

Bipolar Junction Transistor (BJT): invented by Schokley in 1951

Introduction:

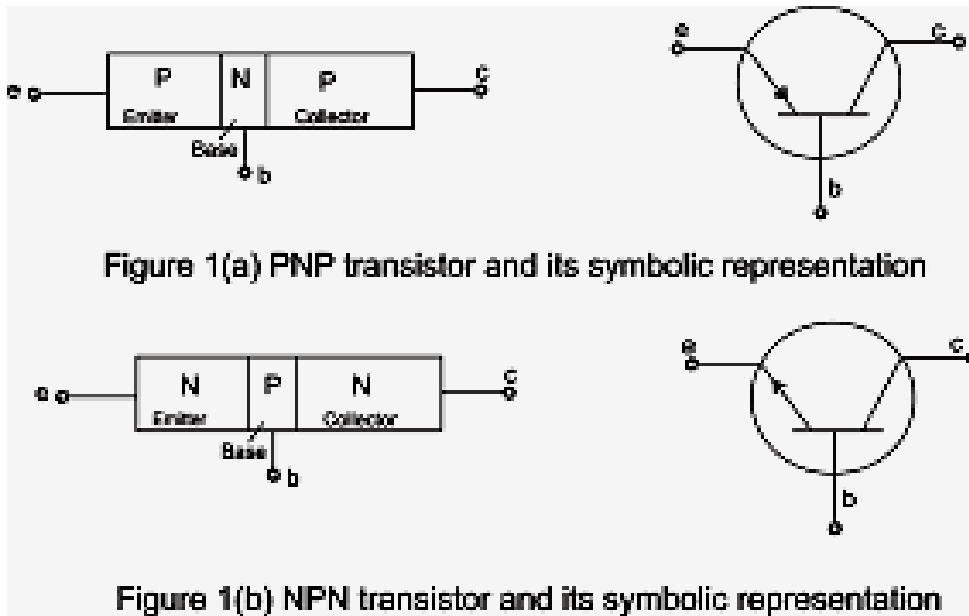


Figure 1(a) PNP transistor and its symbolic representation

Figure 1(b) NPN transistor and its symbolic representation

Bipolar junction transistor can be of two types: p-n-p and n-p-n transistor. If a layer of n-type material is sandwiched between two p-layers then the transistor is referred to as p-n-p type. On the other hand, if a layer of p-type material is sandwiched between two n-layers then the transistor is called n-p-n type.

The central layer is made very thin compared with the outer layers and is called base (B). The doping of base layer is also made considerably lower than that of the outer layers.

The outer layers are called emitter (E) and Collector (C). Their electrical conductivities are made different by adding different amount of impurities. Moreover, in commercial transistors, the area of the collector-base (C-B) junction is made considerably larger than the emitter-base (E-B) junction. This is so because in most cases collector junction has to handle more power than the emitter. So if collector and emitter are interchanged normal transistor action can't be obtained.

For normal operation, E-B junction is forward biased and the C-B junction is reverse biased.

The emitter in p-n-p transistor injects holes into the base and these holes are finally collected by the collector.

Since the E-B junction is forward biased, its dynamic resistance is small and since the C-B junction is reverse biased, its dynamic resistance is large. Thus it is found that an almost same current passes from a low resistor input circuit to a high resistor output circuit. The term transistor has been derived from the words ‘transfer resistor’.

Transistor operation and current components

Operation of p-n-p Transistor:

