Chapter 5
THE BIPOLAR JUNCTION TRANSISTOR (BJT)

5.1 The Transistor Action

The bipolar junction transistor (BJT) was invented in 1948 by Bardeen, Brattain and Shockley at the Bell Laboratories. One year later, Shockley presented the basic theory of operation of the device. Bipolar transistors became widely used in the 1950s and their technology progressed quickly until it became well established by the end of the decade.

The bipolar transistor is an active three terminal device composed of 3 layers of a semiconducting material (silicon). Fig. (5.1) shows a simplified picture of the two basic types of bipolar junction transistors, npn and pnp, along with their symbols when used as circuit elements.

The three regions are referred to as emitter, base and collector. Both types of carriers play an important role in this device; this is why the adjective "bipolar" is used to describe the junction transistor. The arrow on the emitter lead indicates the direction of current flow when the emitter- base junction is forward biased. The emitter current is denoted by I_E, the collector current I_C and the base current I_B, The mechanism of operation i.s essentially the same for both structures.

The bipolar transistor is mainly used in amplification and switching applications. To understand the transistor action, we must first visualize the transistor as two coupled pn junctions put back to back with very thin base width to allow coupling between the two junctions. For amplification to take place, the emitter-base junction should be forward biased, and the collector-base junction reverse biased. This is called the active mode of operation (Fig. 5.2). Consider pnp transistor. The emitter serves as a reservoir of holes, which are injected into the base region to constitute an excess minority carrier concentration. Now, if the base region is sufficiency narrow, most of the holes injected into the base will be able to reach the collector - base junction and will be swept by the electric field which causes the flow of holes towards the p collector. Therefore, despite the collector base junction being reverse biased, a large current flows through it. This is due to the presence of the forward biased emitter -base junction, which supplies plenty of minority carriers to the base region. Therefore, if the emitter base voltage V_{EB} changes by ΔV_{EB}, a small variation occurs in the base current. But this small variation ΔI_B will cause a large change in the emitter current equal to ΔI_E. The variation in I_E is expected to be quite large due to the forward bias of the emitter base junction. If the recombination of holes in the base is neglected. The variation of the hole emitter current ΔI_E will be transferred as it is into the collector current. Thus, a current gain $\Delta I_C / \Delta I_B$ results. A considerable voltage gain can, thus, be obtained if this large variation in the collector current is multiplied by a sufficiently large load resistance.

Fig. (5.3a) shows the potential distribution across a pnp transistor before biasing and Fig. (5.3b) with the emitter junction forward biased and the collector junction reverse biased. Notice that in Fig. (5.3a), the built-in potential ϕ_o is assumed the same for both junctions. In reality, this is not the case since the doping level of the emitter is usually much higher than that of the collector. Fig. (5.3b) shows that the potential barrier of the emitter base junction is lowered by the magnitude of the applied forward bias $|\Delta V_{EB}|$. This potential barrier lowering allows holes to be injected into the base, where they diffuse towards the collector. Once the holes reach the collector-base junction, they are swept across the junction by the action of the electric field in the depletion layer. Note that the potential barrier across this junction is increased by the magnitude of the reverse bias $|\Delta V_{CB}|$, which causes a significant increase in the electric field. It should be mentioned also that the extension of the depletion layer inside the base region increases with increasing reverse bias.

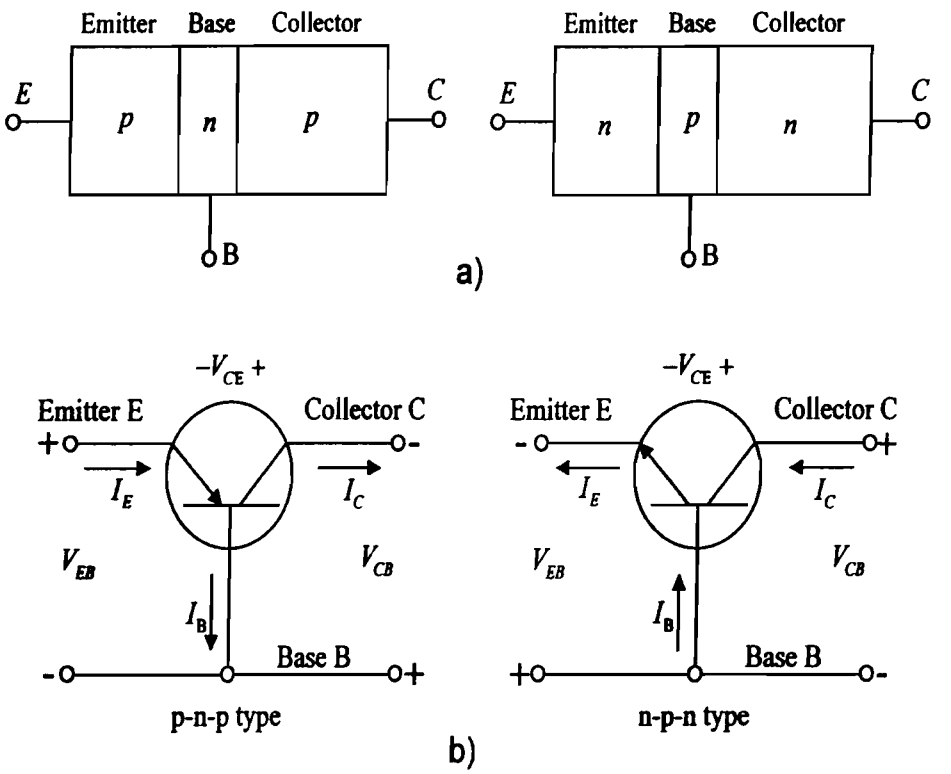


Fig. (5.1) BJT
a) schematic diagram for pnp and npn BJTs
b) circuit symbols of the two types of bipolar transistors

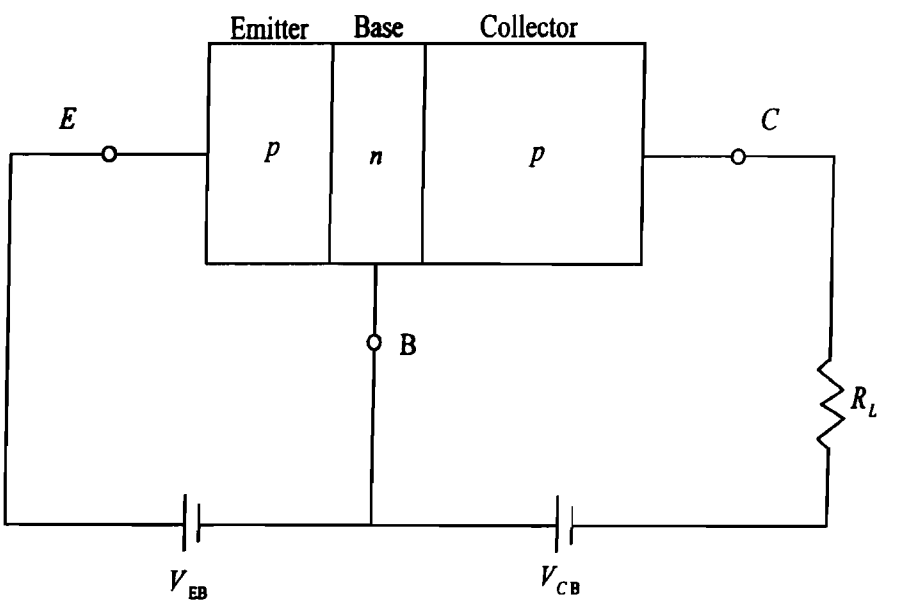


Fig. (5.2) A pnp transistor biased in the active mode of operation

5.2 Current Components and Current Gains

Let us now examine more closely the different current components which flow inside the bipolar transistor and the ratios between them. Uptil now, we have implicitly assumed that the emitter current consists only of holes injected from the emitter into the base. In fact, this component (which we will designate I_{pE}) is not the only one that flows across the forward biased emitter-base junction. Electrons injected from the base into the emitter constitute the second, component of the emitter current (which will be denoted by I_{nE}) as seen in Fig. (5.4). Therefore, the total emitter current I_E is given by

$$I_E = I_{nE} + I_{pE} \tag{5-1}$$

Naturally, bipolar transistor designers would like to make I_{pE} much greater than I_{nE} since electrons injected from the base into the emitter (i.e. I_{nE}) do not contribute to total collector current. For this reason, the emitter doping in commercial transistors is much higher than that of the base. Therefore, increasing the doping of the emitter with respect to that of the base increases the emitter efficiency γ, which is defined as the ratio of the hole emitter current to the total emitter current:

$$\gamma = \frac{I_{pE}}{I_{nE} + I_{pE}} = \frac{I_{pE}}{I_E} \tag{5-2}$$

Let us now follow the holes injected into the base in their course towards the collector. Some of these holes (constituting I_{pE}) will recombine with electrons in the base region and will not reach the collector.

Therefore, the hole collector current I_{pC} is lower than I_{pE} by the amount of the bulk recombination current in the base region. This recombination process increases with increasing base width. For this reason, the base width of a bipolar transistor should be kepi less than the diffusion length of minority carriers in the base. This leads to the definition of the base transport factor B_T, which is the ratio of the hole current reaching the collector junction I_{pC} to the hole current at the emitter junction I_{pE}:

$$B_T = \frac{I_{pC}}{I_{pE}} \tag{5-3}$$

It is evident that the product γB_T gives the ratio of the hole collector current to the total emitter current. The product is called the current transfer ratio and is denoted by α

$$\alpha = \gamma B_T = \frac{I_{pC}}{I_E} \tag{5-4}$$

The factor α should be very close to unity for a good bipolar transistor. This implies that the emitter current should consist mainly of holes (i.e. $\gamma = 1$) and that all these holes traverse the base without recombination to reach the collector (i.e. $B_T = 1$).

