2.2.12 Depletion region and depletion capacitance

It is apparent from figure that the small area of the depletion layer situated to the left of the junction and the small area of situated to the right of the junction are devoid of any mobiles charge but these contain immobile charge due to impurity atoms.

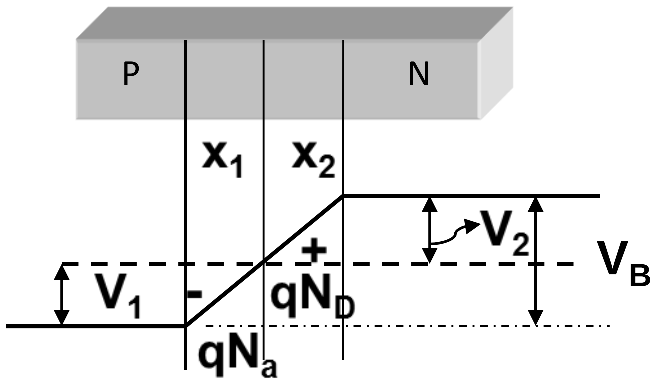


Fig. (2 – 30) Potential barriers in pn junction

On one side of junction (p-region) is a negative immobile charge due to negative ions of the acceptor atoms and on other side of the junction (n-type) there is an exactly equal amount of positive immobile charge due to the positive ions of donor Referring to the figure, the width of the depletion layer near junction is equal to, $X_m = X_p + X_n$, Since charge neutrality of crystal as a whole is maintained, the amount of charges in the area to the left and the right of the junction are equal. It can be rewritten mathematically as,

$$qAN_AX_p = qAN_DX_n (2.48)$$

This equation is valid even when p- and n-regions are doped to different degrees up to certain limits. Thus, when n-region is doped more than p-region, a large area of depletion layer will be situated in the p-region. For example, when density of donor atoms in n-region is ten times larger than that of acceptor atoms in p-region, the width of the depletion layer situated in the p-region will be ten times larger than that in the n-region. The effective area of depletion layer can easily be calculated with the help of Poison's equation which states that the second derivatives of the potential with respect to the distance is proportional to the charge density. In the simple one directional case when the voltage varies only in the x-direction along the length of the crystal, the Poison's equation may be written as,

$$\frac{\partial^2 V\,(x)}{\partial x^2} = \frac{\rho\,(x)}{\epsilon_r \epsilon_o} \quad (2.49)$$

Where, V(x) is the voltage at x, ρ(x) is the charge density at x. In the portion of the depletion layer situated in the p-region near the junction, figure (2-30), the charge density may be given by,

$$\rho = -q\,N_a \quad (2.50)$$

The negative sign is put since the acceptor atoms are negatively ionized. Substitution in equation (2.49) , we can get the depletion layer width in the p-region as,

$$\frac{\partial^2 V(x)}{\partial x^2} = \frac{q\, N_a}{\epsilon_r \epsilon_o} (2.50)$$

In similar way the Poisson's equation may be applied for the portion of the depletion layer situated in the n-region near the junction. Considering the charge density in this region is due to the positively ionized donor atoms, we may write Poisson's equation for this portion of the depletion layer as,

$$\frac{\partial^2 V(x)}{\partial x^2} = \frac{q\, N_d}{\epsilon_r \epsilon_o} (2.51)$$

Equation (2.49) can be solved for the particular charge distribution of interest to obtain the depletion region width as a function of the junction voltage. Two approximations for charge distribution can be used to cover most examples.

The abrupt or step junction approximation is used when there is an abrupt change in impurity concentration at the junction and the impurity concentration on either side of the junction is fairly uniform. Alloy junctions are usually approximated by an impurity concentration by abrupt junction distribution. Diffused junctions are usually approximated by a linear charge distribution in which the charge varies linearly through the depletion region near the junction (linearly graded junction). Solution of Poisson's equation for the abrupt approximation results in,

$$x_p = \sqrt{V \left(\frac{2\, \epsilon_r \epsilon_o}{q}\right) \left(\frac{N_D}{N_A N_D + N_A^2}\right)} (2.52)$$

And,

$$x_n = \sqrt{V\left(\frac{2\,\epsilon_r\epsilon_o}{q}\right)\left(\frac{N_A}{N_A N_D + N_D^2}\right)} \quad (2.53)$$

Adding the two thicknesses in equations 2.52 , and 2.53, we obtain the full width of the depletion regions as,

$$x_m = \sqrt{V\left(\frac{2\,\epsilon_r\epsilon_o}{q(N_A + N_D)}\right)}\left[\sqrt{\frac{N_A}{N_D}} + \sqrt{\frac{N_D}{N_A}}\right] \quad (2.54)$$

For the homogenously doped crystal as in equation before, the depletion region width varies as the square rote of the junction voltage. For a numerical example, let us take a germanium crystal where ϵ_r is to be equal 16 and say $N_D = N_A = 10^{15}$ /cm^3, and ϵ_o is 8.87 x 10^{-15} F/cm. and the built-in voltage is 200 mV at no biasThe width of the depletion region is then 8 x 10^{-5} cm.

