1 III-V Heterojunction Bipolar Transistors

Mark Lundstrom
School of Electrical Engineering
Purdue University
West Lafayette, Indiana, USA

1.1 Introduction

The heterojuction bipolar transistor (HBT) is a device concept almost as old as the bipolar transistor itself. The basic idea was described by Shockley in his patent for the junction bipolar transistor [Sho48]. Some years later, the concept was explored by Kroemer who introduced the notion of *quasi-electric fields* which arise from the variation of material composition in a semiconductor heterostructure [Kro57]. Early HBT's suffered from the primitive fabrication technology for heterojunctions, but with the development of liquid phase epitaxy (LPE) in the 1970's device performance improved (e.g. [Dum72, Bai80, Ben81]). High-performance devices and circuits, however, required control of doping and composition on a nanometer scale as well as uniformity across the wafer. During the early 1980's, new epitaxial growth techniques capable of such control and uniformity became available. Kroemer pointed out that with new growth techniques such as molecular beam epitaxy (MBE) and metal-organic chemical vapor deposition (MOCVD), the potential of HBT's could be realized [Kro82]. A vigorous, world-wide research program directed at realizing the digital and millimeter wave potential of HBT's was initiated in the early 1980's and continues today. When Kroemer pointed out the potential for HBT's in 1982, the highest reported f_T was 1.6 GHz [Ben81]. As I write this in 1990, AlGaAs/GaAs HBT's have achieved an f_T of 171 GHz [Ish90a] and digital switching speeds of 1.9 ps have been reported [Ish88b].

Quantity	*Expression*
Bandgap	$1.424 + 1.247x$ $(0 \leq x \leq 0.45)$
	$1.900 + 0.125x +$
	$+ 0.143x^2$ $(0.45 \leq x \leq 1.0)$
Conduction band discontinuity	$\Delta E_{CO} = 0.65\Delta E_G (0.0 \leq x \leq 0.45)$
Density-of states electron mass:	
Γ valley	$0.067 + 0.083x$
X valley	$0.85 - 0.14x$
Λ valley	$0.56 + 0.1x$
Valence band mass:	
light hole	$0.087 + 0.063x$
heavy hole	$0.62 + 0.14x$
Static dielectric constant	$13.18 - 3.12x$

Table 1.1. Selected material properties of $Al_xGa_{1-x}As$. From [Ada85].

1.3 HBT Fundamentals

The wide bandgap emitter is the central feature of a heterojunction bipolar transistor. The wide bandgap suppresses the back injection of holes into the emitter, so the base can be heavily doped without sacrificing common emitter current gain. The heavy base doping lowers the base resistance, which reduces RC charging times and improves device speed. Another benefit is that it raises the output resistance of the transistor. The wide bandgap also permits a light emitter doping, which lowers the emitter-base junction capacitance. This section introduces the basic concepts relevant to heterojunction bipolar transistors. We'll focus on the simplest HBT, a bipolar transistor with a wide bandgap emitter. More sophisticated compositional profiles, such as graded bases and wide bandgap collectors are possible and will be discussed later.

The Wide Gap Emitter Bipolar Transistor

Displayed in Fig. 1.4a is the energy band diagram of a heterojunction bipolar transistor. This transistor has an $Al_{0.3}Ga_{0.7}As$ N-type emitter, and a GaAs base and collector. By compositionally grading from $Al_{0.3}Ga_{0.7}As$ to GaAs over a distance of about 300Å, the conduction band spike at the emitter-base heterojunction has been removed. When band spikes are absent, the analysis of an HBT is especially simple.

Figure 1.4b shows a simple, one-dimensional sketch of the transistor and defines some carrier fluxes. In the normal, active mode of operation, conventional p-n junction theory gives the electron current injected from the emitter into the base as

$$F_{EN} = D_B \frac{\Delta n(X_{PE})}{W_B} = \frac{D_B n_{iB}^2}{N_{AB} W_B}\left(e^{qV_{BE}/k_B T} - 1\right) = \frac{I_C}{qA_E}, \qquad (1.11)$$

where D_B is the diffusion coefficient of minority carrier electrons in the base, A_E is the emitter area, and all other terms have their usual meanings. Similarly, the hole flux injected into the emitter is

$$F_{EP} = D_E \frac{\Delta p(X_{NE})}{W_E} = \frac{D_E n_{iB}^2}{N_{DE} W_E}\left(e^{qV_{BE}/k_B T} - 1\right). \qquad (1.12)$$

