



<https://theses.gla.ac.uk/>

Theses Digitisation:

<https://www.gla.ac.uk/myglasgow/research/enlighten/theses/digitisation/>

This is a digitised version of the original print thesis.

Copyright and moral rights for this work are retained by the author

A copy can be downloaded for personal non-commercial research or study,
without prior permission or charge

This work cannot be reproduced or quoted extensively from without first
obtaining permission in writing from the author

The content must not be changed in any way or sold commercially in any
format or medium without the formal permission of the author

When referring to this work, full bibliographic details including the author,
title, awarding institution and date of the thesis must be given

Enlighten: Theses

<https://theses.gla.ac.uk/>
research-enlighten@glasgow.ac.uk

Electron Transport in Dry-Etched n^+ -GaAs Nanostructures

Mahfuzur Rahman

Submitted for the degree of Doctor of Philosophy
to the Department of Electronics and Electrical Engineering
at the University of Glasgow

March 1992

© Mahfuzur Rahman, 1992.

ProQuest Number: 11011476

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest 11011476

Published by ProQuest LLC (2018). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code
Microform Edition © ProQuest LLC.

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

Erratum

The widths of the wet-etch wires in Chapter 3 have been miscalculated. References to these wet-etch measurements should be ignored.

ନତିବା

(to Lotika)

*"Why is it that you physicists always require so much expensive equipment?
Now the Department of Mathematics requires nothing but money
for paper, pencils and waste paper baskets, and the
Department of Philosophy is better still. It doesn't even ask
for waste paper baskets."*

– A quote in a philosophy book.

Acknowledgements

This thesis owes its existence to many people.

The late J Lamb and his successor P J R Laybourn are thanked for providing the facilities within the Electronics Department, using which all the work described in this thesis is done.

The (unfortunate?) task of having to supervise the whole of this work, and keep me sane during most of the last few years, fell upon the shoulders of C D W Wilkinson and A R Long.

The trials and tribulations of fabrication were largely learned under the tutelage of A Stark and D Gourlay. Fellow (more hardened and experienced) researchers S Thoms, J Adams, I Thayne, and others, shared their preciously gained knowledge on techniques to ensure that samples survived just that little bit longer before they were lost or destroyed.

The fabrication labs (through all their metamorphoses) have been managed very ably by D MacIntyre, with the help of D Gourlay, H McLelland, S Ferguson, G Hopkins and K M Donnelly. Much of the dry-etching was done with the help of R Darkin, D Clifton and D Lawson. The low-temperature measurements would not have been possible without the help of D Irons, A Ross and J Smith. A special further thank you to S Ferguson, because she asked for a special mention.

The early lack of experience with how to cope with the harrowing problem of noise in measurements was offset somewhat with the help and suggestions of S P Beaumont, P Beton, J Weaver, R Cusco, K Thomas and J G Williamson. The help and finance of D Gourlay (the help), and S P Beaumont (the finance) for trying to overcome the problems of sample bonding are also appreciated.

The work on dry-etch damage characterisation was originally motivated by S

Thoms. Chapter 4 certainly would not exist had it not been for the efforts of A R Long, I K MacDonald and M Kinsler, which resulted in the initial attempts to model the damage by A R Long. This part of the work has benefitted enormously from discussions with C D W Wilkinson, A R Long, M A Foad, N P Johnson, R Cheung, S P Beaumont, C M Sotomayor-Torres, and M R S Taylor. The understanding of the quantum transport properties of n^+ -wires has benefitted enormously from discussions with J H Davies, J G Williamson, C D W Wilkinson, A R Long and S P Beaumont. Most of the GaAs material used in this work was grown by M C Holland.

My colleagues, who have shared the "Ph.D. experience" during these last few years, are also thanked.

The unfortunate task of having to review parts of this thesis fell to the "preselected" victims, C D W Wilkinson, J H Davies, J G Williamson and S Ferguson.

Finally, I must thank the rest of my family for urging me on through the whole of this work, and especially Lotika, were it not for whom I would most certainly have given up in frustration and anguish during the latter stages of writing this thesis, and no doubt ended up in the "home" for the bewildered research student.

Outline of thesis

This thesis contains a study of electrical transport in n^+ -GaAs wires fabricated by electron beam lithography and silicon tetrachloride reactive-ion etching.

The etching is done at ion energies around 300 eV. At these energies, damage is known to be created in the wires being fabricated. To characterise the amount of sidewall damage, the conductances of wires of different widths are measured, and the minimum conducting width (cut-off width) is found. The measurements show that dry-etching gives a higher cut-off width than wet-etching. Dry-etching at energies in the range 280-380 eV seem to give similar cut-off widths, whereas etching for progressively longer periods of time clearly increases the cut-off width.

A quantitative model is developed to account for these measurements. Defect profiles for top surface and sidewall damage are obtained. Comparisons to DLTS profiles indicate that defect diffusion can be ignored to a good approximation, and a simple exponential source function used to model defect formation. Assuming a two-level defect structure, conductances for wires, and sheet conductances for etched epilayers, are calculated by numerically solving Poisson's equation. It is found that a single source function, characteristic of the etch process, gives an excellent fit to all the wire conductance measurements. The model thus accounts for differences in carrier concentration, epilayer thickness and etch time/depth. The analysis indicates that defects are created at a far higher rate at the sidewalls than expected. This is thought to be due to the flux of ricochet and sputtered material from the top surface being etched.

Low temperature measurements of submicron length n^+ -wires in a magnetic field reveal fluctuations in longitudinal, Hall and nonlocal configurations. The phase coherence length is found to decrease as the wire is made narrower. Measurements of weak localisation in a very narrow wire suggest that the boundary scattering in these

wires is predominantly diffuse. Measurements of weak localisation over a range of temperatures indicate three contributions to the phase breaking rate: an electron-phonon term, an electron-electron term, and a temperature independent term. Two electron-electron scattering mechanisms are examined, but cannot be distinguished by the measurements. The temperature independent term, and the dependence of the coherence length upon wire width, are discussed in terms of magnetic scattering between conduction electrons and unpaired electrons trapped in defect states along the edge of the depletion region.

The Hall measurements are shown to probe voltage fluctuations associated with a phase coherent region. This is confirmed by measurements at a cross-junction. It is also shown that the symmetric and antisymmetric components of the conductance fluctuations in a wire can be manifestly decoupled by making the wire width much narrower than the probe width.

The low temperature electrical response of n^+ -wires to a pulse of light is measured in a zero magnetic field. Random telegraph signals are observed in both longitudinal and transverse ("Hall") measurements. These are explained in terms of electron interaction with defects along the edge of the depletion region.

Table of constants

quantity	symbol	value	units
Boltzmann constant	k_B	1.38×10^{-23}	J K ⁻¹
Planck constant	h	6.63×10^{-34}	J s
	\hbar	1.055×10^{-34}	J s
Electron rest mass	m_0	9.11×10^{-31}	kg
Elementary charge	e	1.602×10^{-19}	C
Permittivity in vacuum	ϵ_0	8.854×10^{-12}	F m ⁻¹
Speed of light in vacuum	c	3.0×10^8	m s ⁻¹
GaAs dielectric constant	ϵ_r	13.1	
Pi	π	3.1415926	

Contents

Acknowledgements	v
Outline of thesis	vii
Table of constants	ix
1. Introduction	1
2. Experimental methods	4
2.1 Outline	4
2.2 Device fabrication	5
2.2.1 Electron-beam lithography	5
2.2.2 Resists	6
2.2.3 Metallisation: ohmics and Schottkys	9
2.2.4 Etching	10
2.2.5 Process notes	13
2.3 Measurement methods	15
2.3.1 Measurement of resistance	15
2.3.2 The lock-in technique	17
2.3.3 Interference in measurements	18

2.3.4	Cryogenic systems and superconducting magnets	19
2.3.5	Automation of measurements	20
2.4	References	20
3.	Characterisation of dry-etch damage	22
3.1	Outline	22
3.2	Introduction	23
3.3	Defects in semiconductors	23
3.3.1	Lattice defects	23
3.3.2	Point defects in GaAs	24
3.3.3	The <i>DX</i> -centre and EL2 in GaAs	25
3.4	Dry-etching with SiCl_4	27
3.4.1	Characterisation of dry-etch damage	27
3.4.2	SiCl_4 reactive-ion etching	28
3.4.3	Characterisation of SiCl_4 -etch damage	29
3.5	Experiments on SiCl_4 -etched n^+ -wires	29
3.5.1	The cut-off method	29
3.5.2	Sample fabrication	31
3.5.3	Experimental results on cut-off widths	31
3.5.4	Hot electron effects	36
3.5.5	Photoconductance effects	38
3.6	References	38
4.	Model for SiCl_4-etch damage	41
4.1	Outline	41
4.2	Models for dry-etch damage	42
4.2.1	Requirements of the model	42
4.2.2	Effects of surface	42
4.2.3	The <i>DX</i> -model	42
4.2.4	Use of a two-level defect	45
4.3	Defect distributions in etched structures	46
4.3.1	Defect formation and propagation	46
4.3.2	Defect distribution in one-dimension (top surface damage)	48
4.3.3	Defect distribution in two-dimensions (sidewall damage)	50

4.3.4	Fit to DLTS profiles	51
4.4	Conductances of etched structures	54
4.4.1	Conductances for etched wires	54
4.4.2	Sheet conductances for etched epilayers	56
4.4.3	Comparison with experimental wire conductances	56
4.4.4	Comparison with experimental sheet conductances	61
4.5	References	62
5.	Quantum transport in mesoscopic systems	65
5.1	Outline	65
5.2	Mesoscopic systems	66
5.2.1	Electron coherence	66
5.2.2	Length, time and energy scales	68
5.2.3	Energy averaging	72
5.2.4	Mesoscopic systems	74
5.2.5	The Landauer formula	75
5.3	Quantum interference effects	77
5.3.1	The Aharonov-Bohm effect	77
5.3.2	Weak localisation	80
5.3.3	Universal conductance fluctuations	85
5.4	References	91
6.	Fluctuations and quantum noise phenomena	96
6.1	Outline	96
6.2	The longitudinal measurement	97
6.2.1	Sample fabrication and measurement system	97
6.2.2	Fluctuations in the longitudinal measurement	100
6.2.3	Role of voltage probes in the quantum conductance measurement	106
6.2.4	Use of the autocorrelation function	109
6.2.5	Boundary scattering and classical and quantum size effects	111
6.2.6	Effects of boundary scattering in n^+ -wires	116
6.2.7	Coexistence of localisation and fluctuation effects	129
6.3	Hall and nonlocal measurements	130
6.3.1	The Hall measurement	130

6.3.2	Symmetries of the 4-terminal conductance	135
6.3.3	Nonlocal fluctuations	137
6.4	Photoconductance	141
6.4.1	Random telegraph signals in 2-probe wires	141
6.4.2	Random telegraph signals in Hall geometries	148
6.5	References	151
7.	Summary and conclusions	156
Appendix		161
A.	Derivation of defect distributions	161
A.1	Derivation of top surface damage distribution	161
A.2	Derivation of sidewall damage distribution	165
B.	Numerical methods	169
B.1	Cubic spline interpolation	169
B.2	Integrating by the method of Romberg	170
B.3	Solving the 1-dimensional Poisson equation	170
B.4	Solving the 2-dimensional Poisson equation	171
B.5	The digamma function	173
C.	Program listings	173
C.1	Program listing 1: wire conductances	173
C.2	Program listing 2: sheet conductances	181
D.	References	188

Introduction

A variety of influences – scientific, commercial, military, *etc.* – have driven the evolution of electronics, starting from the point-contact detectors in the very early 1900s, to the "hi-fi" systems and personal computers nearly 90 years on. Vacuum tube technology was largely spawned from investigations into fundamental physical phenomena of interest at the turn of the century. Refinement of the silicon point-contact detector during the war effort led to the development of the first point-contact transistor at the Bell Labs in 1947. Thereafter, growing commercial pressure led to the development of integrated technology, on which the semiconductor industry is currently built.

Exploitation of semiconductor technology has primarily focussed on the use of silicon, coupled with photolithography as the main fabrication tool. Progress has generally been measured by how small minimum features sizes of devices can be made. Since the first monolithic devices appeared around 1958 there has been a steady downward trend in this size. As the limits of scaling are reached, questions pertaining to the further evolution of electronics arise. Several wide-ranging philosophies are already under active investigation.

At the forefront of these is the continued scaling of conventional circuits (employing MOSFETs, MESFETs, *etc.*), with photolithography being supplanted by

higher resolution lithographic techniques, employing electron-beams, ion-beams or X-rays. With ultimate resolution achieved, further evolution using "conventional" (*i.e.* current-day) electronics would rely on a systems-level approach, where the focus of attention is not on the devices themselves but on the way they are connected together. Good examples of this are systems based on neural networks.

More ambitious and radical ideas for the future are based on bioelectronics, where organic matter is expected to perform complex functions, or molecular electronics, where logic functions are performed (in some manner) at an atomic level. Yet further ideas for future development are based on currently accessible lithographic technology, but rely on new physical phenomena for circuit operation. These include quantum electronics, where quantum interference effects are utilised, and single-electron electronics, where single electron effects are utilised.

The 1980s saw developments in the physics of nanostructures, which have given foundation to the hope that new and useful circuit devices will emerge in the areas of quantum and single-electron electronics. These developments have resulted, in part, from the use of high resolution fabrication techniques, such as electron-beam lithography and dry-etching, and from improved material quality, obtained using molecular beam epitaxy. Although practical electronic circuits are still lacking, a large amount of understanding has been gained – and continues to be gained – about all aspects of what goes on in structures at dimensions of a few tens of nanometres.

The work described in this thesis pertains to a specific type of nanostructure, a long thin rib of conducting material, called a "wire." Under special conditions these structures exhibit effects characteristic of a 1-dimensional system. The wires used in this thesis are fabricated from heavily *n*-doped epitaxial gallium arsenide, using electron-beam lithography and dry-etching. Although these wires are not practical circuit devices, the effects studied do exhibit limitations that will be experienced in making any devices of a similar type, as well as yielding much interesting and relevant physics.

Chapter 2 is concerned with methods of fabrication, and techniques for the measurement of conductance, the quantity of primary interest in such structures. The use of dry-etching is known to modify properties of the material being etched. Chapter 3 reviews the current status of experimental studies on this dry-etch "damage," and summarises what is known about simple defects in gallium arsenide. To assess the effect of damage, conductance measurements have been made on dry-etched wires at

room temperature, and the results are also presented in this chapter. In Chapter 4, the information contained in Chapter 3 is used to form a mathematical model of how damage affects the conductance of wires. Numerical solutions to the problem are presented.

The subsequent two chapters are concerned with low temperature quantum effects observed in dry-etched wires. Chapter 5 describes the basic theory required to understand these effects. Much of this theory is now several years old, and (with slight exceptions) is reasonably well established. Chapter 6 contains the results of low temperature measurements made on the above wires. The new effects found are explained in the context of the theory reviewed in Chapter 5. Finally, Chapter 7 contains a summary of the work presented in the thesis.

Experimental methods

2.1 Outline

The first half of the chapter is concerned with device fabrication. The primary tool used is electron-beam lithography. A brief review of electron-beam lithography, describing the electron-beam machine and the main fabrication techniques, is given. In the structures made, an important step during fabrication has been the use of reactive-ion etching which is also explained. The section ends with notes on the precise fabrication steps used.

The second half of the chapter is concerned with methods required to make measurements on completed devices. The basic quantity to be found is the sample resistance (or conductance). The simplest approach to this is to perform a d.c. measurement, and indeed this is done whenever possible. However, circumstances arise where the signal levels that must be monitored are so low that such measurements become exceptionally difficult. Here a.c. methods prevail. The basic theory behind a.c. and d.c. techniques is described. Also included are brief accounts of cryogenic systems and superconducting magnets, since these form an integral part of the measurement apparatus in many cases.

2.2 Device fabrication

2.2.1 Electron-beam lithography

Over the years several methods of writing patterns onto semiconductor substrates have evolved. Traditionally photolithography has been – and still is – the predominant form. However, in endeavours to achieve absolute miniaturisation, photolithography suffers from a severe drawback. Feature sizes on a photolithographic mask are diffraction limited: too small, and the ultra-violet light forms a diffraction image on the photoresist beyond. Several technologies have emerged which allow patterns to be defined, where dimensions are an order of magnitude smaller than those achievable with photolithography [although methods to "beat" the resolution limits of photolithography are continually being sought, *e.g.* Todokoro *et al.* (1991)]. Chief among these are X-ray lithography and electron-beam (*e-beam*) lithography. In terms of flexibility and ease of use, especially when processing small jobs, *e-beam* lithography is perhaps the most well suited of all these methods for defining ultrasmall patterns. Its main drawback, however, is the lack of achievable throughput as compared to photolithography and X-ray lithography. As such, the major current uses of *e-beam* lithography have been limited to mask production for photolithography, and as a research tool in nanolithography.

In the work to be described all processing steps were carried out by *e-beam*. The basic machine is a Philips PSEM 500 scanning electron microscope specially converted for lithography [Mackie (1984), Adams (1990)]. A simplified schematic diagram is shown in Fig. 2.1. The source of the electrons is a heated tungsten filament at the top of the column. The beam of electrons emerging from this source is directed and focussed into a spot at the bottom of the column by a series of magnetic lenses [see Agar (1980) for more information on electron optics]. The sample sits on a motor-driven stage at the bottom of the column. The spot can be scanned over areas of the sample under computer control. This is how patterns are written. The pattern may be repeated on other areas of the sample by moving the stage which is also under computer control.

The pattern written on the sample must somehow be "recorded." That is, the sample must be coated with a medium which is responsive to the electron beam. This medium is referred to as resist. Typically, the areas of the resist exposed by the *e-beam* are dissolved away using some suitable solvent. Thus, an image of the scanned pattern is left on the sample. The lithographic process is then completed by using the

remaining resist as a mask, either for depositing metals on the sample (metallisation), or for removing surfaces not covered by the resist (*e.g.* by wet-etching). Resists, metallisation and etching are the topics of the next few sections.

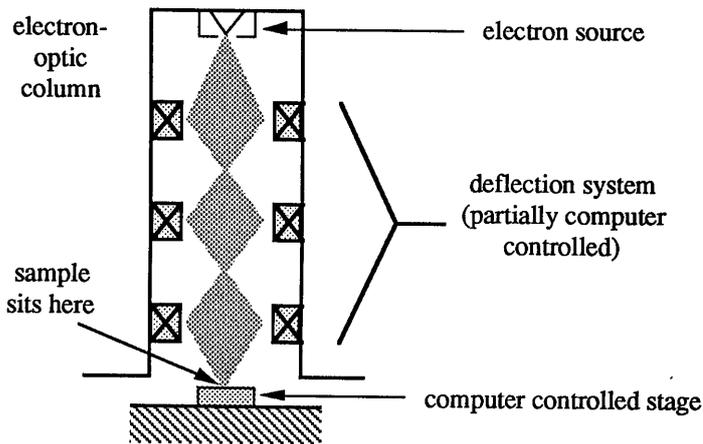


Figure 2.1: Schematic diagram of an *e*-beam machine. The electrons are directed onto the sample by a series of magnetic lenses.

2.2.2 Resists

After the advent of *e*-beam systems (*viz.* electron microscopes), and the realisation that this provided an avenue to vastly reduced lithographic dimensions, suitable *e*-beam resists were sought for. During the mid-1960s it was discovered that PMMA (polymethyl methacrylate – commercially known as Perspex) had many of the qualities desirable of an *e*-beam resist [Haller *et al.* (1968)]. Since then, many other forms of resist have also been developed, with usage dependent on intended application [*e.g.* Wolfstädter *et al.* (1990)]. However, PMMA is the one currently in most widespread use for nanolithography.

Resists fall into two general categories: positive and negative. Depending on the type of resist, electrons interacting with it will either break its constituent long-chain molecules into smaller ones, or induce cross-linking between them [Binnie (1985)]. In the former case, developing the resist in a suitable solvent removes the "broken" molecules, *i.e.* the pattern left in the resist is the pattern scanned by the *e*-beam, as shown in Fig. 2.2(a). This is known as a positive resist. PMMA is an example of a

positive resist. In the other case, developing the resist removes the regions *unexposed* by the *e*-beam, as shown in Fig. 2.2(b). This is known as a negative resist. HRN (high resolution negative resist) is an example of a negative resist. PMMA is usually developed with methyl-isobutyl ketone (MIBK) diluted in isopropyl alcohol (IPA). HRN is developed in undiluted MIBK.

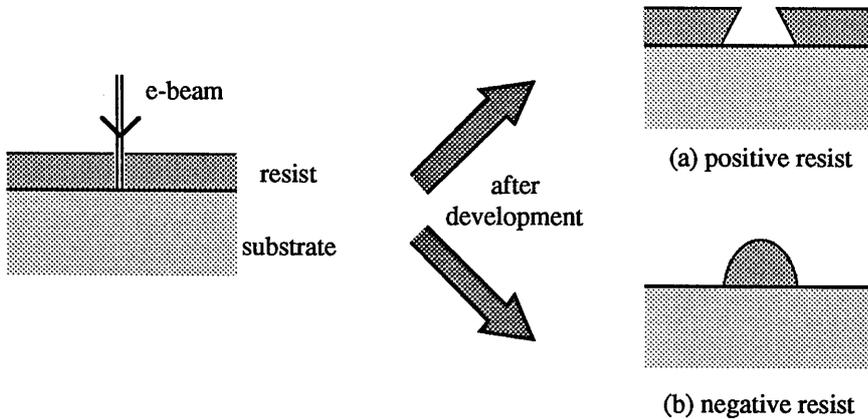


Figure 2.2: The effects of *e*-beam exposure on positive and negative resists. Any exposed positive resist is removed upon development, whereas for negative resist the unexposed regions are removed.

Resist is applied to the surface of a sample by spin-coating. A few drops are deposited at the centre of the chip, following which the chip is revolved on a spinner at several thousand revolutions per minute (rpm). The resist should spread itself over the surface in a uniform layer, with any excess being thrown off during spinning. The thickness of the resist is primarily determined by two factors [Binnie (1985)]: (i) the viscosity of the resist solution, and (ii) the spin speed. Resist thickness is unaffected by the total spin time after an initial transient stage. Less viscous solutions and higher spin speeds result in thinner resist layers. The viscosity depends on the amount of resist there is dissolved in an appropriate solvent. For PMMA the solvent is either chlorobenzene (for thicker layers) or 'o'-xylene (for thinner layers). For HRN it is Microposit thinner. After spinning, the resist is dried by baking in an oven.

The single most important use of PMMA is in the metallisation of samples by the

technique of lift-off, as demonstrated in Fig. 2.3 [Hatzakis (1967)]. For correctly exposed samples, details of the electron-resist-substrate interaction [see Broers (1988) for a brief review of this] cause the developed PMMA to have an overhanging profile [Fig. 2.3(a)]. When metal is evaporated onto the sample from directly above, it deposits on the sample surface and on the resist surface [Fig. 2.3(b)]. Now if the PMMA is dissolved away (usually in acetone), the metal that was previously on the resist is "lifted-off," and only the pattern on the sample remains [Fig. 2.3(c)]. This is how metallisation levels are transferred onto a sample.

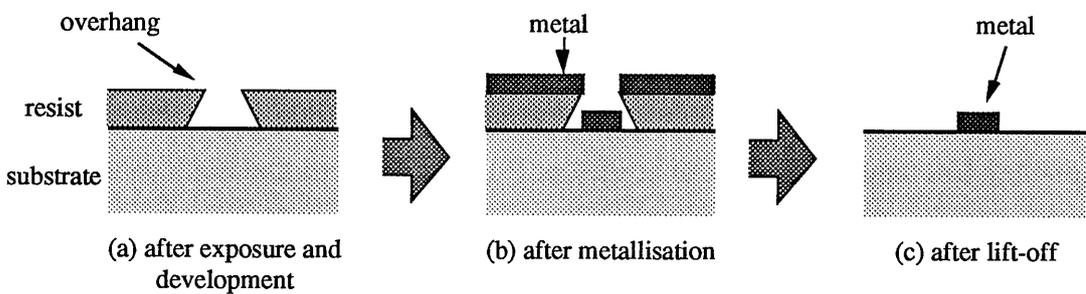


Figure 2.3: The lift-off process using a single layer resist. The process relies on the formation of an overhang in the resist profile prior to metallisation.

The above method works well for thick layers of resist. However, for high-resolution lithography thinner layers are required, and as shown in Fig. 2.4(a) the overhang these give may be insufficient to allow lift-off. The solution to this is to use a bilayer system of resist [Beaumont *et al.* (1981)]. The sensitivity of a resist, *i.e.* how easily it is exposed by the electrons, is dependent on the mean molecular weight of its constituent polymeric chains. A higher molecular weight resist will give smaller linewidths for a given electron dosage as compared to lower molecular weight resists. Thus by using higher molecular weight resist on top of lower molecular weight resist [Fig. 2.4(b)], the profile shown in Fig. 2.4(c) will be obtained after exposure and development. An overhang has been engineered, and lift-off is again possible. For the bilayer systems used below, PMMA of two different molecular weights are used. These are referred to as "BDH" (lower molecular weight), and "Elvacite" (higher molecular weight).

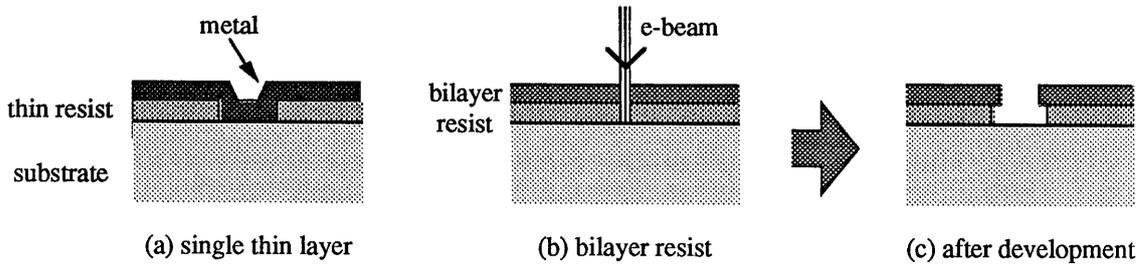


Figure 2.4: The lift-off process with thin bilayer resist. Lift-off is not possible with a single thin layer resist, as shown in (a). With a properly chosen resist combination, an overhang can be artificially induced in thin resist systems, as shown in (b) and (c).

2.2.3 Metallisation: ohmics and Schottkys

The predominant use of the lift-off technique is to pattern ohmic and Schottky contacts. It is also used to pattern interconnect metallisation (low resistance connections between two parts of a circuit) and masks for etching. Ohmic contacts are necessary in most devices, as these provide the electrical connection between the active parts of the semiconductor and the outside world. A contact is said to be a good ohmic if its I - V curve is linear around the origin, and the corresponding resistance is acceptable. Schottky contacts should exhibit diode characteristics, *viz.* high resistance under reverse bias, and lower resistance under forward bias. It is also desirable for Schottky contacts to have a high reverse breakdown voltage. [See Sze (1981), Ch. 5, for more information on the electrical behaviour of ohmics and Schottkys.]

For gallium arsenide (GaAs), the most popular ohmic systems are gold/germanium/nickel (Au/Ge/Ni) alloys [Braslaw (1983)]. Au is a natural choice for an ohmic since it is such a good conductor. However, it is found that almost any metal deposited on GaAs forms a Schottky, and Au is no exception. On the other hand when the underlying GaAs is very heavily doped, the Schottky characteristics deteriorate and will in fact, for sufficient doping, give good ohmic behaviour. Since Ge is a dopant for GaAs, the philosophy behind using a Au/Ge ohmic system is clear: the Ge should make the GaAs beneath the contact very heavily doped, and an ohmic should result.

To ensure that the Ge is incorporated into the GaAs and a proper alloy formed, the sample must be annealed at temperatures high enough to cause alloying of the Au/Ge and GaAs. For GaAs, the temperature at which alloying occurs varies with the ratio of

Au/Ge used. The temperature is found to be a minimum for a ratio of approximately 88%(Au):12%(Ge) by weight. This is the eutectic point, and corresponds to a temperature of about 340-360 C for GaAs. The incorporation of Ge into the GaAs in a straight Au/Ge alloy is not reliable, however. A small quantity of Ni (around 1-2% by weight) is therefore included in the recipe to "prepare" the GaAs for Ge incorporation. The combined Au/Ge/Ni recipe is found to work very well. [A more detailed account of ohmic formation can be found in Ogawa (1980). An alternative ohmic system involving the use of germanium/palladium has been demonstrated to work by Marshall *et al.* (1985).] For extremely heavily doped GaAs the precise ratios of Au/Ge/Ni become less important, and good ohmics are formed over a range of compositions around the eutectic point.

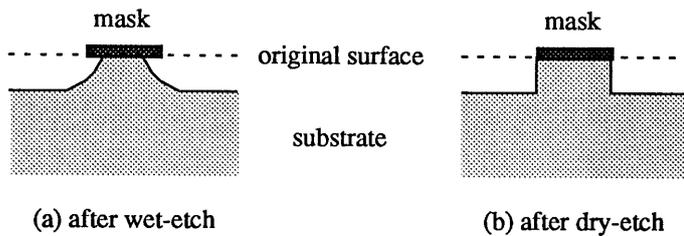


Figure 2.5: The effects of wet- and dry-etching. Wet-etches are generally isotropic, producing rounded profiles, whereas dry-etches can be made anisotropic, with much more vertical etch profiles.

2.2.4 Etching

Etching involves the removal of surface layers of the sample. This is usually done to define the active (*i.e.* the current-carrying) regions of the device. There are two basic methods of etching. Wet-etching involves the use of chemical solutions to remove material via a chain of reactions. In dry-etching, the sample is bombarded by a flux of ions. Surface removal here proceeds via physical sputtering, chemical reactions, or a mixture of both. Fig. 2.5 shows the outcome of typical wet- or dry-etches. The difference is manifest. In the case of the wet-etch, etching occurs in the lateral as well as the vertical direction: it is isotropic. This leads to an undercutting of the mask – highly undesirable, especially if the mask is very narrow. The dry-etch on the other hand occurs predominantly vertically: it is anisotropic. This makes it preferable over wet-

etching for defining very narrow features. However, the sputtering component of the dry-etch mechanism weighs against it. This difficulty does not occur in wet-etching.

Wet-etching can be done with many different kinds of etch solution. A review of some of the possibilities is given by Barnard and Blagden (1984). The basic etching mechanism is fairly simple. One component of the etch solution serves to oxidise the semiconductor surface, another component strips the oxide. The two processes occur in parallel, with a rate determined by the concentration of the various components. Typical GaAs etches will comprise three components: (i) hydrogen peroxide as the oxidising agent; (ii) hydrochloric acid, sulphuric acid or ammonium hydroxide as the deoxidising agent; and, (iii) de-ionised water to dilute the solution. In general, wet-etch profiles are much more complicated than indicated in Fig. 2.5(a). The profile will depend not only on the composition and concentration of the etch, but also on the material being etched, its crystallographic orientation, and sometimes also on the ambient temperature. The same type of etch in varying concentrations can give overcuts, undercuts, steps, *etc.* Refer to Barnard and Blagden (1984) for more details. New etch solutions should be tested before use.

Dry-etching can also be done with a variety of etching species and etching methods. The primary method used below is reactive-ion etching (RIE) [Bollinger *et al.* (1984) give a brief review of this]. A simplified schematic diagram of a RIE system is shown in Fig. 2.6. With the system under vacuum, the etch gas is introduced into the chamber. The top electrode is grounded, and the bottom electrode (the cathode) driven by an r.f. signal. This generates the plasma. The plasma in turn induces a dark space near the cathode, and a voltage drop (the d.c. self-bias) across the dark space. The d.c. self-bias is determined mainly by the strength of the r.f. signal (the etch power). The electric field due to the d.c. self-bias causes a downward flow of ions from the plasma across the dark space. If a sample is placed on the cathode, the ions will strike the sample from a vertical direction. Etching will occur in this direction if the ions react chemically with the semiconductor to form volatile compounds which are then pumped away. In practice, etching will also occur due to sputtering of the sample surface by the ions.

Both wet- and dry-etching require masks to be used to protect regions that are to remain unetched. For wet-etching the mask is usually resist. One major potential problem exists when using resist. If its adhesion to the surface is poor the resist will peel off in the etch, and the sample will be destroyed. Adhesion may be promoted by:

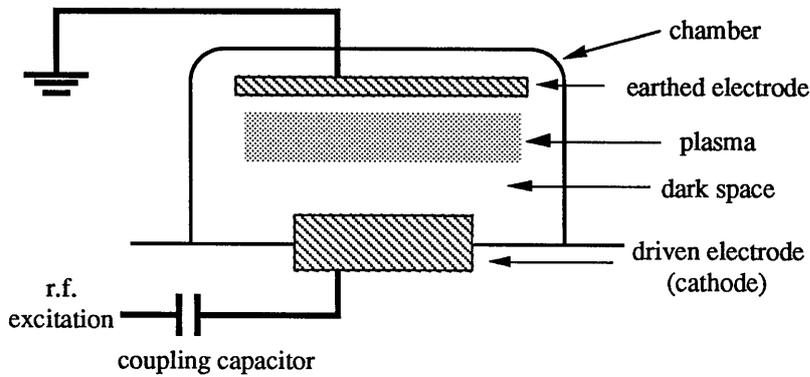


Figure 2.6: Schematic diagram of a reactive-ion etching chamber. Ions are accelerated across the dark space and impinge upon the sample which is placed on the cathode. Here the ions react with the semiconductor resulting in the semiconductor being etched.

- (i) ensuring a clean surface prior to spinning the resist;
- (ii) thoroughly baking the resist;
- (iii) not using strong etch solutions for too long;
- (iv) using the adhesion promoter HMDS (sometimes helps); and,
- (v) ensuring that the air humidity is not too high (important for some resists).

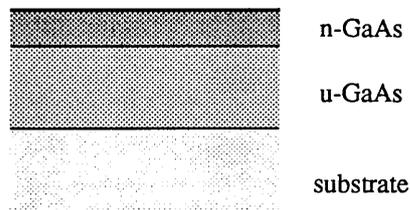
Metal masks can also be used.

The mask for dry-etching too can be resist or metal. The exact choice depends on the etch gas being used. The devices fabricated below have primarily been etched in a silicon tetrachloride (SiCl_4) plasma. Both HRN and nickel (Ni) masks have been used. Nichrome (NiCr) is also possible in place of the Ni. Several factors govern the choice for a particular pattern:

- (i) Ni is deposited using PMMA and lift-off. Lines defined by lift-off tend to be more uniform than lines made with HRN.
- (ii) HRN is easier to process – no metallisation step necessary.
- (iii) Ni is perhaps more robust than HRN for long etches, or etches done under a high d.c. self-bias. This is largely due to the non-rectangular profile of the HRN [*c.f.* Fig. 2.2(b)]. Parts of the HRN near the edges may become etched right through.
- (iv) After a SiCl_4 etch, the Ni reacts with any residual chlorine left in the chamber to form nickel chloride. This has a bubbly appearance, and is useless as a mask for further etching. HRN has no such problems.

- (v) The HRN mask cannot be removed easily after the etch. This will cause problems if the etch depth is required to be measured, and cannot be used if Schottky gates are to be subsequently deposited on the etched structure. The Ni can be stripped using hydrochloric acid.
- (vi) Lift-off has a high failure rate for extremely dense patterns. Here there is little choice but to use HRN. This is the main criterion for the use of HRN in Chapter 6.

Figure 2.7: Schematic diagram showing the structure of the n -GaAs epilayer used.



2.2.5 Process notes

All the devices to be described are made using GaAs, predominantly heavily n -doped epilayer material (Fig. 2.7) grown by the Glasgow University Molecular Beam Epitaxy (MBE) Group. Three lithographic levels are used to make the devices. The smallest features defined are etch masks for the actual wire structure. Ohmics are also defined, to allow current flow into and out of the structure. Finally, to allow easy wire-bonding to the sample, the ohmic lines are extended and terminated with large square pads. Fig. 2.8 shows a typical finished pattern. All steps are done by e -beam. In the notes below, resists are identified by their composition. The percentage value refers to the amount by weight of resist in the dissolving solvent (see §2.2.2).

- (i) Ohmic level. 15% BDH is spin-coated at 5000 rpm for 60 sec. This is then baked at 180 C for several hours. After pattern exposure, the resist is developed in 1:1 IPA:MIBK at 23 C for long enough to ensure that all the resist has been removed from the exposed regions (usually > 30 sec). The sample is then metallised with Au/Ge/Ni, lift-off being in acetone. Prior to metallisation, the surface oxide is removed by a short dip in 4:1 H₂O:HCl. This improves the quality of the ohmics. The ohmics are annealed at around 340 C for 30 sec on an electrically heated strip, in an atmosphere of argon/hydrogen (95%/5% by volume).

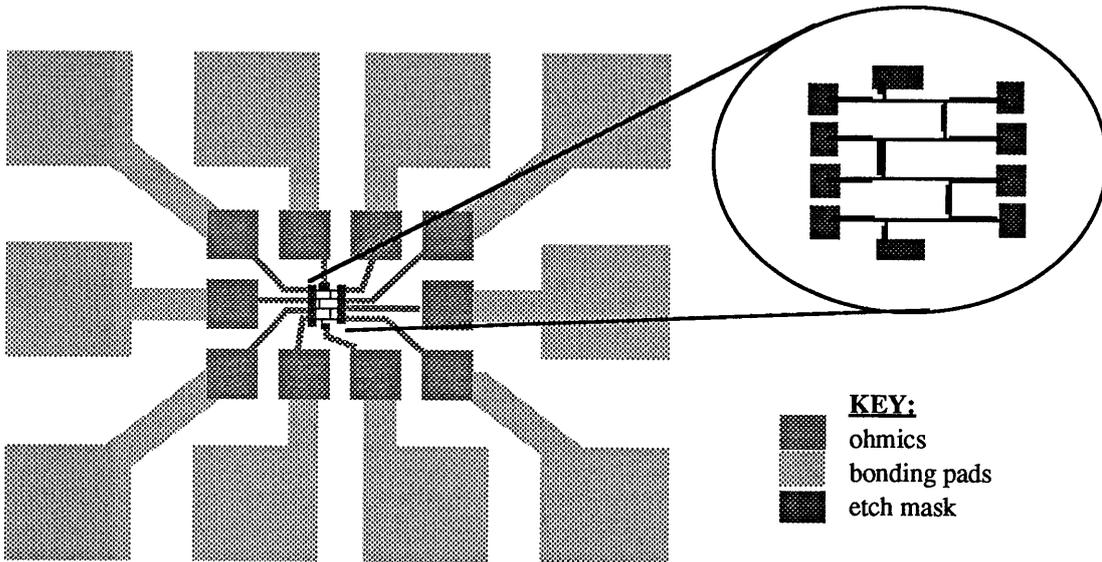


Figure 2.8: Pattern written by *e*-beam lithography. Three levels are defined. The wire level is magnified in the bubble. If Ni is used as the etch mask, it is removed after dry-etching.

- (ii) Dry-etch. Either bilayer PMMA or single layer HRN is used. To define a Ni mask, bilayer 4% BDH / 2.5% Elvacite is spun on at 5000 rpm for 60 sec. The BDH is baked at 180 C for about an hour, with a further bake of several hours after the Elvacite is spun on. The exposed sample is developed in 2.5:1 IPA:MIBK at 23 C for 1 min 20 sec. 30-40 nm Ni is then evaporated onto the sample followed by lift-off in acetone. No deoxidation etch is performed. If HRN is used, it is spun on at 7000 rpm for 60 sec and baked at 120 C for 30 min. It is developed by several short dips in neat MIBK, with thorough intermediate rinses in IPA. This minimises solvent absorption by the HRN, and is continued until the unexposed HRN is cleared. The sample is then dry-etched. If Ni is used, it is subsequently stripped in 1:1 H₂O:HCl. The dry-etching is generally done with d.c. self-bias voltages between 280-380 V (corresponding to etch powers of 100-150 W). The machine is a Plasmatech RIE 80. The chamber is subjected to an oxygen preclean prior to the etch run. The table temperature is maintained at 40 C.
- (iii) Bonding pads. These are deposited as in step (i). The anneal is not necessary here.

Steps (ii) and (iii) above are interchangeable. If the mask is defined as the final lithographic step, the chip may be cleaved prior to etching. This allows comparisons

between different etch conditions to be made using only one chip, thereby saving lithography time.

2.3 Measurement methods

2.3.1 Measurement of resistance

The basic method of measuring a resistance is to use Ohm's Law. Knowing the current (I) through the resistor, and the potential drop (V) across it, the resistance is given by $R = V/I$. The current can usually be measured quite accurately, so the problem reduces to one of finding the voltage. The two ways this can be done are shown in Fig. 2.9.

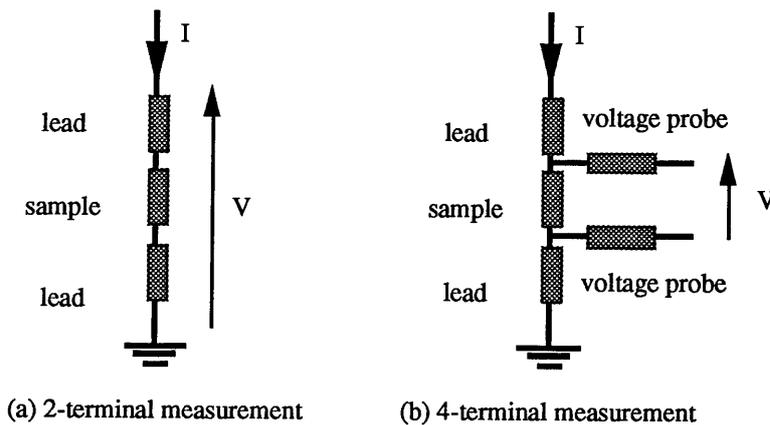


Figure 2.9: Diagrams illustrating 2- and 4-terminal resistance measurements. The 4-terminal measurement is preferred since errors due to lead resistances are minimised.

The first method typically uses stimulus-measurement units (SMUs). These can supply current and simultaneously monitor the voltage at which the current is being supplied. This is known as a 2-terminal resistance measurement. It is the quickest and easiest way to measure the resistance. However, the value it yields is not quite accurate. As seen in Fig. 2.9(a), the potential drop is measured not only across the sample, but also across the leads. Hence the resistance value measured is that of the sample plus

that of the leads. This is no problem if the lead resistance is an insignificant fraction of the sample resistance, but is obviously in error otherwise. The Hewlett Packard 4145B allows such measurements to be made fairly simply. This is used primarily for testing lead continuity and the state of the sample. Serious resistance measurements are rarely made this way, although exceptions do arise.

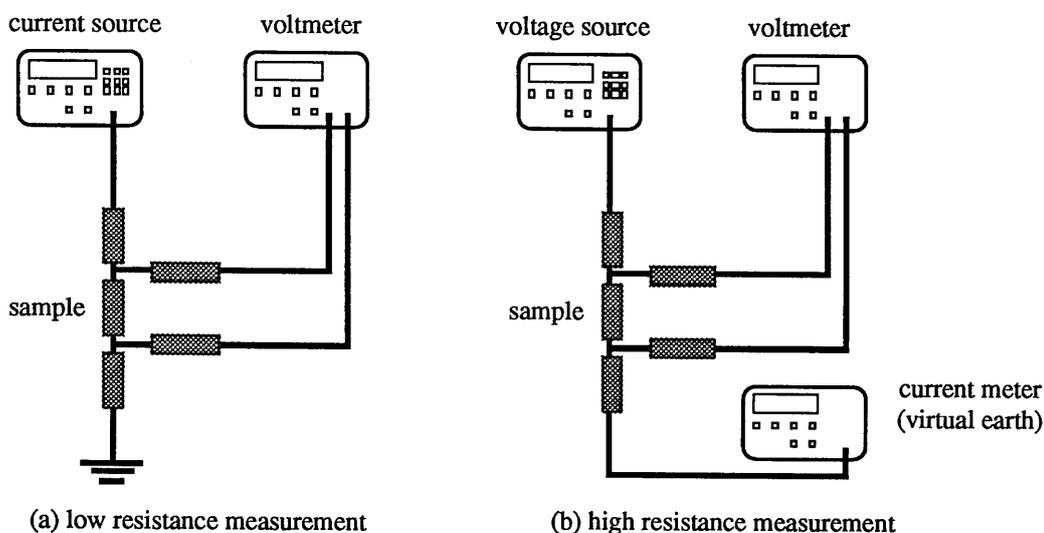


Figure 2.10: The two basic methods of performing 4-terminal resistance measurements. The method used depends on the voltage drop permissible across the sample.

The second method is a 4-terminal resistance measurement [Fig. 2.9(b)]. Here probes are brought out from either side of the sample, and a voltmeter attached to these. If the resistance presented to the circuit by the voltmeter (its input resistance) is sufficiently high, then it will not disturb the currents flowing in the rest of the circuit significantly. In this case, a reasonably accurate measurement of the potential drop across the sample, and hence of the sample resistance, can be made. Keithley and Hewlett Packard produce a range of instruments to supply and measure d.c. currents and voltages, which conform to the requirements for making the above measurements.

There are two ways of making 4-terminal resistance measurements, one appropriate for low resistance samples, the other for high resistance samples. In the former case, the circuit is driven by a current source [Fig. 2.10(a)], and in the latter by a

voltage source [Fig. 2.10(b)]. Adopting this procedure prevents either excessive currents passing through the sample, or excessive voltages being developed across it.

The d.c. measurements described above are fine for voltages much bigger than about $100\ \mu\text{V}$. Smaller voltages are much harder to measure since they tend to become engulfed in noise. Other problems also exist with d.c. measurements. If two dissimilar metals are brought together, a d.c. potential can exist across the junction. When the junction is in the voltage measuring part of the circuit, this leads to errors in the measured voltage. Such a junction also acts as a thermocouple. If the temperature drifts, so will the voltages. This adds further to the measurement errors. [See Keithley *et al.* (1984) for more details on low-level d.c. measurements.] Such problems are circumvented by using a.c. methods.

2.3.2 The lock-in technique

The methods of measuring resistance 2-terminally and 4-terminally using a.c. signals are analogous to the methods mentioned in §2.3.1. When using a.c. the principle being evoked is that of linear response: if a circuit is excited at a particular frequency, its response will also be at that frequency, with the strengths of excitation and response being linearly related. In practice the circuit may also respond at harmonics or sub-harmonics of the exciting frequency. In any case, if the amplitudes of the voltage across the sample and the current through it are known, the ratio gives the resistance. With two modifications any of the circuits shown in Figs. 2.9 and 2.10 can be used for a.c. measurements: the d.c. current/voltage sources should be replaced by signal generators; and the current/voltage monitors should be replaced by lock-in amplifiers (lock-ins).

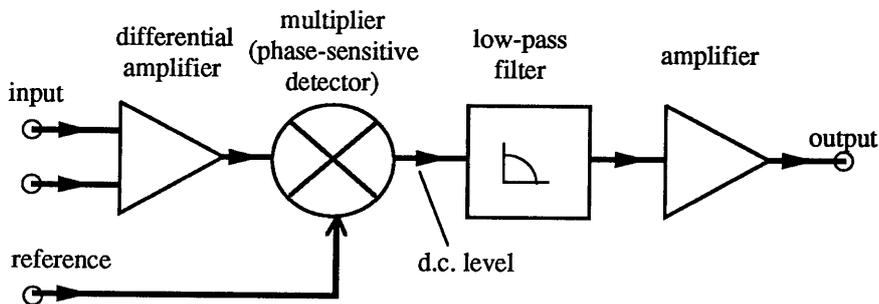


Figure 2.11: Block diagram of the main functional units of a lock-in amplifier.

A lock-in is basically a sensitive a.c. voltmeter. Fig. 2.11 gives a simplified schematic diagram of the main functional units of a lock-in. At the heart of a lock-in is a phase-sensitive detector (PSD) [Horowitz and Hill (1989), Ch.13]. Two signals are fed into the PSD: one is the input signal from the sample; the other is a reference sinewave derived from the excitation to the sample. The input signal will generally be a superposition of the sample's response to the excitation and the unwanted noise affecting the measurement. It is the job of the PSD to separate these two. Its output will be a d.c. level proportional to the amplitude of the sample's response to the excitation. This then goes through some smoothing and further amplification before appearing as a measurement.

Many sources of interference can affect low-level measurements (see next section for more details). In general, the amount of noise depends on the frequency being considered. Thus, certain ranges of frequency will be "quieter" than others. The frequencies at which these ranges lie depend primarily on the environment around the measuring apparatus. The whole idea behind using a.c. excitations and lock-ins is to try to make the measurement at a frequency where the effect of noise at other frequencies is minimised. This is the crux of the lock-in technique, and with care signals down to a few tens of nanovolts can be measured.

2.3.3 Interference in measurements

Two sources of measurement error have been mentioned already in §2.3.1: junction potentials and drift. The former is a d.c. voltage, and the latter can be regarded as extremely low frequency noise. For any reference frequency above a Hertz (Hz) or so, both of these signals will be ignored by the lock-in.

Capacitive and inductive coupling are further sources of noise. Capacitive coupling occurs when changing electric fields in the environment induce changing currents in the measuring circuit. This effect can be reduced somewhat by using coaxial ("coax") cables with BNC connectors. If the outer conductor of the coax is properly grounded, the core should be completely screened from stray electric fields outside. Inductive coupling occurs when changing magnetic fields in the environment induce currents in the measuring circuit. This effect can be minimised by using twisted pairs of coax cables. This reduces the loop area through which the magnetic field acts, and hence lowers the inductance of the circuit.

A potentially troublesome source of interference in the measurements is

microphonic noise. If the measuring circuit vibrates in any way, changing capacitances in the circuit induce changing voltages. The BNC connections tend to be the biggest source of microphonics, although the main body of the coax cable will also generate a small amount. This noise form can be minimised by ensuring that all cables are secured to a solid surface. A large component of microphonic noise can be eliminated by ensuring that the circuit is properly grounded.

Incorrect grounding is the most important remaining source of interference in low-level measurements. If the circuit is earthed at more than one point, ground-loops are formed. For low-level measurements the two ground points are actually at slightly different potentials, leading to errors in the measurement. Thus low-level measurements should be done with at most one ground point. Ground-loops are avoided by decoupling the ground in the measurement circuit from those of all the instruments (so the measuring circuit is completely floating), and grounding to an independent (not mains) earth.

Other sources of interference also exist. These include $1/f$ and shot noise which arise in any measurements. Either of these can affect the measurements, and must be "smoothed" out using the time-constant settings on the lock-ins. Mains noise, electromagnetic noise from other equipment in the lab, *etc.* should not present a problem provided a sensible excitation frequency is chosen (although excessive amounts of such noise can cause electron heating which is detrimental to low-temperature measurements).

2.3.4 Cryogenic systems and superconducting magnets

Many of the measurements require the sample to be at very low temperatures. Several methods are used to achieve this. The simplest is just to immerse the sample in a liquid helium (4 K) bath. This has been done for some of the measurements made in Chapter 6. More sophisticated methods require the use of cryostats. Here the temperature of the sample may be varied to an extent by balancing the flow of cold helium gas and the heat generated from a small heating element near the sample. Base temperatures are achieved by switching the heater off.

Some of the cryostats used incorporate medium-size superconducting coils in their design. The coils sit in the same helium that cools the sample. With this arrangement, moderate magnetic fields up to around 7 tesla can be applied to the sample. Magnetoresistance measurements can thus be made.

2.3.5 Automation of measurements

All the measurements are automated. The instruments are connected to a computer via IEEE (also called GPIB) or RS232 interfaces. The instruments can then be interrogated via software, and the data stored digitally for later analysis. The use of x-y recorders is generally cumbersome, and becomes unmanageable when several instruments have to be monitored simultaneously.

2.4 References

- ADAMS J A, Ph.D. Thesis, Glasgow University (1990).
- AGAR A W, R H Alderson, D Chescoe, *Principles and Practice of Electron Microscope Operation*, North Holland Publishing Co., Amsterdam (1980).
- BARNARD H A, G Blagden, "Wet chemical etching of Si, Ga(Al)As and Ga(In)As(P)," Library Search Report, GEC Hirst Research Centre (1984).
- BEAUMONT S P, T Tamamura, C D W Wilkinson "Two-layer resist system for efficient lift-off in very high resolution electron-beam lithography," *Microcircuit Engineering 80*, Elsevier, Amsterdam (1981).
- BINNIE C E, Ph.D. Thesis, Glasgow University (1985).
- BOLLINGER D, S Iida, O Matsumoto, "Reactive ion-etching: its basis and future," *Solid State Technology* May, 111 (1984).
- BRASLAU N, "Ohmic contacts to GaAs," *Thin Solid Films* **104**, 391 (1983).
- BROERS A N, "Resolution limits for electron-beam lithography," *IBM J. Res. Develop.* **32**, 502 (1988).
- HALLER I, M Hatzakis, R Srinivasan, "High resolution positive resist for electron beam exposure," *IBM J. Res. Develop.* **12**, 251 (1968).
- HATZAKIS M, "Semiconductor metallizing process," *IBM Tech. Disclosure Bull.* **10**, 494 (1967).
- HOROWITZ P, W Hill, *The Art of Electronics*, Cambridge University Press, Cambridge (1989).
- KEITHLEY J F, J R Yeager, R J Erdman, *Low-Level Measurements*, Keithley Instruments Inc., Cleveland, Ohio (1984).
- MACKIE W S, Ph.D. Thesis, Glasgow University (1984).

- MARSHALL E D, W X Chen, C S Wu, S S Lau, "Non-alloyed ohmic contact to n -GaAs by solid phase epitaxy," *Appl. Phys. Lett.* **47**, 298 (1985).
- OGAWA M, "Alloying behaviour of Ni/Au-Ge films on GaAs," *J. Appl. Phys.* **51**, 406 (1980).
- SZE S M, *Physics of Semiconductor Devices*, John Wiley & Sons, New York (1981).
- TODOKORO Y, H Watanabe, Y Hirai, N Nomura, M Inoue, "Self-aligned phase shifting mask for contact hole fabrication," in ed. G Declerck, L Van den Hove, F Coopmans, *Microcircuit Engineering 90*, Elsevier, Amsterdam (1991).
- WOLFSTÄDTER B, A Colquhoun, K Wegener, "Electron beam lithography of 100 nm T-gates for GaAs MESFETs," in ed. H Ahmed, *Microcircuit Engineering 89*, Elsevier, Amsterdam (1990).

Characterisation of dry-etch damage

3.1 Outline

Dry-etch damage is not completely understood. Although many techniques have evolved to characterise this damage, *i.e.* to say whether one etch process is better than another, the precise mechanisms of damage formation, and the exact nature of the damage itself, are not yet known. This chapter looks at some of the experimental aspects of dry-etch damage, and lays the groundwork for the development of a model for etch damage in Chapter 4. After introducing some basic facts about native defects in GaAs, some of the damage characterisation techniques that have been developed are described. The information some of these techniques yield on SiCl_4 -etch damage in particular is highlighted. The cut-off method is a straightforward technique of characterising sidewall damage. The method is applied to the case of SiCl_4 -etched n^+ -GaAs wires, yielding useful information. Some additional low temperature measurements on these wires give further supporting information.

3.2 Introduction

At some stage during device fabrication the active regions of the device need to be defined. This may be done either by damaging the surrounding areas (*e.g.* by boron implantation), or by etching. Both wet- and dry-etching can be used. However, when very small features need to be defined the isotropic behaviour of wet-etches can cause problems (see §2.2.4). For this reason much effort has gone into investigating the suitability of dry-etch methods for the fabrication of nanostructures. The main problem here is the electrical and optical damage that may be incurred by the material during etching. This is believed to be caused by the physical component of the etching mechanism (see §2.2.4).

That dry-etching, especially at high energies, can "damage" the underlying material is not in doubt. Many different techniques suggest that structural and compositional changes can occur in the material as a result of the etching. Although the body of knowledge is growing, a detailed understanding is as yet lacking on exactly how such damage propagates into the material, or indeed what form the damage takes. Such an understanding is desirable in certain instances, since the effects on device performance can then be assessed with greater confidence. When modelling transport through nanostructures, for example, it would be very useful to know exactly how the damage affects parameters such as mobility, or the mean free path. More realistic models may then be developed, leading to a better understanding of experimental observations.

In the next few sections, the current status of thinking on dry-etch damage is surveyed, and new experimental data on dry-etched n^+ -wires is presented. This, together with evidence from recent DLTS and surface-etching measurements, allows a phenomenological model for the distribution of defects in a dry-etched structure to be developed in the next chapter.

3.3 Defects in semiconductors

3.3.1 Lattice defects

Defects come in many different forms. Before attempting to understand better the nature of dry-etch damage, a summary of current knowledge on defects and defect states is given. Defects can be separated into two broad categories: point defects and

extended defects. The difference is simply in spatial extension, with point defects being very much more localised.

Point defects include vacancies, interstitials, and substitutional impurities such as the Si-donor in GaAs. Their presence in the host lattice can have a substantial effect on the material properties. The aforementioned Si impurity in GaAs is easily ionised, thereby enhancing electrical conduction. Cr-doping works in reverse, trapping electrons and making the GaAs highly resistive. Optically, defect levels within the energy gap may act as non-radiative recombination centres (*e.g.* Au-doped silicon), or as radiative recombination centres (*e.g.* N-doped GaAsP). Some of these effects are desirable, others are detrimental.

The properties of point defects are usually understood by analysing the defect potential. If the defect potential is weak, electron states centred on the defect well extend for several lattice spacings in all directions. In such cases, the effective mass Schrödinger equation describes the defect properties very well [see for example Stoneham (1975) for further details of the effective mass approximation]. If the defect potential is weak at long distances, but significant in the unit cell within which the defect is situated, the effective mass Schrödinger equation must be augmented by a "central cell correction." As the defect potential becomes more singular, other techniques must be used to study the defect properties. Such defect states are generally termed deep levels, even though their exact energy may not be "deep" [see Jaros (1982)].

Extended defects include many types of dislocations and faults in the lattice. Dislocation loops, for example, may form if a small region of crystal is inserted into the lattice (or removed from it). Hayes and Stoneham (1985), Ch. 3, and Kittel (1986), Ch. 20, give further examples. Extended defects of this nature act as sinks for point defects and free carriers. As such, their presence in significant numbers will usually have deleterious consequences on material properties. Detailed modelling of extended defects is much harder than that of point defects, and only general properties can usually be calculated [extreme examples of which may be found in Klienert (1989)].

3.3.2 Point defects in GaAs

There are six types of intrinsic point defect in GaAs: Ga and As vacancies, interstitials or antisites. The usual notation for these is to write X_Y where X is the defect type, and Y is its location. Thus, V_{As} denotes a vacancy on the arsenic sublattice, Ga_i a gallium interstitial, and As_{Ga} an arsenic on the gallium sublattice (arsenic antisite).

Interaction between intrinsic point defects will produce point defects with different characteristics [see Bourgoin *et al.* (1988) for a review]. Thus, for example, the complex $\text{As}_{\text{Ga}}\text{-As}_i$ (arsenic antisite coupled to an arsenic interstitial) is currently believed to be a possible structure for the deep trap EL2 (see §3.3.3).

The properties of the intrinsic point defects and their complexes have been extensively studied over the years, since they can be produced in a controlled fashion by bombarding the GaAs with an energetic beam of electrons [Pons and Bourgoin (1985)]. An impinging electron, with energy up to 1 MeV, dissipates its energy to the lattice in a series of collisions, producing Frenkel pairs (interstitial coupled to a vacancy, *e.g.* $V_{\text{As}}\text{-As}_i$). The density of defects created in this way can be calculated (theoretically) as a function of incoming electron energy. It has been found that the gallium interstitial tends to be quite mobile, and that the Frenkel pair $V_{\text{Ga}}\text{-Ga}_i$ readily recombines. Electron irradiation thus generally creates defects on the arsenic sublattice. Lim *et al.* (1987) have directly measured the energy required to create a Frenkel pair to be around 10 eV. This is in accord with theoretical estimates.

Intrinsic point defects have been identified and characterised by a variety of methods [Bourgoin *et al.* (1988)]. Of these, the twin techniques of electron paramagnetic resonance (EPR) and electron nuclear double-resonance (ENDOR) can provide a unique signature for the defect. Both rely on the electron-nuclear spin-spin interactions, within the defect (EPR), and between the defect and its surroundings (ENDOR); and since the interactions are sensitive to the exact structure of the defect, the constituent components of the defect can in principle be deduced [see also Hayes and Stoneham (1985)]. The electrical properties of the defect can be found from a variety of techniques, the best of which is perhaps deep-level transient spectroscopy (DLTS) [see Lang (1974) for a description of this], which can give information such as defect concentration, energy level and capture cross-section, and identify whether the defect is a majority or minority carrier trap. These and other methods have together allowed a reasonable level of understanding to be attained about the simple defects in GaAs [Bourgoin *et al.* (1988)].

3.3.3 The DX-centre and EL2 in GaAs

The two interesting deep-level traps in GaAs are the DX-centre and EL2. The DX-centre was first observed in low temperature Hall measurements in AlGaAs

[Nelson (1977)]. Many models for this defect have been proposed, and its exact origin is still debated [Baraff (1991)]. The *DX*-centre has several distinguishing characteristics. Nelson (1977) observed that at temperatures of less than about 60 K, electrons excited out of the defect by optical pumping were not readily recaptured upon removal of the light (persistent photoconductivity, or PPC). It is found that electrons trapped in *DX*-centres can be emitted by thermal excitations of ~ 0.1 eV, or optical excitations of ~ 1.0 eV. Since thermal excitations involve lattice vibrations, whereas optical excitations do not (Franck-Condon transition), a large lattice relaxation is believed to occur when the *DX*-centre becomes filled [Lang and Logan (1977)]. This *et al.* (1985) observed that the *DX*-centre must be related to the donor impurity, *i.e.* the usual shallow donor state becomes a deep level when it is *DX*-like. Finally, the *DX* ground state gives the appearance of being related to the band structure, since it is only observed when the subsidiary *L*-valley in AlGaAs (or GaAs under pressure) becomes the minimum [*e.g.* Mizuta *et al.* (1985)]. However, Chadi and Chang (1988) have argued that this is only an illusion, showing that this apparent dependence can have another cause.

