusing from n^+ to *p* side. (b) The applied bias reduces V_o and thereby allows electrons to diffuse or be injected into the *p*-side. Recombination around the junction and within the diffusion length of the electrons in the *p*-side leads to photon emission.
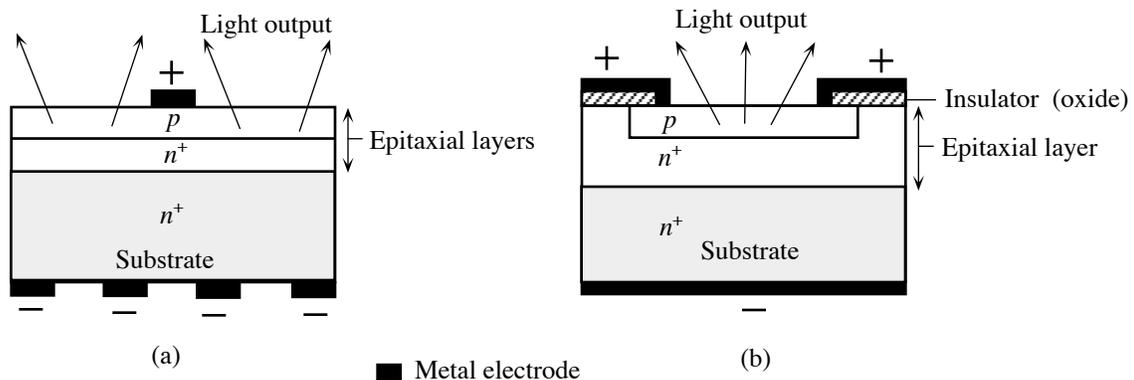**Figure 9**

　　　As soon as a forward bias *V* is applied, this voltage drops across the depletion region since this is the most resistive part of the device. Consequently, the built-in potential V_o is reduced to $V_o - V$ which then allows the electrons from the n^+ side to diffuse, or become injected, into the *p*-side as depicted in Figure 9 (b). The hole injection component from *p* into the n^+ side is much smaller than the electron injection component from the n^+ to *p*-side. The recombination of injected electrons in the depletion region as well as in the neutral *p*-side results in the spontaneous emission of photons. Recombination primarily occurs within the depletion region and within a volume extending over the diffusion length L_e of the

electrons in the *p*-side. This recombination zone is frequently called the active region[7]. The phenomenon of light emission from EHP recombination as a result of minority carrier injection as in this case is called injection electroluminescence. Because of the statistical nature of the recombination process between electrons and holes, the emitted photons are in random directions; they result from spontaneous emission processes in contrast to stimulated emission. The LED structure has to be such that the emitted photons can escape the device without being reabsorbed by the semiconductor material. This means the *p*-side has to be sufficiently narrow or we have to use *heterostructure* devices as discussed below.

5.2. Device Structures

In their simplest technological form, LEDs are typically fabricated by *epitaxially* growing doped semiconductor layers on a suitable substrate (*e.g.* GaAs or GaP) as depicted in Figure 10 (a). This type of planar *pn* junction is formed by the epitaxial growth of first the *n*-layer and then the *p*-layer. The substrate is essentially a mechanical support for the *pn* junction device (the layers) and can be of different crystal. The *p*-side is on the surface from which light is emitted and is therefore made narrow (a few microns) to allow the photons to escape without being reabsorbed. To ensure that most of the recombination takes place in the *p*-side, the *n*-side is heavily doped (n^+). Those photons that are emitted towards the *n*-side become either absorbed or reflected back at the substrate interface depending on the substrate thickness and the exact structure of the LED. The use of a segmented back electrode as in Figure 10 (a) will encourage reflections from the semiconductor-air interface. It is also possible to form the *p*-side by diffusing dopants into the epitaxial n^+-layer which is a diffused junction planar LED as illustrated in Figure 10 (b).

A schematic illustration of typical planar surface emitting LED devices. (a) *p*-layer grown epitaxially on an n^+ substrate. (b) First n^+ is epitaxially grown and then *p* region is formed by dopant diffusion into the epitaxial layer.

Figure 10

If the epitaxial layer and the substrate crystals have different crystal lattice parameters, then there is a lattice mismatch between the two crystal structures. This causes lattice strain in the LED layer and hence leads to crystal defects. Such crystal defects encourage radiationless EHP recombinations. That is, a defect acts as a recombination center. Such defects are reduced by lattice matching the LED epitaxial layer to the substrate crystal. It is therefore important to lattice-match the LED layer to the substrate crystal. For example, one of the AlGaAs alloys is a direct bandgap semiconductor that has a bandgap in the red-emission region. It can be grown on GaAs substrates with excellent lattice match which results in high efficiency LED devices.

Figure 10 (a) and (b) both show the planar *pn* junction based simple LED structures. Not all light rays reaching the semiconductor-air interface, however, can escape because of total internal reflection

---

[7]The "active region" term is probably more appropriate for laser diodes in which there is a photon amplification in this region.

(TIR). Those rays with angles of incidence greater than the critical angle θ_c become reflected as depicted in Figure 11 (a). For the GaAs-air interface, for example, θ_c is only 16° which means that much of the light suffers TIR. It is possible to shape the surface of the semiconductor into a dome, or hemisphere, so that light rays strike the surface at angles less than θ_c and therefore do not experience TIR as illustrated in Figure 11 (b). The main drawback, however, is the additional difficult process in fabricating such domed LEDs and the associated increase in expense. An inexpensive and common procedure that reduces TIR is the encapsulation of the semiconductor junction within a transparent plastic medium (an epoxy) which has a higher refractive index than air and also has a domed surface on one side of the *pn* junction as shown in Figure 11 (c). Many individual LEDs are sold in similar types of plastic bodies.
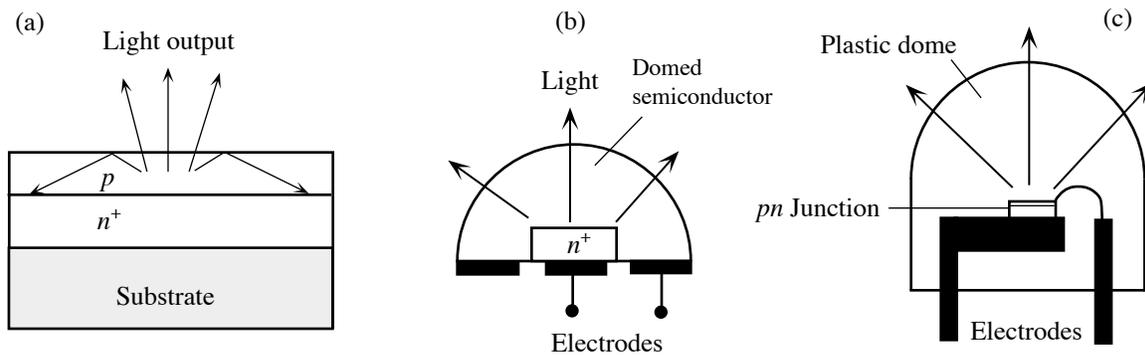
(a) Some light suffers total internal reflection and cannot escape. (b) Internal reflections can be reduced and hence more light can be collected by shaping the semiconductor into a dome so that the angles of incidence at the semiconductor-air surface are smaller than the critical angle. (b) An economic method of allowing more light to escape from the LED is to encapsulate it in a transparent plastic dome.

