The breakdown voltage is calculated using $E_{BR} = E_{BHE}$ in (1.24) to give

$$BV = \frac{q(N_A + N_D)E_b}{2qN_A N_D}$$

$$= \frac{1.04 \times 10^{-12} \times 15 \times 10^{15}}{2 \times 1.6 \times 10^{-19} \times 5 \times 10^{13} \times 10^{16}} \times 9 \times 10^{14} V$$

$$= 88 V$$

1.3 Large-Signal Behavior of Bipolar Transistors

In this section, the large-signal or dc behavior of bipolar transistors is considered. Large-signal models are developed for the calculation of total currents and voltages in transistor circuits, and such effects as breakdown voltage limitations, which are usually not included in models, are also considered. Second-order effects, such as current-gain variation with collector current and Early voltage, can be important in many circuits and are treated in detail.

The sign conventions used for bipolar transistor currents and voltages are shown in Fig. 1.5. All bias currents for both npn and pnp transistors are assumed positive going into the device.

1.3.1 Large-Signal Models in the Forward-Active Region

A typical npn planar bipolar transistor structure is shown in Fig. 1.6a, where collector, base, and emitter are labeled C, B, and E, respectively. The method of fabricating such transistor structures is described in Chapter 2. It is shown there that the impurity doping density in the base and the emitter of such a transistor is not constant but varies with distance from the top surface. However, many of the characteristics of such a device can be predicted by analyzing the idealized transistor structure shown in Fig. 1.6b. In this structure the base and emitter doping densities are assumed constant, and this is sometimes called a "uniform-base" transistor. Where possible in the following analyses, the equations for the uniform-base analysis are expressed in a form that applies also to nonuniform base transistors.

A cross section $AA'$ is taken through the device of Fig. 1.6b and carrier concentrations along this section are plotted in Fig. 1.6c. Hole concentrations are denoted by $p$ and electron concentrations by $n$ with subscripts $p$ or $n$ representing p-type or n-type regions. The n-type emitter and collector regions are distinguished by subscripts $E$ and $C$, respectively. The carrier concentrations shown in Fig. 1.6c apply to a device in the forward-active region. That is, the base–emitter junction is forward biased and the base–collector junction is reverse biased. The minority-
carrier concentrations in the base at the edges of the depletion regions can be
calculated from a Boltzmann approximation to the Fermi-Dirac distribution
function to give

\[ n_p(0) = n_p \exp \left( \frac{-V_{BE}}{V_T} \right) \]  
\[ (1.27) \]

\[ n_p(W_b) = n_p \exp \left( \frac{V_{BC}}{V_T} \right) = 0 \]  
\[ (1.28) \]

where \( W_b \) is the width of the base from the base–emitter depletion layer edge to the
base–collector depletion layer edge and \( n_p \) is the equilibrium concentration
of electrons in the base. Note that \( V_{BC} \) is negative for an npn transistor in the forward-active
region and thus \( n_p(W_b) \) is very small. Low-level injection conditions are
assumed in the derivation of (1.27) and (1.28). This means that the minority-carrier
concentrations are always assumed much smaller than the majority-carrier
concentration.

If recombination of holes and electrons in the base is small, it can be shown
that the minority-carrier concentration \( n_p(x) \) in the base varies linearly with
distance. Thus a straight line can be drawn joining the concentrations at \( x = 0 \) and
\( x = W_b \) in Fig. 1.6c.

For charge neutrality in the base, it is necessary that

\[ N_A + n_p(x) = p_p(x) \]  
\[ (1.29) \]

and thus

\[ p_p(x) - n_p(x) = N_A \]  
\[ (1.30) \]

where \( p_p(x) \) is the hole concentration in the base and \( N_A \) is the base doping density
that is assumed constant. Equation 1.30 indicates that the hole and electron
concentrations are separated by a constant amount and thus \( p_p(x) \) also varies
linearly with distance.

