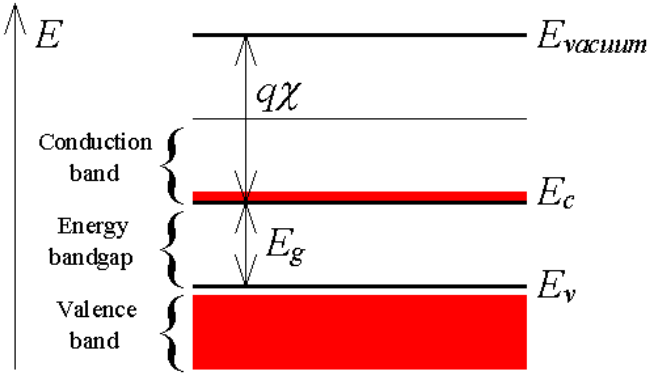
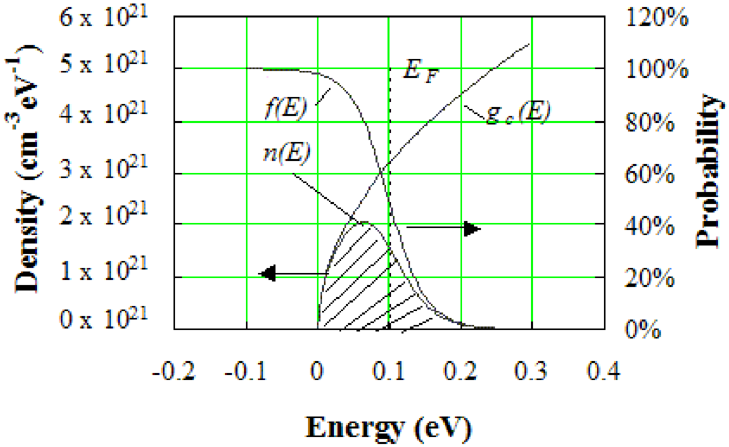
Review

Energy Bands



Carrier Density & Mobility



© Bart Van Zeghbroeck 2007

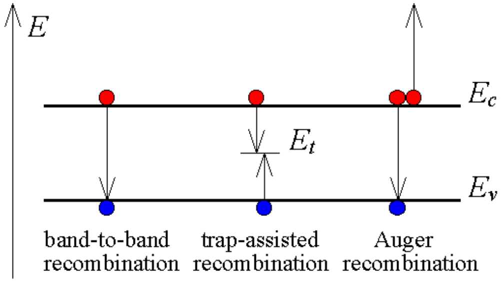
Carrier Transport

$$J_n = qn\mu_n \mathcal{E} + qD_n \frac{dn}{dx}$$

$$J_p = -qD_p \frac{dp}{dx}$$

$$\mu = \frac{q\tau_c}{m^*}$$

$$\sigma = \frac{\Delta J}{\mathcal{E}} = q(n\mu_n + p\mu_p)$$



$$G_{p,light} = G_{n,light} = \alpha \frac{P_{opt}(x)}{E_{ph}A}$$

Current Transport: Diffusion, Thermionic Emission & Tunneling

For Diffusion current, the depletion layer is large compared to the mean free path, so that the concepts of drift and diffusion are valid. The current depends exponentially on the applied voltage, V_a, and the barrier height, ϕ_B.

$$J_n = \frac{q^2 D_n N_c}{V_t} \sqrt{\frac{2q(\phi_i - V_a) N_d}{\varepsilon_s}} \exp(-\frac{\phi_B}{V_t})[\exp(\frac{V_a}{V_t}) - 1]$$

Electric-field at MS Junction:

$$\mathcal{E}_{\max} = \sqrt{\frac{2q(\phi_i - V_a) N_d}{\varepsilon_s}}$$

The thermionic emission theory assumes that electrons, with an energy larger than the top of the barrier, will cross the barrier provided they move towards the barrier. The actual shape of the barrier is ignored.

$$J_{MS} = A^* T^2 \, e^{-\phi_B / V_t} (e^{V_a / V_t} - 1)$$

ϕ_B is the Schottky barrier height.

Richardson's Constant

$$A^* = \frac{4 \pi q m^* k^2}{h^3}$$

and velocity

$$v_R = \sqrt{\frac{kT}{2 \pi m}}$$

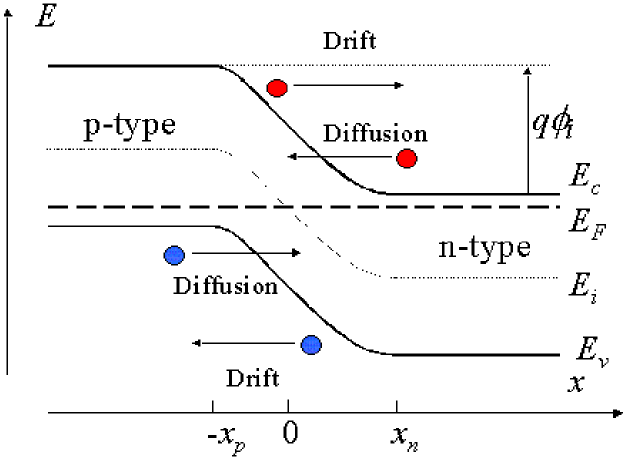
For tunneling, the carrier density equals the density of available electrons, n, multiplied with the tunneling probability, Θ, yielding:

$$J_n = q \, v_R \, n \, \Theta$$

$$\Theta = \exp\left(-\frac{4}{3} \frac{\sqrt{2qm^*}}{h} \frac{\phi_B^{3/2}}{\mathcal{E}} \right)$$

where $\varepsilon = \phi_B / L$

The p-n Junction vs MS Junction



$$\phi_i = V_t \ln \frac{N_d N_a}{n_i^2} \qquad \phi = \phi_i - V_a$$

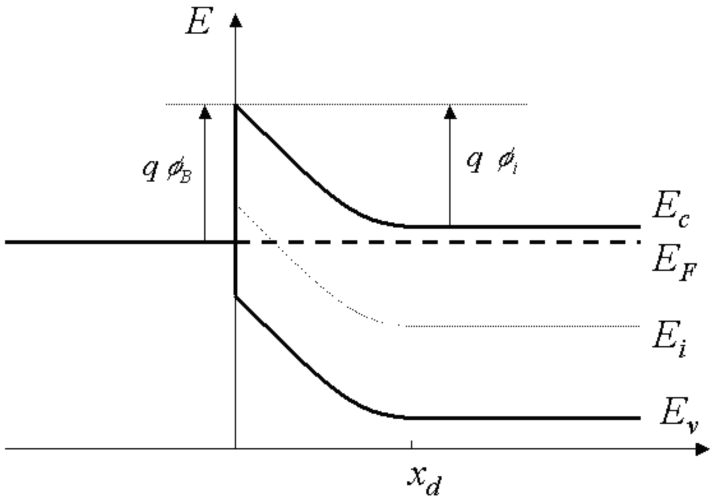
At Full Depletion:

$$Q_n = qN_d x_n \qquad \mathcal{E}(x=0) = -\frac{qN_a x_p}{\varepsilon_s} = -\frac{qN_d x_n}{\varepsilon_s}$$

$$Q_p = -qN_a x_p \qquad N_d x_n = N_a x_p$$

$$x_d = x_n + x_p$$

$$x_d = \sqrt{\frac{2\varepsilon_s}{q}\left(\frac{1}{N_a} + \frac{1}{N_d}\right)(\phi_i - V_a)}$$



$\phi_B = \Phi_M - \chi$, for an n-type semiconductor

$\phi_B = \frac{E_g}{q} + \chi - \Phi_M$, for a p-type semiconductor

$\phi_i = \Phi_M - \chi - \frac{E_c - E_{F,n}}{q}$, n-type

$\phi_i = \chi + \frac{E_c - E_{F,p}}{q} - \Phi_M$, p-type

$$\phi_i = \phi_B - V_t \ln \frac{N_c}{N_d}$$

$$\phi(x=\infty) - \phi(x=0) = \phi_i - V_a$$

The p-n Junction: Capacitance

Any variation of the charge within a p-n diode with an applied voltage variation yields a capacitance, which must be added to the circuit model of a p-n diode. This capacitance related to the depletion layer charge in a p-n diode is called the junction capacitance.

$$C(V_a) \overset{\Delta}{=} \left| \frac{dQ(V_a)}{dV_a} \right|$$

$$Q_n = qN_d x_n$$

$$Q_p = -qN_a x_p$$

$$x_n = \sqrt{\frac{2\varepsilon_s}{q} \frac{N_a}{N_d} \frac{1}{N_a + N_d} (\phi_i - V_a)}$$

$$C_j = \sqrt{\frac{q\varepsilon_s}{2(\phi_i - V_a)} \frac{N_a N_d}{N_a + N_d}}$$

Looks like a parallel plate capacitor, namely: $C_j = \dfrac{\varepsilon_s}{x_d}$

except that the depletion layer width and hence the capacitance is voltage dependent.