Figure 11

6. LED Materials

There are various direct bandgap semiconductor materials that can be readily doped to make commercial *pn* junction LEDs which emit radiation in the red and infrared range of wavelengths. An important class of commercial semiconductor materials which cover the visible spectrum is the **III-V ternary alloys** based on alloying GaAs and GaP which are denoted as $GaAs_{1-y}P_y$. In this compound, As and P atoms from group V are distributed randomly at normal As sites in the GaAs crystal structure. When $y < 0.45$, the alloy $GaAs_{1-y}P_y$ is a direct bandgap semiconductor and hence the EHP recombination process is direct which is depicted in Figure 12 (a). The rate of recombination is directly proportional to the product of electron and hole concentrations. The emitted wavelengths range from about 630 nm, red, for $y = 0.45$ ($GaAs_{0.55}P_{0.45}$) to 870 nm for $y = 0$, GaAs.

$GaAs_{1-y}P_y$ alloys (which includes GaP) with $y > 0.45$ are indirect bandgap semiconductors. The EHP recombination processes occur through recombination centers and involve lattice vibrations rather than photon emission. However, if we add **isoelectronic impurities** such as nitrogen (in the same group V as P) into the semiconductor crystal then some of these N atoms substitute for P atoms. Since N and P have the same valency, N atoms substituting for P atoms form the same number of bonds and do not act as donors or acceptors. The electronic cores of N and P are, however, different. The positive nucleus of N is less shielded by electrons compared with that of the P atom. This means that a conduction electron in the neighborhood of a N atom will be attracted and may become trapped at this site. N atoms therefore introduce localized energy levels, or electron traps, E_N, near the conduction band edge as depicted in Figure 12 (b). When a conduction electron is captured at E_N, it can attract a hole (in the valence band) in its vicinity by Coulombic attraction and eventually recombine with it directly and emit a photon.

The emitted photon energy is only slightly less than E_g as E_N is typically close to E_c. As the recombination process depends on N doping, it is not as efficient as direct recombination. Thus, the efficiency of LEDs from N doped indirect bandgap $GaAs_{1-y}P_y$ semiconductors is less than those from direct bandgap semiconductors. Nitrogen doped indirect bandgap $GaAs_{1-y}P_y$ alloys are widely used in inexpensive green, yellow and orange LEDs.

(a) $GaAs_{1-x}P_x$ (b) N doped GaP   (c) Al doped SiC

$x < 0.45$
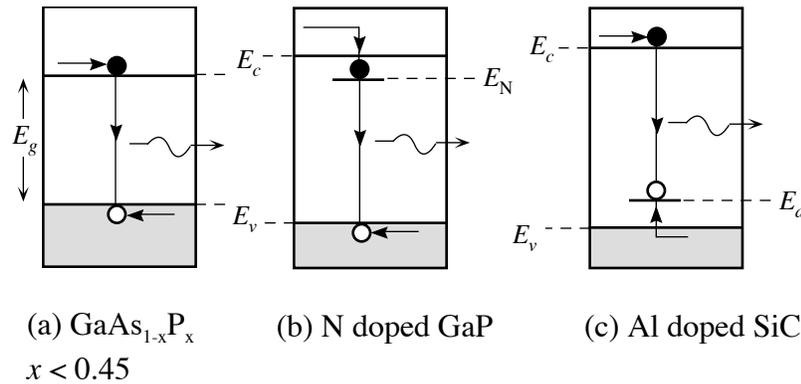
(a) Photon emission in a direct bandgap semiconductor. (b). GaP is an indirect bandgap semiconductor. When doped with nitrogen there is an electron trap at E_N. Direct recombination between a trapped electron at E_N and a hole emits a photon. (c) In Al doped SiC, EHP recombination is through an acceptor level like E_a.
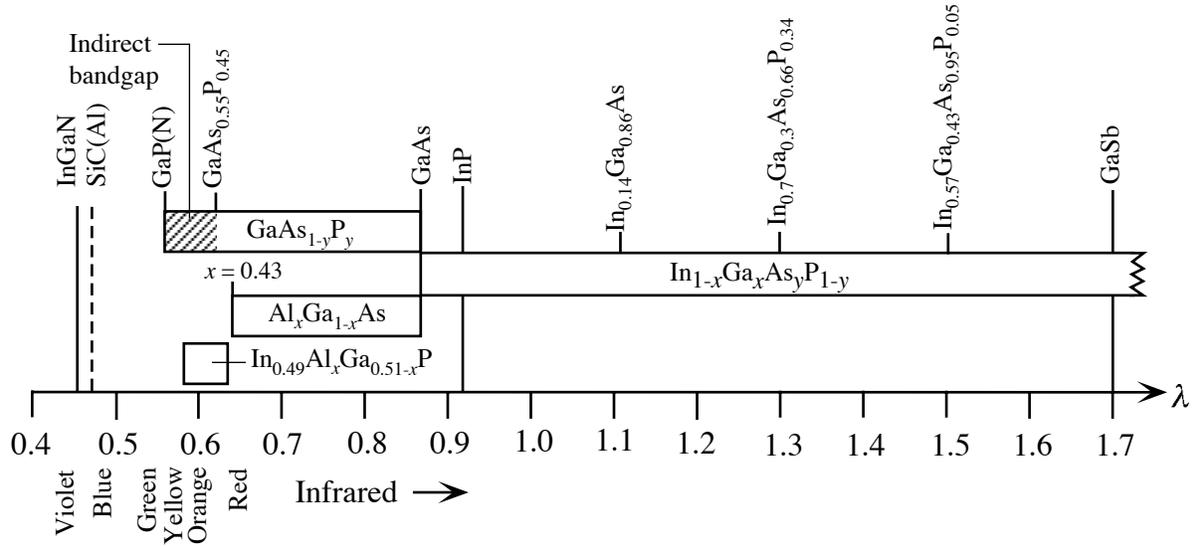
Figure 12

There are two types of blue LED materials. GaN is a direct bandgap semiconductor with an E_g of 3.4 eV. The blue GaN LEDs actually use the GaN alloy InGaN which has a bandgap of about 2.7 eV which corresponds to blue emission. The less efficient type is the Al doped *silicon carbide* (SiC), which is an indirect bandgap semiconductor. The acceptor type localized energy level captures a hole from the valence band and a conduction electron then recombines with this hole to emit a photon as schematically shown in Figure 12 (c). As the recombination process is not direct and therefore not as efficient, the brightness of blue SiC LEDs is limited. Recently there has been considerable progress made towards more efficient blue LEDs using direct bandgap compound semiconductors such as II-VI semiconductors, for example ZnSe (Zn and Se are in groups II and VI in the Periodic Table). The main problem in using II-VI compounds is the current technological difficulty in appropriately doping these semiconductors to fabricate efficient *pn* junctions.

There are various commercially important direct bandgap semiconductor materials that emit in the red and infrared wavelengths which are typically **ternary** (containing three elements) and **quarternary** (four elements) alloys based on III and V elements, so called **III-V alloys**. For example, GaAs with a bandgap of about 1.43 eV emits radiation at around 870 nm in the infrared. But ternary alloys based on $Al_{1-x}Ga_xAs$ where $x < 0.43$ are direct bandgap semiconductors. The composition can be varied to adjust the bandgap and hence the emitted radiation from about 640 - 870 nm, from deep red light to infrared.

In-Ga-Al-P is a quarternary III-V alloy (In, Ga, Al from III and P from V) that has a direct bandgap variation with composition over the visible range. It can be lattice matched to GaAs substrates when in the composition range $In_{0.49}Al_{0.17}Ga_{0.34}P$ to $In_{0.49}Al_{0.058}Ga_{0.452}P$. Recent high intensity LEDs have been based on this material which is likely to eventually dominate the high-intensity visible LED range.

