Fabrication and Characterization of Ultra-small Tunnel Junctions for Single Electron Devices.

Robert Leiper Nicol.

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

September 1997.

To Irene,

and in memory of

Dad.

SUMMARY

This thesis presents research into the development and characterization of fabrication processes for high temperature single electron devices. Some background to single electron effects and the fabrication processes used in this work is presented. Analysis, development and characterization work on the fabrication processes for ultrasmall single electron devices is also presented. Electrical measurements on the films used and the single electron structures fabricated are then presented and discussed. The work on the fabrication processes is then discussed in the light of these electrical results. Conclusions are drawn with regard to the practicalities and technologies required for true high temperature operational single electron devices.

Work on the fabrication processes has shown that traditional tunnel junction formation techniques result in structure sizes which are too large to provide the high temperature effects required. Where lithographic techniques alone are used to shrink pattern dimensions, the processes become unreliable. In the case of the suspended mask shadow evaporation process used here, a limiting reliable overlap width of 40nm is expected and experienced. Attempts to fabricate structures below this size resulted in unreliable tunnel junction formation. The second technique investigated, the crossed track technique, suffered from serious problems arising from the angled evaporation process and from step coverage difficulties. The third fabrication technique attempts to control the placement of grains within the aluminium film. This technique has the advantages of simplicity and ability to form the smallest tunnel junctions with the material system used here. This system was chosen as the main fabrication process for investigation of high temperature single electron devices in this work.

Measurements of resistivity and resistivity temperature dependence of the aluminium films were used to characterize the film types. The temperature dependence and magnitude of the resistivity have shown the films to be very conductive, or metallic. By virtue of this high conductivity, the structure behaviour should be dominated by the device, or tunnel junction, properties. The results obtained from the devices at 4.2K do

not show the presence of single electron effects. However, the fabricated structures did demonstrate tunnelling behaviour. The absence of single electron effects has been attributed to the structure sizes. Despite being among the smallest possible in aluminium metallizations, these granular structures are apparently too large. The explanation for this is derived from the presence of stray capacitances between the grains forming the tunnel junctions. This raises the junction capacitance and therefore reduces the charging energy of the junction and the temperature of operation.

Despite the failure to observe high temperature single electron effects, this work has shown that it is possible to control the placement of aluminium grains within a film to form tunnel junction structures. Were a self organizing system, such as this, to be produced on a smaller scale, it would be possible to significantly increase the temperature of operation of single electron devices.

ACKNOWLEDGEMENTS

I would like to thank the staff and students of the Department of Electronics and Electrical Engineering, University of Glasgow, for helping make my research an enjoyable experience and this thesis possible. I would like to thank my supervisor, Professor Steve P. Beaumont, for his help with this project and for his patience during the writing of this thesis. Many thanks are due to the excellent support given to this work from the members of the Nanoelectronics Research Centre, the staff of the Mechanical and Electrical workshops and the Vacuum Laboratories. Special thanks are needed for Dave, Susan, Robert and Helen for many hours spent interpreting instructions on how to run yet another of my Beamwriter jobs. Thanks are also due to Steve Thoms and Douglas MacIntyre for their help and technical expertise. Thanks also go to John Weaver, Mahfuzur Rahman, Iain Thayne and Nigel Cameron for helping me find all kinds of electrons and showing me what to do with them.

Finally, I would like to thank my wife, Irene, and my parents for their love, support and understanding throughout my academic career. Special thanks go to Irene for correcting my chemistry, grammar and spelling and for many, many cups of coffee.

CONTENTS

SUMMARY III

ACKNOWLEDGEMENTS V

CHAPTER ONE INTRODUCTION 1

1.1 Developments in electronics 1
1.1.1 The history of the transistor and integrated circuit 2
1.1.2 Miniaturization or scaling 5
1.1.3 Limitations to device scaling 11
1.2 Quantum Effect Devices 13
1.2.1 Reduced dimensionality 13
1.2.2 Quantum devices 16
1.3 Motivation and aims of this work 18
1.4 Organization of this thesis 19
1.5 References 23

CHAPTER TWO TUNNELLING AND SINGLE ELECTRON DEVICES 25

2.1 Introduction 25
2.2 Tunnelling 26
2.2.1 Tunnelling Probability 28
2.2.2 Tunnel Current 33
2.2.3 WKB approximation 36
2.3 Single electron tunnelling 38
2.3.1 Coulomb blockade of tunnelling 40
2.3.2 Double tunnel junctions 46
2.3.3 One Dimensional Arrays 58
2.3.4 Two dimensional arrays and granular films 61

2.3.5 Development of single electron devices 63
2.4 References 65

CHAPTER THREE FABRICATION TECHNIQUES 69

3.1 Introduction 69
3.2 Sample preparation 70
3.2.1 Organic contamination clean 71
3.2.2 Inorganic contamination clean 71
3.2.3 Oxidation 71
3.2.4 Deoxidation or oxide etch 72
3.3 Lithography 72
3.3.1 Electron beam lithography 72
3.3.2 Photolithography 83
3.4 Metallization and lift-off 86
3.4.1 Lift-off processing 86
3.4.2 Interconnect metallization 87
3.4.3 Device metallization 89
3.5 Evaporation systems 90
3.5.1 Plassys MEB450 90
3.5.2 High vacuum evaporator 91
3.5.3 Ultra High Vacuum Evaporator 92
3.6 Suspended mask fabrication 96
3.6.1 Mask layers 97
3.6.2 Mask patterning 97
3.6.3 Mask suspension 98
3.6.4 Mask removal 98
3.7 Pattern details 98
3.7.1 Main interconnect layer 99
3.7.2 Thin film interconnect layer 99
3.7.3 Bonding layer 100
3.7.4 Photolithography masks and sample layout 101
3.8 Process details 102
3.8.1 Lead frame ($5X5MM^2$ OR $2X2MM^2$) 103
3.8.2 High resolution resist processes 105
3.8.3 Process for "single layer device and interconnect" 106
3.9 Conclusion 107
3.10 References 108

CHAPTER FOUR SINGLE ELECTRON DEVICE FABRICATION 111

4.1 Introduction 111
4.2 Suspended mask structures 112
4.2.1 Angled evaporation device fabrication 113
4.2.3 Control over device fabrication 114
4.3 Crossed track structures 120
4.3.1 Crossed track technique 120
4.3.2 Pattern design and development 121
4.3.3 Resist structure development 123
4.3.4 Fabrication problems 126
4.3.5 Rejection of crossed track process 131
4.4 Patterned granular film structures 132
4.4.1 Patterned thin film structure 133
4.4.2 Nucleation and growth of thin films 134
4.4.3 Characterization of film structure 139
4.4.4 Characterization of pattern and structure development 149
4.5 Conclusions 158
4.6 References 160

CHAPTER FIVE THIN FILM ELECTRICAL CHARACTERIZATION 161

5.1 Introduction 161
5.2 Metallic film resistivity 162
5.2.1 Metallic conduction 162
5.2.2 Conduction Processes in thin metal films 168
5.2.3 Conduction processes in granular films 175
5.2.4 Conduction effects observed in aluminium 179
5.3 Resistivity measurement techniques 181
5.3.1 Sheet resistance 183
5.3.2 Frequency "lock-in' techniques 185
5.4 Resistivity measurement system 187
5.4.1 Cryostat 188
5.4.2 Measurement Circuit 190
5.4.3 The HP4145 semiconductor parameter analyser 192
5.5 Resistivity measurements 193
5.5.1 High vacuum deposited films. 193