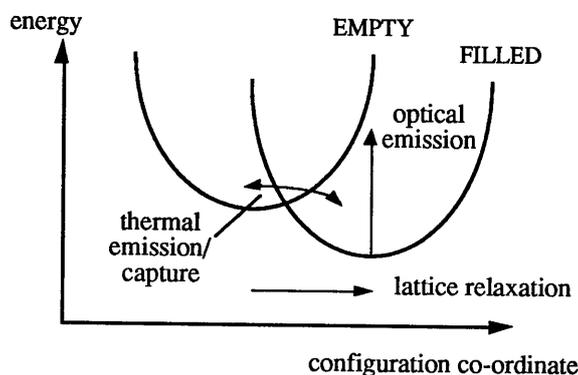


Figure 3.1: Simplified configuration co-ordinate diagram for the *DX*-centre. When the defect fills, a change in the configuration co-ordinate results. This represents the lattice relaxation.

The properties of many defects can be summarised using configuration co-ordinate diagrams [see Hayes and Stoneham (1985), Ch. 2]. Lang and Logan (1977) drew the first configuration co-ordinate diagram for the *DX*-centre, with an associated large lattice relaxation as shown in Fig. 3.1. This shows that the *DX*-centre has a barrier to both emission and capture. Lang *et al.* (1979) explained this with a model

where the DX -centre was a donor-vacancy complex (*e.g.* Si-V_{As}). However, recent models treat the DX as arising solely from the Si-donor, arguing [*e.g.* Morgan (1986)] that a donor-defect complex is inconsistent with measurements of DX -centres in GaAs under hydrostatic pressure. Currently favoured models stem from the work of Chadi and Chang (1988). In these models, the DX -centre is formed when a Si-donor in a substitutional site is displaced in the direction of an interstitial site, resulting in *two* electrons becoming trapped (*i.e.* the DX is negatively charged, not neutral as previously thought). Although problems still exist, the evidence in favour of such a model is mounting [see Mooney (1991)].

The structure of the defect EL2 is also still debated, currently thought to be either the As_{Ga} antisite, or the $\text{As}_{\text{Ga}}\text{-As}_i$ complex. Bourgoin *et al.* (1988) give its configuration co-ordinate diagram and a summary of its properties. Like the DX -centre, EL2 is donor-like with an associated large lattice relaxation. Its main distinguishing characteristic is its presence in bulk (not MBE-grown) GaAs in concentrations of $10^{19}\text{-}10^{20}\text{ m}^{-3}$ where it is believed to be responsible for the semi-insulating nature of the material.

3.4 Dry-etching with SiCl_4

3.4.1 Characterisation of dry-etch damage

Identification of the defects involved in dry-etch damage has proved to be very difficult, and no clear picture is as yet available. The situation is complicated by the fact that many different types of etching gases and mechanisms are used, so the type of defects introduced may vary from case to case. A great variety of techniques have been used to study this problem, with many yielding the simple qualitative observation that the material has slightly modified properties once it has been etched. In some cases, phenomenological models have been developed to try to explain the observations, and indeed this is the basic approach adopted in the next chapter. None of these models, however, gives a detailed microscopic understanding of the defects themselves.

Electrical, radiation and imaging techniques have been used to characterise damage. Taneya *et al.* (1989) have used photoluminescence to show that etching can introduce non-radiative recombination centres to a significant depth ($\sim 0.7\ \mu\text{m}$ in their study). Raman spectroscopy, which utilises the phonon-photon interaction in crystals,

has been used by Lishan *et al.* (1989) to show that the surface depletion depth in GaAs can be increased by the etching. Watt *et al.* (1988) have also used Raman spectroscopy to show that the crystal structure at the etched surface is slightly modified.

Modification of the etched surface is also indicated by measurements of barrier height, leakage current and breakdown voltage of Schottky diodes [*e.g.* Pang *et al.* (1983)]. Thoms *et al.* (1986) have shown that depletion regions at etched sidewalls are larger than would be expected from a knowledge of the free-carrier concentration in the material. Finally, using transmission electron microscopy (TEM) Cheung *et al.* (1990) have directly observed that etching modifies the stoichiometry of the material near etched sidewalls.

The above methods have largely been used to assess the performance of different types of etch with respect to how much damage they create. It has been found that damage is greatest for sputtering type etches (such as those involving the noble gases), and decreases as the etch is made more "reactive" [see, for example, Lishan *et al.* (1989)]. With reactive-ion etching the damage is found to decrease as the energy of the bombarding ions is reduced, being almost negligible for the recently developed ECR technique [*e.g.* Cheung *et al.* (1989)].

3.4.2 SiCl₄ reactive-ion etching

The basic configuration used for silicon tetrachloride (SiCl₄) etching is shown in Fig. 2.6. Chlorine radicals formed in the plasma are transported across the dark space, whereupon they react with the GaAs to form volatile gallium and arsenic chlorides (GaCl₃ and AsCl₃). The products are pumped away and new etch gas is fed into the system in a continuous cycle.

Since the basic chemical reaction relies on the presence of the chlorine, other chlorine-containing gases (such as boron tetrachloride or carbon tetrachloride) can also be used. Yet further reactive-ion processes use gases such as methane/hydrogen, Freon 12, *etc.* The underlying theme behind the use of all of these processes is that some volatile gallium or arsenic compound is formed. Silicon tetrachloride provides a relatively fast, highly anisotropic etch, removing typically 200 nm per minute of GaAs, and producing vertical sidewalls and a smooth (though not always perfectly flat) etched surface. The etching in this thesis is done exclusively with SiCl₄.

3.4.3 Characterisation of SiCl_4 -etch damage

In addition to the chemical reactions mentioned above, it is also possible for the incoming ions to collide with the GaAs surface, and dissipate energy to it. The resulting damage has been characterised by many of the techniques outlined in §3.4.1. For later reference, it is useful to collect together specific pieces of information on SiCl_4 -etching obtained from these investigations.

TEM observations by Foad *et al.* (1991) show that the defect density decreases roughly monotonically going away from the etched sidewall. XPS measurements by Foad (1992a) show that the etching causes arsenic depletion in the top few nanometres of the material, which appears to tie in with the observations of non-stoichiometry from the TEM. Measurements of sheet resistance as a function of etch depth by Foad *et al.* (1992b) appear to show a slight difference between wet-etching and SiCl_4 dry-etching.

Perhaps the most revealing information has been obtained by Johnson *et al.* (1992) using DLTS on SiCl_4 -etched n -GaAs surfaces. Five distinct defect levels were resolved, labelled N1-N5, all being majority-carrier (electron) traps. Measurements of defect density show a roughly exponential dependence with depth for most of the traps. None of the detected traps corresponds to any of the simple defects discussed in §3.3.2, although these DLTS measurements were not capable of detecting very shallow traps (*e.g.* V_{As}).

3.5 Experiments on SiCl_4 -etched n^+ -wires

3.5.1 The cut-off method

Thoms *et al.* (1986) demonstrated a ready technique of measuring sidewall depletion depths. Consider the conductance of the rib of conducting material ("wire") shown in Fig. 3.2. This has been formed from epitaxial n^+ -GaAs on an undoped GaAs buffer, etched to give the required shape (see §3.5.2).

The conductance of the wire is measured by applying a current through the sample, and measuring the voltage. The measured value is related to the dimensions of the wire and the depletion depths by the resistance formula $R = \rho L/A$. From Fig. 3.2(b), the conducting cross-sectional area is, $A = (w - 2x_d) \times (t - x_0)$, where t is the epilayer thickness, x_d is the sidewall depletion depth, and x_0 is the top surface depletion depth (assumed to be unaffected by the etching). The depletion region at the interface is

almost entirely in the μ -GaAs, and so may be ignored (as is affirmed by the numerical simulations in the next chapter). Knowing the sheet resistance, R_s , of the material gives the conductance,

$$G = \frac{1}{R_s L} (w - 2x_d), \quad R_s = \frac{1}{ne\mu(t - x_0)}, \quad (3.1)$$

where n is the carrier concentration, and μ the electron mobility. In practice R_s may be obtained directly from a Van der Pauw measurement. An estimate of x_d is then obtained by measuring the conductances of wires of varying widths, and fitting Eq. (3.1) to the data.

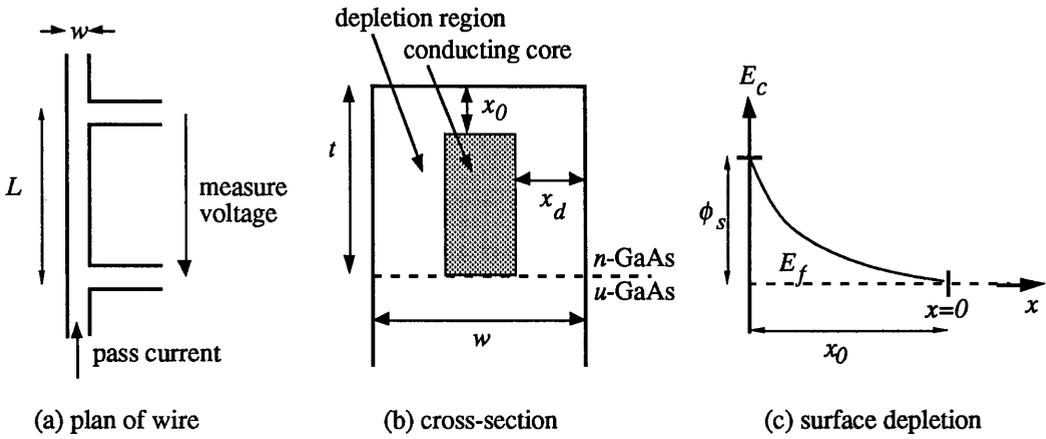


Figure 3.2: The wire conductance is found as shown in (a). The conducting part of the wire is the n -doped GaAs not including the depletion region, as shown in (b). (c) shows the profile of the depletion region caused at the surface of GaAs by the pinning of the Fermi level.

In Eq. (3.1), x_d is a phenomenological parameter. In principle, this distance could either represent a true depletion depth (obtained by solving the Poisson equation), or resistivity changes near the etched sidewall could be giving the *appearance* of a decreased conducting width. In practice, the number of defects necessary to give the required resistivity change is much bigger than the number required to give a true change in the depletion depth (according to the model to be developed in Chapter 4). Since the number of defects necessary for the latter is already of the same order of magnitude as the measured defect concentration (see Chapter 4), it will be assumed that resistivity changes are not the principle cause of the effects observed. Thus, x_d is

assumed to be obtainable from the Poisson equation.

In a completely undamaged sample $x_d = x_0$, where x_0 is obtained from the one-dimensional Poisson equation as shown in Fig. 3.2(c):

$$\frac{d^2 E_c(x)}{dx^2} = \frac{eN_D}{\epsilon}; \quad E_c(x_0) = \phi_s, \quad \left(\frac{dE_c}{dx} \right)_{x=0} = 0, \quad (3.2)$$

where E_c is the conduction band edge energy (in eV), N_D is the doping density, $\epsilon = \epsilon_0 \epsilon_r$, and ϕ_s is the surface potential arising from surface pinning [see Rhoderick and Williams (1988) for more details of surface effects]. In GaAs, the Fermi level is pinned near mid-gap at the surface, so $\phi_s \approx 0.7$ eV at room temperature. Solving for x_0 gives,

$$x_0 = \left(\frac{2\epsilon\phi_s}{eN_D} \right)^{1/2}. \quad (3.3)$$

Thus, an undamaged or wet-etched sample is expected to obey Eq. (3.1) but with x_d replaced by x_0 . In measured dry-etched wire samples x_d is found to be larger than x_0 . Mechanisms by which this is possible are discussed in Chapter 4, but the effect is attributed to dry-etch damage. The distance $2x_d$ is called the cut-off width, since it is the width at which the wire conductance should go to zero.

3.5.2 Sample fabrication

The samples are fabricated as outlined in §2.2.5. The etch mask consists of a set of four wires of length $L = 10 \mu\text{m}$ and of different widths (Fig. 2.8). The material used is heavily doped (around 10^{24}m^{-3}) MBE-grown *n*⁺-GaAs of thickness 50-85 nm. The material was predominantly grown by the Glasgow University MBE Group, although some layers were also obtained from IMEC. Some wet-etched wires were also prepared. Here the dry-etch step is replaced by two wet-etch steps, the first for large area isolation, the second to define the wires. PMMA was used as the mask, and 200:2:1 H₂O: NH₄OH:H₂O₂ as the etch solution. The sample was etched for about 3 minutes with the solution at 7 C. SEMs of the wires are shown in Fig. 3.3.

3.5.3 Experimental results on cut-off widths

The two quantities to be measured are the width of the wire, and its conductance. Since the depletion depth being inferred is only a few tens of nanometres in extent, both the width and the conductance must be measured with care. Errors greater than about 10-15% in either quantity would lead to unacceptable scatter in the data points, and consequently reflect poorly on the fitted value of x_d in Eq. (3.1).

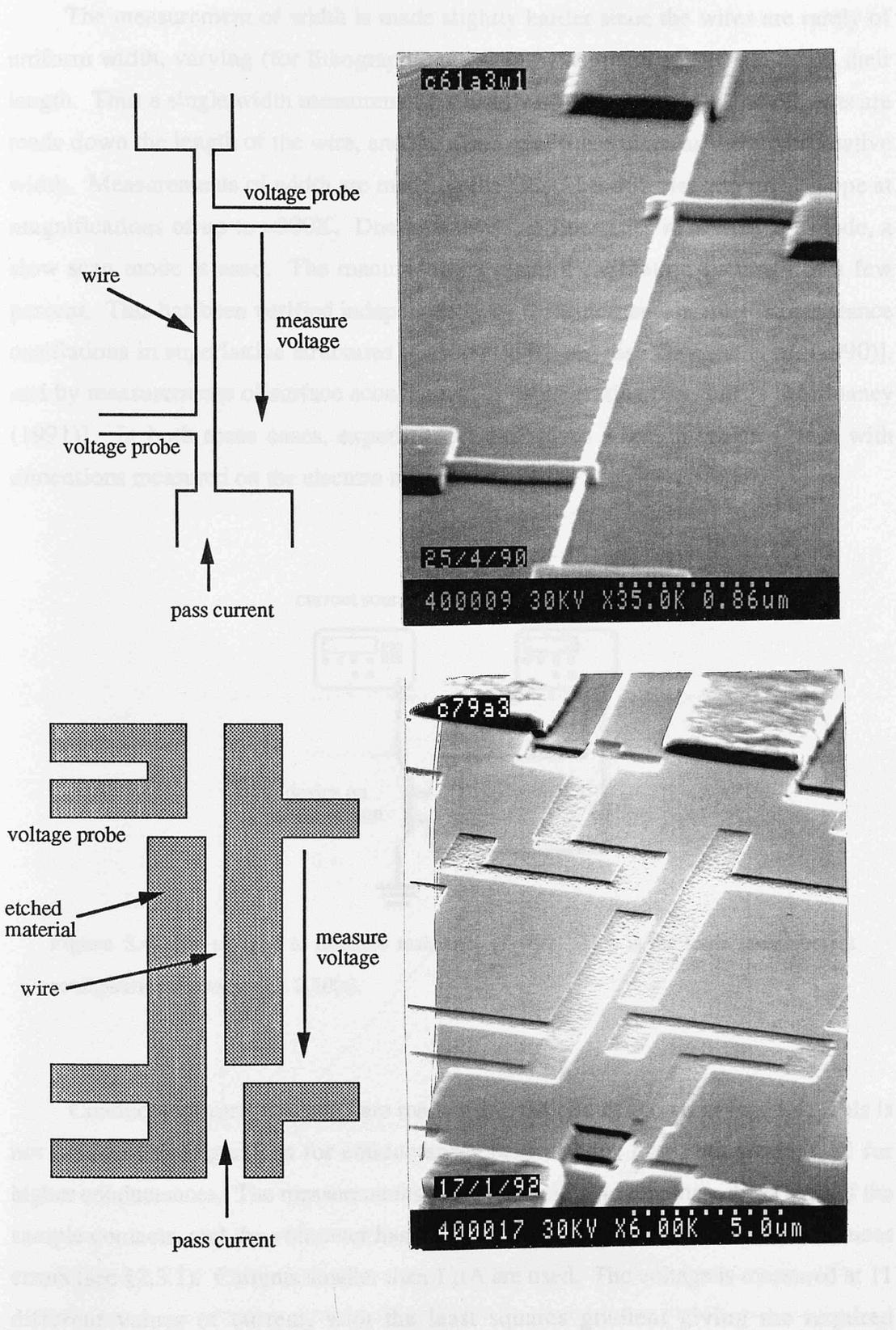


Figure 3.3: Above is an SEM of a dry-etched wire, below an SEM of a wet-etched wire. The patterns used allow 4-terminal measurements of the conductance to be made.

The measurement of width is made slightly harder since the wires are rarely of uniform width, varying (for lithographic reasons) by as much as ± 10 nm down their length. Thus a single width measurement is insufficient. Instead 11 measurements are made down the length of the wire, and the average of these taken as the representative width. Measurements of width are made on the Hitachi S-900 electron microscope at magnifications of up to $\times 300\text{K}$. Due to screen non-linearities in normal TV mode, a slow scan mode is used. The manufacturers claim a calibration accuracy of a few percent. This has been verified independently by measurements of magnetoresistance oscillations in superlattice structures [Cusco (1990); see also Davison *et al.* (1990)], and by measurements of surface acoustic waves using grating transducers [McEnaney (1991)]. In both these cases, experimental data gives a length which agrees with dimensions measured on the electron microscope.

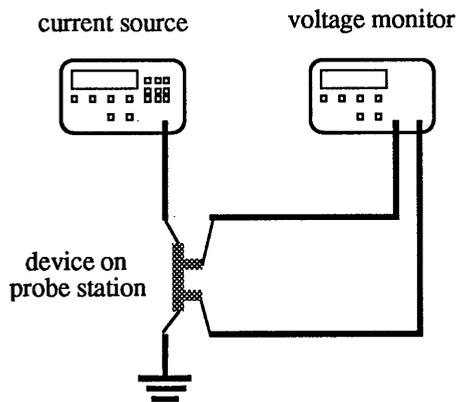


Figure 3.4: Set-up used to measure resistance of wire. This is the same measurement configuration shown in Fig. 2.10(a).

Conductance measurements are made using the circuit shown in Fig. 3.4. This is not the ideal configuration for conductances below about $1\ \mu\text{S}$, but works well for higher conductances. The measurements are four-terminal to eliminate the effects of the sample contacts, and the voltmeter has a high input impedance, which further reduces errors (see §2.3.1). Currents smaller than $1\ \mu\text{A}$ are used. The voltage is measured at 11 different values of current, with the least squares gradient giving the required conductance. The current is maintained sufficiently small to ensure that the I - V curve is always linear, *i.e.* negligible sample heating occurs.

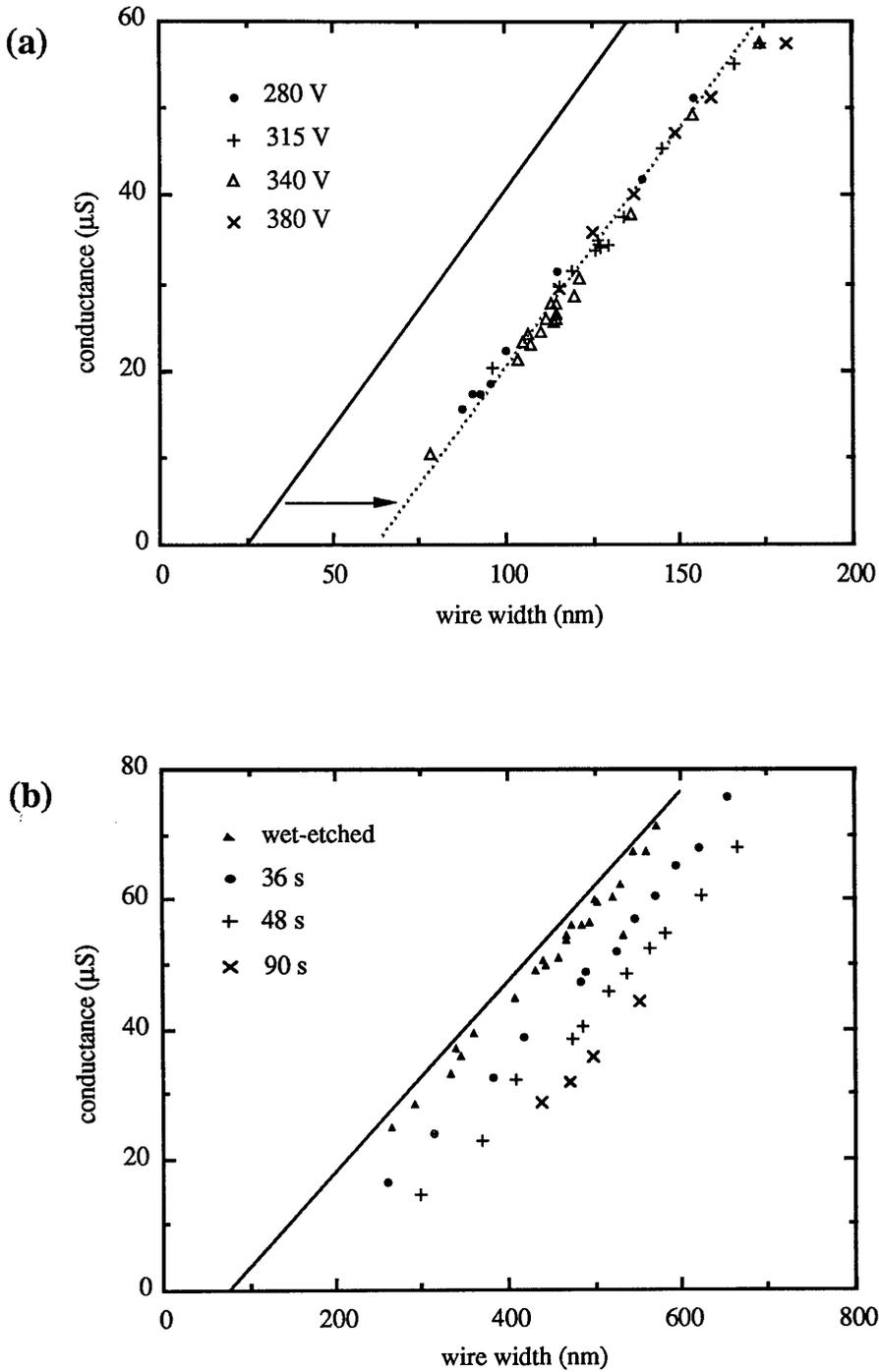


Figure 3.5: Effects of dry-etching on wire conductances. The solid lines represent Eq. (3.1) with $x_d = x_0$, where x_0 is given by Eq. (3.3) for the appropriate material doping density. Dry-etching is seen to reduce the conductance of a wire from its expected zero damage (wet-etch) value.

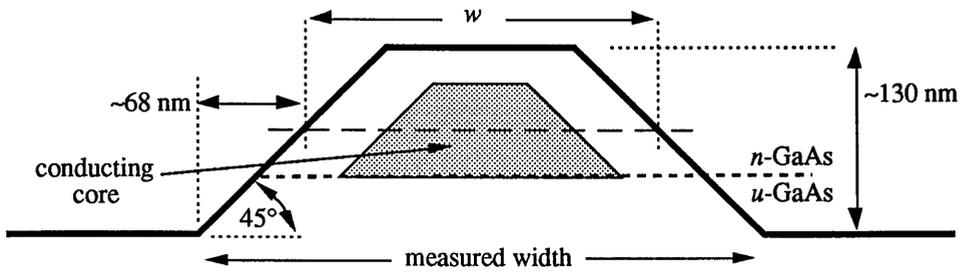


Figure 3.6: Schematic cross-section (not to scale) of a wet-etch wire. The epilayer thickness is 85 nm. The top surface depletion depth (zero damage) is ~ 40 nm. The width w is taken to be the width at the middle of the conducting part of the epilayer. This means that ~ 136 nm must be subtracted from the measured wire width to give w .

Fig. 3.5 shows experimental (room temperature) measurements of wire conductances for two different epilayers. Fig. 3.5(a) shows results for a 50 nm thick GaAs epilayer with carrier concentration $6.5 \times 10^{24} \text{ m}^{-3}$ and mobility $0.14 \text{ m}^2 \text{ V}^{-1} \text{ s}^{-1}$; and Fig. 3.5(b) shows results for a 85 nm thick epilayer with carrier concentration $6.5 \times 10^{23} \text{ m}^{-3}$ and mobility $0.31 \text{ m}^2 \text{ V}^{-1} \text{ s}^{-1}$. All the wires are $10 \mu\text{m}$ long. Fig. 3.5(b) shows the conductances of a set of wet-etched wires. Fig. 3.6 show the cross-sectional profile of one of these wet-etched wires. The sides of the wire are overcut by roughly 45° , and the etch depth is about 130 nm. The measured width (averaged over 11 separate measurements) is the width at the base of the wire, which is clearly not a good representative wire width. Instead, the wire width, w , is taken to be the mean width of the conducting part of the epilayer, *i.e.* ~ 136 nm less than the measured width. The solid lines in Fig. 3.5(a) and 3.5(b) represent Eq. (3.1) with $x_d = x_0$, *i.e.* the "zero damage" case. It is seen that the conductances of the wet-etch wires lie very close to this line. The difference is due in part to experimental uncertainty in the exact etch depth and angle of overcut, and in part to the fact that Eq. (3.1) is a 1-dimensional formula applied to a 2-dimensional problem. See also §4.4.3, where the 2-dimensional problem is solved in full.

Four sets of experimental points for dry-etched wires are shown in Fig. 3.5(a), obtained for self-bias voltages of 280 V, 315 V, 340 V and 380 V. The corresponding etch powers lie in the range 100-150 W. The etch time was 30 s for all the samples [with an etch rate of around $200(\pm 10) \text{ nm min}^{-1}$]. The change in bombarding energy

from 280 eV to 380 eV seems to have little effect on the measured conductance within experimental error, since all four sets of points lie roughly on the same line, cutting off at around 60 nm. The choice of Ni or HRN as the etch mask does not affect this result. Fig 3.5(b) shows wires etched at 310 V, but for times 36 s, 48 s and 90 s. In contrast to the case where the bombardment energy was varied, changes in the etch time *do* have a marked effect on the wire conductance, the cut-off width rising from 160 nm at 36 s, to about 250 nm at 90 s.

The fact that the only apparent difference between the lines representing Eq. (3.1) and the corresponding lines for the dry-etched wires is a shift along the x-axis (with little change in gradient), indicates that the sidewall depletion depth is increased as a result of dry-etching. The unchanged gradient vindicates the earlier assertion that the top surface depletion depth (and hence R_s) is unaffected by the etching. Thus, the dry-etching affects only the regions adjacent to the sidewalls.

From the two sets of data given in the two graphs, where the carrier concentration was varied by an order of magnitude, the relation,

$$x_d \approx (2.2 \pm 0.2) x_0 , \quad (3.4)$$

seems to approximately hold for an etch time of ~ 30 s and energies around 300 eV. Thus, it appears as though the ratio of the measured sidewall depletion depth (*i.e.* after dry-etching) to the nominal [Eq. (3.3)] depletion depth is approximately independent of the carrier concentration. Measurements on two-terminal wires at several intermediate doping concentrations also indicate that relation (3.4) approximately holds [A R Long, M Kinsler, I K MacDonald, J J Thompson, unpublished]. In the next chapter, this observation together with the measurements of Fig. 3.5, are explained in terms of a quantitative model.

3.5.4 Hot electron effects

As the above wires are cooled, their transport behaviour exhibits new features [Long *et al.* (1990)]. If the wire is pulsed with a high electric field for several seconds, the subsequent low field conductance is found to be reduced from its original value. The conductance then recovers back towards its original value with a rate dependent on the temperature of the sample. Fig. 3.7 shows typical traces for a 230 nm wide wire.

The important piece of information given by this measurement is that the predominant defect introduced by the etching is an electron trap, in agreement with DLTS measurements (§3.4.3). In principle, it may also be possible to deduce the trap

depth from measurements done at different temperatures. In the steady state, the free carriers are described by an equilibrium Fermi-Dirac distribution $f_0(E)$, as is the distribution of filled defect states. In a strong electric field, the conduction electrons in this heavily doped material are described by a new distribution, $f(E)$, with an electron temperature T_{el} different from the lattice temperature. [Nag (1980), Ch. 11, gives further details of the description of hot electron distribution functions.]

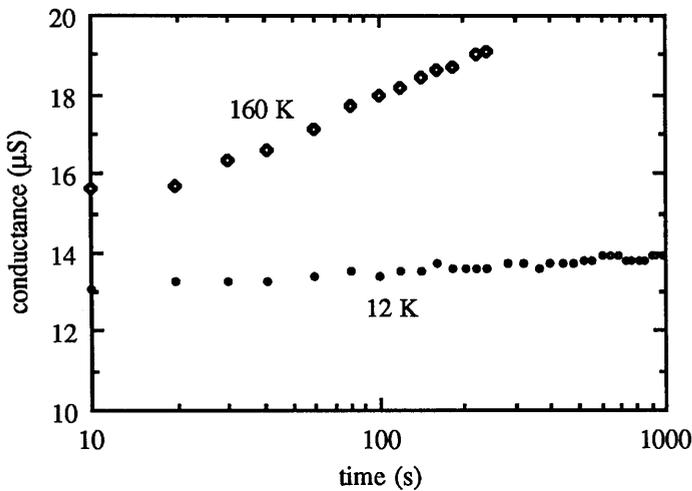


Figure 3.7: Conductance of a 230 nm wide wire (fabricated from $\sim 3.0 \times 10^{24} \text{ m}^{-3}$ carrier concentration material) as a function of time, after a $\sim 3 \times 10^5 \text{ V m}^{-1}$ field pulse had been applied across it for 60 s. The rate of recovery decreases at lower temperatures.

If the field is maintained for any length of time, the distribution of trapped electrons in the defect states, $f_t(E)$, will start relaxing towards the new distribution $f(E)$. If the field is then removed, the *conduction* electron distribution will relax to the steady state value much faster [Main *et al.* (1990) estimate that this occurs in a few nanoseconds for n^+ structures] than the *trapped* electron distribution. Since the trapped charge affects the depletion depth and hence the conductance, curves similar to that of Fig. 3.6 are expected. Also, since lower temperatures affect emission and capture rates from defect levels, the observed rate of recovery should also change. A more quantitative analysis of the effect is possible using a Shockley-Read-Hall type of analysis [Shockley and Read (1952)] which assumes negligible lattice relaxation associated with the defects, but this is not presented here. The main problem with the

analysis is that the exact heated electron distribution is unknown.

3.5.5 Photoconductance effects

If the above wires are pulsed with light instead of a high field, then the conductance transient is observed to go the opposite way, *i.e.* the conductance is initially increased from its original value, then decays back towards the steady state [Long *et al.* (1990)]. At low temperatures, the decay back to the steady state is found to take an exceedingly long time. The effect is reminiscent of PPC associated with DX-centres (see §3.3.3), but a large lattice relaxation need not necessarily be invoked (although it is certainly possible). Instead, the fact that empty traps are in the depletion region spatially separated from the free carriers is sufficient to provide a barrier against recapture. However, as carriers slowly do become trapped the conductance decays back towards the steady state. These measurements, as in §3.5.4, demonstrate that the predominant defect levels are electron traps.

3.6 References

- BARAFF G A, "Comparison of three DX structural calculations presented at Thessaloniki," *Semicond. Sci. Technol.* **6**, B9 (1991).
- BOURGOIN J C, H J von Bardeleben, D Stiévenard, "Native defects in gallium arsenide," *J. Appl. Phys.* **64**, R65 (1988).
- CHADI D J, K J Chang, "Theory of the atomic and electronic structure of DX centers in GaAs and AlGaAs alloys," *Phys. Rev. Lett.* **61**, 873 (1988).
- CHEUNG R, Y H Lee, C M Knoedler, K Y Lee, T P Smith III, D P Kern, "Sidewall damage in n^+ -GaAs quantum wires from reactive ion etching," *Appl. Phys. Lett.* **54**, 2130 (1989).
- CHEUNG R, A Birnie, J N Chapman, S Thoms, C D W Wilkinson, "Evaluation of dry etch damage in nano-structures by direct transmission electron microscopic examination," *Microelect. Eng.* **11**, 591 (1990).
- CUSCO R, unpublished (1990).
- DAVISON M, R P Taylor, E S Alves, P H Beton, M Dellow, P C Main, L Eaves, M Henini, O H Hughes, S P Beaumont, C D W Wilkinson, "Temperature and angular dependence of magnetoresistance oscillations in a 2DEG subjected to a

periodic potential," *Physica B* **165&166**, 867 (1990).

FOAD M A, S Hefferman, J N Chapman, C D W Wilkinson, "Evaluation of sidewall damage in nanostructures etched in CH_4/H_2 and SiCl_4 , by direct TEM observation," *GaAs and Related Compounds 1990*, Adam Hilger Ltd., Bristol (1991).

FOAD M A, Ph.D. thesis, Glasgow University (1992).

FOAD M A, S Thoms, C D W Wilkinson, "A new technique for dry-etch damage assessment in GaAs," submitted (1992).

HAYES W, A M Stoneham, *Defects and Defect Processes in Nonmetallic Solids*, John Wiley & Sons, New York (1985).

JAROS M, *Deep Levels in Semiconductors*, Adam Hilger Ltd., Bristol (1982).

JOHNSON N, M A Foad, C D W Wilkinson, "Electron traps in SiCl_4 dry etched GaAs," unpublished (1992).

KITTEL C, *Introduction to Solid State Physics*, John Wiley & Sons, New York (1986).

KLEINERT H, *Gauge Fields in Condensed Matter Vol. II Stresses and Defects*, World Scientific, Singapore (1989).

LANG D V, "Deep-level transient spectroscopy: a new method to characterize traps in semiconductors," *J. Appl. Phys.* **45**, 3023 (1974).

LANG D V, R A Logan, "Large-lattice relaxation model for persistent photoconductivity in compound semiconductors," *Phys. Rev. Lett.* **39**, 635 (1977).

LANG D V, R A Logan, M Jaros, "Trapping characteristics and a donor-complex (DX) model for the persistent-photoconductivity trapping center in Te-doped AlGaAs," *Phys. Rev. B* **19**, 1015 (1979).

LIM H, H J von Bardeleben, J C Bourgoin, "Experimental determination of the energy stored by a vacancy-interstitial pair in GaAs," *Phys. Rev. Lett.* **58**, 2315 (1987).

LISHAN D G, H F Wong, D L Green, E L Hu, J L Merz, D Kirillov "Dry etch induced damage in GaAs investigated using Raman scattering spectroscopy," *J. Vac. Sci. Technol. B* **7**, 556 (1989).

LONG A R, I K MacDonald, M C Vernon, M Rahman, J J Thompson, M Kinsler, S P Beaumont, C D W Wilkinson, C R Stanley, "Field-induced resistivity and persistent photo-conductivity in quantum wires," unpublished (1990).

MAIN P C, R P Taylor, L Eaves, S Thoms, S P Beaumont, C D W Wilkinson, "Conduction in n^+ -GaAs wires," in ed. S P Beaumont, C M Sotomayor-Torres, *Science and Engineering of One- and Zero-Dimensional Semiconductors*,

Plenum Press, New York (1990).

McENANEY K, Ph.D. Thesis, Nottingham University (1991).

MIZUTA M, M Tachikawa, H Kukimoto, S Minomura, "Direct evidence for the *DX* center being a substitutional donor in AlGaAs alloy system," *Jap. J. Appl. Phys.* **24**, L143 (1985).

MOONEY P M, "Donor-related levels in GaAs and AlGaAs," *Semicond. Sci. Technol.* **6**, B1 (1991).

MORGAN T N, "Theory of the *DX* center in AlGaAs and GaAs crystals," *Phys. Rev. B* **34**, 2664 (1986).

NAG B R, *Electron Transport in Compound Semiconductors*, Springer-Verlag, Berlin (1980).

NELSON R J, "Long-lifetime photoconductivity effect in *n*-type GaAlAs," *Appl. Phys. Lett.* **31**, 351 (1977).

PANG S W, G A Lincoln, R W McClelland, P D DeGraff, M W Geis, W J Piacentini, "Effects of dry etching on GaAs," *J. Vac. Sci. Technol. B* **1**, 1334 (1983).

PONS D, J C Bourgoin, "Irradiation-induced defects in GaAs," *J. Phys. C* **18**, 3839 (1985).

RHODERICK E H, R H Williams, *Metal-Semiconductor Contacts*, Clarendon Press, Oxford (1988).

SHOCKLEY W, W T Read, "Statistics of the recombination of holes and electrons," *Phys. Rev.* **87**, 835 (1952).

STONEHAM A M, *Theory of Defects in Solids*, Clarendon Press, Oxford (1975).

TANEYA M, Y Sugimoto, K Akita, "Characterization of subsurface damage in GaAs processed by Ga⁺ focused ion-beam-assisted Cl₂ etching using photoluminescence," *J. Appl. Phys.* **66**, 1375 (1989).

THEIS T N, T F Kuech, L F Palmateer, P M Mooney, "Far-infrared spectroscopy of silicon donors in AlGaAs," *GaAs and Related compounds 1984*, Adam Hilger Ltd., Bristol (1985).

THOMS S, S P Beaumont, C D W Wilkinson, J Frost, C R Stanley, "Ultrasmall device fabrication using dry-etching of GaAs," *Microelect. Eng.* **5**, 249 (1986).

WATT M, C M Sotomayor-Torres, R Cheung, C D W Wilkinson, H E G Arnot, S P Beaumont, "Raman scattering investigations of the damage caused by reactive-ion etching of GaAs," *J. Mod. Optics* **35**, 365 (1988).

Model for SiCl₄-etch damage

4.1 Outline

The information presented in Chapter 3 clearly suggests that depletion regions in *n*-GaAs are extended as a result of SiCl₄-etching (at energies greater than about 200 eV). In this chapter, it is shown how this can happen if deep electron traps are introduced into the sample as a result of energy/momentum transfer from the bombarding ions to the lattice. Expressions for the distributions of these defects are derived for both top surface etching and sidewall etching. It appears that to a good approximation defect diffusion can be ignored when calculating these distributions.

Assuming these defects are deep electron traps, the conductances of etched structures can be calculated by solving Poisson's equation. This is done for both wires and epilayers. It is found that the model can quantitatively account for all the experimental data on wire conductances obtained in the previous chapter, when a single defect source function is defined for SiCl₄-etching of GaAs at a particular energy. Qualitative comparisons with experiments on surface etched epilayers indicate that the model may be extensible to forms of etch other than SiCl₄.

4.2 Models for dry-etch damage

4.2.1 Requirements of the model

Clearly it is the purpose of any model to explain the observed effects as consistently (internal and external) as possible. In the case being considered here, DLTS indicates the presence of at least five electron traps (hole traps, shallow states, and levels resonant with the conduction band are other possibilities) whose microscopic structure is uncertain. Therefore, there is insufficient evidence for a microscopic understanding of SiCl₄-etch damage formation and propagation, and any model must be built on the basis of phenomenology.

The basic purpose of the model will be to try to explain the experimental results of §3.5.3 on wire conductances, in as quantitative a manner as possible. This is non-trivial, since several quite different parameters were varied in the experiments. The model must explain the variation in cut-off width not only with carrier concentration, but also with etch time. The epilayer thickness was also varied in the experiments, but this is perhaps less significant. Since the evidence (see §§3.4.1 and 3.5.3) points to depletion regions in *n*-GaAs extending as a result of dry-etching, several mechanisms by which this is possible are studied in turn.

4.2.2 Effects of surface

The main surface parameter that affects the depletion depth, given by Eq. (3.3), is the surface potential ϕ_s . In the experiments of §3.5.3, the depletion depth is observed to increase by a factor of 2 or more. This requires a change in ϕ_s of a factor of 4 or more, *i.e.* from 0.7 eV to 2.8 eV, putting the Fermi level at the surface of the GaAs well outside the bandgap. This is clearly unphysical, both in relation to the band structure, and since it implies an enormous hole concentration at the surface of the sample [as a very rough estimate, $p \sim (2m_h^*(E_f - E_v)/\hbar^2)^{3/2} / 3\pi^2 \sim 10^{27} \text{ m}^{-3}$]. An increase in ϕ_s cannot therefore by itself explain the increase in depletion depth caused by dry-etching.

4.2.3 The DX-model

The initial interest in this model arose from the observations of PPC-like behaviour in §3.5.5 [Long *et al.* (1990)]. The apparent dependence of measured depletion width on doping density [*c.f.* Eq. (3.4)] added credence to the model. However, despite

providing a natural explanation for these observations, the model has several faults which become apparent after further analysis.

As a starting point, some form of DX -centre is required in the etched wire. According to current models of DX -centres in Si-doped GaAs (see §3.3.3), the DX level is not the ground state of the Si-donor in GaAs at atmospheric pressure, the level being resonant with the conduction band, about 0.26 eV above the conduction band edge. Thus, this type of DX -centre cannot be responsible for the effects of dry-etching on the conductance. Instead, the XPS data of arsenic depletion, and TEM observations of non-stoichiometry near the etched sidewalls (see §3.4.2) suggest – albeit inconclusively – that arsenic vacancies are introduced into the sample in significant numbers by the etching. Recalling the model of Lang *et al.* (1979), it is possible that the required DX -behaviour could be obtained from a $Si-V_{As}$ complex (see §3.3.3).

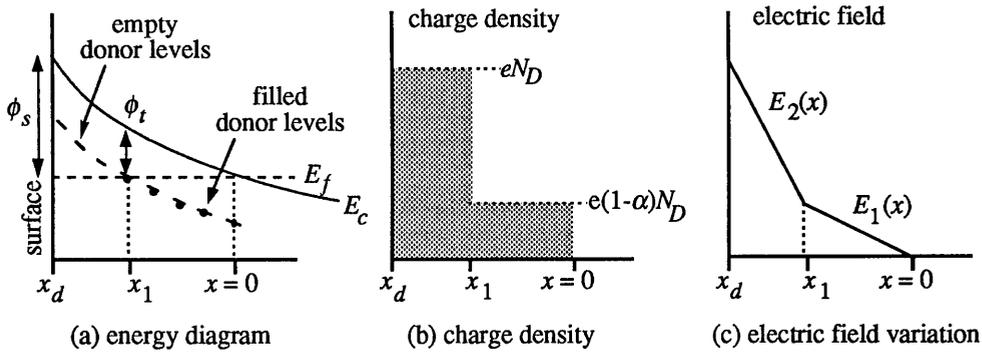


Figure 4.1: Illustration of how a DX -type defect would affect the surface depletion region at an etched sidewall. The defect energy levels in (a) are only schematic – the effect of lattice relaxation is not shown.

The apparent doping dependence of the measured depletion width seems to follow most naturally if the DX -type defect has the same charge states as the Si-donor. Thus, it would be positively charged when empty, but neutral when occupied by an electron. With this being the case, some simple analysis yields the depletion depth that would be expected. Consider Fig. 4.1(a), which shows the distribution of filled and empty defect levels near the sidewall. Such diagrams are ubiquitously used in studies of deep traps [see for example Kimerling (1974), or Rhoderick and Williams (1988)]. Let α denote

the fraction of Si-donors that have become coupled to an arsenic vacancy, and have become "deep." In the steady state, all such deep levels lying beneath E_f are assumed to be filled, as shown in the diagram. Hence these donors will become neutral upon capturing an electron, and the charge distribution will be as indicated in Fig. 4.1(b), with the corresponding field profile in Fig. 4.1(c). The fields (in units eV m⁻¹) are given by:

$$E_1(x) = \frac{e(1-\alpha)N_D}{\epsilon} x; \quad E_2(x) = \frac{eN_D}{\epsilon} (x-x_1) + E_1(x_1), \quad (4.1)$$

where N_D is the doping density, and $\epsilon = \epsilon_0\epsilon_r$. The potential is simply given by the area beneath these curves. Taking the Fermi level as the zero of energy gives conditions on the conduction band edge: $E_c(x_1) = \phi_t$ and $E_c(x_d) = \phi_s$, where ϕ_t is the trap depth, and ϕ_s is the surface potential. The areas are easily calculated giving,

$$\phi_t = \frac{e(1-\alpha)N_D}{2\epsilon} x_1^2; \quad \phi_s = \phi_t + \frac{e(1-\alpha)N_D}{\epsilon} x_1 x_2 + \frac{eN_D}{2\epsilon} x_2^2. \quad (4.2)$$

Solving for x_1 and x_2 gives the depletion depth:

$$x_d = x_1 + x_2 = \left(\frac{2\epsilon}{eN_D} \right)^{1/2} \left[\alpha \left(\frac{\phi_t}{1-\alpha} \right)^{1/2} + (\phi_s + \alpha\phi_t)^{1/2} \right]. \quad (4.3)$$

It is easily seen that the ratio x_d/x_0 , where x_0 is the "damage free" depletion depth given by Eq. (3.3), is independent of the doping density, which is the desired result.

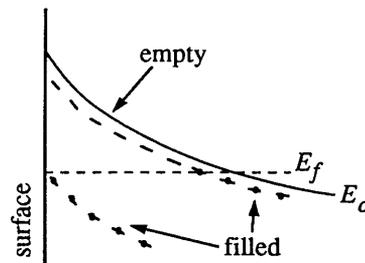
After this success, any attempts at further analysis exposes the weaknesses of the model. Eq. (4.3) for the depletion depth has two "adjustable" parameters (*i.e.* whose values are unknown): the fraction of donors that are coupled to arsenic vacancies, α , and the trap depth, ϕ_t . Although there would appear to be great freedom in what values these parameters might adopt, this is not the case. It is found that for almost any value of ϕ_t in the upper half of the energy gap, α takes values in excess of 0.8. In other words, in excess of 80% of the donors develop deep levels and are not ionised, even at room temperature. Such high values of α give cause for concern.

The formation of Si- V_{As} for such a large proportion of donors requires a high degree of correlation between the donor and vacancy positions. Assuming that the donors and vacancies are distributed randomly through the sample, and that the complex only forms for nearest-neighbour coupling, a rough estimate may be made of the density of V_{As} required to give the factor α . Since the Si-donor sits on the gallium sublattice, it is surrounded by four arsenic sites. If the probability of one of these sites being vacant is α , then roughly $\alpha/4$ of the arsenic lattice sites in the whole sample are vacant. For $\alpha > 0.8$, this involves around 20% of the arsenic in the crystal. Although overestimated, this number still seems unrealistically large.

A further weakness of the model lies in the fact that any of the above types of

defect lying within the conducting core of the wire, would trap electrons too. With such a large number of defects, a marked change in conductivity (or sheet resistance) should result from the etching, in disagreement with the observations of §3.5.3. The fact that α must be constant to several tens of nanometres [at least 120 nm from each sidewall for the 90 s data in Fig. 3.5(b)] is also not plausible, since this requires a *constant* density of V_{As} to exist over this depth, and there appears to be no mechanism by which this can occur as a result of dry-etching. For such reasons, a dry-etch damage model based on a donor-defect DX -type centre seems untenable.

Figure 4.2: Schematic drawing of energy levels of defect used to explain the effects of dry-etch damage.



4.2.4 Use of a two-level defect

The main problem with the DX -model is that the charge transition resulting from electron trapping, *i.e.* $Si^+ \rightarrow Si_{DX}^0$, does not affect the depletion region as strongly as required. This means that α must be unrealistically large to give the measured conductance values. A better solution to the problem is to consider the defect being introduced by the dry-etching as existing independently of the Si-donors, and having energy levels roughly as shown in Fig. 4.2.

Two energy levels are postulated, serving a two-fold purpose: (i) the lower energy level primarily determines the sidewall depletion depth, and (ii) the upper energy level is the one giving rise to the warm electron effects (§3.5.4) and gives the DLTS signal. The numerical analysis in the next few sections will show that such a defect energy level structure allows the results of §3.5.3 to be modelled very well. It should also be remembered that five defect levels are measured in the DLTS of $SiCl_4$ -etched GaAs (see §3.4.3), not just one. Thus, the defect in Fig. 4.2 is meant to be *representative* of the defects introduced by the dry-etching. That is, the postulated defect represents the

combined effect of all the defects that may actually be present.

Having postulated an energy level structure for the defect, conductance calculations further require the defect distribution in the etched structures to be known. Since defects can diffuse, the distributions will initially be obtained by solving the diffusion equation. However, by comparing the formulae to profiles measured by DLTS, it will subsequently be argued that defect diffusion can be ignored (to a good approximation) in the problem.

4.3 Defect distributions in etched structures

4.3.1 Defect formation and propagation

It was stated in §4.2.4 that a defect distribution will initially be sought from amongst the solutions to the diffusion equation. To find the desired solution, certain assumptions must be made about defect formation and propagation into the material (specification of boundary conditions, and source terms in the diffusion equation). The problem of defect formation has been widely studied previously. Radiation damage in the form of intrinsic point defects [Bourgoin *et al.* (1988)] has already been mentioned in §3.3.2. Formation of intrinsic point defects was also the favoured explanation of Dubonos and Koveshnikov (1990) in explaining data obtained by 2 keV bombardment of GaAs using argon ions. Typical energies used in SiCl_4 etching are around 300 eV (*c.f.* §3.5.3). Since Frenkel pair formation requires around 10 eV [Lim *et al.* (1987)], and line dislocation formation typically requires 10 eV per interatomic spacing [Hayes and Stoneham (1985), Ch. 3], the low energy bombardment used in SiCl_4 -etching is more likely to create point defects than extended ones. The problem of defect formation and propagation will be mathematically stated in two different ways, corresponding to two slightly different physical pictures.

In one picture, bombarding ions are envisaged to transfer some energy to the crystal surface leading to the formation of Frenkel pairs *solely* at the surface. These then dissociate, and the defects migrate into the crystal by diffusion. This case is treated as a nonhomogeneous boundary condition to the diffusion equation, *viz.* a non-zero defect flux pointing into the crystal is specified at the surface being etched. In the second picture, defects are created not only at the surface, but also some distance within the material. A number of studies [Taneya *et al.* (1989); Dubonos and Koveshnikov

(1990); Johnson *et al.* (1992)] appear to show an exponential distribution of defects going away from the etched surface. So this case is treated by including an exponential source term in the diffusion equation. The source term relates to the transfer of energy/momentum from the bombarding ion to the lattice, and may find some microscopic justification in future analyses. The precise mathematical formulation of the two cases, and the corresponding solutions, are given in the next two sections. Fits to DLTS profiles will then be used to determine which (if either) is more physical.

Of the two cases stated above, diffusion plays a much bigger role in the first to drive defects into the sample. Although Dubonos and Koveshnikov (1990) appear to observe significant defect diffusion, even at room temperature, the reproducibility of the width-conductance curves of §3.5.3 seems to suggest otherwise. If this is so, it would appear that the effects of diffusion should be negligible. However, this need not be so. During etching, any transfer of energy from the bombarding ions, which is below the threshold for creating defects, will dissipate as heat to the lattice (or the electronic states). Chadderton (1965), Ch. 3, shows that this may be treated as a problem in heat diffusion, and will result in the creation of a transient temperature field in the structure during etching. Since diffusion processes usually have a strong temperature dependence [see, for example, Boltaks (1963), Ch. 2], this represents a possible mechanism for enhancing diffusion during dry-etching. In solving for the defect distribution, however, the diffusion coefficient is assumed to be constant.

In addition to diffusion, other processes for defect motion exist which are ignored or are insignificant in the problem to be solved. The projected range of 300 eV ions is at most a few Ångstroms [*e.g.* Sze (1985), Ch. 10], so ion penetration can be safely ignored. Taneya *et al.* (1989) show that ion channelling (albeit for 10 keV Ga⁺ ions) should also be insignificant in such problems, although they do not discount secondary channelling effects. Other effects may not be so insignificant. If the diffusing defects are *charged* (as has been postulated in §4.2.4), then the electric field due to the donor ions and other defects may assist the diffusion process ["stress-assisted diffusion" – see, for example, Shewmon (1963), Ch. 1]. This would add a drift term to the diffusion equation, and would require the diffusion and Poisson equations to be solved in a self-consistent manner. The defect may also change charge state as its energy level crosses E_f , adding further complications. In principle, defects can interact (*e.g.* recombination of a vacancy-interstitial pair), requiring an entire series of coupled diffusion equations to be solved. Such a generalised approach has been used, for example, by Orłowski

(1988) to study impurity diffusion in silicon. All these latter complications are ignored.

In §§4.3.2-4.3.3 defect distributions are obtained both for etched top surfaces, and for etched sidewalls. The former can be fitted to DLTS profiles (which are measured on etched top surfaces), as well as being used in sheet conductance calculations for etched epilayers [*c.f.* measurements of Foad *et al.* (1992) - see §3.4.3]. The sidewall defect distributions are used in conductance calculations for dry-etched wires. The discussions later in this chapter will highlight important differences in damage formation mechanisms between top surface etching and sidewall etching (§4.4.3).

4.3.2 Defect distribution in one-dimension (top surface damage)

Two methods by which defect generation and migration into the etched sample can occur have been described in §4.3.1. It is the purpose of this and the subsequent sections to find which (if either) describes experiments better, primarily by comparing to the DLTS profiles measured by Johnson *et al.* (1992). In this section, the defect distribution for an etched epilayer is found.

If the finite size of the etched sample is ignored, its symmetry reduces the problem to that of solving the one-dimensional diffusion equation. However, this is not a standard boundary value problem since the surface at which the boundary conditions (*viz.* conditions on the defect flux) are applied, is moving. This is very similar to problems involving impurity redistribution during epitaxial growth or oxide growth (on Si) *etc.*, except that different boundary conditions are used [Runyan (1965)]. Analogous problems also exist in thermal physics, such as the melting of ice or laser ablation of metals, collectively known as Stefan problems [see, for example, Crank (1981), or Ockendon and Hodgkins (1975)], although the problem at hand is simpler to solve. In what follows, the defect density for top surface damage is written $\Gamma(y,t)$.

(a) In the first case, defects are created solely at the etched surface, and migrate into the material through diffusion. The surface is assumed to be being removed at a constant rate v (which is roughly true for SiCl₄, apart from a short induction time at the start). The moving boundary value problem is thus,

$$\left(D \frac{\partial^2}{\partial y^2} - \frac{\partial}{\partial t}\right) \Gamma(y,t) = 0; \quad y \geq vt, \quad t \geq 0, \quad (4.4a)$$

$$\Gamma(y,0) = 0, \quad -D \left(\frac{\partial \Gamma}{\partial y}\right)_{y=vt} = F_1, \quad \lim_{y \rightarrow \infty} \Gamma(y,t) = 0, \quad (4.4b)$$

where y is the vertical distance from the original surface of the sample [see Fig. 4.3(a)],

t is the time, and D is the defect diffusion coefficient. The restriction $y \geq vt$ allows for the motion of the surface. The flux is given by $-D \nabla \Gamma$ according to Fick's first law of diffusion (the diffusion equation being Fick's second law). The defects are assumed to be diffusing in from the surface with a constant flux of magnitude F_1 . The solution is obtained by transforming Eq. (4.4) into a drift-diffusion equation, and Laplace transforming in time (Appendix A.1):

$$\Gamma(Y, \tau) = F_1 \int_0^\tau dt' \left\{ \frac{1}{\sqrt{\pi D t'}} e^{-(Y+vt')^2/4Dt'} - \frac{v}{2D} \operatorname{erfc}\left(\frac{Y+vt'}{2\sqrt{Dt'}}\right) \right\}, \quad (4.5)$$

where $Y = y - vt'$ is measured from the etched surface [see Fig. 4.3(a)], and τ is the etch time.

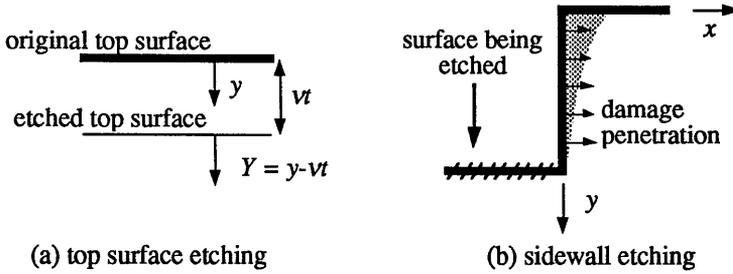


Figure 4.3: Co-ordinate systems used to calculate defect distributions. In sidewall etching, damage only penetrates through the region of sidewall exposed to the bombarding ion flux.

(b) In the second case, defect creation within the sample is allowed, with the defect flux at the etched surface being zero. Defect creation is described by the exponential source function in Eq. (4.6a) The problem here is stated as:

$$\left(D \frac{\partial^2}{\partial y^2} - \frac{\partial}{\partial t} \right) \Gamma(y, t) = -G_1 e^{-(y-vt)/\lambda}; \quad y \geq vt, \quad t \geq 0, \quad (4.6a)$$

$$\Gamma(y, 0) = 0, \quad -D \left(\frac{\partial \Gamma}{\partial y} \right)_{y=vt} = 0, \quad \lim_{y \rightarrow \infty} \Gamma(y, t) = 0. \quad (4.6b)$$

Again, the restriction $y \geq vt$ allows for the motion of the etched surface. G_1 and λ are phenomenological parameters describing the rate of defect creation and the mean creation depth (assuming a steady bombarding ion flux). The solution is obtained in a similar manner to the previous case (Appendix A.1):

$$\Gamma(Y, \tau) = \frac{G_1}{\omega} e^{-Y/\lambda} (1 - e^{-\omega\tau}) - \frac{DG_1}{\omega\lambda} \int_0^\tau dt' [1 - e^{-\omega(\tau-t')}] \times \left\{ \frac{1}{\sqrt{\pi Dt'}} e^{-(Y+vt')^2/4Dt'} - \frac{v}{2D} \operatorname{erfc}\left(\frac{Y+vt'}{2\sqrt{Dt'}}\right) \right\}, \quad (4.7a)$$

where,

$$\omega = \frac{v}{\lambda} - \frac{D}{\lambda^2}. \quad (4.7b)$$

Again, Y is measured from the etched surface [see Fig. 4.3(a)], and τ is the etch time.

Eqs. (4.5) and (4.7) for the defect distributions both contain a number of parameters. The etch depth can be measured on an SEM, or by Talystep. The induction time for SiCl₄ is usually very small, so the etch time is also known. This leaves the "free" parameters F_1 and D in Eq. (4.5), and G_1 , λ and D in Eq. (4.7). When calculating conductances in later sections it will be found that the trap depth, ϕ_t , is also a parameter. However, calculations show that the exact value of ϕ_t makes little difference to the conductance. This is because ϕ_t only affects the amount of charge trapped in the upper level of the defect shown in Fig. 4.2, which is not very significant.

4.3.3 Defect distribution in two-dimensions (sidewall damage)

To solve for the sidewall damage in a wire, a few simplifying assumptions will be made concerning the structure geometry. The defect distribution will be found for a single sidewall (an etched step). The diffusion equation will be applied over the quarter plane $x \geq 0, y \geq 0$. This entails ignoring the etched surface, as shown in Fig. 4.3(b). All the defects are assumed to enter the material from the section of sidewall exposed to the bombarding ions. The flux of bombarding ions is assumed to be constant, so there is a uniform generation of defects over the exposed sidewall. The etch rate, v , is assumed to be constant, with the induction time being negligible (as in §4.3.2). The two cases discussed in §4.3.1 are treated separately. The sidewall defect distribution is denoted $\Gamma(\mathbf{r}, t)$.

(a) In the first case, all defects are generated at the surface, so the defect flux over the exposed sidewall is a non-zero constant. The problem may be stated as:

$$\left(D \nabla^2 - \frac{\partial}{\partial t} \right) \Gamma(\mathbf{r}, t) = 0; \quad x \geq 0, \quad y \geq 0, \quad (4.8a)$$

$$\Gamma(\mathbf{r}, t) = 0, \quad -D \left(\frac{\partial \Gamma}{\partial y} \right)_{y=0} = 0, \quad \lim_{r \rightarrow \infty} \Gamma(\mathbf{r}, t) = 0, \quad (4.8b)$$

$$-D \left(\frac{\partial \Gamma}{\partial x} \right)_{x=0} = F_1 [1 - \theta(y - vt)], \quad (4.8c)$$

where x and y are as shown in Fig. 4.3(b), t is the time, ∇ is the two-dimensional gradient operator, and F_1 is the magnitude of the defect flux. The θ -function appearing in the boundary condition (4.8c) is defined,

$$\theta(y) = \begin{cases} 0, & y < 0 \\ 1, & y > 0 \end{cases} \quad (4.9)$$

Its presence describes the downward motion of the etched surface, exposing larger and larger sections of sidewall. An integral representation for $\Gamma(\mathbf{r}, t)$ can be obtained from Eqs. (4.8) using the Green function method (see Appendix A.2):

$$\Gamma(\mathbf{r}, \tau) = F_1 \int_0^\tau dt' \frac{e^{-x^2/4D\Delta t'}}{2\sqrt{\pi D\Delta t'}} \left[\operatorname{erfc}\left(\frac{y-vt'}{2\sqrt{D\Delta t'}}\right) - \operatorname{erfc}\left(\frac{y+vt'}{2\sqrt{D\Delta t'}}\right) \right], \quad (4.10)$$

where τ is the etch time, v the etch rate, and $\Delta t' = \tau - t'$. This gives the defect distribution when defects are created purely at the surface.

(b) In the second case, the defect fluxes are zero on the surfaces of the sample being etched, but a source term is added to the diffusion equation to allow for defect generation a small distance away from the etched sidewall. The diffusion equation now becomes:

$$\left(D\nabla^2 - \frac{\partial}{\partial t} \right) \Gamma(\mathbf{r}, t) = -G_1 e^{-x/\lambda} [1 - \theta(y - vt)]; \quad x \geq 0, y \geq 0, \quad (4.11a)$$

$$\Gamma(\mathbf{r}, 0) = 0, \quad -D \left(\frac{\partial \Gamma}{\partial x} \right)_{x=0} = 0, \quad -D \left(\frac{\partial \Gamma}{\partial y} \right)_{y=0} = 0, \quad \lim_{r \rightarrow \infty} \Gamma(\mathbf{r}, t) = 0, \quad (4.11b)$$

where G_1 and λ are phenomenological parameters describing the rate of defect creation and the mean depth to which they are created. As before, the θ -function describes the downward motion of the etched surface, and again the problem may be solved using the Green function method (Appendix A.2):

$$\Gamma(\mathbf{r}, \tau) = \frac{G_1}{4} \int_0^\tau dt' e^{D\Delta t'/\lambda^2} \left\{ e^{-x/\lambda} \operatorname{erfc}\left(\frac{\sqrt{D\Delta t'}}{\lambda} - \frac{x}{2\sqrt{D\Delta t'}}\right) + e^{x/\lambda} \operatorname{erfc}\left(\frac{\sqrt{D\Delta t'}}{\lambda} + \frac{x}{2\sqrt{D\Delta t'}}\right) \right\} \\ \times \left[\operatorname{erfc}\left(\frac{y-vt'}{2\sqrt{D\Delta t'}}\right) - \operatorname{erfc}\left(\frac{y+vt'}{2\sqrt{D\Delta t'}}\right) \right], \quad (4.12)$$

where τ is the etch time, v the etch rate, and $\Delta t' = \tau - t'$.

