Chapter 4
PN Junction

4.1 Junction Structure

The pn junction refers to a piece of semiconductor say p-type. Part of the material is exposed to n-doping so that the final structure is p-type in one side and n-type on the other. It is never meant to be two pieces of materials one n-type and one p-type glued or connected together. The micro scale contact between the p-region and the n-region is what is called junction. The junction surface is an imaginary plane, where we can identify the p-region from the n-region. There are two basic structures. One is called abrupt and the other is graded. In the abrupt junction the material switches types abruptly at the junction plane. In the graded structure the doping changes gradually from one side to the other. In the following analysis we will use the abrupt junction model. (Fig. 4.1).

4.2 pn Junction at Thermal Equilibrium

In the p-region the hole concentration is high while in the n-region the hole concentration is low. Therefore, hole diffusion takes place from the p-region to the n-region. This diffusion causes acceptor ions near the junction plane to be uncovered since some holes have been removed to the n-side. Similarly, some electrons in the n-side diffuse to the p-region leaving behind uncovered donor ions.

This results in the formation of a dipole layer (called transition region or depletion region). The width of this layer will not increase indefinitely because an electric field will soon develop within the depletion region directed from the donor ions to the acceptor ions. This electric field will act to thwart further diffusion of holes from the p-side to n-side, or electrons from the n-side to the p-side. At thermal equilibrium we will have a hole diffusion current counteracted by a hole drift current in the opposite direction. Also, an electron diffusion current is counteracted by an electron drift current. The net hole current is zero and the net electron current is zero. The existence of an internal (built-in) electric field results in the setup of an internal (built-in) potential difference (barrier potential or contact potential) ϕ_0 as given from eqn. (3-18) by

$$\frac{p_{p_0}}{p_{n_0}} = e^{|e|\frac{\phi_0}{V_{th}}} \qquad (4\text{-}1)$$

Where p_{p_0} is the hole concentration in the p-region at thermal equilibrium (equal to N_A) and p_{n_0} is the hole concentration in the n-region at thermal equilibrium (equal to n_i^2/N_D). Thus, eqn. (4-1) becomes

$$\phi_0 = \Delta V = V_{th}\ell n\frac{p_{p_0}}{p_{n_0}} = V_{th}\ell n\frac{N_A N_D}{n_i^2} \qquad (4\text{-}2)$$

This is the internal built-in voltage which leads to band bending, but it cannot be measured directly by a voltmeter. Fermi level at thermal equilibrium E_{F_0} must be constant since the net current is zero, whether the device is open circuit (oc) or short circuit (sc). The transition region forms a space charge region in the bulk around the junction plane. But away from the junction the material is neutral on both sides, where there is no charge and no net electric field. The band diagram of a pn junction at thermal equilibrium is shown (Fig. 4.2).

We also plot the carrier concentration distribution (Fig. 4.3). The hole concentration in the p-region is $p_{p_0} = N_A$ throughout the neutral region. It then drops very rapidly within the depletion region to $p_{n_0} = n_i^2/N_D$ at the border of the neutral region of the n-region.

4.3 The Potential Barrier

We must realize that at thermal equilibrium the diffusion currents (and hence the drift currents) are quite large (Prob. 4.2). This is due to the fact that the concentration gradients are quite high. The concentration drops fast over a very small width. Also, the electric field is quite large due to the fact that the potential drops in a very small distance. We can now visualize the flow of two opposing currents. The forward current represents holes diffusing from the p-region to the n-region and electrons from the n-region to the p-region. The reverse current - mainly drift of holes being returned by the internal electric field flows from the n-region to the p-region and electrons from the p-region to the n-region. We note that the potential barrier is like an obstacle in the path of the diffusion process.

Thus, the forward current represents those carriers with sufficient energy to surmount the barrier which acts like a hill. The reverse current represents carriers flowing back in the opposite direction, or carriers falling off the top of the hill. We can use the Maxwellian tail (Fig. 4.4) to explain how carriers surmount the barrier, noting that the band diagram is drawn for electrons which means that the band diagram for holes has opposite polarity, i.e., the barrier impedes holes diffusing from p-region to n-region. At thermal equilibrium, both forward current I_F and reverse current I_r are equal and given by

$$I_F = I_r = A_o e^{-|q|\phi_0/kT} \qquad (4\text{-}3)$$

where A_o is a constant, since only carriers with sufficient energy $|q|\phi_0$ can surmount the barrier which separates E_{c_p} from E_{c_n}

4.4 The I-V Characteristics of a pn Diode

We call the pn junction a diode and is given the symbol shown (Fig. 4.1b). When an external battery is connected to the pn junction such that the positive side of the battery is connected to the p-region and the negative side of the battery to the n-region, we call this connection forward bias. In this case, the net electric field inside the depletion region is decreased and the potential barrier height is reduced from ϕ_0 to $\phi_0 - V_d$, where V_d is the diode voltage. At thermal equilibrium there is a balance between two currents; forward current and reverse current.

The forward current is driven by diffusion, whereas the reverse current is basically drift within the depletion region. The forward current as given by eqn. (4-3) must now be expressed as

$$I_F = Ae^{-|q|(\phi_0 - V_d)/kT} \qquad (4\text{-}4)$$

In forward bias we call $V_d = V_f$ (and in the reverse bias V_r). The barrier height is now $\phi_0 - V_f$, i.e., the diffusion current (forward current) increases due to the lowering of the obstacle, namely, the potential barrier (Fig. 4.5). In other words, the Maxwellian tail now acts to drive in more carriers. However the reverse current does not change, since it represents carriers driven back, which are not really affected by the height at the barrier. It is easy to see that sliding down the height does not depend much on how high the barrier is, hence called saturation current I_o. So the reverse current remains as before

$$I_r = I_o = A_o e^{-|q|\phi_0/kT} \qquad (4\text{-}5)$$

Thus, there is a net current

$$\begin{aligned} I_{net} &= I_F - I_r \\ &= A_o e^{-|q|(\phi_0 - V_d)/kT} - A_o e^{-|q|\phi_0/kT} \\ &= A_o e^{-|q|\phi_0/kT}\left[e^{|q|V_d/kT} - 1\right] \qquad (4\text{-}6) \end{aligned}$$

$$I_D = I_{net} = I_o\left[e^{|q|V_d/kT} - 1\right] \qquad (4\text{-}7)$$

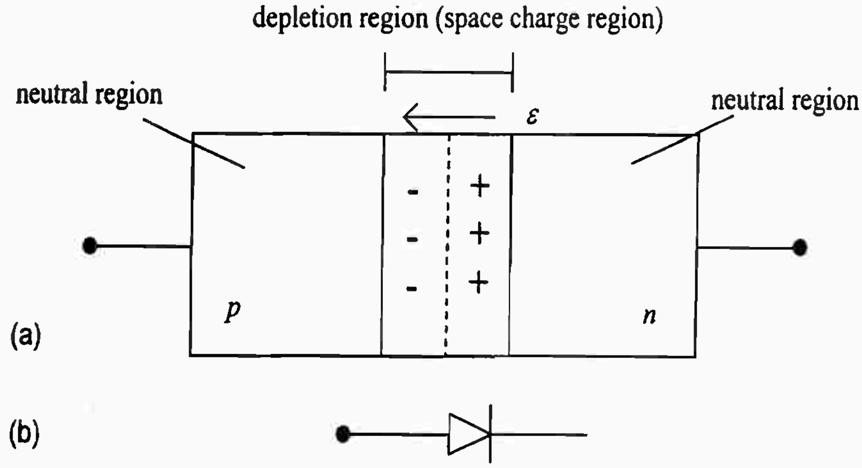


Fig. (4.1) Abrupt pn junction
a) structure b) symbol

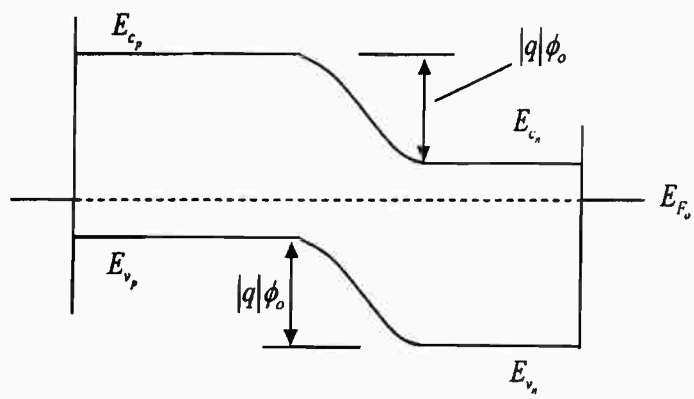


Fig. (4.2) Band diagram of a pn junction at thermal equilibrium

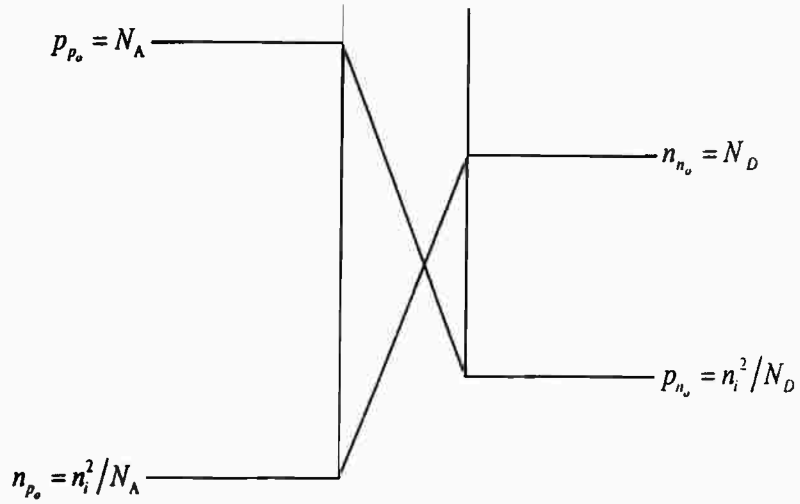


Fig. (4.3) Carrier distribution in a pn junction at thermal equilibrium

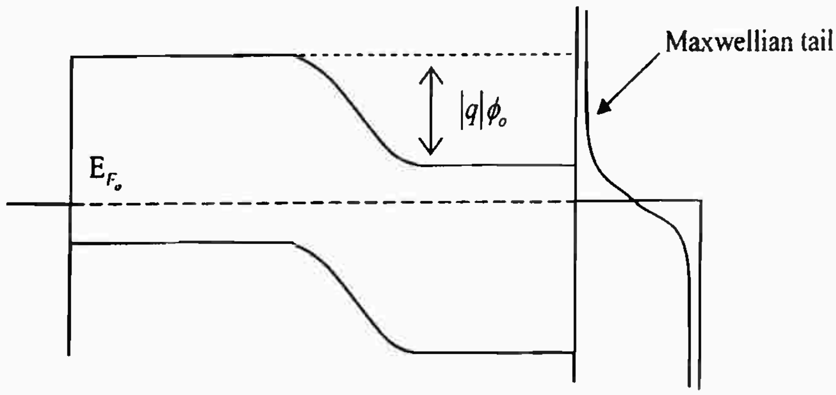


Fig. (4.4) Maxwellian tails for carriers in a pn junction at thermal equilibrium

We may rewrite eqn. (4-7)

$$I_D = I_o\left[e^{V_d/V_{th}} - 1\right] \tag{4-8}$$

where V_{th} is the thermal voltage and is equal to 0.025V at room temperature. Eqn. (4-8) is called the I-V characteristic of the pn diode (Fig. 4.6). It is seen that for forward bias the current I_d increases exponentially with diode voltage drop V_d, whereas for reverse bias (the positive side of the battery connected to the n- region) the current is basically $-I_o$. This means that the I-V characteristic is asymmetrical. It allows current to flow in the forward direction, whereas it practically blocks the current in the reverse direction.

4.5 Load Line

Consider a diode in series with a resistor and an external voltage is applied across the combination. We note that the current in a series circuit must be the same throughout the circuit. To calculate this current analytically we have a problem, since the I-V characteristic is nonlinear. We cannot simply apply Ohm's law as in a combination of two series resistors. A graphical method has been derived (Fig. 4.7) to calculate the current in this case. We locate on the V axis a point at V_{BB}, where V_{BB} in the battery voltage. Then we draw a line (called load line) with a slope $-1/R_L$, where R_L is the series resistor. The point of intersection is called the Q point (Quiescent point) or biasing point. We then read I_d which flows in both the diode and the resistor. The voltage or the diode V_d is also read on the V axis. This is justified since

$$V_{BB} = V_d + V_R \tag{4-9}$$

$$= V_d + IR \tag{4-10}$$

We may use as an approximation a linearized model of the pn diode in which we assume the IV characteristic a straight line with a solpe $1/R_F$ where R_F is called the forward resistance and V_γ is called the cutin voltage (Fig. 4.8). We may define the incremental conductance g_d

$$g_d = \frac{dI_d}{dV_d} \tag{4-11}$$

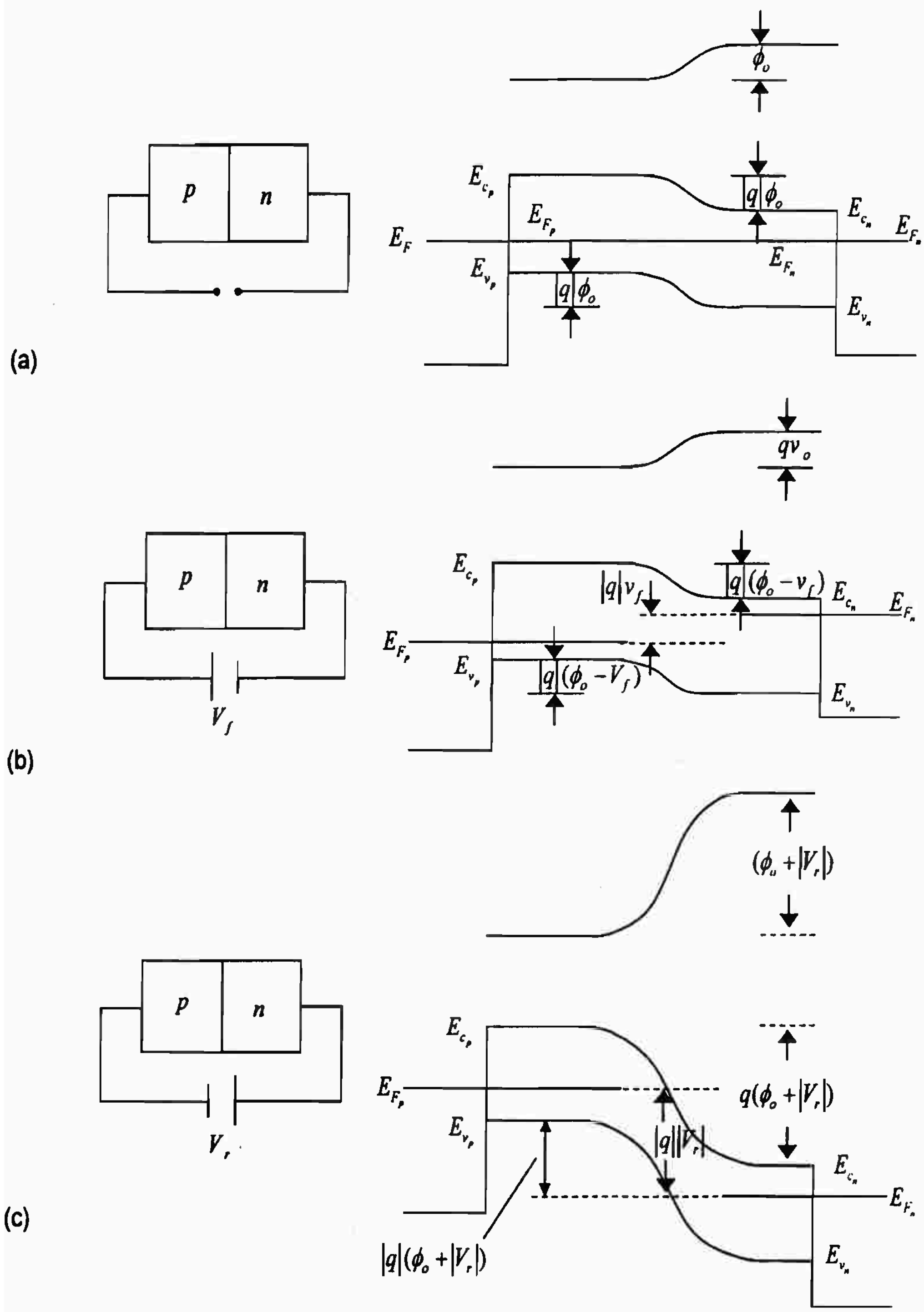


Fig (4.5) Effect of bias on the band diagram of a pn junction under
a) thermal equilibrium b) forward bias c) reverse bias

