CHAPTER 6

Nonequilibrium electron transport in heterostructure bipolar transistors

6.1 INTRODUCTION

The Heterostructure Bipolar Transistor (HBT) concept, introduced by Shockley in the late 1940's [1], was promoted in the 1950's [2]. However, at that time it was not possible to implement the HBT, even as an experimental device which could be tested in a laboratory, because suitable semiconductor crystal growth and processing techniques had not yet been invented.

The advent of Molecular Beam Epitaxy (MBE) and Metalorganic Chemical Vapor Deposition (MOCVD) crystal growth changed this by providing the needed monolayer control of semiconductor composition. In the 1980's it became possible to form heterostructures which featured precise control of conduction and valence band potential profile in the crystal growth direction. In addition, the same methods could precisely control n and p-type impurity profiles. This combination of composition and doping profile control significantly increases the number of device concepts that could be realized in the laboratory. In the following the possibility of exploiting these developments to create high-performance HBT devices whose behavior is determined by non-equilibrium electron transport will be explored.

6.1.1 Advantages of HBTs

Advances in communication and information processing systems require implementation of very high-performance electronic circuits. For example, establishing efficient data paths between subsystems involves sophisticated digital, analog and sometimes optoelectronic circuitry. In recent years, the HBT has emerged as a leading contender to satisfy some of these needs. The reason for this is apparent when one considers the potential advantages of HBTs compared to competing technologies. The device engineering considerations include the use of wide bandgap heterostructure emitters allowing design of HBTs with very high ($\sim 10^{20}$ cm^{-3}) base doping levels which can significantly reduce base resistance. The availability of epitaxially grown material such as $In_{0.53}Ga_{0.47}As$ which have high electron mobility and can exhibit velocity overshoot are also factors which enable high-speed device performance. The ability to scale to submicron dimensions while maintaining high-speed performance but significantly reducing power consumption is an additional advantage of HBTs. Circuit designers can utilize the intrinsically high transconductance of HBTs to make very efficient circuits. The high transconductance of HBTs allows greater current handling capability than Field Effect Transistors (FETs). From a manufacturing point of view HBTs have the advantage that voltage threshold for current flow is controlled by the built-in base-emitter junction potential so that very uniform and reproducible device characteristics can be achieved. Both the high current handling capability and device threshold unifor-

mity make HBTs suitable for high-performance analog circuitry. Finally, compared to FETs the HBT is well shielded from surface traps which contribute to $1/f$ noise and trap induced output conductance.

An important potential competitive technology is based on FETs with a very small sub-0.1 μm gate length. Such FETs can also integrate digital functionality and achieve high frequency operation. Although, sub-micron Si-based digital FET circuitry is a very important electronic technology, it is difficult to design FET devices which *simultaneously* maintain precise control of turn-on threshold, high-current operation, and efficient high-frequency operation. Hence, HBT circuits are often the technology of choice for applications with the most stringent high-performance analog circuit requirements.

The purpose of this chapter is to describe the physical processes which result in intrinsically high-performance HBT devices and to explore the electron transport phenomena which limit this performance.

6.2 THE SEMICLASSICAL APPROACH

In Chapter 5 the notion of quasi-particles as the elementary excitations of a many-body system was introduced. Electron transport in semiconductors was then discussed using a semiclassical picture in which an injected electron scatters by exciting quasi-particles. This chapter makes use of the framework developed in Chapter 5 to explore the physics determining minority carrier electron transport in p-type semiconductors and its application to bipolar transistors. One can develop an understanding of scaling in devices with nonequilibrium electron transport by calaculating scattering rates. Many of the results that can be obtained are general because they are based on powerful phase-space arguments. Following the establishment of a theory, it is natural to proceed to compare expectations with the results of experiments.

In the previous chapter we applied our knowledge of elastic and inelastic scattering rates for a nonequilibrium conduction band electron in a n-type semiconductor to help design a nonequilibrium electron transistor that operated at room temperature. Unfortunately, the nonequilibrium electron scattering rates are directly related to the common-emitter current gain, limiting the usefulness of the device. Ideally, one would like to capture the advatages of fast electron motion inherent to nonequilibrium electron transport while, at the same time, decoupling current gain from nonequilibrium scattering rates. A good way to approach this is to *separate* the nonequilibrium (minority) electrons in the conduction band from the majority carriers. A device that satisfies this criteria is the *n-p-n* HBT with a wide bandgap emitter designed to launch nonequilibrium conduction band electrons into the transistor base. As with nonequilibrium electron transistors it would be helpful to calculate elastic and inelastic scattering rates for nonequilibrium conduction band electrons in p-type semiconductors.

6.2.1 Quasi-particle excitation spectra

The semiconductor bandstructure will determine the kinematic constraints on quasi-particle excitation. Within the effective mass approximation, the valence band maxima for a typical direct bandgap bulk semiconductor such as GaAs consists of a heavy hole (hh) band and a light hole (ℓh) band which is degenerate in energy at the Γ-symmetry point. The effective mass of charge carriers in the heavy hole band is $m_{\rm hh}$ and the effective mass of charge carriers in the light hole band is $m_{\ell \rm h}$. As a starting point, one

would like to determine the distribution of occupied states in each isotropic band. For convenience, one may choose the low temperature limit for which $\mu \to E_F$. To calculate E_F, consider a bulk-doped majority p-type carrier concentration $p = p_{hh} + p_{\ell h}$, where p_{hh} is the concentration of heavy holes and $p_{\ell h}$ is the concentration of light holes. The heavy hole Fermi wave vector is

$$k_{F_{hh}} = (3\pi^2 p_{hh})^{1/3} \tag{6.1}$$

and light hole Fermi wave vector is

$$k_{F_{\ell h}} = (3\pi^2 p_{\ell h})^{1/3} \tag{6.2}$$

The Fermi energy is $E_F = \hbar^2 k_{F_{hh}}^2 / 2m_{hh} = \hbar^2 k_{F_{\ell h}}^2 / 2m_{\ell h}$ so that $k_{F_{hh}}^2 / m_{hh} = k_{F_{\ell h}}^2 / m_{\ell h}$. It follows that the heavy hole and light hole carrier concentration are related via the ratio of effective masses by $p_{hh} = (m_{hh}/m_{\ell h})^{3/2} p_{\ell h}$. Since total carrier concentration $p = p_{hh} + p_{\ell h}$, the concentration of heavy holes can be written

$$p_{hh} = p/(1 + (m_{\ell h}/m_{hh})^{3/2}) \tag{6.3}$$

and the concentration of light holes is

$$p_{\ell h} = p/(1 + (m_{hh}/m_{\ell h})^{3/2}) \tag{6.4}$$

Hence, the Fermi energy for total majority carrier concentration p is simply

$$E_F = \hbar^2 k_{F_{hh}}^2 / 2m_{hh} = \hbar^2 k_{F_{\ell h}}^2 / 2m_{\ell h} = \frac{\hbar^2}{2m_{hh}} \left(\frac{3\pi^2 p}{1 + (m_{\ell h}/m_{hh})^{3/2}} \right)^{2/3} \tag{6.5}$$

As a specific example, consider GaAs. The heavy hole effective mass is $m_{hh} = 0.5 m_0$ and the light hole effective mass is $m_{\ell h} = 0.082 m_0$ giving a ratio $(m_{\ell h}/m_{hh})^{3/2} = 0.0664$. In this case the Fermi energy is related to the p-type carrier concentration via $E_F = 6.99 \times 10^{-15}\ p^{2/3}$ eV. Hence, for p-type carrier concentration $p = 2 \times 10^{18}$ cm^{-3} one obtains

$k_{F_{hh}} = 3.815 \times 10^6$ cm^{-1}

$k_{F_{\ell h}} = 1.545 \times 10^6$ cm^{-1}

$p_{hh} = 1.875 \times 10^{18}$ cm^{-3}

$p_{\ell h} = 1.246 \times 10^{17}$ cm^{-3}

$E_F = 11.09$ meV.

For p-type carrier concentration $p = 2 \times 10^{19}$ cm^{-3} one obtains

$k_{F_{hh}} = 8.219 \times 10^6$ cm^{-1}

$k_{F_{\ell h}} = 3.329 \times 10^6$ cm^{-1}

$p_{hh} = 1.875 \times 10^{19}$ cm^{-3}

$p_{\ell h} = 1.246 \times 10^{18}$ cm^{-3}

$E_F = 51.48$ meV.

The Fermi energy for $p = 2 \times 10^{19}$ cm^{-3} in GaAs is about the same as for an *n*-type carrier concentration of $n = 1 \times 10^{18}$ cm^{-3}. This is a consequence of the difference in effective masses m_{hh} and m_e. The presence of a heavy hole (hh) and a light hole (ℓh) band (we ignor the $j = 1/2$ split-off band) places constraints on the energy and momentum transfer associated with single particle excitations. For simplicity, one may consider these valence band excitations at $T = 0$ K. An injected nonequilibrium conduction band electron with initial wave vector $\mathbf{k}$ and energy E above the conduction band CB_{min} in the transistor base may interact with excitations of the system and scatter, changing momentum by $-\mathbf{q}$ and loosing energy $\hbar\omega$ as illustrated in Fig. 6.1(a) and (b). The Fig. illustrates nonequilibrium electron dynamics in (a) real-space and (b) $\mathbf{k}$-space for the forward biased emitter-base of an abrupt-junction of an *n-p-n* HBT

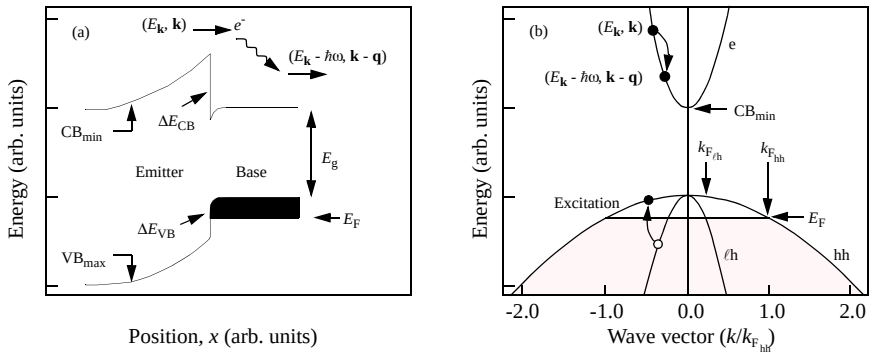
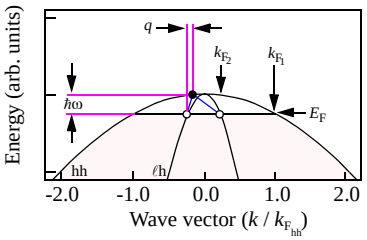


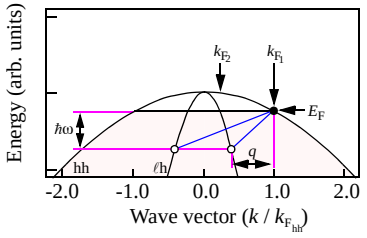
Fig. 6.1. (a) Schematic band diagram of a forward biased emitter-base junction of a *n-p-n* HBT. The wide bandgap emitter has a conduction band offset ΔE_{CB} and a valence band offset ΔE_{VB}. The bandgap E_g of the base, the conduction band minimum CB_{min}, the valence band maximum VB_{max}, and hole quasi-Fermi level (E_F) are indicated. Electrons are injected into the base with initial kinetic energy $E_\mathbf{k}$ and wave vector $\mathbf{k}$. As the electrons traverse the base they can scatter into new states. (b) Semiconductor band structure of the HBT base in (a) showing valence heavy-hole band *hh* with Fermi wave vector $k_{F_{hh}}$, light-hole band ℓh with Fermi wave vector $k_{F_{\ell h}}$, and conduction band e. Electrons in the conduction band can scatter by creating an excitation in the valence band of energy $\hbar\omega$ and wave vector $\mathbf{q}$.

One may determine the range of single particle excitations allowed by energy and momentum conservation in the valence band. Pictorially, the limits of allowed single particle transitions are shown as sketches in Fig. 6.2 where energy and wave vector have been normalized to E_F and $k_{F_{hh}}$ such that $Y = \hbar\omega/E_F$ and $X_{hh} = q/k_{F_{hh}}$. Extrema exist where the transition originates or terminates at the Fermi energy.



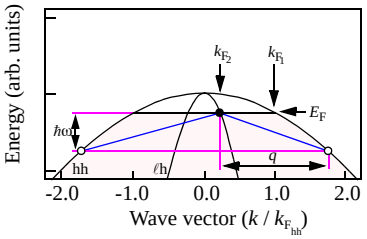
1. Heavy hole to light hole transition to the light hole Fermi energy

$$\hbar\omega = E_F - (\hbar^2/2m_{hh})(k_{F_{hh}}\sqrt{m_{\ell h}/m_{hh}} - q)^2$$
$$Y = 1 - m_{\ell h}/m_{hh} \pm 2X_{hh}\sqrt{m_{\ell h}/m_{hh}} - X_{hh}^2$$
$$Y = 1 - (\sqrt{m_{\ell h}/m_{hh}} \mp X_{hh})^2$$
$$Y_{min} = 1 - (m_{\ell h}/m_{hh}) \text{ at } X_{hh} = 0$$



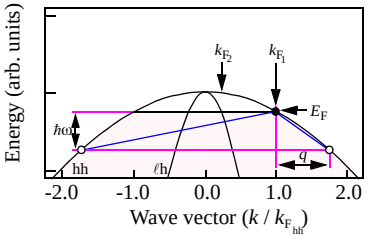
2. Heavy to light hole transition from the heavy hole Fermi energy

$$\hbar\omega = (\hbar^2/2m_{hh})(k_{F_{hh}} - q)^2 m_{\ell h}/m_{hh} - E_F$$
$$Y = (m_{\ell h}/m_{hh}) \pm 2X_{hh}(m_{\ell h}/m_{hh}) + X_{hh}^2(m_{\ell h}/m_{hh}) - 1$$
$$Y = (m_{\ell h}/m_{hh})(X_{hh} \mp 1)^2 - 1$$
$$Y_{max} = (m_{hh}/m_{\ell h}) - 1 \text{ at } X_{hh} = 0$$
$$Y = 0 \text{ when } X_{hh} = 1 \pm \sqrt{m_{\ell h}/m_{hh}}$$



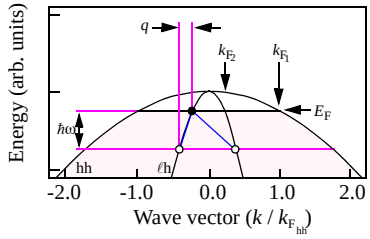
3. Light to heavy hole transition from the light hole Fermi energy

$$\hbar\omega = (\hbar^2/2m_{hh})(k_{F_{hh}}\sqrt{m_{\ell h}/m_{hh}} + q)^2 - E_F$$
$$Y = (m_{\ell h}/m_{hh}) \pm 2X_{hh}\sqrt{m_{\ell h}/m_{hh}} + X_{hh}^2 - 1$$
$$Y = (X_{hh} \pm \sqrt{m_{\ell h}/m_{hh}})^2 - 1$$
$$Y = 0 \text{ when } X_{hh} = 1 \pm \sqrt{m_{\ell h}/m_{hh}}$$



4. Heavy hole to heavy hole transition

$$\hbar\omega = (\hbar^2/2m_{hh})(k_{F_{hh}} + q)^2 - E_F$$
$$Y = X_{hh}^2 \pm 2X_{hh}$$
$$Y = X_{hh}(X_{hh} \pm 2)$$
$$Y = 0 \text{ when } X_{hh} = 0 \text{ and } X_{hh} = 2$$



5. Light hole to light hole transition

$$\hbar\omega = (\hbar^2/2m_{\ell h})(k_{F_{\ell h}} + q)^2 - E_F$$
$$\omega = (m_{\ell h}/m_{hh})(k_{F_{hh}}^2 + q^2 - k_{F_{hh}}^2) + 2k_{F_{\ell h}}q\sqrt{m_{\ell h}/m_{hh}}$$
$$Y = X_{hh}^2 \pm 2X_{hh}\sqrt{m_{hh}/m_{\ell h}}$$
$$Y = 0 \text{ when } X_{hh} = 0 \text{ and } X_{hh} = 2\sqrt{m_{\ell h}/m_{hh}}$$

Fig. 6.2. Schematic diagram of the limits of single particle excitations of energy $\hbar\omega$ and momentum q near the Γ symmetry point in the valance band of a p -type semiconductor at temperature $T = 0$ K. The heavy hole effective electron mass is m_{hh} and the light hole effective electron mass is $m_{\ell h}$.