4.3.4 Fit to DLTS profiles

Defect concentrations in SiCl_4 -etched GaAs have been measured by both Lootens *et al.* (1991) and Johnson *et al.* (1992). The latter used the same machine and etch conditions that were used to fabricate the wires of §3.5.3, so Eqs. (4.5) and (4.7) have been fitted to their measurements. Unfortunately, these measurements have been

plagued by order of magnitude sample-to-sample variations in measured defect concentrations, the reason for which is uncertain. Nevertheless, the measurements reveal up to five defect levels, all with concentrations varying roughly exponentially with depth from the etched surface.

Typical DLTS profiles obtained by Johnson *et al.* (1992) are shown in Fig. 4.4. The five defect levels are labelled N1-N5. Since N2 is the most abundant, Eqs. (4.5) and (4.7) will be fitted to its profile. Thus, the assumption is that N2 is the primary defect responsible for depletion region extension. Since the charge states of N2 are unknown, this defect may indeed *not* be responsible for the decrease in conductance after dry-etching, but for the sake of argument, its profile is nevertheless used to extract some useful information. In the numerical analysis to follow, it also will be assumed that the depth of the upper level in Fig. 4.2 is ~ 0.34 eV, the measured depth of the level N2. The lower level of Fig. 4.2 is assumed to be filled always.

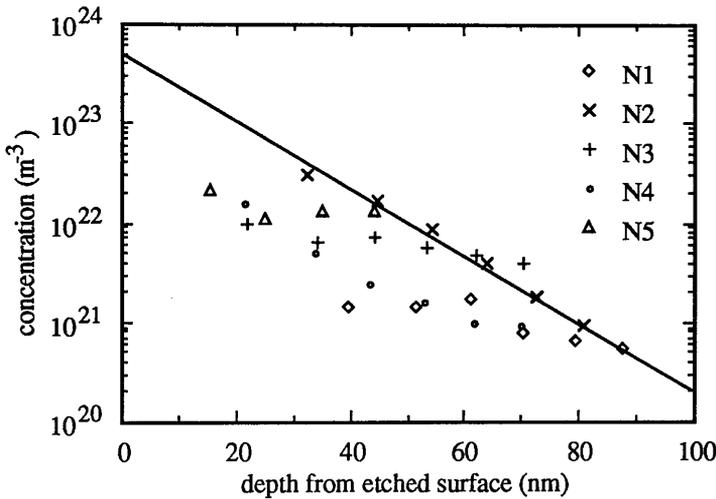


Figure 4.4: DLTS measurements of defect profiles in SiCl_4 -etched GaAs [Johnson *et al.* (1992)]. The straight line is a fit of Eq. (4.13) to the profile of the level N2. The corresponding parameters are $\lambda = 13.5$ nm and $G_1 = 1.02 \times 10^{23} \text{ m}^{-3} \text{ s}^{-1}$.

With the parameters $v = 200 \text{ nm min}^{-1}$ and $\tau = 30 \text{ s}$, Eq. (4.5) can be fitted with $F_1 \sim 10^{16} \text{ m}^{-2} \text{ s}^{-1}$ and $D \sim 10^{-17} \text{ m}^2 \text{ s}^{-1}$ (these are order of magnitude fits, which is sufficient for the present analysis), and Eq. (4.7) with $G_1 \sim 10^{23} \text{ m}^{-3} \text{ s}^{-1}$, $\lambda \sim 10 \text{ nm}$

and $D \sim 10^{-22} \text{ m}^2 \text{ s}^{-1}$ (this latter fit is not unique, but is typical for small values of D). It is immediately evident from these numbers that a fit to Eq. (4.5) requires a very high value of diffusion coefficient. For many defects in GaAs, such high values of D are unreasonable except at very high temperatures (perhaps several hundred Celsius). If such heating were to occur during etching, serious consequences on etching characteristics would result [e.g. Pearton *et al.* (1989)]. On the basis of this, it would seem that the first method of defect formation and propagation discussed in §4.3.1 is unrealistic, and so should not be pursued further.

Numerical evaluation of the integrals in (4.7) and (4.12) show that a further simplification can be made in the analysis. As D is made smaller, both these expressions approach limiting forms. It is easily shown that putting $D = 0$ in Eqs. (4.6a) and (4.11a), and integrating with respect to time gives,

$$\Gamma(Y, \tau) = \frac{\lambda G_1}{\nu} e^{-Y/\lambda} (1 - e^{-\nu\tau/\lambda}), \quad (4.13)$$

for the top surface etching, and

$$\Gamma(x, y, \tau) = G_1 e^{-x/\lambda} \left(\tau - \frac{y}{\nu} \right), \quad (4.14)$$

for sidewall etching. Eqs (4.7) and (4.13) agree to better than 1% for $D < 10^{-19} \text{ m}^2 \text{ s}^{-1}$, and Eqs. (4.12) and (4.14) agree to better than 1% for $D < 10^{-22} \text{ m}^2 \text{ s}^{-1}$. As D is increased by a few orders of magnitude above these values, the error only increases to ~10%. At the same time, the corresponding change in the calculated wire conductance is less than 1 μS . Thus, it would seem that provided the defect diffusion coefficient is not extremely large, D can more or less be ignored from the problem altogether, with very little error. Since Eqs. (4.13) and (4.14) are basically obtained by integrating the source function $G_1 e^{-x/\lambda}$ over the etch time, this corresponds to a picture where defects remain roughly where they are created. Eqs. (4.13) and (4.14) are used in all analyses to follow. It will be shown later that Eq. (4.14) gives an excellent description of the etching results of §3.5.3.

Exponential defect distributions (one-dimensional) have been used in several previous models on dry-etch damage [e.g. Mullins and Brunnschweiler (1976), and Taneya *et al.* (1989)]. However, Eq. (4.13) is somewhat more fundamental than any of these in that it tries to relate the distribution to actual etching parameters. Eq. (4.13) is also a concise mathematical expression of the graphical analysis presented by Ide *et al.* (1989) on the effects of etch rate on the amount of damage observed in dry-etched top surfaces.

A least squares fit of Eq. (4.13) to the profile of N₂ yields $\lambda = 13.5$ nm and $G_1 = 1.02 \times 10^{23}$ m⁻³ s⁻¹, where $v = 200$ nm min⁻¹ and $\tau = 30$ s were used. The fit is the straight line in Fig. 4.4. Although these values of λ and G_1 will not be used directly, they can be usefully compared to the corresponding values obtained for sidewall etching when wire conductances are calculated later in this chapter. This gives possible insights into differences between damage formation at top surfaces and sidewalls.

4.4 Conductances of etched structures

4.4.1 Conductances for etched wires

Using the expressions for the defect density, and the defect energy level structure discussed in §4.2.4, the conductance of an etched wire may be computed. This is done by first integrating the two-dimensional Poisson equation to find the conduction band edge, $E_c(x,y)$. This gives the free carrier density which is then integrated to give the conductance.

In addition to the defect charge, the Poisson equation will also contain the donor density, N_D , and the free electron and hole concentrations, $n(x,y)$ and $p(x,y)$. The free electron concentration is assumed to be given by,

$$n(x,y) = \int_{E_c(x,y)}^{\infty} dE g_e(E) f_0(E), \quad (4.15)$$

where $g_e(E)$ is the free electron density of states for a parabolic band, and $f_0(E)$ is the equilibrium Fermi-Dirac distribution:

$$g_e(E) = \frac{1}{2\pi} \left(\frac{2m_e^*}{\hbar^2} \right)^{3/2} (E - E_c)^{1/2}. \quad (4.16)$$

$$f_0(E) = \frac{1}{1 + e^{(E-E_f)/k_B T}}, \quad (4.17)$$

where E_f is the Fermi level (chemical potential), and $m_e^* = 0.067m_0$. The density of states function above is not strictly accurate for such a heavily doped system. A better expression would take into account such effects as band non-parabolicity, the density of states tail and other corrections appropriate for the high doping density [Shklovskii and Efros (1984)]. The free hole concentration is assumed to be given by,

$$p(x,y) = \int_{-\infty}^{E_v(x,y)} dE g_h(E) [1 - f_0(E)], \quad (4.18)$$

where $g_h(E)$ is the hole density of states summed over the light and heavy hole bands:

$$g_h(E) = \frac{1}{2\pi} \left[\left(\frac{2m_{lh}^*}{\hbar^2} \right)^{3/2} + \left(\frac{2m_{hh}^*}{\hbar^2} \right)^{3/2} \right] (E_v - E)^{1/2}. \quad (4.19)$$

Here $m_{lh}^* = 0.082m_0$, and $m_{hh}^* = 0.45m_0$. The valence band edge is given by,

$$E_v(x,y) = E_c(x,y) - E_g, \quad E_g = 1.52 - \frac{5.4 \times 10^{-4} T^2}{T + 204}, \quad (4.20)$$

where the energy gap varies with temperature [Sze (1985), Ch. 1].

If the defect level has a depth ϕ_t below the conduction band edge, the density of trapped charge in the upper level of Fig. 4.2 will be given by $\Gamma(x,y,t)f_0(E_c(x,y)-\phi_t)$. Thus, the total charge density associated with the defects from the left sidewall is,

$$N_T(x,y)_{ls} = \Gamma(x,y,t)_{ls} \left[1 + f_0(E_c(x,y)_{ls} - \phi_t) \right]. \quad (4.21)$$

There will be a similar contribution from the right sidewall. As has been discussed in §4.3.4, it will be assumed that $\phi_t = 0.34$ eV, the depth measured for the N2 defect by Johnson *et al.* (1992).

Putting all the expressions together gives the Poisson equation for $E_c(x,y)$:

$$\nabla^2 E_c(x,y) = \frac{e}{\epsilon} \left[N_D - N_c \mathcal{F}_{1/2} \left(\frac{E_c(x,y) - E_f}{k_B T} \right) + N_v \mathcal{F}_{1/2} \left(\frac{E_f - E_v(x,y)}{k_B T} \right) - N_T(x,y)_{ls} - N_T(x,y)_{rs} \right], \quad (4.22)$$

where all energies are measured in eV, and the electron and hole densities have been rewritten in a more convenient form using the Fermi-Dirac integral of order $\frac{1}{2}$ [Blakemore (1962)]:

$$\mathcal{F}_{1/2}(\eta) = \frac{2}{\sqrt{\pi}} \int_0^\infty \frac{z^{1/2} dz}{1 + e^{z+\eta}}. \quad (4.23)$$

N_c and N_v are the effective densities of states at the conduction and valence band edges respectively. If desired, the Fermi-Dirac integral can be approximated by one of several analytical expressions: the Boltzmann factor for $\eta < -3$ [Sze (1985), Ch. 1]; Sommerfeld's asymptotic expansion for $\eta > 3$ [Blakemore (1962)]; or the implicit expressions due to Joyce and Dixon (1977) for intermediate values of η . Blakemore (1962) considers yet further approximations. For the purposes of solving the Poisson equation, however, the integral will be evaluated numerically. The Poisson equation is solved over the rectangular cross-section of the wire, with the potential being fixed at 0.7 eV on the top surface and sidewalls, and the field set to zero at the substrate (these are the boundary conditions). The method of solution is discussed further in Appendix B.4.

Once the conduction band edge is obtained by solving the Poisson equation, the free electron concentration, $n(x,y)$, is given by evaluating the Fermi-Dirac integral. The conductance is then,

$$G(w) = \frac{e\mu}{L} \int_0^w dx \int_0^{v\tau} dy n(x,y), \quad (4.24)$$

where μ is the electron mobility, L and w are the wire length and width, and the double integral is performed over the cross-sectional area of the wire [see Appendix B.1]. The contribution to the conductance from the hole population is negligible and is ignored.

4.4.2 Sheet conductances for etched epilayers

Calculation of the sheet conductance of an epilayer proceeds in an analogous manner to that of the wire conductance, except the corresponding one-dimensional formulae are used. First the one-dimensional Poisson equation is solved to give the conduction band edge energy, $E_c(y)$. From this the electron concentration, $n(y)$, is found, which is then integrated to give the sheet conductance.

The Poisson equation with defect charge $N_T(y)$ arising from surface damage [*c.f.* Eq. (4.21)] is,

$$\frac{d^2 E_c(y)}{dy^2} = \frac{e}{\epsilon} \left[N_D - N_c \mathcal{F}_{1/2} \left(\frac{E_c(y) - E_f}{k_B T} \right) + N_v \mathcal{F}_{1/2} \left(\frac{E_f - E_v(y)}{k_B T} \right) - N_T(y) \right], \quad (4.25)$$

where all energies are measured in eV. This equation is solved with the potential being fixed at 0.7 eV on the etched surface, and the field set to zero at the substrate. The electron concentration is found from $E_c(y)$ using the Fermi-Dirac integral. The sheet conductance is then given by,

$$G(v\tau) = e \mu \int_0^\infty n(Y) dY \quad (4.26)$$

where Y is measured from the surface of the etched sample, and $v\tau$ is the etch depth. The method of solution of the Poisson equation, and the method for evaluating the integral in Eq. (4.26), are outlined in Appendices B.1-B.3.

4.4.3 Comparison with experimental wire conductances

Using the methods presented above, the conductances of n^+ -wires can be calculated with and without damage. Fig. 4.5 shows the results of calculations made for zero damage for the 50 nm $6.5 \times 10^{24} \text{ m}^{-3}$ GaAs epilayer, with a set of experimental

points measured for dry-etched wires included for comparison. For zero damage, the curves given by Eqs. (3.1) and (4.24) differ slightly. Although both give roughly the same sheet resistance (same gradient), the two-dimensional calculation appears to cut-off 7-8 nm above $2x_0$, where x_0 is given by Eq. (3.3).

The reason for this can be seen by studying the two-dimensional potential profile within the wire. Fig. 4.6 shows the variation in $E_c(x,y)$ over the cross-section of a 50 nm wide wire, as obtained from the Poisson equation, Eq. (4.22). Compare this with Fig. 3.2(b), which was used to derive Eq. (3.1). Note the curvature of the depletion region at the corners of the epilayer in Fig. 4.6. The curvature effectively adds a few nanometres to the nominal cut-off width of an undamaged wire, and is the reason for the difference noted above. This difference also explains why the wet-etch points in Fig. 3.5(b) are slightly shifted from the line given by Eq. (3.1) with $x_d = x_0$. The two-dimensional calculation of wire conductance versus width also shows a curvature at low conductances, evident in Fig. 4.5. Again, this is due to the two-dimensional nature of the problem.

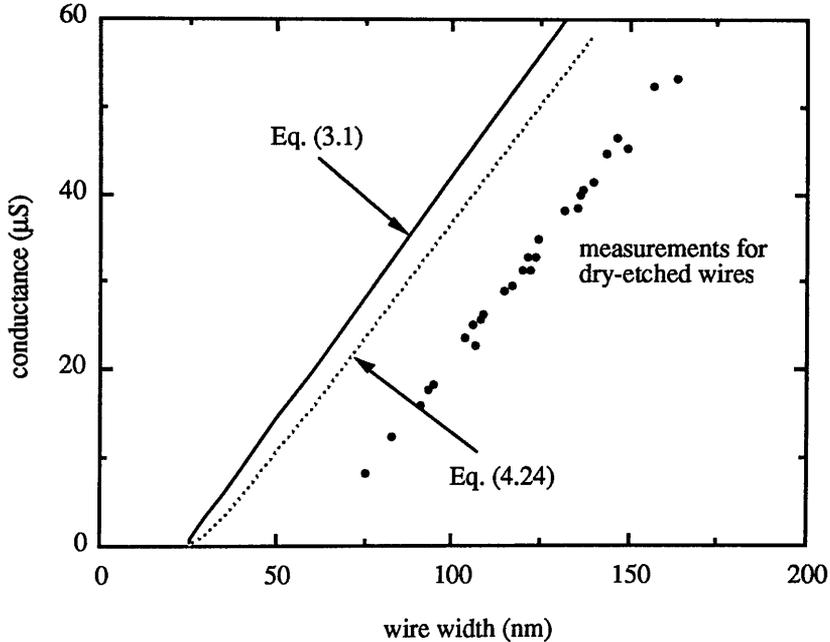


Figure 4.5: The lines are calculated conductances for zero damage. The points are measured for wires fabricated from the $6.5 \times 10^{24} \text{ m}^{-3}$ material, using 300 V etching for 30 s, with HRN as the etch mask.

Fig. 4.7 shows all the relevant experimental data of §3.5.3, together with calculated conductances for the damaged wires. The parameter values used [for Eq. (4.14)] were $\lambda = 9(\pm 1)$ nm, $G_1 = 5.5(\pm 0.2) \times 10^{23}$ m⁻³ s⁻¹, and $v = 200(\pm 10)$ nm min⁻¹, and it is evident that the fit is excellent (the errors given above indicate the ranges over which these parameters can be varied before the fit deviates by a more than a few percent). The result is significant. It shows that given the source function $G_1 e^{-x/\lambda}$ [c.f. Eq. (4.11)] with the parameter values above, the effects of SiCl₄-etching of GaAs at energies of ~300 eV on the conductances of wires are *completely determined*, with respect to variations in carrier concentration, epilayer thickness, and etch time/depth.

```

meshSize           = 2.50 nm
wireWidth          = 50.00 nm
etchDepth          = 100.00 nm
surfacePotential   = 0.70 eV

```

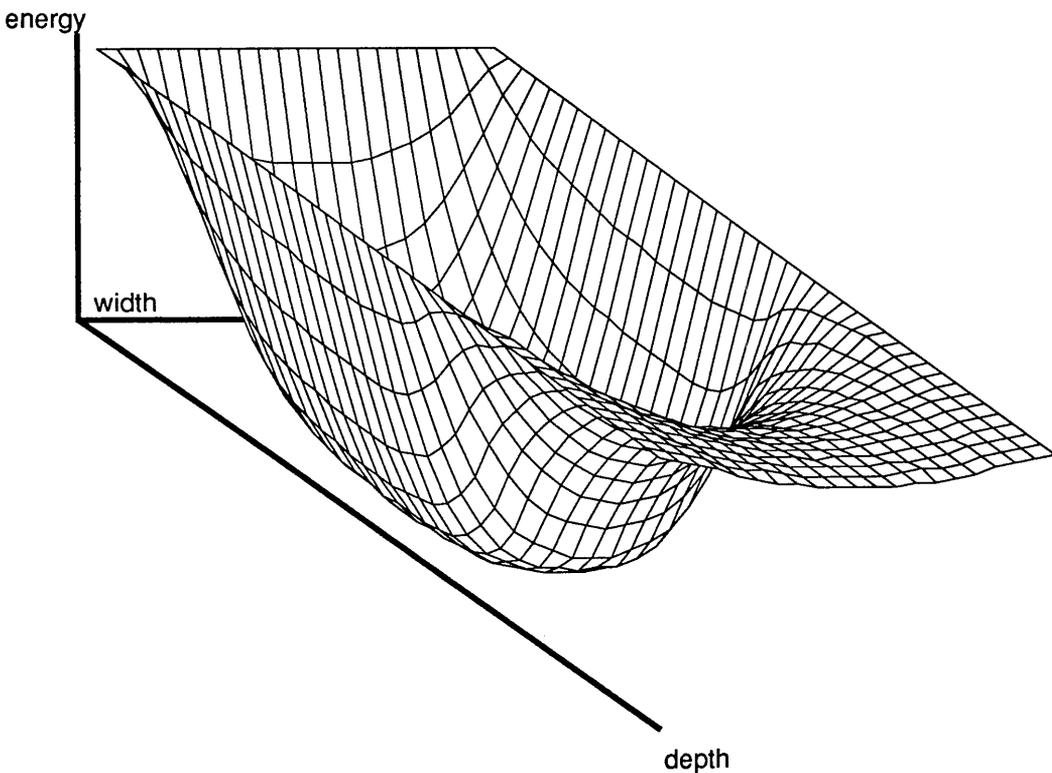


Figure 4.6: Two-dimensional profile of conduction band edge as calculated by solving the Poisson equation. The wire is 50 nm wide, and is made from the 6.5×10^{24} m⁻³ material.

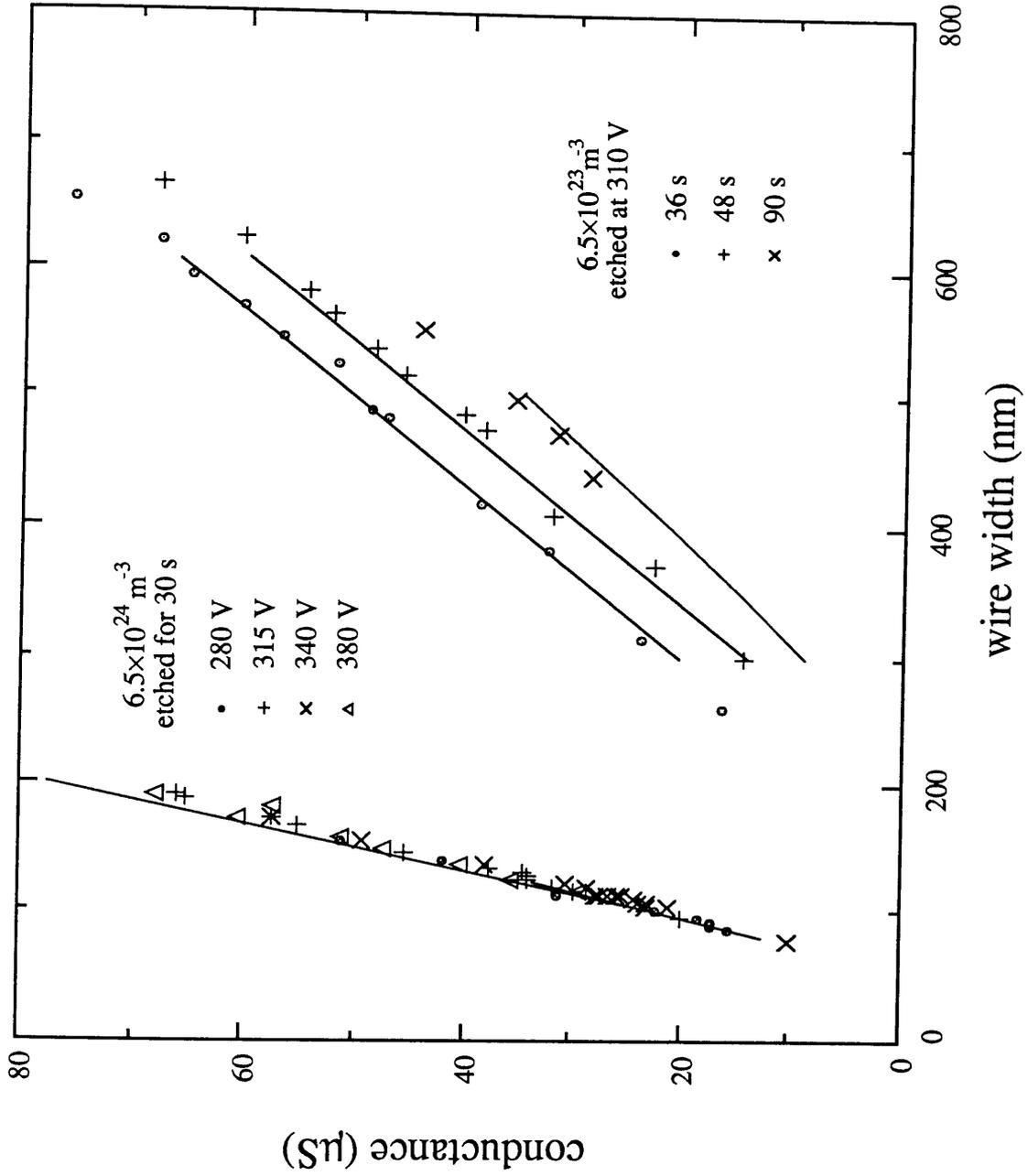
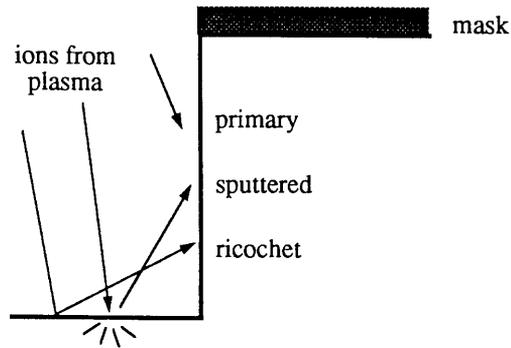


Figure 4.7: The experimental points here are taken from Figs. 3.5(a) and 3.5(b), for materials with different doping and epilayer thickness. The solid lines are fits, calculated from the model described in the text, using the parameters $\lambda = 9 \text{ nm}$ and $G_1 = 5.5 \times 10^{23} \text{ m}^{-3} \text{ s}^{-1}$. This one set of parameters fits all the data, reproducing the effects of etch time too.

Figure 4.8: Components of the bombarding flux impinging upon a sidewall as it is being etched.



Comparing the above parameters with those obtained for top surface damage (see §4.3.4), where $\lambda \sim 13.5 \text{ nm}$ and $G_1 \sim 1.02 \times 10^{23} \text{ m}^{-3} \text{ s}^{-1}$, also reveals something interesting. Since ion collisions with the top surface are predominantly "head-on," whereas those with the sidewall are mainly "glancing," sidewall damage would be expected to be far less severe than top surface damage. Davis and Tiwari (1991) [see also Davis (1991)] have pointed out that the image potentials at the sidewall may deflect incoming ions away from their near-vertical trajectories, so that the collisions may in fact be less glancing than originally anticipated. Nevertheless, λ and G_1 would still be expected to be considerably less for sidewall damage than for top surface damage. Although λ is slightly smaller, it is seen that G_1 is slightly *greater* for sidewalls than for top surfaces. Thus, defects appear to be created at a much higher rate at sidewalls than would be intuitively expected.

There are several possible reasons for this. The most interesting is that it is the effect of a secondary ion flux impinging upon the sidewall, in addition to the primary ion flux from the plasma. This flux most likely comprises ricochet ions and sputtered material from the surface being removed, as shown in Fig. 4.8. The flux density arising from such processes must be significant compared to the primary ion flux density to give the observed creation rate.

In addition to the above, inadequacies in the model [*i.e.* Eq. (4.14)], or inaccuracies in the DLTS measurements could be further explanations of why the rate of defect creation at the sidewalls appears to be so large compared with the corresponding value for top surfaces. Until further developments are made, these possibilities must also be born in mind.

4.4.4 Comparison with experimental sheet conductances

Foad *et al.* (1992) have measured the sheet conductance of a 140 nm thick $3.0 \times 10^{24} \text{ m}^{-3}$ GaAs epilayer as a function of etch depth for various forms of etch, including wet-etch and SiCl_4 -RIE at 290 V, amongst others. Two sets of measurements using SiCl_4 -etching were made. One indicated that SiCl_4 and wet-etching were roughly similar, the other that SiCl_4 did indeed give some damage, resulting in a 10-15 nm extension in the surface depletion depth. Using the parameter values obtained from the DLTS fit ($\lambda = 13.5 \text{ nm}$ and $G_1 = 1.02 \times 10^{23} \text{ m}^{-3} \text{ s}^{-1}$) gives a calculated conductance versus etch depth curve which indicates a roughly 3-4 nm extension in the surface depletion depth. This lies somewhere between the two curves measured experimentally, and the comparison is inconclusive. A better comparison requires further work to be done.

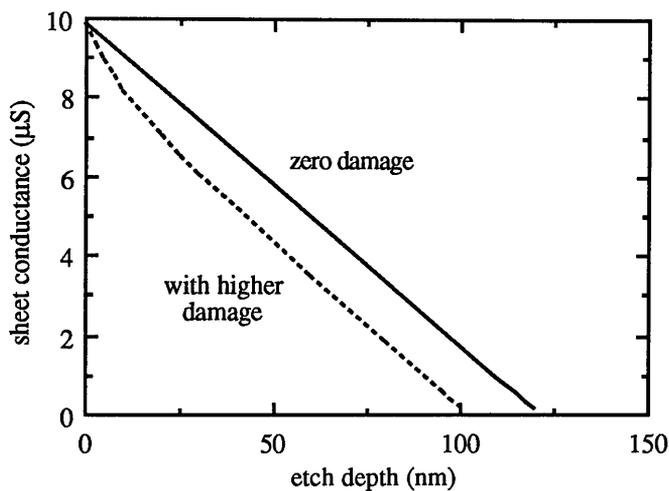


Figure 4.9: Calculation of sheet conductance of a dry etched epilayer of thickness 140 nm and carrier concentration $3.0 \times 10^{24} \text{ m}^{-3}$. Two curves are shown, one for etching with zero damage, the other for etching with damage. See text for details.

Perhaps more interesting are the calculated curves shown in Fig. 4.9. The straight line is the zero damage case. The other line is calculated with $\lambda = 14 \text{ nm}$, $G_1 = 10^{24} \text{ m}^{-3} \text{ s}^{-1}$ and $v = 170 \text{ nm s}^{-1}$. It is seen that the amount of damage appears to saturate after a short period. This form of curve is very similar to what has been observed for CH_4/H_2 etching, and what is expected for dry-etching with the inert gases, such as Ar,

Ne, *etc.* [Foad *et al.* (1992)]. Eq. (4.13) shows that decreasing v or increasing λ has a similar effect on the prefactor as increasing G_1 . Taneya *et al.* (1989) using Ga⁺ FIB/Cl₂ etching, and Dubonos and Koveshnikov (1990) using Ar-plasma etching believe that various types of point defect are introduced into their etched material by the etching process. With this being the case, the defect distribution Eq. (4.13) derived in §4.3.4 should also be applicable to these forms of etches too. However, a complete set of data is as yet lacking for any of these forms of etch to allow model parameters to be fitted uniquely.

4.5 References

- BLAKEMORE J S, *Semiconductor Statistics*, Pergamon Press, Oxford (1962).
- BOLTAKS B I, *Diffusion in Semiconductors*, Infosearch Ltd., London (1963).
- BOURGOIN J C, H J von Bardeleben, D Stiévenard, "Native defects in gallium arsenide," *J. Appl. Phys.* **64**, R65 (1988).
- CHADDERTON L T, *Radiation Damage in Crystals*, Methuen & Co. Ltd., London (1965).
- CRANK J, "How to deal with moving boundaries in thermal problems," in ed. R W Lewis, K Morgan, O C Zienkiewicz, *Numerical Methods in Heat Transfer*, John Wiley & Sons Ltd., Chichester (1981).
- DAVIS R J, "Image potentials and the dry etching of submicron trenches with low-energy ions," *Appl. Phys. Lett.* **59**, 1717 (1991).
- DAVIS R J, S Tiwari, "Image potentials and dry etch directionality in integrated optoelectronics," LEOS Topical Meeting on Microfabrication for Optical Devices, Newport Beach, California (1991).
- DUBONOS S V, S V Koveshnikov, "Formation of highly mobile defects in GaAs under Ar-plasma etching," *Phys. Stat. Sol. A* **120**, 77 (1990).
- FOAD M A, S Thoms, C D W Wilkinson, "A new technique for dry-etch damage assessment in GaAs," submitted (1992).
- HAYES W, A M Stoneham, *Defects and Defect Processes in Nonmetallic Solids*, John Wiley & Sons, New York (1985).
- IDE Y, N Takado, K Asakaw, "Optical characterization of reactive ion beam etching induced damage using GaAs/AlGaAs quantum well structures," in ed. H

- Watanabe, *GaAs and Related Compounds 1989*, IOP Publishing Ltd., Bristol (1989).
- JOHNSON N, M A Foad, C D W Wilkinson, "Electron traps in SiCl₄ dry etched GaAs," unpublished (1992).
- JOYCE W B, R W Dixon, "Analytic approximations for the Fermi energy of an ideal Fermi gas," *Appl. Phys. Lett.* **31**, 354 (1977).
- KIMERLING L C, "Influence of deep traps on the measurement of free-carrier distributions in semiconductors by junction capacitance techniques," *J. Appl. Phys.* **45**, 1839 (1974).
- LANG D V, R A Logan, M Jaros, "Trapping characteristics and a donor-complex (DX) model for the persistent-photoconductivity trapping center in Te-doped AlGaAs," *Phys. Rev. B* **19**, 1015 (1979).
- LIM H, H J von Bardeleben, J C Bourgoin, "Experimental determination of the energy stored by a vacancy-interstitial pair in GaAs," *Phys. Rev. Lett.* **58**, 2315 (1987).
- LONG A R, I K MacDonald, M C Vernon, M Rahman, J J Thompson, M Kinsler, S P Beaumont, C D W Wilkinson, C R Stanley, "Field-induced resistivity and persistent photo-conductivity in quantum wires," unpublished (1990).
- LOOTENS D, P Van Daele, P Demeester, P Clauws, "Study of electrical damage in GaAs induced by SiCl₄ reactive ion etching," *J. Appl. Phys.* **70**, 221 (1991).
- MULLINS F H, A Brunnschweiler, "The effects of sputtering damage on the characteristics of Molybdenum-Silicon Schottky barrier diodes," *Solid-State Electron.* **19**, 47 (1976).
- OCKENDON J R, W R Hodgkins, ed. *Moving Boundary Value Problems in Heat Flow and Diffusion*, Clarendon Press, Oxford (1975).
- ORLOWSKI M, "Unified model for impurity diffusion in silicon," *Appl. Phys. Lett.* **53**, 1323 (1988).
- PEARTON S J, A B Emerson, U K Chakrabarti, E Lane, K S Jones, K T Short, Alice E White, T R Fullowan, "Temperature dependence of reactive ion etching of GaAs with CCl₂F₂:O₂," *J. Appl. Phys.* **66**, 3839 (1989).
- RHODERICK E H, R H Williams, *Metal-Semiconductor Contacts*, Clarendon Press, Oxford (1988).
- RUNYAN W R, *Silicon Semiconductor Technology*, McGraw-Hill Book Co., New York (1965).
- SHEWMON P G, *Diffusion in Solids*, McGraw-Hill Book Company Inc., New York

(1963).

SHKLOVSKII B I, A L Efros, *Electronic Properties of Doped Semiconductors*, Springer-Verlag, Berlin (1984).

SZE S M, *Semiconductor Devices*, John Wiley & Sons, New York (1985).

TANEYA M, Y Sugimoto, K Akita, "Characterization of subsurface damage in GaAs processed by Ga⁺ focused ion-beam-assisted Cl₂ etching using photoluminescence," *J. Appl. Phys.* **66**, 1375 (1989).

Quantum transport in mesoscopic systems

5.1 Outline

The general problem of quantum transport in crystalline systems has been studied for many decades. Use of electron-beam lithography through the 1980s opened up new areas of research for experimental physicists. At the same time, theorists were uncovering new approaches for understanding the nature of transport in the systems being studied experimentally. These investigations still continue.

Of particular interest are two regimes of low temperature quantum transport. The first involves the use of very pure samples, where the electron mean free path is very long and electron waveguiding effects are at the fore. This is the quantum ballistic regime. As more scatterers are introduced into the system, the electron mean free path decreases, and the quantum ballistic phenomena are replaced by quantum diffusive phenomena. This is the second regime of interest, often called the mesoscopic regime. Electron transport in mesoscopic systems are described using the laws of quantum mechanics, but due to the diffusive nature of the problem, some complex and interesting effects result. It is the purpose of this chapter to outline the basic ideas of quantum transport in the mesoscopic regime. These ideas will then be used in Chapter 6 to understand the effects being measured in n^+ -GaAs wires at low temperatures.

5.2 Mesoscopic systems

5.2.1 Electron coherence

Quantum interference effects in nanostructures typically manifest themselves in conductance measurements. On a qualitative level, such effects can be understood in terms of single electron states (wavefunctions). Electron-lattice, electron-impurity, and electron-electron interactions all affect the motion of an electron through the semiconductor. An electron starting off in a particular state will end up in a different state as a result of these interactions – a process called scattering. The types of scattering predominating during a particular measurement will determine whether interference effects are observable, or whether they are destroyed, *i.e.* averaged away.

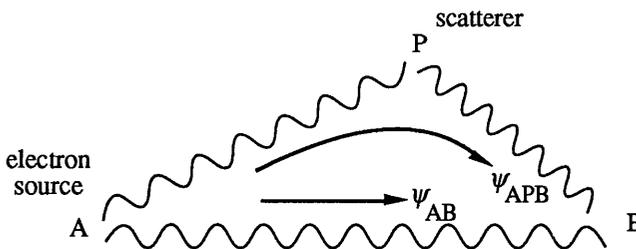


Figure 5.1: Electron motion from point A to point B can either be direct, or indirect via scattering off point P. The type of scattering at P can be either coherent or incoherent.

Fig. 5.1 shows how an electron "wave" can propagate from point A to point B by two different paths. One path goes directly from A to B, whereas the other is scattered at P. The scattering at P can be either coherent ("preserving" the phase of the electron wave), or incoherent ("destroying" the phase of the electron wave). In coherent scattering, the phases of the electron wave before and after scattering will have a definite, fixed relationship. The phase could remain completely unchanged, or there could be a known phase shift as a result of the scattering. [A study of phase shifts in the partial wave analysis of scattering experiments was much beloved of particle physicists several decades ago. Although less grand, the scattering in the present context can be thought of in an analogous manner – see, for example, Mahan (1981), Ch. 4] In incoherent

scattering, the electron phase becomes completely undefined after the scattering event – the electron loses its phase "memory" as a result of the scattering. This is called phase breaking.

Now, let ψ_{AB} be the amplitude at B of the wave travelling directly from A to B, and ψ_{APB} be the corresponding amplitude of the wave that has been scattered at P. In real samples there are typically a large number of scattered paths, not just one. Thus, in addition to ψ_{APB} , there will also be amplitudes ψ_{AQB} , ψ_{ARB} , ... arising from scattering at Q, R, If the scattering is incoherent, then by direct analogy with the addition of incoherent sources of light,

$$|\psi_B|^2 = |\psi_{AB}|^2 + |\psi_{APB}|^2 + |\psi_{AQB}|^2 + \dots \quad (5.1)$$

where ψ_B is the wavefunction amplitude at B. The net result of this sum is that interference effects are "smoothed out," and the wave nature of the electron motion is hidden. This corresponds to the classical limit of electron transport, described by classical electron diffusion.

The situation is altered drastically when the scattering is predominantly coherent. In this case, it is the wavefunction amplitudes that add (not the "intensities"):

$$\psi_B = \psi_{AB} + \psi_{APB} + \psi_{AQB} + \dots = \psi^{\text{direct}} + \psi^{\text{scatt}} \quad (5.2)$$

where $\psi^{\text{direct}} = \psi_{AB}$, and $\psi^{\text{scatt}} = \psi_{APB} + \psi_{AQB} + \dots$ are the direct and scattered components contributing to ψ_B . Two cases are of interest here. If $\psi^{\text{scatt}} \rightarrow 0$ in the sample, then the electron transport is said to be quantum ballistic. The boundaries of the sample can then "guide" the electron waves (boundary scattering is not included in ψ^{scatt}), much as in microwave waveguides. The low scattering limit can be realised in nanostructures fabricated using ultra pure MBE-grown GaAs/AlGaAs heterolayers. The last 4-5 years have seen the discovery of a number of ballistic phenomena, leading to a hive of activity in this area. Examples include the quantised conductance of a quantum point contact [van Wees *et al.* (1988); Wharam *et al.* (1988)], the quenching of the Hall effect [Roukes *et al.* (1987)], and coherent electron focussing in a perpendicular magnetic field [van Houten *et al.* (1989)]. For a detailed review, see Beenakker and van Houten (1991).

When ψ^{scatt} in Eq. (5.2) is not so small (but still the result of coherent scattering), interference effects are still observable, although their description and analysis is much more complex. This corresponds to the regime of quantum diffusion, where the electrons scatter (perhaps frequently) in going from A to B, but their wave nature remains intact. The resulting interference effects require changes to the classical

descriptions of their diffusive motion. Some of the effects are discussed later in this chapter. This will lay the groundwork for attempting to understand the measurements on n^+ -GaAs wires in Chapter 6. In the next few subsections, the above arguments on coherence are quantified slightly.

5.2.2 Length, time and energy scales

Quantum interference is affected by a number of factors, *e.g.* scattering, temperature, applied magnetic field, *etc.* The effect of these factors on the quantum interference may be quantified by introducing a number of parameters. The relative values of these parameters helps in the understanding of quantum interference effects. First, the various scattering processes affecting electron motion are dealt with.

Impurity scattering of electrons is generally elastic, *i.e.* energy conserving, since the impurities are much more massive than the electron, and will therefore gain a negligible amount of kinetic energy during the collision. [It will be assumed that energy is not transferred to electrons bound to the impurity. This is a good approximation provided the energy of the conduction electron is not too high.] Typically, impurity scattering will preserve the electron phase, and interference effects will still be possible. The mean distance between elastic scattering events is denoted l , the elastic mean free path. [Strictly speaking, a distinction should be drawn between small angle and large angle scattering, but for the present argument such complications are ignored.] The mean time between elastic collisions is $\tau = l/v_F$, where v_F is Fermi velocity. Such collisions usually determine the diffusion coefficient of the electrons at the Fermi surface,

$$D = \frac{v_F l}{d}, \quad (5.3)$$

where d is the dimensionality of the sample. [Scattering at the system boundaries can also affect D . See §6.2.5.]

The electron-lattice interaction, via electron-phonon scattering, is very deleterious to interference effects. At high¹ temperatures it is the dominant scattering mechanism, and interference effects are hard to observe. Energy is invariably transferred in such scattering processes. If the energy transferred is much smaller than $k_B T$, then the scattering is effectively elastic (or *quasi*-elastic); if the energy transferred is comparable to $k_B T$, the scattering is inelastic. [Since the Debye temperature, θ_D , generally

¹ The exact meaning of "high" here depends on a number of material parameters, such as impurity density, carrier concentration, *etc.*

determines the highest phonon energies, electron-phonon scattering will tend to be quasi-elastic if $k_B T > k_B \theta_D$, and inelastic otherwise. See, for example, Blatt (1968), Ch. 5, or Seeger (1984), Ch. 6.] The number of inelastic electron-phonon collisions per second (scattering rate) may be written τ_{ph}^{-1} , where τ_{ph} is the mean time between collisions. In reality there may be several different types of inelastic electron-phonon interaction, so τ_{ph}^{-1} represents the net scattering rate due to all of these.

Inelastic electron-phonon scattering is not the only contribution to the inelastic scattering rate, τ_{in}^{-1} . Electron-electron scattering also contributes to τ_{in}^{-1} since one electron can lose energy to another (their indistinguishability does not affect this). Thus, like electron-phonon scattering, electron-electron scattering is phase breaking. Denoting the electron-electron scattering rate τ_{ee}^{-1} , the combined inelastic scattering rate is the sum,

$$\frac{1}{\tau_{in}} = \frac{1}{\tau_{ph}} + \frac{1}{\tau_{ee}}. \quad (5.4)$$

The current-carrying electron states are those diffusing at the Fermi surface. The effective (mean) distance travelled by these electrons between inelastic collisions is defined,

$$L_{in} = (D \tau_{in})^{1/2}, \quad (5.5)$$

called the inelastic scattering length, where D is the diffusion coefficient.

Slight complications can arise in impurity scattering due to the spin of the conduction electron. If high atomic number nuclei are present in the sample, then the spin-orbit interaction will be important. The spin-orbit Hamiltonian is given by [Zeiger and Pratt (1973), Ch. 1; Chakravarty and Schmid (1986)],

$$H_{SO} = -\frac{e\hbar}{4m^*c^2} \sigma \cdot [\nabla V(\mathbf{r}) \times \dot{\mathbf{r}}], \quad (5.6)$$

where the components of σ are the Pauli spin matrices, $V(\mathbf{r})$ is the impurity potential, $\dot{\mathbf{r}}$ is the electron velocity, and m^* is its effective mass. The effect of the spin-orbit interaction is to change the phase of the spin part of the electron wavefunction [Bergmann (1984)]. This is an elastic process, and it does not destroy the phase coherence of the electron. The spin-orbit scattering length is $L_{SO} = (D \tau_{SO})^{1/2}$, where the spin-orbit scattering rate, τ_{SO}^{-1} , contributes to the net elastic scattering rate, τ^{-1} . In metals, these two rates can be related by [Meservey and Tedrow (1978)],

$$\frac{1}{\tau_{SO}} = (\alpha Z)^4 \frac{1}{\tau}, \quad (5.7)$$

where $\alpha = e^2/4\pi\epsilon_0\hbar c \approx 1/137$ is the fine structure constant, and Z is the atomic number.

If the impurity has a net spin, then the exchange interaction between the

conduction electron and the impurity electrons will be important. The interaction can be written [Mahan (1981), Ch. 1; Chakravarty and Schmid (1986)],

$$H_s = -J \boldsymbol{\sigma} \cdot \mathbf{S} , \quad (5.8)$$

where $\boldsymbol{\sigma}$ is the electron spin, \mathbf{S} is the net impurity spin, and J is the coupling constant which is proportional to the exchange integral. Like the impurity interactions discussed above, the interaction represented by H_s is elastic. However, unlike the ones above, it is also phase breaking. The interaction is variously referred to as magnetic scattering, or spin-flip scattering. The magnetic scattering length is $L_s = (D \tau_s)^{1/2}$, where the magnetic scattering rate, τ_s^{-1} , contributes to both the elastic scattering rate, and the phase breaking rate. In magnetic scattering, the spin on the conduction electron rotates by a certain amount, as does the spin on the impurity [Bergmann (1984)]. This randomises the phase in the spin part of the electron wavefunction, and leads to the loss of coherence [Stern *et al.* (1990) give a more detailed argument].

The above are the most important scattering processes affecting quantum interference measurements. The various processes affect either the elastic scattering length, l , or the coherence (phase breaking) length, $L_\phi = (D \tau_\phi)^{1/2}$, where the phase breaking rate is τ_ϕ^{-1} . In the literature, L_{in} is often used in place of L_ϕ , but this is strictly only true when magnetic scattering is insignificant in the sample being considered. The coherence length also goes under the name Thouless length (L_{Th}). In Aharonov-Bohm experiments (§5.3.1) it is found that [van Haesendonck (1991)],

$$\frac{1}{\tau_\phi} = \frac{1}{\tau_{in}} + \frac{1}{\tau_s} , \quad (5.9)$$

but in weak localisation experiments (§5.3.2) the relation [Hikami *et al.* (1980); van Haesendonck (1991)],

$$\frac{1}{\tau_\phi} = \frac{1}{\tau_{in}} + \frac{2}{\tau_s} , \quad (5.10)$$

applies. Measurements of weak localisation can often provide a useful method of deducing τ_ϕ^{-1} and τ_{SO}^{-1} for the sample under study. However, it is much harder to deconvolute τ_{in}^{-1} and τ_s^{-1} from τ_ϕ^{-1} [Bergmann (1984)]. It is generally found that $\tau_\phi^{-1} \propto T^p$, and τ_{SO}^{-1} is temperature independent, where p depends on the scattering mechanisms prevalent in the sample, and also on the sample dimensionality.

Two energy scales are of considerable importance. Finite temperatures smear out the Fermi surface over an energy range $\sim k_B T$. This scale ultimately controls the observability of all quantum effects, since if $k_B T$ is high enough, all such effects disappear. In terms of a "scattering" length, $k_B T$ becomes the thermal diffusion length,

$L_T = (D \tau_T)^{1/2} = (\hbar D/k_B T)^{1/2}$. The thermal diffusion length is discussed further in the next subsection, in the context of energy averaging. Also important is the energy, eV , gained by an electron in traversing a region across which there is a potential difference, V . If $eV > k_B T$, then the electrons often behave as if they were at an effective temperature, $T_{el} \sim eV/k_B > T$. In this case, it is T_{el} (not T) which enters as the relevant temperature in expressions relating to the electron gas. This can either be a good or bad thing, depending on the requirements of the experiment.

The strength of the magnetic field sets two further scales which can affect quantum interference. Electrons moving in a plane with a perpendicular applied magnetic field, will tend to undergo a cyclotron motion with frequency $\omega_c = eB/m^*$. If the path of these orbits lies within a phase coherent region, then Landau levels are formed [*e.g.* Ando *et al.* (1982)],

$$\psi_N(\mathbf{r}) \propto \exp\left(-\frac{r^2}{4L_B^2}\right), \quad (5.11)$$

where $L_B = (\hbar/eB)^{1/2}$ is the magnetic length, and N is the Landau index. The Landau wavefunction thus has a Gaussian envelope, with approximate radius $\sim L_B$ (for $N = 0$). The length scale L_B is therefore identified as the quantum cyclotron radius. [The expression for L_B is readily obtained by equating the rotational kinetic energy of a classical particle undergoing cyclotron motion, $\frac{1}{2} m^* \omega_c^2 L_B^2$, to the energy of the $N = 0$ Landau level, $\frac{1}{2} \hbar \omega_c$. For higher Landau levels the quantum cyclotron radius becomes $(2N+1)^{1/2} L_B$. The *classical* cyclotron radius is obtained by equating the centrifugal force, $m^* \omega_c^2 r_c$, to the Lorentz force, $e v_F B$, *i.e.* $r_c = \hbar k_F / eB$.] Landau levels form when $L_B < L_\phi$, but only become experimentally observable when their width in energy (due to thermal broadening, disorder broadening, *etc.*) is much smaller than $\hbar \omega_c$.

An applied magnetic field will also tend to lift any spin degeneracy of conduction electrons in the sample under study. The magnetic moment of an electron may be written,

$$\boldsymbol{\mu} = -\frac{g_s \mu_B}{\hbar} \mathbf{S}, \quad (5.12)$$

where $\mu_B = e \hbar / 2m^*$ is the Bohr magneton, g_s is the spin g-factor, and \mathbf{S} is the spin angular momentum of the electron. The orientational potential energy of the electron is $-\boldsymbol{\mu} \cdot \mathbf{B}$, so the Zeeman splitting between the spin up and spin down electrons is [Zeiger and Pratt (1973), Ch. 1],

$$\Delta E_Z = -\mu_\downarrow \cdot \mathbf{B} + \mu_\uparrow \cdot \mathbf{B} = g \mu_B B, \quad (5.13)$$

where g is the effective Landé g-factor in the material, and $\mathbf{S} \cdot \mathbf{B} = \pm \frac{1}{2} \hbar B$. The Zeeman

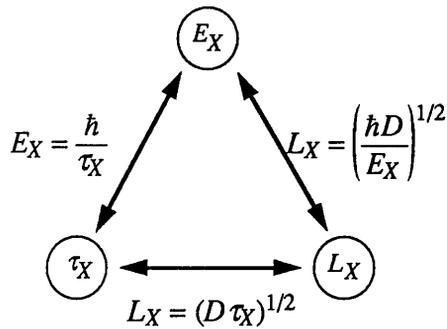
splitting provides an energy scale which can be compared, for example, to $k_B T$. If $g\mu_B B > k_B T$, then it may be possible to resolve the two spin states experimentally. If $g\mu_B B < k_B T$, then the two spin states are still effectively degenerate.

In general, the processes above can be compared in one of several equivalent ways. Three of these ways (as length, time and energy scales) are shown in Fig. 5.2, and have been used above. Bergmann (1984) also uses equivalent magnetic field scales to compare the different scattering mechanisms:

$$B_X = \frac{\hbar}{4eD\tau_X} . \tag{5.14}$$

This is equivalent to comparing the scattering rate, since $B_X \propto \tau_X^{-1}$.

Figure 5.2: Diagram showing the equivalence between using length, time or energy scales to compare different processes.



5.2.3 Energy averaging

The phase breaking processes described above are not the only ways that interference effects can be destroyed. Quantum effects should be strongest at the absolute zero of temperature, where all the contributing electrons (*viz.* those at the Fermi surface) have the same energy, and hence are described by the same probability density function (or wave function, up to an arbitrary phase factor). When the temperature is raised, the Fermi surface is smeared out, and electrons in different energy levels can contribute to the interference process. Since electrons in different energy levels are expected to have different wavefunctions (strictly modulus squared of the wavefunction amplitudes), it may be thought that the more distinct energy levels there are, the greater will be the averaging of the overall interference effect. However, this is only partially correct.

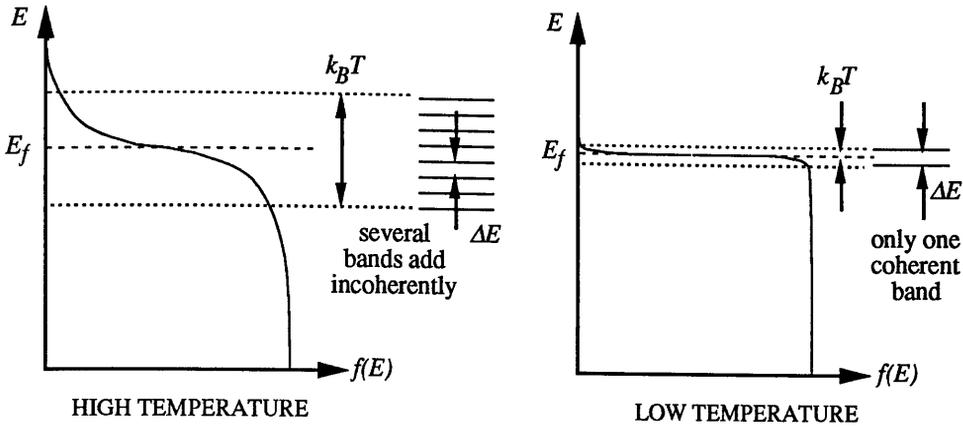


Figure 5.3: Illustration of energy averaging. Electrons within an energy range ΔE contribute to the same interference pattern. If $k_B T > \Delta E$ then there are several bands which contribute differently to the interference. $f(E)$ is the Fermi-Dirac distribution function.

If T is the electron temperature, electrons within a range of energies $k_B T$ around the Fermi level, E_f , will contribute to the conduction process. However, only electrons within a band of energies ΔE can be regarded as having the same probability density function. When $\Delta E > k_B T$ then all the current-carrying electrons will contribute to the same interference pattern. But if $\Delta E < k_B T$, there will effectively be several bands of electrons (see Fig. 5.3) whose interference patterns are essentially uncorrelated. The interference pattern thereby produced is some average over these bands, and is diminished from the effect obtained from just a single band. This is called energy averaging. It may be naively thought that ΔE is just the spacing between the distinct energy levels of the system. However, this is not so: ΔE can be much much larger than the level spacing. It is in fact related to the uncertainty in an energy level due to the time taken by an electron in diffusing across the sample, as the following argument due to Thouless (1977) shows.

Consider a sample of size L within which an electron is diffusing. If the diffusion coefficient is D , then the mean time taken for the electron to cross L is,

$$\tau = \frac{L^2}{D}. \quad (5.15)$$

If the electron phase is coherent across L , then the electron entering and leaving L can

be regarded as scattering into and out of an electron state, with mean lifetime τ . This finite lifetime introduces an uncertainty in the energy of this state, called the correlation (or Thouless) energy:

$$E_C = \frac{\hbar}{\tau}. \quad (5.16)$$

All electrons within a range of energy E_C of one another can effectively be regarded as having the same energy, *i.e.* they will have the same probability density function. In other words, E_C can be identified with ΔE of Fig. 5.3. Thus, for energy averaging not to occur requires $E_C > k_B T$, or, using Eqs. (5.15) and (5.16):

$$L < \left(\frac{\hbar D}{k_B T} \right)^{1/2}. \quad (5.17)$$

The right hand term of Eq. (5.17) is defined as L_T , the thermal diffusion length (though L_T is sometimes defined with h rather than \hbar). Thus, for no energy averaging to occur within a phase coherent sample, its size must be less than L_T , *i.e.* the inequality $L < L_T$ must be satisfied.

5.2.4 Mesoscopic systems

Based on the length scales introduced above, the sample under study can be placed into one of three regimes, depending on the sample size, L . (i) The quantum ballistic regime, $L < l, L_\phi, L_T$. Here the electron traverses the sample without suffering any form of scattering, elastic or inelastic. This generally requires very pure samples and very low temperatures. Some of the effects observed in this regime have already been mentioned in §5.2.1. (ii) The quantum diffusive regime, $l < L < L_\phi, L_T$. Here the electron is permitted to suffer elastic collisions in traversing the sample, but no phase breaking is allowed to occur. The problem is thus diffusive in nature. However, since the electron state remains phase coherent during all the scattering events, this is not classical diffusion, but quantum diffusion. (iii) The classical regime, $L \gg l, L_\phi, L_T$. This represents the normal, everyday world, where quantum interference effects have been completely washed away. Electron transport is described by classical diffusion.

The remainder of this thesis is concerned primarily with case (ii) above, where quantum diffusive effects dominate the transport behaviour of the systems being measured. This will be referred to as the mesoscopic regime. Ballistic phenomena are extremely difficult to obtain in the n^+ -GaAs used. However, the dimensions of structures fabricated in n^+ -GaAs by e -beam lithography and dry-etching are small enough that they do fall within the mesoscopic regime.

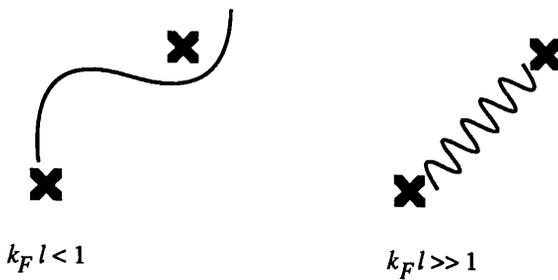


Figure 5.4: Highly disordered systems can be characterised by the condition $k_F l < 1$, where the electron wavelength is comparable with the scattering length. For weak disorder, $k_F l \gg 1$, where the electron wavelength is much smaller than the scattering length.

Quantum electronic systems can also be characterised in terms of the value of the product $k_F l$, where k_F is the Fermi wavevector, and l the elastic mean free path. The limit $k_F l < 1$ is usually satisfied in the presence of strong disorder (small l). As seen in Fig. 5.4, this corresponds to the case where the wavelength of the electron is bigger than the distance between scatterers. The electron state becomes localised, and no conduction can take place. The inequality $k_F l \gg 1$, where the electron manages to propagate some distance between scattering events, is called the weak disorder limit. The special significance attached to the value of the product $k_F l$ is that $k_F l \gg 1$ corresponds to the limit where the quantum corrections to classical transport parameters are (*in principle*) small, and can be calculated using perturbative expansions in the small parameter $(k_F l)^{-1}$. Note that quantum effects will arise irrespective of the value of $k_F l$ provided $L_\phi > l$, but it is the case $k_F l \gg 1$ that will be relevant to the systems studied in Chapter 6.

5.2.5 The Landauer formula

The fact that electrons behave like waves for sufficiently small samples ($L < L_\phi$) has required a reappraisal of the usual methods of calculating conductances (*e.g.* via the Boltzmann equation). A relatively concise formulation of the problem has been achieved by Büttiker (1986), extending an idea originally due to Landauer [for reviews, see Datta and McLennan (1990), or Stone and Szafer (1988)].

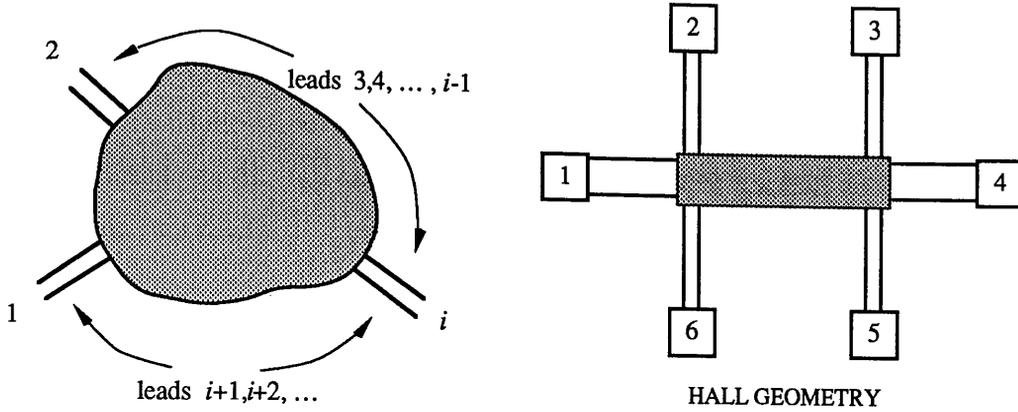


Figure 5.5: The Landauer formula relates the conductance of an arbitrary mesoscopic system to the electron transmission probabilities between the different leads attached to it. The Hall geometry has six such leads.

Consider the system shown in Fig. 5.5, which shows a conducting region with a number of leads attached. For example, this could be a Hall geometry, with six leads. The leads can be used either for passing current or for measuring voltage. Voltage probes are defined by the condition that no net current should flow into them.

A quantity $R_{ij,mn}$ may be defined, which is the resistance measured across leads m and n (*i.e.* m and n are the voltage probes) with current being driven through leads i and j . In the Hall geometry of Fig. 5.5, $R_{14,23}$ would describe an ordinary 4-probe longitudinal resistance measurement, and $R_{14,26}$ a Hall measurement. If $t_{kl}(E)$ is the probability amplitude of an electron of energy E being transmitted from lead l into lead k , and μ_k is the chemical potential of the electrons in lead k , then a Landauer formula relates the current in any particular lead to the $t_{kl}(E)$ and μ_k [see Datta and McLennan (1990)]. The measured resistance may then be written in the form,

$$R_{ij,mn} = \frac{h}{e^2} \frac{\mu_m - \mu_n}{S_{ij}}, \quad (5.18)$$

where S_{ij} is some function of the $t_{kl}(E)$ and μ_k , and $(\mu_m - \mu_n)/e$ is the potential difference between leads m and n . This method of calculating the quantum conductance of a phase coherent sample is called the Landauer-Büttiker formalism.

The interesting quantity in Eq. (5.18) which brings out the quantum mechanical nature of the problem, is S_{ij} . The quantity S_{ij} relates to how well the sample scatters the electron waves entering it into the output leads, and therefore encompasses all the

quantum interference effects that are occurring. Alternatively, if the wavefunction could be found for the electron in the sample, S_{ij} would be related to the modulus square of the amplitude of the wavefunction in leads i and j . Explicit expressions for S_{ij} can be found in simple cases. For example, this has been done for the Hall geometry of Fig. 5.5 by Büttiker (1988).