The bandgap of quarternary alloys $In_{1-x}Ga_xAs_{1-y}P_y$ can be varied with composition (x and y) to span wavelengths from 870 nm (GaAs) to 3.5 μm (InAs) which includes the optical communication

wavelengths of 1.3 and 1.55 μm. Figure 13 summarizes some typical wavelengths that can be emitted for a few selected semiconductor materials over the range from 0.4 to 1.7 μm.

Free space wavelength coverage by different LED materials from the visible spectrum to the infrared including wavelengths used in optical communications. Hatched region and dashed lines are indirect E_g materials.

Figure 13

Table 1

Selected LED semiconductor materials. Optical communication channels are at 850 nm (local network) and at 1.3 and 1.55 μm (long distance). D = direct, I = Indirect bandgap. DH = Double heterostructure. $\eta_{external}$ is typical and may vary substantially depending on the device structure.

Semiconductor	Substrate	D or I	λ (nm)	$\eta_{external}$ (%)	Comment
GaAs	GaAs	D	870 - 900	10	Infrared LEDs
Al_xGa_{1-x} As (0< x < 0.4)	GaAs	D	640 - 870	5 - 20	Red to IR LEDs. DH
$In_{1-x}Ga_xAs_yP_{1-y}$ ($y \approx 2.20x$, 0 < x < 0.47)	InP	D	1 - 1.6 μm	> 10	LEDs in communications
InGaN alloys	GaN or SiC Saphire	D	430 - 460	2	Blue LED
			500 - 530	3	Green LED
SiC	Si; SiC	I	460 - 470	0.02	Blue LED. Low efficiency
$In_{0.49}Al_x\,Ga_{0.51-x}$ P	GaAs	D	590 - 630	1 - 10	Amber, green red LEDs
$GaAs_{1-y}P_y$ (y < 0.45)	GaAs	D	630 - 870	< 1	Red - IR
$GaAs_{1-y}P_y$ (y > 0.45) (N or Zn, O doping)	GaP	I	560 - 700	< 1	Red, orange, yellow LEDs
GaP (Zn-O)	GaP	I	700	2 - 3	Red LED
GaP (N)	GaP	I	565	< 1	Green LED

     External efficiency η_{external} of an LED quantifies the efficiency of conversion of electrical energy into an emitted external optical energy. It incorporates the "internal" efficiency of the radiative recombination process and the subsequent efficiency of photon extraction from the device. The input of electrical power into an LED is simply the diode current and diode voltage product (*IV*). If P_{out} is the optical power emitted by the device then

$$\eta_{\text{external}} = \frac{P_{out}(\text{Optical})}{IV} \times 100\% \qquad\qquad \textit{External quantum efficiency} \qquad (1)$$

and some typical values are listed in Table 1. For indirect bandgap semiconductors η_{external} is generally less than 1%, where as for direct bandgap semiconductors with the right device structure, η_{external} can be substantial.

7. Heterojunction High Intensity LEDs

A junction (such as a *pn* junction) between two differently doped semiconductors that are of the same material (same bandgap E_g) is called a **homojunction**. A junction between two different bandgap semiconductors is called a **heterojunction.** A semiconductor device structure that has junctions between different bandgap materials is called a **heterostructure device** (HD). The refractive index of a semiconductor material depends on its bandgap. A wider bandgap semiconductor has a lower refractive index. This means that by constructing LEDs from heterostructures, we can engineer a dielectric waveguide within the device and thereby channel photons out from the recombination region.
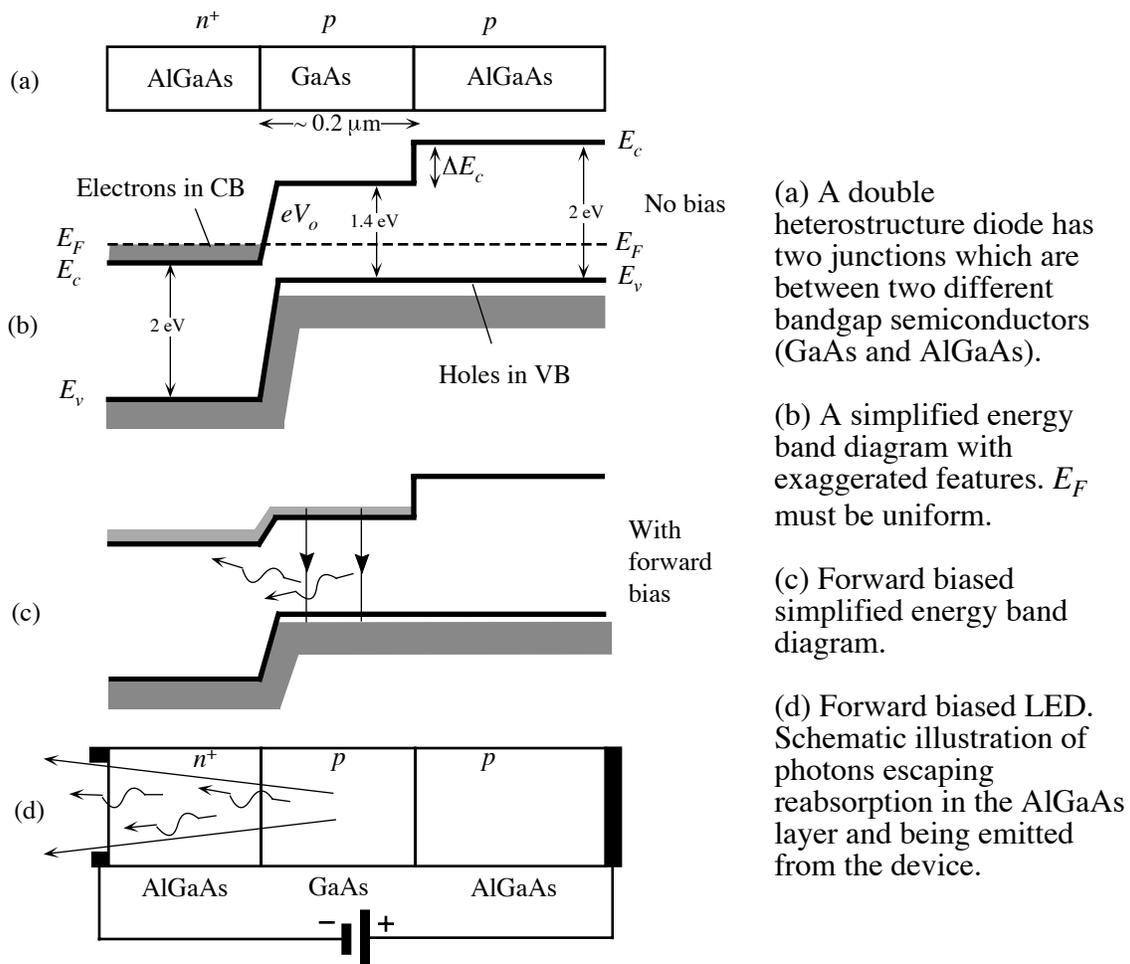

(a) A double heterostructure diode has two junctions which are between two different bandgap semiconductors (GaAs and AlGaAs).

(b) A simplified energy band diagram with exaggerated features. E_F must be uniform.

(c) Forward biased simplified energy band diagram.