One can use the equations developed in Fig. 6.2 to sketch the dispersion relation for single particle excitations. Unfortunately, the dimensionless units Y and $X_{\rm hh}$ do not allow one to plot the results for all semiconductors because the equations in Fig. 6.2 contain the ratio of effective masses $m_{\rm hh}/m_{\ell \rm h}$. This forces one to choose a specific semiconductor. As a concrete example, consider GaAs with $m_{\rm hh}/m_{\ell \rm h} = 0.5/0.082 = 6.1$. The shaded region of Fig. 6.3 shows the range of allowed single particle excitations. It is immediately apparent that an important *qualitative difference* between excitations in a single band and excitations in multiple bands is the possibility of vertical ($q = 0$) transitions with energy loss ($\hbar\omega \neq 0$). This fact directly impacts the dielectric response of the semiconductor. The allowed vertical transitions for $q = 0$ have energy loss whose value can be in the range from $Y_{\rm min} = 1 - (m_{\ell \rm h}/m_{\rm hh})$ to $Y_{\rm max} = (m_{\rm hh}/m_{\ell \rm h}) - 1$.

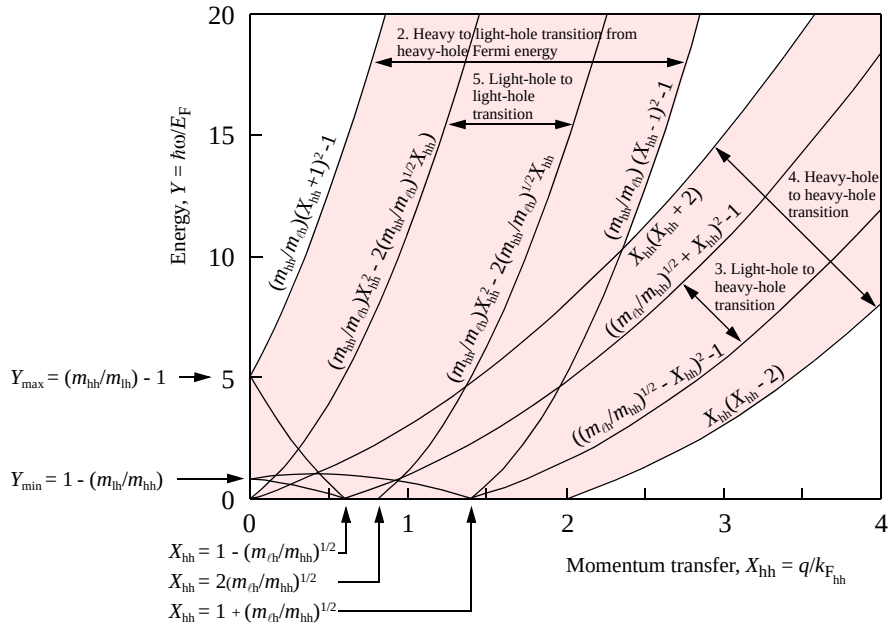


Fig. 6.3. Dispersion relation for single particle excitations in p-type GaAs with $m_{\rm hh} = 0.5m_0$ and $m_{\ell \rm h} = 0.082m_0$. The allowed single particle transitions are indicated by the shaded regions.

So far we have discussed the single particle exciations and seen that vertical transitions are possible. Such vertical transitions are potentially important because Coulomb scattering is weighted to $q \to 0$. Another consequence is that single particle excitations and regions of Landau damping dominate the dielectric response. However, collective coupled plasmon / phonon excitations are not entirely damped by these single particle processes. To gain further insight, consider p-type GaAs with $p = 2 \times 10^{18}\ {\rm cm}^{-3}$, $m_1 = 0.5m_0$, $m_{\ell \rm h} = 0.082m_0$, $k_{\rm F_{hh}} = 3.82 \times 10^6\ {\rm cm}^{-1}$. In Fig. 6.4(a) results of evaluating the real and imaginary parts of $\varepsilon(q, \omega)$ for $T \to 0$ K, $\gamma = 0.09$ and a small value of scattered wave vector $q = 0.02k_{\rm F_{hh}}$. The rather small, unphysical value of γ is used purely to illustrate the embedded structure of $\varepsilon(q, \omega)$. It

is clear from the Fig. that $\varepsilon(q, \omega)$ is quite rich in features. The real part of the dielectric function ε_1 has zeros at energy $\hbar\omega_{TO}$ and $\hbar\omega_C$. The former is due to the tranverse polar-optic phonon and the later is due to collective vertical interband transitions. The imaginary part of the dielectric function is ε_2. In the Fig., a broad range of single-particle excitations exist from energy Y_{min} to Y_{max}. There is also a large peak in ε_2 at energy $\hbar\omega_{TO}$. Fig. 6.4(b) shows the correponding loss function $\mathrm{Im}(-1/\varepsilon(q, \omega))$. As expected, the scattering strength (or spectral weight) of single particle excitations extends in energy from Y_{min} to Y_{max}. The zero in ε_1 at energy $\hbar\omega_{TO}$ is strongly suppressed by the peak in ε_2. However, the zero in ε_1 due to collective interband excitations at energy $\hbar\omega_C$ is not strongly damped by single-particle excitations and there is a large peak in the spectral function. The dielectric function for p-type zinc-blende semiconductors is quite complex and has been studied by Yevik and Bardyszewski [3]. In particular, because the value of the Fermi energy for $p = 2\times10^{18}\ \mathrm{cm}^{-3}$ in GaAs has the relatively small value of $E_F = 11.09\ \mathrm{meV}$ there is a quite strong temperature dependence to the loss function. While this changes many details [4], the overall scattering strength at this carrier concentration is not very temperature sensitive for typical vaues of $E_{\mathbf{k}}$. This is because calculation of the total inelastic scattering rate involves an integral over a large portion of the spectral weight which tends to average out local peaks in oscillator strength.

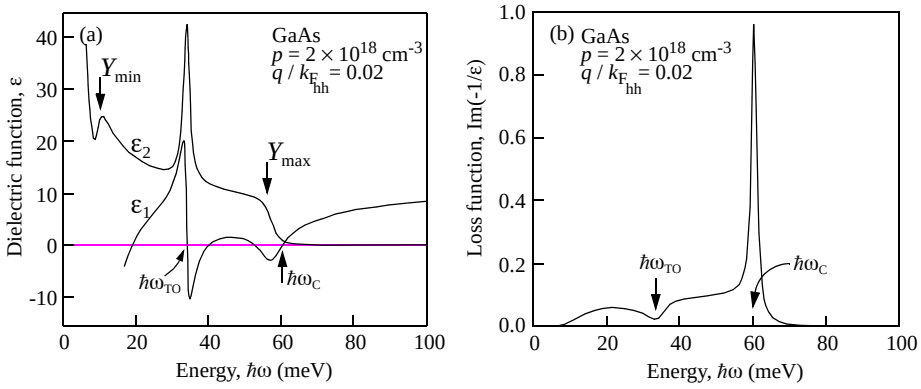


Fig. 6.4. a) Plot of the real ε_1 and imaginary ε_2 part of the low temperature dielectric response function for a fixed value of $q = 0.02\times k_{F_{hh}}$ as a function of energy loss $\hbar\omega$ at a carrier concentration of $p = 2\times10^{18}\ \mathrm{cm}^{-3}$ in GaAs. The Fermi energy $E_F = 11.09\ \mathrm{meV}$ and heavy hole Fermi wave vector $k_{F_{hh}} = 3.82\times10^{6}\ cm^{-1}$. (b) The corresponding loss function $\mathrm{Im}(-1/\varepsilon(q, \omega))$ for (a). Parameters used in the calculation are $m_{hh} = 0.5m_0$, $m_{\ell h} = 0.082m_0$, $\varepsilon_\infty = 11.1$, $\hbar\omega_{LO} = 36.3\ \mathrm{meV}$, $\hbar\omega_{TO} = 33.3\ \mathrm{meV}$, and $\gamma = 0.09$.

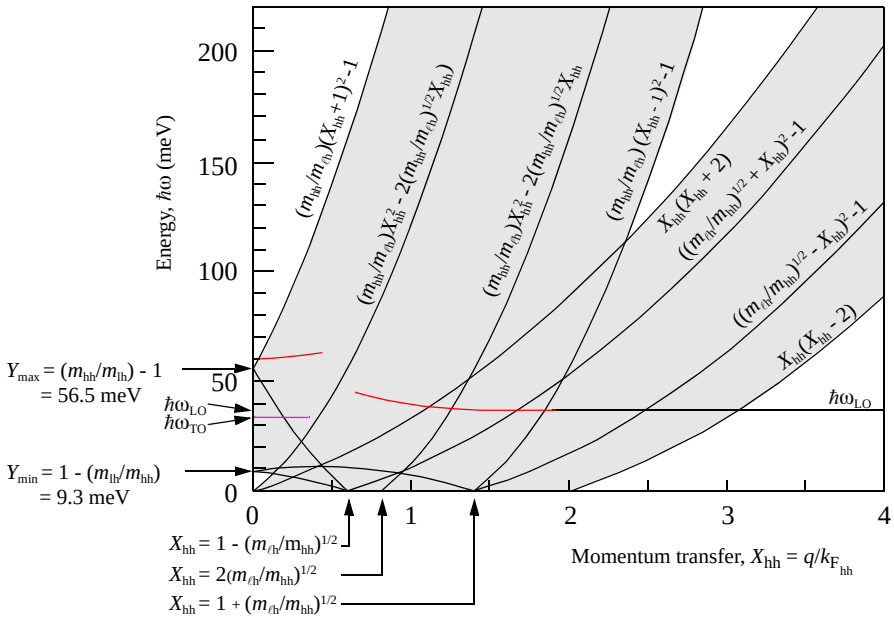


Fig. 6.5. Dispersion relation for single particle and collective excitations in p-type GaAs with $p = 2\times10^{18}$ cm^{-3}, $m_{lh} = 0.082m_0$, $m_{lh}/m_{hh} = 0.164$, $E_F = 11.09$ meV, $k_{F_{hh}} = 3.82\times10^6$ cm^{-1}, and $\varepsilon_\infty = 11.1$. The collective excitation at $\hbar\omega_C = 60$ meV associated with interband excitations is Landau damped in the single particle exciation region. Also shown is the contribution of longitudinal polar-optic phonon at energy $\hbar\omega_{LO} = 36.3$ meV. The energy loss function is function is suppressed for small q near the transverse polar-optic phonon energy $\hbar\omega_{TO} = 33.3$ meV (see Fig. 6.4(b)).

In Fig. 6.5 the dispersion realtion for p-type GaAs with carrier concentration $p = 2\times10^{18}$ cm^{-3} is shown.

6.2.2 Anticipated trends for electron dynamics

Electron dynamics are, in part determined by the inelastic scattering rate. The inelastic scattering rate in the low temperature limit is calculated via the integral

$$1/\tau_{in} = \frac{2m_3e^2}{\pi\hbar^2k}\int \mathrm{Im}\left\{\frac{-1}{\varepsilon(q,\omega)}\right\}\frac{dq}{q}d\omega \tag{6.6}$$

One needs to be careful to conserve energy and momentum in the scattering process. For a *conduction band electron* energy $E_{\mathbf{k}}$, wave vector $\mathbf{k}$, the scattered electron has energy $E_{\mathbf{k}} - \hbar\omega$ and wave vector $\mathbf{k} + \mathbf{q}$. Electrons are scattered from one state to another *within* the conduction band. Using an effective mass model for the bands and considering the $T = 0$ K limit gives rise to a parabola of integration for the kinematically allowed transitions between states. Only transitions within the parabola contribute to the inelastic scattering rate. The parabola is

$$Y = Y_{\mathbf{k}} - \frac{m_{hh}}{m_e}\left(X_{hh} - \sqrt{\frac{m_e}{m_{hh}}}Y_{\mathbf{k}}\right)^2 \tag{6.7}$$

where $Y_{\mathbf{k}} = E_{\mathbf{k}}/E_{\mathrm{F}}$ is the initial kinetic energy of the electron in the conduction band. Fig. 6.6 shows parabolae of integration for $Y_{\mathbf{k}} = 10$ and $Y_{\mathbf{k}} = 20$ for GaAs. By comparing to the example for p-type GaAs with carrier concentration $p = 2\times10^{18}\ \mathrm{cm}^{-3}$ shown in Fig. 6.5, it is clear that the parabola does not include spectral weight from collective excitations of energy $\hbar\omega_{\mathrm{C}}$ near $q = 0$. In fact, this kinematic constraint, combined with the fact that typically most spectral weight is contained within heavy-hole to heavy-hole single-particle excitation due to their greater density of states, results in single-particle exciation usually dominating inelastic processes. This has lead some to develop a damped single-pole plasmon approximation which, as pointed out in Ref. 3, leaves out many scattering processes and significant errors in calculated scattering rates.

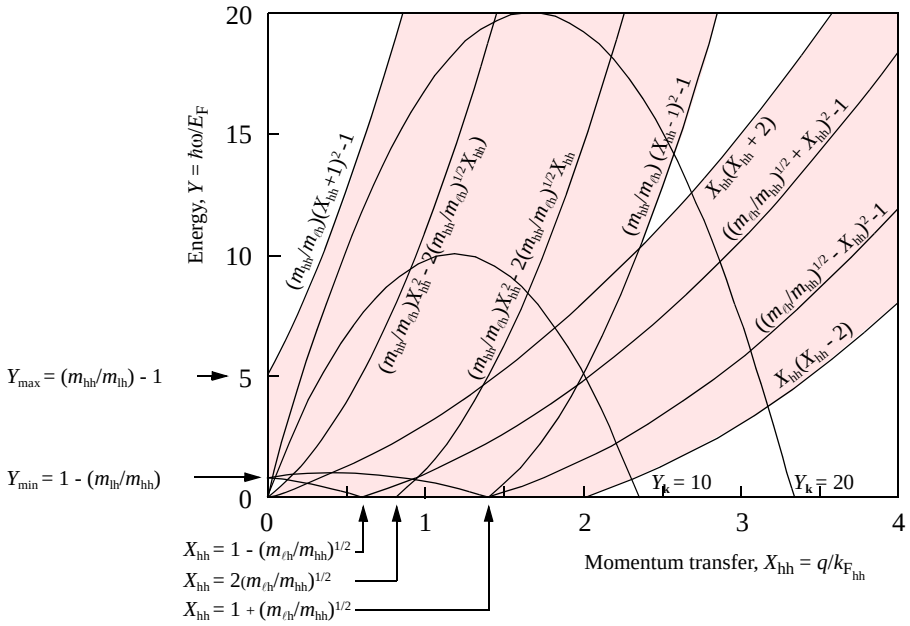


Fig. 6.6. Dispersion relation for single-particle excitations in p-type GaAs with $m_{\mathrm{hh}} = 0.5m_0$ and $m_{\ell\mathrm{h}} = 0.082m_0$. Paraboloas of integration for a nonequilibrium conduction band electron of energy $E_{\mathbf{k}} = 10\times E_{\mathrm{F}}$ and $E_{\mathbf{k}} = 20\times E_{\mathrm{F}}$ are shown. In GaAs, the conduction band electron has an effective electron mass $m_{\mathrm{e}} = 0.07m_0$.

The kinematic constraints represented by the parabola of integration has some further consequences. For a given E_{F}, if the energy of the nonequilibrium electron in the conduction band is increased we can integrate over more of the excitation phase space and over more small-q excitations. If the p-type carrier concentration is increased, then E_{F} increases, and the parabola of integration using the normalized units in Fig. 6.6 shrinks in size. Hence, for fixed $E_{\mathbf{k}}$, with increasing p a reduced portion of phase space with fewer low-q vaues is integrated. Even though the total spectral intensity can increase for large p, the integrated spectral intensity in a small region of phase space can decrease to give a reduced inelastic scattering rate. This result is general because it is based on a phase-space argument. Hence, for high carrier concentrations used in

devices, a fundamantal limit to nonequilibriumm device performance (the inelastic scattering rate) can actually improve!

One might expect some qualitative differences in the elastic and inelastic scattering rates for minority conduction band electrons scattering off majority carrier valence band holes compared to the previous case described in Chapter 5 which involved conduction band electrons scattering off majority carriers in the conduction band.

First, conduction band electrons with energies approaching CB_{min} can scatter both elastically and inelastically. The Pauli exclusion principle does not provide a forbidden energy range for a dilute density of minority carriers in the conduction band.

Second, the difference in effective mass between a conduction band electron of mass m_e and heavy holes of mass m_{hh} will influence elastic and inelastic scattering rates via the dielectric function. The parabolas of integration and the single particle excitations depicted in Fig. 6.6 are defined interms of ratios of effective electron mass.

Elastic scattering rates are in large part determined by the dielectric screening. We expect the heavy hole to statically screen ionized impurities much more effectively than a conduction band electron. This is because the relatively large value of m_{hh} allows the holes to get very close to the ionized impurity before its energy becomes too great.