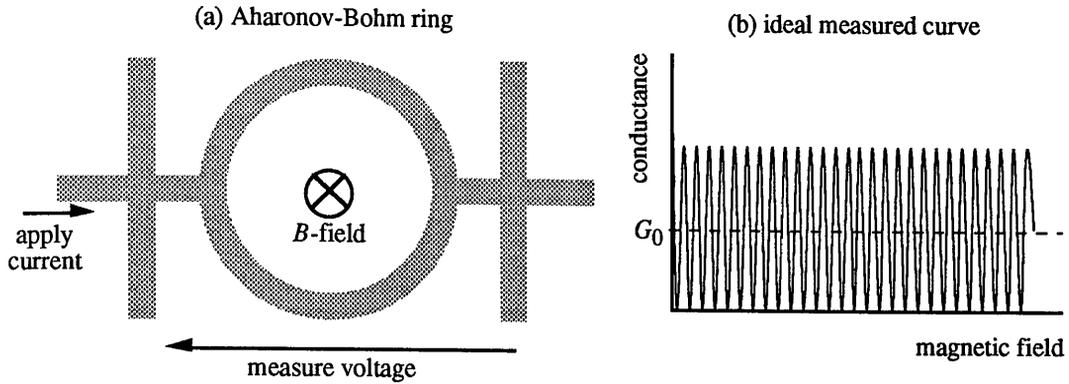


Figure 5.6: Device used to observe the Aharonov-Bohm oscillations. These oscillations occur in the conductance as a function of applied perpendicular magnetic field.

5.3 Quantum interference effects

5.3.1 The Aharonov-Bohm effect

Although the Aharonov-Bohm effect is not studied experimentally in this thesis, it does provide a useful frame of reference for explaining effects which are (see §5.3.2 and §5.3.3). One way of understanding the Aharonov-Bohm effect is demonstrated in Fig. 5.6. Current is passed through a loop, and the voltage measured between two sidearms [Fig. 5.6(a)]. If the electron wavefunction is coherent across the device, then application of a magnetic field through the loop causes its conductance to oscillate about a mean value, G_0 , with a period:

$$\Delta B = \frac{h}{e S}, \quad (5.19)$$

where S is the area of the loop. This is shown in Fig. 5.6(b).

The implications of Eq. (5.19) are discussed, for example, in Aharonov and Bohm (1959), or in Olariu and Popescu (1985). A qualitative explanation of the effect is

readily obtained. The electron propagates as a wave in both the upper and lower arms of the loop, the waves interfering upon reaching the output of the device. If the interference at the output is constructive, then the electrons pass through and the conductance is high; if it is destructive, then the electrons cannot get through and the conductance is low. A magnetic field passing through the loop affects the phases of the electron waves in the upper and lower arms differently. The net phase difference between the interfering waves is given by,

$$\delta\phi = \frac{e}{\hbar} \int_{loop} \mathbf{A} \cdot d\mathbf{l} = \frac{e}{\hbar} \Phi, \quad (5.20)$$

where \mathbf{A} is the vector potential, Φ is the flux passing through the loop, and the line integral is performed around the loop. The line integral is converted into a surface integral using Stokes' Theorem. Eq. (5.19) for the period of the conductance oscillations is recovered by requiring that $\delta\phi = 2\pi$.

The Aharonov-Bohm effect as described above (strictly speaking, this is only an Aharonov-Bohm effect if \mathbf{B} is confined to the middle of the loop) was first observed by Chambers (1960) using a magnetised iron whisker in the column of an electron microscope. The geometry of Fig. 5.6 was used to observe the solid-state version of this effect many years later, by Webb *et al.* (1985). They used a gold loop, fabricated by *e*-beam lithography. MilliKelvin temperatures were required to observe the oscillations which had a very small amplitude ($< 0.5\%$ of the mean). Since then, the effect has also been observed in semiconductor [Timp *et al.* (1987)] as well as other metallic systems. The conductance-field curves are usually further complicated by the presence of lower-frequency aperiodic oscillations, arising from the disorder in the system (see §5.3.3). The h/e oscillations can usually be picked out by Fourier transforming the data with respect to the field variable.

The h/e Aharonov-Bohm effect will be utilised in the explanation of universal conductance fluctuations in §5.3.3. There also exists an $h/2e$ effect (as well as higher harmonics) as found experimentally by Sharvin and Sharvin (1980). This is illustrated in Fig. 5.7. In the h/e effect the two electron paths depart from P and interfere at Q; in the $h/2e$ effect, the two paths start *and end* at the same point. In the h/e effect, the area is encircled by a single charge e , giving Eq. (5.20) for the phase change, and (5.19) for the periodicity; for the $h/2e$ effect, the single charge is replaced by a double charge, so in Eqs. (5.19) and (5.20) the substitution $e \rightarrow 2e$ should be made. The $h/2e$ oscillations in fact comprise coherent backscattering (where the two paths differ only in

sense – see §5.3.2) as well as the first harmonic Aharonov-Bohm oscillations (where the two paths can differ completely). The h/e and $h/2e$ oscillations can be found in the same sample, and are distinguished by Fourier transforming the data [e.g. Ford *et al.* (1988)].

Mention should also be made of the significance of the Aharonov-Bohm effect to physics. In performing the experiment on the device shown in Fig. 5.6(a), the magnetic field is applied normal to the plane of the page. In the true Aharonov-Bohm effect, the magnetic field would be non-zero *only* inside the loop, such that the electrons in the loop experience *zero magnetic field*. Then, changing the field at the centre of the loop would still cause oscillations at the output, *even though the electrons in the device never experience any magnetic field*. In classical physics, the potentials are largely a mathematical aid to calculating the fields: the electrostatic and vector potentials (ϕ, \mathbf{A}) are in some respects superfluous to the classical description of nature, and what really matters are a knowledge of the electric and magnetic field strengths (\mathbf{E} and \mathbf{B}). However, the Aharonov-Bohm effect shows that at a quantum level, \mathbf{E} and \mathbf{B} do not fully describe nature, and in fact it is ϕ and \mathbf{A} that are the important quantities.

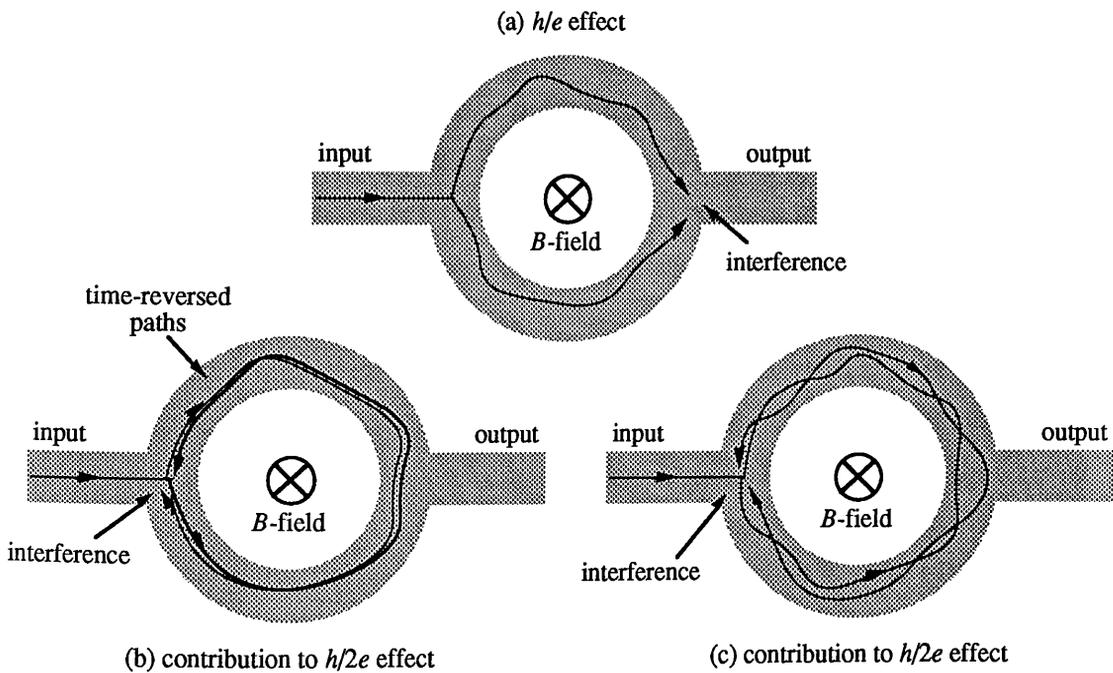


Figure 5.7: (a) Normal h/e Aharonov-Bohm effect. Two mechanisms contribute to the $h/2e$ effect: (b) coherent backscattering (§5.3.2), and (c) a true harmonic.

The Aharonov-Bohm effect is not isolated to magnetic fluxes. An electrostatic version also exists [Washburn *et al.* (1987)]. Indeed, these two effects can be combined using the 4-vector potential $A^\mu = (\phi, \mathbf{A})$. The gravitational Aharonov-Bohm effect has also been observed [Colella *et al.* (1975)]. Wu and Yang (1975) have shown that in theory the effect can occur for nuclear, in addition to electromagnetic and gravitational, interactions.

5.3.2 Weak localisation

Since impurity scattering is usually elastic (non phase breaking), quantum effects can be expected even when the system contains large numbers of such scatterers. The two limits of weak and strong disorder should be distinguished. If the density of scatterers is extremely high, it is found that electron states which would classically be expected to diffuse, actually become "trapped" (localised) within a certain region of the material when quantum effects are taken into account [Anderson (1958)]. In the opposite limit of weak disorder, scattering of the electron waves from the impurity potentials gives rise to quantum corrections which modify the classical conductivity [see, for example, Ashcroft and Mermin (1976), Ch. 1, for a discussion of the Drude model],

$$\sigma_0 = \frac{ne^2\tau}{m^*}, \quad (5.21)$$

where n is the electron concentration, τ is the mean free time between elastic scattering events, and m^* is the effective mass of the electron. These corrections usually result in a decrease in the conductivity of the sample. In contrast to the previous (strong) form of localisation, this is called weak localisation. It should be noted that the conductivity can also be decreased via a modification to the normal electron-electron interaction due to the disorder (see later).

The existence of weak localisation was first noted by Abrahams *et al.* (1979). A physical picture of the effect has been given by Bergmann (1984). The effect of weak localisation on the conductivity may be understood from Fig. 5.8. This shows two electron paths emanating from a point P, and returning to the *same* point. Both electron waves follow the *same* contour, but in opposite directions. The paths are said to be time-reversed.² This is basically the same as the loop shown in Fig. 5.7(b) for

² The time-reversal operator is defined $T: (\mathbf{r}, \mathbf{p}, \sigma) \rightarrow (\mathbf{r}, -\mathbf{p}, -\sigma)$, where \mathbf{r} and \mathbf{p} are the position and momentum of the electron, and σ is its spin. Applying T to the Schrödinger equation, $i\hbar\dot{\psi} = H\psi$, gives $i\hbar\dot{\psi} = TH\psi$. For $B = 0$, and for zero coupling to impurity spins, the Hamiltonian is time-reversal invariant, *i.e.* $TH = H$. Thus, the Schrödinger equation becomes $-i\hbar\dot{\psi} = H\psi$, and the only difference

the $h/2e$ Aharonov-Bohm effect. The arguments used here are similar to those used to explain coherence in §5.2.1. The wavefunction amplitudes of the two scattered electron waves in Fig. 5.8 are written ψ_+ and ψ_- . Whether interference occurs or not as a result of the electron waves traversing the loop and meeting back at P again depends on the type of scattering present in the sample. If the scattering is incoherent, then the wave "intensities" add:

$$|\psi_P|^2 = |\psi_+|^2 + |\psi_-|^2 = 2|\psi|^2, \quad (5.22)$$

where ψ_P is the wavefunction amplitude of the electron once it has returned to P, and $|\psi|^2 = |\psi_+|^2 = |\psi_-|^2$. There is no interference. This corresponds to the case of classical diffusion, so Eq. (5.22) is the classical result.

If on the other hand the scattering is coherent, then interference occurs at P, and,

$$|\psi_P|^2 = |\psi_+ + \psi_-|^2 = 4|\psi|^2, \quad (5.23)$$

where the superposition at P is phase coherent, so the amplitudes of the two electron waves are equal ($\psi_+ = \psi_- = \psi$). Eq. (5.23) is the quantum equivalent of Eq. (5.22). The probability of an electron returning to its point of departure is thus seen to be twice as high for quantum diffusion, as compared to classical diffusion. This is the phenomenon of weak localisation. The quantum diffusion coefficient (and hence the conductivity) is therefore lower than the value expected classically.

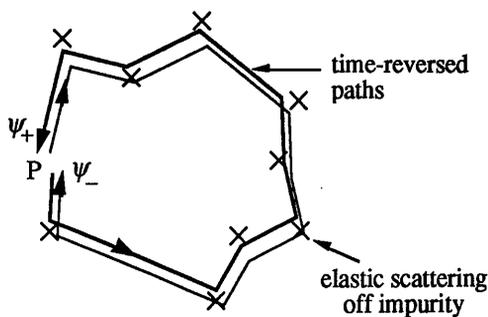


Figure 5.8: The origin of weak localisation lies in the interference of two time-reversed paths.

The crosses represent elastic impurity scattering. With no magnetic field applied, the interference at P is completely phase coherent.

is a change in the sign of the time [Zeiger and Pratt (1973), App. 9]. Saying that two electron trajectories are time-reversed means that one can be transformed into the other simply by reversing the direction of the arrow.

The process depicted in Fig. 5.8, where the two paths are time-reversed, is called coherent backscattering [Bergmann (1984)]. Such processes can significantly modify the Drude conductivity [as originally noticed by Langer and Neal (1966)]. The calculated conductivity depends on the dimensionality of the sample [Abrahams *et al.* (1979); Lee and Ramakrishnan (1985)],

$$\sigma_{3D}(L) = \sigma_0 + \frac{e^2}{\hbar\pi^3} \frac{1}{L}, \quad (5.24a)$$

$$\sigma_{2D}(L) = \sigma_0 - \frac{e^2}{\hbar\pi^2} \ln \frac{L}{l}, \quad (5.24b)$$

$$\sigma_{1D}(L) = \sigma_0 - \frac{e^2}{\hbar\pi} L, \quad (5.24c)$$

where σ_0 is the Drude conductivity (5.21), L is the sample size, and l is the elastic mean free path. An interesting point to note about these expressions is their sample size dependence. *Conductances*, being extensive quantities, are expected to show a size dependence. *Conductivities*, on the other hand, are classically intensive quantities, and should show no such dependence. Eqs. (5.24) show the surprising result that quantum conductivities *do* show a dependence on the sample size.

The above expressions apply at the absolute zero of temperature, where electron coherence is expected to extend over an infinite volume. For nonzero temperatures, the effective phase coherent region is reduced by phase breaking scattering processes. In this case, the sample size L in Eqs. (5.24) should be replaced by the coherence length, L_ϕ . An explicit temperature dependence can be obtained by writing $L_\phi = (D\tau_\phi)^{1/2}$. Assuming a power-law variation of the dephasing rate with temperature [$\tau_\phi^{-1} \propto T^p$] gives $L_\phi = aT^{-p/2}$, where a is a proportionality factor. This then gives [Abrahams and Ramakrishnan (1980); Lee and Ramakrishnan (1985)]:

$$\sigma_{3D}(T) = \sigma_0 + \frac{e^2}{\hbar\pi^3} \frac{1}{a} T^{p/2}, \quad (5.25a)$$

$$\sigma_{2D}(T) = \sigma_0 + \frac{p}{2} \frac{e^2}{\hbar\pi^2} \ln \left(\frac{T}{T_0} \right), \quad (5.25b)$$

$$\sigma_{1D}(T) = \sigma_0 - \frac{ae^2}{\hbar\pi} T^{-p/2}, \quad (5.25c)$$

where T_0 is a constant.

Eq. (5.25b) shows that the correction to the two-dimensional conductivity is logarithmic in temperature. Experiments to investigate this effect, by Dolan and Oshraff (1979), showed that there was indeed a $\ln T$ dependence in the conductivity of thin films, but the magnitude of the prefactor did not agree with the theory [Anderson *et al.*

(1979)]. Soon afterwards, Al'tshuler *et al.* (1980) [see also Finkel'shtein (1983), and Al'tshuler and Aronov (1983)] showed that a $\ln T$ dependence in the conductivity also arises from the electron-electron interaction as modified by the disorder [Lee and Ramakrishnan (1985)]:

$$\sigma_{2D}(T) = \sigma_0 + \frac{e^2}{4\hbar\pi^2} \left[2 - \frac{3}{2}F \right] \ln(T\tau), \quad F \ll 1, \quad (5.26)$$

where the two terms in the square brackets represent the exchange and direct interactions, and τ is the mean time between elastic scattering events. To lowest order, the localisation and interaction effects are additive [Lee and Ramakrishnan (1985)], so the logarithmic dependence from the interaction term, in a sense, disguises the true weak localisation effect. In dimensions other than 2, a value of $p = 1$ in Eqs. (5.25) leads to a similar indistinguishability in the temperature dependences of the two effects.

It would appear that the two effects cannot always be distinguished using simple conductance measurements. Fortunately, however, the localisation and interaction effects can be distinguished by another means. Coherent backscattering exists only because of the totally constructive interference occurring at point P between the two time-reversed paths. If the strict phase relationship between the two electron waves is broken in any way, the effect is immediately diminished. Application of a magnetic field does just that. By analogy with the $h/2e$ Aharonov-Bohm effect, any magnetic flux passing through the loop of Fig. 5.8 will induce a phase difference between the two paths. Since the interference at P can no longer be completely constructive, so the effects of weak localisation begin to disappear, and the conductivity rises. This results in characteristic positive slopes at low fields in the magnetoconductance curves of weakly disordered samples.

As a function of the magnetic field, the conductivity of a two-dimensional weakly disordered sample can be written [Al'tshuler *et al.* (1980); Hikami *et al.* (1980)]:

$$\sigma_{2D}(B,T) = \sigma_{2D}(0,T) - \frac{e^2}{2\hbar\pi^2} \left[\psi\left(\frac{1}{2} + \frac{\hbar}{4eBD\tau}\right) - \psi\left(\frac{1}{2} + \frac{\hbar}{4eBD\tau_\phi}\right) \right], \quad (5.27)$$

where τ_ϕ is the mean time between phase breaking collisions, $\psi(x)$ is the digamma function [see Abramowitz and Stegun (1965) for its properties], and the magnetic field B is applied normal to the sample. The term $\sigma_{2D}(0,T)$ is the combined conductivity arising from both localisation and interaction effects. In some limits, the above expression can be simplified a little by using the asymptotic ($x \gg 1$) approximation for the digamma function, $\psi\left(\frac{1}{2} + x\right) \approx \ln x$. In the limit $L_B \gg l$, the first digamma

function in Eq. (5.27) can be approximated by a logarithm. For very small fields, the condition $L_B \gg L_\phi$ holds, and the second digamma function can also be approximated by a logarithm. In this case, Eq. (5.27) basically reduces to Eq. (5.24b).

For thin wires, the result is different, given by [Al'tshuler and Aronov (1981)]:

$$\sigma_{1D}(B, T) = \sigma_{1D}(0, T) - \frac{e^2}{\hbar\pi} \left(\frac{1}{L_\phi^2} + \frac{1}{D\tau_B} \right)^{-1/2}, \quad (5.28)$$

where $D\tau_B = 3L_B^4/w^2$ for a wire of rectangular cross-section, w is the wire width, $L_B = (\hbar/eB)^{1/2}$ is the magnetic length, and the field is applied in a direction perpendicular to both the length and the width of the wire. In the limit $B \rightarrow 0$, Eq. (5.28) reduces to Eq. (5.24c).

Continuing the analogy with the $h/2e$ Aharonov-Bohm effect, a simple estimate may be made of the approximate field strength required to completely suppress the weak localisation. Fig. 5.7(b) shows that passing a flux of $h/2e$ through the area encircled by the two electron paths induces a phase change of 2π between these paths. Thus, coherent backscattering between the time-reversed paths in Fig. 5.8 should be completely suppressed by passing a flux of roughly $h/2e$ through the area traced out by the paths. Thus, the positive magnetoconductance will last up to field of roughly,

$$\Delta B_{WL} \sim \frac{h}{2e} \frac{1}{L_\phi^2}, \quad (5.29)$$

where L_ϕ^2 is the area over which the electron wave is coherent. [Up to a factor of 2π , this is the same as saying that weak localisation is suppressed when the magnetic length becomes smaller than the coherence length, $L_B < L_\phi$, where $L_B = (\hbar/2eB)^{1/2}$ is the magnetic length for a particle with charge $2e^-$.] Since the interaction contribution to the conductivity is unaffected by low magnetic fields, the two causes of the $\ln T$ dependence can thus be distinguished. The interaction effect is modified by a stronger magnetic field. This happens when the spin degeneracy of the time-reversed paths is lifted, and strong spin scattering exists in the material ($g\mu_B B \gg k_B T$, and $g\mu_B B \gg \hbar\tau_{SO}^{-1}$ or $\hbar\tau_s^{-1}$).

Bergmann (1984) has shown that the formulae (5.27) and (5.28), expressing the weak localisation correction to the conductivity as a function of the applied magnetic field, can be very useful in extracting the coherence length, L_ϕ , in the sample. All the parameters in Eqs. (5.27) and (5.28) are known, with the exception of L_ϕ . This means that L_ϕ can be used as an "adjustable" parameter to fit the theory to the experiment (*i.e.* fit the equations to the positive magnetoconductance at small field). Values of L_ϕ obtained in this way have been found to agree very well with estimates obtained by other

methods. Having fitted the weak localisation effect, parameters relating to the interaction effect can then be deduced [Lee and Ramakrishnan (1985)].

It should be noted that interference between the two paths in Fig. 5.8(a) need not lead to weak localisation and positive magnetoconductance. When there is strong spin-orbit scattering in the sample (due to the presence of heavy-nucleus impurities, *e.g.* Au) then the interference between the two paths is actually slightly destructive, *i.e.* the conductivity is slightly higher than the Drude value. This is called weak *anti*-localisation. Application of a magnetic field now causes a *negative* magnetoconductance. An expression analogous to Eq. (5.27) can be derived, which includes the effects of spin-orbit scattering [Hikami *et al.* (1980); Bergmann (1984)]. This allows both L_ϕ and L_{SO} to be extracted from the measurements.

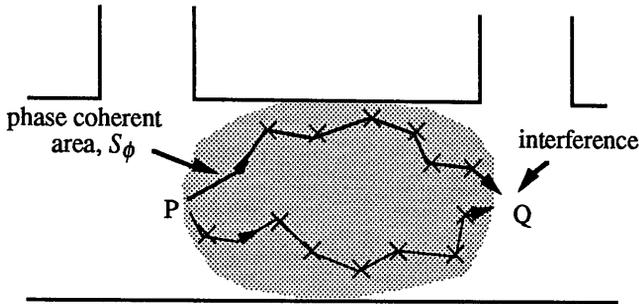


Figure 5.9: Typical electron trajectories across a phase coherent sample. Interference between such trajectories leads to fluctuations in the conductance.

5.3.3 Universal conductance fluctuations

It was mentioned in connection with the Landauer formula, Eq. (5.18), that the quantum conductance of a phase coherent sample was determined by how it scattered incoming electron waves. Measurements of weak localisation above can also be thought of in terms of the Landauer formula, where the incoming electron waves are backscattered to exactly the same point from which they entered the sample (the word "sample" here is used in the same sense as in Fig. 5.5, *i.e.* not including the leads). However, in the actual measurement of the longitudinal conductance, there are normal scattered trajectories going from one current lead to the other, as shown in Fig. 5.9, in addition to the backscattered trajectories. For phase coherent samples, these trajectories

are found to significantly affect the quantum transport behaviour.

The interference arising from backscattered trajectories is simple in the sense that the two electron waves always add coherently at the return point. The interference arising from the trajectories of Fig. 5.9 are much more complex. In numerical studies, Stone (1985) showed that the interference depended in a crucial way on the impurity configuration in the sample. The changes in the interference patterns led to changes in the calculated conductance of the sample, as given by the Landauer formula: different impurity configurations generally gave different quantum conductances.

The extreme sensitivity of the conductance to the impurity configuration is the central characteristic of the type of interference. The argument can be stated in a more quantitative way. Consider an extremely large number of phase coherent samples (an ensemble) which are identical in every respect, except in the positions of their impurities. Define G to be the conductance of a typical sample, and $G_0 \equiv \langle G \rangle$ to be the ensemble average of the conductance, where ensemble averaging (averaging over all impurity configurations) is denoted by the angular brackets. The fluctuation of the conductance of a particular sample from the mean conductance is $\delta G = G - G_0$. The root mean square fluctuation in conductance is thus,

$$\delta G_{rms} = \langle \delta G^2 \rangle^{1/2} = [\langle G^2 \rangle - G_0^2]^{1/2}. \quad (5.30)$$

Al'tshuler (1985), and Lee and Stone (1985) showed the remarkable result that the root mean square (rms) fluctuation of the conductances of the samples in this ensemble, is a constant: $\delta G_{rms} \sim e^2/h$. This result is independent of both the detailed form of the impurity potential, and the exact geometry of the sample. The only requirement is that the sample be phase coherent.

The distribution function for small fluctuations is approximately Gaussian [Al'tshuler *et al.* (1986)]:

$$f(\delta G) \propto \exp\left(-\frac{\delta G^2}{2\delta G_{rms}^2}\right), \quad \delta G < \frac{e^2}{h} \ll G_0, \quad (5.31)$$

i.e. $f(\delta G)\Delta G$ is proportional to the number of samples with a conductance lying in the range $G_0 + G \pm \frac{\Delta G}{2}$ (or $G_0 - G \pm \frac{\Delta G}{2}$). Al'tshuler *et al.* (1986) have shown that for large δG , the distribution function deviates from the Gaussian, developing a $(\ln)^2$ dependence within the exponential. From the arguments following Eq. (5.30), it is understood that the *same* distribution (5.31) applies, irrespective of the value of G_0 , provided $G_0 \gg e^2/h$.

The above result shows that any two samples randomly chosen from the ensemble

will have conductances which differ on average by $\sim e^2/h$. Feng, Lee and Stone (1987) have shown that two such samples need differ by the position of just a *single* impurity. They show that moving a strong scattering centre a distance of roughly k_F^{-1} is sufficient to change the conductance (on average) by $\sim e^2/h$. This shows just how sensitive the quantum conductance is on what appear to be minute details.

The theoretical studies described above were initiated in an effort to understand some very complex low temperature magnetoresistance curves measured in the experiments of Umbach *et al.* (1984). These experiments had initially set out to find the periodic Aharonov-Bohm oscillations in very small metal rings. Instead they found a complex *aperiodic* structure, where the conductance varied apparently randomly *but as a precise function* of the applied magnetic field, B . The curves were completely reproducible over long periods of time, provided the sample was maintained at low temperatures. Experiments on Silicon MOS inversion layers [*e.g.* Kwasnick *et al.* (1984); Licini *et al.* (1985)] revealed similarly complex conductances curves, but this time as a function of the Fermi level, E_f (which is related to the gate voltage). Again, these curves were completely reproducible provided the sample was maintained at a low temperature. Lee and Stone (1985) coined the term "universal conductance fluctuations" (UCFs) to describe these effects. These are "universal" in the sense that the fluctuations are insensitive to the degree of disorder, or the precise geometry of the sample, provided the sample is phase coherent.

It is not immediately apparent how the theory, where the impurity positions are changed, relates to the experiment, where either B or E_f is varied (but the impurity configuration does not change³). In fact, Al'tshuler *et al.* (1986) have shown that the above are all equivalent, and that changing B or E_f by a certain amount corresponds to completely changing the impurity configuration. This is the basis of the "ergodic hypothesis" which states exactly the above equivalence. [The ergodic hypothesis is fundamental to the ensemble description of statistical thermodynamic systems. As an example, consider the time-averaged behaviour of such a statistical system. Rather than study the dynamics of the system as a function of time, the Gibbs prescription of "hiding" the dynamics by considering an infinite ensemble of similar systems, is followed. The ensemble in effect contains all possible states that the system could be in as it evolves with time. The ergodic hypothesis states that the final result should be the

³ This point has been studied in detail by Taylor *et al.* (1988) for n^+ -GaAs wires. The impurity configuration in these samples was found to change significantly only if the sample temperature was raised above ~ 80 K.

same for both descriptions. This need not hold in general, however. See Landau and Lifshitz (1986), Ch. 1. In the present case, three parameters can be varied. Al'tshuler *et al.* (1986) effectively showed that the *same* ensemble can be used to describe the "evolution" of the phase coherent sample as any of these parameters is varied.] Physically, as either B or E_f is changed, the potential "landscape" seen by the electrons at the Fermi surface is also changed. This then leads to the fluctuations in the conductance. By the ergodic hypothesis, these fluctuations should also have an rms value $\sim e^2/h$.

The study of fluctuations is assisted considerably using the correlation function $F(\Delta E, \Delta B)$, introduced by Lee and Stone (1985):

$$\begin{aligned} F(\Delta E, \Delta B) &= \langle [G(E_f, B) - G_0] \times [G(E_f + \Delta E, B + \Delta B) - G_0] \rangle \\ &= \langle G(E_f, B) G(E_f + \Delta E, B + \Delta B) \rangle - G_0^2, \end{aligned} \quad (5.32)$$

where $G_0 = \langle G(E_f, B) \rangle$ is the mean conductance, and the angular brackets denote ensemble averaging. It further turns out [Lee *et al.* (1987)] that F is *not* a function of E_f , but *is* a function of B . The correlation function gives an indication of how the conductances at two different values of field compare. Thus, $F(0, B_0) = 1$ indicates perfect correlation with periodicity B_0 in field, *i.e.* $G(B) = G(B + B_0)$ always, whereas $F(0, B_0) = 0$ indicates no correlation. The expected form of the correlation function is shown in Fig. 5.10 [Lee and Stone (1985)]. At $\Delta E, \Delta B = 0$, the correlation is perfect. However, for larger differences in energy or field, the correlation decays, and eventually tends to zero (as is expected for "random" signals).

The correlation function evaluated at $\Delta E = \Delta B = 0$, gives the variance of the conductance fluctuations, so the rms value is just the square root:

$$\delta G_{rms} = \sqrt{F(\Delta E=0, \Delta B=0)}, \quad (5.33)$$

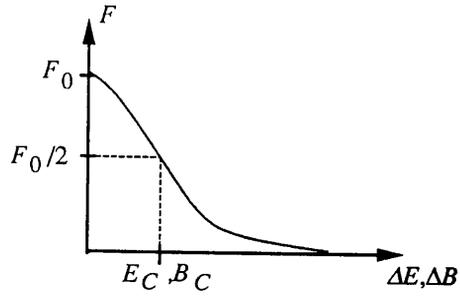
For zero magnetic field (*e.g.* in Si MOS inversion layer systems), the corresponding rms fluctuations turn out to be [Lee *et al.* (1987)],

$$\frac{\delta G_{rms}}{e^2/h} = \begin{cases} 1.088, & d = 3 \\ 0.862, & d = 2 \\ 0.729, & d = 1 \end{cases} \quad \text{for } B = 0, \quad (5.34)$$

where d is the dimensionality of the phase coherent region (assumed to be a cube, square, or wire). At stronger magnetic fields, the rms fluctuations are different [Lee *et al.* (1987)]:

$$\frac{\delta G_{rms}}{e^2/h} = \begin{cases} 0.78, & d = 3 \\ 0.61, & d = 2 \\ 0.52, & d = 1 \end{cases} \quad \text{for } B > \Delta B_{WL}, \quad (5.35)$$

Figure 5.10: Expected correlation function for a UCF measurement. The variance of the fluctuations, and the correlation energy and field can be obtained from the curve.



where ΔB_{WL} is defined in Eq. (5.29). These differ from the $B = 0$ rms fluctuations by a factor of $\frac{1}{\sqrt{2}}$. Spin-orbit scattering, magnetic scattering, and Zeeman splitting of the electron spins also modify δG_{rms} [Lee *et al.* (1987); Feng (1989)].

The value $F = F_0/2$ (see Fig. 5.10) also has a special significance in that it defines the correlation energy, E_C , and the correlation field, B_C :

$$F(E_C, 0) = F(0, B_C) = \frac{F_0}{2}. \quad (5.36)$$

The correlation terms, E_C and B_C , are the characteristic scales between dips or peaks in the fluctuation curves. The correlation energy turns out to be [Lee and Stone (1985)]:

$$E_C \approx \frac{\pi^2 \hbar D}{L^2}, \quad (5.37)$$

where D is the electron diffusion coefficient, and L is the sample size. Up to a factor of π^2 , this is the same as the correlation energy that was defined in Eq. (5.16). In both cases, E_C is the change in electron energy required to take the electron "out" of the coherent band [*c.f.* Fig. 5.3] it is currently in, and place it in an uncorrelated band.

The correlation field is [Lee *et al.* (1987)]:

$$B_C \sim \frac{\hbar}{e S_\phi}, \quad (5.38)$$

where S_ϕ is the phase coherent area normal to the magnetic field. [This statement requires further elaboration when energy averaging is expected to occur. For details, see Lee *et al.* (1987).] This value for correlation field can be obtained by a simple heuristic argument. By analogy with the h/e Aharonov-Bohm effect, a phase change of 2π between the two paths shown in Fig. 5.9 would be sufficient to completely change the interference pattern. This would correspond to a flux of h/e passing through the area of the loop, which for typical pairs of paths is the sample area, S_ϕ . A flux of h/e through an area S_ϕ gives the field B_C in Eq. (5.38).

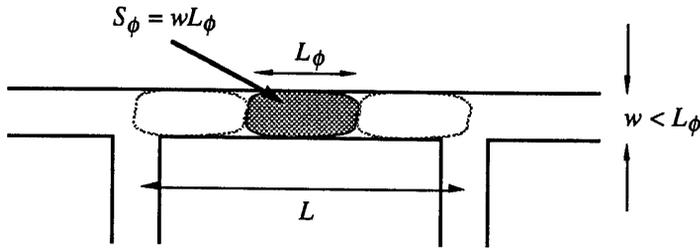


Figure 5.11: In a long wire, the fluctuations from several phase coherent regions add incoherently. The result is an rms conductance fluctuation smaller than expected.

The values of rms fluctuations given above are valid for $T = 0$, and for a phase coherent sample, $L \sim L_\phi$. At finite temperatures, or for $L > L_\phi$, averaging will occur and the value of δG_{rms} will be reduced. For simplicity consider the quasi-1-dimensional wire of Fig. 5.11, which has width $w < L_\phi$, and length L . If $L > L_\phi$, then the sample is no longer phase coherent, there being roughly $N = L/L_\phi$ coherent regions within the sample. Let R_ϕ be the resistance of one of these phase coherent regions, and R be the resistance of the whole wire, so that $R_\phi = R/N$. The rms conductance fluctuation of one of these regions is $\delta G_\phi = e^2/h$. The corresponding resistance fluctuation is found by differentiating $R_\phi = 1/G_\phi$, i.e. $\delta R_\phi = \delta G_\phi R_\phi^2$. If the conductance of each phase coherent region is fluctuating independently of the others, the total fluctuation will be the rms fluctuation of N resistors in series:

$$\begin{aligned} \delta R_{rms} &= \sqrt{(\delta R_1)^2 + (\delta R_2)^2 + \dots + (\delta R_N)^2} \\ &= \sqrt{N (\delta G_\phi R_\phi^2)^2} = \frac{1}{N^{3/2}} \frac{e^2}{h} R^2 . \end{aligned} \quad (5.39)$$

Hence,

$$\delta G_{rms} = \frac{\delta R_{rms}}{R^2} \sim \left(\frac{L_\phi}{L}\right)^{3/2} \frac{e^2}{h} , \quad L_\phi \ll L_T < L . \quad (5.40)$$

Thus, as N is made progressively larger, δG_{rms} becomes smaller, $\delta G_{rms} \propto N^{-3/2}$, and in the limit $N \rightarrow \infty$, the classical conductance for the sample is recovered.

More generally, the result (5.40) changes depending on the relative values of L , L_ϕ and L_T (the latter is the thermal diffusion length which describes energy averaging). The exact dependences can be found from the correlation function [Al'tshuler and Khmel'nitskii (1985); Lee *et al.* (1987)]:

$$\frac{\delta G_{rms}}{e^2/h} \sim \frac{L_T}{L} \left(\frac{L_\phi}{L}\right)^{1/2}, \quad d=1, \quad L_T \ll L_\phi < L, \quad (5.41a)$$

$$\frac{\delta G_{rms}}{e^2/h} \sim \left(\frac{L_T}{L}\right)^{(4-d)/2}, \quad d=2,3, \quad L_T \ll L_\phi < L, \quad (5.41b)$$

$$\frac{\delta G_{rms}}{e^2/h} \sim \left(\frac{L_\phi}{L}\right)^{(4-d)/2}, \quad d=1,2,3, \quad L_\phi \ll L_T < L, \quad (5.41c)$$

where d is the dimensionality of the sample with respect to interference effects. Eq. (5.40) is recovered from (5.41c) by putting $d = 1$. Dependences with L_{SO} , L_s and L_Z [where $L_Z = (\hbar D/g\mu_B B)^{1/2}$ – see §5.2.2] are discussed in Lee *et al.* (1987).

In cases where $L < L_\phi$, it is experimentally found that the measured rms fluctuation can exceed e^2/h [Benoit *et al.* (1987); Skoçpol *et al.* (1987)]. However, this does not mean that the above theory is incorrect. Instead, the theory is to be modified slightly, calling into question the role of the voltage probes in the quantum conductance measurement. This introduces ideas of "nonlocality" in these types of measurements. Nonlocality will be discussed further in Chapter 6.

5.4 References

- ABRAHAMS E, P W Anderson, D C Licciardello, T V Ramakrishnan, "Scaling theory of localisation: absence of quantum diffusion in two dimensions," *Phys. Rev. Lett.* **42**, 673 (1979).
- ABRAHAMS E, T V Ramakrishnan, "Scaling theory of localisation and non-ohmic effects in two dimensions," *J. Non-Cryst. Solids* **35**, 15 (1980).
- ABRAMOWITZ M, I A Stegun, *Handbook of Mathematical Functions*, Dover Publications Inc, New York (1965).
- AHARONOV Y, D Bohm, "Significance of electromagnetic potentials in the quantum theory," *Phys. Rev.* **115**, 485 (1959).
- AL'TSHULER B L, A G Aronov, P A Lee, "Interaction effects in disordered Fermi systems in two dimensions," *Phys. Rev. Lett.* **44**, 1288 (1980).
- AL'TSHULER B L, A G Aronov, "Magnetoresistance of thin films and of wires in a longitudinal magnetic field," *JETP Lett.* **33**, 499 (1981).
- AL'TSHULER B L, A G Aronov, "Fermi-liquid theory of the electron-electron interaction effects in disordered metals," *Solid State Comm.* **46**, 429 (1983).
- AL'TSHULER B L, "Fluctuations in the extrinsic conductivity of disordered

- conductors," *JETP Lett.* **41**, 648 (1985).
- AL'TSHULER B L, D E Khmel'nitskii, "Fluctuation properties of small conductors," *JETP Lett.* **42**, 359 (1985).
- AL'TSHULER B L, V E Kravtsov, I V Lerner, "Statistics of mesoscopic fluctuations and instability of one-parameter scaling," *Sov. Phys. JETP* **64**, 1352 (1986).
- ANDERSON P W, "Absence of diffusion in certain random lattices," *Phys. Rev.* **109**, 1492 (1958).
- ANDERSON P W, E Abrahams, T V Ramakrishnan, "Possible explanation of nonlinear conductivity in thin-film metal wires," *Phys. Rev. Lett.* **43**, 718 (1979).
- ANDO T, A B Fowler, F Stern, "Electronic properties of two-dimensional systems," *Rev. Mod. Phys.* **54**, 437 (1982).
- ASHCROFT N W, N D Mermin, *Solid State Physics*, Holt, Reinhart and Winston, New York (1976).
- BEENAKKER C W J, H van Houten, "Quantum transport in semiconductor nanostructures," *Adv. Phys.* **44**, 1 (1991).
- BENOIT A, C P Umbach, R B Laibowitz, R A Webb, "Length-independent voltage fluctuations in small devices," *Phys. Rev. Lett.* **58**, 2343 (1987).
- BERGMANN G, "Weak localisation in thin films," *Phys. Rep.* **107**, 1 (1984).
- BLATT F J, *Physics of Electronic Conduction in Solids*, McGraw-Hill Book Co., New York (1968).
- BUTTIKER M, "Four-terminal phase-coherent conductance," *Phys. Rev. Lett.* **57**, 1761 (1986).
- BUTTIKER M, "Absence of backscattering in the quantum Hall effect in multiprobe conductors," *Phys. Rev. B* **38**, 9375 (1988).
- CHAKRAVARTY S, A Schmid, "Weak localisation: the quasiclassical theory of electrons in a random potential," *Phys. Rep.* **140**, 193 (1986).
- CHAMBERS R G, "Shift of an electron interference pattern by enclosed magnetic flux," *Phys. Rev. Lett.* **5**, 3 (1960).
- COLELLA R, A W Overhauser, S A Werner, "Observation of gravitationally induced quantum interference," *Phys. Rev. Lett.* **34**, 1472 (1975).
- DATTA S, M J McLennan, "Quantum transport in ultrasmall electronic devices," *Rep. Prog. Phys.* **53**, 1003 (1990).
- DOLAN G J, D D Osheroff, "Nonmetallic conduction in thin metal films at low temperatures," *Phys. Rev. Lett.* **43**, 721 (1979).

- FENG S, P A Lee, A D Stone, "Sensitivity of the conductance of a disordered metal to the motion of a single atom: implications for $1/f$ noise," *Phys. Rev. Lett.* **56**, 1960, 2772(E) (1986).
- FENG S, "Mesoscopic conductance fluctuations in the presence of spin-orbit coupling and Zeeman splitting," *Phys. Rev. B* **39**, 8722 (1989).
- FINKEL'SHTEIN A M, "Influence of Coulomb interaction on the properties of disordered metals," *Sov. Phys. JETP* **57**, 97 (1983).
- FORD C J B, T J Thornton, R Newbury, M Pepper, H Ahmed, C T Foxon, J J Harris, C Roberts, "The Aharonov-Bohm effect in electrostatically defined heterojunction rings," *J. Phys. C* **21**, L325 (1988).
- van HAESSENDONCK C, H Vloeberghs, Y Bruynseraede, "Conductance oscillations and phase coherence in submicron metal films," in ed. B Kramer, *Quantum Coherence in Mesoscopic Systems*, Plenum Press, New York (1991).
- HIKAMI S, A I Larkin, Y Nagaoka, "Spin-orbit interaction and magnetoresistance in the two dimensional random system," *Prog. Theor. Phys.* **63**, 707 (1980).
- van HOUTEN H, C W J Beenakker, J G Williamson, M E I Broekaart, P H M van Loosdrecht, B J van Wees, J E Mooij, C T Foxon, J J Harris, "Coherent electron focusing with quantum point contacts in a two-dimensional electron gas," *Phys. Rev. B* **39**, 8556 (1989).
- KWASNICK R F, M A Kastner, J Melngailis, P A Lee, "Nonmonotonic variations of the conductance with electron density in ~ 70 nm-wide inversion layers," *Phys. Rev. Lett.* **52**, 224 (1984).
- LANDAU L D, E M Lifshitz, *Statistical Physics*, Pergamon Press, Oxford (1986).
- LANGER J S, T Neal, "Breakdown of the concentration expansion for the impurity resistivity of metals," *Phys. Rev. Lett.* **16**, 984 (1966).
- LEE P A, T V Ramakrishnan, "Disordered electronic systems," *Rev. Mod. Phys.* **57**, 287 (1985).
- LEE P A, A D Stone, "Universal conductance fluctuations in metals," *Phys. Rev. Lett.* **55**, 1622 (1985).
- LEE P A, A D Stone, H Fukuyama, "Universal conductance fluctuations in metals: effects of finite temperature, interactions, and magnetic field," *Phys. Rev. B* **35**, 1039 (1987).
- LICINI J C, D J Bishop, M A Kastner, J Melngailis, "Aperiodic magnetoresistance oscillations in narrow inversion layers in Si," *Phys. Rev. Lett.* **55**, 2987 (1985).

- MAHAN G D, *Many-Particle Physics*, Plenum Press, New York (1981).
- MESERVEY R, P M Tedrow, "Surface relaxation times of conduction-electron spins in superconductors and normal metals," *Phys. Rev. Lett.* **41**, 805 (1978).
- OLARIU S, I I Popescu, "The quantum effects of electromagnetic fluxes," *Rev. Mod. Phys.* **57**, 339 (1985).
- ROUKES M L, A Scherer, S J Allen Jr, H G Craighead, R M Ruthen, "Quenching of the Hall effect in a one-dimensional wire," *Phys. Rev. Lett.* **59**, 3011 (1987).
- SEEGER K, *Semiconductor Physics*, Springer-Verlag, Berlin (1985).
- SHARVIN D Yu, Yu V Sharvin, "Magnetic-flux quantization in cylindrical film of a normal metal," *JETP Lett.* **34**, 272 (1982).
- SKOŃPOL W J, P M Mankiewich, R E Howard, L D Jackel, D M Tennant, A D Stone, "Nonlocal potential measurements of quantum conductors," *Phys. Rev. Lett.* **58**, 2347 (1987).
- STERN A, Y Aharonov, Y Imry, "Phase uncertainty and loss of interference: a general picture," *Phys. Rev. A* **41**, 3436 (1990).
- STONE A D, "Magnetoresistance fluctuations in mesoscopic wires and rings," *Phys. Rev. Lett.* **54**, 2692 (1985).
- STONE A D, A Szafer, "What is measured when you measure a resistance? – The Landauer formula revisited," *IBM J. Res. Develop.* **32**, 384 (1988).
- TAYLOR R P, P C Main, L Eaves, S P Beaumont, S Thoms, C D W Wilkinson, "Aperiodic conductance fluctuations as a probe of changes in the microscopic scattering configuration in n^+ -GaAs:Si wires," in ed. W Zawadzki, *Proc. 19th Int. Conf. on the Physics of Semiconductors*, Institute of Physics, Polish Academy of Sciences, Warsaw (1988).
- THOULESS D J, "Maximum metallic resistance in thin wires," *Phys. Rev. Lett.* **39**, 1167 (1977).
- TIMP G, A M Chang, J E Cunningham, T Y Chang, P Mankiewich, R Behringer, R E Howard, "Observation of the Aharonov-Bohm effect for $\omega_c \tau > 1$," *Phys. Rev. Lett.* **58**, 2814 (1987).
- UMBACH C P, S Washburn, R B Laibowitz, R A Webb, "Magnetoresistance of small quasi-one-dimensional, normal metal rings and lines," *Phys. Rev. B* **30**, 4048 (1984).
- WEBB R A, S Washburn, C P Umbach, R B Laibowitz, "Observation of h/e Aharonov-Bohm oscillations in normal-metal rings," *Phys. Rev. Lett.* **54**, 2696

- (1985).
- van WEES B J, H van Houten, C W J Beenakker, J G Williamson, L P Kouwenhoven, D van der Marel, C T Foxon, "Quantized conductance of point contacts in a two-dimensional electron gas," *Phys. Rev. Lett.* **60**, 848 (1988).
- WHARAM D A, T J Thornton, R Newbury, M Pepper, H Ahmed, J E F Frost, D G Hasko, D C Peacock, D A Ritchie, G A C Jones, "One-dimensional transport and the quantisation of the ballistic resistance," *J. Phys. C* **21**, L209 (1988).
- WU T T, C N Yang, "Concept of nonintegrable phase factors and global formulation of gauge fields," *Phys. Rev. D* **12**, 3845 (1975).
- ZEIGER H J, G W Pratt, *Magnetic Interactions in Solids*, Clarendon Press, Oxford (1973).

Fluctuations and quantum noise phenomena

6.1 Outline

In this chapter, the low temperature conductance of submicron length SiCl_4 -etched n^+ -GaAs wires is studied. In these wires, strong conductance fluctuations are observed as a function of applied magnetic field, not only in the longitudinal measurement discussed in the theory of Chapter 5, but also in Hall and nonlocal configurations. The measurements of weak localisation and UCFs suggest that although electron scattering from the edges of the sidewall depletion region is predominantly diffuse elastic, a contribution to the phase breaking rate also arises. This is discussed in the context of the model on dry-etch damage in Chapter 4. The Hall measurement is shown to probe fluctuations associated with the phase coherent region at the junction.

The low temperature time-dependent response of SiCl_4 -etched n^+ -wires to pulse of light appears to show random jumps in the conductance. These random telegraph signals, as they are called, have been observed in both two-terminal and four-terminal wires, and also in transverse ("Hall") measurements. The theory for this is based on the sensitivity of the quantum conductance to the exact impurity configuration within the sample.

6.2 The longitudinal measurement

6.2.1 Sample fabrication and measurement system

The wires used here are fabricated in the way described in Chapter 2. The material is grown by MBE, and comprises a 50 nm n^+ -GaAs epilayer on top of an undoped GaAs buffer (see Fig. 2.7). This is the same material that was used in the experiments of Chapter 3. Van der Pauw measurements give a material carrier concentration of about $\sim 6.5 \times 10^{24} \text{ m}^{-3}$, and a mobility $\sim 0.14 \text{ m}^2 \text{ V}^{-1} \text{ s}^{-1}$ at room temperature. The etching is done in SiCl_4 at 300 V for 30 s, the mask being HRN. Note that the experiments discussed in Chapter 3 showed no difference (within experimental error) between using HRN and Ni as the etch mask. The reasons for using HRN have been summarised in §2.2.4. SEMs of typical structures that have been fabricated are shown in Fig. 6.1. After fabrication, the samples are bonded onto non-magnetic TO5 headers, and mounted in a ^4He cryostat equipped with a $\sim 6 \text{ T}$ superconducting magnet.

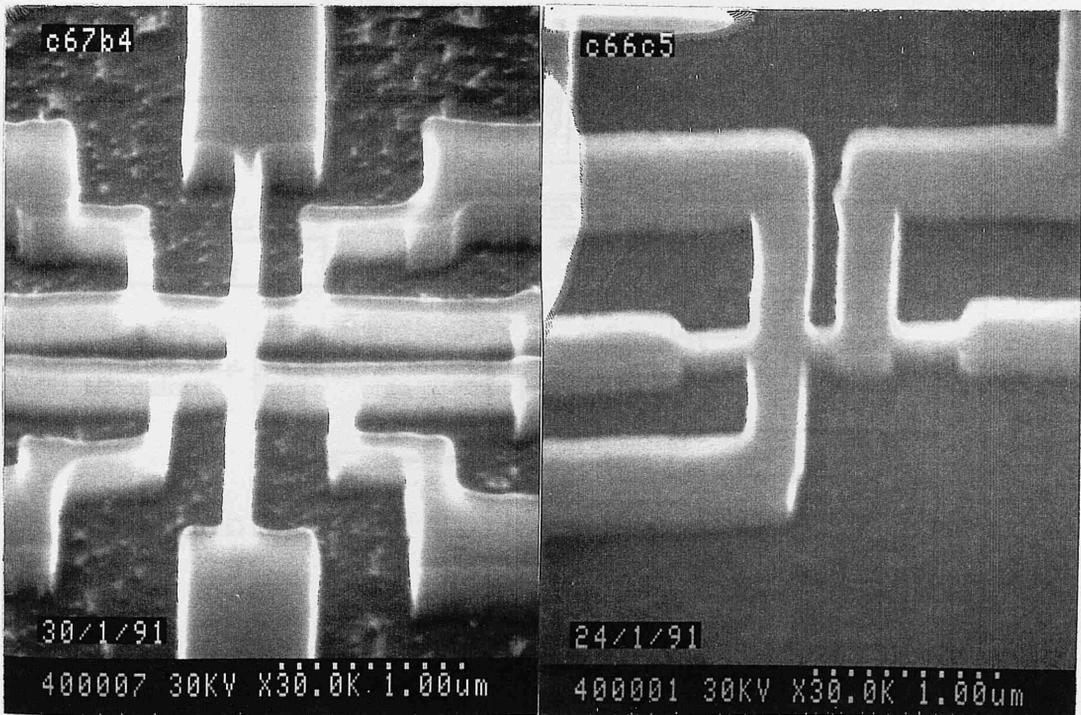


Figure 6.1: Typical structures fabricated for studying conductance fluctuations. Both are based on Hall geometries. See text for further details.

Several important parameters are given in Table 6.1 for 1-, 2- and 3-dimensional Fermi surfaces. For the material being used, electron diffusion occurs in 3 dimensions (see also §6.2.2), so only the last column of Table 6.1 is relevant. It is seen that $k_F l \gg 1$ irrespective of the dimensionality of any prospective structure, so the material satisfies the weak disorder condition required by the theory of Chapter 5. Generally for the wires studied in this section, the conductances are much greater than e^2/h , so the system is said to be metallic. [The material is also metallic in the sense that the electron gas is degenerate: because of screening the electrons do not freeze out, and the sample retains its Fermi surface even at the lowest temperatures.]

Fig. 6.2 shows the circuit used to perform the measurements. A -40 dB (divide by 100) attenuator is used to obtain low excitation voltages. One lock-in is used to monitor the voltage across a 560 k Ω test resistor placed in series with the sample. This allows the current through the sample to be known. Two further lock-ins are used to monitor voltages in the sample itself. The lock-ins are read at regular intervals by a computer, which stores the data. The computer is also connected to the magnet power supply (not shown), which allows magnetic field readings to be taken.

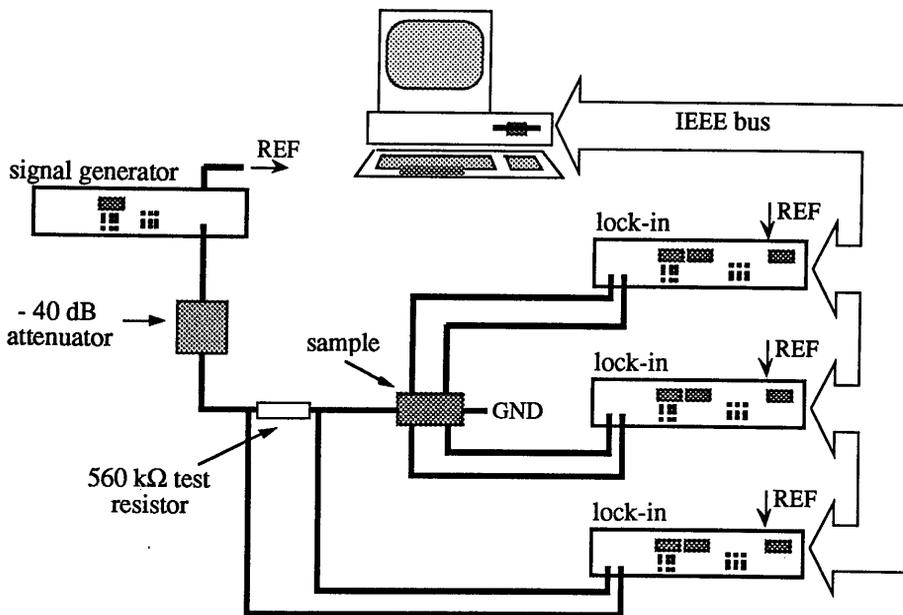


Figure 6.2: Configuration used to measure sample conductances. A low frequency ac excitation is used. The test resistor allows the current through the sample to be determined.

	1-dimensional	2-dimensional	3-dimensional
carrier density	$N_1 = N_D w_1 t_1 = 4.4 \times 10^{10} \text{ m}^{-1}$	$N_2 = N_D t_1 = 2.4 \times 10^{17} \text{ m}^{-2}$	$N_3 = N_D = 6.5 \times 10^{24} \text{ m}^{-3}$
Fermi wavevector	$k_1 = \frac{\pi N_1}{2} = 6.9 \times 10^{10} \text{ m}^{-1}$	$k_2 = (2\pi N_2)^{1/2} = 1.2 \times 10^9 \text{ m}^{-1}$	$k_3 = (3\pi^2 N_3)^{1/3} = 5.8 \times 10^8 \text{ m}^{-1}$
Fermi wavelength	$\lambda_1 = \frac{2\pi}{k_1} = 0.091 \text{ nm}$	$\lambda_2 = \frac{2\pi}{k_2} = 5.2 \text{ nm}$	$\lambda_3 = \frac{2\pi}{k_3} = 10.8 \text{ nm}$
Fermi velocity	$v_1 = \frac{\hbar k_1}{m^*} = 1.2 \times 10^8 \text{ m s}^{-1}$	$v_2 = \frac{\hbar k_2}{m^*} = 2.1 \times 10^6 \text{ m s}^{-1}$	$v_3 = \frac{\hbar k_3}{m^*} = 1.0 \times 10^6 \text{ m s}^{-1}$
elastic mean free path	$l_1 = v_1 \tau = 6.4 \text{ } \mu\text{m}$	$l_2 = v_2 \tau = 111 \text{ nm}$	$l_3 = v_3 \tau = 53 \text{ nm}$
diffusion coefficient	$D_1 = v_1 l_1 = 838 \text{ m}^2 \text{ s}^{-1}$	$D_2 = \frac{1}{2} v_2 l_2 = 0.12 \text{ m}^2 \text{ s}^{-1}$	$D_3 = \frac{1}{3} v_3 l_3 = 0.017 \text{ m}^2 \text{ s}^{-1}$
thermal diffusion length (4 K)	$L_{T1} = \left(\frac{\hbar D_1}{k_B T} \right)^{1/2} = 40 \text{ } \mu\text{m}$	$L_{T2} = \left(\frac{\hbar D_2}{k_B T} \right)^{1/2} = 480 \text{ nm}$	$L_{T3} = \left(\frac{\hbar D_3}{k_B T} \right)^{1/2} = 182 \text{ nm}$
$k_F l$	$k_1 l_1 = 4.8 \times 10^5$	$k_2 l_2 = 140$	$k_3 l_3 = 30$

Table 6.1: Material parameters for 50 nm n^+ -GaAs epilayer with carrier concentration $N_D = 6.5 \times 10^{24} \text{ m}^{-3}$ and mobility $0.14 \text{ m}^2 \text{ V}^{-1} \text{ s}^{-1}$ assuming a 1-, 2- or 3-dimensional Fermi surface. The conducting thickness is $t_1 = 37.5 \text{ nm}$, and the elastic mean free time is $\tau = 5.3 \times 10^{-14} \text{ s}$. [For the 1D entries, the conducting width is taken as $w_1 = 180 \text{ nm}$.]

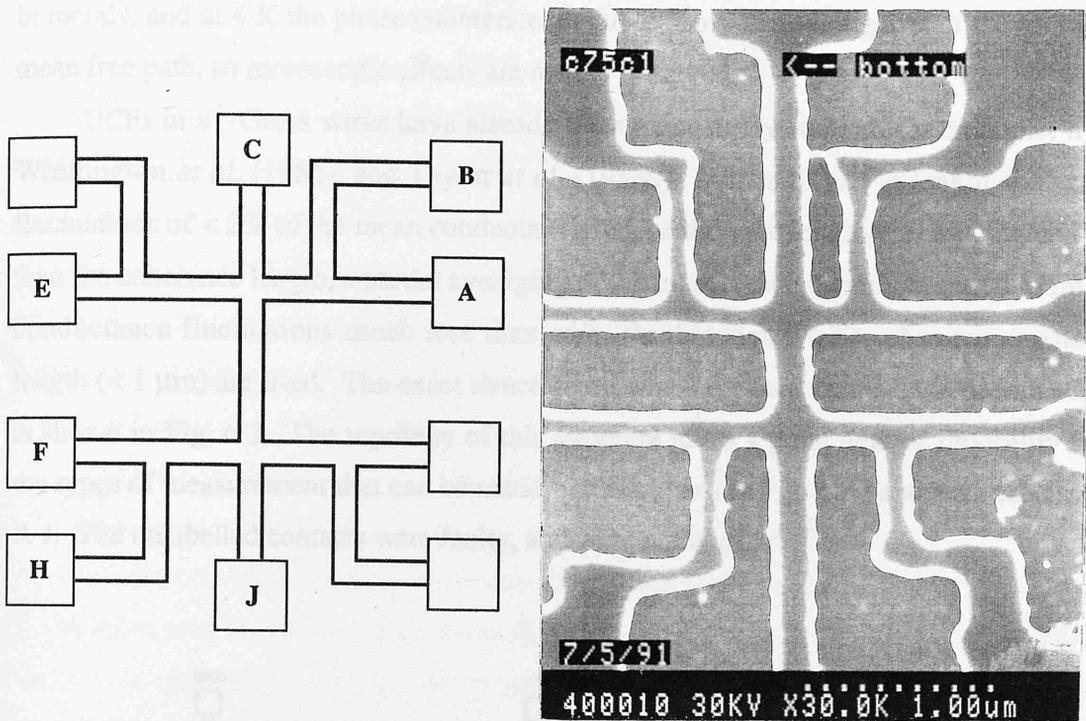


Figure 6.3: The device used in the measurements of conductance fluctuations. The working contacts are labelled A-J. The topology of the structure allows a variety of measurements to be made using just the one sample.

6.2.2 Fluctuations in the longitudinal measurement

The origins of universal conductance fluctuations (UCFs) were explained in §5.3.3. In metals, UCFs are observed at milliKelvin temperatures, since only then is the coherence length significant compared to the sample size (the elastic mean free path is very short in metals – of the order of the interatomic spacing). In addition to this, the very high conductivity of metals means that e^2/h fluctuations tend to be a very small fraction of the mean conductance ($< 0.5\%$). On the other hand, in high mobility materials (*e.g.* high purity GaAs/AlGaAs heterostructures) the elastic mean free path can be comparable to the coherence length, and electron transport in these systems is not strictly quantum diffusive. The theory of Chapter 5 must then be modified [see, for

example, Beenakker and van Houten (1988), or Bird *et al.* (1990)]. The n^+ -GaAs wires of Chapter 3 lie between these two limits. The conductivity is not as high as it is in metals, and at 4 K the phase coherence length is many times longer than the elastic mean free path, so mesoscopic effects are readily observed.

UCFs in n^+ -GaAs wires have already been extensively studied, for example, by Whittington *et al.* (1986), and Taylor *et al.* (1988a). Wires 10 μm in length showed fluctuations of $< 2\%$ of the mean conductance. Since the wire length was much greater than the coherence length, a partial averaging of the interference patterns resulted in rms conductance fluctuations much less than e^2/h . In this thesis wires of much smaller length ($< 1 \mu\text{m}$) are used. The exact structure on which measurements have been made is shown in Fig. 6.3. The topology of this structure gives a great deal of flexibility in the types of measurement that can be made. The contacts in Fig. 6.3 have been labelled A-J. The unlabelled contacts were faulty, and were not used.