(d) Forward biased LED. Schematic illustration of photons escaping reabsorption in the AlGaAs layer and being emitted from the device.

Figure 14

The homojunction LED shown in Figure 11 (a) has two drawbacks. The *p*-region must be narrow to allow the photons to escape without much reabsorption. When the *p*-side is narrow, some of the injected electrons in the *p*-side reach the surface by diffusion and recombine through crystal defects near the surface. This radiationless recombination process decreases the light output. In addition, if the recombination occurs over a relatively large volume (or distance), due to long electron diffusion lengths, then the chances of reabsorption of emitted photons becomes higher; the amount of reabsorption increases with the material volume.

LED constructions for increasing the intensity of the output light make use of the double heterostructure (DH) structure. Figure 14 (a) shows a **double-heterostructure** (DH) device based on two junctions between different semiconductor materials with different bandgaps. In this case the semiconductors are AlGaAs with $E_g \approx 2$ eV and GaAs with $E_g \approx 1.4$ eV. The double heterostructure in Figure 14 (a) has an n^+p heterojunction between n^+-AlGaAs and *p*-GaAs. There is another heterojunction between *p*-GaAs and *p*-AlGaAs. The *p*-GaAs region is a thin layer, typically a fraction of a micron and it is lightly doped.

The simplified energy band diagram for the whole device in the absence of an applied voltage is shown in Figure 14 (b). The Fermi level E_F is continuous through the whole structure. There is a potential energy barrier eV_o for electrons in the CB of n^+-AlGaAs against diffusion into *p*-GaAs. There is a bandgap change at the junction between *p*-GaAs and *p*-AlGaAs which results in a step change ΔE_c in E_c between the two bands of *p*-GaAs and *p*-AlGaAs. This, ΔE_c, is effectively a potential energy barrier that prevents any electrons in the CB in *p*-GaAs passing to the CB of *p*-AlGaAs.

When a forward bias is applied, majority of this voltage drops between the n^+-AlGaAs and *p*-GaAs and reduces the potential energy barrier eV_o, just as in the normal *pn* junction diode. This allows electrons in the CB of n^+-AlGaAs to be injected (by diffusion) into *p*-GaAs as shown in Figure 14 (c). These electrons, however, are *confined* to the CB of *p*-GaAs since there is a barrier ΔE_c between *p*-GaAs and *p*-AlGaAs. The wide bandgap AlGaAs layers therefore act as **confining layers** that restrict injected electrons to the *p*-GaAs layer. The recombination of injected electrons and the holes already present in this *p*-GaAs layer results in spontaneous photon emission. Since the bandgap E_g of AlGaAs is greater than GaAs, the emitted photons do not get reabsorbed as they escape the active region and can reach the surface of the device as depicted in Figure 14 (d). Since light is also not absorbed in *p*-AlGaAs, it can be reflected to increase the light output. Another advantage of the AlGaAs/GaAs heterojunction is that there is only a small lattice mismatch between the two crystal structures and hence negligible strain-induced interfacial defects (*e.g.* dislocations) in the device compared with the defects at the surface of the semiconductor in conventional homojunction LED structure shown in Figure 11 (a). The DH LED is much more efficient than the homojunction LED.

8. LED Characteristics

The energy of an emitted photon from an LED is not simply equal to the bandgap energy E_g because electrons in the conduction band are distributed in energy and so are the holes in the valence band. Figure 15 (a) and (b) illustrate the energy band diagram and the energy distributions of electrons and holes in the conduction band (CB) and valence band (VB) respectively. The electron concentration as a function of energy in the CB is given by $g(E)f(E)$ where $g(E)$ is the density of states and $f(E)$ is the Fermi-Dirac function (probability of finding an electron in a state with energy E). The product $g(E)f(E)$ represents the electron concentration per unit energy or the concentration in energy, and is plotted along the horizontal axis in Figure 15 (b). There is a similar energy distribution for holes in the VB.

The electron concentration in the CB as a function of energy is asymmetrical and has a peak at $\frac{1}{2}k_BT$ above E_c. The energy spread of these electrons is typically about $\sim 2k_BT$ from E_c as shown in Figure 15 (b). The hole concentration is similarly spread from E_v in the valence band. We should recall the rate of direct recombination is proportional to both the electron and hole concentrations at the energies involved. The transition, which is identified as 1 in Figure 15 (a), involves the direct recombination of an electron at E_c and a hole at E_v. But the carrier concentrations near the band edges are very small and hence this type of recombination does not occur frequently. The relative intensity of light at this photon energy $h\upsilon_1$ is small

as shown in Figure 15. The transitions that involve the largest electron and hole concentrations occur most frequently. For example, the transition 2 in Figure 15 (a) has the maximum probability as both electron and hole concentrations are largest at these energies as shown in Figure 15 (b). The relative intensity of light corresponding to this transition energy, $h\upsilon_2$, is then maximum, or close to maximum, as indicated in Figure 15 (c)[8]. The transitions marked as 3 in Figure 15 (a) that emit relatively high energy photons, $h\upsilon_3$, involve energetic electrons and holes whose concentrations are small as apparent in Figure 15 (b). Thus the light intensity at these relatively high photon energies is small. The fall in light intensity with photon energy is shown in Figure 15 (c). The relative light intensity vs. photon energy characteristic of the output spectrum is shown in Figure 15 (c) and represents an important LED characteristic. Given the spectrum in Figure 15 (c) we can also obtain the relative light intensity vs. wavelength characteristic as shown in Figure 15 (d), since $\lambda = c/\upsilon$. The **linewidth** of the output spectrum, $\Delta\upsilon$ or $\Delta\lambda$, is defined as width between half-intensity points as defined in Figure 15 (c) and (d).