(To keep the algebra simple, we have assumed that the quasi-neutral emitter and base regions are short in comparison to a minority carrier diffusion length.) From these two expressions, we evaluate the emitter injection efficiency as

$$\gamma = \frac{F_{EN}}{F_{EN} + F_{EP}} \qquad (1.13)$$

and estimate the common emitter current gain as

$$\beta_{dc} = \frac{\gamma}{1-\gamma} = \frac{F_{EN}}{F_{EP}} = \frac{(D_B / W_B) N_{DE} n_{iB}^2}{(D_E / W_E) N_{AB} n_{iE}^2}. \qquad (1.14)$$

Since the ratio, n_{iB}^2 / n_{iE}^2 goes as $\exp[(E_{GE}-E_{GB})k_BT]$, we conclude that

$$\beta_{dc} \sim \frac{N_{DE}}{N_{AB}} e^{\Delta E_G / k_B T}. \qquad (1.15)$$

where ΔE_G is the difference in bandgap between the emitter and base. For a conventional bipolar transistor, $E_{GE} = E_{GB}$, so we must make $N_{DE} \gg N_{AB}$ to insure high gain. For an HBT, however, $\exp[\Delta E_G/k_BT]$ is typically very large, so high gain can be achieved even when the doping is inverted so that $N_{AB} \gg N_{DE}$. In a typical $Al_{0.3}Ga_{0.7}As$:GaAs HBT, for example $\exp[\Delta E_G/k_BT]$ is about 10^5 at room temperature. The result is that that base doping may be selected to minimize the base resistance; it has little effect on the dc gain.

Effects of Recombination

Equation (1.14) greatly overestimates the gain because carrier recombination was not considered. Let's consider first the effects of recombination in the quasi-neutral base. The base transport factor, α_T, is

$$\alpha_T = 1 - F_{BR} / F_{EN}, \tag{1.16}$$

where F_{BR} is the flux of carriers recombining in the base. A good estimate for F_{BR} is

$$F_{BR} = \frac{\Delta n(X_{PE}) W_B}{2\tau_B} = \frac{n_{iB}^2 W_B}{2N_{AB}\tau_B}\left(e^{qV_{BE}/k_BT} - 1\right), \tag{1.17}$$

where τ_B is the minority carrier lifetime in the base. With (1.11) and (1.17) inserted in (1.16) we find the base transport factor as

$$\alpha_T = 1 - \frac{W_B^2}{2L_B^2} = 1 - \frac{\tau_b}{\tau_B}, \tag{1.18}$$

where L_B is the minority carrier diffusion length in the base and $\tau_b = W_B^2 / 2D_B$ is the base transit time. The common emitter current gain is readily evaluated from

$$\beta_{dc} = \frac{\gamma\alpha_T}{1 - \gamma\alpha_T}, \tag{1.19}$$

but the wide gap emitter makes γ essentially one, so the base transport factor limits the gain. We conclude that $\beta_{dc} \simeq \alpha_T / (1 - \alpha_T) = 2L_B^2 / W_B - 1 \simeq \tau_B/\tau_b$. Using the typical device and material parameters for an AlGaAs/GaAs HBT as listed in Table 1.2, we estimate $\beta_{dc} \simeq 1{,}000$ which is a far higher gain than is typically observed. This treatment also predicts that the gain is independent of the collector current, which occurs for silicon bipolar transistors and for InP/InGaAs HBT's, but not for AlGaAs/GaAs HBT's.

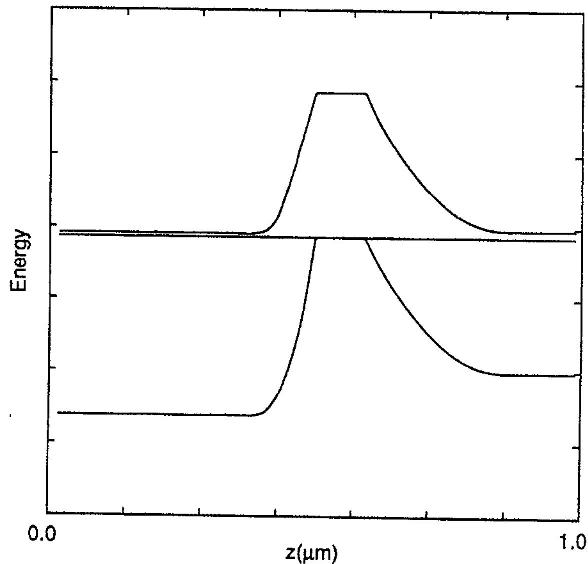


Figure 1.4a. *Energy band diagram for a graded-junction, $Al_{0.3}Ga_{0.7}As$:GaAs, N-p-n, heterojunction bipolar transistor. In the N-type emitter, the composition is linearly graded from x = 0 to x = 0.3 for a distance of 300Å beginning at the emitter-base junction.*