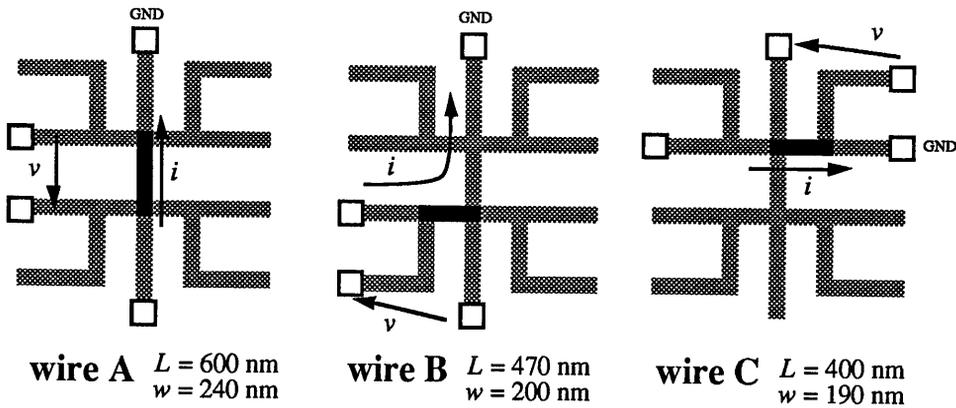


Figure 6.4: Three different wire segments across which 4-terminal measurements of conductance fluctuations are made. The lengths and widths are estimates from Fig. 6.3.

The standard 4-terminal conductance measurement discussed in Chapter 2 is used to study the UCFs in this structure. The conductances of three different wire segments can be measured, as shown in Fig. 6.4. These have lengths (A) 600 nm, (B) 470 nm, and (C) 400 nm, where the lengths are measured between the centres of the voltage probes. The respective *physical* widths are roughly (A) 240 nm, (B) 200 nm, and (C) 190 nm. It must be remembered that the *conducting* width is roughly 60 nm less than

the physical width due to sidewall depletion (see Chapter 3).

Fig. 6.5 shows UCF traces measured for the three wire segments A, B, and C, with the sample being at a temperature of 4 K. The magnetic field was varied from 0 to 3 T. Two traces are shown for wire A. These curves look almost like noise, except for one important point – the two traces for wire A were taken at different times (roughly 40 min apart), and yet are more or less identical. Thus, although appearing like random noise, the complex arrangement of peaks and troughs are, in fact, perfectly reproducible allowing for noise in the measuring circuit. Many similar such measurements, on this and other samples, show that these curves are reproducible only if the sample is maintained at a low temperature. Cycling the sample to room temperature and back to 4 K always produces a different trace.

Taylor *et al.* (1988b) have done a careful study of the effects of reproducibility in n^+ -GaAs wires. They find that the curve (also called a magnetofingerprint) remains reproducible provided the sample is maintained at temperatures of much less than ~ 100 K. In these studies, the authors argued that charge trapped in the resonant DX -level in GaAs was responsible for the observed effects. In general, however, any changes associated with *any* defect present in the sample would lead to the same effects. Thus, redistribution of charges trapped in defects introduced by dry-etching, or indeed diffusion of these defects, cannot be precluded. By the same token, the presence of mobile defects inherent in the sample (interstitial impurities, *etc.*) will also contribute to changes in the UCF trace. In addition to thermal cycling, redistribution of trapped charge can also occur through the application of light pulses [*e.g.* Davison *et al.* (1990)], or voltage spikes [*e.g.* Mailly *et al.* (1989)].

The root mean square fluctuation, δG_{rms} , in the conductances of the curves shown in Fig. 6.5 are roughly $0.42 e^2/h$, $0.60 e^2/h$ and $0.82 e^2/h$ for the 600 nm, 470 nm and 400 nm long wires respectively. The fluctuations are calculated after subtracting off a linear background rise in the conductance with field. This linear rise remains at high temperatures, and is due to the *classical* Hall effect manifesting itself in the longitudinal measurement. See also §6.3.2. The change in δG_{rms} with length would suggest that averaging or scaling of the fluctuations is taking place with length. A log-log plot of wire length, L , versus δG_{rms} is shown in Fig. 6.6. A least squares fit gives $\delta G_{rms} \propto L^{-1.66}$. This is very close to the $L^{-3/2}$ relation expected from a chain of 1-dimensional (1D) phase coherent regions [see Eq. (5.41)]. The curve is not exactly $L^{-3/2}$ because it is actually an admixture of L^{-2} as well as $L^{-3/2}$ – see also §6.2.3.

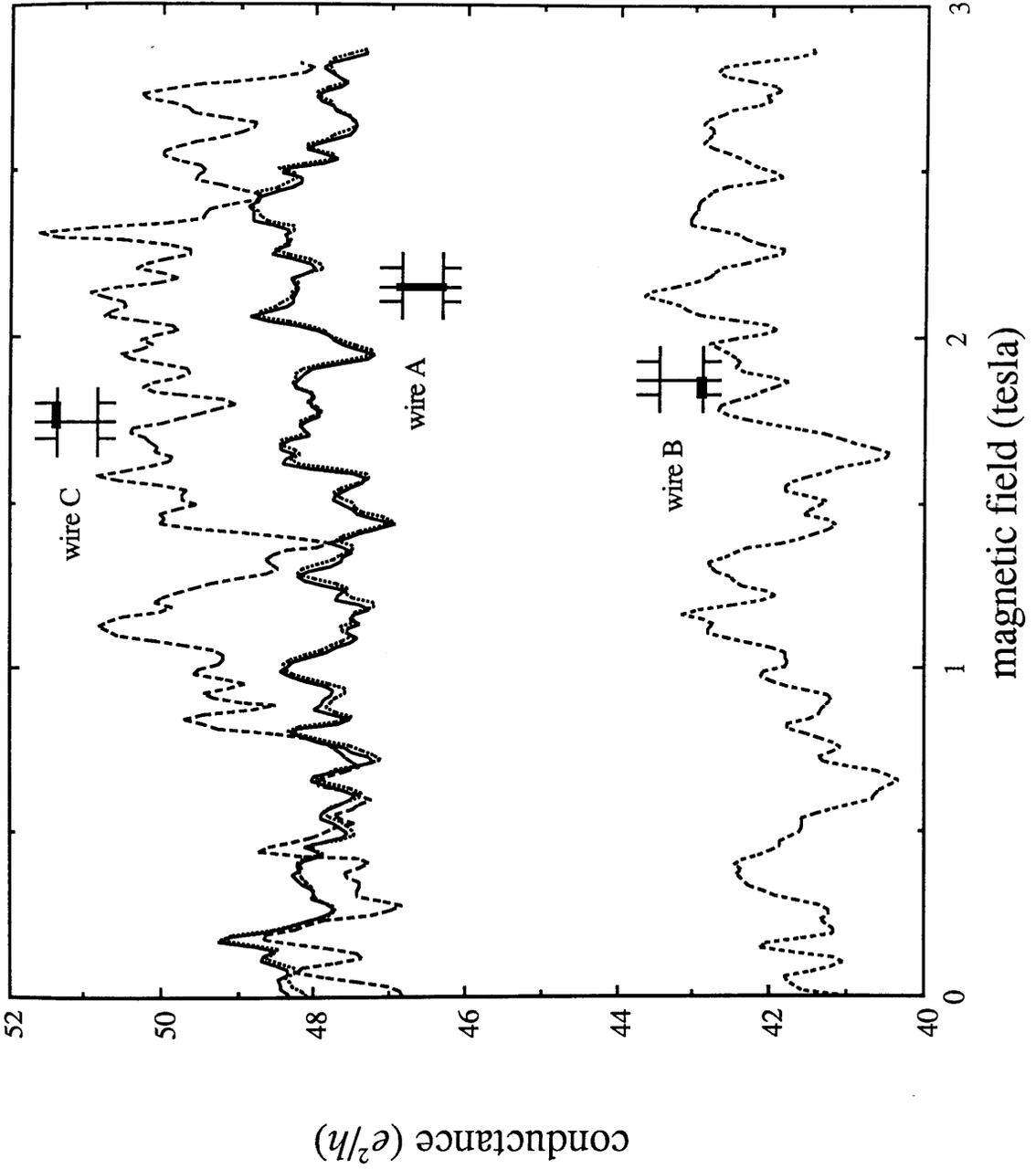


Figure 6.5: Measured conductance fluctuations of wires A, B and C in units of e^2/h . The two measurements on wire A show the reproducibility. The weak localisation in all these curves is largely buried in the fluctuations.

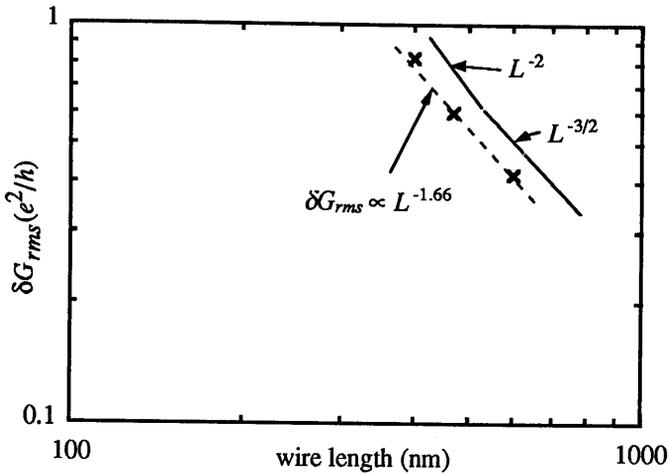


Figure 6.6: Variation in measured rms fluctuation (in units of e^2/h) with wire length. The dotted line is a least squares fit. The wires appear to behave 1-dimensionally with respect to fluctuation effects.

For a 1D system, Eq. (5.35) shows that an rms conductance fluctuation of $0.52 e^2/h$ is expected. The values for wires B and C are, in fact, larger than this. Why this can happen is discussed in §6.2.3, where estimates will be made of L_ϕ in these wires. For wire A, the rms fluctuation is less than $0.52 e^2/h$. For 1D averaging and $L_\phi < L_T$, Eq. (5.40) gives,

$$L_\phi = \left(\frac{1}{0.52} \frac{\delta G_{rms}}{e^2/h} \right)^{2/3} L. \quad (6.1)$$

Thus for wire A, $L_\phi \approx 520$ nm. A further estimate of L_ϕ will be made in §6.2.4 using the correlation function of Lee and Stone (1985). The weak localisation effect in the measurements of Fig. 6.5, although present, are strongly affected by fluctuations. This cannot therefore be used to give yet another estimate of L_ϕ .

The "1D" above does not necessarily refer to the actual dimensionality seen by the electrons (*i.e.* the dimensionality of the Fermi surface). Main *et al.* (1990) measured the dependence of the UCF curve on the orientation of n^+ -GaAs wires with respect to the direction of the applied magnetic field. The conducting core of the wire was roughly 35×35 nm, and UCFs were observed for all orientations. This implies that the electron trajectories within the wire are 3-dimensional, not 1-dimensional. In the present case, the conducting core is of a roughly similar size, so the trajectories here should be 3-dimensional also. Thus, the electrons behave 3-dimensionally with respect to diffusion,

but the wire is 1-dimensional with respect to the fluctuation effects. The 3D nature of the Fermi surface is also inferred from the fact that there are several electrically quantised subbands [with transverse energies given by $E_{mn} \sim \hbar^2(k_m^2 + k_n^2)/2m^*$] occupied within this wire.

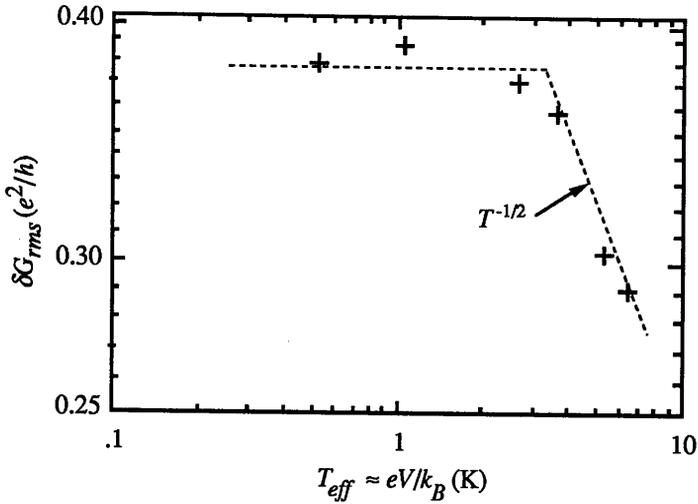


Figure 6.7: Effect of increasing the excitation voltage. For $T_{eff} > 4$ K, the fluctuation amplitudes are reduced.

The value of phase coherence length, L_ϕ , obtained above is of the same order of magnitude as the thermal diffusion length for electrons diffusing in 3-dimensions, $L_T \sim 180$ nm, obtained from Table 6.1. However, the large fluctuations in wires B and C imply that $L_\phi, L_T > 470$ nm (see §6.2.3), so it is possible that the value of L_T in Table 6.1 is slightly underestimated. An estimate of L_T can be made from Fig. 6.7 which shows the effect of Joule heating on the rms fluctuations of wire A. The rms conductance fluctuation in wire A was measured for different excitation voltages. A voltage V across the sample corresponds to a temperature $T_{eff} = eV/k_B$. The figure shows that for $T_{eff} > 4$ K (the lattice temperature), the fluctuation amplitudes are reduced. This can happen because of changes in either L_T or L_ϕ (the latter through the temperature dependence of electron-electron scattering).

Weak localisation measurements of L_ϕ in §6.2.6 (albeit for a different wire), suggest that L_ϕ may be almost temperature independent near 4 K in these dry-etched

wires. In this case, the decrease in δG_{rms} seen in Fig. 6.7 is primarily due to L_T . From Eq. (5.41a), $\delta G_{rms} \propto L_T \propto T^{-1/2}$ for $L_T < L_\phi$ and a temperature independent L_ϕ . Fitting to Fig. 6.7 gives $L_T \sim 420$ nm at 4 K. This value is slightly larger¹ than that in Table 6.1. It seems possible that in wires A, B and C the magnitudes of the UCFs is being limited not by L_ϕ , but by L_T . However, the important conclusion from this section and the next is that $L_\phi \gtrsim 500$ nm for wires A, B and C.

6.2.3 Role of voltage probes in the quantum conductance measurement

The above estimate for L_ϕ in wire A is larger than the probe-to-probe spacings, L , for both wires B and C. Indeed, the large fluctuations measured for these latter wires is also indicative of $L_\phi > L$. Benoit *et al.* (1987), and Skoçpol *et al.* (1987) have shown that in 4-terminal measurements where L is decreased below L_ϕ , the measured conductance does *not* saturate at $\sim e^2/h$. Instead, it is the rms fluctuation in measured *voltage* that saturates, *i.e.* as the probe spacing goes to zero, the measured voltage *does not go to zero*, but saturates at the value it attained for $L \sim L_\phi$. Denote $\delta V_\phi = IR_\phi^2 \delta G_\phi$ as the rms voltage fluctuation when $L \sim L_\phi$, where I is the current, $R_\phi = R_s L/w_1$ is the resistance of the phase coherent region, $\delta G_\phi = e^2/h$ is its rms conductance fluctuation, R_s is the average sheet resistance of the material, and w_1 is the conducting width of the wire. Then for $L < L_\phi$, the following relation is found to hold experimentally:

$$\delta G_{rms} = \frac{I}{V^2} \delta V_\phi \sim \frac{e^2}{h} \left(\frac{L_\phi}{L} \right)^2, \quad (6.2)$$

where δG_{rms} is the measured rms conductance fluctuation. Thus, instead of the dependence $\delta G_{rms} \propto L^{-3/2}$ as it was for $L < L_\phi$ [Eq. (5.40)], the dependence becomes $\delta G_{rms} \propto L^{-2}$ for $L > L_\phi$. So, the fluctuation amplitude of the conductance actually diverges as the probe spacing is made smaller. This explains why $\delta G_{rms} > 0.52 e^2/h$ for wires B and C.

The cause of this phenomenon is the 4-terminal nature of the conductance measurement [Maekawa *et al.* (1987); Baranger *et al.* (1988); Kane *et al.* (1988)]. This is illustrated in Fig. 6.8. The theory of Chapter 5 was concerned with the *two-terminal* conductance, where the current leads were being used as the voltage probes. Thus, the voltage "probes" are effectively always connected to a sample of size $L \approx L_\phi$, as shown in Fig. 6.8(a), and a "universal" fluctuation of $\delta G_{rms} \sim e^2/h$ is always

¹ Note that there are different conventions for defining L_T (*e.g.* bar on h or not). Values of L_T should therefore generally be treated as estimates.

measured for the phase coherent region. This is fine when comparing with experiments where the voltage probes are spaced further apart than the coherence length, but not so otherwise.

When the probes are within a coherence length of one another, then the phase coherent region can extend into the voltage probes as shown in Fig. 6.8(b). The voltage probes then no longer measure the potential at the point at which they are connected to the wire, but a different potential as determined by the electron interference. Theoretical analysis shows that in this case the rms fluctuation in the measured voltage should become independent of the probe-to-probe spacing. This is an interesting demonstration of the nonlocal nature of the quantum conductivity [as defined by Maekawa *et al.* (1987)].

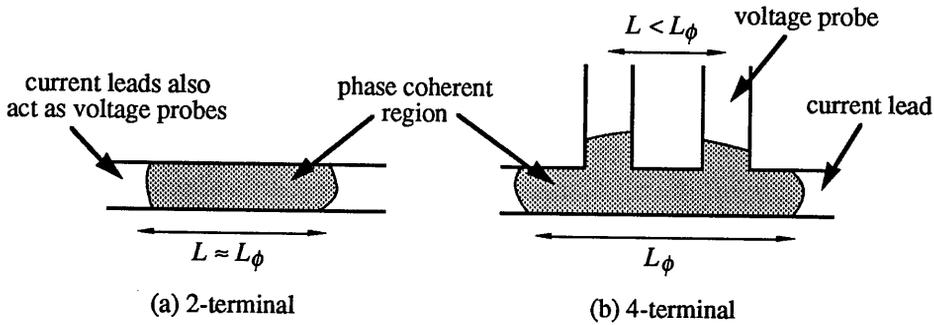


Figure 6.8: Difference between 2-terminal and 4-terminal quantum conductance measurements. In the 4-terminal case, the voltage probes will not necessarily measure the potential at the point of contact with the main wire.

Using Eq. (6.2), estimates can be made of the coherence lengths in wires B and C. Since the rms conductance fluctuation in a 1D wire is $0.52 e^2/h$, Eq. (6.2) gives,

$$L_\phi = \left(\frac{1}{0.52} \frac{\delta G_{rms}}{e^2/h} \right)^{1/2} L . \quad (6.3)$$

With values of L of 470 nm and 400 nm, wires B and C give coherence lengths of 505 nm and 502 nm respectively. These are similar to the values obtained for wire A which is roughly 40 nm wider. As noted before, these values may in fact be influenced by L_T . However, the important point is that $L_\phi \gtrsim 500$ nm.

On the question of the role of the voltage probes in the quantum conductance measurement, a relevant measurement may be made using the structure of Fig. 6.3. The

longitudinal conductance for wire A ($L \approx 600$ nm) above was measured using contacts E and F as the voltage probes. Instead, it is also possible to use contacts E and H. Fig. 6.8 shows two traces of the conductance of wire A, one obtained using voltage probes E-F, the other using voltage probes E-H. It is clear that although the two traces are very similar, the correlation is perhaps not as good as that between the two traces shown in Fig. 6.5. This indicates that there are a small number of phase coherent electron trajectories extending at least 400-500 nm down these probes from the main wire. Clearly, the correlation between the two traces of Fig. 6.5 would be expected to decrease as the probe H is moved closer to the main body of the wire. The measurements in Fig. 6.9, although not very significant, do indicate a small effect. [Note that the UCFs of Fig. 6.5 and 6.9 are different because the sample was cycled up to room temperature and back down to 4 K (the mean conductances are also slightly different).]

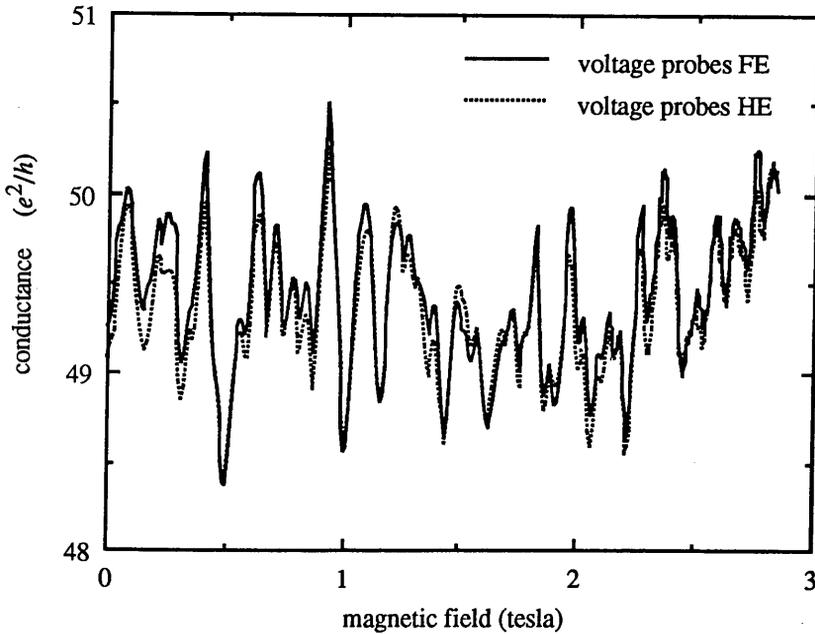


Figure 6.9: UCF traces obtained for wire A using two different voltage probe configurations. The correlation is good, but not as good as for the two measurements using the same set of probes, as shown in Fig. 6.5.

6.2.4 Use of the autocorrelation function

The correlation function, $F(\Delta E, \Delta B)$, introduced by Lee and Stone (1985), was shown in §5.3.3 to be useful in extracting information from UCF curves. Eq. (5.32) defined the correlation function as an ensemble average over all impurity configurations. For the experimental curves of this chapter, the correlation function must be written as an average over all values of magnetic field. In discrete form this is,

$$F_i = \frac{1}{Z} \frac{1}{N} \sum_{j=0}^{N-i} [G_j - \bar{G}] \times [G_{j+i} - \bar{G}] , \quad i = 1, \dots, N, \quad (6.4)$$

where Z is a normalisation constant, and \bar{G} is the mean conductance averaged over all values of field:

$$\bar{G} = \frac{1}{N} \sum_{i=1}^N G_i . \quad (6.5)$$

In the above, N is the number of points measured, G_j is the measured conductance at field B_j , and F_i is the value of the correlation function for the field difference $\Delta B = B_{j+i} - B_j$ (where the B_j 's are assumed to be evenly spaced), *i.e.* F_i is equivalent to $F(\Delta B)$. By the ergodic hypothesis, the correlation function defined in Eq. (6.4) for $Z = 1$ is entirely equivalent to the definition in Eq. (5.32). For the purposes of this chapter, the correlation function will be normalised so that $F(0) = 1$, *i.e.* $Z = (\delta G_{rms})^2$. With this definition, $F(\Delta B)$ is just the autocorrelation function for the UCF curve.

Several precautions are required when taking autocorrelations of experimental data such as those in Fig. 6.5. The UCFs typically consist of "rapid" fluctuations superimposed on much slower changes (in Fig. 6.5, for example, there are slow background rises in the conductance as a function of field). By definition, the autocorrelation function is related to the Fourier power spectrum of the data by,

$$F(\Delta B) = \int \delta G(B+\Delta B) \delta G(B) dB = \mathfrak{F}^{-1} \{ |\delta \tilde{G}|^2 \}, \quad (6.6)$$

where $\delta G = G - \bar{G}$ denotes the deviation from the mean conductance, $\delta \tilde{G}$ is the Fourier transform of δG , and \mathfrak{F}^{-1} is the inverse Fourier transform operator. [The above follows simply from the convolution theorem.] Clearly, any systematic trends in δG (*e.g.* slowly rising/falling background) will contribute significantly to lower frequency Fourier amplitudes. However, these only have meaning when the data includes several periods of this lower frequency background. In the case at hand, the UCFs measured in Fig. 6.5 and in later sections sometimes show a slow linear background rise in conductance with field (due to the Hall effect). To find the correlation between the

fluctuations, this background rise must be subtracted off. If this is not done, then the autocorrelation function will simply reflect the fact that there is a slow background rise, and give little information besides.

By the same token, autocorrelations cannot be taken over very small field ranges, where there is too little information. Suitable criteria for judging what field range to take are that it should be several times larger than B_C , the correlation field; that the range should not include "slow" background trends unless these are the objects of study; and that varying the field range slightly does not affect the resulting value of B_C .

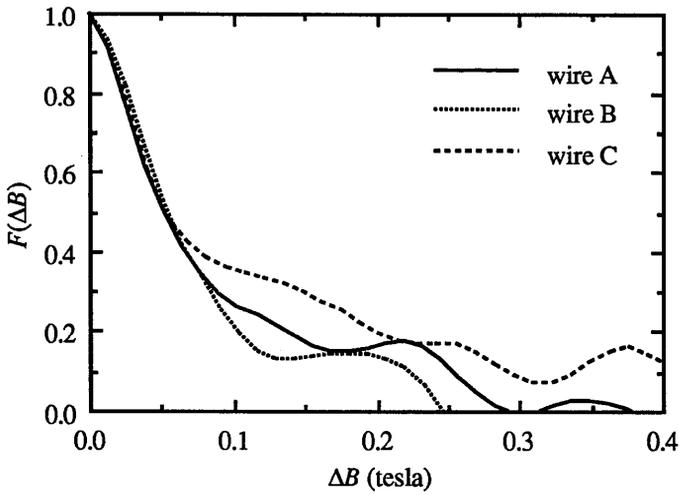


Figure 6.10: Autocorrelation functions for the UCF traces of Fig. 6.5. These give roughly similar correlation fields for the wires A, B and C.

			wire A	wire B	wire C
probe spacing,	L	(nm)	600	470	400
physical width,	w	(nm)	240	200	190
conducting width,	$w_1 \approx (w-60)$	(nm)	180	140	130
correlation field,	B_C	(T)	0.050	0.054	0.050
coherence length,	$L_\phi \approx h/eB_C w_1$	(nm)	460	547	637
estimate of L_ϕ from δG_{rms}		(nm)	520	505	502

Table 6.2: Estimates of phase coherence length made for wires A, B and C using the autocorrelation function defined in the text. Note that $L < L_\phi$ for wires B and C.

The autocorrelation functions of the curves shown in Fig. 6.5 were calculated over the full range of field 0-3 T, after subtracting off the linear background rise. The resulting autocorrelation functions are shown in Fig. 6.10. The correlation field is defined by the condition $F(B_C) = \frac{1}{2}$, from which estimates of the coherence length for each wire may be found using Eq. (5.38):

$$L_\phi \sim \frac{h}{e} \frac{1}{B_C w_1} . \quad (6.7)$$

where w_1 is the *conducting* width of the wire, assumed to be 60 nm less than the physical width. The resulting values of L_ϕ are shown in Table 6.2.

Table 6.2 shows that the correlation field for each of the wires is roughly the same, corresponding to a phase coherent area $S_\phi = h/eB_C \approx 8 \times 10^{-14} \text{ m}^2$. Clearly, for wires B and C this phase coherent area "spills" into the voltage probes, since the resulting estimates of L_ϕ are greater than the probe spacing. Allowing for the errors this introduces into the estimate of L_ϕ , the values of L_ϕ obtained for all three wires are roughly in line with the magnitudes of the rms fluctuations observed in the previous two sections.

6.2.5 Boundary scattering and classical and quantum size effects

Scattering from the edges (or boundaries, or surfaces) of a sample are an important consideration in experiments with structures where two boundaries are within an elastic mean free path of one another. This leads to size effects, *i.e.* effects which occur purely because of the influences of the sample boundaries. Two limits of *elastic* boundary scattering exist, as shown in Fig. 6.11. In the diagram, the edge of the sample occurs where the Fermi level crosses the conduction band edge, creating a potential barrier to electron motion. If the boundary is perfectly smooth, the scattering is specular: the angle of incidence of the electron wave equals the angle of "reflection." If the boundary is perfectly rough, the scattering is diffuse: the electron wave is scattered with equal probability in all directions.² In the general case there will be a mixture of both specular and diffuse scattering.

Electron scattering from rough surfaces has been studied by several phenomenological approaches. Ziman (1960), Ch. 11, for example, considers the surface as having features of a certain average height, η , and width, Λ . As would be expected, when $\lambda_F \gg \eta, \Lambda$ (where λ_F is the Fermi wavelength) the electron does not "see" these

² In the diffuse reflection of *light*, an observer would see the same amount of reflected light independent of the angle of viewing.

features, and the scattering is specular. On the other hand, if $\lambda_F < \eta, \Lambda$ then the electron can resolve the surface roughness, and the scattering becomes more diffuse.³ A particularly simple way of characterising surface scattering involves the single parameter p , called the specularity coefficient or Fuchs parameter, which gives the probability of an electron being scattered specularly from a surface [Fuchs (1938); Ziman (1960), Ch. 11]. Thus, $p = 1$ corresponds to perfectly specular scattering, while $p = 0$ corresponds to perfectly diffuse scattering.

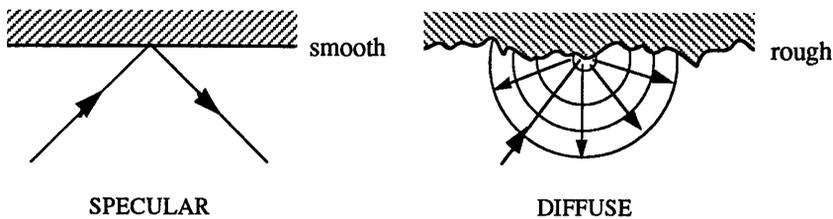


Figure 6.11: Illustrations of specular and diffuse boundary scattering. Both forms of scattering are elastic.

In general, there would appear to be no wide consensus as to which method of fabrication gives what type of boundary scattering. The electron focussing experiments of van Houten *et al.* (1989) show that electron scattering from the potential underneath a Schottky gate in a high mobility 2DEG structure is predominantly specular ($p \sim 1$). In electron focussing experiments where the boundary was defined by ion implantation or by wet-etching, the scattering was found to be more diffuse, with values of p down to about 0.35 [Nakamura *et al.* (1990); Nihey *et al.* (1990)]. Analysis of magneto-resistance curves for GaAs/AlGaAs wires created by removing a very small amount of material by dry-etching, and for wires created by ion implantation, appeared to show that here the scattering was predominantly specular [van Houten *et al.* (1986); van Houten *et al.* (1988); Hiramoto *et al.* (1989)], although it must be stressed that the theory being applied was only developed for the two limits $p = 0$ and $p = 1$, so intermediate

³ The whole argument about specular or diffuse scattering surfaces is just a restatement of something familiar from optics, microscopy, nuclear and particle physics, *etc.*, that to resolve finer detail in an object requires a probe with a shorter wavelength.

values of p could not have been distinguished. However, wires created by ion-beam damage appear to show more diffuse scattering [Thornton *et al.* (1990)].

How boundary scattering affects such things as mobility, resistivity, *etc.* depends on whether it is specular or diffuse. This is illustrated in Fig. 6.12 which shows a typical electron trajectory undergoing either specular or diffuse collisions with the boundary. In both cases, interactions within the sample are assumed to be negligible, so only boundary scattering occurs. In specular scattering, the forward component of the electron momentum is always preserved. Since the forward motion is unaffected by the presence of the boundaries, so the mobility, diffusion coefficient, *etc.*, will be unaffected by specular scattering. However, in the diffuse scattering case, there is a good chance of the electron trajectory being backscattered in the direction from which it came. In this case, the mobility and diffusion coefficient will both decrease from their values in the bulk (far from any boundaries). These will have corresponding effects on the resistivity, *etc.*

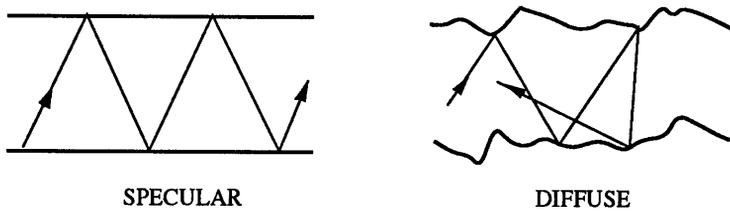


Figure 6.12: In specular scattering, the forward component of the momentum is unaffected by the scattering, so the mobility is unchanged by the presence of the boundary. In the diffuse case, the electron can be backscattered, so the mobility drops.

The classical effects of boundary scattering (*i.e.* without consideration of interference between electron waves) have been extensively studied in metals for many decades now [see Pippard (1989), Ch. 6]. In "large" samples, the electrons suffer many more collisions with impurities in the bulk of the sample than they do with the surface. Hence, the main parameter affecting the resistivity of the sample is the elastic mean free path, l (here defined as the mean distance between backscattering events):

$$\rho_0 = \frac{\hbar k_F}{ne^2 l} . \quad (6.8)$$

This is just the inverse of the Drude conductivity, Eq. (5.12). As the sample is made smaller, the electrons interact more and more with the boundaries. If the scattering is

perfectly diffuse, then the effect of the boundary will be to reduce the effective mean free path, and hence increase the resistivity in the manner described previously. Invoking Matthiessen's rule, the resistivity can be written,

$$\rho_{eff} \approx \frac{\hbar k_F}{ne^2} \left(\frac{1}{l} + \frac{1}{l_b} \right), \quad (6.9)$$

where l_b is the mean distance travelled by the electron before it is backscattered (diffusely) by the boundary, called the boundary scattering length. The term within the parentheses in Eq. (6.9) just comes from the sum of two elastic scattering rates.

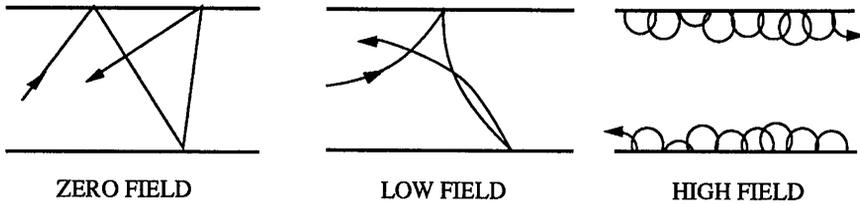


Figure 6.13: Classical size effect of an applied magnetic field on diffusely scattered trajectories. At intermediate fields, backscattered trajectories still exist. At high fields, backscattering is suppressed, and electrons far from the boundary are not influenced by its presence.

If w is the width of a wire, then the electron will typically make w/l_b specular collisions with the boundary (assuming $l \gg w$, and $l_b \gg w$) before a diffuse backscattering event occurs. In other words, there are $\sim w/l_b$ specular scattering events for each diffuse scattering event, so the probability that a particular scattering event will be diffuse is roughly $1/(1+w/l_b) \approx l_b/w$ (assuming $l_b \gg w$). In terms of the specular coefficient, the probability that a particular scattering event will be diffuse is $(1-p)$. Equating the two probabilities gives,

$$l_b = \frac{w}{1-p}, \quad (6.10)$$

This relation has been used to extract approximate values of p in 2DEG wires (see, for example, Thornton *et al.* (1990)). In metal wires, it is found that $l_b \approx w$, *i.e.* the scattering is perfectly diffuse ($p \sim 0$) [see Pippard (1989), Ch. 6].

The effect of the backscattered trajectories of Fig. 6.12 in increasing the resistivity of a sample, is reminiscent of the effect of coherent backscattering discussed in §5.3.2

(although it must be stressed that the former is a *classical* effect, the latter a quantum interference effect). Indeed, the analogy does not stop there. Fig. 6.13 shows how the classical backscattered trajectories can be affected by the application of a perpendicular magnetic field [Pippard (1989), Ch. 6; Thornton *et al.* (1990)]. In weak magnetic fields (when the classical cyclotron diameter is larger than the wire width, $2r_c > w$), an electron will tend to "skip" along one edge of the sample until it interacts with a bulk impurity which transfers it to the other edge. This also constitutes a backscattered trajectory. Whether the conductance rises or falls as a result, depends on whether the total number of backscattered trajectories is increased or decreased when the field is applied (which depends on the relative values of l and w). In strong magnetic fields (when $2r_c < w$), electrons near one edge of the sample will skip on paths closer to that edge. The probability of an electron being scattered from one edge of the sample to the other becomes very small. Electron motion away from the edges is then governed by the elastic mean free path (the boundaries have little effect), and the bulk value of diffusion coefficient is restored. Thus, as the increasing magnetic field suppresses the number of backscattered trajectories, a positive magnetoconductance is observed, very similar to that seen due to the suppression of weak localisation

In addition to the classical size effects discussed above, there are also quantum size effects. In §5.3.2 it was shown how pairs of closed time-reversed trajectories led to the phenomenon of weak localisation. Beenakker and van Houten (1988a) have shown that closed time-reversed paths which scatter elastically off sample boundaries, can modify the weak localisation theory. For 1-dimensional wires, the Al'tshuler-Aronov formula, Eq. (5.28), is modified by the effects of boundary scattering to give:

$$\sigma_{1D}(B,T) = \sigma_0(0,T) - \frac{e^2}{\pi\hbar} \left[\left(\frac{1}{L_\phi^2} + \frac{1}{D\tau_B} \right)^{-1/2} - \left(\frac{1}{L_\phi^2} + \frac{1}{D\tau_B} + \frac{1}{l^2} \right)^{-1/2} \right], \quad (6.11a)$$

where L_ϕ is the coherence length, l is the elastic mean free path, and,

$$\tau_B = \frac{C_1 L_B^4}{w^3 v_F} + \frac{C_2 L_B^2 \tau}{w^2}. \quad (6.11b)$$

Here $L_B = (\hbar/eB)^{1/2}$ is the magnetic length, w is the wire width, v_F is the Fermi velocity, and C_1 and C_2 are parameters dependent on the specularity of the boundary scattering. For diffuse scattering in a wire, $C_1 = 4\pi$ and $C_2 = 3$, and for specular scattering, $C_1 = 9.5$ and $C_2 = 4.8$. Note that putting $D\tau_B = 3(\hbar/weB)^2$ and taking the limit $l \ll L_\phi$ in Eq. (6.11a) recovers the Al'tshuler-Aronov formula.

Although Eqs. (5.28) and (6.11) can generally be fitted to the same positive

magnetoconductance curves (albeit with different parameter values), their realms of applicability are very different. The Al'tshuler-Aronov formula was derived strictly for the limit $l \ll w$, *i.e.* where the effects of the boundary are completely negligible. The Beenakker-van Houten formula above is slightly more general in that it can be applied to the limit $l \gg w$ also [a limit previously studied by Dugaev and Khmel'nitskii (1984) for perfectly diffuse scattering], and that the effects of specular or diffuse boundary scattering can be incorporated via the relation (6.11b). It is also useful in studying the case when $l \sim L_\phi$, as occurs in high mobility 2DEG structures [see van Houten (1988)]. In applying either formula, it is required that the sample be sufficiently long that the motion is diffusive along the length of the wire, not ballistic. Beenakker and van Houten (1988b) have shown that boundary scattering may also affect UCFs (although they only apply their theory to study a 2D channel).

6.2.6 Effects of boundary scattering in n^+ -wires

In analysing their UCF data, Taylor *et al.* (1988a) noted that some form of surface scattering appeared to be contributing to the phase breaking rate, τ_ϕ^{-1} , and that the effect was larger for a narrower wire. In earlier sections, wires of conducting width between 180 nm down to 130 nm appeared to give similar values of coherence length, $L_\phi \gtrsim 500$ nm, with no evidence for effects due to surface scattering. In this section, however, narrower wires are studied, and there *is* evidence here for a width-dependent influence on the coherence length.

Fig. 6.14 shows the narrowest wire studied. This structure is different from that shown in Fig. 6.3, being less flexible. The central wire in the structure is roughly 90 nm wide, with the distance between the centre of one probe to the centre of the other being 500 nm. The conducting width is roughly 30 nm (although there may be some uncertainty in this – see later). This will be called wire E. Another wire that is studied, wire D, has the same geometry as wire E, but with length 590 nm and physical width 160 nm. Wire D has a similar length to wire A, but has roughly half the *conducting* width (~ 100 nm as opposed to ~ 180 nm).

Fig. 6.15 shows a 4 K longitudinal measurement on wire D. Comparing this with the corresponding measurements of Fig. 6.5 reveals two immediate differences. Firstly, wire D shows a much larger weak localisation effect at zero field relative to the fluctuations, as compared to Fig. 6.5. Secondly, the rms conductance fluctuation is smaller, $\delta G_{rms} \approx 0.21 e^2/h$ (calculated for the range $B > 0.5$ T, after subtracting off

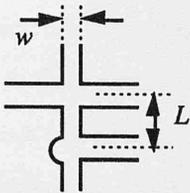
the linear background rise). Three separate estimates may be made for the coherence length in this wire.

A fit of the low field positive magnetoconductance gives several pieces of information. The correction terms in Eqs. (5.27) and (5.28) for 2-dimensional (2D) films and 1-dimensional (1D) wires may be written in terms of conductances as,

$$\Delta G_{2D}(B) = \frac{e^2}{2\pi^2\hbar} \left[\Psi \left(\frac{1}{2} + \frac{\hbar}{4eBl^2} \right) - \Psi \left(\frac{1}{2} + \frac{\hbar}{4eBL_\phi^2} \right) \right], \quad (6.12)$$

$$\Delta G_{1D}(B) = \frac{e^2}{\pi\hbar} \frac{1}{L} \left[\frac{1}{L_\phi^2} + \frac{1}{3} \left(\frac{w_1 e B}{\hbar} \right)^2 \right]^{-1/2}, \quad (6.13)$$

where L is the probe-probe spacing, w_1 is the conducting width, l is the elastic mean free path, Ψ is the digamma function, and L_ϕ is the "free" parameter to be adjusted to fit the experimental curves. A method of evaluating the digamma function is given in Appendix B.5. The asymptotic approximation for the first digamma function in Eq. (6.12) can only be used when $B \ll 0.06$ T, which is too small a field range in the present context.



wire	L (nm)	w (nm)
D	590	160
E	500	90

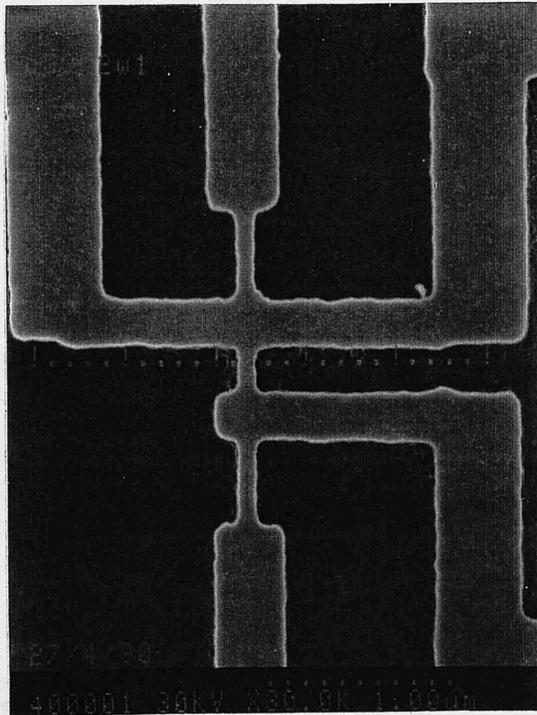


Figure 6.14: SEM of narrowest wire studied, wire E. Wire D has the same geometry, but with slightly larger dimensions.

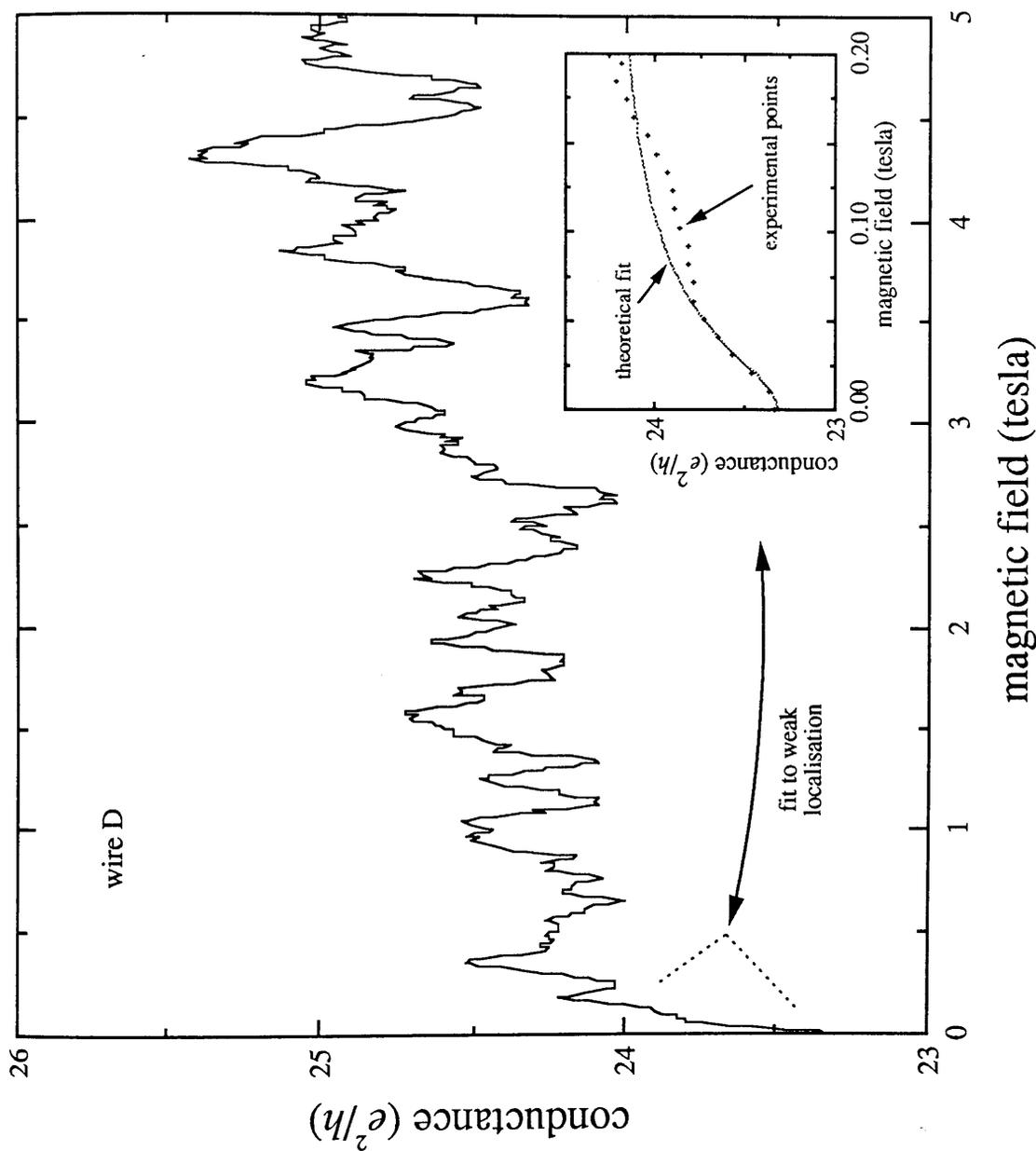


Figure 6.15: Measured conductance fluctuations of wire D in units of e^2/h . The rms conductance fluctuation is less than it was for wires A, B or C, so the weak localisation effect is more prominent. This weak localisation is fitted to the Al'tshuler-Aronov equation as shown in the inset. The parameters used were $L = 590$ nm, $w_1 = 100$ nm and $L_\phi = 300(\pm 10)$ nm.

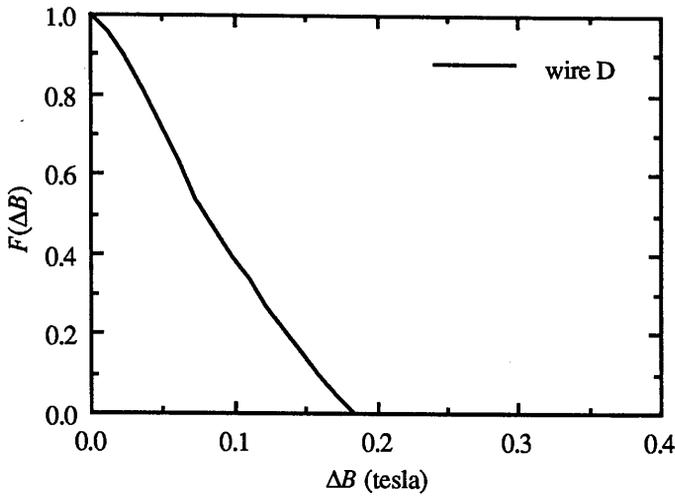


Figure 6.16: Autocorrelation function for UCF trace of wire D in Fig. 6.15.

The positive magnetoconductance for wire D cannot be fitted to the 2D formula. Not only is the shape of the calculated curve wrong, but the zero field conductance correction cannot be obtained for realistic values of coherence length.⁴ However, an excellent fit can be obtained using the 1D formula, as shown in the inset of Fig. 6.15, using the parameters $L = 590$ nm, $w_1 = 100$ nm, and $L_\phi = 300(\pm 10)$ nm. This gives an estimate for the coherence length in wire D, as well as showing that this wire is 1D with regard to the weak localisation effect. [Note that weak localisation measures L_ϕ and not L_T , since it is insensitive to energy averaging. Weak localisation relies on the phase of the electron wave being unchanged upon its return to the point of departure. This does not depend on what energy the electron is at, so energy averaging has no effect. See Lee *et al.* (1987).] There will be an error in this value of L_ϕ since fluctuations are expected to affect the weak localisation correction.

Since $\delta G_{rms} \approx 0.21 e^2/h$ for wire D is less than the $0.52 e^2/h$ expected for a 1D wire, averaging must be occurring. Assuming the averaging of a chain of

⁴ This is easily seen by writing Eq. (5.24b) in terms of conductance, and rearranging the correction part to give the coherence length:

$$L_\phi = l \exp\left(\frac{L}{w_1} \frac{\pi^2 \hbar}{e^2} \Delta G_{2D}\right).$$

Using $L = 590$ nm, $w_1 = 100$ nm, $l = 53$ nm, and $\Delta G_{2D} \approx 39 \mu S$ gives $L_\phi \approx 600$ nm which would imply $\delta G_{rms} \approx 5 \times 10^5 e^2/h$.

independently fluctuating phase coherent regions, gives $L_\phi \approx 322$ nm from Eq. (6.1). Finally, the autocorrelation function for the UCFs of wire D is calculated (for $B > 0.5$ T) after subtracting off the linear background trend, as shown in Fig. 6.16. The correlation field is $B_C = 0.080$ T, giving $L_\phi \approx 520$ nm from Eq. (6.7). This is slightly higher than the previous two estimates.

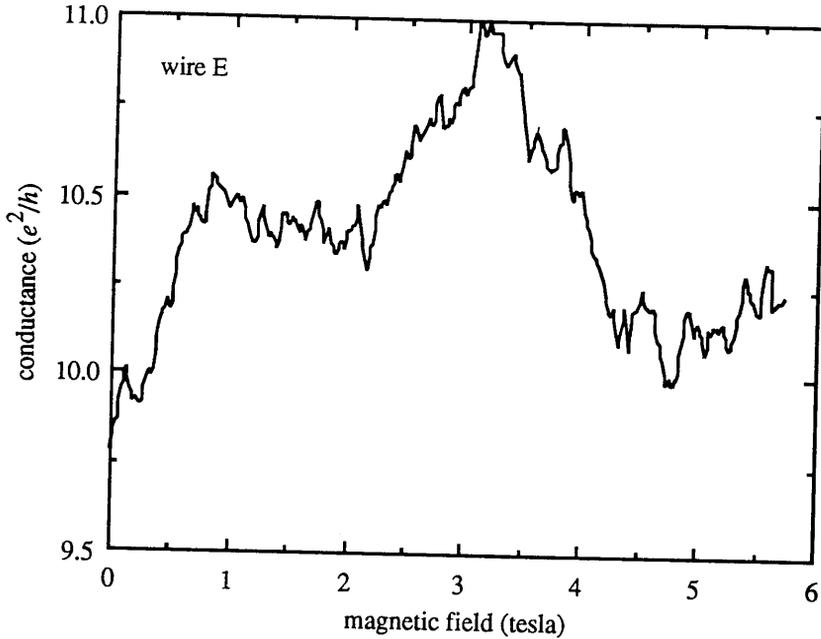


Figure 6.17: UCF trace for the narrowest wire, wire E. This shows "little" fluctuations, as well as much longer period fluctuations.

Fig. 6.17 shows a 4 K UCF trace obtained for wire E ($L = 500$ nm, $w_1 \sim 30$ nm). This curve is qualitatively even more different from the curves in Fig. 6.5 than was the case for wire D. The weak localisation is not suppressed till much higher fields, in addition to which the fluctuations are more complex than the ones observed for the wider wires. These UCFs have a high frequency component, as well as a much lower frequency component.⁵ The former "small" fluctuations will be studied later in §6.3.2 in the context of fluctuations in the Hall voltage. The UCF trace as a whole has an rms

⁵ Unfortunately in measurements on this device, the sample was being affected either by poor electrical screening or by stray light (the cryostat used in this instance had windows which were perhaps not adequately sealed), since the UCF trace would change completely after a period of 1-2 hours. However, two consecutive runs did give reproducible UCFs, although the correlation was poorer than that in measurements on other samples.

conductance fluctuation $\delta G_{rms} \approx 0.25 e^2/h$, giving $L_\phi \approx 307$ nm from Eq. (6.1). The form of the autocorrelation function for this trace, shown in Fig. 6.18, is primarily determined by the lower frequency fluctuations. The correlation field is $B_C \approx 0.53$ T, giving $L_\phi \approx 260$ nm.

The fit to the low field magnetoconductance for wire E is done for several temperatures. The measured curves are shown in Fig. 6.19. These curves cannot be fitted to the 2D formula, Eq. (6.11). With $L = 500$ nm, and conducting width $w_1 = 30$ nm, these curves cannot be fitted to the 1D formula, Eq. (6.12), either. Instead, it is found that very good fits *can* be achieved using $w_1 = 15$ nm, *i.e.* a higher field is required to suppress the localisation than would be expected for $w_1 = 30$ nm. Eq. (6.12) was derived by Al'tshuler and Aronov (1981a) with the assumption that elastic impurity scattering was the dominant diffusion mechanism in the wire. However, for wire E, $w_1 \sim 30$ nm, but the elastic scattering length is $l \approx 53$ nm. Since $l > w_1$, a certain amount of boundary scattering is inevitable, and the Al'tshuler-Aronov formula would not be expected to hold. Thus, boundary scattering must be affecting the field required to suppress the localisation.

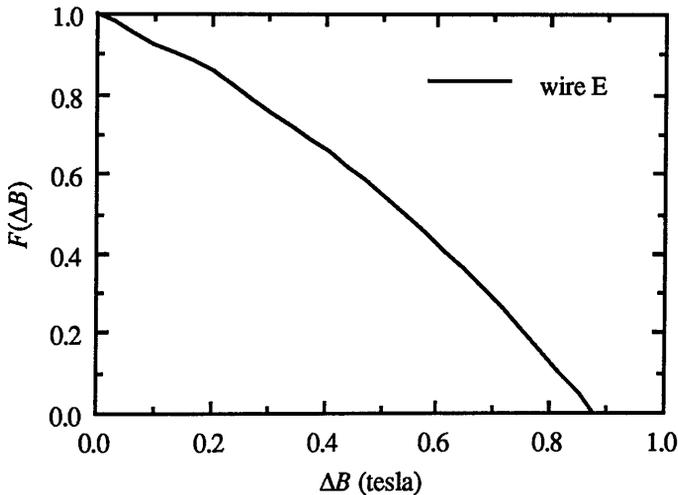


Figure 6.18: Autocorrelation function for the UCF trace of wire E. The shape of this is primarily determined by the "slow" fluctuations in the UCF trace shown in Fig. 6.17.

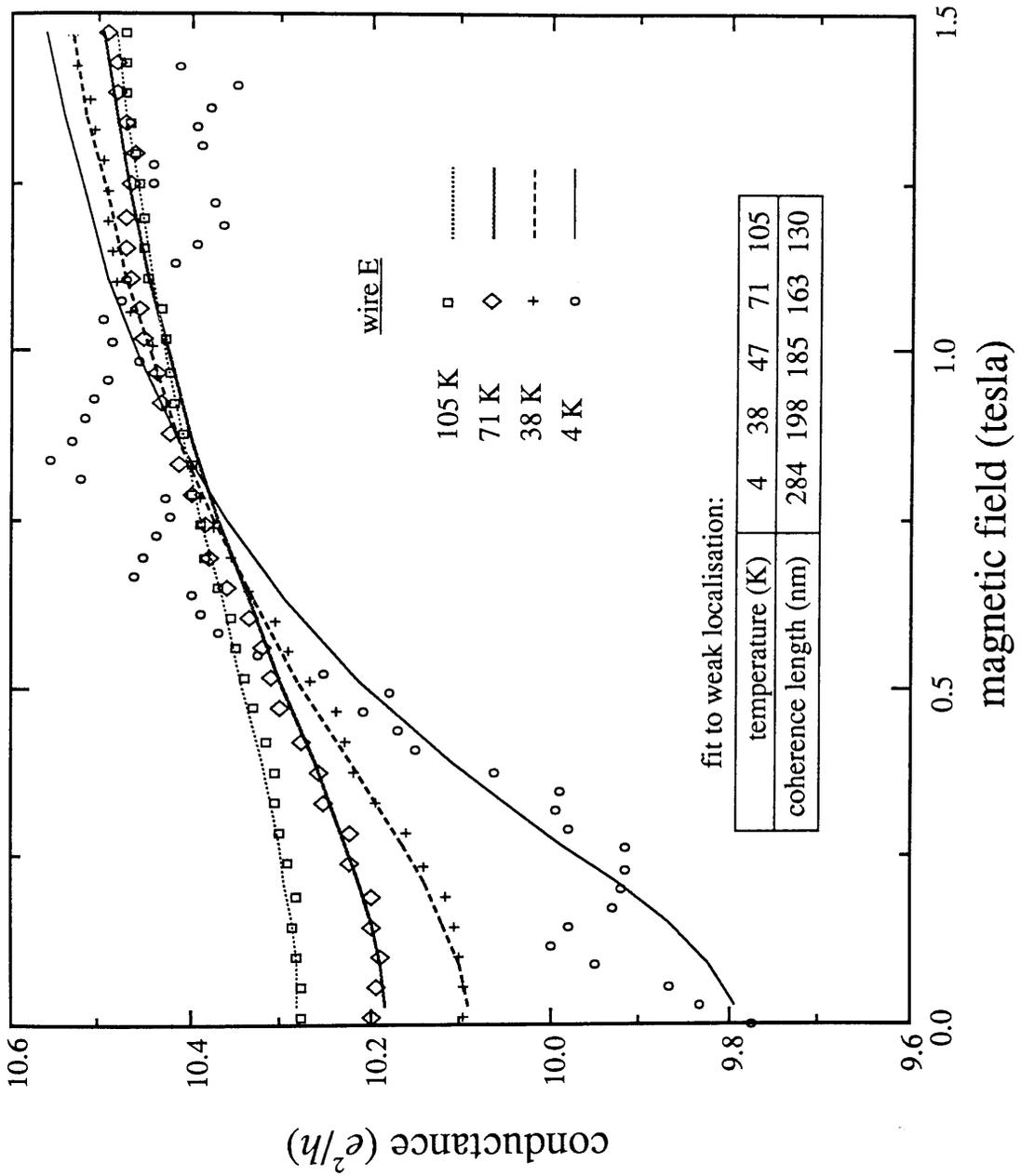


Figure 6.19: Measurements of weak localisation in wire E at several different temperatures. Wire E is expected to be affected by boundary scattering, so the experimental points are fitted to the Beenakker-van Houten equation. The 4 K curve is strongly affected by fluctuations. The measurements for the 47 K curve is not shown. The coherence lengths used to obtain the fit are given in the small table. The parameters used were $L = 500$ nm, $w_1 = 20$ nm, with the assumption of perfectly diffuse boundary scattering.

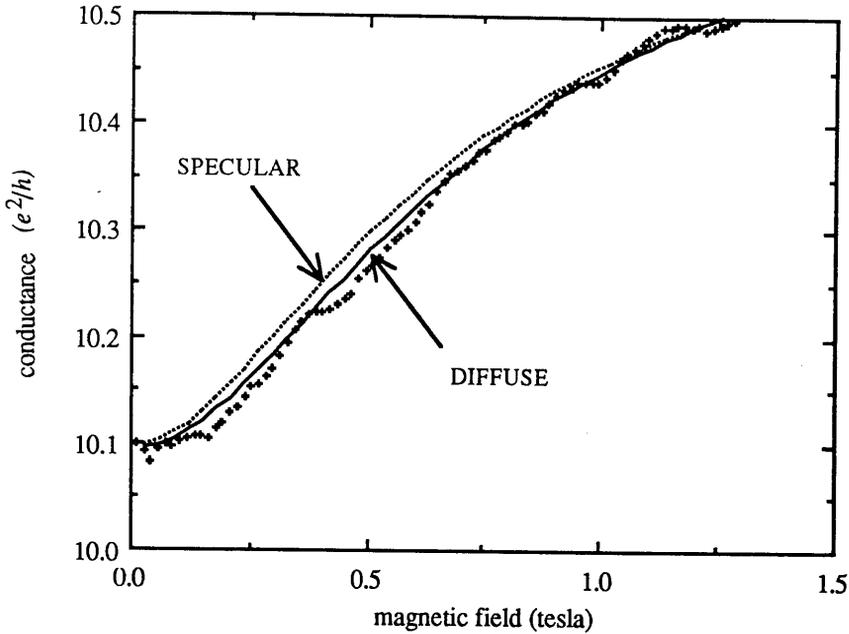


Figure 6.20: Comparison between specular and diffuse fits to the $T = 38$ K weak localisation measurement for wire E. The diffuse fit is a little better.