5.5.2 Ultra-high vacuum deposited films. 201
5.5.3 Thin film interconnect measurements. 206
5.6 Conclusions 209
5.7 References 211

CHAPTER SIX SINGLE ELECTRON DEVICE MEASUREMENTS 213

6.1 Introduction 213
6.2 Device structures 214
6.2.1 Fabrication test circuit 214
6.3 Measurement system design 216
6.3.1 System Overview 216
6.3.2 Signal generator 218
6.3.3 Filter stage 218
6.3.4 Variable dc. reference voltage 225
6.3.5 Mixer circuit 226
6.3.6 Driving circuit 227
6.3.7 Pre-amplifier 231
6.3.8 Instrumentation amplifier 241
6.3.9 Power circuitry 242
6.3.10 Mechanical construction 243
6.3.11 System operational characteristics 249
6.4 Device measurement results 249
6.4.1 Suspended mask samples 248
6.4.2 Granular "SET" electrometers 251
6.4.3 Granular single tunnel junctions 254
6.5 Conclusions 255
6.6 References 258
Appendicies for Chapter 6

CHAPTER SEVEN DISCUSSION 259

7.1 Introduction 259
7.2 Structure formation and control 260
7.3 Structure capacitance 264
7.4 Further development of the granular technique 269
7.5 Possibilities for high temperature SET devices 271
7.6 Conclusions and further work 272

7.6.1 Conclusions 273
7.6.2 Further work. 275
7.7 References 277

CHAPTER ONE

Introduction

1.1 DEVELOPMENTS IN ELECTRONICS

The modern electronics industry is exhibiting unprecedented levels of growth and development, with the semiconductor sector displaying up to 40% per annum growth in market value[1]. The growth in demand and applications for existing products results in a continual need for product and process development to meet the requirements of customers while maintaining profitability. The speed of development of electronic devices has arisen from the ability to integrate and miniaturize devices and systems. Currently it is possible to fabricate 1 to 3 million components onto a single square centimetre of silicon. This provides multifunctionality in a physically small area with good communication between system units and reduction in cost through mass production and economies of scale.

The ability to miniaturize is an overwhelming driving force for the industry which has, in some respects, become self-fulfilling. The need to continually scale processes in order to maintain profitability allows for increases in functionality and complexity. This trend is now assumed as standard and places requirements on the industry to continue a high rate of development and growth. The consumer electronics and

computer markets exemplify this situation. With product lifetimes in these markets being as short as three months to one year, relentless revisions increase the functionality, power and complexity of products, with no real cost increase. Indeed, in most cases the cost is effectively reduced when the complexity or functionality gain is considered.

The following sections of this chapter give a brief discussion of the development of the industry from its early beginnings. The current miniaturization trend and its future implications are also discussed. Research into new modes of device operation and new devices, of which this work is a part, are discussed. From this discussion the motivation and aims of this work are set out. Finally, the organization of this thesis is discussed.

1.1.1 THE HISTORY OF THE TRANSISTOR AND INTEGRATED CIRCUIT

It is tempting to think of the electronics and semiconductor industries emerging from the development of the transistor by J. Bardeen and W.H. Brattain[2] and the initial developments of transistor theory by W. Shockley[3], J.L Moll and others. This view, however, ignores the long history of development in the field of electrical technology and electronics.

In 600BC, Thales of Miletus[4] discovered the ability of rubbed pieces of amber to attract straw. This represents one of the first documented experiments with electricity in any form. Indeed, the root word for the subject, electron, is derived from the Greek *elektron*, meaning amber. In 1751, Benjamin Franklin described electrical current as a fluid however, the basic ideas of atomistic and particulate theories of matter had been suggested by Robert Boyle nearly a century before. In his *The origins of forms and quantities* of 1666, Boyle describes his "corpuscular" theory of matter. This theory was adopted and modified by Issac Newton who added to the particulate theory of matter the possibility of attraction between particles[5]. Laws governing the interaction between charged particles were quantitatively described in the late eighteenth century through many experiments, most notably those of Charles Augustine Coulomb in 1785.

The early nineteenth century saw the subjects of electrical technology and engineering develop. In 1800, Alessandro Volta developed the Voltaic pile, or battery. This provided a means of storing electrical charge and providing a source of current for experiments into the nature and applications of electricity. In 1820, the hitherto separate subjects of electricity and magnetism were combined in Hans Christian Oersted's experiment using a wire carrying electrical current to deflect a compass

needle. This demonstrated that an electric current generates a magnetic field. Michael Faraday's experiments from 1821[6] into electricity and electromagnetic induction lead very quickly to the development of electrical power generating equipment, transformers and telegraphy for long distance communications. However, the fundamental natures of electrical current and light were and in some respects still are, subjects of contention.

Observations by Faraday in 1846 and Gustav Robert Kirchoff in 1857 suggested a connection between light and electromagnetism. James Clerk Maxwell then described light as a form of radiation containing both electric and magnetic components. His theory on *The nature and propagation of electromagnetic radiation* of 1864[7], addressed the wave nature of light and suggested the existence of other forms of electromagnetic radiation. In order to prove the theories of Maxwell, Heinrich Hertz conducted experiments in the propagation of what was to be termed radio-waves and by 1887 had produced equipment capable of transmitting and receiving electromagnetic waves. Experiments carried out by Joseph J. Thomson, in the late 1800's, on cathode rays lead to the discovery of the electron as the first sub-atomic particle. The ability to deflect free high energy electrons, or cathode rays, by means of electric and magnetic fields resulted from this work by 1887. In 1899, he measured the charge to mass ratio, e/m, of the electron. Further experiments, completed in 1913 by Robert A Millikan, produced an accurate measurement of the electron charge. The link between the particulate and wave nature of light, and matter, was provided by the development of Quantum Mechanics in the early 1900's[8].

After the discovery of the electron, Drude proposed a theory for metallic conduction in 1900[9]. This theory, which provided a good model for metallic conduction, was modified by Sommerfeld to include electron velocity distributions based on Fermi-Dirac statistics, rather than the classical Maxwell-Boltzman statistics. This development arose from the Pauli exclusion principle, a quantum mechanical argument. Further modifications to theories of conduction and greater use of quantum mechanics lead eventually to the band theory of solids[10]. This was an essential and fundamental step in the development of the current understanding of solid state electronic devices.

By the end of the 1920's, the subject of electronics was firmly developed. Advances in materials technology and the application of the theories and discoveries noted above lead to the invention of thermionic valves (diodes, triodes etc.), capable of rectifying or amplifying electrical signals. The main application for these "electronic" devices lay in the new fields of radio transmission and radio location, detection and ranging (RADAR). As such, by 1945, a large array of electronic circuits such as timers, flip-flops, oscillators, amplifiers (including feedback amplifiers[11,12]),

modulators/demodulators, signal detectors etc, all existed using valve based circuits. Electrical circuit theories and techniques were also well developed, to such an extent that developments in these electronic devices could be understood quickly and applied.

In 1948 the breakthrough in solid state electronics was provided by the point contact transistor, developed by Brattain and Bardeen. Some of these early developments had their root in earlier suggestions and work. For example, field effect transistors (FETs) are based on concepts first proposed in 1930 and 1935 by J. Lilienfeld and O. Heil respectively[13]. Prior to the development of the transistor, semiconductor based devices did exist but were limited to thermistors for temperature dependent applications and diodes for current rectification. The transistor, being a three terminal device, allowed control of the device state through the third terminal.