For elastic scattering of a conduction band electron off ionized holes screened by heavy holes we have

$$1/\tau_{el} = \frac{2\pi p m_e e^4}{\hbar^3 k^3}\int \frac{\eta\, d\eta}{(\eta^2 \varepsilon(2k\eta, 0))^2} \tag{6.8}$$

where $\eta = \sin(\theta/2)$, θ is the scattering angle between $\mathbf{k}$ and $\mathbf{k}'$, and p is the hole impurity concentration. Assuming only heavy holes in the valence band, then we expect for static screening ($\omega = 0$)

$$\varepsilon(X, 0) = \frac{9}{4}\frac{m_{hh}}{a_0}\frac{\pi}{k_{F_{hh}}^3}\frac{1}{X^3}\left(X + \left(1 - \frac{X^2}{4}\right)\ln\left|\frac{X+2}{X-2}\right|\right) \tag{6.9}$$

where $X = q/k_{F_{hh}}$, $X = 2k\eta$, and $a_0 = 4\pi\hbar^2/e^2 m_0$.

The static dielectric function has a prefactor that scales as m_{hh} so a heavy effective hole mass statically screens better than a lighter effective mass. For GaAs the ratio of electron and heavy-hole effective mass is $m_{hh}/m_e = 0.5/0.07 = 7.1$.

The inelastic scattering rate is determined by the dielectric loss function. One would expect the heavy hole effective mass to result in poor dynamic screening $\omega \neq 0$ leading to increased inelastic scattering rates. The heavy hole is less able to respond to a charge density fluctuation than a light hole. The heavy hole has greater mass than the light hole and so is nimble. The resulting poor screening exposes more bare coulomb potential, leading to greater inelastic scattering rates.

As we learn more about the details of the dielectric function for *p*-type semiconductors we will see that, infact, the situation is not quite so simple. There are some surprises awaiting us!

6.2.3 Preliminary HBT design considerations

Suppose one wishes to design a HBT which makes use of extreme nonequilibrium electron transport. The thickness of the base is x_B. Clearly one will be concerned with increasing the minority carrier electron mean free path in the base, λ_B such that

$\lambda_B / x_B > 1$. From our previous work one will probably be interested in materials such as $In_{0.53}Ga_{0.47}As$ which have relatively small scattering rates and a high electron mobility. Apart from choosing a high mobility, low effective electron mass material for the base, electrons must be injected in a high velocity beam, implying electron injection into high excess kinetic energy states. In addition, since total scattering may be reduced by decreasing the base thickness one expects base thickness to have a small value.

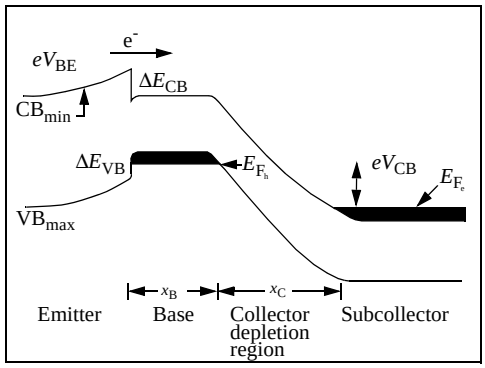


Fig. 6.7. Schematic band diagram of a *n-p-n* HBT under a bias V_{BE} between base and emitter and V_{CB} between collector and base. The wide bandgap emitter has a conduction band offset ΔE_{CB} and a valence band offset ΔE_{VB}. Base thickness x_B, collector depletion region thickness x_C, conduction band minimum CB_{min} and electron (hole) quasi-Fermi level E_{F_e} (E_{F_h}) are also indicated.

Fig. 6.7 shows a schematic band diagram of a *n-p-n* HBT of base thickness x_B, base doping p and collector depletion region thickness x_C. The wide bandgap emitter has a conduction band offset ΔE_{CB} and a valence band offset ΔE_{VB}. The bias voltage V_{BE} causes conduction band electrons to flow across the abrupt emitter-base junction and be injected into the base in the **x** direction with excess kinetic energy, E. An approximate measure of the spread in injection angles is given by $\Delta\theta = \tan^{-1}(k_B T/E)^{1/2}$ and so reasonably beam-like electron injection at room temperature occurs when $E > 200$ meV and $\Delta\theta \leq 20°$. To make use of nonequilibrium electron transport in bipolar devices, we want to optimize conditions to ensure that $x_B < \lambda_B$. Part of this optimization process is the correct choice of materials to give an appropriate heterostructure band offset, ΔE_{CB}. Table 6.1 shows the low-temperature heterostructure band offsets for a number of combinations of lattice matched III-V compound semiconductors. A typical choice of materials is an InP emitter and an $In_{0.53}Ga_{0.47}As$ base with a conduction heterostructure offset $\Delta E_{CB} = 0.26$ eV. To ensure $x_B < \lambda_B$ we can make x_B as small as possible while simultaneously not allowing the base resistance become too high (see section 6.2.4.1) and chose a majority carrier p-type impurity level such that the total minority carrier electron scattering rate in the base is not too great (see section 6.2.4.2).

To utilize the HBT's high current drive capability (useful for high-speed applications) we use thermionic emission to inject electrons from the emitter into the base. Space charging effects at high current density are avoided by doping the emitter and

collector depletion regions to a density $n > j_e / e\upsilon_{av}$ where υ_{av} is an appropriate average electron velocity in the depletion region and j_e is the electron current density. For example, if $\upsilon_{av} = 10^7$ cm s^{-1} then a current density of $j_e \sim 10^5$ A cm^2 can be supported in a region with impurity concentration $n = 10^{17}$ cm^{-3}.

TABLE 6.1. *Low-temperature heterostructure band offsets for some lattice-matched III-V compound semiconductors.*

Material 1	Material 2	E_{g1} (eV)	E_{g2} (eV)	ΔE_{CB} (eV)	ΔE_{VB} (eV)
$Al_{0.48}In_{0.52}As$	$In_{0.53}Ga_{0.47}As$	1.6	0.81	0.54	0.25
InP	$In_{0.53}Ga_{0.47}As$	1.42	0.81	0.26	0.35
InP	$Al_{0.48}In_{0.52}As$	1.42	1.6	0.18	0.0
AlSb	InAs	1.68	0.41	1.3	-0.03
GaSb	InAs	0.81	0.41	0.8	-0.4
AlSb	GaSb	1.68	0.81	0.55	0.32
AlAs	GaAs	2.23	1.52	0.16	0.55

In addition to classical transport effects, when appropriate, one must also consider quantum phenomena such as quantum mechanical reflection. In this case it is important to inject electrons in a narrow range of energies close to E. Optimum device performance occurs when quantum reflections from any abrupt potential step $\phi_{\alpha\beta}$ in the path of the injected electron beam is minimized. This may only be achieved for a limited range of injection energies. Reflections from an abrupt change in potential $\phi_{\alpha\beta}$ (for example at the base-collector junction) are minimized when the nonequilibrium electron group velocity (the slope $\partial\omega/\partial k$ at energy E) is the same either side of the potential step [4, 5]. This impedance matching condition is $m_\alpha / m_\beta = E/(E - \phi_{\alpha\beta})$, where m_α and m_β are the effective electron masses either side of the potential step. It is worth mentioning that at a heterostructure interface potential step, the impedance matching condition also requires that the character of the electron wave function in the two regions be similar (for example sp Γ-like states at energy E) [6].

In the extreme situation where the mean free path through base *and* collector λ_{BC} becomes comparable to $x_B + x_C$, then the mechanisms by which the device operates are qualitatively different compared to the conventional situation. Recall that a typical bipolar transistor operates by transforming a low impedance port (the emitter) into a high impedance port (the collector). Electrons injected from the forward biased emitter-base diode traverse the base and are accelerated in the electric field of the reverse biased base-collector diode. Irreversible scattering processes, such as longitudinal polar-optic phonon emission, take place in the collector depletion region and reverse current flow against the electric field in the depletion region is small. In this way the transformation to high impedance occurs over the length of the collector depletion region. The situation is different for a transistor in which at least a portion of current flow through the base and collector is via coherent "ballistic" transport (i.e., $\lambda_{BC} > x_B + x_C$). As illustrated in Fig. 6.8, one may imagine a transfer Hamiltonian

description [7] in which electrons are scattered into a state, Ψ, in the emitter and scattered out of that state in the degenerately doped n^+ subcollector contact. In such a transistor the impedance transformation mainly takes place in the subcollector contact. Clearly, in this operating condition irreversible scattering in the subcollector can influence the emitter current and thus change the fundamental impedance transforming properties of the transistor.

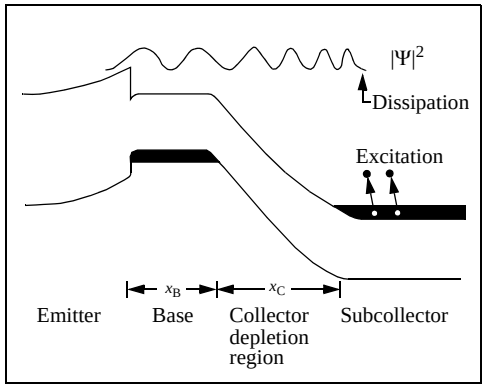


Fig. 6.8. Transfer Hamiltonian description of extreme nonequilibrium electron transport in a n-p-n HBT. Electrons in the emitter scatter into the state Ψ and scatter out of that state by, for example, exciting an electron-hole pair in the n^+ subcollector contact.

6.2.4 Effect of reducing x_B on electron transport

6.2.4.1 Nonequilibrium electron spectroscopy

Thin film semiconductor crystal growth techniques, such as molecular beam epitaxy, allow fabrication of atomically precise layered compositional and doping profiles. In the mid 1980s it became apparent that such abrupt profiles could be used to study nonequilibrium electron transport in semiconductors. There already exist reviews [8, 9] of these early studies so we only need describe the basic concepts. The inset in Fig. 6.9(a) shows the conduction band edge profile of a bipolar nonequilibrium electron spectrometer. When suitably biased, electrons are tunnel injected from the emitter into the conduction band of the p-type base region. Initially, the injected electrons have average excess kinetic energy E. At first, electron momentum is predominantly in the **x** direction (i.e., perpendicular to the plane of the injector). Later, while traversing the base, the injected electrons may experience elastic (impurity) and inelastic collisions, which result in a redistribution of kinetic energy and directional motion. After transiting the base, electrons impinge on the collector barrier. The energy of this barrier, ϕ_{BC}, may be controlled by adjusting the potential, V_{BC}, applied between the collector and base of the spectrometer. It follows that measurement of dI_C/dV_{BC} as a function of V_{BC} gives information related to a projection of the electron momentum distribution perpendicular to the plane of the collector barrier.

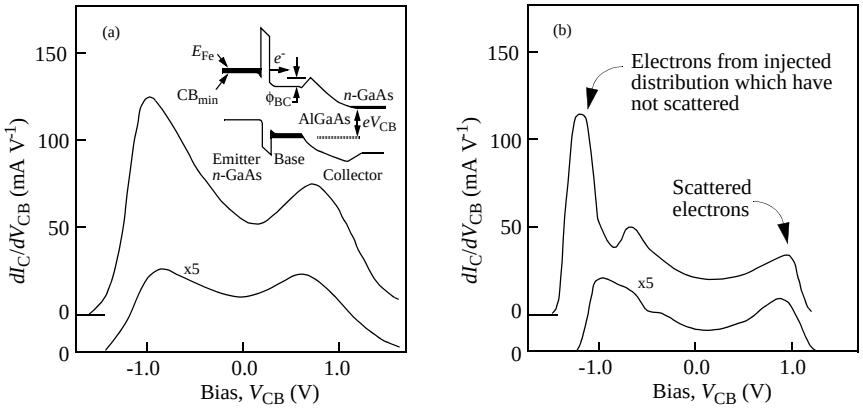


Fig. 6.9. (a) Derivative of the collector current, dI_C/dV_{CB}, with collector voltage, V_{CB}, for the AlGaAs/GaAs bipolar transistor sketched in the inset. The lower curve (5 times vertical scale) is for an injection energy E = 170 meV, the upper curve for E = 220 meV. The peak at around V_{CB} = –1 V contains electrons from the initial injected distribution which have not scattered while the peak at V_{CB} = 0.7 V is due to electrons which have suffered a number of collisions. The measurements were performed at temerpature T = 4.2 K (from [10]). (b) Numerical simulation of the spectra shown in (a). The numerically simulated spectra were broadened by 30 meV (from [11]).

Nonequilibrium electron spectroscopy has been used to demonstrate the existence of nonequilibrium electron transport in the base of an AlGaAs/GaAs HBT in which x_B is comparable to the electron mean free path, λ_B. Fig. 6.9(a) shows two experimental spectra for electrons tunnel injected at the emitter-base junction with initial energies E = 170 and 220 meV and electron velocities 8×10^7 cm s^{-1} and 9×10^7 cm s^{-1} respectively. The p-type GaAs base has an impurity concentration $p = 3 \times 10^{18}$ cm^{-3} and a thickness x_B = 26 nm [10]. Fig. 6.9(b) shows the results of using a semiclassical Monte Carlo algorithm designed to simulate the experimental situation [11]. For E = 220 meV the "ballistic" peak at the injection energy is the dominant feature in the spectrum at a collector-base bias $V_{BC} \sim -1.2$ V. With increasing V_{BC} (decreasing ϕ_{BC}), there is a feature at $V_{BC} \sim -0.7$ V corresponding to the single scattering average energy loss $\langle \hbar\omega \rangle$ = 60 meV, scattering time $\tau \sim 39$ fs ($\gamma \sim 8.5$ meV), mean free path, $\lambda_B \sim 35$ nm, and average scattering angle $\langle \theta \rangle \sim 35^o$. At a bias $V_{BC} \sim 1$ V, $\phi_{BC} \sim 0$ eV, and both spectra show a peak associated with those electrons which have suffered multiple collisions while traversing the base. The overall agreement between the experimental (Fig. 6.9(a)) and calculated (Fig. 6.9(b)) spectra is fairly reasonable. However, the features in the calculated spectra are significantly sharper than in the measured data. This may be due to limitations in the resolution of the spectrometer or it may arise from contributions to scattering that have not been included in the calculations. It is possible that the strength of the Coulomb interaction used underestimates the actual situation. For example, the contribution to injected electron scattering of scattered electrons which accumulate near the Fermi energy has not been included. These electrons form long-lived states with a *nonequilibrium distribution* which should not be calculated using the Lindhard formalism.

6.2.4.2 Nonequilibrium electron scattering rates

Experimental evidence proving the existence of extreme nonequilibrium electron transport in a HBT does not necessarily result in a faster transistor. For example, only modest high-speed performance is obtained for a device similar to the HBT spectrometer described above [12]. Resistance and capacitance effects (*RC* time constants) have to be considered. In a useful device, a reduction in x_B requires an increase in p, the majority charge carrier density in the base, to keep the base sheet resistance R_B small, thereby minimizing *RC* time constants. For example, GaAs with a p-type carrier concentration of $p = 2 \times 10^{20}\ \mathrm{cm}^{-3}$ has a mobility $\mu \sim 50\ \mathrm{cm}^{-2}\ \mathrm{V}^{-1}\ \mathrm{s}^{-1}$, so that a sheet resistance of $R_B \leq 200\ \Omega$ per square requires $x_B > 30\ \mathrm{nm}$. The sublinear decrease in majority carrier mobility with increasing carrier concentration is the underlying reason why this strategy succeeds.

Now consider the effect increasing the *p*-type majority charge carrier density in the base has on nonequilibrium minority carrier electron transport in the base. As sketched in Fig. 6.10(a), a conduction band electron of initial wave vector **k** and energy E above the conduction band minimum is injected into the p-type base. This electron may scatter inelastically, losing energy $Y = \hbar\omega / E_F$ and changing momentum by $X = q / k_{F_{hh}}$, where E_F is the Fermi energy of the majority *p*-type carriers in the base and $k_{F_{hh}}$ is the Fermi wave vector of the heavy-hole band. The electron may also scatter elastically from statically screened ionized impurities in the base. The relevant scattering rates may be determined by appropriate use of the dielectric response function, $\varepsilon(\mathbf{q}, \omega)$ [13-14]. For example, to calculate the total inelastic scattering rate $1/\tau_{in}$ for a fixed injection energy E, one needs to integrate over the spectral weight ($\sim -\mathrm{Im}/\varepsilon(\mathbf{q}, \omega)$) of the coupled majority carrier and longitudinal optic phonon excitations in *p*-type GaAs [14]. Within effective mass theory, energy and momentum conservation (in the $T = 0$ K limit) results in a parabola of integration which, for the indicated p-type carrier concentration, leads to the broken lines plotted in Fig. 6.10(b) (note the normalization of scales in the figure). Calculation of $1/\tau_{in}$ requires integration of the spectral weight within the parabola, that is,

$$1/\tau_{in} = \frac{2m_3 e^2}{\pi \hbar^2 k} \int \mathrm{Im}\left\{\frac{-1}{\varepsilon(q, \omega)}\right\} \frac{dq}{q} d\omega \qquad (6.10)$$

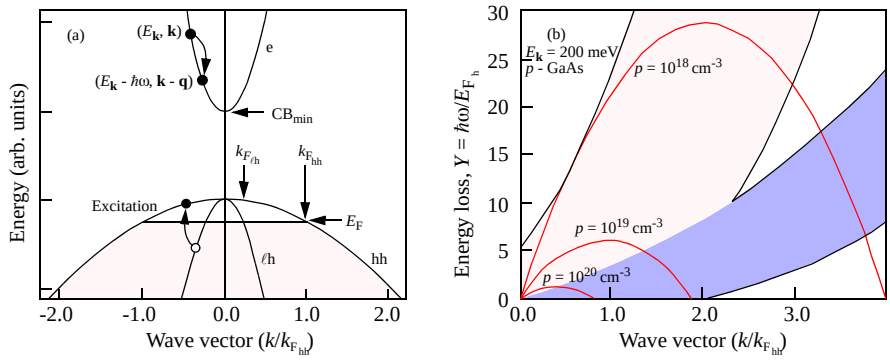


Fig. 6.10. (a) Schematic diagram of a typical III-V semiconductor band structure showing valence heavy-hole band hh with Fermi wave vector $k_{F_{hh}}$, light-hole band ℓh with Fermi wave vector $k_{F_{\ell h}}$, and conduction band e. An electron with initial energy $E_{\mathbf{k}}$ and wave vector $\mathbf{k}$ in the conduction band can inelastically scatter by creating an excitation in the valence band of energy $\hbar\omega$ and wave vector $\mathbf{q}$. The Fermi energy in the valence band is E_F (from [15]). (b) Dispersion relation for single-particle excitations in p-type GaAs at temperature $T = 0$ K. The parabolas of integration are for three different p-type carrier concentrations and for conduction band electron injection energy $E = 200$ meV (from [15]).