Taking a look at Fig. (5.4) we realize that I_{pC} is not the only component constituting the collector current I_C. Indeed, the collector - base junction, being reverse biased, passes a reverse saturation current I_{Co} in addition to I_{pC}, arriving from the emitter base junction. Therefore, the total collector current is given by:

$$I_C = I_{pC} + I_{Co} = \alpha I_E + I_{Co} \tag{5-5}$$

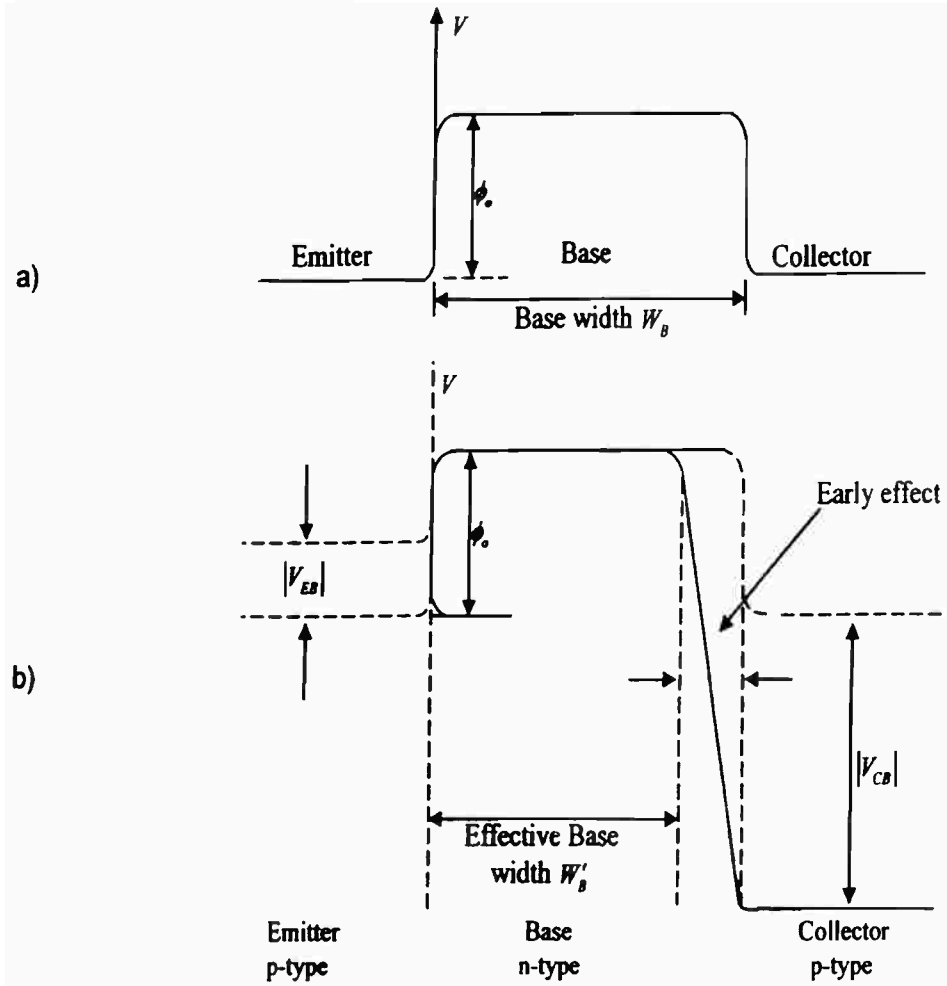


Fig. (5.3) Potential distribution across a pnp transistor
a) without bias b) with forward bias

From Fig. (5.4), the base current I_B consists of two components: The first is emitter base junction forward electron current I_{nE} and the second component is $(1-B_T)I_{pE}$, which represents the holes lost by recombination in the base region. Under steady state condition, an electron current equal to the hole recombination current is supplied by the base terminal. It is evident that the sum of the three terminal currents should be zero. Therefore, the base current is governed by:

$$I_E = I_C + I_B \qquad (5\text{-}6)$$

Making use of eqn. (5-5) and neglecting the reverse saturation current I_{C_o}, the ratio I_C / I_B is given by

$$\frac{I_C}{I_B} = \frac{\alpha}{1-\alpha} = \beta \qquad (5\text{-}7)$$

Eliminating I_E between eqns. (5-5) and (5-6) and using eqn. (5-7)

$$I_C = \beta I_B + (\beta + 1) I_{Co} \qquad (5\text{-}8)$$

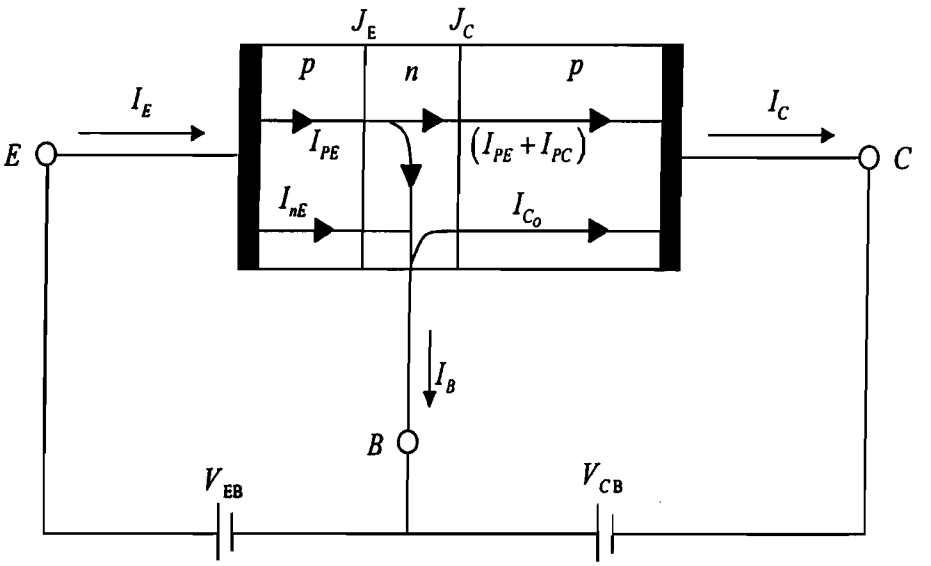


Fig. (5.4) A schematic diagram of a pnp transistor showing the different components flowing through it

Since α is only slightly less than unity, the factor β, which is defined as the base to collector dc current gain can be very large. It is this high value of β that gives rise to a large voltage variation in the output current due to a small variation in the base current, and is the cause of amplification in the transistor.

5.3 Current - Voltage Characteristics of an Ideal Transistor

To simplify the derivation of the I-V characteristics of a bipolar transistor, we are going to make the following assumptions:

1- The doping concentration is uniform in all three regions of the transistor.
2- The base current is carried only by diffusion.
3- The injection level into the base is low.
4- Neglect the electron current in the emitter by making the emitter heavily doped (p^+)

Consider a pnp configuration. From eqn. (4-71) and the discussions on the p-n junction theory, the emitter electron current just at the emitter - base junction i.e., at $x = 0$ (Fig. 5-5) is given by

$$I_{nE}(0) = \frac{|q| A D_n n_{E_o}}{L_E} (e^{V_E/V_{th}} - 1) \qquad (5\text{-}9)$$

where V_E is the voltage directly across the emitter junction with L_E and n_{E_o} being the minority carrier diffusion length and equilibrium concentration in the emitter region, respectively. As for the hole current injected into the base across the emitter-base junction, the situation is quite different. This current is not identical to that found from ordinary p-n junction theory, since the width of the base region is small and coupling occurs between the two junctions. In this case, we have to solve the continuity equation for minority carriers (holes) in the base region.

From eqns. (4-44) and (4-62)

$$I_{pE} = I_{p_n} = -|q| A D_p \frac{dp_n(x)}{dx} \tag{5-10}$$

Recalling eqn. (4-42)

$$D_p \frac{d^2 \Delta p_n(x)}{dx^2} - \frac{\Delta p_n(x)}{\tau_p} = 0 \tag{5-11}$$

The general solution for $p_n(x)$ is given by:

$$p_n(x) - p_{n_o} = K_1 e^{x/L_B} + K_2 e^{-x/L_B} \tag{5-12}$$

where K_1 and K_2 are constants to be determined by the boundary conditions, while L_B is the minority carrier diffusion length in the base.

At $x = 0$ and $x = W_B$ (W_B being the width of the base), the minority carrier concentration $p_n(0)$ and $p_n(W_B)$ are given from eqn. (4-56) by

$$p_n(0) = p_{n_o} e^{V_E/V_\phi} \tag{5-13}$$

$$p_n(W_B) = p_{n_o} e^{V_C/V_\phi} \tag{5-14}$$

where V_E is the emitter junction voltage and V_C is the collector junction voltage.

For the active mode of operation, the collector - base junction is highly reversed biased and $p_n(W_B)$ is nearly equal to zero. This is understandable since the holes arriving at the collector - base junction are swiftly swept by the electric field across the depletion layer into the collector. It is not difficult to obtain the exact values of K_1 and K_2, however, since the width of the base region is made very narrow with respect to L_B in order to minimize minority carrier recombination, we are going to simplify the general solution given by eqn. (5-12) to the following form:

$$p_n(x) - p_{n_o} = K_3 + K_4 x \tag{5-15}$$

Therefore, the excess minority carrier concentration varies linearly with distance inside the base region. Consequently, the diffusion current is independent of distance inside the base, which implies that recombination is negligible (i.e. $B_T = 1$). Therefore, the current $I_{pB}(0) = I_{pn}(0)$ is given by

$$I_{p_n}(0) = -|q| D_p A K_4 = \text{constant}, \quad o < x < W_B \tag{5-16}$$

Making use of the boundary conditions to determine K_4, the hole current $I_{p_n}(0)$ becomes

$$I_{p_n}(0) = -\frac{|q| A D_p p_{n_o}}{W_B} (e^{V_C/V_\phi} - e^{V_E/V_\phi}) \tag{5-17}$$

The total emitter current I_E is the sum of $I_{pn}(0)$ and $I_{nE}(0)$:

$$\begin{aligned} I_E &= I_{p_n}(0) + I_{nE}(0) \\ &= \frac{|q| A D_p p_{n_o}}{W_B} (e^{V_C/V_\phi} - e^{V_E/V_\phi}) + \frac{|q| A D_n A p_{n_o}}{L_E} (e^{V_E/V_\phi} - 1) \end{aligned} \tag{5-18}$$

It is convenient to rewrite the emitter current I_E in the following form

$$I_E = |q| A \left(\frac{D_p p_{n_o}}{W_B} + \frac{D_p p_{E_o}}{L_E} \right) (e^{V_E/V_\phi} - 1) - \frac{|q| A D_p p_{n_o}}{W_B} (e^{V_C/V_\phi} - 1) \tag{5-19}$$

