

# Appendix A

## Appendix A

**Physics, SPICE Parameters & Transistor Models**

# Semiconductors

## Fundamentals

All transistors, diodes, BJT's, FET's, etc are built using some kind of semiconductor material. The most common base material is Silicon. Pure silicon, also called intrinsic silicon, is not a good conductor because there are exactly four electrons in the outer shell. In crystalline form, each silicon atom bonds covalently with another silicon atom which effectively ties up all the free electrons. It requires a lot of energy to strip an electron out of a covalent bond and make it available for charge (and therefore current) movement. What free electrons do exist are due to thermal energy and this is not enough energy to strip many electrons out of their bonds and move them into the "conduction band" as its called. Hence for pure silicon, the resistivity is very high. The density of free carriers available for movement is modeled by a parameter called  $n_i$  where:

$$\text{Equation: A-1 } n_i^2 = KT^3 e^{\frac{-E_{go}}{kT}}$$

K is a constant.

k is Boltzmann's constant.

$E_{go}$  is the energy gap extrapolated to absolute zero. It is 1.2 electron volts for silicon.

T is the temperature in degrees Kelvin.

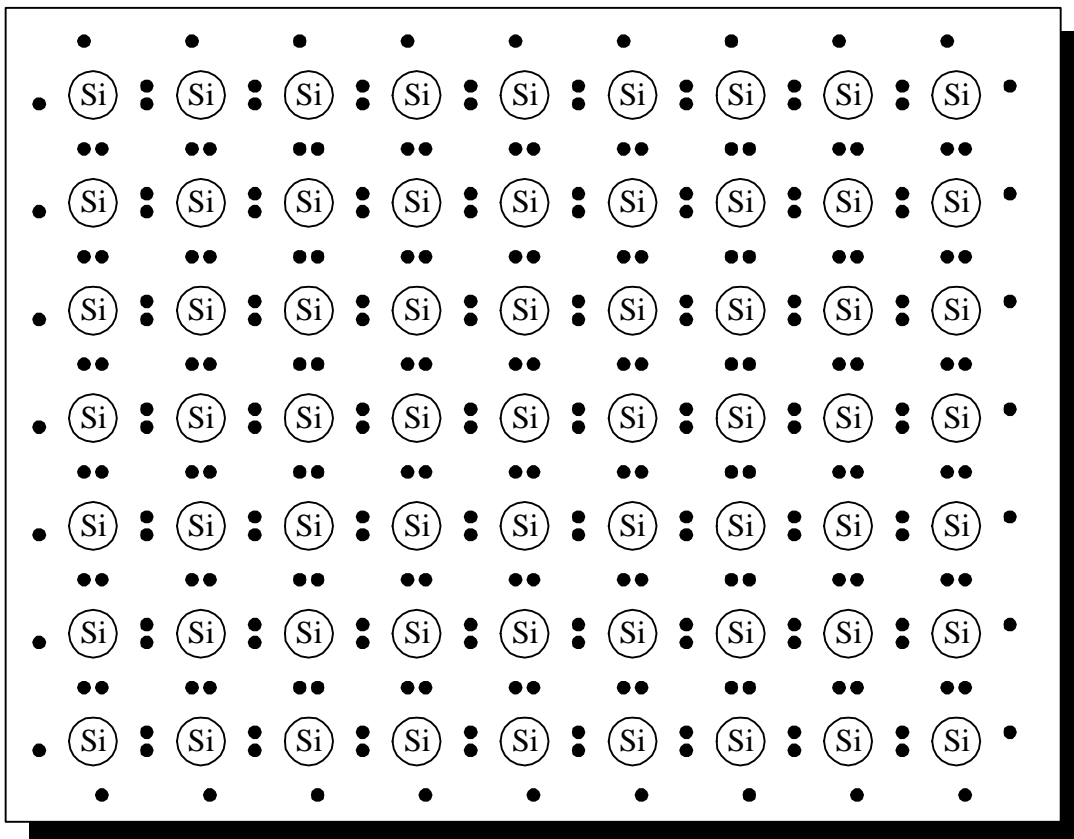


Figure 6-1 Intrinsic Silicon

## Doping And N-Type Material

To make silicon conductive, a process called "doping" is used. To "dope" silicon, we diffuse into the silicon crystal either a chemical from the group 3 or the group 5 section of the periodic table. These are chemicals that have either 3 or 5 electrons in their outer shell. Commonly used elements are indium from

the group 3 section and arsenic from the group 5. If a group 5 element (arsenic) is added, the crystalline structure now has an “extra” electron. This extra electron requires very little energy to move into the conduction band, and in fact, for each group 5 atom added, at room temperature, a free electron is added. For this reason, we say that group 5 atoms are “donor” atoms because they “donate” free electrons to the material. This makes the material very conductive. The concentration of donor atoms is called  $N_d$ , and is a process parameter that is controlled very tightly. When silicon is doped with donor material, we call it “N Type” material, the N standing for negative (electrons). When an electron does free up to the conduction band, it leaves behind a positive charge because of the extra proton in the nucleus that is no longer balanced by the fifth electron. We call this a fixed charge because, being bound to the nucleus, it is not free to move.

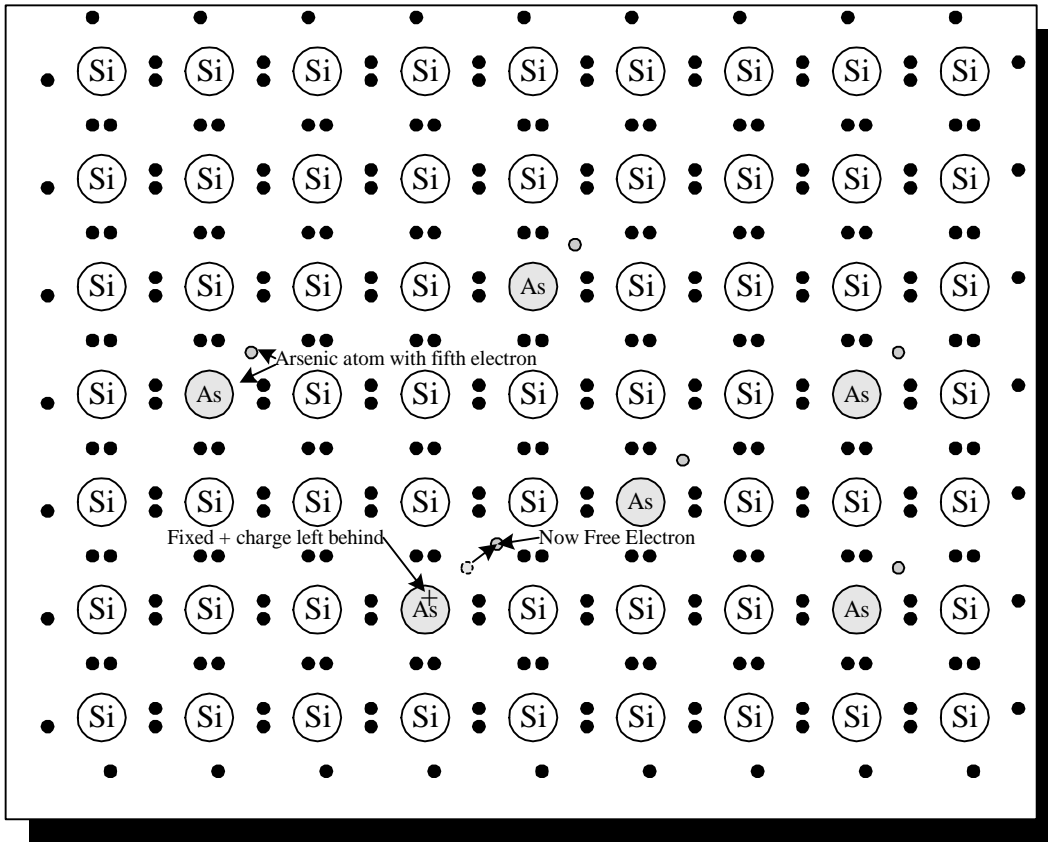
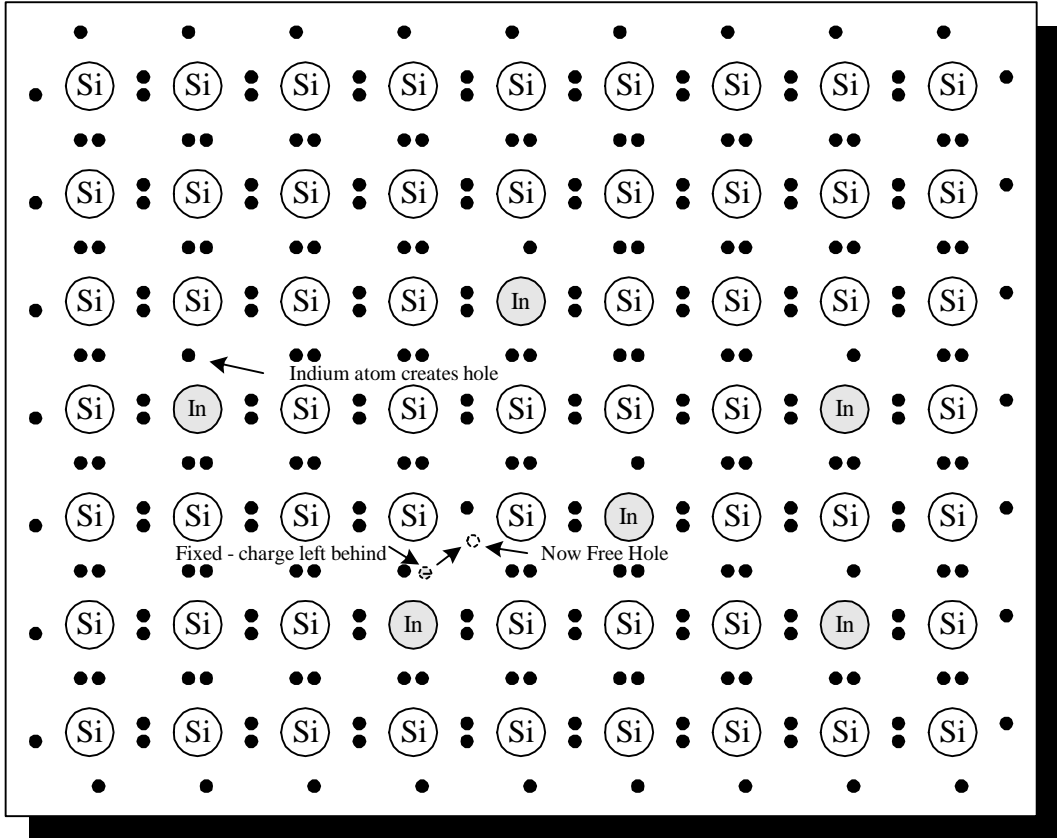


Figure 6-2: N Type Material

## P Type Material

When we dope with a group 3 atom, we create a different type of material (called P Type) that's equally conductive as N type, but for a different reason. A group 3 atom has only 3 electrons in its outer shell. This allows for the easy movement of an electron into the shell structure. Of course when an electron does move into the vacancy in the outer shell created by the addition of the group 3 atoms, it leaves behind a similar vacancy in the crystal structure of an adjacent set of atoms. These vacancies are called “holes”, and they allow for a current flow. Holes are artificial structures that behave just like an electron, except they have a positive charge. We can even assign a mass to them (which isn't too different from an electron mass). Of course, actual current flow is still due to the movement of electrons, but for P type material, it is much easier to talk in terms of the movement of holes. At room temperature, for each group 5 atom added, a hole is created. Group 5 atoms are called “acceptor” atoms because they “accept” electrons. The concentration of acceptor atoms is called  $N_a$ , and is another process parameter that is very tightly controlled. When a hole does free up to the conduction band, it leaves behind a negative charge because of the extra electron that is not balanced by a proton in the nucleus. We call this a fixed charge because it is not free to move.