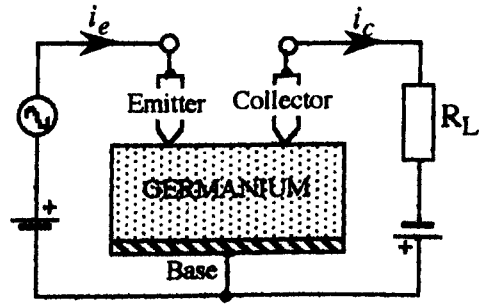


Figure 1.1. Bardeen & Brattain's Point Contact Transistor[2].

The device created by Bardeen and Brattain, depicted in figure 1.1, consisted of a specially prepared n-type germanium block of around 10Ωcm resistivity. The emitter and collector contacts were of the point contact rectifying type while the base of the structure, the third terminal, was formed by the germanium block. This device operated as a current amplifying device, the bipolar transistor, and represented the birth of the modern semiconductor industry. Since then field effect transistors or unipolar devices, high frequency and microwave devices, high power devices and photonic devices have been developed providing a vast array of devices to choose from to fit the application at hand.

Developments in process technologies for crystal growth and wafer fabrication such as doping, etching, patterning, passivating and bonding have resulted in the planar monolithic integrated circuit. In this integrated circuit, only a few microns of the surface of the semiconductor are used to form resistors, capacitors, isolation, active devices and the interconnections. The small dimensions involved allow circuits to be fabricated on a small surface area of the semiconductor. This ability to integrate components together into a small area has resulted in the enormous success and rapid

development of the semiconductor industry.

1.1.2 MINIATURIZATION OR SCALING

Scaling is a controlled form of miniaturization. Scaling requires the reduction of all device dimensions by a common factor, which can then be used to model the miniaturized device behaviour[15]. The ability of the semiconductor industry to continually increase the complexity and functionality of its products is a result of the ability to reliably scale structures. This process is driven by two basic needs; firstly the need to increase integration or packing densities of components in the circuits, secondly the need to increase the speed of operation of the circuits.

1.1.2.1 INCREASED INTEGRATION DENSITY

As we have already discussed, the need to increase the integration density arises from a trade off between profitability and increased functionality. The same circuit when scaled utilizes a smaller surface area of the semiconductor thus providing more devices or circuits per unit area. This gain in device count comes at little extra cost as much the same process steps are required regardless of the circuit size. The cost of manufacture of a unit area of semiconductor is largely the same regardless of the number of components integrated on that area. This increased integration, in many cases, arises from customer demands for significant cost reductions during a product life cycle.

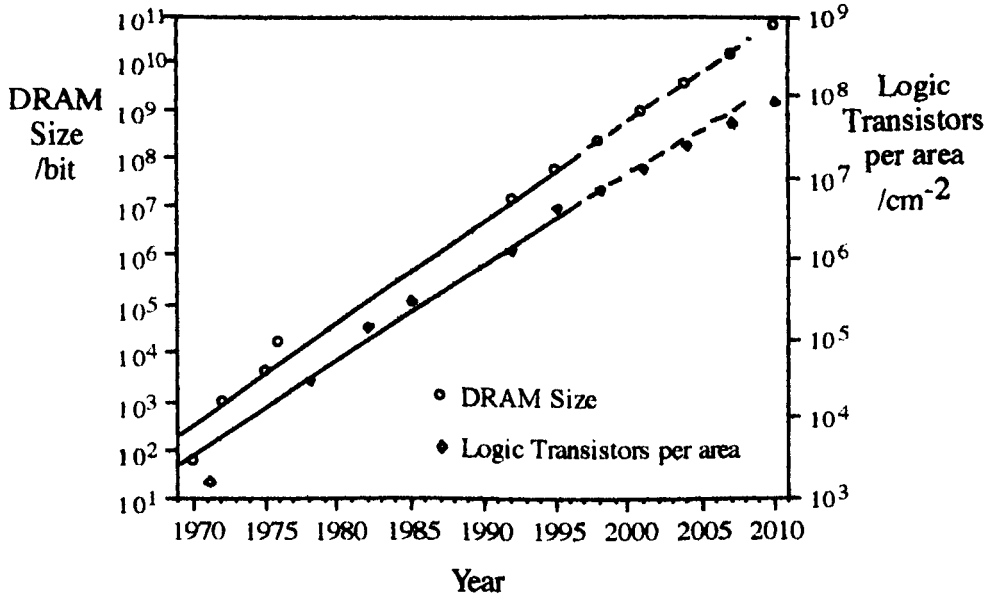


Figure 1.2. Graph of integration density vs. time.

The graph displayed in figure 1.2 gives an indication of the rate of increase in

integration density with time. Two graphs are plotted against time here; dynamic random access memory (DRAM) size and, the number of logic transistors per unit area. The points on the graphs represent the year of first introduction as a production component. DRAM size is generally taken to indicate the state of the industry in terms of raw integration capability. These circuits usually represent the highest device packing densities using the lowest geometries available at the time. The graph of logic transistor integration density is more representative of the bulk of the industry and gives an idea of the ability to integrate functional blocks, such as arithmetic logic units, I/O devices, encoders, decoders, multiplexers, analogue circuits, etc., as opposed to simple storage elements which form the bulk of the memory market.

1.1.2.2 INCREASED OPERATIONAL SPEED

The need to increase the operational speed of devices also arises from the increasing functionality. Higher computational power, data handling rates and more demanding applications require higher operating frequencies from the devices making up the systems. The operational speed of microprocessors, memory chips, input/output devices, amplifiers etc. requires a decrease in device response times to take advantage of higher frequency bandwidth applications and faster information handling rates. For example, increased microprocessor power generally requires some increase in operation frequency to cope with higher information processing rates. This has a corresponding impact on access times for memories and data transfer speeds for input/output devices. Similarly, tele- and data- communication applications require speed increases to take advantage of unused higher frequency bandwidths. As digital circuits enter the telecommunications market, still higher device operational frequencies are required to adequately process and drive digital signals while still taking advantage of high frequency bandwidths, typically around 2.6GHz for current mobile telephone systems.

Speed increase within a transistor type is, to the greatest extent, a result of decreasing feature size. Reduction in feature size results in shorter electron paths and therefore shorter transit times. The operational frequency range of the device, i.e. the bandwidth over which useful gain exists, can be increased through device scaling. Figure 1.3 graphs cut-off frequencies for a number of FET technologies against the gate length to demonstrate this scaling effect. It is clear from this graph that speed increases can also be gained by changing the transistor material technology. For example, the higher electron mobility and lower electron effective mass in gallium arsenide based semiconductors results in a correspondingly higher operation frequency than in silicon based devices.

