FLUCTUATIONS DUE TO OBJECT DISCRETENESS

Thesis by
Wesley Boudville

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1988

(Submitted 3 May 1988)

ACKNOWLEDGMENTS

John Beahan

David Chow

Dr Reuben Collins

Dr Bob Hauenstein

Dr Steve Hetzler

Mike Jackson

Matthew Johnson

Dr Thomas C. McGill, my advisor

Richard Miles

Dr Arati Prabhaker

Ted Woodward

Dr George Wu

Dr Amikam Zur

I have been the recipient of the following fellowships:

Earle B Anthony Fellowship, 1983-4, from Caltech

Hackett Studentship, 1984-7, from the University of Western Australia

IBM Predoctoral Fellowship, 1987-8, from IBM Corporation

The US Office of Naval Research has sponsored part of my research.

Abstract

In this thesis I discuss the effect of randomness or fluctuations in topics drawn from three different areas of condensed matter physics: Metal-semiconductor ohmic contacts, the planar-doped barrier transistor and two dimensional continuum percolation.

Chapter 1 contains an introduction to the thesis. It outlines and summarises the rest of the thesis.

Chapters 2 and 3 describe the work done on metal-semiconductor ohmic contacts. The motivation was to find a possible limit to the fabrication of Very Large Scale Integrated circuits (VLSI). These chips are powered by current entering the chip along metal lines. A metal line ultimately makes ohmic contact to a semiconductor. This is done by doping the semiconductor heavily. In general, it is desirable for the contacts to have as low a resistance as possible, in order to reduce voltage drops across the contact. Furthermore, the conductances per unit area of such contacts should be the same, otherwise quite different currents could flow through two contacts of nominally the same specific conductance. In modelling such a contact, it has previously been assumed that in the depletion region of the semiconductor, at the junction, the dopants form an ionised continuum. However, dopants are discrete. Hence, over the small distance ($\sim 100\,\text{Å}$) of the depletion region of a heavily doped semiconductor, the current will see a spatially randomly varying potential. This causes the resistance of a contact to vary due to the random configuration of dopants in the contact.

In Chapter 2, a continuum model of the junction is presented. This is considered an improvement over that used in the literature for ohmic contacts. Chapter 3 describes how the discreteness of the dopants and their random distribution is taken into account. Simulations are made to find the resistance fluctuations of a contact, given the size of the contact and the doping in the semiconductor. I

investigate how the fluctuations scale as a function of the contact size and doping. Also considered is the effect of making the current trajectories three dimensional, instead of restricting them to being normal to the junction. It is found that there is little modification to the one dimensional nature of the trajectories. (For the purposes of obtaining actual numbers, the semiconductor was chosen to be n-type GaAs.) Simulations indicate that the dopant discreteness will not be a problem in ohmic contacts in VLSI. Rather, it will be a problem for Ultra Large Scale Integrated circuits (ULSI). The resistance fluctuations become significant for contact sizes on the order of 1000 Å. Currently, the semiconductor industry is just at the submicron level. It will probably be at least the mid 1990s before the device sizes approach 1000 Å.

Chapter 4 contains suggestions of future research based on the results of Chapters 2 and 3.

In Chapter 5 an analytic means of estimating the probability distribution of barrier heights in the barrier regions of a planar doped barrier transistor is derived. The barrier heights vary spatially due to the random distribution of discrete dopants in the barrier regions, much as in the metal-semiconductor contact of the earlier chapters. The resultant distribution is compared to one found from a three dimensional finite element simulation in the literature. The agreement is good. The analytic results presented go beyond this by showing the dependence of the barrier fluctuations on the doping and the thickness of the barrier regions.

Chapter 6 describes finite size effects in two dimensional anisotropic continuum percolation. Continuum percolation is where the objects that percolate are placed randomly in a given region, as distinct from percolation on a lattice. The anisotropy refers to the objects having a preferred average orientation. In two (and three) dimensions it has been found that the critical lengths for the onset of percolation are different for percolation along the average object orientation or

transverse to this. There is a universal behaviour to this dependence of critical lengths on the number of objects in the sample and on their degree of orientation. I have developed a theory to explain this. It agrees well with simulation results.

Parts of this thesis have been or will be published under the following titles:

Chapter 2:

Ohmic Contacts to n-type GaAs

WJB and T. C. McGill, *J. Vac. Sci. and Tech. B* **3**, 1192 (1985)

Chapter 3:

Resistance Fluctuations in Ohmic Contacts Due to Discreteness of Dopants

WJB and T. C. McGill, *Appl. Phys. Lett.* **48**, 791 (1986)

Scale Dependence of Resistance Fluctuations at Metal-Semiconductor Junctions

WJB and T. C. McGill, *J. Appl. Phys.* (*in press*)

Chapter 5:

An Efficient Calculation of Barrier Height Fluctuations in Planar-Doped Transistors

WJB and T. C. McGill, submitted to *J. Appl. Phys.*

Chapter 6:

Finite-Size Effects in Two Dimensional Continuum Percolation

WJB and T. C. McGill, to be submitted to *Phys. Rev. B*

Contents

List of Figures

Chapter 1

Thesis Outline

1.1 Introduction

This chapter is an outline of the thesis. I present here a summary of research done in topics drawn from three areas of condensed matter physics. These areas are: The metal-semiconductor ohmic contact; the planar-doped barrier transistor; and two dimensional continuum percolation. One unifying theme exists in this work. Each research topic discusses the effects of randomness in the distribution of objects, on some physical property of a structure containing these objects. For metal-semiconductor contacts, the objects are dopant atoms in the semiconductor, and the property investigated is the zero bias resistance of the contact. In the case of the planar-doped barrier transistor, the objects are also dopants. The dopants are located randomly in the barrier regions of the transistor. The resultant property affected is the barrier height. A distribution of barrier heights is produced, rather than a single barrier height. For two dimensional continuum percolation, the objects are widthless one dimensional sticks, placed randomly in a two dimensional region. The percolation properties of collections of these objects are studied.

In the sections that follow, I will attempt to motivate and summarise the thesis.

1.2 Summary of Chapter 2

Chapters 2 and 3 describe the work done on metal-semiconductor ohmic contacts. The semiconductor is taken to be GaAs. The impetus for this work is to find a possible limit to the fabrication of Very Large Scale Integrated (VLSI) chips. These chips are powered by current entering the chip along metal lines. A metal line eventually makes contact to a semiconductor.[1] The energy diagram of such a contact is shown in Fig. 1.1. The diagram is given for the case of the semiconductor being doped n-type. The contact shown has an applied bias of V volts. Also shown are the conduction band and valence band in the semi-

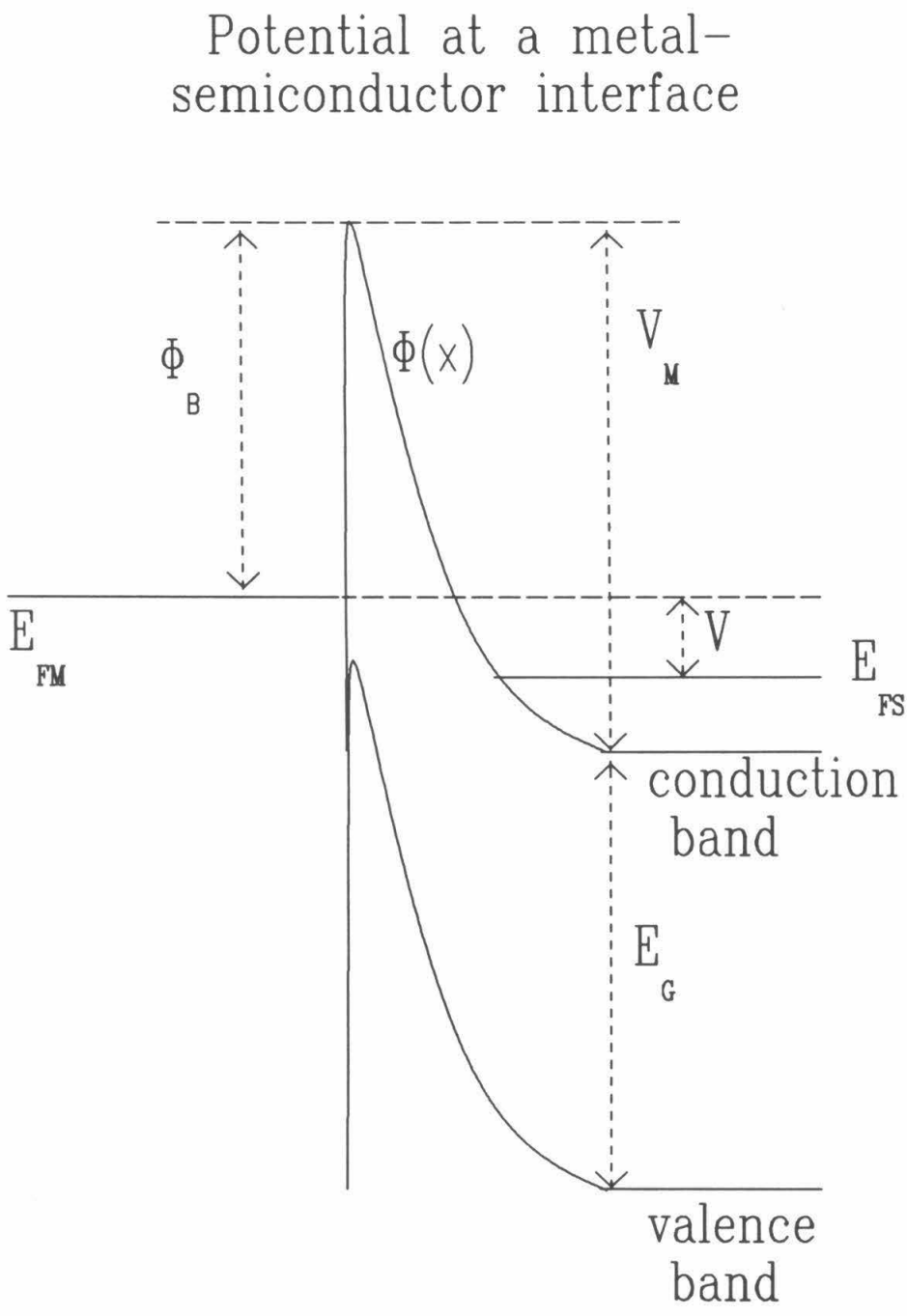

Figure 1.1: Diagram of a metal–n–semiconductor junction for an applied bias V. E_g is the band gap. Φ_B is the barrier height. E_{FM} and E_{FS} are the metal and semiconductor Fermi levels.

conductor. Usually, when such a contact is made, there exists a potential barrier to the flow of current (i.e., electrons) across the contact. Hence, the electrons have to either have sufficient energy to cross over the maximum energy in the barrier (thermionic emission) or, if not, they must tunnel through the barrier (field emission or thermionic-field emission).[2] At room temperature, the probability for thermionic emission is small. The alternative mechanism for current flow is tunnelling. Because of the well-known rapid attenuation of tunnelling probability with increasing distance through which to tunnel,[3] it is necessary to reduce the width of the barrier region in order to obtain a significant probability for tunnelling. This can be done by doping the semiconductor heavily, with dopings of $\sim 10^{19}\, cm^{-3}$ or more. Such large dopings reduce the width of the barrier to $\sim 100\,\text{Å}$ or less. In VLSI technology, this method is used to make an 'ohmic contact' of the metal to the semiconductor. The term 'ohmic contact' means that if the current density-voltage (J-V) relationship of the contact is measured, one would find a straight line through the origin, as in Ohm's law. The reciprocal of the J-V slope at the origin is called the specific resistance or the zero-bias resistance, R_c, measured in $\Omega - cm^2$. It may be used to characterise the contact.

In general, it is desirable for R_c to be as small as possible. This would mean that for a given current coming from the metal power lines, the resultant voltage drop across the junction would be minimised. This is good, in that one would like voltage drops to occur across active devices in the chip, rather than across the passive ohmic contacts. (In passing, note that reducing R_c also reduces the heating losses at the junction.) Furthermore, on a chip several ohmic contacts are made. For different contacts, the R_c should be the same otherwise either the voltage drops will vary across different contacts with the current being the same, or different amounts of current will flow through different contacts. Both possibilities could result in bad device performance if the chip has been designed assuming

same nominal characteristics for the ohmic contacts. In modelling such contacts, it has previously been assumed that in the depletion region of the semiconductor at the junction, the dopants form an ionised continuum. However, dopants are discrete. Hence, over the small distance ($\sim$ 100 Å) of the depletion region of a heavily doped semiconductor, electrons carrying the current will see a randomly varying potential in the depletion region. This arises from the random distribution of the discrete dopants, rather than from an ionised continuum of dopants. The resultant R_c of the contact will depend on the particular configuration of dopants in the contact. Since different contacts will have different random configurations of dopants, R_c will vary from contact to contact.

In attempting to understand this, a continuum model of the junction is first studied. This is done in Chapter 2. The GaAs is heavily doped so that tunnelling dominates the current. The WKB approximation is used to give the transmission probability for tunnelling by an electron of energy E,

$$T(E) = \begin{cases} exp\left(-2i \int_{x_1}^{x_2} k(x,E)\,dx\right), & 0 \leq E \leq qV_m, \\ 1, & E > qV_m, \end{cases}$$

where V_m is the height of the potential barrier, and x_1, x_2 are the turning points of the forbidden region for the wavevector k, and the x direction is normal to the interface. In the depletion region of the semiconductor, the tunnelling through the barrier is characterised by an imaginary wave vector versus energy relationship. Different choices may be made for the dependence of k on the energy. In this work, k is taken to be given by the two-band model from $\mathbf{k} \cdot \mathbf{p}$ theory.[4] This gives the dependence of the tunnelling on both the conduction and valence bands. Previously, the most common theory of ohmic contacts was due to Chang, Fang and Sze.[5] They used an alternative parabolic (one band) relationship for the energy versus wavevector to find the transmission by tunnelling through the depletion region. The two band model is considered to give a more accurate representation

of the tunnelling. Included in the calculation is the image charge in the metal of the tunnelling electron. Also, for heavily doped n-GaAs, the Fermi level in the semiconductor lies in the conduction band. This is known as degenerate doping. It produces a potential energy versus distance relationship in the semiconductor that is non-parabolic. The effect of this on the transmission probability is taken into account. Furthermore, as the doping increases, the L point states start being occupied. Then, to find the semiconductor Fermi level, the filling of both the Γ and the L states is considered.

Results are found for three different barrier heights of the n-GaAs on the metal: 0.6, 0.8 and $1.0\,eV$. The results are presented in the form of zero bias resistance R_c versus the doping in the semiconductor. These are compared to the published results of Chang, Fang and Sze. It is found that the present treatment predicts up to an order of magnitude less R_c at dopings $\geq 10^{19} cm^{-3}$. The difference is mainly due to the improved treatment of the transmission probability through the use of the two band model.

1.3 Summary of Chapter 3

Having now found a more accurate model of the zero bias resistance as a function of doping, a question then arises about a key assumption that is made. Namely, within the depletion region, the dopants are considered to form an ionised continuum. In Chapter 3, I investigate what happens when this assumption is relaxed. The method used to do this is simple. Consider a square metal-n-GaAs contact, with a side of length l. The GaAs is heavily doped, with a specified doping n. The contact is divided into a grid of points. The zero bias resistance is found at each point in this grid, assuming that the tunnelling current flows through the depletion region and normally to the contact. Simulations are performed where

dopants are placed randomly at lattice sites in a face-centred-cubic sublattice of the GaAs zincblende lattice, in a region centred on the contact. The probability of a dopant occupying a lattice site is given by the doping.

Let me next describe how the zero bias resistance is found along a particular trajectory through the depletion region that is normal to the metal-semiconductor junction. The potential energy of the tunnelling electron needs to be calculated. The contributions to this potential are taken to come from three sources. First, and simplest to find, is that due to the image in the metal of the tunnelling electron. Next, a cylinder is constructed in the depletion region, whose axis is the trajectory and with a radius R. The length of the cylinder is x_d, the depletion length. The relationship $R = 2x_d$ is chosen. The second contribution to the potential comes from any discrete dopants in the simulation that are found inside this cylinder. (Of course, the effect of the dopants' images in the metal are also considered.) Finally, the contribution from dopants in the unbounded region outside the cylinder, and in the depletion region, is added to the potential. This last contribution is taken to be from an ionised continuum of dopants. By elementary electrostatics, it is possible to solve analytically for this continuum dopant contribution. Hence, the effect of randomness enters only from dopants found inside the cylinder. Naturally, the question arises as to the choice of R made above, and to the adequacy of this in fully accounting for the effects of dopant randomness. In principle, given a particular set of dopants simulated within a depletion region, one should make such a cylinder of construction as large as possible in order to include more dopants and thus to more correctly handle the effect of randomness. However, to keep the calculation computationally feasible, it is desirable to make R as small as possible. It is shown from a simple statistical argument that the choice of $R = 2x_d$ is, for all practical purposes, equivalent to choosing an infinitely wide cylinder. This choice also enables computations to be made in reasonable time. Having now

found the potential energy along the trajectory, this is then input into the WKB approximation and the two band model to find the transmission as is done in Chapter 2. Hence, the zero bias resistance for the trajectory may be calculated. This method is repeated at all points in the grid that makes up the contact. Given the resultant set of resistances, the reciprocals are averaged to find the characteristic conductance, or resistance, of the entire contact.

Naturally, the average R_c of a contact depends on the configuration of dopants in the semiconductor. It is then possible, at each doping and contact size, to perform simulations to find the amount of variations in R_c. For example, at a contact size of $l = 200\,\text{Å}$, up to an order of magnitude variation can be found in R_c at dopings $> 10^{19}\,cm^{-3}$. The variation in R_c is studied through simulations to see how it scales with doping and contact size. It is compared to a theory developed in this chapter. Good agreement is found. Also, it is possible to extend the calculations by considering current trajectories through the depletion region that are not simply normal to the contact. Little modification is found in the current paths. Most current trajectories are seen to be indeed normal to the contact.

The simulations indicate that fluctuations in resistance arising from dopant discreteness should not be a problem in ohmic contacts as devices get smaller until contact sizes reach the 0.1 micron scale. Thus, this issue would probably be of concern only to Ultra Large Scale Integration.

1.4 Summary of Chapter 4

In Chapter 4, I offer suggestions for future research to extend the work in Chapter 3. A major assumption made in Chapters 2 and 3 was the use of the WKB approximation in finding the tunnelling transmission probability through the depletion region. It is found that the potential in the depletion region due to

the random placement of dopants varies over small length scales of $\sim 10\,\text{Å}$. This suggests that the WKB approximation only provides semiquantitative accuracy in finding the transmission. How can this be improved? The alternative is a full quantum mechanical treatment of the time dependent Schrodinger equation in the depletion region,

$$i\hbar \frac{\partial \psi}{\partial t} = H\psi = \left[-\frac{\hbar^2}{2m^*}\nabla^2 + \phi\right]\psi. \tag{1.1}$$

Here ϕ is the potential, m^* is the effective mass of an electron in the current and ψ is the wavefunction of the electron. Unfortunately, a ϕ arising from a particular configuration of dopants will vary randomly compared to a ϕ from another configuration. There do not exist spatial symmetries in ϕ that enable the simplification of Eq. 1.1, or its time independent version, through group theoretic methods.

The only alternative is a direct time stepping solution of Eq. 1.1. However, this would need to be done in three spatial dimensions, necessitating large scale computations. A recent proposal by D. Kosloff and R. Kosloff may make this practical and accurate.[6] The time derivative in Eq. 1.1 is replaced by a finite difference. This allows time stepping. The problem is in finding $\nabla^2\psi$ at each time step. If the fast Fourier transform (FFT) of ψ is found, to transform ψ to wavevector or $\mathbf{k}$ space, then $\nabla^2\psi(\mathbf{x})$ transforms to $-k^2\psi(\mathbf{k})$ in this space. This is just a multiplication at all points in $\mathbf{k}$ space. Then, an inverse FFT is made on the result to move back to real space, where the result is added to $\phi(\mathbf{x})\psi(\mathbf{x})$. Due to the efficiency of the FFT, this was shown by Kosloff and Kosloff to involve less computation that a finite differencing approach to $\nabla^2\psi$, assuming comparable accuracy.

Thus, the above technique may to used to verify the semiclassical results of Chapters 2 and 3, by finding more accurate transmission probabilities. Also, this could enable the study of time dependent phenomena in the passage of current through the depletion region. Eq. 1.1 is written for a single free carrier in the

current. By extending Eq. 1.1 to handle several particles in the current, it should be possible to study time dependent phenomena such as the noise in an ohmic contact.

1.5 Summary of Chapter 5

Chapter 5 contains the second topic of the thesis. It examines the variation in barrier heights in the barrier regions of a planar doped barrier transistor. The barrier heights vary due to the fluctuations in the potential in the barrier regions caused by the dopant atoms being discrete and randomly located. This is similar to the case of the ohmic contact discussed above. Recently, Arnold and Hess[7] performed a three dimensional simulation whereby dopants were placed randomly in a barrier region. The resultant potential was found from solving Poisson's equation by a finite element method. From this, they obtained a distribution of barrier heights. Computationally, this was rather involved, necessitating as it did the use of a Cray computer to solve for the potential. As an alternative to this, I derive analytically a probability distribution for the barrier heights, using the same assumptions as Arnold and Hess. This distribution is compared to that found by them for a particular simulation. The agreement is good, while my method is numerically trivial to evaluate. Furthermore, the analytic model is more general. It shows the dependence of the barrier distribution on the doping and the thickness of the barrier regions.

1.6 Summary of Chapter 6

Chapter 6 discusses the third topic of this thesis. It describes finite size effects in two dimensional anisotropic continuum percolation. Most of the published work on percolation concerns percolation on a lattice. Distinct from this, however, is

continuum percolation where the objects that percolate can lie anywhere in a given region. The anisotropy refers to the objects having an orientation that is random, with the orientation coming from a probability distribution that gives a preferred average orientation. The objects studied here are widthless one dimensional line segments (sticks) located in a unit square in two dimensions. Recently,[8,9] work has been done on finding the critical lengths for the onset of percolation along and transverse to the average stick orientation as a function of the number of sticks in a sample, and of the degree of alignment of the sticks. The critical lengths were found through simulations.

When there are a finite number of sticks in a sample, it is found that the longitudinal and tranverse critical lengths are different. From simulations, the probability distributions for the lengths and the orientations of the sticks can be varied, and there is found the following universal behaviour: Considered as functions of the number of sticks and the degree of alignment of the sticks, the longitudinal and transverse critical lengths are (to good approximation) largely independent of the distributions. In the published results of Refs. 8 and 9, this was seen. However, neither in these or other references to date has there been an attempt at explaining this. Chapter 6 contains both simulations and a derivation of a simple theory that explains the universal behaviour of the critical lengths. The simulations are more extensive than those in Ref. 8. Good agreement is found between the theory and simulations.

****

The interested reader should refer to the appropriate chapters for a more detailed discussion of the above topics.

References

1. P. B. Ghate, *Handbook of Advanced Semiconductor Technology*, ch. 6 (Van Nostrand, New York, 1988).
2. R. Stratton, *Tunnelling Phenomena in Solids*, ch. 8 (Plenum Press, New York, 1969).
3. L. I. Schiff, *Quantum Mechanics* (McGraw-Hill, Tokyo, 1968).
4. E. O. Kane, *Physics of III-V Compounds* (Academic Press, New York, 1966).
5. C. Y. Chang, Y. K. Fang and S. M. Sze, Solid State Elect. **14**, 541 (1971).
6. D. Kosloff and R. Kosloff, J. Comp. Phys. **52**, 35 (1983).
7. D. Arnold and K. Hess, J. Appl. Phys. **61**, 5178 (1987).
8. I. Balberg and N. Binenbaum, Phys. Rev. B **28**, 3799 (1983).
9. I. Balberg, Phys. Rev. B **31**, 4053 (1985).

Chapter 2

Ohmic Contacts to n–Type GaAs

2.1 Introduction

The development of GaAs technology has led to a continuing search for methods of making ohmic contacts to n-type GaAs. The main obstacle to this has been the barrier height of the metal on the n-GaAs. Typically, one has $\phi_B \sim 0.8eV$ in the band diagram of Fig. 1.1. This provides a large barrier to tunnelling by the electrons in the current. Given that such a barrier height exists, the tunnelling probability can be increased by reducing the distance through which the electrons tunnel. This can be done by increasing the doping in the GaAs. Hence, one of the very interesting recent developments in ohmic contacts to n-GaAs has been the growth of heavily doped layers using Molecular Beam Epitaxy.[1] Furthermore, the most widely used theory of ohmic contacts is that due to Chang, Fang and Sze (CFS),[2] in which a metal-semiconductor contact is treated. Attempts to compare theory with experiment have been mainly based on this theory.[3,4,5]

In CFS's model, it is assumed that the tunnelling through the Schottky barrier is characterised by an imaginary wavevector versus energy that is a simple parabolic (one–band) relationship. In this chapter, the results are presented of a more realistic treatment of the contact resistance for electrons tunnelling through the barrier. The biggest correction is due to the improvement in the model of the imaginary wavevector versus energy relation. A two–band model is assumed that uses the correct energy gap and conduction band effective mass for GaAs. The theory also includes the effects of image charge, negative charge at the interface and the presence of the L–point minimum in the GaAs. Results are found for the contact resistance which differ by as much as an order of magnitude from those of CFS in the doping ranges of interest for device use.

This chapter is organised in the following way. Section 2.2 presents the theoretical model. In Section 2.3, the results are given for the model, and these are compared with CFS. In Section 2.4, a brief mention is made about fluctuations in

the barrier potential due to the discrete nature of the doping. This is a prelude to a more extended discussion in the next chapter. Finally, in Section 2.5, the conclusions of the chapter are presented.

2.2 Theory

For a metal–n–semiconductor junction in reverse bias, with the semiconductor being degenerate, the energy diagram is shown in Fig. 1.1 of the previous chapter.

The corrections due to the image force,[6] negative charge at the interface[7] and non–parabolic corrections to the potential due to the presence of conduction electrons[8] in part of the depletion region are all included. The metal is modelled as a degenerate electron gas, with the radius of the Fermi sphere (the Fermi energy) being $\sim$ 7–8 eV. (This is not shown in Fig. 1.1, as it would be offscale.)

In Fig. 1.1, the energy is measured upwards from the conduction band edge far from the depletion region; E_{FS}, E_{FM} are the Fermi levels in the semiconductor and metal, respectively, and V is the applied bias. Then, following Ref. 9, the current density is given by

$$J = \frac{m^* k_B T q}{2\pi^2 \hbar^3} \int_0^\infty dE\, T(E)\, ln\left(\frac{exp[-(E - E_{FS})/k_B T] + 1}{exp[-(E - E_{FM})/k_B T] + 1}\right), \qquad (2.1)$$

where E is the energy of electron normal to interface, m^* is the conduction band effective mass at the Γ point, and $T(E)$ is the transmission probability. Eq. 2.1 neglects phonon–assisted processes, i.e., the transverse wavevector k is assumed to be conserved for a tunnelling particle. The large Fermi radius of the metal in reciprocal space, relative to the distribution of free carriers in reciprocal space for doped GaAs, or indeed most semiconductors, makes this a good approximation.[10] In the derivation of Eq. 2.1, a number of other approximations have been made. These approximations are detailed in Ref. 11. The approximations are very commonly made in treating a Schottky barrier, and are not likely to result in large

errors. One major assumption is used to guarantee that zero current will result when zero voltage is applied, by assuming that the boundary conditions on the current from the semiconductor to the metal are appropriate for the current in the other direction.

By the WKB approximation, the transmission is given as

$$T(E) = \begin{cases} exp\left(-2i\int_{x_1}^{x_2} k(x)\,dx\right), & 0 \le E \le qV_m, \\ 1, & E > qV_m. \end{cases} \tag{2.2}$$

where x_1, x_2 are the turning points of the forbidden region for **k**. Assuming a direct–gap semiconductor, a two–band model, from $\mathbf{k}\cdot\mathbf{p}$ theory, gives, for k everywhere within the forbidden region,[12]

$$k = -i\left[\frac{2m^*}{\hbar^2 E_g}\Big(\phi(x) - E\Big)\Big(E_g + E - \phi(x)\Big)\right]^{1/2}. \tag{2.3}$$

Let

$$\frac{1}{R_c} \equiv \frac{\partial J}{\partial V} \equiv \int_0^\infty dE\, G(E). \tag{2.4}$$

R_c is the specific contact resistance of the interface, measured in $\Omega\, cm^2$, and $G(E)$ is the conductance distribution function. Then, at zero bias one finds

$$\frac{1}{R_c} = \frac{m^* q^2}{2\pi^2\hbar^3}\int_0^\infty dE\,\frac{T(E)}{exp[(E - E_{FS})/k_B T] + 1}. \tag{2.5}$$

R_c and $G(E)$ are found from Eq. 2.5 by numerical methods. To obtain E_{FS} in Eq. 2.5, the free carrier concentration n is compared with N_C, the effective density of states in the conduction band. It is assumed that the temperature is high enough to neglect carrier freezeout. Then

$$E_{FS} = \begin{cases} k_B T\, ln(n/N_C), & n \le N_C, \\ \dfrac{\hbar^2}{2m^*}(3\pi^2 n)^{2/3}, & n > N_C. \end{cases} \tag{2.6}$$

For carrier concentrations in the GaAs sufficiently large such that the L–point states are occupied, the semiconductor Fermi level is given by

$$n = \frac{1}{3\pi^2}\left[\left(\frac{2m^*}{\hbar^2}\right)^{3/2} E_{FS}^{3/2} + 4\left(\frac{2m_L}{\hbar^2}\right)^{3/2}(E_{FS} - E_{\Gamma L})^{3/2}\right] \tag{2.7}$$

where m_L is the effective mass at the L–point, and $E_{\Gamma L}$ is the energy separation between the L–point and the Γ–point. The factor of 4 in the second term on the right hand side of Eq. 2.7 arises because each Brillouin zone has a net of four L–points.

