MONTE CARLO STUDIES OF NONLINEAR ELECTRON TRANSPORT IN III-V SEMICONDUCTORS

Ki Wook Kim

UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

Approved for Public Release. Distribution Unlimited.
Electron transport in III-V semiconductors, especially the GaAs/AlGaAs material system, is studied in various nonequilibrium situations. Throughout the study, a Monte Carlo simulation method is used for the analysis of transport properties in the semiclassical Boltzmann transport picture. The present work essentially consists of two aspects. The first topic is hot electron transport in GaAs, focusing on the electron impact ionization effects. The dependence of impact ionization rates on the details of the band structure is investigated by using two (local and nonlocal) pseudopotential methods. The spatial evolution of the ionization rate and the average electron energy are studied in nonuniform fields characteristic of p+-n junctions. The effects of field fluctuations due to the random distribution of dopants are studied as well. The possibility of new GaAs electron-emitting diodes is explored numerically and compared with the corresponding Si devices.

The second aspect deals with the effects of conduction band discontinuities on the electron transport. In particular, one-dimensional heterostructures are modeled to study the nonlinear transport across heterointerfaces. First, two heterostructure avalanche (Continued on back)
photodiodes are studied. It is found that overheating, enhanced energy relaxation, and carrier confinement as a consequence of the structure in real space have a pronounced influence on the energy and momentum distribution. As a result, the energy distribution can directly reveal the band structure of the material. The dependence of the impact ionization rate on the band structures of the neighboring layers is also addressed.

The effects of a nonequilibrium phonon distribution on the electron transport are studied as well. The phonon distribution can be considerably perturbed when a large number of carriers propagate across an interface, experience an abrupt energy gain, and subsequently relax through strong phonon emission. An algorithm is developed for the microscopic analysis of phonon dynamics. It is observed that the hot phonons change the scattering rate significantly, and heat the electron energy distribution.

To investigate quasiballistic electron motion, tunneling hot electron transfer amplifier structures are studied at 4.2 K. The numerical results demonstrate the existence of nearly ballistic transport in the base and in the collector barrier, and confirm that the experiments can indeed measure the energy distribution of injected ballistic electrons. The device characteristics, such as transfer ratio and transit time, are also investigated in detail.
MONTE CARLO STUDIES OF NONLINEAR ELECTRON TRANSPORT IN III-V SEMICONDUCTORS

BY

KI WOOK KIM

B.S., Seoul National University, 1983
M.S., University of Illinois, 1985

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Electrical Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 1988

Urbana, Illinois
Electron transport in III-V semiconductors, especially the GaAs/AlGaAs material system, is studied in various nonequilibrium situations. Throughout the study, a Monte Carlo simulation method is used for the analysis of transport properties in the semiclassical Boltzmann transport picture. The present work essentially consists of two aspects. The first topic is hot electron transport in GaAs, focusing on the electron impact ionization effects. The dependence of impact ionization rates on the details of the band structure is investigated by using two (local and non-local) pseudopotential methods. The spatial evolution of the ionization rate and the average electron energy are studied in nonuniform fields characteristic of p^+-n junctions. The effects of field fluctuations due to the random distribution of dopants are studied as well. The possibility of new GaAs electron-emitting diodes is explored numerically and compared with the corresponding Si devices.

The second aspect deals with the effects of conduction band discontinuities on the electron transport. In particular, one-dimensional heterostructures are modeled to study the nonlinear transport across heterointerfaces. First, two heterostructure avalanche photodiodes are studied. It is found that overheating, enhanced energy relaxation, and carrier confinement as a consequence of the structure in real space have a pronounced influence on the energy and momentum distribution. As a result, the energy distribution can directly reveal the band structure of the material. The dependence of the impact ionization rate on the band structures of the neighboring layers is also addressed.
The effects of a nonequilibrium phonon distribution on the electron transport are studied as well. The phonon distribution can be considerably perturbed when a large number of carriers propagate across an interface, experience an abrupt energy gain, and subsequently relax through strong phonon emission. An algorithm is developed for the microscopic analysis of phonon dynamics. It is observed that the hot phonons change the scattering rate significantly, and heat the electron energy distribution.

To investigate quasiballistic electron motion, tunneling hot electron transfer amplifier structures are studied at 4.2 K. The numerical results demonstrate the existence of nearly ballistic transport in the base and in the collector barrier, and confirm that the experiments can indeed measure the energy distribution of injected ballistic electrons. The device characteristics, such as transfer ratio and transit time, are also investigated in detail.
ACKNOWLEDGEMENTS

The author would like to express his sincere gratitude to his thesis advisor, Professor Karl Hess. His continual guidance, insight, and encouragement have been invaluable throughout the study.

The author would like to thank Professors J. P. Leburton, N. Holonyak, Jr., G. Stillman, and J. Eden for the helpful discussions and interest. He is also grateful to Professor Kyekyoon Kim for his early guidance.

Special thanks go to J. Higman, D. Arnold, and Dr. B. Mason. Some of the work presented in this thesis have been done jointly with them.

For their cooperation, valuable discussions, and friendship, the author is indebted to I. Kizilyalli, D. Bailey, Drs. C. Stanton, M. Artaki, T. Wang, Professor U. Ravaioli, and the rest of the group.

The author is especially thankful to his wife, Seehyang, for her love and continuous support.

Finally, the author would like to dedicate this work to his parents. Without their love, patience, and encouragement, this work would never have been possible.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>2. MONTE CARLO SIMULATION METHOD</td>
<td>5</td>
</tr>
<tr>
<td>2.1 Introduction</td>
<td>5</td>
</tr>
<tr>
<td>2.2 Collective Effects</td>
<td>6</td>
</tr>
<tr>
<td>2.2.1 Coupled plasmon/phonon scattering</td>
<td>6</td>
</tr>
<tr>
<td>2.2.2 Pauli exclusion principle</td>
<td>14</td>
</tr>
<tr>
<td>2.3 Tunneling</td>
<td>15</td>
</tr>
<tr>
<td>2.4 Band Bending</td>
<td>18</td>
</tr>
<tr>
<td>2.5 Collision Broadening</td>
<td>20</td>
</tr>
<tr>
<td>3. STUDY OF ELECTRON IMPACT IONIZATION RATE IN GaAs</td>
<td>31</td>
</tr>
<tr>
<td>3.1 Introduction</td>
<td>31</td>
</tr>
<tr>
<td>3.2 Band-Structure Dependence</td>
<td>32</td>
</tr>
<tr>
<td>3.3 Ionization Rates in Nonuniform Fields</td>
<td>35</td>
</tr>
<tr>
<td>3.4 Effects of Random Distribution of Dopants</td>
<td>45</td>
</tr>
<tr>
<td>3.5 Avalanche Electron Emitting Diodes</td>
<td>53</td>
</tr>
<tr>
<td>4. ELECTRON TRANSPORT ACROSS HETEROJUNCTION BAND DISCONTINUITIES IN STRONG ELECTRIC FIELDS</td>
<td>59</td>
</tr>
<tr>
<td>4.1 Introduction</td>
<td>59</td>
</tr>
<tr>
<td>4.2 Description of Model</td>
<td>59</td>
</tr>
<tr>
<td>4.3 Results and Discussion</td>
<td>64</td>
</tr>
<tr>
<td>4.4 Conclusion</td>
<td>85</td>
</tr>
<tr>
<td>5. ELECTRON HEATING BY HOT PHONON EFFECTS IN HETEROSTRUCTURES</td>
<td>86</td>
</tr>
<tr>
<td>5.1 Introduction</td>
<td>86</td>
</tr>
<tr>
<td>5.2 Method of Analysis</td>
<td>87</td>
</tr>
<tr>
<td>5.3 Results and Discussion</td>
<td>89</td>
</tr>
<tr>
<td>5.4 Conclusion</td>
<td>97</td>
</tr>
</tbody>
</table>
CHAPTER 1

INTRODUCTION

Recent advances in semiconductor crystal growth techniques such as molecular beam epitaxy\(^1,2\) (MBE) and metalorganic chemical vapor deposition\(^3\) (MOCVD) have made possible the fabrication of ultrasmall devices, and have opened new possibilities for "band gap engineering" of semiconductor heterostructures with spatially modulated gaps.\(^4\) By using the variability of the boundary conditions which can be imposed on the electronic wavefunction, electrical and optical responses in these structures can be tailored virtually at will. Most of the effort in this area has been devoted to III-V compound semiconductor materials, especially the GaAs/Al\(_x\)Ga\(_{1-x}\)As material system. Along with the possibility of forming latticed-matched artificial structures with Al\(_x\)Ga\(_{1-x}\)As, GaAs itself is known to have electronic properties which are in many ways superior to Si.\(^5-8\) Compared to Si, GaAs shows a significantly higher electron mobility in the low field regime due to the smaller electron mass and a considerably enhanced drift velocity at high fields in short channel devices resulting from the overshoot phenomenon. Because of the relatively weak momentum relaxation processes at low electron energies through polar optical phonon scattering (in contrast to deformation potential scattering with large momentum change in Si), electrons in GaAs can initially gain considerable energy from the field without significant relaxation in momentum distribution and have a velocity overshoot persisting over a considerable time and distance. These properties and the direct band gap make GaAs a suitable material for high speed electronic and optoelectronic devices.

Some of the devices which rely on these properties include superlattice structures,\(^9\) modulation-doped structures,\(^10\) quantum-well heterostructure lasers,\(^11\) heterostructure avalanche photodiodes,\(^12,13\) real-space transfer switching,\(^14\) and ballistic transfer structures.\(^15,16\) In many of
these devices, electrons often experience a reduced dimensionality in the motion (size quantiza-
tion effects) and the displacement in the momentum distribution due to the differences in materi-
als between adjacent layers. A wide range of novel effects such as impact ionization enhance-
ment, real-space transfer, and hot phonon effects has been observed in addition to stan-
dard textbook phenomena such as the particle in a box problem (e.g., red light out of GaAs
quantum wells), the Aharonov-Bohm effect, and various tunneling effects. These effects are
based on electron transport far from equilibrium, and are essential in the operation of the dev-
ices mentioned above.

Since experimental studies of the nonlinear transport are often very difficult due to its
microscopic nature, numerical studies with a microscopic point of view are necessary for better
understanding. Conventional device modeling techniques based on the drift-diffusion equations
are inadequate to correctly treat the nonlinear transport problems since these macroscopic
models cannot accurately account for the electron motion (or electron distribution) in momen-
tum space which is essential in understanding the nonlinear effects. Instead, a technique which
has shown great success for the study of nonlinear transport is the Monte Carlo simulation
method. By stochastically tracing electron movements in both real space and momentum
space, the Monte Carlo method can generate the exact numerical solution of the Boltzmann
transport equation without solving the complicated integro-differential equation directly. The
fact that the simulations are based upon the simple first principles gives the Monte Carlo tech-
nique great flexibility in the inclusion of complicated details such as a realistic band structure,
many scattering mechanisms, complicated boundary conditions, and time- and space-dependent
parameters. The many-particle scheme proposed by Lebwohl and Price is the most suitable
when the collective effects are appreciable and the transient information is of prime concern.
The self-consistent electric field can naturally be included in the simulation by solving the Pois-
son equation simultaneously.
It is the purpose of this thesis to study the electron transport in various nonequilibrium situations in III-V semiconductors, especially the GaAs/AlGaAs material system. Throughout this thesis, a Monte Carlo simulation method has been used for the first-principle analysis of transport properties under the semiclassical Boltzmann transport picture. In Chapter 2, various models to implement complicated boundary conditions into a Monte Carlo method will be described. The treatments of coupled plasmon/phonon scattering, the Pauli exclusion principle, tunneling, and band bending will be discussed in detail. A quantum mechanical correction (collision broadening) to the Boltzmann transport equation and its influence on the electron transport will also be analyzed.

Chapter 3 includes the study of hot electron transport in GaAs. The dependence of impact ionization rate on the details of the band structure will be addressed. The effects of nonuniform electric field in p*-n junctions will be investigated along with the fluctuation in the dopant distribution. The results will be compared with the hot electron characteristics of Si to explore the possibility of new GaAs electron-emitting devices.

Chapters 4 through 6 are devoted to the modeling of GaAs/AlGaAs one-dimensional heterostructures. The effects of the conduction band discontinuity on the electron transport will be discussed along with the models used in the analysis. In Chapter 4, the nonlinear transport across heterointerfaces will be studied in two heterostructure avalanche photodiodes. The evolution of the electron energy and momentum distribution will be investigated both in time and space. The enhancement of impact ionization in these structures will also be studied in conjunction with the dependence on the band structures. Chapter 5 will address hot phonon effects which can occur when a large current flows into a well layer. An algorithm to incorporate phonon dynamics will be described.
To investigate the quasiballistic nature of electron motion and to examine the validity of hot electron spectroscopy, the tunneling hot electron transfer amplifier structures proposed by Heiblum et al.\textsuperscript{16} will be studied in Chapter 6. Special attention will be paid to the transport in the collector barrier, a region where the Heiblum's device differs from the similar planar doped barrier transistor.\textsuperscript{15} The switching characteristics will also be discussed.
CHAPTER 2

MONTE CARLO SIMULATION METHOD

2.1 Introduction

The Monte Carlo method consists of simulating the motion of one or more electrons inside of a crystal, subject to the action of the forces (e.g., an electric field) and of given scattering mechanisms. The duration of carrier free flight time and the scattering events involved in the simulation are selected stochastically, based on the given probabilities describing the microscopic processes. In real devices, especially heterostructure devices, electrons also experience boundary conditions such as reflections at the interface. The accuracy of the calculation is strongly dependent on the physical model of these microscopic processes.

This chapter is devoted to a description of models used in the simulation. The basic Monte Carlo scheme and the treatment of conventional scattering processes (such as electron-phonon interactions and ionized impurity scatterings) have been well documented in the literature,\textsuperscript{25–27} and will not be repeated here. Particular attention will be paid to more complex situations which occur in heavily-doped or novel structures. In Section 2.2, the treatment of collective effects, which are important for a degenerate electron gas, will be discussed. Emphasis will be on two topics, coupled plasmon/phonon scattering and the Pauli exclusion principle. A widely used treatment to implement tunneling, the transfer matrix method, will be discussed in Section 2.3 along with techniques to obtain the effective mass at energies within the band gap. In Section 2.4, the band bending due to the formation of accumulation and depletion regions at the interfaces will be discussed. The last section of this chapter will deal with a quantum mechanical correction to the semiclassical Boltzmann transport equation. The spectral density technique will be described for the inclusion of collision broadening into a Monte Carlo simulation.
2.2 Collective Effects

As the electron concentration increases, interactions between electrons become nonnegligible. Also, the effect of band filling (or the Pauli exclusion principle) which has been neglected in the calculation of scattering rates cannot be disregarded. Normally, Coulombic interactions between electrons are categorized into two scattering mechanisms: the interaction between two independent electrons with short wavelength (electron-electron scattering) and the interaction between an electron and a collective excitation of an electron gas at long wavelength (plasmon scattering). A cutoff wavevector $q_c$ between the individual and the collective excitation of an electron gas is not well defined. Roughly, one half of the Thomas-Fermi wavevector has been used as $q_c$ where the collective excitation starts to damp strongly. As the plasmon frequency approaches to the longitudinal optical phonon energy, the coupling between these modes and the Fröhlich polaron (polar optical phonon scattering) becomes significant. In this section, a theoretical model of these coupled modes will be discussed. Inclusion of the Pauli exclusion principle will also be explored.