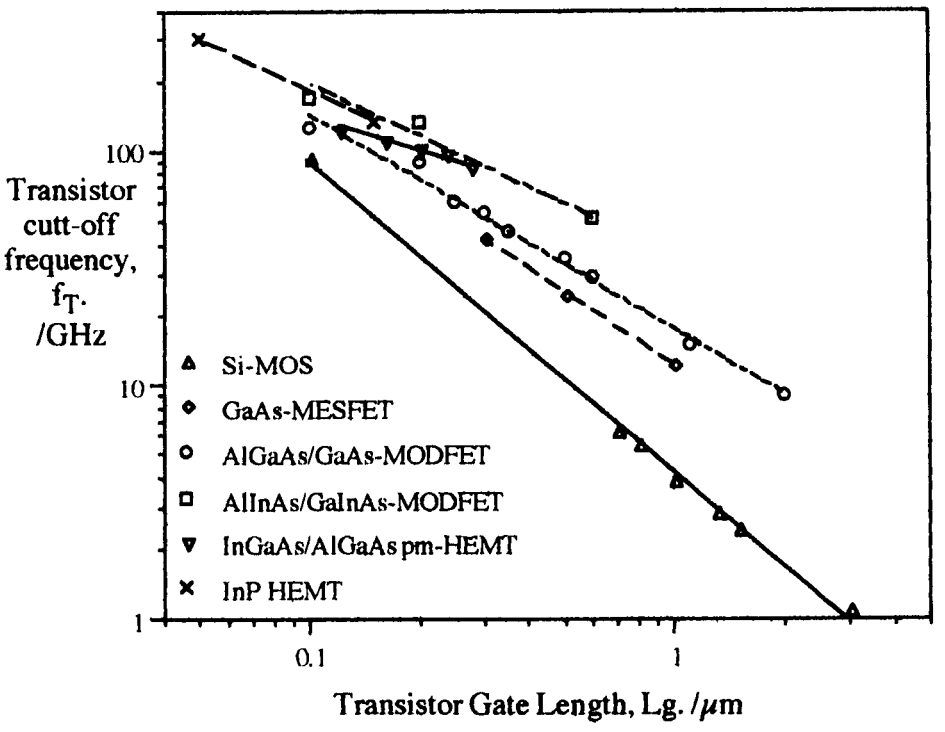


Figure 1.3. Graph of reported cut-off frequencies for differing technologies[15-17].

Reduction in the feature size reduces capacitances within the device leading to a reduction in the characteristic delay times. As device operational speed is generally limited through resistor-capacitor time constants, reduction of capacitances is a major goal of the scaling process. However, reduction in device size may not necessarily result in proportional scaling of parasitic capacitance elements and may result in resistance increase within the device. Therefore, careful process and device development is required to ensure the benefits of miniaturization are gained while minimizing the effect on device performance.

1.1.2.3 SIZE REDUCTION TREND

The current industry direction is a continuation of the scaling phenomena of the last two decades. Figure 1.4 displays the reduction in two critical dimensions related to silicon metal oxide semiconductor (Si-MOS) structures. This graph shows scaling against time for both minimum design rule and gate oxide thickness. The year given is taken as the year of first introduction of a production process. Note that the gate oxide thicknesses given are derived from high speed devices and are around 60% thinner than typical production processes for the given technology design rule. The minimum design rule can be read as the minimum feature size which can be fabricated by a technology or process. This usually represents the gate length in Si-MOS devices. The thickness of the gate oxide in a MOS structure defines the gate to channel capacitance which, with

the gate resistance, is a speed limiting factor. The rate at which the gate oxide can be charged and discharged contributes to the definition of the channel current modulation upper frequency limit. Control of this oxide thickness is a fundamental process parameter, defining the uniformity of device characteristics across a wafer.

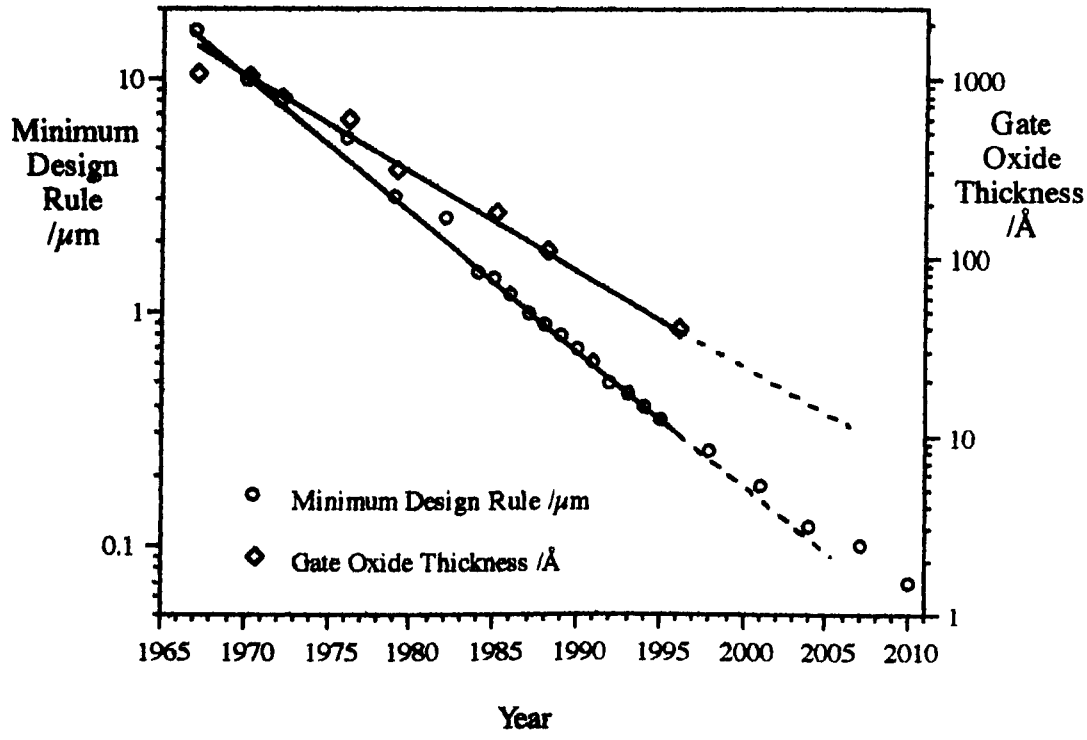


Figure 1.4. Graph of critical feature size reduction vs. time.

1.1.2.4 THE FUTURE FOR SCALING

The Semiconductor Industry Association (SIA) published its "road-map" to the year 2010 in 1995[18]. This is designed to provide structure, focus and direction for the industry to overcome the hurdles which currently prevent mass production of 70nm minimum feature devices. To achieve this degree of size reduction, processing and lithography techniques must be developed beyond the current, largely experimental, state. Size reduction of this scale is not simply related to the patterning process. There are a number of technological hurdles to overcome in working at these small dimensions. For example, to make use of the projected 7 to 8 layer metal processes required, significant improvements in planarization technologies must be achieved. This is necessary not only to aid definition of the interconnect pattern but to help ensure reliability of the wiring levels by prevention of topology induced electromigration.