There is a special case of doping levels, which actually occur often in semiconductor devices. In many cases one side of the junction has a much higher impurity concentration than the other (denoted n^+-p or p^+-n) as is true for an alloy junction. Equations (2.52) and (2.53) can be simplified as follow, to give an approximate value for the depletion region width x_m.

For n^+-p junction, where $N_D > N_A$,

$$x_m = x_p \cong \sqrt{V\left(\frac{2\,\epsilon_r\epsilon_o}{qN_A}\right)} \quad (2.55)$$

For p^+-n junction, where $N_A > N_D$,

$$x_m = x_n \cong \sqrt{V\left(\frac{2\,\epsilon_r\epsilon_o}{qN_D}\right)} \quad (2.56)$$

Note that, when one side of junction is heavily doped, the depletion region expands almost entirely into the lightly doped region. In case of linearly graded junction approximation depletion region width given by,

$$x_m = x_p + x_n = \left[\frac{12\,V\,\epsilon_r\epsilon_o}{q\,a}\right]^{1/3} \quad (2.57)$$

Where, a, is the slope of the linear impurity distribution. The capacitance of the depletion region is given by the same equation that is used to determine the capacitance of a parallel plate capacitor, because for a small voltage increase we add charges at the boundaries. This is the so-called junction capacitance.

$$\frac{\partial Q}{\partial V} = C'_T = \frac{\epsilon_r\epsilon_o}{x_m} \quad (2.58)$$

Where, C'_T is the depletion region capacitance per unit area F/m^2, this capacitance depends on the applied voltage, dielectric constant, area and doping levels. From equations (2.49) and (2.54), we have,

$$C'_T = \sqrt{\left(\frac{q\,\epsilon_r\epsilon_o(N_A + N_D)}{2V}\right)}\left[\sqrt{\frac{N_A}{N_D}} + \sqrt{\frac{N_D}{N_A}}\right] \quad (2.59)$$

Note that, for homogeneously doped crystals 1/C^2 is a linear function of the applied voltage. A plot of 1/C^2 versus applied voltage is linear and extrapolates to V_T. This is shown in figure (2-31). We noted that for forward biases, a net voltage V_T = - V_o , and it seems that one could reach zero depletion region width and infinite capacitance, and it is not possible because in the forward direction the current becomes very large and the forward bias remains less than the built-in voltage. The junction voltage, V, is given by,

$$V = V_T - V_o \qquad (2.60)$$

Where, V_o is the applied junction voltage and V_T is the zero bias built-in voltage which calculated from,

$$V_T = \frac{KT}{q}\ ln\ \frac{N_D N_A}{n_i^2}$$

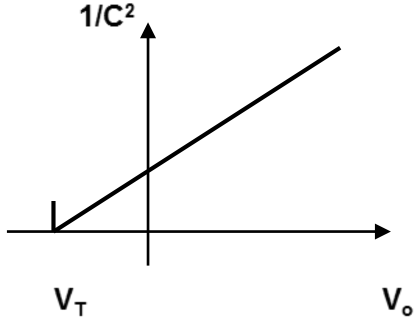


Fig. (2 - 31) A plot of junction capacitance versus reverse voltage

Junction capacitance decreases as reverse bias increases and as the impurity concentration decreases. The capacitance of heavily doped, abrupt junction is obtained by substituting equations (2.55) , and (2.56) into equation (2.58) . For n⁺-p junction, where $N_D >> N_A$,

$$C'_T = \sqrt{\left(\frac{q\ N_A \epsilon_r \epsilon_o}{2V}\right)} \, (2.61)$$

For p⁺-n junction, where $N_A > N_D$,

$$C'_T = \sqrt{\left(\frac{q\ N_D \epsilon_r \epsilon_o}{2V}\right)} \, (2.62)$$

The capacitance of a diffused junction can be approximated by substituting equation (2.59) into equation (2.58), giving,

$$C'_T = \left[\frac{q\ a\ \epsilon_r \epsilon_o}{12\ V}\right]^{1/3} (2.63)$$

2.2.13 Transition capacitance

The charge density distribution and the potential curves which are used to calculate the transition capacitance as shown in figure (2-32)

$$C_T = \frac{\partial Q}{\partial V_j} - \frac{\partial Q}{\partial V_B} = \frac{-\partial Q}{\partial W_B} - \frac{\partial W_B}{\partial V_B} \quad (2.64)$$

Where Q is the charge stored on either side of the junction, we have

$$\frac{-\partial Q}{\partial W_B} = - q\, N_A \quad (constant\ wrt\ W_p) \qquad (2.65)$$

Differentiation for equation (2.65) , we have,

$$\frac{\partial W_p}{\partial V_B} = \left[\frac{2\, \epsilon}{q\, N_A \left(1 + {}^{N_A}/_{N_D}\right)}\right]^{1/2} (V_B)^{-1/2} (2.66)$$

Combining equations (2.64) , (2.65) , and (2.66) , we obtain,

$$C_T = \left(\frac{q\, \epsilon\, N_A\, N_D}{2}\right)^{1/2} (N_D + N_A)^{-1/2} (V_B)^{-1/2} (2.67)$$