The potential energy of the electron is given by

$$\phi(x) = -\frac{q^2}{16\pi\epsilon x} + \phi_1(x) \tag{2.8}$$

where, for degenerate doping, ϕ_1 satisfies the implicit relationships

$$\frac{d^2\phi_1}{dx^2} = \frac{q^2}{\epsilon}\Big[n - n_e(x)\Big] \tag{2.9}$$

$$n_e(x) = \begin{cases} \frac{1}{3\pi^2}\left(\left[\left(\frac{2m^*}{\hbar^2}\right)(E_{FS} - \phi_1)\right]^{3/2} + \right. \\ \left. 4\left[\left(\frac{2m_L}{\hbar^2}\right)(E_{FS} - \phi_1 - E_{\Gamma L})\right]^{3/2}\right), & E_{FS} - \phi_1 \geq E_{\Gamma L}, \\ \frac{1}{3\pi^2}\left[\left(\frac{2m^*}{\hbar^2}\right)(E_{FS} - \phi_1)\right]^{3/2}, & E_{FS} - \phi_1 < E_{\Gamma L}. \end{cases} \tag{2.10}$$

For nondegenerate doping, where $E_{FS} < 0$, let $n_e \equiv 0$. Eqs. 2.9 and 2.10 are solved for ϕ_1 by the method of successive integration.[13] In doing this, the boundary condition of $\phi_1(x) = 0$ at $x = 1.2\times$ the depletion length is chosen to ensure convergence. The complete solution for ϕ_1 would tend to 0 asymptotically for $x \to \infty$. Thus the boundary condition lowers ϕ_1 near the edge of the depletion region. However, this occurs at large tunnelling distances, relative to the tunnelling distances at higher energies, and so will have little effect on the results.

The effect of negatively-charged surface states can be included by adding the following term to the right hand side of Eq. 2.8:

$$\phi_{surf}(x) = -\frac{dq^2N}{\epsilon} exp(-x/d), \tag{2.11}$$

where N is the area density of surface states and d is the penetration length of the states.

2.3 Results and discussion

The following values for the parameters in the theory have been used: $T = 300K$, $E_G = 1.42eV$, $E_{\Gamma L} = 0.284eV$, $m^* = 0.063$, $m_L = 0.55$, $\epsilon = 12.85$ and $N_C = 4.21 \times 10^{17} cm^{-3}$. The value of E_G is taken from Ref. 14, while the other values are taken from Ref. 15.

To illustrate the contributions to the conductance per unit area per unit energy, for various doping concentrations, Fig. 2.1 plots the barrier shapes and values of $G(E)$ for four different dopings: 10^{17}, 10^{18}, 10^{19} and $10^{20} cm^{-3}$. These are at $\phi_{B0} = 0.8eV$, $T = 300K$, and $V = 0\,volt$. Negative surface states are included, with the choice of $N = 5 \times 10^{14} cm^{-2}$, $d = 5\,\text{Å}$. $G(E)$ is plotted on a log scale due to its large variation when the energy is varied from 0 to $1eV$. For example, at a doping of $10^{19} cm^{-3}$, $G(E)$ ranges over ten orders of magnitude. Notice that $G(E)$ has a discontinuity in its slope at $E = qV_m$. This is due to Eq. 2.2, where the transmission below the barrier is given by the WKB approximation, while the transmission above the barrier was taken to be 1.

For a doping of $10^{17} cm^{-3}$, the maximum in $G(E)$ occurs at qV_m, and the total conductance has roughly equal contributions from carriers going over the barrier in thermionic emission and Fowler-Nordheim tunnelling.[16] Observe that the tunnelling contribution is significant only for energies down to $\sim 0.2\,eV$ below qV_m, since the tunnelling length increases strongly for decreasing energy. There is no direct tunnelling, since $E_{FM} = E_{FS} < 0$. At a doping of $10^{18} cm^{-3}$, the maximum in $G(E)$ increases, and it occurs at $E = 0.42\,eV$, which is below qV_m. Due to the narrowing of the barrier, significant conductance is found over most of the tunnelling energies. Fowler-Nordheim tunnelling dominates the conductance, while some direct tunnelling occurs, for $E < E_{FS} = 0.06\,eV$. When the doping is raised to $10^{19} cm^{-3}$, $G(E)$ increases strongly, with its maximum occurring at $E = 0.27\,eV \simeq E_{FS}$. Now direct tunnelling is comparable to Fowler- Nordheim

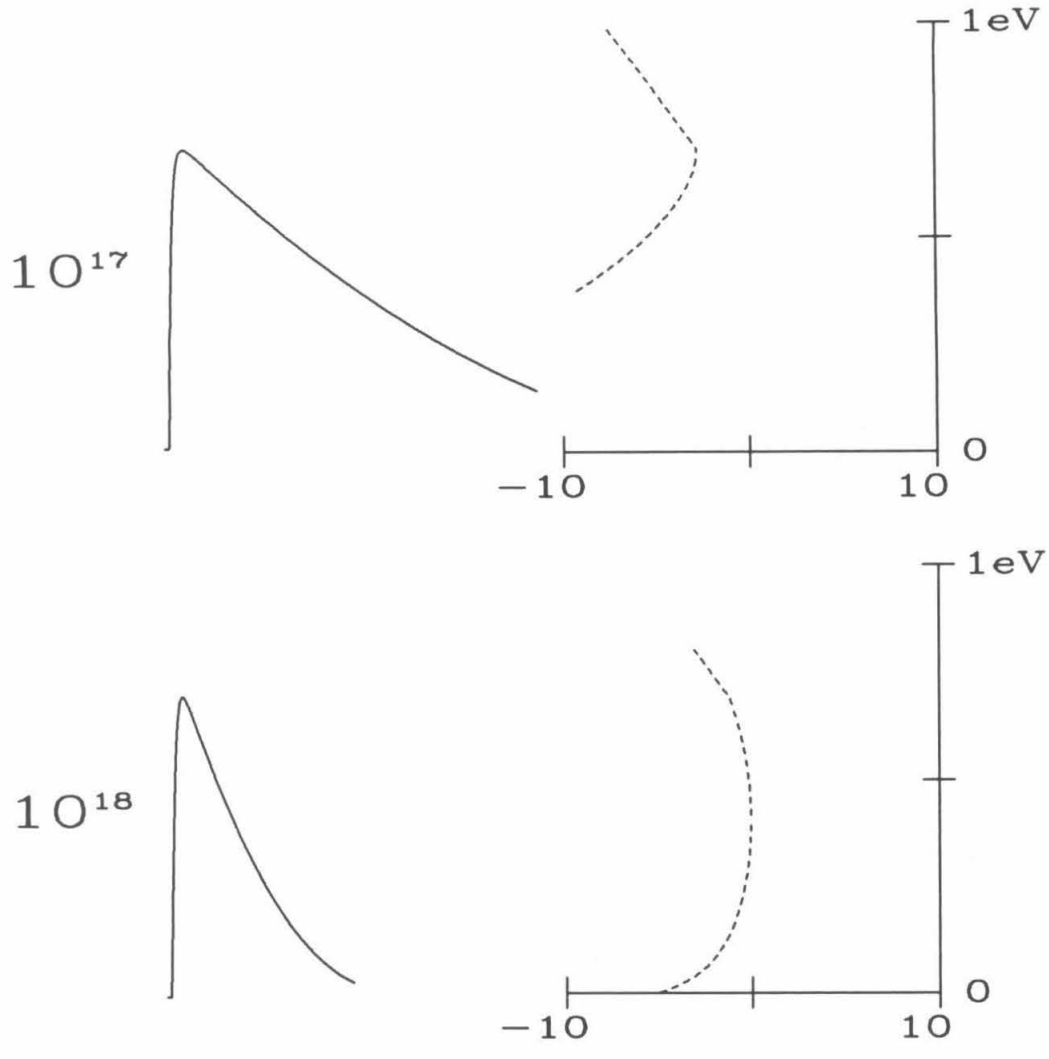

Figure 2.1: The distributions of the specific contact conductance $G(E)$ as a function of energy. To illustrate the energy position on the barrier, the potential energy normal to the junction is plotted using the same energy scale on the vertical axis. The log of the conductance is shown, with the conductance being measured in $\Omega^{-1}m^{-2}J^{-1}$. Results at dopings of 10^{17} and $10^{18}cm^{-3}$ are shown.

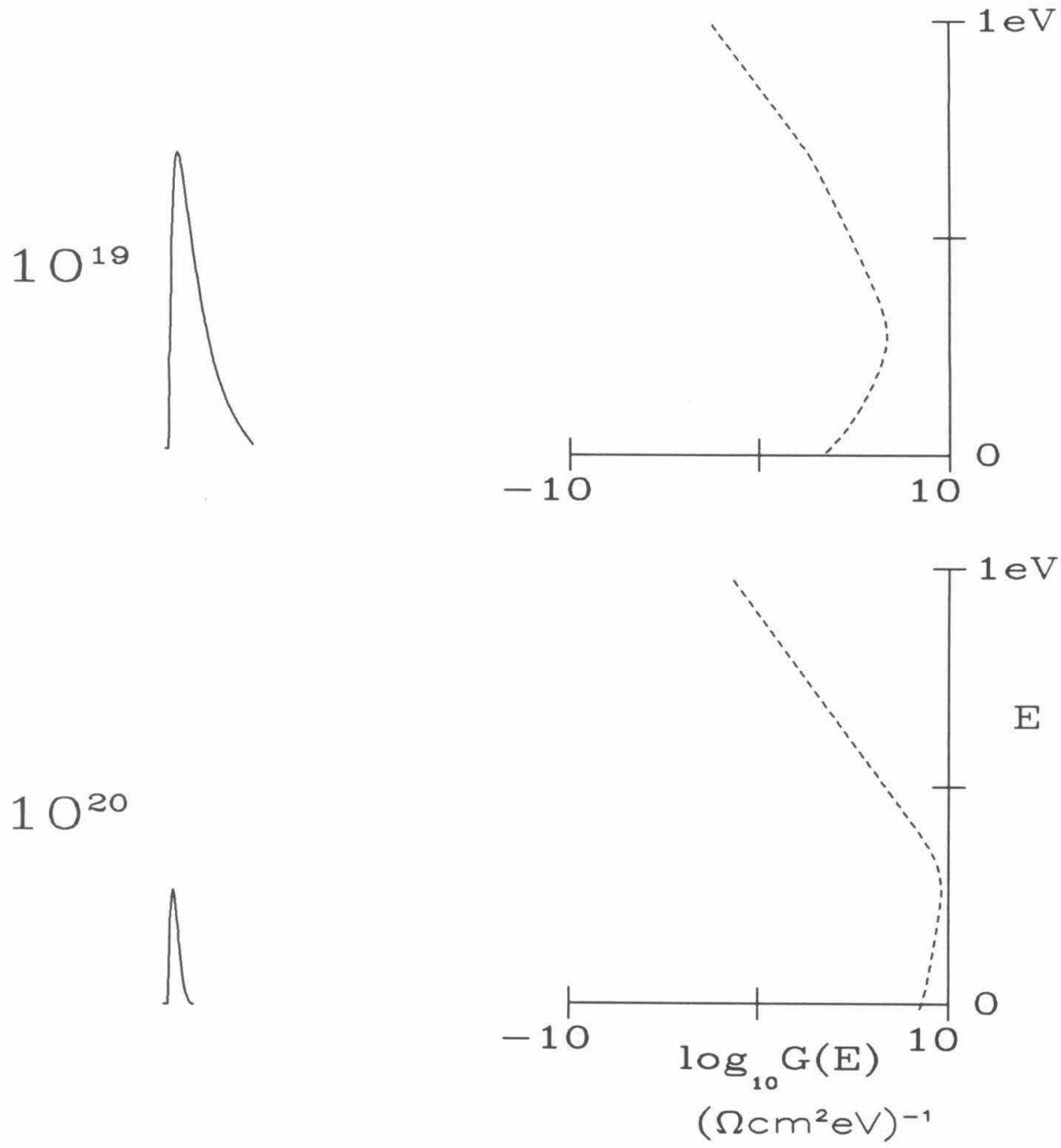

Figure 2.1: (Cont.) Results at dopings of 10^{19} and $10^{20} cm^{-3}$ are shown.

tunnelling. Finally, for a doping of $10^{20} cm^{-3}$, V_m falls considerably, after being approximately constant at lower dopings. This is due to the negative surface states 'annulling' most of the thin barrier. Here $E_{FS} = 0.33\, eV$, and so direct tunnelling is seen to dominate the conductance.

The specific contact resistance as a function of doping is presented for three different intrinsic barrier heights: 0.6, 0.8 and $1.0\, eV$, in Figs. 2.2, 2.3 and 2.4, respectively. The results include all of the effects described above. For comparison the results have been plotted assuming no negative surface charge, and also those from CFS. In Fig. 2.3, where $\phi_{B0} = 0.8\, eV$, the parabolic result is also shown. This occurs if in Eq. 2.10, n_e is set equal to 0, regardless of whether or not the semiconductor is degenerately doped.

Comparison of the results in these figures allows a number of conclusions to be drawn. First are described those which are common to both CFS's model and the model of this chapter. The contact resistance is seen to increase as the barrier height increases, at fixed doping. For example, at a doping of $10^{19} cm^{-3}$, in going from $\phi_{B0} = 0.6$ to $1.0\, eV$, the contact resistance increases by about two orders of magnitude in both models. For dopings between 10^{18} and $10^{19} cm^{-3}$, the contact resistance falls steeply, by about seven orders of magnitude. Recall from Fig. 2.1 and the discussion above, that over this doping range, for $\phi_{B0} = 0.8\, eV$, tunnelling increased to dominate the total conductance, and $G(E)$ increased strongly over all the tunnelling energies. Similar results are seen in this model for $\phi_{B0} = 0.6$ and $1.0\, eV$. Thus the doping range from 10^{18} to $10^{19} cm^{-3}$ may be considered as a transition from a Schottky barrier to an ohmic contact. As a final point of similarity, note that CFS's results and the nonparabolic curves (with no surface charge) tend to merge for dopings near $10^{18} cm^{-3}$. This is to be expected, since the two models differ in their expressions for the tunnelling transmission, and hence should yield the same results when the tunnelling contribution is small.

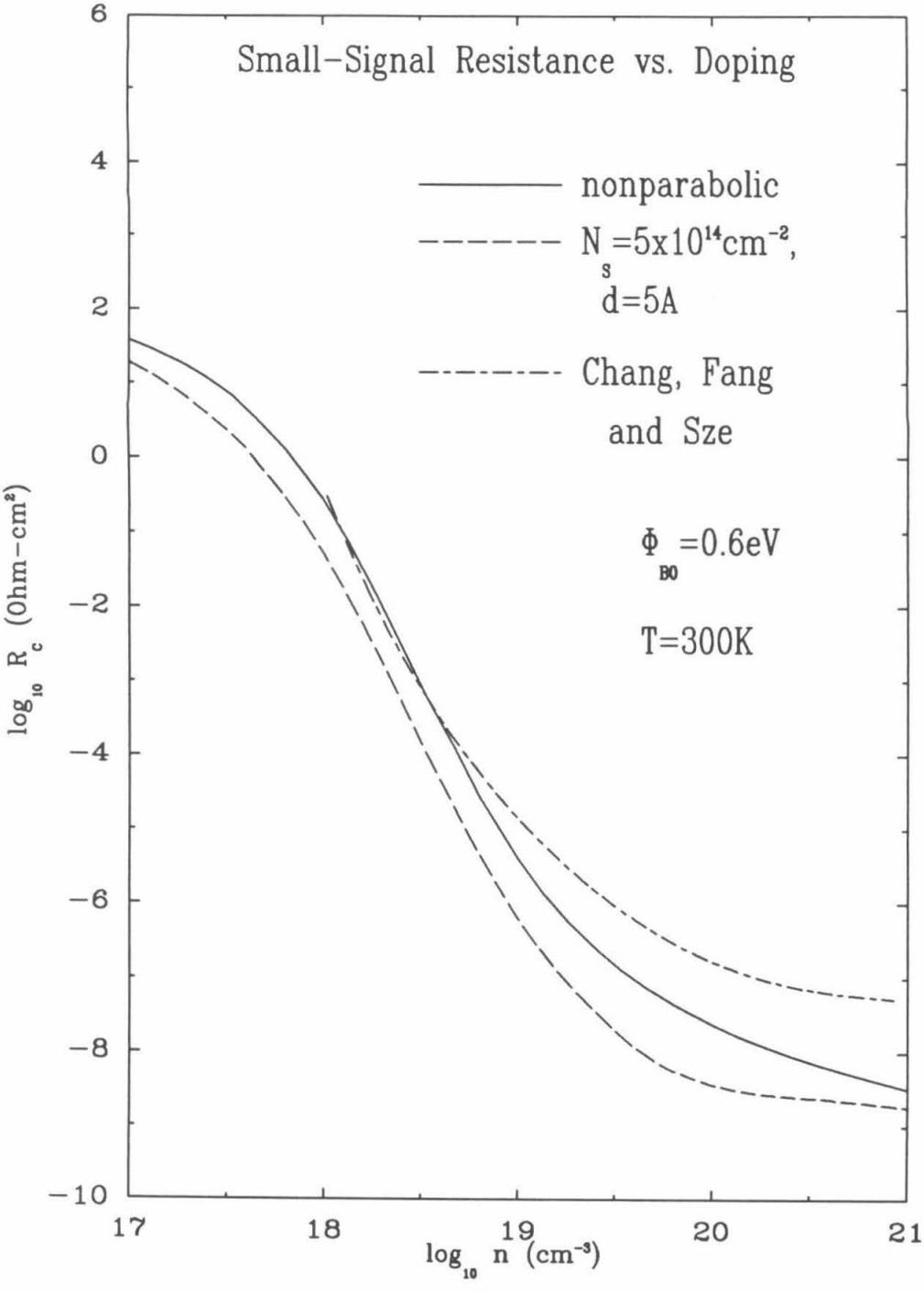

Figure 2.2: The specific contact resistance vs. doping, for $\phi_B = 0.6\,eV$ and $V = 0\,volt$, with and without negative surface charge, with the results from CFS.

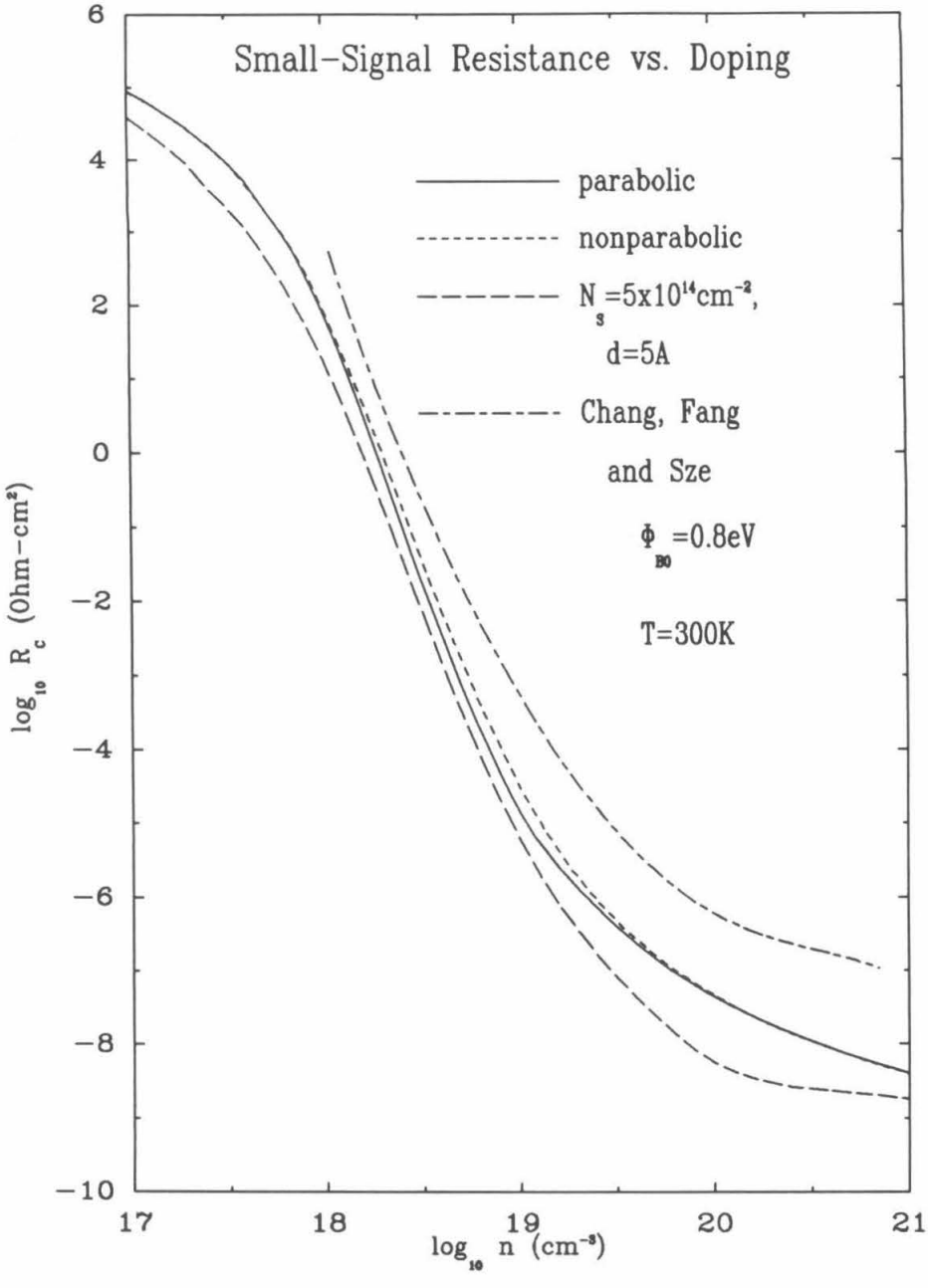

Figure 2.3: The specific contact resistance vs. doping, for $\phi_B = 0.8\,eV$ and $V = 0\,volt$, with and without negative surface charge, with the results from CFS. For comparison, the results assuming a parabolic conduction band in the depletion region are shown.

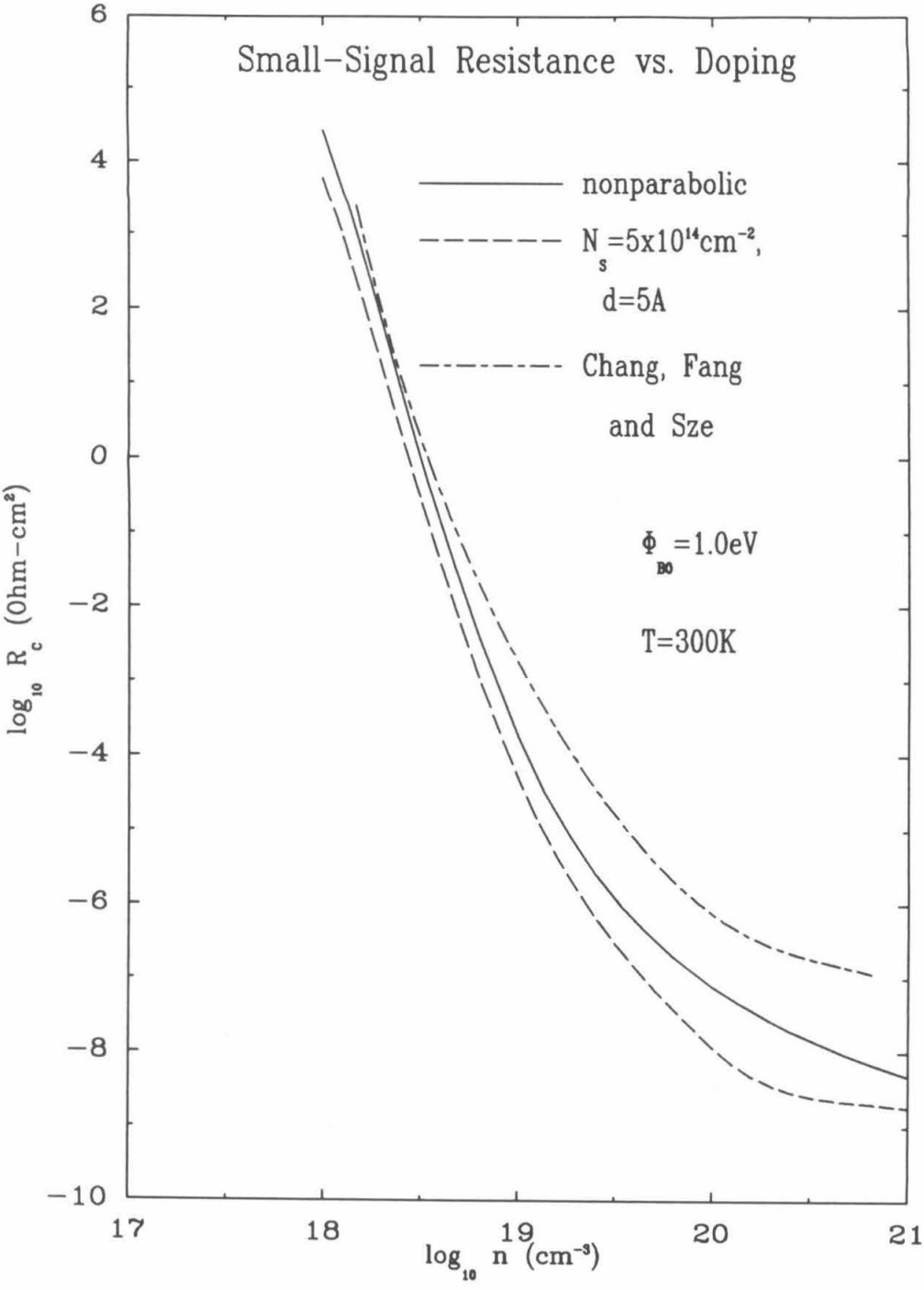

Figure 2.4: The specific contact resistance vs. doping, for $\phi_B = 1.0\,eV$ and $V = 0\,volt$, with and without negative surface charge, with the results from CFS.

Next, consider the differences between the results of the two models. The nonparabolic curves tend to lie below CFS's curves for dopings greater than $10^{18}cm^{-3}$. Between dopings of 10^{18} and $10^{19}cm^{-3}$, the nonparabolic curves fall with steeper slope, and continue doing so up to $10^{21}cm^{-3}$, whereas by $10^{20}cm^{-3}$, the slopes on CFS's curves are much reduced. Thus the gap between the nonparabolic and CFS's curves increases with increasing doping. That the differences between the curves increases with doping is expected, since the tunnelling contribution dominates the conductance, and the change in the function relating the imaginary part of the wavevector to the energy plays an important role, and accounts for most of the changes. Thus, at high dopings ($> 10^{19}cm^{-3}$) it may be possible to obtain metal-semiconductor junctions with substantially lower contact resistance that previously thought feasible.

To illustrate the relative importance of the various factors in the theory, consider Fig. 2.3. For dopings below $10^{18}cm^{-3}$ or greater then $10^{20}cm^{-3}$, there is negligible difference between the parabolic and nonparabolic results. Indeed, between these doping limits, the difference is less than 0.2 on the log scale. (That the parabolic curve falls below the nonparabolic curve is obvious from Eq. 2.9.) By the choice of parameters for the surface charges, greater difference is seen between this curve and the nonparabolic curve than between that and the parabolic curve. Qualitatively, including surface charges causes little difference, except at dopings greater than $10^{20}cm^{-3}$, where the curve flattens. The reason for this can be seen from Fig. 2.1. At $10^{20}cm^{-3}$, the negative surface charges have greatly reduced the barrier. Effectively, there is no barrier. So further increases in doping will have little effect. Quantitatively, it may be said that the presence of any negative surface charge will accentuate the difference between these results and CFS's results. The effect of neglecting image-force lowering is slight; it causes the curves to shift upwards slightly by ~ 0.2.

2.4 Fluctuations

One significant effect that has not been included in this theory is fluctuations in the potential due to the discreteness in the doping. In the next chapter, this idea shall be developed in some detail. A simple Monte Carlo model will be used to find the fluctuations from simulations. The effects of the fluctuations on the conductance as a function of doping and contact size will be calculated, and compared these to the results presented in this chapter.