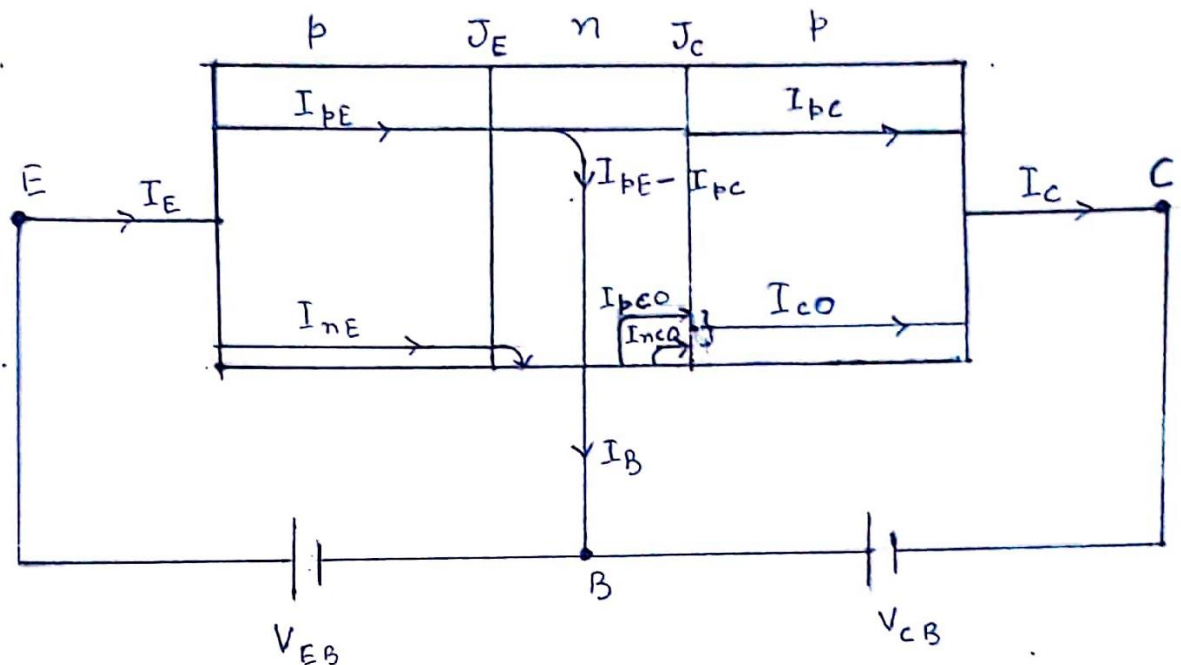


Fig.2 Current components in a p-n-p Transistor with emitter junction forward biased and collector junction reverse biased

To understand the basic operation of transistor, we consider a p-n-p transistor with its E-B junction (J_E) forward biased and C-B junction (J_C) reverse biased as shown in Fig.2

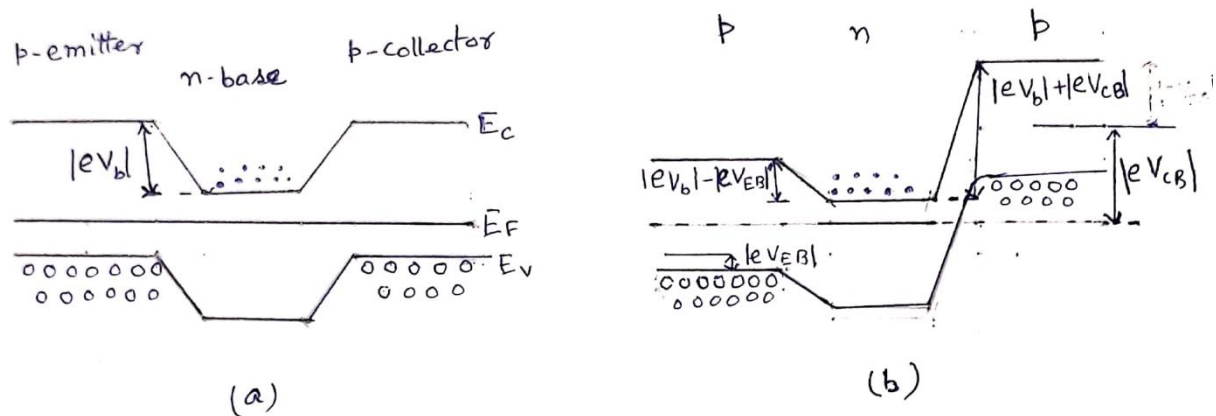


Fig.3 Energy band diagram of a p-n-p transistor (a) unbiased (b) biased

The forward biasing voltage V_{EB} reduces the emitter–base potential. This permits diffusion of holes from the emitter to the base and diffusion of electrons from base to emitter. These two flows constitute the emitter current as $I_E = I_{pE} + I_{nE}$, where I_{pE} is the emitter current due to holes moving from the emitter to the base and I_{nE} is the emitter current due to electrons moving from the base to the emitter. Usually, doping of emitter is much higher than that of base. So, $I_{pE} \gg I_{nE}$ and the emitter current is almost due to holes only. The injected holes diffuse through the base region towards the collector junction. While diffusing through the base few holes (1-3%) are lost due to their recombination with the electrons in the base regions.

The holes reaching the collector junction fall down the potential barrier and are collected by the collector. This gives rise to the component I_{pC} of the collector current. I_{pC} is slightly less than the I_{pE} and the difference ($I_{pE} - I_{pC}$) constitutes a part of the base current I_B .