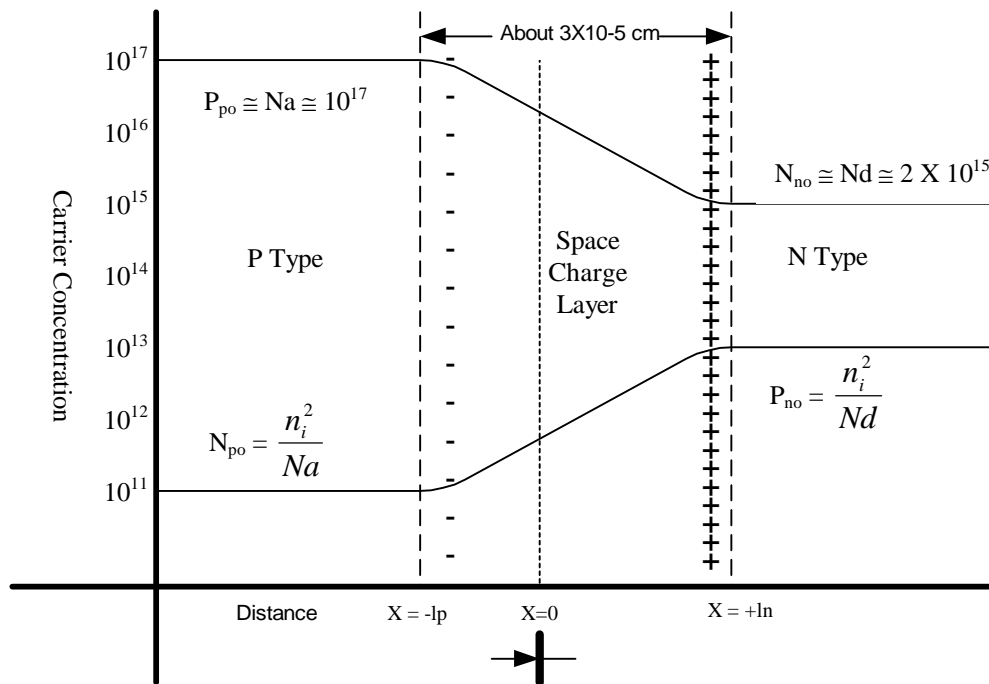
**Figure 6-3: P Type Material**

Material where  $N_a$ , or  $N_d$  is much greater than  $n_i$  is called extrinsic material. In such material the number of free carriers is approximately equal to the doping level (either  $N_a$  or  $N_d$ ).

# The Diode

A diode is formed by creating a junction of N type and P type material without disrupting the crystalline structure.

N type material has a large number,  $N_d$ , of free electrons while P type has a large number,  $N_a$ , of free holes. When a junction of these two materials is formed, a large imbalance of holes and electrons exists across the junction. This sets up a very large flow of electrons into the P type material and a very large flow of holes into the N type material due to diffusion. This current flow is eventually stopped because the flow of electrons leaves behind a fixed positive charge and the flow of holes leaves behind a fixed negative charge. This fixed charge creates an electric field that acts to counter the flow of diffusion current. Eventually equilibrium is established with the diffusion current being exactly canceled due to the electric field established by the fixed charge. A region in the semiconductor that contains this electric field is called the Space Charge Layer (SCL). It is also called the depletion region, because it is "depleted" of free carriers. As soon as a free carrier enters this junction, it is immediately swept to the other side by the extremely high electric fields existing in this region. The SCL is typically very small, on the order of  $10^{-5}$  cm. This creates very large electric fields on the order of  $10^5$  volts per cm. This electric field creates a voltage called the "built in voltage" or "contact potential". This voltage is designated  $\psi_0$  and is usually between .3 volts and a little over 1 volt. Because the amount of diffusion current is directly proportional to the slope of the free carrier profile, the higher the doping level, the greater the diffusion current, the greater the  $\psi_0$ .



**Figure 6-4: The Junction Of A PN Diode With No Applied Voltage**

- $N_a$**  is the acceptor doping level in P type material.
- $N_d$**  is the donor doping level in N type material.
- $P_{po}$**  is the majority (holes) carrier concentration in P type material.
- $N_{po}$**  is the minority (electron) carrier concentration in P type material.
- $N_{no}$**  is the majority (electron) carrier concentration in N type material.
- $P_{no}$**  is the minority (hole) carrier concentration in P type material.
- $n_i$**  is the intrinsic carrier concentration.

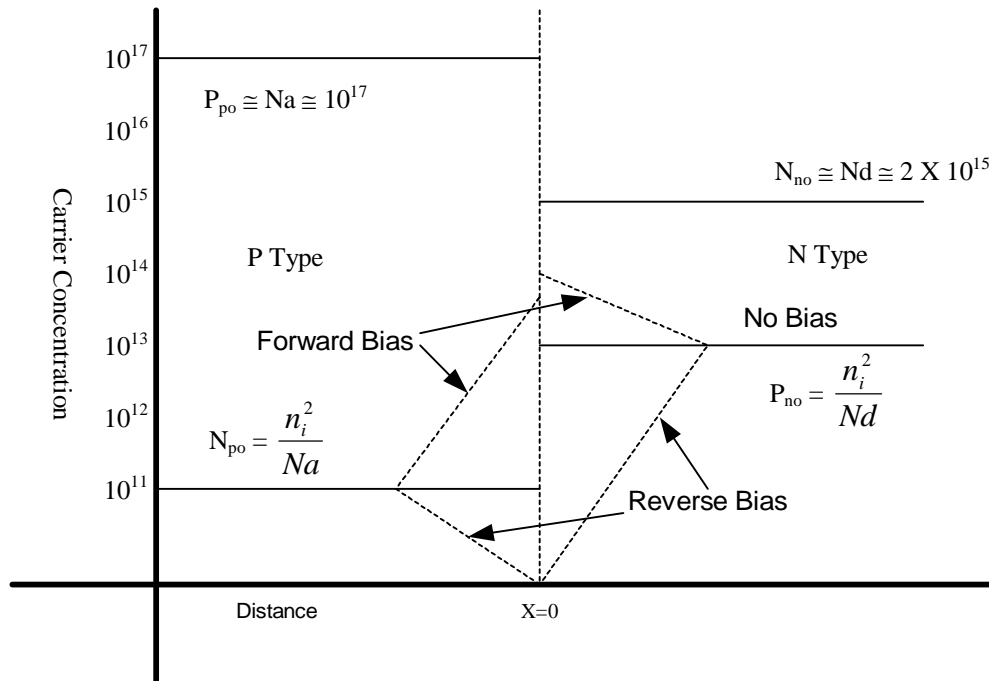


Figure 6-5: The Junction of a PN diode with Compressed Axis

### Forward bias

What happens to a diode under forward bias conditions? A diode is forward biased when an external voltage is applied to the diode in such a way as to reduce or subtract from the built in voltage. Under these conditions, equilibrium is established only with a current flow. The built in voltage has been reduced so the flow of carriers across the junction must increase in order to rebuild the fixed charge to a sufficiently high level to again come to equilibrium. The interesting thing is to examine what is happening on either side of the SCL and on into the neutral regions. Forward bias “injects” minority carriers into both sides of the SCL effectively increasing the concentration of minority carriers well above background at these points. The actual concentration is exponentially related to the applied forward voltage. The background concentration of minority carriers is given by:

**Equation 6-1**  $N_{po} = \frac{n_i^2}{Na}$  and  $P_{no} = \frac{n_i^2}{Nd}$

The concentration of minority carriers is much higher at the boundaries of the SCL over that at a great distance from the SCL. This creates a gradient of minority carriers close to the SCL. Current flows away from the SCL because of diffusion caused by this gradient. As the minority carriers diffuse outward, they gradually recombine with the majority carriers. This converts the current flow from a diffusion of minority carriers to a flow of majority carriers. Any amount of current can flow, because the concentration of minority carriers next to the SCL can be set to any amount. This in turn sets the diffusion slope to be whatever is required to carry the current.

### Reverse Bias

Under reverse bias condition, the applied voltage is additive to the built in voltage. Minority carrier concentrations are reduced at the boundaries of the SCL, and minority carriers diffuse toward the SCL to be swept across the junction. This works exactly the same as in the forward bias case until the minority carrier concentration at the SCL boundary is reduced to zero. Because the current in this region is carried by diffusion, and because the diffusion current is directly proportional the slope of the minority carrier charge concentration, current cannot increase beyond a small amount ( $I_s$ ) because the slope cannot be increased once zero concentration is reached. The slope is defined by the points  $N_{po}$  ( $P_{no}$ ) and zero.

Current cannot be increased beyond what this slope allows no matter how much the reverse voltage is increased. What does happen is the SCL gets wider as more and more of the neutral region of the device is depleted of carriers. This does not result in any substantial increase in current, but it does change other diode parameters such as junction capacitance. It will also cause other parameter changes in a Bipolar Junction Transistor. This will be discussed shortly.

## Width Of The SCL

The width of the SCL is such an important parameter, it is worthy of a few more words. Define the SCL width to be  $L$ . then it can be shown that  $L$  is equal to:

$$\text{Equation 6-2 } L = \sqrt{\frac{2\varepsilon(\psi_0 - V_{\text{applied}})}{q} \left( \frac{1}{Na} + \frac{1}{Nd} \right)}$$

$\varepsilon$  is the dielectric constant for Silicon.

$q$  is the charge on an electron

$Na$ , and  $Nd$  are the acceptor and donor doping concentrations.

$V_{\text{applied}}$  is the forward biased applied voltage

$\psi_0$  equals the built in voltage and is equal to:

$$\text{Equation 6-3 } \psi_0 = V_t * \ln \left( \frac{NaNd}{n_i^2} \right)$$

$$\text{Equation 6-4 } L = |lp| + |ln|$$

Where  $lp$  and  $ln$  are defined in Figure 6-5

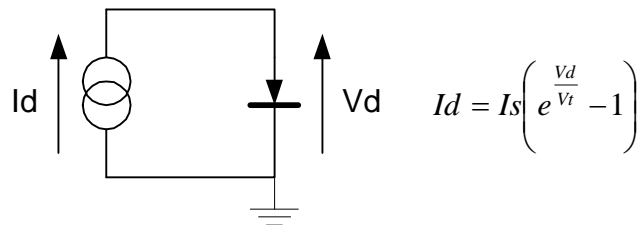
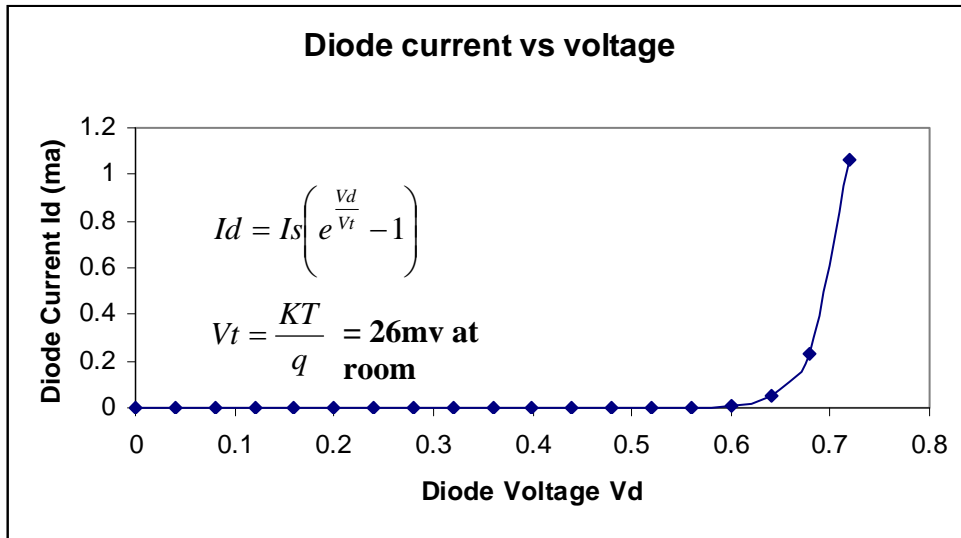
It can be shown that:

$$\text{Equation 6-5 } ln = \frac{Na * L}{Na + Nd}$$

And

$$\text{Equation 6-6 } lp = \frac{Nd * L}{Na + Nd}$$

These last two equation are interesting because they show that for highly asymmetrical doping levels ( $Na \gg ND$  or the other way around), most of the SCL will exist in the region of lightly doped material.



**Figure 6-7: A Forward Biased Diode**

Emission Coefficient

A third parameter, called the emission coefficient ( $ne$ ) is also a critical parameter in the diode equation. With this parameter the diode equation becomes:

$$\text{Equation 6-7 } I_d = I_s \left( e^{\frac{V_d}{neV_t}} - 1 \right)$$

$$\text{Equation 6-8 } V_t = \frac{KT}{q} \cong 26 \text{ mvolts at room temperature}$$

$K$  is Boltzmann's Constant

$T$  is temperature in degrees Kelvin

$q$  is the charge on an electron

$I_s$  is the diode saturation current and is typically  $10^{-19}$  amps

To measure  $I_s$  and  $ne$ , it is necessary to take  $I_d$  vs  $V_d$  data and then plot the  $\text{Log}_e$  of  $I_d$  vs  $V_d/V_t$  as shown below.