In §6.2.5 both classical and quantum contributions to the magnetoconductance were described. In that section, it was shown that classical size effects are dependent on l , whereas quantum size effects depend (as ever) on L_ϕ . In n^+ -material, l is largely insensitive to temperature (this is not so in 2DEGs, as indicated by the high mobilities that are achieved at low temperatures), whereas L_ϕ is much more strongly temperature dependent. The fact that the low field magnetoconductance in Fig. 6.19 grows with decreasing temperature is indicative of quantum interference. Further, classical contributions to the magnetoconductance would primarily be expected when the classical cyclotron diameter is smaller than the conducting width of the wire. For $w_1 \sim 30$ nm, this would only happen for $B > 6$ T. These considerations indicate that the measurements of Fig. 6.19 do indeed represent weak localisation effects, and that the Beenakker-van Houten theory should be used in their interpretation.

The Beenakker-van Houten formula, Eq. (6.11), can be applied either for perfectly diffuse or perfectly specular scattering. As with the Al'tshuler-Aronov formula, the magnetoconductance curve for wire E cannot be fitted using a conducting width $w_1 = 30$ nm. Fig. 6.20 shows the best fits possible for the $T = 38$ K curve using parameters $L = 500$ nm and $w_1 = 20$ nm. A coherence length $L_\phi = 198(\pm 1)$ nm is required in

both cases [*i.e.* L_ϕ is fitted to the nearest nanometre in the diffuse case, and a different value of L_ϕ does not improve the fit in the specular case]. Although there is only a slight difference between the specular and diffuse scattering fits [only a small difference is expected – see Beenakker and van Houten (1988a)], the diffuse fit is better. The fit with specular scattering at $T = 38$ K can be improved by using $w_1 = 15$ nm (this is as for the Al'tshuler-Aronov formula, except a different value of L_ϕ is required). At higher temperatures a good fit for specular scattering can only be achieved using $w_1 = 20$ nm. On the other hand, the diffuse scattering values of C_1 and C_2 give consistent fits at all temperatures using $L = 500$ nm and $w_1 = 20$ nm always, as shown by the solid lines in Fig. 6.19. It should be noted that for these fits the diffusion coefficient was maintained at its bulk value, $D = 0.017$ m² s⁻¹, although technically it should be less than this for the reasons outlined in §6.2.5 [see Beenakker and van Houten (1988a)]. If a smaller value of D is used, then the conducting width, w_1 , required for a good fit would have to be reduced.

In Fig. 6.19, the low field magnetoconductance of the 4 K curve is strongly perturbed by fluctuations, so the "correct" value of L_ϕ cannot be determined from just one curve. A better estimate can be found by taking the zero field correction, $\Delta\bar{G}_{1D}$, averaged over several different UCF traces (see Footnote 5). [Such an averaging procedure was demonstrated to be useful by Mailly and Sanquer (1991). The averaging "removes" the fluctuations, but leaves the weak localisation corrections.] This gives $\Delta\bar{G}_{1D} \approx 44$ μ S (averaged over 7 traces). So, using Eq. (5.24c) gives,

$$L_\phi \approx \frac{\pi\hbar}{e^2} \Delta\bar{G}_{1D} L \approx 284 \text{ nm} , \quad (6.13)$$

which in turn gives the 4 K fit shown in Fig. 6.19. [The 4 K experimental points (ignoring the fluctuations) in Fig. 6.19 are fitted reasonably well by this value of L_ϕ . However, other measured 4 K curves would have shown a poorer fit.]

The values of L_ϕ obtained from the fits in Fig. 6.19 are plotted as a function of temperature in Fig. 6.21. Using $\tau_\phi^{-1} = D/L_\phi^2$, the points in Fig. 6.21 can be plotted in terms of scattering rate as a function of temperature. This is done in Fig. 6.22. The dependence looks almost linear, but this is illusory. [Black *et al.* (1979) have derived a linear T dependence for the 3D scattering rate due to the presence of "two-level systems." Despite the fact that the defects introduced by the dry-etching are imagined to have two levels, one of these levels is always occupied (according to the model), so these defects are not really "two-level systems" at all. It would be wrong, therefore, to try to invoke this mechanism to explain the data of Fig. 6.21.]

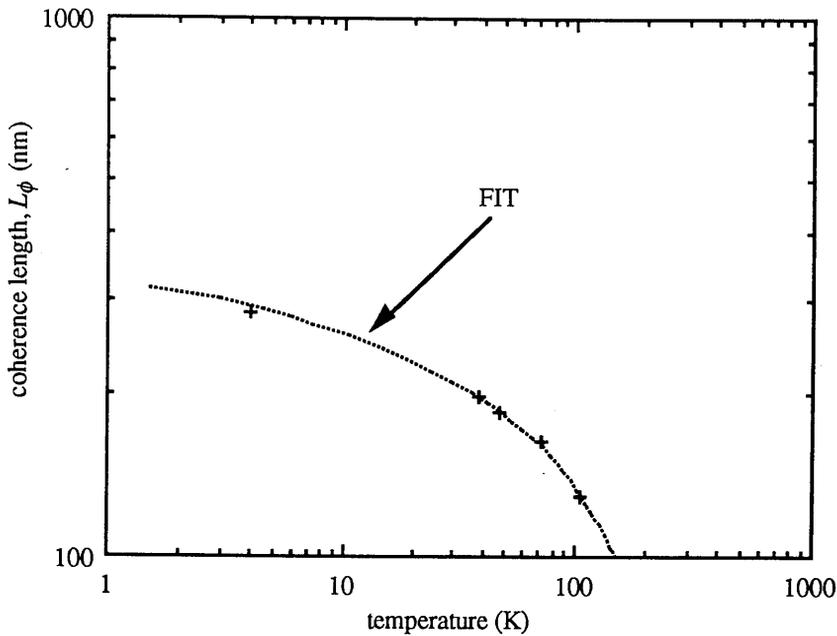


Figure 6.21: Temperature dependence of coherence length in wire E. The fit is made by including electron-phonon and electron-electron scattering rates in the phase breaking rate. The measured coherence lengths also require a temperature independent term to be included.

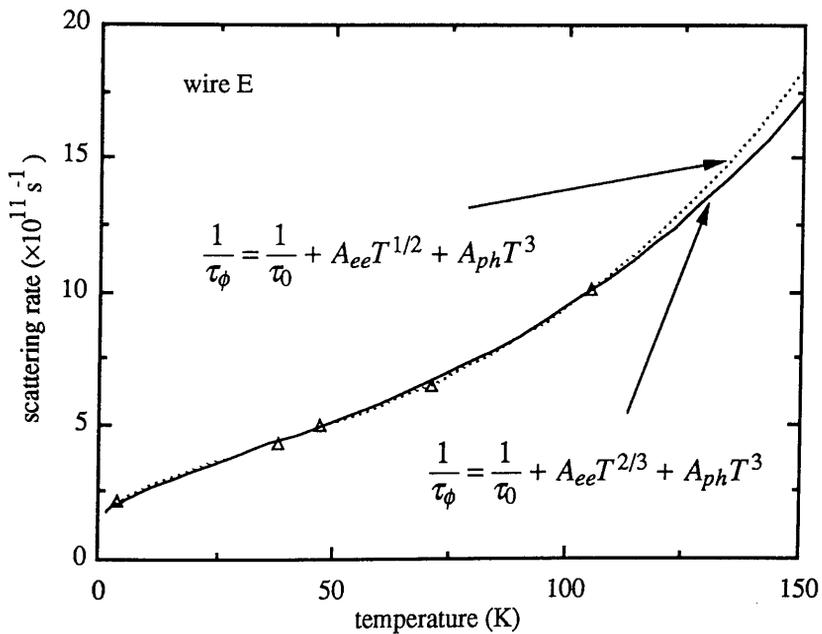


Figure 6.22: Temperature dependence of phase breaking rate as determined from Fig. 6.21 using $\tau_\phi^{-1} = D/L_\phi^2$, using $D = 0.017 \text{ m}^2 \text{ s}^{-1}$. The data can be fitted using either disorder modified electron-electron scattering ($T^{1/2}$), or electron-electron scattering with small energy transfers ($T^{2/3}$).

At low temperatures the dominant phase breaking mechanism is expected to be electron-electron scattering. Now, wire E is 1D with respect to electron-electron interactions (but has a 3D Fermi surface), since its conducting width, $w_1 \sim 30$ nm, and conducting thickness, $t_1 = 37.5$ nm (epilayer thickness minus top surface depletion depth), are both less than L_ϕ or L_T in this wire. The temperature dependence of electron-electron scattering in 1D can be described by a power law, $\tau_{ee}^{-1} \propto T^p$, where $p = 2, \frac{2}{3},$ or $\frac{1}{2}$. In clean metallic systems (no disorder, or $l \gg L_T$), electron-electron scattering is described by the Landau-Baber dependence, $\tau_{ee}^{-1} \propto T^2$, for a 3D Fermi surface [Kaveh and Wiser (1984)]. In the presence of disorder ($l < L_T$), this becomes modified. Al'tshuler and Aronov (1981b) show that for 1D systems the disorder modified electron-electron scattering rate is given by,

$$\frac{1}{\tau_{ee}} \sim \frac{e^2 D}{\hbar GL} \left(\frac{k_B T}{\hbar D} \right)^{1/2}, \quad (6.14)$$

where D is the diffusion coefficient, G is the wire conductance, and L its length. On the other hand, Al'tshuler *et al.* (1982) show that phase breaking can also occur when the electron interacts with fluctuations in the externally applied electromagnetic field (which is equivalent to electron-electron scattering with energy transfers $\delta E \ll k_B T$).

This gives the scattering rate,

$$\frac{1}{\tau_N} = \left(\frac{\sqrt{2D} e^2}{GL} \frac{k_B T}{\hbar^2} \right)^{2/3}, \quad (6.15)$$

called the Nyquist rate. In the n^+ -GaAs samples being used, the electron-electron contribution to τ_ϕ^{-1} is expected to take one of these latter two forms.

Apart from the electron-electron contribution to τ_ϕ^{-1} , there are two other contributions. The first is the electron-phonon scattering rate. At very low temperatures, $T \ll \theta_D$ (where $\theta_D \sim 240$ K is the Debye temperature), the wavevector of thermal phonons is given by $q \approx k_B T / \hbar c$, where $c \sim 3 \times 10^3$ m s⁻¹ is the velocity of sound [see Blakemore (1982) for a review of the major properties of GaAs]. The corresponding wavelength is ~ 36 nm at 4 K. The wire is thus 3D as far as the phonons are concerned. Also $ql \sim 9$, which corresponds to the "clean" limit ($ql \gg 1$). For 3D electron-phonon scattering in this limit, Thouless (1977) gives the approximation,

$$\frac{1}{\tau_{ph}} \approx 22 (\Gamma - 1) \frac{300}{\theta_D} \frac{v_F}{lk_F^3} \left(\frac{k_B T}{\hbar c} \right)^3, \quad (6.16)$$

where v_F is the Fermi velocity, and Γ is the resistance ratio [$\Gamma = \rho_{300}/\rho_0$, where ρ_{300} is the resistivity at 300 K, and ρ_0 is the residual resistivity].

Fig. 6.22 also indicates that at $T = 0$, the phase breaking rate is non-zero, *i.e.*

there is a temperature independent contribution⁶ to τ_ϕ^{-1} . This is denoted τ_0^{-1} . Hence the net phase breaking rate is given by,

$$\frac{1}{\tau_\phi} = \frac{1}{\tau_0} + A_{ee}T^p + A_{ph}T^3, \quad (6.17)$$

where $A_{ee}T^p$ is given by one of Eqs. (6.14) or (6.15).

Fig. 6.22 shows fits of Eq. (6.17) to the experimental points. As it turns out, the data can be fitted equally well for both $p = \frac{2}{3}$ and $p = \frac{1}{2}$. [This problem has been noted previously by Heraud *et al.* (1987).] For $p = \frac{2}{3}$ the fit is given by $\tau_0^{-1} = 1.4 \times 10^{11} \text{ s}^{-1}$, $A_{ee} = 2.5 \times 10^{10} \text{ s}^{-1} \text{ K}^{-2/3}$ and $A_{ph} = 2.6 \times 10^5 \text{ s}^{-1} \text{ K}^{-3}$. For $p = \frac{1}{2}$ the corresponding values are $\tau_0^{-1} = 1.1 \times 10^{11} \text{ s}^{-1}$, $A_{ee} = 5.0 \times 10^{10} \text{ s}^{-1} \text{ K}^{-1/2}$ and $A_{ph} = 3.3 \times 10^5 \text{ s}^{-1} \text{ K}^{-3}$. The values of A_{ee} obtained here can be compared to the theoretical estimates of Eqs. (6.14) and (6.15). For wire E, $G \sim 0.4 \text{ mS}$, $L = 500 \text{ nm}$ and $D = 0.017 \text{ m}^2 \text{ s}^{-1}$. The theoretical and experimental values of A_{ee} are summarised in Table 6.3. Unfortunately, there appears to be approximate agreement for either case. Thus, the method of fitting scattering rates does not seem to be able to distinguish which type of electron-electron scattering mechanism (if either) dominates the low temperature phase breaking rate [this was also the conclusion reached by Heraud *et al.* (1987)]. Experiments by others [*e.g.* Thornton *et al.* (1986) for 2DEG wires] suggest that electron-electron scattering with small energy transfers should in fact dominate at low temperatures (*i.e.* $\tau_\phi^{-1} \approx \tau_N^{-1}$).

	$p = \frac{1}{2}$ (disorder)	$p = \frac{2}{3}$ (Nyquist)
experimental fit	5.0×10^{10}	2.5×10^{10}
theory	1.1×10^{11}	9.5×10^{10}

Table 6.3: Comparison of values of A_{ee} obtained experimentally and theoretically, with the experimental fits being done to Eq. (6.17). The measured data cannot distinguish between the two forms of scattering.

⁶ In light of Footnote 5, it could be argued that the saturation of L_ϕ in Fig. 6.21 is actually caused by electrical interference, giving an electron temperature of $\sim 10\text{-}20 \text{ K}$. However, the observation of a $T^{-1/2}$ dependence in Fig. 6.7 characteristic of energy averaging is in itself evidence of a temperature independent L_ϕ [Washburn *et al.* (1985)] in those measurements, where the electron temperature is known to be $\sim 4 \text{ K}$. This would suggest that the curve of Fig. 6.21 for wire E is more real than illusory, and that the electron temperature is nearer 4 K than $10\text{-}20 \text{ K}$.

The temperature independent term τ_0^{-1} in these wires is of considerable interest. Such a term is frequently observed in different systems, and a variety of causes have been suggested. Previously, Taylor *et al.* (1988) [see also Main *et al.* (1990)] in measurements on n^+ -GaAs wires, noted a width dependence of the temperature independent term, and concluded that some form of surface scattering was the cause. In measurements of fluctuations in wires A, D and E, the fact that the rms conductance fluctuation decreases with decreasing wire width clearly suggests a dependence of τ_ϕ^{-1} on the conducting width. [The $\Delta B = 0$ value of the Lee-Stone correlation function is unaffected (directly) by boundary scattering, so the rms fluctuations are not reduced because of this. See Beenakker and van Houten (1988b).] If the phase breaking rate in each of these wires could be described by Eq. (6.17) with the *same* values of A_{ee} and A_{ph} , then the width dependence of τ_ϕ^{-1} requires that τ_0^{-1} is a function of the wire width. [A_{ee} and A_{ph} are not expected to be strongly affected by the presence of the boundaries.]

In light of the model on dry-etch damage in Chapter 4, a temperature independent phase breaking rate could be related to the defect states near the etched sidewalls. Two possible mechanisms are evident. Inelastic scattering of free electrons by empty defect states within the wire is one possibility, although whether this would have a significant temperature dependence or not is uncertain. Another possibility is magnetic scattering due to the interaction between the spin of the conduction electron and the unpaired spin of an electron trapped in a defect state. Magnetic scattering should be temperature-independent over the range of temperatures used (in spin glasses the Kondo effect gives a slight temperature dependence at very low temperatures). The defect used in Chapter 4 had two levels. If both levels are filled, the net spin should be zero. Hence, only the defects placed within the depletion region, which have trapped only one electron, will contribute to a magnetic scattering rate. This clearly emphasises the role of electron scattering from the edges of the depletion region, since only then would the conduction electron and defect be close enough for scattering to occur.

The scattering time $\tau_0 \sim 7$ ps gives a length $L_0 \sim 350$ nm. Comparing this to the length of wire E, $L = 500$ nm, shows that there are roughly 1-2 such scattering events as the electron diffuses down the wire. From the sidewall defect distribution obtained in Chapter 4, a simple estimate may be made of the number of defects contributing to the scattering mechanism in question. Assume that the sidewall defect distribution is given (very roughly) by $\Gamma(x, \nu\tau/2, \tau)$ [from Eq. (4.14)], where x is the distance from the

sidewall, $v = 200 \text{ nm min}^{-1}$ is the etch rate, and $\tau = 30 \text{ s}$ is the etch time. The edge of the sidewall depletion region is at $x = x_d \approx 30 \text{ nm}$. The conducting thickness of the epilayer is $t_1 = 37.5 \text{ nm}$. If the uncertainty in energy of the electron is $k_B T$, then the uncertainty in the position of the edge of the depletion region is $\Delta x \sim (2\epsilon_0\epsilon_r k_B T / e^2 N_D)^{1/2} = 0.28 \text{ nm}$, where $N_D = 6.5 \times 10^{24} \text{ m}^{-3}$ is the donor density. [Up to a factor, Δx is just the Debye length corresponding to the transition region between the depletion region and the conducting core.] The electrons will primarily interact with the defects lying within Δx of the "edge" of the depletion region. The total number of such defects between the voltage probes in wire E is given by,

$$N_0 \sim 2\Delta x t_1 L \times \Gamma\left(x, \frac{v\tau}{2}, \tau\right) = G_1 \Delta x t_1 L \tau e^{-x_d/\lambda} \sim 3, \quad (6.18)$$

where $L = 500 \text{ nm}$, $G_1 = 5.5 \times 10^{23} \text{ m}^{-3} \text{ s}^{-1}$ is the defect creation rate at the sidewall, and $\lambda = 9 \text{ nm}$ is the mean depth to which defects are created (these values were obtained in §4.4.3). Thus, it appears that the electron interacts with ~ 3 scatterers as it diffuses from one voltage probe to the other.

The effect of these scatterers on the electron coherence depends on the relative values of l and w_1 . If $w_1 < l$, then the electron spends more time interacting with the edge of the depletion region than it does with impurities within the wire. In this case, the electron is more likely to encounter one of the defect scatterers. As w_1 is increased, the electron interacts less with the edge, and the corresponding effects are smaller. These arguments are identical to those used in discussing size effects in §6.2.3, with L_0 being used in analogy to the boundary scattering length, l_b . In essence then, this argument says that the temperature independent contribution to the phase breaking rate is due to a size effect, and (in principle) should not be present in large samples. The top surface depletion region does not contain such defects, so its presence should not affect the phase breaking rate.

6.2.7 Coexistence of localisation and fluctuation effects

The UCF traces for wires A, D and E show that the rms conductance fluctuation decreases with wire width. However, the magnitude of the corresponding zero field weak localisation correction remains unchanged at roughly $\Delta G \sim e^2/h$. This observation shows that localisation and fluctuation effects must be distinct, which is what the usual theory says. Taylor *et al.* (1988) have proposed a simple model which suggests that fluctuations are a manifestation of weak localisation. The model relies on closed time-reversed paths as in weak localisation. However, such paths only exist in

weak magnetic fields, such that $L_B > l$, where the curvature of the electron trajectories can be ignored. At sufficiently high fields, time-reversed paths will no longer exist, and one of the basic assumptions of the model no longer holds. Such a model would not therefore seem able to account for fluctuations at any significant values of field, and the intended link between localisation and fluctuation effects is lost.

It should be stated that localisation and fluctuation effects are only distinct in the limit $k_F l \gg 1$. As the sample becomes more disordered, it is found that localisation *does* affect the rms conductance fluctuation. Signs of this appear in the distribution function of the fluctuations [Eq. (5.31)], which becomes progressively more non-Gaussian as the sample is made more disordered. See, for example, Lerner (1991).

6.3 Hall and nonlocal measurements

6.2.2 The Hall measurement

The classical Hall effect, observed in a transverse measurement configuration (*e.g.* using the Hall geometry shown in Fig. 5.5), arises from a balancing of the Lorentz force, $-e\mathbf{v} \times \mathbf{B}$, and the force due to the Hall field, $-e\mathbf{E}_H$. The measured I - V curve shows a linear dependence of the Hall voltage, V_H , with applied magnetic field strength,

$$V_H = \frac{I}{n_{2D}e} B, \quad (6.19)$$

where n_{2D} is the sheet carrier concentration, and B is the component of the field pointing in a direction normal to the plane of the device. In 2DEGs and Si inversion layers at low temperatures, when the coherence length becomes sufficiently long, Landau quantisation leads to plateaux forming in the Hall conductance, I/V_H , measured as a function of field. This is known as the (integer) quantum Hall effect. [A fractional version also exists, which is a many-body effect associated with the electron gas in a magnetic field.] In n^+ -GaAs, the elastic mean free path is sufficiently small that the quantum Hall effect is not observed (although weak Shubnikov-de Haas oscillations can be seen at high enough fields).

With the working contacts, only a single Hall configuration can be measured using the device shown in Fig. 6.3. This is shown in Fig. 6.23. This Hall resistance is written $R_{JC,EA}$, which denotes current flow between contacts J-C, and a voltage measurement between contacts E-A. At high temperatures, a measurement of $R_{JC,EA}$

yields the linear relation expected from Eq. (6.19). The gradient of this measurement gives a material carrier concentration of $\sim 5.0 \times 10^{24} \text{ m}^{-3}$, not the $6.5 \times 10^{24} \text{ m}^{-3}$ measured by Van der Pauw. The reason for this difference may be due to fluctuations in doping density across the wafer. However, the measured cut-off widths in Chapter 3 were always found to lie within a small range of values around 60 nm, despite a large number of samples being made from the material. Thus, the differences in doping density across the wafer do not seem to affect the cut-off width beyond this small range. The parameter values given in Table 6.1 only change by a few percent if the carrier concentration is $5.0 \times 10^{24} \text{ m}^{-3}$ instead of $6.5 \times 10^{24} \text{ m}^{-3}$.

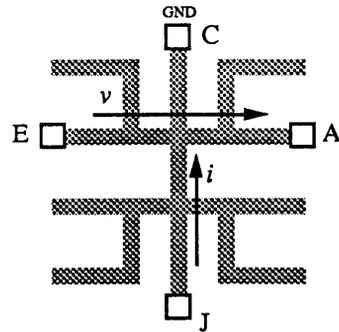


Figure 6.23: The Hall measurement

$R_{JC,EA}$ made using the device of Fig. 6.3.

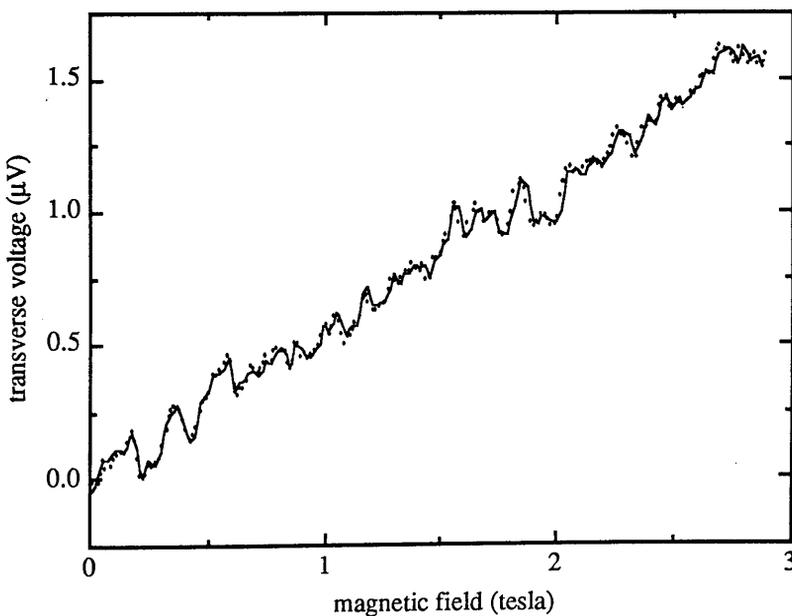


Figure 6.24: 4 K measurement of a Hall voltage using the configuration shown in Fig. 6.23.

Superimposed on the classical linear Hall voltage are reproducible fluctuations.

At 4 K, the measured curve changes. Two measurements of $R_{JC,EA}$ taken at 4 K are shown in Fig. 6.24. Superimposed on the classical linear Hall rise are definite, reproducible fluctuations. Qualitatively, the fluctuations look just like the ones observed in the longitudinal measurement, although the magnitudes cannot be compared directly. Fluctuations in a direct Hall measurement have been studied experimentally by Hansen *et al.* (1990), and theoretically by Ma and Lee (1987). As for the longitudinal measurement, the cause of these fluctuations lies in the interference between the many electron trajectories that exist in the sample. This interference results in electric fields with random magnitude and direction being created near the voltage probes, giving the "random" fluctuations [Al'tshuler and Khmel'nitskii (1985); Kane *et al.* (1988)].

A comparison between the fluctuations in the longitudinal and Hall measurement cannot be made using conductance, since the Hall conductance diverges as $B \rightarrow 0$. Instead, the discussion in §6.2.3 suggests a different method by which these two quantities may be compared. In that section it was stated that voltage probes attached to a phase coherent region always measure an rms voltage fluctuation that is independent of the probe spacing. Here it will be shown that experimentally this is also true for the case where the voltage probes are in a Hall – as opposed to a longitudinal – configuration. In other words, the fluctuations associated with the Hall measurements are just the rms voltage fluctuations associated with a phase coherent region.

The wire has length L , conducting width w_1 , and resistance R . The coherence length in the wire is L_ϕ . Since $L > L_\phi$, the measured longitudinal voltage scales with both wire width and probe spacing. The expectation is that the fluctuations in the Hall measurement are just fluctuations associated with a phase coherent region, so the longitudinal rms voltage fluctuation must be scaled to a wire of length L_ϕ and width w_p , where w_p is the conducting width of the voltage probe. Now $\delta V_{xx} = IR^2 \delta G_{xx}$ is the voltage fluctuation in the longitudinal measurement. If $L > L_\phi$, then $\delta G_{xx} = (L_\phi/L)^{3/2} \delta G_\phi$, where δG_ϕ is the rms conductance fluctuation of a phase coherent region. Also if R_ϕ is the resistance of a phase coherent region of width w_p , where $w_p < L_\phi$, then $R = (w_p L / w_1 L_\phi) R_\phi$. Hence, the equivalent rms voltage fluctuation of a wire of length L_ϕ and conducting width w_p is,

$$\delta V_{eq} = IR_\phi^2 \delta G_\phi = \left(\frac{L_\phi}{L}\right)^{1/2} \left(\frac{w_1}{w_p}\right)^2 \delta V_{xx} . \quad (6.20)$$

This relation was also derived by Hansen *et al.* (1990). Hansen *et al.* (1990) further show that a more accurate expression can be derived:

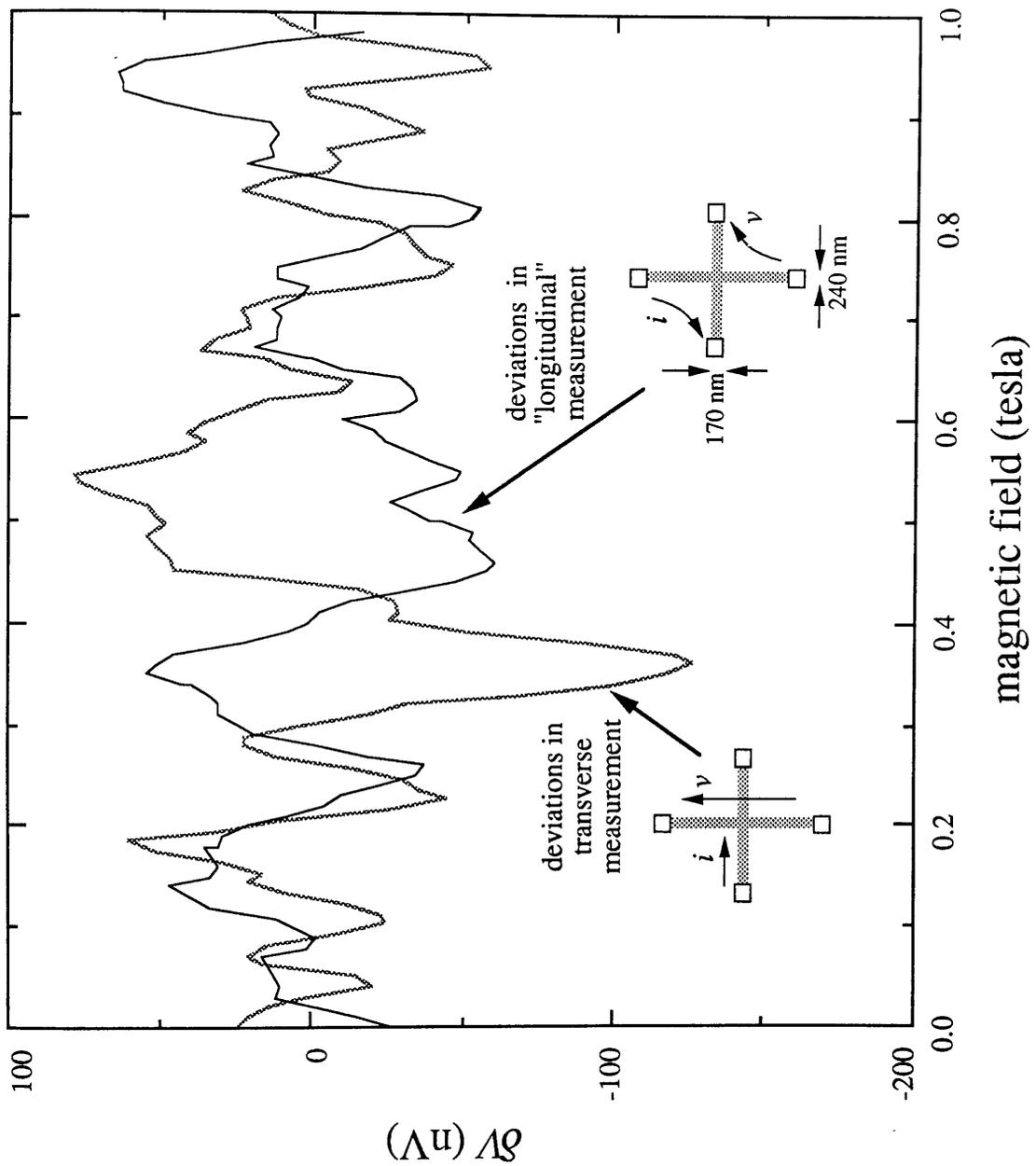


Figure 6.25: Fluctuations measured at a cross-junction. A linear background has been subtracted to give δV in the figure. The magnitudes of the fluctuations are similar for the "longitudinal" and Hall configurations.

$$\delta V_{eq} = \left(\frac{L_\phi}{L}\right)^{1/2} \frac{w_1}{w_p} \left[\frac{2w_1w_p + 2w_p^2}{(w_1 + w_p)^2} \right]^{1/2} \delta V_{xx} . \quad (6.21)$$

If the assertion concerning voltage probes connected to phase coherent regions is correct, then the relation $\delta V_{eq} \approx \delta V_{xy}$ is expected to hold. Here δV_{xy} is the rms voltage fluctuation in the Hall measurement once the classical voltage given by Eq. (6.19) is subtracted off.

The rms voltage fluctuation in the Hall measurement of Fig. 6.24 is $\delta V_{xy} = 56$ nV, where the applied current is $I = 16.72$ nA. The corresponding rms fluctuation in wire A (*c.f.* Figs. 6.4 and 6.22) for the same current level is $\delta V_{xx} = 78$ nV. Using Eq. (6.21) and $L = 600$ nm, $w_1 = 180$ nm, $w_p = 130$ nm, and $L_\phi \approx 520$ nm, gives $\delta V_{eq} = 92$ nV. Allowing for uncertainty in the parameter values used, this is in reasonable agreement with δV_{xy} . Hall measurements on wire E (see Fig. 6.14) gave $\delta V_{xy} = 221$ nV and $\delta V_{xx} = 3.25$ μ V for a current $I = 32.67$ nA. Using $L = 500$ nm, $w_1 \sim 30$ nm, $w_p = 180$ nm, and $L_\phi \sim 300$ nm, gives $\delta V_{eq} = 397$ nV. Again there is reasonable agreement. The agreement found for these wires is better than that found by Hansen *et al.* (1990), where δV_{xy} and δV_{eq} only agreed to within an order of magnitude or so. It is possible that the differences between δV_{xy} and δV_{eq} are due to uncertainties in L_ϕ within the probe and wire, and perhaps additional geometrical effects. Hansen *et al.* (1990) also acknowledge the fact that $w_p \sim L_\phi$ in one of their devices, which would undoubtedly affect δV_{xy} . In the limit $w_p \gg L_\phi$, the fluctuations in the Hall measurement would be expected to average to zero.

As one further test, the fluctuations at the cross-junction itself can be measured, as shown in Fig. 6.25. Two measurements can be performed using the four contacts, one where the voltage probes are on the same side of the current (giving δV_{xx}), the other where the probes are on opposite sides (giving δV_{xy}). Both measurements in Fig. 6.25 were made using the same excitation voltages, and give $\delta V_{xx} = 50$ nV, and $\delta V_{xy} = 35$ nV. Allowing for geometrical effects, these values are in very good agreement, and show that the voltage fluctuations measured by probes attached to a phase coherent region (in this case the junction itself), are independent of the position of the probes with respect to the current path. Any differences arise for reasons mentioned previously.

One last point of interest to note about these direct Hall measurements is the $B = 0$ value of measured Hall voltage. The results of Hansen *et al.* (1990) and those above show that at low temperatures a finite "Hall" voltage is measured despite the field being

zero. This follows naturally from the idea that the measured voltage is a superposition of the true Hall voltage, and of a fluctuating term (which can be non-zero at $B = 0$) caused by electron interference [Al'tshuler and Khmel'nitskii (1985)].

6.3.2 Symmetries of the 4-terminal conductance

In the measurement shown in Fig. 6.17 of the longitudinal fluctuations of wire E, it was noted that the curve appeared to comprise a small, short period component, as well as a larger much longer period component. The longer period component was attributed to fluctuations in the longitudinal conductance. Fig. 6.26 shows the short-period component in terms of a voltage with the linear background subtracted off. The rms voltage fluctuation is $\delta V_{xx} = 296$ nV (where the current is $I = 32.68$ nA). Also shown in the figure are the fluctuations in the Hall measurement, with the classical Hall voltage subtracted off. The rms value of these is $\delta V_{xy} = 221$ nV. It is seen that the magnitudes of the rms fluctuation are very similar for the two measurements.

The similarity does not end there. The autocorrelation function for each curve is also shown in Fig. 6.26. The respective values for the correlation field are $B_C^{xx} \approx 0.035$ T, and $B_C^{xy} \approx 0.062$ T. These values of B_C^{xx} and B_C^{xy} are very similar (allowing for errors due to the small field range over which they were computed), and are much smaller than the value ~ 0.53 T obtained for the longitudinal fluctuations taken as a whole. These comparisons point to the conclusion that the *same* fluctuations are manifesting themselves in the two separate measurements. It will be asserted below that this is indeed the case, and that the measured fluctuations are associated with fluctuations in the antisymmetric part of the conductance.

Early measurements of fluctuations [*e.g.* by Umbach *et al.* (1984)] revealed that the 4-terminal conductance was not symmetric with respect to magnetic field, *i.e.* $G(-B) \neq G(B)$. The reason for this remained a puzzle for a while until Büttiker (1986) showed that it was in fact due to the 4-terminal nature of the measurement. He showed that the symmetry is restored if the current and voltage leads are swapped when the field is inverted, *i.e.* $G_{ij,lm}(-B) = G_{lm,ij}(B)$. This is just a statement of the Onsager-Casimir symmetry relations familiar from the study of classical conductors [more generally from the study of reversible processes – see Landau and Lifshitz (1960), Ch. 3], and shows that the 4-terminal quantum conductance measurement is analogous to a 4-terminal measurement of a classical conductor in this respect.

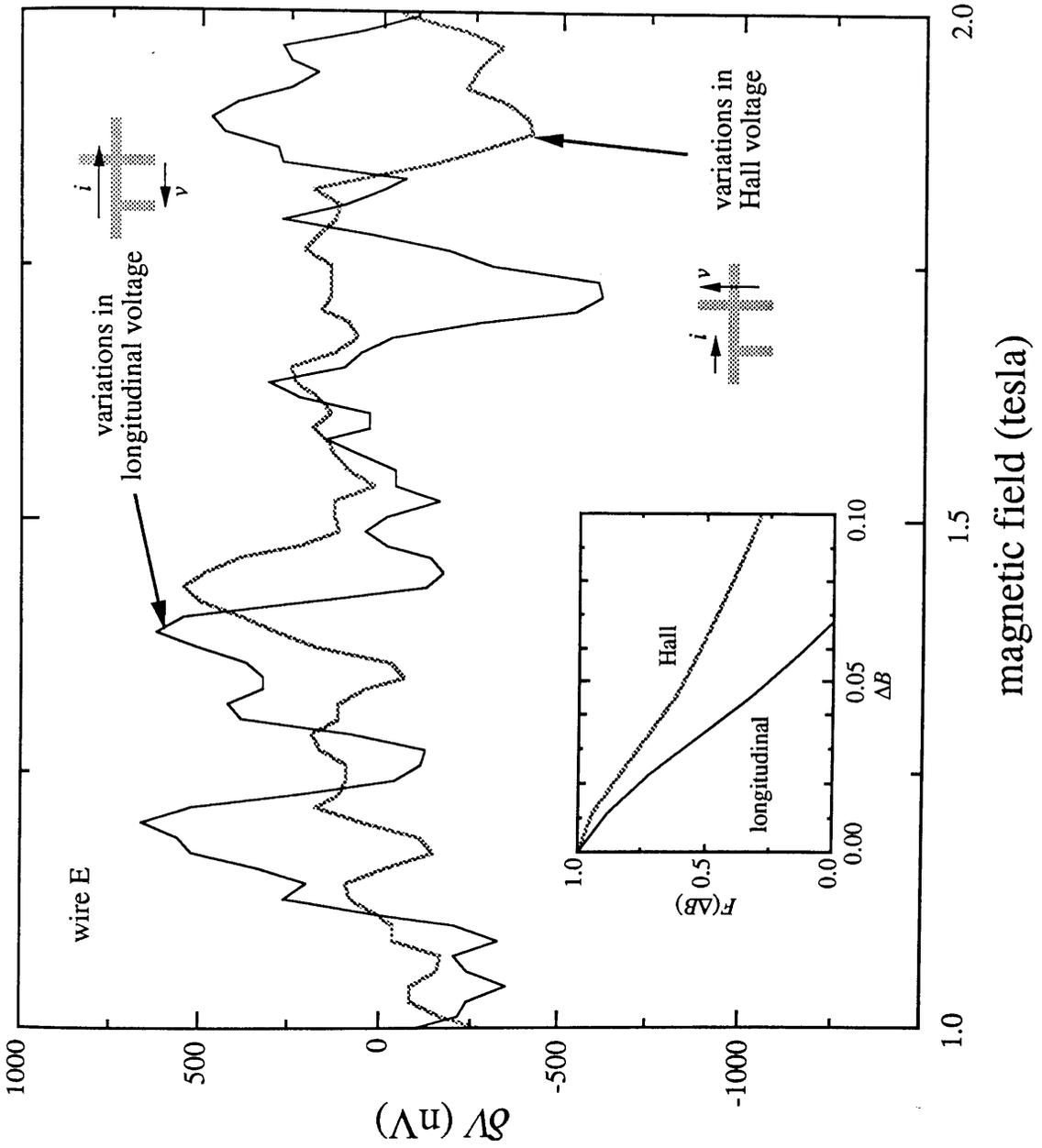


Figure 6.26: Comparison of the "small" fluctuations seen in wire E (Fig. 6.17), and a simultaneous Hall measurement. A linear background has been subtracted to give δV in the figure. The corresponding autocorrelation functions for the two measurements are given in the inset. The similarities between the two sets of measurements suggest they both arise from fluctuations around the regions of the voltage probes.

Following the classical analogy, Benoit *et al.* (1986) showed that the conductance fluctuations could be split into a symmetric part, $G_S(B) = G_S(-B)$, and an antisymmetric part, $G_A(B) = -G_A(-B)$, so,

$$G(B) = G_S(B) + G_A(B) . \quad (6.22)$$

They further showed that the symmetric part was associated with the wire itself, whereas the antisymmetric part was associated with fluctuations around the regions of the probes. This is also born out by theoretical calculations [*e.g.* Kane *et al.* (1988)]. This means that the magnitude of the antisymmetric component of the voltage fluctuations is largely independent of the probe spacing, whereas the symmetric component will decay [confirmed experimentally by Benoit *et al.* (1987) and Hauke *et al.* (1990)]. However, since the measured voltage across the wire increases with wire length, so the relative magnitude of the antisymmetric component becomes smaller and smaller, until for very long wires, these fluctuations cannot be distinguished from the background noise.

The above would suggest an explanation for the effect observed in Fig. 6.26. The Hall measurement probes the fluctuations *only* in the region of the voltage probes. Thus, the fluctuations in G_A are measured. The longitudinal measurement, on the other hand, probes the fluctuations not only near the voltage probes (G_A), but also along the length of the wire itself (*i.e.* G_S). This would therefore suggest that the small fluctuations seen in Fig. 6.26 are in fact fluctuations associated with G_A , *i.e.* fluctuations in the region of the voltage probes. In the measurements of Benoit *et al.* (1986), the conducting widths in the wire and the probe were similar, $w_1 \sim w_p$, so G_S and G_A had to be decoupled by making more than one measurement. However, for wire E above, $w_1 \ll w_p$, and the two different contributions are manifestly separated in the single longitudinal measurement (Fig. 6.17).

6.3.3 Nonlocal fluctuations

The longitudinal and Hall measurements have already been discussed. Apart from these two, there exists a third more intriguing possibility, as shown in Fig. 6.27. The voltage is effectively measured at a single point on the current path. Intuitively, a zero voltage would always be expected,⁷ and the measurement seems somewhat pointless.

⁷ In fact, due to the finite width of the wire, a small classical voltage is expected according to the Van der Pauw formula. However, as the length L is made longer, this classical voltage should tend to zero roughly as e^{-L/w_p} , where w_p is the probe width.

However, in the quantum regime this idea is radically altered, and a zero voltage is no longer found.

Fig. 6.27 also shows the two different ways this measurement can be done using the device of Fig. 6.3. In one case $L = 400$ nm, in the other $L = 470$ nm, where L is roughly the distance separating the voltage probes from the classical current path. Two traces of the measured voltage as a function of applied magnetic field for $L = 470$ nm are shown in Fig. 6.28. Reproducible fluctuations are seen, very similar to those observed for the longitudinal and Hall measurements. These fluctuations are uncorrelated with the fluctuations measured in the other configurations, showing that the electron paths contributing to these are not the same. The fluctuations are superimposed on a constant non-zero background voltage. This background voltage scales with current and remains at high temperatures, and is attributed to a classical (Van der Pauw) voltage. The measurements for the $L = 400$ nm case gives similar fluctuations except with a larger rms value, and a higher constant background voltage.

Since the fluctuations arise in measurements away from the classical current path, the effect is called nonlocal (or remote). Such nonlocal fluctuations have been extensively studied experimentally in Sb, Au and Si-inversion layer wires [Benoit *et al.* (1987); Skoçpol *et al.* (1987); Hauke *et al.* (1990); Washburn (1991)], and are well understood theoretically [DiVincenzo and Kane (1988); Baranger *et al.* (1988)]. The measurements made on n^+ -wires above adds nothing new [Geim *et al.* (1991) have also observed nonlocal fluctuations in n^+ -wires], so the present discussion is limited to a few brief comments.

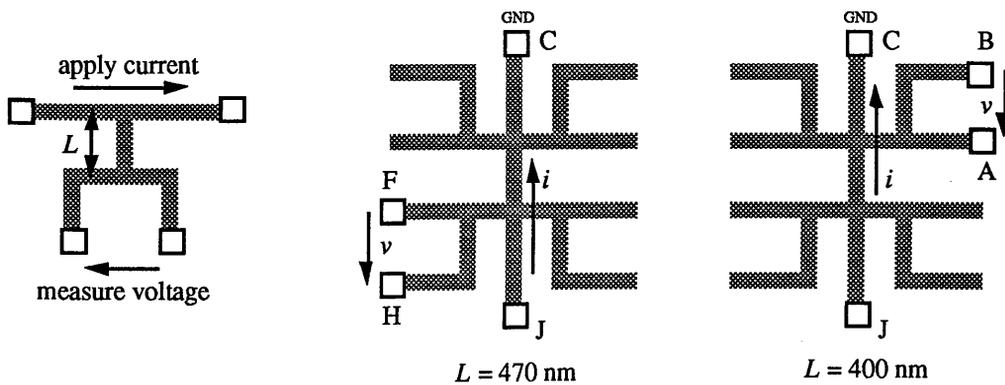


Figure 6.27: In nonlocal measurements the voltage probes are "remote" from the classical current path. The two measurements possible using the structure of Fig. 6.3 are shown.

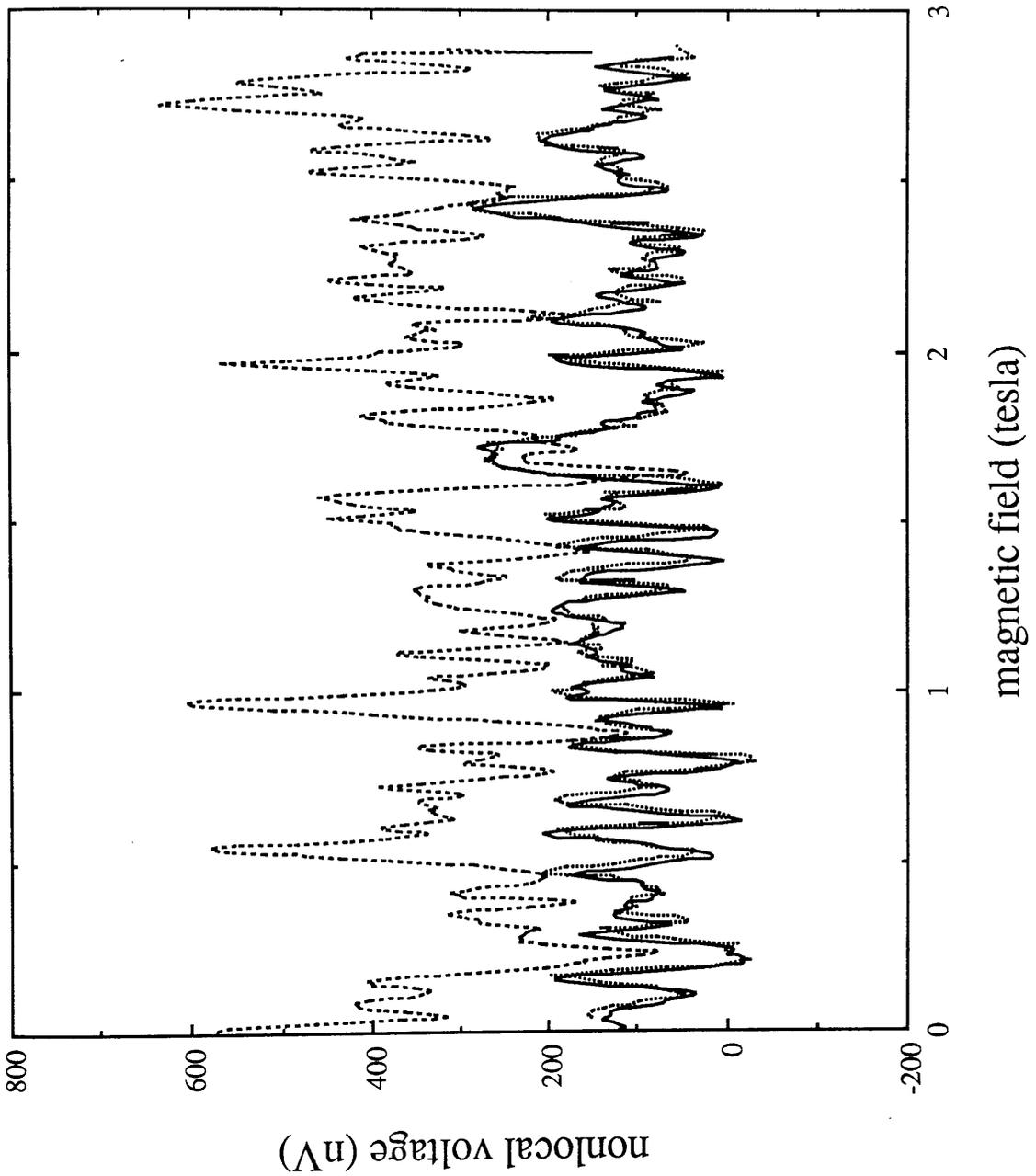


Figure 6.28: Measurements of nonlocal voltage fluctuations made for lengths $L = 400$ nm (upper curve) and $L = 470$ nm (lower curve). The two curves in the $L = 470$ nm case show the reproducibility. The rms magnitude of the fluctuations is larger for the $L = 400$ nm case. The offset is a classical (Van der Pauw) voltage, which is also larger for the $L = 400$ nm case.

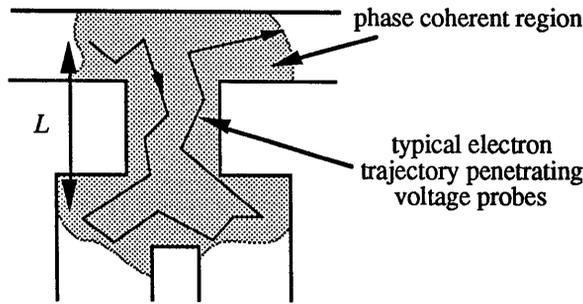


Figure 6.29: Mechanism for nonlocal fluctuations. Coherent electron trajectories extend from the classical current path towards the remote voltage probes. This creates fluctuating electric fields at the voltage probes, which are detected as a fluctuating voltage measurement.

Nonlocal fluctuations occur for exactly the same reasons that voltage fluctuations saturate at a finite value for voltage probes attached to a phase coherent region (see §6.2.3). This is illustrated in Fig. 6.29. Coherent electron trajectories extend from the classical current path into the region of the voltage probes. Since the resultant interference patterns at the mouths of the two voltage probes are different, the electric fields at these points are also different. As the magnetic field is swept, the magnitudes of the electric fields fluctuate, leading to fluctuations in the measured voltage. Since there is a finite probability of a trajectory losing its coherence at any point along the length L , the number of such trajectories is exponentially damped. More detailed theoretical analysis yields a nonlocal voltage fluctuation [DiVincenzo and Kane (1988)]:

$$\delta V_{rem} = \delta V_{\phi} e^{-aL/L_{\phi}} , \quad (6.23)$$

where δV_{ϕ} is the voltage fluctuation measured when $L \rightarrow 0$ (*c.f.* §6.2.3). Theory yields the value $a \approx 1.1$, and $a \rightarrow 1$ as $L \rightarrow \infty$. The value of a has a slight dependence on the exact geometry used. Experimentally, the value $a \approx 1.2 \pm 0.3$ is found for narrow wires [Hauke *et al.* (1990)].

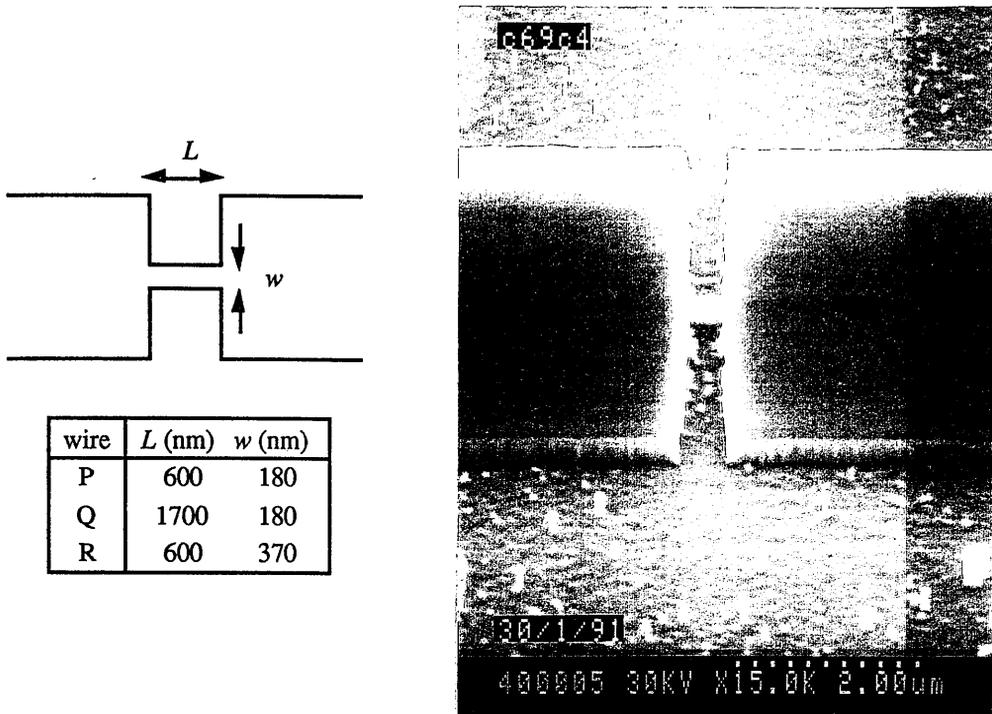


Figure 6.30: Two-probe wires used to study random telegraph signals in response to light pulses.

6.4 Photoconductance

6.4.1 Random telegraph signals in 2-probe wires

It has already been mentioned that the conductance of a mesoscopic sample is very sensitive to the exact configuration of the scatterers in the sample. Al'tshuler and Spivak (1985), and Feng *et al.* (1986) have shown that displacing so much as one scatterer is sufficient to change the quantum conductance on average by $\sim e^2/h$. Classically, no change in the conductance is expected, since the classical conductivity depends only on the total number of scatterers rather than on their positions.

From the work in Chapters 3 and 4, it is known that SiCl_4 dry-etching at around 300 V produces wires with a distribution of defect levels near the sidewalls. Since the charge states of these levels can change by trapping or detrapping carriers, it would be expected that this may lead to time-dependent changes in the measured conductance.

Such changes would therefore reflect the dynamics of electron interaction with the defect levels. Two methods may be used to study such time-dependent effects in these n^+ -wires. The first is to apply high voltages, which have the effect of heating the electron gas in the manner described in §3.5.4, thereby encouraging electron interaction with defect states [a method used, for example, by Mailly *et al.* (1989)]. The second is to apply a light pulse to the sample and measure the time-dependent response [*e.g.* Bykov *et al.* (1988)]. This is the method used in this part of the present work.

The time-dependent response to a light pulse in these n^+ -wires generally contains "pulses" in the conductance versus time. These random telegraph signals (RTS), as they are called, have been observed in many other types of structures, *e.g.* ultra-small Si MOSFETs [Ralls *et al.* (1984)], dry-etched GaAs/AlGaAs wires [Ochiai *et al.* (1988); Mailly and Sanquer (1990)], quantum point contacts [Cobden *et al.* (1991)], and others. This section and the next both study RTS. In some respects, these fluctuations are similar to the magnetic field fluctuations. However, the relative magnitude of these fluctuations is much smaller, $\delta G_{rms} \ll e^2/h$. This means that they tend to be experimentally observable only for low conductance wires, *i.e.* wires very close to cut-off. Enough measurements could not be made to prove conclusively that these are quantum fluctuations, but the wires are in a regime where such effects are expected.

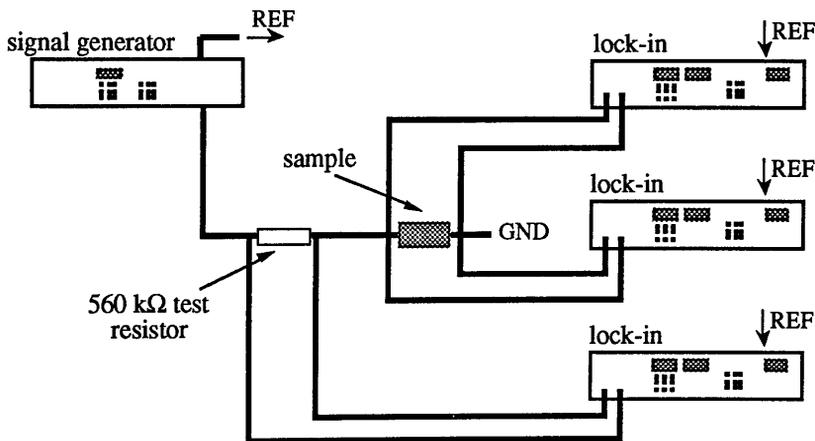


Figure 6.31: Measuring circuit used with two-probe wires. The measurement is 4-terminal, but the resulting conductance includes a contribution of leads as well as from the wire.

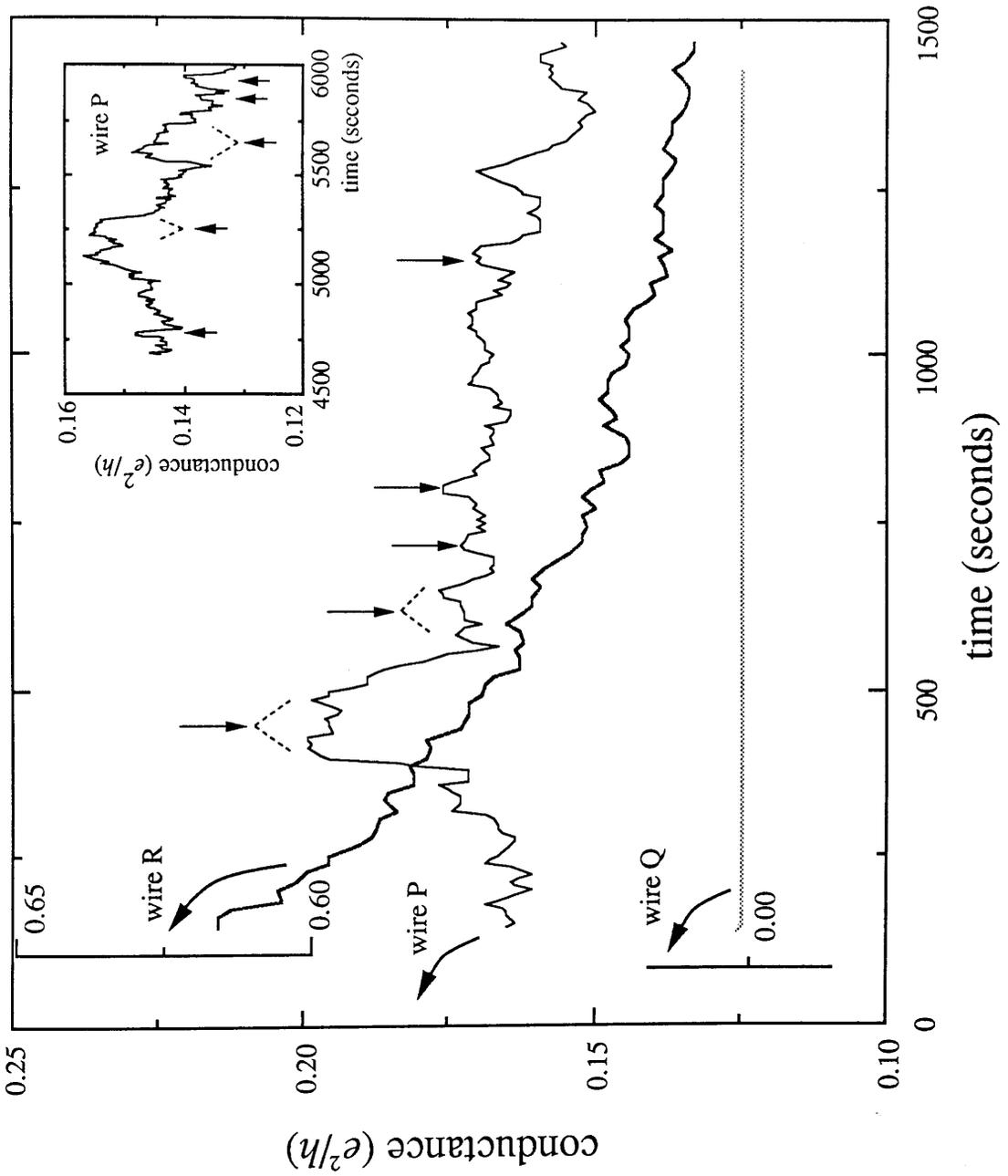


Figure 6.32: Response to a light pulse for wires P, Q and R. Wire Q is quickly cut-off, whilst wire R only shows system noise. Wire P appears to show "pulses," which continue long after the illumination has been removed (see inset). These random telegraph signals are attributed to charge motion between defects within the wire.

The devices used here were fabricated from MBE-grown epitaxial n^+ -GaAs with a carrier concentration of $9.0 \times 10^{23} \text{ m}^{-3} \text{ s}^{-1}$ and a mobility $0.20 \text{ m}^2 \text{ V}^{-1} \text{ s}^{-1}$. The cut-off width for dry-etched wires made from this material is around 170 nm, and $k_{Fl} \approx 12$ in the bulk. Initial studies were made on four probe wires similar to those used in §6.2. Unfortunately, in these wires the voltage probes were generally so resistive that their response to the light pulse was affecting the measurements. Instead, two probe wires were used for these experiments. A typical device is shown in Fig. 6.30. The measuring circuit is shown in Fig. 6.31. The devices are two probe, but a 4-terminal conductance measurement is made. This means that the total measured conductance is a combination of the wire conductance, and the lead conductances.

Three wire dimensions were used: (P) $L = 600 \text{ nm}$, $w \approx 180 \text{ nm}$; (Q) $L = 1.7 \mu\text{m}$, $w \approx 180 \text{ nm}$; and (R) $L = 600 \text{ nm}$, $w = 370 \text{ nm}$. The widths are physical widths. The conducting widths are roughly 170 nm less. Under illumination, the measured conductances were $G_P = 21.3 \mu\text{S}$, $G_Q = 6.1 \mu\text{S}$, and $G_R = 40.5 \mu\text{S}$. The measurement was made at a temperature of 4 K. The wires were illuminated with red light for several seconds, and the conductance was monitored after removal of the light. Fig. 6.32 shows the measured conductances as a function of time for each of the wires. Of the three curves, only wire P appears to show any significant noise. Wire Q becomes cut-off after 100-200 seconds, and wire R shows only noise from the measuring circuit. Wire P has a width near the cut-off width (in the steady state), and a comparatively short length.