It is clear from the figure that small values of p involve an integration over a large portion of the heavy-hole intraband single-particle excitations (dark shaded region in Fig. 6.10(b)) which carry most of the inelastic scattering strength. With increasing p, a *reduced* portion of phase space is integrated. Although, for large values of p, the maximum scattering strength increases, over the small region of phase space in which the integration takes place the scattering strength can decrease. A consequence of this is that, with increasing carrier concentration, $1/\tau_{in}$ increases, reaches a maximum and then decreases. This fact is illustrated in Fig. 6.11(a), in which results of calculating $1/\tau_{in}$ for GaAs as a function of p are given for three different values of E [15]. For low values of $p < 10^{17}$ cm^{-3}, longitudinal polar-optic phonon scattering dominates. For $p \sim 2 \times 10^{19}$ cm^{-3} the scattering rate reaches a maximum and at very high values of $p \geq 5 \times 10^{20}$ cm^{-3} $1/\tau_{in}$ decreases, becoming less than the bare longitudinal polar-optic phonon scattering rate due to efficient screening. Thus, because with decreasing x_B one must increase p, a *fundamental* limit to device performance, $1/\tau_{in}$, can decrease. In addition, since the physics underlying the decrease in $1/\tau_{in}$ at high p is based on a phase space argument, the results are general and apply to other materials systems such as $In_{0.53}Ga_{0.47}As$ and InAs. Fig. 6.12(a) shows results of calculating $1/\tau_{in}$ as a function of p for InAs. For the sake of clarity, in the above discussion we adopted the $T = 0$ K limit. However, for large values of p, similar results apply at temperature $T = 300$ K, the only significant difference being a modest $< 30\%$ increase in the inelastic scattering rate [16]. We are therefore led to a somewhat unexpected result that, in general, inelastic minority carrier scattering rates can decrease with increasing p-type majority carrier concentration.

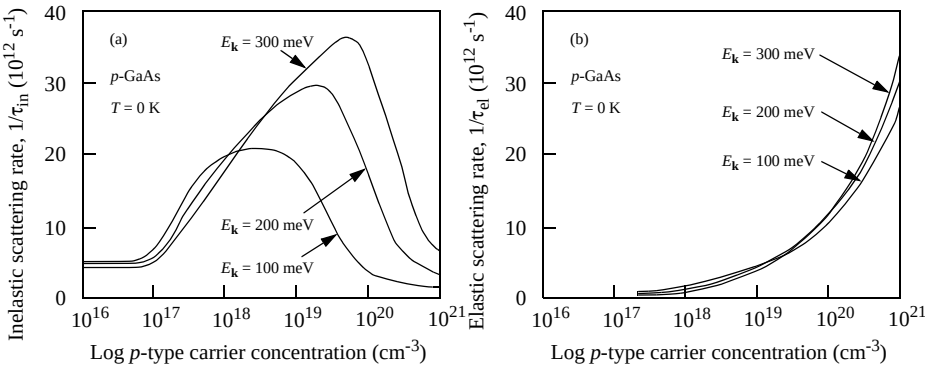


Fig. 6.11. (a) Total inelastic scattering rate $1/\tau_{in}$ as a function of *p*-type carrier concentration in GaAs for three indicated values of electron injection energy $E_{\mathbf{k}}$ (from [15]). (b) Total elastic electron scattering rate $1/\tau_{el}$ due to random ionized impurities as a function of *p*-type carrier concentration in GaAs for three indicated values of electron injection energy *E* (from [15]).

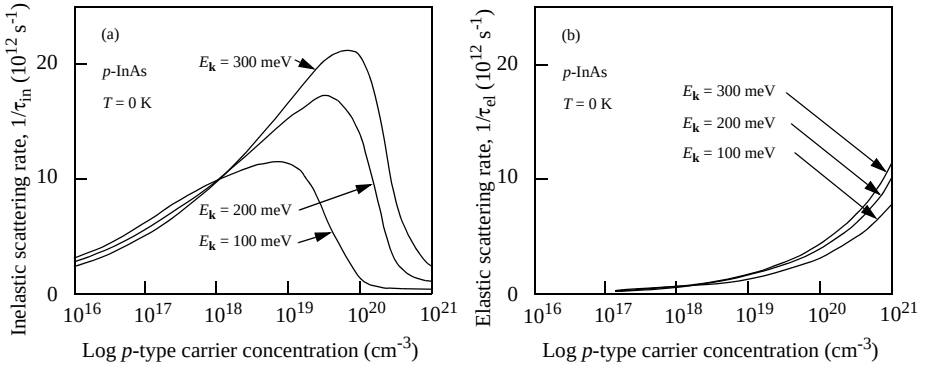


Fig. 6.12. (a) Total inelastic electron scattering rate $1/\tau_{in}$ as a function of *p*-type carrier concentration in InAs for three indicated values of electron injection energy $E_{\mathbf{k}}$. (b) Total elastic electron scattering rate $1/\tau_{el}$ due to random ionized impurities as a function of *p*-type carrier concentration in InAs for three indicated values of electron injection energy *E*.

If the high carrier concentration in the base is created using randomly positioned substitutional impurities, then it is necessary to consider the elastic scattering rate due to those ionized impurities, $1/\tau_{el}$. The results of calculating $1/\tau_{el}$ as a function of p for the same values of $E_{\mathbf{k}}$ are shown for GaAs in Fig. 6.11(b) and InAs in Fig. 6.12(b). It is clear from the figures that $1/\tau_{el}$ is the dominant scattering process for $p \geq 5 \times 10^{20}$ cm^{-3}. This contribution from $1/\tau_{el}$ may be minimized by imposing spatial correlations in dopant atomic positions to create, for example, a periodic superlattice of impurities [17]. One possible approach towards achieving this might utilize atomic layer epitaxy techniques such as delta doping. In fact it has been demonstrated that an ordered plane of high mobility *p*-type dopant atoms may be buried in Si [18-19]. Future advances in thin-film epitaxial crystal growth or nanofabrication technology may eventually provide the needed spatial control of dopant atoms giving rise to dra-

matic reduction in $1/\tau_{el}$. Another approach to minimixing the contribution of $1/\tau_{el}$ is to use modulation doping.

In addition to these considerations, there are limits on the amount of *p*-type impurities that can be placed in the base of a *n-p-n* HBT before E_F becomes larger than ΔE_{VB} and there is a significant back flow of charge carriers from the base into the emitter. This situation has been studied experimentally for different HBTs and is discussed in [20]. For abrupt junction InP/$In_{0.53}Ga_{0.47}As$ HBTs a practical limit appears to be $p \leq 5 \times 10^{20}$ cm^{-3}.

We conclude this section by considering the semiclassical transit time, τ_B, associated with nonequilibrium electron base transport. We have already mentioned that in a *n-p-n* HBT base sheet resistance may be maintained at suitably low values, $R_B < 200\ \Omega$ per square, for high *p*-type carrier concentrations, $p \geq 10^{20}$ cm^{-3}, and modest base thickness, $x_B \geq 30$ nm. It can be shown using semiclassical arguments that a GaAs HBT with injection energy $E \geq 220$ meV, $x_B \leq 30$ nm, and $p \geq 3 \times 10^{20}$ cm^{-3} has an insignificant base transit delay $\tau_B \geq 0.05$ ps. For the present we may feel comfortable that base transport does not present a significant limitation to the high-speed performance of an appropriately designed HBT.

6.2.4.3 Influence of nonequilibrium electron transport on current gain

In a classical *n-p-n* homojunction (and graded-junction) bipolar transistor electrons introduced from the emitter diffuse across the base. If the current gain is base recombination limited, base current $I_B = Q_B/\tau_n$, where Q_B is the electron charge in the base and $1/\tau_n$ is an appropriate electron recombination rate. Because base thickness is typically much less than the minority carrier diffusion length but more than the electron mean free path, $Q_B \propto x_B$ and so the base current scales as $I_B \propto x_B$. The collector current I_C is limited by electron diffusion across the base with an effective electron velocity $\upsilon_{eff} \propto D/x_B$ where *D* is the diffusion constant. Therefore, common emitter current gain $\beta = I_C/I_B \propto 1/x_B^2$. In an abrupt heterostructure bipolar transistor collector current is limited by injection at the emitter-base conduction band spike ΔE_{CB} when $\upsilon_{therm} \exp(-\Delta E_C/k_B T) \ll D/x_B$. In this expression $\upsilon_{therm} < 1 \times 10^7$ cm s^{-1} is the *x*-directed average thermal velocity in the emitter and $k_B T = 0.025$ eV is the thermal energy at room temperature. Thus, for example, the *collector current* in an abrupt junction HBT with $\Delta E_{CB} > 0.2$ eV, $x_B \geq 10$ nm and $D = 25$ cm^2 s^{-1} does not directly depend on base thickness x_B *even* if base transport is diffusive [21]. In such a transistor the base thickness dependence of β arises solely from the x_B dependence of base current, I_B.

Measured room-temperature current gain β for abrupt junction $Al_{0.48}In_{0.52}As$/$In_{0.53}Ga_{0.47}As$ HBTs with $\Delta E_{CB} = 0.47$ eV, $I_C = 100\ \mu$A, and $V_{CB} = 0$ *V* as a function of base thickness x_B is shown in Fig. 6.13. The devices used in these experiments [22] have base thicknesses in the range $20\ \text{nm} \leq x_B \leq 400$ nm, base doping level is

$p = 1.5 \times 10^{19}$ cm^{-3}, and the collector space charge region is $x_C = 300$ nm. It is clear from the data that for $x_B < 100$ nm the current gain β scales as $1/x_B$ whereas for $x_B > 100$ nm β varies approximately as $1/x_B^2$. This behavior is related to nondiffusive electron transport in the base. Naively, we expect a $1/x_B$ dependence when charge transport in the base is dominated by extreme nonequilibrium electron motion. In this simplistic picture the collector current, I_C, is independent of base thickness but the base current, I_B, is proportional to the volume of the neutral base. As expected, it is observed that, for a given base-emitter bias V_{BE}, the difference in current gain for devices with $x_B = 20$ nm and $x_B = 40$ nm is solely determined by changes in base current.

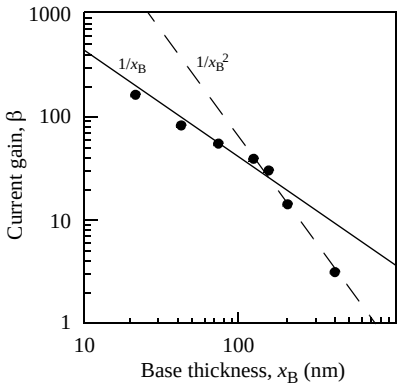


Fig. 6.13. Plot of measured room-temperature common emitter current gain, β, as a function of base thickness, x_B, for an abrupt junction *n-p-n* $Al_{0.48}In_{0.52}As/In_{0.53}Ga_{0.47}As$ HBT[19] with impurity concentration $p = 1.5 \times 10^{19}$ cm^{-3}. Bias conditions are $I_C = 100$ μA and $V_B = 0$ V. The broken line is the behavior expected for $\beta \propto 1/x_B^2$ and the solid line is for $\beta \propto 1/x_B$ (from [22]).

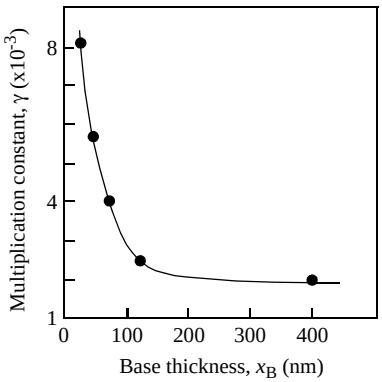


Fig. 6.14. Measured room-temperature value of the avalanche multiplication constant γ versus base thickness x_B for abrupt junction $Al_{0.48}In_{0.52}As/In_{0.53}Ga_{0.47}As$ *n-p-n* HBTs. γ was obtained by measuring the variation in current from $V_{CB} = 0.2$ V to $V_{CB} = 1.0$ V (from [22]).

The presence of extreme nonequilibrium electron transport at the base-collector junction should enhance avalanche multiplication in the collector. The reason for this is nonequilibrium electrons at the base-collector junction are already in high energy states which are energetically relatively close to the avalanche threshold energy. Enhanced avalanche multiplication in the collector has been observed in devices with small x_B thereby independently confirming the existence of extreme nonequilibrium transport in this regime. Fig. 6.14 shows the measured avalanche multiplication constant γ as a function of base thickness, x_B [22]. Here, $\gamma = I_{av}/(I_C - I_{av})$ where $I_{av} = \Delta I_C - \Delta I_B$ is the avalanche current. As may be seen in Fig. 6.14, extreme nonequilibrium electron transport in the base enhances γ when $x_B < 100$ nm. This observation correlates remarkably well with the measured dependence of current gain on

base thickness shown in Fig. 6.13 indicating that the $1/x_B$ behavior of β is related to the presence of extreme nonequilibrium electron transport at the base-collector junction.

In addition to the $1/x_B$ scaling for $x_B < 100$ nm, there exists another regime $100\ \text{nm} \le x_B \le 400\ \text{nm}$ were β might be characterized by a $1/x_B^2$ dependence (see Fig. 6.13). However, as will be demonstrated in section 6.2.2.4, electron dynamics in the base cannot be ascribed to purely diffusive charge transport. This is to be distinguished from the above mentioned classical homojunction transistor scaling behavior in which a single minority carrier diffusion constant describes devices of differing x_B and consequently $I_C \propto 1/x_B$ and $I_B \propto x_B$ giving the familiar $\beta = I_C/I_B \propto 1/x_B^2$.

6.2.4.4 Influence of nonequilibrium electron transport on forward delay

In this section measurements and numerical simulations of the intrinsic forward delay as a function of base thickness in abrupt junction $Al_{0.48}In_{0.52}As/In_{0.53}Ga_{0.47}As$ *n-p-n* heterostructure bipolar transistors will be discussed. This structure is of interest because the large conduction band off-set $\Delta E_{CB} = 0.47$ eV between the $Al_{0.48}In_{0.52}As$ emitter and the $In_{0.53}Ga_{0.47}As$ base results in extreme nonequilibrium electron transport in the base. In the previous section it was pointed out that this fact results in nonclassical scaling behavior in the static characteristics of the device. In particular, break-down voltage in the collector depends on base thickness, x_B, and current gain scales as $1/x_B$ for devices with a thin base [22]. The physics governing these phenomena are intimately related to the temporal and spatial evolution of the nonequilibrium electron distribution initially injected into the base. The purpose of the following is to explore the influence nonequilibrium electron dynamics has on the forward delay time of a transistor as a function of base thickness, x_B.

Fig. 6.15(a) shows measured total forward delay time $\tau_F = \tau_B + \tau_C$ as a function of base thickness $20\ \text{nm} \le x_B \le 400\ \text{nm}$ for an abrupt junction $Al_{0.48}In_{0.52}As/In_{0.53}Ga_{0.47}As$ HBT with base doping $p = 1.5 \times 10^{19}\ \text{cm}^{-3}$ and collector space charge region thickness $x_C = 300$ nm [23]. Here, τ_B and τ_C are the base and collector transit delays respectively. The delay time is obtained by fitting measured s-parameters to the equivalent circuit model shown in the inset of Fig. 6.15(a). Although the absolute value of τ_F may slightly depend on the equivalent circuit model or the extraction method, at present we are only concerned with the relative change (i.e., the scaling behavior) of τ_F for different x_B.