Solving Equation 6-7 for this new form gives:

$$\text{Equation 6-9 } \ln(I_d) = \ln(I_s) + \frac{V_d}{neV_t}$$

$$\text{Let } y = \ln(I_d) \text{ and } x = \frac{V_d}{V_t}$$

then we can generate the linear equation

$$\text{Equation 6-10 } y = mx + b$$

$$\text{Where } m = \frac{1}{ne}, \text{ and } b = \ln(I_s)$$

## Series Resistance

A diode will have some resistance in the neutral region (region far from the junction). We call this resistance the series bulk resistance,  $R_s$ .

$$\text{Let } R_s = \text{Series Bulk resistance} = \frac{\Delta v}{\Delta i}$$

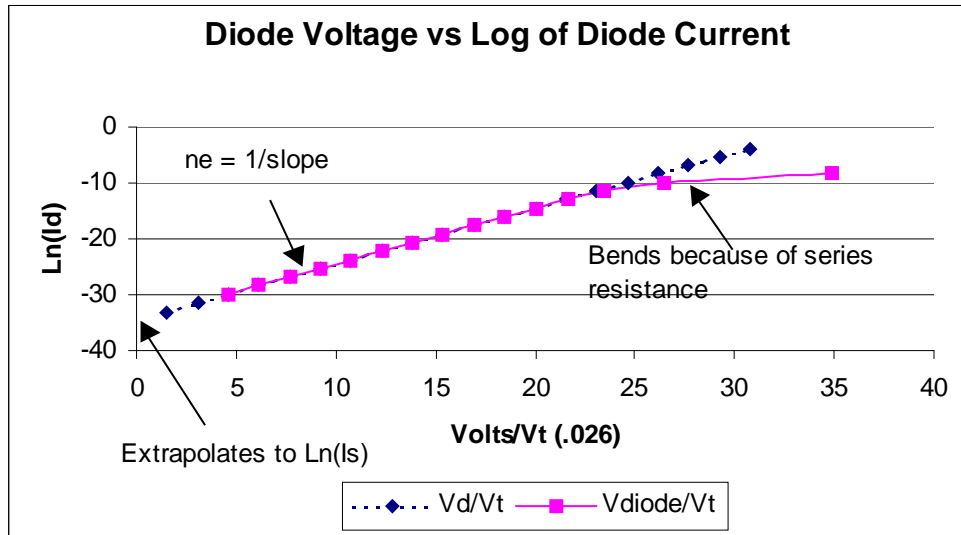


Figure 6-8: Plot Of Diode Voltage Vs Log Of Diode Current With And Without Series Resistance.

## Saturation Current, $I_s$

Saturation current is a very important parameter for both diodes and transistors and as such it is worthy of more discussion. Right at the edge of the SCL, current flow is due to the diffusion of minority carriers. On the P side, the current is due to the diffusion of holes and on the N side, current is due to diffusion of electrons. The total current flow is the sum of these two diffusion currents. It can be shown that  $I_s$  is equal to:

$$\text{Equation 6-11 } I_s = A * q * n_i^2 \left[ \frac{D_n}{L_n * N_a} + \frac{D_p}{L_p * N_d} \right]$$

Where

$A$  is the emitter area

$q$  is the charge on an electron

$D_n$  is the diffusion constant for electrons

$D_p$  is the diffusion constant for holes

$L_n$  is the diffusion length for electrons

$L_p$  is the diffusion length for holes

$N_a$  is the acceptor doping level in the P type material

$N_d$  is the donor doping level in the N type material.

$n_i$  is the number of broken bonds per cubic cm due to thermal action.

The total current then is equal to:

$$\text{Equation 6-12 } I_t = I_n + I_p = \frac{A * q * n_i^2 * D_n * \left( e^{\frac{V_{be}}{V_t}} - 1 \right)}{L_n * N_a} + \frac{A * q * n_i^2 * D_p * \left( e^{\frac{V_{be}}{V_t}} - 1 \right)}{L_p * N_d}$$

If one looks at the ratio of  $I_n$  to  $I_p$ , one finds:

$$\text{Equation 6-13 } \frac{I_n}{I_p} = \left( \frac{D_n * L_p}{L_n * D_p} \right) \left( \frac{N_d}{N_a} \right)$$

It can be shown that:

$$\text{Equation 6-14 } \frac{Dn * Lp}{Ln * Dp} \approx 1$$

Then

$$\text{Equation 6-15 } \frac{In}{Ip} \approx \frac{Nd}{Na}$$

Notice that if the donor doping  $Nd$  is much greater than  $Na$ , then  $In$  is much greater than  $Ip$ , and  $I_t = I_n$

This says that in a diode that has a heavily doped N type material and a lightly doped P type material, most of the current will be due to electrons diffusing away from the junction on the P side. The minority carriers (holes on the N side) will be very small in comparison. In effect, for a diode with a heavily doped N region and a lightly doped P region, and with an applied forward bias, many more electrons will be injected into the P side than holes in the N side.

This ratio will become very important in the transistor because any minority carriers injected from the base into the emitter shows up as base current. The ratio of minority carriers injected from the emitter into the base to the total number of minority carriers injected from the base to the emitter and from the emitter to the base is called the emitter efficiency,  $\gamma$ .  $\gamma$  approaches one in the limit.  $\alpha_f$ , the ratio of collector current to emitter current, cannot exceed  $\gamma$ . This of course, places an upper limit on  $\beta_f$ . It is for this reason that emitters are very heavily doped relative to the base region on a bipolar junction transistor.

## Junction Capacitance

As discussed previously, the SCL is a region inside the semiconductor that is devoid (depleted) of free charge. Across this region is an electric field that generates a voltage,  $\psi_0$ , called the built in voltage. There is fixed charge residing on the boundaries of the SCL. Moving away from the SCL boundaries, free carriers in the form of electrons or holes are found in abundance. In effect we have created a capacitor. The width of the capacitor plates is equal to the width of the SCL (on the order of  $10^{-5}$  cm). However, this width changes with applied voltage effectively making the capacitance dependent on the voltage applied externally to the diode. If a larger reverse voltage is applied, the SCL gets wider, reducing the value of the capacitance. Spice labels the junction capacitance as  $C_j$  and models  $C_j$  as follows:

$$\text{Equation 6-16 } C_j = \frac{C_{j0}}{\left[1 - \frac{V_d}{\Phi_c}\right]^m}$$

$\Phi_c$  is junction potential (equal to  $\psi_0$  and in Spice is referred to as  $V_j$ ) and is about 1 volt for Silicon

$I_s$  is the saturation current and is about  $10^{-14}$  amps

$m$  is the grading coefficient and is between .5 and .333 (Geometry dependent. For an abrupt junction,  $m$  is closer to .5, for a linearly graded junction,  $m$  is closer to .333)

$V_d$  is the voltage across the diode.

$C_{j0}$  is the zero bias junction capacitance.

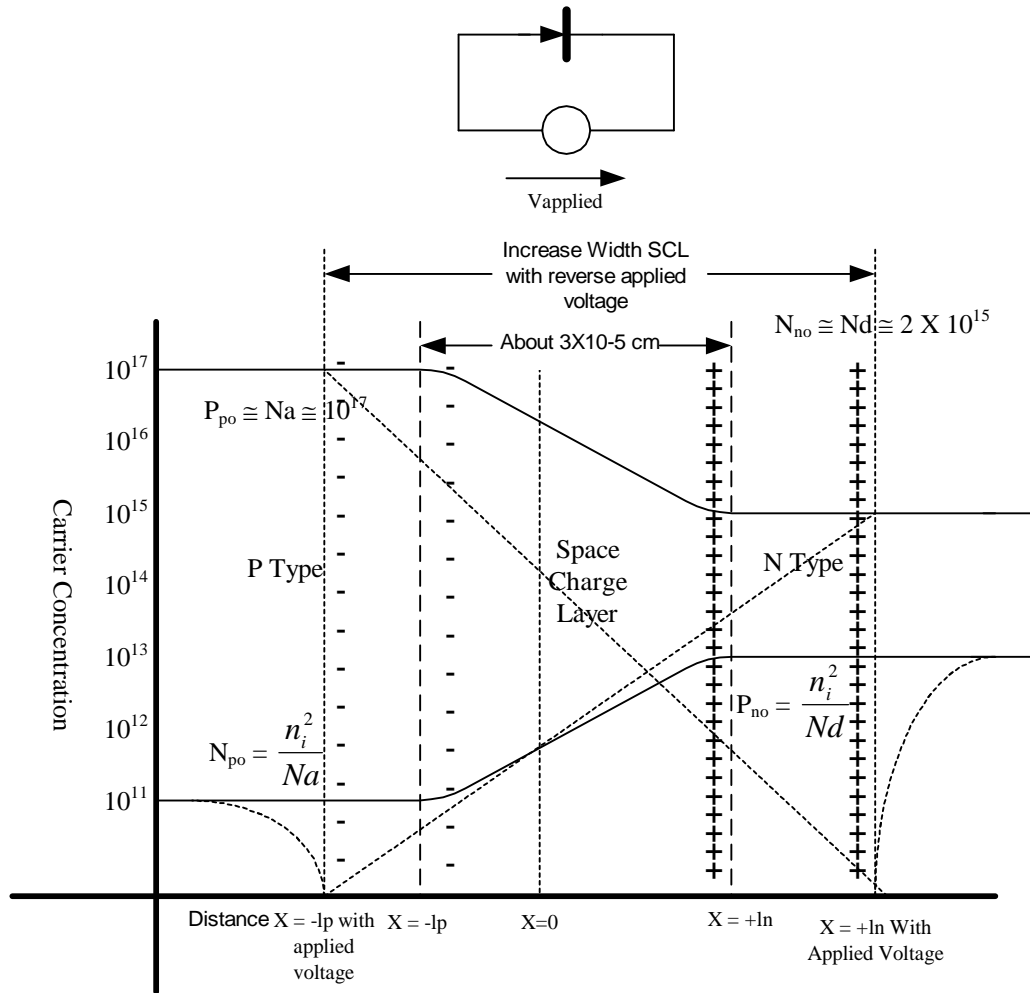


Figure 6-9: Diode Junction With A Reverse Applied Voltage Increasing The Width Of The SC

### Diffusion Capacitance

During forward conduction there are free carriers moving inside the SCL across the junction. For whatever time it takes to transit the junction, these carriers constitute a charge. This makes the junction look like a capacitor (different from the earlier described junction capacitance). The value of this capacitor is dependent on the time the carriers are in the SCL, and the number of carriers present. The time is modeled in Spice with a parameter called the Transit Time and is on the order of 100 psec. Spice uses a label for this parameter of  $T_t$ . Because I'm going to use the same label to represent something quite different later in this book, I'm going to label the Transit Time to be  $T_{transit}$ . The diffusion capacitance is labeled  $C_d$  and is equal to:

Equation 6-17  $C_d = T_{transit} * G_p$

Where

Equation 6-18  $G_p = \frac{I_d}{V_t}$  and is the small signal conductance of the diode.

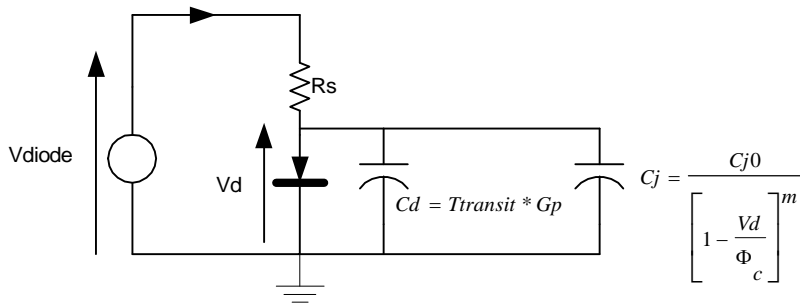


Figure 6-10 : A Large Signal Model For A Diode

### High Frequency Diode Model.

In Figure 6-10.,  $C_j$  dominates in the reverse biased region of the diode because  $I_d$  is zero which makes  $G_p$  zero. Therefore, for reverse bias,  $C_d$  is zero. In the forward region,  $C_d$  begins to dominate. For example: Assume that  $T_t = 100$  psec,  $I_d = 10$  ma, and  $V_t$  is 26 mv, then

#### Equation 6-19

$$\text{eq4: } C_d = \frac{T_t I_d}{V_t} = 38 \text{ pf}$$

This is much larger than the junction capacitance for a typical diode. A plot of the total diode capacitance is shown in Figure 6-11.