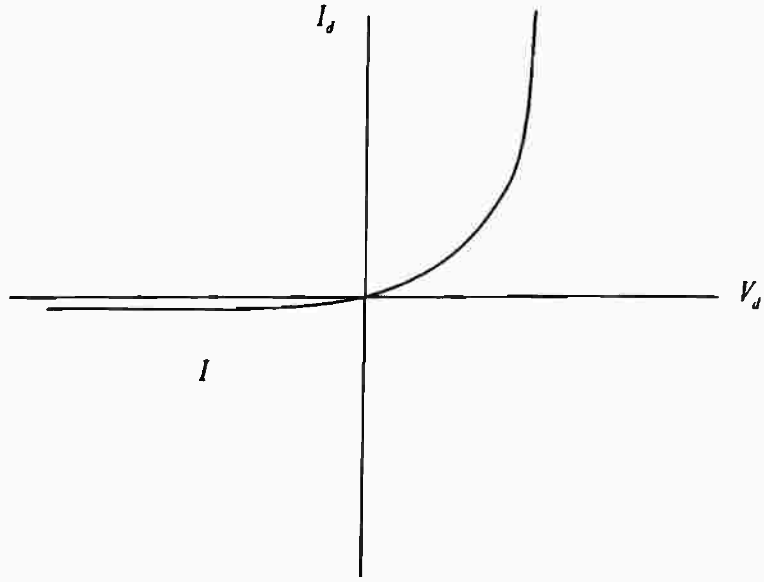


Fig. (4.6) I-V characteristic of a pn diode

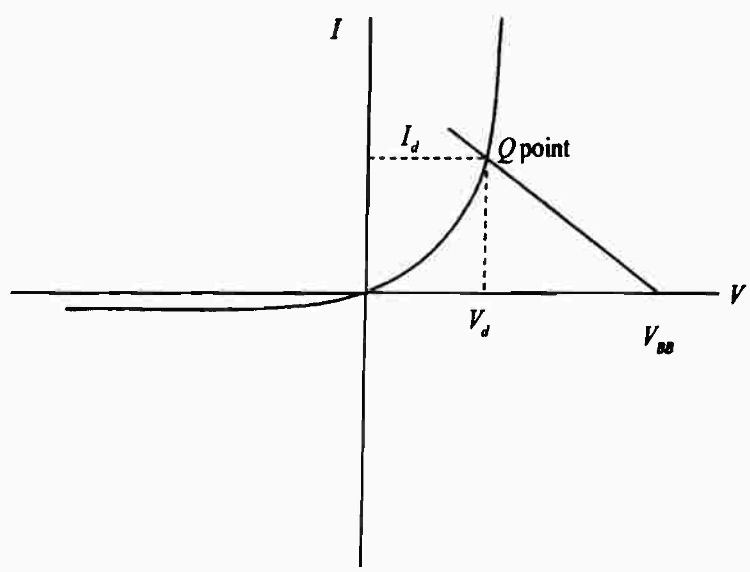


Fig. (4.7) Load line

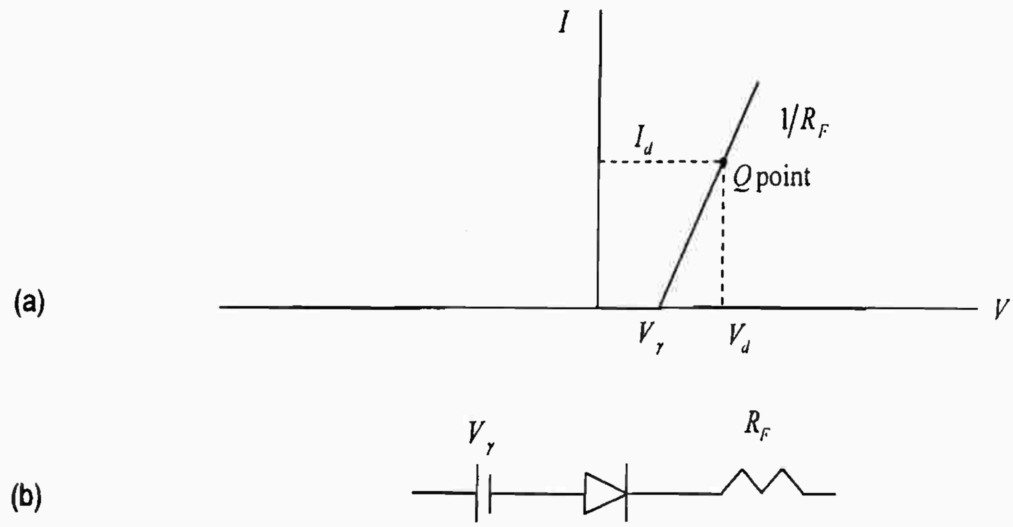


Fig. (4.8) Linearized model of a pn diode
a) IV characteristic b) circuit model

From eqn. (4-8) we have

$$g_d = \frac{I_o}{V_{th}} e^{V_d/V_{th}} \tag{4-12}$$

Which approximately given by

$$g_d = \frac{I_d}{V_{th}} \tag{4-13}$$

Hence,
$$r_d = \frac{1}{g_d} = \frac{V_{th}}{I_d} \tag{4-14}$$

where I_d is the current at the Q point and r is assumed to be constant around that particular Q point by taking

$$r_d = R_F \tag{4-15}$$

Is taken as the representative forward resistance of the diode. Of course at reverse bias we find that r_d or R_r (reverse resistance) is almost infinity. Thus, the diode presents a small forward resistance in the forward bias and an infinite reverse (backward) resistance in the reverse bias. Hence, a pn diode may be regarded as an ideal switch, which is short circuit in the forward direction and open circuit in the reverse direction.

Ex. 4.1

Derive expressions for the electric field under nonuniform doping for large and small electric field intensities at thermal equilibrium.

Solution

From electromagnetic field theory (Poisson's equation)

$$\nabla \cdot \bar{D} = \varsigma \tag{4-16}$$

$$\nabla \cdot \bar{\varepsilon} = \frac{\varsigma}{\epsilon}$$

Where $\bar{D}$ is displacement vector and ς is the charge density

Thus,

$$\nabla \bullet \bar{\varepsilon} = \frac{|q|}{\epsilon}\left(N_D^+ + p - n - N_A^-\right) \tag{4-17}$$

For one dimensional geometry and n-type material, neglecting the hole density,

$$\frac{d\varepsilon}{dx} = \frac{|q|}{\epsilon}\left(N_D^+(x) - n(x)\right) \tag{4-18}$$

From eqn. (3-4) at thermal equilibrium, $J_n = 0$. Thus,

$$|q|\mu_n n(x)\varepsilon(x) + |q|D_n \frac{d_n(x)}{dx} = 0 \tag{4-19}$$

We want to distinguish between two cases:

a) $N_0^+(x) >> n(x)$ or $p(x)$. This is the case of the depletion (or space charge) region

Thus,

$$\frac{d\varepsilon}{dx} = \frac{|q|N_D^+(x)}{\epsilon} \tag{4-20}$$

or

$$\varepsilon(x) = \int_0^x \frac{|q|N_D^+(x)dx}{\epsilon} \tag{4-21}$$

After we obtain $\varepsilon(x)$, we may obtain $n(x)$ from eqn (4-19)

$$\frac{dn(x)}{n(x)} = -\frac{\mu_n}{D_n}\varepsilon(x)dx \tag{4-22}$$

Integrating eqn. (4-22)

$$\int_0^x \frac{dn(x)}{n(x)} = -\int_0^x \frac{\varepsilon(x)dx}{V_{th}} \tag{4-23}$$

$$\ell n \frac{n(x)}{n(x_o)} = \frac{V(x) - V(x_o)}{V_{th}}$$

At x_0 (the boundary), $V(x_0) = 0$

$$n(x) = n(x_o)e^{V(x)/V_{th}} \tag{4-24}$$

which is in agreement with eqn. (3-20)

b) in this case, we assume that under nonuniform doping the electric field is small enough that charge neutrality still holds, so that.

$$\frac{d\varepsilon}{dx} = 0 \tag{4-25}$$

$$n(x) = N_D^+(x) \tag{4-26}$$

Now, the small electric field may be obtained from eqn. (4-19)

$$\varepsilon(x) = -\frac{D_n}{\mu_n}\frac{1}{n(x)}\frac{dn(x)}{dx}$$

$$= -\frac{V_{th}}{N_D^+(x)}\frac{dN_D^+(x)}{dx} \tag{4-27}$$

Ex. 4.2

Obtain expressions for the current densities in terms of pseudo Fermi levels.

Solutions

Recalling eqns. (3-3), (3-4) and (3-8), we may write under non equilibrium conditions

$$n = n_i e^{(F_n - E_{F_i})/kT} \tag{4-28}$$

$$p = n_i e^{(E_{F_i} - F_p)/kT} \tag{4-29}$$

where F_n and F_p are pseudo Fermi levels, i.e., Fermi levels under non equilibrium conditions which are separate for electrons and holes.

Thus,

$$\frac{dn}{dx} = \frac{n_i}{kT} e^{(F_n - E_{F_i})/kT} \left(\frac{dF_n}{dx} - \frac{dE_{F_i}}{dx} \right)$$

$$= \frac{n}{kT} \left(\frac{dF_n}{dx} - \frac{dE_{F_i}}{dx} \right) \tag{4-30}$$

$$\frac{dp}{dx} = \frac{p}{kT} \left(\frac{dE_{F_i}}{dx} - \frac{dF_p}{dx} \right) \tag{4-31}$$

Thus,

$$J_p = \mu_p p \frac{dE_{F_i}}{dx} - \frac{qD_p}{kT} p \left(\frac{dE_{F_i}}{dx} - \frac{dF_p}{dx} \right) \tag{4-32}$$

From eqn. (3-10), eqn. (4-30) reduces to

$$J_p = \mu_p p \frac{dF_p}{dx} \tag{4-33}$$

$$J_n = \mu_n n \frac{dF_n}{dx} \tag{4-34}$$

At thermal equilibrium, $\frac{dF_p}{dx} = \frac{dF_n}{dx} = 0$ since $J_n = J_p = 0$, thus $F_n = F_p$ and is constant.

4.6 The Continuity Equation

Recombination must be included in the diffusion process. Consider a semiconductor specimen which has a hole current density $J_p(x)$. Consider a differential length Δx and a cross sectional area A. The current density $J_p(x+dx)$ may be larger or smaller that $J_p(x)$, depending on the generation and recombination of carriers within the differential volume. The net rate of increase in hole concentration with respect to time $\frac{\partial p(x,t)}{\partial t}$ is the difference between the hole flux per unit volume entering and leaving minus the recombination rate.

The flux of holes entering the differential volume per unit volume is $\frac{1}{|q|}\frac{J_p(x)}{\Delta x}$

The flux of holes leaving the differential volume per unit volume is $\frac{1}{|q|}\frac{J_p(x+\Delta x)}{\Delta x}$

The rate of recombination is $-\frac{\delta p(x,t)}{\tau_r}$ where τ_r is the excess carrier (hole or electron) recombination lifetime. Thus,

$$\frac{\partial p}{\partial t}=\frac{1}{|q|}\frac{J_p(x)-J_p(x+\Delta x)}{\Delta x}-\frac{\delta_p}{\tau_r} \quad (4\text{-}35)$$

As Δx approaches zero, eqn. (4-35) reduces to

$$\frac{\partial p(x,t)}{\partial t}=\frac{\partial \delta p}{\partial t}=-\frac{1}{q}\frac{\partial J_p}{\partial x}-\frac{\delta p}{\tau_r} \quad (4\text{-}36)$$

This expression is called the continuity equation for holes.
Similarly, for electrons

$$\frac{\partial \delta n}{\partial t}=\frac{1}{|q|}\frac{\partial J_n}{\partial x}-\frac{\delta n}{\tau_r} \quad (4\text{-}37)$$

A minus sign is taken into consideration since the electron is a negative charge.
Noting that $p=p_0+\delta p$, where p_0 is constant, we have

$$\frac{\partial \delta p}{\partial t}=-\mu_p\left(p\frac{\partial \varepsilon}{\partial x}+\varepsilon\frac{\partial p}{\partial x}\right)+D_p\frac{\partial^2 \delta p}{\partial x^2}-\frac{\delta p}{\tau_r} \quad (4\text{-}38)$$

and

$$\frac{\partial \delta n}{\partial t}=\mu_n\left(n\frac{\partial \varepsilon}{\partial x}+\varepsilon\frac{\partial n}{\partial x}\right)+D_n\frac{\partial^2 \delta n}{\partial x^2}-\frac{\delta n}{\tau_r} \quad (4\text{-}39)$$

Under conditions of negligible electric field, where ε and $\frac{\partial \varepsilon}{\partial x}$ tend to zero, i.e., negligible drift currents and purely diffusion currents, we have

$$\frac{\partial \delta p}{\partial t}=D_p\frac{\partial^2 \delta p}{\partial x^2}-\frac{\delta p}{\tau_r} \quad (4\text{-}40)$$

and

$$\frac{\partial \delta n}{\partial t}=D_n\frac{\partial^2 \delta n}{\partial x^2}-\frac{\delta n}{\tau_r} \quad (4\text{-}41)$$

In many problems, the steady state distributions of electrons and holes are sought, in which case

$$\frac{\partial \delta p}{\partial t}=\frac{\partial \delta n}{\partial t}=0$$

Thus, at steady state we replace the symbol δ by Δ

$$\frac{d^2\Delta p}{dx^2}=\frac{\Delta p}{D_p\tau_r}=\frac{\Delta p}{L_n^2} \quad (4\text{-}42)$$

$$\frac{d^2\Delta n}{dx^2}=\frac{\Delta n}{D_n\tau_r}=\frac{\Delta n}{L_n^2} \quad (4\text{-}43)$$

where $L_p=\sqrt{D_p\tau_p}$ is called the hole diffusion length and $L_n=\sqrt{D_n\tau_n}$ is called the electron diffusion length.

The physical meaning of the diffusion length can be best understood by considering a situation where the injection of excess holes takes place in a semiconductor bar. Recombination takes place

along the bar and is responsible for continually decreasing Δp from an initial value $\Delta p(0)$ finally to zero, where the effect of recombination has completely wiped out the injection effect.
A reasonable solution for eqn. (4-42)

$$\Delta p(x) = \Delta p(0)e^{-x/L_p}$$

where $\Delta p(o)$ is the value of $\Delta p(x)$ at $x = 0$
In absence of the electric field $\varepsilon = 0$

$$J_p(x) = -|q| D_p \frac{dp}{dx} = -|q| D_p \frac{d\Delta p(x)}{dx} = |q| \frac{D_p}{L_p} \Delta p(0)e^{-x/L_p}$$

$$= |q| \frac{D_p}{L_p} \Delta p(x) \qquad (4\text{-}44)$$

Thus, the diffusion length L_p represents the distance at which the excess hole distribution is reduced to $1/e$ of its initial value (Fig. 4.9) which is $\Delta p(o)$ (the diffusion current reduces by the same value).

We should note that in order to keep this steady state distribution stable in the face of recombination, continued injection must go on. If injection is ceased, the entire distribution dies off gradually according to the continuity eqn. (4-40). We can also show that L_p is the average distance a hole diffuses before recombining.

The probability that a hole injected at $x = 0$ will recombine in a given dx between x and $x + dx$ is P_x, which is the product of two probabilities $P_1 P_2$. Defining P_1 as the probability that a hole injected at $x = 0$ reaches x without recombination and P_2 as the probability of a hole at x recombining in the subsequent interval dx, we have

$$P_1 = \frac{\Delta p(x)}{\Delta p(0)} = e^{-x/L_p} \qquad (4\text{-}45)$$

In so doing, we have taken the probability as the ratio between the differential concentration at x and the initial value at $x = 0$. Similarly, P_2 is the ratio of the differential change in Δp to the initial value at the entrance to the element Δx, which is Δp, and from eqn. (4-43)

$$P_2 = \frac{\Delta p - \Delta p(x + dx)}{\Delta p} = \frac{\frac{d\Delta p}{dx} dx}{\Delta p} = \frac{1}{L_p} dx \qquad (4\text{-}46)$$

Thus, $$P_x(x) = P_1 P_2 = \frac{1}{L_p} e^{-x/L_p} dx \qquad (4\text{-}47)$$

The average of a quantity $f(x)$ is given by $< f(x) >$ as

$$< f(x) > = \int_{-\infty}^{\infty} x f(x) P_x(x) dx \qquad (4\text{-}48)$$

$P_x(x)$ is the probability distribution, which is P_x in this case. Thus,

$$< x > = \int_{o}^{\infty} x \frac{1}{L_p} e^{-x/L_p} dx = L_p \qquad (4\text{-}49)$$

We should note that the quantity to be averaged here is x. The probability distribution is $P_x(x)$ where $< x >$ is the statistical average of distance a hole travels before recombining which is equal to L_p .