1.1.2.5 LITHOGRAPHY FOR 70NM PRODUCTION

While other process problems require to be solved, we still require the ability to pattern

at and below the 70nm scale in a production environment. Development of photolithographic tools to meet these requirements is underway. The high industry investment in the photolithographic process and the great deal of experience within the industry tends to keep the pressure and focus on this patterning technique. The major light source in photolithography is still the mercury arc lamp. The spectral response of this lamp is well characterized and produces four usable spectral bands with the following band definitions and wavelengths; g-line (436nm), h-line (404nm), i-line (365nm) and broad-band (240 to 255nm, centre 248nm). Other sources becoming available are formed from excimer lasers which provide monochromatic light with wavelengths dependent on the gas used in the laser. For example KrF and ArF sources produce light at 248nm and 193nm respectively while F_2 produces light at 157nm wavelength.

$$\text{Resolution}, \quad R = \frac{k\lambda}{NA}. \qquad [1.1]$$

$$\text{Depth of focus}, \quad \text{DoF} = \frac{k\lambda}{NA^2}. \qquad [1.2]$$

Equations 1.1 and 1.2 give the relationships between wavelength and lens numerical aperture (*NA*) in terms of resolution and depth of focus. The *k* value in these equations represent a process dependent constant, typically taken at 0.8 for manufacturing technologies. Current lithographic practice, for 0.35μm production, utilizes i-line and DUV (248nm) light, with NA values of 0.55 to 0.7, producing ±0.6μm to ±0.25μm depth of focus with 0.36μm to 0.28μm resolution. Resists used are still conventional Novolac resin based resists, with diazide as the photo active compounds. Prototyping of phase shifting masks, off axis illumination systems, and various wave-front engineering methods to gain resolution and contrast enhancements from the current exposure processes are already underway[19,20]. Meanwhile, the introduction of alternative light sources, such as excimer lasers, is designed to reduce the wavelength and therefore reduce printable feature sizes. However, development of new resists and processes is becoming essential to meet the demands of these advanced photo-lithographic systems.

Alternatively, the fields of X-ray, electron and ion beam lithography and also scanning probe lithography may yet provide solutions to these processing needs. However these are currently slow and largely experimental processes which do not fit within high volume semiconductor production environments. X-ray lithography has demonstrated feature dimensions of less than 20nm but mask fragility and manufacture is a serious problem for this technology[21]. With ion beam lithography, ultimate resolution is poor,

and may not reliably supply feature sizes below the 0.1μm range. More encouragingly, electron beam lithography is capable of producing feature sizes into the sub 10nm regime but it is hampered by a slow, sequential, exposure process and slow resists, particularly at these small structure sizes. Finally, experimental scanning probe lithographic techniques are too slow to be considered as viable, production worthy technologies in their current state.

1.1.2.6 TRENDS IN WAFER SIZE

One further feature of the scaling process is the effect on wafer size used within a process. As control of process parameters, such as film thicknesses, device dimensions, oxide thicknesses etc. becomes more and more critical, uniformity across a wafer becomes more important. Figure 1.5 depicts the trend towards increasing wafer size within the industry. This trend arises from two main sources; reduction of process edge effects and compensation for increasing die size. Most processes employed in device fabrication exhibit variations across the wafer diameter. These variations are generally greatest at the wafer extremities. High process variability at the wafer edge results in low device yield within these areas. In order to decrease the area affected by this type of variation, larger wafer diameters can be employed to provide a larger central wafer area, and comparatively smaller edge area. Thus increasing the circuit yield.

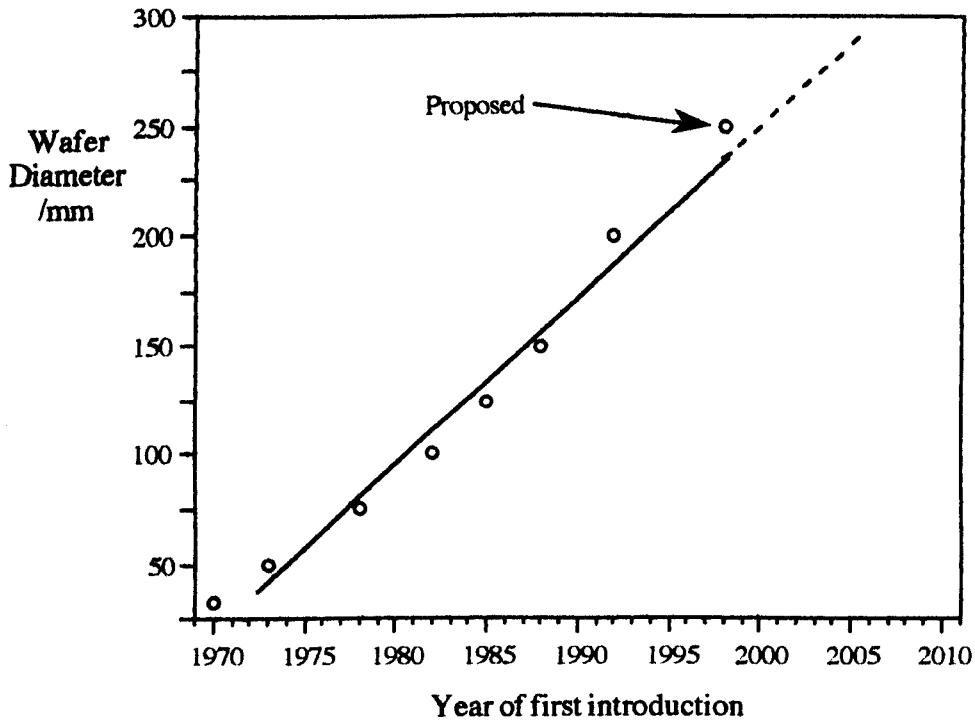


Figure 1.5. Wafer diameter increase with time.

In terms of the size increases, larger wafers are required to maintain profitability through the working die count at the end of the process. In terms of yield, there exists an exponential relationship between maximum possible yield and the die area[22]. Increasing the die area reduces the probable yield and therefore requires more silicon area to produce the same number of devices. The introduction of larger wafer diameters is mainly driven by the increases in die size. For example prototype 1GBit DRAMs have been fabricated with die sizes of 18 x 25mm^2, and current microprocessor sizes are typically of the order of 15 x 15mm^2 or greater. The Digital Equipment Corporation Alpha processor, for example, is around 22 x 22mm^2. As these are using minimum geometries of 0.5μm to 0.35μm, the devices are sensitive to particle diameters down to at least half the minimum geometry[23], that is 0.25μm to 0.12μm diameter. These device dimensions demand at least 200mm, if not 250mm, diameter wafers to maintain profitable working die counts per wafer.

1.1.3 LIMITATIONS TO DEVICE SCALING

The trends discussed in the previous section have been driven predominantly by the technologies behind the silicon MOSFET. However, questions arise related to the fundamental limits of device scaling from two main areas. The first is simply in relation to the deep sub-micron dimensions being employed, while, secondly, we are concerned about the increasing electric field strengths which exist in smaller devices.

Reduced device dimensions require increased control of both linewidth and pattern registration. This not only places tighter requirements on the lithographic tooling, in terms of resolution, stability and registration accuracy, but also on the resist process, in terms of dimensional stability, contrast, resolution and process latitude. As device dimensions shrink, increased doping levels are required in smaller areas to reduce the size of depletion regions within the device. This leads to a degradation of the carrier mobility, resulting in relatively slower devices. As semiconductor-semiconductor junction dimensions become smaller and shallower, with bipolar technologies in particular, reduced thermal budgets come into force in the fabrication process. Each thermal cycle that a device is exposed to results in diffusion of all dopants in the structure. If careful process design and control are not carried out, this leads to excessive broadening of device regions. Again, as the devices get smaller the contact and spreading resistances within the device can limit the advantages gained by scaling.

Coupled with reduced dimensions is the question of increased field strength. Increasing the field strengths can result in electron velocity saturation, inversion layer broadening, lower voltage breakdown of pn junctions, lower electromigration limits and tunnelling

through the thin gate insulator. These effects result in reduction of the speed increase on scaling, lower operating voltages, lower current carrying and power handling capacity, leakage between gate and channel portions of the device and higher susceptibility to noise.