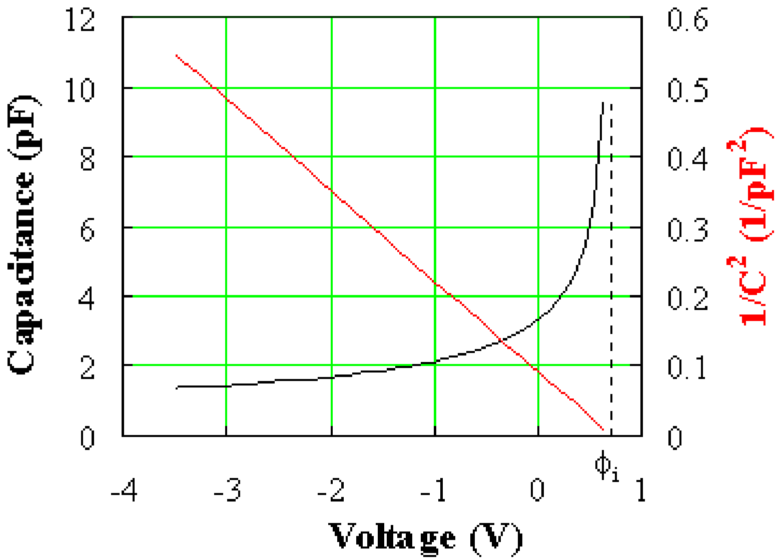
The p-n Junction: Capacitance

$$C_j = \sqrt{\frac{q\,\varepsilon_s}{2(\phi_i - V_a)}\frac{N_a N_d}{N_a + N_d}}$$

Express as: $$\frac{1}{C_j^2} = \frac{2}{q\,\varepsilon_s}\frac{N_a + N_d}{N_a N_d}(\phi_i - V_a)$$ where $$\frac{d(1/C_j^2)}{dV_a} = -\frac{2}{q\,\varepsilon_s}\frac{N_a + N_d}{N_a N_d}$$

A capacitance versus voltage measurement can be used to obtain the built-in voltage and the doping density of a one-sided p-n diode. The built-in voltage is obtained at the intersection of the $1/C^2$ curve and the horizontal axis, while the doping density is obtained from the slope of the curve.

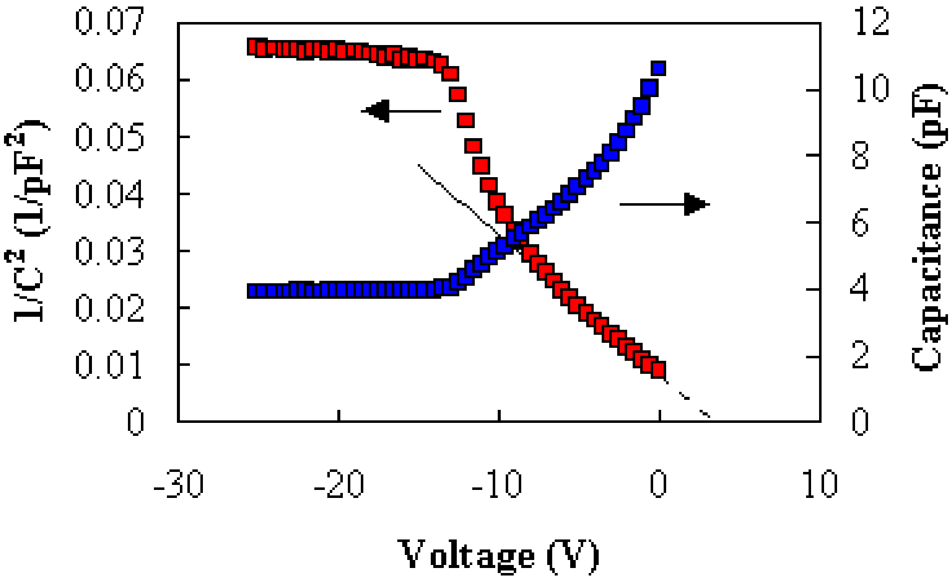
$$\frac{d(1/C_j^2)}{dV_a} = -\frac{2}{q\,\varepsilon_s}\frac{N_a + N_d}{N_a N_d}$$



The p-n Junction: Capacitance

A capacitance-voltage measurement also provides the doping density profile of one-sided p-n diodes.

$$N_d = -\frac{2}{q\,\varepsilon_s}\frac{1}{\frac{d(1/C_j^2)}{dV_a}}, \text{if } N_a >> N_d$$



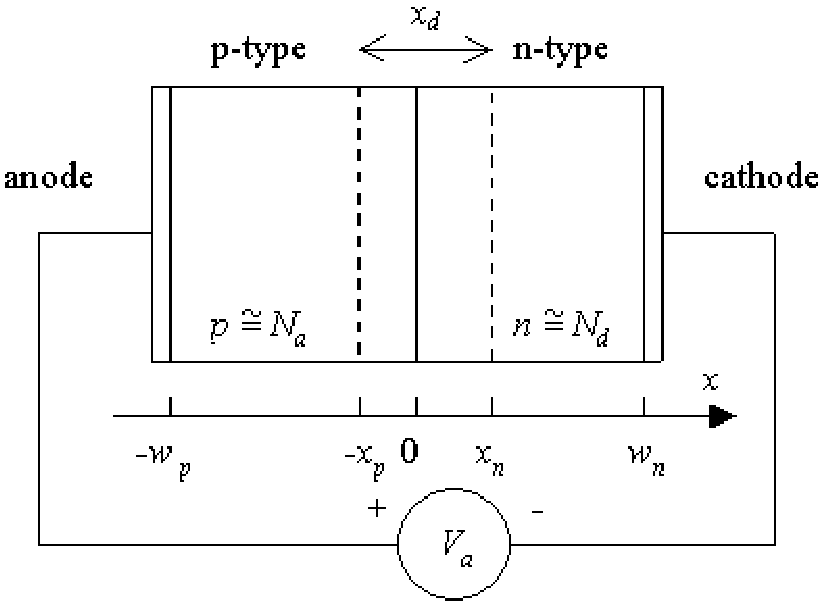
This diode consists of a highly doped p-type region on a lightly doped n-type region on top of a highly doped n-type substrate. The dotted line forms a reasonable fit at voltages close to zero from which one can conclude that the doping density is almost constant close to the p-n interface. The capacitance becomes almost constant at large negative voltages, which corresponds according to equation above at high doping density.

The p-n Junction Current

The current in a p-n diode is due to carrier recombination or generation somewhere within the p-n diode structure.

Under forward bias, the diode current is due to recombination. This recombination can occur within the quasi-neutral regions, within the depletion region or at the metal-semiconductor Ohmic contacts.

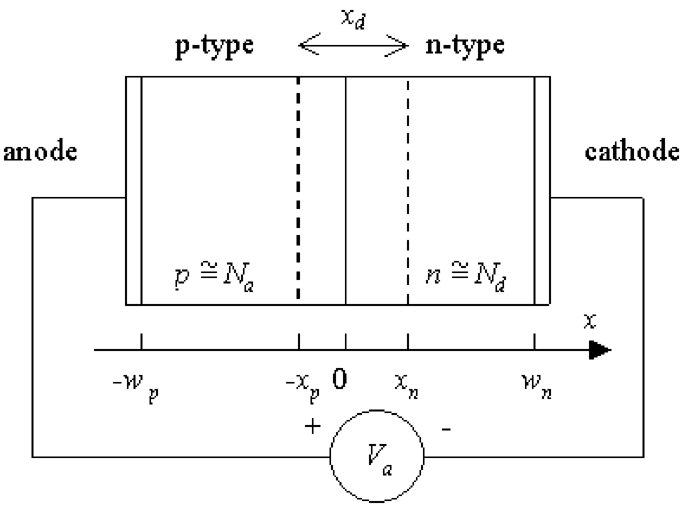
Under reverse bias, the current is due to generation. Carrier generation due to light will further increase the current under forward as well as reverse bias.