Note that, $V_B = V_D - V_J$, and for the reverse bias condition V_J is negative, while for the forward bias condition $V_J < V_D$, thus the junction capacitance does not become excessively large in the forward bias case. The case of graded junction may be worked out exactly as explained for the abrupt junction. The geometry of the junction is shown in figure (2-32). The charge density varies with distance in a linear manner. The result of this case is,

$$C_T = \epsilon^{2/3} \left(\frac{a}{12\, V_B}\right)^{1/3} \quad (2.68)$$

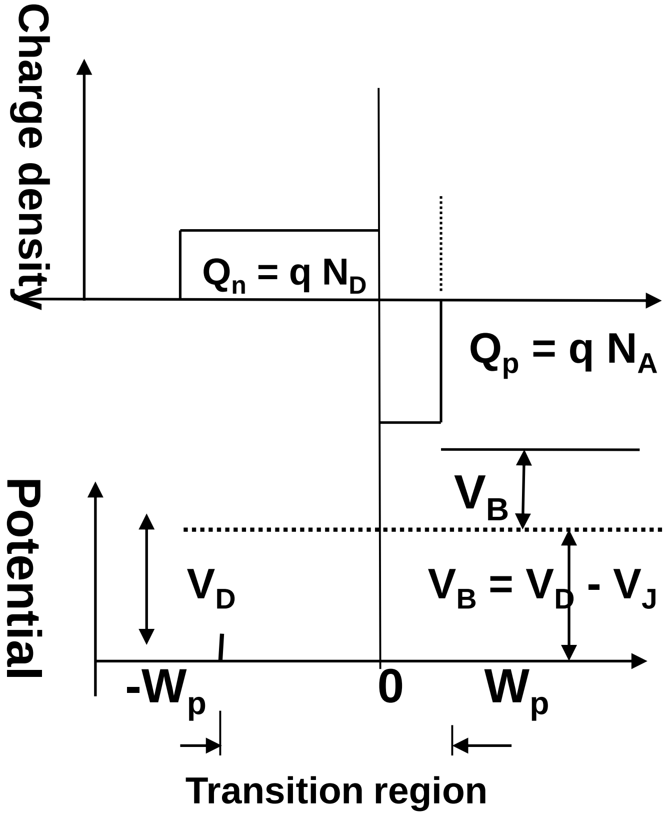


Fig. (2 - 32) Charge density distribution and the potential curves which are used to calculate transition capacitance

Where,

$$a = q\left(\frac{N_A + N_D}{W_I}\right) \quad (2.69)$$

And,

$$\rho\,(x) = a\,x \qquad (2.70)$$

Notice that, the junction is symmetrical, so that

2.2.14 Diffusion capacitance

When p-n junction is biased in the forward direction, the transition capacitance becomes larger, it is true but there is another effect which is more important, this is called storage or diffusion capacitance. With forward bias a large density of minority carriers is injected into the sample from the opposite side of the junction. This density is high near the junction and trails off as one move away from the transition region. The variations of carrier density with current are as shown in figure (2-33). There will be a tail of holes extending into the n-type region and tail of electrons extending into the p-type region.

Expression which describes this distribution in case of excess electrons is, from equation (2.72)

$$n_p(x) = n_{po}\left(exp\frac{qV_j}{KT} - 1\right)exp\frac{-x}{L_n} \quad (2.72)$$

Where, n_p (x) is excess carrier density only. It is total charge in this tail, which must be neutralized by majority carriers from external circuit, that contribute to storage or diffusion

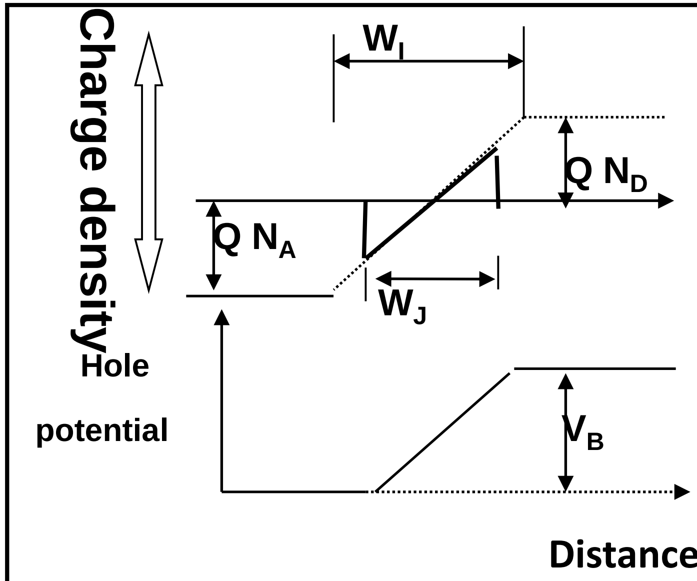


Fig. (2 - 33) Charge density and hole potential for a graded junction

capacitance, we may integrate charge density in the diffusion tail to obtain total charge as follows,

$$Q_n = -q\int_0^\infty n_p^+(x)dx = -q\int_0^\infty n_{po}\left(exp\frac{qV_j}{KT} - 1\right)exp\frac{-x}{L_n}dx = qn_{po}L_n\left(exp\frac{qV_j}{KT} - 1\right)exp\frac{-x}{L_n}dx\Big|_0^\infty$$

$$Q_n = -qn_{po}L_n\left(exp\frac{qV_j}{KT} - 1\right) \quad (2.73)$$

If the voltage is varied, we observe a differential capacitance due to the change in electron stored charge defined by,

$$C_n = -\frac{\partial Q_n}{\partial V_j} (2.74)$$

$$C_D = \frac{q^2}{KT}(n_{po}L_n + p_{no}L_p) exp\frac{qV_j}{KT} (2.75)$$

Note that the storage capacitance increases rapidly with forward bias. It is, in fact directly proportional to the current. We see that the capacitive effect which dominates in the forward bias case is due to the motion and storage of charge carriers, while the transition capacitance is due to a true displacement current just as in an ordinary capacitor the result for diffusion capacitance will be used when transistor equivalent circuits are discussed. It is therefore more convenient to put this result into the following form. C_n can be written as,