The base current is due to flow of electrons from battery to the base to maintain the charge neutrality of the base region.

Since the collector junction is reverse biased there is a small collector current even with the emitter open. This current consisting of two components: I_{nCO} due to minority electrons flowing from the p-type collector to the n-type base and I_{pCO} due to minority holes moving from the n-type base to the p-type collector. The resultant current ($I_{nCO} + I_{pCO}$) is denoted by I_{CO} and is called leakage current or reverse collector saturation current with emitter open. I_{CO} is very much temperature sensitive. Thus we have the relations.

$$I_C = I_{pC} + I_{CO}$$

and $I_E = I_B + I_C$

Operation of n-p-n transistor

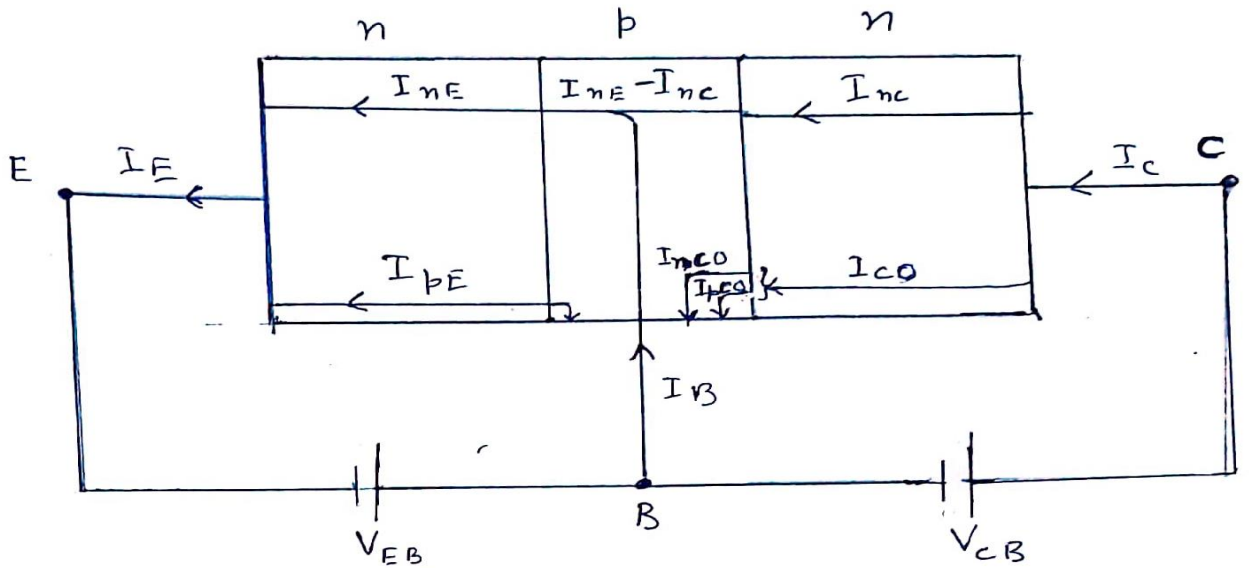


Fig.4 Current components in a n-p-n Transistor with emitter junction forward biased and collector junction reverse biased

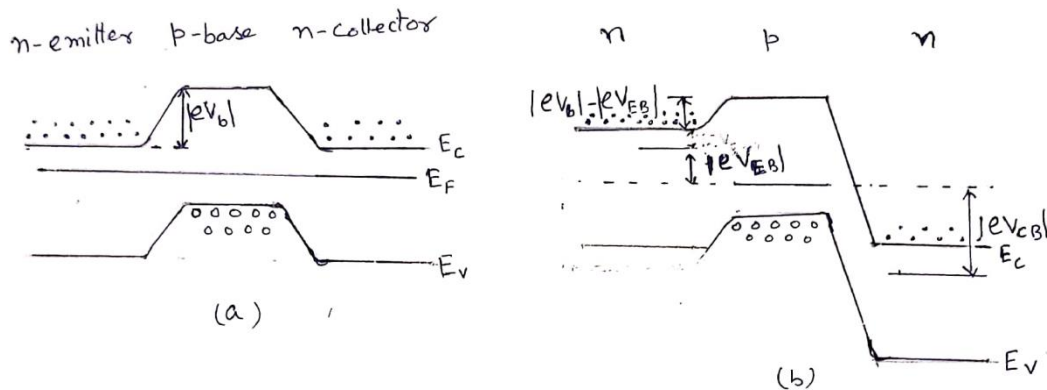


Fig.5 Energy band diagram of a n-p-n transistor (a) unbiased (b) biased

Here, the majority carrier electrons from the emitter are injected into the base and majority carrier holes from the base are injected to the emitter region. These two constitute the emitter current as

$$I_E = I_{nE} + I_{pE}$$

Since the doping of the emitter region is much higher than that of the base, $I_{nE} \gg I_{pE}$ and we have $I_E \approx I_{nE}$.