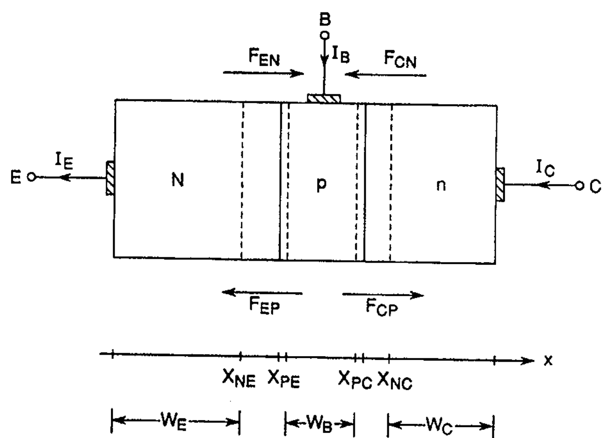


Figure 1.4b. *One-dimensional sketch of an HBT with particle fluxes and terminal currents defined.*

Quantity	Emitter	Base	Collector
doping(cm^{-3})	1.0 x 10^{17}	1.0 x 10^{19}	1.0 x 10^{17}
thickness(μm)	0.5	0.075	0.5
n_i(cm^{-3})	1.11x10^4	2.24x10^6	2.24x10^6
μ_n(cm^2/V-s)	2000	1000	6000
τ(ns)	1	1	1

Table 1.2. *Typical material parameters for an Npn $Al_{0.3}Ga_{0.7}As$:GaAs HBT operating at room temperature. τ is the minority carrier lifetime.*

When carriers recombine in the emitter-base space-charge region (SCR,) the emitter injection efficiency becomes

$$\gamma = \frac{F_{EN}}{F_{EN} + F_{EP} + F_{SCR}}. \qquad (1.20)$$

The flux of carriers recombining in the emitter-base space-charge region is

$$F^b{}_{SCR} = \int_{X_{NE}}^{X_{PE}} R(x)dx = \hat{R}W_{eff}, \qquad (1.21)$$

where $\hat{R}$ is the peak value of the recombination rate. The recombination rate is sharply peaked where n ≃ p and has an effective width of W_{eff}. From conventional diode theory we estimate the peak recombination rate as [Sah57]

$$\hat{R} = \frac{\hat{n}_i}{2\tau_o}\left(e^{qV_{BE}/2k_BT} - 1\right). \qquad (1.22)$$

Since $\hat{R}$ occurs in the compositionally graded SCR, $\hat{n}_i$ refers to the value of n_i where the recombination rate is the largest. The lifetime, τ_o, is assumed to be equal for electrons and holes. The effective width over which recombination occurs is given by [Sah57]

$$W_{eff} = \frac{k_BT/q}{\hat{\xi}}, \qquad (1.23)$$

where $\hat{\xi}$ the electric field where the recombination rate peaks. From (1.21)-(1.23) we find the carrier flux associated with recombination in the emitter-base space charge region to be

$$F^b{}_{SCR} = \frac{\hat{n}_i}{2\tau_o}\frac{k_BT/q}{\hat{\xi}}(e^{qV_{BE}/2k_BT} - 1). \qquad (1.24)$$

Recombination in the emitter-base space-charging region produces a current-dependent emitter injection efficiency which gives β a current dependence, but the treatment as outlined above still over-estimates β. It also suggests that the gain is independent of device size, but AlGaAs/GaAs HBT's commonly display a size dependence to β. Small devices have lower gain than do larger devices [Nak85].

Figure 1.5 shows a more realistic structure for an HBT. Note that the emitter-base junction is typically defined by mesa etching which exposes the emitter-base junction. A GaAs surface typically contain a very high density of recombination centers, so it is recombination at the exposed surface which dominates [Tiw90]. Estimating the surface recombination current of the forward-biased emitter-base junction by arguments similar to those for bulk recombination, we find

$$I^P = \hat{R}_s L_s P, \qquad (1.25)$$

where I^P is the total current, $\hat{R}_s$ is the peak surface recombination rate, L_s is the effective width over which surface recombination occurs, and P is the length of the perimeter of the emitter-base junction. By converting to a flux per unit area and using arguments like those leading to (1.24), we find [Hen78]

$$F^P{}_{SCR} = (\hat{n}_i s_o L_s)(P/A)\left(e^{qV_{BE}/2k_BT} - 1\right), \qquad (1.26)$$

where s_o is the recombination velocity of the surface. (Equation (1.26) applies under modest forward bias because for high biases, the ideality factor decreases from two and approaches one [Tiw89].)