The response to the light pulse has two components. One is the falling background conductance apparent in all the devices for $t < 200 \text{ s}$, which is due to the recombination of excess carriers generated by the light. The second, which is prominent in wire P, is a pulse like structure. Some of the more "definite" pulses (without becoming too subjective) in Fig. 6.32 are marked with an arrow. These pulses are still observed $1\frac{1}{2}$ hours after the initial illumination, as evident in the inset of Fig. 6.32. The step changes in conductance are explained in the theory of Al'tshuler and Spivak (1985), and of Feng *et al.* (1986) as being caused by a sudden change in the impurity configuration within the wire. This is analysed in more detail below.

The rms fluctuation measured for wire P (after the background is subtracted off) is $\delta G_P = 0.011 e^2/h$. When the conductance was remeasured roughly $1\frac{1}{2}$ hours after the initial illumination (see inset to Fig. 6.32), the rms fluctuation was slightly less, δG_P

$= 0.007 e^2/h$. The mean conductance was about $5.5 \mu\text{S}$. In analogy with the fluctuations with respect to magnetic field, a time autocorrelation function, $F(\Delta t)$, can be computed. This is shown in Fig. 6.33. The condition $F(t_c) = \frac{1}{2}$ defines the correlation time t_c , which is the time scale over which the fluctuations occur. In the initial measurement on wire P, the autocorrelation function gives $t_c \approx 70$ s. In the second measurement ($\frac{1}{2}$ hours after the initial illumination) the correlation time is similar.

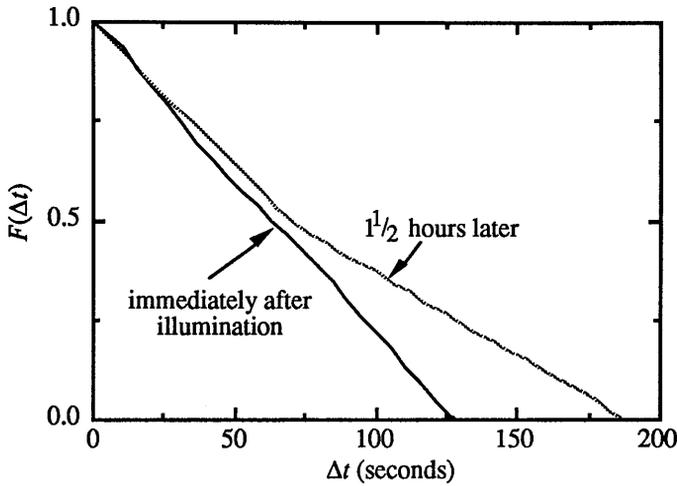


Figure 6.33: Time autocorrelation function for the time-dependent measurements of Fig. 6.32.

The correlation time remains roughly the same $\frac{1}{2}$ hours after the initial illumination.

Since its conducting width is roughly $w_1 \approx 10$ nm, wire P is expected to behave 1-dimensionally (1D) with respect to quantum interference effects (although it is still 3D with respect to electron diffusion). Unfortunately, weak localisation measurements were not made on this material, so the coherence length, L_ϕ , is not well known. However, from the values obtained in §6.2 (albeit for a different material), and from the trend in L_ϕ with conducting width, L_ϕ for wire P would be expected to lie somewhere in the range 100-200 nm. From the mobility, the elastic mean free path is $l \sim 39$ nm. Since $l > w_1$, some size effects are expected.

According to Feng *et al.* (1986), the rms change in conductance expected from the displacement of a scatterer through a distance δr in a wire of length L_ϕ is,

$$\delta G_\phi \sim \frac{e^2}{h} \frac{1}{k_F l} \left(\frac{L_\phi l}{A} \right)^{1/2} \left[1 - \left(\frac{\sin k_F \delta r / 2}{k_F \delta r / 2} \right)^2 \right]^{1/2}, \quad (6.24)$$

where A is the cross-sectional (conducting) area of the wire, and k_F is the Fermi wavevector. Since wire P has a length $L > L_\phi$, the rms fluctuation δG_P is related to δG_ϕ through the averaging relation, Eq. (5.40): $\delta G_P = (L_\phi/L)^{3/2} \delta G_\phi$, where $L = 600$ nm. Eq. (6.24) can be used to obtain an estimate for δr . Take $L_\phi \sim 200$ nm. For a carrier concentration of $9.0 \times 10^{23} \text{ m}^{-3}$, the top surface depletion depth is 33.5 nm, $k_F = 3.0 \times 10^8 \text{ m}^{-1}$, and $k_F l \approx 12$. Thus, solving Eq. (6.24) for $k_F \delta r$ gives, $k_F \delta r \approx 0.34$, or $\delta r \sim 1$ nm. In other words, a scattering centre would need to be displaced roughly 1 nm to give the observed changes in conductance in Fig. 6.32.

Using the defect distributions obtained in Chapter 4, a rough estimate may be made of the number of defects within the length of wire P, which are involved in the measured effects. Take the defect distribution at an etched sidewall to be $\Gamma(x, v\tau/2, \tau)$ [from Eq. (4.14)], where x is the distance from the sidewall, $v = 200 \text{ nm min}^{-1}$ is the etch rate, and $\tau = 30$ s is the etch time. If $w_c \approx 170$ nm is the measured cut-off width for the material, then in the steady state the sidewall depletion depth is roughly $\sim w_c/2$. Under illumination, deep traps are emptied and the sidewall depletion depth is closer to the nominal depletion depth, $x_0 = (2\epsilon_0\epsilon_r\phi_s/eN_D)^{1/2} = 33.5$ nm, where $\phi_s = 0.7$ eV is the surface potential, and $N_D = 9.0 \times 10^{23} \text{ m}^{-3}$ is the doping density. Thus, allowing for the two sidewalls, the total number of defects emptied is roughly,

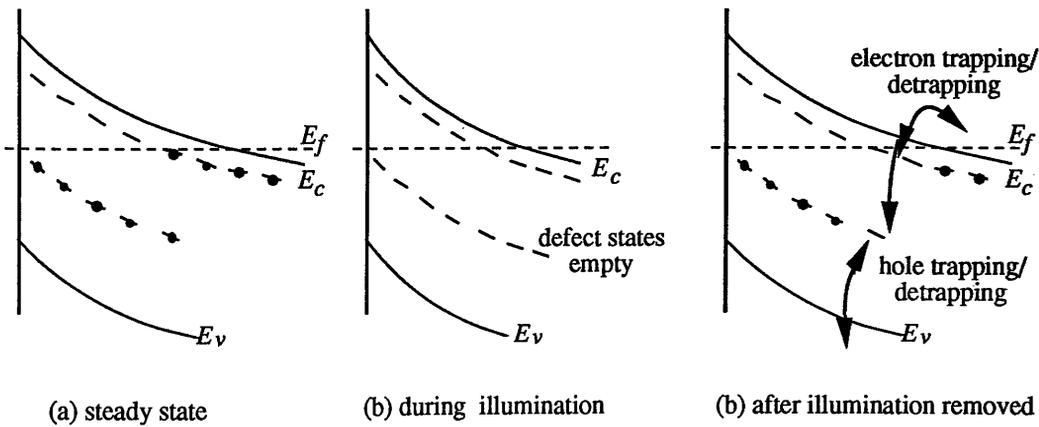


Figure 6.34: Model for the random telegraph signals in response to illumination. The RTS arise from charge trapping/detrapping from defect states at the edge of the depletion region once the illumination has been removed.

$$\begin{aligned}
 N_i &\sim 2(t-x_0)L \int_{x_0}^{w_c/2} dx \Gamma\left(x, \frac{v\tau}{2}, \tau\right) \\
 &= G_1 \tau(t-x_0)L \int_{x_0}^{w_c/2} dx e^{-x/\lambda} \sim 36.
 \end{aligned} \tag{6.25}$$

Here $t = 50$ nm is the epilayer thickness, $L = 600$ nm is the length of wire P, $G_1 = 5.5 \times 10^{23} \text{ m}^{-3} \text{ s}^{-1}$ is the defect creation rate at the sidewall, and $\lambda = 9$ nm is the mean depth to which defects are created (the values of G_1 and λ for SiCl_4 etching at 300 V were obtained in §4.4.3). Eq. (6.25) shows that around 36 defects are involved in giving the effects of Fig. 6.32.

A model can now be built to try to explain the fluctuations observed in the above wires. The two-level defect used in Chapter 4 is assumed to be introduced by the dry-etching. Before illumination, all defect levels below E_f are filled [Fig. 6.34(a)]. During illumination, many of these levels will empty either through electron emission, or hole capture, so the conductance is higher than without illumination [Fig. 6.34(b)]. Once the illumination is removed, many of the ~ 36 defect levels fill within 100-200 seconds. A few defects which are closer to the edge of the depletion region may start trapping and detrapping carriers from the conducting core of the wire in a nonequilibrium manner [Fig. 6.34(c)]. The changes in scattering potential may be local changes in the sidewall depletion depth, which "emulate" the effect of a scatterer moving ~ 1 nm. As one defect traps then detraps an electron, a "pulse" is observed in the measured conductance. These effects can, in principle, continue for ever. In the wider wire R, electrons interact less with the boundaries (see §6.2.4), so the fluctuations are far smaller.

If the times the electron spends in the two states are denoted τ_1 and τ_2 , then assuming the level occupation is determined by a Boltzmann factor gives,

$$\frac{\tau_1}{\tau_2} = \exp\left(-\frac{E_1 - E_2}{k_B T}\right), \tag{6.26}$$

where E_1 and E_2 are the energies of the two levels measured with respect to the Fermi level. As an estimate, the system spends roughly 3-4 times as long in the lower conductance state than it does in the higher conductance state, so $\tau_1 \sim 4\tau_2$. From Eq. (6.26) this implies, $(E_2 - E_1) \sim 19$ meV. This either means two close defect levels separated by this energy, or a defect level that is ~ 19 meV above or below the Fermi level. In one case the electron is moving between two defects, in the other, electrons are being exchanged between the defect and the conducting core of the wire.

Several further measurements would have proved that the above fluctuations have a quantum origin. First, since L_ϕ decreases with increasing temperature, the fluctuations should decay with temperature according to a power law dependent on the dominant phase breaking mechanism. This has been done for example by Bykov *et al.* (1989) who also studied photoconductance fluctuations. Measurement of the UCF trace with respect to magnetic field before and after illumination should, in principle, be different since the scattering configuration will have changed [*e.g.* Davison *et al.* (1990)]. Finally, measurement of the above fluctuations in a magnetic field should yield an rms fluctuation which is a factor of $\sqrt{2}$ less than that measured without the field [*e.g.* Mailly *et al.* (1989)]. This is the same factor of $\sqrt{2}$ reduction mentioned in §5.3.3.

It should also be mentioned that other mechanisms exist whereby defect motion can cause changes in the measured conductance. In one theory, changes in conductivity are brought about by changes in effective scattering cross-section when one defect moves in the vicinity of another. This latter theory applies for the limit $l > L_\phi$ (rather than $L_\phi > l$ as in the UCF theory), and would not be expected to apply for the wires above. See Pelz and Clarke (1987), or Hershfield (1988). For very small samples, the observed conductance changes could also be a (classical) size effect.

6.4.2 Random telegraph signals in Hall geometries

The measurements of conductances in §6.4.2 showed random telegraph signals (RTS) in a two-probe wire. In the current section, RTS is observed not only in the longitudinal, but also in the transverse measurement. The device is similar to the Hall geometry shown in Fig. 6.14. It is made from the $9.0 \times 10^{23} \text{ m}^{-3}$ carrier concentration material, has a physical width $w = 180 \text{ nm}$, centre-to-centre probe spacing $L = 530 \text{ nm}$, and a probe width of 240 nm (*i.e.* these wires allow 4-terminal measurements, and the probes are sufficiently wide that their photoconductance does not significantly affect the measurements). The response of this structure to illumination was measured both in the longitudinal and transverse ("Hall") configurations, at 4 K in zero magnetic field. The primary initial aim was to try to find correlations (or lack of) between these two measurements. To this end, a voltage was applied across the sample, and two lock-ins used to simultaneously monitor the longitudinal and transverse voltages. [A third lock-in to record the current was not available at the time, so the conductances are not accurately known, only the voltages.] Thus, the measuring circuit corresponds to that of Fig. 6.2, with the attenuator and test resistor removed.

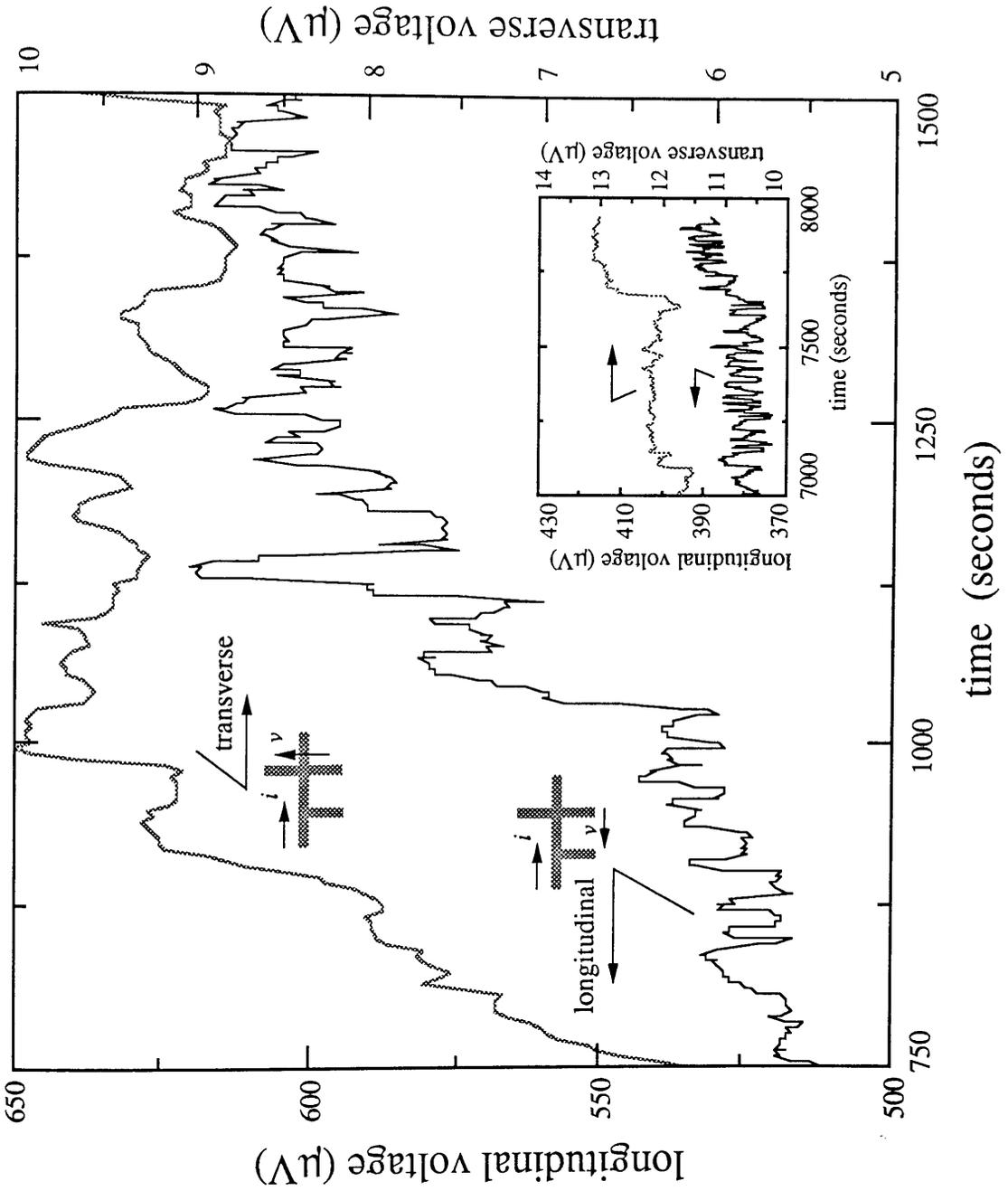


Figure 6.35: Random telegraph signals in simultaneous longitudinal and transverse measurements made in a Hall geometry. The longitudinal measurement shows a pronounced regular structure which continues long after the initial illumination (see inset). The transverse measurement shows fewer steps, and is largely uncorrelated with the longitudinal voltage, indicating that different defects contribute to the two cases.

Fig. 6.35 shows the longitudinal and transverse measurements made after removal of the illumination from the sample (at time $t = 0$). Both show clear steps. The structure in the longitudinal measurement is very distinctive in that the voltage switches between two definite levels with a period of 20-30 s. The new thing in the current set of measurements is that a step structure is also observed in the transverse measurement, where classically no voltage would be expected (for zero magnetic field).

The step structure in the longitudinal measurement is far more regular than that observed in §6.4.1. However, the basic mechanism for the step change is the same, *viz.* a sudden change in the scattering potential. The fact that the conductance switches so regularly between two levels indicates that the scattering potential is also switching regularly between two configurations. From Fig. 6.35 it is evident that the electron spends roughly the same amount of time in either level. In Eq. (6.26) this means $\tau_1 \sim \tau_2$, and hence $E_1 \sim E_2$. In other words, the electron appears to move between two nearby defect levels at the same energy. Every now and again, the electron escapes into the conduction band, leading to a break in the structure. But then another electron becomes trapped, and the RTS starts up again.

The step structure in the transverse measurement is more erratic, showing little correlation with the longitudinal measurement. Al'tshuler and Khmel'nitskii (1985) have pointed out that zero correlation should be expected between transverse and longitudinal UCF measurements. However, in the present case, the steps are due to sudden changes in the scattering potential. Thus, if the *same* defect motion is causing the steps in the longitudinal and transverse voltages, a correlation *is* expected in the times at which the steps occur (although not so in the magnitude or sense). It is clear from Fig. 6.35 that the defects responsible for the longitudinal effect have no influence on the transverse measurement. However, it would be expected that the defects affecting the transverse measurement *should* affect the longitudinal measurement, but whether this is true or not is difficult to judge. The mean transverse voltage is also seen to be offset from zero. This is the effect of a finite "Hall" voltage at $B = 0$ due to fluctuation effects, as mentioned in §6.3.1.

As a final point of interest, it is worth noting that if the RTS in the longitudinal measurement is caused by electron motion in the potential of two defect states, then the electron must be localised *either* near one defect, *or* near the other. This is perhaps an obvious statement to make, until the related problem is considered of an electron in a

double well potential. If the electron always maintains phase coherence (*i.e.* it is in an energy eigenstate), then the time-independent Schrödinger equation may be solved to obtain a wavefunction for the electron in this potential. Now, this wavefunction is spread across both wells, so quantum mechanically speaking it cannot be known whether the electron is in one well or the other. However, the measurement of Fig. 6.35 definitely asserts that the electron is *either* in one defect, *or* in the other. This indicates that the electron must lose phase coherence in moving from one defect to the other (*i.e.* there is some time dependent perturbation acting on the electron state), so presumably this motion is phonon-assisted. The mean lifetime of the electron in one of these defects is then an indication of the strength of the electron-phonon-defect coupling.

6.5 References

- ABRAHAMS E, P W Anderson, P A Lee, T V Ramakrishnan, "Quasiparticle lifetime in disordered two-dimensional metals," *Phys. Rev. B* **24**, 6783 (1981).
- AL'TSHULER B L, A G Aronov, "Magnetoresistance of thin films and of wires in a longitudinal magnetic field," *JETP Lett.* **33**, 499 (1981a).
- AL'TSHULER B L, A G Aronov, "Electron density of states and energy relaxation time in magnetic fields," *Solid State Comm.* **38**, 11 (1981b).
- AL'TSHULER B L, A G Aronov, D E Khmel'nitskii, "Effects of electron-electron collisions with small energy transfers on quantum localisation," *J. Phys. C* **15**, 7367 (1982).
- AL'TSHULER B L, D E Khmel'nitskii, "Fluctuation properties of small conductors," *JETP Lett.* **42**, 359 (1985).
- AL'TSHULER B L, B Z Spivak, "Variation of the random potential and the conductivity of samples of small dimensions," *JETP Lett.* **42**, 447 (1985).
- BARANGER H U, A D Stone, D P DiVincenzo, "Resistance fluctuations in multiprobe microstructures: length dependence and nonlocality," *Phys. Rev. B* **37**, 6521 (1988).
- BEENAKKER C W J, H van Houten, "Boundary scattering and weak localisation of electrons in a magnetic field," *Phys. Rev. B* **38**, 3232 (1988a).
- BEENAKKER C W J, H van Houten, "Flux-cancellation effect on narrow-channel magnetoresistance fluctuations," *Phys. Rev. B* **37**, 6544 (1988b).

- BENOIT A D, S Washburn, C P Umbach, R B Laibowitz, R A Webb, "Asymmetry in the magnetoconductance of metal wires and loops," *Phys. Rev. Lett.* **57**, 1765 (1986).
- BENOIT A, C P Umbach, R B Laibowitz, R A Webb, "Length-independent voltage fluctuations in small devices," *Phys. Rev. Lett.* **58**, 2343 (1987).
- BIRD J P, A D C Grassie, M Lakrimi, K M Hutchings, J J Harris, C T Foxon, "Conductance fluctuations and non-diffusive motion in GaAs/AlGaAs heterojunction wires," *J. Phys.: Condensed Matter* **2**, 7847 (1990).
- BLACK J L, B L Gyorffy, J Jäckle, "On the resistivity due to two-level systems in metallic glasses," *Phil. Mag. B* **40**, 331 (1979).
- BLAKEMORE J S, "Semiconducting and other major properties of gallium arsenide," *J. Appl. Phys.* **53**, R123 (1982).
- BÜTTIKER M, "Four-terminal phase-coherent conductance," *Phys. Rev. Lett.* **57**, 1761 (1986).
- BYKOV A A, G M Gusev, Z D Kvon, D I Lubyshev, V P Migal, "Interband photoconductivity of a mesoscopic GaAs sample," *JETP Lett.* **49**, 135 (1989).
- COBDEN D H, N K Patel, M Pepper, D A Ritchie, J E F Frost, G A C Jones, "Noise and reproducible structure in a GaAs/AlGaAs one-dimensional channel," submitted (1991).
- DAVISON M, R P Taylor, P C Main, P H Beton, S P Beaumont, C D Wilkinson, "Mesoscopic charge mapping by conductance fluctuations," *Physica B* **165&166**, 865 (1990).
- DiVINCENZO D P, C L Kane, "Voltage fluctuations in mesoscopic metal rings and wires," *Phys. Rev. B* **38**, 3006 (1988).
- DUGAEV V K, D E Khmel'nitskii, "Magnetoresistance of metal films with low impurity concentrations in a parallel magnetic field," *Sov. Phys. JETP* **59**, 1038 (1984).
- FENG S, P A Lee, A D Stone, "Sensitivity of the conductance of a disordered metal to the motion of a single atom: implications for $1/f$ noise," *Phys. Rev. Lett.* **56**, 1960, 2772(E) (1986).
- FUCHS K, "The conductivity of thin metallic films according to the electron theory of metals," *Proc. Camb. Phil. Soc.* **34**, 100 (1938).
- GEIM A K, P C Main, P H Beton, P Streda, L Eaves, C D W Wilkinson, S P Beaumont, "A new non-local magnetoresistance effect at the crossover between the classical and quantum transport regimes," submitted (1991).

- HANSEN W, T P Smith III, D P DiVincenzo, K Y Lee, Y H Lee, R Cheung, "Quantum fluctuations in n^+ GaAs wires with invasive probes," *Surf. Sci.* **229**, 321 (1990).
- HAUKE H, S Washburn, A D Benoit, C P Umbach, R A Webb, "Universal scaling of nonlocal and local resistance fluctuations in small wires," *Phys. Rev. B* **41**, 12454 (1990).
- HERAUD A P, S P Beaumont, C D W Wilkinson, P C Main, J R Owers-Bradley, L Eaves, "Inelastic electron scattering in pseudo-one-dimensional metal wires," *J. Phys. C* **20**, L249 (1987).
- HERSHFIELD S, "Sensitivity of the conductance to impurity configuration in the clean limit," *Phys. Rev. B* **37**, 8557 (1988).
- van HOUTEN H, B J van Wees, M G J Heijman, J P André, "Submicron conducting channels defined by shallow mesa etch in GaAs-AlGaAs heterojunctions," *Appl. Phys. Lett.* **49**, 1781 (1986).
- van HOUTEN H, C W J Beenakker, B J van Wees, J E Mooij, "Boundary scattering modified one-dimensional weak localisation in submicron GaAs/AlGaAs heterostructures," *Surf. Sci.* **196**, 144 (1988).
- van HOUTEN H, C W J Beenakker, J G Williamson, M E I Broekaart, P H M van Loosdrecht, B J van Wees, J E Mooij, C T Foxon, J J Harris, "Coherent electron focusing with quantum point contacts in a two-dimensional electron gas," *Phys. Rev. B* **39**, 8556 (1989).
- KANE C L, P A Lee, D P DiVincenzo, "Resistance fluctuations in multilead devices," *Phys. Rev. B* **38**, 2995 (1988).
- KAVEH M, N Wiser, "Electron-electron scattering in conducting materials," *Adv. Phys.* **33**, 257 (1984).
- LANDAU L D, E M Lifshitz, *Electrodynamics of Continuous Media*, Pergamon Press, Oxford (1960).
- LEE P A, A D Stone, "Universal conductance fluctuations in metals," *Phys. Rev. Lett.* **55**, 1622 (1985).
- LEE P A, A D Stone, H Fukuyama, "Universal conductance fluctuations in metals: effects of finite temperature, interactions, and magnetic field," *Phys. Rev. B* **35**, 1039 (1987).
- LERNER I V, "Rigorous perturbation results in the theory of mesoscopic fluctuations: distribution functions and time-dependent phenomena," in ed. B Kramer, *Quantum Coherence in Mesoscopic Systems*, Plenum Press, New York (1991).

- MA M, P A Lee, "Transverse conductance fluctuations and magnetic field asymmetry," *Phys. Rev. B* **35**, 1448 (1987).
- MAEKAWA S, Y Isawa, H Ebisawa, "Conductance fluctuations in small metallic wires," *J. Phys. Soc. Jpn.* **56**, 25 (1987).
- MAILLY D, M Sanquer, J L Pichard, P Pari, "Reduction of quantum noise in a GaAlAs/GaAs heterojunction by a magnetic field: an orthogonal-to-unitary Wigner statistics transition," *Europhys. Lett.* **8**, 471 (1989).
- MAILLY D, M Sanquer, "Reduction of the quantum noise in a GaAlAs/GaAs heterojunction and n^+ GaAs wires of mesoscopic size by a magnetic field," *Surf. Sci.* **229**, 260 (1990).
- MAILLY D, M Sanquer, "Sensitivity of quantum conductance fluctuations to time reversal symmetry," in ed. B Kramer, *Quantum Coherence in Mesoscopic Systems*, Plenum Press, New York (1991).
- MAIN P C, R P Taylor, L Eaves, S Thoms, S P Beaumont, C D W Wilkinson, "Conduction in n^+ -GaAs wires," in ed. S P Beaumont, C M Sotomajor-Torres, *Science and Engineering of One- and Zero-Dimensional Semiconductors*, Plenum Press, New York (1990).
- NAKAMURA K, D C Tsui, F Nihey, H Toyoshima, T Itoh, "Electron focusing with multiparallel one-dimensional channels made by focused ion beam," *Appl. Phys. Lett.* **56**, 385 (1990).
- NIHEY F, K Nakamura, M Kuzuhara, N Samoto, T Itoh, "Electron focusing with multiparallel GaAs-AlGaAs wires defined by damageless processing," *Appl. Phys. Lett.* **57**, 1218 (1990).
- OCHIAI Y, K Ishibashi, S Ishida, M Mizuno, K Gamo, M Kawabe, K Murase, S Namba, "Time dependent resistance changes in a narrow AlGaAs-GaAs wire," *Superlatt. Microstruct.* **6**, 337 (1989).
- PELZ J, J Clarke, "Quantitative 'local-interference' model for $1/f$ noise in metal films," *Phys. Rev. B* **36**, 4479 (1987).
- PIPPARD A B, *Magnetoresistance in Metals*, Cambridge University Press, Cambridge (1989).
- RALLS K S, W J Skoçpol, L D Jackel, R E Howard, L A Fetter, R W Epworth, D M Tennant, "Discrete resistance switching in submicrometer silicon inversion layers: individual interface traps and low-frequency ($1/f$?) noise," *Phys. Rev. Lett.* **52**, 228 (1984).

- TAYLOR R P, M L Leadbeater, G P Whittington, P C Main, L Eaves, S P Beaumont, I McIntyre, S Thoms, C D W Wilkinson, "Universal conductance fluctuations in the magnetoresistance of submicron-size n^+ -GaAs wires and laterally confined n^- -GaAs/(AlGa)As heterostructures," *Surf. Sci.* **196**, 52 (1988a).
- TAYLOR R P, P C Main, L Eaves, S P Beaumont, S Thoms, C D W Wilkinson, "Aperiodic conductance fluctuations as a probe of changes in the microscopic scattering configuration in n^+ -GaAs:Si wires," in ed. W Zawadzki, *Proc. 19th Int. Conf. on the Physics of Semiconductors*, Institute of Physics, Polish Academy of Sciences, Warsaw (1988b).
- THORNTON T J, M Pepper, H Ahmed, D Andrews, G J Davies,, "One-dimensional conduction in the 2D electron gas of a GaAs-AlGaAs heterojunction," *Phys. Rev. Lett.* **56**, 1198 (1986).
- THORNTON T J, M L Roukes, A Scherer, B P van der Gaag, "Magnetoresistance and boundary scattering in ballistic wires," in ed. B Kramer, *Quantum Coherence in Mesoscopic Systems*, Plenum Press, New York (1991).
- THOULESS D J, "Maximum metallic resistance in thin wires," *Phys. Rev. Lett.* **39**, 1167 (1977).
- SKOŹPOL W J, P M Mankiewich, R E Howard, L D Jackel, D M Tennant, A D Stone, "Nonlocal potential measurements of quantum conductors," *Phys. Rev. Lett.* **58**, 2347 (1987).
- UMBACH C P, S Washburn, R B Laibowitz, R A Webb, "Magnetoresistance of small quasi-one-dimensional, normal metal rings and lines," *Phys. Rev. B* **30**, 4048 (1984).
- WASHBURN S, C P Umbach, R B Laibowitz, R A Webb, "Temperature dependence of the normal-metal Aharonov-Bohm effect," *Phys. Rev. B* **32**, 4789 (1985).
- WASHBURN S, "Resistance fluctuations in small samples: be careful when playing with Ohm's Law," in ed. B Kramer, *Quantum Coherence in Mesoscopic Systems*, Plenum Press, New York (1991).
- WHITTINGTON G P, P C Main, L Eaves, R P Taylor, S Thoms, S P Beaumont, C D W Wilkinson, C R Stanley, J Frost, "Universal conductance fluctuations in the magnetoresistance of submicron n^+ GaAs wires," *Superlatt. Microstruct.* **2**, 381 (1986).
- ZIMAN J M, *Electrons and Phonons*, Clarendon Press, Oxford (1960).

Summary and conclusions

The work described in this thesis has focussed on electrical transport in dry-etched n^+ -GaAs wires, both at room temperature and at much lower temperatures. The wires were fabricated using electron-beam lithography and SiCl_4 reactive-ion etching. The materials used comprised heavily Si-doped GaAs epilayers grown by molecular beam epitaxy. The narrowest wires studied had widths of about 70-80 nm.

Characterisation of the dry-etch damage in the sidewalls of these wires was done in Chapter 3, by measuring the wire conductance as a function of the wire width. Materials with carrier concentrations $6.5 \times 10^{24} \text{ m}^{-3}$ and $6.5 \times 10^{23} \text{ m}^{-3}$ were used, with the etching being done at energies between 280-380 eV. The minimum conducting width (the cut-off width) in dry-etched wires was found to be larger than the corresponding width for wet-etched wires. The "extra" sidewall depletion was attributed to sidewall damage incurred as a result of the etching.

Sidewall damage was found to be insensitive to a variation in the d.c. self bias of the etch over a range 280-380 V for an etch time of 30 s, using the $6.5 \times 10^{24} \text{ m}^{-3}$ material. Over this range, the width-conductance curve remained unchanged within experimental error. However, wires etched for different times were found to give very different width-conductance curves, the cut-off width increasing from about 160 nm for a 36 s etch, to about 250 nm for a 90 s etch, using the $6.5 \times 10^{23} \text{ m}^{-3}$ material. For

equivalent etch times the ratio of the measured cut-off width to the nominal cut-off width (as calculated from the carrier concentration) was found to be roughly independent of the carrier concentration (which was varied by an order of magnitude).

A model for the dry-etch damage has been developed in Chapter 4, which describes very well the above measurements on n^+ -wires. The dry-etch damage is assumed to comprise charged defects. Initially, defect distributions for top surface and sidewall damage were obtained by solving the one- and two-dimensional diffusion equations. Two cases were examined. In one, defects are created solely at the surface being etched, following which they migrate into the structure by diffusion. In the other, defects are created a small distance from the surface, following which they can again diffuse. Fits to DLTS-measured defect profiles indicated that diffusion affects the form of the defect profiles significantly only for unrealistically high values of defect diffusion coefficient. For values of diffusion coefficient smaller than about $10^{-20} \text{ m}^2 \text{ s}^{-1}$, the derived expressions tend to limiting forms.

Very simple analytical formulae for top surface and sidewall defect distribution were obtained by assuming an exponential defect source function, and ignoring defect diffusion altogether. The theory has two parameters, the defect creation rate at the surface being etched, and the mean depth to which defects are created. The defect distribution for an etched sidewall was then used to calculate the conductances of dry-etched wires. This was done by assuming a two-level defect, with the lower level being permanently occupied. Numerical solutions of the two-dimensional Poisson equation then gave the required wire conductances. It was found that using a single source function gave an excellent fit to all of the width-conductance measurements made previously. The fit worked despite variations in carrier concentration, epilayer thickness and etch time/depth. It thus appears as if the source function is characteristic of the etch process being used.

The analysis seems to show that sidewall damage is created at a significant rate (of the same order of magnitude) compared to top surface damage as measured by DLTS. Since bombarding ions strike the sidewall at a predominantly grazing angle, whereas the top surface is struck head-on, it would be expected that the creation rate at sidewalls should be much less. The fact that it is not indicates the presence of an additional bombarding flux. This is thought to be the flux of ricochet and sputtered material leaving the surface of the etched GaAs with significant energies. These damage creation processes should be much reduced with low energy etching, for example using electron

cyclotron resonance technology.

The low temperature quantum transport properties of n^+ -wires have also been measured, as described in Chapter 6. Reproducible fluctuations were observed in longitudinal, transverse and nonlocal measurement configurations. These are all caused by interference between the many electron trajectories existing in the sample. The observed effects were analysed using the theories of weak localisation and universal conductance fluctuations reviewed in Chapter 5.

The scaling of the fluctuations in the longitudinal measurement with length suggested that the wires were 1-dimensional (1D) with respect to interference effects (even though they have a 3D Fermi surface). Fits to weak localisation also indicated 1D behaviour. Wires wider than ~ 130 nm showed the largest fluctuations, with narrower wires giving progressively smaller fluctuations. This was shown to be due to a decreasing coherence length with decreasing wire width. The fluctuation amplitude for the wider wires, and analysis of the autocorrelation function, indicated a phase coherence length ≥ 500 nm. A more precise determination could not be made since the thermal diffusion length in these wires had a similar value. The narrower wires gave a phase coherence length ~ 300 nm, independent of the thermal diffusion length.

The narrowest wire studied had a conducting width of ~ 30 nm. Since the elastic mean free path in the material is ~ 50 nm, the issue of boundary scattering was raised. The weak localisation observed in this wire was fitted to the theory of Beenakker and van Houten, which is a modification of the usual Al'tshuler-Aronov theory to allow for the presence of the boundaries. Fits of the experimental curves to the theory indicated that the scattering from the edge of the depletion region in these wires is predominantly diffuse.

Measurements of weak localisation in this wire at temperatures in the range 4-105 K gave the corresponding phase breaking rates. There were three separate contributions to the phase breaking rate: an electron-phonon term, an electron-electron term, and a temperature independent term. Two different electron-electron scattering rates were considered: the disorder modified electron-electron scattering rate, and the rate due to electron-electron collisions with small energy transfers (the Nyquist rate). It was found that the experimental points could be fitted to either mechanism equally well. Theory asserts that one or other should be the dominant electron-electron phase breaking mechanism, but this could not be discerned from the measurements.

The fact that the coherence length decreases for narrower wires suggests that

boundary scattering may be the cause of the temperature independent phase breaking rate observed above. This has been discussed in the context of the model developed for dry-etch damage in Chapter 4. One possibility is magnetic scattering between the conduction electrons and unpaired electrons trapped in defect states near the edge of the depletion region. Such a model is consistent with all the observation made above, but a more systematic study is required to definitely confirm this.

A brief study has also been made of fluctuations in the Hall voltage. Measurements on Hall geometries indicated that voltage probes placed in a Hall configuration do in fact measure voltage fluctuations associated with a phase coherent region. This was affirmed by measurements made at a cross-junction, where the magnitude of the voltage fluctuations was shown to be independent of which pair of leads was used to inject current, and which pair was used as the voltage probes.

Longitudinal measurements on the narrowest wire (~30 nm conducting width) showed long period fluctuations as well as much shorter period fluctuations. The long period fluctuations were attributed to interference between electron trajectories traversing the length of the wire. Hall measurements on this wire gave fluctuations which were very similar to the small fluctuations observed in the longitudinal measurement. It was argued that both, in fact, had the same origin, *viz.* interference between electron trajectories in the vicinity of the voltage probes. In most longitudinal measurements the contribution of electron trajectories travelling the length of the wire, and those confined to the region of the voltage probes, can only be deconvoluted by making more than one measurement. However, in the present case since the main wire was much narrower than the voltage probes, the two components were manifestly separated in the same longitudinal measurement.

Low temperature measurements have also been made of the time-dependent response of narrow n^+ -wire to brief periods of illumination. A two-probe wire showed pulses in the conductance one the illumination had been removed. These random telegraph signals were attributed to charge trapping and detrapping from defect levels at the edge of the depletion region. Using the expression for sidewall defect density obtained in Chapter 4, it was shown that only a handful of defects were involved in giving the observed effects.

Random telegraph signals were also observed in simultaneous longitudinal and transverse measurements on a 5-probe wire in a zero magnetic field. Very pronounced regular structure was observed in the longitudinal measurement. The steps occurring

the transverse measurement were less frequent. The lack of correlation between the longitudinal and transverse measurements indicated that two set of defects were probably responsible for the measured curves.

Thus, electrical transport has been characterised for dry-etched n^+ -GaAs wires. The fact that a quantitative theory has been found, which describes remarkably well the effects of dry-etching on the conductance of samples, may prove useful in other areas of nanofabrication where dry-etching is involved. The study could beneficially be extended to other etch gases and dry-etch methods. In this case, a quantitative analysis could be carried out, for example, on the effects of damage as a result of the dry-recessing of nanometre-scale MESFETs, or on the effects of the non-radiative recombination centre distribution on the luminescence of quantum dots and wires.

The quantum transport measurements on the n^+ -structures have little significance for practical quantum electronic devices. However, they do illustrate some of the difficulties that must be overcome if reliable (wholly deterministic) devices are to be made. Fluctuation phenomena are observed in many semiconductor nanostructures, and are currently viewed as a serious impediment to the successful exploitation of other quantum interference phenomena, such as Aharonov-Bohm type effects. On the positive side, the study of fluctuation phenomena has led to a much better understanding of transport in weakly disordered mesoscopic systems.

Appendix

A. Derivation of defect distributions

A.1 Derivation of top surface damage distribution

The equation that will be solved here is,

$$\left(D \frac{\partial^2}{\partial y^2} - \frac{\partial}{\partial t} \right) \Gamma(y,t) = -G_1 e^{-(y-vt)/\lambda}; \quad y \geq vt, \quad t \geq 0, \quad (\text{A.1a})$$

$$\Gamma(y,0) = 0, \quad -D \left(\frac{\partial \Gamma}{\partial y} \right)_{y=vt} = F_1, \quad \lim_{y \rightarrow \infty} \Gamma(y,t) = 0. \quad (\text{A.1b})$$

This is treated as a mathematical problem, but the equation is not reduced to dimensionless form. The domain of the problem, $y \geq vt, t \geq 0$, does not allow the Green function method to be easily applied. Instead the problem is solved by the method of Laplace transforms. To apply a Laplace transform in time, the domain $y \geq 0, t \geq 0$ is required. This is achieved by using the co-ordinate transformation $(y,t) \rightarrow (y-vt,t)$ which gives the drift-diffusion equation,

$$\left(D \frac{\partial^2}{\partial y^2} + v \frac{\partial}{\partial y} - \frac{\partial}{\partial t} \right) \Gamma(y,t) = -G_1 e^{-y/\lambda}; \quad y \geq 0, \quad t \geq 0, \quad (\text{A.2a})$$

$$\Gamma(y,0) = 0, \quad -D \left(\frac{\partial \Gamma}{\partial y} \right)_{y=0} = F_1, \quad \lim_{y \rightarrow \infty} \Gamma(y,t) = 0. \quad (\text{A.2b})$$

Multiplying both sides of each equation in (A.2) by e^{-st} , and integrating with respect to time from 0 to ∞ (and using integration by parts to transform the time derivative):

$$\left(D \frac{\partial^2}{\partial y^2} + v \frac{\partial}{\partial y} - s\right) \tilde{\Gamma}(y,s) = -\frac{G_1}{s} e^{-y/\lambda}; \quad y \geq 0, \quad (\text{A.3a})$$

$$-D \left(\frac{\partial \tilde{\Gamma}}{\partial y}\right)_{y=0} = \frac{F_1}{s}, \quad \lim_{y \rightarrow \infty} \tilde{\Gamma}(y,s) = 0, \quad (\text{A.3b})$$

where,

$$\tilde{\Gamma}(y,s) = \int_0^\infty \Gamma(y,t) e^{-st} dt. \quad (\text{A.4})$$

is the Laplace transform of $\Gamma(y,t)$. Eq. (A.3a) is a second order ordinary differential equation in y , and may be solved as such. The general solution is given by the sum of the complementary function and the particular integral. The complementary function is a solution to the associated homogeneous problem (source term set to zero). Using a trial solution $\tilde{\Gamma}_{cf} = Ae^{-ay}$ in Eq. (A.3a), with $G_1 = 0$, gives the quadratic equation,

$$Da^2 - va - s = 0. \quad (\text{A.5})$$

Solving for a gives,

$$\tilde{\Gamma}_{cf} = A \exp\left[-\frac{y}{2D} (v + \sqrt{v^2 + 4Ds})\right]. \quad (\text{A.6})$$

Only the positive root of Eq. (A.5) has been retained in the complementary function. The negative root is discarded by the requirement that $\tilde{\Gamma}$ be bounded at infinity [second condition in Eq. (A.3b), which is actually more stringent than simply boundedness]. The particular integral for Eq. (A.3) is given by any particular solution, obtained by any means. Using the trial solution $\tilde{\Gamma}_{pi} = Be^{-y/\lambda}$ in Eq. (A.3a), this time retaining the source term, gives,

$$\frac{DB}{\lambda^2} - \frac{vB}{\lambda} - Bs = -\frac{G_1}{s}. \quad (\text{A.7})$$

Solving for B gives the particular integral,

$$\tilde{\Gamma}_{pi} = \frac{G_1}{s(s+\omega)} e^{-y/\lambda}, \quad (\text{A.8})$$

where,

$$\omega = \frac{v}{\lambda} - \frac{D}{\lambda^2}. \quad (\text{A.9})$$

Thus, the most general solution to Eq. (A.3a), which is also bounded at infinity, is,

$$\tilde{\Gamma}(y,s) = \tilde{\Gamma}_{cf} + \tilde{\Gamma}_{pi} = A \exp\left[-\frac{y}{2D} (v + \sqrt{v^2 + 4Ds})\right] + \frac{G_1}{s(s+\omega)} e^{-y/\lambda}. \quad (\text{A.10})$$

The parameter A is found by applying the boundary condition on $\partial \tilde{\Gamma} / \partial y$ [first condition in Eq. (A.3b)]. This gives the unique solution to Eq. (A.3), which may be written as the sum of three terms:

$$\tilde{\Gamma} = \tilde{\Gamma}_1 + \tilde{\Gamma}_2 + \tilde{\Gamma}_3, \quad (\text{A.11})$$

where,

$$\tilde{\Gamma}_1(y,s) = \frac{G_1}{s(s+\omega)} e^{-y/\lambda} , \quad (\text{A.11a})$$

$$\tilde{\Gamma}_2(y,s) = \frac{2F_1}{s(v+\sqrt{v^2+4Ds})} \exp\left[-\frac{y}{2D}(v+\sqrt{v^2+4Ds})\right] , \quad (\text{A.11b})$$

$$\tilde{\Gamma}_3(y,s) = -\frac{2DG_1}{\lambda s(s+\omega)(v+\sqrt{v^2+4Ds})} \exp\left[-\frac{y}{2D}(v+\sqrt{v^2+4Ds})\right] . \quad (\text{A.11c})$$

Each of these terms may be inverted in turn. Eq. (A.11a) is easily inverted by applying the Laplace inversion formula, and integrating using the method of residues (alternatively using tables):

$$\Gamma_1(y,t) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \tilde{\Gamma}_1(y,s) e^{st} ds = \frac{G_1}{\omega} e^{-y/\lambda} [1-e^{-\omega t}] . \quad (\text{A.12})$$

The second term is inverted with a little help from the convolution theorem. Applying the inversion formula, and rearranging:

$$\begin{aligned} \Gamma_2(y,t) &= \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \tilde{\Gamma}_2(y,s) e^{st} ds \quad (\text{A.13}) \\ &= \frac{F_1}{\sqrt{D}} e^{-yv/2D} \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} ds e^{st} \left(\frac{v}{2\sqrt{D}} + \sqrt{\frac{v^2}{4D} + s}\right)^{-1} \frac{1}{s} \exp\left[-\frac{y}{\sqrt{D}} \sqrt{\frac{v^2}{4D} + s}\right] . \end{aligned}$$

Now, using the change of variables,

$$u = \frac{v^2}{4D} + s, \quad (\text{A.14})$$

gives,

$$\begin{aligned} \Gamma_2(y,t) &= \frac{F_1}{\sqrt{D}} e^{-yv/2D} e^{-v^2t/4D} \frac{1}{2\pi i} \int_{c'-i\infty}^{c'+i\infty} du e^{ut} \frac{1}{(u-v^2/4D)} \frac{e^{-y\sqrt{u/D}}}{(v/2\sqrt{D} + \sqrt{u})} \\ &= \frac{F_1}{\sqrt{D}} e^{-yv/2D} e^{-v^2t/4D} \frac{1}{2\pi i} \int_{c'-i\infty}^{c'+i\infty} du e^{ut} \tilde{F}(u) \tilde{G}(u) , \quad (\text{A.15}) \end{aligned}$$

where,

$$\tilde{F}(u) = \frac{1}{(u-v^2/4D)} , \quad (\text{A.15a})$$

$$\tilde{G}(u) = \frac{e^{-y\sqrt{u/D}}}{(v/2\sqrt{D} + \sqrt{u})} . \quad (\text{A.15b})$$

$\tilde{F}(u)$ is inverted by integrating using the method of residues:

$$f(t) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \tilde{F}(u) e^{ut} du = e^{v^2t/4D} . \quad (\text{A.16})$$

$\tilde{G}(u)$ may be inverted using a result from Abramowitz and Stegun (1965), formula

(29.3.88), which states,

$$\tilde{H}(s) = \frac{e^{-k\sqrt{s}}}{a+\sqrt{s}} \Leftrightarrow h(t) = \frac{e^{-k^2/4t}}{\sqrt{\pi t}} - a e^{ak} e^{a^2 t} \operatorname{erfc}\left(a\sqrt{t} + \frac{k}{2\sqrt{t}}\right), \quad (\text{A.17})$$

where erfc is the complementary error function. This gives,

$$g(t) = \frac{e^{-y^2/4Dt}}{\sqrt{\pi t}} - \frac{v}{2\sqrt{D}} e^{yv/2D} e^{v^2 t/4D} \operatorname{erfc}\left(\frac{v}{2}\sqrt{\frac{t}{D}} + \frac{y}{2\sqrt{Dt}}\right). \quad (\text{A.18})$$

Applying the convolution theorem,

$$\frac{1}{2\pi i} \int_{c'-i\infty}^{c'+i\infty} du e^{ut} \tilde{F}(u) \tilde{G}(u) = \int_0^t dt' f(t-t') g(t'), \quad (\text{A.19})$$

finally gives,

$$\Gamma_2(y,t) = F_1 \int_0^t dt' \left\{ \frac{1}{\sqrt{\pi Dt'}} e^{-(y+vt')^2/4Dt'} - \frac{v}{2D} \operatorname{erfc}\left(\frac{y+vt'}{2\sqrt{Dt'}}\right) \right\}. \quad (\text{A.20})$$

The procedure used for Γ_2 may be repeated to invert $\tilde{\Gamma}_3$. Applying the inversion formula, and using the change of variables (A.14):

$$\begin{aligned} \Gamma_3(y,t) &= -\frac{\sqrt{D} G_1}{\lambda} e^{-yv/2D} e^{-v^2 t/4D} \frac{1}{2\pi i} \int_{c'-i\infty}^{c'+i\infty} \frac{du e^{ut}}{(u-v^2/4D)(u-v^2/4D+\omega)} \frac{e^{-y\sqrt{u/D}}}{(v/2\sqrt{D}+\sqrt{u})} \\ &= -\frac{\sqrt{D} G_1}{\lambda} e^{-yv/2D} e^{-v^2 t/4D} \frac{1}{2\pi i} \int_{c'-i\infty}^{c'+i\infty} du e^{ut} \tilde{F}(u) \tilde{G}(u), \end{aligned} \quad (\text{A.21})$$

where,

$$\tilde{F}(u) = \frac{1}{(u-v^2/4D)(u-v^2/4D+\omega)}, \quad (\text{A.21a})$$

$$\tilde{G}(u) = \frac{e^{-y\sqrt{u/D}}}{(v/2\sqrt{D}+\sqrt{u})}. \quad (\text{A.21b})$$

Integrating by the method of residues,

$$f(t) = \frac{1}{\omega} e^{v^2 t/4D} [1 - e^{-\omega t}], \quad (\text{A.22})$$

and $g(t)$ is given by Eq. (A.18) again. The convolution theorem now yields,

$$\Gamma_3(y,t) = -\frac{DG_1}{\omega\lambda} \int_0^t dt' [1 - e^{-\omega(t-t')}] \left\{ \frac{1}{\sqrt{\pi Dt'}} e^{-(y+vt')^2/4Dt'} - \frac{v}{2D} \operatorname{erfc}\left(\frac{y+vt'}{2\sqrt{Dt'}}\right) \right\}, \quad (\text{A.23})$$

Thus, the desired solution of Eq. (A.1) is given by,

$$\Gamma(y,t) = \Gamma_1(y,t) + \Gamma_2(y,t) + \Gamma_3(y,t). \quad (\text{A.24})$$

The solution to the problem where defects are created solely at the surface, Eq. (4.5), is found by taking $G_1 = 0$, and using the replacements $y \rightarrow Y$ and $t \rightarrow \tau$. The solution to the problem where defect creation occurs within the material, Eq. (4.7), is found by

taking $F_1 = 0$, and again using the above replacements.

A.2 Derivation of sidewall damage distribution

The problem to be solved here may be written in the general form,

$$\left(D\nabla^2 - \frac{\partial}{\partial t}\right)\Gamma(\mathbf{r},t) = -g(\mathbf{r},t); \quad x \geq 0, y \geq 0, \quad (\text{A.25a})$$

$$\Gamma(\mathbf{r},0) = 0, \quad -D\left(\frac{\partial\Gamma}{\partial x}\right)_{x=0} = f(y,t), \quad -D\left(\frac{\partial\Gamma}{\partial y}\right)_{y=0} = 0, \quad \lim_{r \rightarrow \infty} \Gamma(\mathbf{r},t) = 0, \quad (\text{A.25b})$$

where $\mathbf{r} = (x,y)$. This will be solved using the Green function method [see Mackie (1989)]. Usually, this entails inverting equations such as (A.25a) by some transform method to give the solution in the form,

$$\Gamma \sim - \int Gg + \text{boundary terms}, \quad (\text{A.26})$$

where G is the Green function for the problem. In the present case, the diffusion operator is not self-adjoint, and a variant of the above technique must be used. It turns out that the solution to Eq. (A.25) can indeed be written in the form (A.26), but with the Green function, G , replaced by the adjoint Green function, G^+ . This adjoint Green function will turn out to have the same functional form as G , but will satisfy different boundary conditions.

To find the adjoint Green function, G must first be found for the associated problem with zero boundary conditions. The Green function for the quarter plane $x \geq 0, y \geq 0$ is most easily found by first working out the Green function for the infinite domain, $-\infty < x,y < \infty$, and then using the method of images (it can also be found using the Fourier cosine transform directly on the quarter plane). This latter Green function is specified by,

$$\left(D\nabla_r^2 - \frac{\partial}{\partial t}\right)G_0(\mathbf{r},t;\mathbf{r}',t') = \delta^2(\mathbf{r}-\mathbf{r}') \delta(t-t'); \quad -\infty < x,y < \infty, \quad (\text{A.27a})$$

$$\lim_{r \rightarrow \infty} G_0 = 0, \quad \lim_{r \rightarrow \infty} (\nabla_r G_0) = 0. \quad (\text{A.27b})$$

Multiplying both sides of Eq. (A.27) by $e^{-i\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')}$, and integrating with respect to \mathbf{r} (with an integration by parts to transform the Laplacian) gives the transformed equation,

$$\left(-Dk^2 - \frac{\partial}{\partial t}\right)\tilde{G}_0(\mathbf{k};t,t') = \delta(t-t'), \quad (\text{A.28})$$

where,

$$\tilde{G}_0(\mathbf{k};t,t') = \int d^2\mathbf{r} G_0(\mathbf{r},t;\mathbf{r}',t') e^{-i\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')} , \quad (\text{A.29})$$

is the Fourier transform of G_0 , and the integration is performed over all space. Eq. (A.28) is a first order ordinary differential equation:

$$\frac{\partial \tilde{G}_0}{\partial t} + Dk^2 \tilde{G}_0 = -\delta(t-t') . \quad (\text{A.30})$$

The left hand side of this may be written as a total derivative by multiplying the equation by the integrating factor $e^{Dk^2 t}$:

$$\frac{d}{dt} (\tilde{G}_0 e^{Dk^2 t}) = -e^{Dk^2 t} \delta(t-t') , \quad (\text{A.31})$$

from which,

$$\tilde{G}_0 = -e^{-Dk^2(t-t')} . \quad (\text{A.32})$$

Hence the Green function for the infinite domain is,

$$\begin{aligned} G_0(\mathbf{r}, t; \mathbf{r}', t') &= \frac{1}{(2\pi)^2} \int d^2\mathbf{k} \tilde{G}_0(\mathbf{k}; t, t') e^{i\mathbf{k} \cdot (\mathbf{r} - \mathbf{r}')} \\ &= - \int_{-\infty}^{\infty} \frac{dk_x}{2\pi} e^{-[Dk_x^2 \Delta t' - ik_x(x-x')] } \int_{-\infty}^{\infty} \frac{dk_y}{2\pi} e^{-[Dk_y^2 \Delta t' - ik_y(y-y')] } \\ &= \frac{-1}{4\pi D \Delta t'} e^{-[(x-x')^2 + (y-y')^2]/4D \Delta t'} , \end{aligned} \quad (\text{A.33})$$

where $\Delta t' = t - t'$, and the definite integrals have been performed by "completing the squares." Hence, the Green function for the infinite domain is seen to be none other than the two-dimensional Gaussian centred on the point (x', y') . Now, the Green function for the quarter plane is specified by,

$$\left(D \nabla_r^2 - \frac{\partial}{\partial t} \right) G(\mathbf{r}, t; \mathbf{r}', t') = \delta^2(\mathbf{r} - \mathbf{r}') \delta(t - t') ; \quad x \geq 0, y \geq 0 , \quad (\text{A.34a})$$

$$G(\mathbf{r} \neq \mathbf{r}', t=0) = 0 , \quad -D \left(\frac{\partial G}{\partial x} \right)_{x=0} = 0 , \quad -D \left(\frac{\partial G}{\partial y} \right)_{y=0} = 0 , \quad (\text{A.34b})$$

$$\lim_{r \rightarrow \infty} G = 0 , \quad \lim_{r \rightarrow \infty} (\nabla_r G) = 0 . \quad (\text{A.34c})$$

The solution to this is given simply by summing four Gaussians of the form (A.33) centred on the points $(\pm x', \pm y')$ [this is the method of images, by analogy to corresponding problems in electrostatics involving point charges in restricted domains]:

$$\begin{aligned} G(\mathbf{r}, t; \mathbf{r}', t') &= \frac{-1}{4\pi D \Delta t'} \left[\exp \left\{ -\frac{(x-x')^2}{4D \Delta t'} \right\} + \exp \left\{ -\frac{(x+x')^2}{4D \Delta t'} \right\} \right] \\ &\quad \times \left[\exp \left\{ -\frac{(y-y')^2}{4D \Delta t'} \right\} + \exp \left\{ -\frac{(y+y')^2}{4D \Delta t'} \right\} \right] . \end{aligned} \quad (\text{A.35})$$

Now, the *adjoint* Green function, G^+ , for Eq. (A.34) is the Green function for the adjoint diffusion operator. Thus, G^+ has an equation of motion,

$$\left(D \nabla_r^2 + \frac{\partial}{\partial t} \right) G^+(\mathbf{r}, t; \mathbf{r}'', t'') = \delta^2(\mathbf{r} - \mathbf{r}'') \delta(t - t'') ; \quad x \geq 0, y \geq 0 , \quad (\text{A.36})$$

with boundary conditions that have yet to be determined. To find the functional form of G^+ , and the boundary conditions it satisfies, multiply Eq. (A.34a) on both sides by $G^+(\mathbf{r}, t; \mathbf{r}'', t'')$, and integrate over \mathbf{r} and t :

$$\int d^2\mathbf{r} dt G^+ \left(D\nabla_r^2 - \frac{\partial}{\partial t} \right) G = \int d^2\mathbf{r} dt G^+ \delta^2(\mathbf{r}-\mathbf{r}') \delta(t-t') \\ = G^+(\mathbf{r}',t';\mathbf{r}'',t'') . \quad (\text{A.37})$$

Integrating the diffusion operator by parts:

$$G^+(\mathbf{r}',t';\mathbf{r}'',t'') = \int d^2\mathbf{r} dt G \left(D\nabla^2 + \frac{\partial}{\partial t} \right) G^+ \\ + \int d^2\mathbf{r} dt \left\{ \nabla \cdot (G^+ D\nabla G - GD\nabla G^+) + \frac{\partial}{\partial t} (-G^+ G) \right\} \\ = G(\mathbf{r}'',t'';\mathbf{r}',t') - \int dx dy (-G^+ G)_{t=0} + \int dx dy (-G^+ G)_{t=\tau} \\ - \int dy dt \left(G^+ D \frac{\partial G}{\partial x} - GD \frac{\partial G^+}{\partial x} \right)_{x=0} + \int dy dt \left(G^+ D \frac{\partial G}{\partial x} - GD \frac{\partial G^+}{\partial x} \right)_{x=\infty} \\ - \int dx dt \left(G^+ D \frac{\partial G}{\partial y} - GD \frac{\partial G^+}{\partial y} \right)_{y=0} + \int dx dt \left(G^+ D \frac{\partial G}{\partial y} - GD \frac{\partial G^+}{\partial y} \right)_{y=\infty} , \quad (\text{A.38})$$

where Gauss' theorem has been used to convert the volume integral over \mathbf{r},t into surface integrals, and for convenience the restriction $0 \leq t \leq \tau$ is placed on the time, where the surfaces are defined by the six hyperplanes $x = 0, y = 0, x = \infty, y = \infty, t = 0$ and $t = \tau$. Using the conditions on G given in Eqs. (A.34b) and (A.34c), the above reduces to,

$$G^+(\mathbf{r}',t';\mathbf{r}'',t'') = G(\mathbf{r}'',t'';\mathbf{r}',t') - \int dy dt \left(-GD \frac{\partial G^+}{\partial x} \right)_{x=0} \\ - \int dx dt \left(-GD \frac{\partial G^+}{\partial y} \right)_{y=0} + \int dx dy (-G^+ G)_{t=\tau} . \quad (\text{A.39})$$

The adjoint Green function is *defined* such that its functional form is,

$$G^+(\mathbf{r}',t';\mathbf{r}'',t'') = G(\mathbf{r}'',t'';\mathbf{r}',t') , \quad (\text{A.40})$$

i.e. its functional form is exactly that of G , except with the parameters swapped. To satisfy this definition, the adjoint boundary conditions are defined in such a way that the "extra" terms in Eq. (A.39) vanish, *i.e.*,

$$(G^+)_{t=\tau} = 0, \quad -D \left(\frac{\partial G^+}{\partial x} \right)_{x=0} = 0, \quad -D \left(\frac{\partial G^+}{\partial y} \right)_{y=0} = 0 . \quad (\text{A.41})$$

The requirement that G^+ be bounded (actually zero) as $x,y \rightarrow \infty$ is automatically satisfied by the boundedness of G , through Eq. (A.39) and the limiting conditions on G in Eq. (A.34c). Hence, Eqs. (A.36) and (A.41) together with the conditions for boundedness, give a unique specification of the adjoint Green function, the solution being Eq. (A.35) with the swap $(\mathbf{r},t) \leftrightarrow (\mathbf{r}',t')$.