2.2.1 Coupled plasmon/phonon scattering

The scattering rate of the coupled plasmon/phonon modes can be calculated by

$$\frac{1}{\tau_{pl-poh}(k)} = \frac{2\pi}{\hbar} \int \frac{d^3q}{(2\pi)^3} \int \frac{d\hbar \omega}{2\pi} \frac{f_k (1 - f_{\bar{k}-q}) (1 + n_{\omega_q})}{\epsilon_0 q^2} \frac{2e^2}{\epsilon_0 q^2} S(q,\omega) \delta(\hbar \omega + E_{\bar{k}-q} - E_k)$$

(2.1)

and

$$S(q,\omega) \equiv -\text{Im}[\frac{1}{\epsilon_T(q,\omega)}]$$

(2.2)

where $\epsilon_o$ represents the dielectric constant in free space. Since $T = 0$ K is assumed in this formalism, only the emission events are considered. Thus, $f_k$ is either 0 or 1 and $n_{\omega_q}$ is 0 in the cal-
The total dielectric function calculated in the random phase approximation can be written

\[ \epsilon_T(q, \omega) = \epsilon_L(q, \omega) + \epsilon_e(q, \omega) - 1, \]  
\[ \text{where the dielectric function of the lattice takes the form} \]

\[ \epsilon_L(q, \omega) = 1 - \epsilon_\infty \frac{\omega_{LO}^2 - \omega_F^2}{\omega^2 - \omega_F^2}, \]  
\[ \text{and the electronic dielectric function is denoted} \]

\[ \epsilon_e(q, \omega) = [\epsilon_\infty + \epsilon_e^L(q, \omega) - 1] + i [\epsilon_e^L(q, \omega)]. \]  

Then, the total dielectric function can be simplified as

\[ \epsilon_T(q, \omega) = [\epsilon_\infty + \epsilon_e^L(q, \omega) - 1] + i [\epsilon_e^L(q, \omega)]. \]  

The real and imaginary parts of the Lindhard function are written

\[ \epsilon_L^{el}(q, \omega) = 1 + D \left( \frac{\epsilon_0^2 m^* k_F^2}{\hbar^2 2\pi^2 q^2} \right) \left\{ 1 + \frac{1}{2q^2} \left[ 1 - \frac{1}{q^2} \left( 1 - \frac{1}{q^2} \right)^2 \ln \left| \frac{1 + \frac{1}{q^2} \frac{1}{n} \frac{1}{n'} \left( \frac{n}{n'} - \frac{n}{n'} \right)} {1 - \frac{1}{q^2} \left( 1 - \frac{1}{q^2} \right)^2} \ln \left| \frac{1 + \frac{1}{q^2} \frac{1}{n} \frac{1}{n'} \left( \frac{n}{n'} - \frac{n}{n'} \right)} {1 - \frac{1}{q^2} \left( 1 - \frac{1}{q^2} \right)^2} \right| \right] \right\}, \]

\[ \epsilon_L^{el}(q, \omega) = \begin{cases} 
D \left( \frac{\epsilon_0^2}{\hbar^2 2\pi^2 q^2} \right) \left[ \epsilon_\infty \left( 1 - \frac{1}{q^2} - \frac{1}{q^2} \right)^2 \right], & 0 < \hbar \omega < \hbar^2 2m \left( q^2 - 2qk_F \right) \\
D \left( \frac{\epsilon_0^2}{\hbar^2 2\pi^2 q^2} \right) \left[ 1 - \left( \frac{1}{q^2} - \frac{1}{q^2} \right)^2 \right], & \hbar \omega > \hbar^2 2m \left( q^2 - 2qk_F \right) \\
0, & \hbar \omega > \hbar^2 2m \left( q^2 + 2qk_F \right) \\
0, & q > 2k_F, \hbar \omega < \hbar^2 2m \left( q^2 - 2qk_F \right) 
\end{cases} \]
where $q' = q/k_F$ and $\nu = m^* \hbar \omega / \hbar^2 k_F^2$. The parameter $D$ is the number of equivalent valleys (i.e., 1 for the $\Gamma$ valley), and $k_F$ is the Fermi wavevector. When $\epsilon^{el}_i=0$, Im$[1/\epsilon_T(q,\omega)]$ has resonance frequencies known as the dispersion relation and is dominated by these resonance frequencies in the integration over the frequency $\omega$. By application of the Dirac formula

$$\frac{1}{x-x'+i\eta} = P\left[\frac{1}{x-x'}\right] - i\pi \delta(x-x')$$

(2.9)

and the use of following relationship

$$\delta(f(x)) = \sum x_0 \frac{\delta(x-x_0)}{\partial f(x)/\partial x}_{x_0}$$

(2.10)

for the zeros $x_o$ of $f(x)$, the scattering rate in the region where $\epsilon^{el}_i=0$ can be obtained by

$$\frac{1}{\tau_{pl-pop}(k)} = \sum \omega_o \int \frac{d^3q}{(2\pi)^3} \frac{2e^2\pi}{\epsilon_0 q^2} \frac{1}{\partial \epsilon_T(q,\omega)/\partial \omega}_{\omega_o}$$

(2.11)

where $\omega_o$'s are the solutions of the equation $\epsilon_T(q,\omega)=0$. For the region where $\epsilon^{el}_i$ is nonzero, the resonance frequencies are broadened out and should be integrated over the possible energy-momentum phase space.

The resulting dispersion relation gives the following four coupled modes as shown in the Fig. 2.1: two phonon-like modes at small and large wavevectors, a plasmon-like mode at small wavevectors, and the Landau damping mode which results in the excitation of electrons from the Fermi sea through the damping of collective motion. Depending on the plasmon frequency $\omega_p^0 = (n e^2 / \epsilon_0 \epsilon_m^* m^*)^{1/2}$, the $\omega_+$ and $\omega_-$ branches in the figure respectively become plasmon-like and phonon-like for $\omega_p^0 > \omega_{LO}$, and vice versa for $\omega_p^0 < \omega_{LO}$. The Landau damping mode corresponds to the region where $\epsilon^{el}_i$ is nonzero. In a different approach, the Landau damping mode can be treated as individual pair electron-electron scatterings as briefly discussed previously. To distinguish the collective plasmon interaction and the individual pair interaction, a cutoff
Figure 2.1  A schematic plot of the dispersion relation of various modes contributing to the coupled plasmon/phonon scattering processes. A degenerate electron density of $1 \times 10^{18}$ cm$^{-3}$ and $E_F \approx 0.13$ eV have been assumed.
wavevector has been introduced based upon the singularities of the dielectric function in such treatments. The approach which uses the total dielectric function is more rigorous.

The calculated scattering rates by these coupled modes at \( n = 1 \times 10^{18} \text{cm}^{-3} \) are shown in Fig. 2.2 as a function of energy. For plasmon- and phonon-like modes, the scattering rates have been obtained in the limit of either short or long wavevectors. By making the short wavevector approximation, the Lindhard function becomes

\[
\epsilon^\prime \omega(q, \omega) = 1 - \epsilon_\infty \frac{\omega^2_\omega(q)}{\omega^2}, \quad \omega^2_\omega(q) = \omega^2_\omega(1 + \frac{3}{5} \frac{\hbar^2 k_F^2}{m^2 \omega^2_\omega q^2}),
\]

(2.12)

which gives the following dispersion relation:

\[
\omega^2_\omega(q) = \frac{[\omega^2_\omega(q) + \omega^2_{LO}] \pm \sqrt{[\omega^2_\omega(q) + \omega^2_{LO}]^2 - 4\omega^2_\omega(q) \omega^2_{LO}}}{2}.
\]

(2.13)

In actual calculations, \( \omega_{\pm}(q=0) \) has been used. The phonon-like mode at long wavevectors has been treated as uncoupled polar optical phonon scattering (\( q \to \infty \) approximation). As can be seen in the figure, the scattering rate is zero below the Fermi energy and sharply increases afterwards until it reaches the maximum. The Landau damping is the most dominant mechanism at low energies, and the plasmon-like mode (\( \omega_+ \)) becomes important as the energy increases. The contribution by the phonon-like modes (\( \omega_{LO} \) and \( \omega_- \)) is negligible. Figure 2.3 shows the coupled plasmon/phonon scattering rate as a function of electron density for an electron of \( E = 0.3 \) eV. It is observed that as the density increases, the scattering rate increases up to a point and then decreases. This decrease is due to the shift in the Fermi energy toward higher energies with the increase in density, thereby shifting the onset of scattering. Although the maximum scattering rate increases with density, the scattering rate with large density can be still in the increasing stage and relatively small at \( E = 0.3 \) eV as shown in Fig 2.4.
Figure 2.2 Scattering rates due to the various coupled plasmon/phonon modes as a function of energy with a degenerate electron density of $1\times10^{18} \text{cm}^{-3}$. 
Figure 2.3  Coupled plasmon/phonon scattering rates as a function of electron density for an electron of $E_F = 0.3$ eV
Figure 2.4 Coupled plasmon/phonon scattering rates as a function of energy for different electron densities.
2.2.2 Pauli exclusion principle

When the electron gas is degenerate, the assumption that the occupation probability \( f(\vec{k}) \) is much less than one is no longer valid. The \([1-f(\vec{k})]\) factor neglected in the standard Monte Carlo simulation is expected to be of considerable importance. Inclusion of the fermionic nature of electrons (the Pauli exclusion principle), however, requires a knowledge of the distribution function. Attempts have been made to incorporate degeneracy into a single particle Monte Carlo method with only limited success.\(^{34}\) The ensemble Monte Carlo method, for which the distribution can be obtained naturally at all times, is the appropriate choice.\(^{35}\)

The procedure generally used to include the Pauli exclusion principle is as follows. The occupation probability \( f(\vec{k}) \) is updated during the simulation by dividing the number of electrons in each mesh cell in momentum space by the number of total states available. Electron drift is treated as in the nondegenerate case. When a scattering occurs, the final state is selected first without considering the exclusion principle. Based upon \( f(\vec{k}) \), the rejection method is then applied to test if the chosen final state is unoccupied. If the transition is rejected, the scattering event is treated as a self-scattering and the state of electron is not changed.

The size of the mesh cells in momentum space should be chosen carefully, depending on the number of electrons in the ensemble. While it is preferable to keep the size of the mesh small for better resolution of the distribution function, it must be large enough so as not to introduce a significant statistical error when the maximum occupancy of a cell by the ensemble electrons is rounded off to an integer. It is important to ensure that calculations are not dependent on the number of electrons in the ensemble and the size of the mesh cells.

The method discussed above, however, requires a significant amount of memory in space dependent simulations. Instead an approximate, but much simpler, method can be used in device simulations since the degeneracy effects are important only in heavily doped regions. In these
regions, large charge density results in low electric fields and low electron temperature. Strong electron-electron interactions tend to restore the Fermi-Dirac distribution. Therefore, the occupation probability function can be validly approximated as

\[
f(\tau, \mathbf{r}, t) \approx \frac{1}{1 + \exp\left(\frac{E_F(t) - E_F(\tau, t)}{k_B T_e(\tau, t)}\right)}
\]

where the Fermi energy \(E_F(t)\) is obtained from the local electron density and the electron temperature \(T_e(\tau, t)\) is from the local average energy.

### 2.3 Tunneling

As the control of epilayer growth becomes more accurate, quantum mechanical effects such as tunneling should be included at the interfaces along with classical transfer. In general, these effects can be incorporated into the Monte Carlo simulation from an independent-particle point of view. Inclusion of many particle effects such as interference is inherently difficult due to the stochastic nature of the Monte Carlo method. Although the effective mass approximation is no longer valid at abrupt interfaces, it is commonly used due to the lack of a simple alternative. More precise results can be obtained by using an empirical Bloch wavefunction approach. In the calculation of tunneling probability two methods, the WKB approximation and the transfer matrix method, are widely used under the effective mass approximation. In this section, the transfer matrix method, which is the more accurate of the two, is described in detail.

Figure 2.5 shows the schematic of an arbitrary one-dimensional potential barrier. Instead of using a continuously varying potential energy, the barrier is approximated as a summation of constant potential energy segments. It is assumed that energy and parallel momentum are conserved throughout the calculation. The wavefunction \(\psi_j\) in the \(j\)th region is given by
Figure 2.5 Energy-band diagram (solid line) and approximated potential function (dashed line) for an arbitrary one-dimensional potential.
\[ \psi_j(x) = A_j \exp(i k_j x) + B_j \exp(-i k_j x) \] (2.15)

where \( k_j \) is obtained from the energy-momentum conservation assumption. From the continuity of \( \psi_j(x) \) and \( (1/m_j^*) \left( \frac{d\psi_j}{dx} \right) \) at each boundary,\(^{44} \) determining \( A_n \) and \( B_n \) can be reduced to the multiplication of the following \( n \) \((2 \times 2)\) matrices:

\[
\begin{pmatrix} A_0 \\ B_0 \end{pmatrix} = M_1 \begin{pmatrix} A_1 \\ B_1 \end{pmatrix} = \prod_{j=1}^{n} M_j \begin{pmatrix} A_n \\ B_n \end{pmatrix} \equiv \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} \begin{pmatrix} A_0 \\ B_0 \end{pmatrix}.
\] (2.16)

By setting \( B_n = 0 \), the tunneling probability \( P(k_0) \) can be written as

\[
P(k_0) = \left( \frac{k_n/m_n^*}{k_0/m_0^*} \right) \frac{|A_n|^2}{|A_0|^2} = \left( \frac{k_n/m_n^*}{k_0/m_0^*} \right) |M_{11}|^{-2}.
\] (2.17)

If nonparabolicity is included, the effective mass is a function of kinetic energy (thus, position in real space) even in the homogeneous material system. The effective mass in the forbidden gap (i.e., in the barrier region) can be defined by using the extension of a \( \vec{k} \vec{p} \) method.\(^{45,46} \)

The model discussed above describes the tunneling process in which electrons stay in the same valley in momentum space, and does not include the elastic intervalley transfer between band minima. Negative differential resistance has been reported in a single barrier structure through the \( X \) minima in the GaAs-AlAs system\(^{47,48} \) and the valence band in the CdTe-HgCdTe system.\(^{49} \) The tight binding method can be used to estimate these effects of the mixing between valleys.\(^{41,50} \)

Although generally used, the assumption of parallel momentum conservation does not have a rigorous theoretical justification other than the hand-waving argument of lattice homogeneity in the direction parallel to the interfaces. Due to the position dependent effective mass, the total energy (kinetic plus potential) in the direction of propagation is not conserved even when scattering is not present. This seems to suggest that tunneling should be treated three-dimensionally in principle. Further study is needed to clarify this point.
2.4 Band Bending

In the tunnel structures, tunneling currents are extremely sensitive to the actual shape of the barrier through which charge carriers tunnel, and to the voltage drop in the electrodes as well as in the barrier. Thus, it is essential to perform self-consistent calculations of the energy-band diagram which includes the effects of band bending and the buildup of space charge in the quantum wells. In some cases, energy-band profiles even indicate that the dominant carrier transport mechanisms are quite different from those expected when these effects are neglected. An algorithm to obtain realistic energy-band profiles is discussed in this section, primarily in single- and double-barrier structures.

Figure 2.6 illustrates a schematic of the potential profile of a single-barrier structure with an undoped AlGaAs layer sandwiched between heavily doped GaAs contact layers. For the calculation of band profile in this structure, the following boundary conditions are generally used: (i) Far away from the barrier layer, the semiconductor properties are bulk-like. (ii) There is no interface charge. (iii) Under no bias, the Fermi level remains constant throughout the entire structure. Under an external bias $V_a$, the quasi-Fermi level in each contact layer is constant up to the interface with the barrier. These quasi-Fermi levels must differ by $eV_a$. With these boundary conditions, the Poisson equation is solved self-consistently. Starting from an initial charge distribution, the potential profile and the charge distribution are iterated self-consistently. Since the quasi-Fermi level is always known (conditions (i) and (iii) above), the charge distribution which corresponds to a given potential profile can be readily obtained. In double-barrier resonant tunneling structures, the Schrödinger equation is solved simultaneously with the Poisson equation to include the charge buildup in the quantum-well layer. The quasi-Fermi level of the contact layer, from which most of the carriers in the well layer come, is generally used in the well layer also.
Figure 2.6  A schematic of the potential profile of a single-barrier structure with an undoped AlGaAs layer sandwiched between heavily doped GaAs layers. The solid line includes the band bending, and the dashed line does not.
2.5 Collision Broadening

The development of methods to calculate the electron transport properties of materials far away from equilibrium conditions is an unsolved problem in physics today. An understanding of the behavior of electrons in high electric fields is becoming increasingly important in semiconductor structures which approach angstrom sizes. Such effects as collision broadening, finite collision lifetimes, the intra-collisional field effect, and extremely high electron scattering rates must be considered in studies of these nonequilibrium problems. For systems which are not highly perturbed from equilibrium and where such effects are of only limited importance, semiclassical calculations based on the Boltzmann transport equation prove to be very useful in predicting physical properties. As mentioned previously, Monte Carlo methods in particular are very flexible and powerful for transport calculations in this semiclassical regime. Unfortunately, these methods fail in the high-field, high-scattering rate environments which are of increasing interest in microstructured semiconductor devices. In situations where the quantum mechanical nature of the system becomes important, a technique which has the power to calculate electron transport has not yet been found. Although a reasonably complete theoretic framework has been established, quantum transport has, to this point, proven to be computationally intractable.

Collision broadening is the spread of the electron energy-momentum relation due to scattering. This quasiparticle nature of the electron is best described by the finite width of the spectral function of an electron interacting with its environment. In a seminal paper, Chang et al. explored the idea of including collision broadening into Monte Carlo calculations of electron transport in semiconductors. Their approach was to include the broadening of the electronic states using the ideas developed by Barker in his work on quantum transport and the intracollisional field effect. This paper by Chang et al. has received much comment and has spurred further research into the problem. Recently, Lugli et al. included the broadening of the
electronic states in a model Monte Carlo calculation using a more precise method, which makes use of the many-body spectral function. Their work used a single parabolic band with no upper band edge and the Born approximation for scattering.