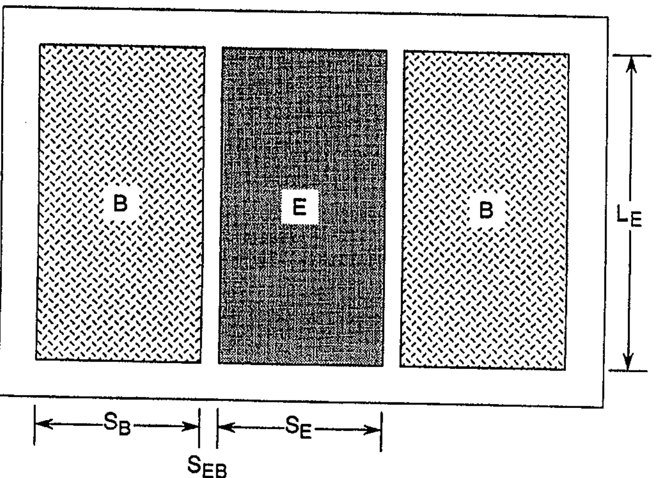
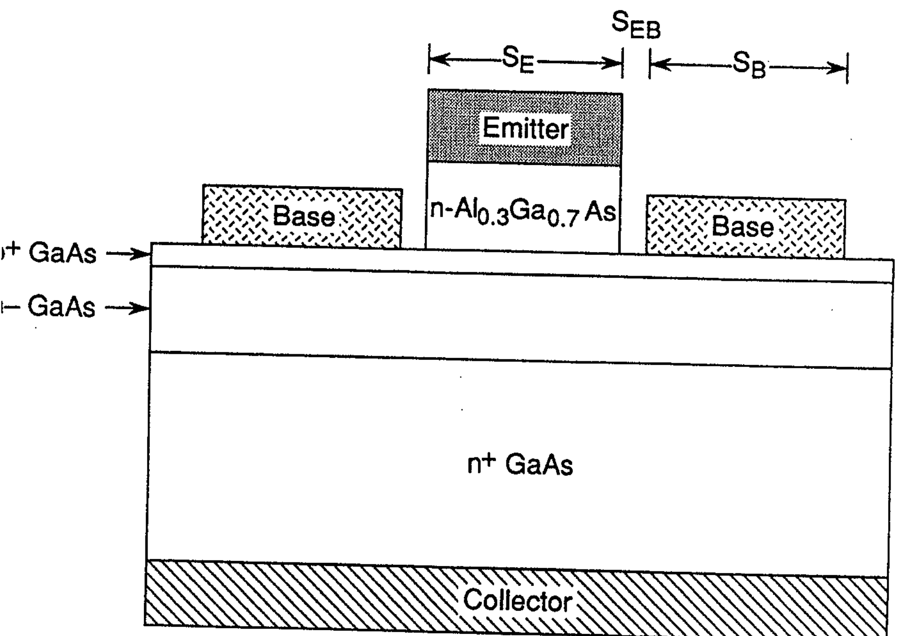


Figure 1.5a. *Top view showing the physical layout for a microwave HBT.*



Figur 1.5b. *Cross-sectional sketch showing the device structure of a typical HBT.*

Assuming that recombination in the perimeter SCR dominates, we evaluate the emitter injection efficiency from (1.20) using (1.11) and (1.26). Using the result to obtain an expression for β, we find.

$$\beta_{dc} = \frac{n_{iB}^2 D_B A}{\hat{n}_i s_o L_s N_{AB} W_B P} e^{qV_{BE}/2k_B T}, \tag{1.27}$$

which shows that the gain show a strong bias dependence. Equation (1.27) also illustrates the size dependence to β. Large devices have a large A/P ratio, which results in a higher gain. By noting the $J_C \simeq qF_{EN}$, we can express β as a function of collector current. The result is

$$\beta_{dc} = \left[\frac{A}{P}\frac{n_{iB}}{\hat{n}_i}\frac{1}{s_o L_s}\sqrt{\frac{D_B}{qN_{AB}W_B}}\right]\sqrt{J_C}. \tag{1.28}$$

Equation (1.28) does a good job of describing some important features of the β vs. I_C characteristic of AlGaAs/GaAs HBT's. It shows that β has a strong dependence on collector current; a log-log plot of β vs. I_C will show a slope of one-half. The current dependence also produces a difference in the dc and ac gains. Defining $\beta_{dc} = I_C / I_B$ and $\beta_{ac} = \partial I_C / \partial I_B$, one can readily show that

$$\beta_{ac} = 2\beta_{dc} \tag{1.29}$$

when recombination in the space-charge region dominates. Equation (1.28) also shows that the gain of small devices as the ratio A/P decreases.