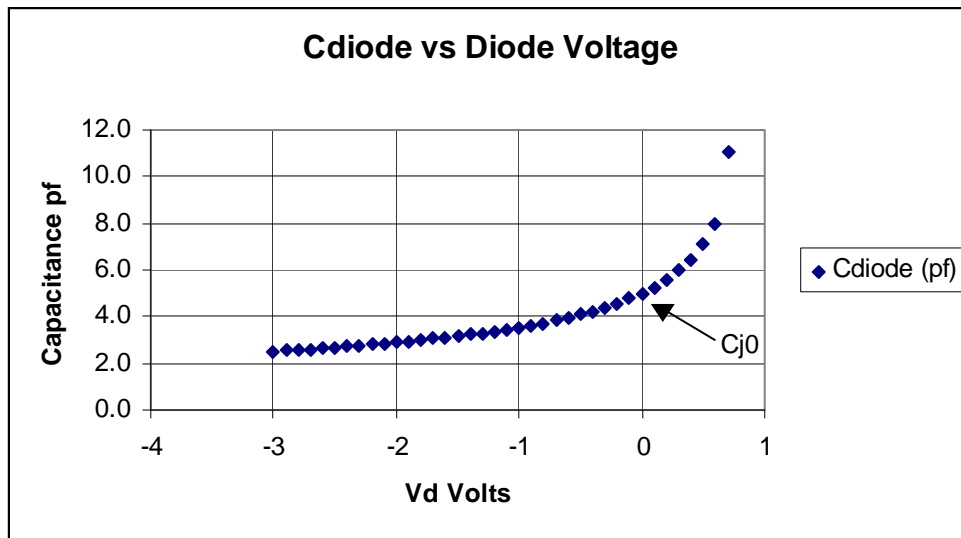
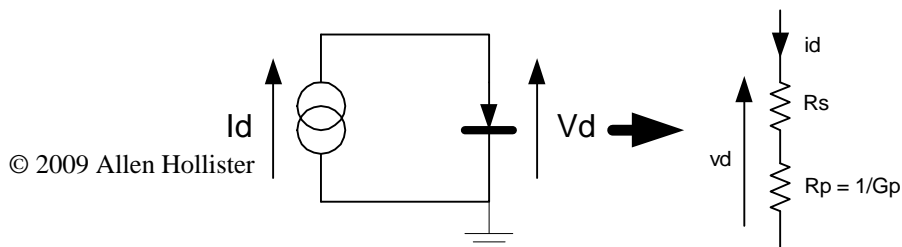


Figure 6-11: Diode Capacitance vs. Applied Voltage.

### A Small Signal, Low Frequency Model For The Diode.

For the forward biased condition, the diode looks like a resistor at low frequencies. The resistance is highly variable with diode current and can be derived in the following way.



### Figure 6-12: A Small Signal Model For The Diode

For forward bias,  $I_d$  is approximately equal to:

**Equation 6-20:**  $I_d = I_s * e^{\frac{V_d}{V_t}}$

To find the small signal conductance, difference this equation with respect to  $V_d$ , then let  $I_d$  be equal to the bias point.

**Equation 6-21:**  $G_p = \left. \frac{dI_d}{dV_d} \right|_{I_d=I_b} = \frac{I_s * e^{\frac{V_d}{V_t}}}{V_t} = \left. \frac{I_d}{V_t} \right|_{I_d=I_b} = \frac{I_b}{V_t}$

## Junction Breakdown

Diode junctions “breakdown” or suddenly go from high resistance in the reverse region to very low resistance due to two mechanisms, 1) Avalanche breakdown and Zener Breakdown.

### Avalanche Breakdown

Avalanche breakdown occurs when the electric fields in the space charge layer is large enough so that carriers traversing the region acquire sufficient energy to break covalent bonds in collisions with the crystal structure. Every such ionizing collision produces a hole and an electron, each of which is accelerated by the field and has the possibility of producing another ionizing collision before it leaves the SCL. Neglecting recombination in the layer, all of the carriers produced will contribute to the total reverse current. In this manner, the very small reverse current,  $I_s$ , is greatly amplified for a very small increase in reverse voltage. Avalanche breakdown typically occurs for voltages greater than about five volts for silicon.

### Zener Breakdown

Zener breakdown is the direct disruption of interatomic bonds in the SCL by very high electric fields (greater than  $10^6$  volts/cm), which produces mobile hole-electron pairs. Zener breakdown occurs in abrupt junctions between highly doped regions and depends only on the strength of electric field present. Avalanche breakdown occurs when the acceleration of carriers in the SCL is great enough to cause ionizing collisions with atoms. Avalanche breakdown depends on both the electric field and on the distance available to accelerate the carriers.

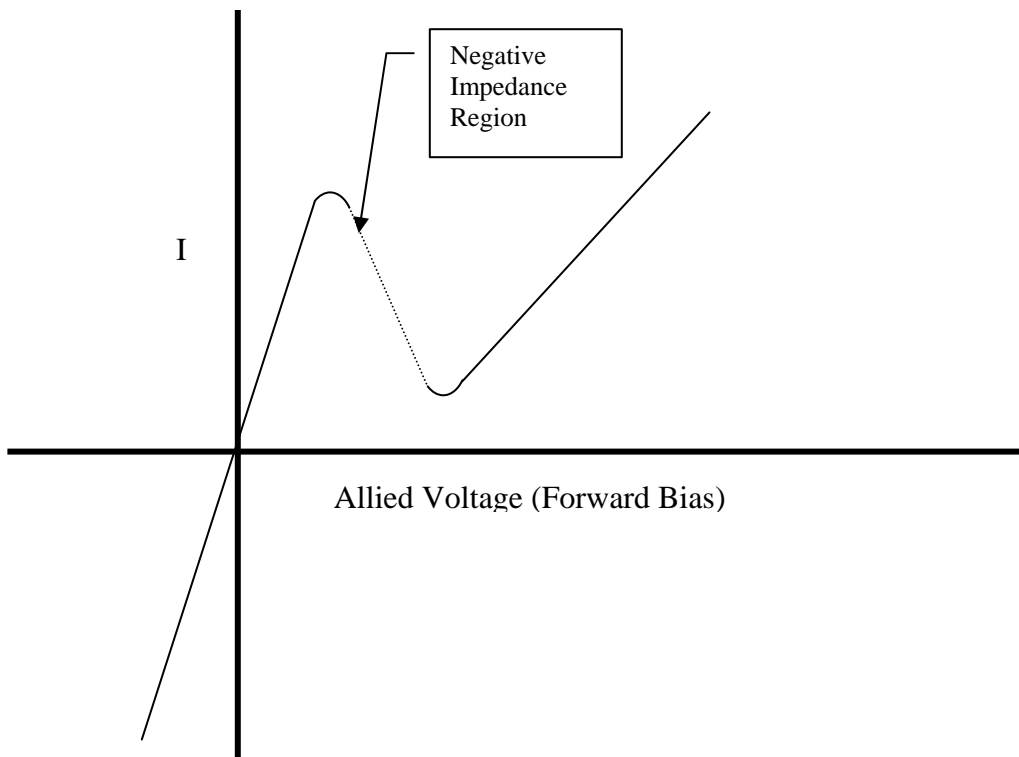
## Tunnel Diode

A reasonable question is can the doping levels ( $N_a$  and  $N_d$ ) be increased to the point where the built in voltage,  $\Psi_0$ , is sufficient to cause Zener breakdown? Repeating Equation 6-2 and Equation 6-3 here, it is obvious that the SCL width ( $L$ ) gets smaller and the built in voltage gets larger with higher doping concentrations. It is entirely possible to increase the doping concentrations to a level where the built in voltage is sufficient to cause Zener breakdown. (Notice that it won't cause avalanche breakdown, because the  $L$  becomes less with doping concentrations reducing the distance available for carriers to be accelerated to a point where they can knock carriers out of their bonds through kinetic energy.)

Equation 6-2  $L = \sqrt{\frac{2\epsilon(\Psi_0 - V_{applied})}{q} \left( \frac{1}{N_a} + \frac{1}{N_d} \right)}$

Equation 6-3  $\Psi_0 = V_t * \ln \left( \frac{N_a N_d}{n_i^2} \right)$

When this occurs, we create a new kind of device called a tunnel diode. In this diode the device is conductive even with a small applied reverse voltage because  $\Psi_0$  is sufficiently high to cause the device to be in zener breakdown. As we apply a forward voltage, current increases because the device is in breakdown. It also subtracts from  $\Psi_0$ , and at some point the combination voltage is insufficient to support zener breakdown, and the current immediately drops to a much lower value as the resistance increases. Continuing to increase the applied voltage cause current to begin to increase again, and at this point, it follows the standard diode equation. In the region where the device suddenly falls out of zener breakdown, is a negative impedance region. In other words as the voltage increases, the current decreases. This allows for amplification (or bistable operation), and in fact is a very fast. The switch from high current to low current in a tunnel diode is one of the fastest semiconductor switches known, and was used for many years in oscilloscopes as the trigger detector.



**Figure 6-13: V/I Characteristic for a Tunnel Diode**

# The Bipolar Junction Transistor

Figure 6-14 and Figure 6-15 show the construction of an NPN Transistor. Notice that it consists of a region of highly doped N type material called the emitter, a region of lightly doped P type material called the base, and a region of moderately doped N type material called the collector. All of these regions are in the same crystalline structure without any breaks, tears, or disruptions. The reason for the high doping level in the emitter and the low doping level in the base was described in the section on Saturation Current. Any minority carriers that are injected from the base into the emitter shows up as base current. We want all of the current flow to be due to minority carriers injected from the emitter into the base. This is done by making the ratio of emitter doping to base doping as high as possible.

## Regions of Operation

In this section, I am going to mainly discuss the transistor in the “normal” mode of operation. In this mode, the emitter base junction is forward biased, and the collector base junction is reversed biased. To be complete, the reader must be aware that there are three other permutations that are equally valid and used. However, this is the most common mode of operation, and once this mode is explained, the other modes will be largely obvious. To be complete, the other modes of operation are shown in Table 6-1

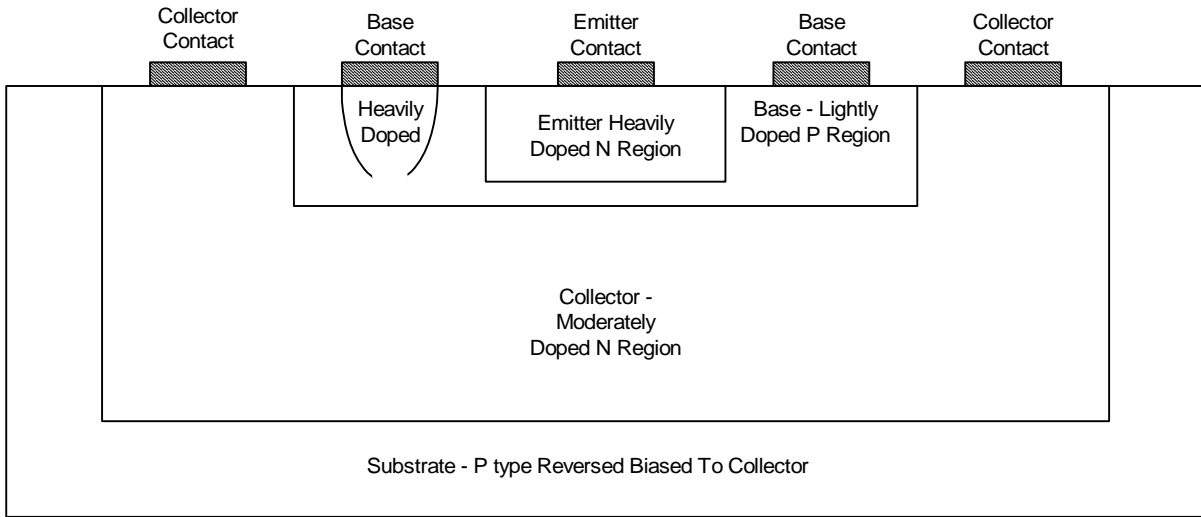
	<b>Emitter – Base Junction</b>	<b>Collector – Base Junction</b>
Normal Region	Forward Biased	Reversed Biased
Saturation Region	Forward Biased	Forward Biased
Inverse Region	Reverse Biased	Forward Biased
Cutoff Region	Reverse Biased	Reverse Biased

**Table 6-1: The Four Regions Of Operation For A Bipolar Junction Transistor.**

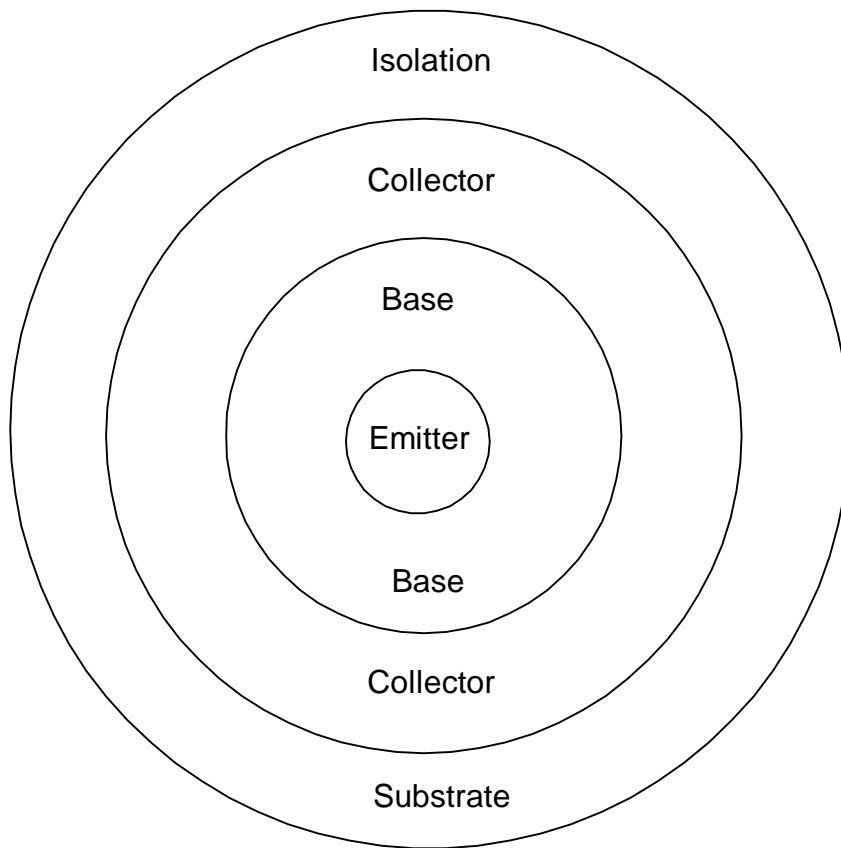
Much of the theory needed to understand transistor operation has already been developed from the discussion on diodes. So rather than developing new mathematics for the transistor, the equations already developed for the diode will be used. The only difference is that the transistor has an extra junction.