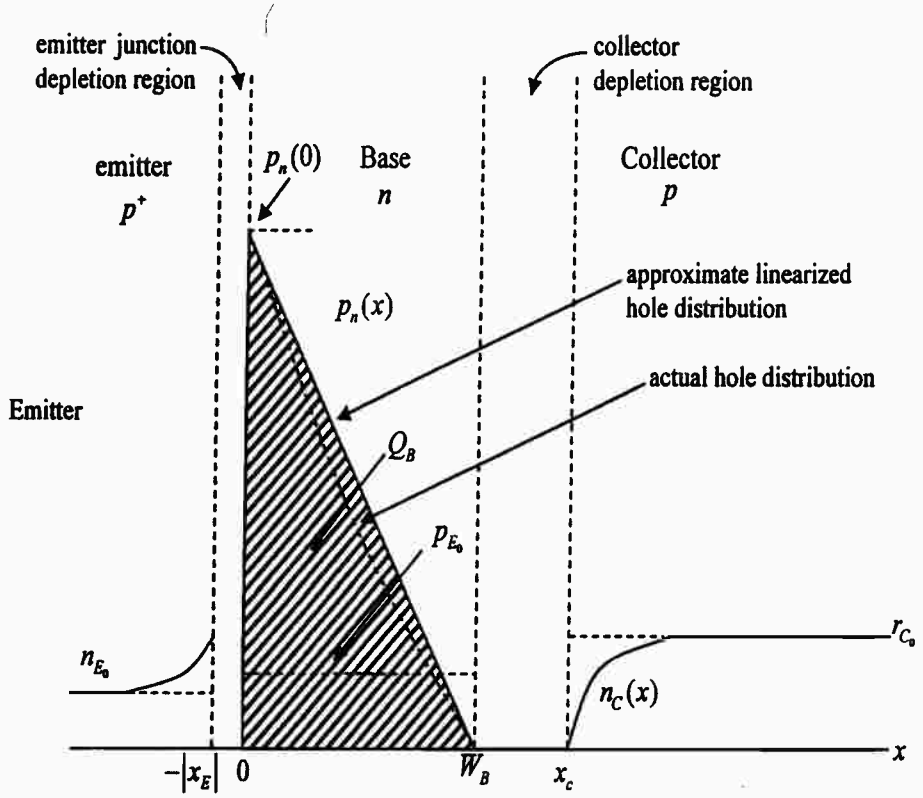


Fig. (5.5) Excess minority carrier distribution around the forward biased emitter – base junction of a pnp transistor

If the transistor is operating in the active mode (i.e., collector – base junction reverse biased and emitter – base junction forward biased) then I_E will be given by

$$I_E = -|q|A\left(\frac{D_p p_{n_o}}{W_B} + \frac{D_p n_{E_o}}{L_E}\right)\left(e^{V_E/V_{th}} - 1\right) + \frac{|q|AD_p p_{n_o}}{W_B} \qquad (5\text{-}20)$$

The emitter efficiency γ can then be approximately given by

$$\gamma = \frac{I_{p_E}}{I_{p_E} + I_{n_E}} = \frac{\dfrac{D_p p_{n_o}}{W_B}}{\dfrac{D_p p_{n_o}}{W_B} + \dfrac{D_n n_{E_o}}{L_E}}$$

$$= \frac{1}{1 + \dfrac{D_n N_D W_B}{D_p N_A L_E}} = \frac{1}{1 + \sigma_B W_B / \sigma_E L_E} \qquad (5\text{-}21)$$

where N_D and N_A are the base and emitter doping concentrations, respectively. The emitter injection efficiency can be improved by increasing the ratio of the doping concentration in the emitter to that in the base. A typical value of this ratio is 1000.

EX. 5.1

A silicon n^+pn transistor has a base width of $2\mu m$ and base resistivity of $0.001\Omega m$. Take $I_E = 1\mu m$ and the emitter resistivity $0.00001\Omega m$. Calculate the injection efficiency of holes for the emitter of this transistor.

Solution

The injection efficiency for npn transistor is given from

$$\gamma = \frac{1}{1+\sigma_p W_B/\sigma_n L_E}$$

$$= \frac{1}{1+\dfrac{(1/0.001)(2\times10^{-6})}{(1/0.00001)(1\times10^{-6})}} = \frac{1}{1+2+10^{-2}} = 0.98 \text{ or } 98\%$$

Up till now we have neglected any recombination inside the base region. However, an estimation of the recombination current is not difficult as the total minority carrier charge Q_B inside the base can be easily found by integrating its linear distribution of Fig. (5.5) and following the steps of the charge control model of eqns. (4-75) through (4-80)

$$Q_B = |q|\frac{p_n(0)W_B}{2}A \qquad (5\text{-}22)$$

If the average hole lifetime inside the base is τ_p, then the hole recombination current which is I_B given by

$$I_B = |q|A\frac{p_n(0)W_B}{2\tau_p} = \frac{|q|AW_B}{2\tau_p}p_{n_o}e^{V_E/V_o} \qquad (5\text{-}23)$$

The hole current reaching the collector (I_{pc}) is, therefore, obtained by substracting the above recombination current lost in the base from the initial emitter hole current I_{pE}:

$$I_{pc} = I_{pE} - I_B$$

which is approximately given by

$$I_{pc} = I_{pE}\left(1-\frac{W_B^2}{2L_B^2}\right), \qquad (5\text{-}24)$$

where $L_B = \sqrt{D_p\tau_p}$ is the diffusion length of holes in the base. The transport factor is given by

$$B_T = \frac{I_{Pc}}{I_{pE}} = 1-\frac{W_B^2}{2L_B^2} \qquad (5\text{-}25)$$

Thus, the transistor alpha in the active mode is given by

$$\alpha = \gamma B_T = \frac{1-W_B^2/2L_B^2}{1+\dfrac{D_n N_D W_B}{D_p N_A L_E}} \qquad (5\text{-}26)$$

We notice that in the calculation of the base minority current we have assumed that the excess hole distribution inside the base is linear. Indeed, this approximation is quite accurate in the case of narrow base widths ($W_B << L_B$). However, according to this same approximation, the diffusion current inside the base is constant and the collector hole current is exactly the same as the emitter hole current. This implies that the base current is zero, which is not true. In fact, the distribution of excess holes inside the base is never perfectly linear even for very small base widths.

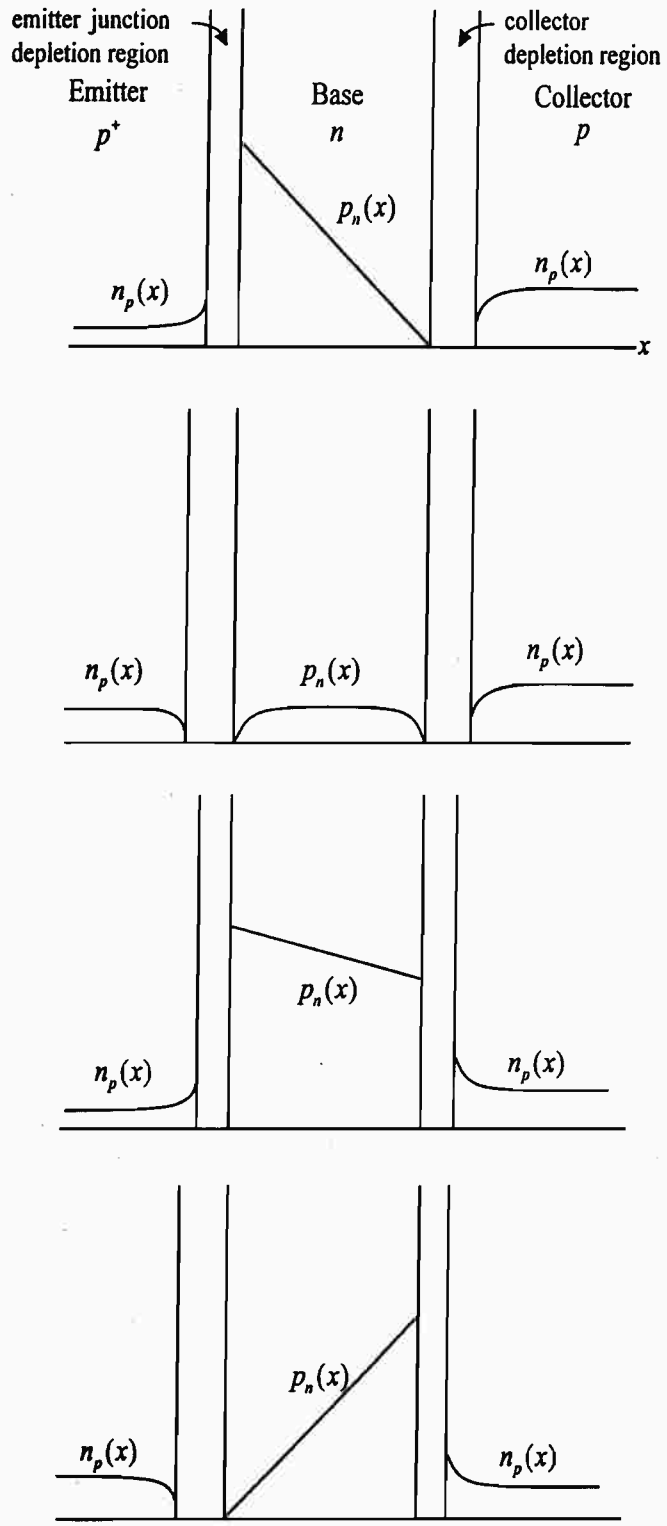


Fig. (5.6) Minority carrier distributions across a pnp transistor for the four different modes of operation.

a) active mode b) cutoff mode c) saturation mode d) inverse mode

The dashed curve of Fig. (5.5) shows the real situation, where the slope of the distribution varies slightly from the emitter to collector. The deviation from linearity will satisfy the condition that the collector hole current is different from the emitter hole current. On the other hand, the value of the calculated base current will not be significantly affected, since the area under the curve is nearly the same.

5.4 Modes of Operation

So far, we have been concerned only with a single mode of operation, in which the emitter-base junction is forward biased and the collector-base junction reverse biased. As mentioned before, this is called the active or normal mode and it is in this mode that a bipolar transistor is usually employed in almost all analog circuits. In practice, however, there are three other modes of operation, depending on the voltage polarities of the two transistor junctions. Fig. (5.6) shows the minority carrier concentration profiles for the different modes of operation of a pnp bipolar transistor.