The wavelength for the peak intensity and the linewidth $\Delta\lambda$ of the spectrum are obviously related to the energy distributions of the electrons and holes in the conduction and valence bands and therefore to the density of states in these bands (and hence to individual semiconductor properties). The photon energy for the peak emission is roughly $E_g + k_BT$ inasmuch as it corresponds to peak-to-peak transitions in the energy distributions of the electrons and holes in Figure 15 (b). The linewidth $\Delta(h\upsilon)$ is typically between $2.5k_BT$ to $3k_BT$ as shown in Figure 15 (c).
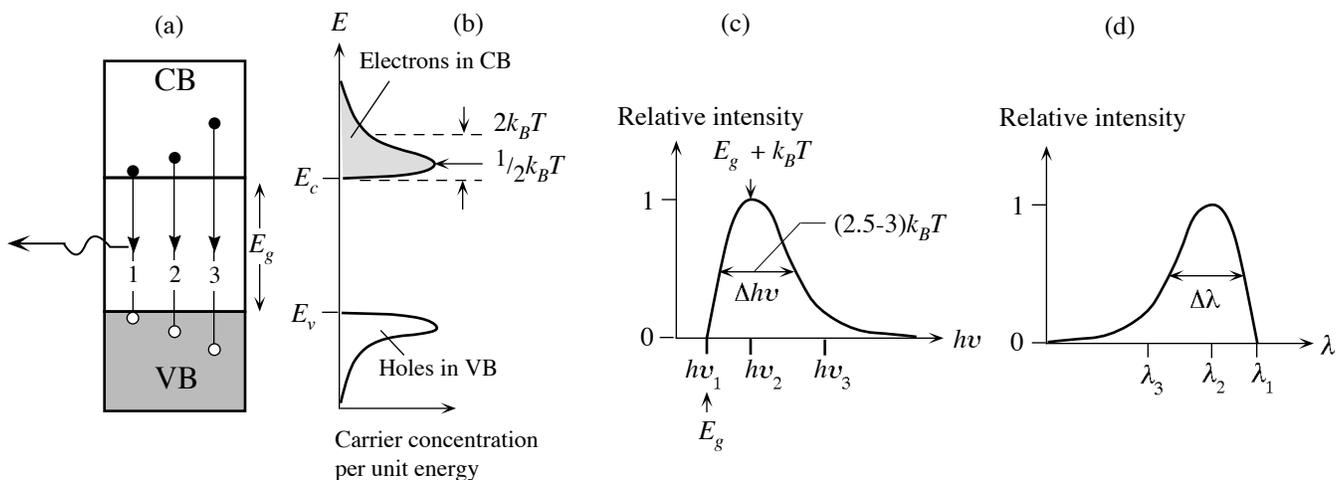
(a) Energy band diagram with possible recombination paths. (b) Energy distribution of electrons in the CB and holes in the VB. The highest electron concentration is $(1/2)k_BT$ above E_c. (c) The relative light intensity as a function of photon energy based on (b). (d) Relative intensity as a function of wavelength in the output spectrum based on (b) and (c).
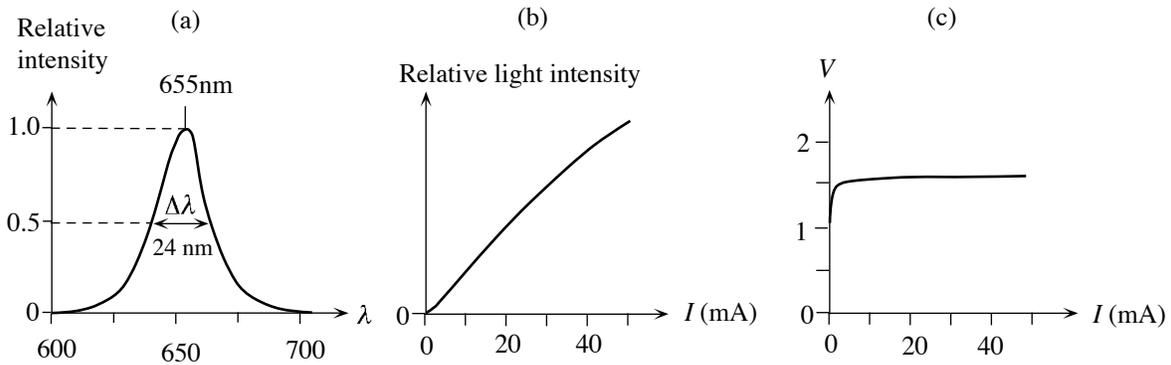
Figure 15

The output spectrum, or the relative intensity vs. wavelength characteristics, from an LED depends not only on the semiconductor material but also on the structure of the *pn* junction diode, including the dopant concentration levels. The spectrum in Figure 15 (d) represents an idealized spectrum without including the effects of heavy doping on the energy bands. For a heavily doped *n*-type semiconductor, there are so many donors that the electron wavefunctions at these donors overlap to generate a narrow impurity band centered at E_d, but extending into the conduction band. Thus the donor impurity band

---

[8]The intensity is not necessarily maximum when both the electron and hole concentrations are maximum, but it will be close.

overlaps the conduction band and hence effectively lowers E_c. The minimum emitted photon energy from heavily doped semiconductors is therefore less than E_g and depends on the amount of doping.

Typical characteristics of a red LED (655 nm) as an example are shown in Figure 16 (a) to (c). The output spectrum in Figure 16 (a) exhibits less asymmetry than the idealized spectrum in Figure 15 (d). The width of the spectrum is about 24 nm which corresponds to a width of about $2.7k_BT$ in the energy distribution of the emitted photons. As the LED current increases, so does the injected minority carrier concentration, thus the rate of recombination and hence the output light intensity. The increase in the output light power is not linear with the LED current, however, as apparent in Figure 16 (b). At high current levels, strong injection of minority carriers leads to the recombination time depending on the injected carrier concentration and hence on the current itself; this leads to a nonlinear recombination rate with current. Typical current-voltage characteristics are shown in Figure 16 (c) where it can be seen that the **turn-on** or the **cut-in voltage** is about 1.5 V from which point the current increases very steeply with voltage. The turn-on voltage depends on the semiconductor and generally increases with the energy bandgap E_g. For example, for a blue LED it is typically about 3.5 - 4.5 V, for a yellow LED, it is about 2 V, and for a GaAs infrared LED it is around 1 V.

(a) A typical output spectrum (relative intensity vs. wavelength) from a red GaAsP LED. (b) Typical output light power vs. forward current. (c) Typical *I-V* characteristics of a red LED. The turn-on voltage is around 1.5 V.

Figure 16

8.1. Example: LED Output Spectrum

Given that the width of the relative light intensity vs. photon energy spectrum of an LED is typically around $\sim 3k_BT$, what is the linewidth $\Delta\lambda_{1/2}$ in the output spectrum in terms of wavelength?

Solution

We note that the emitted wavelength λ is related to the photon energy E_{ph} by

$$\lambda = c/\upsilon = hc/E_{ph}$$

If we differentiate λ with respect to photon energy E_{ph} we get

$$\frac{d\lambda}{dE_{ph}} = -\frac{hc}{E_{ph}^2}$$

We can represent small changes or intervals (or Δ) by differentials, *e.g.* $\Delta\lambda/\Delta E_{ph} \approx |d\lambda/dE_{ph}|$, then

$$\Delta\lambda \approx \frac{hc}{E_{ph}^2}\Delta E_{ph}$$

We are given the energy width of the output spectrum, $\Delta E_{ph} = \Delta(h\upsilon) \approx 3k_BT$. Then, using the latter and substituting for E_{ph} in terms of λ we find,

$$\Delta\lambda \approx \lambda^2 \frac{3kT}{hc}$$

Thus, at $\lambda = 870$ nm, $\Delta\lambda = 47$ nm

$\lambda = 1300$ nm, $\Delta\lambda = 105$ nm

$\lambda = 1550$ nm, $\Delta\lambda = 149$ nm

These linewidths are typical values and the exact values depend on the LED structure.

Questions and Problems

1. **Direct bandgap *pn* junction** Consider a GaAs *pn* junction which has the following properties. $N_a = 10^{16}$ cm^{-3} (*p*-side), $N_d = 10^{16}$ cm^{-3} (*n*-side), $B = 7.21 \times 10^{-16}$ m^3 s^{-1}, cross-sectional area $A = 0.1$ mm^2. What is the diode current due to diffusion in the neutral regions at 300 K when the forward voltage across the diode is 1 V? Note: $\mu_e = 7000$ cm^2 V^{-1} s^{-1}, and $\mu_e = 310$ cm^2 V^{-1} s^{-1}.