The measured characteristics of AlGaAs/GaAs HBT's are well-described by (1.28). As illustrated in Fig. 1.6a, the log β_{dc} vs. log I_C characteristic does indeed display a slope of 1/2 under low and moderate biases. For higher biases, recombination in the quasi-neutral base becomes significant, but the perimeter recombination current is always substantial so β always displays a current dependence. Figure 1.6b shows that the size effect is strong. It is, however, possible to suppress perimeter recombination. One approach is to chemically treat

the exposed perimeter [San87, Chu90]; another is to leave a depleted layer of AlGaAs on the exposed base [Hay90].

Ebers-Moll Model for Graded-Junction HBT's

We have concentrated so far on describing important features of HBT's in the normal active region of operation. To describe device operation more generally, we need to develop an Ebers-Moll model for the HBT [Lee86, Lun86]. When compositional grading is employed to remove band spikes, the model can be derived using the same arguments that apply for homojunction transistors - we just need to use the appropriate value of n_i for the emitter, base, and collector regions. Referring again to Fig. 1.4b, we can express the terminal emitter and collector currents as

$$I_C = qA_E\left(F_{EN} - F_{CN} - F'_{BR} - F_{CP} - F_{CSCR}\right) \qquad (1.30a)$$

and

$$I_E = qA_E\left(F_{EN} - F_{CN} - F''_{BR} + F_{EP} + F_{ESCR}\right), \qquad (1.30b)$$

where F_{ESCR} and F_{CSCR} denote the recombination currents associated with the emitter-base and collector-base junctions respectively. As Fig. 1.5 illustrates, the emitter-base and base-collector junction areas may be quite different. In (1.30), A_E is the area of the emitter-base junctions; the difference in junction areas will be accounted for when the fluxes are defined.

The net flux of carriers injected into the base from the two junctions is

$$F_{EN} - F_{CN} = \frac{D_B n_{iB}^2}{W_B N_{AB}}\left[\left(e^{qV_{BE}/k_BT} - 1\right) - \frac{A_C}{A_E}\left(e^{qV_{BC}/k_BT} - 1\right)\right], \qquad (1.31)$$

which is a generalization of (1.11) for finite base-collector voltages. Similarly, the flux of holes injected from the base into the collector is

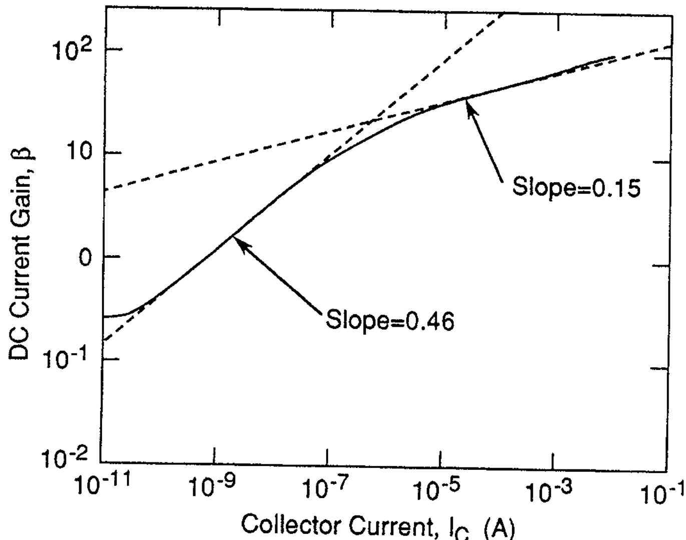


Figure 1.6a. *Measured β vs. I_C characteristic of a typical $Al_{0.3}Ga_{0.7}As$:GaAs HBT. (From [Kim90].)*

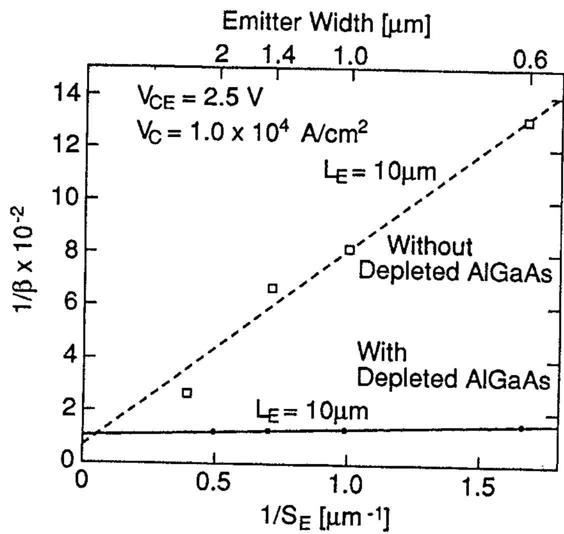


Figure 1.6b. *Experimental results demonstrating the size effect, which causes β to decrease for small AlGaAs/GaAs HBT"s. (From [Hay90].)*

$$F_{CP} = \frac{D_C n_{iC}^2}{W_C N_{DC}} \frac{A_C}{A_E}\left(e^{qV_{BC}/k_BT} - 1\right). \tag{1.32}$$