In the cutoff mode (Fig. 5.6b), both junctions are reverse biased. The collector current is nearly zero and a virtual open circuit exists between the collector and the emitter terminals. No excess minority carrier charge is stored in the base and the base current approaches zero. This mode corresponds to the OFF state of the transistor when used as a switch. On the other hand, the saturation mode (Fig. 5.6c) corresponds to the ON state of the transistor in digital applications. In this mode, both junctions are forward biased and the resultant collector to emitter voltage V_{CE} approaches zero. As a result, the transistor acts as a short circuit and the collector current is controlled by the external circuit elements.

When the emitter-base junction is reverse biased and the collector-base junction is forward biased, the transistor is said to be operating in the inverse mode. In other words, the emitter serves as a collector and the collector as an emitter. In this mode, the transistor alpha (denoted α_I) is smaller than the normal current transfer ratio (α_n). This is because the injection efficiency of the collector is much lower than that of the emitter, due to the low doping level of the collector with respect to that of the base. A plot of I_C versus V_{BE} (Fig. 5.7) shows that up to $V_{BE} = 0.5\text{V}$ the collector current is very small. This is called cutin region. Of course for V_{BE} negative we have cutoff. When $V_{BE} \cong 0.7\text{V}$ this is the active region. Where $V_{BE} > 0.7\text{V}$ we have saturation.

5.6 Transistor Configurations

There are three different configurations for the transistor as a circuit element. So far, we have only referred to the common base (CB) configuration (Fig. 5.8a) in which [he base terminal is common between the input and output circuits. Fig. (5.8b) and (5.8c) show the common-emitter and common-collector configurations.

A) The Common- Base Configuration

The action of the common-base configuration as a voltage amplifier has been demonstrated at the beginning of this chapter. In this configuration, the input signal is introduced between the collector and base terminals. It is clear that this configuration does not produce current amplification since the output current (I_C) is the same as the input current I_E or slightly less.

Fig. (5.9a) shows the relation between V_{EB} and I_E with V_{CB} as a parameter. This family of curves is known as the input characteristics.

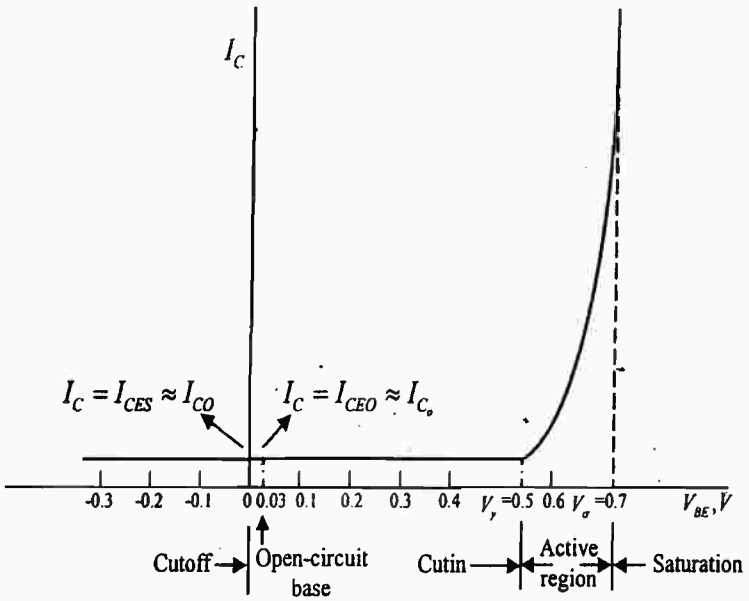


Fig. (5.7) Plot of collector current against base-to-emitter voltage (I_C is not shown to scale)

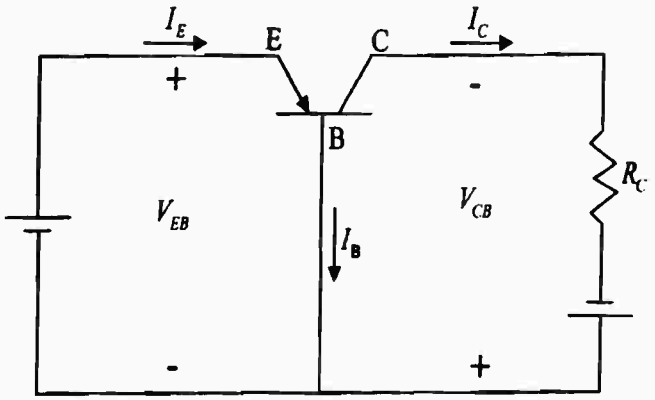
When the collector-base junction is open circuited, the relation between V_{EB} and I_C is nothing but that of a normal pn junction. However, when the collector-base junction is short circuited or reverse biased, the emitter current at a given value of V_{EB} becomes greater. This is due to the reduction of the active base width as the depletion layer extends into the base region and the effective sweeping action of the collector junction which is called Early effect. Such a reduction is the value of the active base width makes the gradient of minority carriers in the base more steep and consequently, the emitter current increases. On the other hand, Fig. (5.9b) shows the output characteristics of a pnp transistor in the common-base configuration. This is a relation between I_C and V_{CB}, with the input current I_E as a parameter. When the emitter base junction is forward biased and the collector-base junction is reverse biased (i.e., V_{CB} is negative), the transistor is said to be in the active region and $I_E = I_C$. If, however, the emitter-base junction is reversed, so that $I_E = 0$, while the other junction is also reversed, the transistor is in the cut off mode. Hence, the condition for cut off is $I_E = 0$.

In the saturation region, forward bias of the collector causes large increase in collector current for small changes in the collector voltage. The collector current then increases exponentially with voltage as in a diode characteristic.

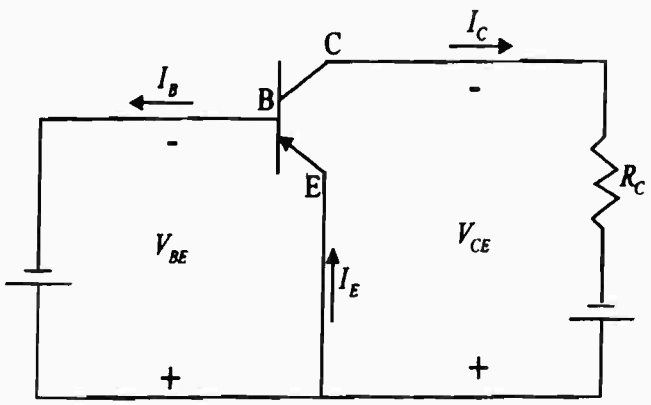
B) The Common-Emitter Configuration

The common-emitter configuration is the most frequently used of all three. As shown in Fig. (5.8b) the input signal is applied between the base and the emitter terminals and the output is taken across a load connected between the collector and the emitter. When the transistor is connected in this configuration, it can amplify both input current or voltage signals. The collector current is controlled by the base current and any small change in I_B is amplified to produce a large change in I_C.

a)



b)



c)

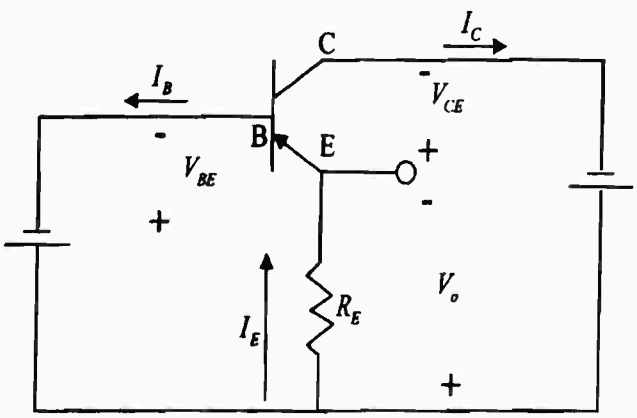
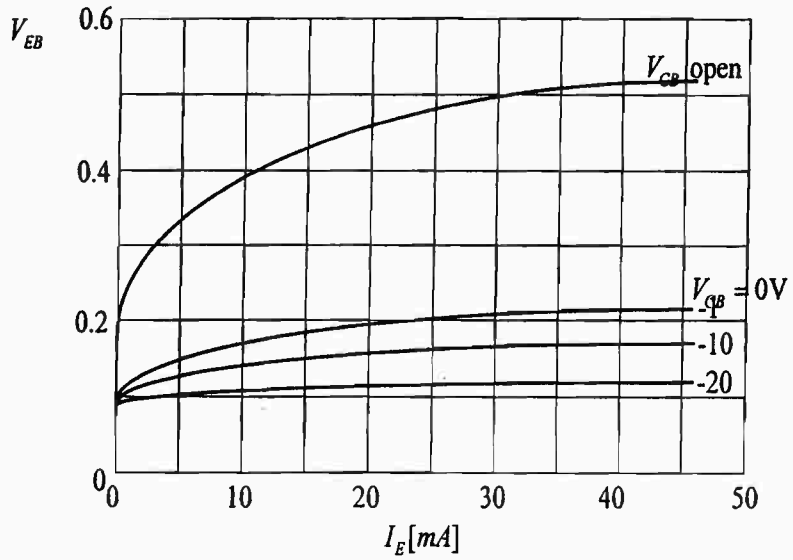


Fig (5.8) Circuit diagrams showing the three different pnp transistor configurations:
a) common-base b) common-emitter c) common-collector

a)



b)

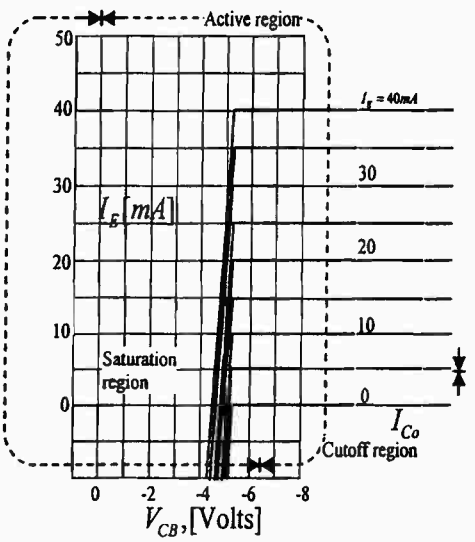


Fig. (5.9) Plots of I-V characteristics of a pnp bipolar transistor in the common-base configuration

a) input relation b) output relation

In the common-emitter configuration, the base current is controlled by the outer circuit elements. Fig. (5.10a) shows the input characteristics of a pnp transistor in the common-emitter configuration. Each curve gives the relation between the base current and the base to emitter voltage for a given value of the output voltage, that is the collector to emitter voltage V_{CE}. When $V_{CE} = 0$, the collector and emitter regions are tied together and the resultant is a parallel combination of the two pn junctions. Therefore, the I-V curve in this case is the same as that of a normal pn junction diode. If V_{BC} is zero, I_B will be zero. Increasing $|V_{CE}|$ with constant V_{BE} causes a decrease in base width (Early effect) which decreases the recombination base current.