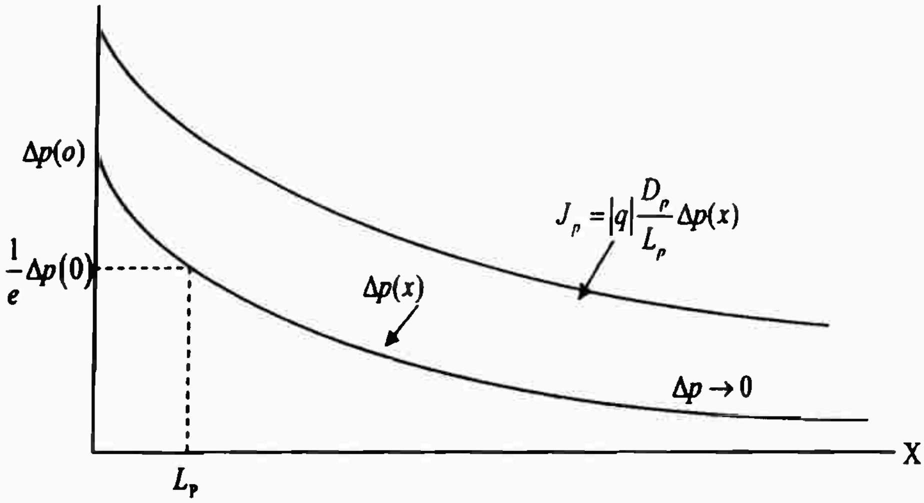


Fig. (4.9) Meaning of the diffusion length

Ex. 4.3

Calculate the diffusion potential for a pn junction in which $N_A = 10^{14} cm^{-3}$ and $N_D = 10^{12} cm^{-3}$. Hence show how ϕ_o changes with temperature, noting that $n_i(300^o k) = 10^{10} cm^{-3}$ and $E_g = 1eV$

Solution

We have
$$\phi_o = 0.025\ \ell n\frac{10^{26}}{10^{20}} = 0.025\ \ell n 10^6$$
$$= 0.025 \times 6 \times 2.3$$
$$= 0.345\text{V}$$

$$\frac{n_i^2(T)}{n_i^2(300^o K)} = \left(\frac{T}{300}\right)^3 e^{-(V_G/V_{th} - V_G/0.025)} \qquad (4\text{-}50)$$

Where,

$$V_{th} = 0.025\left(\frac{T}{300}\right) \qquad (4\text{-}51)$$

and V_G is the bandgap in eV

$$\phi_o(T) = 0.025\left(\frac{T}{300}\right)\left[\ell n(N_A N_D) - \ell n\left(n_i^2(300)\left(\frac{T}{300}\right)^3\right) + \left(\frac{V_G}{0.025}\frac{300}{T} - \frac{V_G}{0.025}\right)\right]$$
$$= 0.025\left(\frac{T}{300}\right)\left[\ell n\left(\frac{N_A N_D}{n_i^2(300)}\right) - 3\ell n\left(\frac{T}{300}\right)\right] + V_G - \left(\frac{T}{300}\right)V_G \qquad (4\text{-}52)$$

Taking the derivative term by term

$$\frac{d\phi_o}{dT} = \frac{0.025}{300}\ell n\left(\frac{N_A N_D}{n_i^2(300)}\right) - 0.025\left(\frac{T}{300}\right)\times 3\times\left(\frac{300}{T}\right)\times\frac{1}{300} - \frac{0.025}{300}\times 3\ell n\left(\frac{T}{300}\right) - \frac{V_G}{300}$$

$$= \left[\frac{0.025}{300}\ell n\left(\frac{N_A N_D}{n_i^2(300)}\right) - \frac{0.025}{100} - \frac{V_G}{300}\right] - \frac{0.025}{100}\ell n\left(\frac{T}{300}\right)$$

$$= \left[\frac{0.345}{300} - \frac{0.025}{100} - \frac{1}{300}\right] - \frac{0.025}{100}\ell n\left(\frac{T}{300}\right)$$

$$= -2.4\times10^{-3} - 2.5\times10^{-4}\ell n\left(\frac{T}{300}\right) \qquad (4\text{-}53)$$

We see from eqn. (4-53) that the rate of variation $\phi_o(T)$ with T is negative, i.e., ϕ_o decreases as temperature increases, and near room temperature the rate becomes temperature dependent to lesser extent. Thus,

$$\frac{d\phi_o(T)}{dT} \simeq -2.4\times10^{-3}\ \text{V/°K} \qquad (4\text{-}54)$$

4.7 Injection and Current Components in a pn Junction

So far we have considered the current components within the transition region on a phenomenological basis. The current in a biased pn junction must be continuous. We want now to analyze the current components as they exist in the neutral region outside the depletion region. We still maintain that the field is zero in the neutral region, but excess holes and electrons are dumped (or injected) into these regions, by lowering the potential barrier for a forward biased junction (Fig. 4.10). At thermal equilibrium, the hole densities on both sides of the junction are related by eqn. (4.1)
Rewriting eqn. (4-1) under forward bias,

$$\frac{p_{p_o}}{p(x_{n_o})} = e^{|q|(\phi_o - V_f)/kT} \qquad (4\text{-}55)$$

This equation is based on the assumption that the current under forward bias is still zero (similar to thermal equilibrium). The verification of this statement comes about from the fact that the variation of the concentrations of holes and electrons is very fast in a small distance (width of depletion region). Thus, the diffusion currents and the drift currents are very large. The net current which is the difference, however, is very small. Thus, we may neglect the net current with respect to both components as we did in the case of thermal equilibrium. Hence, we may use the same result as in eqn. (4-1)
Dividing eqn. (4-1) by eqn. (4-55) we get

$$\frac{p(x_{n_o})}{p_{n_o}} = e^{|q|V_f/kT} \qquad (4\text{-}56)$$

Defining $\Delta p(x_{n_o})$ as the initial injected excess hole density at x_{n_o} in the n-region, and noting that for small forward voltages (or low-level injection), $\Delta p(x_{n_o}) << p_{p_o}$, we find, using eqn. (4-56) that,

$$\Delta p(x_{n_o}) = p(x_{n_o}) - p_{n_o}$$

$$= p_{n_o}(e^{|q|V_f/kT} - 1) \qquad (4\text{-}57)$$

Similarly,

$$\Delta n\left(-\left|x_{p_o}\right|\right) = n_{p_o}(e^{|q|V_f/kT} - 1) \qquad (4\text{-}58)$$

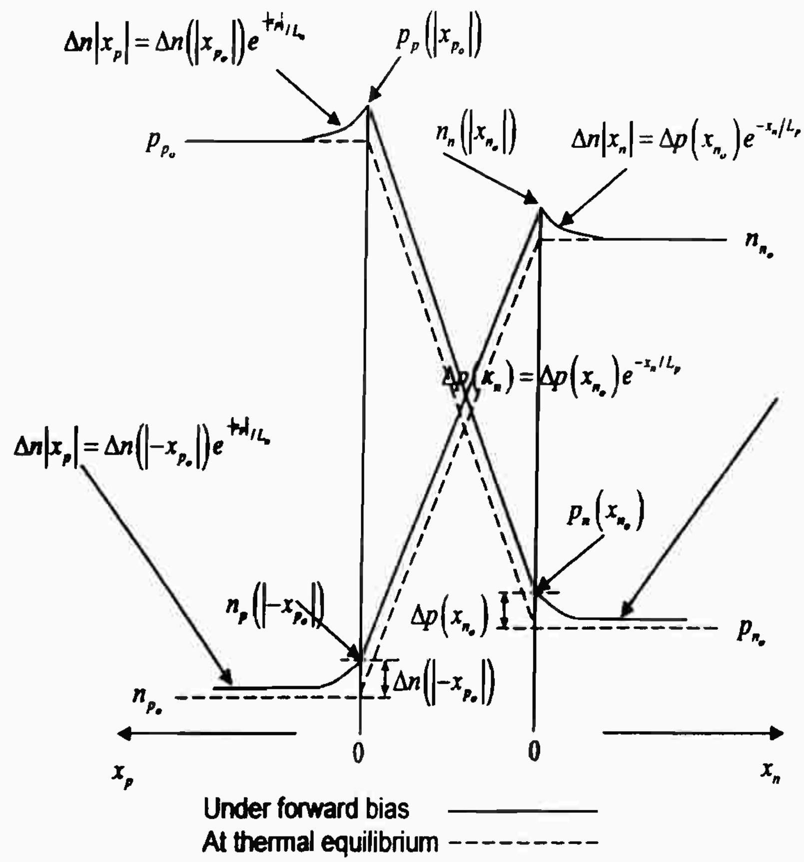


Fig. (4.10) Spatial distributions of electrons and holes at thermal equilibrium (dotted) and under forward bias (solid)

where x_{p_o} represents the beginning of the neutral region away from the junction on the p-side, and x_{n_o} the beginning of the neutral region away from the junction on the n-side. The injected holes and the injected electrons follow steady state distributions given by

$$\Delta p(x_n) = p_{n_o}(e^{qV_f/kT} - 1)e^{-x_n/L_p} \qquad (4\text{-}59)$$

$$\Delta n(x_p) = n_{p_o}(e^{qV_f/kT} - 1)e^{-x_p/L_n} \qquad (4\text{-}60)$$

where x_n and x_p measure from x_{n_o} and $-|x_{p_o}|$, respectively

We assume that the neutral regions have infinite conductivity (or very large majority carrier concentrations). Current can flow in this case with a very small electric field, that is why we can regard the neutral regions as field-free, and hence the ohmic voltage drop there is zero.

In the n-material the hole current has two components: diffusion and drift. Likewise, in the p-material, the electron current has two components: diffusion and drift (Fig. 4.11). We shall neglect the drift of the minorities on both sides. Thus, we are left with the hole diffusion current in the n-side and the electron diffusion current in the p-side. The majority current has both drift and diffusion components. We neglect any recombination or generation in the transition region.

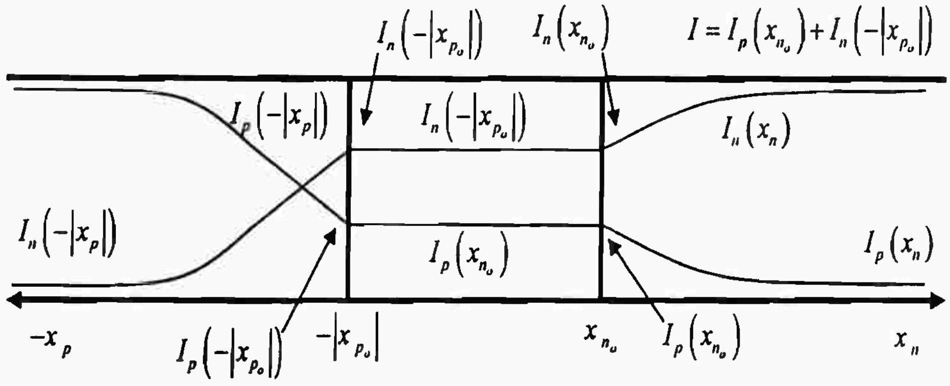


Fig. (4.11) Current components in a pn junction under forward bias

Therefore, the hole and electron currents must be constant with distance in the depletion region. Thus, the electron current $I_n(x_{n_o})$ at x_{n_o} is actually $I_n\left(-\left|x_{p_o}\right|\right)$, which is the electron diffusion current in the p-side at $-\left|x_{p_o}\right|$. Similarly, $I_p\left(\left|-x_{p_o}\right|\right)$, which is the hole current at $-\left|x_{p_o}\right|$, must be equal to the hole diffusion current at x_{n_o}, which is $I_p(x_{n_o})$.

Thus, the total current at x_{n_o} or $-\left|x_{p_o}\right|$ (or anywhere in the junction) must be the sum of the values of the minority diffusion currents at the borders of the depletion region. From eqn. (4-59) and (4-60)

$$I_p(x_n) = -|q|AD_p\frac{d\Delta p(x_n)}{dx_n}$$

$$= \frac{|q|AD_p}{L_p}p_{n_o}(e^{|q|V_f/kT}-1)e^{-x_n/L_p} \quad (4\text{-}61)$$

$$I_p(x_{n_o}) = \frac{|q|AD_p}{L_p}p_{n_o}(e^{|q|V_f/kT}-1) \quad (4\text{-}62)$$

Similarly,

$$I_n\left(-\left|x_{p_o}\right|\right) = \frac{|q|AD_n}{L_n}n_{p_o}(e^{|q|V_f/kT}-1) \quad (4\text{-}63)$$

Note that the electron current must be in the same direction as the hole current, adding to it. Thus,

$$I = \left(\frac{|q|AD_p}{L_p}p_{n_o} + \frac{|q|AD_n}{L_n}n_{p_o}\right)(e^{|q|V_f/kT}-1) \quad (4\text{-}64)$$

$$= I_o(e^{|q|V_f/kT}-1) \quad (4\text{-}65)$$

where I_o, the reverse (saturation) current is given by

$$I_o = \left(\frac{|q|AD_p}{L_p}p_{n_o} + \frac{|q|AD_n}{L_n}n_{p_o}\right) \quad (4\text{-}66)$$

We should note that to maintain the condition of charge neutrality, the holes injected into the n-region must be surrounded by electrons of equal distribution.

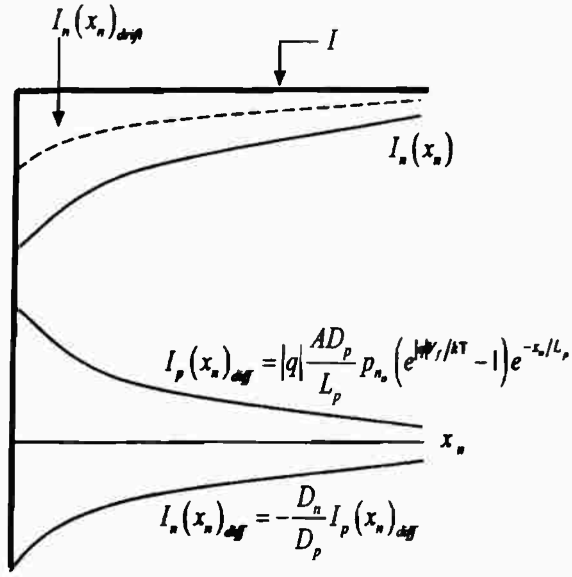


Fig. (4.12) Detailed current components in the n-side of a forward biased pn junction

$$\Delta n(x_n) = \Delta p(x_{n_o})e^{-X_n/L_p} \tag{4-67}$$

This gives an electron diffusion current component

$$I_n(x_n)_{diff} = \frac{-|q|AD_n}{L_n} p_{n_o}(e^{|q|V_f/kT} - 1)e^{-X_n/L_p} \tag{4-68}$$

We note that the ratio of the electron to hole diffusion currents is equal to $\left(-D_n/D_p\right)$

Thus, the total electron current in the n-side is

$$I_n(x_n) = I - I_p(x_n)$$

$$= I - \frac{|q|AD_p}{L_p} p_{n_o}(e^{|q|V_f/kT} - 1)e^{-X_n/L_p} \tag{4-69}$$

The electron diffusion current is given by eqn. (4-68) and the electron drift current is given by

$$I_n(x_n)_{drift} = I_n(x_n) - I_n(x_n)_{diff}$$

$$= I - \frac{|q|AD_p}{L_p} p_{n_o}(e^{|q|V_f/kT} - 1)e^{-X_n/L_p} + \frac{|q|AD_n}{L_p} p_{n_o}(e^{|q|V_f/kT} - 1)e^{-X_n/L_p}$$

$$= I - \frac{|q|A\left(D_p - D_n\right)}{L_p} p_{n_o}(e^{|q|V_f/kT} - 1)e^{-X_n/L_p} \tag{4-70}$$

This situation is shown in Fig. (4.12)

We should note that a few diffusion lengths away from the junction, near the contacts with the metal, the current in the n-side is wholly an electron drift component. Similarly, in the p-side near the metal contact, the current is a hole drift component. Actually, no holes can flow in the wire but electrons simply move out from the valence band into the wire and, hence, holes are said to flow in. More precisely, we may say that vacancies in the Maxwellian tail of the metal draw valence band electrons out of the semiconductor at the contact with the p-side. Then, these electrons are reintroduced back into the conduction band of the semiconductor at the contact with the n-side from the Maxwellian tail of the metal there.

We should note from the study of the direction of current flow (Fig. 4.13) that electrons at point (1) flow into the conduction band of the n-side. Near the junction. This electron flow helps neutralize the hole infiltration and supply the drift electron component, which replenishes electrons lost in the recombination process. At the junction itself, the remnant electron current becomes itself an infiltration thrust into the p-side (injection), establishing what amounts to an electron beachhead, which is beleaguered by hole defences. In short, we may eliminate for a while the concept of a hole and talk about an electron flow at point

(1). The recombination process brings this electron down into the valence band, thus, eliminating a free electron and a free hole, then emitting a photon. The bound electrons can still move toward point (2) and exit the semiconductor there. Thus, we can still think of the current flow as a perpetual or continuous electron motion under forward bias in both the conduction band and the valence band.

Although we have assumed the bulk regions to be field-free, yet a drift current exists, since the conductivity is assumed to be very high (or the majority carrier densities are very high). Therefore, it is quite legitimate to neglect the minority drift currents. We should note that the net current is constant throughout the device. Therefore we may calculate it within the transition region (just as good a region as anywhere else). The net current must be constant throughout the entire structure. It is our prerogative to choose the point at which we calculate the current.