In this section, a method which includes one particular quantum effect, collision broadening, into the semiclassical Monte Carlo method by applying the spectral function technique is described in detail. Comparisons will be made with the treatment of Chang et al.\textsuperscript{59} and Lugli et al.\textsuperscript{60} It should be emphasized that this treatment does not constitute a complete quantum transport method. It incorporates only one of many quantum phenomena into a semiclassical transport calculation. A full quantum treatment is still a problem for the future.

The method discussed here is based on the semiconductor model that was employed by Chang et al.\textsuperscript{59} The full band structure of the bulk GaAs conduction band is calculated using an empirical nonlocal pseudopotential method.\textsuperscript{62} The band structure is needed for a quantitative and even qualitative understanding of high-field transport in semiconductors because of the large effect of the density of states on the scattering rates. Our model still simplifies the electron dynamics by approximating the full electron scattering at high-electron energies from all the scattering mechanisms by an interaction with a single nonpolar optical phonon mode with frequency $\omega_p$. The coupling between the electrons and the phonons is chosen so that the cross section at an energy of 0.6 eV above the conduction band edge is equal to the total cross section of electrons scattering from all the scatterers in GaAs.\textsuperscript{63} Thus, the well-known analytic formulas\textsuperscript{23} are used up to 0.6 eV and the method described below is used above 0.6 eV for the calculation of scattering rates.

The electron scattering rate due to the phonons is calculated in the Fock approximation\textsuperscript{64} by self-consistently solving the equation,
for the electronic self-energy $\Sigma$ which gives the perturbation of electronic states. The scattering rate can be obtained from the relation

$$\frac{1}{\tau} = -\frac{2}{\hbar} \text{Im} \Sigma(E(\vec{k})).$$

In Eq. (2.18), $g(\vec{k})$ is the coupling between the electrons and the phonons, and $\epsilon(\vec{k})$ is the dispersion relation for the electrons as given by the band structure calculation. The use of Eq. (2.18) for the self-energy includes the approximation that phonon absorption processes are negligible.

At the temperature of 77 K which was used in this calculation, the phonon occupation coefficient for absorption is of the order $10^{-2}$ while the emission coefficient is of order 1. In the case of non-polar optical phonons, $g(\vec{k})$ is well approximated by a constant. As pointed out by Chang et al.,\textsuperscript{59} when the electron-phonon coupling is independent of the wavevector, the self-energy is a function of the electron energy alone, $\Sigma(\vec{k},E) = \Sigma(E)$.

The electron scattering rates calculated by the conventional analytic method and the method discussed above are shown in Fig. 2.7 as a function of energy. The phonon energy $\omega_0 = 29 \text{ meV}$ and the electron-phonon coupling $g = 0.2 \text{ eV}$ have been used. Due to the decrease in the density of states, the scattering rate by the self-energy method starts to deviate from the rate by the conventional method at high energies. Figure 2.8 shows the self-energy as calculated self-consistently from Eq. (2.18) and the self-energy calculated by the Born approximation (where the $\Sigma$ on the right-hand side of Eq. (2.18) is replaced by an imaginary infinitesimal $\delta$). In the Born approximation, the imaginary part of the self-energy $\text{Im} \Sigma(E)$ is proportional to $\rho_0(E - \hbar \omega_0)$ where $\rho_0$ is the density of states. Thus one can see from Fig. 2.8 that the conduction band extends to approximately 4 eV. It is also evident that the Fock approximation results in considerably reduced scattering near the center of the band where the density of states is high.
Figure 2.7 Electron scattering rates at 77 K as a function of energy. The solid line is obtained by the conventional method using the analytic formulas, and the dashed line by the self-energy method.
Figure 2.8  Electronic self-energy in the Fock approximation, \(-\text{Im}\Sigma(E)\) and \(\text{Re}\Sigma(E)\) (inset), plotted as a function of the electronic energy. Also shown is \(-\text{Im}\Sigma(E)\) calculated using the Born approximation (dashed line).
approximation is therefore needed for a more accurate calculation of the scattering rates at high energies.

Collision broadening is introduced into Monte Carlo calculations by the inclusion of a distribution for the electron energy after each collision. In the usual semiclassical transport calculation, the rate of transition of the electron state from before to after a phonon collision is given by the golden rule:

$$W(\vec{k}_f, \vec{k}_i) \sim |V|^2 \delta(\epsilon(\vec{k}_f) - \epsilon(\vec{k}_i) - \hbar \omega_0) \quad (2.20)$$

where $V$ is the effective electron-phonon potential. The final state is given exactly by the initial state and the phonon energy. As described in some detail by Reggiani et al.,\(^6\) the quasiparticle nature of the interacting electron can be included by replacing the delta function in the golden rule by a joint spectral density $K$. The joint spectral density $K$ is a convolution over the initial and final state spectral functions,

$$K(\vec{k}_f, \vec{k}_i) = \int \frac{dE}{2\pi} A(\vec{k}_f, E - \hbar \omega_0) A(\vec{k}_i, E). \quad (2.21)$$

The transition rate in this case is given by

$$W(\vec{k}_f, \vec{k}_i) \sim |V|^2 K(\vec{k}_f, \vec{k}_i). \quad (2.22)$$

In Eq. (2.21), the spectral function for the interacting electron is calculated from the self-energy by

$$A(\vec{k}, E) = \frac{-2 \text{Im} \Sigma(\vec{k}, E)}{[E - \epsilon(\vec{k}) - \text{Re} \Sigma(\vec{k}, E)]^2 + [\text{Im} \Sigma(\vec{k}, E)]^2}. \quad (2.23)$$

The final electron state after scattering by a phonon is given by the joint spectral density distribution $K(\vec{k}_f, \vec{k}_i)$ reflecting the spread in the quasiparticle energy. It is obvious that this method of including the broadening of the electronic states reduces to the golden rule when the bare electron spectral function, $A(\vec{k}, E) = 2\pi \delta(E - \epsilon(\vec{k}))$, is used in Eqs. (2.21) and (2.22).
Since the self-energy is a function of the electron energy alone, $K$ is a function of only the initial and final electron energies, $K(\epsilon_f, \epsilon_i) = K(\epsilon(\epsilon_f), \epsilon(\epsilon_i))$. The quasiparticle energies $\epsilon$ are integrated out in Eq. (2.21). In Fig. 2.9, the joint spectral density $K(\epsilon_f, \epsilon_i)$ is plotted as a function of the final energy $\epsilon_f$ for three values of the initial energy $\epsilon_i$. As can be seen in the figure, $K$ is asymmetric with an enhanced tail which extends into the region of the band where the density of states is largest. This tail is due to the peak in the density of states near the band center. For $\epsilon_i$ below the peak in the density of states, this tail in $K$ extends out for $\epsilon_f > \epsilon_i$, while for $\epsilon_i$ above the peak, this tail extends out for $\epsilon_f < \epsilon_i$.

It is also important to note that the integral $\int d\epsilon_f K(\epsilon_f, \epsilon_i) = 1$. This, of course, is needed for conservation of the electrons. This integral over $K$ gives a check on the numerical calculation. Finally it should be pointed out that this model includes the real part of $\Sigma(\epsilon)$, unlike the model of Lugli et al., which gives the renormalization of the electronic-band energies due to the phonons. This makes the model completely consistent with the quasiparticle picture of the electrons.

To study the importance of collision broadening, Monte Carlo calculations of the electron energy distribution in a uniform electric field have been performed in GaAs using the model described above at 77 K. Polar optical phonons and intervalley scattering have been included for energy loss of low-energy electrons and intervalley electronic motion respectively. The possibility of impact ionization has also been considered. Figure 2.10 shows the steady-state electron distribution for an external field of 500 kV/cm. Included in this figure are results of calculations using three different forms for the transition rate. The solid curve in Fig. 2.10 gives the result for a simulation using the golden rule transition rate as given by Eq. (2.20). The dotted curve gives the electron distribution with the full collision broadening as described above included. The dashed curve is a simulation using a Lorentzian approximation to the broadening similar to
Figure 2.9  Joint spectral density $K(\epsilon_f, \epsilon_i)$ plotted as a function of the final energy $\epsilon_f$ for three different initial energies.
Figure 2.10  Steady-state electron energy distribution in GaAs for golden rule (nonbroadened) calculation (solid line), initial state (Lorentzian) broadening (dashed line), and full collision broadening (dotted line) with a constant electric field of 500 kV/cm
that used by Chang et al. In this approximation, the width of the state before the scattering 
\( \text{Im}\Sigma(\varepsilon_i) \) has been used to broaden the transition rate. For this initial state broadening, the 
spectral density distribution is

\[
K(\varepsilon_f, \varepsilon_i) = \frac{1}{\pi} \frac{-2 \text{Im}\Sigma(\varepsilon_i)}{(\varepsilon_f - \varepsilon_i - \text{Re}\omega)^2 + (\text{Im}\Sigma(\varepsilon_i))^2}.
\]  

(2.24)

This approximation simplifies the calculation of the joint spectral density \( K(\varepsilon_f, \varepsilon_i) \) considerably 
while including an approximate broadening of the electronic states. This Lorentzian state 
broadening reproduces many of the features of the full broadening calculation. It should be pos-
sible to improve on this approximation by including the spread in both the initial and final 
states as suggested by the work of Barker.

It is easily seen from Fig. 2.10 that the collision broadening has a profound effect on the 
calculated electronic distribution. Although the average energies of all three models are equal 
within the accuracy of the calculation, the simulations which included broadening have a much 
larger spread in the electronic distribution with a maximum at a somewhat lower energy and a 
particularly pronounced enhancement in the high-energy tail. This is because with the collision 
broadening, the electrons at the energy range where the density of states is high (here, between 1 
eV and 2 eV) tend to be scattered away from the original energies due to the large spread in the 
spectral density. Thus, the broadened distributions have a very nonthermal (nonexponential) 
high-energy tail.

Unlike this result, Lugli et al. found a run-away phenomenon when collision broadening 
was included. The average electron energy of their broadened distribution was significantly 
higher than that calculated using the golden rule. The inclusion of the precise band structure 
inhibits this run-away. The reason for this can be seen by considering the function \( K \) plotted in 
Fig. 2.9. High in the band the scattering probability is skewed towards lower energies due to
the larger density of states there. For the strictly parabolic dispersion relation used by Lugli et al., the density of states always increases. In our results, the electron energy distribution is larger in both the high- and low-energy regions, than the golden rule results with little or no change in the average energy. There is no run-away when the full band structure of the semiconductor is considered in the electronic model.

The high-energy electron population enhancement due to broadening is in the region of the band where processes such as impact ionization begin to occur. This significant change in the electronic distribution will have a large effect on the calculated rates of any high-energy effects. The accurate calculation of these processes would seem to require, therefore, the inclusion of the broadening of the electronic states. However, caution is needed for the numerical application of this spectral density function technique. Since the spectral density is not a classical probability density, one often has to make approximations, such as limiting the final energy within certain energy ranges or making the spectral density always positive definite, in order to implement this quantum mechanical quantity into a semiclassical calculation. It is important to make sure that the results are not significantly affected by these approximations.
CHAPTER 3

STUDY OF ELECTRON IMPACT IONIZATION RATE IN GaAs

3.1 Introduction

Carrier multiplication by impact ionization is an essential mechanism in semiconductor devices such as photodiodes and IMPATT diodes, and has been subject to many theoretical and experimental investigations. Most of the analytical theories contain several adjustable parameters whose physical meaning are not easy to understand. Recently, a Monte Carlo algorithm which does not contain adjustable parameters and includes a realistic band structure was developed and applied by Shichijo and Hess to investigate high field transport in GaAs. More refined Monte Carlo calculations of the impact ionization rates over a wide range of electric fields have been carried out by Brennan and Hess. The data of these calculations are in good agreement with experimental data, for example, given by Bulman et al. In these calculations, a steady-state model for uniform electric fields has been used. The steady-state model is appropriate for analyzing the data measured in the p-i-n diodes or punch-through p-n diodes with long depletion regions. However, most of the rate measurements have been carried out either in abrupt p+-n junctions or Schottky barriers where the electric fields are not constant. Also, the random distribution of dopants results in fluctuations of the electric fields and affects electron transport. It is, therefore, desirable to study the impact ionization in nonuniform electric fields, which can includes nonlocal effects, as has been pointed out by Capasso et al.

In this chapter, the electron impact ionization in GaAs is studied in various situations with an empirical pseudopotential band structure. Included scattering mechanisms are polar optical phonon scattering and equivalent intervalley scattering as well as nonequivalent intervalley
scattering. Section 3.2 deals with the dependence of the impact ionization rate on the details of the band structure. In Section 3.3, the spatial evolution of the ionization rate and the average electron energy will be discussed for field variations typical for p+-n junctions or Schottky barriers. The effects of field fluctuation due to the random distribution of dopants will be studied in Section 3.4. The final section is devoted to the study of a new type of device (avalanche electron emitting diodes\textsuperscript{77,78}) which is heavily dependent on the impact ionization process.

### 3.2 Band-Structure Dependence

In the previous studies,\textsuperscript{63,72} the empirical local pseudopotential method (as developed by Cohen and Bergstresser\textsuperscript{79}) was used to obtain the band structure. This band-structure calculation can be improved by using the empirical nonlocal pseudopotential method of Chelikowsky and Cohen\textsuperscript{62} in which the effect of nonlocality or energy dependence of the pseudopotential is included. This difference in method brings changes in the band structure, and affects the calculation of impact ionization rates. By carefully studying these changes, one can get important information on the basic characteristics of impact ionization.

Figure 3.1. shows the first conduction band of GaAs calculated by the local and nonlocal pseudopotential method. Compared with the local band structure, the nonlocal band structure has slightly lower energies for all values of $k$. This difference in energy between the two bands is more pronounced at high energies. A comparison of the density of states computed with both methods is shown in Fig. 3.2. As can be seen, the profile of the nonlocal method exhibits a higher peak in the density of states. The states are also shifted to lower energies in the nonlocal calculation. Since the scattering rate by phonons is proportional to the final density of states, one can expect that the energy loss of electrons will be higher for the nonlocal case and the average energy will be lower. Accordingly the impact ionization rate should be smaller.
Figure 3.1  $E(k)$ relation of the first conduction band for GaAs. The data of local are calculated based upon the empirical local pseudopotential method of Cohen and Bergstresser,\textsuperscript{79} while the data for nonlocal are calculated based upon the empirical nonlocal pseudopotential method of Chelikowsky and Cohen.\textsuperscript{82}
Figure 3.2  Total density of states of the first conduction band of GaAs as a function of electron energy. The data of local are obtained from the local band structure in Fig. 3.1, while the data of nonlocal are obtained from the nonlocal band structure in Fig. 3.1.
The impact ionization rates calculated by a Monte Carlo method with these two band structures are shown in Fig. 3.3. The data for the local method are directly from Brennan and Hess.\textsuperscript{72} As expected the impact ionization rate is considerably lower for the nonlocal method than for the local method in the entire range of electric fields. This difference is relatively small at high fields and increases as the field decreases. The reason for this is that the high-energy tail of the distribution is sensitively affected by the details of the band structure and especially by the increase in the density of states. At low electric fields, it is this high-energy tail (the lucky and nearly lucky electrons) which contributes most to the ionization rate. At high fields, the distribution function is heated so much that the bulk of the spherically symmetric part of the energy distribution function contributes to the ionization, and the band structure has less influence.

Note also that the increased scattering for the nonlocal band structure leads to the disappearance of the small anisotropy of $\alpha$ which is observed for the local band structure. In other words, increased intervalley scattering shifts the anisotropy to much lower fields for the nonlocal case.

### 3.3 Ionization Rates in Nonuniform Fields

There are two widely used ways to define the impact ionization rate. One is based on the rate of the change of the particle density, and the other on the rate of the change of the current density. Beni and Capasso\textsuperscript{80} have shown that these two definitions are in general not the same and the difference is proportional to the derivative of the drift velocity. Of course, in homogeneous steady-state problems, these two definitions give the same result. However, in the spatially transient case, the difference may be significant. It is generally accepted that the definition based on the current density must be used in the transient cases.\textsuperscript{80} Therefore, the connection between these two definitions and the microscopic model used in an actual Monte Carlo...
Figure 3.3  Calculated electron impact ionization rate as a function of inverse electric field. The data of local are from Ref. 72.
simulation (the inverse mean free path between ionizing collisions) should be investigated for the appropriate evaluation of the impact ionization rate in nonlocal problems.