The **"long" diode** expression applies to p-n diodes in which recombination/generation occurs in the quasi-neutral region only. This is the case if the quasi-neutral region is much larger than the carrier diffusion length.

The **"short" diode** expression applies to p-n diodes in which recombination/generation occurs at the contacts only. In a short diode, the quasi-neutral region is much smaller than the diffusion length.

The p-n Junction Current



The electric field and potential are obtained by using the full depletion approximation. Assuming that the quasi-Fermi energies are constant throughout the depletion region, one obtains the minority carrier densities at the edges of the depletion region, yielding:

$$p_n(x = x_n) = p_{n0}\, e^{V_a / V_t}$$

$$n_p(x = -x_p) = n_{p0}\, e^{V_a / V_t}$$

The carrier density at the metal contacts is assumed to equal the thermal-equilibrium carrier density. This assumption implies that excess carriers immediately recombine when reaching either of the two metal-semiconductor contacts. This results in the following set of boundary conditions:

$$p_n(x = w_n) = p_{n0}$$

$$n_p(x = -w_p) = n_{p0}$$

The p-n Junction Current

The general expression for the ideal diode current is obtained by applying the boundary conditions to the general solution of the diffusion equation for each of the quasi-neutral regions. L_n (and L_p) are the diffusion lengths.

$$L_n = \sqrt{D_n \tau_n}$$

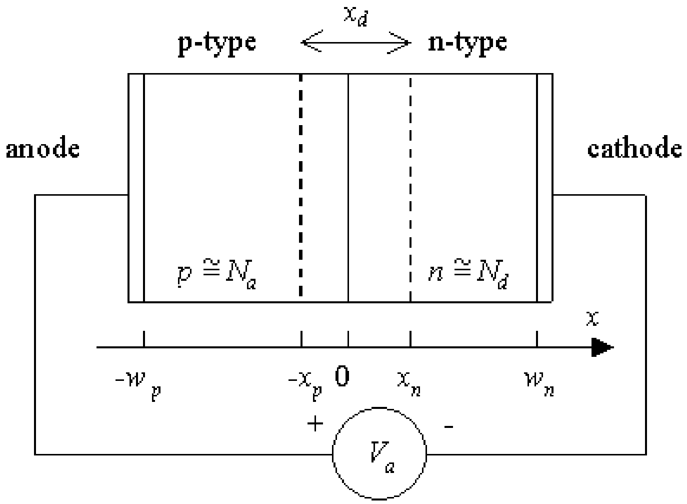
$$\frac{\partial n(x,t)}{\partial t} = D_n \frac{\partial^2 n_p(x,t)}{\partial x^2} - \frac{n_p(x,t) - n_{p0}}{\tau_n}$$

$$n_p(x \le -x_p) = n_{p0} + Ce^{-(x+x_p)/L_n} + De^{(x+x_p)/L_n}$$

$$p_n(x \ge x_n) = p_{n0} + Ae^{-(x-x_n)/L_p} + Be^{(x-x_n)/L_p}$$

$$n_p(x \le -x_p) = n_{p0} + Ce^{-(x+x_p)/L_p} + De^{(x+x_p)/L_p}$$

where A, B, C and D are constants whose value remains to be determined.



$$p_n(x \ge x_n) = p_{n0} + p_{n0}(e^{V_a/V_t} - 1)\left[\cosh\frac{x - x_n}{L_p} - \coth\frac{w_n'}{L_p}\sinh\frac{x - x_n}{L_p}\right]$$

$$w_n' = w_n - x_n$$

The p-n Junction Current

$$p_n(x \geq x_n) = p_{n0} + p_{n0}(e^{V_a/V_t} - 1)[\cosh\frac{x - x_n}{L_p} - \coth\frac{w_n'}{L_p}\sinh\frac{x - x_n}{L_p}] \qquad w_n' = w_n - x_n$$

$$n_p(x \leq -x_p) = n_{p0} + n_{p0}(e^{V_a/V_t} - 1)[\cosh\frac{x + x_p}{L_n} + \coth\frac{w_p'}{L_n}\sinh\frac{x + x_p}{L_n}] \qquad w_p' = w_p - x_p$$

$$\begin{aligned} J_p(x \geq x_n) &= -qD_p\frac{dp}{dx} \\ &= -\frac{qD_p p_{n0}}{L_p}(e^{V_a/V_t} - 1)[\sinh\frac{x - x_n}{L_p} - \coth\frac{w_n'}{L_p}\cosh\frac{x - x_n}{L_p}] \end{aligned}$$

$$\begin{aligned} J_n(x \leq -x_p) &= qD_n\frac{dn}{dx} \\ &= \frac{qD_n n_{p0}}{L_n}(e^{V_a/V_t} - 1)[\sinh\frac{x + x_p}{L_n} + \coth\frac{w_p'}{L_n}\cosh\frac{x + x_p}{L_n}] \end{aligned}$$

The total current then equals the sum of the maximum electron current in the p-type region, the maximum hole current in the n-type regions and the current due to recombination within the depletion region. We ignore the recombination in the depletion region.

$$I = A[J_n(x = -x_p) + J_p(x = x_n) + J_r] \cong I_s(e^{V_a/V_t} - 1) \qquad I_s = qA[\frac{D_n n_{p0}}{L_n}\coth(\frac{w_p'}{L_n}) + \frac{D_p p_{n0}}{L_p}\coth(\frac{w_n'}{L_p})]$$

The p-n Junction Current: "long" quasi-neutral region

A diode with a "long" quasi-neutral region has a quasi-neutral region which is much larger than the minority-carrier diffusion length in that region, or $w_{n'} > L_p$ and $w_{p'} > L_n$.

$$\coth x = \frac{1}{\tanh x} \cong \frac{1}{x} \text{ ,for } x << 1$$

$$J_p(x \geq x_n) = \frac{qD_p p_{n0}}{L_p}(e^{V_a/V_t} - 1)\exp\frac{-(x - x_n)}{L_p} \qquad J_n(x \leq -x_p) = \frac{qD_n n_{p0}}{L_n}(e^{V_a/V_t} - 1)\exp\frac{x + x_p}{L_n}$$

$$I_s = qA[\frac{D_n n_{p0}}{L_n} + \frac{D_p p_{n0}}{L_p}] = qA[\frac{n_{p0} L_n}{\tau_n} + \frac{p_{n0} L_p}{\tau_p}]$$

Let’s now assume that the recombination rate is constant in the depletion region. To further simplify the analysis we will consider only the recombination in the n-type region. The current due to recombination in the depletion region is then given by:

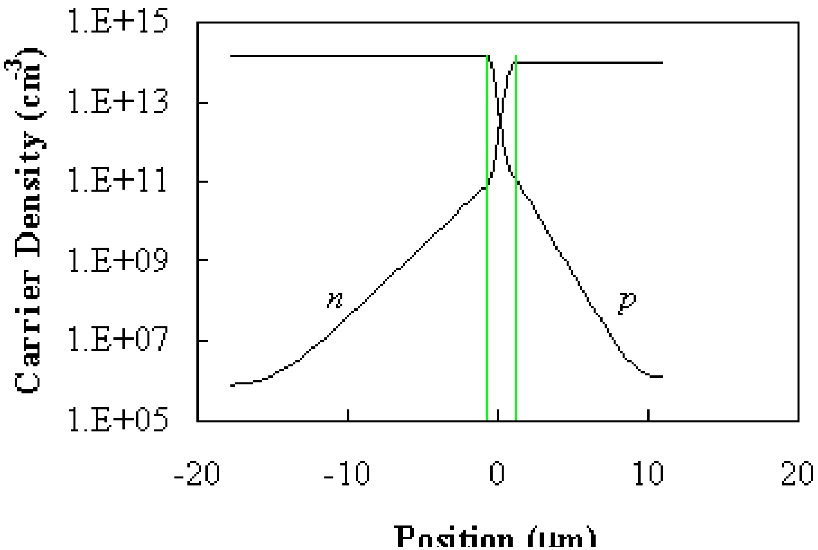
$$I_r \geq qA\frac{p_{n0} x_n}{\tau_p}(e^{V_a/V_t} - 1)$$