## Diffusion Of Minority Carriers In The Base Region

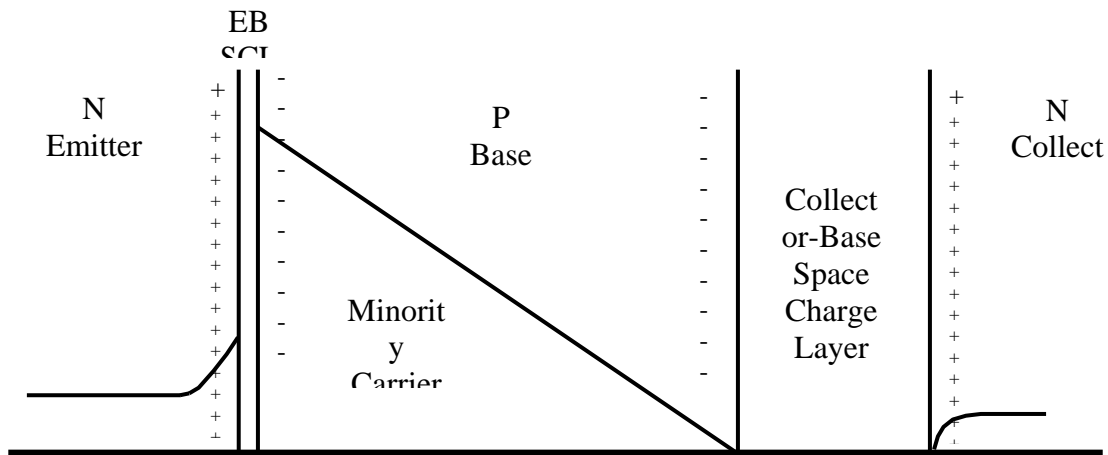
A minority carrier diagram for a transistor operated in the normal region of operation is shown in Figure 6-17. Notice that minority carriers are injected into the base from the emitter by applying a forward voltage ( $V_{be}$ ) across the emitter-base junction. The collector-base junction is reverse biased which causes the minority carriers in the base region at the collector-base junction to go to zero. This sets two points on the minority carrier concentration in the base region, a high concentration at the emitter-base junction, and zero at the collector-base junction. It is assumed that a straight line connects these two points. In reality this line isn't completely straight because some of the minority carriers recombine with majority carriers before they reach the collector.



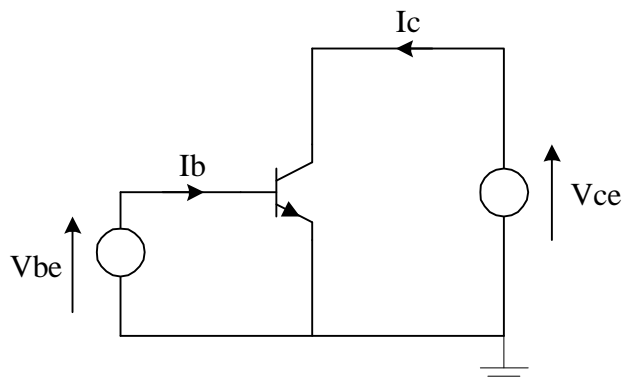
**Figure 6-14: Planer Transistor, Side View**



**Figure 6-15: Transistor, Top View.**



**Figure 6-16: A minority carrier diagram for an NPN transistor biased in normal region of operation.**



**Figure 6-17: An NPN Transistor Biased In The Normal Region Of Operation.**

Any recombination that occurs shows up as base current, an unwanted effect. Because of the uneven concentration, minority carriers diffuse across the base region until they hit the collector-base junction where they are swept across the junction into the collector region by the high electric field that exists in this region.

### **Emitter, Base, Collector Doping Levels**

Looking at Figure 6-16 more closely, one notices that the width of the base-emitter junction is relatively narrow. This is because of the extremely high level of doping in the emitter vs the very low value of doping in the base. High values of doping (on either side) make the SCL narrow. The downside of making low values of doping is that the resistivity of the material increases. This makes for poorer high frequency performance because of high values of base resistance (or emitter resistance or collector resistance). The downside of high values of doping is a lower breakdown voltage. Most emitter-base junctions breakdown at around five volts or less because of the very high doping concentrations in the emitter. The collector is usually doped at a moderate level as a compromise between these two problems. We usually want to have more collector-base breakdown voltage, but we also want a fairly low value for collector resistance.

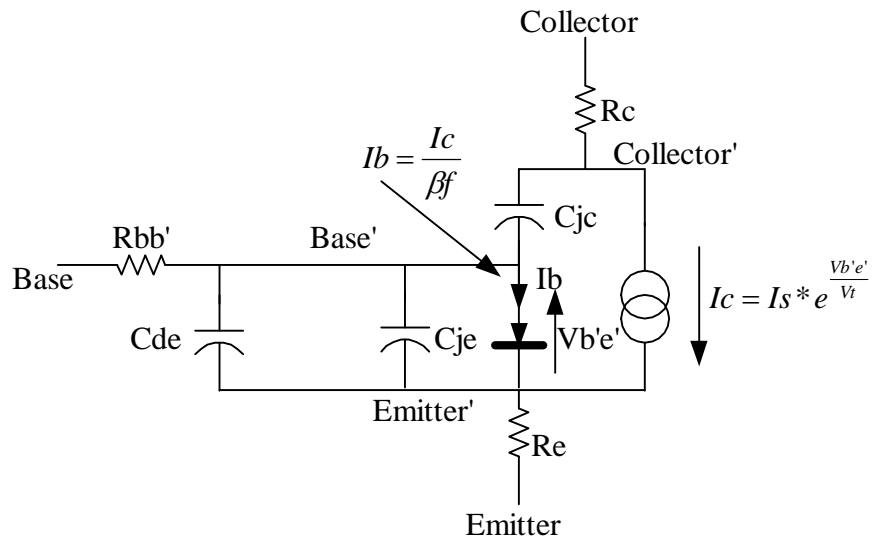
## Base Width

The distance between the emitter-base junction and the collector-base junction is called the base width. It is desirable to make this distance as short as possible. It takes time for the minority carriers to diffuse from the emitter-base junction to the collector-base junction, (a parameter called  $\tau_f$ ). The longer this time, the higher the probability they will run into a majority carrier and recombine. When this happens, base current is generated lowering  $\beta_f$ . This is very undesirable.

## Base Current

Base current is never intended. Contrary to popular opinion, a BJT is not a current controlled device, it is a voltage controlled device. If it were possible, we would always like  $\beta_f$  to be infinity. So what are the components of base current? Base current,  $I_b$ , comes from three primary sources. The first is minority carriers injected into the emitter from the base. We have already talked about this source of base current. It is managed by making the emitter doping very high relative to the base doping. The second source is due to the leakage current ( $I_{sc}$ ) of the base-collector junction. This is usually very small. The third source is due to minority carriers in the base recombining with the majority carriers, rather than making it across the base region to be swept into the collector region. Recombination in the base occurs just due to random chance, i.e. the longer the carrier spend in the base region, the higher the probability it will encounter a majority carrier to recombine with. It is for this reason that base regions are made as narrow as possible. A narrow base region means less recombination means a higher  $\beta_f$ . Recombination in the base region also occurs because of crystal defects (which occur principally at the surface of the semiconductor, but can occur anywhere). These crystal defects form "traps" that trap any free carrier that happens to get near them. This is the principle cause of the decrease in  $\beta_f$  at very low currents.

## Large Signal Model For The BJT Operated In The Normal Region Of Operation



**Figure 6-18: Large Signal Model Of Transistor Biased In The Normal Region Of Operation.**

**Equation 6-22** 
$$C_{jc} = \frac{C_{jc0}}{\left(1 - \frac{V_{bc}}{V_{jc}}\right)^{m_c}}$$

And

**Equation 6-23** 
$$C_{je} = \frac{C_{je0}}{\left(1 - \frac{V_{be}}{V_{je}}\right)^{m_e}}$$

### Junction Capacitance

$C_{jc}$  and  $C_{je}$  are the junction capacitance's. These are the same as the junction capacitor that was in the diode, except now there are two.  $V_j$  are the junction voltages,  $m_e$  and  $m_c$  are the grading coefficients (between .5 and .333 in value), and  $C_{jc0}$  and  $C_{je0}$  are the zero bias junction capacitance's.

### Diffusion Capacitance

Notice that in Figure 6-16 there are a considerable number of free minority carriers in the base region. This charge is basically "stored" in the base region for the time it takes to transit the base. This makes the base look like a capacitor. For the normal region of operation, we model this capacitor with a parameter called the forward transit time,  $\tau_f$ . The forward transit time is the time required for the minority carriers to diffuse from the emitter to the collector. (If the device is operated in the reverse mode, there is a parameter called the reverse transit time. The forward transit time is not equal to the reverse transit time.)

In Spice,  $\tau_f$  is usually called TF, and the reverse transit time is called TR.  $C_{de}$  is the diffusion capacitance and is equal to:

**Equation 6-24**  $C_{de} = \tau_f * g_m$

Where

**Equation 6-25**  $g_m = \frac{I_c(\text{BiasPoint})}{V_t}$

$I_s$  is the diode saturation current and is measured in the same way as the as it was for the single diode.  $R_c$ ,  $R_e$ , and  $R_{bb'}$  are the bulk resistances in the collector, emitter, and the base. These can be very difficult to measure, especially  $R_{bb'}$ . Also  $R_{bb'}$  can have a large effect on circuit performance.

Notice that  $\beta_f$  doesn't enter into these equations. In an ideal device,  $\beta_f$  would equal  $\infty$ , and the models would work just fine. Unfortunately, there is always some base current flow, and  $\beta_f$  does model this flow based on a given collector current. But it is not the underlying physical model that we use to model the BJT. Some small circuit models do use a current controlled current source model. In this case a small signal  $\beta_f$  is used. Please note that the small signal  $\beta_f$  is not the same as the large signal  $\beta_f$ !

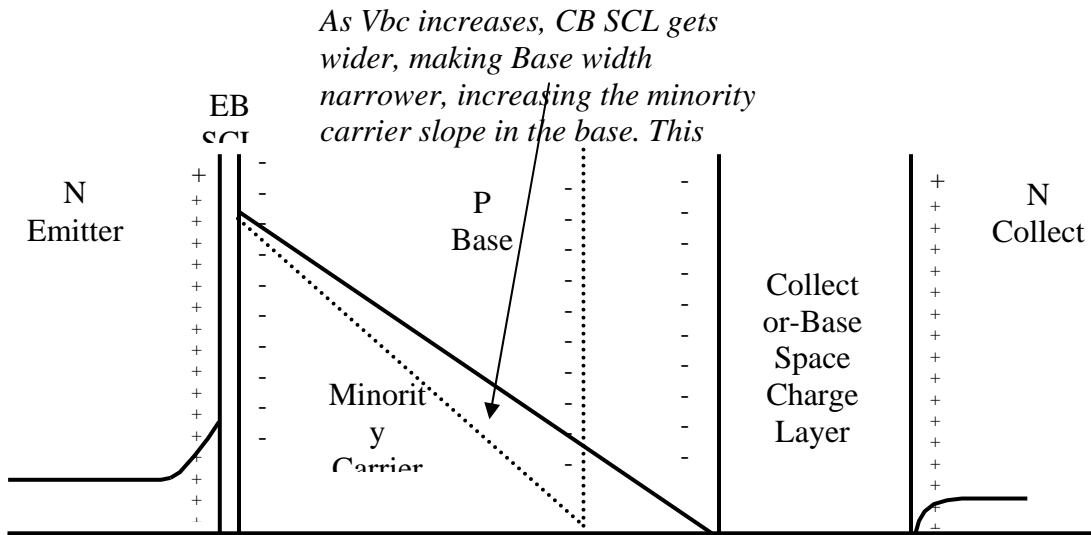
The small signal  $\beta_f$  is the derivative of the collector current with respect to the base current;  $\frac{dI_c}{dI_b}$ . This is not equal

to  $\frac{I_c}{I_b}$ !