Thornber has derived the ionization rate from the Boltzmann equation, and has shown its equivalence to the definition of the ionization rate in terms of the current density. In his derivation, the electron impact ionization rate is give by

$$
\alpha_e(\mathbf{x}, t) = \frac{e\hbar e}{\int d^3p_1 \int d^3p_j \int d^3p_k \int d^3p_l R_{ijkl}(\mathbf{p}_i, \mathbf{p}_j, \mathbf{p}_k, \mathbf{p}_l) f_i(\mathbf{p}_i, t) / J_e(\mathbf{x}, t)}
$$

(3.1)

where \( J_e(\mathbf{x}, t) \) is the electron particle flux density and \( R_{ijkl}(\mathbf{p}_i, \mathbf{p}_j, \mathbf{p}_k, \mathbf{p}_l) \) is the rate at which a \((1, \mathbf{p}_i)\) carrier scatters to \((j, \mathbf{p}_j)\), creating an electron-hole pair \((i, \mathbf{p}_i) - (k, \mathbf{p}_k)\). This equation can be written as

$$
\alpha_e(\mathbf{x}, t) = \frac{n(\mathbf{x}, t)}{J_e(\mathbf{x}, t)} \sum_i \int d^3p_1 \bar{f}_i(\mathbf{p}_i, t) \sum_{ijkl} \int d^3p_j \int d^3p_k \int d^3p_l R_{ijkl}(\mathbf{p}_i, \mathbf{p}_j, \mathbf{p}_k, \mathbf{p}_l)
$$

(3.2)

$$
= \frac{1}{v_a(\mathbf{x}, t)} \sum_i \int d^3p_1 R_i(\mathbf{p}_i) \bar{f}(\mathbf{p}_i, t)
$$

where \( v_a(\mathbf{x}, t) \) is the electron drift velocity. The function \( \bar{f}(\mathbf{p}_i, t) \) is the probability distribution of the conduction band and \( R_i(\mathbf{p}_i) \) is the rate at which a \((1, \mathbf{p}_i)\) electron ionizes. The rate per unit time of creating a pair by impact ionization is defined as

$$
\frac{1}{r_a(\mathbf{x}, t)} = \sum_i \int d^3p_1 R_i(\mathbf{p}_i) \bar{f}(\mathbf{p}_i, t).
$$

(3.3)

Then Eq. (3.2) becomes

$$
\alpha_e(\mathbf{x}, t) = \frac{1}{v_a(\mathbf{x}, t)r_a(\mathbf{x}, t)} = \frac{1}{\lambda_a(\mathbf{x}, t)}
$$

(3.4)

which is precisely the inverse mean free path between ionizing collisions. Hence, the microscopic model used in Monte Carlo calculations, Thornber's definition, and therefore the definition of the impact ionization rate in terms of the current density are all equivalent. This means that we
can directly compare the measured data with the calculated results by the Monte Carlo simulation. Thornber's definition (and therefore the microscopic model) is valid as long as the underlying Boltzmann equation picture is correct and applies to nonlocal problems. More detailed discussions on the applicability of this model can be found in Ref. 81.

For the microscopic model, it is essential to know $R_{ijk}(\mathbf{p}_i, \mathbf{p}_j, \mathbf{p}_k; \mathbf{p}_l)$ and $\tilde{f}(\mathbf{p}_i, t)$ precisely. The Keldysh formula has been used to compute $R_i(\mathbf{p}_i)$. However, the use of the Keldysh formula makes it difficult to find the precise final states of electrons and holes. Approximations used to restart these particles at zero energy and neglect created (or secondary) electrons are good for steady-state situations. This is because the actual energies of final states are close to zero in general, and the nonlocal effects of secondary electrons are not in the steady-state calculations by definition (device length is assumed to be infinite). However, for the transient case, the created electrons which may have an energy distribution quite different from the launched electrons have nonnegligible effects. It is then necessary to keep track of the secondary electrons to obtain the exact nonlocal effects. The importance of secondary electrons becomes significant when the impact ionization rate is high.

The electric field profiles considered in this section are shown in the inset of Fig. 3.4. Under the depletion approximation, these profiles describe the electric fields in abrupt $p^+-n$ junctions or Schottky barriers which are commonly used to measure the impact ionization rate. The zero of the coordinate system $x=0$ is located at the junction interface at which the Monte Carlo electrons are launched. Electrons are launched with zero energy. In the calculation, the empirical nonlocal pseudopotential band structure shown in the previous section is used. The material parameters, such as the ionization threshold and coupling constant for phonon scatterings, are the same as in Ref. 72. Due to the very high electric field and low doping density ($\approx 10^{16} \text{cm}^{-3}$) in the $n$ region, ionized impurity scattering is not included.
Figure 3.4 Electron impact ionization rates in \(<100>\) direction with the corresponding field profiles in the inset. The length of the dead space is approximately 700 Å for (a) and (b), and 500 Å for (c) and (d).
Figure 3.4 shows the electron impact ionization rate in GaAs with the corresponding field profiles. As can be seen, the ionization rate is very low for a considerable distance (the dead space) before it reaches steady state. Similar characteristics are observed for the average energy as shown in Fig. 3.5. Since the electrons are launched at zero energy, a certain distance is needed for the electron gas to gain enough energy to impact ionize. The dead space is one of the typical nonlocal effects. Grant expected this distance to be the distance needed for electrons to be ballistically heated to the threshold energy. Since, however, electrons lose energy through inelastic collisions while they gain energy from the electric field, longer distances than those obtained from the simple ballistic transport theory are computed. A precise definition of the dead space is not straightforward in general. Here, we define it as the distance needed to reach the maximum of the impact ionization rate. Since the transient rates in the steady regime after the dead space are in good agreement with the steady-state rates, the distance as defined above is equivalent to the distance needed to reach the steady-state ionization rate. As can be seen from Fig. 3.4, the length of the dead space decreases as the maximum field increases since the electrons are heated up more quickly. This length corresponds to the distance which is needed to ballistically gain about 3.4 eV (which then corresponds to approximately twice the impact ionization threshold energy and two and one half times the band gap energy for electrons in GaAs). The rate measured in p+-n junctions are usually plotted vs. 1/E with a dead space correction estimated by Grant's method.

Figure 3.6 shows the impact ionization rates for electric field profile (a) of Fig. 3.4 pointing in two different crystallographic directions (〈100〉 and 〈111〉). For comparison, the steady-state ionization rates are included. An anisotropy seems to develop at low fields. The agreement of transient and steady-state data after the dead space means that the simple steady-state method with a uniform field can be used to calculate the ionization rate in this range.
Figure 3.5 Average electron energy in $<100>$ direction. The electric field profile of each case is given in Fig. 3.4.
Figure 3.6  Comparison of electron impact ionization rates between the transient calculation and the steady-state calculation. The electric field profile of the transient calculation is profile (a) of Fig. 3.4.
In these calculations, the electrons have been launched at the maximum field. In p+n junctions, the field does not increase infinitely fast to its maximum strength because of the small depletion region on the p+ side. A similar slower increase will occur in a Schottky barrier because of image force effects. If a factor of 100 is assumed for the ratio between doping levels \( N_A/N_D \) (which is typical for p+n diodes), an additional 100 Å is added to the depletion region for profile (a) of Fig. 3.4. Figure 3.7 shows the result including this correction. Electrons are launched at \( x=100 \) Å and the appropriate field profile is included between \( x=-100 \) Å and \( x=0 \) along with profile (a) of Fig. 3.4 for positive x values. As expected, the ionization rate is higher at a given x in the dead space and reaches the peak faster for this case. In the steady regime, the rates of two cases are in good agreement.

The correction by including secondary electrons has also been estimated, and is shown in Fig. 3.7. Since the ionizing and secondary electrons are indistinguishable after the ionization event, one need only follow a single electron after the event, with proper weighting, to model the entire ensemble. In this calculation, the excess energy of an ionizing electron is randomly distributed between possible final states by the energy conservation law, and the two final electrons are assumed to have identical energy and momentum for simplicity. Thus, if a sample electron experiences an ionizing collision, all the subsequent contributions by that electron are weighted by the factor of 2. This correction has a negligible effect on the dead space since only a very small number of secondary electrons are created in this region. However, as the number of secondary electrons grows by reaching the regime with the high ionization rate, the ionization rate becomes smaller than that without the correction because the secondary electron distribution is cooler than the launched electrons with a dead space for each secondary electron. To accurately account for this effect, the Keldysh formula must be replaced by a more precise theory which can give \( R_{ijkl}(\vec{p}_i,\vec{p}_j,\vec{p}_k,\vec{p}_l) \) rather than \( R_i(\vec{p}_i) \), as suggested by Kane.
Figure 3.7 Effects of the corrections. The no correction case corresponds to profile (a) of Fig. 3.4.
3.4 Effects of Random Distribution of Dopants

The effects of field fluctuations due to the random distribution of dopants on impact ionization were first addressed by Shockley in his celebrated paper on p-n junctions. Shockley assumed that dopants are incorporated in the junction independently, and therefore follow the Poisson distribution. This assumption is justified because dopants are usually incorporated during growth or processing under the condition of high temperature. At these temperatures, the thermally generated intrinsic electron and hole concentrations are high enough to screen the electrostatic interactions between dopants. Portraying dopants as a continuum with fluctuating density, Shockley calculated the root-mean-square voltage deviation between two points separated by a distance \( d \),

\[
\delta V = \left[ \frac{q^2}{4\pi \epsilon^2} (N_D + N_A) d \right]^{1/2}.
\]  

(3.5)

Here \( q \) is the magnitude of the electronic charge, \( \epsilon \) is the permittivity, and \( N_D \) and \( N_A \) are average donor and acceptor concentrations. Equation 3.5 is derived by considering an infinite depleted semiconductor with uniform average doping. Since only the high field region of the p-n junction is of interest in connection with impact ionization phenomena, it is valid for Shockley to neglect the effects of the edge of the depletion region. From Eq. (3.5), Shockley proceeds to derive an effective electric field for impact ionization,

\[
F_{\text{eff}} = \frac{F}{1 - \left( \frac{F_f}{F} \right)^{1/2}}.
\]  

(3.6)

Here \( F \) is the electric field and \( F_f \) is a quantity representing the fluctuations in the field which is given by

\[
F_f = \frac{(\delta V)^2/d}{E_{\text{th}}/q}.
\]  

(3.7)
where $E_{th}$ is the threshold energy for impact ionization. In deriving Eqs. (3.6) and (3.7), Shockley assumes that the average electric field is constant, which is a rather weak approximation. Since the effective field is larger than the average field, Shockley predicted strong enhancement in the impact ionization rate due to the field fluctuations.

More accurate results can be obtained with the use of a computer. With the location of each discrete dopant randomly chosen in a p-n junction, the realistic field fluctuations can be calculated by solving the Poisson equation in three dimensions using a method such as the finite difference SOR algorithm. Treating the impurities as infinitesimal point charges would result in singularities in the solution of the Poisson equation. These singularities would contribute to the calculated fluctuations (which are still finite) in an unphysical way if included into the field fluctuations without modification, since the effect of the core electrons is totally neglected. Thus, it is more appropriate to use a pseudopotential and therefore cut off the potential variations within the first Bohr-radius. For simplicity of calculation in this model, it is assumed that the carriers which contribute to the current do not significantly affect the potential energy profile. It is also assumed that the applied bias affects only the average potential energy profile, but not the fluctuations.

The particular case considered is a GaAs junction doped with $3.6 \times 10^{18}$ cm$^{-3}$ acceptors on one side and $3.6 \times 10^{16}$ cm$^{-3}$ donors on the other side. The root-mean-square voltage fluctuations between two points in space is plotted in Fig. 3.8 as a function of the distance between two points. Shockley's result, given by Eq. (3.5), is shown also for comparison. The curves show reasonably close agreement. In the calculation, the spatial coordinate normal to the p-n junction has been denoted by $x$, and the transverse coordinates by $y$ and $z$. The normal component of the electric field $F_x(x,y,z)$ is obtained by numerically differentiating the potential. Throughout this section, the normal component of the electric field will be referred to as the field in the
Figure 3.8 Root-mean-square voltage fluctuations between two points as a function of the distance between the points. The points are in the n side of a p⁺-n junction. Shockley's result is given by Eq. (3.5).
The fluctuations in the field are given by

\[ \delta F_x(x,y,z) = F_x(x,y,z) - \bar{F}_x(x). \]  

(3.8)

Here \( \bar{F}_x(x) \) is the field averaged over \( y \) and \( z \). The fluctuations are, then, superimposed on the field, averaged over \( y \) and \( z \), for a \( p^+\)-\( n \) junction under high reverse bias in order to calculate the electron impact ionization rates. This average field is calculated from the depletion approximation, which gives excellent agreement with more exact calculations of the average electric field except near the edge of the depletion region. Since the latter is a low-field region, it is not important to impact ionization. Statistical analysis is performed on one-dimensional cuts of the field as a function of \( x \), with \( y \) and \( z \) held constant. The probability density of the root-mean-square (rms) field fluctuations for the one-dimensional segments can be computed from

\[ \delta F(y,z) = \left[ \frac{1}{N-1} \sum_{i=1}^{N} \left( F_x(x_i,y,z) - \bar{F}_x(x_i) \right)^2 \right]^{1/2}. \]  

(3.9)

From Eq. (3.9), one can see how much the field fluctuates in the normal direction to the junction.

To save computer time the electron impact ionization rates are not calculated for the entire three-dimensional volume. Two one-dimensional electric field profiles are chosen from the probability density of the rms fluctuations of Eq. (3.9): one for which the standard deviation is greater than 50% of those for one-dimensional cuts in the sample mesh, and one for which the standard deviation is greater than 90%. The former represents a typical example and the latter an extreme case. These two electric field profiles are shown in Fig. 3.9 along with the average electric field given by the depletion approximation for a reverse bias of 24 volts.

The impact ionization rates in GaAs for the electric fields given in Fig. 3.9 are shown in Fig. 3.10. The same transport model and the same empirical band structure are used as given in Section 3.3. The effect of the secondary electrons is not included since it is negligible as
Figure 3.9  Electric field profiles in a $p^*-n$ junction. 90% represents the case for which the standard deviation is greater than 90% of those for one-dimensional cuts in the sample mesh, and 50% represents the case for which the standard deviation is greater than 50%.
Figure 3.10  Electron impact ionization rates in GaAs in \(<100>\) direction. The electric field profile for each case is given in Fig. 3.9.
discussed in the previous section. In the dead space where the ionization rates are very low in
spite of the high field, the ionization rates are almost the same in the three cases even though
the local fields have quite different values. This is because the electron distribution is too cold to
respond to the variation in the local field sensitively and, therefore, the nonlocal effects are dom-
inant. When the electrons gain sufficient energy by reaching the steady region after the dead
space, the ionization rates become affected by the local fluctuations resulting in a spatially
fluctuating ionization rate. When the electric field varies slowly in space with respect to the
mean free path of an electron, the transient ionization rates depend only on the local field and
not explicitly on the position. As discussed in the previous section, such is the case in the
depletion approximation. However, the field fluctuations shown in Fig. 3.9 are not slowly vary-
ing. The response of electrons to these large field gradients is delayed. Therefore, adjacent posi-
tive and negative field fluctuations tend to cancel. Hence, the big fluctuations such as observed
near $x = 1500 \text{ Å}$ in the 90% case in Fig. 3.9 result in a delayed and relatively small change in
the ionization rate near $x = 2000 \text{ Å}$ (Fig. 3.10).

This nonlocal effect is clearly shown in Fig. 3.11 for the average electron energy. To gain
energy an electron must travel a distance through a field while suffering from inelastic collisions.
Since the fluctuations are very rapid when compared to the ionization mean free path, the
energy gain of the electrons from the ups and downs of the field are virtually averaged out and
the average electron energy in Fig. 3.11 shows rather small changes. These small changes in
average energy also result in small fluctuations of the ionization rate as discussed above. There-
fore, a significant increase in the ionization rate due to the field fluctuation is not observed,
unlike in Shockley's steady-state analysis. This is different from the case of impact ionization
in a superlattice where the changes in potential energy are large and vary less rapidly compared
to the mean free path. No averaging-out will occur if electrons gain a considerable amount of
energy within a very short time period ($< \tau_{ph}$ where $\tau_{ph}$ is the phonon scattering time constant).
Figure 3.11 Average electron energy in GaAs in \textless 100 \textgreater  direction. The electric field profile for each case is given in Fig. 3.9.
3.5 Avalanche Electron Emitting Diodes

The recent report of efficient silicon cold cathodes by van Gorkom et al. has triggered renewed interest in such solid-state electron sources. Current densities as high as 1000 A/cm² and reduced brightnesses of 900 A/(cm²·sr·V) for a 1 μm diameter emitting area have been reported, making these devices attractive for use in numerous applications. Although p-n emitters were proposed long ago, only recently have they been proven practical with high-current density and efficiency. The device consists of a shallow p⁺-n++ junction fabricated at the semiconductor surface with the n++ channel at the surface as shown in the inset of Fig. 3.12. Cesium is typically absorbed onto the surface in order to lower the work function to approximately 1.7 eV. Reverse biased in avalanche breakdown, electrons entering the depletion region from the p-type side of the junction gain energy from the field and impact ionize, creating free electron-hole pairs. Due to the very high electric field, the avalanche multiplication (i.e., impact ionization) is an essential part of the device operation in this structure. Electrons which have sufficient energy upon reaching the surface may then overcome the work function and be emitted into the vacuum. The efficiency is defined as \( \eta = j_{\text{vac}}/j_{\text{junc}} \) where \( j_{\text{vac}} \) is the current density in the vacuum and \( j_{\text{junc}} \) is the current density through the junction, normal to the surface. The maximum field in the junction is typically 1.5×10⁶ V/cm.