so that I_r can be ignored if: $I_r << I$, for $x_n << L_p$

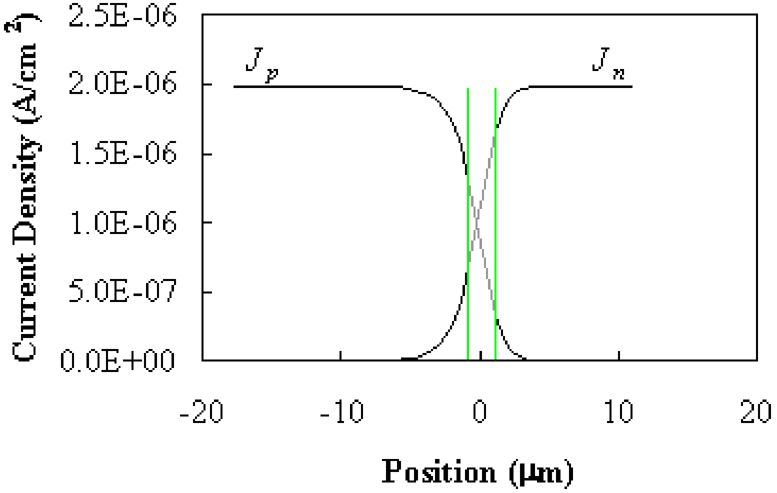
A necessary, but not sufficient requirement is therefore that the depletion region width is much smaller than the diffusion length for the ideal diode assumption to be valid.

The p-n Junction Current: An example

Lets consider a silicon p-n diode with N_a = 1.5 x 10^{14} cm^{-3} and N_d = 10^{14} cm^{-3} . The minority carrier lifetime was chosen to be very short, namely 400 ps, so that most features of interest can easily be observed. We start by examining the electron and hole density throughout the p-n diode.



The majority carrier densities in the quasi-neutral region simply equal the doping density. The electron and hole current density as calculated using above equations The current due to recombination in the depletion region was assumed to be constant.



The p-n Junction Current: "short" quasi-neutral region

A "short" diode is a diode with quasi-neutral regions, which are much shorter than the minority-carrier diffusion lengths. As the quasi-neutral region is much smaller than the diffusion length one finds that recombination in the quasi-neutral region is negligible so that the diffusion equations are reduced to:

$$0 = D_n \frac{d^2 n_p}{dx^2}, \text{ and } 0 = D_p \frac{d^2 p_n}{dx^2}$$

The resulting carrier density varies linearly throughout the quasi-neutral region:

$$n_p = A + Bx \text{ , and } p_n = A + Bx$$

Applying the same boundary conditions at the edge of the depletion region as above and requiring thermal equilibrium at the contacts yields:

$$p_n = p_{n0} + p_{n0}(\mathrm{e}^{V_a / V_t} - 1)\left(1 - \frac{x - x_n}{w_n'}\right) \qquad n_p = n_{p0} + n_{p0}(\mathrm{e}^{V_a / V_t} - 1)\left(1 + \frac{x + x_p}{w_p'}\right)$$

$$I = A[J_n(x = -x_p) + J_p(x = x_n) + J_r] \cong I_s(e^{V_a / V_t} - 1) \qquad I_s = qA[\frac{D_n n_{p0}}{w_p'} + \frac{D_p p_{n0}}{w_n'}]$$

The "short" and "long" diode expressions are the same except for the **use of the diffusion length or the quasi-neutral region width in the denominator, whichever is smaller.**

The p-n Junction Current: An example

An abrupt silicon p-n junction ($N_a = 10^{16}$ cm^{-3} and $N_d = 4 \times 10^{16}$ cm^{-3}) is biased with $V_a = 0.6$ V. Calculate the ideal diode current assuming that the n-type region is much smaller than the diffusion length with $w_{n'} = 1$ μm and assuming a "long" p-type region. Use $\mu_n = 1000$ cm^2/V-s and $\mu_p = 300$ cm2/V-s. The minority carrier lifetime is 10 ms and the diode area is 100 μm by 100 μm.

The current is calculated from: $$I = qA[\frac{D_n n_{p0}}{L_n} + \frac{D_p p_{n0}}{w_n'}](e^{V_a/V_t} - 1)$$

with

$D_n = \mu_n V_t = 1000 \times 0.0258 = 25.8$ cm^2/V-s

$D_p = \mu_p V_t = 300 \times 0.0258 = 7.75$ cm2/V-s

$$L_n = \sqrt{D_n \tau_n} = \sqrt{25.8 \times 10^{-5}} = 161\ \mu m$$

$n_{p0} = n_i^2/N_a = 10^{20}/10^{16} = 10^4$ cm^{-3}

$p_{n0} = n_i^2/N_d = 10^{20}/ 4 \times 10^{16} = 2.5 \times 10^3$ cm^{-3}

yielding I = 40.7 mA

Note that the hole diffusion current occurs in the "short" n-type region and therefore depends on the quasi-neutral width in that region. The electron diffusion current occurs in the "long" p-type region and therefore depends on the electron diffusion length in that region.

The p-n Junction Current: Band-to-band and SRH

The recombination/generation current due to band-to-band recombination/generation is obtained by integrating the net recombination rate, $U_{b\text{-}b}$, over the depletion region:

$$J_{b-b} = q \int_{-x_p}^{x_n} U_{b-b}\, dx \qquad J_{b-b} = q \int_{-x_p}^{x_n} n_i^2 (e^{V_a / V_t} - 1)\, dx = q n_i^2 b w (e^{V_a / V_t} - 1)$$

The current due to band-to-band recombination has therefore the same voltage dependence as the ideal diode current and simply adds an additional term to the expression for the saturation current.

The current due to trap-assisted recombination in the depletion region is also obtained by integrating the trap-assisted recombination rate over the depletion region width:

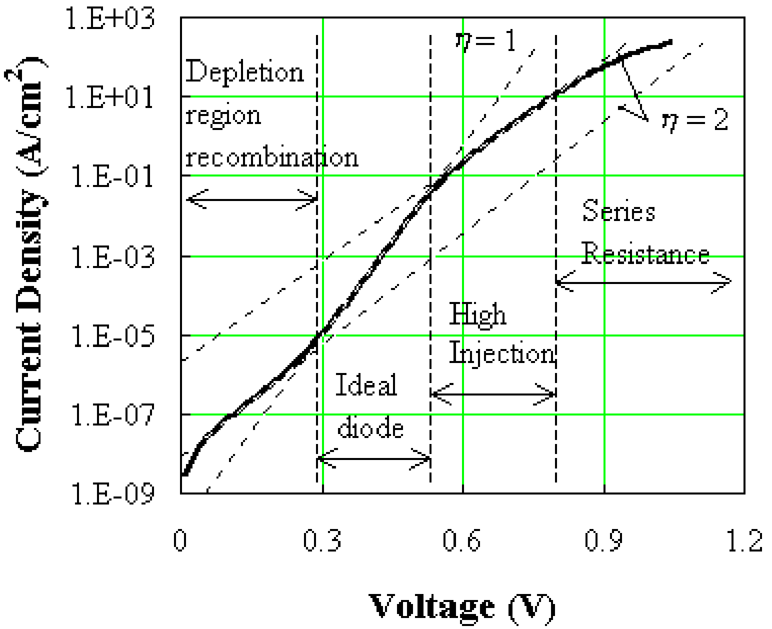
$$J_{SHR} = q \int_{-x_p}^{x_n} U_{SHR}\, dx \qquad J_{SHR} = \frac{q n_i x'}{2\tau} (e^{V_a / 2V_t} - 1)$$

This does not provide an actual solution since the effective width, x', still must be determined by performing a numeric integration. Nevertheless, the above expression provides a way to obtain an upper estimate by substituting the depletion layer width, x_d, as it is always larger than the effective width.