Collector current is produced by minority-carrier electrons in the base
diffusing in the direction of the concentration gradient and being swept across the
collector-base depletion region by the field existing there. The diffusion current
density due to electrons in the base is

\[ J_e = qD_e \frac{dn_p(x)}{dx} \]  
\[ (1.31) \]

where \( D_e \) is the diffusion constant for electrons. From Fig. 1.6c

\[ J_e = -qD_e \frac{n_p(0)}{W_b} \]  
\[ (1.32) \]

If \( I_c \) is the collector current and is taken as positive flowing into the collector, it
follows from (1.32) that

\[ I_c = qAD_e \frac{n_p(0)}{W_b} \]  
\[ (1.33) \]

where \( A \) is the cross-sectional area of the emitter. Substitution of (1.27) into (1.33)
gives

\[ I_c = \frac{qAD_e n_p}{W_b} \exp \left( \frac{V_{BE}}{V_T} \right) \]  
\[ (1.34) \]

\[ = I_s \exp \left( \frac{V_{BE}}{V_T} \right) \]  
\[ (1.35) \]

where

\[ I_s = \frac{qAD_e n_p}{W_b} \]  
\[ (1.36) \]

and \( I_s \) is a constant used to describe the transfer characteristic of the transistor in the
forward-active region. Equation 1.36 can be expressed in terms of the base
doping density by noting that \( n_p = \frac{n_i}{N_A} \) (see Chapter 2)

\[ n_p = \frac{n_i}{N_A} \]  
\[ (1.37) \]

and substitution of (1.37) in (1.36) gives

\[ I_s = \frac{qAD_e n_i^2}{W_b N_A} = \frac{qAD_e n_i^2}{Q_b} \]  
\[ (1.38) \]

where \( Q_b = W_b N_A \) is the number of doping atoms in the base per unit area of
the emitter and \( n_i \) is the intrinsic carrier concentration in silicon. In this form (1.38)
applies to both uniform and nonuniform base transistors and \( D_e \) has been replaced by
\( D_x \), which is an average effective value of the electron diffusion constant: in the
base. This is necessary for nonuniform base devices because the diffusion constant
is a function of impurity concentration. Typical values of \( I_s \) as given by (1.38) are
\( 10^{-14} - 10^{-15} \) A.

Equation 1.35 gives the collector current as a function of base–emitter voltage.
The base current \( I_s \) is also an important parameter and, at moderate current levels,
consists of two major components. One of these \( (I_{31}) \) represents recombination
of holes and electrons in the base and is proportional to the minority-carrier charge
\( Q_s \) in the base. From Fig. 1.6c, the minority-carrier charge in the base is

\[ Q_s = \frac{1}{2} n_p(0) W_b q \]  
\[ (1.39) \]

and we have

\[ I_{31} = \frac{Q_s}{\tau_s} = \frac{1}{2} \frac{n_p(0) W_b q}{\tau_s} \]  
\[ (1.40) \]
where $t_b$ is the minority-carrier lifetime in the base. $I_{B1}$ represents a flow of majority holes from the base lead into the base region. Substitution of (1.27) in (1.40) gives

$$I_{B1} = \frac{1}{2} \frac{n_{pe} W_q q A}{r_b} \exp \frac{V_{BE}}{V_T} \tag{1.41}$$

The second major component of base current (usually the dominant one in integrated-circuit $n$-type devices) is due to injection of holes from the base into the emitter. This current component depends on the gradient of minority carrier holes in the emitter and is

$$I_{B2} = \frac{q AD_e}{L_e} p_{ae}(0) \tag{1.42}$$

where $D_e$ is the diffusion constant for holes and $L_e$ is the diffusion length (assumed small) for holes in the emitter. $p_{ae}(0)$ is the concentration of holes in the emitter at the edge of the depletion region and is

$$p_{ae}(0) = p_{ae} \exp \frac{V_{BE}}{V_T} \tag{1.43}$$

If $N_D$ is the donor atom concentration in the emitter (assumed constant) then

$$p_{ae} \approx n_i^2 \frac{N_D}{N_p} \tag{1.44}$$

The emitter is deliberately doped much more heavily than the base, making $N_p$ large and $p_{ae}$ small, so that the base-current component, $I_{B2}$, is minimized.