One point of clarification is due here. We considered earlier, on a phenomenological basis, that the current is the difference between the forward current $I_o e^{|q|V_f/kT}$ and the reverse current I_o. What is the nature of these currents? If we choose to calculate the current inside the depletion region, then we talk of two large components of current; one diffusion and the other drift. The difference is the net current, which is very small compared to either. It is not easy to calculate the net current this way. We choose to calculate the net current at the points x_{n_0} and $-|x_{p_o}|$. There, the hole current is the diffusion component $I_{p_{dif}}(x_{n_o})$ and the electron current is the diffusion component $I_{n_{dif}}\left(-|x_{p_o}|\right)$, and the total net current is the sum of these two components. Since in the depletion region we assume no recombination or generation.

$$I_{n_{dif}}\left(-|x_{p_o}|\right) = I_n\left(x_{n_o}\right) \tag{4-71}$$

Similarly,

$$I_{p_{dif}}\left(x_{n_o}\right) = I_p\left(-|x_{p_o}|\right) \tag{4-72}$$

Thus, neglecting the drift components of the minorities, the total current is given by

$$I = I_P\left(-|x_{P_o}|\right) + I_n\left(x_{n_o}\right) \tag{4-73}$$

We see that this current is given by

$$\begin{aligned} I &= I_{o_p} e^{|q|V_f/kT} - I_{o_p} + I_{o_n} e^{|q|V_f/kT} - I_{o_n} \\ &= \left(I_{o_p} + I_{o_n}\right) e^{|q|V_f/kT} - \left(I_{o_p} + I_{o_n}\right) \\ &= I_o e^{|q|V_f/kT} - I_o \\ &= I_o\left(e^{|q|V_f/kT} - 1\right) \end{aligned} \tag{4-74}$$

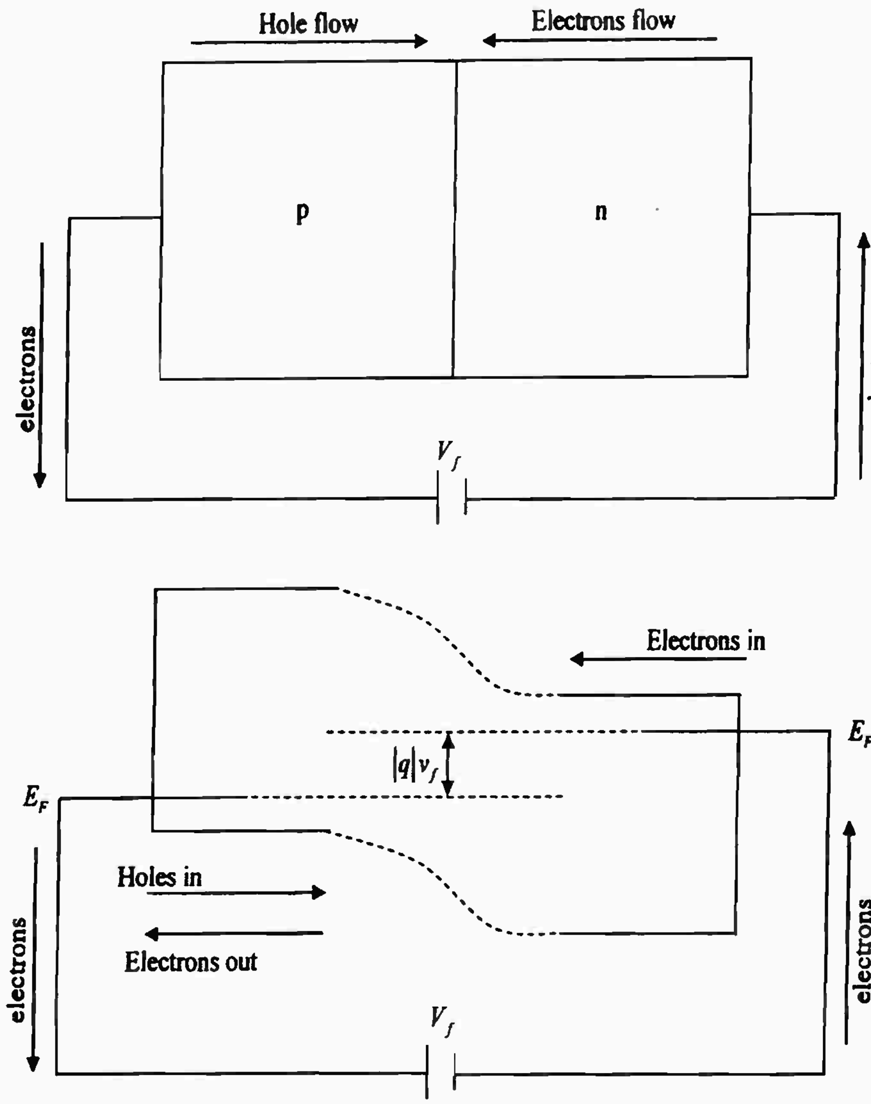


Fig. (4.13) Directions of current flow in a pn junction under forward bias

The forward component $I_{o_p} e^{|q|V_f/kT}$ represents holes injected into the n-region, and $I_{o_n} e^{|q|V_f/kT}$ electrons injected into the p-region. The forward current sometimes is referred to, as the diffusion current. We find this terminology both confusing and misleading. The forward current is not strictly the diffusion current. Also, the reverse components I_{o_p} and I_{o_n} represent current components of carriers sliding down the barrier, i.e., by the action of the barrier. This is why this current is often times referred to as drift current. This is also misleading, since the origin of this current in the bulk is rather the diffusion of minority carriers toward the junction while it is drift in nature inside the depletion region.

It is convenient to say that diffusion current inside the transition region is directed from p to n and the drift current inside the transition region is directed from n to p. The net current is the difference between the diffusion and drift components, just as it is the difference between the forward and the reverse components.

Therefore, it is tempting to relate the diffusion current in the transition region to the forward current and the drift current in the transition region to the reverse current, for the respective hole and electron components. This is a false argument. The diffusion and drift components in the transition region are spatially dependent, while the net current is constant throughout the structure independent of x. Which is the difference between the diffusion and the drift components, which is also equal to the difference between the forward and the reverse currents. But the forward current is not the diffusion current in the transition region nor is the reverse current the drift current in the transition region.

We should also note that as electrons are ultimately supplied on one side as a drift current in the n-type region, they exit as holes in the p-type side regardless of the difference in concentrations. This is because the current must be continuous. In general, $J = \sigma_p \varepsilon_p = \sigma_n \varepsilon_n$, where σ_p is the conductivity in the p-material and σ_n is the conductivity in the n-material, ε_p is the small electric field in the neutral region of the p-material and ε_n is the small electric field in the neutral region of the n-material. Therefore, the conductivities σ_p and σ_n will be different and ε_p and ε_n are adjusted such that the current is constant and small ohmic voltage drops appear in the neutral (almost field-free) regions.

It is interesting to note that the pn junction characteristic has been predetermined by the minority carriers injected into both sides of the junction, not the majority carriers. This is true for all junction devices, in which an injection mechanism takes place. We call such devices bipolar since both types of carriers are important, acting as minorities on both sides of the junction.

From the IV characteristic (Fig. 4.6), we note that the current of a forward biased junction is very small, until a critical value of voltage is reached at which the current increases dramatically with voltage. It is possible to idealize the characteristic (Fig. 4.8). Here, the characteristic is linerazed, such that up to V_γ (called cut in-voltage) the current is practically zero. Then, the current increases linearly at a slope $1/R_F$, where R_F is called the forward resistance.

We should note that this V_γ is very close to ϕ_0. The limiting value of the forward bias in the contact potential, at which point the junction is washed out. We should note that ϕ_0 as given by eqn. (4-2) has an upper limit of the bandgap itself (when E_{F_n} coincides with E_{c_n} and E_{F_p} coincides with E_{v_p}) which case is called heavily doped pn junction.

One has to distinguish between the steady state distributions of excess carriers and the current flow. In fact, the currents may be calculated as the rates of supply of charges necessary to keep the steady state distribution independent of time. One may calculate the current by considering steady state conditions. To maintain the steady state carrier distributions, the rate of recombination must be exactly balanced out by the rate of supply. Holes invading the n-side will be annihilated by recombination. To maintain the beach head without increase or decrease, the rate of injection of holes injected across the barrier (which is the depletion region) must be exactly equal to the rate by which the distribution decays through recombination. From eqns. (4-59) and (4-60)

$$I_p = \frac{Q_p}{\tau_{r_p}} = \frac{|q| A}{\tau_{r_p}} \int_{x_{n_o}}^{\infty} \Delta p(x_n) dx_n = \frac{|q| A L_p}{\tau_{r_p}} p_{n_o} (e^{|q| V_f / kT} - 1) \quad (4\text{-}75)$$

$$I_n = \frac{Q_n}{\tau_{r_n}} = \frac{|q| A}{\tau_{r_n}} \int_{-|x_{p_o}|}^{-\infty} \Delta n(x_p) dx_p = \frac{|q| A L_n}{\tau_{r_n}} n_{p_o} (e^{|q| V_f / kT} - 1) \quad (4\text{-}76)$$

where τ_{r_p} is the hole recombination lifetime in the n-region and τ_{r_n} is the electron recombination lifetime in the p-region. Thus, the total hole and electron currents across the channel are given by

$$I = \left(\frac{|q| A L_p}{\tau_{r_p}} p_{n_o} + \frac{|q| A L_n}{\tau_{r_n}} n_{p_o} \right) (e^{|q| V_f / kT} - 1) \quad (4\text{-}77)$$

$$= I_o (e^{|q| V_f / kT} - 1) \quad (4\text{-}78)$$

where

$$I_o = \left(\frac{|q| A L_p}{\tau_{r_p}} p_{n_o} + \frac{|q| A L_n}{\tau_{r_n}} n_{p_o} \right) \quad (4\text{-}79)$$

Using eqn. (4-42) and (4-43)

$$= |q| A n_i^2 \left(\frac{D_p}{L_p N_D} + \frac{D_n}{L_n N_A} \right) \quad (4\text{-}80)$$

This method of analysis is called the charge control method. It is quite instructive in calculating the current in a pn junction, provided recombination or generation is neglected in the depletion region.

4.8 Transition Capacitance

The transition (depletion or space charge) region is a dipole layer in which positive charge is accumulated in the form of positive ions on the n-side nd negative ions on the p-side. Charge carrier densities may be neglected in comparison with the fixed ion concentrations. To analyze for the electric field distribution inside the depletion region, we need to consider Poisson's equation.

$$\frac{d^2 V}{dx^2} = -\frac{\varsigma}{\epsilon} \quad (4\text{-}81)$$

Where ς is the charge density and ϵ is the permittivity. Alternatively, since $\varepsilon = -\frac{dV}{dx}$, we get

$$\frac{d\varepsilon}{dx} = \frac{\varsigma}{\epsilon} \quad (4\text{-}82)$$

The positively charged side of the depletion region extends a distance w_n into the n-side and w_p into the p-side, which correspond to points x_{n_o} and $-|x_{p_o}|$, respectively. These points, form the border lines with the neutral regions. Within w_p, the fixed ion concentration is N_A and within w_n the fixed ion concentration is N_D. From the consideration of the overall charge neutrality, the total positive charge must be balanced out by the total negative charge. Thus,

$$|q| A N_D w_n = q A N_A w_p \quad (4\text{-}83)$$

Therefore, we conclude that the transition region extends farther into the side of lighter doping. If the n-side is heavily doped $(N_D > N_A), w_n < w_p$ and the space charge is almost entirely in the p-side. By neglecting charge carrier densities n and p in the depletion region, with respect to $N_A \, or \, N_D, \rho$ is constant and the intergration of eqn. (4-82) yields a linear relation for (x)

Thus, within w_p, we have

$$\frac{d\varepsilon}{dx} = -\frac{|q| N_A}{\epsilon}, \quad -|x_{p_o}| < x < o \quad (4\text{-}84)$$

And within w_n, we have

$$\frac{d\varepsilon}{dx}=\frac{qN_D}{\in} \quad , o < x < x_{n_o} \tag{4-85}$$

These results are shown (Fig. 4.14). At the proper junction point, the electric field calculations on both sides must coincide, giving a maximum value for the electric field ε_{o_m}

$$\varepsilon_{o_m}=\frac{|q|N_A}{\in}w_p=\frac{|q|N_D}{\in}w_n \tag{4-86}$$

Which is in agreement with eqn. (4-83). We should note that the electric field is negative, which means that it is directed opposite to the positive x direction. Within w_p, the field is zero at $-|x_{p_o}|$ and builds up with negative values at a slope of $-\frac{|q|N_A}{\in}$ toward ε_{o_m}. At the maximum point, the slope switches positive to $\frac{|q|N_A}{\in}$, and ε decreases in value from a maximum value of ε_{o_m} toward zero at x_{n_0}. Thus, at the border lines of the depletion region on both sides, the electric field must be zero, since these border lines indicate the beginnings of the neutral field-free region.

The potential profile $V(x)=-\int_0^x \varepsilon(x)dx$, and therefore the total potential difference ϕ_o for a pn junction at equilibrium is given by the area under the $\varepsilon(x)$ curve. Due to the linear dependence of the electric field, the area underneath is triangular and is given by

$$|\phi_o|=\frac{1}{2}|\varepsilon_{o_m}|W \tag{4-87}$$

where W is the width of the depletion region, given by:

$$W=w_n+w_p \tag{4-88}$$

From eqn (4-83) we have

$$w_p=\frac{N_D}{N_A}w_n \tag{4-89}$$

Substituting eqn. (4-89) into eqn. (4-88)

$$W=w_n+\frac{N_D}{N_A}w_n=\left(1+\frac{N_D}{N_A}\right)w_n \tag{4-90}$$

Thus,

$$w_n=\frac{W}{1+\frac{N_D}{N_A}} \tag{4-91}$$

and

$$w_p=\frac{N_D}{N_A}\frac{W}{1+\frac{N_D}{N_A}}=\frac{W}{1+\frac{N_A}{N_D}} \tag{4-92}$$

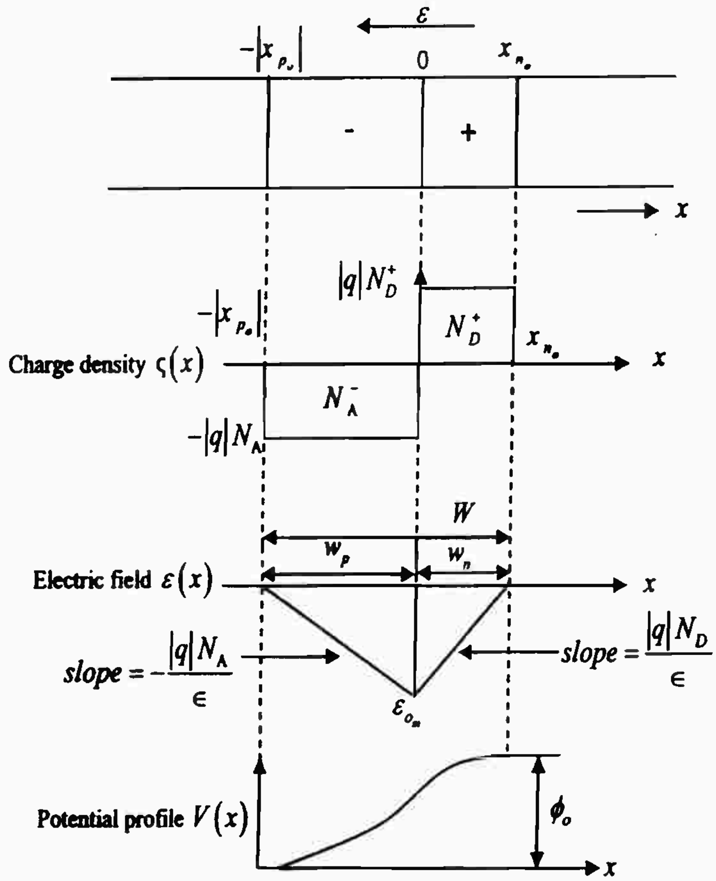


Fig. (4.14) Space charge density, electric field and voltage profiles in the transition region

From eqns. (4-87) (4-86) and (4-91) we may obtain an expression for W in terms of ϕ_o

$$\phi_o = \frac{1}{2}\frac{|q|N_D}{\epsilon}\frac{W^2}{1+\frac{N_D}{N_A}}$$

$$= \frac{1}{2}\frac{|q|}{\epsilon}\frac{N_D N_A}{N_A + N_D}W^2 \qquad (4\text{-}93)$$

or

$$W = \left[\frac{2\epsilon}{|q|}\phi_o\left(\frac{1}{N_A}+\frac{1}{N_D}\right)\right]^{1/2} \qquad (4\text{-}94)$$

So far we have considered the width of the depletion region at thermal equilibrium, and we have seen from eqn. (4-94) that it is proportional to the square root of the contact potential. We want to consider next the situation when a reverse voltage V_r is applied across the pn junction. In this case, the width of the depletion region increases, since more uncompensated fixed ions must be uncovered on both sides of the junction. This is needed to account for an increased barrier $\phi_o + |V_r|$ instead of ϕ_o, and hence, the triangle representing ε must be enlarged by the same proportion to account for an increased area.