The p-n Junction: Real I-V Characteristics

1) Ideal diode region where the current increases by one order of magnitude as the voltage is increased by 60 mV. This region is referred to as having an ideality factor, n, of one. This ideality factor is obtained by fitting a section of the curve to the following expression for the current:

$$J = J_s e^{V_a/\eta V_t} \qquad \eta = \frac{\log(e)}{V_t \text{ slope}} = \frac{\frac{1}{\text{slope}}}{59.6 \text{ mV/decade}}$$



2) The ideality factor is 2, and the current is dominated by the trap-assisted recombination

3) The current becomes limited by high injection effects and by the series resistance. High injection occurs in a forward biased p-n diode when the injected minority carrier density exceeds the doping density. High injection will therefore occur first in the lowest doped region of the diode since that region has the highest minority carrier density.

$$V_a = 2V_t \ln \frac{N_d}{n_i}$$

4) For higher Va, the current increases linearly due to the series resistance of the diode. This series resistance can be due to the contact resistance between the metal and the semiconductor. This series resistance increases the external voltage, V_{a^*}, relative to the internal voltage, V_a, considered so far.

$$V_a^* = V_a + IR_s$$

The p-n Junction: Breakdown

The maximum reverse bias voltage that can be applied to a p-n diode is limited by breakdown. Breakdown is characterized by the rapid increase of the current under reverse bias. The corresponding applied voltage is referred to as the breakdown voltage. The breakdown voltage is a key parameter of power devices.

Two mechanisms can cause breakdown:

- avalanche multiplication
- quantum mechanical tunneling of carriers through the bandgap

Neither of the two breakdown mechanisms is destructive. However heating caused by the large breakdown current and high breakdown voltage causes the diode to be destroyed unless sufficient heat sinking is provided.

Imperical formula for breakdown in Si:

$$\left|\mathcal{E}_{br}\right| = \frac{4x10^{5}}{1-\frac{1}{3}\log(N/10^{16})} \quad \text{V/cm}$$

Assuming a one-sided abrupt p-n diode, the corresponding breakdown voltage is:

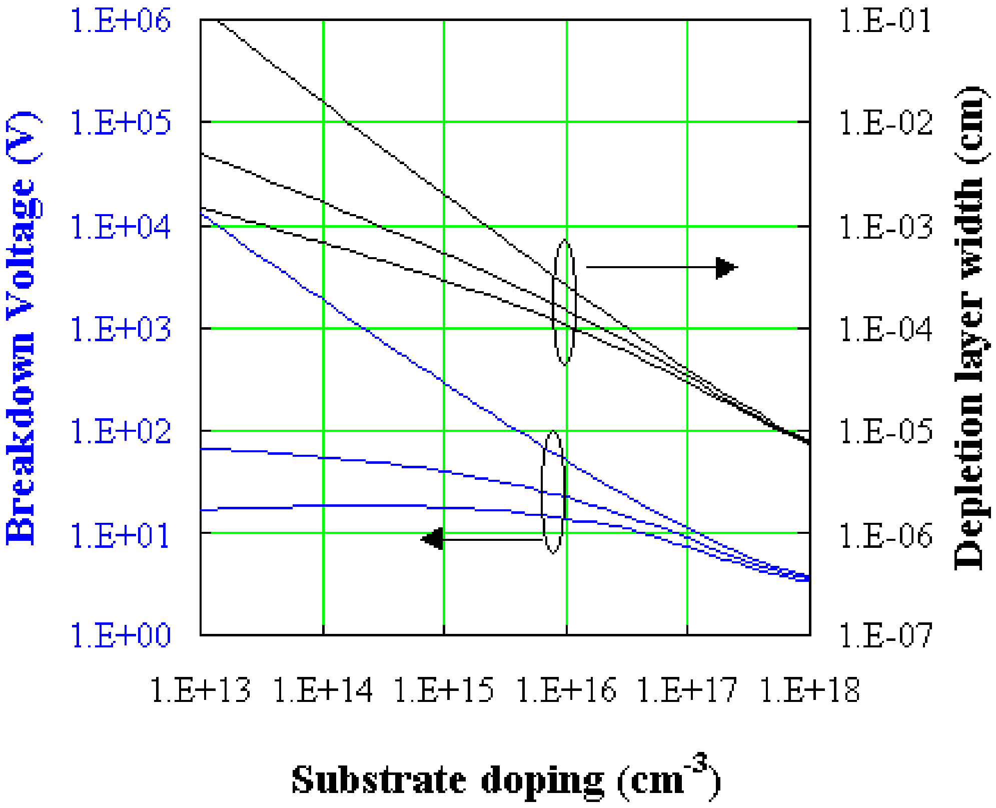
$$\left|V_{br}\right| = -\phi_i + \frac{\left|\mathcal{E}_{br}\right|^{2}\varepsilon_s}{2qN}$$

The corresponding depletion layer width equals:

$$w_{br} = \frac{\left|\mathcal{E}_{br}\right|\varepsilon_s}{qN}$$

The p-n Junction: Breakdown Edge effects

Few p-n diodes are truly planar and typically have higher electric fields at the edges. Since the diodes will break down in the regions where the breakdown field is reached first, one has to take into account the radius of curvature of the metallurgical junction at the edges. Most doping processes including diffusion and ion implantation yield a radius of curvature on the order of the junction depth. The breakdown voltages and depletion layer widths are plotted below as a function of the doping density of an abrupt one-sided junction.



Breakdown voltage and depletion layer width at breakdown versus doping density of an abrupt one-sided p-n diode. Shown are the voltage and width for a planar (top curves), cylindrical (middle curves) and spherical (bottom curves) junction with 1 mm radius of curvature.

The p-n Junction: Avalanche Breakdown

Avalanche breakdown is caused by impact ionization of electron-hole pairs. When applying a high electric field, carriers gain kinetic energy and generate additional electron-hole pairs through impact ionization. The ionization rate is quantified by the ionization constants of electrons and holes, α_n and α_p.

$$dn = \alpha_n n\, dx$$

The ionization causes a generation of additional electrons and holes.
Assuming that the ionization coefficients of electrons and holes are the same,
the multiplication factor M, can be calculated from:

$$M = \frac{1}{1 - \int_{x_1}^{x_2} \alpha\, dx}$$

The integral is taken between x_1 and x_2, the region within the depletion layer where the electric field is assumed constant and large enough to cause impact ionization. Outside this range, the electric field is assumed to be too low to cause impact ionization.

$$M = \frac{1}{1 - \left|\frac{V_a}{V_{br}}\right|^n}, \text{ where } 2 < n < 6$$

The p-n Junction: Zener Breakdown

Quantum mechanical tunneling of carriers through the bandgap is the dominant breakdown mechanism for highly doped p-n junctions. The analysis is identical to that of tunneling in a metal-semiconductor junction where the barrier height is replaced by the energy bandgap of the material.

The tunneling probability equals:

$$\Theta = \exp\left(-\frac{4}{3}\frac{\sqrt{2m^{*}}}{q\hbar}\frac{E_g^{3/2}}{\mathcal{E}}\right)$$

where the electric field equals = $\varepsilon_g/(qL)$

The tunneling current is obtained from the product of the carrier charge, velocity and carrier density. The velocity equals the Richardson velocity, the velocity with which on average the carriers approach the barrier while the carrier density equals the density of available electrons multiplied with the tunneling probability, yielding:

$$J_n = q\, v_R\, n\, \Theta \qquad v_R = \sqrt{\frac{kT}{2\pi m}}$$

The tunneling current therefore depends exponentially on the bandgap energy to the 3/2 power.

Bipolar Junction Transistors

Invented it in 1948 by Bardeen, Brattain and Shockley, at Bell Laboratories, as part of a post-war effort to replace vacuum tubes with solid-state devices.