The original problem, Eq. (A.25) is now solved as follows. Multiply Eq. (A.25)

by G^+ and integrate over \mathbf{r} and t :

$$\begin{aligned} - \int d^2\mathbf{r} dt G^+ g &= \int d^2\mathbf{r} dt G^+ \left(D\nabla^2 - \frac{\partial}{\partial t} \right) \Gamma \\ &= \int d^2\mathbf{r} dt \Gamma \left(D\nabla^2 + \frac{\partial}{\partial t} \right) G^+ + \int d^2\mathbf{r} dt \left\{ \nabla \cdot (G^+ D\nabla\Gamma - \Gamma D\nabla G^+) + \frac{\partial}{\partial t} (-G^+ \Gamma) \right\} \\ &= \Gamma - \int dy dt \left(G^+ D \frac{\partial \Gamma}{\partial x} \right)_{x=0} - \int dx dt \left(G^+ D \frac{\partial \Gamma}{\partial y} \right)_{y=0} - \int dx dy (-G^+ \Gamma)_{t=0} , \quad (\text{A.42}) \end{aligned}$$

where in the last step Gauss' theorem has been used to convert the volume integral into a surface integral [*c.f.* Eq. (A.38)], and the adjoint boundary conditions have been used.

Using the boundary conditions on Γ then gives,

$$\Gamma = - \int d^2\mathbf{r} dt G^+ g + \int dy dt (G^+ f)_{x=0} . \quad (\text{A.43})$$

This is exactly in the form (A.26) except with G replaced by G^+ . With,

$$f(y,t) = F_1 [1 - \theta(y - vt)] , \quad (\text{A.43a})$$

$$g(\mathbf{r},t) = G_1 e^{-x/\lambda} [1 - \theta(y - vt)] , \quad (\text{A.43b})$$

Eq. (A.43) may be written,

$$\begin{aligned} \Gamma(\mathbf{r}',t') &= -G_1 \int_0^\tau dt' \int_0^\infty dx' \int_0^{vt'} dy' G^+(x,y,\tau;x',y',t') e^{-x'/\lambda} \\ &\quad - F_1 \int_0^\tau dt' \int_0^{vt'} dy' G^+(x,y,\tau;x'=0,y',t') . \quad (\text{A.44}) \end{aligned}$$

Substituting in the appropriate expression for G^+ , and after much tedious algebra, the final answer is arrived at:

$$\begin{aligned} \Gamma(\mathbf{r},\tau) &= \frac{G_1}{4} \int_0^\tau dt' e^{D\Delta t'/\lambda^2} \left\{ e^{-x/\lambda} \operatorname{erfc} \left(\frac{\sqrt{D\Delta t'}}{\lambda} - \frac{x}{2\sqrt{D\Delta t'}} \right) + e^{x/\lambda} \operatorname{erfc} \left(\frac{\sqrt{D\Delta t'}}{\lambda} + \frac{x}{2\sqrt{D\Delta t'}} \right) \right\} \\ &\quad \times \left[\operatorname{erfc} \left(\frac{y - vt'}{2\sqrt{D\Delta t'}} \right) - \operatorname{erfc} \left(\frac{y + vt'}{2\sqrt{D\Delta t'}} \right) \right] \\ &\quad + F_1 \int_0^\tau dt' \frac{e^{-x^2/4D\Delta t'}}{2\sqrt{\pi D\Delta t'}} \left[\operatorname{erfc} \left(\frac{y - vt'}{2\sqrt{D\Delta t'}} \right) - \operatorname{erfc} \left(\frac{y + vt'}{2\sqrt{D\Delta t'}} \right) \right] , \quad (\text{A.45}) \end{aligned}$$

where erfc is the complementary error function:

$$\operatorname{erfc}(z) = \frac{2}{\sqrt{\pi}} \int_z^\infty e^{-u^2} du . \quad (\text{A.46})$$

Now identifying τ as the etch time, and putting $\Delta t' = \tau - t'$, gives Eq. (4.10) when $G_1 = 0$, and Eq. (4.12) when $F_1 = 0$.

B. Numerical methods

B.1 Cubic spline interpolation

Interpolated functions are used in a number of places in the numerical calculations. The Fermi-Dirac integral, for instance, is required to be evaluated many hundreds of times during the solution of the Poisson equation, which slows the calculation down considerably. To circumvent this problem, a cubic spline interpolation function is built to represent the integral. Likewise, once the conduction band edge energy, E_c , has been calculated by the Poisson solver, an integral needs to be performed to find the conductance. Since E_c is only known at the mesh points, the integral could be done using a crude sum over the mesh points. However, a little more accuracy is achieved by fitting a smooth cubic spline interpolation function over the mesh points, and integrating this *function* (by the method of Romberg – see below).

The method of interpolating using cubic splines is explained in more detail in Press *et al.* (1989), Ch. 3. What follows is a *very* brief summary. Let the set of ordered pairs $\{(x_i, y_i)\}$, $i = 1, \dots, N$, represent the known values of a function y , *i.e.* $y_i = y(x_i)$. Further, let the x_i be evenly spaced, *i.e.* $x_{i+1} - x_i = h$ [cubic splines can be defined without this restriction]. Then the cubic spline interpolation function is defined,

$$y = y(x) = Ay_i + By_{i+1} + \frac{1}{6}(A^3 + A)h^2y_i'' + \frac{1}{6}(B^3 - B)h^2y_{i+1}'' , \quad (\text{B.1})$$

where,

$$A = \frac{x_{i+1} - x}{h} , \quad B = \frac{x - x_i}{h} , \quad (\text{B.2})$$

and,

$$x_{i+1} \leq x \leq x_i . \quad (\text{B.3})$$

The parameters y_i'' are determined below. Thus, for x given between two "mesh points" as in Eq. (B.3), the interpolated value of y is given by Eq. (B.1). The parameters y_i'' are defined by the set of equations,

$$y_1'' = 0, \quad (\text{B.4a})$$

$$\frac{1}{6}y_{i-1}'' + \frac{2}{3}y_i'' + \frac{1}{6}y_{i+1}'' = \frac{1}{h^2}[y_{i+1} - 2y_i + y_{i-1}], \quad i = 2, \dots, N-1 , \quad (\text{B.4b})$$

$$y_N'' = 0 . \quad (\text{B.4c})$$

This system of equations can be written in matrix form and solved as such, to give the y_i'' . Knowing these, all the parameter values on the right hand side of Eq. (B.1) are known, so Eq. (B.1) represents a *function* which interpolates the original data set. The set of equations in Eq. (B.4) ensures that this interpolation function is continuous up to

the first derivative.

B.2 Integrating by the method of Romberg

At several stages during the calculations, definite integrations are performed of known functions (the Fermi-Dirac integral, the diffusion equation solutions for the defect distribution, and the interpolation function representing the conduction band edge when calculating the conductance). These integrations are performed using the method of Romberg. The algorithm is described in Press *et al.* (1989), Ch. 4, to which the reader is referred. The basic idea is that the function is summed over successively finer meshes, with extrapolations to zero mesh size being made at each step. The process is brought to an end once a specified tolerance is reached. [This is an extremely crude description which does no justice to the algorithm. Several subtleties in the method of extrapolation ensure that convergence is reached much more rapidly than in other integration algorithms which also use extrapolations to zero mesh size. Details in Press *et al.* (1989).] Note that the method works for a defined *function*, not tabulated values (for which, *e.g.*, Simson's rule may be used).

B.3 Solving the 1-dimensional Poisson equation

The 1-dimensional Poisson equation,

$$\frac{d^2\phi(y)}{dy^2} = -\rho(y), \quad \phi(0) = \phi_s, \quad \left(\frac{d\phi}{dy}\right)_{y=L} = 0, \quad (\text{B.5})$$

where L is the system length, and ϕ_s is the surface potential, is very easy to solve numerically. Let $\phi_i = \phi(y_i)$ and $\rho_i = \rho(y_i)$ be the values of ϕ and ρ on the mesh points y_i , $i = 0, \dots, N$ which are assumed to be uniformly spaced, *i.e.* $y_{i+1} - y_i = h$. Then Eq. (B.5) can be replaced by a finite difference equivalent [*e.g.* Gerald (1970), Ch. 9]:

$$\phi_0 = \phi_s, \quad (\text{B.6a})$$

$$\frac{1}{h^2}[\phi_{i+1} - 2\phi_i + \phi_{i-1}] = -\rho_i, \quad i = 1, \dots, N-1, \quad (\text{B.6b})$$

$$\frac{1}{2h}[\phi_N - \phi_{N-2}] = 0. \quad (\text{B.6c})$$

Since the ρ_i are known, the system of equations represented by Eq. (B.6) may be written in matrix form, and solved as such.

The actual 1-dimensional Poisson equation that must be solved, Eq. (4.25), is considerably more complicated than Eq. (B.5). The equation required to be solved is of the form,

$$\frac{d^2\phi(y)}{dy^2} = -R[\phi(y)] , \quad \phi(0) = \phi_s , \quad \left(\frac{d\phi}{dy}\right)_{y=L} = 0 , \quad (\text{B.7})$$

where R is a function of $\phi(y)$. This equation is solved by a relaxation method, whereby a zeroth order ("order" means "step") estimate $\phi^{(0)}$ is successively refined towards the solution, $\phi^{(k)} \rightarrow \phi$ for k sufficiently large. At each order k , the replacement,

$$R[\phi^{(k)}(y)] \rightarrow \rho^{(k+1)}(y) \quad (\text{B.8})$$

is made in Eq. (B.7). The equation is then of the form (B.5), and can be solved to give the next order estimate, $\phi^{(k+1)}$. Actually, to ensure stability of the method, a slight modification must be made to this procedure. Thus, the method is as follows.

(1) The zeroth order estimate is made by assuming a simple charge distribution:

$$\rho^{(0)}(y) \sim \begin{cases} eN_D/\epsilon , & y < y_0 \\ 0 & , y > y_0 \end{cases} \quad (\text{B.9})$$

where N_D is the donor density, $y_0 = (2\epsilon\phi_s/eN_D)^{1/2}$ is the nominal ("damage free") depletion depth, and energies are measured in eV. Inserting Eq. (B.9) into Eq. (B.5) and solving, gives $\phi^{(0)}$. This then gives a first order estimate for the charge, $\rho^{(1)}(y) = R[\phi^{(0)}(y)]$.

(2) Using this estimate of $\rho^{(k)}(y)$, Eq. (B.5) is again solved to give ϕ_{new} . Rather than setting $\phi^{(k)} = \phi_{new}$, which is unstable, the next order estimate is taken as,

$$\phi^{(k)} = a\phi_{new} + (1-a)\phi^{(k-1)} , \quad (\text{B.10})$$

where a is some constant chosen to ensure convergence at a reasonable rate. Then $\rho^{(k+1)}(y) = R[\phi^{(k)}(y)]$.

(3) The above process is repeated until the difference between $\phi^{(k-1)}$ and $\phi^{(k)}$ is within some predefined tolerance:

$$\|\phi^{(k)} - \phi^{(k-1)}\| < \epsilon , \quad (\text{B.11})$$

where ϵ is the tolerance, and $\|\dots\|$ is a suitably defined norm.

(4) Once this convergence has been achieved, the electron density in the structure is found by evaluating the Fermi-Dirac integral, and the conductance in turn found from this. This gives the conductance for a particular mesh size. The whole calculation is then repeated for several successively smaller mesh sizes. The final calculated value of conductance is taken as the extrapolation to zero mesh size.

B.4 Solving the 2-dimensional Poisson equation

Solving the 2-dimensional Poisson equation is analogous to the 1-dimensional

case, except relaxation is used twice, so to speak. The 2-dimensional Poisson equation,

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) \phi(x,y) = -\rho(x,y) , \quad (\text{B.12a})$$

$$\phi(0,y) = \phi(w,y) = \phi(x,0) = \phi_s , \quad \left(\frac{\partial \phi}{\partial y} \right)_{y=v\tau} = 0 , \quad (\text{B.12b})$$

where ϕ_s is the surface potential, w is the wire width, and $v\tau$ is the etch depth, is solved using the method of successive over relaxation, as described in Press *et al.* (1989), Ch. 17. An initial estimate of ϕ is successively refined towards a convergent solution of Eq. (B.12), with the process being terminated once some specified tolerance is reached. If $\phi_{ij} = \phi(x_i, y_j)$, $i = 0, \dots, M$ and $j = 0, \dots, N$, then the derivative boundary condition is specified by imposing,

$$\phi_{i,N} = \phi_{i,N-2} , \quad i = 0, \dots, M , \quad (\text{B.13})$$

at the start of each iteration of the relaxation process.

As in the 1-dimensional case, the actual 2-dimensional Poisson equation that must be solved, Eq. (4.22), is much more complicated than Eq. (B.12). This is of the form,

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) \phi(x,y) = -R[\phi(x,y)] , \quad (\text{B.14})$$

with boundary conditions given by Eq. (B.12b). R is a function of $\phi(x,y)$. As before, the equation is solved by a relaxation method, where R is replaced by a simple function at each step [*c.f.* Eq. (B.8)] to give an equation of the form (B.12). The method is as follows.

(1) A zeroth order estimate of ϕ is made by assuming the simple charge distribution,

$$\rho^{(0)}(x,y) \sim \begin{cases} eN_D/\epsilon , & \text{near sidewalls within epilayer} \\ 0 , & \text{otherwise} \end{cases} \quad (\text{B.15})$$

where N_D is the doping density, and energies are measured in eV. This is fed into Eq. (B.12), which is solved by successive over relaxation to give $\phi^{(0)}$. This gives $\rho^{(1)}(x,y) = R[\phi^{(0)}(x,y)]$.

(2) Eq. (B.12) is again solved to yield ϕ_{new} . As for the 1-dimensional case, the next order estimate is taken as,

$$\phi^{(k)} = a\phi_{new} + (1-a)\phi^{(k-1)} , \quad (\text{B.16})$$

where a is suitably chosen to ensure convergence. Then $\rho^{(k+1)}(x,y) = R[\phi^{(k)}(x,y)]$.

(3) The process is repeated until some tolerance is reached.

(4) Once convergence is reached, the resulting conduction band edge gives the

wire conductance. The whole calculation is repeated for several successively smaller mesh sizes. The calculated conductance is taken to be the extrapolation to zero mesh size.

B.5 The digamma function

The digamma function is defined [see Abramowitz and Stegun (1965), Ch. 6],

$$\psi(z) = \frac{d}{dz} \ln \Gamma(z) , \quad (\text{B.17})$$

where,

$$\Gamma(z) = \int_0^{\infty} u^{z-1} e^{-u} du , \quad (\text{B.18})$$

is the gamma function. The function is required for fitting the low field weak localisation magnetoconductance of 2-dimensional films. An efficient numerical method for evaluating $\ln \Gamma(z)$ is given by Press *et al.* (1989), Ch. 6. The digamma function may then be calculated using,

$$\psi(z) = \frac{1}{2h} [\ln \Gamma(z+h) - \ln \Gamma(z-h)] , \quad (\text{B.19})$$

for a small constant h .

C. Program listings

C.1 Program listing 1: wire conductances

This program is used to solve the wire conductance problem as defined in §4.4.1. Wire dimensions and material parameters are defined at the start of the program. The mesh size is also specified here. It must be ensured that all dimensions are in multiples of this mesh size. The output of the program is the wire conductance. The program is run several times for a particular wire, with successively smaller mesh sizes. The desired conductance is then the extrapolation to zero mesh size.

```
// -----
// Description: Solves the 2D Poisson equation to find the conduction band edge Ec in
//              a dry-etched wire structure. An analytically obtained defect distribution
//              function can be included, the defects being electron traps with charge
//              states 1- (empty) to 2- (full). A double integral is then
//              performed to give the wire conductance.
// Language:    Domain/C++ implementation of Release 2.0 AT&T C++ Translator.
```

```

// Machine used: HP 9000 (Apollo) Series 400 workstation.
// Author:      M Rahman, Glasgow University.
// Started:    22 May 1991.
// -----

#include <stdio.h> // for standard input and output
#include <math.h>  // for fabs(),erf(),erfc() etc

#define YES 1
#define NO 0

// -----
// Physical constants and parameters.
// -----

#define CHARGE 1.602e-19
#define PERMITTIVITY 8.854e-12
#define ETA_RELATIVE 13.1
#define ELECTRON_MASS 9.11e-31
#define hBAR 1.055e-34
#define BOLTZMANN_CONSTANT 1.38e-23

double topSurfacePotential=0.7,
       sidewallPotential=0.7,
       meshSize=10.0e-9,
       wireLength=10e-6,
       wireWidth=50e-9,
       epiThickness=50e-9,
       dopingDensity=6.5e24,
       interfaceDensity=dopingDensity,
       mobility=0.14,
       mass_e=0.067*ELECTRON_MASS,
       mass_h=0.47*ELECTRON_MASS, // using hh=0.45m; lh=0.082m
       temperature=300.0,
       mix=0.015, // mixing factor for relaxation procedure
       F1=0.74e15, // defect current density from surface (if used) units: num/area/sec
       D=1e-22, // defect diffusion coefficient
       Etr=0.34, // trap depth measured from conduction band edge
       Lx=9e-9, // damage source = G1*exp(-X/Lx) (if used)
       G1=5.5e23, // strength of damage source at surface (if used) units: num/volume/sec
       etchDepth=100e-9,
       etchRate=3.33e-9, // =3.33 nm/s usually
       etchTime=etchDepth/etchRate;
int constantFlux=NO, // =NO if exponential source is used instead of constant flux
    includeDamage=YES;

const double e_div_eta = CHARGE/(PERMITTIVITY*ETA_RELATIVE);
double kT = BOLTZMANN_CONSTANT*temperature,
       beta = CHARGE/kT,
       Nc = 2.0*exp(1.5*log(mass_e*kT/(2.0*PI*hBAR*hBAR))),
       Nv = 2.0*exp(1.5*log(mass_h*kT/(2.0*PI*hBAR*hBAR))),
       Nc_2_div_sqrtPI = Nc*2.0/sqrt(PI),
       Nv_2_div_sqrtPI = Nv*2.0/sqrt(PI);

// -----
// Main declarations.
// -----
// Nd[][] = Charge density in wire. Also acts as the source term for the Poisson eqn.
// Ec[][] = Conduction band edge. The 2D Poisson equation is solved to obtain this.
// Gd[][] = Defect density in wire. Assumed to be a solution to the diffusion equation.
// eD,eD2,hD,hD2: Used for cubic spline interpolation of electron and hole densities.
// Ele,...: Determines range over which spline interpolation is carried out.
// -----

#define MAX_ARRAY_SIZE 270
#define NUM_SPL 200

```

```

typedef double Array[MAX_ARRAY_SIZE][MAX_ARRAY_SIZE];
typedef double (*funcPtr)(double);
typedef double splArray[NUM_SPL];

const double E1e=-1.0, E2e=2.0, dEe=(E2e-E1e)/NUM_SPL,
            E1h=-2.4, E2h=0.6, dEh=(E2h-E1h)/NUM_SPL;

Array Nd={0}, Ec={0}, Ec_old={0}, Gd={0};
splArray eD={0}, eD2={0}, hD={0}, hD2={0};

// -----
// Function prototypes.
// -----
void ConvertToChargeDensity(Array f, int jmax, int lmax);
double residual(Array f, Array u, int j, int l);
void solve(Array f, Array u, int jmax, int lmax, double rjac, int mits, double eps);
void integrate(funcPtr func, double a, double b, double& ss, int proper);
void buildSpline(double* y, double* y2, int n, double dx);
double splineValue(splArray y, splArray y2, double x, double xlo, double dx);

// -----
// Miscellaneous routines.
// -----
// DESCRIPTION:
//   These routines do small manipulations on the data.
// FUNCTIONS:
//   fitToMesh()    - Ensures that the passed parameter is a multiple of the mesh size.
//   MakeConstant() - Set the entire array to a constant value.
// -----

void fitToMesh(double& aLength) { aLength=meshSize*int(aLength/meshSize+0.5); }

void MakeConstant(double* a, int jmax, int lmax, double setValue) {
    for (int j=0; j<=jmax; j++) for (int l=0; l<=lmax; l++) a[jmax*j+l] = setValue;
}

// -----
// Routines to calculate carrier concentrations.
// -----
// DESCRIPTION:
//   Calculate the electron and hole densities by numerically evaluating the Fermi-Dirac
//   integral of order 1/2, represented as F_1_2 below.
// FUNCTIONS:
//   Eg()          - Returns the GaAs energy gap as a function of temperature.
//   FDintegrand() - Integrand of the Fermi-Dirac integral. The integral is evaluated at
//                   FDarg, which is set before the integrating routine is called.
//   F_1_2()       - Evaluates the Fermi-Dirac integral at the specified value.
//   Nelectron()   - Returns the number of electrons given the local conduction band
//                   energy, by calling F_1_2.
//   Nhole()       - As for Nelectron but for holes.
//   electronDensity() - Returns the electron density, by calling Nelectron() or by
//                   calling the previously calculated spline function.
//   holeDensity() - As for electronDensity but for holes.
//   BuildSplineForElectronDensity() - Fits a cubic spline to the electron density as a
//                   function of Ec. Using the spline function instead of continually
//                   evaluating the Fermi-Dirac integral speeds up the relaxation
//                   procedure for the Poisson equation.
//   BuildSplineForHoleDensity() - Fits a cubic spline to the hole density.
//   Build
// -----

double Eg(double T) { return (1.52-(5.4e-4*T*T)/(T+204.0)); }

double FDarg=0.0; // to be set before calling FD_integrand

```

```

double FD_integrand(double z) { return sqrt(z)/(1.0+exp(z+FDarg)); }

double F_1_2(double x) {
    double result=0.0;
    FDarg=x;
    integrate(funcPtr(FD_integrand),0.0,15.0,result,YES);
    return result;
}

double N_electron(double E) { return Nc_2_div_sqrtPI*F_1_2(beta*E); }
double N_hole(double E) { return Nv_2_div_sqrtPI*F_1_2(-beta*E); }

double electronDensity(double E) {
    return (E<Ele || E>E2e-dEe) ? N_electron(E) : splineValue(eD,eD2,E,Ele,dEe);
}

double holeDensity(double E) {
    return (E<Elh || E>E2h-dEh) ? N_hole(E) : splineValue(hD,hD2,E,Elh,dEh);
}

void BuildSplineForElectronDensity() {
    double p=Ele;
    for (int i=0; i<NUM_SPL; i++, p+=dEe) eD[i]=N_electron(p);
    buildSpline(eD,eD2,NUM_SPL,dEe);
}

void BuildSplineForHoleDensity() {
    double p=Elh;
    for (int i=0; i<NUM_SPL; i++, p+=dEh) hD[i]=N_hole(p);
    buildSpline(hD,hD2,NUM_SPL,dEh);
}

// -----
// Routines to calculate defect concentrations.
// -----
// DESCRIPTION:
// Calculate the defect density Gd[j][1], arising from each sidewall by numerically
// evaluating either of two integrals, depending on the status of the variable
// constantFlux: constantFlux=YES causes damage_constantDefectFlux() to be evaluated;
// constantFlux=NO causes damage_exponentialSource to be evaluated.
// FUNCTIONS:
// damage_constantDefectFlux() - Represents an integral solution of the homogeneous
// diffusion equation in the quarter plane {x>=0,y>=0}, with non-homogeneous
// boundary conditions, -D.dG/dx=F1 at the sidewalls, where G is
// written as Gd[j][1] in the program.
// damage_exponentialSource() - Represents an integral solution of the nonhomogeneous
// diffusion equation in the quarter plane {x>=0,y>=0}, with source term
// g=G1.exp(-x/Lx), and homogeneous boundary conditions.
// BuildDefectDensityFunction() - Evaluates one of the above integrals at each mesh
// point in the wire. The result is stored in array Gd[j][1].
// BuildDefectDinsityFunction2() - assigns the array Gd[j][1] with values obtained from
// the simplified form of the defect distributions.
// -----

double X=0.0, Y=0.0; // set before integrating damageIntegrand
double damage_constantDefectFlux(double t) {
    double _4Dt=4.0*D*(etchTime-t), sqrt_4Dt=sqrt(_4Dt),
        x_dependentPart = exp(-X*X/_4Dt),
        y_dependentPart = erfc((Y+etchRate*t)/sqrt_4Dt)-erfc(Y+etchRate*t)/sqrt_4Dt;
    return x_dependentPart*y_dependentPart/sqrt_4Dt;
}

double damage_exponentialSource(double t) {
    double Dt=D*(etchTime-t), sqrtDt=sqrt(Dt), sqrt_4Dt=2.0*sqrtDt,
        sqrtDt_div_Lx=sqrtDt/Lx, X_div_Lx=X/Lx, X_div_sqrt_4Dt=X/sqrt_4Dt,
        x_dependentPart = exp(-X_div_Lx)*erfc(sqrtDt_div_Lx-X_div_sqrt_4Dt)
            + exp(X_div_Lx)*erfc(sqrtDt_div_Lx+X_div_sqrt_4Dt),

```

```

    y_dependentPart = erfc((Y-etchRate*t)/sqrt_4Dt)-erfc((Y+etchRate*t)/sqrt_4Dt);
return exp(sqrtDt_div_Lx*sqrtDt_div_Lx)*x_dependentPart*y_dependentPart;
}

void BuildDefectDensityFunction(Array Gd, int jmax, int lmax) {
    double N_defect=1e20, preFactor = (constantFlux) ? F1/sqrt(PI) : 0.25*G1;
    funcPtr damageDistribution = (constantFlux) ?
        funcPtr(damage_constantDefectFlux) : funcPtr(damage_exponentialSource);
    MakeConstant((double*)Gd,jmax,lmax,0.0);
    for (int j=0; j<=jmax; N_defect=1e20, j++) for (int l=0; l<=lmax && N_defect>1.0; l++) {
        X=double(j)*meshSize; Y=double(l)*meshSize;
        integrate(damageDistribution,0.0,etchTime,N_defect,NO);
        Gd[j][l] += (N_defect * preFactor); Gd[jmax-j][l] += N_defect;
    }
}

void BuildDefectDensityFunction2(Array Gd, int jmax, int lmax) {
    double N_defect=0.0;
    MakeConstant((double*)Gd,jmax,lmax,0.0);
    for (int j=0; j<=jmax; j++) for (int l=0; l<=lmax; l++) {
        X=double(j)*meshSize; Y=double(l)*meshSize;
        N_defect = G1*exp(-X/Lx)*(etchTime-Y/etchRate);
        Gd[j][l] += N_defect; Gd[jmax-j][l] += N_defect;
    }
}

// -----
// Routines to calculate the total charge concentration.
// -----
// DESCRIPTION:
// These routines provide the source term for the 2D Poisson equation.
// FUNCTIONS:
// CalculateChargeConcentration() - Sums the electron, hole, defect, and trapped-electron
// densities, with -ve charge for defects and trapped electrons. Electrons are
// assumed to be trapped with the equilibrium Fermi-Dirac distribution.
// ConvertToChargeDensity() - Scales the number densities above to give charge density.
// -----

void CalculateChargeConcentration(Array Nd, int jmax, int lmax) {
    int ldepth = int(lmax*epiThickness/etchDepth+0.5);
    for (int j=0; j<=jmax; j++) for (int l=0; l<=lmax; l++) {
        Nd[j][l] = -electronDensity(Ec[j][l]) + holeDensity(Ec[j][l]-Eg(temperature));
        if (includeDamage) Nd[j][l] -= Gd[j][l]*(1.0+1.0/(1.0+exp(beta*(Ec[j][l]-Etr))));
    } for (j=0; j<=jmax; j++) for (l=0; l<=ldepth; l++) Nd[j][l] += dopingDensity;
}

void ConvertToChargeDensity(Array f, int jmax, int lmax) {
    double scaled_e_div_eta = e_div_eta*meshSize*meshSize;
    for (int j=0; j<=jmax; j++) for (int l=0; l<=lmax; l++) f[j][l] *= scaled_e_div_eta;
}

// -----
// Numerical routines.
// -----
// DESCRIPTION:
// C++ implementations of routines taken from Numerical Recipes (Pascal). Page numbers
// below refer to the book, which should be consulted for further details.
// FUNCTIONS:
// residual() - Evaluates the residual for the diffusion equation.
// totalResidual() - Finds the sum of the residuals over the entire mesh.
// solve() - Solves Poisson equation with Eg=const on top and sides, and
// dEg/dx=0 on the bottom ('sor' p717).
// extrapolate() - Used for polynomial extrapolations ('polint' p92).
// trapzd() - Used for closed interval definite integrals ('trapzd' p125).
// midpnt() - Used for open ended definite integrals ('midpnt' p132).
// integrate() - Integrates the given function using the method of Romberg. If

```

```

//          proper=YES then integrates over a closed interval ('qromb' p129),
//          otherwise over an open interval ('qromo' p133).
//  buildSpline() - Builds an interpolating cubic spline using the given data points
//                  ('spline' p99).
//  splineValue() - Returns the interpolated spline fit at value x ('splint' p100).
// -----

double residual(Array f, Array u, int j, int l) {
    return u[j+1][l]+u[j-1][l]+u[j][l+1]+u[j][l-1]-4*u[j][l]-f[j][l];
}

double totalResidual(Array f, Array u, int jmax, int lmax) {
    double anorm=0.0;
    for (int j=1; j<jmax; j++) for (int l=1; l<lmax; l++) anorm+=fabs(residual(f,u,j,l));
    return anorm;
}

void solve(Array f, Array u, int jmax, int lmax, double rjac, int mits, double eps) {
    double anormf=totalResidual(f,u,jmax,lmax), omega=1.0, anorm=anormf+1.0, resid=0.0;
    for (int n=1; n<=mits && (anorm>eps*anormf || n<3); n++) {
        for (int j=0; j<jmax; j++) Ec[j][lmax]=Ec[j][lmax-2];
        for (j=1, anorm=0.0; j<jmax; j++) for (int l=(n+j)%2+1; l<lmax; l+=2) {
            anorm += fabs(resid=residual(f,u,j,l));
            u[j][l] += omega*resid/4.0;
        }
        omega = 1.0/(1.0-rjac*rjac*((n==1)?0.5:0.25*omega));
    }
}

void extrapolate(double *xa, double *ya, int n, double x, double& y, double& dy) {
    double *c = new double[n], *d = new double[n];
    int ns=1;
    double dif = fabs(x-xa[0]), dift=0.0;
    for (int i=0; i<n; i++) {
        if ((dift=fabs(x-xa[i]))<dift) { ns=i; dif=dift; }
        c[i] = d[i] = ya[i];
    }
    y=ya[ns--];
    for (int m=1; m<n; m++) {
        for (i=0; i<n-m; i++) {
            double ho=xa[i]-x, hp=xa[i+m]-x, den=(c[i+1]-d[i])/(ho-hp);
            d[i]=hp*den; c[i]=ho*den;
        }
        y += (dy=(2*ns<n-m)?c[ns+1]:d[ns--]);
    }
    delete c; delete d;
}

void trapz(funcPtr func, double a, double b, double& s, int n, int& TrapzIt) {
    if (n==1) { TrapzIt=1; s = 0.5*(b-a)*(func(a)+func(b)); }
    else {
        double del=(b-a)/TrapzIt, x=a-0.5*del, sum=0.0; // NB here x=a+0.5*del-del
        for (int j=0; j<TrapzIt; j++) sum += func(x+=del);
        s = 0.5*s+(b-a)*sum/(TrapzIt*2);
    }
}

void midpnt(funcPtr func, double a, double b, double& s, int n, int& MidpntIt) {
    if (n==1) { MidpntIt=1; s = (b-a)*func(0.5*(a+b)); }
    else {
        double del=(b-a)/(3.0*MidpntIt), ddel=del+del, x=a-0.5*del, sum=0.0;
        for (int j=0; j<MidpntIt; j++) { sum+=func(x+=del); sum+=func(x+=ddel); }
        s = s/3.0+(b-a)*sum/(MidpntIt*3);
    }
}

void integrate(funcPtr func, double a, double b, double& ss, int proper) {
    const double eps=(proper)?1e-3:0.01, df=(proper)?4.0:9.0;
    const int jmax=20, jmaxp=21, k=5;
    double *h = new double[jmaxp], *s = new double[jmaxp];
    double *c = new double[k], *d = new double[k];
}

```

```

h[0]=1.0; double dss=ss+1.0; int iter=0;
for (int j=0; j<jmax && fabs(dss)>eps*fabs(ss); j++) {
  if (proper) { trapzd(func,a,b,s[j],j+1,iter); }
  else { midpnt(func,a,b,s[j],j+1,iter); }
  if (j>k-1) {
    for (int i=0; i<k; i++) { c[i]=h[j-k+i]; d[i]=s[j-k+i]; }
    extrapolate(c,d,k,0.0,ss,dss);
  } s[j+1]=s[j]; h[j+1]=h[j]/df;
} delete d; delete c; delete s; delete h;
}

void buildSpline(double* y, double* y2, int n, double dx) {
  double *u = new double[n], p=0.0;
  y2[0]=-0.5, u[0]=0.0;
  for (int i=1; i<n-1; i++) {
    y2[i] = -0.5/(p=0.5*y2[i-1]+2.0);
    u[i] = (y[i+1]+y[i-1]-2.0*y[i])/dx;
    u[i] = (3.0*u[i]/dx-0.5*u[i-1])/p;
  } y2[n-1] = (-0.5*u[n-2])/(0.5*y2[n-2]+1.0);
  for (i=n-2; i>=0; i--) y2[i] = y2[i]*y2[i+1]+u[i];
  delete u;
}

double splineValue(double* y, double* y2, double x, double xmin, double dx) {
  int klo=int((x-xmin)/dx);
  double xlo=double(klo)*dx+xmin, a=(xlo+dx-x)/dx, b=(x-xlo)/dx;
  return a*y[klo]+b*y[klo+1]+((a*a*a-a)*y2[klo]+(b*b*b-b)*y2[klo+1])*dx*dx/6.0;
}

// -----
// Main calculational routines.
// -----
// DESCRIPTION:
// These routines allow the 2D Poisson equation with an arbitrary source term to be
// solved using a relaxation method. This gives the conduction band edge Ec[j][1].
// A numerical double integral over the resulting local electron density then gives
// the conductance.
// FUNCTIONS:
// SetBoundaryValues() - Set the boundary values of Eg[[]].
// FindZerothOrderSolution() - Artificially puts some charge into the epilayer region
// of the wire and integrates the Poisson equation to find a zeroth
// order solution for the full problem.
// FindFinalSolution() - Iterates solutions of the Poisson equation, starting from the
// zeroth order solution, until convergence is reached. A relaxation
// procedure is used to iterate successive solutions: (1) A solution
// is obtained to Ec[[]] by calling solve(). (2) This gives a new
// electron distribution which is used in the next iteration. (3) The
// values of Ec[[]] used in the next iteration are weighted averages
// of the current solution with the previous solution. (4) Repeat
// until the totalResidual is within the specified tolerance. This
// gives the desired solution.
// Nelectron_dx() - Used for evaluating the x-integral in the double integral for the
// conductance. This forms the integrand for Nelectron_dx_dy().
// Nelectron_dx_dy() - Used for evaluating the double integral for the conductance.
// WireConductance() - Evaluates the wire conductance by first interpolating the
// conduction band edge using a mesh of cubic splines, and then
// integrating the resulting spline function using Nelectron_dx_dy().
// SolveProblem() - Calls the above functions to get the solution. First find a
// zeroth order solution to start the relaxation procedure with.
// Solve the Poisson equation to obtain the conduction band edge.
// Finally use this to obtain the electron density and hence the
// conductance of the wire.
// -----

void SetBoundaryValues(Array a, int jmax, int lmax) {
  for (int j=0; j<=jmax; j++) a[j][0] = topSurfacePotential;
  for (int l=0; l<=lmax; l++) a[0][l] = a[jmax][l] = sidewallPotential;
}

```

```

}

void FindZerothOrderSolution(Array Nd, Array Ec, int jmax, int lmax, double rjac) {
    MakeConstant((double*)Ec, jmax, lmax, 0.0);
    MakeConstant((double*)Nd, jmax, lmax, 0.0);
    SetBoundaryValues(Ec, jmax, lmax);
    double depletionEstimate = sqrt(2.0*sidewallPotential/(e_div_eta*dopingDensity));
    int jdepl = int(jmax*depletionEstimate/wireWidth+0.5),
        ldepl = int(lmax*depletionEstimate/etchDepth+0.5),
        ldepth = int(lmax*epiThickness/etchDepth+0.5);
    jdepl = (jdepl<2)?2:jdepl; ldepl = (ldepl<2)?2:ldepl;
    for (int j=0; j<=jmax; j++) for (int l=0; l<=ldepth; l++) Nd[j][l] = 0.2*dopingDensity;
    for (j=0; j<jdepl; j++) for (l=0; l<ldepth; l++)
        Nd[j][l] = Nd[jmax-j][l] = dopingDensity;
    for (l=0; l<ldepl; l++) for (j=0; j<=jmax; j++) Nd[j][l] = 2.0*dopingDensity;
    for (j=0; j<=jmax; j++) Nd[j][ldepth]=0.3*interfaceDensity;
    ConvertToChargeDensity(Nd, jmax, lmax);
    solve(Nd, Ec, jmax, lmax, rjac, 100, 0.05);
}

void FindFinalSolution(Array Nd, Array Ec, int jmax, int lmax, double rjac, double mix) {
    double eps=1e-4;
    double anorm_bef=0.0, anorm_aft=anorm_bef+1.0;
    int ldepth = int(lmax*epiThickness/etchDepth+0.5);
    for (int iter=1; iter<500 && anorm_aft>eps*anorm_bef; iter++) {
        CalculateChargeConcentration(Nd, jmax, lmax);
        ConvertToChargeDensity(Nd, jmax, lmax);
        for (int j=0; j<=jmax; j++) for (int l=0; l<=lmax; l++) Ec_old[j][l]=Ec[j][l];
        if (iter==1) anorm_bef=totalResidual(Nd, Ec, jmax, lmax);
        solve(Nd, Ec, jmax, lmax, rjac, 200, 0.1);
        for (j=0; j<=jmax; j++) for (l=0; l<=lmax; l++)
            Ec[j][l] = (1.0-mix)*Ec_old[j][l]+mix*Ec[j][l];
        anorm_aft=totalResidual(Nd, Ec, jmax, lmax);
    }
}

double *spl_j, *spl_j2;
double Nelectron_dx(double x) {
    return splineValue(spl_j, spl_j2, x, 0.0, meshSize); }

int hj;
double *Ec12[ NUM_SPL ];
double Nelectron_dx_dy(double y) {
    for (int j=0; j<hj; j++) spl_j[j]=splineValue((double*)Ec[j], Ec12[j], y, 0.0, meshSize);
    buildSpline(spl_j, spl_j2, hj, meshSize);
    double halfResult=0.0;
    integrate(Nelectron_dx, 0.0, wireWidth/2.0, halfResult, YES);
    return 2.0*halfResult;
}

double WireConductance(Array Ec, int jmax, int lmax) {
    int half_jmax=int(jmax/2)+1;
    double *Ecj = new double[half_jmax], *Ecj2 = new double[half_jmax], conductance=0.0;
    for (int j=0; j<half_jmax; j++)
        buildSpline((double*)Ec[j], Ec12[j], lmax, meshSize);
    spl_j=Ecj, spl_j2=Ecj2, hj=half_jmax;
    integrate(Nelectron_dx_dy, 0.0, etchDepth, conductance, YES);
    delete Ecj; delete Ecj2;
    for (j=half_jmax-1; j>=0; j--) delete Ec12[j];
    return conductance*CHARGE*mobility/wireLength;
}

void SolveProblem() {
    fitToMesh(wireWidth); fitToMesh(epiThickness); fitToMesh(etchDepth);
    int jmax=int(wireWidth/meshSize+0.5), lmax=int(etchDepth/meshSize+0.5);
    double rjac = (cos(PI/(jmax+1))+cos(PI/(lmax+1)))/2.0;
    FindZerothOrderSolution(Nd, Ec, jmax, lmax, rjac);
    printf("integrating for nel ... \n"); BuildSplineForElectronDensity();
}

```

```

printf("integrating for hol ...\n"); BuildSplineForHoleDensity();
if (includeDamage) {
    printf("building defect density ... \n"); BuildDefectDensityFunction(Gd, jmax, lmax); }
FindFinalSolution(Nd, Ec, jmax, lmax, rjac, mix);
printf("Calculating conductance ... \n");
printf("CONDUCTANCE = %e\n", WireConductance(Ec, jmax, lmax));
}

// -----
void main() {
    SolveProblem();
}

```

C.2 Program listing 2: sheet conductances

This program is used to solve the epilayer sheet conductance problem as stated in §4.4.2. The parameter 'simulationDepth' determines the length of the system over which the 1-dimensional Poisson equation is solved. The program gives the sheet conductance for the structure. The program is run several times with successively smaller mesh sizes. The desired sheet conductance is the extrapolation to zero mesh size.

```

// -----
// Description: Solves the 1D Poisson equation to find the conduction band edge Ec in
//              a dry-etched epilayer. An analytically obtained defect
//              distribution function can be included, the defects being electron
//              traps with charge states 1- (empty) to 2- (full).
// Language:    Domain/C++ implementation of Release 2.0 AT&T C++ Translator.
// Machine used: HP 9000 (Apollo) Series 400 workstation.
// Author:      M Rahman, Glasgow University.
// Started:     28 July 1991.
// -----

#include <stdio.h> for standard input and output
#include <math.h> for fabs(), erf(), erfc() etc

#define YES 1
#define NO 0

// -----
// Physical constants and parameters.
// -----

#define CHARGE 1.602e-19
#define PERMITTIVITY 8.854e-12
#define ETA_RELATIVE 13.1
#define ELECTRON_MASS 9.11e-31
#define hBAR 1.055e-34
#define BOLTZMANN_CONSTANT 1.38e-23

double surfacePotential=0.7,
       meshSize=10.0e-9,
       epiThickness=140e-9,
       simulationDepth=300e-9,
       dopingDensity=3.0e24,
       interfaceDensity=dopingDensity,
       mobility=0.17,

```

```

mass_e=0.067*ELECTRON_MASS,
mass_h=0.47*ELECTRON_MASS, // using hh=0.45m; lh=0.082m
temperature=300.0,
mix=0.07, // mixing factor for relaxation procedure
F1=0.74e15, // defect current density from surface (if used) units: num/area/sec
D=1e-19, // defect diffusion coefficient
Etr=0.34, // trap depth measured from conduction band edge
Ly=14e-9, // damage source = G1*exp(-Y/Ly) (if used)
G1=1.0e24, // strength of damage source at surface (if used) units: num/volume/sec
etchDepth=135e-9,
etchRate=2.83e-9, // =3.33 nm/s usually
etchTime=etchDepth/etchRate;
int constantFlux=NO, // =NO if exponential source is used instead of constant flux
includeDamage=YES;

```

```

const double e_div_eta = CHARGE/(PERMITTIVITY*ETA_RELATIVE);
double kT = BOLTZMANN_CONSTANT*temperature,
beta = CHARGE/kT,
Nc = 2.0*exp(1.5*log(mass_e*kT/(2.0*PI*hBAR*hBAR))),
Nv = 2.0*exp(1.5*log(mass_h*kT/(2.0*PI*hBAR*hBAR))),
Nc_2_div_sqrtPI = Nc*2.0/sqrt(PI),
Nv_2_div_sqrtPI = Nv*2.0/sqrt(PI);

```

```

// -----
// Main declarations.
// -----
// Nd[] = Charge density in wire. Also acts as the source term for the Poisson eqn.
// Ec[] = Conduction band edge. The 2D Poisson equation is solved to obtain this.
// Gd[] = Defect density in wire. Assumed to be a solution to the diffusion equation.
// eD,eD2,hD,hD2: Used for cubic spline interpolation of electron and hole densities.
// Ele,...: Determines range over which spline interpolation is carried out.
// -----

```

```

#define MAX_ARRAY_SIZE 2000
#define NUM_SPL 200

```

```

typedef double Array[MAX_ARRAY_SIZE];
typedef double (*funcPtr)(double);
typedef double splArray[NUM_SPL];

```

```

const double Ele=-1.0, E2e=2.0, dEe=(E2e-Ele)/NUM_SPL,
Elh=-2.4, E2h=0.6, dEh=(E2h-Elh)/NUM_SPL;

```

```

Array Nd={0}, Ec={0}, Ec_old={0}, Gd={0};
splArray eD={0}, eD2={0}, hD={0}, hD2={0};

```

```

// -----
// Function prototypes.
// -----
void ConvertToChargeDensity(Array f, int jmax);
double residual(Array f, Array u, int j, int l);
void solve(Array f, Array u, int jmax, int mits, double eps);
void integrate(funcPtr func, double a, double b, double& ss, int proper);
void buildSpline(double* y, double* y2, int n, double dx);
double splineValue(splArray y, splArray y2, double x, double xlo, double dx);

// -----
// Miscellaneous routines.
// -----
// DESCRIPTION:
// These routines do small manipulations on the data.
// FUNCTIONS:
// fitToMesh() - Ensures that the passed parameter is a multiple of the mesh size.
// MakeConstant() - Set the entire array to a constant value.
// -----

```

```

void fitToMesh(double& aLength) { aLength=meshSize*int(aLength/meshSize+0.5); }

void MakeConstant(double* a, int jmax, double setValue) {
    for (int j=0; j<=jmax; j++) a[j] = setValue;
}

// -----
// Routines to calculate carrier concentrations.
// -----
// DESCRIPTION:
// Calculate the electron and hole densities by numerically evaluating the Fermi-Dirac
// integral of order 1/2, represented as F_1_2 below.
// FUNCTIONS:
// Eg() - Returns the GaAs energy gap as a function of temperature.
// FDintegrand() - Integrand of the Fermi-Dirac integral. The integral is evaluated at
// FDarg, which is set before the integrating routine is called.
// F_1_2() - Evaluates the Fermi-Dirac integral at the specified value.
// Nelectron() - Returns the number of electrons given the local conduction band
// energy, by calling F_1_2.
// Nhole() - As for Nelectron but for holes.
// electronDensity() - Returns the electron density, by calling Nelectron() or by
// calling the previously calculated spline function.
// holeDensity() - As for electronDensity but for holes.
// BuildSplineForElectronDensity() - Fits a cubic spline to the electron density as a
// function of Ec. Using the spline function instead of continually
// evaluating the Fermi-Dirac integral speeds up the relaxation
// procedure for the Poisson equation.
// BuildSplineForHoleDensity() - Fits a cubic spline to the hole density.
// -----

double Eg(double T) { return (1.52-(5.4e-4*T*T)/(T+204.0)); }

double FDarg=0.0; // to be set before calling FD_integrand
double FD_integrand(double z) { return sqrt(z)/(1.0+exp(z+FDarg)); }

double F_1_2(double x) {
    double result=0.0;
    FDarg=x;
    integrate(funcPtr(FD_integrand),0.0,15.0,result,YES);
    return result;
}

double N_electron(double E) { return Nc_2_div_sqrtPI*F_1_2(beta*E); }
double N_hole(double E) { return Nv_2_div_sqrtPI*F_1_2(-beta*E); }

double electronDensity(double E) {
    return (E<E1e || E>E2e-dEe) ? N_electron(E) : splineValue(eD,eD2,E,E1e,dEe);
}

double holeDensity(double E) {
    return (E<E1h || E>E2h-dEh) ? N_hole(E) : splineValue(hD,hD2,E,E1h,dEh);
}

void BuildSplineForElectronDensity() {
    double p=E1e;
    for (int i=0; i<NUM_SPL; i++, p+=dEe) eD[i]=N_electron(p);
    buildSpline(eD,eD2,NUM_SPL,dEe);
}

void BuildSplineForHoleDensity() {
    double p=E1h;
    for (int i=0; i<NUM_SPL; i++, p+=dEh) hD[i]=N_hole(p);
    buildSpline(hD,hD2,NUM_SPL,dEh);
}

```

```

// -----
// Routines to calculate defect concentrations.
// -----
// DESCRIPTION:
// Calculate the defect density Gd[j][1], arising from each sidewall by numerically
// evaluating either of two integrals, depending on the status of the variable
// constantFlux: constantFlux=YES causes damage_constantDefectFlux() to be evaluated;
// constantFlux=NO causes damage_exponentialSource to be evaluated.
// FUNCTIONS:
// damage_constantDefectFlux() - Represents an integral solution of the homogeneous
// diffusion equation in the quarter plane {x>=0,y>=0}, with non-homogeneous
// von-Neumann boundary conditions, -D.dG/dx=F1 at the sidewalls, where G is
// written as Gd[j][1] in the program.
// damage_exponentialSource() - Represents an integral solution of the nonhomogeneous
// diffusion equation in the quarter plane {x>=0,y>=0}, with source term
// g=G0.exp(-x/Ly), and homogeneous von-Neumann boundary conditions.
// BuildDefectDensityFunction() - Evaluates one of the above integrals at each mesh
// point in the wire. The result is stored in array Gd[j].
// BuildDefectDensityFunction2() - evaluates the defect density using the simplified
// expression for top surface damage distribution.
// -----

double Y=0.0, T=0.0, w=etchRate/Ly-D/(Ly*Ly), D_div_PI=D/PI;
double damage_convolutionIntegral(double t) {
    double _4Dt=4.0*D*t, Y_plus_vt=Y+etchRate*t,
           f_Tt=exp(w*(t-T))-1.0,
           h1_t=0.5*etchRate*erfc((Y_plus_vt)/sqrt(_4Dt)),
           h2_t=sqrt(D_div_PI/t)*exp(-Y_plus_vt*Y_plus_vt/_4Dt);
    return f_Tt*(h1_t-h2_t);
}

double damage_constantFlux(double t) {
    double _4Dt=4.0*D*t, Y_plus_vt=Y+etchRate*t,
           h1_t=exp(-Y_plus_vt*Y_plus_vt/_4Dt),
           h2_t=0.5*etchRate*erfc(Y_plus_vt/sqrt(_4Dt));
    return sqrt(D_div_PI/t)*h1_t-h2_t;
}

void BuildDefectDensityFunction(Array Gd, int jmax) {
    if (etchTime) {
        double convolution_result=0.0;
        for (int j=1; j<=jmax; j++) {
            Y=double(j)*meshSize;
            if (constantFlux) {
                integrate(damage_constantFlux,0.0,T=etchTime,convolution_result,NO);
            } else {
                integrate(damage_convolutionIntegral,0.0,T=etchTime,convolution_result,NO);
            }
            Gd[j] = G1*(exp(-Y/Ly)*(1.0-exp(-w*T))+convolution_result/Ly)/w;
        }
    } else { MakeConstant(Gd,jmax,0.0); }
}

void BuildDefectDensityFunction2(Array Gd, int jmax) {
    for (int j=0; j<=jmax; j++) {
        Y=double(j)*meshSize;
        Gd[j] = G1*Ly*exp(-Y/Ly)*(1.0-exp(-etchRate*etchTime/Ly))/etchRate;
    }
}

// -----
// Routines to calculate the total charge concentration.
// -----
// DESCRIPTION:
// These routines provide the source term for the 2D Poisson equation.
// FUNCTIONS:

```

```
// CalculateChargeConcentration() - Sums the electron, hole, defect, and trapped-electron
// densities, with -ve charge for defects and trapped electrons. Electrons are
// assumed to be trapped with the equilibrium Fermi-Dirac distribution.
// ConvertToChargeDensity() - Scales the number densities above to give charge density.
// -----
```

```
void CalculateChargeConcentration(Array Nd, int jmax) {
    int jdepth = int(jmax*(epiThickness-etchDepth)/simulationDepth+0.5);
    for (int j=0; j<=jmax; j++) {
        Nd[j] = -electronDensity(Ec[j]) + holeDensity(Ec[j]-Eg(temperature));
        if (includeDamage) Nd[j] -= Gd[j]*(1.0+1.0/(1.0+exp(beta*(Ec[j]-Etr))));
    } for (j=0; j<=jdepth; j++) Nd[j] += dopingDensity;
}
```

```
void ConvertToChargeDensity(Array f, int jmax) {
    double scaled_e_div_eta = e_div_eta*meshSize*meshSize;
    for (int j=0; j<=jmax; j++) f[j] *= scaled_e_div_eta;
}
```

```
// -----
// Numerical routines.
// -----
// DESCRIPTION:
// C++ implementations of routines taken from Numerical Recipes (Pascal). Page numbers
// below refer to the book, which should be consulted for further details.
// FUNCTIONS:
// totalResidual() - Finds the sum of the residuals over the entire mesh.
// solve() - Solves Poisson equation with Eg=const on the surface, and dEg/dx=0
// near the substrate.
// extrapolate() - Used for polynomial extrapolations ('polint' p92).
// trapzd() - Used for closed interval definite integrals ('trapzd' p125).
// midpnt() - Used for open ended definite integrals ('midpnt' p132).
// integrate() - Integrates the given function using the method of Romberg. If
// proper=YES then integrates over a closed interval ('qromb' p129),
// otherwise over an open interval ('qromo' p133).
// buildSpline() - Builds an interpolating cubic spline using the given data points
// ('spline' p99).
// splineValue() - Returns the interpolated spline fit at value x ('splint' p100).
// -----
```

```
double totalResidual(Array Nd, Array Ec, int jmax) {
    double resid=0.0;
    for (int j=1; j<jmax; j++) resid += fabs(Ec[j+1]-2.0*Ec[j]+Ec[j-1]-Nd[j]);
    return resid;
}
```

```
void solve(Array Nd, Array Ec, int jmax) {
    Nd[1]=-(Ec[0]=surfacePotential), Ec[1]=-2.0, Nd[jmax]=0;
    for (int j=1; j<jmax-1; j++) { Nd[j+1] -= (Nd[j] / Ec[j]); Ec[j+1]=-2.0-1.0/Ec[j]; }
    Nd[jmax] += Nd[jmax-2]-(Nd[jmax-1]/Ec[jmax-1])/Ec[jmax-2];
    Ec[jmax] = Nd[jmax]/(1.0-1.0/(Ec[jmax-1]*Ec[jmax-2]));
    for (j=jmax-1; j>0; j--) Ec[j]=Nd[j]-Ec[j+1]/Ec[j];
}
```

```
void extrapolate(double *xa, double *ya, int n, double x, double& y, double& dy) {
    double *c = new double[n], *d = new double[n];
    int ns=1;
    double dif = fabs(x-xa[0]), dift=0.0;
    for (int i=0; i<n; i++) {
        if ((dift=fabs(x-xa[i]))<dift) { ns=i; dif=dift; }
        c[i] = d[i] = ya[i];
    } y=ya[ns--];
    for (int m=1; m<n; m++) {
        for (i=0; i<n-m; i++) {
            double ho=xa[i]-x, hp=xa[i+m]-x, den=(c[i+1]-d[i])/(ho-hp);
            d[i]=hp*den; c[i]=ho*den;
        } y += (dy=(2*ns<n-m)?c[ns+1]:d[ns--]);
    }
```

```

} delete c; delete d;
}

void trapzd(funcPtr func, double a, double b, double& s, int n, int& TrapzdIt) {
  if (n==1) { TrapzdIt=1; s = 0.5*(b-a)*(func(a)+func(b)); }
  else {
    double del=(b-a)/TrapzdIt, x=a-0.5*del, sum=0.0; // NB here x=a+0.5*del-del
    for (int j=0; j<TrapzdIt; j++) sum += func(x+=del);
    s = 0.5*s+(b-a)*sum/(TrapzdIt*=2);
  }
}

void midpnt(funcPtr func, double a, double b, double& s, int n, int& MidpntIt) {
  if (n==1) { MidpntIt=1; s = (b-a)*func(0.5*(a+b)); }
  else {
    double del=(b-a)/(3.0*MidpntIt), ddel=del+del, x=a-0.5*del, sum=0.0;
    for (int j=0; j<MidpntIt; j++) { sum+=func(x+=del); sum+=func(x+=ddel); }
    s = s/3.0+(b-a)*sum/(MidpntIt*=3);
  }
}

void integrate(funcPtr func, double a, double b, double& ss, int proper) {
  const double eps=(proper)?1e-3:0.01, df=(proper)?4.0:9.0;
  const int jmax=20, jmaxp=21, k=5;
  double *h = new double[jmaxp], *s = new double[jmaxp];
  double *c = new double[k], *d = new double[k];
  h[0]=1.0; double dss=ss+1.0; int iter=0;
  for (int j=0; j<jmax && fabs(dss)>eps*fabs(ss); j++) {
    if (proper) { trapzd(func,a,b,s[j],j+1,iter); }
    else { midpnt(func,a,b,s[j],j+1,iter); }
    if (j>k-1) {
      for (int i=0; i<k; i++) { c[i]=h[j-k+1]; d[i]=s[j-k+1]; }
      extrapolate(c,d,k,0.0,ss,dss);
    } s[j+1]=s[j]; h[j+1]=h[j]/df;
  } delete d; delete c; delete s; delete h;
}

void buildSpline(double* y, double* y2, int n, double dx) {
  double *u = new double[n], p=0.0;
  y2[0]=-0.5, u[0]=0.0;
  for (int i=1; i<n-1; i++) {
    y2[i] = -0.5/(p=0.5*y2[i-1]+2.0);
    u[i] = (y[i+1]+y[i-1]-2.0*y[i])/dx;
    u[i] = (3.0*u[i]/dx-0.5*u[i-1])/p;
  } y2[n-1] = (-0.5*u[n-2])/(0.5*y2[n-2]+1.0);
  for (i=n-2; i>=0; i--) y2[i] = y2[i]*y2[i+1]+u[i];
  delete u;
}

double splineValue(double* y, double* y2, double x, double xmin, double dx) {
  int klo=int((x-xmin)/dx);
  double xlo=double(klo)*dx+xmin, a=(xlo+dx-x)/dx, b=(x-xlo)/dx;
  return a*y[klo]+b*y[klo+1]+((a*a*a-a)*y2[klo]+(b*b*b-b)*y2[klo+1])*dx*dx/6.0;
}

// -----
// Main calculational routines.
// -----
// DESCRIPTION:
// These routines allow the 2D Poisson equation with an arbitrary source term to be
// solved using a relaxation method. This gives the conduction band edge Ec[j][1].
// A numerical double integral over the resulting local electron density then gives
// the conductance.
// FUNCTIONS:
// SetBoundaryValues() - Set the boundary values of Eg[[]].
// FindZerothOrderSolution() - Artificially puts some charge into the epilayer region
// of the wire and integrates the Poisson equation to find a zeroth

```

```

// order solution for the full problem.
// FindFinalSolution() - Iterates solutions of the Poisson equation, starting from the
// zeroth order solution, until convergence is reached. A relaxation
// procedure is used to iterate successive solutions: (1) A solution
// is obtained to Ec[][] by calling solve(). (2) This gives a new
// electron distribution which is used in the next iteration. (3) The
// values of Ec[][] used in the next iteration are weighted averages
// of the current solution with the previous solution. (4) Repeat
// until the totalResidual is within the specified tolerance. This
// gives the desired solution.
// Nelectron_dx() - Used for evaluating the x-integral in the double integral for the
// conductance. This forms the integrand for Nelectron_dx_dy().
// Nelectron_dx_dy() - Used for evaluating the double integral for the conductance.
// sheetConductance() - Evaluates the sheet conductance by first interpolating the
// conduction band edge using a mesh of cubic splines, and then
// integrating the resulting spline function using Nelectron_dx_dy().
// SolveProblem() - Calls the above functions to get the solution. First find a
// zeroth order solution to start the relaxation procedure with.
// Solve the Poisson equation to obtain the conduction band edge.
// Finally use this to obtain the electron density and hence the
// conductance of the wire.
// -----

```

```

void FindZerothOrderSolution(Array Nd, Array Ec, int jmax) {
    MakeConstant((double*)Ec, jmax, 0.0);
    MakeConstant((double*)Nd, jmax, 0.0);
    double depletionEstimate = sqrt(2.0*surfacePotential/(e_div_eta*dopingDensity));
    int jdepl = int(jmax*depletionEstimate/simulationDepth+0.5),
        jdepth = int(jmax*(epiThickness-etchDepth)/simulationDepth+0.5);
    jdepl = (jdepl<2)?2:jdepl;
    for (int j=0; j<jdepl; j++) Nd[j] =dopingDensity;
    ConvertToChargeDensity(Nd, jmax);
    solve(Nd, Ec, jmax);
}

```

```

void FindFinalSolution(Array Nd, Array Ec, int jmax, double mix) {
    double eps=1e-4;
    double anorm_bef=0.0, anorm_aft=anorm_bef+1.0;
    int jdepth = int(jmax*(epiThickness-etchDepth)/simulationDepth+0.5);
    for (int iter=1; iter<2000 && anorm_aft>eps*anorm_bef; iter++) {
        CalculateChargeConcentration(Nd, jmax);
        ConvertToChargeDensity(Nd, jmax);
        for (int j=0; j<=jmax; j++) Ec_old[j]=Ec[j];
        if (iter==1) { anorm_bef=totalResidual(Nd, Ec, jmax); }
        else { anorm_aft=totalResidual(Nd, Ec, jmax); }
        solve(Nd, Ec, jmax);
        for (j=0; j<=jmax; j++) Ec[j] = (1.0-mix)*Ec_old[j]+mix*Ec[j];
    }
}

```

```

double *splE, *splE2;
double Nelectron_dy(double y) {
    return electronDensity(splineValue(splE, splE2, y, 0.0, meshSize)); }

```

```

double sheetConductance(Array Ec, int jmax) {
    double Nintegral=0.0, *Ec2;
    buildSpline(Ec, Ec2 = new double[jmax], jmax, meshSize);
    splE=Ec, splE2=Ec2;
    integrate(Nelectron_dy, 0.0, simulationDepth, Nintegral, YES);
    delete Ec2;
    return CHARGE*mobility*Nintegral;
}

```

```

void SolveProblem() {
    fitToMesh(simulationDepth); fitToMesh(epiThickness); fitToMesh(etchDepth);
    int jmax=int(simulationDepth/meshSize+0.5);
    FindZerothOrderSolution(Nd, Ec, jmax);
    printf("integrating for nel ... \n"); BuildSplineForElectronDensity();
}

```

```
printf("integrating for hol ...\n"); BuildSplineForHoleDensity();
if (includeDamage) {
    printf("building defect density ...\n"); BuildDefectDensityFunction(Gd, jmax); }
FindFinalSolution(Nd, Ec, jmax, mix);
printf("integrating for conductance ...\n");
printf("SHEET CONDUCTANCE = %e\n", sheetConductance(Ec, jmax));
}

// -----
void main() {
    SolveProblem();
}
```

D. References

- ABRAMOWITZ M, I Stegun, *Handbook of Mathematical Functions*, Dover Publications Inc., New York (1965).
- GERALD C F, *Applied Numerical Analysis*, Addison-Wesley Publishing Co., Reading, Massachusetts (1970).
- MACKIE A G, *Boundary Value Problems*, Scottish Academic Press, Edinburgh (1989).
- PRESS W H, B P Flannery, S A Teukolsky, W T Vetterling, *Numerical Recipes in Pascal*, Cambridge University Press, Cambridge (1989).