2.5 Conclusion

A model has been presented for the calculation of contact resistance and conductance as a function of energy of the tunnelling electron, for a metal– n–GaAs junction, where the GaAs is heavily doped, so that tunnelling is significant. This model uses the WKB approximation and the two–band model to obtain a non-analytic expression for the transmission probability. The contact resistance at zero bias is compared with that from CFS's model. It is suggested that this model determines the transmission probability more accurately than CFS's model, which assumes constant transmission for all tunnelling energies.

References

1. P. A. Barnes and A. Y. Cho, Appl. Phys. Lett. **33**, 651 (1978).
2. C. Y. Chang, Y. K. Fang and S. M. Sze, Sol. St. Elect. **14**, 541 (1971).
3. J. G. Werthen and D. R. Scifres, J. Appl. Phys. **52**, 1127 (1981).
4. R. L. Mozzi, W. Fabian and I. J. Piekarski, Appl. Phys. Lett. **35**, 337 (1979).
5. N. Braslau, J. Vac. Sci. and Tech. **19**, 803 (1981).
6. S. M. Sze, *Physics of Semiconductor Devices*, ch. 5, (Wiley, New York, 1981).
7. G. H. Parker, T. C. McGill, C. A. Mead and D. Hoffman, Solid St. Elect. **11**, 201 (1968).
8. J. W. Conley and G. D. Mahan, Phys. Rev. **161**, 681 (1967).
9. D. R. Fredkin and G. H. Wannier, Phys. Rev. **128**, 2054 (1962).
10. J. W. Conley, C. B. Duke, G. D. Mahan and J. J. Tiemann, Phys. Rev. **150**, 466 (1966).
11. D. J. BenDaniel and C. B. Duke, Phys. Rev. **152**, 683 (1966). See also Ref. 9.
12. E. O. Kane, *Physics of 3–5 Compounds, Vol. 1*, ch. 3 (Academic Press, New York, 1966).
13. K. S. Kunz, *Numerical Analysis*, (McGraw-Hill, New York, 1957).
14. C. D. Thurmond, J. Electro. Soc. **122**, 1133 (1975).
15. J. S. Blakemore, J. Appl. Phys. **53** ,R123 (1982).
16. R. H. Fowler and L. W. Nordheim, Proc. Roy. Soc. (London) **A119** 173 (1928).
17. W. J. Boudville and T. C. McGill, Appl. Phys. Lett. **48**, 791 (1986).

Chapter 3

Resistance Fluctuations in Ohmic Contacts

3.1 Introduction

One of the standard models of ohmic contacts is a metal-semiconductor junction in which the semiconductor is sufficiently heavily doped that tunneling is the major transport mechanism through the Schottky barrier.[1] This ideal model of the ohmic contact has been realised in recent years. Ohmic contacts to n–GaAs have been made by molecular beam epitaxy, where the layers of GaAs have been heavily doped with Sn[2] or Si,[3] with dopings in the range of 10^{19} to $10^{20} cm^{-3}$. At these dopings, the current through a metal–n–GaAs junction thus formed is dominated by tunnelling. The model of ohmic contacts by Chang, Fang and Sze (CFS)[1] has been most commonly used to analyse results such as those by Barnes and Cho[2] and Kirchner *et al.*[3]

In the previous chapter an improved theory was presented. Here the theory is extended to include the effects of fluctuations in the potential due to the randomness of the doping within the depletion region. These fluctuations could be important in the range of doping where ohmic contacts are formed. For example, at dopings around $10^{19} cm^{-3}$, where the average interdopant separation is ~ 30 Å, the depletion length is ~ 120 Å. Thus an electron traversing through the depletion region would be likely to encounter, not a potential due to a continuum of ionised dopants, but a potential influenced strongly by a few dopants, and hence varying considerably.

Section 2 describes how to compute the effect of dopant fluctuations on the average resistance of a metal-semiconductor ohmic contact. Results are given for 200 Å square contacts and for three possible barrier heights of the junction. Section 3 extends this by investigating how the resistance fluctuations scale as a function of contact size and doping. A theory is derived which is found to have good agreement with simulations. In Sections 2 and 3, the assumption is made that the trajectories of the free carriers through the depletion region were taken to be normal to the

junction. The results of Sections 2 and 3 indicate that resistances can vary by an order of magnitude over different parts of a given contact. Hence, there exists the possibility that 'hot spots' (regions of relatively low resistance) will conduct most of the current through a contact. This could result in current trajectories that are not normal to the junction, as they flow preferentially through the hot spots. In Section 4, a simple extension of classical mechanics is used to find current trajectories through the depletion region that are three dimensional in nature. It will be found from simulations in Section 4 that most of the current flow is indeed normal to the junction. This justifies the assumption made in the earlier sections.

Finally, Section 5 presents the conclusion for this chapter. The next chapter describes the possible direction of future research suggested by the results of this chapter.

3.2 Modelling Resistance Fluctuations

In this section, I will explain how the fluctuations in the resistance of ohmic contacts are modelled. Simulation results are presented for three different barrier heights of the semiconductor on the metal. These are for a 200 Å square contact of the metal on the semiconductor.

3.2.1 Calculating Resistance Fluctuations

This section describes how to calculate the effect of the random dopant configurations on the zero bias resistance of an ohmic contact.

As before, the WKB approximation is used to find the transmission through the depletion region,

$$T(E) = \begin{cases} exp\left(-2i\int_{x_1}^{x_2} k(x)dx\right), & 0 \leq E \leq qV_m, \\ 1, & E > qV_m, \end{cases} \tag{3.1}$$

where V_m is the height of the potential barrier with respect to the conduction band outside the depletion region, and x_1, x_2 are the turning points of the forbidden region for k, where the x direction is normal to the interface. Within the forbidden region, the wavector is given by the two-band model,[4]

$$k = -i\left[\frac{2m^*}{\hbar^2 E_g}\Big(\phi(x) - E\Big)\Big(E_g + E - \phi(x)\Big)\right]^{1/2}, \tag{3.2}$$

where E_g is the band gap, and $\phi(x)$ is the potential energy in the depletion region. At zero bias, the small-signal resistance is given by

$$\frac{1}{R_c} = G_c = \frac{m^* q^2}{2\pi^2\hbar^3}\int_0^\infty dE\,\frac{T(E)}{exp[(E - E_{FM})/k_B T] + 1}. \tag{3.3}$$

In the previous chapter, ϕ was found by including nonparabolic corrections and the L point states. The depletion region was assumed to consist of a uniform continuum of ionised dopants.

To include the effects of fluctuations in the potential, a different approach has been taken. In principle one could put down the dopants in a random fashion, solve for the electrostatic potential, and then solve for the transport through the potential barrier. However, this problem is much too difficult and some approximations are made.

Consider a square box in the depletion region. The square is in the plane of the interface, and the other side is normal to the interface, and of length x_d. See Fig. 3.1. For now, the length of a side of the square in the plane of the interface is taken to be 200 Å. To compute the potential in this box, a number of simplifying approximations have been made. First, it is assumed that the width of the depletion region does not vary in the plane of the interface. The depletion width is taken to be given by the standard continuum approximation,[8]

$$x_d = \sqrt{\frac{2\epsilon}{nq}(\phi_{B0} + E_{FS})}, \tag{3.4}$$

where ϕ_{B0} is the intrinsic barrier height, n is the density of donors, and E_{FS} is the Fermi level in the semiconductor. E_{FS} is found from Eqs. 2.6 and 2.7 in the previous chapter.

Second, the contributions to the potential seen by the electron at a point (x, y, z) in the depletion region is divided into three parts,

$$\phi(x, y, z) = \frac{-q^2}{16\pi\epsilon x} + \phi_1(x) + \phi_2(x, y, z). \tag{3.5}$$

The first term on the right hand side is the image term of the electron in the metal. The term ϕ_1 is due to dopants in the depletion region that are far from the electron, and, hence, should cause small fluctuations in the potential. The term ϕ_2 is the potential due to ionised dopants in the depletion region near the electron, and it is responsible for most of the fluctuations in the potential.

To be specific, the regions contributing to the potential seen by an electron are illustrated in Fig. 3.2. This is for an electron travelling along a particular trajectory normal to the junction. The trajectory is contained within the box of Fig. 3.1. The depletion region is divided into two parts: A cylinder of radius R and depth x_d, and the region outside the cylinder. The radius of the cylinder is taken to be twice the thickness of the depletion layer. ($R = 2x_d$ in Eq. 3.6 below.) In Section 3, this particular choice of the cylinder radius will be shown to be adequate in sufficiently accounting for the effect of dopant fluctuations. The contribution to the potential from the ionised donors outside the cylinder is calculated by assuming that the charge density is continuous. Fig. 3.1 shows a differential volume element dV outside the cylinder. The image of dV in the metal is shown; however, the image potential due to the undepleted GaAs is not included because of the small electron density when compared to a metal.[5] By integrating over the volume element in the depletion region and outside the cylinder, it is found that the continuum charge

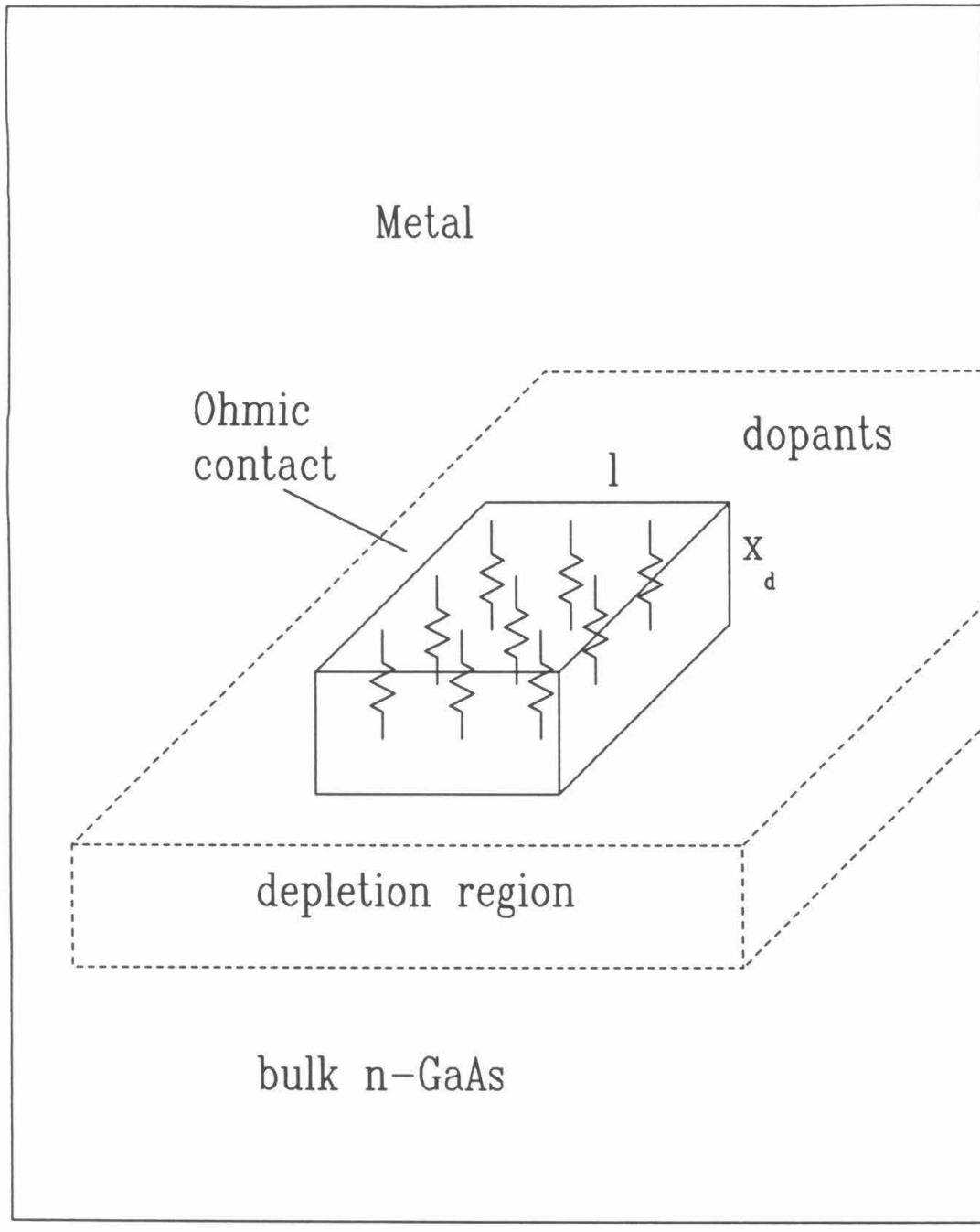

Figure 3.1: A square metal-semiconductor contact. The length of a side is l. The solid box is the contact; considered as a array of parallel resistors. The depletion length is x_d. The dopants are put randomly in the dotted box.

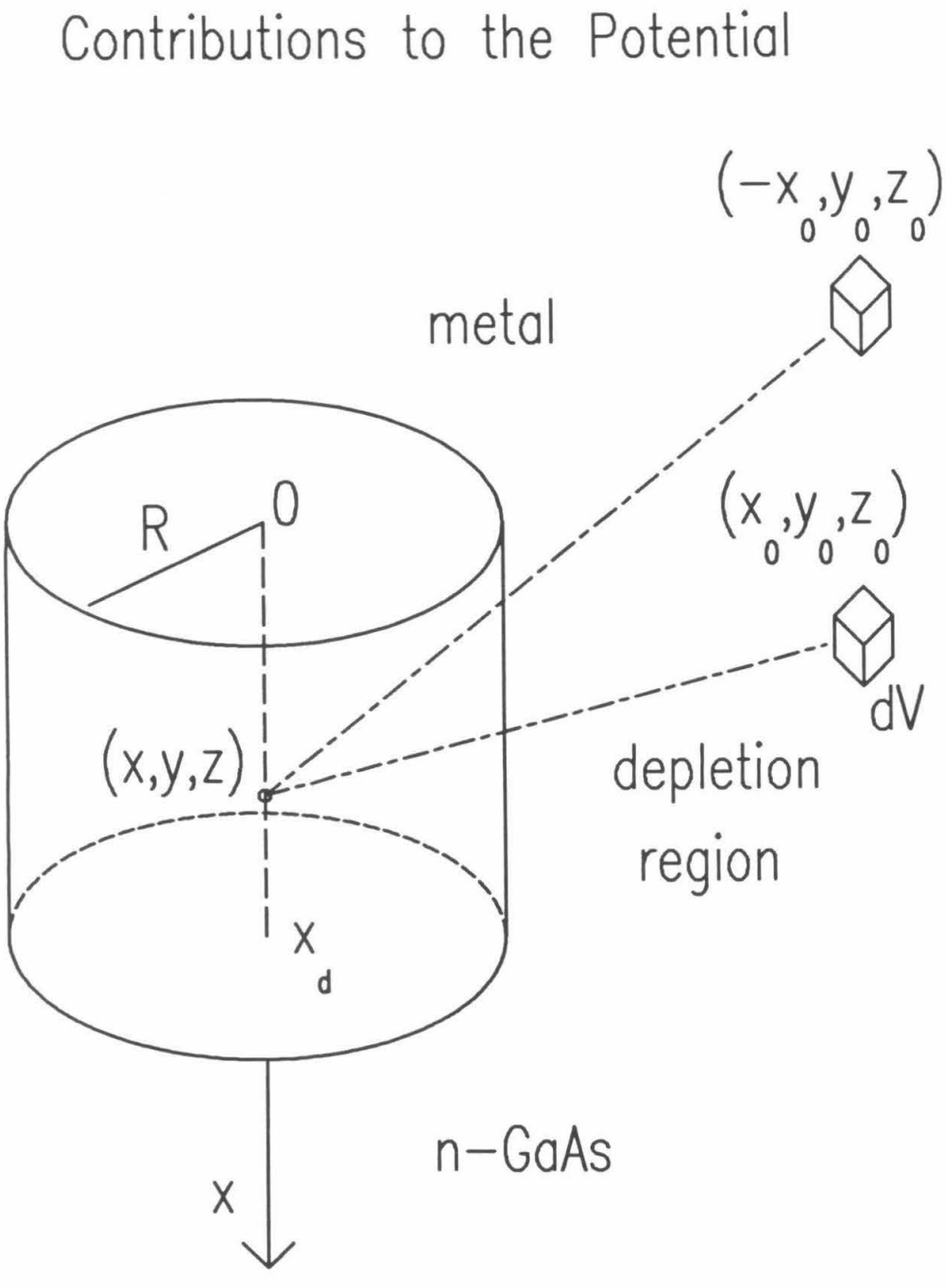

Figure 3.2: The contribution to the potential at (x, y, z) in the depletion region, from a differential charge at (x_0, y_0, z_0) outside the cylinder. Also shown is the contribution from the image at $(-x_0, y_0, z_0)$.

contribution is

$$\phi_1(x) = \phi_{m0} + \frac{nq^2R^2}{4\epsilon}\left(\frac{2x}{R}\sqrt{1+\left(\frac{x}{R}\right)^2} - \frac{(x+x_d)}{R}\sqrt{1+\left(\frac{x+x_d}{R}\right)^2}\right.$$
$$+\frac{(x_d-x)}{R}\sqrt{1+\left(\frac{x_d-x}{R}\right)^2} - 2\,ln\left[\sqrt{1+\left(\frac{x}{R}\right)^2} - \frac{x}{R}\right]$$
$$\left.-ln\left[\left(\sqrt{1+\left(\frac{x_d-x}{R}\right)^2} - \frac{(x_d-x)}{R}\right)\Big/\left(\sqrt{1+\left(\frac{x+x_d}{R}\right)^2} - \frac{(x+x_d)}{R}\right)\right]\right) \quad (3.6)$$

where $\phi_{m0} = \phi(0)$ in the absence of image force lowering.

The charged donors inside the cylinder are taken to be discrete and their contribution to the potential is calculated in detail, including their image in the metal. (The image contribution from the undepleted GaAs is neglected again.) Of course, this contribution to the potential will depend in detail on the spatial arrangement of the ionised donors found in the depletion region. The spatial configuration of the ionised donors is given by using a random number generator to distribute the ionised dopants on one of the two face-centred-cubic sublattices of a zincblende lattice. For n type doping, this corresponds to placing dopants at cation sites. The density is set by the doping. The atoms are inside a rectangular slab of GaAs whose thickness equals x_d. The dimensions of the rectangular slab are chosen to be large enough to include the box of Fig. 3.1 within which ϕ is to be found, and also to include cylinders centred on any trajectory within the box. The slab is also illustrated in Fig. 3.1.

The contribution to conductance per unit area for a given cross-sectional area is calculated using Eqs. 3.1–3.3. The transport is calculated as if the potential depended only on the x direction. The formulas are evaluated on an equally spaced grid with a grid spacing of 40 Å. To find the zero bias resistance of the contact, the conductances found from the grid of trajectories are averaged. From Fig. 3.1, the contact is considered to be equivalent to a parallel array of resistors between the

metal and the bulk semiconductor. Hence, to find the characteristic resistance of a particular contact, the natural quantities to average are the conductances along the different trajectories.

In a square contact of size l, the number of trajectories is proportional to l^2. Hence, due to computational necessity, the spacings between trajectories had to be made as large as possible without averaging out the amount of fluctuations between neighbouring trajectories. To investigate this, several dopant configurations were generated, and conductances were averaged along trajectories separated by different spacings. This was done at dopings of 10^{18}, 10^{19} and $2.5 \times 10^{19} cm^{-3}$. Little differences were found between spacing trajectories 5 Å and 40 Å apart. Accordingly, all the simulations described below are taken at the latter trajectory spacing. The trajectory spacing could perhaps be increased further still, but a qualitative upper bound is provided by the average interdopant separation. At the highest doping considered, $10^{20} cm^{-3}$, the average interdopant separation is 20 Å. If one wishes to consider effects of dopant discreteness, it is useful not to increase the trajectory spacing too much above this distance.

To examine the variations in potential and transport that would be obtained for other configurations, the random number generator was used to generate a number of different spatial arrangements of atoms. The entire transport calculation was carried out for each case. The results of simulations are presented in the next section.

3.2.2 Simulation Results

This section describes the resistance fluctuations in 200 Å size square contacts, found from random arrangements of dopants. The method of Section 2.1 is used to compute the fluctuations.

First, Fig. 3.3 shows an example of a potential found by this method, at a

doping of $8.0 \times 10^{19}\,\mathrm{cm}^{-3}$ and $\phi_{b0} = 0.8\,\mathrm{eV}$. The contour lines are lines of constant z coordinate, spaced 5 Å apart, where the z direction is in the plane of the interface. It is seen that the potential fluctuates significantly. Thus, at a given tunnelling energy, the length of the forbidden region can vary widely, especially if a dopant is nearby, causing the potential to fall. From Eqs. 3.1, 3.2 and 3.3, it is seen that this variation in length can cause a large variation in R_c.

In Figs. 3.4, 3.5 and 3.6, the results of the calculations are shown for barrier heights of 0.6, 0.8 and 1.0 eV along with the results for various models. The x direction has been chosen to be the $< 100 >$ direction of the lattice. The results are presented as specific contact resistance (the reciprocal of the conductance per unit area) versus the doping. Each dot in the figures is obtained from a different random configuration of dopants, and represents G_c, averaged over a square contact in the plane of the interface, with a side of length $l = 200$ Å. Individual G_c s are found at in a grid of y and z coordinates, spaced 40 Å apart. Five simulations are made at each doping value. Having obtained an averaged G_c for each dopant configuration, these averaged results are in turn averaged at each doping value. (These are not displayed in Figs. 3.4-3.6.) A least square polynomial is then fitted to the resultant $\{(log_{10}n, log_{10}R_c)\}$ at each barrier height. The curves are shown in the figures under the label 'Monte Carlo $< 100 >$'. For comparison, the results from CFS and the results from the theory developed in the previous chapter (labelled 'nonparabolic') are also displayed. It will be recalled that the results from the last chapter are obtained by assuming a continuum distribution of charged dopants within the depletion region, rather than the discrete case considered here.

A number of important points can be obtained by studying the results presented in these figures. From the figures, it is seen that for dopings $< 10^{18} cm^{-3}$, the Monte Carlo results exhibit small fluctuations about the continuum curve. However, as the doping is increased, the average resistance falls about half an

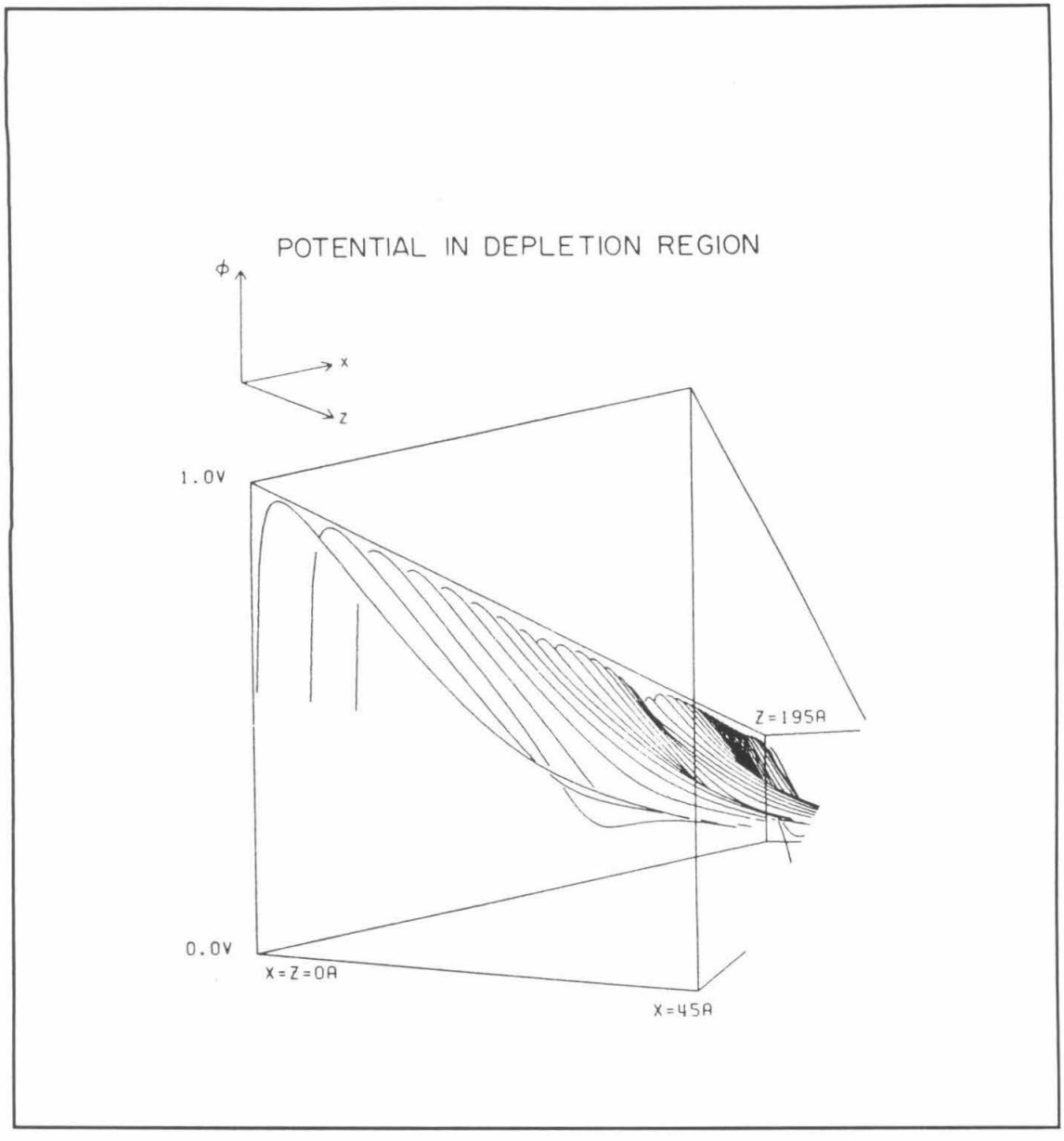

Figure 3.3: The random potential for one configuration of ionised donors in a plane in the depletion region. The average doping is $8.0 \times 10^{19} cm^{-3}$. The Schottky barrier height is 0.8eV. The applied bias is zero. The x direction is normal to the interface, and is the $< 100 >$ direction of the GaAs lattice. The z direction is along the interface.

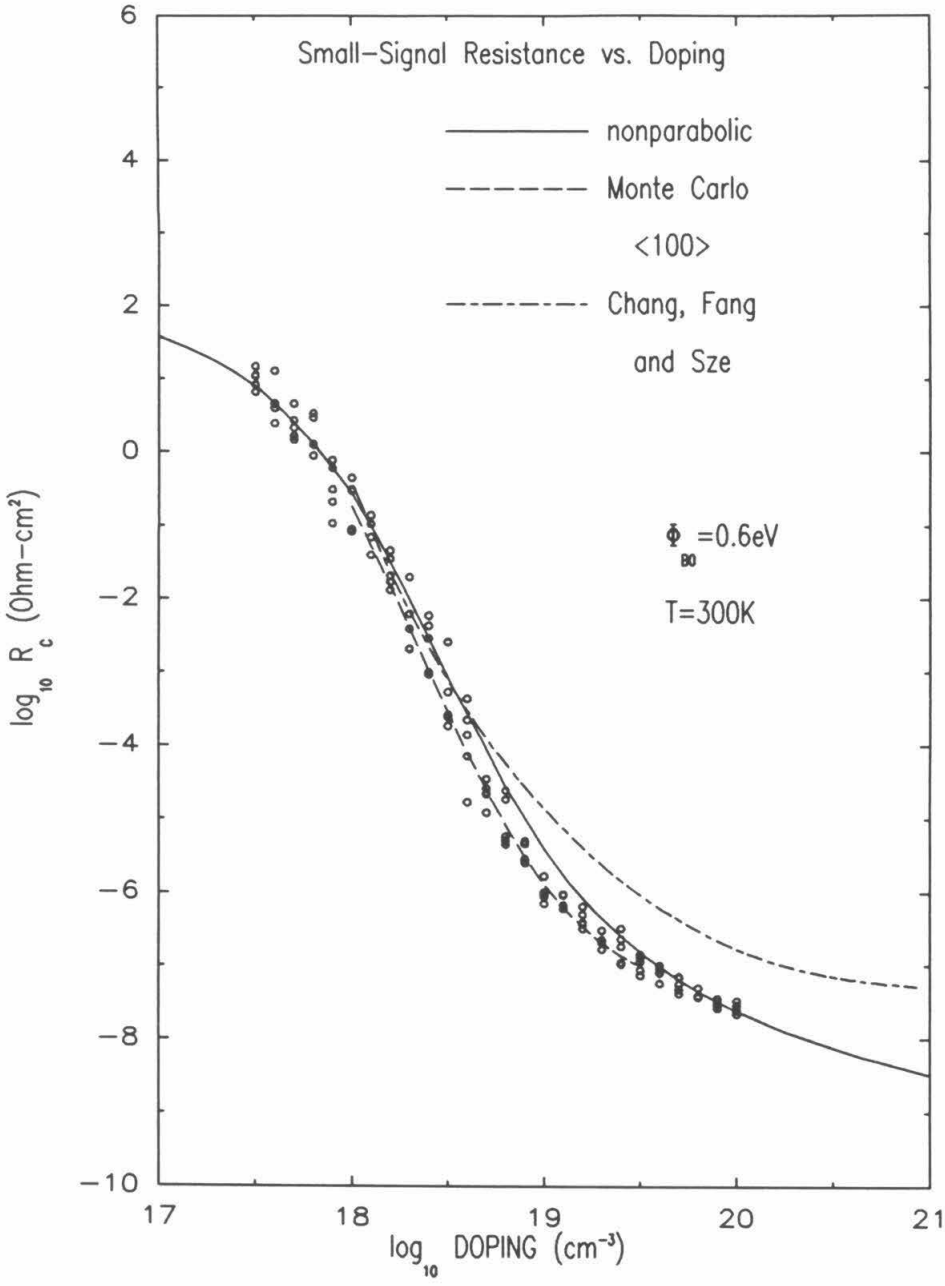

Figure 3.4: Specific contact resistance versus doping. The barrier height is 0.6 eV. The applied bias is zero. The points are simulations. The dashed line is fitted to the simulations. The solid line is from the theory in Ch. 2. The dash-dotted line is from CFS.