The flux of carriers recombining in the base, due to injection from the emitter-base junction, is

$$F'_{BR} = \frac{W_B n_{iB}^2}{2\tau_B N_{AB}}\left(e^{qV_{BE}/k_BT} - 1\right) \tag{1.33}$$

and that due to electrons injected from the base-collector junction is

$$F''_{BR} = \frac{W_B n_{iB}^2}{2\tau_B N_{AB}} \frac{A_C}{A_E}\left(e^{qV_{BC}/k_BT} - 1\right) \tag{1.34}$$

We describe the flux of carriers due to recombination in the space-charge regions as

$$F_{ESCR} = \frac{I_{ES2}}{qA_E}\left(e^{qV_{BE}/2k_BT} - 1\right) \tag{1.35}$$

and

$$F_{CSCR} = \frac{I_{CS2}}{qA_E}\left(e^{qV_{BC}/2k_BT} - 1\right), \tag{1.36}$$

where the saturation current densities, I_{ES2} and I_{CS2} describe recombination in both the bulk and perimeter space-charge regions.

Having defined the carrier fluxes, it is a simple matter to insert them in (1.30) to obtain the Ebers-Moll equations as

$$I_C = \alpha_F I_{ES1}\left(e^{qV_{BE}/k_BT} - 1\right) - \frac{I_{CS1}}{\gamma_{C12}}\left(e^{qV_{BC}/k_BT} - 1\right) \tag{1.37a}$$

and

$$I_E = \frac{I_{ES1}}{\gamma_{E12}}\left(e^{qV_{BE}/k_BT} - 1\right) - \alpha_R I_{CS1}\left(e^{qV_{BC}/k_BT} - 1\right). \tag{1.37b}$$

In these equations,

$$\alpha_F = \frac{\left(1 - W_B^2 / 2L_B^2\right)}{1 + \dfrac{(D_E / W_E) N_{AB} n_{iE}^2}{(D_B / W_B) N_{DE} n_{iB}^2}}, \tag{1.38a}$$

$$\alpha_R = \frac{\left(1 - W_B^2 / 2L_B^2\right)}{1 + \dfrac{(D_C / W_C) N_{AB} n_{iC}^2}{(D_B / W_B) N_{DC} n_{iB}^2}}, \tag{1.38b}$$

and the saturation current densities are given by

$$I_{ES1} = qA_E\left[\frac{D_B n_{iB}^2}{W_B N_{AB}} + \frac{D_E n_{iE}^2}{W_E N_{DE}}\right] \tag{1.39a}$$

and

$$I_{CS1} = qA_C\left[\frac{D_B n_{iB}^2}{W_B N_{AB}} + \frac{D_C n_{iC}^2}{W_C N_{DC}}\right]. \tag{1.39b}$$

The factors, γ_{E12} and γ_{C12} are the ratios of the n = 1 current components to the total currents for each junction [Lee85]. For example,

$$\gamma_{E12} = \frac{I_{ES1}\left(e^{qV_{BE}/k_BT} - 1\right)}{I_{ES1}\left(e^{qV_{BE}/k_BT} - 1\right) + I_{ES2}\left(e^{qV_{BE}/2k_BT} - 1\right)}. \tag{1.40}$$

The Ebers-Moll relations for graded-junction HBT's very nearly those for homojunction transistors, we simply use the appropriate value of n_i for the emitter, base, and collector. Nevertheless, the dc current vs. voltage characteristics they predict show some differences from homojunction transistors.