Thus,

$$\phi_o + |V_r| = \frac{1}{2}|\varepsilon_{r_m}|W \qquad (4\text{-}95)$$

where ε_{r_m} is the maximum electric field at the junction point proper under reverse bias V_r, and is given by

$$|\varepsilon_{r_m}| = \frac{|q|N_A}{\epsilon}\frac{1}{1+\frac{N_A}{N_D}}W$$

$$= \frac{|q|}{\epsilon}\frac{1}{\left[\frac{1}{N_A}+\frac{1}{N_D}\right]}W \qquad (4\text{-}96)$$

From eqn. (4-88) under the condition of reverse bias, we have

$$W = \left[\frac{2\epsilon}{|q|}\left(\frac{1}{N_A}+\frac{1}{N_D}\right)\right]^{1/2}\left(\phi_o + |V_r|\right)^{1/2} \qquad (4\text{-}97)$$

We may now look into the total positive charge Q^+, and the total negative charge Q^- in the transition region. From eqns. (4-83), (4-91) and (4-92)

$$|Q^+| = |Q^-| = |q|AN_D\frac{W}{1+\frac{N_D}{N_A}} \qquad (4\text{-}98)$$

$$= |q|A\frac{W}{\left[\frac{1}{N_D}+\frac{1}{N_A}\right]}$$

Substituting eqn. (4-94) into eqn. (4-95)

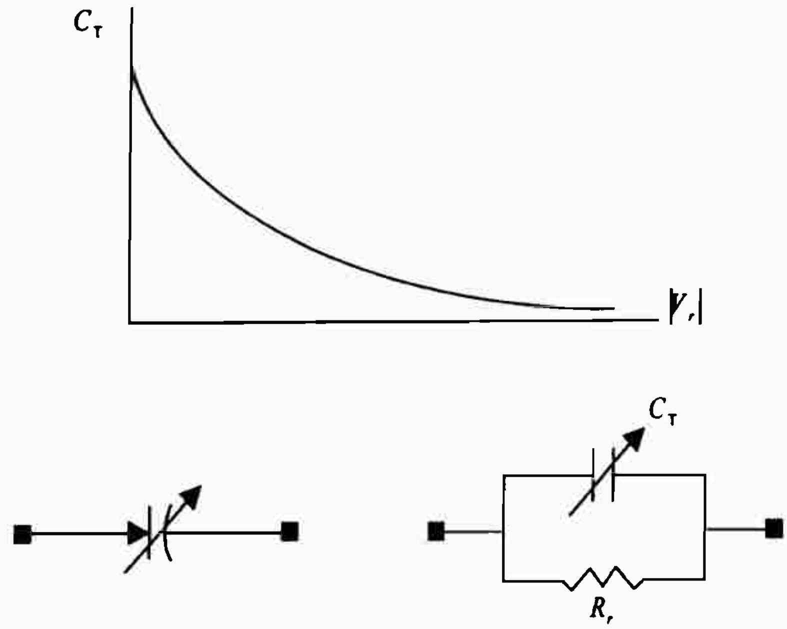
$$Q = |Q^+| = |Q^-| = \frac{|q|A\left[\frac{2\epsilon}{|q|}\left(\frac{1}{N_A}+\frac{1}{N_D}\right)\right]^{1/2}}{\left[\frac{1}{N_D}+\frac{1}{N_A}\right]}\left(\phi_o + |V_r|\right)^{1/2}$$

$$= A\left(\frac{2\epsilon|q|}{\frac{1}{N_D}+\frac{1}{N_A}}\right)^{1/2}\left(\phi_o + |V_r|\right)^{1/2} \qquad (4\text{-}99)$$

We see here that as $|V_r|$ increases, more space charge is built up in the transition region, this is a capacitive effect. The transition capacitance C_T appears under ac conditions* for a small signal, superimposed on an applied dc reverse bias V_r, and is given , using eqn (4-99)

* By ac, we do not necessarily mean sinusoidal variations, but any dynamic variations from the dc (biasing or quiescent conditions) regardless of the time waveform of the signal.

a)



(b) (c)

Fig. (4.15) Varactor
a) characteristic b) symbol c) equivalent circuit

$$C_T = \frac{dQ}{dV_r} = \frac{A}{2}\left(\frac{2\in|q|}{\frac{1}{N_D}+\frac{1}{N_A}}\right)^{1/2}(\phi_o+|V_r|)^{1/2} \qquad (4\text{-}100)$$

It is interesting to rewrite eqn. (4-100)

$$C_T = \frac{\in A}{\left[\frac{2\in}{|q|}\left(\frac{1}{N_D}+\frac{1}{N_A}\right)\right]^{1/2}(\phi_o+|V_r|)^{1/2}} \qquad (4\text{-}101)$$

From eqn. (4-97)

$$C_T = \frac{\in A}{W} \qquad (4\text{-}102)$$

This is the well known formula for the capacitance of a parallel plate capacitor of an area A and separation W.

The voltage dependence of the transition capacitance has given rise to a device called varactor (variable reactance), which is a capacitor whose capacitance can be changed by changing a reverse bias applied to it, This device is used in resonant circuits (tank circuits) in radio and TV receivers, where tuning is achieved by using a potentiometer to change the value of the capacitance (Ex. 4.5). This is called electronic tuning and is far superior to mechanical methods used in the past to change the common area of parallel plate capacitors. In general,

$$C_T \text{ is proportional to } |V_r|^{-n}\ , \qquad |V_r| > \phi_o \tag{4-103}$$

where $n = 1/2$ for an aprupt junction and $1/3$ for a graded junction. The voltage dependence of C_T is shown in Fig. (4.15). Also shown is an equivalent circuit of the varactor, where R_T is very high resistance, representing the reverse bias portion of the IV characteristic.

Ex. 4.4

a) Obtain expressions for the carrier distribution, potential profile and current components in the depletion region of a pn junction at thermal equilibrium. Hence, show that the total current must be zero.

b) Show what happens when a forward bias is applied. Take $N_A = 10^{16} cm^{-3}$, $N_D = 10^{14} cm^{-3}$, $A = 0.1 cm^2$, $D_p = 13 cm^2/s$ and $D_n = 34 cm^2/s$, $n_i = 10^{10} cm^{-3}$, $\in_r = 12$, $V_{BB} = 0.3V$.

Solution:

a) From eqns. (3-3) and (3-4) we neglect the net current in comparison with the diffusion and drift components (an assertion to be tested shortly). Thus, we obtain the following expressions for carrier distributions in the transition region in harmony with eqns. (3-19) and (3-20):

$$p(x) = p_{p_o} e^{-\Delta V(x)/V_{th}} \tag{4-104}$$

$$n(x) = n_{p_o} e^{-\Delta V(x)/V_{th}} \tag{4-105}$$

where x is the distance measured from $-|x_{p_o}|$ into the positive x direction (Fig. 4.16) the value p_{p_o} and n_{p_o} are those at $-|x_{p_o}|$ or $x = 0$. Note that $\Delta V(x)$ is the potential measured from zero at $x = 0$.

From eqn. (4-84), Fig. (4-14) and Fig. (4-16)

$$\varepsilon(x) = -\frac{|q| N_A}{\in} x \tag{4-106}$$

$$\Delta V(x) = \int_{-\infty}^{x} \varepsilon(x) dx$$

$$= \frac{1}{2}\frac{|q| N_A}{\in} x^2 \quad , \quad |x_{p_o}| > |x| > 0 \tag{4-107}$$

For $W > x > |x_{p_o}|$, we have from Fig. (4.16)

$$\varepsilon(x) = -\frac{|q| N_D}{\in}(W - x) \tag{4-108}$$

Integrated the area under $\varepsilon(x)$, we get

$$\Delta V(x) = \phi_o - \frac{1}{2}\frac{|q| N_D}{\in}(W - x)^2 , \quad W > x > |x_{p_o}| \tag{4-109}$$

Substituting eqns. (4-107) and (4-109) into eqns. (4-108) and (4-105), we get

$$p(x) = p_{p_o} e^{\frac{1}{2}\frac{|q| N_A x^2}{\in V_{th}}} \quad , \quad |x_{p_o}| > |x| > 0 \tag{4-110}$$

$$p(x) = p_{p_o} e^{\phi_o/V_{th}} e^{\frac{1}{2}\frac{|q| N_D}{\in}\frac{(W-x)^2}{V_{th}}} , \quad W > |x| > |x_{p_o}| \tag{4-111}$$

Similarly,

$$n(x) = n_{p_o} e^{\frac{1}{2}\frac{|q|N_A x^2}{\epsilon V_{th}}} \quad , |x_{p_o}| > |x| > 0 \tag{4-112}$$

$$n(x) = n_{p_o} e^{\psi_o/V_{th}} e^{\frac{1}{2}\frac{|q|N_D (W-x)^2}{\epsilon V_{th}}} \quad , \mathrm{W} > |x| > |x_{p_o}| \tag{4-113}$$

We see the $p(x)$ and $n(x)$ fall off very sharply with x, which is in agreement with the notion that the charge in the depletion region is basically fixed ions.

The four components of current inside the depletion region may be expressed, using eqns. (4-110) through (4-113) as

$$I_{p_{diff}} = -|q| AD_p \frac{dp}{dx}$$

$$= \left(\frac{|q|AD_p}{V_{th}} p_{p_o}\right)\left(\frac{|q|N_A}{\epsilon}\right) x e^{\frac{1}{2}\frac{|q|N_A x^2}{\epsilon V_{th}}} , |x_{p_o}| > |x| > 0 \tag{4-114}$$

$$I_{p_{diff}} = \left(\frac{|q|AD_p}{V_{th}} p_{p_o}\right) e^{-\psi_o/V_{th}} \left(\frac{|q|N_D}{\epsilon}(W-x)\right) e^{\frac{1}{2}\frac{|q|N_D (W-x)^2}{\epsilon V_{th}}} , \mathrm{W} > |x| > |x_{p_o}| \tag{4-115}$$

$$I_{p_{drift}} = |q| \mu_p A p(x) \varepsilon(x)$$

$$= -\left(|q|\mu_p A p_{p_o}\right)\left(\frac{|q|N_A x}{\epsilon}\right) e^{\frac{1}{2}\frac{|q|N_A x^2}{\epsilon V_{th}}} , |x_{p_o}| > |x| > 0 \tag{4-116}$$

$$I_{p_{drift}} = -|q|\mu_p A p_{p_o} e^{-\psi_o/V_{th}} \left(\frac{|q|N_D}{\epsilon}(W-x)\right) e^{\frac{1}{2}\frac{|q|N_D (W-x)^2}{\epsilon V_{th}}} , \mathrm{W} > |x| > |x_{p_o}| \tag{4-117}$$

Similarly,

$$I_{n_{diff}} = -|q| AD_n \frac{dn}{dx}$$

$$= \left(|q| A \frac{D_n}{V_{th}} n_{p_o}\right)\left(\frac{|q|N_A x}{\epsilon}\right) e^{\frac{1}{2}\frac{|q|N_A x^2}{\epsilon V_{th}}} , |x_{p_o}| > |x| > 0 \tag{4-118}$$

$$\mathrm{I}_{n_{diff}} = \left(|q| \mathrm{A} \frac{D_n}{V_T} n_{p_o}\right) e^{\psi_o/V_T} \left(\frac{|q|N_D}{\epsilon}(W-\mathrm{X})\right) e^{\frac{1}{2}\frac{|q|N_D (W-\mathrm{X})^2}{\epsilon V_T}} , \mathrm{W} > |x| > |x_{p_o}| \tag{4-119}$$

$$I_{n_{drift}} = |q| \mu_n A n(x) \varepsilon(x)$$

$$= -\left(|q|\mu_n A n_{p_o}\right)\left(\frac{|q|N_A x}{\epsilon}\right) e^{\frac{1}{2}\frac{|q|N_A x^2}{\epsilon V_{th}}} , |x_{p_o}| > |x| > 0 \tag{4-120}$$

$$I_{n_{drift}} = -\left(|q|\mu_n A n_{p_o}\right) e^{\psi_o/V_{th}} \left(\frac{|q|N_D}{\epsilon}(W-x)\right) e^{\frac{1}{2}\frac{|q|N_D (W-x)^2}{\epsilon V_{th}}} , \mathrm{W} > |x| > |x_{p_o}| \tag{4-121}$$

Using eqn (3-12), and comparing eqn.(4-114) with eqn. (4-116) and eqn.(4-109) with eqn. (4-117), we note that $I_{p_{drift}} + I_{p_{diff}} = 0$ at thermal equilibrium for all values of x. Similarly, we can show that $I_{n_{drift}} + I_{n_{diff}} = 0$ at thermal equilibrium for all values of x.

b) when an external battery is connected in forward bias such that V_f is applied on the pn junction, the triangle of the electric field distribution shrinks in size (Fig. 4.16), such that the new area underneath is given by

$$\phi_o - V_f = \frac{1}{2}\varepsilon'_m W' \qquad (4\text{-}122)$$

where ε'_m and W' are the new values of the maximum electric field and width of the depletion region under forward bias V_f, given from eqns. (4-97), (4-96) and (4-86)

$$W' = \left[\frac{2\epsilon}{|q|}\left(\frac{1}{N_A}+\frac{1}{N_D}\right)\right]^{1/2}\left(\phi_o - V_f\right)^{1/2} \qquad (4\text{-}123)$$

$$\varepsilon'_m = \frac{|q|N_A}{\epsilon}\left|x'_{p_o}\right| = \frac{|q|N_D}{\epsilon}x'_{p_o} = \frac{|q|}{\epsilon}\frac{1}{\left[\frac{1}{N_A}+\frac{1}{N_D}\right]}W' \qquad (4\text{-}124)$$

Where x'_{n_o} and $\left|x'_{p_o}\right|$ are the new values of x_{n_o} and $\left|x_{p_o}\right|$ under forward bias (Fig. 4.16) and are given for eqns. (4-91) and (4-92) by

$$x'_{n_o} = \frac{W'}{1+\frac{N_D}{N_A}} \qquad (4\text{-}125)$$

$$\left|x'_{p_o}\right| = \frac{W'}{1+\frac{N_A}{N_D}} \qquad (4\text{-}126)$$

Under these conditions, we may rewrite eqns. (4-96) and (4-97) for $\left|x'_{p_o}\right| > x' > 0$ as

$$\varepsilon'(x') = -\frac{|q|N_A}{\epsilon}x', \; x'_{p_o} > x' > 0 \qquad (4\text{-}127)$$

$$\Delta V'(x') = \frac{1}{2}\frac{|q|N_A}{\epsilon}(x')^2, \; \left|x'_{p_o}\right| > x' > 0 \qquad (4\text{-}128)$$

and we may rewrite eqns. (4-108) and (4-109) for $W > x' > \left|x'_{p_o}\right|$ as

$$\varepsilon'(x') = -\frac{|q|N_D}{\epsilon}(W'-x'), \; x' > \left|x'_{p_o}\right| \qquad (4\text{-}129)$$

$$\Delta V'(x') = \left(\phi_o - V_f\right) - \frac{1}{2}\frac{|q|N_D}{\epsilon}\left(W'-x'\right)^2, \; x' > x'_{p_o} \qquad (4\text{-}130)$$

We then proceed to calculate the various components of current (Prob. 4.2).

Ex. 4.5

Propose circuits using varactors in

a) electronic tuning
b) generation of FM signals

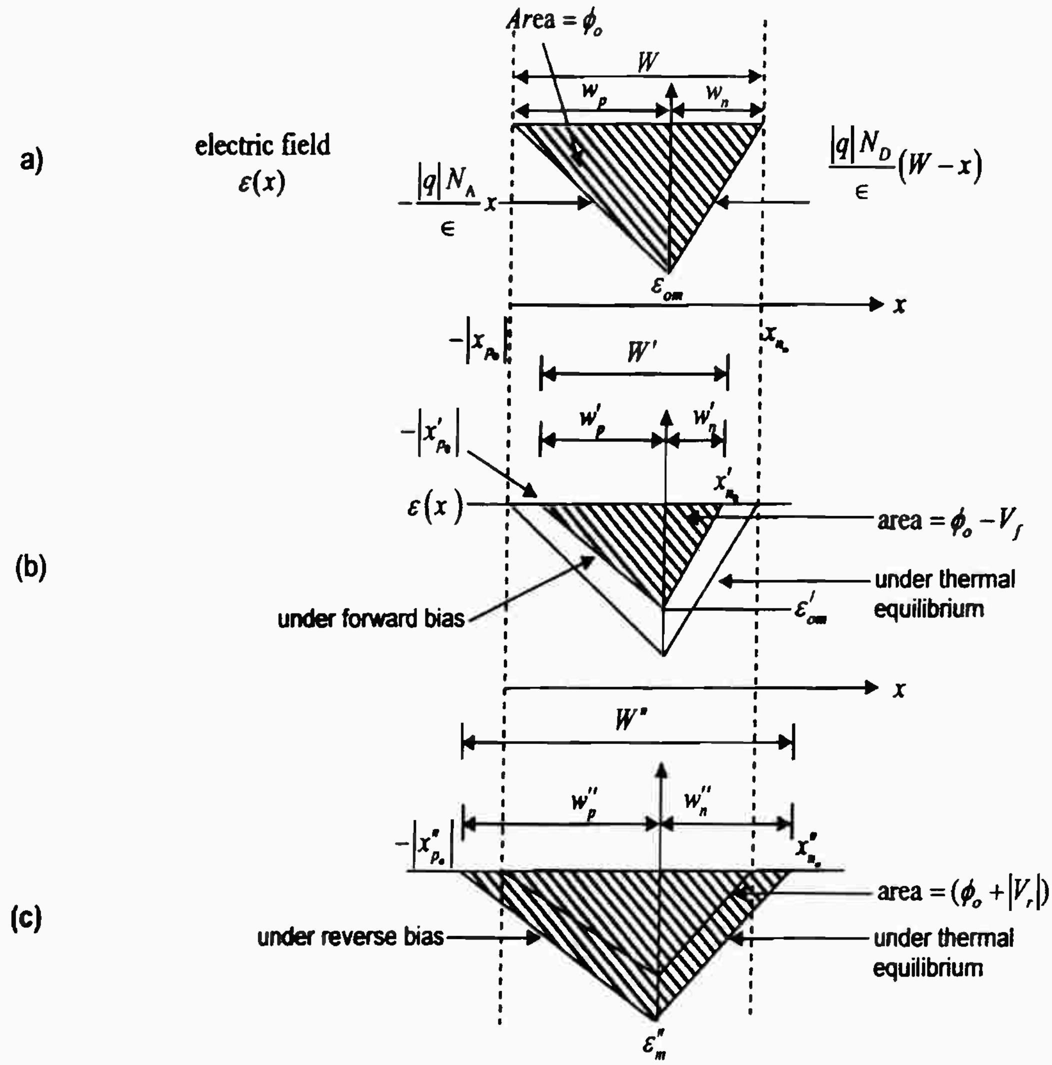


Fig. (4.16) Electric field distribution

a) under thermal equilibrium
b) under forward bias
c) under reverse bias

Solution

a) The circuit shown (Fig. 4.17a) is a tunable resonator used as an electronic tuning system. The capacitor C_c is called blocking capacitor. It is open circuited at dc (blocking dc) and is short circuited for high frequency input voltages or currents. From the signal point of view, the circuit reduces to a parallel resonant (tank) circuit (Fig. 4.17b). The value of the capacitance is C_T and the resonant frequency f_o

$$f_o = \frac{1}{2\pi}\frac{1}{\sqrt{LC_T}} \qquad (4\text{-}131)$$

The value of C_T and hence f_o, may be varied by the potentiometer. The voltage applied to the varactor varies in the ratio R_1 / R_t. The value of R_s is usually $\gg R_t$. Thus, as the voltage is varied across the varactor, the resistance of the tank circuit is constant. The equivalent circuit is shown (Fig. 4.17b) with $R \cong R_s$. An FM modulator is shown (Fig. 4.17c).

b) From eqn. (4-101), we see that

$$C_T = KV^{-n} \tag{4-132}$$

Where K is a constant and $1 > n > 0$. In the case of an abrupt junction., $n = 1/2$. We assume that C_c is a short circuit at high frequencies.