Other limits to the scaling process arise from considering the energetics of device operation and quantum mechanical effects where small dimensions are involved. In terms of the energy relations, a major limit is imposed by the thermal energy of carriers in a device structure. The energy difference between logic states must be large when compared with this energy. This limit is particularly relevant when we are considering low voltage device operation in order to overcome high field effects. Where we have an energy difference between logic states of ΔE_s, the probability of a thermal fluctuation of equal energy within the circuit or device response time, τ, is given by the Boltzmann relation [24];

$$P = exp - \frac{\Delta E_s}{k_B T}, \qquad [1.3]$$

resulting in a failure rate determined by:

$$\Gamma_f = \frac{1}{\tau} exp - \frac{\Delta E_s}{k_B T}, \qquad [1.4]$$

expressed in failures per unit time per device. In order to ensure a low probability of error due to thermal fluctuations, energies of $\Delta E_s \approx 100 k_B T$ are required. Thus, classical device operation must use state voltage differences of at least 2.6V in order to operate with low error rates at room temperature. Similarly, when examining new device phenomena, the energy scale of the process involved must have this magnitude of energy associated with it for useful operation at the temperature in question.

Similarly, quantum limitations on device operation are energy scale dependent. As with thermal limitations, we require the characteristic energy of the quantum process to be greater than other energy scales in the system before these effects become significant. However, it is possible for classical MOSFET structures to fail through tunnelling. In particular, if the gate oxide becomes thin enough tunnelling between the gate and channel can easily provide current of the same order of magnitude as the device channel current. In this instance severe modification to device operation results. It is important to note that on 0.35μm MOS processes the gate oxide thickness is some 4nm, see figure 1.4. This dimension is within reach of the limit at which the gate oxide becomes thin enough to allow significant gate to channel leakage by quantum mechanical tunnelling.

1.2 QUANTUM EFFECT DEVICES

Developments in the fields of molecular beam epitaxy (MBE) and metal-organic chemical vapour deposition (MOCVD) have resulted in a great deal of investigation into the properties of ultra thin semiconductor films. The use of advanced lithographic techniques to laterally pattern these structures and other materials to nanometre dimensions has resulted in a vast amount of work into nanometre scale devices and structures. This work has looked at a number of subjects; firstly the characteristics of conduction in specially fabricated, pure, ultra-thin materials and layers, secondly the basic fundamentals of the conduction processes in low dimensional structures or high purity materials and, thirdly, the possibility of utilizing these materials or conduction processes for new electronic devices.

MBE and MOCVD techniques allow interfaces between different materials to be defined on an atomic scale, producing well controlled, abrupt junctions between differing material types. Such control of the purity and composition of the material exerts control over the electrical and optical properties of the fabricated structures. In this way, techniques such as band-gap engineering[25] can be employed to determine the fundamental physical characteristics of a semiconducting material.

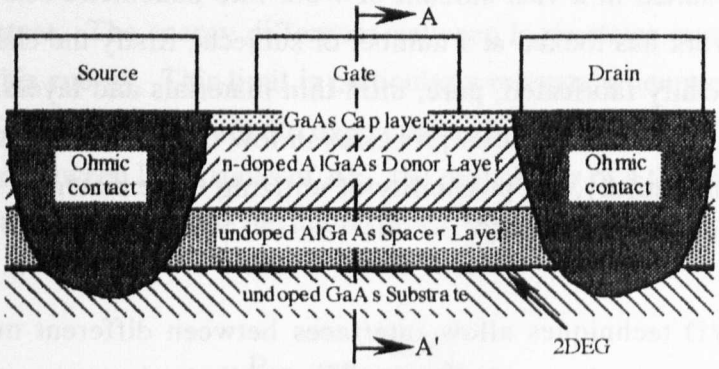
These techniques are in direct contrast with more traditional techniques for manipulating semiconductor-semiconductor interfaces, such as diffusion, ion-implantation, etching (both wet and dry), chemical and physical vapour deposition and sputtering. Interfaces produced by these techniques are neither atomically defined nor abrupt. Such interfaces tend to be diffuse, even where abrupt junctions are the intention, with the junction region extending over a number of inter-atomic spacings, typically tens of Angstroms at best. Similarly, metal-oxide, metal-semiconductor and oxide-semiconductor interfaces are neither uniform, well controlled, abrupt nor pure. These deficiencies result in a reduction in carrier mobility through crystal imperfections, surface roughness and strain. The amorphous nature of some of the deposited films used in semiconductor technologies results in further disorder within the device system and further degradation of the charge transport properties.

1.2.1 REDUCED DIMENSIONALITY

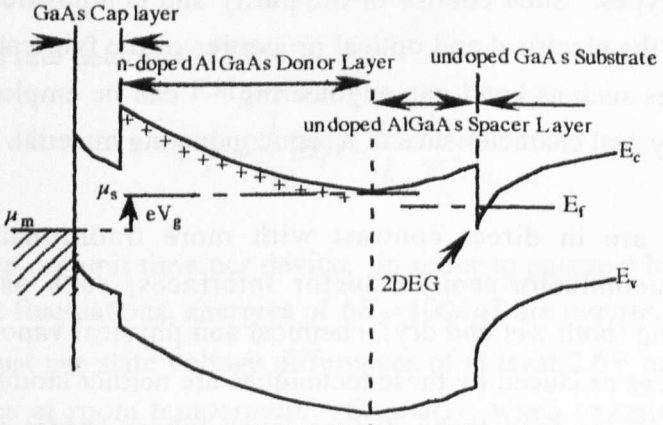
Reducing the dimensions of an electronic structure leads not only to changes in the density of states, but also to the splitting of energy levels in the band. This becomes an important effect when the structure size approaches or becomes less than the electron

wavelength in that material. At this point the energy levels in the band structure separate, and the dimension which has been reduced is effectively removed from the device. In this way, a three dimensional material can be turned into a two dimensional electronic system, i.e. an electron plane - or gas. Further dimensional constraints reduce this to either a one-dimensional systems, a wire, or a zero-dimensional system, a dot.

1.2.1.1 2 DIMENSIONAL SYSTEMS



Physical device structure



Energy band diagram sketch, through section AA'

Figure 1.6. Two dimensional electron gas formation, and the MODFET

Two dimensional electron systems are, perhaps, the most common of reduced dimensionality systems. Removal of one dimension, typically that perpendicular to the surface, leaves what is termed a two-dimensional electron gas, or 2DEG, in the plane parallel to the surface. These structures are commonly used as a basis for studies in high purity, long electron mean free path materials and low dimensional device physics. Other uses for this type of material system have been in the development of high electron mobility transistors (HEMTs) for high speed electronic systems. Figure 1.3 displays cut off frequencies for four types of 2DEG transistor technologies, two formed as modulation doped FETs (MODFETs) and two as HEMT structures.

The electronic systems which form the basis of studies into quantum effects in semiconductors are generally derived from MBE grown heterostructure layers. Within these layers, two dimensional electron gas formation is achieved through arrangement of the layer structure. The MBE layered approach can also be used to form quantum well structures as the layers are grown through the use of large and small energy gap materials. It is possible to utilize these layers through both perpendicular and parallel conduction. Where parallel conduction is employed, the device types are generally conventional operation transistors such as HEMTs, MODFETs, or laterally patterned low dimensional structures.