Thus the emitter current is almost entirely due to the electrons moving from the emitter to the base. The injected electrons diffuse through the base towards the collector junction. A few of the injected electrons are lost due to the recombination with the holes in the base. The electrons reaching the collector junction are collected by the collector and

give rise to the current component I_{nC} . The difference ($I_{nE} - I_{nC}$) constitutes a part of the base current. Since the collector junction is reverse biased there is a reverse saturation current, I_{CO} through this junction even when the emitter is open. It consists of two parts: I_{nC0} , due to movement of minority electrons from the base to the collector, and I_{pC0} , due to the movement of minority holes from the collector to the base. Thus:

$$I_C = I_{nC} + I_{CO}$$

and $I_E = I_B + I_C$

Three Modes of Connection of a Transistor:

There are three different configurations or modes of connection namely, common base (CB), common emitter and common collector. The configuration in which the base terminal is common to both the input and output circuits is known as the common base (CB) or grounded base mode. A p-n-p transistor connected in CB mode is shown in Fig 6(a). In common emitter or grounded emitter mode, emitter is common to both the input and output circuits. Fig 6(b) shows a p-n-p transistor in CE mode. When the collector terminal is common to both the input and the output circuits, the transistor is said to be in common collector (CC) mode or grounded collector mode. CC mode of a p-n-p transistor is shown in Fig.6(c).

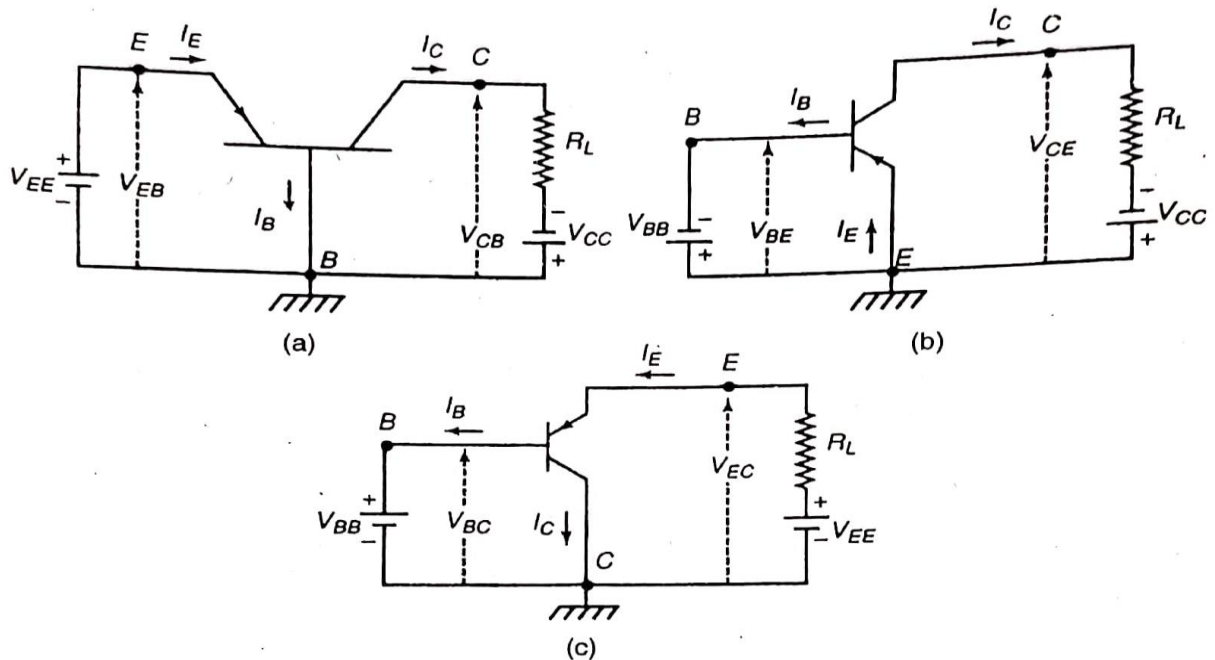


Fig. 6 p-n-p transistor connected in (a) CB, (b) CE and (c) CC mode

n-p-n transistor in different modes are shown in Fig. 7. These are obtained simply by changing the polarities of the batteries in Fig.6.