From eqns. (4-131) and (4-122), with $V_d = |V_r| + v_m$, where V_r is a dc reverse voltage and υ_m is a small ac signal, we get

$$f_o = \frac{(|V_r| + v_m)^{n/2}}{2\pi\sqrt{LK}} = \frac{|V_r|^{n/2}\left(1 + \frac{v_m}{|V_r|}\right)^{n/2}}{2\pi\sqrt{LK}} \tag{4-133}$$

We want f_o to vary linearly with v_m. We assume $\left|\frac{v_m}{V_r}\right| << 1$, and using Taylor's series, we get

$$f_o = \frac{|V_r|^{n/2}}{2\pi\sqrt{LK}}\left[1 + \frac{nv_m}{2|V_r|} + \frac{n}{4}\left(\frac{n}{2} - 1\right)\left(\frac{v_m}{|V_r|}\right)^2 + \dots\right] \tag{4-134}$$

For $\frac{v_m}{|V_r|} << 1$, we get

$$f_o = \frac{|V_r|^{n/2}}{2\pi\sqrt{LK}}\left[1 + \frac{n}{2}\frac{v_m}{|V_r|}\right] \tag{4-135}$$

For $v_m = V_m sin\,\omega_m t$, then f_o becomes an instantaneous frequency f_i, given by

$$f_i = \frac{|V_r|^{n/2}}{2\pi\sqrt{LK}}\left(1 + \frac{n}{2}\frac{V_m}{|V_r|}\sin\omega_m t\right) \tag{4-136}$$

Thus, the tank circuit has become one of a time varying resonant frequency. Such a tank circuit may be used to generate a sinusoidal signal whose frequency is itself a sinusoidal function of time controlled by an input signal v_m. This scheme is called frequency modulation (FM). The control signal is called the modulating signal and the circuit is called an FM modulator.

Ex. 4.6

Sketch the pseudo Fermi levels of a forward biased pn junction, and hence obtain an expression for the product np.

Solution

We rewrite eqns. (4-28) and (4-29)

$$n = n_i e^{(F_n - E_{F_i})/kT} \tag{4-137}$$

$$p = n_i e^{(E_{F_i} - F_p)/kT} \tag{4-138}$$

From eqns. (4-33) and (4-34)

$$J_p = \mu_p p \frac{dF_p}{dx} \tag{4-139}$$

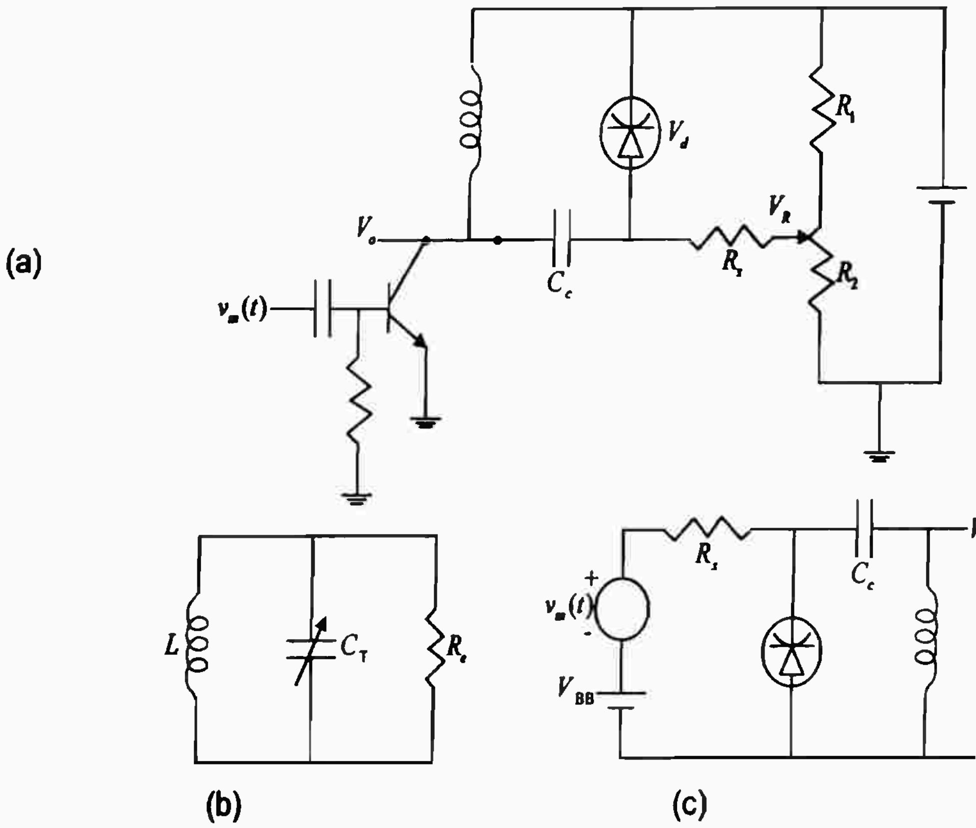


Fig. (4.17) Varactor circuits
a) electronic tuning circuit b) equivalent tank circuit c) FM modulator

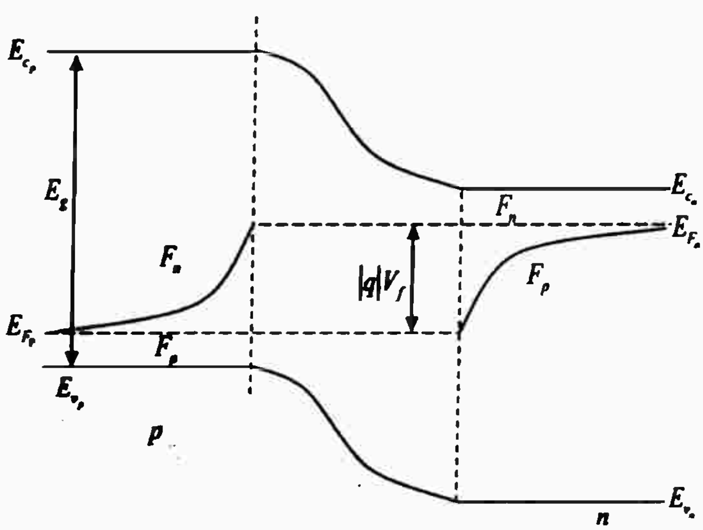


Fig. (4.18) Spatial dependence of the pseudo Fermi levels of a forward biased pn junction

$$J_n = \mu_n n \frac{dF_n}{dx} \quad (4\text{-}140)$$

We define

$$F_p = -|q|\psi_p \quad (4\text{-}141)$$

$$F_n = -|q|\psi_n \quad (4\text{-}142)$$

where ψ_p *and* ψ_n indicate the potential corresponding to the pseudo Fermi levels in the p and n regions, respectively. The product np is given by

$$np = n_i^2 e^{(F_n - F_p)/kT} \quad (4\text{-}143)$$

$$= n_i^2 e^{|q|(\psi_p - \psi_n)/kT}$$

For a forward bias junction $\psi_p - \psi_n > 0$ and $np > n_i^2$. We may rewrite eqns. (4-139) and (4-140) as (Prob. 4.10)

$$J_p = -|q|\mu_p p \frac{d\psi_p}{dx} = -\sigma_p \frac{d\psi_p}{dx} \quad (4\text{-}144)$$

$$J_n = -|q|\mu_n n \frac{d\psi_n}{dx} = -\sigma_n \frac{d\psi_n}{dx} \quad (4\text{-}145)$$

This formalism bears a resemblance to the ohmic relation

$$J = \sigma\varepsilon = -\sigma \frac{dV}{dx} \quad (4\text{-}146)$$

Thus ψ plays the same role as V, but takes into account the diffusion effect. At thermal equilibrium $\psi_p = \psi_n =$ constant and

$$J_n = J_p = 0 \quad (4\text{-}147)$$

Inside the depletion region both p and n vary by many orders of magnitude. Since J_p and J_n must stay at very low values, the only way to guarantee this is by keeping $\frac{dF_p}{dx}$(or $\frac{d\psi_p}{dx}$) and $\frac{dF_n}{dx}$(or $\frac{d\psi_n}{dx}$) close to zero. This is equivalent to saying F_n $(or \psi_n)$ are nearly constant neglecting generation and recombination in the depletion region. Outside the depletion region, F_p(or ψ_p) and F_n(or ψ_n) exhibit a spatial dependence, such that J_n and J_p are kept at small values. Fig. (4.18) shows this dependence. If the voltage on the pn junction is V_f, then

$$\psi_p - \psi_n = V_f \quad (4\text{-}148)$$

From eqns. (4-143) and (4-148) we get

$$np = n_i^2 e^{|q|V_f/kT} \quad (4\text{-}149)$$

4.9 Breakdown and Zener Diode

From the IV characteristic of a pn junction, we have seen that under a reverse voltage, a reverse current I_o flows. Since this current is very small, it is usually assumed that the pn junction is practically an open circuit (very large resistance). When the reverse voltage is sufficiently increased, a point is reached at which breakdown takes place. There are two mechanisms for breakdown. The first is called Zener breakdown. In this mechanism, the maximum electric field (as given by eqn. (4-86)) reaches the point of the dielectric strength of the material. Covalent bonds, at the weakest point in the transition region (where the electric field is maximum) are torn asunder. Electrons and holes are violently generated by the action of the electric field. Once breakdown occurs, it propagates in the depletion region. The depletion region all of a sudden becomes a source of electron-hole pairs generated within. Therefore, I_o increases dramatically almost without any further increase in external voltage. As a result, the current increases without limit at a constant value of voltage, called Zener voltage V_Z. Unless an external limiting (series) resistance is used, the diode burns out.

The second mechanism for breakdown (called avalanche breakdown) entails a chain reaction of electrons liberated in breakdown, causing a secondary emission of electrons and holes from other bonds by imparting a large kinetic energy to electrons and holes bound in the bonds of atoms in the transition region. This action is multiplicative, since the librated carriers cause further breakdown.

The two mechanisms are usually concurrent and Zener voltage is indicative of both mechanisms. Such a breakdown does not cause permanent damage in the pn junction, unless the limiting resistor is left out. At breakdown, the internal resistance of the diode (called Zener resistance R_Z is very small). We cannot consider the diode under breakdown conditions as a short circuit, though, because of the presence of a constant dc voltage V_Z.In fact, a reasonable equivalent circuit for a diode under breakdown conditions must include an internal voltage source V_Z and an internal Zener resistance R_Z which may often times be considered zero. Fig. (4.19) shows an equivalent circuit together with the IV characteristic under breakdown.

A diode which is operative specifically under breakdown conditions is called Zener diode. Before breakdown is reached, a Zener diode is just a reverse biased pn junction. At breakdown, a Zener diode behaves practically as a constant voltage source. Current increases indefinitely and the voltage cannot be appreciably increased above V_Z. The Zener diode has ratings, which specify the maximum allowable power dissipation. One of the important applications of Zener diodes is the diode voltage regulator. In the circuit shown (Fig. 4.20), a Zener diode is used to regulate the voltage across a load resistor R_L. A series resistor R_s is used to limit the current in the circuit at breakdown.

It is important to look into the transfer function of the circuit. The transfer function is a relation between the output voltage V_o and the input voltage V_i. For a small positive input voltage, the Zener diode is considered an open circuit and the current flows in R_s and R_Z in series (Fig. 4.20).

Thus,

$$V_o = \frac{R_L}{R_L + R_S} V_i \quad \text{for} \quad V_o < V_Z \qquad (4\text{-}150)$$

For large positive values of V_i

$$V_o = V_Z \quad \text{for} \quad V_o \geq V_Z \qquad (4\text{-}151)$$

From eqns. (4-150) and (4-151), we see that a critical value of the input voltage V_{ic} exists, at which breakdown occurs, given by

$$V_{ic} = \left(1+\frac{R_S}{R_L}\right)V_Z \tag{4-152}$$

We see that the transfer characteristic must be linear for $V_o < V_Z$, as given by eqn. (4-150) and then saturates at V_Z for $V_i \geq V_{ic}$ Fig. (4.21)

For negative values of V_i, in excess of the cut-in voltage V_γ the Zener diode is forward biased.

We can use this transfer characteristic to obtain the output voltage as a function of time if the input voltage is sinusoidal. We may draw the input voltage versus time, with the time axis along the negative V_o direction. The output voltage is obtained point for point for two cases: $V_i < V_{ic}$ and $V_i > V_{ic}$.

In this model, we have neglected V_γ (Fig. 4.22). In the case when $V_i < V_{ic}$, we have only positive half cycles. For $V_i > V_{ic}$, those half cycles are clipped. We now analyze the current in this circuit.

$$I_R = \frac{V_i}{R_S + R_L} \quad , \quad V_i < V_{ic} \tag{4-153}$$

For $V_i > V_{ic}$, $V_0 = V_Z$, and hence

$$I_L = \frac{V_Z}{R_L} \quad \text{for} \quad V_o \geq V_Z \tag{4-154}$$

and

$$I_R = \frac{V_i - V_Z}{R_S} \quad \text{for} \quad V_o \geq V_Z \tag{4-155}$$

$$\begin{aligned} I_Z &= I_R - I_L \\ &= \frac{V_i - V_Z}{R_S} - \frac{V_Z}{R_L} \\ &= \frac{V_i}{R_S} - V_Z\left(\frac{1}{R_S}+\frac{1}{R_L}\right) \quad \text{for} \quad V_o \geq V_Z \end{aligned} \tag{4-156}$$

We conclude that the Zener diode will take up the difference between I_R and I_L and thus, the Zener diode diverts the excess current. As such, this scheme is considered as a protective circuit, while R_L represents a sensitive device such as an ammeter coil or a transducer.

We may conclude by considering the factors which determine the values of V_Z.

We should note from eqns. (4-90) and (4-97), that

$$|\varepsilon_{m}| = \left(\frac{2|q|}{\epsilon}\right)^{1/2} \frac{(\phi_o + |V_r|)^{1/2}}{\left(\frac{1}{N_A}+\frac{1}{N_D}\right)^{1/2}} \tag{4-157}$$

or

$$(\phi_o + |V_r|) = \left(\frac{1}{N_A}+\frac{1}{N_D}\right)\left(\frac{\epsilon}{2|q|}\right)\varepsilon_{rm}^2 \tag{4-158}$$

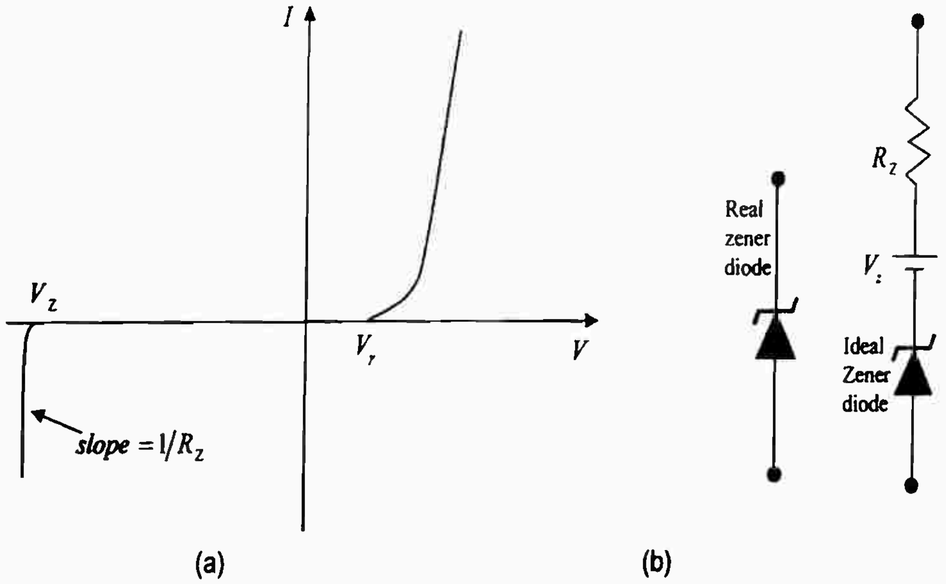


Fig. (4.19) Zener diode
a) IV characteristic b) symbol and equivalent circuit

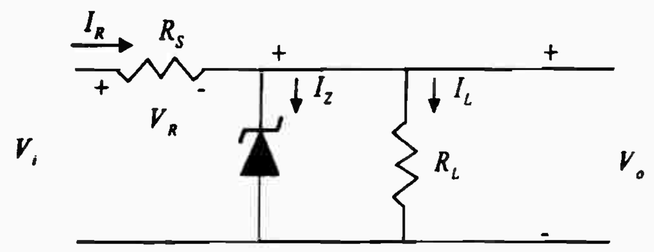


Fig. (4.20) Zener diode voltage regulator

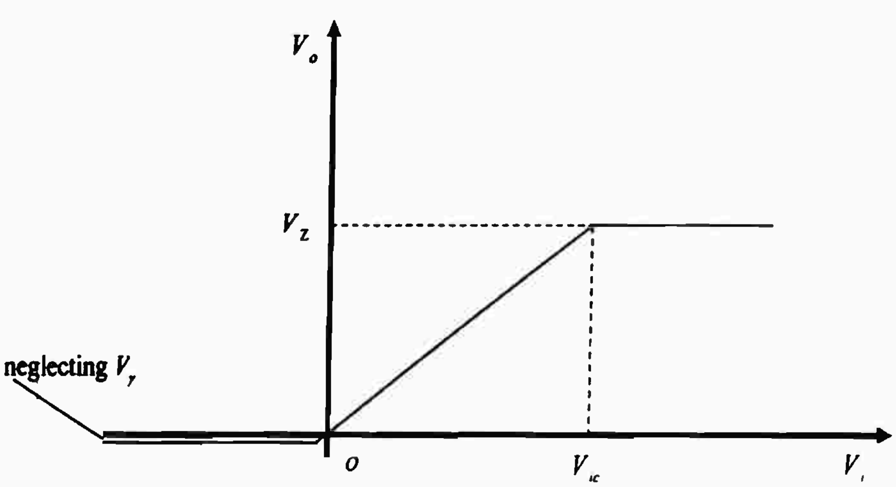


Fig. (4.21) Transfer characteristic for Zener diode voltage regulator

(a)