Where the charge transport is arranged to be perpendicular through the layer structure, a large number of transistor actions can be developed. Heterojunction bipolar transistors and hot electron transistors are commonly formed on these layers, as are resonant tunnelling diodes and transistors. With the resonant tunnelling devices, the ability to controllably form quantum wells within the material layers defines the resonant states within the device. The use of resonant tunnelling devices provides the possibility of device operation into the terahertz regime due to the strong negative differential resistance regions created in the device current-voltage curves through the resonant regions[26].

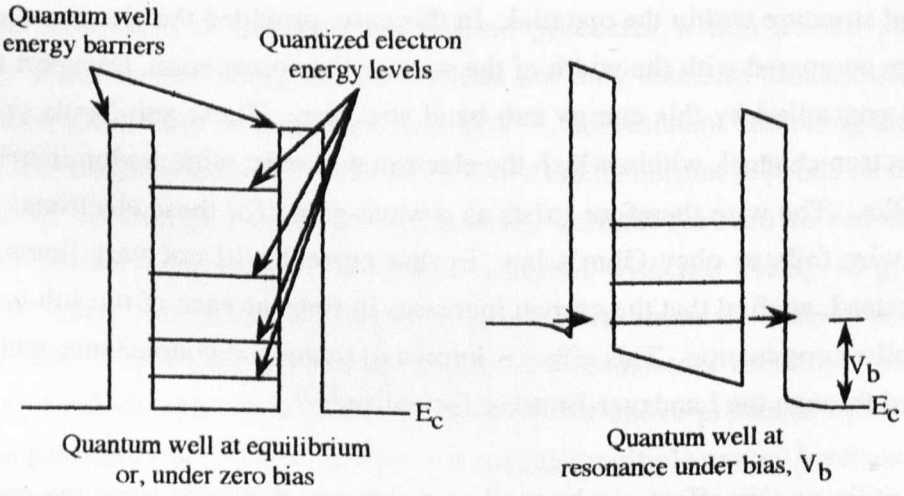


Figure 1.7. Quantum well layer formation and band diagram.

1.2.1.2 1 DIMENSIONAL AND ZERO DIMENSIONAL SYSTEMS

Reduction of either the x or y dimension results in sub-band creation in the remaining dimension and the creation of a one dimensional conduction channel. These systems are exemplified by quantum point contacts (QPCs) and 1D wires. QPCs like 1D wires exhibit resistance quantization and can be used as resistance standards or alternatively

to define zero dimensional and other device types. Quantum wires have been used in studies on ballistic electron conduction[27] and localization[28]. However, particularly with QPCs, fluctuation effects which are manifestations of sample dependent universal conductance fluctuations[29] can be exhibited due to temperature effects, impurities and device non-idealities. These effects tend to limit the fabrication repeatability and therefore usefulness of such devices.

Zero dimensional systems provide resonant tunnelling and single electron tunnelling effects. In particular with zero dimensional systems coupling of energy into the structure is achievable through the energy levels which exist within the quantum wells created.

1.2.2 QUANTUM DEVICES

Quantum effect devices are those whose operation and characteristics depend in some manner on quantum mechanical effects. These effects can arise from reduction of device dimensions or from quantum mechanical processes occurring with sufficient energy to become the dominant process in the device operation.

As previously mentioned, reduction in device geometries leads to a splitting of the energy band structure within the material. In this case, provided the electron mean free path is large compared with the width of the wire or the constriction, transport through the wire is controlled by this energy sub-band structure. These sub-bands provide a ballistic electron channel, within which the electron will exist with predominately wave characteristics. The wire therefore exists as a wave-guide for these electrons. At this point, the wire fails to obey Ohm's law, in that current will not vary linearly with voltage. Instead, we find that the current increases in steps as each of the sub-bands are opened to allow conduction. This effect is known as quantized conductance which may be described through the Landauer-Büttiker formalism[30].

Structures utilizing this effect can be used as resistance standards since the resistance value depends on physical constants i.e. $R_q=h/e^2$ per channel. Here we see a single conduction channel has a resistance of 25.8kΩ, or 12.9kΩ where there is no electron spin selection within the channel. This enforces a current limitation of 77μA/V on any conduction channel defined in such devices, in the absence of magnetic effects. Another feature of this sub-band structure is the ability to use the constriction as an energy filter. The conduction channels that exist through the structure have given energies. In order for an electron to traverse the constriction or wire it must have the same energy as the channel. Therefore, transmitted electrons will exhibit a tightly

defined energy spectrum. In a similar vein, it is possible to use these types of structure to directly couple light of given wavelengths into an electrical circuit through the use of detectors formed from these structures.

The behaviour of structures in this quantum ballistic regime is termed non-local, in that the presence of a structure which causes an interference effect will affect the conductance behaviour in other parts of the structure[31]. This is exemplified by the Ahronov-Böhm[32] effect. In this effect, the interference of electron wave-functions which have traversed different but equal length paths (around a ring in most fabricated devices) modifies the transmission coefficients of the structure. Where the electron energy is changed, by application of a magnetic field in the Ahronov-Böhm effect, oscillations in this transmission coefficient can be observed. In the Ahronov-Böhm effect, these modifications are periodic in the magnetic flux threading the loop. This is a manifestation of a change in the interference behaviour of the electron paths. Other types of interference devices are proposed and under investigation, the most successful of which are superconducting quantum interference devices which provide very sensitive measurement of magnetic fields. In order to simulate optical elements and interferometers, work has also been carried out into electrical analogues of lenses and interferometers[33]in ballistic electron conduction materials.

One of the most utilized quantum mechanical processes within device physics is tunnelling. Electron tunnelling devices include not only standard tunnelling devices, such as tunnel diodes and tunnel base transistors[34], but resonant tunnelling diodes and transistors and single electron effect devices. As a tunnel current depends on the width and height of an energy barrier, it is relatively straight-forward to envisage and understand device applications. It is this, coupled with the ability to grow thin controllable layers into material structures and pattern on small dimensions, which makes the use of resonant tunnelling diodes and transistors attractive for use in electronics. Such devices are capable of operation far above present frequencies and provide the possibility of multi-state logic, through the multiple energy levels within the device[35]. Single electron devices, which are dealt with in Chapter 2, also are promising in the possibility of large bit storage capacities and the ability to operate at low current and power levels. Another feature of these devices is the ability to operate through the use of large confinement potentials defined in small structures. In such conditions, higher energies can be utilized in the device operation, allowing high temperature operation, whereas the temperature of operation for semiconductor based quantum ballistic electron devices may well be limited to below 77K

An important aspect of quantum devices is their ability to be used in metrology

applications. As the device operation depends in some form or another on fundamental constants, it is possible to utilize these devices to great effect and accuracy in measurements of voltages, currents, resistances, magnetic fields etc. Therefore, the usefulness should not be ruled out on the basis of the low temperatures required for their operation.

1.3 MOTIVATION AND AIMS OF THIS WORK

An enormous amount of research and development work is being carried out around the world into reduction of current device dimensions and development of new device types and modes of operation. All of these projects have similar goals in that they aim to better understand the conduction processes, the operation of current and future device types and to provide technologies, processes, and devices for the future electronics industry. Significant progress has been made in understanding conduction processes and the implications of quantum effects within low dimensional structures. Other areas where significant progress has been made are in the fabrication and operation of devices utilizing these principles. However, a large number of new devices and operation modes have limited applicability through two main constraints; low current operation and low temperature of operation. The low operating temperature is by far the most limiting factor since buffering may be employed to overcome low current operation, although this is not an ideal solution.