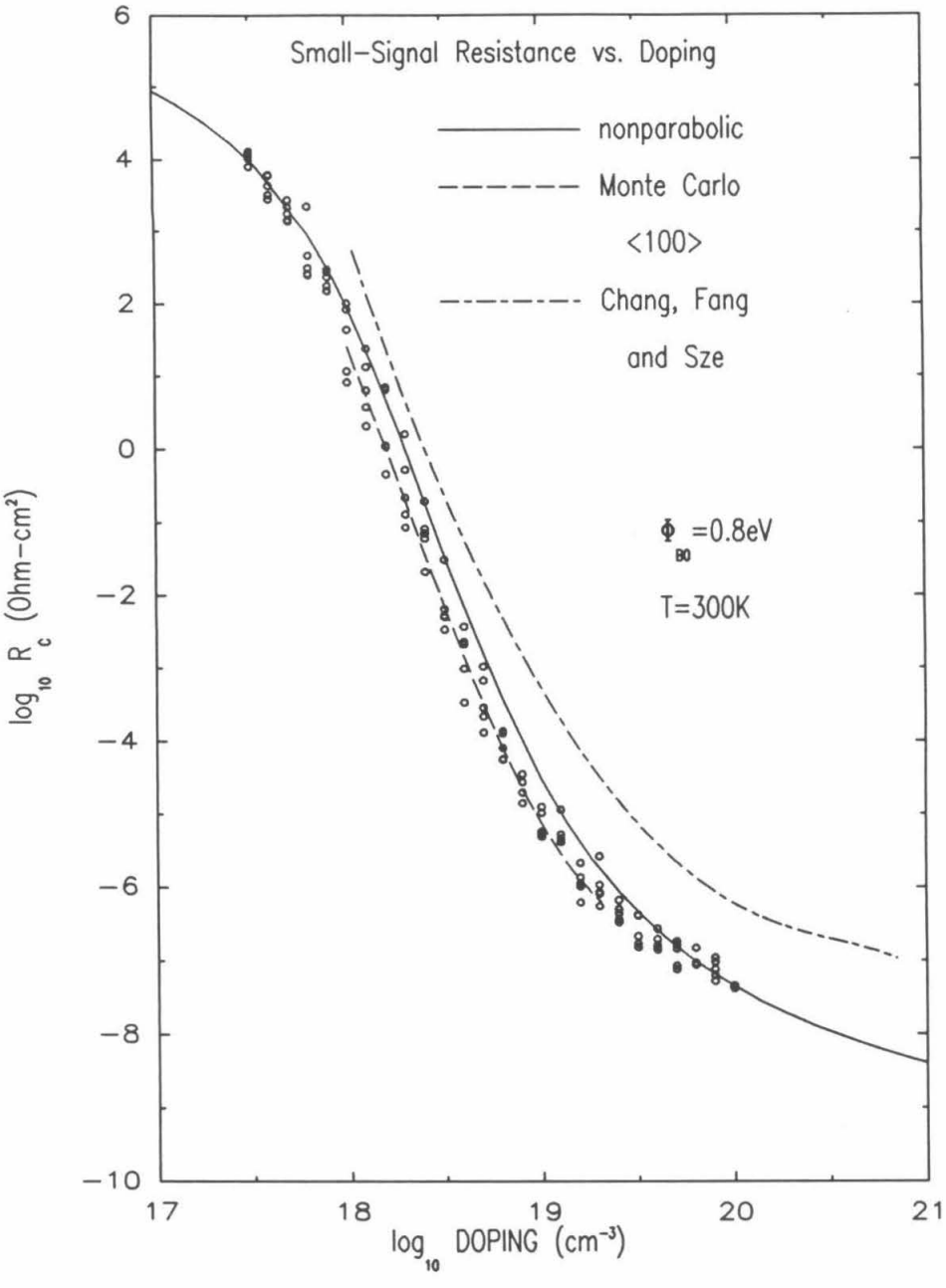

Figure 3.5: Specific contact resistance versus doping. The barrier height is 0.8 eV. The applied bias is zero. The points are simulations. The dashed line is fitted to the simulations. The solid line is from the theory in Ch. 2. The dash-dotted line is from CFS.

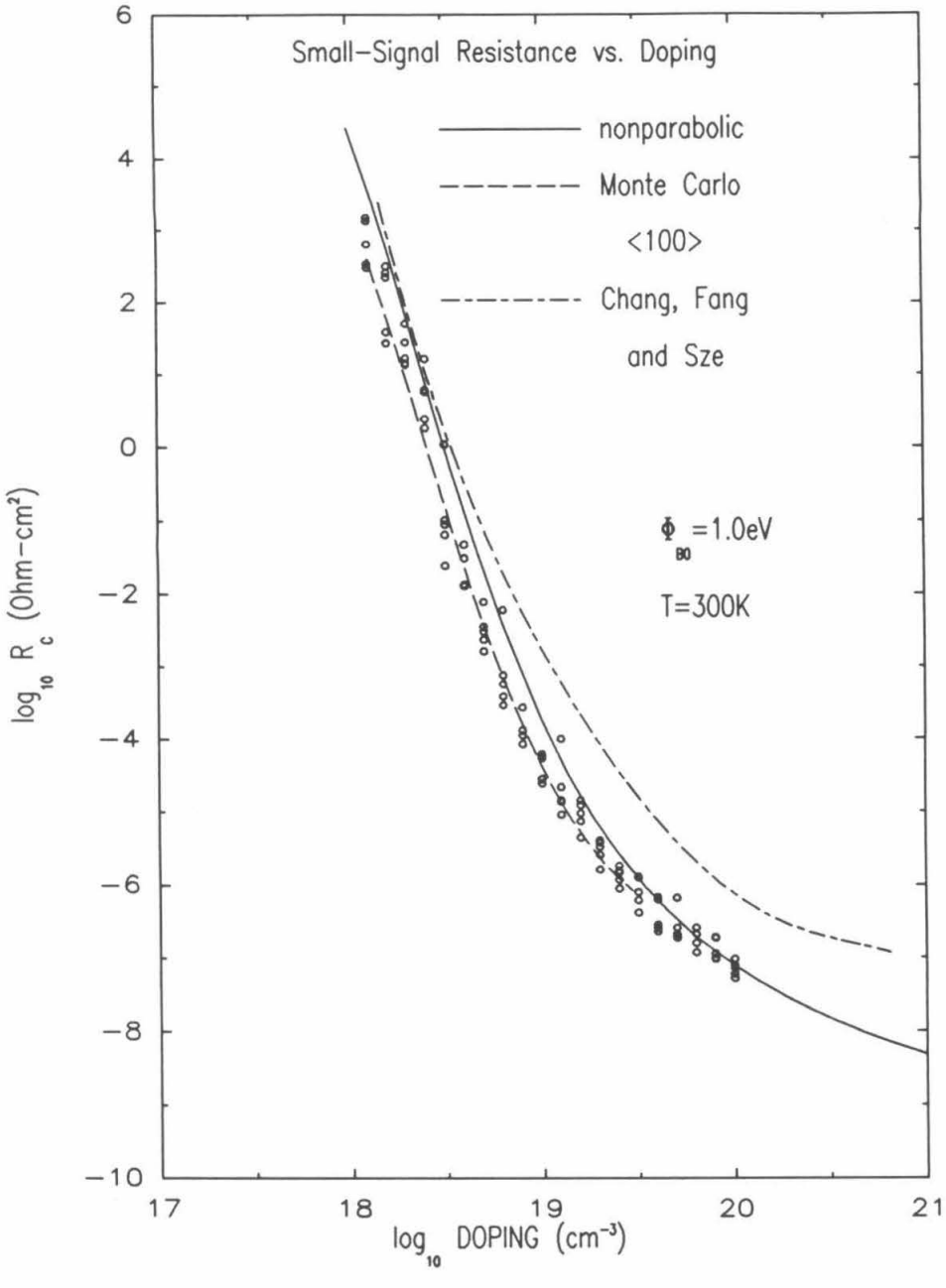

Figure 3.6: Specific contact resistance versus doping. The barrier height is 1.0 eV. The applied bias is zero. The points are simulations. The dashed line is fitted to the simulations The solid line is from the theory in Ch. 2. The dash-dotted line is from CFS.

order of magnitude below the continuum curve. Furthermore, the fluctuations increase to as much as an order of magnitude between results from different dopant configurations at the same doping and barrier height. Thus current through the junction may flow preferentially through 'hot spots': regions of high conductance. Figs. 3.7 and 3.8 show simulations that have been performed for $< 111 >$ and $< 110 >$ orientations of the semiconductor on the metal. The barrier height in these figures has been taken as $\phi_B = 0.8\,eV$. This is the typical barrier height for a metal-n-GaAs contact.[8] For the same random seed, the particular dopant configuration will of course vary, depending on the lattice orientation. This will lead to different resistances. But for a set of simulations at a different lattice orientation, the resultant contact resistances display the same behaviour as those plotted in Fig. 3.5 for the $< 100 >$ orientation and the $0.8\,eV$ barrier height.

3.2.3 Remarks

It has been found that the discreteness of the ionised dopants, within the depletion region of a metal–n–GaAs junction, causes a significant lowering in the small–signal resistance at zero bias, as a function of doping. The small–signal resistance can be up to half an order of magnitude less than those resulting from assuming a continuous charge distribution, and up to an order–and–a–half less than the results obtained by CFS.[1] More importantly, large fluctuations in the conductance are found.

From the results at the three different barrier heights, it is apparent that resistance fluctuations will be found at an ohmic contact, regardless of the particular barrier height of the semiconductor on the metal. Also, such fluctuations will occur regardless of the lattice orientation of the semiconductor on the metal.

These fluctuations may have serious implications for the transport properties of metal-semiconductor junctions. The results indicate that for a semiconductor

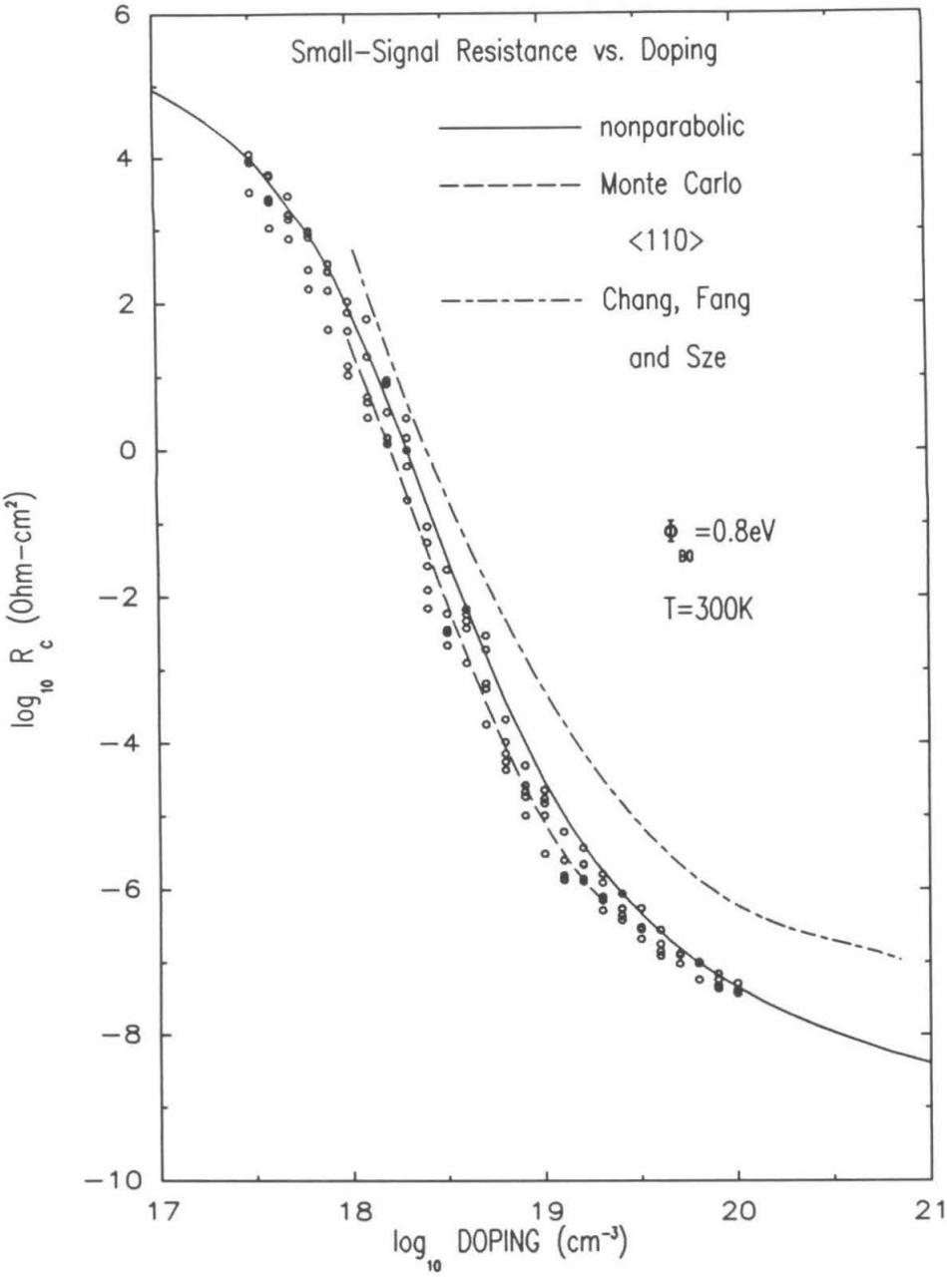

Figure 3.7: Specific contact resistance versus doping. The lattice is oriented at $< 110 >$. The barrier height is 0.8 eV. The applied bias is zero. The points are simulations. The dashed line is fitted to the simulations. The solid line is from the theory in Ch. 2. The dash-dotted line is from CFS.

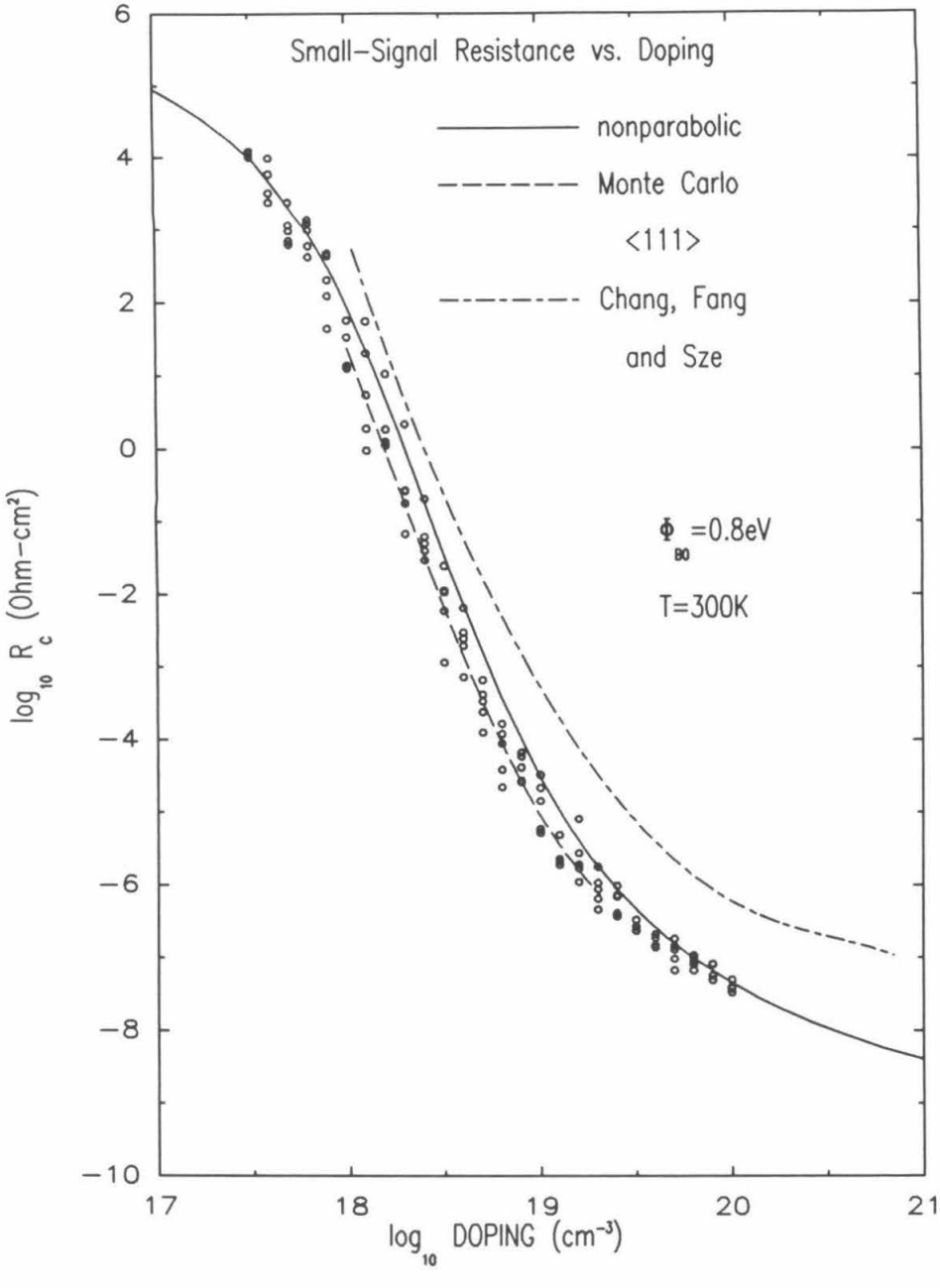

Figure 3.8: Specific contact resistance versus doping. The lattice is oriented at $< 111 >$. The barrier height is 0.8eV. The applied bias is zero. The points are simulations. The dashed line is fitted to the simulations. The solid line is from the theory in Ch. 2. The dash-dotted line is from CFS.

which is uniformly doped on a macroscopic scale, the conductance, and hence the current, can differ by as much as an order of magnitude, over different spatial regions of the junction. This large spatial variation in the current may be a significant contribution to the noise of a metal-semiconductor junction, and hence may adversely affect the performance of micron or sub-micron size devices employing such junctions.

3.3 Scaling of Fluctuations

Here, the work of Section 2 is extended to include investigating the dependence of the resistance fluctuations on the cross sectional area, in the plane of the interface. The intention here is to see how quickly the fluctuations, which were simulated in Section 2 at a cross sectional length of 200 Å, average out as larger cross sections are considered, and as a function of doping. In Subsection 3.1, the results of the simulations are presented. In Subsection 3.2, I describe the theoretical predictions of the variation of the fluctuations in contact resistance with the size of the dot and the doping. A comparison between the results of Subsections 3.1 and 3.2 is made in Subsection 3.3. Finally, Subsection 3.4 contains the conclusion for Section 3.

3.3.1 Simulation Results

In this subsection, results are given for simulations made of metal-semiconductor contacts, where the size of the contacts and the doping in the semiconductor are varied.

The resistance fluctuations are considered as a function of cross sectional area, in the plane of the metal-n-GaAs interface. Dopants are randomly placed at substitutional sites in the depletion region, and the conductance is then found along

one dimensional trajectories, normal to the interface. The conductances are then averaged over a cross sectional area. This method was described in detail in Section 2. Here, the conductances normal to the interface are found in a square grid lying in the plane of the interface, with sides of length l. It is of interest to now find how the conductance fluctuations scale, as l is increased.

Simulations were performed at three dopings: 10^{18}, 10^{19} and $2.5 \times 10^{19} cm^{-3}$. Square devices were simulated, with the conductances being found in a square grid and then averaged, to characterise the device. For the dopings of 10^{19} and $2.5 \times 10^{19} cm^{-3}$, runs were made for 10 different square cross sectional areas, with the length of a side of a device ranging in size from 200 Å to 2000 Å. While, at a doping of $10^{18} cm^{-3}$, six different cross sectional areas were considered, ranging from 200 Å to 1600 Å. For each device size and doping, 20 simulations were performed. In each simulation, a different random configuration of dopants was generated to find an average conductance for the device. Given a trajectory spacing of 40 Å, the number of trajectories varied from 25 at a device size of 200 Å to 2500 at a device size of 2000 Å.

The results of a representative subset of these runs are displayed in Fig. 3.9. For the sake of clarity, not all the cross sectional results are shown, and for those which are, only a representative quarter of the runs are illustrated. Also plotted are the average conductance and the 1σ lines (where the 1σ lines are the mean plus or minus one standard deviation), as found from the full set of 20 Monte Carlo runs at each cross section.

It can be seen that at a doping of $2.5 \times 10^{19} cm^{-3}$, the fluctuations (as measured by the 1σ lines) converge to less than a tenth of an order of magnitude by 2000 Å. While at a doping of $10^{19} cm^{-3}$, the fluctuations converge similarly from about half of an order of magnitude at 200 Å, to a tenth of an order of magnitude at 2000 Å. Visibly, though, the fluctuations are greater than at $2.5 \times 10^{19} cm^{-3}$. At a lower

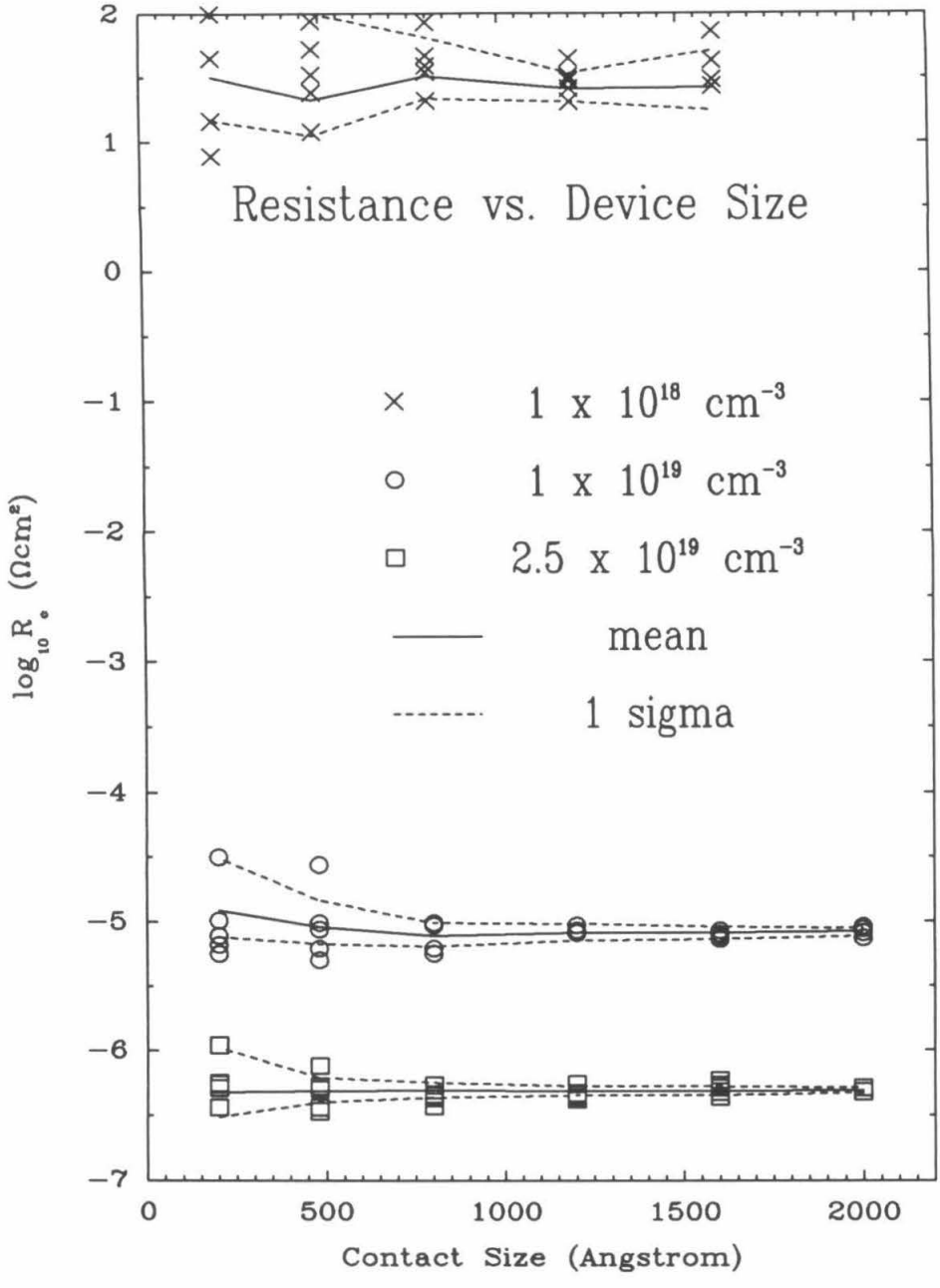

Figure 3.9: The dependence of resistance fluctuations on contact size in the plane of the junction at dopings of $10^{18}, 10^{19}$ and $2.5 \times 10^{19} cm^{-3}$. The solid lines are the means, and the dashed lines are the 1σ lines, found from the simulations. A typical subset of the runs is shown.

doping of $10^{18}cm^{-3}$, convergence is slower still; going only to about a quarter of an order of magnitude at 1200 Å. This would suggest that the contributions of resistance fluctuations to the noise properties of the junction would be more significant at $10^{18}cm^{-3}$ that at 10^{19} or $2.5 \times 10^{19}cm^{-3}$, even over device sizes of less than a thousand Angstroms.

3.3.2 Theoretical Scaling Dependence

Here, scaling arguments are derived for the fluctuations in the conductance of metal-semiconductor contacts. The analysis in this section is outlined as follows. First, start on a microscopic level with the fluctuations in the presence or absence of dopants at lattice sites. These dopants are ionised in the depletion region, so an expression can be found for the fluctuations in the potential energy produced by the dopants, along a trajectory that is normal to the interface. Next, these fluctuations are related to the resultant fluctuations in the conductance along the trajectory. Finally, the conductances are averaged over a square grid of trajectories, for a contact size of l, and the fluctuations in this averaged conductance are investigated.

Consider first the variation in the potential energy, ϕ, along a trajectory normal to the interface. Let this be the x direction. Let the trajectory be at $y = z = 0$, where y and z are in the plane of the interface. At $(x', 0, 0)$, where $x' < x_d$ =the depletion length, the contribution to the potential due to a dopant at a lattice site (x, y, z) in the depletion region is given by

$$\phi(x', 0, 0; x, y, z) = \frac{q}{4\pi\epsilon} g(x, y, z) \left[\frac{1}{\sqrt{(x - x')^2 + r^2}} - \frac{1}{\sqrt{(x + x')^2 + r^2}} \right], \quad (3.7)$$

where $r = \sqrt{y^2 + z^2}$ and $g(x, y, z)$ is the dopant occupation number. It is assumed that the $\{g\}$ at different lattice sites are independent random variables, where

$$g(x, y, z) = \begin{cases} 1, & \text{with probability } p \\ 0, & \text{with probability } 1 - p, \end{cases}$$

and p is found from the average doping, n. Let $V[\alpha]$ denote the variance of an arbitrary random variable α. Thus, at a lattice site (x, y, z) the binomial distribution gives[7]

$$V[g(x,y,z)] = p(1-p). \tag{3.8}$$

With this, the fluctuations in the potential energy at the point $(x', 0, 0)$ will now be found. As was done in Section 2, a cylinder is erected. It is of length x_d and radius R, such that its axis is normal to the interface and coincides with the trajectory $y = z = 0$. The fluctuations in $\phi(x', 0, 0)$ are attributed to the dopants within this cylinder. An obvious question now arises: How can a value of R be chosen such that there are enough dopants inside the cylinder to adequately model the potential fluctuations, and yet have the simulations be amenable to reasonable computational effort? This issue is addressed below.