$C_n = \frac{g_n L_n^2}{D_n}$ **Where,**

$$g_p = \frac{qD_p p_{no}}{L_p}\left(\frac{q}{KT} exp\frac{qV_j}{KT}\right) = \frac{qL_p}{KT} (2.77)$$

Thus we can write,

$$C_D = \frac{g_n L_n^2}{D_n} + \frac{g_p L_p^2}{D_p} (2.78)$$

For simplicity consider abrupt junction, p-side is heavily doped, so that total current $I = I_{pn(0)}$,and excess minority charge Q will exists only in the n-side .

$$Q = \int_0^{\infty} qAp_{n(o)}\, e^{-x/L_p} dx = q\,A\,L_p p_{n(o)} \quad Eq. 2.79$$

And, $C_D = \frac{\partial Q}{\partial V} = qAL_p \frac{\partial p_{n(o)}}{\partial V} \qquad Eq. 2.80$

The hole current is $I = I_{pn(x)}$ at x = 0 , so:

$$I = \frac{q\,A\,D_p p_{n(o)}}{L_p} \quad Eq. 2.81$$

And,

$$\frac{\partial p_{n(o)}}{\partial V} = \frac{L_p}{q\,A\,D_p}\,\frac{\partial I}{\partial V}$$

$$= \frac{L_p}{q\,A\,D_p}\;g \quad Eq.2.82$$

Combining equations 2.80, 2.82, gives,

$$C_D = \frac{g\,L_p^2}{D_p} \qquad Eq.2.83$$

And since, $\tau_p = \frac{L_p^2}{D_p} = \tau \qquad Eq.2.84$

Or, $C_D = \tau_n g_n + \tau_p g_p (85)$

It will often be true that one current will dominate at the junction and hence the value of C_D will be simplified. If the hole current dominates, for example, we have,

$$C_D = \tau_p g_p = \tau_p g_{ac} (86)$$

From equations 2.83, 2.84, results:

$$C_D = g\,\tau \qquad Eq.2.87$$

From equation 2.64, we have,

$$\frac{1}{R} = g = \frac{I}{\eta\,V_T}$$

Substitute in equation 2.87 by 2.64, we get,

$$C_D = \frac{\tau\,I}{\eta\,V_T} \qquad Eq.2.88$$

From equation 2.55, $C_D \alpha$ I, and we have assumed that I is due to I_p only, then for accurate assumption, C_D in Eq.2.88 is C_{Dp} and similar expression can be determined for C_{Dn}, so, in general,

$$C_D = C_{Dp} + C_{Dn} \qquad Eq.2.89$$

For reverse bias, g is very small, and then C_D is neglected with respect to C_T. For forward bias, C_D is greater than C_T by millions of times, so C_T is neglected.

Example (59)

Calculate the diffusion capacitance of the diode at zero bias. Use μ_n= 1000 cm²/V-s, μ_p = 300 cm²/V-s, w_p' = 1 µm and w_n' = 1 mm. The minority carrier lifetime equals 0.1 ms.

A- For the same diode, find the voltage for which the junction capacitance equals the diffusion capacitance.

Solution

The diffusion capacitance at zero volts equals

$$C_{d(o)} = \frac{I_{s(p)}\, r_p}{V_t} + \frac{I_{s(n)}\, t_{r(n)}}{V_T} = 1,73\ x\ 10^{-19}\ F \quad ,$$

Using

$$I_{s(p)} = q\,\frac{A\, p_{no} D_p}{L_p} \quad \text{and} \quad I_{s(n)} = q\,\frac{A\, n_{po} D_n}{w_p^!}$$

Where the "short" diode expression was used for the capacitance associated with the excess charge due to electrons in the p-type region. The "long" diode expression was used for the capacitance associated with the excess charge due to holes in the n-type region0

The diffusion constants and diffusion lengths equal

$D_n = \mu_n \times V_t$ = 25.8 cm²/s

$D_p = \mu_p \times V_t$ = 7.75 cm²/s

$$L_p = \sqrt{D_p \tau_p}$$

And the electron transit time in the p-type region equals

$$t_{\gamma n} = \frac{w_p^2}{2\, D_n} = 193\ ps$$

The voltage at which the junction capacitance equals the diffusion capacitance is obtained by solving

$$\frac{C_{jo}}{\sqrt{1 - \frac{V_a}{\gamma}}} = C_{do}\, e^{V_a/V_T}$$

Yielding V_a = 0.442 V

2.2.15 Switching Time in p-n Diodes

When a diode is driven from reverse to forward state, an interval of time is required before the diode return to its steady state. Such time is called Forward Recovery Time t_{fr} and it is the time difference between the 10% of the diode voltage and the time when the voltage reaches to 90% of its final value. In case of forward bias, the densities of both minorities at the junction are the largest as shown in figure (2.34a). When the diode is reversed, the steady state of densities is zero at the junction figure (2-34b). The minority concentration reaches to its thermal equilibrium value $p_{n(0)}$, $n_{p(0)}$ far from the junction , it reaches such value after recovery time t_{rr} (Reverse Recovery Time) from