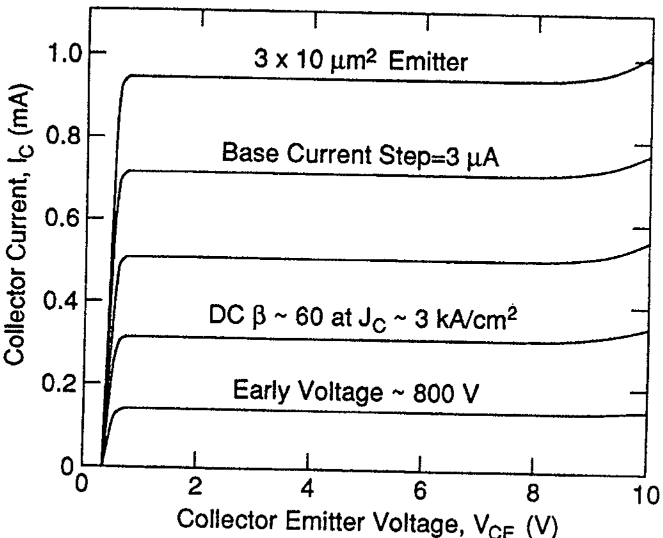


Figure 1.7a. *Measured common emitter characteristics for typical AlGaAs/GaAs HBT. (From [Kim90]).*

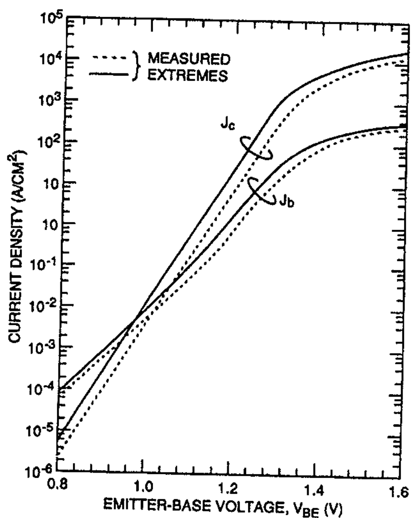


Figure 1.7b. *Measured Gummel plot for a typical AlGaAs/GaAs HBT. (A Gummel plot is a plot of the collector and base currents versus emitter-base voltage for a fixed base-collector voltage.) From [Haf90]).*

dc Characteristics of Graded-Junction HBT's

Figure 1.7 shows typical characteristics measured for graded-junction, AlGaAs/GaAs HBT's. Several features deserve comment. First note the collector current goes through zero at a finite value of V_{CE} which is termed the offset voltage, V_{OS}. Note also that the output conductance of the HBT is especially low and that the I_C and I_B vs. V_{BE} characteristics (the so-called Gummel plot) display quite different ideality factors.

Homojunction bipolar transistors also display an offset voltage, but it is usually too small to notice. By setting the collector current to zero in the Ebers-Moll equation, we find the offset voltage of the HBT to be

$$V_{OS} = \frac{k_B T}{q} \log\left[\frac{I_{CS1}}{\alpha_F \gamma_{C12} I_{ES1}}\right]. \qquad (1.41)$$

By noting that $\alpha_F I_{ES1} = \alpha_R I_{CS1} A_E / A_C$, we can express (1.41) as

$$V_{OS} = \frac{k_B T}{q} \log\left[\frac{A_C / A_E}{\alpha_R \gamma_{C12}}\right]. \qquad (1.42)$$

Equation (1.42) shows that substantial offset voltages occur when the collector-base junction area is much larger than the emitter-base junction, when the reverse α is small, or when the quality of the base-collector junction, as measured by γ_{C12}, is poor. The parasitic emitter, base, and collector resistances, which we've neglected, can also affect V_{OS} [Won89]. Figure 1.8 illustrates the physical mechanism responsible for the offset voltage. If a base current is injected when V_{CE} is zero, then equal parts must be injected into the emitter and collector if I_C is to be zero. Because of the wide bandgap, holes tend to be preferentially injected into the base-collector space-charge region where they recombine which forward biases the base-collector junction and produces a negative I_C. By applying a sufficient bias, V_{CE}, the energy barrier for holes at the base-collector junction is raised which forces the current to recombine in the emitter-base space-

charge region thereby forward-biasing the emitter-base junction and placing the transistor in the normal, active operating region.

A typical Gummel plot for an HBT is displayed in Fig. 1.7b. The collector current vs. V_{BE} characteristic displays an ideality factor of almost exactly one. In fact, when non-unity ideality factors are observed for I_C vs. V_{BE}, it usually means that the junction wasn't properly graded. The base current shows an $n \simeq 2$ ideality factor because it is dominated by recombination in the perimeter space-charge region.

ac Characteristics of Graded-Junction HBT's

We conclude this introductory look at HBT's by examining their high frequency performance. For analog applications, two important figures of merit are f_T, the current gain cutoff frequency, and f_{max}, the maximum frequency of oscillation. The well-known expression,

$$1/2\pi f_T = R_e\left(C_{je}+C_{jc}\right)+\tau_b+\tau_c, \qquad (1.43)$$

relates f_T to the RC charging times and to the transit times. The first term describes the charging of the emitter-base and base-collector capacitances through the dynamic emitter resistance, $R_e = dI_E/dV_{BE}$. For a uniformly doped base, the base transit time is

$$\tau_b = W_B^2/2D_B, \qquad (1.44)$$

and the collector delay time is

$$\tau_c = W_C/2\upsilon_s. \qquad (1.45)$$

Equation (1.45) assumes that electrons travel across the collector-base space-charge region at υ_s, the saturated velocity. The collector delay time is smaller than the transit time W_C/υ_s, because electrons image charge on the n^+-subcollector, so collector current flows as soon as electrons enter the collector-base space-charge region [Pri67].