The study of these devices is interesting from a practical point of view in an effort to design more efficient devices, and also from the point of view of high-field transport physics. In this section, the potential performance of GaAs devices is studied. The results are compared to the calculated and the experimental data for Si. The transport model used in these calculations is essentially the same as described in the previous sections. Unlike GaAs, the first two conduction bands are included in the Si model, since the second conduction band is degenerate with the first at the X point and can therefore have a significant population at high electric
Figure 3.12 Calculated and experimental data for the efficiency $\eta$ vs. the work function $\phi$. The experimental data are represented by the best-fit exponential function given in Ref. 77. The calculated data for Si are from Ref. 87. The inset shows the cross-section of the device as fabricated.
fields. In GaAs, the second conduction band lies at least 1 eV above the first and would have a correspondingly small population. Even though the devices are heavily doped on both the n- and p-type sides of the junction, ionized impurity scattering is not included. Since only the high-energy electrons are of interest, and for these electrons Coulombic scattering prefers small angles, the momentum distribution will not be disturbed. It is assumed that the effect of the holes on the avalanche process is to simply create additional electrons near the p-type side of the junction, which will behave identically to the thermally generated electrons leaving the p-type region. Under these assumptions, the electrons created by hole-initiated ionization are accounted for in the model. Band-to-band tunneling has not been included in the model, though a 10% tunneling contribution has been estimated using an analytic model. In this section, the secondary electrons created by impact ionization events are accounted for, as described in Section 3.3, to include the effect of avalanche multiplication on both the number and distribution of electrons reaching the barrier at the surface. In addition, a numerical enhancement of the high-energy tail of the distribution has been used to speed convergence of the calculation of η.

Finally, the transmission of the electrons over the barrier into the vacuum must be treated. For the classical case (i.e., parabolic energy bands), one can define a directional component of the energy such as $E_x \propto k_x^2$. The quantum treatment of the transmission would require a Bloch function description of the electrons in the semiconductor with the added complications of surface effects. Instead, a simple, analytic model for the transmission coefficient has been used in this treatment. In order to account for the fact that not all electrons with total energy $E_T$ can surmount the barrier, we have used an effective transmission probability

$$T_{\text{eff}}(E_T) = 1 - \left( \frac{\phi}{E_T} \right)^\eta, \quad E_T > \phi, \quad \frac{\partial E_T}{\partial k_\perp} > 0$$

(3.10)

where $\phi$ is the work function and $k_\perp$ is the component of the crystal momentum normal to the
surface. This treatment is exact in the limit of a spherically symmetric distribution function and parabolic energy bands. Qualitative results such as the comparison of different materials and calculated trends should not be seriously affected by assuming Eq. (3.10). Comparison with experiment shows that a constant-factor scaling of the efficiency is sufficient to obtain good agreement with experiment.

In Fig. 3.12, results of the Monte Carlo calculations are compared with published experimental data\textsuperscript{7} for $\eta$ vs. $\phi$. For the experimental data, the best-fit function $\eta = 0.6\exp(-\phi/0.41)$ (where $\phi$ is in eV's) has been used. For GaAs the device parameters used in the calculation are $N_A = 1 \times 10^{18} \text{ cm}^{-3}$ and $N_D = 5 \times 10^{18} \text{ cm}^{-3}$. The Si devices are assumed to have $N_A = 1 \times 10^{18} \text{ cm}^{-3}$ and $N_D = 5 \times 10^{19} \text{ cm}^{-3}$. In each case, the device is assumed to be biased such that the maximum electric field is $1.5 \times 10^6 \text{ V/cm}$. Corresponding depletion widths are $W_p = 1070 \text{ Å}$ and $W_n = 210 \text{ Å}$ for GaAs, and $W_p = 980 \text{ Å}$ and $W_n = 20 \text{ Å}$ for Si. The undepleted $n^{++}$ channel width $d$ is $70 \text{ Å}$ for both the GaAs and Si devices. These values of $d$ are chosen based on the silicon emitters which have been demonstrated. The results show that GaAs has a much smaller calculated efficiency than Si. The calculated data for Si follow a curve which is nearly parallel to the experimental data. The difference in magnitude may be partially attributed to the assumption of Eq. (3.10) for the transmission coefficient. However, the major reason for the difference appears to be of experimental origin: it is known that the measured efficiency is highly sensitive to the surface contamination. In a more recent experiment, emitters with clean surfaces show an efficiency of $4\%$ to $5\%$ and up to $8\%$ at $\phi = 1.7 \text{ eV}$\textsuperscript{90} The efficiency is also strongly dependent on the actual dopant distribution. Diffusion of the dopant distribution results in lower maximum electric field and, thus, smaller experimental efficiencies.\textsuperscript{90} Considering these two points, we conclude that the calculated data represent an upper boundary value. For GaAs, the calculated efficiency is also nearly parallel to the experimental Si data. We
would expect that the difference between the measured efficiencies would be consistent with the
difference in the calculated values, so the Monte Carlo simulation indicates roughly an order of
magnitude greater efficiency for comparable Si sources.

The calculated efficiency is plotted as a function of \( d \) for both GaAs and Si in Fig. 3.13.
Note that for both materials, the efficiency is, to a good approximation, exponential in \( d \). The
steeper decrease of the efficiency for GaAs is indicative of the higher scattering rates (larger den-
sity of states). The control of \( d \) in fabricating the device is therefore even more important for
GaAs devices. Note that the higher n-type doping levels possible in Si allow one to make the
channel narrower than for GaAs while maintaining the same current-carrying capability.

For both Si and GaAs, the work function is larger than the threshold energy for impact
ionization, \( \phi \geq E_{th} \). Thus, it is clear that impact ionization is at once essential for, and detri-
mental to, the operation of the device. The junctions must be biased in avalanche breakdown
for a sufficient number of electrons to reach the surface to obtain practical values of \( j_{\text{vac}} \), and yet
by impact ionizing, many electrons lose nearly all of their energy and are then unable to sur-
mount the barrier. A soft threshold for impact ionization is necessary to enable a significant
number of electrons to remain in the high-energy \( (E_F \geq \phi) \) tail of the distribution. From various
investigations,\(^{63,88,91}\) it can be seen that the threshold in Si is softer than in GaAs. Also, the
higher density of states, and therefore higher scattering rates, at energies near \( \phi \) in GaAs
decrease the fraction of electrons in the high-energy tail. These effects both contribute to the
lower efficiency of GaAs as indicated in Figs. 3.12 and 3.13. The device performance could be
also enhanced by relocating the generation of secondary electrons nearer to the p-type side of
the junction. Present devices have not been optimized in this respect.
Figure 3.13  Calculated data for the efficiency $\eta$ vs. the undepleted $n^{++}$ width $d$. The data for Si are from Ref. 87.
CHAPTER 4

ELECTRON TRANSPORT ACROSS HETEROJUNCTION BAND DISCONTINUITIES IN STRONG ELECTRIC FIELDS

4.1 Introduction

There have been significant theoretical efforts to understand the effects of heterointerfaces on the carrier transport. Most of the efforts have been devoted to the transport properties in a direction parallel to the interfaces such as quasi-two-dimensional transport in field effect devices. However, the study of carrier transport across heterointerfaces is also essential for the understanding of many interesting nonlinear effects such as ionization enhancement in heterostructure avalanche photodiodes (APD's).

In this chapter a detailed study of nonlinear electron transport across heterointerfaces in strong electric fields is presented, based on an ensemble Monte Carlo simulation method. Unlike most of the previous studies on this subject, particular attention has been devoted to microscopic features such as the evolution of the energy and momentum distribution function in space and time, from which one can deduce macroscopic results. It is shown that the combined effects of real-space and momentum-space structure bring pronounced peaks and anisotropy in energy and momentum distribution respectively. The dependence of the ionization rate on the band structure of the neighboring layer is also discussed.

4.2 Description of Model

Monte Carlo simulations of electron transport have been performed for various configurations and material compositions of the AlGaAs/GaAs system. Figure 4.1 shows two single-quantum-well heterostructures in which detailed studies have been carried out. The structures shown in Fig. 4.1(a) and (b) may represent part of the i region of a reverse-biased
Figure 4.1 Energy-band diagram of AlGaAs/GaAs single-well heterostructures for (a) quantum-well APD and (b) doped quantum-well APD.
quantum-well APD\textsuperscript{12} and the p\textsuperscript{+}-i-n\textsuperscript{+}-i region of a reverse-biased doped quantum-well APD\textsuperscript{13} respectively. Generally superlattice structures are studied in connection with novel versions of these devices. However, only one period has been simulated in this chapter in order to clarify the fundamental physics of nonlinear transport.

Each layer in Fig. 4.1 is assumed to be ideal and free of defects and oriented along the \textlangle100\textrangle direction. Electrons are generated in a region prior to layer 1 by an external source such as light and injected into the layers that are depleted by a reverse bias. To rule out the possible nonlocal effects in the energy distribution when electrons enter the well layer from layer 1, a length of 1500 Å has been chosen for the width of layer 1. The energy and momenta of the generated electrons are assumed to obey a Maxwell-Boltzmann distribution before they are accelerated by the high fields. Since the semiclassical transport is intended to be studied, the well layer is rather long (a width of 400 Å is assumed for the GaAs layer unless otherwise specified). Therefore, size quantization and resonance effects are of minor importance. The actual size of layer 3 is unimportant for the results which we report. Unlike in Fig. 4.1(a), the electric field in Fig. 4.1(b) is not uniform. The electric field in layer 1 varies from 500 kV/cm to 100 kV/cm due to the change in the build-in-field, while the field in layers 2 and 3 is fixed to 100 kV/cm. Such field distributions could be obtained in multiquantum-well structures by means of an appropriate periodic doping profile.\textsuperscript{13} The Al mole fraction for layers 1 and 3 is chosen to be 0.36 unless otherwise specified. To get the conduction band offset, the 70/30 rule\textsuperscript{97} has been used. The empirical band structure in GaAs is calculated as in Chapter 3, while for AlGaAs the "analytical" band structure as described in Ref. 98 is used. Since the electrons in the AlGaAs layers are hot due to the high electric field, inclusion of a more realistic (but currently not available) band structure for AlGaAs may be of importance. The material parameters used in this study are summarized in Table 4.1.
Table 4.1 Material parameters for the simulation program

<table>
<thead>
<tr>
<th>Material</th>
<th>Valley</th>
<th>Effective mass ratio ((m/m_0))</th>
<th>Nonparabolicity ((eV^{-1}))</th>
<th>Valley separation relative to the (\Gamma)-valley minimum ((eV))</th>
</tr>
</thead>
<tbody>
<tr>
<td>GaAs</td>
<td>(\Gamma ) valley</td>
<td>0.069</td>
<td>0.67</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(L) valley</td>
<td>1.47 (longitudinal) 0.12 (transverse)</td>
<td>0.44 0.29</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(X) valley</td>
<td>1.58 (longitudinal) 0.24 (transverse)</td>
<td>0.36 0.48</td>
<td></td>
</tr>
<tr>
<td>(\text{Al}<em>x\text{Ga}</em>{1-x}\text{As})</td>
<td>(\Gamma ) valley</td>
<td>0.069+0.103*(x)</td>
<td>0.67-0.94*(x)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(L) valley</td>
<td>1.47 (longitudinal) 0.12+0.036*(x) (transverse)</td>
<td>0.44-0.038*(x) 0.29-0.647*(x)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(X) valley</td>
<td>1.58+1.21*(x) (longitudinal) 0.24 (transverse)</td>
<td>0.36 0.48-1.156*(x)</td>
<td></td>
</tr>
</tbody>
</table>
Included scattering mechanisms are polar optical and deformation potential scattering as well as impact ionization. Electron-electron scattering is also included in the well layer to account for the charge accumulation which can occur. Impact ionization rates are calculated by using the estimator proposed by Thornber as discussed in Chapter 3. The ionization threshold energy and the $p$ factor (of the Keldysh formula) for GaAs are chosen to be 1.7 eV and 0.5 respectively. For AlGaAs, the threshold energy has been chosen proportional to the increase in the energy gap over GaAs. In the calculation, the effect of secondary electrons has also been included for some extent. For simplicity, the ionizing and ionized (secondary) electrons are given zero energy after ionization occurs.

At the interfaces, electron transfer for the L and X valley electrons is treated classically by conserving electron energy and parallel momentum. With classical, it is meant that an electron can cross the interface only when it has high enough energy to overcome the barrier. This simplification is motivated by the fact that the time constants for tunneling are relatively long compared to the short transients due to the high field. For the $\Gamma$ valley electrons, the tunneling process is included along with the classical transfer. By conserving energy and particle flux, the WKB approximation is used to calculate the tunneling probability. The reflections at the interfaces are assumed to be specular. The quantum interference effects by the reflections are unimportant since the inelastic mean free path is much shorter than the well width.

In the calculation, a constant time step ($2 \times 10^{-15}$ sec) discretization scheme is used. This allows the tracking of the time evolution of electron transport parameters. For the electric field profile, the self-consistent correction is not included. At high fields such as used in these calculations, the band distortion by the trapping effects in the well is not important.
4.3 Results and Discussion

While crossing the interface and propagating into the GaAs layer, electrons gain excess energy from the conduction band offset and, therefore, exhibit an overheated energy distribution. Figure 4.2 shows the electron energy distributions with an electric field of 500 kV/cm in structure (a) of Fig. 4.1. The energy distributions in layers 1 and 3 are calculated with 40 Å from the interface while the displayed distribution for layer 2 represents an average over the whole layer. In this figure, the energy distribution in layer 2 is shifted to higher energy (by the amount of the band edge discontinuity) as expected, and exhibits two peaks, one at very low energy near the Γ valley minimum and the other at high energy. The reason for the low-energy peak is two-fold. The shift in energy distribution enhances the impact ionization and creates cooled electrons in the Γ valley. Also, the heterointerface barriers tend to confine (by reflection) cold electrons in the well region. However, most of the cold electrons are reheated quickly due to the high field applied and, thus, the low-energy peak is reduced in spite of the high ionization rate. As the field decreases, the low-energy peak also decreases along with the ionization rate as shown in Fig. 4.3. The confinement effect is negligible in this case.

The overheating causes strongly enhanced electron-phonon interactions. Therefore, electrons in layer 2 relax to the steady-state distribution rather rapidly. As shown in Fig. 4.4, this quick relaxation results in nearly the same distribution as in the situation illustrated in Fig. 4.2, although the initial distribution in layer 1 is different due to the different Al mole fraction. The small peaks for layers 1 and 3 are due to the low energetic electrons which overcome or tunnel through the potential barrier.

For the simulation of structure (b) in Fig. 4.1, however, electrons experience strong relaxation and confinement due to the relatively small electric field in the GaAs layer (100 kV/cm). In this case, a smooth distribution enters the GaAs layer from layer 1 and relaxes to a distribution...
Figure 4.2  Electron energy distributions in structure (a) of Fig. 4.1. The electric field in each layer is 500 kV/cm.
Figure 4.3 Electron energy distributions in layer 2 for structure (a) of Fig. 4.1. The electric field varies from 500 kV/cm to 300 kV/cm.
Figure 4.4  Electron energy distributions in structure (a) of Fig. 4.1. The electric field is 500 kV/cm. The Al mole fraction is 0.12.
with three steep peaks which are clearly related to the $\Gamma$, L, and X valley energies as shown in Fig. 4.5. The process of relaxation can be understood from the study of the time development of the energy distribution which is shown in Fig. 4.6. The zero of the time coordinate is set when electrons start to be injected into layer 1. However, the actual time when electrons start to enter layer 2 is about 2 psec later since they have to travel through layer 1 first. As shown in the figure, the electron distribution is initially very hot ($t=2.5$ psec) while a peak begins to develop near the $\Gamma$ valley minimum due to the strong impact ionization. As the electrons experience increasingly the enhanced electron-phonon interactions, the distribution cools down to lower energy and the peaks near L and X valley minima start to form. By the time $t=6$ psec, electrons have reached the layer 2-layer 3 interface and exhibit a distribution with three pronounced peaks. Finally, the peak near the $\Gamma$ valley minimum grows as the distribution reaches the steady state.