(b)

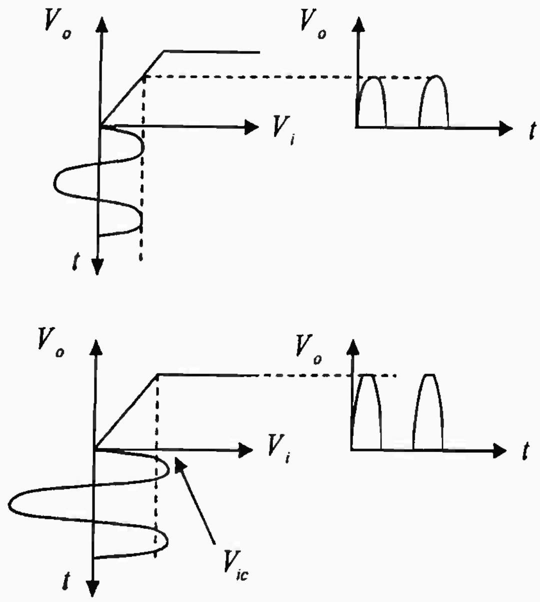


Fig. (4.22) Using the transfer characteristic to obtain the clipped output voltage as a function of time

a) $V_i < V_{ic}$ b) $V_i > V_{ic}$

At breakdown, V_r is V_Z and ε_m is the dielectric strength ε_{BD}, which is a constant of the material. Thus,

$$V_Z = \frac{\epsilon}{2|q|}\left(\frac{1}{N_A} + \frac{1}{N_D}\right)\varepsilon_{BD}^2 \quad \text{for} \quad V_Z > \phi_o \qquad (4\text{-}159)$$

Hence, the actual value of V_Z can be determined by controlling the doping concentrations of both sides of the junction. This explains why there are so many different values of V_Z for Si diodes, which all have the same breakdown field or dielectric strength.

4.10 Diode Circuits

One of the basic applications of pn diodes is rectification. Fig. (4.23) shows a half wave rectifier (HWR), a full wave rectifier (FWR) and a bridge rectifier (BR). Such circuits may be used in ac measuring instruments. The rms value is to be measured. Therefore, we use a FWR to rectify the ac signal. We measure the average current and obtain the rms by conversion (Prob. 4.15).

Another important circuit, called the rectifier capacitor input filter, is shown (Fig. 4.24). This is used to obtain a dc voltage from the rectified output waveform. The diodes D_1 and D_2 conduct alternately. As the input voltage increases toward the peak of the sine wave, a point is reached at which D_1 conducts while D_2 is still off. The capacitor is charged to the peak input voltage. As the input voltage decreases, the capacitor voltage cannot follow fast variations of the input. Thus, the anode voltage is higher than the cathode voltage, and D_1 becomes reverse biased (D_1 is off with D_2 still off). Now, the capacitor discharges into R rather slowly. In the negative half cycle, the input voltage increases in the negative direction, but the rectified output increases in the positive direction. A point is reached when diode D_2 conducts and the capacitor recharges.

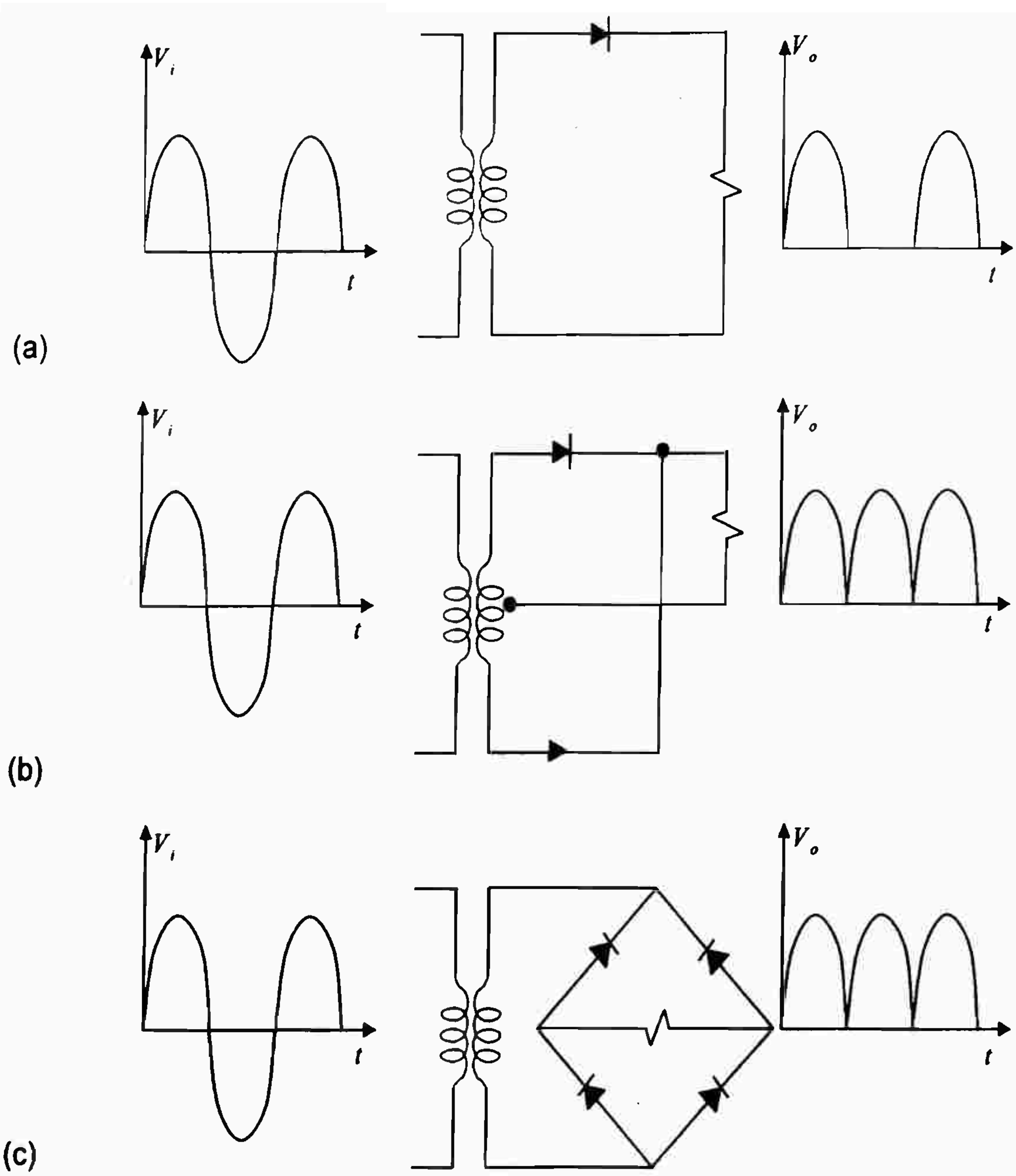


Fig. (4.23) Simple rectifier diode circuits
a) HWR b) FWR c) BR

The output waveform appears as an approximate sawtooth. The fluctuation around the mean voltage V_{dc} is called the ripple voltage V_{rip}. A ripple factor γ_r is defined as $V_{ripple\ rms}/V_{dc}$. The rms of the ripple voltage is given by (Prob. 4.13)

$$V_{ripple\ rms} = \frac{\Delta V_{rip}}{2\sqrt{3}} \qquad (4\text{-}160)$$

Where ΔV_{rip} is the full ripple swing.

We may now consider the dynamics of the charging and discharging of the capacitor. The charge gained by the capacitor within T_1 is $C\Delta V_{rip}$. The charge lost by the capacitor within T_2 is $I_{dc}T_2$. Where I_{dc} is the dc current due to the discharge in the load resistor and is the rate of the charge leakage from the capacitor into the load resistor.

For the waveform, to maintain a steady state, a balance must be maintained between the charge gained and the charge lost. Thus,

$$C\Delta V_{rip} = I_{dc}T_2 \tag{4-161}$$

$$\Delta V_{rip} = \frac{I_{dc}T_2}{C} \tag{4-162}$$

Since $T_2 \gg T_1$ we may assume $T_2 = \frac{T}{2} = \frac{1}{2f}$. The factor 2 comes about due to the fact that two rectified half cycles exist per period. We get

$$\Delta V_{rip} = \frac{I_{dc}}{2fC} \tag{4-163}$$

The average dc voltage V_{dc} is given in terms of the maximum input voltage V_m by

$$V_{dc} = V_m - \frac{\Delta V_{rip}}{2} = V_m - \frac{I_{dc}}{4fC} \tag{4-164}$$

This equation may be represented by a dc Thevenin's equivalent circuit (Fig. 4.24e)

The output resistance of the capacitor input filter is $\frac{1}{4fC}$. Thus,

$$I_{dc} = \frac{V_m}{R_L + \frac{1}{4fC}} \tag{4-165}$$

and

$$V_{dc} = I_{dc}R_L \tag{4-166}$$

$$= V_m \frac{R_L}{R_L + \frac{1}{4fC}} \tag{4-167}$$

or

$$V_m = \left(1 + \frac{1}{4fCR_L}\right)V_{dc} \tag{4-168}$$

We define the voltage regulation VR as

$$VR = \frac{V_m - V_{dc}}{V_{dc}} \tag{4-169}$$

$$= \frac{\Delta V_r/2}{V_{dc}} = \frac{I_{dc}}{4fCV_{dc}}$$

$$= \frac{1}{4fCR_L} \times 100\% \tag{4-170}$$

Thus, the regulation, which measures the drop in the output voltage as current is drawn, is deceased by increasing the time constant CR_L. We calculate now the ripple factor γ_r

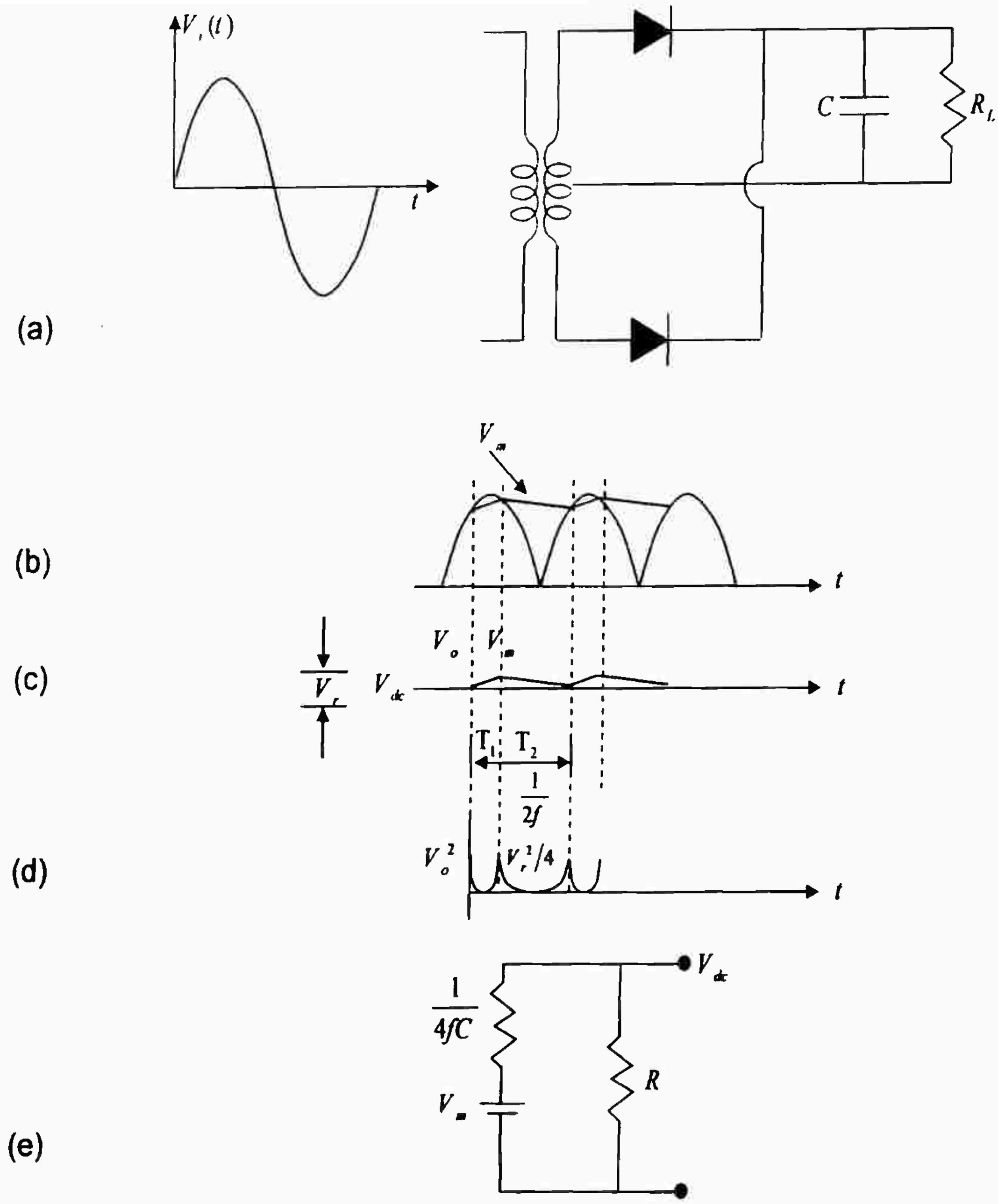


Fig. (4.24) Capacitor input filter

a) basic circuit
b) output waveform
c) linearized output voltage
d) square of the voltage
e) equivalent output circuit

$$\gamma_r = \frac{1}{\sqrt{3}}\frac{\Delta V_{rip}}{V_{dc}} = \frac{I_{dc}}{4\sqrt{3}fCV_{dc}} \tag{4-171}$$

$$= \frac{1}{4\sqrt{3}fCR_L} \times 100\% \tag{4-172}$$

As C is increased, γ_r is reduced, and we get a smoother dc voltage. The relation between the ripple factor and voltage regulation is given by

$$VR = \sqrt{3}\gamma_r \tag{4-173}$$

To smooth out the ripple, we may put an LC filter, called L section (or smoothing LC filter) (Fig. 4-25a). The output voltage between points a and a' is a approximately as given before. This voltage is now applied on the LC_2 section. The ripple is now reduced by a factor of $1/4\omega^2 LC_2$ for $2\omega L > 1/2\omega C_2$ and $1/2\omega C_2 \ll R$, noting that we are working with twice the frequency for a full wave rectifier. The new ripple factor becomes (Prob. 4.13).

$$\gamma_r = \frac{1}{4\sqrt{3}fC_1R_L\left(4\omega^2 LC_2\right)} \tag{4-174}$$

More than one section may be used to obtain an even smoother output. Another method for smoothing employs a Zener diode (Fig. 4.23b).
We must note that V_Z in this case must be lower than V_{min}

$$V_{min} = V_m - \Delta V_{rip}$$

$$= V_m - \frac{I_{dc}}{2fC}$$

$$= V_m - \frac{V_{dc}}{2CR_L} \tag{4-175}$$

It can be shown (Prob. 4.13) that the following condition must apply

$$V_Z < V_{dc}\left[1 - \frac{1}{4fCR_L}\right] \tag{4-176}$$

Another application for pn diodes is clipping circuits (Fig. 4.26a). From the transfer characteristic, we note that when the anode voltage exceeds the cathode voltage the diode is ON, otherwise it is OFF. The output voltage $V_o = 5V$ for as long as $V_i < 5V$ otherwise V_o follows V_i. For the input voltage shown (Fig. 4.26b) we get a clipped output waveform. There are different clipping circuits for different output waveforms. This is an important application called wave shaping, i.e., changing the shape of an input waveform to obtain tailored waveforms from a sinusoidal source.

Finally, a clamping circuit (Fig. 4.24c) may be included in this overview of diode applications. This circuit uses a capacitor and a diode. The capacitor charges to the peak of the input voltage. Thus, a dc voltage is produced on the capacitor, which serves as a level shifter. Thus, the output voltage is the same as the input voltage, shifted by the amount of the dc voltage developed on the capacitor. The dc shifting of the input waveform is useful in certain applications, such as oscilloscopes.