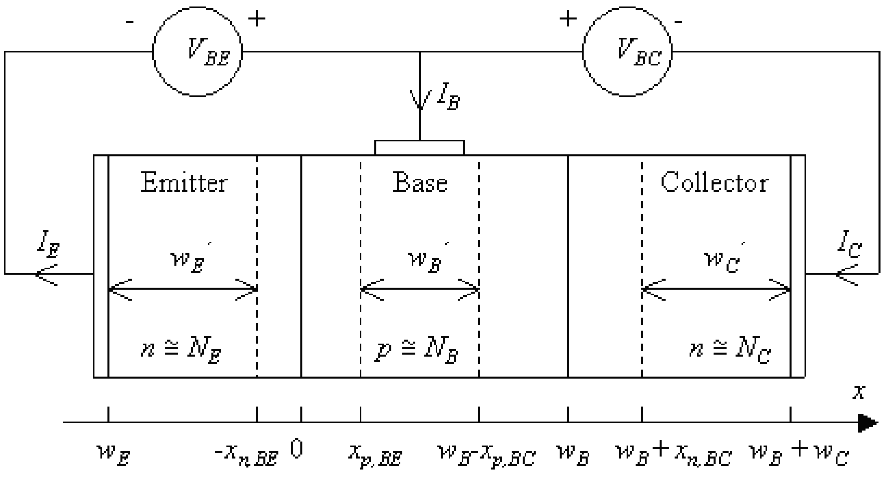
1956 Nobel Prize in Physics

While almost all logic circuits, microprocessor and memory chips contain exclusively MOSFETs, bipolar transistors remain important for some devices:

- ultra-high-speed discrete logic circuits such as emitter coupled logic (ECL)
- power-switching applications
- microwave power amplifiers
- Heterojunction bipolar transistors (HBTs) for cell phone amplifiers

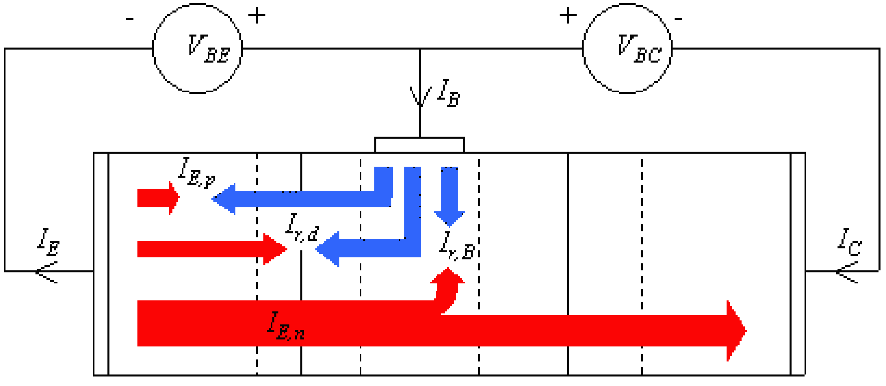
Bipolar Junction Transistors

A bipolar junction transistor consists of two back-to-back p-n junctions, who share a thin common region with width, w_B. Contacts are made to all three regions, the two outer regions called the emitter and collector and the middle region called the base. The device is called “bipolar” since its operation involves both types of mobile carriers, electrons and holes.



(a)

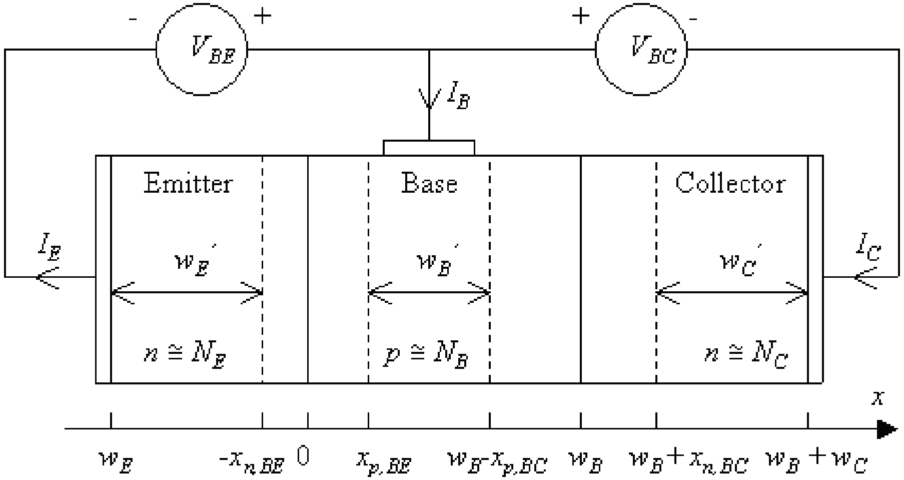
(a) Structure and sign convention of a npn bipolar junction transistor.



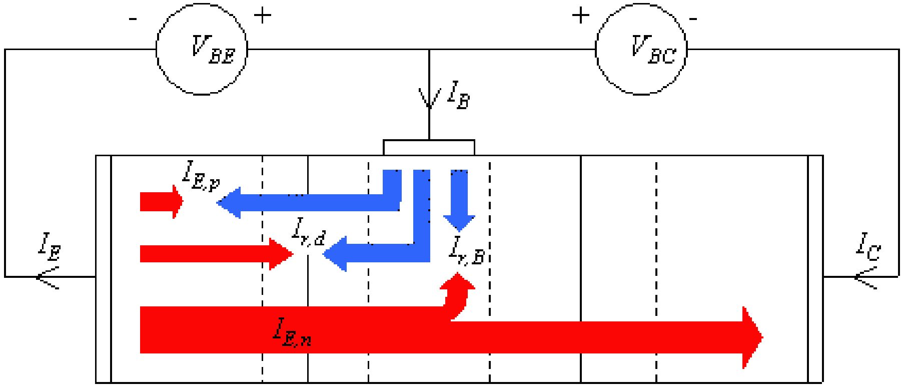
(b)

(b) Electron and hole flow under forward active bias, VBE > 0 and VBC = 0.

Bipolar Junction Transistors



(a)



(b)

Since the device consists of two back-to-back diodes, there are depletion regions between the quasi-neutral regions.

$$w_E' = w_E - x_{n,BE}$$

$$w_B' = w_B - x_{p,BE} - x_{p,BC}$$

$$w_C' = w_C - x_{n,BC}$$

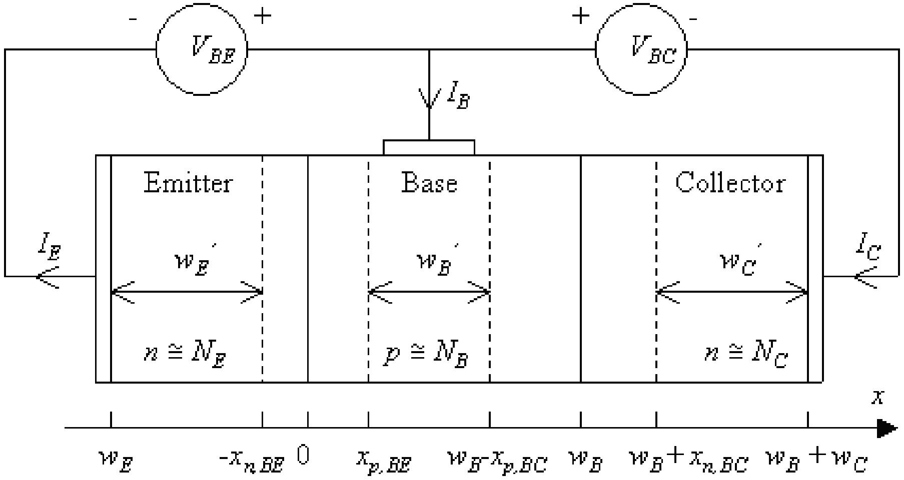
$$x_{n,BE} = \sqrt{\frac{2\varepsilon_s(\phi_{i,BE} - V_{BE})}{q}\frac{N_B}{N_E}\left(\frac{1}{N_B + N_E}\right)}$$

$$x_{p,BE} = \sqrt{\frac{2\varepsilon_s(\phi_{i,BE} - V_{BE})}{q}\frac{N_E}{N_B}\left(\frac{1}{N_B + N_E}\right)}$$

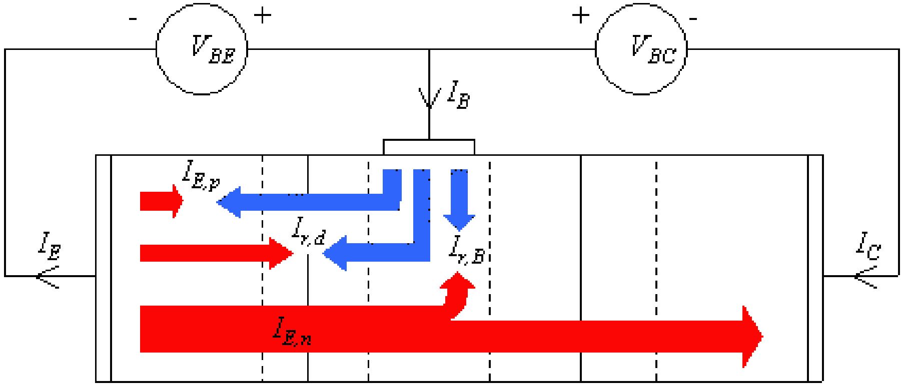
$$x_{p,BC} = \sqrt{\frac{2\varepsilon_s(\phi_{i,BC} - V_{BC})}{q}\frac{N_C}{N_B}\left(\frac{1}{N_B + N_C}\right)}$$

$$x_{n,BC} = \sqrt{\frac{2\varepsilon_s(\phi_{i,BC} - V_{BC})}{q}\frac{N_B}{N_C}\left(\frac{1}{N_B + N_C}\right)}$$