2. **The Si *pn* junction** Consider a long *pn* junction diode with an acceptor doping, N_a, of 10^{18} cm^{-3} on the *p*-side and donor concentration of N_d on the *n*-side. The diode is forward biased and has a voltage of 0.6 V across it. The diode cross-sectional area is 1 mm^2. The minority carrier recombination time, τ, depends on the dopant concentration, N_{dopant} (cm^{-3}), through the following approximate relation

$$\tau = \frac{5 \times 10^{-7}}{\left(1 + 2 \times 10^{-17} N_{dopant}\right)}$$

a Suppose that $N_d = 10^{15}$ cm^{-3}. Then the depletion layer extends essentially into the *n*-side and we have to consider minority carrier recombination time, τ_h, in this region. Calculate the diffusion and recombination contributions to the total diode current given that when $N_a = 10^{18}$ cm^{-3}, $\mu_e \approx 250$ cm^2 V^{-1} s^{-1}, and when $N_d = 10^{15}$ cm^{-3}, $\mu_h \approx 450$ cm^2 V^{-1} s^{-1}. What is your conclusion?

b. Suppose that $N_d = N_a$. Then W extends equally to both sides and, further, $\tau_e = \tau_h$. Calculate the diffusion and recombination contributions to the diode current given that when $N_a = 10^{18}$ cm^{-3}, $\mu_e \approx 250$ cm^2 V^{-1} s^{-1}, and when $N_d = 10^{18}$ cm^{-3}, $\mu_h \approx 130$ cm^2 V^{-1} s^{-1}. What is your conclusion?

Relative spectral output power

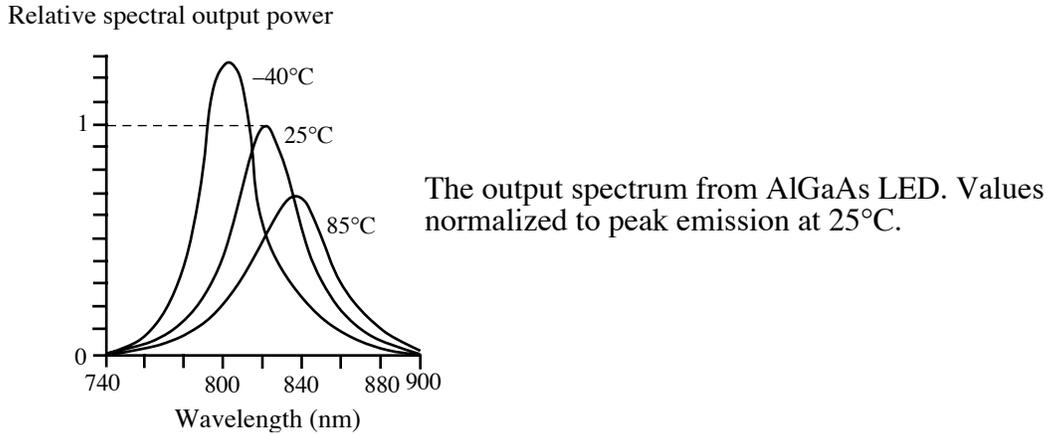The output spectrum from AlGaAs LED. Values normalized to peak emission at 25°C.

Figure 17

3.   **AlGaAs LED emitter** An AlGaAs LED emitter for use in a local optical fiber network has the output spectrum shown in Figure 17. It is designed for peak emission at 820 nm at 25 °C.

a   What is the linewidth $\Delta\lambda$ between half power points at temperatures −40 °C, 25 °C and 85 °C? What is the empirical relationship between $\Delta\lambda$ and T given three temperatures and how does this compare with $\Delta(h\upsilon) \approx 2.5k_BT - 3k_BT$?

b   Why does the peak emission wavelength increase with temperature?

**c**   Why does the peak intensity decrease with temperature?

d   What is the bandgap of AlGaAs in this LED?

e   The bandgap, E_g, of the ternary alloys $Al_xGa_{1-x}As$ follows the empirical expression,

$$E_g(eV) = 1.424 + 1.266x + 0.266x^2$$

What is the composition of the AlGaAs in this LED?

f   When the forward current is 40 mA, the voltage across the LED is 1.5 V and the optical power that is coupled into a multimode fiber through a lens is 25 μW. What is the efficiency?
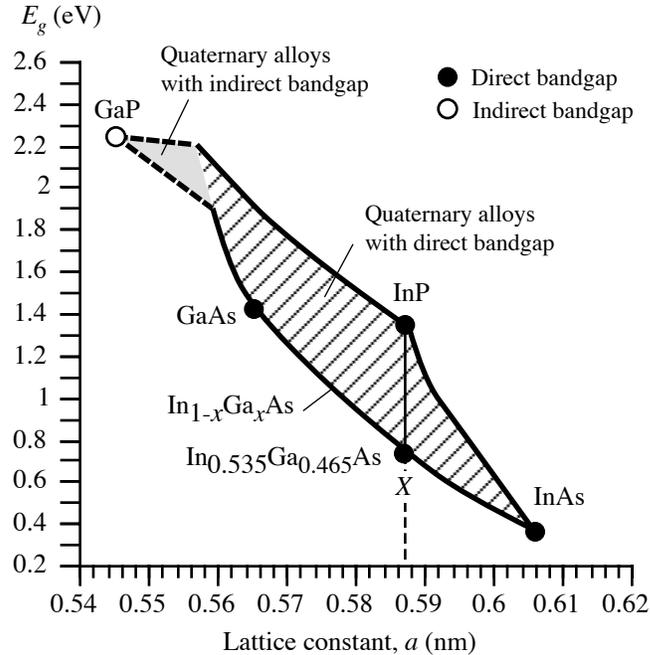
Bandgap energy E_g and lattice constant a for various III-V alloys of GaP, GaAs, InP and InAs. A line represents a ternary alloy formed with compounds from the end points of the line. Solid lines are for direct bandgap alloys, whereas dashed lines are for indirect bandgap alloys. Regions between lines represent quaternary alloys. The line from X to InP represents quaternary alloys $In_{1-x}Ga_xAs_{1-y}P_y$ made from $In_{0.535}Ga_{0.465}As$ and InP which are lattice matched to InP.

Figure 18

4.   III-V compound semiconductors in optoelectronics Figure 18 represents the bandgap E_g and the lattice parameter a in the quaternary III-V alloy system. A line joining two points represents the changes in E_g and a with composition in a ternary alloy composed of the compounds at the ends of that line. For example, starting at GaAs point, $E_g = 1.42$ eV and $a = 0.565$ nm, and E_g decreases and a increases as GaAs is alloyed with InAs and we move along the line joining GaAs to InAs. Eventually at InAs, $E_g = 0.35$ eV and $a = 0.606$ nm. Point X in Figure 18 is composed of InAs and GaAs and it is the ternary alloy $In_xGa_{1-x}As$. It has $E_g = 0.7$ eV and $a = 0.587$ nm which is the same a as that for InP. $In_xGa_{1-x}As$ at X is therefore lattice matched to InP and can hence be grown on an InP substrate without creating defects at the interface.

Further, $In_xGa_{1-x}As$ at X can be alloyed with InP to obtain a quaternary alloy $In_xGa_{1-x}As_yP_{1-y}$ whose properties lie on the line joining X and InP and therefore all have the same lattice parameter as InP but different bandgap. Layers of $In_xGa_{1-x}As_yP_{1-y}$ with composition between X and InP can be grown epitaxially on an InP substrate by various techniques such as liquid phase epitaxy (LPE) or molecular beam epitaxy (MBE).

The shaded area between the solid lines represents the possible values of E_g and a for the quaternary III-V alloy system in which the bandgap is direct and hence suitable for direct recombination.

The compositions of the quaternary alloy lattice matched to InP follow the line from X to InP.

a   Given that the $In_xGa_{1-x}As$ at X is $In_{0.535}Ga_{0.465}As$, show that quaternary alloys $In_xGa_{1-x}As_yP_{1-y}$ are lattice matched to InP when $y = 2.15x$.

b   The bandgap energy E_g, in eV for $In_xGa_{1-x}As_yP_{1-y}$ lattice matched to InP is given by the empirical relation,

$$E_g \text{ (eV)} = 1.35 - 0.72y + 0.12\, y^2$$

Find the composition of the quarternary alloy suitable for an emitter operating at 1.55 μm.