Substitution of (1.43) and (1.44) in (1.42) gives

$$I_{B2} = \frac{q AD_e}{L_e} \frac{n_i^2 N_D}{N_p} \exp \frac{V_{BE}}{V_T} \tag{1.45}$$

The total base current, $I_B$, is the sum of $I_{B1}$ and $I_{B2}$

$$I_B = I_{B1} + I_{B2} = \left( \frac{1}{2} \frac{n_{pe} W_q q A}{r_b} + \frac{q AD_e n_i^2}{L_e N_D} \right) \exp \frac{V_{BE}}{V_T} \tag{1.46}$$

Although this equation was derived assuming uniform base and emitter doping, it gives the correct functional dependence of $I_B$ on device parameters for practical double-diffused nonuniform base devices. Second-order components of $I_B$, which are important at low current levels, are considered later.

Since $I_c$ in (1.35) and $I_B$ in (1.46) are both proportional to $\exp(V_{BE}/V_T)$ in this analysis, the base current can be expressed in terms of collector current as

$$I_B = \frac{I_c}{\beta_f} \tag{1.47}$$

where $\beta_f$ is the forward current gain. An expression for $\beta_f$ can be calculated by substituting (1.34) and (1.46) in (1.47) to give

$$\beta_f = \frac{q AD_e n_i^2}{W_q} \left( \frac{1}{2} \frac{n_{pe} W_q q A}{r_b} + \frac{q AD_e n_i^2}{L_e N_D} \right) \exp \frac{V_{BE}}{V_T} \tag{1.48}$$

where (1.37) has been substituted for $n_{pe}$.

Equation (1.48) shows that $\beta_f$ is maximized by minimizing the base width $W_q$ and maximizing the ratio of emitter to base doping densities $N_p/N_D$. Typical values of $\beta_f$ for $n$-type transistors in integrated circuits are 50 to 100, whereas lateral $pnp$ transistors (to be described in Chapter 2) have values 10 to 100. Finally, the emitter current is

$$I_E = -(I_C + I_B) = -\left( I_C + \frac{I_c}{\beta_f} \right) = -\frac{I_c}{\alpha_f} \tag{1.49}$$

where

$$\alpha_f = \frac{\beta_f}{1 + \beta_f} \tag{1.50}$$

The value of $\alpha_f$ can be expressed in terms of device parameters by substituting (1.48) in (1.50) to obtain

$$\alpha_f = \frac{1}{1 + \beta_f} = \frac{1}{1 + \frac{W_q}{2 \tau_b D_e} + \frac{D_e \frac{W_q N_D}{N_p}}{L_e \frac{N_D}{N_p}} \approx \alpha_f \gamma} \tag{1.51}$$

where

$$\alpha_f = \frac{1}{1 + \frac{W_q}{2 \tau_b D_e} + \frac{D_e \frac{W_q N_D}{N_p}}{L_e \frac{N_D}{N_p}}} \tag{1.51a}$$

$$\gamma = \frac{1}{1 + \frac{D_e \frac{W_q N_D}{N_p}}{L_e \frac{N_D}{N_p}}} \tag{1.51b}$$

The validity of (1.51) depends on $W_q/2 \tau_b D_e < 1$ and $(D_e/L_e)(W_q/N_p) < 1$, and this is always true if $\beta_f$ is large [see (1.48)]. The term $\gamma$ in (1.51) is called the emitter injection efficiency and is equal to the ratio of the electron current ($n$-type transistor) injected into the base from the emitter to the total hole and electron current crossing the base–emitter junction. Ideally $\gamma \rightarrow 1$, and this is achieved by making $N_p/N_D$ large and $W_q$ small. In that case very little reverse injection occurs from base to emitter.

The term $\alpha_f$ in (1.51) is called the base transport factor and represents the fraction of carriers injected into the base (from the emitter) that reach the collector. Ideally $\alpha_f \rightarrow 1$ and this is achieved by making $W_q$ small. It is evident from the above development that fabrication changes that cause $\alpha_f$ and $\gamma$ to approach unity also maximize the value of $\beta_f$ of the transistor.