For base thickness less than approximately 100 nm the total forward delay is dominated by τ_C and is therefore insensitive to variations in x_B. The measured delay time of $\tau_F \sim \tau_C = 0.5 \pm 0.2$ ps corresponds to an average electron velocity in the collector of approximately $\upsilon_C \sim x_C/2\tau_C = 3 \times 10^7\ \text{cm s}^{-1}$. When $x_B \ge 100$ nm the base transit delay becomes important and τ_F increases. For $x_B = 400$ nm measured forward delay is $\tau_F = 6.7$ ps and, since τ_C remains essentially constant, one may estimate $\tau_B = 6.2$ ps corresponding to an average forward electron velocity in the base of $\upsilon_B \sim x_B/\tau_B = 6.4 \times 10^6\ \text{cm s}^{-1}$. Obviously, the dramatic increase in base transit time

delay can only be explained by a large change in average electron velocity with increasing base thickness, x_B.

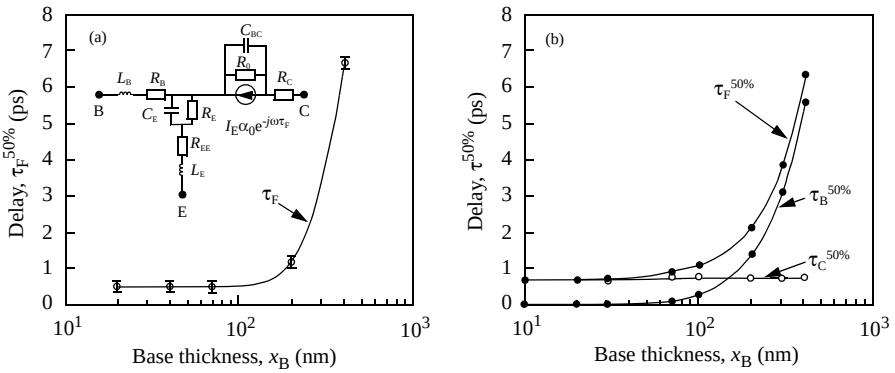


Fig. 6.15. (a) Small signal forward delay, τ_F, for different base thicknesses, x_B. The measurements were carried out at emitter current density of 5×10^4 A cm^{-2} and collector-emitter voltage bias V_{CE} = 1.5 V. The error bars are ± 0.2 ps. The inset shows the equivalent circuit model used to model the device. (b) Results of calculating $\tau_F^{50\%}$, $\tau_B^{50\%}$ and $\tau_C^{50\%}$ as a function of base thickness for the device parameters used in the experiments (from [22]).

To understand the physics underlying this experimental data it is necessary to perform numerical simulations, results of which are described in [23]. The electrons contributing to the emitter current, I_E, are initially thermally distributed in the conduction band of the wide bandgap $Al_{0.48}In_{0.52}As$ emitter. Under forward base-emitter bias, electrons are injected with approximate excess kinetic energy ΔE_{CB} = 0.47 eV into the Γ-valley conduction band of the *p*-type $In_{0.53}Ga_{0.47}As$ base. There is also the possibility that some electrons enter the base, or subsequently scatter, into the subsidiary L-valley 0.55 eV above CB_{min}. The injected electrons may scatter elastically from statically screened ionized p-type impurities or inelastically off excitations of the *p*-type majority carriers ($p = 1.5 \times 10^{19}$ cm^{-3}) while traversing the base. Electrons in the conduction band may also scatter off other conduction band electrons. Having traversed the base, electrons are accelerated in the electric field of the reverse-biased collector space-charge region. Here they may suffer inelastic collisions with phonons and other electrons. Those electrons which gain enough kinetic energy may also transfer from the non-parabolic Γ-valley to the high effective electron mass L- and X-valley conduction band minima.

Fig. 6.15(b) shows results of calculating the forward delay for 50% of step-injected emitter current to flow through the device, $\tau_F^{50\%}$. Also shown are the contributions of $\tau_B^{50\%}$ and $\tau_C^{50\%}$ as a function of base thickness, x_B. The results are in good qualitative agreement with the experimental data shown in Fig. 6.15(a) even though the measured delay, τ_F is not identical to the calculated $\tau_F^{50\%}$. The numerical simulations clearly illustrate that, although base transit delay increases with increasing x_B, τ_B does not become greater than τ_C until $x_B \geq 135$ nm. In contrast, for a similar transistor but with

diffusive base transport and diffusion constant $D = 25\ \text{cm}^2\ \text{s}^{-1}$, τ_B becomes greater than τ_C when $x_B \geq 60\ \text{nm}$. Hence, the abrupt emitter / base junction HBT can have more than double the base thickness of the corresponding graded junction HBT without suffering a penalty in τ_F.

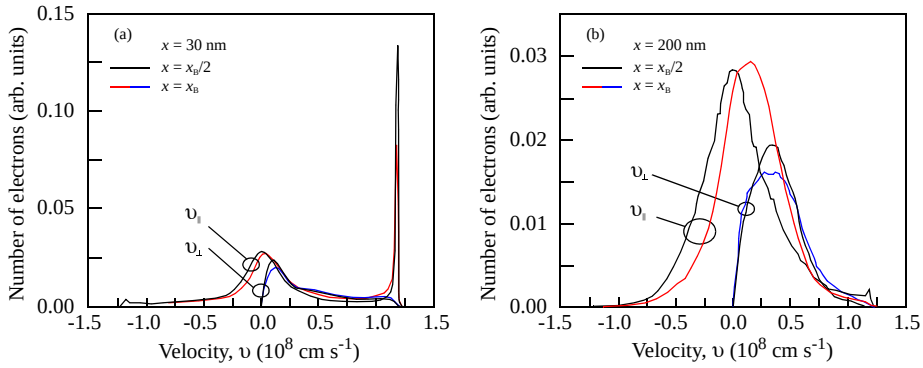


Fig. 6.16. Results of calculating the longitudinal (perpendicular to the plane of the heterostructure) electron velocity, $\upsilon_\parallel$, and transverse electron velocity, $\upsilon_\perp$, distribution in the base of an $Al_{0.48}In_{0.52}As/In_{0.53}Ga_{0.47}As$ *n-p-n* HBT for $x = x_B/2$ and $x = x_B$ with (a) $x_B = 30\ \text{nm}$ and (b) $x_B = 200\ \text{nm}$ (from [23]).

To establish the role played by nonequilibrium electron transport in determining the trends in Fig. 6.15(a) and (b), Fig. 6.16 shows longitudinal and transverse electron velocity distribution in the base at $x = x_B/2$ and $x = x_B$ for (a) $x_B = 30\ \text{nm}$ and (b) $x_B = 200\ \text{nm}$. On the one hand, for a base thickness of $x_B = 30\ \text{nm}$ the peak in longitudinal electron velocity at approximately $1.2 \times 10^8\ \text{cm s}^{-1}$ is due to extreme nonequilibrium electron transport across the base. On the other hand, the absence of any significant high velocity peak for $x_B = 200\ \text{nm}$ is a clear indication that electrons have suffered many elastic and inelastic collisions. Never-the-less, for $x_B = 200\ \text{nm}$ the electron distribution function has not thermalized to the lattice temperature, $T = 300\ \text{K}$. Of course, if the electron distribution could be described by some effective electron temperature T_e, one might consider introducing an effective diffusion constant $D(T_e)$ to describe electron motion. However, as illustrated in Fig. 6.17 such an approach fails because, even for $x_B = 400\ \text{nm}$, the nonequilibrium electron distribution function *cannot* be described by an effective electron temperature, T_e.

Although high energy injection in abrupt junction $Al_{0.48}In_{0.52}As/In_{0.53}Ga_{0.47}As$ HBTs results in nonequilibrium electron transport which does, in fact, reduce base transit delay, it may not be the optimum design in a transistor if device speed is limited by resistive and capacitive parasitics. In such a case it is necessary to carefully evaluate the relative advantages and disadvantages of abrupt versus graded emitter heterostructures. In a graded device the turn on voltage is reduced by approximately ΔE_{CB} resulting in reduced power dissipation. Furthermore, the emitter current ideality factor, *n*,

will be close to unity (compared to around $n = 1.5$ for an abrupt junction $Al_{0.48}In_{0.52}As/In_{0.53}Ga_{0.47}As$ HBT) resulting in a significant reduction in emitter dynamic resistance $R_E = nk_BT/eI_E$, where k_BT is the thermal energy, and e is the electron charge. Therefore, if the base delay for diffusive transport can be tolerated, as usually occurs in a device dominated by parasitics, it may be beneficial to grade the emitter / base junction. On the other hand, it is important to note that the use of an abrupt emitter / base junction in the transistor allows for more than a factor of two thicker base for a given transit delay and this, of course, results in a lower base resistance. Reduction of base resistance is particularly important in devices with relatively large emitter dimensions. Thus, nonequilibrium electron transport provides an opportunity to reduce base resistance without suffering a penalty in base transit time. For further discussion on the relative merits of graded and abrupt junction HBTs see section 6.2.5.

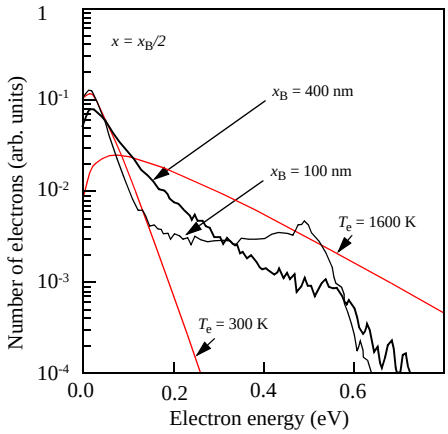


Fig. 6.17. Calculated electron energy distribution function at $x = x_B/2$ for $x_B = 100$ nm and $x_B = 400$ nm. Dotted curves are the thermal electron distribution corresponding to an effective electron temperature, T_e (from [23]).

6.2.5 Effect of reducing x_C on collector transport

This section discusses the effect reducing collector thickness has on electron transport through the collector space charge region of a HBT. First, we note that because current flow through the collector obeys Maxwell's equations, we require $\nabla \cdot (\mathbf{j}_{el} + \partial/\partial t[\mathbf{D}/(4\pi)]) = 0$ where $\mathbf{j}_{el}$ is electron current density and $\mathbf{D}$ is electric displacement field. For electrons moving in the $\mathbf{x}$ direction at constant average velocity υ_0 and in the limit of small signal low frequency response, this leads to the well known result for collector delay $\tau_C = x_C/2\upsilon_0$ [24]. Thus, the characteristic electron transit time is $\tau_{Ce} = 2\tau_C$. The factor of two arises from the electric displacement field's contribution to current flow in the collector.

Since total signal delay in a transistor is $\tau_{EC} \sim \tau_B + \tau_C + R_E(C_E + C_C) + \ldots$, where $R_E = nk_BT/eI_E$ (I_E is emitter current and $n \geq 1$ is the emitter base ideality factor),

C_E is emitter capacitance, and $C_C = a/x_C$ is collector capacitance (a is a constant), one may now vary x_C to optimize τ_{EC}. This requires $\partial\tau_{EC}/\partial x_C = 1/2\upsilon_0 - R_E a/x_C^2 = 0$ so that $x_C = (2R_E a\upsilon_0)^{1/2}$ and hence an expression for the emitter-collector delay is $\tau_{EC} \sim \tau_C + R_E C_C = 2(R_E a/2\upsilon_0)^{1/2}$. One is therefore led to the incorrect conclusion that increasing electron velocity in the collector by an order of magnitude only enhances optimized transistor speed by a factor of approximately three. The error in this type of simplistic scaling argument is obvious. The average velocity at which nonequilibrium electrons move through the collector is, among other things, determined by collector thickness, x_C. This, of course, makes optimization of transistor delay, τ_{EC}, a more complicated (and interesting) procedure.

We now consider the limit of extreme nonequilibrium transport in a collector of thickness x_C. In fact, it has been shown experimentally that under moderate bias conditions and for $x_C \leq 50$ nm collector transit times can be so short that a significant number of electrons transit the collector without substantial inelastic scattering [25]. For small x_C "ballistic" transport depends on the collector-base bias, V_{CB}, and the width in energy of the conduction band, E_{bw}. An electron in the collector arm may only be accelerated to an energy $E \leq E_{bw}$ before interband scattering must take place. To avoid this we require $eV_{CB} \leq E_{bw} - E_g - E_{F_e} - E_{F_h}$, where E_{F_e} is the Fermi energy in the *n*-type subcollector contact, and E_{F_h} is the Fermi energy of holes in the *p*-type base. Ignoring the contribution from E_{F_e} and E_{F_h}, we note that for GaAs at room temperature $E_{bw} = 1.85$ eV and $E_g = 1.4$ eV, giving $eV_{CB}(\max) = 0.45$ eV, and for $In_{0.53}Ga_{0.47}As$ $E_{bw} = 2.15$ eV, $E_g = 0.75$ eV giving a somewhat larger $eV_{CB}(\max) = 1.4$ eV. These examples are indicative of a trend which, as we shall see, predicts that the high frequency performance of $In_{0.53}Ga_{0.47}As$ is superior to that of GaAs.

Monte Carlo techniques have been used to estimate the semiclassical transit time τ_{Ce} ($\sim 2\tau_C$) associated with electron transport in a GaAs collector of thickness x_C [11]. The device simulated in this case has emitter injection energy $E = 220$ meV, $p = 3 \times 10^{20}$ cm^{-3}, and $x_B = 25$ nm. Fig. 6.18 shows the impulse response transit time for 50% of electrons to traverse a collector of thickness $x_C = 100$ nm as a function of collector-base bias, V_{CB}. The electron transit time increases from $\tau_{Ce} \sim 0.4$ ps at $V_{CB} \sim -1$ V to $\tau_{Ce} \sim 1$ ps at $V_{CB} \sim 0$ V. This is because at $V_{CB} \sim -1$ V most electrons transiting the collector remain in the high-velocity Γ-valley, whereas for $V_{CB} \sim 0$ V the majority of electrons scatter into the low-velocity L and X-valleys. In realistic applications we expect the base-collector junction to be reverse biased, i.e., $V_{CB} > 0$ V. However, for $V_{CB} > 0$ V velocity overshoot (velocities in excess of the steady state equilibrium velocity-field characteristic shown in Fig. 2.3) is unimportant and τ_{Ce} is essentially established by the saturation velocity ($\sim 1 \times 10^7$ cm s^{-1}) in the collector [26]. Thus, for $x_C = 100$ nm, electron collector transit time $\tau_{Ce} \sim 1.0$ ps ($\tau_{Ce} \gg \tau_B \sim 0.05$ ps) limits the intrinsic performance of the device. Under these operating conditions, for devices in which collector transit time is important, there is no particular advantage using GaAs in preference to Si which has the same saturation velocity. It is worth noting that, in contrast to GaAs, smaller bandgap semiconductors

such as $In_{0.53}Ga_{0.47}As$ with large intervalley energy separation have a useful operating range of $V_{CB} > 0$ V in which collector velocity overshoot may be used to reduce τ_{Ce} [27]. To illustrate this, Fig. 6.19 shows results of calculating the electron transit time τ_{Ce} as a function of total potential drop in the collector arm, ϕ_{BC}, for GaAs and $In_{0.53}Ga_{0.47}As$. As may be seen, for a reasonable value of $\phi_{BC} \sim 1$ eV, $In_{0.53}Ga_{0.47}As$ has more than a factor of two faster intrinsic transit time than GaAs. Not surprisingly, the fastest HBTs are fabricated from $In_{0.53}Ga_{0.47}As$ rather than GaAs [28].

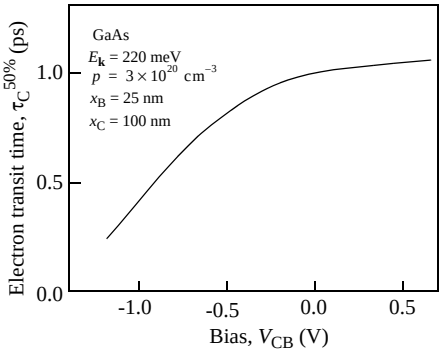


Fig. 6.18. Collector electron transit time, τ_{Ce}, in GaAs as a function of base-collector voltage, V_{CB}, for emitter injection energy $E_{\mathbf{k}} = 220$ meV, $x_B = 25$ nm, $p = 3 \times 10^{20}$ cm^{-3}, $T = 300$ K, and $x_C = 100$ nm (from [11]).

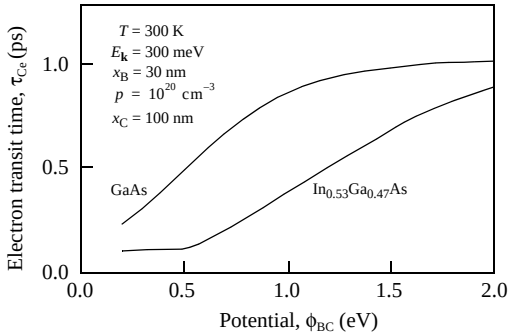


Fig. 6.19. Collector electron transit time, τ_{Ce}, in GaAs and $In_{0.53}Ga_{0.47}As$ as a function of base-collector potential energy drop ϕ_{BC} for emitter injection energy $E_{\mathbf{k}} = 300$ meV, $x_B = 30$ nm, $p = 1 \times 10^{20}$ cm^{-3}, $x_C = 100$ nm, and $T = 300$ K.