The output characteristics of a pnp transistor in the common - emitter configuration are depicted in Fig. (5.10b). Each curve represents the collector current I_C as a function of the collector to emitter voltage V_{CE} for a given value of the base current I_B. Similar to the output characteristics of the common base configuration, this family of curves may be divided into three regions: the active, the cutoff and the saturation regions.

In the active region, the collector current is very sensitive to variations in the base current I_B. It is in this region that a transistor should be biased to act as a current or voltage amplifier. Making use of eqns. (5.5) and (5.6), we obtain:

$$I_C = \frac{I_{C_o}}{1-\alpha} + \frac{\alpha}{1-\alpha} I_B = \frac{I_{C_o}}{1-\alpha} + \beta I_B = (\beta+1) I_{C_o} + \beta I_B \qquad (5\text{-}27)$$

Actually if α were truly constant I_C would be independent of V_{CE}. Due to Early effect α increases with $|V_{CE}|$. Thus an increase in β is more pronounced.

The above eqn. proves that the collector current could be controlled by the base current in a very sensitive manner, since the factor β is normally very high. The collector current increases with increasing $|V_{CE}|$ in the active region due to the subsequent reduction of the active base width and the increase of the transistor α which approaches unity. In other words, the increase of the collector current for a fixed value of I_B in the active region is due to the increase of the factor β as the collector to emitter voltage $|V_{CE}|$ is increased.

In the cutoff region, both junctions are reverse biased. In fact, the condition for cutoff is that $I_E = 0$. In this case, the collector current and the base current are both equal to the reverse saturation current of the collector-base junction I_{C_o}. It is to be noted that reducing the base current to zero is not sufficient to drive the transistor into cutoff.

According to eqn. (5.48), if $I_B = 0$, then $I_C = I_{CE_o} = I_{C_o}/(1-\alpha)$, which is much higher than I_{C_o}. So, it is necessary to reverse the bias of the emitter-base junction, so that α may drop to zero and $I_C = I_{C_o}$.

When both junctions are forward biased, the collector to emitter voltage falls down to a value in the order of a few tenths of a volt. This is called the saturation region, in which case the collector current ceases to increase with increasing base current and is mainly governed by the external circuit elements. From Fig. (5.8b) the collector current is given by:

$$I_C = \frac{V_{CC} - V_{CE}}{R_L} \qquad (5\text{-}28)$$

$$I_{CS} = \frac{V_{CC}}{R_L} \qquad (5\text{-}29)$$

In the case of saturation, as mentioned above, the collector to emitter voltage $V_{CE,sat}$ is very small compared to the supply voltage V_{CC} and the collector saturation current I_{CS} is given by eqn. (5-29).

Consequently, any further increase in the base current will be met with an identical increase in the emitter current, whereas the collector current remains unchanged. A transistor can be driven into saturation by injecting a high current into the base region. This will cause the collector current to increase and V_{CE} to drop down due to the increasing drop across R_L. The ratio $V_{CE,sat}/I_C$ is called CE saturation resistance

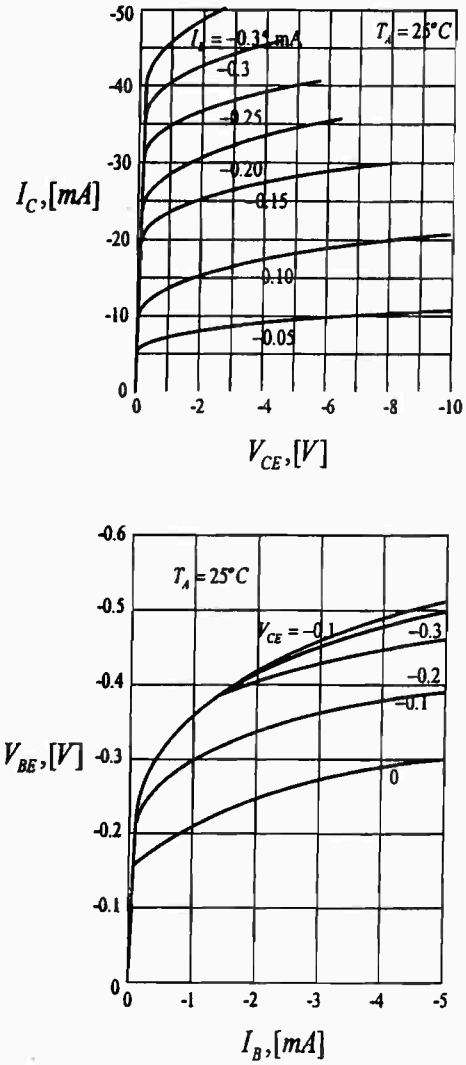


Fig. (5.9) Plots of I-V characteristics of a pnp bipolar transistor in the common emitter configuration
a) input relation b) output relation

(R_{CEsat}). The ratio I_C / I_B in saturation is less than β at dc since an increase is I_B does not produce an equivalent increase in I_C.

C) The Common-Collector Configuration

In this configuration (Fig. 5.8c), the output is taken across a load connected between the emitter terminal and ground. It is evident that such a configuration cannot produce a voltage gain, since the output voltage is the major part of the input signal applied between the base and ground. In fact, the voltage gain of this configuration is only slightly less than unity, since most of the input signal goes across R_L, due to the very small forward resistance of the emitter-base junction. For this reason, a transistor in the common collector configuration is referred to as a voltage follower. Such a configuration is capable of producing current amplification not voltage amplification, and is mostly used as a buffer

between two different circuits. A buffer serves to isolate two circuits connected in cascade. It should have a large input impedance, low output impedance and should provide unity gain. Concerning the transistor operation in the common collector configuration, it is very much similar to that in the common emitter one. In the cutoff region, the emitter current is equal to zero, whereas in the saturation mode the collector to emitter voltage is negligibly small.
To summarize:

A) Active Mode:
The base io emitter voltage is about 0.7V for Si transistors and the collector current is related to the base current by the relation $I_C = \beta I_B$, where I_{C_o} may be neglected.

B) Saturation Mode:
The base to emitter voltage is about 0.8V for Si transistor. The collector to emitter voltage is in the order of a few tenths of a volt 0.2 V for Si. The collector current no longer responds to variations in the base current and is governed by the external circuit elements.

C) Cutoff Mode:
The two junctions are reverse biased and the emitter current I_E is zero. A reverse bias of 0V-0.5V is sufficient to drive the base-emitter junction of a Si transistor into cutoff.

Ex. 5.2
Show that V_{CEsat} is about 0.2 V for a CE Si transistor.

Solution
In the CE configuration of an npn transistor,

$$V_{CE} = V_{BE} + V_{CB}$$

In the active region, V_{BE} for Si transistor is 0.7V, and V_γ of the CB junction is 0.5V. If $V_{CE} = 1\text{V}$, V_{BE} is 0.3V above, V_{CB} and the CB junction is reverse biased. If $V_{CE} = 0.5\text{V}$, V_{CB} becomes - 0.2 V, i.e., the CB junction is forward biased but at a voltage less than the cutin voltage of the CB junction. If $V_{CE} = 0.2\text{V}$, V_{CB} becomes -0.5 V and the CB junction begins
to conduct, i.e., the start of saturation occurs when both EB and CB junctions are forward biased. As the transistor is driven deeper into saturation, V_{BE} becomes 0.8 V and V_{CB} becomes -0.6 V and V_{CE} remains constant at 0.2 V.

Ex. 5.3
The silicon transistor of the circuit shown has $\beta = 100$ and $I_{C_o} = 20\ nA$. Find the minimum value of R_B required to keep the transistor in the active region.

Solution
Assume that the transistor is in the active region. Therefore, $V_{BE} = 0.7\text{V}$.
The transistor will go into saturation if V_{CE} is 0.2 V and

$$I_B \geq \frac{I_C}{\beta}$$

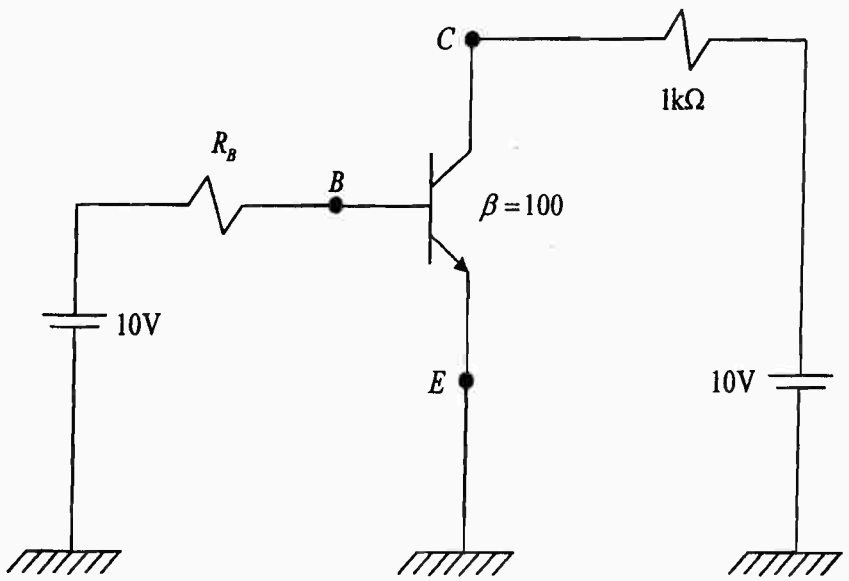


Fig. (Ex.5.3)

or

$$\frac{(10-0.7)\text{V}}{R_B} \geq \frac{(10-0.2)\text{V}}{1k\Omega}\frac{1}{\beta}$$

Therefore,

$$R_B \leq \frac{(10-0.7)}{(10-0.2)}\,100\times10^3$$

$$R_B \leq 100\times10^3\Omega$$

The above value for R_B is a critical value, below which the transistor goes into saturation. Therefore, it is the minimum value that can keep the transistor in the active mode.