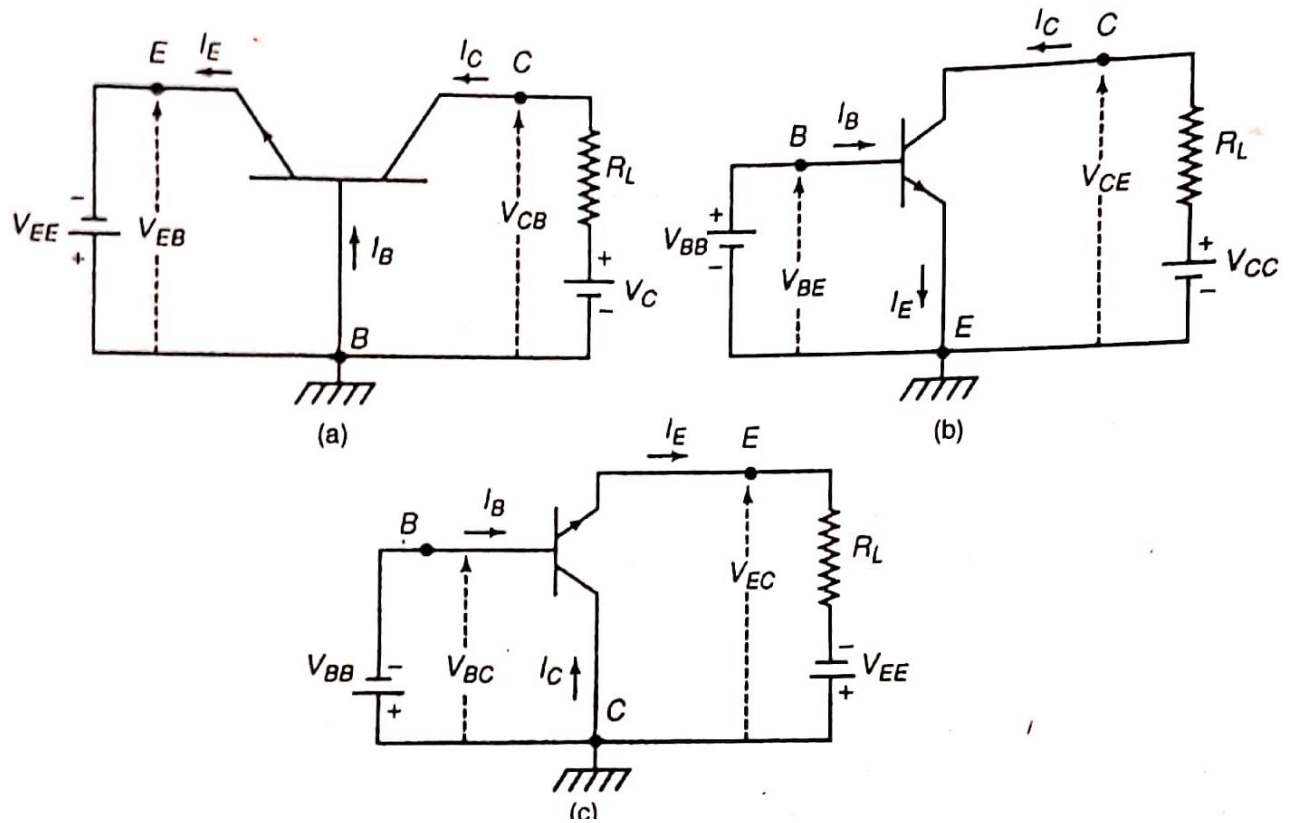


Fig. 7 n-p-n transistor connected in (a) CB, (b) CE and (c) CC mode

Transistor Alpha (α) and Beta (β):

The collector current I_C has two components (i) a part of the emitter current reaching the collector and(ii) a small component due to the reverse saturation current I_{CO} . Thus

$$I_C = \alpha I_E + I_{CO} \quad (1)$$

Here α is called transistor alpha. It represents the fraction of emitter current that can reach the collector. Usually I_{CO} is much smaller than I_C and we can write

$$I_C \approx \alpha I_E$$

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

In d.c. mode of operation we define d.c. alpha $\alpha_{d.c} = \frac{I_C}{I_E}$

Thus, $\alpha = \alpha_{d.c}$.