Integrating over the cylinder, it is found that

$$V[\phi(x',0,0)] = \int_0^R dr \int_0^{x_d} dx \, \frac{2\pi r}{V_c} \left[\frac{\partial \phi(x',0,0;x,y,z)}{\partial g(x,y,z)}\right]^2 V[g(x,y,z)], \tag{3.9}$$

where V_c is the volume per lattice site. So $V_c \propto a^3$, where a is the lattice constant of the semiconductor. (Strictly, Eq. 3.9 is true only when g is a continuous random variable, but, if need be, g may be modelled as such, and appropriate limits taken that reduce to the discrete model of g that is considered here.) It is seen that

$$\begin{aligned} V[\phi(x',0,0)] &= \frac{2\pi p(1-p)x_d}{V_c}\left(\frac{q}{4\pi\epsilon}\right)^2 \\ &\times \int_0^{R/x_d} dr\, r \int_0^1 ds \left[\frac{1}{\sqrt{r^2+(s-x'/x_d)^2}} - \frac{1}{\sqrt{r^2+(s+x'/x_d)^2}}\right]^2 \\ &= \frac{2\pi p(1-p)x_d}{V_c}\left(\frac{q}{4\pi\epsilon}\right)^2 f(x'/x_d; b), \end{aligned} \tag{3.10}$$

where b is the aspect ratio of the cylinder within which the dopants are considered to be discrete, $b \equiv R/x_d$. $f(x'/x_d; b)$ is a dimensionless function of its first

argument, which is the normalised distance inside the depletion region, and f is parametrised by its second argument, the aspect ratio b. It can be shown that f is given by

$$\begin{aligned} f(x;b) \equiv\ & b\left[arctan\left(\frac{1-x}{b}\right) + arctan\left(\frac{1+x}{b}\right)\right] \\ & + \left(\frac{1-x}{2}\right) ln\left(1 + \frac{b^2}{(1-x)^2}\right) + \left(\frac{1+x}{2}\right) ln\left(1 + \frac{b^2}{(1+x)^2}\right) \\ & + \int_0^1 ds \Bigg[ln(|s^2 - x^2| + s^2 + x^2) \\ & - ln\left(\sqrt{b^4 + 2b^2(s^2+x^2) + (s^2-x^2)^2} + b^2 + s^2 + x^2\right)\Bigg]. \end{aligned} \tag{3.11}$$

Fig. 3.10 displays f, for various values of the aspect ratio. The uppermost curve in Fig. 3.10 is for the case of $b = \infty$, when the contribution to the fluctuations is taken from the entire depletion region. The curve for $b = 2$ effectively coincides with this curve. Thus, to adequately model the effect of fluctuations, it suffices to take $b = R/x_d = 2$, and this has been done for the results presented in this chapter.

Having found the fluctuations in the potential, let us now investigate the fluctuations in the conductance. At zero bias, the conductance G is given by Eq. 3.3. It is assumed that the GaAs is heavily doped, so that the tunnelling contribution dominates the thermionic contribution to the conductance. Then, the important contribution to the integral in Eq. 3.3 is over energies E such that tunnelling occurs, with $T(E)$ being given by the WKB approximation,

$$T(E) = exp(-A),$$

with A being defined as

$$A \equiv 2i \int_{x_1}^{x_2} dx\, k(x, E),$$

where x_1 and x_2 are the classical turning points.

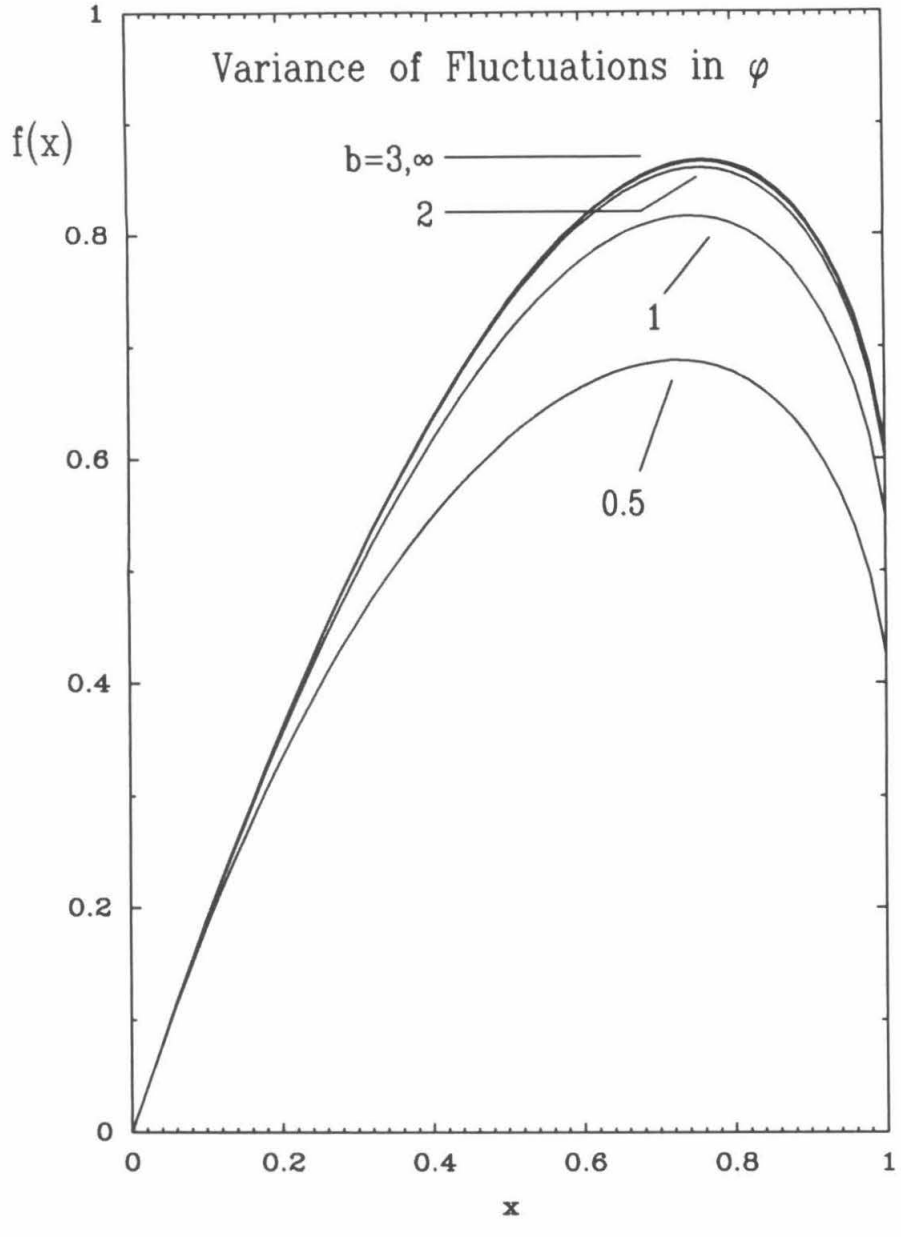

Figure 3.10: The function $f(x;b)$ plotted for several values of b. It models the spatial dependence of the variance in the potential energy fluctuations, from Eq. 3.11. The x coordinate is the normalised distance in the depletion region, from the interface, and b is the aspect ratio of a cylinder around a trajectory normal to the interface.

Eq. 3.2 gives the dependence of k on the potential. Because of the sensitive nature of $T(E)$ at tunnelling energies, it should be expected that the fluctuations in the conductance are dominated by the fluctuations in A.

Given G along a particular trajectory, let it be labelled by (y, z) coordinates. To find the conductance of a device, $\{G\}$ is now averaged over a square grid of trajectories,

$$\overline{G} \equiv \frac{1}{m}\sum_{i=1}^{m} G_i$$

where $G_i \equiv G(y_i, z_i)$, and m is the total number of trajectories in the grid. Let the length of a side of the device be l. So $l = \Delta\sqrt{m}$, where Δ is the spacing between trajectories. Of technological interest is how $\overline{G}$ will vary between devices of the same size and doping. Now for the variance in $\overline{G}$,

$$V[\overline{G}] \simeq \frac{1}{m^2}\sum_{k}\Big[\frac{\partial}{\partial g_k}\Big(\sum_{i} G_i\Big)\Big]^2 V[g_k], \qquad (3.12a)$$

where the sum over k is over all lattice sites, and the partial derivatives are evaluated at $g_k = E[g_k] = p$ for all k. Eq. 3.12a contains the assumption that the variance of $\overline{G}$ can be approximated by retaining only the linear terms in the Taylor expansion of $\overline{G}$ about $g_k = p$.[7] From Eq. 3.3, G can be written as

$$G = \int dE\, h(E)\, e^{-A}.$$

So

$$V[\overline{G}] \simeq \frac{1}{m^2}\sum_{k}\Big[\sum_{i}\int dE_i\, h(E_i)\, e^{-A_i}\,\frac{\delta A_i}{\delta\phi_i}\frac{\partial\phi_i}{\partial g_k}\Big]^2 V[g_k],$$

where $\delta A_i/\delta\phi_i$ indicates a functional derivative, and $\partial\phi_i/\partial g_k$ is evaluated inside the $\int dx_i$ that arises from the functional derivative. So it is found that

$$V[\overline{G}] \simeq \frac{1}{m^2}\sum_{ij}\int dE_i\, dE_j h(E_i)h(E_j)\, e^{-A_i-A_j}\,\frac{\delta A_i}{\delta\phi_i}\frac{\delta A_j}{\delta\phi_j}\sum_{k}\frac{\partial\phi_i}{\partial g_k}\frac{\partial\phi_j}{\partial g_k}\,V[g_k]. \qquad (3.12b)$$

Let us define

$$L \equiv \sum_{k}\frac{\partial\phi_i}{\partial g_k}\frac{\partial\phi_j}{\partial g_k}\,V[g_k].$$

Let i be fixed. Then $\partial\phi_i/\partial g_k \neq 0$ only if k labels a lattice site inside a cylinder of radius R centred on trajectory i. And similarly for $\partial\phi_j/\partial g_k$. Hence, the only non zero terms in L arise when i and j label overlapping cylinders, as may be seen in Fig. 3.11, where B is the volume of intersection. Making a change from a sum over k to an integral over B, it is found that

$$L = \frac{p(1-p)}{V_c}\int_B d^3x\,\frac{\partial\phi_i}{\partial g(x)}\frac{\partial\phi_j}{\partial g(x)},$$

where Eq. 3.7 is used to obtain $\partial\phi/\partial g$.

In the double sum over i and j in Eq. 3.12b, let the sum over i be the outer sum. Then the sum over j is confined to trajectories within a circle of radius $2R$ centred on i. Here, edge effects of the contact are neglected. Eq. 3.12b now needs to be simplified. Let a point on the i trajectory have coordinates (x_1, y_1, z_1) and a point on the j trajectory have coordinates (x_2, y_2, z_2). Choose the origin of the axes to lie on the interface at the i trajectory, and align the y axis to point to the j trajectory. Let t be the distance between the trajectories. Hence, $y_1 = z_1 = z_2 = 0$ and $y_2 = t$. Now

$$\begin{aligned} V[\overline{G}] &\simeq \frac{1}{m^2}\sum_i\int dE_i\,h(E_i)\,e^{-A_i}\,\frac{\delta A_i}{\delta\phi_i}\sum_j\int dE_j\,h(E_j)\,e^{-A_j}\,\frac{\delta A_j}{\delta\phi_j}\,L \\ &= \frac{1}{m^2}\sum_i\int dE_i\,h(E_i)\,e^{-A_i}\,\frac{\delta A_i}{\delta\phi_i}\int dE_j h(E_j)\,e^{-A_j}\,\frac{\delta A_j}{\delta\phi_j}\frac{1}{\Delta^2}\int_0^{2R} t\,dt\int_0^{2\pi} d\psi\,L, \end{aligned} \tag{3.13}$$

where the sum over j was converted to an integral using polar coordinates (t, ψ). Let us make the definition

$$K(x_1, x_2) \equiv \int_0^{2R} t\,dt\int_0^{2\pi} d\psi\,L.$$

By using the above coordinates and changing all variables of integration to dimen-

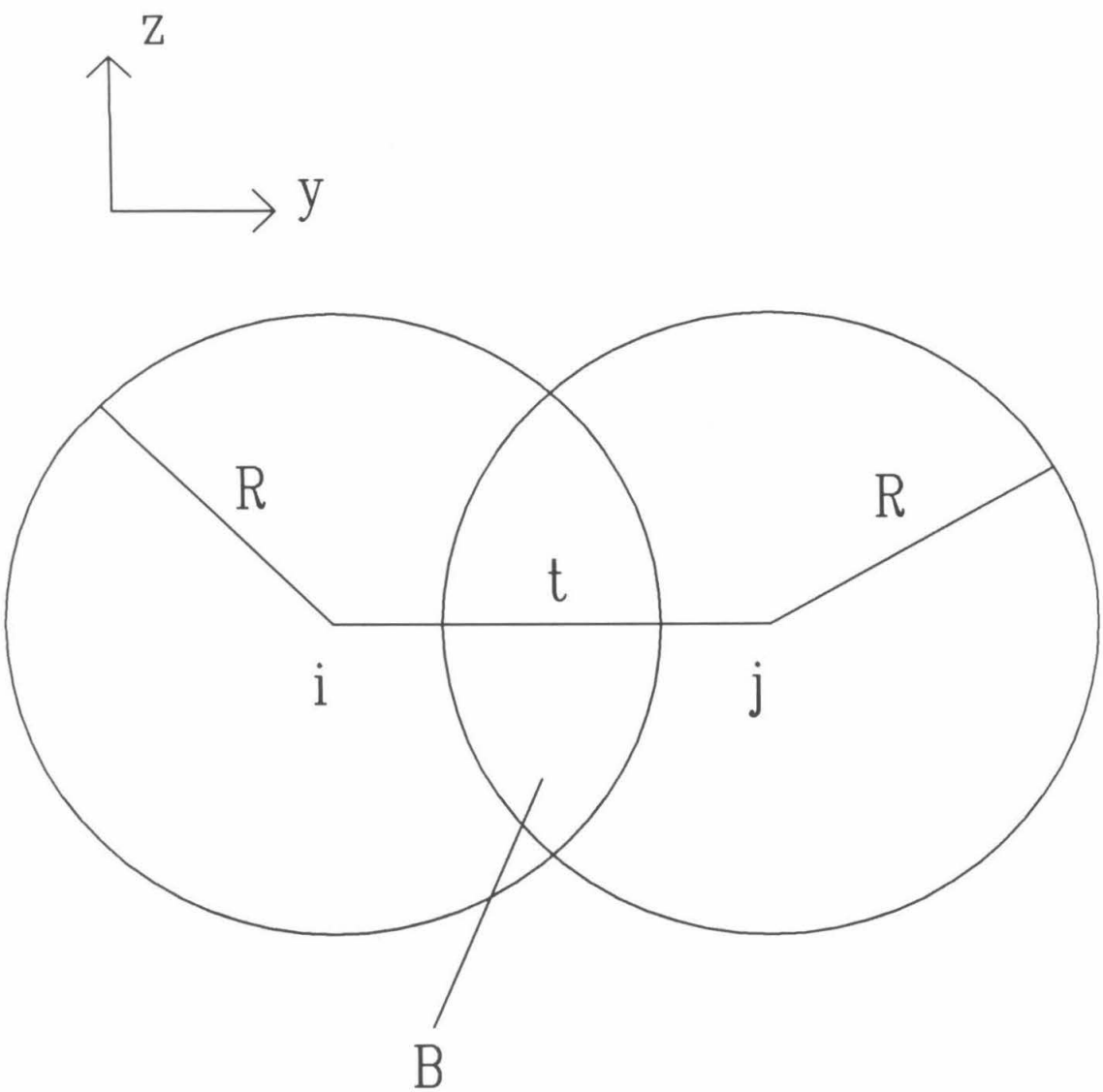

Figure 3.11: The interface, with two trajectories, i and j, separated by a distance t. The random contribution to the potentials along i and j are found from dopants inside the cylinders, shown in cross section. B is the volume of intersection.

sionless quantities, it can be shown that

$$K = \frac{32\pi R^4 p(1-p)}{x_d V_c}\left(\frac{q}{4\pi\epsilon}\right)^2 \int_0^1 dt \int_0^{\sqrt{1-t^2}} dz \int_t^{\sqrt{1-z^2}} dy \int_0^1 dx$$
$$\times t\left[\frac{1}{\sqrt{b^2(y^2+z^2)+(x_1'-x)^2}} - \frac{1}{\sqrt{b^2(y^2+z^2)+(x_1'+x)^2}}\right]$$
$$\times\left[\frac{1}{\sqrt{b^2z^2+b^2(2t-y)^2+(x_2'-x)^2}} - \frac{1}{\sqrt{b^2z^2+b^2(2t-y)^2+(x_2'+x)^2}}\right], \tag{3.14}$$

where $x_1' \equiv x_1/x_d$ and $x_2' \equiv x_2/x_d$, and $R = b\,x_d$ was used. Consider the integrand of Eq. 3.14. The main contribution to K comes when the first denominator in each of the square brackets is small. This occurs when $t, z \to 0+$ and $y \to t+$. Then each square bracket is large when $x = x_1'$ or $x = x_2'$ and the second term in each bracket can be neglected. For maximum contribution, $x_1' = x_2'$ is needed. Also, for small t, the integrand then tends to

$$t\frac{1}{\sqrt{b^2t^2}}\frac{1}{\sqrt{b^2(2t)^2}} \to \frac{1}{t}.$$

This is plausible, since $V[\overline{G}]$ should be strongly influenced by the instances when the trajectories coincide as, from Fig. 3.11, this gives the greatest volume overlap between the trajectories' cylinders.

Returning to Eq. 3.13, it is seen that $(\delta A_i/\delta\phi_i)\,(\delta A_j/\delta\phi_j)$ gives $\int dx_1\,dx_2$. This double integral can then be rescaled to dimensionless variables of integration and hence two powers of x_d may be extracted. Substituting all this into Eq. 3.13 yields

$$V[\overline{G}] \simeq \frac{32\pi p(1-p)R^4 x_d}{m V_c \Delta^2}\left(\frac{q}{4\pi\epsilon}\right)^2 \int dE_i\, h(E_i)\, e^{-A_i}\frac{\delta A_i}{\delta\phi_i}\int dE_j\, h(E_j) e^{-A_j}\frac{\delta A_j}{\delta\phi_j}\,K', \tag{3.15}$$

where K' designates the dimensionless multiple integral in Eq. 3.14, and the sum over i in Eq. 3.13 was over m identical terms, enabling its reduction to the above expression. Now for the average conductance, one obtains

$$E[\overline{G}] = E[G] \simeq \int dE\, h(E)\, e^{-A}.$$

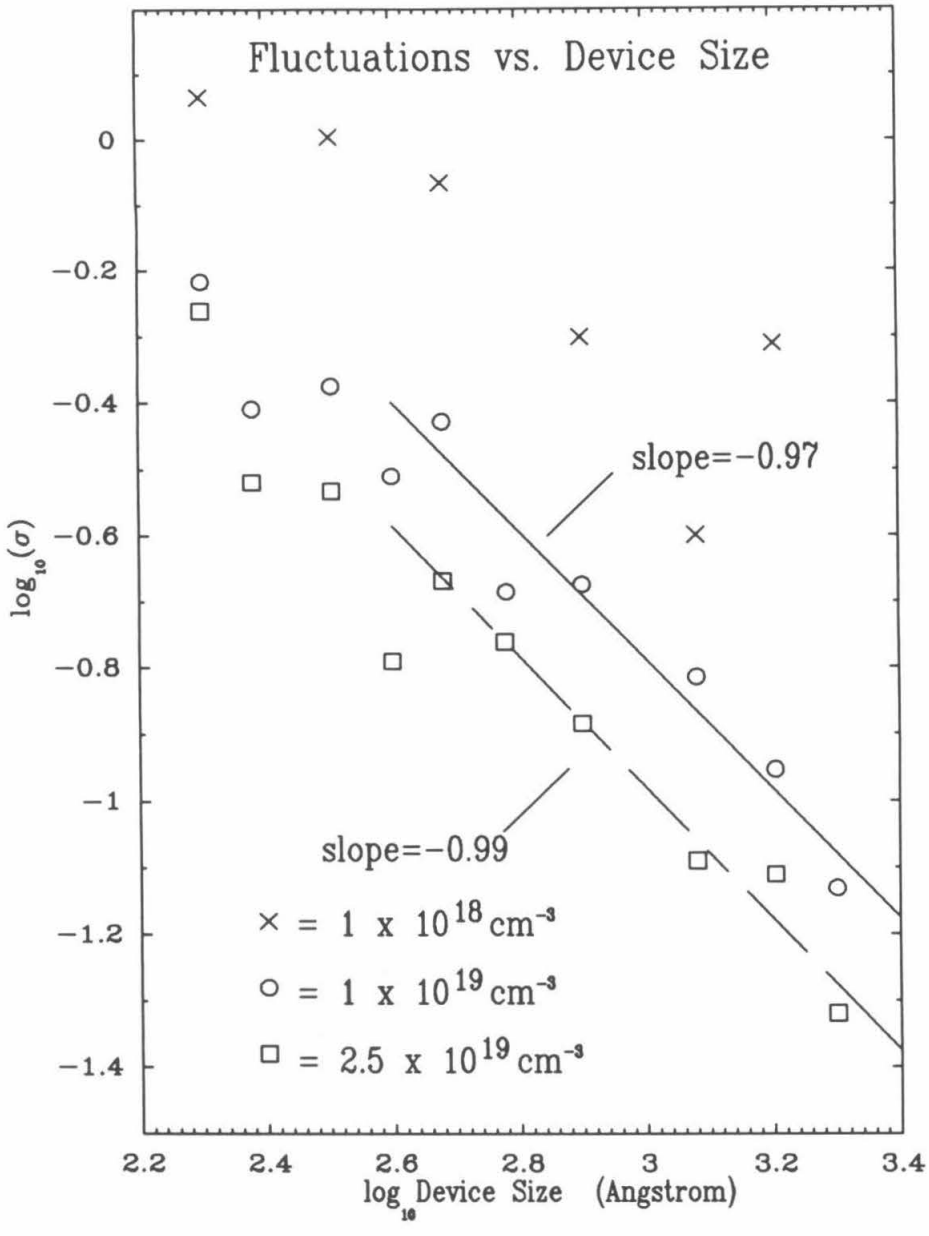

Figure 3.12: The relative fluctuations from Fig. 3.7 versus device size, where the relative fluctuation is the standard deviation divided by the mean. Least-squares lines are fitted to the data at dopings of 10^{19} and $2.5 \times 10^{19} cm^{-3}$, for sizes ≥ 500 Å. The solid line is fitted to the 10^{19} points, and the dashed line is fitted to the 2.5×10^{19} points.

Next, consider the relative fluctuations, defined as

$$\sigma^2 \equiv V[\overline{G}]/E^2[G]. \tag{3.16}$$

From Eq. 3.15 and Eq. 3.2, it can be seen that the $\delta A/\delta\phi$ terms will produce the quantity $2m^*/\hbar^2 E_g$. So for the relative fluctuations the scaling behaviour should be given to a large extent by

$$\sigma^2 \sim \frac{p(1-p)\,R^4 x_d m^*}{m\,\Delta^2\,V_c\,\hbar^2 E_g}\left(\frac{q}{\epsilon}\right)^2.$$

This assumes that the terms contained in the energy integrals in Eq. 3.15 are largely cancelled when the division is made by $E^2[G]$. Indeed, this should be accurate when most of the conductance occurs over a small energy range of the tunnelling electron. This may be seen from Eq. 3.3, where the product of the two factors in the integrand can produce such an effect, at large doping.[1] In this event, the energy integrals in Eq. 3.14 effectively collapse into integrations over a small energy range. With the same effect occurring for $E^2[G]$, the cancellation of the strongly varying terms will occur when σ^2 is found.

Now, recall that $m\,\Delta^2 = l^2$, and $p \propto n$, $p \ll 1$, and $R \propto x_d$. For the depletion length, the usual continuum approximation is used[5] to give Eq. 3.4. This is rewritten below as

$$x_d = \left(\frac{2\epsilon}{nq}\phi_B\right)^{1/2},$$

where $\phi_B = \phi_{B0} + E_{FS}$, ϕ_{B0} is the intrinsic barrier height and E_{FS} is the Fermi level in the semiconductor. E_{FS} depends on the doping, as given by Eqs. 2.6 and 2.7. But since it is dominated by ϕ_{B0}, this dependence can be neglected. Then the key result is arrived at, for the scaling dependence of the variance of fluctuations of a spatially-averaged conductance,

$$\sigma^2 \propto \frac{m^*}{l^2\,\hbar^2 E_g}\sqrt{\frac{\phi_B^5\,\epsilon}{n^3 q}}. \tag{3.17}$$

It can be shown that the above expression is dimensionless. Note that in light of the approximations made in arriving at Eq. 3.17, it should be regarded as a qualitative prediction of the scaling exponents. The main approximation used in obtaining Eq. 3.17 was the neglect of edge effects. Since the relative contribution of the current along the edges as compared to the bulk of the device is $1/l$, this assumption is asymptotically good for large devices.

3.3.3 Comparison of Simulations and Theory

Here, the results of the simulations of Section 3.1 are compared with the scaling predictions of Section 3.2.

To better compare the fluctuations at the different dopings, it helps to normalise the results of Section 3.1 and Fig. 3.9 with respect to the mean. Hence, in Fig. 3.12 the relative fluctuations, σ, found from the simulations are plotted versus the device size, where σ is defined in Eq. 3.16. At dopings of $10^{19} cm^{-3}$ and $2.5 \times 10^{19} cm^{-3}$, it is seen that for cross sections < 500 Å, the relative fluctuations vary widely and in no discernable fashion with cross section. Hence, least-squares lines have been fitted to the log-log data, for cross sections ≥ 500 Å. (Each line is fitted to 6 points.) Thus, at a doping of $10^{19} cm^{-3}$, a slope of -0.97 with a correlation coefficient of -0.96 was found; and at a doping of $2.5 \times 10^{19} cm^{-3}$, the slope was -0.99 and the correlation coefficient was -0.99. This compares well with the $1/l$ power dependence predicted by Eq. 3.17. Furthermore, it is seen that the vertical distance between the two lines is approximately constant, at a value of $\sim 0.19 - 0.20$ over the range of device sizes of 1000 Å to 2000 Å. The doping dependence for σ of of $n^{-3/4}$ predicted by Eq. 3.17 would suggest a vertical distance of 0.3, in reasonable agreement with our simulations.

No attempt has been made to fit a line to the $10^{18} cm^{-3}$ data in Fig. 3.12, as it clearly has considerably greater relative fluctuations, which do not have the

simple dependences on n or l found at the higher dopings. (At least over the cross sections investigated.) Thus, around a doping of $10^{18} cm^{-3}$, Eq. 3.17 breaks down. This is plausible, since to obtain Eq. 3.17, it was assumed that tunnelling dominated the conductance, so that the conductance fluctuations depended primarily on the fluctuations in the WKB exponent. At $10^{18} cm^{-3}$ there is still a considerable thermionic dependence, rendering this assumption invalid.

3.3.4 Remarks

Attempts have been made to quantify the fluctuation dependence of the resistance on cross sectional length, in the plane of the interface, at dopings of $10^{18} cm^{-3}$ and higher. From simulations, the fluctuations have been seen to decrease with increasing cross section. As a rough rule of thumb, for dopings of $10^{19} cm^{-3}$ or greater, the relative fluctuations were found to fall below 10 percent, for devices larger than 1000 Å. The dependences of the relative fluctuations on doping and device size were found to reasonably follow the predictions of Eq. 3.17, at sizes $\geq$ 500 Å, and for dopings $\geq 10^{19} cm^{-3}$.

3.4 Three Dimensional Current Trajectories

Thus far in this chapter, and also in the previous chapter, it has been assumed that the current flows through the depletion region in paths that are normal to the contact. This assumption has also been made in other works involving the potential fluctuations in heterojunctions due to the discreteness of dopants. For example, this was done in a study of the double Schottky barrier at grain boundaries of polycrystalline semiconductors,[9] and in a study of planar-doped barrier transistors.[10] The question arises that if the conductances vary over different regions of a contact, would there be significant three-dimensional current flow through regions of

high conductance? If this were so, then these high conductance regions should dominate the electrical properties of the entire contact. This issue is investigated here.

Section 4.1 describes the method used to find three dimensional trajectories through the depletion region. Section 4.2 gives the results of simulations using this method.

3.4.1 Method

To find the potential fluctuations, the approach of the earlier sections is followed. Dopants are placed randomly at lattice sites in the depletion region and the resultant potential is found in a three dimensional region within the depletion region. We now wish to find the trajectory of an electron from the metal side of the depletion region to the bulk n-GaAs side. The simplest method would be by using classical mechanics. Thus, for the conservation of energy one gets

$$\left(\frac{d\mathbf{r}}{dt}\right)^2 = \frac{2}{m}\Big[E - V(\mathbf{r})\Big], \tag{3.18a}$$

and for the force equation

$$m\frac{d^2\mathbf{r}}{dt^2} = -\nabla V(\mathbf{r}), \tag{3.18b}$$

for a particle of effective mass m and energy E. The problem is that for a given E, there may exist classically forbidden regions for the trajectories, where $E < V(\mathbf{r})$.