5. **External conversion efficiency**   The *external power or conversion efficiency* η_{ext} is defined as

$$\eta_{ext} = \frac{\text{Optical power output}}{\text{Electrical power output}} = \frac{P_o}{IV}$$

One of the major factors reducing the external power efficiency is the loss of photons in extracting the emitted photons which suffer reabsorption in the *pn* junction materials, absorption outside the semiconductors and various reflections at interfaces.

The total light output power from a particular AlGaAs red LED is 2.5 mW when the current is 50 mA and the voltage is 1.6 V. Calculate its external conversion efficiency.

6. **Linewidth of LEDs**   Experiments carried out on various direct bandgap semiconductor LEDs give the output spectral linewidth (between half intensity points as in Figure 17) listed in Table 2. From 8.1, we know that a spread in the wavelength is related to a spread in the photon energy,

$$\Delta\lambda \approx \frac{hc}{E_{ph}^2}\Delta E_{ph} \tag{1}$$

Table 2

Linewidth $\Delta\lambda_{1/2}$ between half points in the output spectrum (Intensity vs. wavelength) of four GaAsP LEDs.

Peak wavelength of emission (λ) nm	650	810	820	890	950	1150	1270	1500
$\Delta\lambda_{1/2}$ nm	22	36	40	50	55	90	110	150
Material (Direct E_g)	AlGaAs	AlGaAs	AlGaAs	GaAs	GaAs	InGaAsP	InGaAsP	InGaAsP

Table 3

Linewidth $\Delta\lambda_{1/2}$ between half points in the output spectrum (intensity vs. wavelength) of various visible LEDs using SiC and GaAsP materials.

Peak wavelength of emission (λ) nm	468	565	583	600	635
$\Delta\lambda_{1/2}$ nm	66	28	36	40	40
Color	Blue	Green	Yellow	Orange	Red
Material	SiC (Al)	GaP (N)	GaAsP (N)	GaAsP (N)	GaAsP

Suppose that we write $E_{ph} = hc/\lambda$ and $\Delta E_{ph} = \Delta(h\upsilon) \approx mk_BT$ where m is a numerical constant. Show that,

$$\Delta\lambda \approx \lambda^2 \frac{mk_BT}{hc} \tag{2}$$

and by appropriately plotting the data in Table 2, and assuming $T = 300$ K, find m.

Table 3 shows the linewidth $\Delta\lambda_{1/2}$ for various visible LEDs. Radiative recombination is obtained by appropriately doping the material. Using $m \approx 3$, $T = 300$ K, in Equation (2) calculate the expected

spectral width for each and compare with the experimental value. What is your conclusion? Do you think E_N in Figure 12 (b) is a discrete level?

7. **Internal quantum efficiency** The internal efficiency η_{int} gauges what fraction of electron hole recombinations in the forward biased *pn* junction are radiative and therefore lead to photon emission. Nonradiative transitions are those in which and electron and a hole recombine through a recombination center such as a crystal defect or an impurity and emit phonons (lattice vibrations). By definition,

$$\eta_{int} = \frac{\text{Rate of radiative recombination}}{\text{Total rate of recombination (radiative and nonradiative)}} \qquad (1)$$

or $\qquad \eta_{int} = \dfrac{\dfrac{1}{\tau_r}}{\dfrac{1}{\tau_r} + \dfrac{1}{\tau_{nr}}} \qquad (2)$

where τ_r is the mean lifetime of a minority carrier before it recombines radiatively and τ_{nr} is the mean lifetime before it recombines via a recombination center without emitting a photon. The total current I is determined by the total rate of recombinations whereas the number of photons emitted per second (Φ_{ph}) is determined by the rate of radiative recombinations.

$$\eta_{int} = \frac{\text{Photons emitted per second}}{\text{Total carriers lost per second}} = \frac{\Phi_{ph}}{I/e} = \frac{P_{op(int)}/h\upsilon}{I/e} \qquad (3)$$

where $P_{op(int)}$ is the optical power generated internally (not yet extracted).

For a particular AlGaAs LED emitting at 850 nm it is found that $\tau_r = 50$ ns and $\tau_{nr} = 100$ ns. What is the internal optical power generated at a current of 100 mA?

NOTATION

A	cross-sectional area
B	direct recombination capture coefficient
c	speed of light (3.0×10^8 m s^{-1})
CB	conduction band
C_{dep}	depletion layer capacitance
D	diffusion coefficient (m^2 s^{-1})
DH	double-heterostructure
e	electronic charge (1.602×10^{-19} C); electron (subscript)
E	energy
E_c, E_v	energy levels of conduction and valence band
E_F	fermi energy level
E_g	bandgap energy
$\mathcal{E}$	electric field (V m^{-1})
$\mathcal{E}_o$	internal electric field (V m^{-1})
EHP	electron-hole pair
$f(E)$	Fermi-Dirac function
$g(E)$	density of states
$G_{thermal}$	thermal generation rate of electron-hole pairs
h	Planck's constant (6.6261×10^{-34} J s); hole (subscript)

I	current (A)
J_D	diffusion current density (A m^{-2})
J_{so}	reverse saturation current density (A m^{-2})
k_B	Boltzmann's constant (1.3807×10^{-23} J K^{-1})
L	diffusion length (m)
LED	light emitting diode
M	metallurgical junction
n	electron concentration (m^{-3})
n_{no}, p_{po}	equilibrium majority carrier concentrations (m^{-3})
n_{po}, p_{no}	equilibrium minority carrier concentrations (m^{-3})
$n_p(0)$	electron concentration just outside the depletion region (m^{-3})
n_i	intrinsic concentration (m^{-3})
N_a	concentration of acceptor ions (m^{-3})
N_c, N_v	effective density of states at the conduction and valence band edges (m^{-3})
N_d	concentration of donor ions (m^{-3})
p	hole concentration (m^{-3})
$p_n(0)$	hole concentration just outside depletion region (m^{-3})
PE	potential energy
q	charge (C)
S	closed surface
SCL	space charge layer
T	temperature
V	voltage (V)
VB	valence band
V_o	built-in voltage (V)
V_r	reverse bias voltage (V)
W_o	total width of depletion region
W_p, W_n	width of p and n depletion regions, respectively
ε	$\varepsilon_o \varepsilon_r$, permittivity of a medium (C V^{-1} m^{-1} or F m^{-1})
ε_o	permittivity of free space or absolute permittivity (8.8542×10^{-12} C V^{-1} m^{-1} or F m^{-1})
ε_r	relative permittivity or dielectric constant
η	diode ideality factor
λ	wavelength
μ	drift mobility (m^2 V^{-1} s^{-1})
ρ_{net}	net space charge density
τ	minority carrier lifetime (s)
τ_g	mean thermal generation time
υ	frequency

USEFUL DEFINITIONS

Acceptor atoms are dopants that have one less valency than the host atom. They therefore accept electrons from the valence band (VB) and thereby create holes in the VB which leads to $p > n$ and hence to a p-type semiconductor.