In principle, there may be several possible reasons for these peaks. To clarify this point, a calculation without the conduction band offset has been performed. In this case, GaAs is used for layers 1, 2, and 3. The electric field of 300 kV/cm is used for layer 1 and 100 kV/cm for layers 2 and 3. The results are plotted in Fig. 4.7. Again, the electron distribution in layer 1 is obtained within 40 Å from the interface while the distribution in layer 2 is calculated as an average of the whole layer of 400 Å width. The results for this case show that an overheated, smooth distribution enters layer 2 (this is true even in the presence of a discontinuity as for Fig. 4.5) and relaxes to a cool, smooth distribution. Steep peaks near the valley minima as observed in Fig. 4.5 do not appear. Therefore, the peaks near the $\Gamma$ and L valley minima are due to the effects of the well. Clearly the enhanced probability of impact ionization and the associated energy loss, and the confinement effect due to the well boundary (the layer 2-layer 3 potential barrier) will lead to an accumulation of cooler electrons at the $\Gamma$ valley minimum. Indeed, a pronounced peak can be found at low energies as shown in Fig. 4.5. The growth of the peak near
Figure 4.5  Electron energy distributions in structure (b) of Fig. 4.1. The electric field in layer 1 is 500 kV/cm, while the field in layers 2 and 3 is 100 kV/cm.
Figure 4.6 Time evolution of electron energy distribution in layer 2 for structure (b) of Fig 4.1. The operating conditions are those of Fig 4.5.
Figure 4.7  Electron energy distributions in layers 1 and 2. GaAs is used for all the layers. The electric field in layer 1 is 300 kV/cm, while the field in layers 2 and 3 is 100 kV/cm.
the $\Gamma$ minimum in Fig. 4.6 (between $t=6$ psec and $t=12$ psec) is due to the further confinement of cold electrons. Due to the potential barrier, the electrons in the L valley minima also experience confinement. However, the band offset for the L valleys is very small (0.07 eV when the Al mole fraction is 0.36), and the confinement by this small barrier alone can not explain the distinctive peak exhibited in Fig. 4.5. These peaks at higher energies close to the L and X minima can be explained by the following two effects. When electrons enter the quantum well and abruptly gain energy from the band offset, the distribution experiences the strong enhancement of intervalley scattering. Hence, due to this enhancement and the differences in the band structures, the L and X valleys are overpopulated at the beginning stage. However, the scattering rates of electrons from the L and X valleys back to the $\Gamma$ valley are small due to the small density of states at the $\Gamma$ minimum (compared to L and X). Combined with the rapid cooling of the distribution as it propagated in the GaAs layer, the electrons at the L and X valley minima encounter a bottleneck for their return to lower energies. Therefore, as the overheated electron gas cools down fast, very pronounced peaks develop at the L and X minima in contrast to the case of homogeneous structure which shows only relatively broad peaks encompassing the energies close to both L and X minima.

These effects occur, of course, only if the well is not too wide. As the well width is increased, the energy distribution gradually approaches the smooth form of the steady-state distribution as shown in Fig. 4.8. For the case of the larger well width, the cold electrons can be reheated by the relatively high electric field (100 kV/cm) while the total number of ionization events is virtually constant. Therefore, the population in the low-energy peak decreases significantly and the bottleneck effects diminish also due to the longer relaxation. The distribution shows similar changes as the electric field applied becomes smaller. Figure 4.9 shows the energy distribution in the GaAs layer with the electric field in layer 1 ranging from 500 kV/cm to 100 kV/cm. Although the distribution entering the GaAs layer is very different in each case,
Figure 4.8  Electron energy distribution in layer 2 for structure (b) of Fig. 4.1. The electric field in layer 1 is 500 kV/cm, while the field in layers 2 and 3 is 100 kV/cm. The width of layer 2 varies from 400 Å to 600 Å.
Figure 4.8  Electron energy distributions in layer 2 for structure (b) of Fig. 4.1. The electric field in layers 2 and 3 is 100 kV cm while the field in layer 1 varies from 500 kV cm to 100 kV cm.
electrons in the GaAs layer have very similar distribution through the rapid relaxation process. Yet the distributions are not fully relaxed and show the dependence on the initial condition. The electrons injected with smaller energy (i.e., smaller field) have a less populated high-energy tail and, therefore, a smaller $\Gamma$ valley peak. Also the peak at the $L$ valley minima becomes less pronounced with smaller injection energy since the less overheated distribution can relax faster.

As the band offset is decreased, the confinement and the bottleneck effects are expected to decrease rapidly. When the offset is reduced to 0.10 eV, the peaks at the $\Gamma$ and $L$ minima virtually disappear as shown in Fig. 4.10. It is also interesting to note that the distribution for the smaller potential energy gain (i.e., smaller band discontinuity) has a larger high-energy tail. This is because with larger band discontinuities the distribution cools down more due to the stronger confinement of cold electrons.

The above effects of electron reflection, enhanced intervalley scattering and impact ionization due to overheating, and the backscattering bottleneck correspond to the band structure of the material in a significant way. Figure 4.11 shows a two-dimensional plot ((001) plane) of the momentum distribution function in structure (b) of Fig. 4.1. In this figure, the momentum distribution clearly reveals the cubic symmetry of the crystal. A pronounced accumulation of electrons in the $\Gamma$-sphere and $L$- and $X$- ellipsoids makes the band structure highly visible. It is interesting to note that the shift due to the electron drift (in the <100> direction) is rather small in spite of the still high electric field of 100 kV/cm. Therefore, the standard approximation of small drift in hot electron theory is justified. However, the picture of a displaced Fermi-sphere is totally incorrect. The constant energy surface is a very special one in this case, and one should rather call it a displaced Boltzmann cube. The cubic symmetry of the crystal is also found in the distribution function for homogeneous GaAs. However, due to the lack of accumulated electrons at the valley minima, the structure is less pronounced and the distribution of electrons is more diffuse, as shown in Fig. 4.12. When a high electric field is applied, the cubic
Figure 4.10 Electron energy distributions in layer 2 for structure (b) of Fig. 4.1. The electric field in layer 1 is 500 kV/cm, while the field in layers 2 and 3 is 100 kV/cm. The Al mole fraction varies from 0.36 (0.31 eV) to 0.12 (0.10 eV).
Figure 4.11 Electron momentum distribution in layer 2 plotted on the (001) plane. The operating conditions are those of the 300 kV/cm case in Fig. 4.9.
Figure 4.12  Electron momentum distribution in layer 2 plotted on the (001) plane. The operating conditions are those of Fig. 4.7
symmetry is not clearly reflected in the momentum distribution as shown in Fig. 4.13. Since the
distribution is so widely scattered, the details of the band structure (i.e., \( \Gamma \)-sphere, \( L \) and \( X \)-
elipsoids) are blurred. The picture of a displaced sphere can be more successfully used in this
case. A larger shift in the distribution is also noted in this case.

We now turn to macroscopic transport effects, in particular, to impact ionization. Impact
ionization in heterostructures is a process of high complexity due to its nonlocal nature and
heavy dependence on (subtle) details of the band structure. The electron impact ionization rate
\( \alpha \) in the GaAs layer is studied below as a function of distance. The zero of the coordinate sys-
tem is chosen at the layer 1-layer 2 interface. Figure 4.14 shows the rate \( \alpha \) in structure (a)
shown in Fig. 4.1. In this figure, \( \alpha \) exhibits what one could term an overshoot at a certain dis-
tance after the electrons enter the GaAs layer. To clarify the reasons for this effect, the evolu-
tion of the energy distribution function has been studied as the electrons leave the AlGaAs layer
and transfer to the GaAs layer experiencing the same field everywhere (i.e., in structure (a)).
This study shows that there are two pronounced effects which influence \( \alpha \). The first is the well-
known enhancement of the ionization due to the band edge discontinuity which overheats the
electron gas and increases, therefore, the ionization probability.\(^{12} \) This effect alone would lead to
a monotonical decrease in \( \alpha \) from its overheated value to the steady-state value.

The second effect arises from the difference in the band structure of the neighboring materi-
als. Depending on the the density of states at high energies, the high-energy tail of the distribu-
tion can be larger or smaller in different materials. If the density of the states is large, the
scattering by phonons increases. As a result, the energy loss becomes high and the high-energy
tail of the distribution is suppressed. In our particular band structure models, the density of
states at and above 2 eV is larger in AlGaAs. Therefore, the high-energy tail in the distribution
is smaller in the AlGaAs layer than in the GaAs layer. As electrons propagate into layer 2, they
Figure 4.13  Electron momentum distribution in layer 2 plotted on the (001) plane. The operating conditions are those of Fig. 4.2.
Figure 4.14  Impact ionization rates in layer 2 for structure (a) of Fig. 4.1. Uniform electric fields ranging from 500 kV/cm to 300 kV/cm are applied throughout the layers (layers 1, 2, and 3). The zero of the coordinate system is chosen at the layer 1-layer 2 interface.
abruptly gain energy from the band edge discontinuity. This enhances the ionization rate. However, in the first portion of the well, the distribution still has a rapidly decreasing high-energy tail, which exhibits the shape typical for AlGaAs as shown in Fig. 4.15. Since the scattering rate at high energies is smaller in GaAs (small density of state), electrons can be accelerated toward the more accessible high-energy tail in GaAs fast enough so that the ionization rate can increase while the overheating of the distribution is relaxed on average. Therefore, $\alpha$ increases over a distance which is needed for the high-energy distribution to adjust itself to the GaAs band structure. Subsequently a decrease to the steady-state value follows. As the electric field in layer 2 is decreased, the two effects competing with each other may still result in a monotonically decreasing $\alpha$. Even though electrons begin to adapt to the band structure with more accessible high energy states, the distribution can be cooled down fast enough by the relaxation to overcompensate for the increase of the tail. The ionization rate for structure (b) of Fig. 4.1 is shown in Fig. 4.16. Due to the smallness of the field in the GaAs layer (100 kV/cm), the relaxation becomes dominant and the ionization rate monotonically decreases as electrons propagate in this case.

As mentioned previously, impact ionization is very sensitive to the details of the band structure. However, the current knowledge of the band structure (especially AlGaAs) is not precise. As a consequence, we do not know if the overshoot of $\alpha$ indeed exists in this material system. However, the physical reasons for the effects are clear and therefore there may exist a material system which will show an effect of this kind.

Finally, even though the electron ionization rate in the narrow gap layer (GaAs) has been enhanced by the heterojunction as shown in Fig. 4.14, this enhancement may not be large enough to compensate for the small ionization rate in the wide gap layer (AlGaAs) where virtually no ionization occurs. In spite of this point, the $\alpha/\beta$ ratio can be still enhanced significantly. The enhancement of $\alpha$ improves as the field becomes low.
Figure 4.15 Spatial evolution of electron energy distribution in layer 2 for structure (a) of Fig. 4.1. The operating conditions are those of Fig. 4.2.
Figure 4.16  Impact ionization rates in layer 2 for structure (b) of Fig. 4.1. The electric field in layer 1 varies from 500 kV/cm to 300 kV/cm, while the field in layers 2 and 3 is fixed to 100 kV/cm.
4.4 Conclusion

A numerical study of nonlinear transport across heterojunction interface in strong electric fields has been performed. Two typical heterostructures have been chosen to analyze the effects of a quantum well on the electron energy and momentum distributions. In the quantum-well APD structure with a high field in the well layer, electrons experience negligible confinement in the well layer in spite of heavy ionization and a potential barrier. On the other hand, the doped quantum-well APD structure (relatively small field in the well layer) shows much stronger relaxation and confinement. Reflections at the heterointerface, carrier overheating at the potential step, and the $L \rightarrow \Gamma$, $X \rightarrow \Gamma$ back scattering bottlenecks produce pronounced structures in the energy and momentum distributions which directly reveal the band structure. It is also found that the ionization rates are heavily dependent on the band structure of the neighboring layer. An ionization overshoot effect may exist in these structures due to the band structure difference between two adjacent layers.
CHAPTER 5

ELECTRON HEATING BY HOT PHONON EFFECTS IN HETEROSTRUCTURES

5.1 Introduction

Recently the effects of the nonthermal phonon distributions known as hot phonon effects have received much attention. The slow energy relaxation rate of photo-excited hot carriers in time-resolved measurements has often been attributed to a nonequilibrium phonon distribution in bulk GaAs and quantum wells. Also, the phonon side band operation in photo-excited quantum-well heterostructure lasers has been explained by nonequilibrium phonon effects. The phonon distribution can be perturbed as well when a large number of carriers propagate across an interface, experience an abrupt energy gain, and subsequently relax through strong phonon emission. As a general rule, the phonon distribution will be disturbed whenever a significant number of phonons (especially zone-center phonons) are emitted in a time scale shorter than the phonon decay time constant.

To understand the physical mechanism of this process, Lugli and coworkers have applied a numerical approach and studied the evolution of the phonon distribution and its effects on the energy relaxation rate when hot electrons are injected by a strong light pulse. No such numerical effort has yet been made to analyze the phonon distribution and its effect on the carrier transport in a heterostructure with charge carriers propagating across an interface. This situation is very important in many problems involving thermionic emission such as quantum-well laser diodes.

This chapter presents a numerical study of the effects of nonthermal phonon distributions on electron transport across heterolayers. For the microscopic analysis, not only the electron dynamics but also the phonon dynamics are studied in detail. In Section 5.2 the model applied
in the analysis is briefly described. The effects of conduction band discontinuity on the phonon distributions are explained and thoroughly discussed along with the influence of a wide range of operating conditions in Section 5.3. The conclusion will be followed at the end.

5.2 Method of Analysis

The specific structure under consideration is a GaAs-AlGaAs single-well heterostructure as shown in the inset of Fig. 5.1, which is essentially the same with the structure discussed in Chapter 4. Each layer is assumed to be undoped with an unintentional impurity concentration of $1 \times 10^{14} \text{cm}^{-3}$. The Al mole fraction used for layers 1 and 3 is chosen to be 0.2 which results in the conduction band discontinuity of 0.17 eV at the interface assuming $\Delta E_c = 0.70 \Delta E_g$. Throughout the simulation, it has been assumed that electrons are injected into the well layer after traveling through layer 1. As discussed in the previous chapter, layer 1 is chosen long enough (1500 Å) to rule out possible nonlocal effects of the electron energy distribution. A 400 Å-thick GaAs layer forms the well layer. In the well layer, size quantization and resonance effects are not included since the width is longer than the inelastic scattering mean-free path. For moderately high electric fields ($\sim 50 \text{kV/cm}$), the scattering mean-free path is approximately 50-100 Å. The specific width for layer 3 is not important for this study. For the simulation, an ensemble Monte Carlo method is used along with a realistic band structure in the GaAs layer. Included scattering mechanisms are polar optical phonon (POP) and intervalley phonon scattering as well as electron-electron scattering. Details of the treatments (including the boundary conditions at the interface) and the material parameters can be found in Chapter 4.

To account for the hot phonon effects, the treatment of POP scattering, which involves the zone-center longitudinal-optical (LO) phonons, has been modified to include a nonequilibrium phonon distribution. Although intervalley scattering also occurs mostly via phonon emission, its contribution to the hot phonon effects is negligible since intervalley scattering mainly involves
Figure 5.1 Nonequilibrium LO-phonon distributions $n_q(x,t)$ at three different locations in the well with the injection current density of $1 \times 10^4$ A/cm$^2$. The inset shows the structure considered in the calculation (with a uniform electric field). The equilibrium phonon distribution $n_q^0 = 1/\exp(\hbar \omega_{LO}/kT) - 1$ is approximately 0.34 at 300 K. The acousticalization constant $2\pi/a$ is $1.11 \times 10^3$ cm$^{-1}$ for GaAs. The spatial mesh size is 50 Å.
near zone-boundary phonons with large phonon wavevectors. When the angular dependence of
the phonon distribution in momentum space $\mathbf{q}$ is neglected, the perturbation in the phonon dis-
tribution $\Delta n_q$ can be obtained by dividing the net number of emitted phonons $\Delta N_q$ by the pho-
non density of states $D_q$ (i.e., $\Delta n_q = \Delta N_q / D_q$). The density of states $D_q$ is proportional to $q^2$.
Thus, the perturbation $\Delta n_q$ by phonons with large $q$ is suppressed rapidly as $q$ increases and
POP scattering (which involves zone-center phonons) can be more effective in perturbing the
phonon distribution than intervalley scattering even though intervalley scattering has a larger
scattering rate. Therefore, only the perturbation in the LO-phonon distribution by POP scatter-
ing is considered. During the simulation, the phonon absorption and emission events by POP
scattering are monitored both in time and space. Based upon this information from POP
scattering and the phonon-phonon relaxation processes which include the decay to zone-
boundary acoustic phonons, the position-dependent LO-phonon distribution $n_q(x,t)$ is recalcu-
lated every 100 fs as a function of the amplitude $q$ of the phonon wavevector. The anisotropy
of $n_q(x,t)$ which may occur in the presence of strong electric fields is not included in this study.
For the phonon-phonon relaxation process, the phenomenological relaxation time $\tau_{\text{ph-pb}}$ has been
measured to be 7 psec at 77 K, and this value is used in absence of direct measurement data
for 300 K. Since the relaxation time $\tau_{\text{ph-pb}}$ is not a strong function of temperature, $\tau_{\text{ph-pb}}$ at
300 K should indeed be fairly close to 7 psec. With the newly calculated phonon distribution
$n_q(x,t)$, the POP scattering rates are also recalculated and the probability distribution of the
final states after POP scattering is modified to reflect this new phonon distribution. Possible
effects of screening have not been included.