To account for tunnelling, the usual procedure previously has been to take one dimensional trajectories normal to the junction and evaluate the attenuation along these trajectories. This is improved upon here by using a semiclassical method recently developed by Das and Mahanty[11] for analysing the tunnelling current in a scanning tunnelling microscope. Within a forbidden region, the time t in the above equations is replaced by an imaginary time $i\tau$, giving

$$\left(\frac{d\mathbf{r}}{d\tau}\right)^2 = \frac{2}{m}\Big[V(\mathbf{r}) - E\Big] \tag{3.19a}$$

and

$$m\frac{d^2\mathbf{r}}{d\tau^2} = \nabla V(\mathbf{r}). \qquad (3.19b)$$

Hence, let a trajectory be given that starts in a classically allowed region. The trajectory is time stepped using Eqs. 3.18. If it reaches the boundary of a forbidden region, then Eqs. 3.19 are used to evolve the trajectory within the forbidden region. (For a more detailed discussion of this procedure, see Ref. 11.)

Having found the trajectory by this method, the WKB approximation in Eq. 3.1 is used to obtain the transmission along the trajectory. As before, the portions of the trajectory that pass through a forbidden region give rise to the attenuation in the transmission. The x direction in Eq. 3.1 now refers to the direction along the trajectory. As was done earlier, the two-band model[4] is used to find the wavevector in the forbidden region, from Eq. 3.2.

Hence given a starting point for the trajectory, and the energy E of the electron, the transmission coefficient can be found. The energy is then varied to find $\{T(E)\}$ over a range of energies. Knowing this, the small-signal resistance at zero bias, R_c, can then be found from Eq. 3.3. Given a metal-semiconductor contact, this procedure can be performed for a grid of starting points of the trajectories over the area of the contact, as was done earlier with the one dimensional trajectories. Using this method, the results of simulations are presented and discussed below.

3.4.2 Simulation Results

An example of trajectories found from a simulation is shown in Fig. 3.13. This depicts 36 trajectories starting at the metal-semiconductor junction. The GaAs is doped to $10^{19} cm^{-3}$. In the diagram, the metal is on the left side and the bulk n-GaAs is on the right side. The trajectories are shown in the depletion region, which is of width $\sim$ 120 Å. The initial spacing between the trajectories is 40 Å. Thus the trajectories span a square contact of side 200 Å. The trajectories are for

a particle energy of 0.6 eV with respect to the zero of energy in the flat part of the conduction band in the bulk n-GaAs. From Fig. 3.13, it is seen that most trajectories are little modified from linear, one dimensional trajectories that are normal to the junction. It should be noted that the trajectories do not necessarily contribute equally to the average conductance of the device. The attenuation along each trajectory determines the weighting of that trajectory's contribution to the total conductance. A trajectory which passes considerably through a forbidden region will be highly attenuated.

Similar results to Fig. 3.13 are found from other simulations. Given a particular dopant configuration, the resultant R_c of a contact as found from one dimensional or three dimensional trajectories differ slightly, as compared to the variation arising from (randomly) changing the dopant configuration.

Within the limitations of the WKB approximation, the simulations suggest that the current through the depletion region of an ohmic contact flows mainly normal to the metal-semiconductor junction.

3.5 Conclusion

In this chapter, the effects of dopant discreteness on the electrical characteristics of an ohmic contact have been investigated. It is found through simulations that the dopant discreteness produces fluctuations in the zero bias resistance of ohmic contacts.

Technologically, this may be of concern, as different resistances of ohmic contacts that are nominally meant to have the same resistance may lead to different voltage drops across the contacts during the operation of a chip containing these contacts. Thus, it was investigated how these fluctuations should scale with the device size l and doping n. A model was developed which predicted that the square

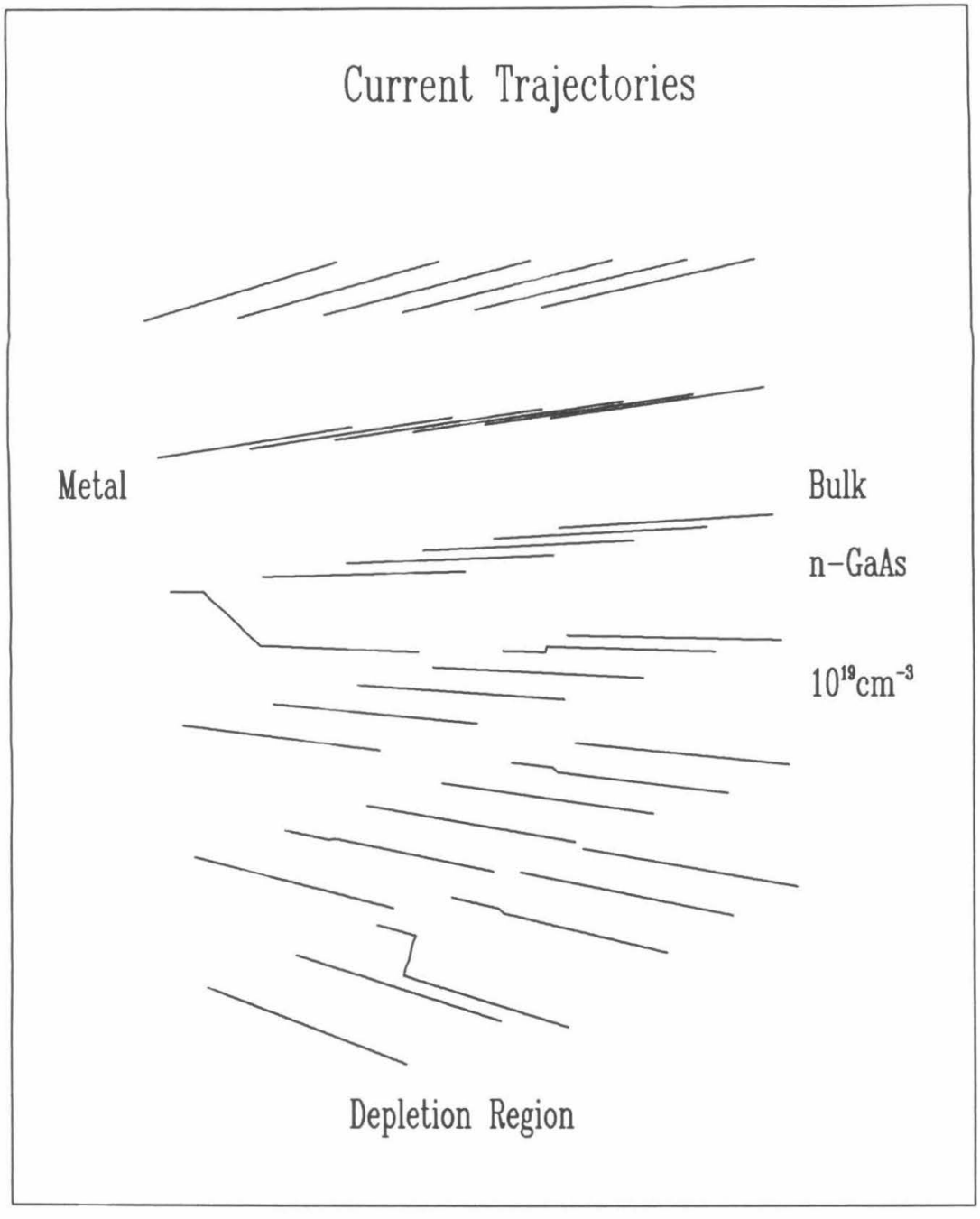

Figure 3.13: An example of three dimensional trajectories through the depletion region of a 200 Å square contact. The trajectories start at the metal-semiconductor junction. The doping is $10^{19}cm-3$. The energy of the electron is $0.6\,eV$.

of the relative resistance fluctuations, σ^2, defined in Eq. 3.16, should scale as

$$\sigma^2 \propto \frac{1}{l^2 n^{3/2}}.$$

This was compared with simulations. Good agreement was found when the GaAs was doped to $10^{19} cm^{-3}$ or higher. The simulations indicate that in this doping range, for contacts of size $l > 1000\,\text{Å}$, the relative fluctuations should fall below 10%. Because of this, it is suggested that the resistance fluctuations at ohmic contacts arising from dopant discreteness should not be a problem for the current GaAs technology.

References

1. C. Y. Chang, Y. K. Fang and S. M. Sze, Sol. St. Elect. **14**, 541 (1971).
2. P. A. Barnes and A. Y. Cho, Appl. Phys. Lett. **33**, 651 (1978).
3. P. D. Kirchner, T. N. Jackson, G. D. Pettit and J. M. Woodall, Appl. Phys. Lett. **47**, 26 (1985).
4. E. O. Kane, *Physics of III–V Compounds* **1**, ch. 3 (Academic Press, New York, 1966).
5. N. W. Ashcroft and N. D. Mermin, *Solid State Physics*, (Holt, Rinehart and Winston, New York, 1976).
6. D. J. BenDaniel and C. B. Duke, Phys. Rev. B **152**, 683 (1966).
7. P. L. Meyer, *Intro. Probability and Statistical Applications*, (Addison-Wesley, Reading, 1977).
8. S. M. Sze, *Physics of Semiconductor Devices*, (Wiley, New York, 1981).
9. G. D. Mahan, J. Appl. Phys. **55**, 980 (1984).
10. D. Arnold and K. Hess, J. Appl. Phys. **61**, 5178 (1987).
11. B. Das and J. Mahanty, Phys. Rev. B **36**, 898 (1987).

Chapter 4

Future Directions

In this chapter I will outline possible further directions of research that are suggested by the work of the two previous chapters. Please note that this discussion is necessarily qualititative and somewhat speculative in nature, pertaining as it does to the future. In Section 4.1, the time dependent Schrodinger equation is suggested for a continued investigation of the properties of an ohmic contact. Section 4.2 discusses other issues that may be pursued. Finally, Section 4.3 gives a conclusion for the chapter.

4.1 Time Dependent Schrodinger Equation

In the earlier chapters, the quantum mechanical scattering of the electron current through the depletion region was treated by the WKB approximation of Eq. 3.1 and the two band model of Eq. 3.2. When a generalisation of the current trajectories from one dimension to three dimensions was made in section 3.4, the WKB approximation was retained. It has the computational virtue of giving the transmission coefficient for scattering through a potential as an explicit and easy to calculate dependence on the potential. However, if one goes through the derivation[1] of the WKB approximation, it is found to apply best when the potential varies slowly compared to the wavelength of the particle. Consider the example of potential that was shown in Fig. 3.3. The potential is given over a two dimensional region of the depletion region, of approximate size $100\,\text{Å} \times 200\,\text{Å}$. The potential varies significantly over distances of $\sim 20\,\text{Å}$ or more. By comparison, an electron with an energy of $0.5\,eV$ with respect to the flat part of the conduction band in the bulk GaAs would have a wavelength $\sim 70\,\text{Å}$. Thus, the results of using the WKB approximation should be regarded as semiquantitative.[2]

Another problem also arises from the use of the two band model in the earlier chapters. This is obtained from the $\mathbf{k} \cdot \mathbf{p}$ method which assumes that the electron

is moving in a periodic potential. Hence, the time independent wavefunction ψ of the electron can be written using Bloch's theorem as

$$\psi(\mathbf{x}) = e^{i\mathbf{k}.\mathbf{x}}\, u_{n\mathbf{k}}(\mathbf{x}),$$

where $u_{n\mathbf{k}}$ has the periodicity of the potential. Since the potential in the depletion region lacks any periodicity, the use of the $\mathbf{k}\cdot\mathbf{p}$ method may be considered somewhat approximate.

How can this be improved? The alternative to the above semiclassical treatment is a full (non-relativistic) quantum mechanical treatment of the time *dependent* Schrodinger equation in the depletion region,

$$i\hbar\frac{\partial\psi}{\partial t} = H\psi = \left[-\frac{\hbar^2}{2m^*}\nabla^2 + \phi\right]\psi. \tag{4.1}$$

Here ϕ is the potential in the depletion region and the effective mass approximation is used to give m^*. The traditional approach to this problem has been to search for stationary states.[1,3] That is, a basis set of time independent wavefunctions $\{\theta_n\}$ is sought that satisfies Eq. 4.1 with the LHS replaced by $E_n\theta_n$. Having (presumably) found the $\{\theta_n\}$, the time dependent function ψ is then expanded in terms of these as

$$\psi(\mathbf{x},t) = \sum_n c_n\,\theta_n(\mathbf{x})\,exp(-iE_nt/\hbar). \tag{4.2}$$

However, the instances where $\{\theta_n\}$ can be found analytically are scarce.[1,3] If computers are available, then the idea of searching for stationary states is very useful and highly developed. For example, this approach is used in quantum chemistry to solve for the bonding in small molecules,[4,5] where the potential in such instances can typically be described by a few parameters. Also, the potential usually exhibits some spatial symmetries which enable group theoretic methods to be used to simplify the problem.[6] The resultant infinite sum over the basis states in Eq. 4.2 can be well approximated by a finite sum over a few of the basis functions. In

the case at hand, the problem is more difficult. Consider again the example of the potential in Fig. 3.3. Due to the random placement of dopants, the potential along trajectories normal to the junction clearly cannot be described in terms of a few parameters. The potential contains no spatial symmetries with which to simplify Schrodinger's equation. Hence, it is highly unlikely that a useful (i.e., small) set of stationary states can be found with which to expand a time dependent solution.

The only remaining possibility of solving Eq. 4.1 is to use a explicit time stepping method. A promising approach is one suggested by D. Kosloff and R. Kosloff.[7] They proposed a fast Fourier transform (FFT) method that can be applied from one to three spatial dimensions. The method is as follows.

Consider Eq. 4.1. A wavefunction is given at $t = 0$. The potential ϕ is specified over the spatial domain of interest. Let ψ_n be the wavefunction after the n-th time step, and let δt be the time interval between time steps. The time derivative in Eq. 4.1 is approximated by the second order finite difference,

$$\frac{\partial \psi_n}{\partial t} \simeq \frac{\psi_{n+1} - \psi_{n-1}}{2\,\delta t}. \tag{4.3}$$

Now the RHS of Eq. 4.1, $H\psi_n$, needs to be found. The effect of the potential is easy to compute: $\phi(\mathbf{x})\psi_n(\mathbf{x})$ is a local operation. But ∇^2 is a nonlocal operator. The usual approach has been to replace it by a finite spatial difference.[8,9] The problem is the resultant loss of significant figures. To compensate for this, a high density of grid points is needed. Unfortunately, the computational effort correspondingly increases, especially with the number of spatial dimensions under consideration. A better method is to perform an FFT on $\psi_n(\mathbf{x})$ to give $\psi_n(\mathbf{k})$ in $\mathbf{k}$ space. In this space ∇^2 transforms to a local operator, $-k^2$. The term $-\nabla^2\psi_n$ is then evaluated as $k^2\psi_n(\mathbf{k})$ over all $\mathbf{k}$ space. An inverse FFT is performed on the result to move it back into real space, where it is added to $\phi(\mathbf{x})\psi_n(\mathbf{x})$. Thus, Eq. 4.1 is replaced by

$$\frac{i\hbar}{2\,\delta t}\Big[\psi_{n+1}(\mathbf{x}) - \psi_{n-1}(\mathbf{x})\Big] = \frac{\hbar^2}{2m^*} F^{-1}\Big\{k^2\, F\{\psi_n(\mathbf{x})\}\Big\} + \phi(\mathbf{x})\psi_n(\mathbf{x}), \tag{4.4}$$

where F designates an FFT and F^{-1} is the inverse FFT. This gives the wavefunction at time $n+1$ in terms of the wavefunctions at times n and $n-1$. Hence, it can be used to time evolve the wavefunction through the potential ϕ.

It may be appreciated that this method has several strong features. It is linear, elegant and simple to understand. The method can be applied easily in one, two or three spatial dimensions. The FFT is an extremely efficient algorithm for computing a Fourier transform, with source code in Fortran and Pascal being readily available.[10] The method was found by Kosloff and Kosloff to have the advantage of being more accurate than the spatial finite differencing approach, given the same computational effort. Using the FFT method, the energy and norm of the initial wavefunction can be conserved, serving as checks on the calculations. They applied this method to the study of the hydrogen reaction in two dimensions $H^+ + H_2$.[11]

I wish to suggest that the FFT method may be applied to the three dimensional problem of scattering of current through the depletion region. A preliminary three dimensional scattering calculation was performed on a Vax 11-785. The intention was to find an estimate of the computational effort required for such a problem. The initial wavefunction for the electron free carrier was taken to be a three dimensional Gaussian function, with a standard deviation of 8 Å. The particle was given an energy of $0.6\,eV$ with respect to the zero of energy in the flat part of the conduction band in the bulk n-GaAs. The particle was started from the bulk n-GaAs and it moved towards the depletion region. The depletion length was 70 Å, corresponding to a doping of $4 \times 10^{19} cm^{-3}$. The space consisted of $(256, 8, 8)$ points along the (x,y,z) directions. Eq. 4.4 was used for 1200 time steps to evolve the wavefunction, after which time most of the wavefunction had reached the metal. The CPU time taken was 1 day 20 hours.

In principle, a calculation similar to the above could be used to find a trans-

mission coefficient $T(E)$ for a particle of energy E incident on a given potential. To find the conductance (as in Eq. 3.3) one needs $T(E)$ over a range of energies. It should also be said that the above example was insufficient to give a $T(E)$. More points need to be added in the y and z directions to properly handle the three dimensional scattering before the transmission coefficient can be estimated. Hence, the above example indicates that a computer substantially more powerful than a Vax 11-785 is needed to attack the three dimensional scattering through a depletion region.

How feasible is this? Within the past three years, there have been minisupercomputers available that offer a speedup of an order of magnitude or better over a Vax 11-785.[12] Let me be specific. For the purposes of benchmarking, a one dimensional implementation of Eq. 4.4 was run on several machines. On the Vax, the CPU time taken was 3300 seconds. By comparison, the Alliant FX/8 took 143 seconds, the Convex C-1 took 92 seconds and the Vax 8700 took 894 seconds. Most impressive of those surveyed, however, was the SCS machine, which performed the computation in 82 seconds. These runtimes show that close to two orders of magnitude in speedup is achievable on some machines. In addition, two new superminicomputers made by Ardent Computer and Stellar Computer promise even faster performance than those machines tested here. (At the time of writing, no runtimes are available.)

Furthermore, given the FFT method, one new computer architecture offers an intriguing potential for high performance. It is the 'Hypercube' geometry which has been proposed and implemented by Fox and Seitz.[13,14] (Commercially, the Intel iPSC/2 VX and the Ametek Series 2010 computers are Hypercube machines.) A computer with this geometry has several processors, where the number of processors is a power of two and the processors are linked to each other as vertices of a cube in Boolean n-space. It so happens that this is a hard-wired FFT machine.

That is, the computations and the internal data flow of an FFT can be shown to map directly onto a Hypercube.[13,15] Hence, in Eq. 4.4, the FFT can be done extremely efficiently on a Hypercube. In addition, the $\phi(\mathbf{x})\psi_n(\mathbf{x})$ dot product term is trivial to compute efficiently. This sum of products can be distributed equally amongst the processors, with the only communication being needed when each processor has computed its subsum. Unfortunately, a runtime for the one dimensional FFT program is unavailable for a Hypercube. But this geometry does offer promising prospects for time stepping Eq. 4.4.

4.2 Other Research

If the algorithm of Eq. 4.4 for solving Schrodinger's equation can indeed be implemented usefully on a new machine then further avenues of investigation arise. The most immediate would be the question of 'hole-burning' that was raised in the previous chapter. That is, with regions of varying conductance in a contact, would current flow preferentially through the regions of high conductance? Hence, would the conductance of these regions then characterise the conductance of the contact? In Chapter 3, the semiclassical generalisation of Newton's equations and the WKB method were used to give a rough 'no' to this. The implementing of Eq. 4.4 would give an important improvement over this. One may look at the 'trajectories' of the wavefunctions through the depletion region and see if these indeed tend to go through regions of high conductance.

Another issue of interest is a multiparticle formalism for the current. As stated in Eq. 4.1, the Schrodinger equation is written assuming an independent particle model for the free carrier. Only the interaction between a single free carrier and the dopants in the depletion region is considered. This may give a valid approximation to the current for low current density. But at some current density, this surely

breaks down. Space charge effects between the free carriers may set an upper limit on current through any 'hot spots' in the potential. Screening effects may arise, when an electron travelling close to a positively charged dopant neutralises its effect on other electrons in the current. This could have the consequence of producing a time varying noise source to the current. If dopants are screened and unscreened by some carriers within the current, the potential seen by the rest of the current will have a (random) time dependence. This may be considered to be an intrinsic noise of the ohmic contact. Thus, by extending Eq. 4.1 to handle several particles in the current, it should be possible to study these phenomena.

4.3 Conclusion

As a continuation of the research of Chapters 2 and 3, I wish to propose solving the three dimensional time dependent Schrodinger equation for the scattering of current through the depletion region of an ohmic contact. The scattering is due to the discreteness and the random positions of the dopants in the depletion region. The FFT method of Eq. 4.4 is suggested as a means of doing this. It could be used to find the transmission coefficient of a free carrier through the depletion region. Thus, by comparing this to the transmission given by the WKB approximation, one can check the validity of the latter. This would be an important extension of the earlier chapters, which were predicated on the use of the WKB approximation to give the transmission through the depletion region.

The performance of new computers should make the FFT method of Eq. 4.4 a feasible one to use. Also, by using an explicit time dependent method, there exists the possibility of studying any inherently time dependent aspects of the noise of an ohmic contact.

References

1. L. I. Schiff, *Quantum Mechanics*, (McGraw-Hill, Tokyo, 1968).
2. B. Das and J. Mahanty, Phys. Rev. B **36**, 898 (1987).
3. G. Baym, *Lectures on Quantum Mechanics*, (Benjamin, Reading, 1977).
4. J. N. Murrell, S. F. A. Kettle and J. M. Tedder, *Valence Theory*, (John Wiley, London, 1965).
5. A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry*, (Macmillan, New York, 1982).
6. E. P. Wigner, *Group Theory*, (Academic Press, New York, 1959).
7. D. Kosloff and R. Kosloff, J. Comp. Phys. **52**, 35 (1983).
8. A. Askar, J. Chem. Phys. **74**, 6133 (1981).
9. A. Goldberg, H. Schey and J. Schwartz, Am. J. Phys. **35**, 177 (1967).
10. W. H. Press, B. P. Flannery, S. A. Teukolsky and W. T. Vetterling, *Numerical Recipes*, (Cambridge University Press, New York, 1986).
11. R. Kosloff and D. Kosloff, J. Chem. Phys. **79**, 1823 (1983).
12. T. Manuel, Electronics **61**, 51 (3 March 1988).
13. G. C. Fox and S. W. Otto, Phys. Today **37**, 50 (May, 1984).
14. C. L. Seitz, J. VLSI and Comp. Sys. **1**, 311 (1986).
15. M. O. Newton, *Analysis of a Parallel Implementation of the FFT*, (Caltech Computer Science internal document, 1982).

Chapter 5

An Efficient Calculation of Barrier Height Fluctuations in Planar-Doped Transistors

5.1 Introduction

Advances in semiconductor device fabrication have made possible the construction of small structures where the mean free path of charge carriers is comparable to the device dimensions. One device which is expected to show such ballistic transport is the planar-doped barrier transistor (PDBT) proposed by Shur and Eastman.[1] This device structure is a majority carrier device consisting of an n+ emitter, base and collector. Fig. 5.1 shows a schematic of the transistor. The base is separated from both the emitter and collector by a barrier structure consisting of an intrinsic layer, a thin p+ region and an intrinsic layer. These barriers act as the controls between the emitter-base and base-collector. In addition to the device applications, the structure can be used to analyse the momentum and energy distribution of carriers emitted over the barrier.

In a typical usage, electrons from the emitter cross the first barrier to reach the base, with some energy and momentum distribution. By varying the base-collector bias, the barrier height in the collector p+ region can be altered. This acts as an analyser of the electron momentum distribution, as the electrons travel from the base to the collector by thermionic transmission over the collector p+ region. Levi *et al*[2] attributed part of the broadness in the resultant measured distribution to be due to fluctuations in the potential in the barrier region caused by the random distribution of acceptors. Using a Markov averaging method, Mahan[3] investigated similar barrier fluctuations at grain boundaries in polycrystalline semiconductors. Recently, Arnold and Hess[4] have investigated this effect of statistical fluctuations in the characteristics of small GaAs PDBTs. They studied through simulations the fluctuations in barrier height due to the randomness in the distribution of acceptors in the barrier region. They found the barrier height variations by solving Poisson's equation in three dimensions, for a given random distribution of acceptors.

In this chapter, a simple statistical treatment of the barrier height fluctuations

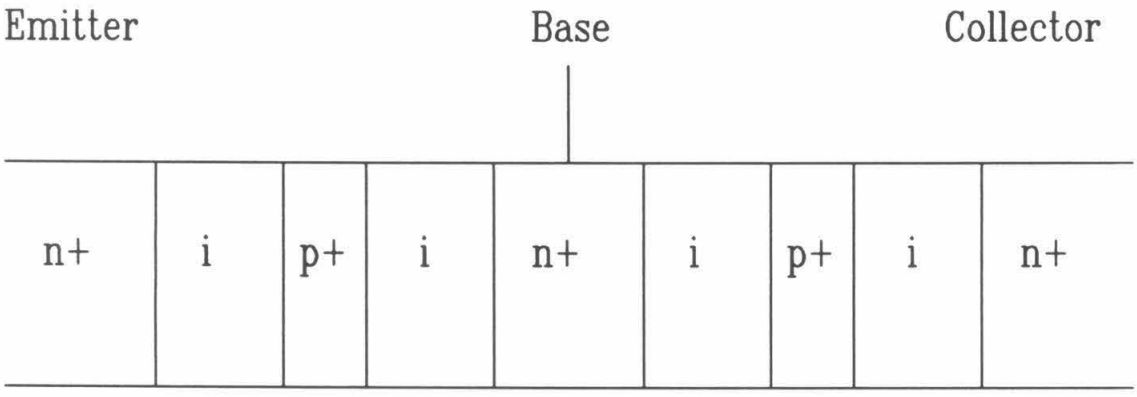

Figure 5.1: A schematic of the planar-doped barrier transistor.

in one barrier region is presented. The treatment is analytic. I will show how the fluctuations depend on the doping and the thickness of the region. In Section 5.2, the theory is described. The results are presented and a comparison is made with the work of Arnold and Hess. in Section 5.3. Section 5.4 contains the conclusion for the chapter.

5.2 Theory

The method is outlined as follows. First, the variance in the potential is derived along a trajectory that is normal to the barrier region, as a function of dopants close to the trajectory. Then, the contribution to the variance from the other dopants in the barrier region is added. Next, under suitable approximations, the probability distribution for a spatial average of barrier heights can be found.

The planar-doped barrier transistor of Ref. 4 included an n+ layer doped to $10^{18} cm^{-3}$, a 250 Å intrinsic layer, a 100 Å p+ barrier layer doped to $10^{18} cm^{-3}$, another 250 Å intrinsic layer, and an n+ layer doped to $10^{18} cm^{-3}$. One n+ layer may be taken as the base, and the other as either the emitter or collector electrode. Note that the total length of the intrinsic/p+/intrinsic sandwich, 600 Å, is large enough to effectively preclude transport by tunnelling. Thus, as mentioned above, the current is considered to consist of only thermionic transmission over the barrier region.

Arnold and Hess modelled the acceptor potential by replacing the $1/r$ potential by a constant, ϕ_0, when $r < R$. R is an approximate measure of the radius of the ionised acceptor. Hence, it is roughly equal to the Bohr radius of the captured electron. This was based on the pseudopotential treatment by Ning and Sah,[5] though they considered ionised donors. For GaAs, one has $R \sim 100$ Å. If the p+ layer thickness, l, is less than $2R$, then one has the situation depicted in Fig. 5.2.

Here, the x direction is normal to the p+ layer, with y and z being in the plane of the layer. A trajectory and a point $(x,0,0)$ on the trajectory are depicted in the figure. To find the contribution to the potential at this point, consider first acceptors that are within a sphere of radius R, centred on $(x,0,0)$. Let $V(x)$ be the volume of intersection between the sphere and the boundaries of the p+ layer: $x = 0, l$. It can be shown that

$$V(x) = \frac{\pi l}{3}(3R^2 - l^2 + lx - x^2). \tag{5.1}$$

As expected, the maximum volume occurs when $x = l/2$.

Within the p+ layer, the acceptors are assumed to be distributed randomly at lattice sites. Let V_c be the volume per lattice site, and p_0 be the probability that a lattice site is occupied by an acceptor. So $p = p_0/V_c$, where p is the doping in the p+ layer.