5.7 The Transistor as a Switch

Bipolar transistors are widely used as switches in digital applications. In switching applications, the transistor is made to operate either in the cut off region or in saturation by the action of a base current control signal. In saturation, the transistor acts as a virtual short circuit between the collector and emitter terminals, since $V_{CE} \simeq 0$, and I_C depends only on the external circuit elements. On the other hand, when the transistor is driven into cutoff, the collector current is nearly equal to zero and a virtual open circuit exists between the collector and the emitter.

A) Steady State Switching

Fig. (5.11) shows a transistor being driven from saturation to cutoff by a pulse waveform applied to the base. This is the basis of using a transistor as an inverter in digital circuits.

When the transistor is in saturation, it is said to be in the ON state, whereas when driven into cutoff it is said to be in the OFF state. In the cutoff region, the following conditions prevail:

1- Both junctions are reversed biased.

1- Both junctions are reversed biased.
2- The emitter current is equal to zero.
3- The collector current is equal in magnitude to the base current, and both are equal to the reverse saturation current of the collector-base junction.

When the input signal V_s is high, the transistor is driven into saturation giving an output voltage of 0.2V which is low. When the input is 0V the transistor is off and the output is V_{cc}, i.e., high. This is the realization of the truth table of an inverter which is:

Input	output
Low	High
High	Low

On the other hand the saturation condition is satisfied as both junctions are forward biased, in which case the collector to emitter voltage is nearly zero and $I_C = I_\alpha V_{CC} R_L$. Fig. (5.12) shows the operating points corresponding to the ON and OFF states on the output characteristics of a transistor in the common emitter configuration. The minimum base current required to drive the transistor into saturation is given by

$$I_B \geq \frac{I_\alpha}{\beta} = \frac{V_{CC}}{\beta R_L} \qquad (5\text{-}30)$$

To test if the transistor is in saturation or not the base and collector currents should be determined independently from the loop equation, after assuming V_{CE} and V_{BE} have their saturation values. If the determined base current is greater than I_C/β then the assumption is correct and the transistor is indeed in saturation. If however I_B turns out to be equal to I_C/β then the transistor is not in saturation.

B) Transient switching

When a device is switched from one state to another it is desirable that such switching take place instantly. As with the pn junction switching from one state to another is accomplished by a change in the carrier distributions around the junctions. These carrier distributions need a finite time in order to be established or be swept away. This is called the switching time. It is useful to summarize switching times as follows in Fig. (5.13) for a CE transistor.

A) Delay Time

The delay time t_d is the time taken by the transistor from the instant of switching on till the collector current reaches 10 percent of its steady state value I_{CS}. The delay lime is determined by:

(1) the time required to charge the transition capacitances,
(2) the time required by the carriers to cross the base and be collected at the collector terminal.

B) Rise Time

It is the time t_r required by the collector current to rise from 10% to 90% of its saturation value I_{CS}. In other words, it is the time necessary to build up the active base charge, such that its gradient reaches 90% of its maximum value.

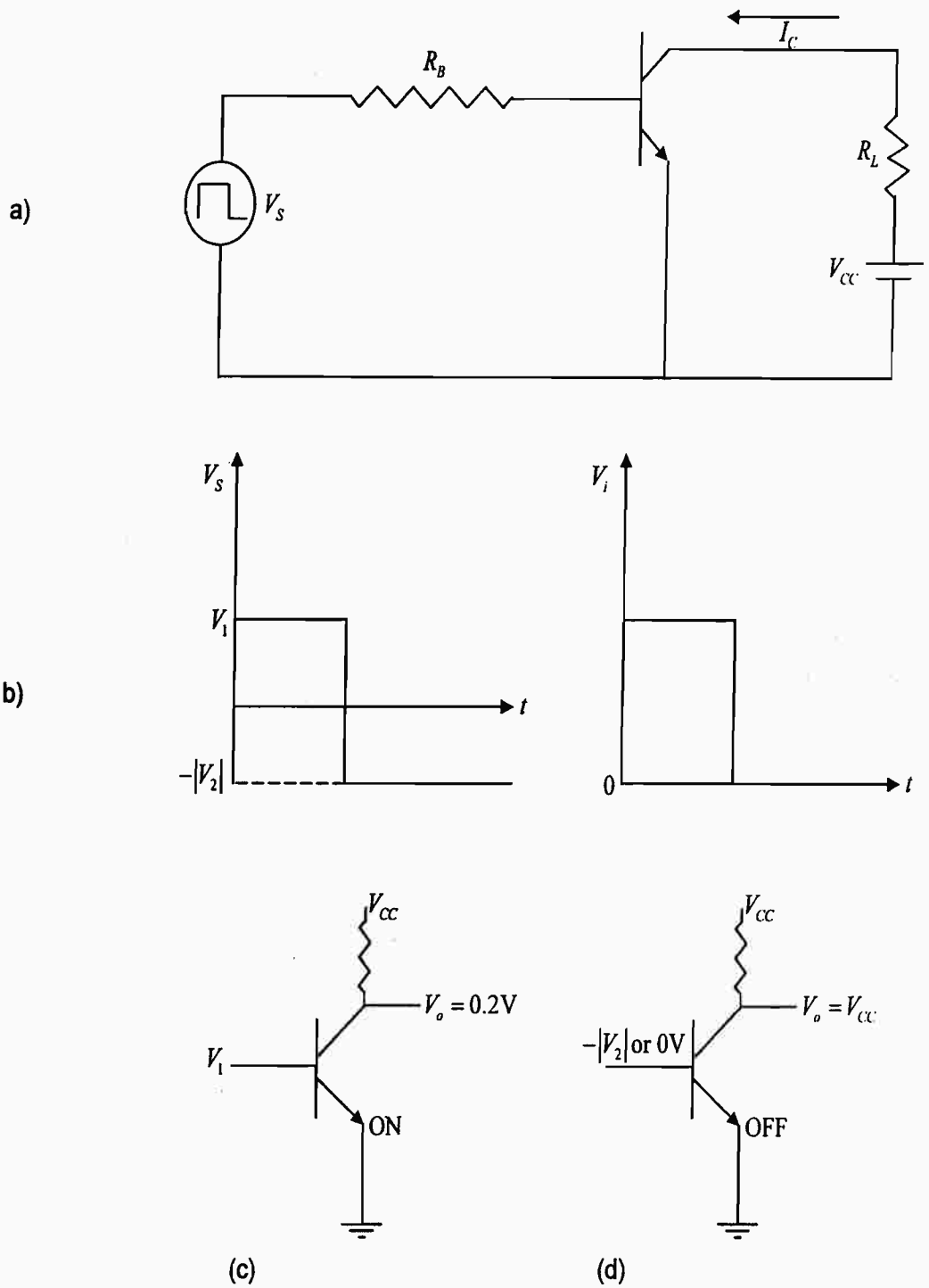


Fig. (5.11) Elementary switching circuit for an npn transistor in the common – emitter configuration

a) circuit b) input waveform

c) inverter ON (low state either $-|V_2|$ or 0V) d) inverter OFF

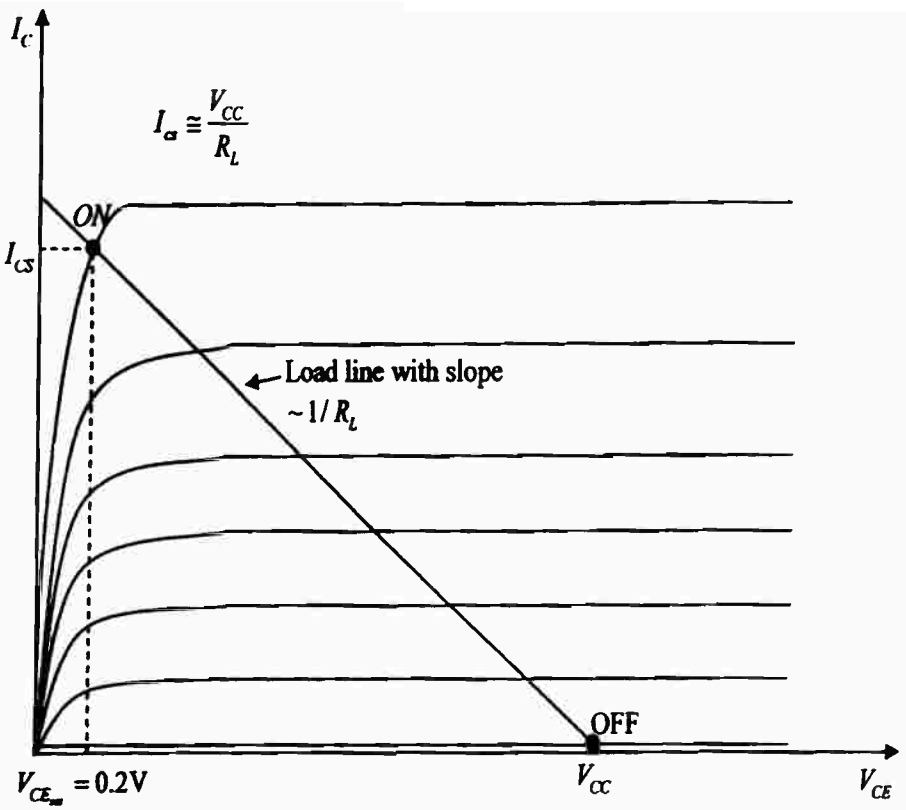


Fig. (5.12) Output characteristics and load line for the circuit of Fig. (5.11)

C) Fall Time

The fall time t_f is defined as time taken for the collector current to decrease from 90% to 10% of its saturation value I_{CS}. During this time, the active base charge or the excess minority carrier gradient in the base is swept off.

D) Storage Time

The storage time t_s is the time lapse between the moment when the base current starts to reverse and the moment when the collector current reaches 90% of I_{CS}. This is the time required to bring the transistor out of saturation. In saturation, an amount of excess carriers, beyond those necessary to maintain the collector current in the active mode, is stored in the base. This excess charge has to be first swept away before the current starts to decay. The time required to evacuate such a charge is the prime limiting parameter to the switching speed of bipolar transistors. This is the storage time t_S.