In a.c. mode of operation we define the small signal short circuit common base current gain ($\alpha_{a.c}$). It is defined as the ratio of the change in the collector current to the emitter current at constant collector to base voltage, i.e.,

$$\alpha_{a.c.} = \frac{\Delta I_C}{\Delta I_E} \Big|_{V_{CB}=\text{constant}}$$

For a good transistor $\alpha \approx \alpha_{d.c.} \approx \alpha_{a.c.}$ and must be ≈ 1 . Usually, in good transistors α ranges from 0.98 to 0.988. However, α is not constant and varies with I_E , V_{CB} and temperature.

In CE mode, input current is base current I_B and output current is I_C . In this case the performance of a transistor is expressed in terms of a parameter called transistor beta (β). Using the relation: $I_E = I_B + I_C$, we get from equation (1)

$$\begin{aligned} I_C &= \alpha(I_B + I_C) + I_{CO} \\ \therefore (1 - \alpha)I_C &= \alpha I_B + I_{CO} \\ \therefore I_C &= \frac{\alpha}{1 - \alpha} I_B + \frac{1}{1 - \alpha} I_{CO} \end{aligned}$$

$$I_C = \beta I_B + (1 + \beta) I_{CO} \quad (2)$$

where $\beta = \frac{1}{1 - \alpha}$ is called transistor beta. In d.c. mode of operation we define $\beta_{d.c.}$ or $h_{FE} = \frac{I_C}{I_B}$.

Usually $I_B \gg I_{CO}$ and hence, $I_C \approx \beta I_B$. Thus $\beta_{d.c.} \approx \beta$.

In commercial transistors $\beta_{d.c.}$ or h_{FE} is in the range 50 to 400. Moreover, h_{FE} depends on I_C and temperature.

For operation with a.c., one defines the small signal short circuit CE current gain $\beta_{a.c.}$ or h_{fe} . It is defined as

$$\beta_{a.c.} \text{ or } h_{fe} = \frac{\Delta I_C}{\Delta I_B} \Big|_{V_{CE}=\text{constant}}$$

h_{fe} is found to be larger than h_{FE} for small I_C and h_{fe} is smaller than h_{FE} for large I_C . However, in most cases h_{fe} differs from h_{FE} by less than 20%.

Relation between α and β :

By definition, $\alpha_{d.c} = \frac{I_C}{I_E}$

and $\beta_{d.c} = \frac{I_C}{I_B}$

Now, $I_E = I_B + I_C$

Dividing both sides by I_C , we get

$$\frac{1}{\alpha_{d.c}} = \frac{1}{\beta_{d.c}} + 1$$

$$\text{or, } \frac{1}{\beta_{d.c}} = \frac{1}{\alpha_{d.c}} - 1$$

$$\text{or, } \beta_{d.c} = \frac{\alpha_{d.c}}{1 - \alpha_{d.c}}$$

Now since $\alpha_{d.c} = \alpha$ and $\beta_{d.c} \approx \beta$, we can write

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

Variation of α with emitter current and frequency

Variation of α with I_E :

Transistor α is low at very low emitter current I_E . It increases with I_E , passes through a maximum and then continues to fall for higher I_E . At very low emitter currents, the recombination in the emitter depletion region may be large compared with the current due to diffusion of minority carriers in the base region. This makes α low. With increasing total emitter current the diffusion current begins to dominate over the recombination and the leakage current. It enhances α . However, at very high emitter currents the injected minority carrier concentration approaches the majority carrier concentration in the base. Due to presence of large number of carriers conductivity in the base region increases and base current also increases i.e., majority carrier concentration in the base region increases. Presence of large number of such carriers decreases the life time of minority carriers. Hence α becomes low at high currents.

Variation of α with frequency:

The carriers injected from the emitter into the base take some time to cross the base. It is called transit time. At low frequencies the time period of signal is large compared to this transit time and α remains independent of frequency. But at high frequencies the time period of the signal may become comparable or even smaller than the transit time. Then a fraction of the carriers injected into the base will fail to reach the collector before the polarity of the signal changes. Therefore, some injected carriers are trapped in the base and rate of recombination increases. Hence, α decreases. The maximum usable frequency is described in terms of alpha cut-off frequency (f_α). It is the frequency at which f_α decreases to $\frac{1}{\sqrt{2}}$ of its maximum value at low frequency.

Since electron mobility is far greater than hole mobility an n-p-n transistor can be operated over higher frequencies than a p-n-p transistor.