Let $\phi_{in}(x)$ be the potential due to acceptors inside $V(x)$. Note that $p_0 \ll 1$ and that usually the total number of lattice sites inside $V(x)$ is $V(x)/V_c \gg 1$. Hence, the number of acceptors inside $V(x)$, call it n, is a random variable with a Poisson distribution.[6] Since each acceptor in $V(x)$ contributes ϕ_0 to $\phi_{in}(x)$, it is seen that ϕ_{in} also has the same Poisson distribution, with the probability that $\phi_{in} = n\phi_0$ being given by

$$P(\phi_{in}(x) = n\phi_0) = \frac{\alpha(x)^n\, e^{-\alpha(x)}}{n!}, \tag{5.2a}$$

where

$$\alpha(x) \equiv p\, V(x). \tag{5.2b}$$

From the properties of the Poisson distribution, one finds for the mean and the variance,

$$E[\phi_{in}(x)/\phi_0] = V[\phi_{in}(x)/\phi_0] = \alpha(x). \tag{5.3}$$

So Eqs. 5.1 and 5.3 indicate that the maximum mean and variance of ϕ_{in} occur at $x = l/2$.

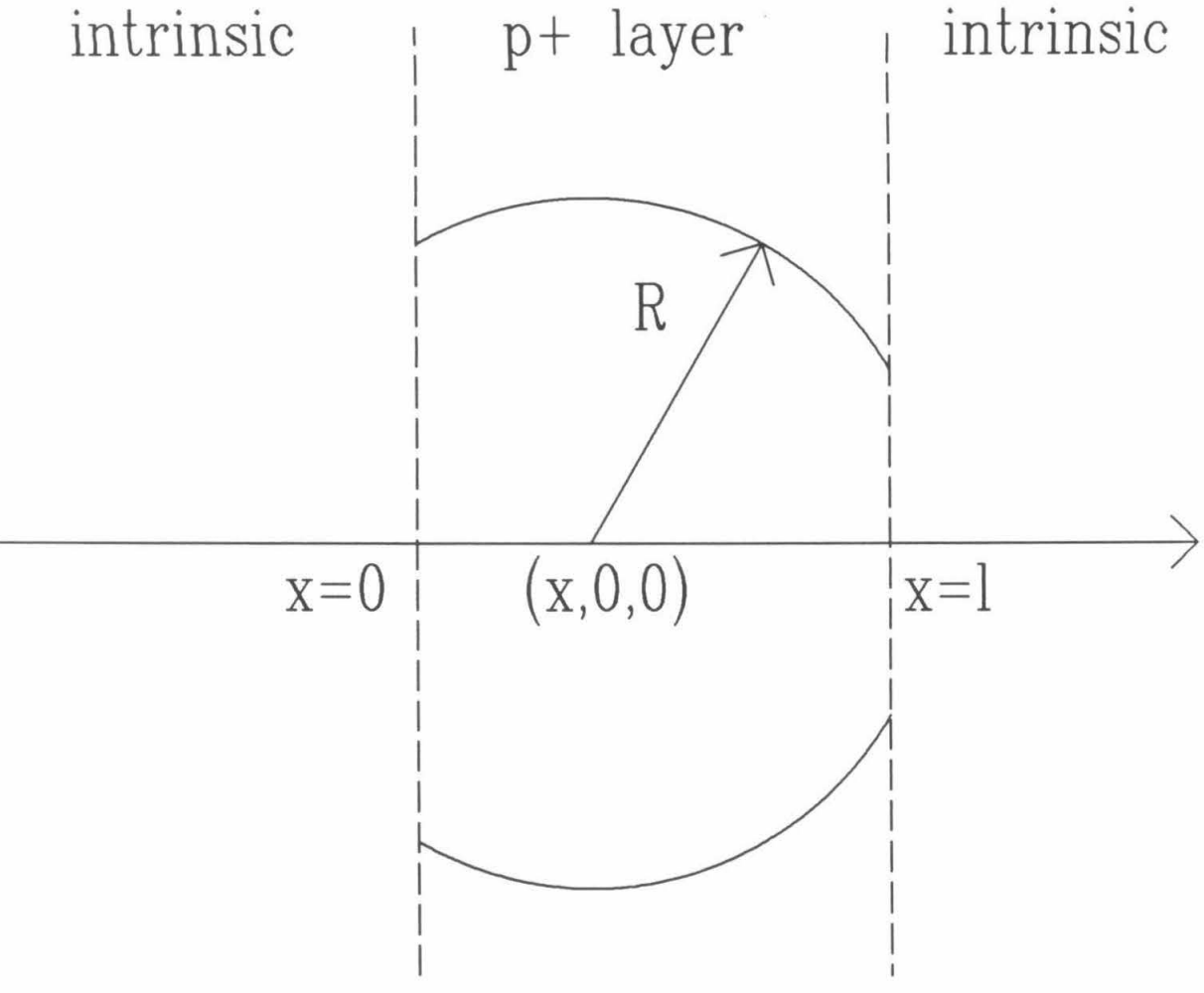

Figure 5.2: The p+ layer between 2 intrinsic regions. The contribution to the potential at (x,0,0) is found from acceptors inside the sphere of radius R, and from those in the p+ layer and outside the sphere. l is the layer thickness.

I now wish to find the contribution to the variance from acceptors outside the sphere of radius R. The Thomas-Fermi screening of the acceptors is used to change the $1/r$ potential to the Yukawa form $exp(-k_0 r)/r$ for $r > R$. Let $\phi_{out}(x)$ be due to acceptors outside $V(x)$. Consider now a sphere of radius $r > R$, centred on $(x, 0, 0)$. The surface area of the sphere, subject to the constraints $0 < x < l$, is given by

$$A = 2\pi l r.$$

The acceptors in a shell of thickness dr on the surface of the sphere contribute equally to ϕ_{out}. And shells of different $r > R$ contain statistically independent numbers of acceptors. Thus, the variance of ϕ_{out} is given by

$$\begin{aligned} V[\phi_{out}(x)] &= \int_R^\infty dr\, A \frac{p_0}{V_c} \left(\frac{q}{4\pi\epsilon}\right)^2 \frac{e^{-2k_0 r}}{r^2} \\ &= \left(\frac{q}{4\pi\epsilon}\right)^2 2\pi l p \int_1^\infty \frac{dt}{t} e^{-2k_0 R t}. \end{aligned} \tag{5.4}$$

Note that this is independent of x. Similarly, one can show that the average value of ϕ_{out} is independent of x. From Eqs. 5.3 and 5.4, the total variance of the potential in the middle of the p+ layer is then found to be

$$V[\phi(x)] = p\,\phi_0^2\, V(x) + \left(\frac{q}{4\pi\epsilon}\right)^2 2\pi\, l\, p \int_1^\infty \frac{dt}{t} e^{-2k_0 R t}. \tag{5.5}$$

Similarly, the mean of ϕ is given by

$$E[\phi(x)] = p\,\phi_0\, V(x) + \frac{q}{4\pi\epsilon} \frac{2\pi l p}{k_0} e^{-k_0 R}. \tag{5.6}$$

For thermionic emission, the quantity of interest is the maximum in ϕ along a trajectory. This is the barrier height. Thus, considering Eq. 5.5 and the comments following Eq. 5.3, the assumption is made that the variance in the barrier height can be approximated by the variance in ϕ at mid-barrier. This assumption should be good since Eq. 5.6 shows that the largest average value of ϕ is at $x = l/2$. For no fluctuations, this would be the barrier height. In the presence of fluctuations,

it may be supposed that $\phi(l/2)$ will still tend to be the maximum in ϕ and hence define the barrier height.

Let us now evaluate Eq. 5.5 and 5.6. Following Ref. 4, the following choice of parameters is made: $l = 100\,\text{Å}$, $R = 100\,\text{Å}$ and $p = 10^{18} cm^{-3}$. Let $\phi_0 = q/4\pi\epsilon R$. For the Yukawa parameter, k_0, one can use[7]

$$k_0 = \left(\frac{4\,m^*}{a_0 m_0}\right)^{1/2} \left(\frac{3\,p}{\pi}\right)^{1/6} \tag{5.7}$$

where m^* is the effective mass and a_0 is the Bohr radius of hydrogen.

It is found that $V[\phi_{out}] = 6 \times 10^{-11} eV^2$ and $V[\phi_{in}] = 3.4 \times 10^{-4} eV^2$, so $\sqrt{V}[\phi(l/2)] = 18meV$. Likewise, one also finds that $E[\phi_{out}] = 0.01meV$ and $E[\phi_{in}(l/2)] = 28meV$. Thus, the contributions to the mean and the variance of the potential are dominated by acceptors close to the trajectory: $V[\phi_{in}(l/2)] \gg V[\phi_{out}]$ and $E[\phi_{in}(l/2)] \gg E[\phi_{out}]$. This is due to the heavy doping of $p = 10^{18} cm^{-3}$, which increases the screening in the Yukawa potential via Eq. 5.7. So whenever the screening causes the second terms of Eqs. 5.5 and 5.6 to be negligible compared to the first terms, ϕ at mid-barrier, and the maximum in ϕ, have approximately a Poisson distribution.

Suppose now it is desired to find the variance in the barrier heights averaged over a cross sectional area B. Let $s \equiv \sqrt{B}$ be the typical linear extent of the area. If $s > 2R$, then trajectories separated by a distance s or more may be considered to be approximately independent, due to the rapid falloff of the Yukawa potential. Let $\overline{\phi_B}$ be the barrier height averaged over an area B. Then for $s > 2R$, the variance of $\overline{\phi_B}$ is roughly

$$V[\overline{\phi_B}] \sim \frac{4\,R^2\,V[\phi(l/2)]}{B}. \tag{5.8}$$

5.3 Results

In this section, the results of the above model are compared with the probability distribution found in Ref. 4 for a particular simulation.

From the remarks above, one can consider $\phi(l/2)$ as having a Poisson distribution. Let $\overline{\phi_B}$ be the average of n independent $\phi(l/2)$ random variables. It can be shown[6] from an elementary statistical argument that the probability distribution of $\overline{\phi_B}$ is given by

$$P\left(\overline{\phi_B} = \frac{m}{n}\phi_0\right) = \frac{[n\alpha(l/2)]^m\, e^{-n\alpha(l/2)}}{m!}, \tag{5.9}$$

where m is a non-negative integer, l is the thickness of the acceptor layer and α is given by Eq. 5.2b. Following Eq. 5.8, let $n \simeq B/4R^2$. Note the close similarity between Eq. 5.9 and the Poisson distribution of Eq. 5.2a.

The probability distribution of Eq. 5.9 is plotted in Fig. 5.3, along with the probability distribution of barrier heights from Fig. 2 of Ref. 4. This latter distribution was for one particular simulation. To facilitate comparison, two things have been done. First, the mean of the distribution is equated to that of Ref. 4, which was $206\,meV$. This amounts to choosing a common zero of energy. The example of Ref. 4 is followed in taking the zero of energy to be in the flat portion of the conduction band in the n+ regions, assuming zero bias. Note that the resultant mean of the distributions shown in Fig. 5.3 is to some extent arbitrary. By varying the doping in the n+ regions and the widths of the intrinsic layers and then solving Poisson's equation, the mean barrier height can be altered (even under zero bias). Second, the number of $\phi(l/2)$ random variables to be averaged over was chosen in order to equate the variances of the two distributions in Fig. 5.3. This resulted in $n = 11$ in Eq. 5.9. From Fig. 5.3 it can be seen that the fit is excellent in view of the number of approximations that been made. The choice of $n = 11$ deserves comment. From Eq. 5.8, this corresponds to an area of at

least $B \simeq 5 \times 10^5 \text{ Å}^2$. Alternatively, if the barrier heights are found only along one direction in the plane of the barrier, then $n = 11$ corresponds to a transverse distance of at least 2200 Å. This latter estimate is consistent with the data in Fig. 1 of Ref. 4. Comparison is awkward since the dimensions of B are not specified in Ref. 4 for the distribution that is shown in its Fig. 2. Hence, the good agreement in Fig. 5.3 may be fortuitous.

5.4 Conclusion

In summary, it has been shown that barrier height fluctuations in a planar-doped transistor may be approximated by the fluctuations in the potential in the centre of the barrier layer. To find the fluctuations averaged over an area, one obtains the Poisson-like distribution of Eq. 5.9. This distribution is computationally very easy to evaluate. By using this instead of three dimensional finite difference simulations, considerable computational effort may be averted. The method is also simpler than the numerical Markov averaging approach of Ref. 3. Furthermore, Eq. 5.9 has the utility of explicitly containing the dependence of the barrier height distribution on the doping and thickness of the barrier layer of the transistor. This was not given in Ref. 4. Thus, I suggest that the results offer an improvement and an extension over those in Ref. 4 for the planar-doped barrier transistor. This may be useful for those designing such devices.

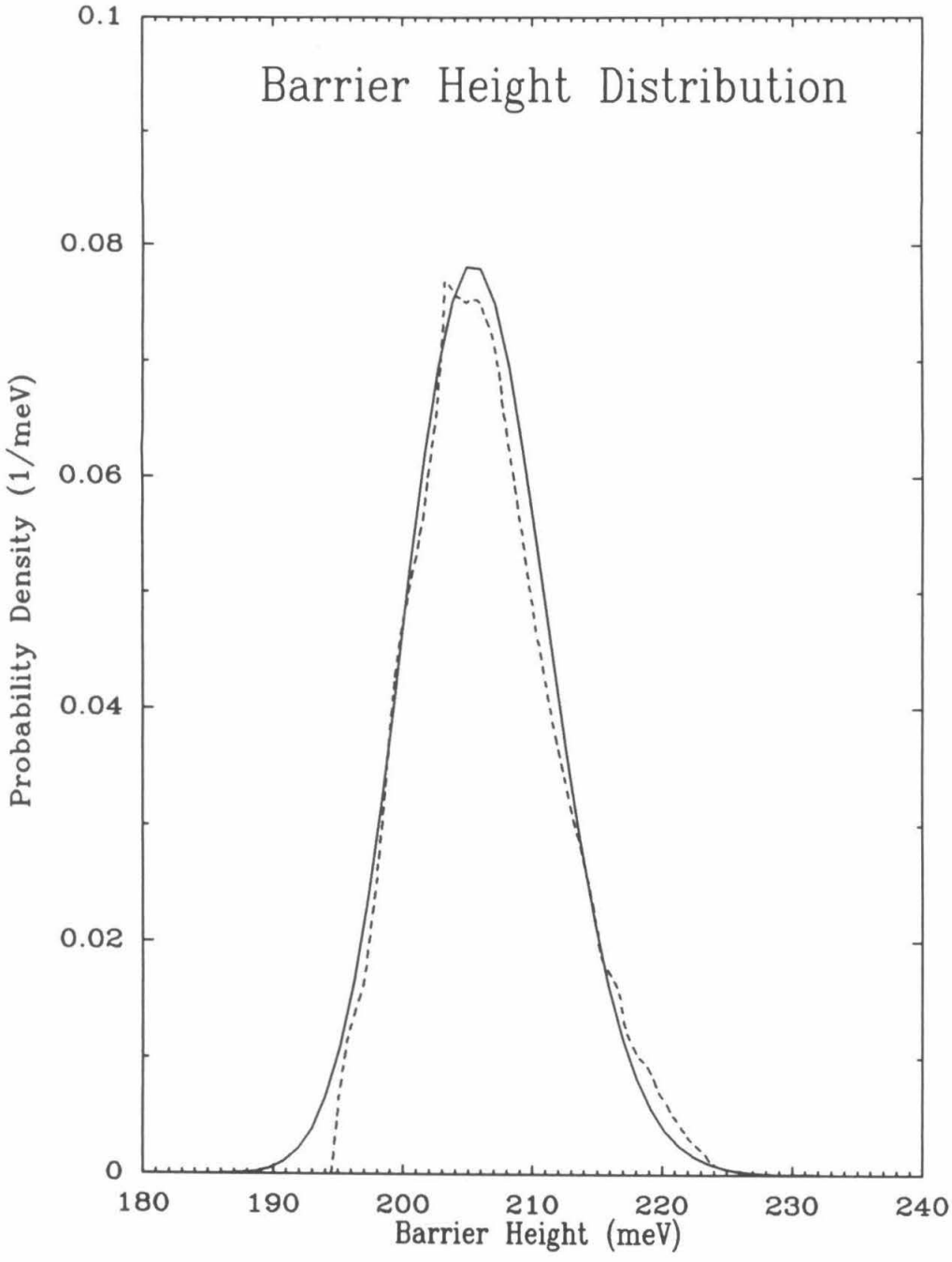

Figure 5.3: Barrier height probability distributions. The solid line is from the discrete probability distribution of Eq. 5.9 in the text. The dashed line is from Fig. 2 of Ref. 3. The layer thickness $l = 100\,\text{Å}$ and the doping $p = 10^{18} cm^{-3}$.

References

1. M. S. Shur and L. F. Eastman, IEEE Trans. Elec. Dev. **ED26**, 1677 (1979).
2. A. F. J. Levi, J. R. Hayes, P. M. Platzman and W. Weigmann, Phys. Rev. Lett. **55**, 2071 (1985).
3. G. D. Mahan, J. Appl. Phys. **55**, 980 (1984).
4. D. Arnold and K. Hess, J. Appl. Phys. **61**, 5178 (1987).
5. T. H. Ning and C. T. Sah, Phys. Rev. B **4**, 3468 (1971).
6. P. L. Meyer, *Intro. Probability and Statistical Applications*, (Addison-Wesley, Reading, 1977).
7. N. W. Ashcroft and N. D. Mermin, *Solid State Physics*, (Holt Rinehart and Winston, New York, 1976).

Chapter 6

Finite-Size Effects in Two Dimensional Continuum Percolation

6.1 Introduction

In recent years there has been interest in the electrical properties of composite materials consisting of conducting fibres or sticks, embedded in an insulating polymer matrix. Typically, the composites are prepared by mixing the fibres and polymer in a liquid state and letting the result cool while it is being poured. This leads to an angular distribution of the fibres in the solid state, about the flow direction of the pour. Experimentally, such composites are found to have a threshold dependence of the electrical conductivity on the fibre length.[1] Various models have been tried to explain such behaviour. These include the effective medium theory.[2] This has not been very successful as there is a large difference between the conductivities of the fibres and the insulating polymer.[3] More promising are percolation simulations. In this chapter, the onset of percolation will be studied in two dimensions. Simulations will be performed and a simple theory will be developed to explain the percolation behaviour.

The first percolation study was done by Pike and Seager.[4] They considered widthless (one dimensional) sticks of constant length, in a two dimensional medium, isotropic on a macroscopic scale. They investigated the case of continuum percolation, where the objects that percolate are placed uniformly in the medium, as distinct from percolation on a lattice,[5] where the objects are constrained to lie on lattice points. Pike and Seager's work was considerably extended by Balberg and Binenbaum[6] to the case of anisotropic systems where, as mentioned above, the fibres have an average preferred orientation. They also considered various distributions of stick length. From their simulations, the critical stick lengths for percolation along the average stick orientation and transverse to this were found as functions of anisotropy (defined below in Section 6.2) and the number of sticks in the sample. However, due to computing constraints, they only considered a few random configurations of sticks. They showed that in the limit of an infinite

ensemble of sticks, the longitudinal and transverse percolation thresholds converge to a common function. Here, I will extend their work by considering more thoroughly the finite size effects on the percolation thresholds of ensembles with small numbers of sticks. In Ref. 6 and other papers[7,8] on continuum percolation this appears to have been a neglected point of interest. A simple model will be derived for the dependences of the thresholds on anisotropy and numbers of sticks and compare this to the results of extensive simulations.

This chapter is arranged as follows. Section 6.2 describes the computer simulations. Section 6.3 contains the model of the percolation thresholds. Section 6.4 compares the results of the model with those of the simulations. Finally, a conclusion is given in Section 6.5.

6.2 Simulations

In this section, a method is outlined for performing simulations to find the critical lengths.[6] First, the basic terms shall be defined. Then, the procedure for obtaining the critical lengths will be given.

To start, let me define some quantities. Consider a set of N widthless sticks. The centres of the sticks are placed uniformly in the unit square in the plane, $[0,1] \times [0,1]$. The sticks have some angular probability distribution $f(\theta)$ about the y-axis, where θ is the angle between a stick and the y direction. The angular distributions that will be considered here all have the feature that

$$f(\theta) = f(-\theta), \quad \theta \in (0, \pi/2) \tag{6.1}$$

and that the average angle is $E[\theta] = 0$. The sticks' lengths are given by another probability distribution, $g(L)$. It is assumed that θ and L are independent. Given a random configuration of N sticks with lengths, angles and centres $\{(L_i, \theta_i, x_i, y_i)\}$,

the following quantities may be defined:

$$P_l \equiv \frac{1}{N} \sum_{i=1}^{N} L_i \, |cos\, \theta_i| \tag{6.2a}$$

and

$$P_t \equiv \frac{1}{N} \sum_{i=1}^{N} L_i \, |sin\, \theta_i|. \tag{6.2b}$$

These are the average longitudinal and transverse stick components with respect to the y direction, respectively. From these the macroscopic anisotropy can be defined as

$$P \equiv \frac{P_l}{P_t}. \tag{6.2c}$$

Clearly, for an isotropic sample $P = 1$. The larger P is, the more oriented the sticks are along the y direction. A given sample of sticks is considered to be percolating along the y direction if a continuous path can be traced between intersecting sticks from $y = 0$ to $y = 1$. Similarly for the x direction. The longitudinal critical length, L_{cl} is the lowest average length that gives the onset of percolation in the y direction. The transverse critical length, L_{ct} is likewise defined for x percolation.

For a given collection of sticks, how is a possible conducting path found between opposite ends of the unit square? First, it is necessary, for each of the N sticks, to look for any intersections with the other $(N-1)$ sticks. Fig. 6.1 shows an example of two sticks, labelled i and j. The sticks are of lengths L_i and L_j and are oriented at angles θ_i and θ_j with respect to the y direction. The centres of the sticks have coordinates (x_i, y_i) and (x_j, y_j). Each stick also has another number associated with it: A cluster number. Initially, before searching for any intersections amongst the sticks, the cluster number of each stick is initialised to the index of the stick. That is, the sticks are numbered from 1 to N. The cluster number is set equal to the number of the stick.

Let d_{ij} be the distance between the centres of the sticks. Clearly, if $d_{ij} > (L_i + L_j)/2$ then there can be no intersection. Otherwise, the following quantities

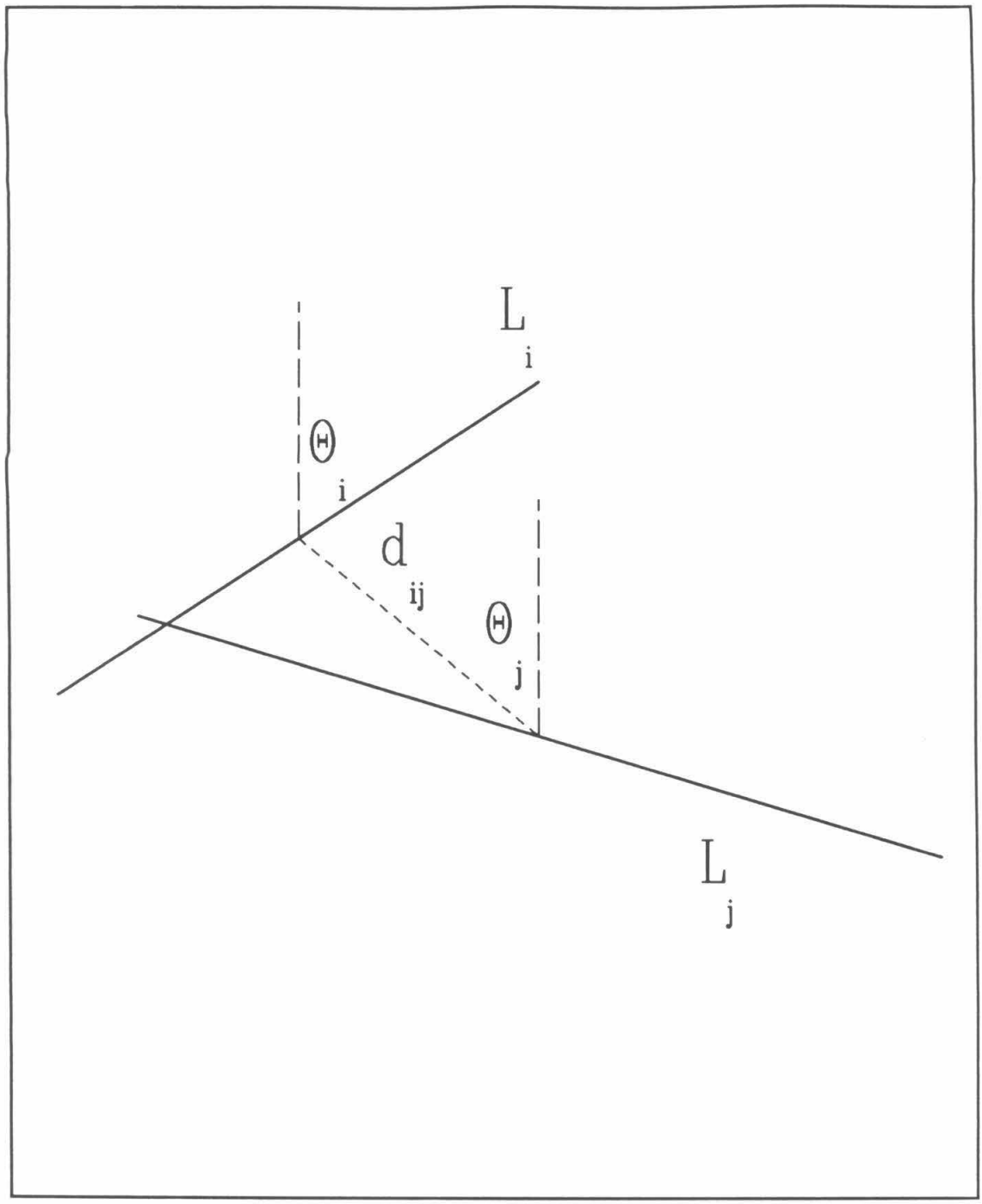

Figure 6.1: Two intersecting sticks i and j, with lengths L_i and L_j and angles θ_i and θ_j. The sticks are separated by the distance d_{ij} between the centres.

are computed:

$$A_i = d_{ij}\,|cos(\theta_j + \gamma)/sin(\theta_j - \theta_i)|$$

$$A_j = d_{ij}\,|cos(\theta_i + \gamma)/sin(\theta_j - \theta_i)|$$

where γ is defined by $tan\gamma = (y_i - y_j)/(x_i - x_j)$. From straightforward trigonometry, it can be shown that if $A_i \leq L_i/2$ and $A_j \leq L_j/2$ then the sticks intersect. If this is the case, then the sticks are considered to be in the same cluster. The (common) cluster number of the sticks is taken to be the minimum of the two prior cluster numbers of the sticks. A search is then made over the other $(N-2)$ sticks to find those sticks with the higher of the two prior cluster numbers. Any such sticks are given the new (lower) cluster number. This is used to merge two clusters that are found to have a common intersection.

By performing the above algorithmn, it is possible to resolve the random collection of sticks into a set of nonintersecting clusters. One also searches for any intersection of the sticks with the boundaries. For any such intersections, the cluster number of the intersecting stick is associated with the boundary. To find an instance of longitudinal percolation, say, it is a simple matter to compare the cluster numbers associated with the $y = 0$ boundary with those associated with the $y = 1$ boundary. If there are any numbers in common, then there exists an instance of longitudinal percolation.

I wish to investigate the dependences of L_{cl} and L_{ct} on the anisotropy, P, and on the number of sticks in the system, N. First, consider systems with a given N. To vary the anisotropy, the variance of $f(\theta)$ is changed. For example, for a normal distribution of angles the angular distribution f is $N(0, \theta_d)$, where θ_d is the standard deviation and the mean of the normal distribution is 0. By progressively reducing θ_d from some initial value, the anisotropy can be correspondingly increased. Consider now that a variance of $f(\theta)$ has been selected. Start with some small average value of L, $E[L]$. With the parameters of f and g chosen, 20

random configurations of sticks are generated, each with N sticks. I then look for any percolation. Twenty random configurations at each value of anisotropy and $E[L]$ were made to increase the statistical reliability of finding accurate critical lengths. This is an improvement over Balberg and Binenbaum who presented critical lengths found from three to five configurations per value of anisotropy. Note that for sufficiently small $E[L]$, few of the sticks will overlap. Systematically, $E[L]$ is incremented until percolation is found. Often this is longitudinal percolation. This should be expected, as it is easier to percolate along a preferred direction than normal to it. (Occasionally, at low anisotropy, $P \leq 1.4$, transverse percolation may first be encountered.) By further incrementing $E[L]$, it is possible to eventually come upon transverse percolation. During the varying of $E[L]$, P is recorded for each random configuration. It is found that to good approximation, P is independent of $E[L]$. My results and a comparison with the theory of Section 6.3 will be presented in Section 6.4.