5.8 The Hybird - π Equivalent Circuit

For small signal deviations around the Q point we may develop an equivalent circuit for BJT which relates various small signals (an ac signal does not necessarily mean sinusoidal) around the Q point. The hybrid - π model is a high frequency small signal equivalent circuit for bipolar transistors connected in the common emitter configuration and biased in the active mode. This model derives its importance from the fact that its elements are directly related to the physical parameters of the transistor.

Moreover, the common-emitter circuit is the most frequently used of the three transistor circuit configurations, as it provides current gain as well as voltage gain. As a result, the hybrid-π model has become the most widely used small signal equivalent circuit of bipolar transistors. The equivalent circuit does not contain the battery or any dc biasing voltages. But if the deviation is dc then it still can be used for corresponding dc deviations or changes of voltages or currents around the Q point.

The hybrid-π, or the Giacoletto model is shown in Fig. (5.14). The resistance r'_{bb} is the base spreading resistance; it extends between the base terminal and a non-accessible node b' inside the base region. The voltage drop between the node b[1] and the emitter terminal is the net small signal voltage change across the base emitter junction. As a result, small variations in the minority carrier charge injected from the emitter to the base are determined by the voltage υ'_{be}. Therefore, the small signal collector current, with the collector dc voltage constant or with the collector short circuited from the point of view of an ac signal is directly proportional to υ'_{be} and the constant of proportionality is the transconductance g_m defined as:

$$g_m = \left.\frac{di_C}{d\upsilon_E}\right|_{V_C = \text{constant}} \tag{5-31}$$

where $\upsilon_s = \upsilon'_{be}$ is the voltage directly impressed on the base-emitter junction. Making use of the collector current expression eqn. (5.29), it is straightforward (Prob. 5.8), to derive.

$$g_m = \frac{|I_C|}{V_{th}} \tag{5-32}$$

As seen from the above relation, the transconductance g_m depends on the dc operating point and is directly proportional to the dc collector current I_C. The transconductance is one of the most important parameters, since it determines the small signal gain of the transistor. Bipolar transistors are characterized by their high transconductance in comparison with other types of transistors.

The emitter capacitance C_e is the overall layer capacitance of the emitter-base junction. It is the sum of diffusion capacitance and the depletion capacitance. The diffusion capacitance accounts for the accumulated charge due to injection while the depletion capacitance is due to uncovered ions. Carriers supplied through the base terminal, compensating for the recombination process give rise to the base current. Therefore, the conductance $g'_{be} = 1/r'_{be}$ is nothing but the ratio of the change in the base current I_B to the change in the emitter junction voltage V_E.

$$r_\pi = g'_{be} = \left.\frac{di_B}{dv_E}\right|_{V_C = \text{constant}} \tag{5-33}$$

We can deduce (Prob. 5-9):

$$g'_{be} = \frac{|I_B|}{V_{th}} \tag{5-34}$$

The capacitance C_{TC}. is the depletion layer capacitance of the reverse biased collector base junction. In parallel with C_{TC}, we find the feedback resistance $r_{b'c}$. This resistance accounts for the variation of the collector current due to the variation of the collector base junction reverse bias and the resulting base width modulation. The resistance r_{ce} is the output resistance between the collector and emitter terminals and is in the order of 80kΩ or higher. Finally $r_{bb'}$ represents the base spread or ohmic resistance.

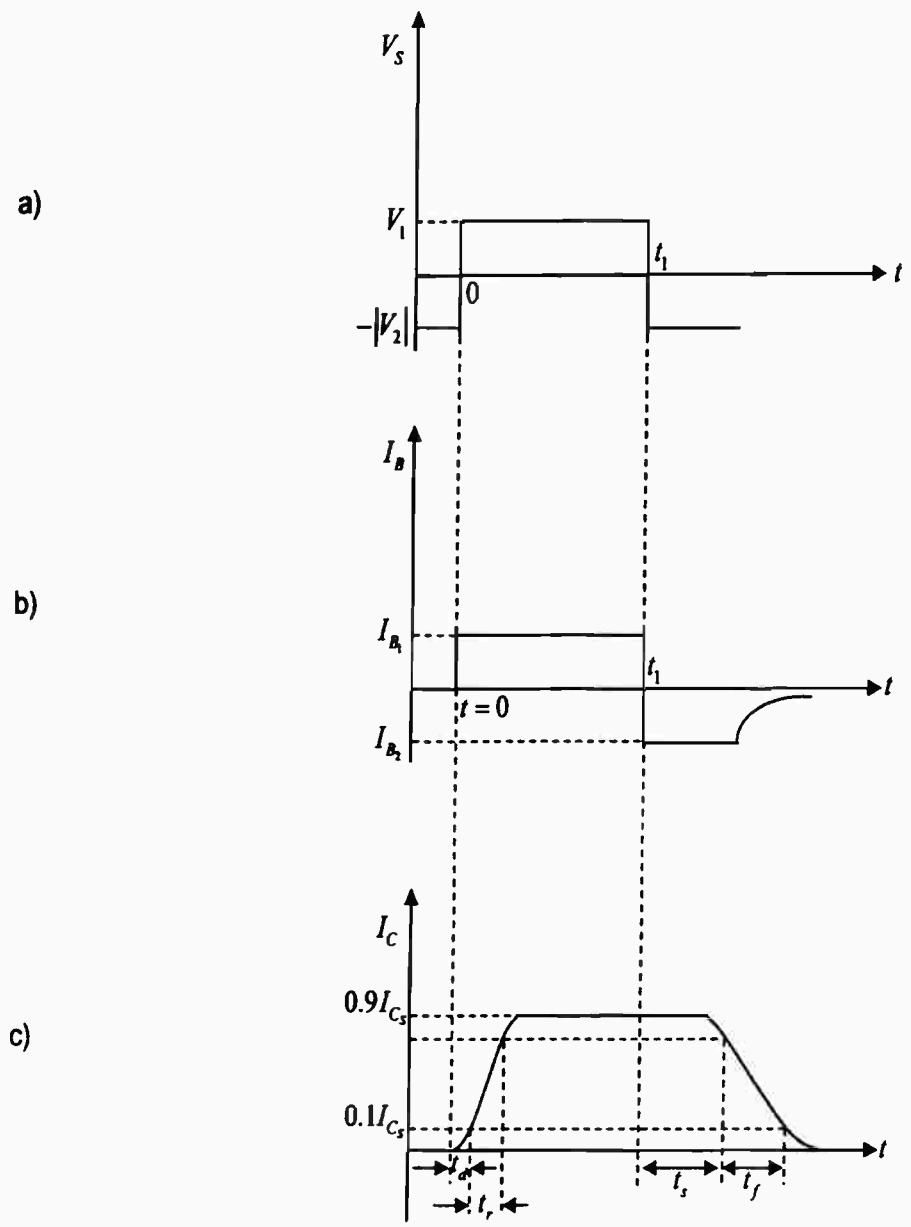


Fig. (5.13) Switching transients in a common-emitter transistor:
a) input voltage waveform b) base current waveform c) collector current waveform

5.9 Simplified h-Model

An equivalent simplified hybrid model is also used (Fig. 5.15), where $r_{b'c}$ is considered very high, assuming no frequency effects, i.e., eliminating all capacitances and neglecting r_{bb} we end up with a simplified h-model (Fig. 5.17), where the input resistance is h_{ie} and the ac current output is i_c and the ac input voltage is $V_{be} = v_i$ and the ac input current is $i_b = i_i$

$$h_{ie} = r_{b'e} \tag{5-35}$$

$$i_c = g_m v_{be} = g_m (i_b h_{ie}) = h_{fe} i_b = \beta i_b \tag{5-36}$$

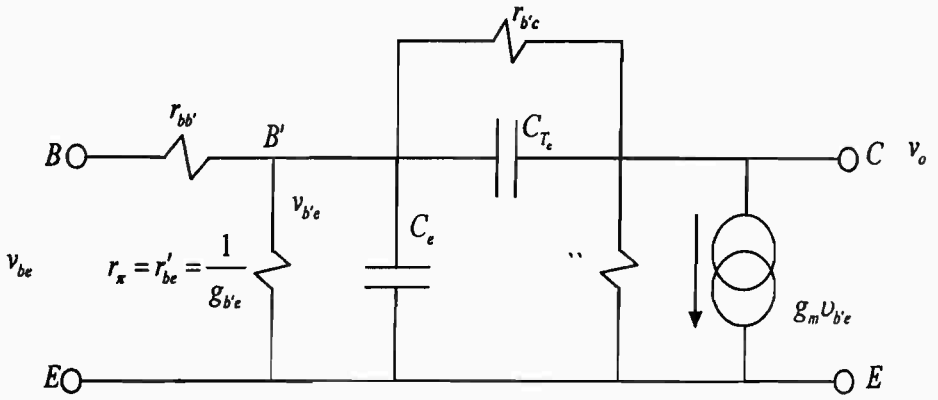


Fig. (5.14) The hybrid π equivalent circuit

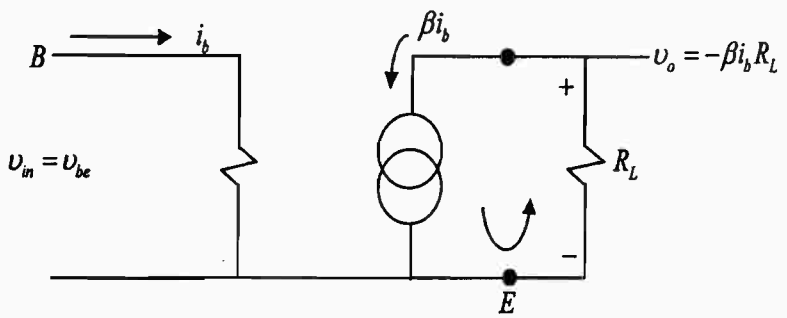


Fig. (5.15) Simplified h-model

We define the current gain as the ratio of small signal (ac) collector current to the small signal (ac) base current. This is β or h_{fe} given by

$$h_{fe} = \beta = g_m h_{ie} \qquad (5\text{-}37)$$

This means that a small variation in base current gives a large variation in the collector current. It is to be noted that h_{fe} varies with dc currents in the circuit as well with temperature (Fig. 5.15). The voltage gain A_V is given by

$$A_V = \frac{v_o}{v_i} = -\frac{\beta i_{in} R_L}{i_{in} h_{ie}} = -\frac{\beta R_L}{h_{ie}} \qquad (5\text{-}38)$$

This is the basis of using a BJT as an amplifier. But where, does this amplification come from noting that energy cannot be generated or destroyed?