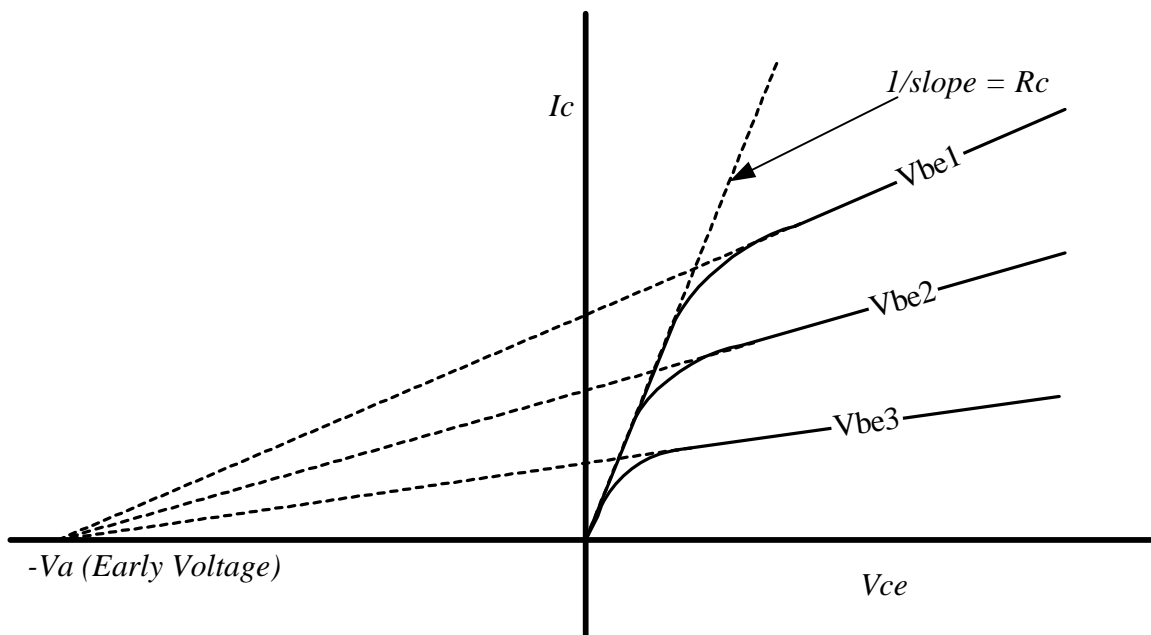
The complete Gummel-Poon model takes into account all four quadrants for the transistor. This generates additional parameters such as a reverse transit time and a reverse  $\beta$ .

## Basewidth Modulation

If the collector-base voltage ( $V_{bc}$ ) is increased (made more reversed biased), the SCL of the collector-base junction gets wider decreasing the width of the base region. The SCL will also punch more into the collector region, but because the collector is doped heavier than the base, most of the increase in width occurs on the base side. This effect is shown in Figure 6-19. This has the effect of increasing the slope of the minority carrier distribution in the base, which, in turn, causes  $I_s$  and  $\beta_f$  to increase, which then increases the collector current. This is why there is a finite output resistance for a transistor. Also, since the base width is more narrow, the transit time  $\tau_f$  is decreased making the device faster. These effects are referred to as the Early effect, and it is modeled by a parameter called the Early Voltage,  $V_a$ . A way to measure Early Voltage is shown Figure 6-20.



**Figure 6-19: A Minority Carrier Diagram For An NPN Transistor Biased In Normal Region Of Operation.**



**Figure 6-20: Measuring Early Voltage and Collector Resistance Using a Curve Tracer**

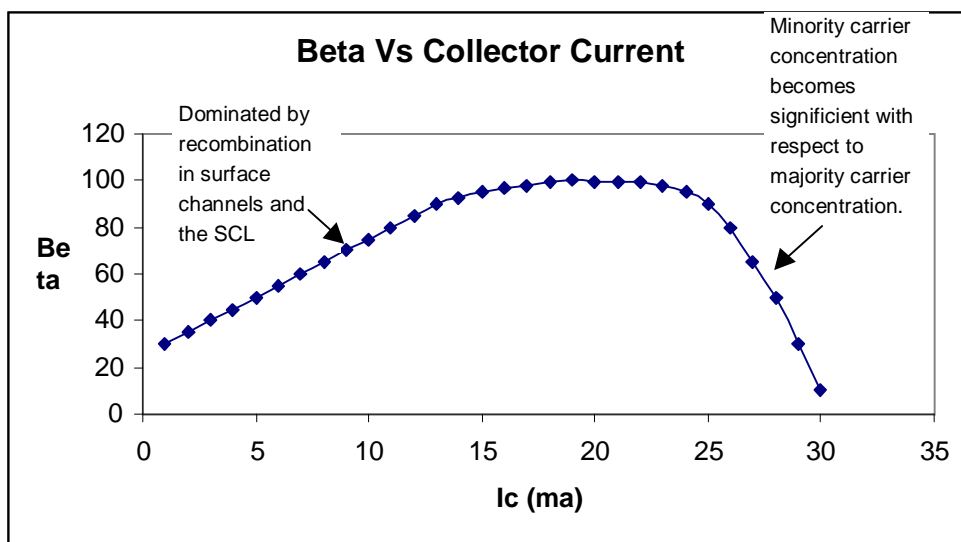
The effects on  $I_s$ ,  $\beta_f$ , and  $\tau_f$  are shown in the following equations. Notice that  $I_s$  and  $\beta_f$  both increase with increasing reverse  $V_{bc}$ , while  $\tau_f$  decreases.

**Equation 6-26** 
$$I_s(V_{bc}) = \frac{I_s(0)}{\left(1 - \frac{V_{bc}}{V_a}\right)}$$

**Equation 6-27** 
$$\beta_f(V_{bc}) = \frac{\beta_f(0)}{\left(1 - \left|\frac{V_{bc}}{V_a}\right|\right)}$$

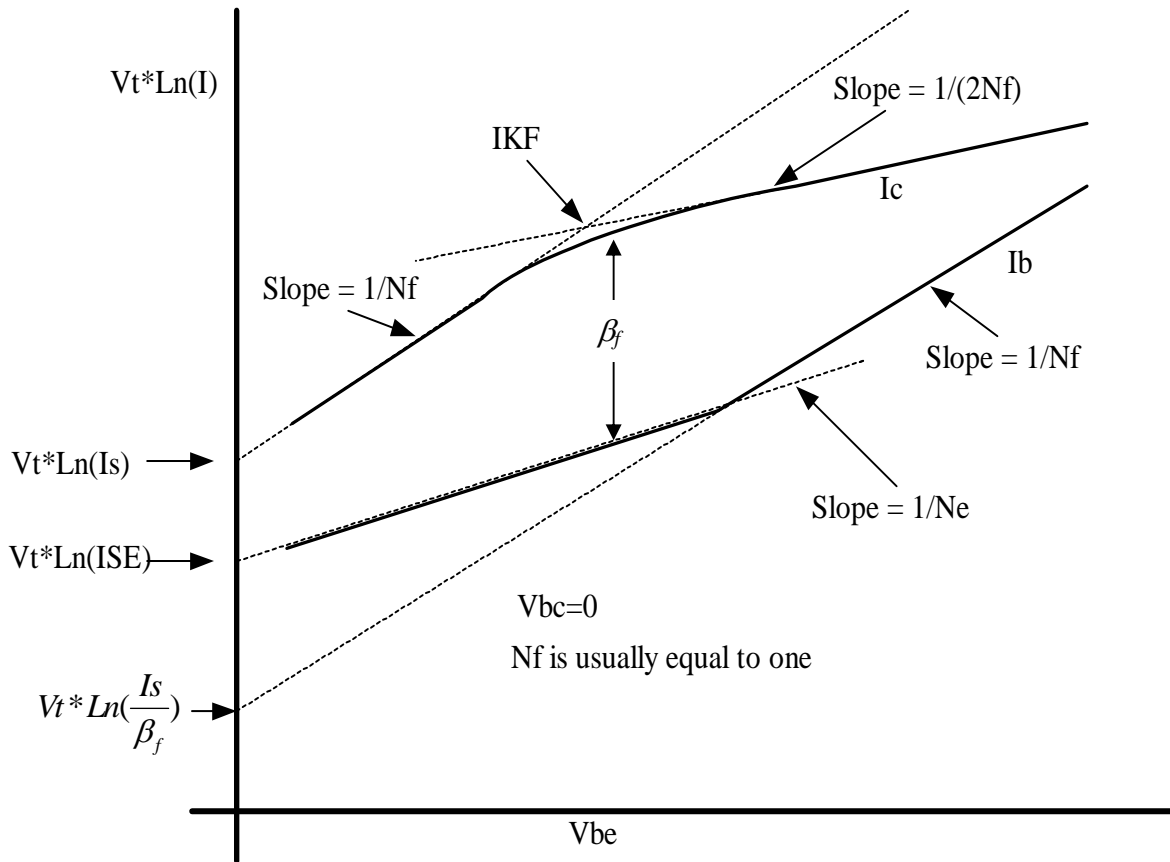
**Equation 6-28** 
$$\tau_f(V_{bc}) = \tau_f(0) * \left(1 - \left|\frac{V_{bc}}{V_a}\right|\right)^2$$

### $\beta_f$ variation with collector current?



**Figure 6-21: Beta As A Function Of Collector Current**

As shown in Figure 6-21,  $\beta_f$  varies quite a lot with collector current. It has an optimal point where it is a maximum, The reason it decreases from this optimum point depend on whether it is operated at very low currents or at very high currents. A better way to look at this effect is shown in Figure 6-23 and is also the technique used to measure the required SPICE parameters.

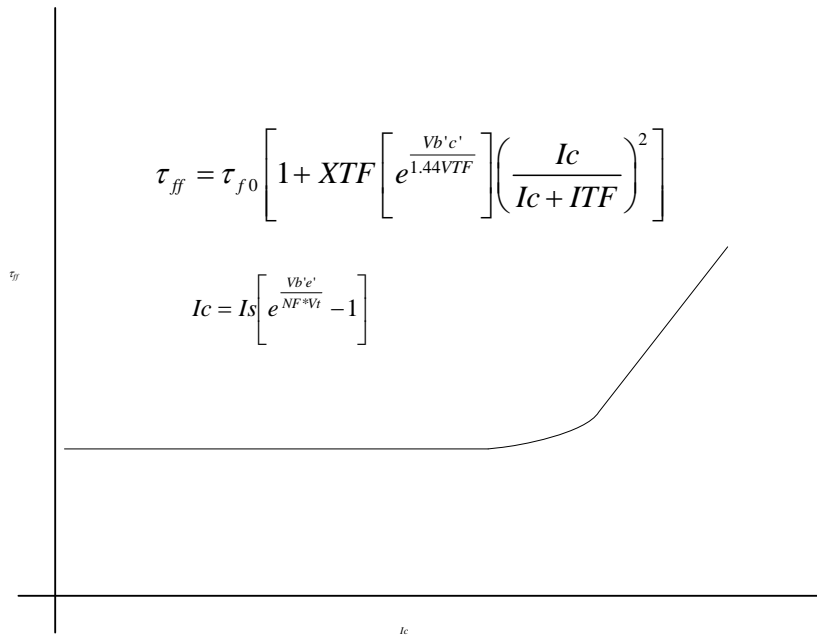


**Figure 6-22:  $I_c$  And  $I_b$  Vs  $V_{be}$  On A Log Plot**

In Figure 6-22  $I_{SE}$ ,  $I_{SC}$ ,  $N_e$ , and  $N_C$  determine the low current rolloff in  $\beta_f$  in both forward and reverse directions.

$I_{KF}$  and  $I_{KR}$  are the forward and reverse knee currents, and determine the high current rolloff with  $I_c$ . From this plot, it is apparent that the rolloff in  $\beta_f$  at low current is due to a non-linear increase in base current due to surface recombination and trapping in the base region. At high currents it's just the opposite as the problem exists with the collector current. At high currents, two new effects become dominant because of the extremely high level of minority carriers in the base region. Basically, their electric charge begins to have an effect, whereas this was negligible at low currents. The first of these effects is called Base Pushout and is an effect where the base region is "pushed" into the collector by the extremely high density of minority carriers thereby widening the base region. Any increase in base width causes an increase in recombination which lowers  $\beta_f$ . The second effect is called crowding. This effect causes minority carriers to extend out into silicon beyond the area of the emitter base junction due to the electric field generated by the high density of minority carriers in the base region. It should be noted that these effects also increase  $\tau_f$  (and decrease the high frequency performance) because of the effective lengthening of the basewidth. This is shown in Figure 6-23.