It is clear that collector transit time depends on the distribution of nonequilibrium electrons impinging on the base-collector junction and is therefore linked to base transport dynamics. To illustrate this, imagine an extreme case in which electrons are injected at energy E and suffer no electron scattering in the base region. In this situation, the momentum distribution of electrons arriving at the base-collector junction is sharply peaked in the forward direction and electrons accelerated in the collector's electric field rapidly reach the energy threshold for Γ-X intervalley scattering, impact ionization, etc. This experimentally verified effect [29], is important in HBTs fabri-

cated using $In_{0.53}Ga_{0.47}As$ for which nonequilibrium electron transport plays a significant role in device performance. An additional consequence is that collector transit time cannot be specified independently of the nonequilibrium base transport parameters E, p, and x_B.

Before concluding this section it is worth mentioning that fluctuations in current flow occur in the device. A current pulse of density 10^4 A cm^{-2} and duration 1 ps flowing through an area 1 μm^2 contains 625 electrons. Semiclassically we might expect a 4% fluctuation in (shot noise) current density. Statistical variations of this type become increasingly important in small devices operating at high-speed. However, in principle, fluctuations can also result in correlations and instabilities in electron current flowing through the active region of the device [30-31]. These instabilities, which are well known in plasma physics [32], have yet to be observed in semiconductor devices. The probable reason for this is a high threshold current density below which the plasma instability will not grow due to the presence of damping (such as phonon emission) in the system.

Another type of instability in the collector arm of a HBT is based on the transferred electron Gunn-Hilsum effect. Simulations indicate that in an appropriately designed device electron density in the collector arm is unstable and oscillatory collector current can occur [33]. Under normal circumstances the Kirk effect suppresses such oscillations by reducing the electric field in the critical domain-forming region of the collector.

6.2.6 High frequency performance

A consequence of the existence of extreme nonequilibrium electron transport in abrupt $Al_{0.48}In_{0.52}As/In_{0.53}Ga_{0.47}As$ and $InP/In_{0.53}Ga_{0.47}As$ HBTs is the fact that base and collector transport are dependent quantities. In this situation the intrinsic collector delay τ_C depends on the base transit delay τ_B and, therefore, the relevant single physical quantity to be measured is the intrinsic (i.e. without parasitic charging times) emitter-collector delay τ_F. Of course, τ_F is itself controlled by electron dynamics in the base and collector, so knowledge about electron scattering mechanisms (see section 6.2.2.2) leads to a more complete understanding of device operation.

Experiments have been performed in which the high frequency current gain h_{21} of abrupt junction $InP/In_{0.53}Ga_{0.47}As$ HBTs were studied in the temperature range 55 K $\leq T \leq$ 340 K [28]. Transistors with an emitter area $A_E = 3.5 \times 3.5$ μm^2 were fabricated from crystals with the layer structure shown in Table 6.2. The ideality of the emitter-base junction is $n = 1.3$ and that of the collector is $n = 1$. Another point worth mentioning is that, because of the high base doping $p = 1 \times 10^{20}$ cm^{-3}, the base sheet resistance is low, $R_B = 400$ Ω per square. This low sheet resistance is important as it allows some design flexibility. In particular, a wider emitter finger can be used without significantly increasing the parasitic resistance. This relaxes lateral scaling limitations and results in a small emitter area to collector area ratio of $1/1.3$.

TABLE 6.2. *InP/$In_{0.53}Ga_{0.47}As$ n-p-n HBT layer structure*

Layer	Material 2	Doping (cm^{-3})	Thickness (nm)
Cap	*n*-$In_{0.53}Ga_{0.47}As$	7×10^{19}	200
Emitter	*n*-InP	1×10^{18}	300
Base	*p*-$In_{0.53}Ga_{0.47}As$	1×10^{20}	50
Collector	*n*-$In_{0.53}Ga_{0.47}As$	2×10^{16}	300
subcollector	*n*-$In_{0.53}Ga_{0.47}As$	5×10^{19}	250
Substrate	Semi-insulating Fe: InP		

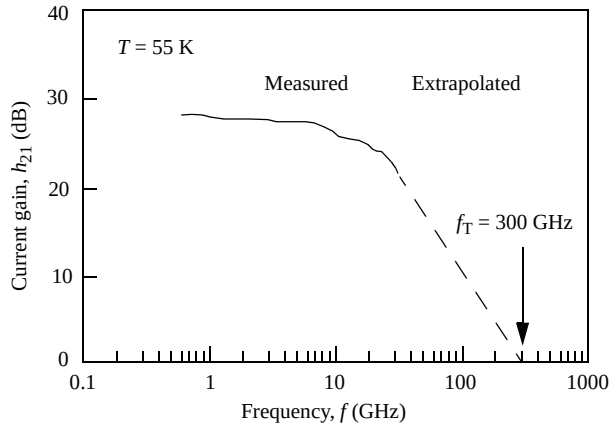


Fig. 6.20. h_{21} as a function of frequency for an abrupt $InP/In_{0.53}Ga_{0.47}As$ *n-p-n* HBT at temperature $T = 55\ K$. The extrapolated cut-off frequency is $f_T = 300\ GHz$. $V_{CE} = 1.2\ V$, $I_C = 15\ mA$, and $A_E = 1.2 \times 10^{-7}\ cm^2$ (from [28]).

The device operates by injecting electrons from the wide bandgap InP emitter into the $In_{0.53}Ga_{0.47}As$ base with an excess kinetic energy of $E \sim 0.26\ eV$ and initial velocity of around $1 \times 10^8\ cm\ s^{-1}$. In this device the static common emitter current gain $\beta \sim 50$ is limited by back injection of holes from the heavily doped $p = 1 \times 10^{20}\ cm^{-3}$ base into the emitter [20]. While traversing the base of thickness each electron has a probability of inelastically or elastically scattering by angle θ and changing kinetic energy by $\hbar\omega$. After traversing the base a nonequilibrium distribution of electrons encounters and accelerates in the electric field of the collector. While in the collector electrons may experience various scattering processes including polar-optic, intervalley and electron-electron collisions. All of these events contribute to the intrinsic collector delay.

The forward current gain h_{21} derived from s-parameters follows a - 6 dB / octave roll-off and extrapolates to the short circuit unity gain cut-off frequency, f_T, at $h_{21} = 0$. The total emitter-collector delay time, τ_{EC} is related to f_T via

$$\frac{1}{2\pi f_T} = \tau_{EC} = \tau_E + \tau_{CC} + \tau_F \tag{6.11}$$

where τ_E is the emitter charging time, τ_{CC} is the collector charging time and τ_F is the intrinsic forward delay. Contributing to τ_F is the base transit delay, τ_B and the collector space charge delay τ_C. The intrinsic forward delay may be accurately determined by equivalent circuit modeling. Fig. 6.20 shows typical measured dependence of h_{21} on frequency at a temperature $T = 55$ K, which extrapolates to f_T greater than 300 GHz.

Fig. 6.21 shows as a function of temperature, *T*, the values of τ_F extrapolated from the measured s-parameter data. The small signal data is that for an abrupt InP/$In_{0.53}Ga_{0.47}As$ *n-p-n* HBT with emitter area $A_E = 1.2 \times 10^{-7}$ cm^2 under bias $V_{CE} = 1.2$ V and $I_C = 20$ mA. At a temperature $T = 340$ K the forward delay is $\tau_F = 0.5 + 0.05$ ps. With decreasing temperature τ_F becomes smaller saturating at a value $\tau_F \sim 0.28 \pm 0.05$ ps for $T \leq 150$ K. As seen from Table 6.3, the dominant temperature dependent elements, as determined by their effect on the delay time, are τ_F, R_E, and R_B. R_E is proportional to the temperature and R_B is controlled by carrier mobility in the heavily doped base.

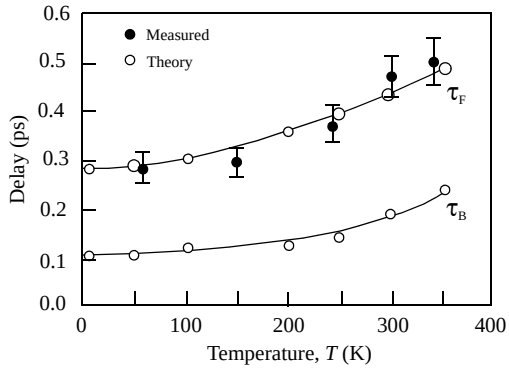


Fig. 6.21. Values of τ_F extracted from the measured s-parameter data (solid dots) and result of numerically simulating (open dots) small-signal delay τ_F and τ_B as a function of temperature, *T*. The device is an abrupt junction $InP/In_{0.53}Ga_{0.47}As$ *n-p-n* HBT. Emitter area is $A_E = 1.2 \times 10^{-7}$ cm^2 and device layer structure is the same as shown in Table 6.3. The solid line in the figure is a least-square fit to the simulated data points (from [28]).

TABLE 6.3. *Small-signal parameters for a InP/$In_{0.53}Ga_{0.47}As$ n-p-n HBT with* $15\ mA \le I_C \le 20\ mA$, $V_{CE} = 1.2\ V$, *and* $A_E = 1.2 \times 10^{-7}\ cm^2$.

Parameter	T = 55 K	T = 150 K	T = 240 K	T = 300 K	T = 340 K
τ_F (ps)	0.27	0.28	0.38	0.45	0.50
τ_{EC} (ps)	0.50	0.88	1.06	1.18	1.24
α	0.98	0.98	0.97	0.97	0.97
C_{je} (fF)	200.1	233.9	227.2	225.6	205.7
C_{jc} (fF)	9.1	13.3	13.9	14.3	12.8
C_{fb} (fF)	15.2	21.9	20.0	19.8	19.5
R_{EE} (Ω)	5.1	9.2	8.3	8.7	8.1
R_{BB} (Ω)	1.60	2.03	1.73	1.23	1.12
R_B (Ω)	48.6	53.7	58.1	61.6	65.3
R_C (Ω)	2.3	4.33	3.41	3.68	3.37
R_E (Ω)	0.29	0.47	1.14	1.32	1.61
R_{FB} (Ω)	7.8	12.0	11.8	10.6	10.5

The open circles in Fig. 6.21 show results of using Monte Carlo methods to calculate total small signal delay, τ_F, in the HBT as a function of temperature, T (see Table 6.4). Also shown in the figure is the small-signal base transit delay τ_B as a function of temperature. The agreement between experiment and calculation is satisfactory. The physics underlying the temperature dependence of τ_F is related to the temperature dependence of $1/\tau_{in}$. At low temperatures ($T \le 50$ K) the base transit delay $\tau_B \sim 0.11$ ps is about one third of the total forward delay $\tau_F \sim 0.30$ ps. The average electron velocity in the base is $\upsilon_B = 4.5 \times 10^7$ cm s^{-1} and in the collector it is $\upsilon_C \sim 8 \times 10^7$ cm s^{-1}. At room temperature ($T = 300$ K) $\tau_B \sim 0.19$ ps is almost half the total delay $\tau_F \sim 0.43$ ps. In this case $\upsilon_B = 2.6 \times 10^7$ cm s^{-1} and $\upsilon_C \sim 6 \times 10^7$ cm s^{-1}. The high average electron velocity in the base arises from non-equilibrium electron transport. A diffusive transport model cannot be justified as it would imply an unreasonably high value for minority carrier mobility of $\mu(T = 50\ \text{K}) = 2.7 \times 10^4$ cm^2 V^{-1} s^{-1} at low temperatures and $\mu(T = 300\ \text{K}) = 2.5 \times 10^3$ cm^2 V^{-1} s^{-1} at room temperature for an $In_{0.53}Ga_{0.47}As$ p-type majority carrier concentration of $p = 1 \times 10^{20}$ cm^{-3}. In addition, attempts to fit the electron distribution function to an effective electron temperature, T_e, greater than the lattice temperature fail. It is therefore inappropriate to introduce an "effective diffusion constant".

The observed temperature dependence of intrinsic forward delay, τ_F, is due to variation in quasi-electron scattering with temperature and carrier density. This may be illustrated by considering the inelastic scattering rate, $1/\tau_{in}$ for a central Γ-valley conduction band quasi-electron of initial kinetic energy $E = 150$ meV. In the neutral

base quasi-electrons may scatter elastically from statically screened ionized p-type impurities or inelastically, changing kinetic energy by $\hbar\omega$ and changing momentum by **q**. Quasi-electron scattering rates may be calculated within the random phase approximation using a method outlined in Chapter 5 or Ref. [36]. Within this formalism, the dielectric response function is

$$\varepsilon(\mathbf{q},\omega) = \varepsilon_\infty\left\{1 - \frac{(\omega_{LO}^2 - \omega_{TO}^2)}{\omega(\omega + i\gamma) - \omega_{TO}^2}\right\} + \chi_e(\mathbf{q},\omega) \tag{6.12}$$

where ε_∞ is the high frequency dielectric constant, **q** is the scattered wave vector, $\hbar\omega$ is the change in energy, ω_{LO} is the longitudinal optic phonon frequency, ω_{TO} is the transverse optic phonon frequency, γ is a collision broadening term, and $\chi_e(\mathbf{q},\omega)$ is the electron contribution. In this case the inelastic scattering rate, $1/\tau_{in}$, is calculated using

$$1/\tau_{in} = \sum_{k'} \frac{8\pi e^2}{\hbar q^2}\mathrm{Im}\left[\frac{-1}{\varepsilon(\mathbf{q},\omega)}\right] g(-\hbar\omega)(1-f(E-\hbar\omega)) \tag{6.13}$$

where the Fermi function $f(E) = 1/(\exp[(E-\mu)/k_BT] + 1)$, the Bose function $g(E) = 1/(\exp[E/k_BT] - 1)$, μ is the chemical potential and k_B is Boltzmann's constant.

To illustrate these ideas, Fig. 6.22(a) shows results of calculating inelastic quasi-particle scattering rate for injected nonequilibrium conduction band quasi-particles of initial energy, $E_\mathbf{k}$ in $In_{0.53}Ga_{0.47}As$ with n-type carrier concentration $n = 1\times10^{17}$ cm^{-3} and the indicated temperatures. Fig. 6.22(b) shows results of calculating the coupled electron-electron / optical phonon scattering rate in $In_{0.53}Ga_{0.47}As$ as a function of temperature for carrier densities $n = 0$ cm^{-3}, $n = 1\times10^{16}$ cm^{-3}, $n = 1\times10^{17}$ cm^{-3} and $n = 1\times10^{18}$ cm^{-3}. In both cases optical phonon scattering is important, however, for $n = 1\times10^{17}$ cm^{-3} the electron-electron interaction is appreciable, leading to a more rapid increase in $1/\tau_{in}$ with increasing temperature. For example, the total inelastic scattering time for a nonequilibrium conduction band quasi-electron of energy $E_\mathbf{k} = 150$ meV in the presence of $n = 1\times10^{17}$ cm^{-3} other electrons in thermal equilibrium is near 50 fs at $T = 300$ K compared to 125 fs for $n = 0$ cm^{-3}.

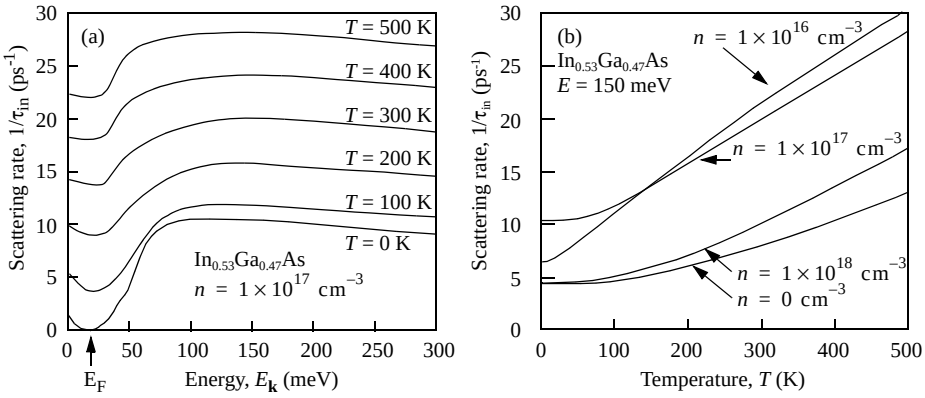


Fig. 6.22. (a) Calculated inelastic quasi-particle scattering rate in $n = 1 \times 10^{17}\ \text{cm}^{-3}$ $In_{0.53}Ga_{0.47}As$ as a function of electron energy $E_{\mathbf{k}}$ and temperature. (b) Temperature dependence of inelastic scattering rate for quasi-particles of energy $E_{\mathbf{k}} = 150\ \text{meV}$ for the indicated carrier densities. Parameters used in the calculations are effective electron mass $m_e = 0.042\ m_0$, high frequency dielectric constant $\varepsilon_\infty = 11.34$, longitudinal optic-phonon energy $\hbar\omega_{LO} = 34.5\ \text{meV}$, and transverse optic-phonon energy $\hbar\omega_{TO} = 31.2\ \text{meV}$.