5.3 Results and Discussion

As the electrons enter the well from the barrier (layer 1), electrons experience an abrupt
energy gain at the heterointerface. The overheated electrons then relax rapidly and cause the
nonthermal phonon population in the well to increase. When electrons are injected with higher energy, more phonons are generated. However, as the average injection energy increases, intervalley scattering becomes more important and, thus, many of the phonons involved in the process become zone-boundary phonons which are, as mentioned, ineffective in disturbing the phonon distribution. Therefore, the injection of electrons with average injection energy well above the Γ-L separation is not desirable for the study of hot phonon effects. Hot phonon effects are also diminished in the presence of strong electric fields within the well since electrons can be quickly reheated and populate the energy range of predominant intervalley scattering. In Fig. 5.1, simulation results are shown for a uniform electric field of 50 kV/cm throughout the layers (1-3) and an injection current density of $1 \times 10^4$ A/cm$^2$. The mesh size used for the calculation of the position-dependent quantities has been 50 Å and the zero of the coordinate system $x=0$ is at the layer 1-layer 2 interface as shown in the inset of Fig. 5.1. The steady-state LO-phonon distribution is shown in Fig. 5.1 at three different positions in the well. The deviation from the equilibrium phonon distribution is relatively small at the layer 1-layer 2 interface and grows as the electrons travel further into the well. This is because at the entrance in the well, the POP emission rate is relatively small due to the high average injection energy and the existence of a short "near ballistic" range for the injected electrons. Near the layer 2-layer 3 interface, the reflections due to the potential barrier lead to an accumulation of low-energy electrons. These electrons with low energies tend to emit phonons within a narrow range of relatively large wavevectors as can be seen from simple considerations of energy and momentum conservation. Therefore, they cause a long tail in the disturbed phonon distribution as is shown in the figure. Similar elongation in the disturbance of the phonon distribution has also been observed in Ref. 108. These effects do not exist in very wide wells since the electrons gradually relax to equilibrium toward the layer 2-layer 3 interface.
The nonequilibrium phonon distribution affects the POP scattering rates considerably. The equilibrium and nonequilibrium POP scattering rates are compared in Fig. 5.2. The rates with hot phonon effects in this figure correspond to the LO-phonon distribution at the layer 2-layer 3 interface in Fig. 5.1, where the phonon perturbation is pronounced. It is important to note that stronger enhancement is observed in the absorption rate compared to the emission rate, especially at low energies. This change in the scattering rate results in the difference in the electron energy distribution. Figure 5.3 shows the electron energy distribution at the layer 2-layer 3 interface. Clearly, inclusion of hot phonons leads to an increase in the electron temperature as shown by the fact that there are more electrons at high energies. Due to the strong enhancement in the phonon absorption rate, cold Γ valley electrons in the well can reabsorb the nonequilibrium phonons which have been emitted by the hot electrons injected over the barrier. Thus, with the inclusion of hot phonons, more electrons transfer to the L valleys and contribute to a more pronounced peak at the L minima. It is interesting to note that the backscattering bottleneck effect\textsuperscript{113} is also observed at the energy of the L valley minima. Another interesting point is that two different procedures which have been used to calculate the hot phonon effects give almost identical energy distributions. Procedure (A) is the method described above in which both the POP scattering rate and the probability distribution of the final states after POP scattering is updated. In procedure (B), the POP scattering rate is recalculated based on the updated phonon distribution, but the probability distribution of the final states for thermal equilibrium is used to choose the final states after POP scattering. Evidently, procedure (B) has a different preference for the phonon wavevectors involved in the scattering and generally results in different final electron states compared to procedure (A). However, since a constant phonon energy is used for POP scattering, the change in electron energy is the same in both procedures.

The heating by the reabsorption of hot phonons reduces the confinement of electrons in the well. Figure 5.4 shows the time evolution of the number of electrons in the well (normalized to
Figure 5.2  Calculated POP scattering rates in equilibrium and nonequilibrium cases. The nonequilibrium rates correspond to the LO-phonon distribution at the layer 2-layer 3 interface in Fig. 5.1.
Figure 5.3  Electron energy distributions at the layer 2-layer 3 interface. For procedure (A) the probability distribution for the final states is updated to reflect the nonequilibrium distribution, while the probability distribution at thermal equilibrium is used for procedure (B).
Figure 5.4 Time evolution of the number of electrons in the well (normalized to the steady-state value) after the injection current is turned off at $t=0$. The inset shows the corresponding development of the average energy in the well. The number of electrons and the average energy are averaged over layer 2.
the steady-state value) after the injection current is turned off at \( t=0 \). Due to the residual electrons in layer 1, the response to the switch-off is delayed. As can be seen in the figure, electrons are emitted from the well about two times faster than without the inclusion of hot phonon effects. The corresponding development of the average electron energy in the well with time is shown in the inset of Fig. 5.4. Due to the emission of electrons with high energies above the barrier, the average energy falls rapidly and then approaches equilibrium in the well. As expected, the difference in average energies (with and without the inclusion of hot phonon effects) decreases with time but is still sizable after several picoseconds due to the presence of hot phonon absorption.

As the electric field in the well decreases, the electrons experience stronger relaxation toward lower energies which results in an energy distribution close to equilibrium. The decrease in the average injection energy also leads to a reduction in the electron temperature in the well. As the electron distribution cools down, the effect of phonon reabsorption from the nonequilibrium distribution grows since an increasing fraction of electrons resides near or below the threshold energy for POP emission where only absorption is allowed. Figure 5.5 shows the electron energy distribution in the well when the electric field is zero in all layers, which is analogous to the operating conditions in quantum-well laser diodes. In contrast to Fig. 5.4, the injection of electrons into the well layer is started at \( t=0 \) from the layer 1-layer 2 interface by thermionic emission in this calculation. The same injection current density as for Fig. 5.1 is assumed. Since the simulation is very time-consuming under these circumstances, electrons are simulated only for the first 12 psec. The distribution displayed in Fig. 5.4 represents an average over the whole layer 2 (400\( \text{Å} \)) at \( t=12 \) psec. As shown in the figure, a significant portion of the electrons near the threshold energy for emission absorbs phonons and moves to higher energy. About 30\% increase in the electron temperature is therefore observed (inset of Fig. 5.5) compared to about 15\% increase in the previous case (inset of Fig. 5.4). The hot phonon distribution will have an even larger effect in steady state since the effect is still growing with time at \( t=12 \) psec.
Figure 5.5 Electron energy distribution in the well at $t=12$ psec with no electric field applied. The time development in the average energy in the well is shown in the inset. The injection current is turned on at $t=-\infty$. The distribution and the average energy are averaged over layer 2.
The injection current density is an important parameter as large current densities enhance hot phonon effects. A smaller but still considerable disturbance of the phonon distribution is seen in Fig. 5.6 with an injection current density of $1 \times 10^3 \text{A/cm}^2$. With very high injection currents, the perturbation of near zone-boundary phonons may also become important. The hot phonon effects will, of course, be more visible at lower temperature. Due to the small phonon occupation at thermal equilibrium, phonon absorption is possible only in nonequilibrium situations at low temperature.

5.4 Conclusion

The effects of the nonequilibrium phonon distribution on electron transport across heterointerfaces have been studied by an ensemble Monte Carlo simulation method. A strongly disturbed phonon distribution is observed as a consequence of the conduction band discontinuity. The hot phonon distribution causes in turn more energetic electron distribution because of phonon reabsorption and reduces, therefore, the confinement of electrons in the well. As the average electron energy decreases in the well, the effect of phonon reabsorption gains significance since more electrons are near or below the threshold energy for POP emission where only absorption is allowed.
Figure 5.5  Nonequilibrium LO-phonon distributions $n_\alpha(x,t)$ at three different locations in the well with the injection current density of $1 \times 10^3 A/cm^2$. 
CHAPTER 6

ELECTRON TRANSPORT IN TUNNELING HOT ELECTRON TRANSFER AMPLIFIERS

6.1 Introduction

The idea of ballistic transport in thin semiconductor layers has received considerable attention. However, experimental verifications of such transport have been rare and indirect. Recently, Levi et al.\textsuperscript{15} and Heiblum et al.\textsuperscript{16} have independently reported direct observation of ballistic transport with the use of hot electron transistors as an electron energy spectrometer. Based on studies of the collector current-voltage characteristics, they have deduced the energy spectra of electrons which travel through the heavily doped short base region, and have shown the existence of a ballistic (or quasiballistic) peak. Since then, a number of theoretical analyses\textsuperscript{114-117} have been performed to investigate the details of this nonequilibrium transport phenomenon and to assess the feasibility of hot electron spectroscopy. Most of these efforts have been devoted to studying transport in the base as a function of the collector barrier height. For the spectrometer with a planar doped barrier structure,\textsuperscript{15} such analyses are appropriate since the collector current is controlled mainly by the collector barrier height (collector-base bias). For the tunneling hot electron transfer amplifier (THETA) structure,\textsuperscript{16} however, the collector current can in principle be influenced not only by the collector-base bias but also by the details of transport in the relatively thick ($\sim 1000 \, \text{Å}$) collector barrier region with a retarding field.

In this chapter, the electron transport in a THETA structure is studied numerically. The evolution of the electron energy distribution (width and location of the peak) in the base and the collector barrier is simulated in detail to examine the validity of hot electron spectroscopy. The device characteristics as a transistor including transfer ratio and transit time are also investigated. The simulation results are compared to the experiments of Heiblum and coworkers.\textsuperscript{16}
6.2 Computational Model

The conduction band profile of the structure under consideration is schematically shown in Fig. 6.1. The structure consists of two undoped AlGaAs layers between three heavily doped GaAs layers \( N_D = 1 \times 10^{18} \text{ cm}^{-3} \), where \( N_D \) represents the donor concentration. The Al composition and the widths of the two AlGaAs layers are 0.3, 120 Å, and 1000 Å, respectively. These conditions are chosen to match the experiment by Heiblum et al.\(^1\) A slightly larger base (400 Å) has been chosen to encompass the width of the grading region. The compositional grading itself is not included in the simulation. For the electron transport study, an ensemble Monte Carlo method has been used. Analytical \( \Gamma-L-X \) band structures with nonparabolicity are incorporated for both GaAs and AlGaAs. Using \( \Delta E_c = 0.68 \Delta E_g \) for the conduction band discontinuity,\(^7\) the particular Al mole fraction, 0.3, corresponds to a bandedge discontinuity of 0.25 eV at the layer boundaries. Material parameters that have been used are described in Chapter 4.

Included scattering mechanisms are polar optical phonon and intervalley phonon scattering for the undoped AlGaAs layers (barriers), and elastic ionized-impurity and coupled plasmon/phonon scattering as well as intervalley phonon scattering for the heavily doped GaAs layer (base). Due to the degenerate electron concentration, the polar optical phonon modes in the base are strongly coupled to the collective plasmon mode and the electron-hole excitation through long-range polarization fields. The scatterings by the coupled plasmon/phonon modes are basically treated as discussed in Chapter 2. In addition to the three-dimensional treatment in Chapter 2, a quasi-two-dimensional correction is also included in this study. As the device dimension shrinks, the plasmon mode is strongly influenced by the reduced dimensionality. The electron gas does not, of course, support a collective motion perpendicular to the layer interface whose wavelength is larger than the layer width. The exact inclusion of such effects involves a
Figure 6.1 Conduction band profile for an AlGaAs/GaAs tunneling hot electron transfer amplifier when a retarding field is applied in the collector barrier. The doping concentration in the $n^+$-GaAs layers is $1 \times 10^{18}$ cm$^{-3}$. The AlGaAs layers are undoped. The Al composition is 0.3.
great deal of complexity and has not been included in this study. Instead, a maximum wavelength limit is simply imposed to the plasmon-like mode based on the base width (device limit). The possibility of interaction with two-dimensional plasmons which propagate parallel to the layer interface still exists. Since, however, the electrons entering the base initially have negligible parallel momentum, the coupling to the two-dimensional plasmon can be neglected due to the tight restriction in the energy and momentum conservation as discussed in the following. Our treatment poses as a first-order approximation to include two-dimensional effects.

Let $k$ be the momentum of initial state and $k'$ the momentum of final state after scattering. Then the energy conservation for plasmon emission gives

$$\frac{\hbar^2 k'^2}{2m^*} + \hbar \omega_p(q) = \frac{\hbar^2 k^2}{2m^*} \tag{6.1}$$

when a parabolic band structure is assumed with the effective mass $m^*$. The plasmon dispersion relation in two dimensions\textsuperscript{118} gives

$$\omega_p^2(q) \simeq \alpha q + \beta q^2 \tag{6.2}$$

where $\alpha = 2e^2 E_F/\hbar^2$ and $\beta = 3E_F/2m^*$. Here, $E_F$ is the Fermi energy. Assuming $k-k'=q$, Eq. (6.1) gives

$$\frac{q}{2k} + \frac{m^*}{\hbar} \frac{(\alpha q + \beta q^2)^{1/2}}{q} = \cos \Theta. \tag{6.3}$$

Therefore, for a given $k$, there is a critical angle $\Theta_c$ above which an appropriate $q$ (in turn, final state) can not be found. Assuming $\alpha q >> \beta q^2$ which is appropriate at long wave length, an estimate of $\Theta_c \sim 40^\circ$ is obtained with $m^*/m_0 = 0.067$ and $E(k) = 0.4$ eV for an electron density of $1x10^{12}$ cm$^{-2}$ at $T = 0$ K. As $E(k)$ decreases, $\Theta_c$ also decreases to a smaller value. This estimate shows that two-dimensional plasmon scattering is an ineffective scattering mechanism unless the electrons have considerable parallel momentum (i.e., $0 \leq \Theta \leq \Theta_c$). Although the conservation of
perpendicular momentum \( \vec{k}_1 = \vec{k}'_1 \) can not be justified in full due to the lack of translational invariance, this estimate shows qualitatively that two-dimensional plasmon scattering is of limited importance.

Some of the calculated inelastic scattering rates are shown in Fig. 6.2 as a function of energy. Two different procedures have been used to obtain the total inelastic scattering rates. Procedure (A) includes the device limit (base width) in the plasmon-like mode as described above, while procedure (B) does not. As can be seen in the figure, these two procedures exhibit sizable differences in the scattering rate. Although the Landau damping mode is the most dominant scattering mechanism and is treated independent of the base width in this study, the difference between the two treatments (A and B above) of the plasmon-like mode affects the outcome of the transport study to a considerable degree. This further emphasizes the importance of precise treatment of the finite base size. For ionized impurity scattering, the well-known Brooks-Herring model has been used along with Thomas-Fermi screening. After the scattering, electron states are randomly chosen based on energy and momentum conservation. Due to the broadened resonance frequencies, the Landau damping mode involves more difficulties in the selection of final states than other mechanisms. In this study, the amount of energy loss \( \hbar \omega \) for a particular scattering event by the Landau damping mode is determined by a precalculated distribution \( F(\omega) \) shown in Fig. 6.3. After choosing \( \hbar \omega \), the change in momentum is decided by another look-up table. The final state of the electron excited from the Fermi sea is subsequently picked based on the changes in the scattered electron. The Pauli exclusion principle (zero temperature) is included for the final-state selection.

Since the temperature considered is very low (4.2 K), it is assumed that electrons are injected into the base mainly by tunneling through the emitter barrier. Assuming further that the emitter barrier tunneling probability \( D(E) \) depends only on longitudinal electron energy \( E_x \),
Figure 6.2  Calculated inelastic scattering rates for Γ valley electrons in the base. Procedure (A) for total inelastic scattering rate includes the device limit (base width) in the plasmon-like mode, while procedure (B) does not.
Figure 6.3 Probability distribution $F(\omega)$ for the amount of energy loss $\hbar \omega$ in a particular scattering event by the Landau damping mode. The distribution $F(\omega)$ is defined as

$$F(\omega) = \int_{q_{\text{min}}}^{q_{\text{max}}} \frac{dq}{q} (\text{Im} \frac{1}{\epsilon(q, \omega)}),$$

and is calculated as described in Chapter 2. Nonparabolicity is included in the calculation. To obtain the probability distribution for a specific initial energy $E$, $F(\omega)$ should be renormalized to the area under $F(\omega)$ from $\hbar \omega = 0$ to $E - E_F$. 