6.3 Theory

Here a simple model is presented that attempts to explain the dependence of the longitudinal and transverse critical lengths on the anisotropy and the number of sticks in the ensemble.

As before, let there be N sticks in the ensemble, with an angular distribution $f(\theta)$ and a stick length distribution $g(L)$. The first approximation is to replace all the stick lengths L_i with the average length $L_0 \equiv \sum L_i/N$. Next, the angular distribution $f(\theta)$ is replaced by one in which half the sticks are oriented at θ' and

the other half at $-\theta'$ with respect to the y axis, where θ' is defined by

$$\begin{aligned} tan\,\theta' &= \frac{\langle|sin\,\theta|\rangle}{\langle|cos\,\theta|\rangle} \\ &= \frac{P_t}{P_l} \end{aligned} \tag{6.3}$$

For definitiveness, consider now the case of longitudinal percolation. That is, the intention is to find a set of overlapping sticks that goes from $y = 0$ to $y = 1$. Consider a stick at an angle θ', labelled $\overline{AC}$ in Fig. 6.2, where B is the centre of the stick. The only sticks which can intersect $\overline{AC}$ with non zero probability are oriented at $-\theta'$. Furthermore, the centres of these sticks must lie in the parallelogram shown in Fig. 6.2. (Balberg *et al*[9] define this as the excluded area of the two sticks.) Suppose, in tracing out a possible cluster, that the cluster is considered to start at $y = 0$ and that $\overline{AC}$ is the highest stick in the cluster, thus far. Let $\overline{AC}$ be the j-th stick in the cluster. To make progress towards $y = 1$, the probability is needed that a stick has a centre in $\triangle HFG$ and is oriented at $-\theta'$. Remembering that the sticks are distributed uniformly, the probability is given by the product of the probability of one stick intersecting $\overline{AC}$ with a higher y centre times the number of such sticks. Thus the probability for intersecting a stick in $\triangle HFG$ is

$$\begin{aligned} P_{j/2} &= Area(\triangle HFG)\left(\frac{N}{2} - \frac{j}{2}\right) \\ &= P_t\,P_l\left(\frac{N}{2} - \frac{j}{2}\right). \end{aligned} \tag{6.4}$$

(Since $P_{j/2}$ is a probability, if the right hand side of Eq. 6.4 is greater than 1, then set $P_{j/2} = 1$.) Furthermore, if such an intersection occurs it can be seen that the average y coordinate of the intersecting stick will be at a distance $P_l/3$ higher than the y coordinate of $\overline{AC}$'s centre. Thus, the average number of intersections for y percolation is

$$n \equiv \frac{3}{P_l}. \tag{6.5}$$

The probability of finding one percolating cluster in the y direction is given by the

product of n terms of the form of Eq. 6.4. Thus, it can be written as

$$P^{l}_{cluster} = \prod_{j=1}^{n/2} P_j^2, \tag{6.6}$$

where from Eq. 6.4, P_j is defined as

$$P_j \equiv \begin{cases} a_j, & a_j \leq 1 \\ 1, & a_j > 1 \end{cases} \tag{6.7}$$

and

$$a_j \equiv P_t P_l \left(\frac{N}{2} - j\right).$$

Note that in Eq. 6.6 it was assumed that n is even. If n is odd, Eq. 6.6 is multiplied by $P_{n/2+1}$.

Eq. 6.6 gives the probability of forming one cluster. However, it may be possible for several clusters to form, especially for longitudinal percolation in highly anisotropic samples. As mentioned above, the y coordinate of the intersecting stick in the cluster is at an average distance of $P_l/3$ higher. For the x coordinate of the intersecting stick, averaging over $\triangle HFG$ gives 0. This comes from

$$Prob(x < x_B) = Prob(x > x_B) = \frac{1}{2}.$$

But, it can be seen that the average x coordinate of the intersecting stick is $P_t/3$, assuming that $x > x_B$, and it is $-P_t/3$ if $x < x_B$. Hence, for longitudinal percolation there is a random walk in the x direction, of average step size $P_t/3$ and equal probability of stepping positive or negative. From Eq. 6.5, it takes n steps to percolate in the y direction. During these steps, the cluster will span an approximate distance[10] in the x direction of

$$v_0 \equiv \sqrt{n}\,\frac{P_t}{3}.$$

Since the domain of interest is bounded by 0 and 1 in the x direction, let

$$v = \begin{cases} v_0, & v_0 \leq 1 \\ 1, & v_0 > 1 \end{cases}.$$

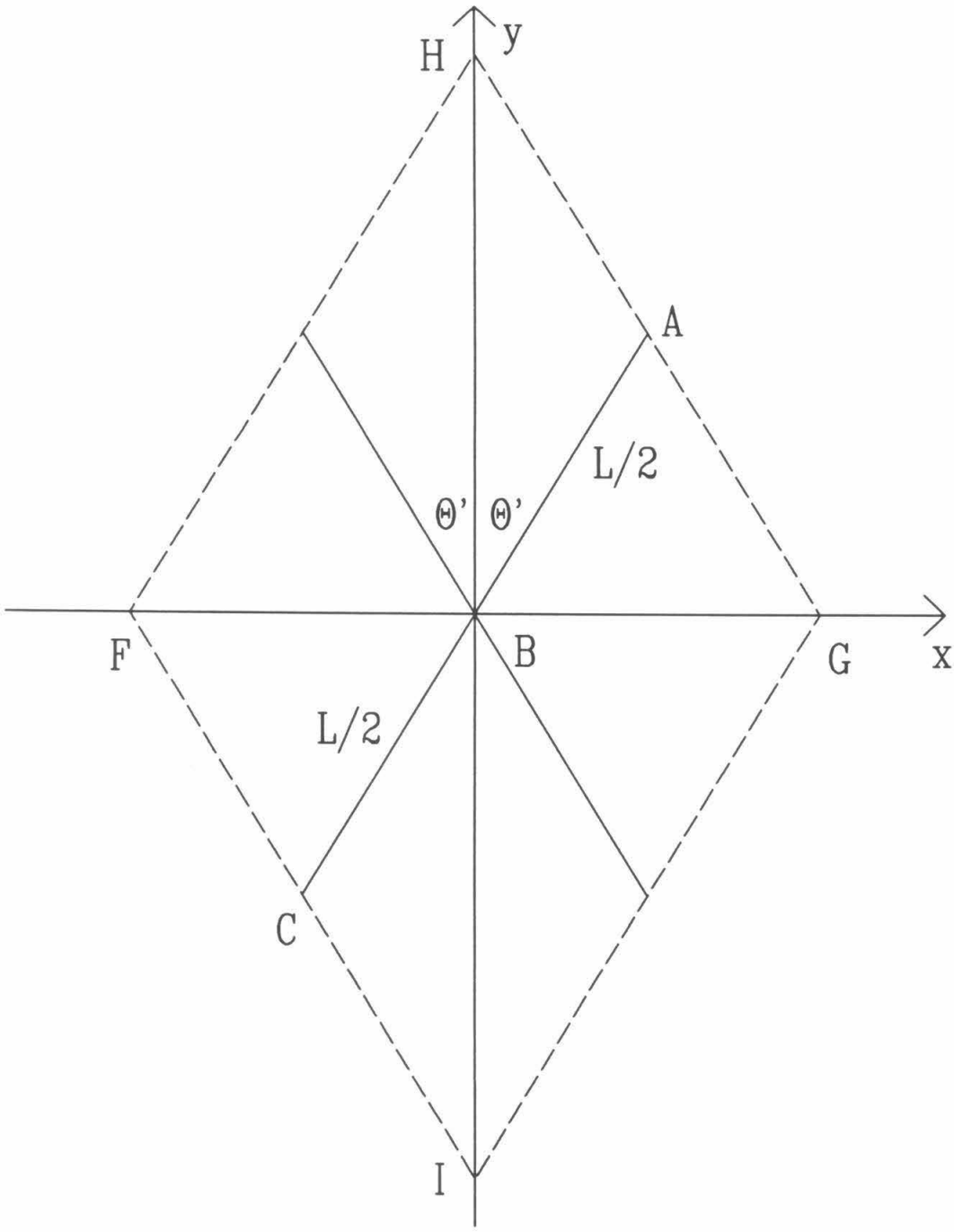

Figure 6.2: A stick $\overline{AC}$ of length L at an angle θ' with respect to the y axis, where θ' is defined in Eq. 6.3. Sticks at an angle $-\theta'$, with centres inside the parallelogram HFIG will intersect $\overline{AC}$.

Then v is the fractional transverse distance covered when longitudinal percolation is being considered. Therefore, the total probability for a longitudinal percolating cluster is given by diving Eq. 6.6 by v to give

$$P_{total} = \frac{1}{v} P^{l}_{cluster}. \tag{6.8}$$

To find the transverse percolation, it suffices to simply interchange P_l and P_t in the above equations. Thus, the derivation given is symmetric in both longitudinal and transverse percolation. Apart from the relative values of P_l and P_t there is no fundamental distinction between the two modes of percolation.

Given values for $\langle|sin\,\theta|\rangle$ and $\langle|cos\,\theta|\rangle$, the above equations can be used to find the critical lengths and the anisotropy.

One further matter needs to be addressed. Given a $\langle|cos\,\theta|\rangle$, which is defined by a particular $f(\theta)$ and a choice of variance of f, there corresponds a unique $\langle|sin\,\theta|\rangle$. But for the set of arbitrary probability distributions f, subject to the constraints of Eq. 6.1, there will be no unique relation between $\langle|cos\,\theta|\rangle$ and $\langle|sin\,\theta|\rangle$. How can an anisotropy be found that is in some sense independent of a particular f? To answer this, note that I have the following inequalities:

$$0 \leq \langle|cos\,\theta|\rangle, \langle|sin\,\theta|\rangle \leq 1 \tag{6.9a}$$

and

$$1 \leq d \leq 2 \tag{6.9b}$$

where

$$d \equiv \langle|cos\,\theta|\rangle + \langle|sin\,\theta|\rangle \tag{6.9c}$$

and the triangle inequality was used to get the lower bound of Eq. 6.9b. The following is done. A value of d is chosen to satisfy Eq. 6.9b. I then vary $\langle|sin\,\theta|\rangle$ in steps, starting from $d/2$ and decrementing to 0 (though not going all the way).

Then $\langle|cos\,\theta|\rangle$ is given by

$$\langle|cos\,\theta|\rangle = \begin{cases} d - \langle|sin\,\theta|\rangle, & d - \langle|sin\,\theta|\rangle \leq 1 \\ 1, & d - \langle|sin\,\theta|\rangle > 1 \end{cases}. \tag{6.10}$$

With these values the anisotropy can be found, and Eq. 6.8 can be used to obtain the critical lengths. The value of 1.5, the midpoint of Eq. 6.9b, is chosen as the most reasonable value of d. It should be realised that the procedure of using Eqs. 6.9c and 6.10 is an approximation. In general, a given f will not yield $\langle|cos\,\theta|\rangle$ and $\langle|sin\,\theta|\rangle$ satisfying the linear relationship of Eq. 6.9c over a range of values of the variance.

A program was written to find the lowest lengths that set $P_{total} \geq 1$ in Eq. 6.8 as a function of anisotropy, for both longitudinal and transverse percolation. These are the critical lengths. The number of sticks in the sample was a parameter in this calculation.

6.4 Results

Here the results of Sections 6.2 and 6.3 are presented and compared with each other. First, consider the predictions of the model of the previous section. The values of critical lengths from Eq. 6.8 are displayed in Figs. 6.3 and 6.4 for 100 and 500 sticks, respectively. In each figure, the solid line is for longitudinal percolation while the dashed line is for transverse percolation. Following Balberg and Binenbaum,[6] the critical lengths have been normalised in units of the average interstick separation, r, where

$$r \equiv \frac{1}{\sqrt{\pi N}}.$$

From the simulations the longitudinal and transverse critical lengths L_{cl}, L_{ct} have been found as functions of anisotropy, for ensembles of 100 and 500 sticks. These are displayed in Figs. 6.3 and 6.4, respectively. For Fig. 6.3, the critical

lengths have been found for three types of ensembles: Those with normal distributions for f and g; with a uniform f and a normal g; and with constant absolute angle and a δ-function for g (ie., all the sticks in an ensemble are the same length). This last pair of distributions corresponds to the simplified choices of distributions made in Section 6.3. Similar remarks hold for Fig. 6.4. The solid objects in the figures are the critical longitudinal lengths. The hollow objects are the critical transverse lengths.

From Figs. 6.3 and 6.4 it can be seen that for both the simulations and the theory the transverse critical lengths lie distinctly above the longitudinal critical lengths. Considering separately the transverse and longitudinal results, notice that the simulations with constant length and absolute angle tend to yield larger critical lengths than the other simulations. This is expected,[6] as a distribution of lengths will cause the sticks with lengths greater than the mean length to contribute preferentially to the percolation. Hence percolation will start sooner than if all the lengths in a sample are constant. Nonetheless, the simulation results for transverse percolation are clustered close enough, and likewise for longitudinal percolation, a universal behaviour of the critical lengths on the anisotropy is found that is largely independent of the choice of distributions. Comparing the theory with the simulations, there is seen to be good agreement for both 100 and 500 sticks. The largest disagreement is for the longitudinal percolation of 100 sticks, with the theory lying above the simulations. But even here, the theoretical curve exhibits the same trends as the simulations and the disagreement is only semiquantitative.

Note that in comparing Figs. 6.3 and 6.4 the longitudinal and transverse results tend to converge together as the number of sticks in a sample increase from 100 to 500. This is in accordance with Balberg and Binenbaum[6] who showed by a topological argument that in the limit of infinite N, the two types of critical lengths coincide.

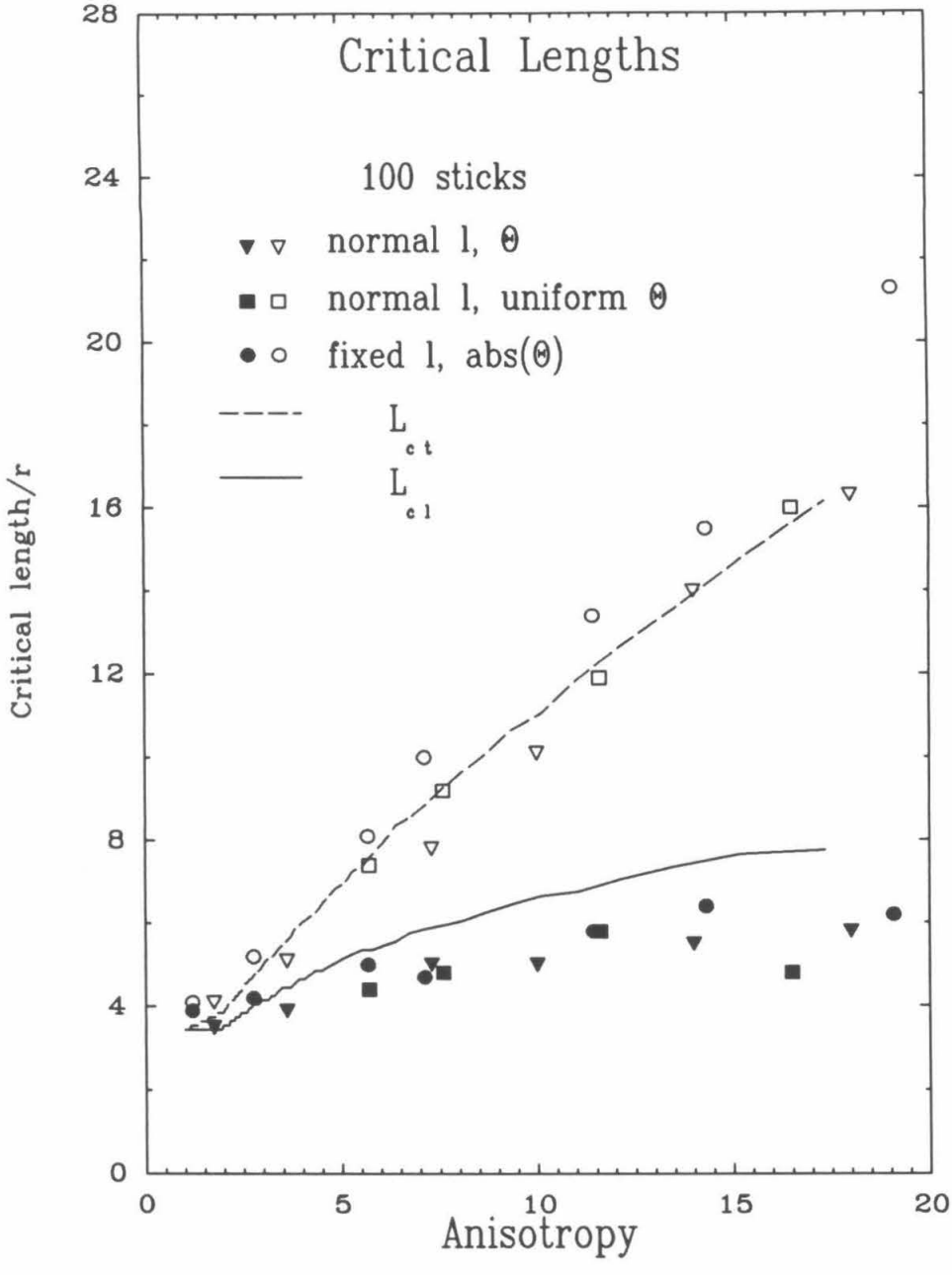

Figure 6.3: The critical lengths as a function of anisotropy for samples of 100 sticks. The anisotropy is defined in Eq. 6.2. The solid and hollow symbols are from simulations for longitudinal and transverse percolation respectively. The lines are from Eq. 6.9.

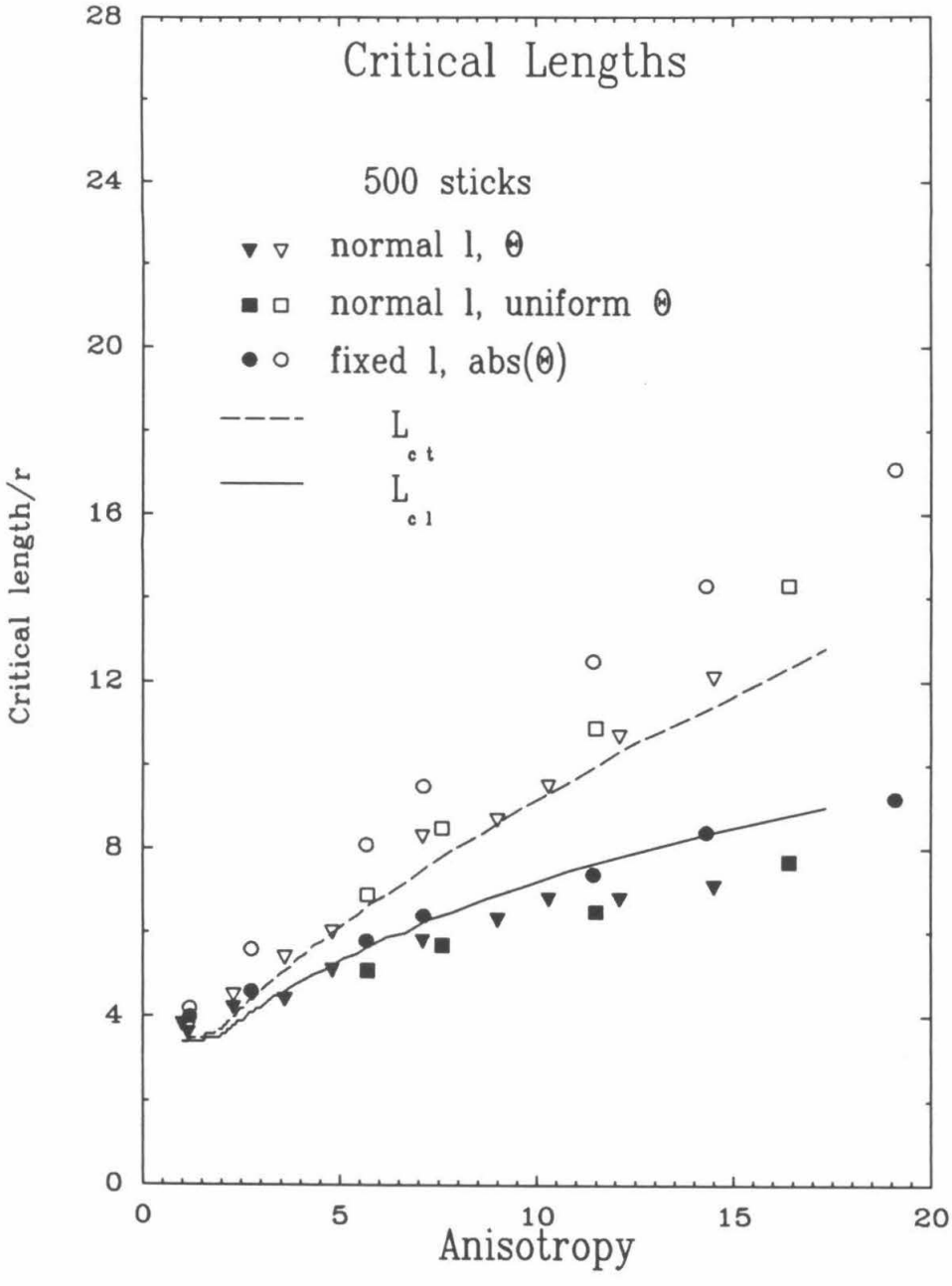

Figure 6.4: The critical lengths as a function of anisotropy for samples of 500 sticks. The anisotropy is defined in Eq. 6.2. The solid and hollow symbols are from simulations for longitudinal and transverse percolation respectively. The lines are from Eq. 6.9.

Let me add a comment about the midrange choice of $d = 1.5$ made in Eqs. 6.9 and 6.10 in the previous section. This was done to give a definite prescription for the anisotropy. If d is set equal to the lower limit, 1, curves are found that start at a critical length around 5 at isotropy, and remain above the displayed curves as the anisotropy is increased. This was found for both longitudinal and transverse percolation. The agreement between the (undisplayed) $d = 1$ curves and the simulations is not as good as for the $d = 1.5$ curves shown here in Figs. 6.3 and 6.4. Alternatively, choosing d equal to the upper limit of 2 gives curves that start and remain slightly lower than the displayed curves. Hence, the choice of $d = 1.5$ is considered to be a reasonable one.

Therefore, it is suggested that the theory of Section 6.3 gives good predictions of critical length. The theory also has the computational advantage of being much faster to run. The curves in Figs. 6.3 and 6.4 took less than a minute each to find, on a Vax 11/785. By contrast, the CPU time for the simulations is measured in hours. This is especially true for the simulations of 500 sticks. The CPU time required scales roughly as the square of the number of sticks, as the most computationally intensive task is to find the possible intersections amongst a collection of N sticks.

6.5 Conclusion

The finite-size effects of anisotropic continuum percolation in two dimensions have been investigated. The elements that percolate are widthless sticks. A simple theory has been developed to explain the dependence of longitudinal and transverse critical lengths on anisotropy and the finite number of sticks in the sample. By comparing the theory to simulations, good agreement is found. It also affords significant computational advantages over performing simulations. I believe that

the theory is the first to explain the finite-size simulation results for anisotropic continuum percolation in two dimensions. Thus, I suggest that the theory offers a useful complement to the running of simulations.

References

1. D. M. Bigg, Poly. Eng. Sci. **19**, 1188 (1979).
2. H. Hatta and M. Taya, J. Appl. Phys. **58**, 2478 (1985).
3. N. Ueda and M. Taya, J. Appl. Phys. **60**, 459 (1986).
4. G. E. Pike and C. H. Seager, Phys. Rev. B **10**, 1421 (1974).
5. J. M. Ziman, *Models of Disorder*, (Cambridge University Press, Cambridge, 1979)
6. I. Balberg and N. Binenbaum, Phys. Rev. B **28**, 3799 (1983).
7. *I.* Balberg, Phys. Rev. B **31**, 4053 (1985).
8. L. W. Noh, S. Lee, Y. Song and J. R. Gaines, Phys. Lett. **114A**, 207 (1986).
9. I. Balberg, C. H. Anderson, S. Alexander and N. Wagner, Phys. Rev. B **30**, 3933 (1984).
10. F. Reif,*Fundamentals of Statistical and Thermal Physics*, (McGraw-Hill, New York, 1965).

Appendix A

Software for Chapters 2 and 3

Listed below is the structure of the Fortran program **ohmic.for** which is used to find the conductance in Chapters 2 and 3. This is intended to guide those who may wish to use or extend this program. Subroutine names are in **boldface** below. If a subroutine is not documented, then its operation is obvious.

---

ohmic.for (main level)

get input parameters; **input**

get results; **method**

end.

---

subroutine method

write input parameters to file *ohmic.dat*;

define constants; **constants**

do vary bias

do vary doping

if no simulations **then**

find conductance; **conduct**

else

do simulations for conductance; **montcrl**

endif

write results to file *ohmic.dat*;

enddo;

enddo;

end;

---

subroutine conduct

initialise variables;

get effective density of states;

get Fermi levels in semiconductor and metal;

get depletion length;

do vary energy of particle

get conductance G(E) at this energy; **condis**

enddo;

get conductance due to thermionic emission;

get total conductance from summing over all energies; **sumexp**

end;

---

function condis(E)

initialise variables;

get intersections of energy E with potential; **finlim**

get transmission through forbidden region;

get conductance G(E) from the transmission;

end;

---

function sumexp

sums a set of numbers, which are input as exponents of 10;

end;

---

subroutine finlim

finds intersections between input energy E and input potential

in array y;

(linear interpolation is assumed between points in y)

---

subroutine montcrl

get thicknesses of volume in which to put dopants;

get contribution to potential from dopant continuum

outside cylinder into array *potcont*; **continm**

put dopants randomly at lattice sites; **lattice**

get conductances for 1d or 3d trajectories; **conduct4**

end;

---

subroutine continm

initialise variables;

do vary x position normal to junction

get potential at x from continuum of dopants outside

cylinder centred on the x direction;

add contribution from image of carrier in metal;

enddo;

end;

```
subroutine lattice
  initialise variables;
  get probability of lattice site being occupied by dopant;
  do sum over unit cells in the given volume
    do sum over 4 fcc positions in a unit cell
      randomly put a dopant at this lattice site;
      enddo;
    enddo;
  end;
```

```
subroutine conduct4
  initialise variables;
  do vary z coordinate of trajectory
    do vary y coordinate of trajectory
      get potential centred on initial trajectory (y,z);        varpot
      do vary energy of particle
        get trajectory and conductance;                         traject4
        enddo;
      get conductance due to thermionic emission;
      get total conductance by summing over all energies;       sumexp
      enddo;
    enddo;
  end;
```

```
subroutine varpot
  (input: trajectory (y,z) coordinates and cylinder radius);
  initialise variables;
  do search over dopants
    if dopant inside cylinder then
      do vary x along trajectory
        add contribution of dopant to potential at x;
        enddo;
      endif;
    enddo;
  end;
```

```
subroutine constants
  initialises useful constants;
  end;
```

```
subroutine traject4
  initialise variables;
  get transmission exponent for 1d trajectory along initial (y,z) coordinates;
  initialise trajectory;
  if within specified volume then
    if reached yz boundary of known potential then
      find potential in volume centred on (y,z);                varpot
      endif;
    get next point on trajectory;
    if in forbidden region then
      add contribution to transmission from present point on
      the trajectory;
      endif;
    endif;
  end;
```

Appendix B

Software for Chapter 6

Listed below is the structure of the Fortran program **sticks.for** which is used to find the critical lengths for the two dimensional anisotropic continuum percolation of Chapter 6. This is intended to guide those who may wish to use or extend this program. Subroutine names are in **boldface** below. If a subroutine is not documented, then its operation is obvious.

---

sticks.for (main level)

get input parameters;	**input**
do vary mean stick length	
do several random samples	
place stick centres randomly in unit square;	**centres**
get angular distribution of sticks;	**angles**
get length distribution of sticks;	**lengths**
check for percolation;	**cluster**
find anisotropy of sample;	**aniso**
store results in file *sticks.out*;	**output**
accumulate statistics;	
enddo;	
enddo;	
store results in file *sticks.outt*;	**out2**
end.	

---

subroutine centres

place centres of dopants uniformly in unit square;

end;

---

subroutine angles

choose angles from different distributions:

constant absolute angle;

uniform distribution;

normal distribution;

end;

---

subroutine lengths

choose lengths from different distributions:

constant length;

uniform distribution;

normal distribution;

end;

---

subroutine aniso

find total longitudinal component of sample;

find total transverse component of sample;

get ratio of longitudinal and transverse components;

end;

```
subroutine cluster
  initialise cluster number of sticks;
  do for all sticks
    do search over all other sticks
      if the two sticks intersect then
        give the sticks the lower of the 2 cluster numbers;
        do for all sticks
          merge the 2 clusters;
          enddo;
        endif;
      enddo;
    enddo;
  find clusters that overlap boundaries;
  if any clusters go to opposite y boundaries then
    have found longitudinal percolation;
    endif;
  if any clusters go to opposite x boundaries then
    have found transverse percolation;
    endif;
  end;
```