Bipolar Junction Transistors



(a)



(b)

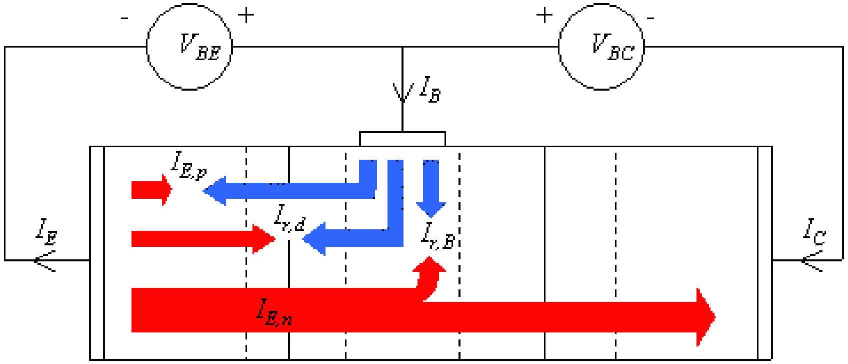
$$\phi_{i,BE} = V_t \ln \frac{N_E N_B}{n_i^2}$$

$$\phi_{i,BC} = V_t \ln \frac{N_C N_B}{n_i^2}$$

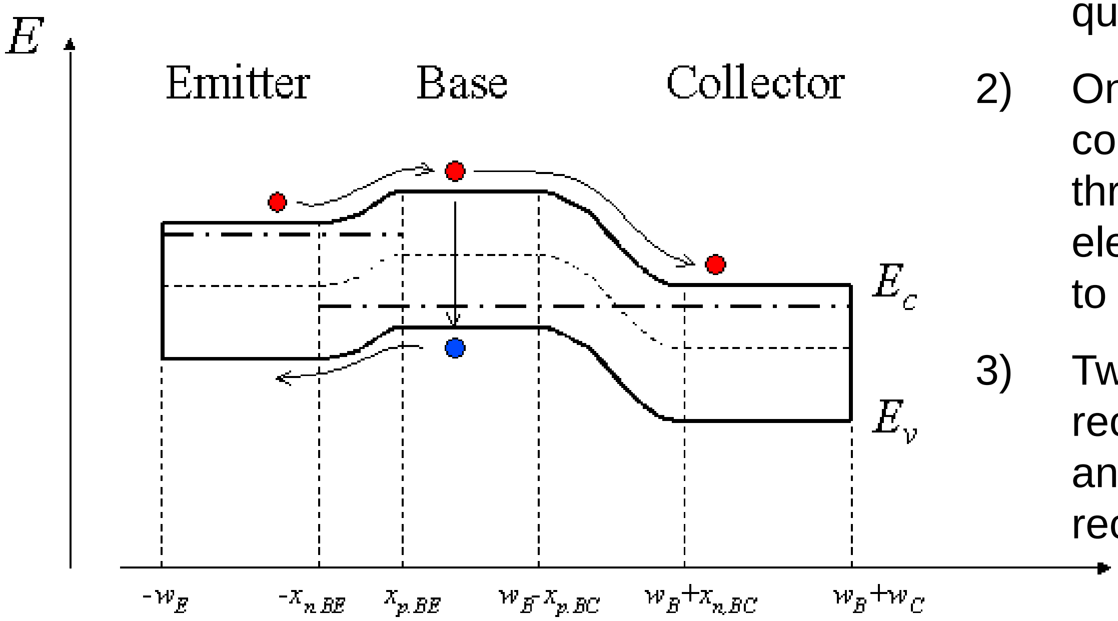
The base and collector current are positive if a positive current goes into the base or collector contact. The emitter current is positive for a current coming out of the emitter contact. This also implies that the emitter current, I_E, equals the sum of the base current, I_B, and the collector current, I_C:

$$I_E = I_C + I_B$$

Bipolar Junction Transistors



(b)



The forward active bias mode of operation is obtained by forward biasing the base-emitter junction and reverse biasing the base-collector junction

1) Electrons diffuse from the emitter into the base and holes diffuse from the base into the emitter. Note that electrons can diffuse as minority carriers through the quasi-neutral base.

2) Once the electrons arrive at the base-collector depletion region, they are swept through the depletion layer due to the electric field. These electrons contribute to the collector current.

3) Two more currents include the base recombination current (vertical arrow), and the base-emitter depletion layer recombination current, $I_{r,d}$.

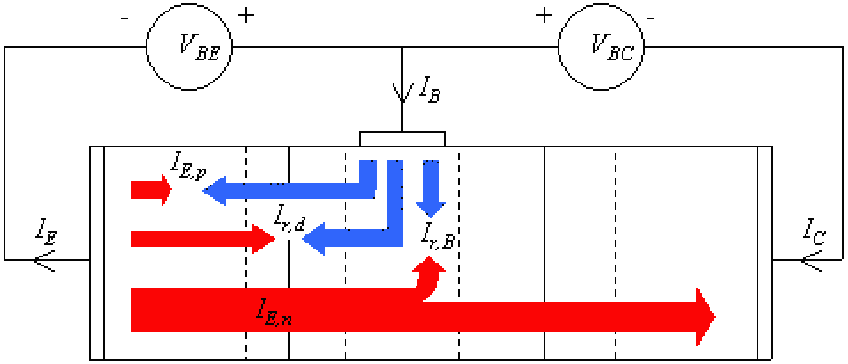
Bipolar Junction Transistors

The total emitter current is the sum of the electron diffusion current, $I_{E,n}$, the hole diffusion current, $I_{E,p}$ and the base-emitter depletion layer recombination current, $I_{r,d}$.

$$I_E = I_{E,n} + I_{E,p} + I_{r,d}$$

$$I_C = I_{E,n} - I_{r,B}$$

$$I_B = I_{E,p} + I_{r,B} + I_{r,d}$$



(b)

The transport factor, α, and current gain β are defined by:

$$\alpha = \frac{I_C}{I_E}$$

If the collector current is almost equal to the emitter current, the transport factor, α, approaches one. The current gain, β, can therefore become much larger than one.

$$\beta = \frac{I_C}{I_B} = \frac{\alpha}{1-\alpha}$$

Rewriting the transport factor, α, as the product of the emitter efficiency, γ_E, the base transport factor, α_T, and the depletion layer recombination factor, δ_r. The depletion layer recombination factor, δ_r, equals the ratio of the current due to electron and hole diffusion across the base-emitter junction to the total emitter current.

$$\alpha = \alpha_T \, \gamma_E \, \delta_r$$

Bipolar Junction Transistors

A bipolar transistor with an emitter current of 1 mA has an emitter efficiency of 0.99, a base transport factor of 0.995 and a depletion layer recombination factor of 0.998. Calculate the base current, the collector current, the transport factor and the current gain of the transistor.

The transport factor and current gain are: $\alpha = \gamma_E \alpha_T \delta_r = 0.99 \times 0.995 \times 0.998 = 0.983$

and $\beta = \dfrac{\alpha}{1 - \alpha} = 58.1$

The collector current then equals $I_C = \alpha I_E = 0.983\ \mathrm{mA}$

And the base current is obtained from: $I_B = I_E - I_C = 17\ \mu\mathrm{A}$

We now have current gain -- how do we switch it on and off?