$N_D = 1 \times 10^{18}$ cm$^{-3}$
it can be shown that the energy distribution of the injected electrons is proportional to 
\((E_F - E_x)D(E_x)\) for \(E_x < E_F\) at 0 K, where \(E_F\) is the Fermi energy in the emitter. In the simulation, electrons are continuously launched at every 10 fsec from the emitter barrier-base interface based on the energy distribution discussed above, and are tracked until they leave the device or join the cold base electron sea. The cold base electrons are not simulated unless excited from the Fermi sea through scatterings by the Landau damping modes. The transfer matrix formalism discussed in Chapter 2 is used in the calculation of tunneling probability \(D(E_x)\). Parallel momentum is initially neglected since it is much smaller than the perpendicular momentum. Coherent interference effects\(^{39}\) are not considered in this study. At the interfaces, electron transfer for the \(L\) and \(X\) valley electrons is treated classically by conserving energy and parallel momentum. An electron can cross the interface when it has high enough energy \(E_x\) to overcome the barrier. The valley to which the electron belongs is assumed to be unchanged during the transfer. For the \(\Gamma\) valley electrons, quantum mechanical reflection is included. The reflections at the interfaces are assumed to be specular. When an electron in the base has smaller energy than the interface barrier height, it is assumed that the electron either reaches the base electrode or recombines with a hole. Since the continuous launching scheme is used as mentioned above, the collector current can be obtained by the number of electrons that reach the collector region per unit time. Finally, the effect of band bending is not included in the simulation.

6.3 Results and Discussion

The electron energy distributions are shown in Fig 6.4 for two different locations and an emitter-base bias \(V_{eb}\) of -0.35 V and a collector-base bias \(V_{cb}\) of -0.10 V. The injected energy distribution is estimated as described above and shows a peak at \(\sim 15\) meV below the Fermi level with a full width of \(\sim 35\) meV at half maximum. The distribution after the base is obtained at the collector barrier side of the interface based on the incoming electrons from the
Figure 6.4  Electron energy distributions at various locations along with $\frac{\partial I_c}{\partial V_{cb}}$ when the emitter-base bias is -0.35 V. The electrons are ballistically injected into the base from the emitter through tunneling. The distribution after the base has been obtained at the collector barrier side of the base-collector barrier interface based on the incoming electrons from the base. The zero of the potential energy is chosen at the beginning of the collector barrier which is $\Delta E_c$ above the base potential. $\frac{\partial I_c}{\partial V_{cb}}$ has been calculated directly from the simulation result of the collector current-voltage characteristics.
base. The electrons traveling back to the base are not considered in the distribution. Since the electrons which are backscattered into the base from the collector barrier can reenter the collector barrier after suffering further scattering in the base, the distribution after the base is a function of $V_{cb}$. This dependence, however, is not prominent in this study due to the relatively large conduction band discontinuity $\Delta E_c$. Hence, the distribution calculated at a particular $V_{cb}$ can reflect the intrinsic base transport characteristics. The zero of the potential energy is chosen at the beginning of the collector barrier which is $\Delta E_c$ above the base potential. As can be seen in the figure, the electrons preserve the shape of the injected distribution after traveling through the base. While the width is slightly widened by the scattering, the peak of the distribution after the base has not been displaced, thus demonstrating the ballistic (or quasiballistic) transport in the base. Indeed, a significant portion ($\sim 55\%$) of the electrons which enter the collector barrier has passed through the base without a collision. Some electrons which suffer small angle scatterings (especially elastic) overcome the potential barrier and contribute to the long tail in the distribution. Contributions by the electrons excited from the Fermi sea is negligible. No distinctive secondary peak due to phonon replica is observed in the tail. As discussed previously, the scattering by the Landau damping mode is the most frequent mechanism in the base. However, the $\Gamma \rightarrow L$ intervalley scattering is more detrimental to the base transport due to its randomizing nature. Electrons which suffer the intervalley scattering in the base have little chance to overcome the barrier and eventually contribute to the base current.

To investigate the experimental evaluation of the distribution, we have calculated the derivative $\partial I_c/\partial V_{cb}$ and plotted it also in Fig. 6.4. It is seen that $\partial I_c/\partial V_{cb}$ exhibits a pronounced peak which can be clearly related to the ballistic peak of the distribution after the base region (thus, the injected distribution). The location and the width of the peaks obtained from these three different curves are identical within the simulation error. Such coincidence implies
that at sufficiently large negative $V_{cb}$ only ballistic (or quasiballistic) electrons contribute to the collector current, and the distribution at the end of the collector barrier still resembles closely the injected distribution. This is clearly shown in Fig. 6.5 where the energy distribution at the collector barrier-collector interface is plotted for various $V_{cb}$. For $V_{cb} = V_{eb} + \Delta E_c$ (here, $V_{cb} = -0.10 \text{ V}$), the distribution reproduces the shape of the injected distribution after traveling through the base and the collector barrier. More than 75% of the transmitted electrons are ballistic in this case. As the negative bias increases, the electrons keep virtually the same distribution, losing only the portion which is reflected by the potential barrier. A slight growth in the peak is also observed, which is due to reduced scattering in the collector barrier. These results prove clearly the concept of ballistic transport and its detection in THETA devices. As $V_{cb}$ becomes less negative, the bulk of the scattered electrons start to enter the collector. Since the collector current in this regime is strongly influenced by the details of scattering in the collector barrier, it is difficult to deduce information on the base transport from $\frac{\partial I_c}{\partial V_{cb}}$. The evolution of the energy distribution and $\frac{\partial I_c}{\partial V_{cb}}$ have also been studied at $V_{eb} = -0.40 \text{ V}$. Due to the increase in injection energy, electrons are subject to stronger scattering in both the base and the collector barrier. However, no qualitative deviation from Fig. 6.4 is observed as shown in Fig. 6.6.

Figure 6.7 illustrates the ballistic transfer ratio in the collector current as a function of $V_{cb}$ for $V_{eb} = -0.35 \text{ V}$ and $-0.40 \text{ V}$. At large negative $V_{cb}$, the ballistic transfer ratio is large and varies slowly with $V_{cb}$. Perfect ballistic transport is not realized even at very high bias due to the elastic small angle scattering by ionized impurities in the base. A drastic decrease in the ballistic ratio follows as $V_{cb}$ becomes less negative. Since the massive contribution of scattered electrons to the collector current causes this decrease, the location of the sudden drop in the ballistic transfer ratio for $V_{eb} = -0.35 \text{ V}$ matches well with the increase in $\frac{\partial I_c}{\partial V_{cb}}$ shown in Fig.
Figure 6.5  Energy distributions at the collector barrier side of the collector barrier-collector interface with various collector-base bias conditions when the emitter-base bias is -0.35 V. Each curve is normalized differently so that the ratio of two distributions at a given energy represents the ratio of collector current with that energy.
Figure 6.6 Electron energy distributions at various locations along with $\frac{dI_c}{dV_{cb}}$ when the emitter-base bias is -0.40 V. The distribution after the base is obtained at the collector-base bias of -0.15 V. Other conditions are the same with Fig. 6.4.
Figure 6.7  Ballistic transfer ratio in the collector current as a function of collector-base bias.
113

6.4, as discussed previously. With an increase in injection energy, ballistic transport becomes less probable for a given $V_{cb}$ due to the increased intervalley scattering rate.

Thus, the results of our numerical study are in excellent overall agreement with the experiment by Heiblum et al.\textsuperscript{16} However, there are points that need consideration for a complete comparison. In the experiment, the width of the distribution obtained from $\partial I_c/\partial V_{cb}$ ($\sim 60 \text{ meV}$) is larger than in the simulation ($\sim 35 \text{ meV}$). Most of the broadening comes, in our opinion, from the band bending at the emitter-emitter barrier interface. Due to the formation of the accumulation region, the injected distribution of the tunneling current can be widened considerably.\textsuperscript{53} Similar band bending occurs at both the interfaces of the base region and may affect the electron transport considerably. To investigate such effects, a one-dimensional analysis as described in Chapter 2 can be used. Due to the finite base resistance, the choice of base potential should be considered carefully when the base width is very small. The reflections at the collector contact, which is less than ideal, can also broaden the distribution in the experiment. The experimental result for $\partial I_c/\partial V_{cb}$ shows that as $V_{cb}$ becomes less negative, a smaller and relatively flat distribution follows the ballistic peak when compared to the numerical result plotted in Figs. 6.4 and 6.6. This suggests that the simulation predicts a larger population of electrons at low energies and, thus, a stronger energy loss in the direction of propagation by scattering than the experiment includes. Considering that the base transfer ratio agrees closely in both cases, the scattering rate in the collector barrier (mostly polar optical phonon) might have been slightly overestimated. In the simulation, about 36% of the injected electrons transfer the base ballistically.

In normal transistor operation, positive $V_{cb}$ is applied to obtain high $I_c/I_s$ ratio. Due to strong intervalley scattering in the collector barrier, ballistic transport through the device is highly improbable with positive $V_{cb}$. However, electrons still travel significant distances ballistically and, thus, have a higher velocity than the saturation velocity. Figure 6.8 shows the total
Figure 6.8  Total transit time as a function of collector-base bias
transit time as a function of $V_{cb}$. There exists a negative $V_{cb}$ range which minimizes the transit time, corresponding to the condition of large ballistic probability and simultaneously relatively little deceleration of the electrons by the applied retarding field. As $V_{cb}$ becomes less negative, the transit time increases sharply at the onset of intervalley scattering and gradually approaches to a saturation value. At $V_{cb} = 1.0$ V, the transit time is approximately 0.71 psec and 0.74 psec for $V_{eb} = -0.35$ V and -0.40 V, respectively, while the base transit time is approximately 0.05 psec and 0.06 psec. It is difficult to define a base transit time with negative $V_{cb}$ due to the backscattering of electrons from the collector barrier. Without considering the backscattered electrons which hardly contribute to the collector current, the base transit time for negative $V_{cb}$ is expected to be close to the value with positive bias. The total transit time is the parameter which gives the ultimate limit of the switching speed. To reduce the switching time, a short base width is desired. However, a short base width brings large base resistance, which also limits the switching speed. Higher doping in the reduced base may be the solution for this problem. While keeping the base resistance low, higher doping in the narrow base may cause considerably smaller increase in the scattering rate of the injected electrons than in the bulk due to the reduced dimensionality in the density of the final states. Figure 6.9 shows the collector current switching characteristics as $V_{cb}$ switches from -1.0 V to 1.0 V with $V_{eb} = -0.35$ V. At $t = 0.75$ psec, the collector current reaches to more than 90% of the steady-state value, which shows good agreement with the total transit time. There is a possibility that the current rise time can be much shorter than the transit time. When a small negative $V_{cb}$ is applied for the off state, mobile electrons can be effectively stored in the collector barrier and contribute to the collector current rapidly when the bias is switched to positive. An overshoot in current is expected in this situation. As the current density increases, the band bending in the collector barrier due to mobile electrons becomes more important. The effect of a self-consistent field has been found to be negligible even at 500 A/cm$^2$. Notice, however, large current density can also bring the
Figure 6.9 Collector current switching characteristics as the collector-base bias switches from 
-1.0 V to 1.0 V at $t = 0$ with the emitter-base bias of -0.35 V. The collector 
current is normalized to the emitter current.
nonequilibrium phonon distribution known as hot phonon effect\textsuperscript{101} in the base and can alter the base transport significantly.

6.4 Conclusion

In this chapter, the electron transport in a THETA device with a base width of 400 Å and the collector barrier width of 1000 Å has been numerically studied at 4.2 K. Using an ensemble Monte Carlo method which includes coupled plasmon-phonon interaction and the Pauli exclusion principle in the heavily doped base region, we have shown the existence of nearly ballistic transport in the base and the collector barrier for a negative \( V_{cb} \). From \( \partial I_c / \partial V_{cb} \), the ballistic peak of the distribution after the base (thus, the injected distribution) has been recovered which confirms the ideas of hot electron spectroscopy in THETA structures by Heiblum and coworkers. With positive \( V_{cb} \), ballistic transport through the device is highly improbable due to strong intervalley scattering in the collector barrier. However, electrons still travel much of the device ballistically and, thus, have a higher velocity than the saturation velocity. The total transit time which gives the ultimate limit of the switching speed is approximately 0.71 psec at \( V_{cb} = -1.0 \) V when \( V_{eb} = -0.35 \) V. To obtain optimum transit time and current gain, the electron injection energy in the base should be lower than the threshold for \( \Gamma \rightarrow L \) phonon emission.
CHAPTER 7

SUMMARY

Electron transport in III-V semiconductors, especially the GaAs/AlGaAs material system, is studied in various nonequilibrium situations. Throughout the study, a Monte Carlo simulation method is used for the first-principle analysis of transport properties under the semiclassical Boltzmann transport picture. Physical models to include such complicated effects as coupled plasmon/phonon scattering, band filling, quantum mechanical transmission, and band bending are studied in detail. The spectral density function technique is also discussed in conjunction with the possible inclusion of collision broadening, a quantum mechanical effect, into a semiclassical Monte Carlo simulation.

The nonlinear transport study in the present work essentially consists of two aspects. The first topic is hot electron transport in GaAs, focusing on the electron impact ionization effects. The dependence of impact ionization rates on the details of the band structure is investigated by using two (local and nonlocal) empirical pseudopotential methods. The spatial evolution of the ionization rate and the average electron energy are studied in nonuniform fields characteristic of p-n junctions and Schottky barriers. An appropriate definition of the impact ionization rate in transient problems is discussed along with the Keldysh model. The effects of field fluctuations due to the random distribution of dopants are studied as well, and are found to have only limited importance on the enhancement of the impact ionization rate. The possibility of new GaAs electron-emitting diodes is explored numerically and compared with the corresponding Si devices.

The second aspect deals with the effects of conduction band discontinuities on the electron transport. In particular, one-dimensional heterostructures are modeled to study the nonlinear transport across heterointerfaces. First, two heterostructure avalanche photodiodes are studied.
Particular attention is paid to microscopic features such as the evolution of the energy and momentum distribution function in space and time, from which one can deduce macroscopic results. It is found that overheating, enhanced energy relaxation, and carrier confinement as a consequence of the structure in real space have a pronounced influence on the energy and momentum distribution. As a result, the energy distribution can have a structure which directly reveals the band structure (of the material). The dependence of the impact ionization rate on the band structures of neighboring layers is also addressed.

The effects of a nonequilibrium phonon distribution on the electron transport are studied. The phonon distribution can be perturbed whenever a significant number of phonons (especially zone-center phonons) are emitted on a time scale shorter than the phonon decay time constant. Such a situation can occur when a large number of carriers propagate across an interface, experience an abrupt energy gain, and subsequently relax through strong phonon emission. An algorithm is developed for the microscopic analysis of phonon dynamics. It is observed that the hot phonons change the scattering rate considerably and heat the electron energy distribution due to the reabsorption of nonequilibrium phonons.

To investigate the quasiballistic nature of electron motion and to examine the validity of hot electron spectroscopy, the tunneling hot electron transfer amplifier structures are studied at 4.2 K. Special attention is paid to the transport in the collector barrier, a region in which the tunneling hot electron transfer amplifier differs from the similar planar doped barrier transistor. Our model includes the effects of coupled plasmon/phonon interaction and the Pauli exclusion principle along with more conventional features. The numerical results demonstrate the existence of nearly ballistic transport in the base and the collector barrier, and confirm that the experiments can indeed measure the energy distribution of injected ballistic electrons. The device characteristics such as transfer ratio and transit time are also investigated in detail, and discussed along with the optimum operating conditions.
REFERENCES


28. See, for example, B. K. Ridley, Quantum Processes in Semiconductors (Clarendon Press, Oxford, 1982).


40. "...for example, P. J. Price, Superlatt. Microstruct. 2, 213 (1986)."


84. E. O. Kane, Phys. Rev. 159, 624 (1967).


87. J. M. Higman, private communication.


90. G. G. P. van Gorkom, private communication.


VITA

Ki Wook Kim He attended the Seoul National University, and graduated with a Bachelor of Science degree in Electronic Engineering in February 1983. He has been attending the University of Illinois at Urbana-Champaign since June 1983. He received a Master of Science degree in Electrical Engineering in May 1985. During 1986-1988, he was awarded an IBM predoctoral fellowship. He is currently a candidate for the degree of Doctor of Philosophy in Electrical Engineering from the University of Illinois at Urbana-Champaign.