Boltzmann statistics describes the behavior of a collection of particles (e.g. gas atoms) in terms of their energy distribution. It specifies the number of particles, $N(E)$, with an energy E through $N(E) \propto \exp(-E/kT)$ where k is the Boltzmann constant and T is the temperature (Kelvins). The description is non-quantum mechanical in the sense that

there is no restriction to how many particles can have the same state (the same wavefunction) with an energy E. It applies when there are only a few particles compared with the number of possible states so that the likelihood of two particles having the same state becomes negligible. This is generally the case for thermally excited electrons in the conduction band of a semiconductor where there are many more states than number of electrons. The kinetic energy distribution of gas molecules in a tank obeys the Boltzman statistics.

Built-in voltage (V_o) is the voltage across a *pn* junction, going from *p* to *n*-type semiconductor, in an open circuit. It is not the voltage across the diode which is made up of V_o as well as the contact potential at the metal electrode to semiconductor junctions.

Conduction band (CB) is a band of energies for the electron in a semiconductor where it can gain energy from an applied field and drift and thereby contribute to electrical conduction. The electron in the CB behaves as if it were a "free" particle with an effective mass, m_e^*.

Depletion (space charge) layer capacitance is the incremental capacitance (dQ/dV) due to the change in the exposed dopant charges in the depletion layer as a result of the change in the voltage across the *pn* junction.

Donor atoms are dopants in the semiconductor that have a valency one more than the host atom. They therefore donate electrons to the conduction band (CB) and thereby create electrons in the CB which leads to $n > p$ and hence to an *n*-type semiconductor (*n* is the electron concentration in the CB and *p* is the hole concentration in the valence band (VB)).

Heterojunction is a junction between different semiconductor materials, for example between GaAs and GaAlAs ternary alloy. There may or may not be a change in the doping. The doping in the wider bandgap semiconductor is denoted with a capital letter, N or P, and that in the narrower bandgap semiconductor with lower case n or p.

Homojunction is a junction between differently doped regions of the same semiconducting material, for example a *pn* junction in the same silicon crystal; there is no change in the bandgap energy, E_g.

Law of the junction relates the injected minority carrier concentration just outside the depletion layer to the applied voltage. For holes in the *n*-side, that is, for minority carriers, it is $p_n(0) = p_{no}\exp(eV/kT)$ where $p_n(0)$ is the hole concentration just outside the depletion layer, p_{no} is the equilibrium hole concentration (both in the *n*-side), V is the external applied voltage, k is the Boltzmann constant and T is the temperature (K).

Light emitting diode (LED) is a semiconductor diode which emits incoherent radiation. LEDs operate on the principle of spontaneous emission resulting from electron-hole pair injection and *direct* recombination under forward bias. Consider what happens when a *p-n+* junction is forward biased. As soon as a forward bias V is applied across this junction, this voltage drops across the depletion region since this is the most resistive part of the device. Consequently, the built-in potential V_o is reduced to $V_o - V$ which then allows the electrons from the n^+ side to diffuse, or become injected, into the *p*-side. The hole injection component from *p* into the n^+ side is much smaller than the electron injection component from the n^+ to *p*-side. The recombination of injected electrons in the depletion region as well as in the neutral *p*-side results in the spontaneous emission of photons. Recombination primarily occurs within the depletion region and within a volume extending over the diffusion length L_e of the electrons in the *p*-side. This recombination zone is frequently called the **active region**. The phenomenon of light emission from electron-hole pair (EHP) recombination as a result of minority carrier injection as in this case is called **injection electroluminescence.** Because of the statistical nature of the recombination process between electrons and holes, the emitted photons are in random directions; they result from spontaneous emission processes in contrast to stimulated emission. The LED structure has to be such that the emitted photons can escape the device without being reabsorbed by the semiconductor material. This means the *p*-side has to be sufficiently narrow or we have to use *heterostructure* devices.

Majority carriers are electrons in an *n*-type and holes in a *p*-type semiconductor.

Metallurgical junction is where there is an effective junction between the *p*-type and *n*-type doped regions in the crystal. It is where the donor and acceptor concentrations are equal or where there is a transition from *n*- to *p*-type doping.

Minority carrier diffusion length (L) is the mean distance a minority carrier diffuses before recombination, $L = \sqrt{[D\tau]}$ where D is the diffusion coefficient and τ is the minority carrier lifetime.

Minority carrier injection is the flow of electrons into the *p*-side and holes into the *n*-side of a *pn* junction when a voltage is applied to reduce the built-in voltage across the junction. It can also refer to an increase in the minority carrier concentration due to photogeneration.

***pn* junction** is a contact between a *p*-type and an *n*-type semiconductor. It has rectifying properties.

Recombination of an electron-hole pair involves an electron in the conduction band (CB) falling in energy down into an empty state (hole) in the valence band (VB) to occupy it. The result is the annihilation of the electron-hole pair. Recombination is direct when the electron falls directly down into an empty state in the VB as in GaAs. Recombination is indirect if the electron is first captured locally by a defect or an impurity, called a recombination center, and from there it falls down into an empty state (hole) in the VB as in Si and Ge.

Recombination centers are defects or impurities in a crystal that facilitate the recombination of electrons and holes, usually (but not always) via phonon emissions, that is, without photon emission.

Relative permittivity (ε_r) or the dielectric constant of a dielectric is the fractional increase in the stored charge per unit voltage on the capacitor plates due to the presence of the dielectric between the plates (the whole space between the plates is assumed to be filled). Alternatively we can define it as the fractional increase in the capacitance of a capacitor when the insulation between the plates is changed from vacuum to a dielectric material keeping the geometry the same.

Reverse bias is the application of an external voltage to a *pn* junction such that the positive terminal is connected to the *n*-side and negative to the *p*-side. The applied voltage increases the built-in potential and hence the internal field in the space charge layer.

Semiconductor is a nonmetallic element (e.g. Si or Ge) that contains both electrons and holes as charge carriers in contrast to an enormous number of electrons only as in metals. A hole is essentially a "half-broken" covalent bond which has a missing electron and therefore behaves effectively as if positively charged. Under the action of an applied field the hole can move by accepting an electron from a neighboring bond thereby passing on the "hole". Electron and hole concentrations in a semiconductor are generally many orders of magnitude less than those in metals thus leading to much smaller conductivities.

Substrate is a single mechanical support which carries active and passive devices. For example, in integrated circuit technology, typically many integrated circuits are fabricated on a single silicon crystal wafer which serves as the substrate.

Thermal generation current is the current that flows in a reverse biased *pn* junction as a result of the thermal generation of electron-hole pairs in the depletion layer which become separated and swept across by the built-in field.

Valence band (VB) is a band of energies for the electrons in bonds in a semiconductor. The valence band is made of all those states (wavefunctions) that constitute the bonding between the atoms in the crystal. At absolute zero of temperature the valence band is full of all the bonding electrons of the atoms. When an electron is excited to the conduction band, then this leaves behind an empty state which is called a hole. It carries positive charge and behaves as it were a "free" positively charged entity with an effective mass of m_h^*. It moves around the VB by having a neighboring electron tunnel into the unoccupied state.

"But the practitioners of science are, after all, human beings. They are not immune to the normal influences of egotism, economic self-interest, fashion, wishful thinking and laziness. A scientist may try to steal credit, knowingly initiate a worthless project for gain, or take a conventional idea for granted instead of looking for a better explanation. From time to time scientists even fudge their results, breaking one of the most serious taboos of their profession."

Murray Gell-Mann (Professor Emeritus, California Institute of Technology; Nobel Laureate 1969)
The Quark and the Jaguar,
Adventures in the simple and the complex
(W.H. Freeman and Company, New York, USA, 1994, p.80)