(a)

(b)

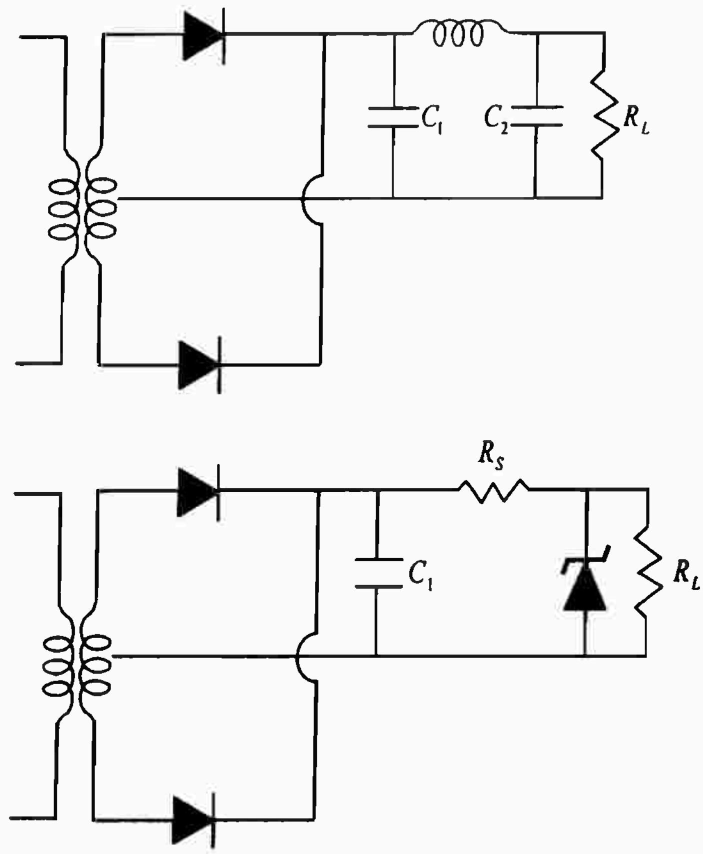


Fig. (4.25) Capacitor input filter, using
a) a smoothing filter
b) a Zener diode voltage regulator

4.11 Diode Ratings

The maximum power rating of a diode has a special significance. We must ensure that the diode does not get too hot, i.e., the power dissipation must be less than the maximum rating. Given the power rating of the diode, can we obtain the maximum current? We draw the hyperbola representing the maximum rating. We draw a tangent from the point V_{BB} to the hyperbola. The shaded area (Fig. 4.27) represents the working area and the tangent represents the highest slope of the load line or minimum R_L. If we decrease R_L any further, the load line will cross the maximum rating hyperbola at two points, which means that it will encounter regions of dissipation higher than the allowed rating of the device.

Thus, the power rating is an important parameter that has to be specified for a diode, and must not be exceeded. All parameters that affect the performance of a device are collected in a data sheet produced by the manufacturer. Typically, such parameters for a pn diode are:

V_{RM}: Peak reverse voltage (or peak inverse voltage). This is the absolute peak of the voltage that may be applied in reverse bias across the diode.

V_{BR}: Reverse breakdown voltage. This is the minimum reverse voltage at which the device may breakdown.

I_F (or I_D): Steady state forward current. This is the maximum current that may be passed continuously through the diode. It is usually specified for 25°C and must be derated for higher temperatures.

$I_{FM_{surge}}$: Peak surge current. This current may be passed for the time period specified through a diode operating below a specified temperature. The surge current is much higher than the maximum forward current and may flow briefly upon turn on.

I_R (or I_o): Static reverse current (also called leakage current). It is the reverse saturation current for a specified reverse bias and maximum device temperature.

V_F (or V_D): Static forward voltage drop. It is the maximum forward voltage drop for a given forward current and device temperatures

P_D : Continuous power dissipation at 25°C. It is the maximum power that the device can safely dissipate on a continuous basis in free air. This rating must be downgraded at higher temperatures and may be upgraded when the device is mounted on a heat sink.

C_t: Total capacitance. It is the maximum capacitance for a forward biased diode at a specified forward current.

t_{rr}: Reverse recovery time. It is the maximum time for the device to switch from ON to OFF. Since charges are stored and can be depleted by diffusion, this recovering time puts a limitation in the switching speed of all devices based on charge transport and diffusion such as pn diode and BJT transistors.

It is very important that these ratings not be exceeded, otherwise failure takes place. For reliability, the absolute maximum ratings should not even be approached. Also, the maximum ratings must be downgraded for operations at higher temperatures. A factor of safety from the upper ratings is usually employed.

4.12 Metal Semiconductor Contact

When a metal semiconductor contact is made, two types of situations arise, namely, rectifying (non ohmic) and ohmic contacts, depending on a number of considerations. Diodes based on metal semiconductor contacts are called Schottky diodes. They are known to have shorter switching times, since the minority carrier diffusion is not a major factor in the injection mechanism hence the charge storage delay is eliminated. Here, the current is basically determined by thermionic emisision of electrons from the abundant metal reservoir into the semiconductor. Thermionic emission accounts for release of electrons from a metal depending on temperature. In this case the Maxiwellian tail dumps electrons directly into the semiconductor (Prob. 4.18). This feature makes these diodes appropriate for high speed switching. A detailed analysis of Schottky barriers is left out and will not be attempted here.

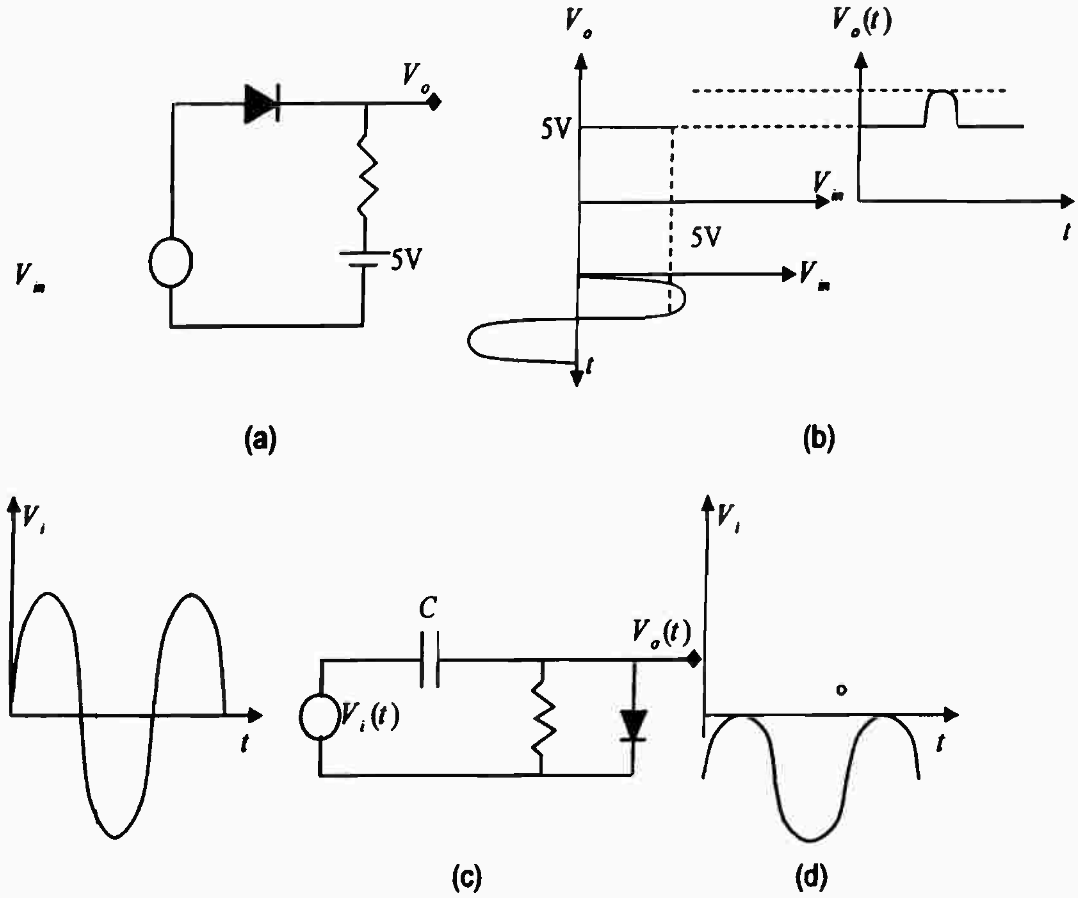


Fig. (4.26) Diode wave shaping circuits
a) clipping circuit b) clipped waveform
c) clamping circuit d) clamped waveform

Ohmic contacts can be ensured in situations where the semiconductor is heavily doped in the contact region. One important consideration has to be remembered, however, namely that at the metal semiconductor junction, Fermi level must be aligned at equilibrium. This brings about local band bending and sets in internal built-in fields and potential distributions. But the battery voltage remains equal to the difference between Fermi levels divided by the electronic charge.

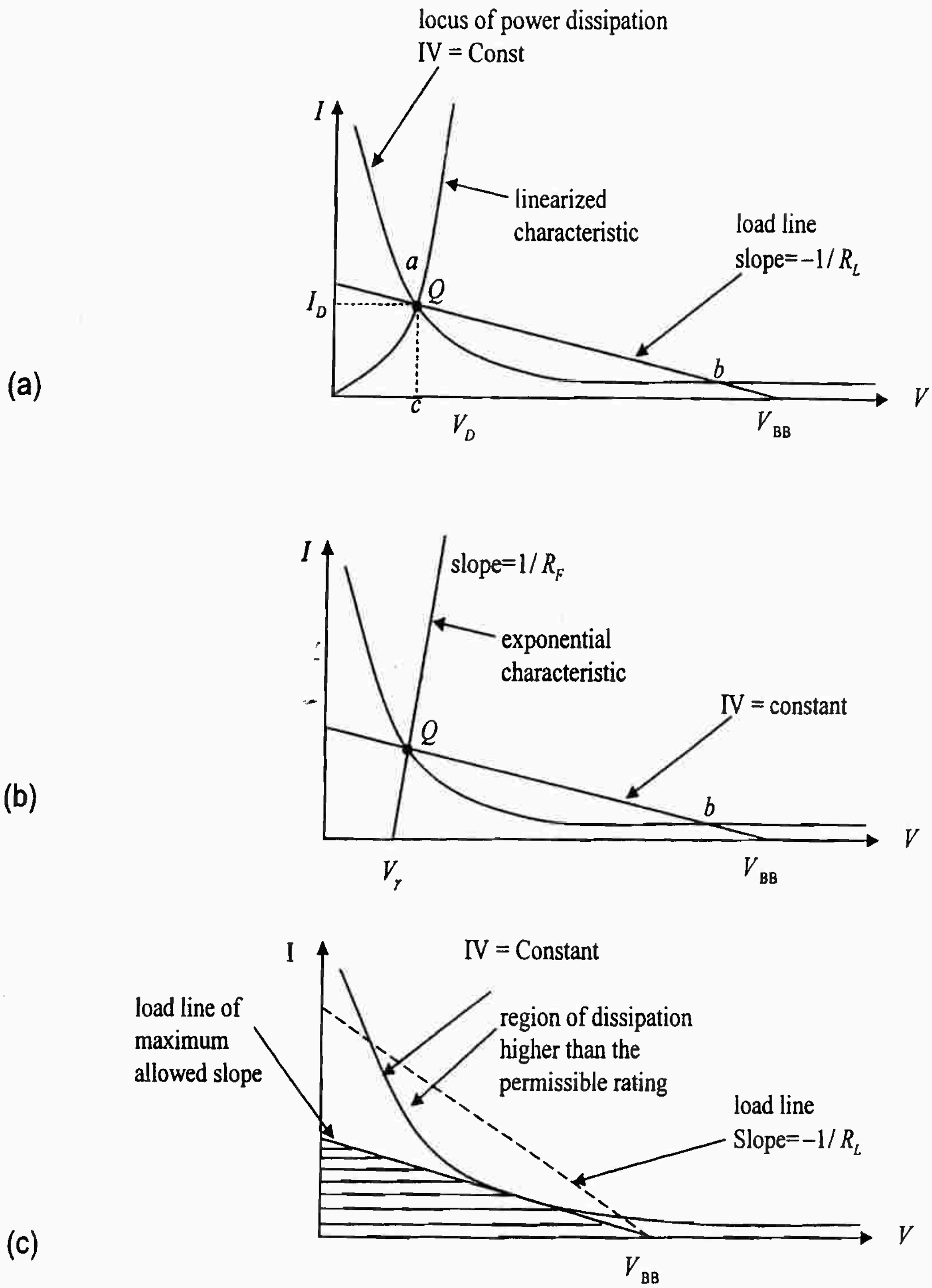


Fig. (4.27) Biasing and power rating of a pn diode and the concepts of load line, Q-point and power rating for

a) load line for exponential characteristic
b) load line for linearized characteristic
c) load line for maximum permissible power dissipation

Problems

1- a) Show that the contact potential in a pn junction is intimately related to the bandgap.

b) The resistivities of the two sides of an abrupt Si diode are $2\Omega cm$ (p-side) and $1\Omega cm$ (n-side). Calculate the height of the potential barrier. Take $\mu_n = 1300\ cm^2/Vs$ and $\mu_p = 500\ cm^2/Vs$.

2- a) Obtain approximate expressions for the hole and electron distributions inside the depletion region under forward bias.

b) Obtain approximate expressions for diffusion and drift current components inside the depletion region under forward bias. Sketch the distributions of the current components.

c) What is the net current in this case?

3- a) Prove that the ratio of hole to electron currents crossing a pn junction is given by $\frac{I_{pn}(o)}{I_{np}(o)} = \frac{\sigma_p L_n}{\sigma_n L_p}$,

Hence, show what happens in a p^+n junction

b) Show that $I_o = AV_{th} \frac{b\sigma_i^2}{(1+b)^2}\left(\frac{1}{L_p\sigma_n} + \frac{1}{L_n\sigma_p}\right)$, $b = \frac{\mu_n}{\mu_p}$

4- a) Obtain expressions for the current components in a p^+n junction and sketch the results.

b) Sketch the steady state carrier distributions and the current components under reverse bias for both pn and p^+n junctions

5- a) Consider a diode biased in the forward direction. Find $\frac{dI}{dT}$

b) Calculale the factor by which the current will be multiplied when the temperature is increased by 1°C and also by increasing temperature from 25°C to 50°C.

6- a) Prove that for a pn junction with $N_A \ll N_D$, $W = \left(\frac{2\varepsilon\mu_p\psi_o}{\sigma_p}\right)^{1/2}$ on which side does the depletion region extend more?

b) Show that under such conditions, $C_T = 2.9\times10^{-4}\left(\frac{N_A}{|V_r|}\right)^{1/2} pF$

7- a) Propose a measurement procedure for hole lifetime, and electron lifetime.

b) Assuming that the injected concentration varies with an angular frequency ω $P_n(x,t) = P_n(x)^{e^{j\omega t}}$.

Write the continuity equation in this case.

c) Show that the ac solution can be obtained by replacing L_p by $L_p(1+j\omega\tau_p)$

8- a) If $N_D = 10^{14}\ cm^3$, and $n_i = 1.5\times10^{10}$, calculate the potential barrier at room temperature

b) Calculate the width of the depletion region, then calculate the diffusion currents and the drift currents at thermal equilibrium. Estimate the electric field in the depletion region at thermal equilibrium. Take $D_n = 34\ \mathrm{cm}^2/s$, $D_p = 13\ \mathrm{cm}^2/s$, $\mu_n = 1300\ \mathrm{cm}^2/Vs$ and $\mu_p = 500\ \mathrm{cm}^2/Vs$.

9- a) Show how the contact potential, bandgap and cut in voltage are related

b) Obtain an expression for C_T and ε_m for a p^+n junction

10- Verify eqn. (4-144).

11- Verify eqn. (4-158).

12- Verify eqn. (4-160).

13- Design a dc power supply using a capacitor input filter for an output voltage of 5V.

a) using an LC smoothing filter

b) using a Zener diode The input is 220 V, 50Hz. The ripple factor is 1%.

14- a) The form factor is the ratio of rms to the average value of the waveform. Calculate the form factor for a FWR output due to a sine wave input.

b) Design a circuit of an ac voltmeter, which actually measures the average value and reads out the rms.

15- a) Design a clipping circuit, which converts a sine wave to a near square wave.

b) Design a clamping circuit, which shifts an ac voltage into all positive values.

16- Analyze the circuits of Fig. (4.17).

17- Show how recovery time limits the speed of switching in a pn diode or a BTT.

18- Using a simple model based on Maxiwellian tail discuss the operation and I-V characteristic of a Shottky diode.

19- Design a circuit that converts a sine wave into a square wave.

20- Propose an experimental method to measure the internal contact potential of a pn junction.

References

1- "Microelectronics. Circuit analysis and Design", D. Neaman, 3rd edition, McGraw Hill, N. Y., 2007.

2- "Semiconductor Devices", S. M. Sze, 3rd edition, Wiley, N. Y., 2006.

3- "Solid state Electronics Devices", B.G. Streetman, S. K. Banerjee, 6th edition. Pearson-Prentice Hall, Upper Saddle River., N. J., 2006.

4- "Electronic Devices and Circuit Theory", R. Boylestad, L. Nashelsky, 8th edition, Prentice Hall, Upper Saddle River, N. J., 2002.

5- "Principles of Electronic Materials and Devices", S. Kasap, 2nd edition, McGraw Hill, N. Y., 2002.

6- "Semiconductor Devices and circuits", T. Bogart, 2nd edition, Maxwell Macmillan, N. Y., 1991.

7- "Physical Foundations of Solid State and Electron Devices", A. M. Ferendeci; Mc Graw Hill, 1991.

8- "Electronic Devices", S. R. Fleeman; Prentice Hall, Englewood Cliffs, N.J., 1990.

9- " Electronic Materials Science " , J. W. Maya; S. S. Lau, Maxwell Macmillan N. Y., 1990.

10- "Semiconductors and Electronic Devices" A. Barlev, Prentice Hall International, London, 1979.

11- "Solid State Physical Electronics" A. Van der Ziel, Prentice Hall, Englewood Cliffs. N. J. 1976.