TABLE 6.4. *Calculated small-signal delay for a InP/$In_{0.53}Ga_{0.47}As$ n-p-n HBT with* $I_C = 15\ mA$, $V_{CE} = 1.2\ V$, *and* $A_E = 1.2 \times 10^{-7}\ cm^2$.

T (K)	τ_F (ps)	τ_B (ps)
5	0.28	0.11
50	0.30	0.11
100	0.31	0.12
200	0.36	0.12
250	0.38	0.13
300	0.43	0.19
350	0.50	0.24

The above discussion of τ_F refers to a collector-base bias voltage of $V_{CB} \sim 0.3\ \text{V}$. In this situation base transport is more temperature sensitive than collector transport. The relative temperature dependent contributions of Γ-L and Γ-X scattering do not dominate. At larger reverse bias (e.g., $V_{CB} = 1\ \text{V}$), more electrons scatter into the X- and L-valleys and the relative importance of the temperature dependence of phonon and electron-electron scattering in these valleys increases. Fig. 6.23 shows how this manifests itself as an enhanced temperature sensitivity of τ_F with increasing temperature and increasing V_{CB}. At values of $V_{CB} \leq 0.2\ \text{V}$ the temperature sensitive collector diffusion capacitance causes the experimentally determined τ_F to increase.

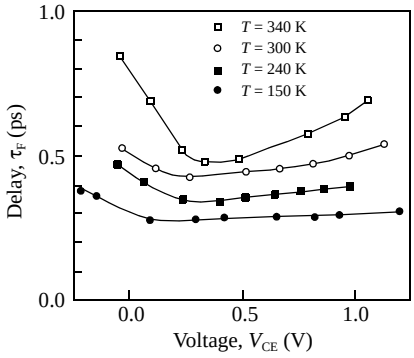


Fig. 6.23. Small signal forward delay, τ_F, extracted from measured s-parameter measurements for various collector-base biases, V_{CB} and various temperatures, T. The device is an abrupt $InP/In_{0.53}Ga_{0.47}As$ *n-p-n* HBT. $15\ mA \le I_C < 20\ mA$, $A_E = 1.2 \times 10^{-7}\ cm^2$, and device layer structure is the same as shown in Table 6.3. The solid line in the figure is to guide the eye (from [28])

6.2.7 Comparison of graded and abrupt junction HBTs

Much progress has been made fabricating high-speed *n-p-n* HBTs lattice matched to InP. Very high-speed circuits have been demonstrated [37-38] as well as discrete devices [30]. There has, however, been little attempt to optimize and quantify the intrinsic merits of various transistor designs. In the following we compare the results of numerically simulating charge transport in graded and abrupt junction *n-p-n* HBTs lattice matched to InP [39].

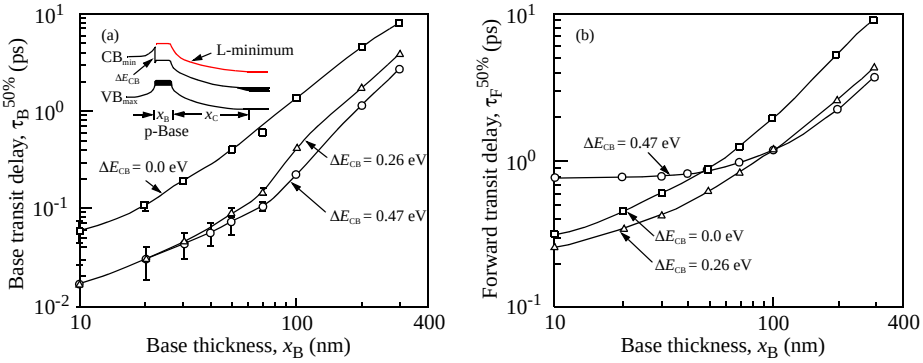


Fig. 6.24. (a) Delay, $\tau_B^{50\%}$, for different base thicknesses, x_B and the indicated values of ΔE_{CB}. The simulation is for HBTs with base doping $p = 1 \times 10^{19}\ cm^{-3}$ operated at temperature $T = 300\ K$, current density $5 \times 10^4\ A\ cm^{-2}$, $x_C = 300\ nm$, the collector space charge region is doped $n = 2 \times 10^{16}\ cm^{-3}$, and collector-base voltage $V_{CB} = 0.25\ V$. The inset shows a schematic band diagram of an abrupt junction $Al_{0.48}In_{0.52}As/In_{0.53}Ga_{0.47}As$ *n-p-n* HBT under forward bias. (b) Results of calculating $\tau_F^{50\%}$, as a function of x_B and ΔE_{CB} for the same device as in (a) (from [39]).

The inset in Fig. 6.24(a) shows a schematic band diagram of a typical $Al_{0.48}In_{0.52}As/In_{0.53}Ga_{0.47}As$ *n-p-n* HBT under forward bias. Qualitatively, we expect nondiffusive

electron transport in such a device to provide a design window in which base thickness may be increased to reduce base resistance without significantly sacrificing intrinsic forward delay time.

To quantify these ideas numerical simulations of electron transport in HBTs with different conduction band off-sets ΔE_{CB}, base thickness x_B, and different p-type impurity concentration in the base were performed. In the model, an initially thermal distribution of electrons in the forward biased wide bandgap emitter-base junction are injected with approximate excess kinetic energy ΔE_{CB} ($\Delta E_{CB} = 0$ eV for graded emitter) into the Γ-valley conduction band of the p-type $In_{0.53}Ga_{0.47}As$ base. For large ΔE_{CB}, some fraction of electrons enter the base, or subsequently scatter, into the subsidiary L-valley 0.55 eV above CB_{min} (see inset Fig. 6.24(a)). In the model elastic and inelastic electron scattering processes such as electron-electron, electron-phonon and intervalley transfer are taken into account. Electron transfer from the non-parabolic Γ-valley to the high effective electron mass L- and X-valley conduction band minima are particularly important in the reverse biased collector space-charge region.

Fig. 6.24(a) shows results of calculating base transit delay, $\tau_B^{50\%}$, for 50% of step-injected emitter current to flow through the base. The simulation is for a HBT operated at temperature $T = 300$ K, with an $In_{0.53}Ga_{0.47}As$ base doped $p = 1 \times 10^{19}$ cm^{-3}, current density $j = 5 \times 10^4$ A cm^{-2}, $x_C = 300$ nm, $In_{0.53}Ga_{0.47}As$ collector space charge region doped $n = 2 \times 10^{16}$ cm^{-3}, and collector-base voltage $V_{CB} = 0.25$ V. The three $\tau_B^{50\%}$ versus x_B curves are for $\Delta E_{CB} = 0.47$ eV, $\Delta E_{CB} = 0.26$ eV and $\Delta E_{CB} = 0$ eV corresponding to an abrupt $Al_{0.48}In_{0.52}As$, InP and graded emitter respectively. The numerical simulations show that $\tau_B^{50\%}$ can be maintained at a small value for a wider range of x_B by using an abrupt emitter-base junction with large ΔE_{CB}. This result is a direct consequence of the existence of extreme nonequilibrium transport in the transistor base for devices with large ΔE_{CB}.

Fig. 6.24(b) shows calculated forward delay $\tau_F^{50\%} = \tau_B^{50\%} + \tau_C^{50\%}$ as a function of base thickness for the same device designs as used in Fig. 6.24(a). Here one sees the influence collector delay $\tau_C^{50\%}$ has on forward delay $\tau_F^{50\%}$. Devices with $\Delta E_{CB} = 0.47$ eV have larger $\tau_F^{50\%}$ for small x_B since, unlike a classical transistor in which base and collector charge transport are decoupled, nonequilibrium electron transport dynamics in the base can increase collector delay, $\tau_C^{50\%}$. This occurs because, for small x_B, electrons arrive at the base-collector junction with significant excess kinetic energy, are readily accelerated to high energies by the electric field in the collector space charge region, and therefore efficiently scatter into low velocity L- and X-valley subsidiary minima. Under these circumstances, collector delay can be greater than in the corresponding diffusive device.

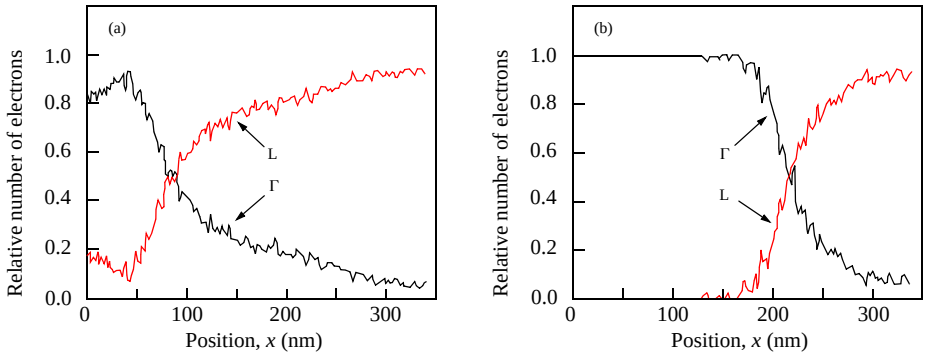


Fig. 6.25. Calculated steady-state Γ - and L-valley population as a function of position, x, for device of Fig. 6.22 with $x_B = 40$ nm and (a) $\Delta E_{CB} = 0.47$ eV, (b) $\Delta E_{CB} = 0$ eV (from [39]).

Fig. 6.25 illustrates this by plotting the steady-state Γ - and L-valley population as a function of position, x, in a device with $x_B = 40$ nm, $p = 1 \times 10^{19}$ cm^{-3} and (a) $\Delta E_{CB} = 0.47$ eV and (b) $\Delta E_{CB} = 0$ eV. The high injection energy used for Fig. 6.25(a) results in a small number of electrons being introduced into the L-valley. In addition, electrons maintain their excess kinetic energy while traversing the thin base so that electrons accelerated in the collector space-charge region rapidly transfer into the low velocity L-valley. By way of contrast, Fig. 6.25(b) shows that for $\Delta E_{CB} = 0$ eV electrons remain in the high velocity Γ-valley for most of the time while traversing the collector space-charge region of the device.

Fig. 6.26(a) and (b) shows results of calculations similar to those used to generate Fig. 6.24 but now for HBTs with $p = 1 \times 10^{20}$ cm^{-3}. Clearly, the underlying trends established previously are maintained for devices with an order of magnitude more p-type doping. There are, however, a number of small differences. For example, $\tau_F^{50\%}$ initially decreases with increasing x_B since slightly more efficient electron cooling in the base results in a reduced $\tau_C^{50\%}$.

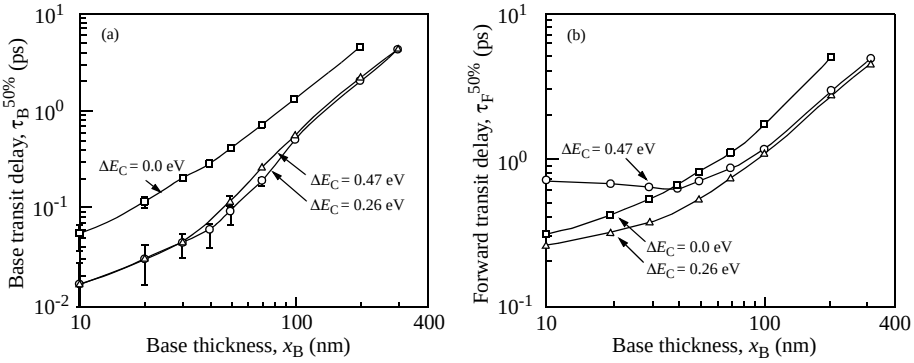


Fig. 6.26. Same as Fig. 6.24(a) and (b) but for $p = 1 \times 10^{20}$ cm^{-3} (from [39]).

Because minority carrier mobility in $In_{0.53}Ga_{0.47}As$ decreases sublinearly with increasing majority carrier concentration it is clear from the above that adopting high p-type doping levels is a good strategy to ensure both small $\tau_F^{50\%}$ and low base resistance in *any* of the transistor designs we have considered. However, the use of an abrupt $Al_{0.48}In_{0.52}As/In_{0.53}Ga_{0.47}As$ emitter-base junction allows for about a factor of two thicker base and lower base resistance compared to a graded device. Never-the-less, if the device is limited by resistive and capacitive parasitics, a graded emitter has important advantages for digital circuit applications. Table 6.5 lists the turn-on voltage, V_{BE}, minimum supply voltage, V_{EE}, and normalized power dissipation for a series gated Current Mode Logic (CML) latch (see Fig. 6.27) which is the standard building block of many bipolar digital circuits. The minimum supply voltage in this architecture is $V_{EE} \sim 3V_{BE} + 2V_S$, where V_S is the voltage swing which is taken to be 0.4 V in this example. Obviously, a significant reduction in power dissipation can be achieved through the use of either an abrupt $InP/In_{0.53}Ga_{0.47}As$ or graded emitter-base junction compared to an abrupt junction $Al_{0.48}In_{0.52}As/In_{0.53}Ga_{0.47}As$ device. It is also worth noting that a graded junction HBT has other advantages such as a large valence band offset and a low emitter dynamic resistance.

TABLE 6.5. *Comparison of supply voltage, V_{EE}, and relative power dissipation of a standard series gated CML latch implemented in different abrupt and graded emitter-base junction $In_{0.53}Ga_{0.47}As$ HBT designs.*

Emitter	$Al_{0.48}In_{0.52}As$	InP	Graded
V_{BE} (V)	1.2	0.9	0.7
V_{EE} (V)	-4.4	-3.5	-2.9
Relative power	1.0	0.79	0.66

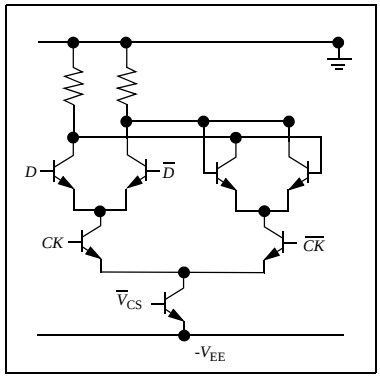


Fig. 6.27. Schematic of a series-gated CML latch.

Clearly there are a number of factors to consider when comparing the relative merits of HBTs with different emitter-base junction designs. Never-the-less, for a given base delay time, implementation of an abrupt junction allows use of a thicker base and thus lower base resistance compared to a graded junction. However, since the abrupt junction increases the device turn-on voltage by approximately ΔE_{CB}, the relatively small conduction band discontinuity of an abrupt junction InP/$In_{0.53}Ga_{0.47}As$ device represents a good compromise between high intrinsic speed, low base resistance and reasonably low power consumption.

6.3 COMMENTS ON SEMICLASSICAL UNDERSTANDING

We have chosen to adopt a semiclassical description of electron transport in various HBT structures and numerical simulations relied heavily on the Monte Carlo method [40 - 43]. It is easy to be convinced of the utility of such an approach simply because we accounted for most of the phenomena presented using semiclassical arguments. However, as discussed in [44], there is still much to learn about nonequilibrium electron transport in scaled semiconductor devices.

Semiclassically, electron motion in band i of a semiconductor is assumed to be described by wave packets formed from the dispersion relation, $E_i(\mathbf{k})$, of the idealized static periodic crystal potential. Typically, one proceeds by appealing to the correspondence principle which in turn leads to concepts of point particle motion, mean free path, and "ballistic" propagation (no additional electron scattering) between instantaneous scattering events. Such an approach is a manifestly inadequate description of nonequilibrium electron dynamics in modern semiconductor structures. In fact, in any practical device, the whole notion of "ballistic" electron transport is totally inappropriate. At first sight, renormalization procedures can be used to recover some of the concepts of "ballistic" motion. For example, calculation of electron self-energy in the presence of polar-optic phonon scattering might be used to estimate a renormalized effective electron mass. However, the electron self-energy introduces a spectral function which, at least in principle, invalidates the use of an exact $E_i(\mathbf{k})$ relationship. There is no longer a definite relationship between energy E and wave vector $\mathbf{k}$ and, in general, no longer an electron state in a definite band of index i. This, of course, is only the first indication that something is wrong with the semiclassical description. However, rather than dwell on generalities, it is convenient to focus on a specific example.