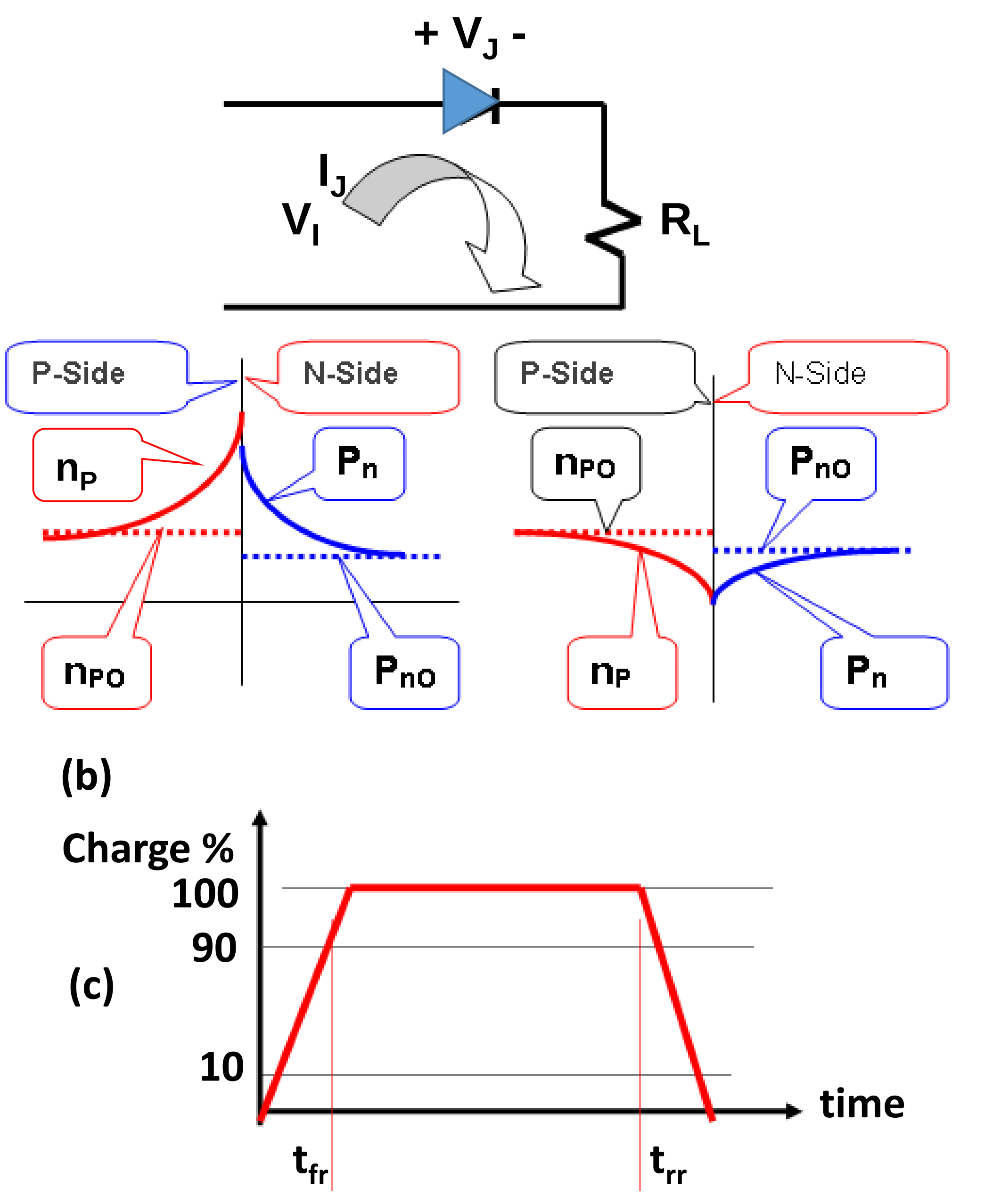
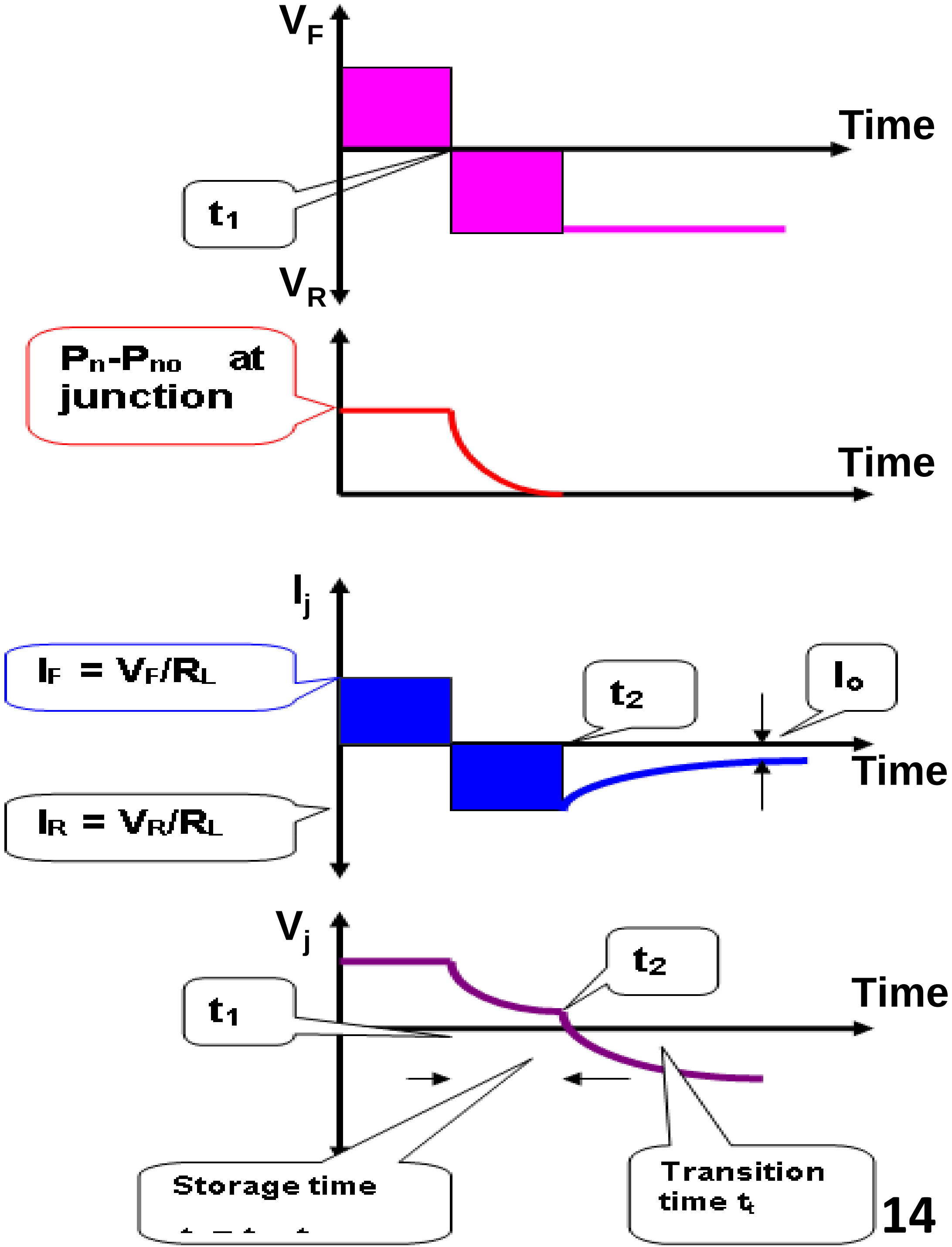