$\tau_f$  as a function of  $I_c$ .



**Figure 6-23: Increase In  $\tau_f$  With Increase In IC**

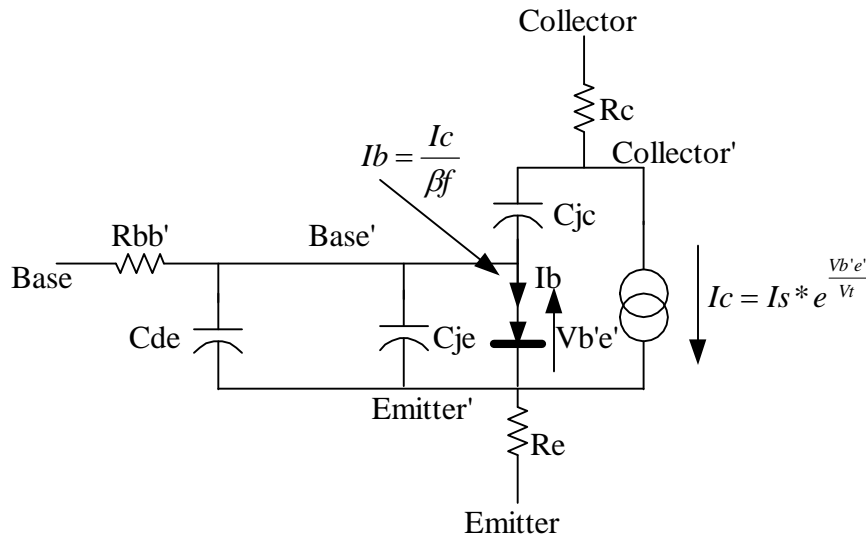
**Equation 6-29** 
$$\tau_f = \tau_{f0} \left[ 1 + XTF \left[ e^{\frac{Vb'c'}{1.44VTF}} \right] \left( \frac{I_f}{I_f + ITF} \right)^2 \right]$$

**Equation 6-30** 
$$I_f = I_s \left[ e^{\frac{Vb'e'}{NF*Vt}} - 1 \right]$$

***XTF*** Coefficient for bias dependence of  $\tau_f$ .  
***ITF*** High current parameter for effect on  $\tau_f$ .  
***VTF*** Voltage describing  $V_{bc}$  dependence of  $\tau_f$ .  
***NF*** Forward current emission coefficient.

Notice that  $\tau_f$  is fairly constant at low currents. It begins to increase at high currents because of crowding and base push out factors. This slows the device down at high currents.

## The Linear Hybrid Pi model for the BJT.



**Figure 6-24 Repeating The Large Signal Model**

The Hybrid Pi linear model is derived from the large signal model by assuming the transistor is biased at some “Q” point ( $I_C$  is a certain number of ma., and  $V_{CE}$  is a certain number of volts), then linearizing the diodes by taking the appropriate derivatives and using the results to give linear gain constants. In Figure 6-25, assume that transistor is operated at a bias point equal to  $V_{ce_q}$ ,  $I_{c_q}$ . Then the gain is linearized as:

$$\text{Equation 6-31 } gm = \left. \frac{dI_C}{dv_{be}} \right|_{I_C=I_{Cq}} = \frac{I_S * e^{\frac{V_{be}}{V_t}}}{V_t} = \left. \frac{I_C}{V_t} \right|_{I_C=I_{Cq}} = \frac{I_{C_q}}{V_t}$$

Where  $V_t = \frac{KT}{q} = 26\text{mv}$  at room temperature.

The base emitter diode is replaced by a resistor,  $R_\pi$  whose value is:

$$\text{Equation 6-32 } R_\pi = \frac{\beta_f}{gm}$$

This equation is justified because the DC current through the base emitter junction is  $\frac{I_{C_q}}{\beta_f}$ . Earlier, we proved that the linearized diode looked like a resistor with a conductance value of:

$$G_p = \left. \frac{dI_d}{dv_d} \right|_{I_d=I_b} = \frac{I_S * e^{\frac{v_d}{V_t}}}{V_t} = \left. \frac{I_d}{V_t} \right|_{I_d=I_b} = \frac{I_b}{V_t}$$

Knowing that  $I_b = \frac{I_{C_q}}{\beta_f}$  allows for the derivation of Equation 17.

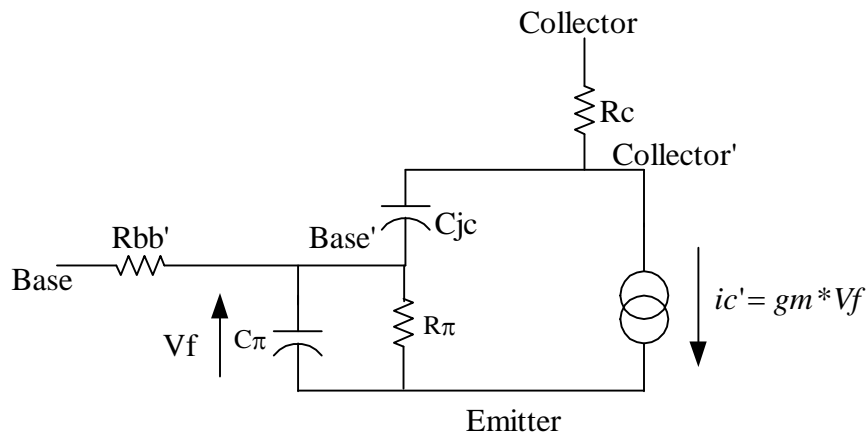
$$\text{Equation 6-33 } C_\pi = C_{de} + C_{je} \text{ operated at the bias point}$$

**Equation 6-34**  $C_{de} = gm * \tau_f$  Where  $\tau_f$  is the forward transit time.

**Equation 6-35**  $C_{je} = \frac{C_{je0}}{\left(1 - \frac{V_{be}}{V_{je}}\right)^{me}}$  Defined earlier

**Equation 6-36**  $C_{jc} = \frac{C_{jc0}}{\left(1 - \frac{V_{bc}}{V_{jc}}\right)^{mc}}$  Defined earlier

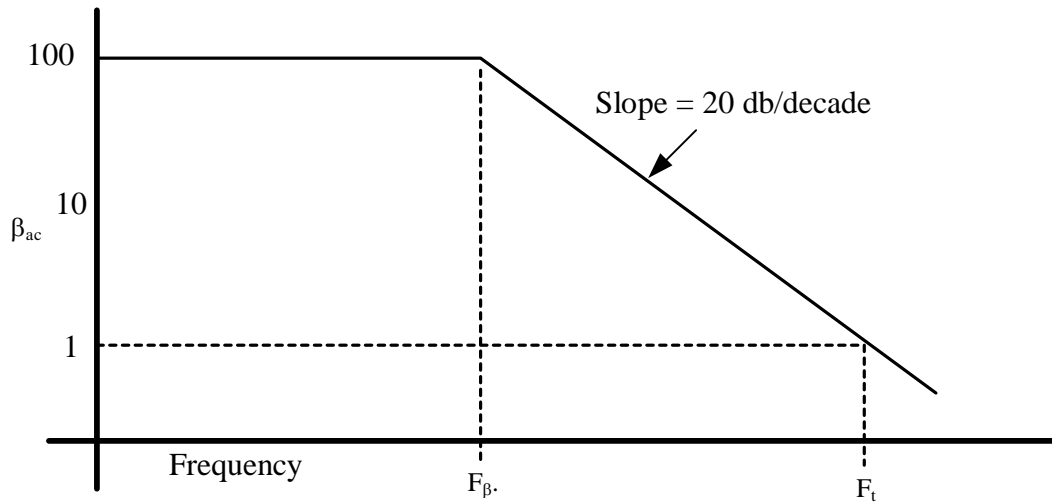
Putting all of this together gives the following result:



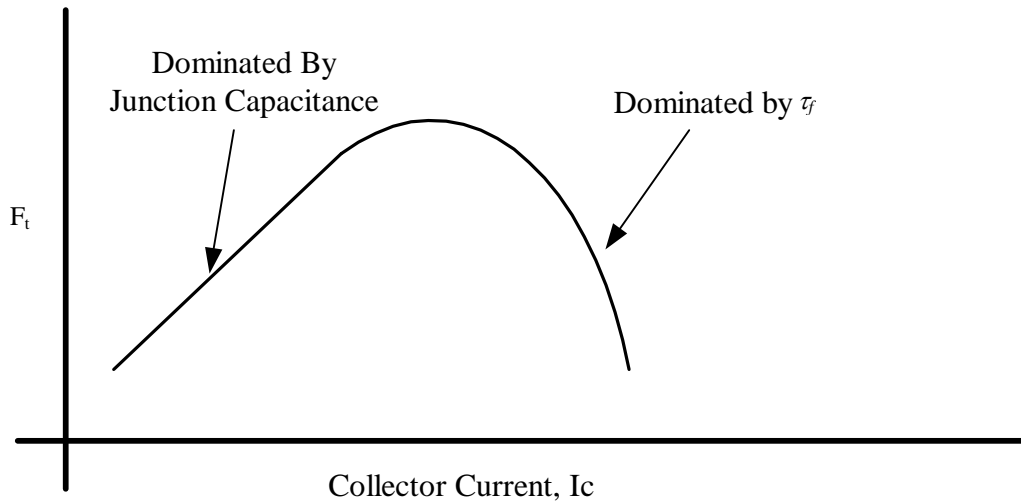
**Figure 6-25: The Linear Hybrid Pi Model For A Bipolar Junction Transistor.**

### $F_t$

The high frequency performance of a transistor is frequently modeled by the parameter  $F_t$ . This parameter specifies the frequency at which the small signal (ac) current gain ( $\beta_{ac}$ ) for the device is unity for a given bias point. It assumes that for a given bias point,  $\beta_{ac}$  has a single pole defined as  $F_{\beta}$ .  $F_t$  is not itself a Spice parameter, but instead is a function of Spice parameters. The Hybrid  $\pi$  model can be used to derive the equation relating the Spice parameters to the  $F_t$ .



**Figure 6-26: Bode Plot Of  $\beta_{ac}$**



**Figure 6-27: A Plot Of  $F_t$  Vs  $I_c$**

### **$F_t$ vs Collector Current**

A plot  $F_t$  vs collector current is shown in Figure 6-27. Notice that  $F_t$  is a strong function of  $I_c$ . At low collector currents,  $F_t$  is dominated by the junction capacitors while at high currents, it is dominated by  $\tau_f$ . At high currents,  $F_t$  decreases with current because  $\tau_f$  increases with current. A derivation of  $F_t$  as a function of Spice parameters will now be done using the Hybrid  $\Pi$  model, and the circuit shown in Figure 6-29.

## Deriving $F_t$ In Terms Of Spice Parameters Using The Hybrid Pi Model.



**Figure 6-28: The Circuit Used To Determine  $F_t$  As A Function Of Spice Parameters.**

In the above circuit, the frequency at which  $\frac{ic}{isig} = 1$  is defined to be  $F_t$ .

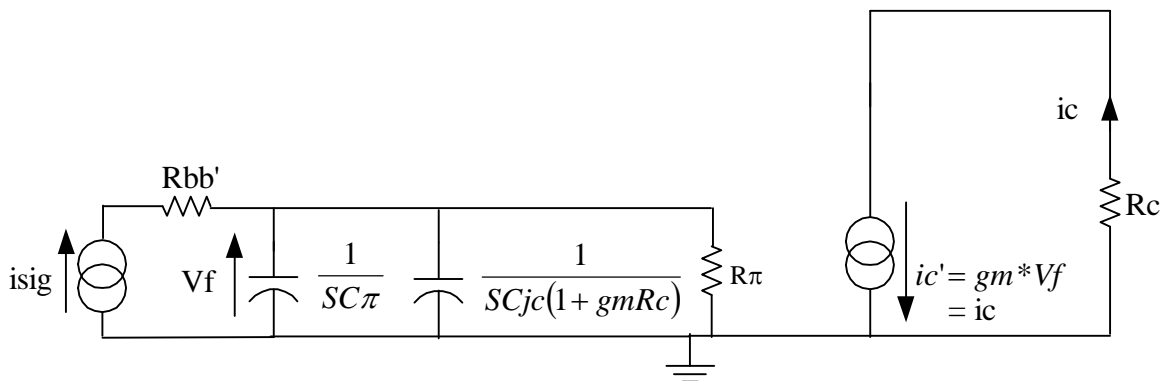
To simplify the problem, we will make use of Miller's theorem which states that the input impedance to a voltage amplifier with feedback is equal to the feedback impedance divided by 1 minus the voltage gain ( $Z_{in} = \frac{Z_f}{1 - A_v}$ ). Assuming that the collector resistance,  $R_c$ , is small, then the voltage gain of the above circuit is approximately:

**Equation 6-37**  $A_v = -gmR_c$

When applied to the above circuit, one would obtain an input impedance due to  $C_{jc}$  of:

**Equation 6-38**  $Z_{in} = \frac{1}{sC_{jc}(1 + gmR_c)}$

Redrawing the circuit with this new assumption yields the following circuit diagram.