5.10 Energy Considerations

We should note that at dc (Q point) we have,

$$V_{CC} = V_{CE} + I_C R_L \tag{5-39}$$

$$V_{CC} I_C = V_{CE} I_C + I_C^2 R_L \tag{5-40}$$

This means that the collector input power is divided between the collector emitter power needed to pump carriers to surmount the barrier height and the power loss across R_L. When I_b is increased by an incremental amount ΔI_b, the collector current I_C changes by a small amount ΔI_C.

$$V_{CC}(I_C + \Delta I_C) = (V_{CE} + \Delta V_{CE})(I_C + \Delta I_C) + (I_C + \Delta I_C)^2 R_L \tag{5-41}$$

But from eqn. (5-39), V_{CC} = constant

$$\Delta V_{CE} = -\Delta I_C R_L \tag{5-42}$$

It can be shown (Prob. 5.18),

$$\Delta V_{CE} \Delta I_C = -\Delta I_C^2 R_L \tag{5-43}$$

Thus, the output power comes from a saving in the power loss in the BTT.

Problems

1- A silicon npn transistor has a base width of $1.5\ \mu m$. Electrons are injected into base forming a uniform concentration gradient, which drops to zero at the collector junction. If the electron concentration at the emitter edge of the base is 1.5×10^{20} electron/m^3, determine the electron diffusion current inside the base under steady state conditions. The electron mobility in silicon is $0.14\ m^2/(V.s)$.

2- Derive the ratio of the electron to hole currents I_{nE}/I_{pE} crossing the emitter of a transistor in terms of the conductivities of the emitter and base regions.

3- Derive the junction voltages V_E and V_C and the collector to emitter voltage V_{CE} in terms of the transistor currents.

4- Obtain the emitter - base junction voltage required to produce cutoff.

5- Find the input resistance and output resistance in the following cases using the simplified h-model.
   a) CE with zero R_E.
   b) CE with R_E.
   c) CC
   d) CB

6- The connection shown is known as the Darlington pair. Assuming that the transistors are identical and have the same β, find the current gain, input and output resistances. Neglect the reverse saturation currents. Show the advantages of this circuit.

7- Analyze the circuit in Prob.5 and find the current gain and voltage gain. Show the advantages of using this circuit.

8- In CE circuit, $R_C=3k\Omega$, $R_B=50k\Omega$.
$V_{BB}=5V$, $V_{CC}=10V$, $\beta=100$. Determine whether or not the silicon transistor is in saturation and find I_B and I_C. Repeat if a $2k\Omega$ emitter resistance is added.

9- Find the transistor currents for the circuit shown. The transistor has $\beta=99$ and $I_{C_o}=20\ nA$ and is made from silicon. Is this transistor saturated? What happen if R_B is changed to $500k\Omega$? What do you conclude? What is the minimum value of R_B to produce cutoff?

10-Repeat the problem above if a $1.8\ k\Omega$ resistor is connected between the emitter terminal and ground.

11-A silicon transistor is used in the circuit shown with $V_{CC}=22.5V$, $R_C=5.6\ k\Omega$, , $R_2=10\ k\Omega$, $R_1=90\ k\Omega$, $\beta=50$. Find the Q point. (Hint: use Thevenin's theorem)

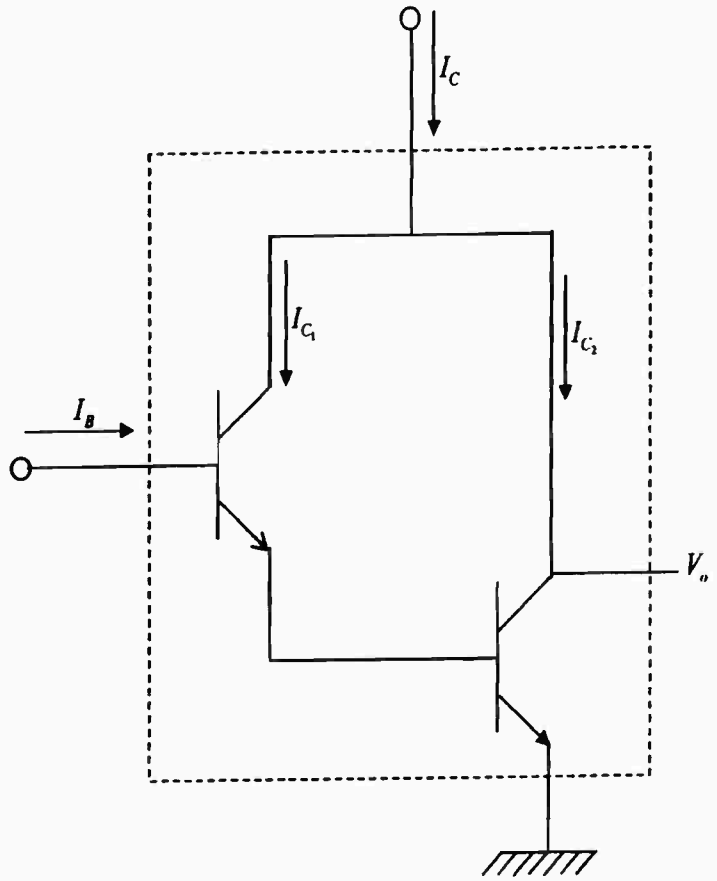


Fig. Prob. (5.5)

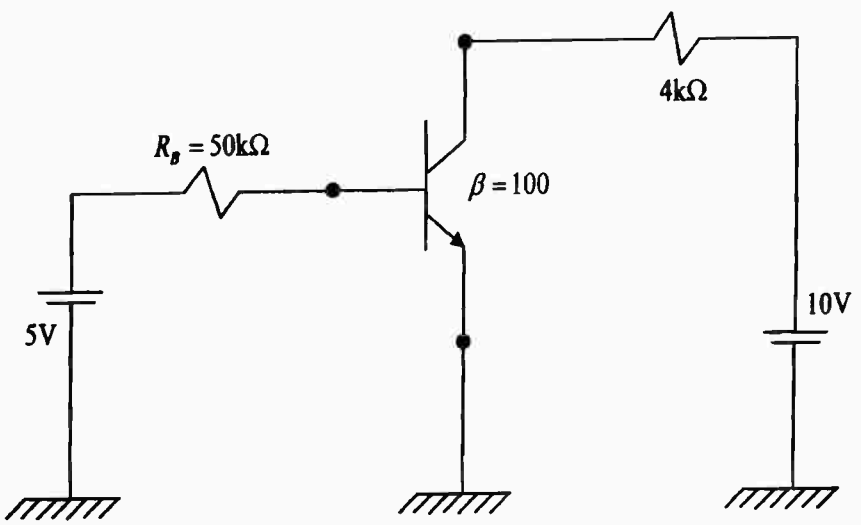


Fig. Prob. (5.7)

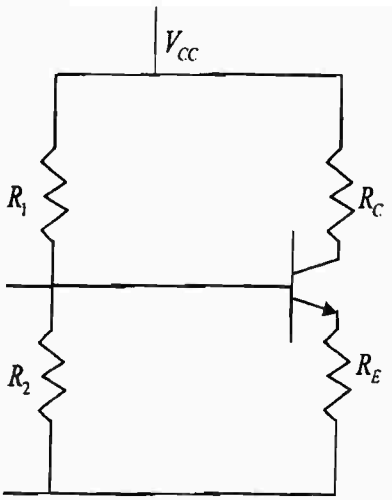


Fig. Prob. (5.9,10)

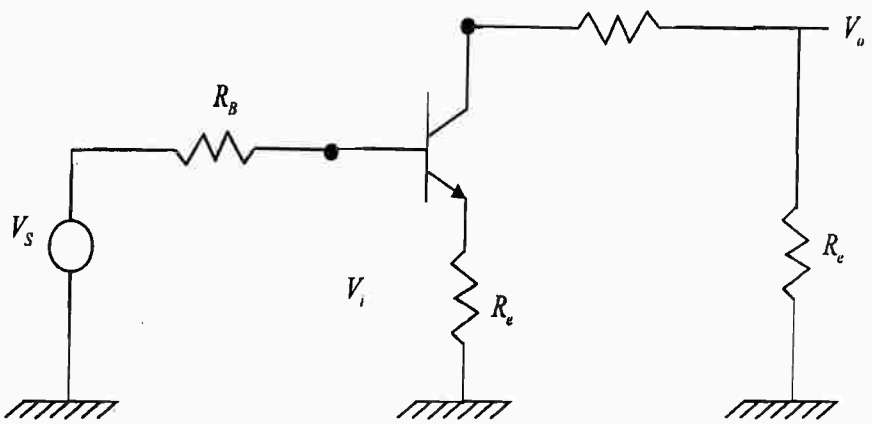


Fig. Prob. (5.11)

12-In the Fig. (Prob.9) circuit, β varies from 36 to 90.
Find R_E, R_1 and R_2, where $R_C = 4\text{k}\Omega$, $V_{CC} = 20\text{V}$, the nominal bias point $V_{CE} = 10\text{V}$, $I_C = 2\text{mA}$. (I_C should be in the range 1.75 to 2.25 mA as β varies from 36 to 90) Neglect I_{C_o}.

13-Find the voltage gain and current gain of the ac circuit shown using the simplified h model. What happens when R_E is zero and when R_E exists in the emitter lead.

14-In the above problem the output is taken across R_E find the voltage gain. This is common collector configuration. Show what it may be used for. Does R_L have any effect on the gain or input resistance? Why?

15-Convert the problem above to common base and find the current gain and voltage gain. What do you conclude?

References

1- "Microelectronics, Circuit Analysis and Design", D. A. Neamen, 3rd edition, Mc Graw Hill, N. Y. 2007.

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

3- "Microtronics, Circuits", A. S. Sedra, K. C. Smith, 5th edition, Oxford University, Press N. Y., 2004.

4- "Device Electronic for Integrated Circuits", R. S. Muller, T. L. Kamins, 3rd edition, Wiley N. Y., 2003.

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

6- Electronic Devices and Circuits", T. F. Bogart, 2nd edition, Maxwell Macmillan. N. Y., 1991.

7- "Electronic Devices", M. S. Said, H. Fikri, H. Haddara, Al-shorouk, Cairo, 1982.