Bipolar Junction Transistors: Bias Modes

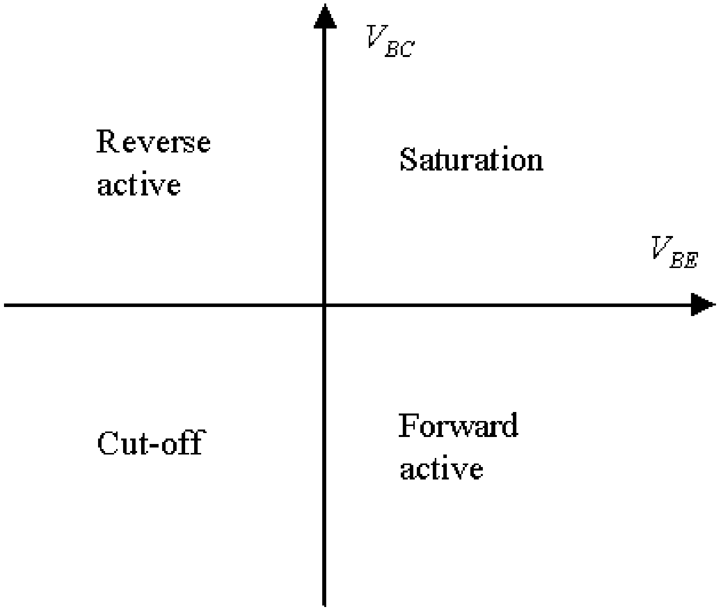
While the forward active mode of operation is the most useful bias mode when using a bipolar junction transistor as an amplifier, one cannot ignore the other bias modes especially when using the device as a digital switch. These bias modes include the forward active mode of operation, the reverse active mode of operation, the saturation mode and the cut-off mode.

The forward active mode is the one where we forward bias $V_{BE} > 0$ and reverse bias $V_{BC} < 0$. This mode is the one used in bipolar transistor amplifiers.

In bipolar transistor logic circuits, one frequently switches the transistor from the “off” state (cut-off) to the low resistance “on” state (saturation)

In the cut-off mode, both junctions are reversed biased, $V_{BE} < 0$ and $V_{BC} < 0$, so that very little current goes through the device. This corresponds to the “off” state of the device.

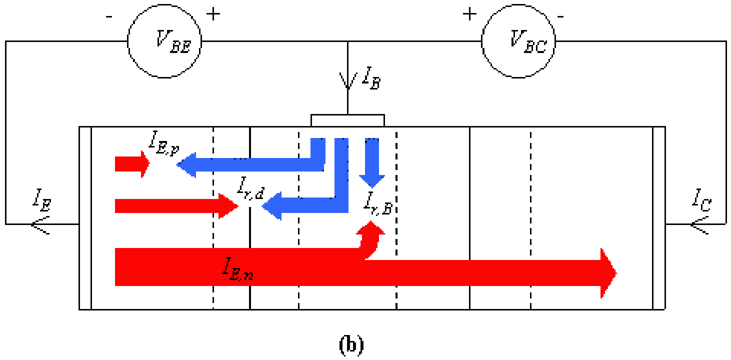
In the saturation mode, both junctions are forward biased, $V_{BE} > 0$ and $V_{BC} > 0$. This corresponds to the low resistance “on” state of the transistor.



Bipolar Junction Transistors: Forward Bias

The ideal transistor model is based on the ideal p-n diode model:

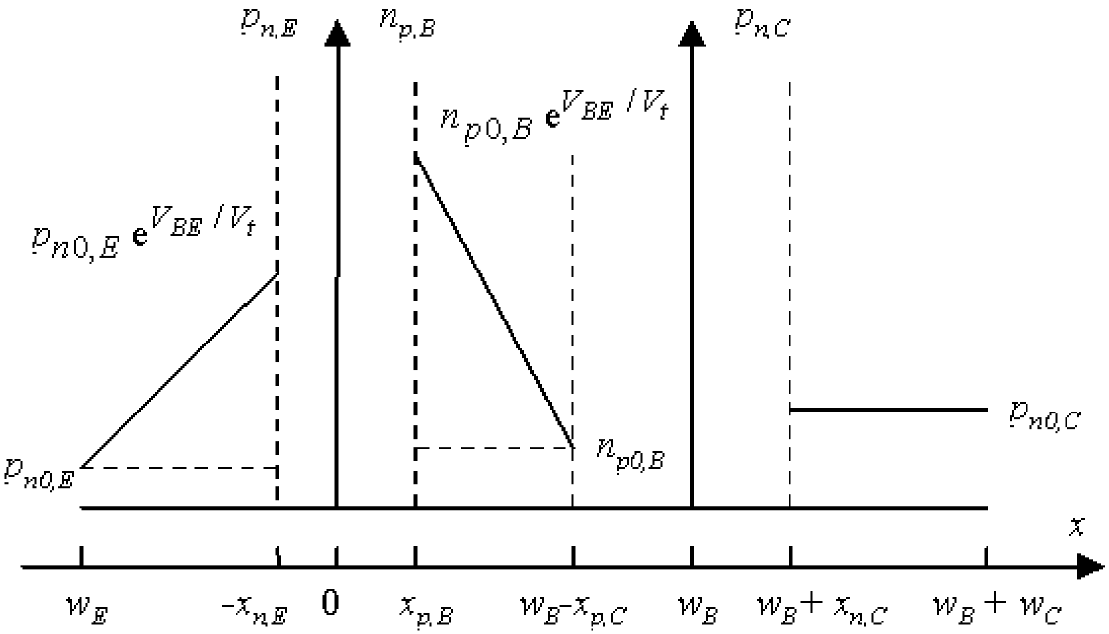
- all quasi-neutral regions are much smaller than the minority-carrier diffusion lengths.
- no recombination within the depletion regions is taken into account.



The forward active mode is obtained by forward-biasing the base-emitter junction.

The base-collector junction current is eliminated by setting $V_{BC} = 0$.

The minority-carrier distribution in the quasi-neutral regions of the bipolar transistor is graphed.



Bipolar Junction Transistors: Forward Bias

The emitter current due to electrons and holes are obtained using the "short" diode expressions, yielding:

$$I_{E,n} = q n_i^2 A_E \left(\frac{D_{n,B}}{N_B w_B'} \right) \left(\exp(\frac{V_{BE}}{V_t}) - 1 \right) \qquad I_{E,p} = q n_i^2 A_E \left(\frac{D_{p,E}}{N_E w_E'} \right) \left(\exp(\frac{V_{BE}}{V_t}) - 1 \right)$$

This charge is proportional to the triangular area in the quasi-neutral base:

$$\Delta Q_{n,B} = q A_E \int_{x_{p,E}}^{w_B - x_{p,C}} n_p(x) - n_{p0} \, dx$$

For “short” diode: $$\Delta Q_{n,B} = q A_E \frac{n_i^2}{N_B} \left(\exp(\frac{V_{BE}}{V_t}) - 1 \right) \frac{w_B'}{2}$$

So, $$I_{E,n} = \frac{\Delta Q_{n,B}}{t_r}$$

Bipolar Junction Transistors: Forward Bias

Next, we need to find the emitter efficiency and base transport factor. It is typically the emitter efficiency, which limits the current gain in transistors made of silicon or germanium.

$$\gamma_E = \frac{1}{1 + \frac{D_{p,E} N_B w'_B}{D_{n,B} N_E w'_E}}$$

The long minority-carrier lifetime and the long diffusion lengths in those materials justify the exclusion of recombination in the base or the depletion layer. The resulting current gain, under such conditions, is:

$$\beta \cong \frac{D_{n,B} N_E w'_E}{D_{p,E} N_B w'_B}, \quad \text{if} \quad \alpha \cong \gamma_E$$

A typical current gain for a silicon bipolar transistor is 50 - 150.

The base transport factor equals:

$$\alpha_T = 1 - \frac{t_r}{\tau_n} = 1 - \frac{w'^2_B}{2 D_{n,B} \tau_n}$$

BiPolar Junction Transistors (BJT): Other Effects

- **Base-width modulation:** As the voltages applied to the base-emitter and base-collector junctions are changed, the depletion layer widths and the quasi-neutral regions vary as well. This causes the collector current to vary with the collector-emitter voltage.

- **Recombination in the depletion region:** as in a p-n diode, the recombination in the depletion region causes an additional diode current.

- **High injection effects:** as in p-n diode, igh injection effects occur in a BJT

- **Base spreading resistance and emitter current crowding:** Large area BJTs can have a very non-uniform current distribution due to the resistance of the base layer. This resistance causes a voltage variation across the base region. This voltage variation then causes a variation of the emitter current density, especially since the emitter current density depends exponentially on the local base-emitter voltage.

- **Temperature dependence:** BJT typically have only weak T-dependence. The base transport reduces with temperature, primarily because the mobility and recombination lifetime are reduced with increasing temperature. Occasionally the transport factor initially increases with temperature, but then reduces again.

- **Breakdown:** The breakdown mechanisms of BJTs are similar to that of p-n junctions. Since the base-collector junction is reversed biased, it is this junction where breakdown typically occurs. Just like for a p-n junction the breakdown mechanism can be due to either avalanche multiplication as well as tunneling.