Consider what happens if one places an abrupt potential barrier in the collector arm of a HBT. Obviously, one expects quantum mechanical reflection to play an important role in determining device performance. Fig. 6.28 shows results of measuring common emitter current gain β with bias V_{CE} for the AlAs/GaAs structure sketched in the inset [45]. It is apparent that there are strong resonances in β associated with quantization of energy levels in the 40 nm thick base-collector region. Semiclassical arguments suggest that both elastic (impurity) and inelastic scattering in the collector and delta doped p-type base are responsible for the resonance minima not decreasing to values closer to $\beta = 0$. Earlier, our semiclassical description of nonequilibrium electron transport led us to believe that inelastic scattering rates in the base present a *fundamental* limit to transistor performance. Thus, in an effort to quantify this limit, one might try to calculate scattering rates for perpendicular nonequilibrium electron transport across the two dimensional plane of p-type dopants in the base. In addition to various other complications, these scattering rates cannot be evaluated independently of collec-

tor bias, V_{CB}. In fact, since base and collector may no longer be separated one might, as a first step, assume coupling through the barriers is weak and appeal to the transfer Hamiltonian formalism schematically depicted in Fig. 6.8. We then calculate how the density of states between the two barriers shown in Fig. 6.28 is renormalized by elastic (impurity) and inelastic scattering. Of course, this technique, which has been used to qualitatively describe the measured large renormalization effects in resonant tunnel junctions [46 - 47], presupposes we know how to calculate the appropriate electron self-energy due to elastic (impurity) and inelastic interactions. Needless to say, a further (incorrect) assumption is that the current-voltage characteristic of the device is solely related to the spectral properties of the renormalized density of states.

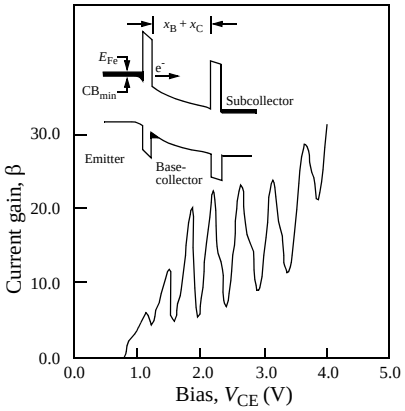


Fig. 6.28. Measured current gain, β with bias, V_{CE} for the AlAs/GaAs HBT sketched in the inset (from [45]). Lattice temperature is $T = 4.2$ K, emitter area is 7.8×10^{-5} cm^2 and base current is $I_B = 0.1$ mA. The AlAs tunnel emitter is 8 nm thick, the *p*-type base is delta doped with a Be sheet concentration of 6×10^{13} cm^{-2}, the AlAs collector barrier is 5 nm thick and emitter-collector barrier separation is $x_B + x_C = 40$ nm.

Conventional Boltzmann and Monte Carlo theories assume that collisions may be treated as isolated local events. The assumption of a local collision rate follows from the quantum kinetic theory of nonequilibrium Green's functions in slowly varying space and time potentials [50 - 52] and is often invalid in devices where scattering rates are large and / or memory effects are important. A further assumption is that scattering events are independent of the driving field / potential profile. This is invalid if before completion of each individual collision other scattering processes occur or an electron gains so much energy from the field during a collision that the collision process itself is altered. The latter is known as the intracollisional field effect [53 - 55] and typically becomes important for fields greater than 2×10^5 V cm^{-1}. As mentioned previously, when the characteristic time scale (the carrier transit time) is comparable to the collision time, as can occur in "ballistic" transport, collisional broadening can no longer be ignored. Electron energy and wave vector become independent variables and the relationship between them is described by a spectral density function. In fact, in most modern devices, the assumptions underlying the Boltzmann and Monte Carlo theories are invalid.

In addition to the above, one expects quantum interference to become important in small scale devices when significant changes in the carrier velocity occur on a length scale comparable to the particle wave length, but short compared to its mean free path.

In low effective electron mass semiconductors this occurs for potential changes of greater than ~ 0.2 eV and length scales less than ~ 10 nm (i.e. electric fields greater than $\sim 2 \times 10^5$ V cm^{-1}). The spatial potential profile, in which the electron moves, must then be treated quantum mechanically, thereby altering the states and scattering rates. Another important requirement becomes apparent by simply recognizing the fact that any model of electron transport in microstructures should not violate Maxwell's equations. Thus, one needs to be able to deal with the possibility of large dynamic space charging effects due to quantum mechanical reflection as, for example, occurs in a resonant tunnel diode. Not only can elastic quantum mechanical reflections from an abrupt change in potential strongly influence inelastic scattering rates, dissipative processes, such as phonon scattering, can in turn modify quantum reflection and transmission rates. This type of feedback is driven by unitarity [59]. Ultimately, what is needed is a new approach to describing quantum transport in modern devices to replace the old methods which are no longer either adequate or valid. Such an approach should naturally describe electron interactions on very short time and length scales where transient polarization effects are expected to play an important role.

One possible approach might be to adopt techniques which have been used to model very fast optical processes such as photon absorption and subsequent generation of transient far infrared radiation (FIR) near the surface of *n*-type semiconductors [65]. The common explanation of such FIR generation is in terms of a two step process in which first electron-hole pairs are created by photon absorption and vertical transitions in real space. Following this, electrons are accelerated by the semiconductor's surface depletion field while the holes remain confined near the surface. The resulting current surge $\partial \mathbf{J}/\partial t$ gives rise to FIR, where total current density $\mathbf{J} = (\mathbf{j}_{el} + \partial/\partial t[\mathbf{D}/(4\pi)])$ is proportional to an average carrier density n and an average mobility μ. Of course, on very short time scales, it is the displacement current $\partial/\partial t[\mathbf{D}/(4\pi)]$ which makes the most important contribution to $\mathbf{J}$. A considerably simpler and natural approach to the problem is to introduce the concept of non-vertical transitions in real space [65]. Now, the spatial separation of final electron and hole states leads to an instantaneous FIR polarization $\mathbf{P}_0$. The second time derivative of this polarization $\partial^2 \mathbf{P}_0/\partial t^2$ determines the radiated signal. In this case the behavior of $\mathbf{P}_0$ is intimately related to the nonlinear optical susceptibility of the semiconductor. Thus, a very convenient stationary-state picture can be used to describe otherwise complex transient electron dynamics in semiconductors. It may be possible to use similar ideas to describe very fast transient phenomena in devices such as HBTs and then make use of appropriate dynamical equations to describe the time evolution of the electron states.

Part of the appeal of adopting the above mentioned approach is that it avoids the difficulties associated with separating optical and electronic processes on short time scales. The significance of this becomes apparent when one recognizes that electron scattering is mediated via emission and absorption of a photon so that on the very shortest time scales one ultimately cannot distinguish between optical and electronic phenomena.

6.4 CONCLUDING REMARKS

The majority of this chapter has discussed nonequilibrium electron transport in vertically scaled heterostructure bipolar transistors. The existence of extreme nonequilib-

rium or "ballistic" electron transport has been demonstrated experimentally and a description of electron scattering developed to describe the observations. The Landau quasi-particle concept and the associated phase-space constraits enabled development of a theory of electron scattering in which the underlying physics could easily be identified and understood. However, despite these important advances, some basic inadequacies of our understanding of electron transport in modern devices have become apparent, illustrating the need for new theoretical and experimental studies. As high-speed devices such as HBTs decrease in size and increase in speed of operation, extreme nonequilibrium electron transport will play an increasing role in determining device performance. Knowledge of what is and is not understood concerning nonequilibrium electron transport in HBTs is only one aspect of HBT *technology.* Usually devices and circuits are manufactured and sold based on actual system performance and price advantages relative to alternatives.

Of course, to demonstrate and establish the usefulness and manufacturability of a technology based on HBTs is a substantial undertaking. One key to future applications will involve reduction in lateral dimensions to reduce HBT power consumption while maintaining the high-speed performance of appropriately vertically scaled structures. As long ago as 1993, work in Japan demonstrated f_T = 100 GHz operation at I_C = 600 μA drive current in a device with emitter area $A_E = 2 \times 10^{-8}$ cm^2 (1×2 μm^2) [66]. Further lateral scaling should result in practical HBTs with similar high-speed performance at only a few μ*A* drive current. While impressive from an individual device point of view, such progress is only one of many technical advances which establish an advanced technology. In this Chapter we have chosen to ignor many detailed technical issues in favor of describing the basic electron transport physics determining device behavior.

REFERENCES

1. W. Shockley, U. S. Patent 2,569,347 June, 1948.
2. H. Kroemer, *Proc. IRE* **45**, 1535 (1951).
3. D. Yevick and W. Bardyszewski, Phys. Rev. **B39**, 8605 (1989).
4. T. Kaneto, K. W. Kim, and M. A. Littlejohn, J. Appl. Phys. **72**, 4139 (1992).
5. A. F. J. Levi and T. H. Chiu, *Appl. Phys. Lett*. **51**, 984 (1987), T. H. Chiu and A. F. J. Levi, *Appl. Phys. Lett*. **55**, 1891 (1989).
6. M. D. Stiles and D. R. Hamann, *Phys. Rev. B* **38**, 2021 (1988).
7. J. Bardeen, *Phys. Rev. Lett*. **6**, 57 (1961).
8. J. R. Hayes and A. F. J. Levi, *IEEE J. Quantum Electron*. **QE-22**, 1744 (1986).
9. K. Hess and G. J. Iafrate, *Proc. IEEE* **76**, 519 (1988).
10. K. Berthold, A. F. J. Levi, J. Walker, and R. J. Malik, *Appl. Phys. Lett*. **52**, 2247 (1988).
11. P. H. Beton and A. F. J. Levi, *Appl. Phys. Lett*. **55**, 250 (1989).
12. A. F. J. Levi, R. N. Nottenburg, Y. K. Chen, and J. E. Cunningham, *Appl. Phys. Lett*. **54**, 2250 (1989).
13. D. Pines and P. Nozieres, "The Theory of Quantum Liquids." Benjamin, New York (1966).
14. A. F. J. Levi and Y. Yafet, *Appl. Phys. Lett*. **51**, 42 (1987).
15. A. F. J. Levi, *Electron. Lett*. **24**, 1273 (1988).
16. W. Bardyszewski and D. Yevick, *Appl. Phys. Lett*. **54**, 837 (1989).
17. A. F. J. Levi, S. L. McCall, and P. M. Platzman, *Appl. Phys. Lett*. **54**, 940 (1989).
18. R. L. Headrick, B. E. Weir, A. F. J. Levi, D. J. Eaglesham, and L. C. Feldman, *Appl. Phys. Let*t. **57**, 2779 (1990).
19. R. L. Headrick, A. F. J. Levi, H. S. Luftman, J. Kovalchick, and L. C. Feldman, *Phys. Rev. B* **43**, 14711 (1991).
20. B. Jalali, R. N. Nottenburg, A. F. J. Levi, R. A. Hamm, M. B. Panish, D. Sivco, and A. Y. Cho, *Appl. Phys. Lett*. **56**, 1460 (1990).
21. M. S. Lundstrom, *Solid-State Electronics* **29**, 1173 (1986).
22. A. F. J. Levi, B. Jalali, R. N. Nottenburg, and A. Y. Cho, *Appl. Phys. Lett*. **60**, 460 (1992).
23. J. A. Baquedano, A. F. J. Levi, B. Jalali, and A. Y. Cho, *Appl. Phys. Lett*. **63**, 2231 (1993).
24. See e.g. R. G. Meyer and R. S. Muller, *IEEE Trans. Electron Devices* **ED-34**, 450 (1987).
25. K. Berthold, A. F. J. Levi, J. Walker, and R. J. Malik, *Appl. Phys. Lett*. **54**, 813 (1989).
26. P. I. Rockett, *IEEE Trans. Electron Devices* **ED-35**, 1573 (1988).

27. L. W. Massengill, T. H. Glisson, J. R. Hauser, and M. A. Littlejohn, *Solid State Electron*. 29, 725 (1986).
28. J. Laskar, R. N. Nottenburg, J. A. Baquedano, A. F. J. Levi, and J. Kolodzey, *IEEE Trans. Electron Devices* **41**, 1942 (1993).
29. R. N. Nottenburg, A. F. J. Levi, B. Jalali, D. Sivco, D. A. Humphrey, and A. Y. Cho, *Appl. Phys. Lett*. **56**, 2660 (1990).
30. V. Gruzinskis, R. Mickevicius, J. Pozhela, and A. Reklaitis, *Europhys. Lett*. **5**, 339 (1988).
31. V. Bareikis, R. Katilius, J. Pozhela, S. V. Gantsevich, and V. L. Gurevich, "Spectroscopy of Nonequilibruim Electrons and Phonons." Chapt. 8, Vol. 35 of Modern Problems in Condensed Matter Sciences, general editors, V. M. Agranovich and A. A. Maradudin, volume editors C. V. Shank and B. P. Zakharchenya, North Holland, Amsterdam (1992).
32. J. R. Pierce, *J. Appl. Phys*. **15**, 721 (1944).
33. V. A. Posse, B. Jalali, and A. F. J. Levi, Appl. Phys. Lett. **66**, 3319 (1995).
36. G. D. Mahn, "Many Particle Physics." Plenum Press, New York, (1986).
37. For example, J. Jensen, M. Hafizi, W. Stanchina, R. Metzeger, D. Rensch, Technical Digest of 1992 GaAs IC Symposium, Miami Beach, Florida, October 4-7, 1992, p. 101 (IEEE cat. #92CH3197-1).
38. Y. Matsuoka, S. Yamahata, S. Yamaguchi, K. Murata, E. Sano, and T. Ishibashi, *IEICE Trans. Electron*. **E76-E**, 1392 (1993).
39. J. A. Baquedano, A. F. J. Levi, and B. Jalali, *Appl. Phys. Lett*. **64**, 67 (1994).
40. W. Faucett, A. D. Boardman, and A. D. Swain, *J. Phys. Chem. Solids* 31, 1963 (1970).
41. C. Jacoboni and L. Reggiani, *Rev. Mod. Phys*. 55, 645 (1983)
42. L.W. Massengill, T.H. Glisson, J.R. Hauser and M.A. Littlejohn, *Solid-State Electronics* **29**, 725 (1986).
43. M. V. Fischetti and S. E. Laux, *Phys. Rev. B* **38**, 9721 (1988)
44. A. F. J. Levi and S. Schmitt-Rink, "Spectroscopy of Nonequilibruim Electrons and Phonons." Chapt. 6, Vol. 35 of Modern Problems in Condensed Matter Sciences, general editors, V. M. Agranovich and A. A. Maradudin, volume editors C. V. Shank and B. P. Zakharchenya, North Holland, Amsterdam (1992).
45. A. F. J. Levi, unpublished (1990).
46. G. S. Boebinger, A. F. J. Levi, S. Schmitt-Rink, A. Passner, L. N. Pfeiffer, and K. W. West, *Phys. Rev. Lett*. **65**, 235 (1990).
47. G. S. Boebinger, A. F. J. Levi, A. Passner, L. N. Pfeiffer, and K. W. West, *Phys. Rev. B* **47**, 16608-16611 (1993).
48. J. F. Müller, A. F. J. Levi, and S. Schmitt-Rink, *Phys. Rev. B* **38**, 9843 (1988).
49. L. Liang and C. S. Lent, *J. Appl. Phys.* **68**, 1741 (1990).

50. L. P. Kadanoff and G. Baym, “Quantum Statistical Mechanics.” Benjamin, New York (1962).
51. L. V. Keldysh *Zh. Eksp. Teor. Fiz*. **47**, 1515 (1964).
52. J. R. Barker and D. K. Ferry, *Solid State Electron*. **23**, 519 (1980).
53. J. R. Barker and D. K. Ferry, *Phys. Rev. Lett*. **42**, 1779 (1979).
54. F. S. Khan, J. H. Davies, and J. W. Wilkins, *Phys. Rev. B* **36**, 2578 (1987).
55. R. Bertoncini, A. M. Kriman, and D. K. Ferry, *Phys. Rev. B* **41**,1390 (1990).
56. W. R. Frensley, *Phys. Rev. B* **36**, 1570 (1987).
57. W. R. Frensley, *Phys. Rev. B* **37**, 10379 (1988).
58. N. C. Kluksdahl, A. M. Kriman, D. K. Ferry, and C. Ringhofer, *Phys. Rev. B* **39**, 7720 (1989).
59. L. D. Landau, and E. M. Lifshitz, “Quantum Mechanics.” Pergamon, Oxford (1977).
60. A. J. Leggett, S. Chakravarty, A. T. Dorsey, M. P. A. Fisher, A. Garg, and W. Zwerger, *Rev. Mod. Phys*. **59**, 1 (1987).
61. R. Bruinsma and P. M. Platzman, *Phys. Rev. B* **35**, 4221 (1987).
62. M. D. Stiles, J. W. Wilkins, and M. Persson, *Phys. Rev. B* **34**, 4490 (1986).
63. B. Y. Gelfand, S. Schmitt-Rink, and A. F. J. Levi, *Phys. Rev. Lett*. **62**, 1683 (1989).
64. W. Cai, T. F. Zheng, P. Hu, B. Yudamin, and M. Lax, *Phys. Rev. Lett*. **63**, 418 (1989).
65. S. L. Chuang, S. Schmitt-Rink, B. J. Green, P. N. Saeta, and A. F. J. Levi, *Phys. Rev. Lett*. **68**, 102, (1992).
66. H. Nakajima, K. Kurishima, S. Yamahata, T. Kobayashi, and Y. Matsuoka, *Electon. Lett*. **29**, 1887 (1993).