**Figure 6-29: Equivalent Circuit Using Miller's Theorem For Computing  $F_t$**

Figure 6-29 is a simplified circuit using the assumption that that Miller's theorem can be applied using a voltage gain equal to  $-gm \cdot R_c$ . As long a  $R_c$  is small, this is a good assumption.

The current gain for this circuit becomes:

$$\text{Equation 6-39 } A_i = \frac{ic}{isig} = gm * Zin$$

Where  $Zin$  is the impedance of the parallel combination of  $C\pi$ ,  $R\pi$ , and  $Cjc(1+gm)$ .

This yields:

$$\text{Equation 6-40 } A_i = \frac{gmR\pi}{sR\pi[C\pi + Cjc(1 + gmRc)] + 1}$$

$F_t$  is the frequency at which  $A_i$  becomes unity. At this frequency, it is a good assumption that:

$$\text{Equation 6-41 } sR\pi (C\pi + Cjc(1 + gmRc)) \geq 1$$

This allows us to set:

$$\text{Equation 6-42 } sR\pi (C\pi + Cjc(1 + gmRc)) = gmR\pi$$

Next substitute  $j\omega$  for  $S$  and solve for the  $\omega$ .

$$\text{Equation 6-43 } \omega_t = \frac{gm}{C\pi + Cjc(1 + gmRc)}$$

$$\text{Equation 6-44 } \omega_t = 2\pi f_t$$

Substituting

$$\text{Equation 6-45 } f_t = \frac{1}{2\pi \left( \frac{C\pi}{gm} + \frac{Cjc}{gm} + CjcRc \right)}$$

Remember

$$\text{Equation 6-46 } C\pi = Cde + Cjc \text{ and } Cde = gm\tau_f$$

Substituting

$$\text{Equation 6-47 } f_t = \frac{1}{2\pi \left( \tau_f + \frac{Cje}{gm} + \frac{Cjc}{gm} + CjcRc \right)}$$

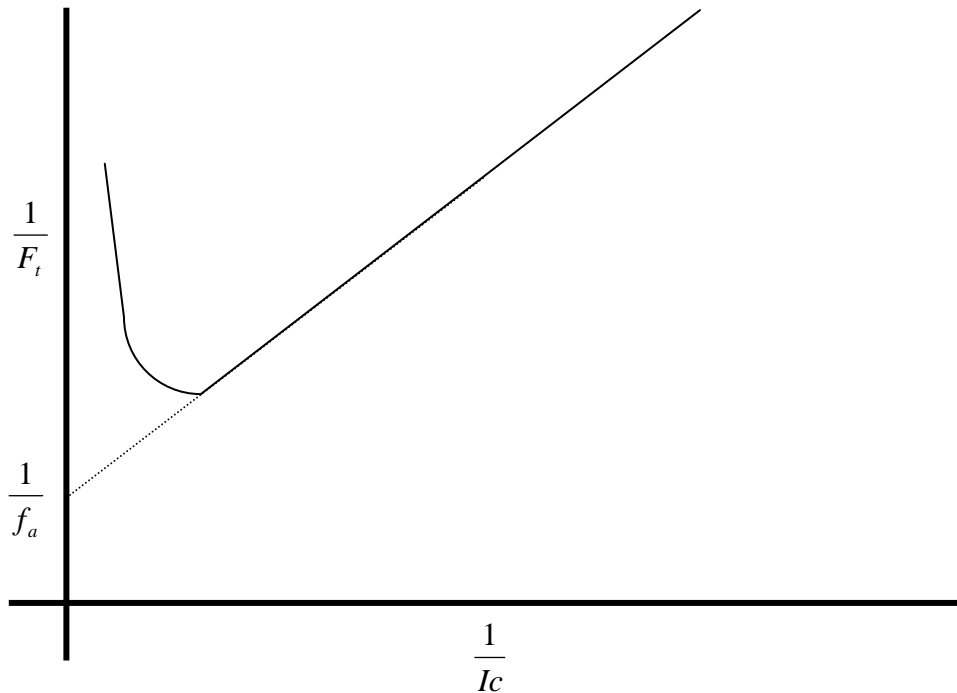
Or

$$\text{Equation 6-48 } f_t = \frac{1}{2\pi \left( \tau_f + \frac{Cje}{gm} + Cjc \left( \frac{1}{gm} + Rc \right) \right)}$$

Equation 6-48 can be solved for  $\tau_f$  in terms of  $F_t$  giving us Equation 6-49

$$\text{Equation 6-49 } \tau_f = \frac{1}{2\pi F_t} - \frac{1}{gm} [Cje + Cjc(1 + gmRc)]$$

$F_t$  can be used to determine  $\tau_f$  as shown in Figure 6-30



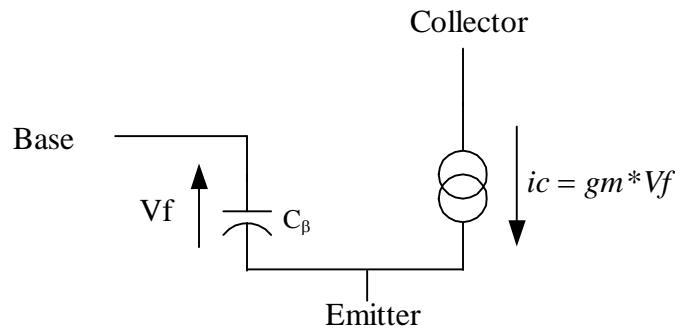
**Figure 6-30: A Typical Plot Of  $(1/F_t)$  As A Function Of  $(1/I_c)$**

The intercept  $(1/f_a)$  in Figure 6-30 of the extrapolated straight line at  $1/I_c = 0$  is related to  $\tau_f$  by:

**Equation 6-50** 
$$\tau_f = \frac{1}{2\pi f_a} - C_{jc}(V_{b'c'})r_c$$

### Simplifying The Hybrid Pi Model

While the hybrid pi model is very useful, it is too complicated if one is only interested in low frequency or high frequency effects. The hybrid pi is valid across the entire range of frequencies, and this is just too much complexity for easy use. It is easy to simply the model for low frequency use; simply remove the capacitors. For high frequency modeling, things aren't quite that intuitive. The current hybrid pi model has a pole at  $F_\beta$  as shown in Figure 6-26. To make an effective model at high frequencies, assume that  $\beta$  is infinity so that there is no breakpoint. This forces a model of a pure single pole frequency dependent source. This is accomplished by removing the resistor,  $R_\pi$ , for the circuit shown in Figure 6-29.

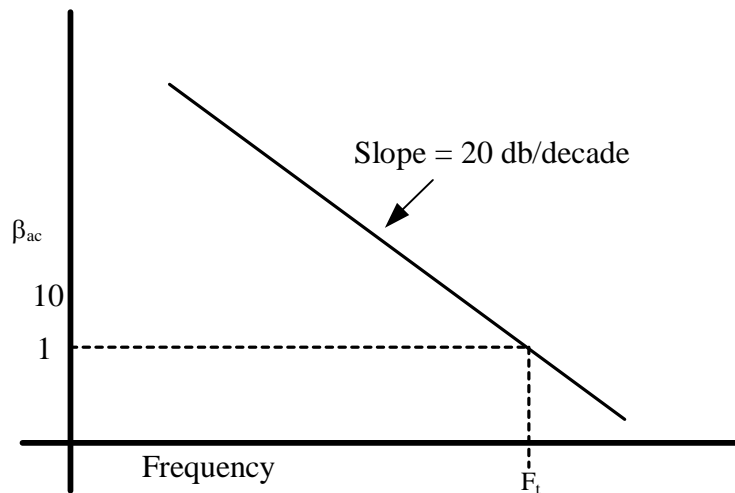


**Figure 6-31: A High Frequency Model For The BJT Derived From The Hybrid Pi Model**

Figure 6-31 shows the high frequency model that results from removing  $R\pi$ . In this model:

**Equation 6-51**  $C_{\beta} = C_{je} + gm * \tau_f + C_{jc}(1 + gmRc)$

A bode plot for this model is shown in Figure 6-32



**Figure 6-32: A Bode Plot Of The High Frequency Model.**

This model is valid any time the frequency of interest lies above  $F_{\beta}$ . Even if the signal range is well below this frequency, this model can still be valid for computing things like stability. A 100 MHz oscillator doesn't care if it was only supposed to amplify frequencies up to 10KHz!

One obtains:

**Equation 6-52**  $C_{\beta} = \frac{gm}{2\pi f_t}$

This is a very interesting and useful result as it relates the high frequency performance to a single element that is a function of a single fundamental transistor parameter and the transistor operating point.

### **T<sub>t</sub>, A Very Useful Parameter.**

Define the following parameter  $T_t$  to be:

**Equation 6-53**  $T_t = \frac{1}{2\pi f_t}$

Then

$$\text{Equation 6-54 } C_{\beta} = gm * T_t$$

It is important to note that  $T_t$  is not the transit time of the device. In the Spice world  $T_t$  is defined to be the transit time in a diode. This is not how this parameter is used in this book.

Remember

$$\text{Equation 6-55 } gm = \frac{Ic}{V_t}, \text{ Ic being the bias point}$$

The base current is

$$\text{Equation 6-56 } i_b = V_f * S * C_{\beta}$$

Substituting for  $C_{\beta}$

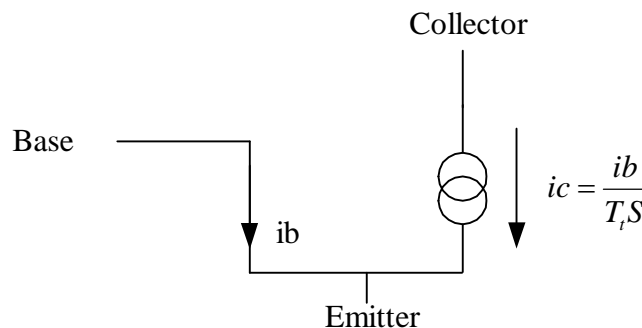
$$\text{Equation 6-57 } i_b = V_f * S * gm * T_t$$

## A Current Gain Version Of The High Frequency Model

The current gain is defined to be

$$\text{Equation 6-58 } A_i = \frac{Ic}{Ib} = \frac{gm * V_f}{V_f * S * gm * T_t} = \frac{1}{T_t * S}$$

This equation generates the next model.



**Figure 6-33: A Very Simple High Frequency Model For The BJT**

The model of Figure 6-33 is extremely useful for calculating high frequency effects in transistors.

## Summary

These two models both work very well at high frequencies, and are equivalent to each other. They are used to explain many effects of high frequency circuits, including oscillation. Neither work at DC ( $\beta_f$  is assumed to be  $\infty$  at DC). Both include the effects of  $C_{jc}$  and  $R_c$  (as long as the collector is at AC ground). If additional resistance is added to the collector (so that the collector is not at AC ground), then  $C_{\pi}$  must be increased to compensate for this additional resistance. Finally,  $R_{bb'}$  is not taken into account. However, the driving source usually has a finite output resistance, and it is frequently possible to simply

add  $R_{bb'}$  to this external resistor. The choice of which one to use is somewhat dependent on the actual circuit; for the BJT the current drive version is usually a good choice, but not always!

I have gone to great lengths in this chapter to show that these models are based in physics; in particular, the physics of the bipolar junction transistor. Interestingly, if the starting point had been a FET, the end result would have been the same (assuming that a single pole rolloff adequately describes the device). One might think that the capacitor version would be a better choice for a FET, but in reality, it really does not matter much as these models are limited to high frequency performance only and both model that effect. Again, there are times that the capacitor version is a better choice, if for no other reason than equation simplification, but both models have their place even in the FET world.

In fact, these models can be used to model the high frequency performance of almost any device assuming it has a single pole high frequency rolloff. In chapter 6, I use this a variant of this model to show the high frequency performance of op-amps (where high frequency is definitely a relative term).

In future chapters, these models will be used to predict high frequency performance of particular circuits. Additional circuit elements will be added as necessary.

# **Chapter 1**

## **Appendix B**

### **Scattering Parameters & Smith Charts**