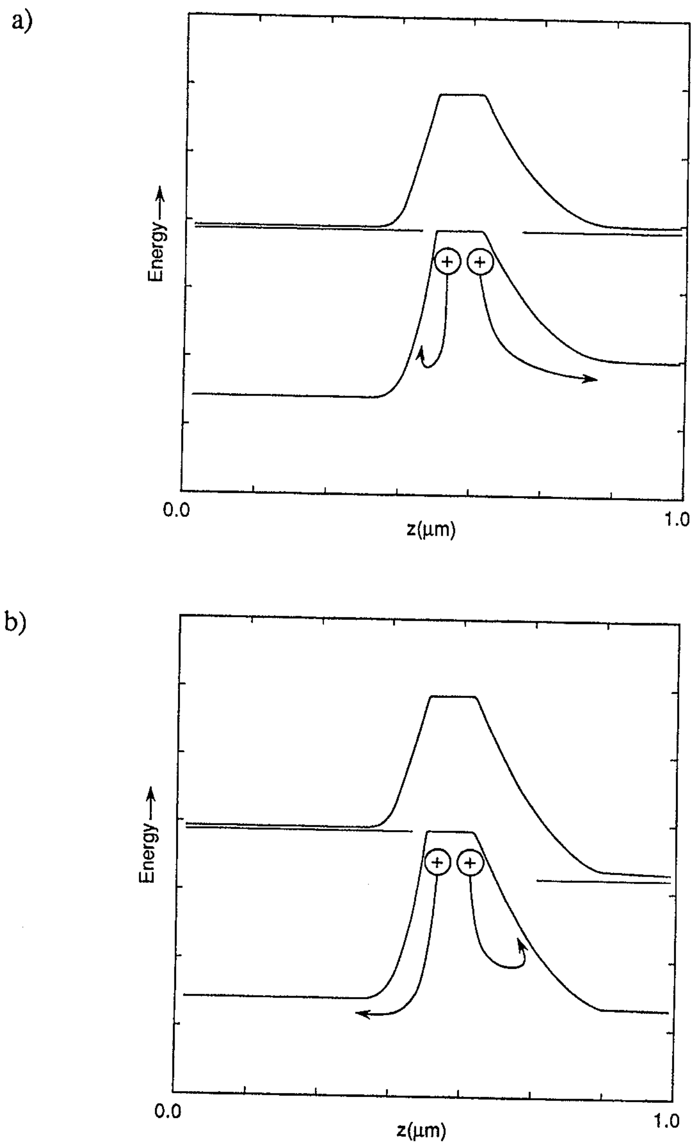


Figure 1.8. Illustration of the physical cause of the offset volatage, V_{OS}. a) $V_{CE} = 0$ and b) $V_{CE}>0$.

III-V HBT's offer higher f_T's than conventional silicon bipolar transistors for several reasons. Because of the emitter-base heterojunction, the emitter doping can be lowered without sacrificing current gain. The lower emitter doping lowers the base-emitter junction capacitance which reduces the R_eC_{je} charging time. The transit time contributions to f_T also tend to be low in III-V HBT's. The high electron mobilities typical of III-V semiconductors reduce the base transit time. The saturation velocity, υ_s, is similar for III-V semiconductors and silicon, but strong velocity overshoot can occur in III-V semiconductors. Devices can be optimized to exploit velocity overshoot in order to reduce the collector delay time [Maz86b, Ish88a].

The current gain cutoff frequency is evaluated by short-circuiting the output and driving the base with a current source. As a result, f_T, is insensitive to the base resistance, R_B. Another figure of merit,

$$f_{max} \sim \sqrt{f_T / 8\pi R_B C_{jc}}, \qquad (1.46)$$

is the unity power gain frequency or maximum frequency of oscillation. For HBT's, the low base resistance associated with the heavy base doping translates into a very high f_{max}. Transistors with high f_T and f_{max} are required for high-frequency performance, and very similar considerations apply to digital circuits. Compared to Si bipolar transistors, III-V HBT's posses advantages in both figures of merit, so III-V HBT's are especially suited to high-speed applications.