Fig. (2 – 34) Charge and Discharge through Junction

Fig. (2 – 35) Conduct diode sequence process

Figure (2-35) shows sequence accompanies reverse biasing of a conduct diode. If V_F is applied for time up to t_1, voltage across R_L is large compared to voltage across diode and equal $I_F = V_F/R_L$. At $t = t_1$, voltage is reversed $V = -V_R$, current I does not drop to zero, but it reversed and remain $i = - V_R/R_L = - I_R$ until time $t = t_2$. At $t = t_2$, injected minority at $x = 0$ has reached its equilibrium state. At t_1, diode voltage falls by $(I_f+I_R)\ R_L$ but does nor reversed. At $t = t_2$, voltage begins to reverse and current increased. Time t_1 to t_2 is called Storage Time t_s and time t_2 to the time when diode has recovered is called Transition Time t_t. Diode reaches its recovery time t_t when C_T



2.2.16 Reverse bias breakdown

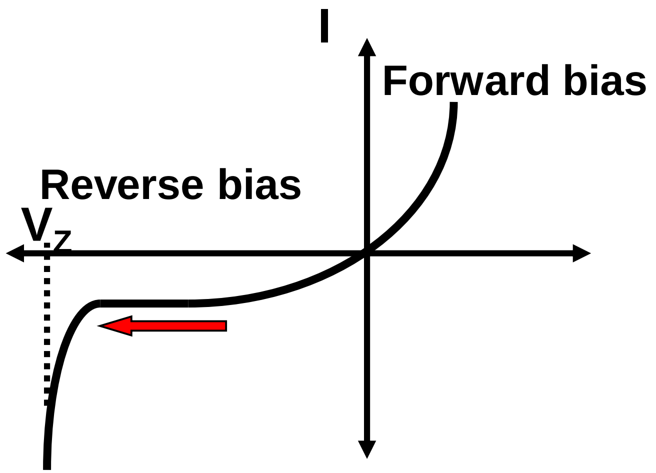


Fig. (2 - 36) diode Characteristics showing breakdown point

Ideal diode equation predicts a saturation current ($-I_o$) regardless of magnitude of reverse voltage. There are two factors which cause deviation from this theoretical saturated value. First is leakage current; is almost ohmic component of current due to leakage surface edges of junction. It will add to the saturation current, so that there will be a gradual increase in current as reverse bias is increased. The second factor is junction breakdown, which is shown in figure (2-36). Electrically, breakdown is observed as a sudden increase in reverse current at some voltage V_R. The sharpness of knee at breakdown and slope of the characteristics vary usually, current is limited only by external circuit resistance. There are two distinct mechanisms which can be used to explain breakdown. One of these is zener breakdown, and the other is avalanche breakdown.

Both will be discussed because in case of wider junction, avalanche breakdown is evident. It is possible to distinguish two forms, because in zener breakdown, breakdown voltage decreases as temperature increases, and in avalanche breakdown, breakdown voltage increases as the temperature increases. At first, zener effect was thought to be cause of all junction breakdowns. zener effect is described as internal field

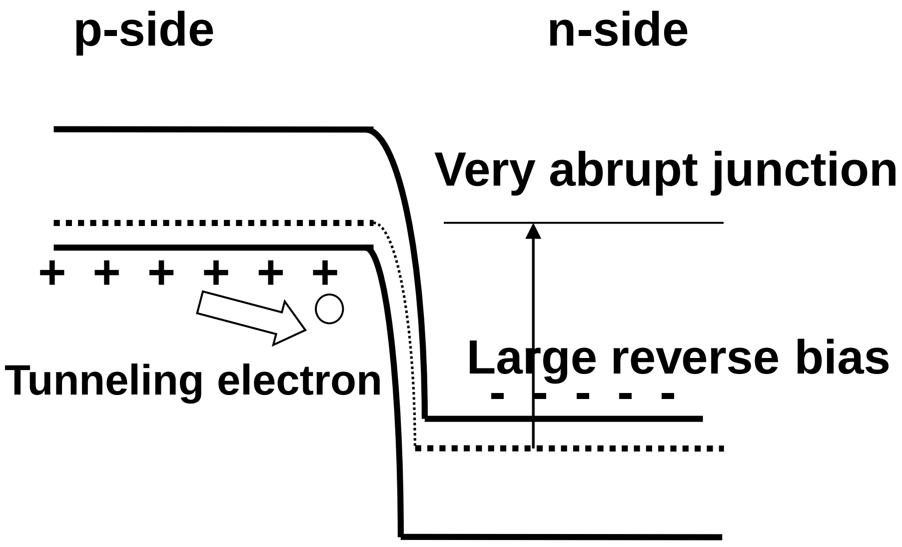


Fig. (2 - 37) electron in the valance band is shown undergoing auto ionization, this phenomenon is known as zener effect. The electron may be considered to tunnel through the potential barrier

emission. This means in a very thin junction where field may become large with only small applied reverse bias voltage, bands in the transition region are steeply tilted and it is possible for electrons to jump the forbidden gap. This situation is shown in figure (2-37). It can be shown that the application of a voltage V to an abrupt junction in Ge produce a field ε across the junction of:

$$\varepsilon = 2\ x\ 10^4 \left(\frac{2V}{\rho}\right)^{1/2} \quad v/cm$$

Where, ρ is the resistivity on the other side of the junction. One assumes resistivity of 5 ohm-cm and an applied voltage of 10 volts; the field across the junction is in the order of 40000 V/cm. Such a field is theoretically high enough to cause considerable emission across the gap. The effect of the reverse field is of course to make the step in the transition region high. In the presence of high field, any electron free in the valance band can be accelerated by the field. Such electrons can escape from the valance band across the gap as shown because among other reasons, little or no energy increase is required to move it from the one such filleted band to the other. Thus in a narrow junction a relatively small external voltage seems to be able to cause direct emission from the valance band to conduction band. This process is a true Zener, or internal field emission effect. It turns out that the critical field necessary for this effect is about 200,000 V/cm in Ge. Such a field can only ever be established across a reverse bias p-n junction. Since Zener breakdown occurs at a given field strength it follows that the actual

breakdown voltage depends on the width of the depletion layer. This in turn is fixed by the distribution of charge impurities across the junction. For an abrupt type where a sudden transition from p to n occurs, we find the Zener voltage given by,

$$V_z = 100\,\rho_n + 50\,\rho_p (in\ Ge),$$
$$= 40\,\rho_n + 8\,\rho_p (in\ Si) (2.90)$$

Where, (ρ_n) and (ρ_p) are the resistivity's on either side of the junction, V_Z in volts if (ρ_n) and (ρ_p) are in ohm-cm. In most case of practical interest either ρ_n or ρ_p is negligible, and then V_Z may be calculated in terms of the other. The constants in the above equation were found for particular group of junctions and probably vary slightly for different qualities of Ge or Si, since they depend on the carrier mobility values. If the transition from p to n-type is graded, so that the net concentration across the junction follows a low of, $N_D - N_A = a\,x \qquad (2.91)$

Where, x is the distance in cm, then it turns out that the zener voltage for Ge is,

$$V_Z = \frac{5\,x\,10^{11}}{\sqrt{a}}$$

Where, V_Z in volts and (a) in cm^{-4}.

An increase in temperature causes the forbidden gap to become somewhat narrower, and thus the probability of zener emission increases and zener breakdown occurs at a lower voltage. In wider junctions, it is found that breakdown occurs at much lower voltage than can be explained by zener breakdown theory, the phenomena of avalanche breakdown is believed to be responsible. In order to understand avalanche breakdown, we must realize that, there are two processes necessary to cause breakdown. In the first place there must be a primary ionization process. In the case of junction this is due to acceleration of electrons from p to n-side of the transition region. The electrons will normally have scattering collisions which do not allow enough energy to build up in one free path, so that ionization of the lattice may occur. If however, the veiled is very large, an electron may be gain enough energy in one free path so that it can excite another electron out of its bounded state and raise it into the conduction band. The amount of kinetic energy needed to cause this is called pair production and it is about 0,72eV. Pair production creates an electron in the conduction band and a hole in the valance band, both free to move.

When this process occurs in the transition region, both electron and hole are accelerated by the field. The electron can go on to ionize further electrons. The hole enters into the secondary process, which is necessary for breakdown to occur. The hole is accelerated in the opposite direction from the electron and it is also ionizing or cause pair production. Hence we see that there is a feedback process which gives us added electron which goes on to cause pair production, etc. The process is, in other hand, self-maintaining so long as the field exists. Two details of the process are of interest. One is that there is a multiplication region near breakdown in which electrons are causing ionizing processes but holes are not yet causing enough ionization of the breakdown voltage. At higher temperature, the mean free paths of electrons and holes are shorter. Therefore, larger is needed to cause ionization and a higher breakdown voltage is observed.

Example (60)

An abrupt pn junction has doping densities of N_A = 5x10^{15} atoms/cm^3 and N_D = 10^{16} atoms/cm^3. Calculate the breakdown voltage if E_{crit} = 3x10^5 V/cm.

Solution

$$V_R = \frac{\varepsilon_{si}\,(N_A + N_D)}{2\,q\,N_A N_D}\,E_{max}^2$$

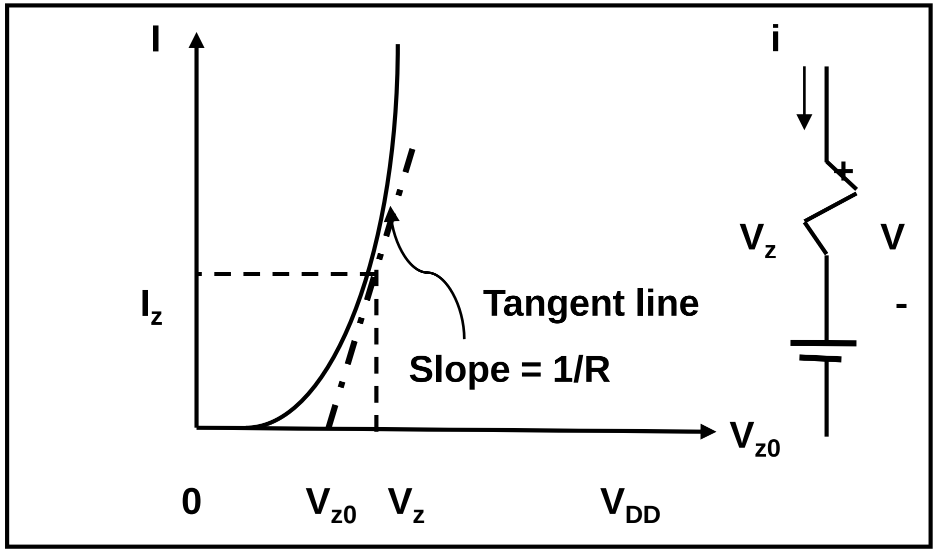
$$= \frac{1.04\,x\,10^{-12}\,(5\,x\,10^{15} + 10^{16})}{2\,x\,1.6\,x\,10^{-19}\,x\,5\,x\,10^{15}\,x\,10^{16}} = 88\,V$$

Example (62)

A Zener diode at room temperature (V_T = 0.0259 V) has the specifications V_Z = 10 V, I_Z = 10 mA , and r_z = 20 Ω. Calculate

(a) the Zener breakdown ideality factor η, (b) the voltage V_{Z0} in the linear circuit shown, and

(c) the voltage at which the breakdown current is I_Z / 10.



Solution

(a) Equation $r_z = {}^{\eta_z V_T}/_{I_z}$ be used to solve for ηz to obtain $\eta_z = \frac{I_z r_z}{V_T} = \frac{10\,m\,x\,20}{25.9\,m} 7.72$

(b) The voltage V_{Z0} is calculated from Eq. $V_{z0} = V_z - I_z r_z$ to obtain

$$V_{z0} = V_z - I_z r_z = 10 - 10\,m\,x\,20 9.8\,V$$

(c) If I_S << I_Z, Eq.

$$i = I_s\left[exp\left(\frac{V}{\eta\,V_T}\right) - 1\right] - I_z\,exp\left(-\frac{v + V_z}{\eta_z V_T}\right)$$

Can be solved for the voltage at which i = −1 mA as follows: $v = -V_z - \eta_z\,V_T\,ln\left(\frac{i}{-I_z}\right) -$

$$10 - 7.72\,x\,25.9\,m\,x\,ln\,(0.1) = -9.54\,V$$