The device dimensions currently being employed in semiconductor systems tend to result in quantum mechanically derived effects being observable only where noise, lattice vibrations and subsequent scattering are eliminated through reduction of their relevant energy scales below that of the effect under observation. Within semiconductor systems, significant difficulties arise when attempting to push device dimensions into the sub-40nm regime. At this point, the rate of change of electric field within these systems can become too low to provide adequate confinement potentials. Where the confinement potential is defined by grown structures (MBE or MOCVD), the problem is less acute as the material builds in these potentials. Where patterning is employed to make a functional device from this material, the confinement potentials become blurred and structures sizes tend to be larger than expected. For example, reactive ion etching of structures introduces damage resulting in non-abrupt surface potentials. These potentials can reach some way into the device and damage the electrical structure being defined[36]. Similarly, where electric fields are being used to define regions within a device, the dielectric constants result in active area dimensions too large to achieve adequate confinement potentials at high temperature.

The work presented in this thesis examines a system which may be defined on a small enough scale to allow high temperature operation of single electron devices. These devices operate, as their name suggests, on the effect of only one electron. The potentials set up within the device can result in state switching controlled by the presence or absence of a single electronic charge. The benefits of such a device technology are those of ultra low power operation, multi-state logic, hyper sensitive current measurement and, potentially, enormous bit storage densities due to the very small device sizes required for this technology. Currently, around 10^4 to 10^5 electrons are used to store 1 bit of information within a DRAM memory element. In the case of single electron logic, the possibility of using one electron to store that bit, and the use of multi-state logic opens up tremendous logic and information handling possibilities for the future.

Although single electron effects have been observed with characteristic temperatures well above room temperature, these systems have generally been uncontrollable and difficult to make use of within an engineered system. Where devices have been specifically engineered, useful device applications are prohibited by low operation temperatures. The aims of this work are therefore to investigate the fabrication techniques for metal-insulator-metal tunnel junction based single electron systems with a view to decreasing the structure size and thereby increasing the temperature of operation. The goal for these aims is to achieve a fabricated single electron transistor, or electrometer, operating at 4.2K or above. The goal of 4.2K is to be seen as a first step in raising the operation temperature above room temperature.

1.4 ORGANIZATION OF THIS THESIS

The work presented in this thesis divides into three main parts. Firstly, a review of the theory and development of single electron devices is given in chapter two. Secondly, fabrication techniques are presented in chapters three and four. Thirdly, chapters five and six present the work carried out on electrical measurements of the materials used and single electron device structures fabricated during this work. To conclude this thesis, chapter seven reviews and discusses the work in terms of its aims and results and presents areas for further work and investigation.

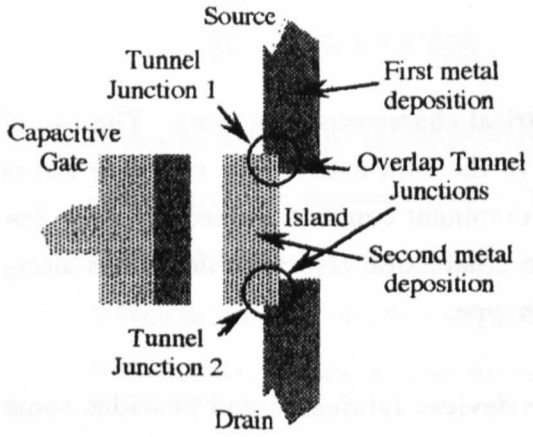
Discussion of the theory of single electron effects, given in chapter two, divides into three main parts. Firstly, an introduction to the theory of tunnelling is given, with a discussion of metal based devices. The theory of single electron tunnelling devices is reviewed, with the emphasis being on normal metal, as opposed to superconducting, single electron devices. Finally, a review of the work carried out in the development of

usable single electron structures is given.

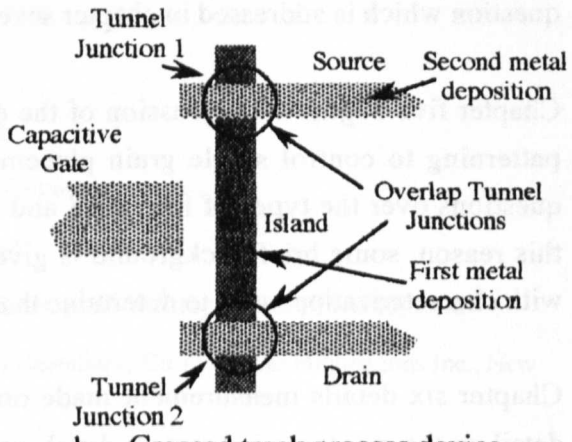
The experimental work documented in this thesis is presented in chapter three to six, and the order of the work requires some description and background at this point. Due to the small scale of the structures being fabricated in this work, it is necessary to fabricate lead frame structures to electrically connect to the devices. While not directly relevant to the devices of interest this is an enabling technology without which the devices could not be measured. Chapter three, therefore, discusses the general fabrication techniques used in this work with specific reference to the interconnect lead frame as an example. This allows presentation of the general fabrication details without confusion with the techniques involved in ultrasmall device fabrication. Chapter three concludes with a description of the process used to fabricate suspended masks and a detailed description of all process flows used in this work.

Chapter four deals with the detailed fabrication and characterization work for the three device technologies under investigation. The first technology creates overlay tunnel junctions using shadow mask techniques and is described in detail in section 4.2. The shadow mask can be used with multiple angled evaporations to selectively deposit areas of the pattern and produce an overlap between two layers of metal. Provided the underlying layer is oxidized to form a tunnel barrier, a tunnel junction is formed in the overlap. Figure 1.8a depicts the device structure formed by this technique where a single electron transistor (gated island, double tunnel junction) is to be formed. The circuit diagram for a single electron transistor is given in figure 1.8d for reference.

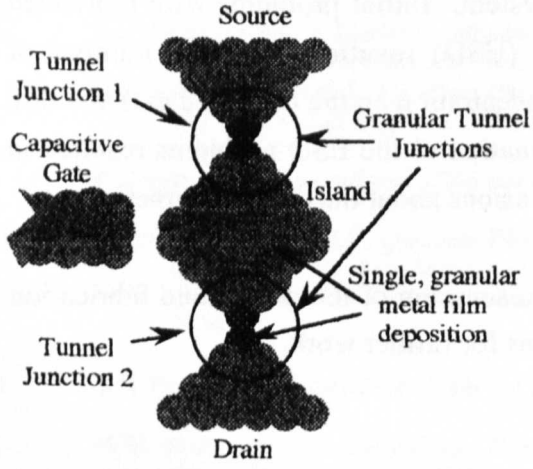
The suspended mask technique is a standard technique and was adopted in this work to provide a known working process for development of overlap tunnel junction. The process does, however, limit the flexibility of the pattern and produces unwanted areas of film deposition. To attempt to remove the difficulties with the suspended mask process, the crossed track process was investigated. In the crossed track process, multiple angled evaporations are again used, but with resist shadowing is employed to produce overlap tunnel junctions where the resist tracks cross. The structures to be formed with this process are depicted in figure 1.8b. This process provides a planar structure with the only pattern constraint being that tunnel junctions are formed by the overlap, or crossing, of perpendicular metal tracks. Again oxidation between the deposition of the two metal layers results in the formation of a tunnel barrier. This process exhibited a number of significant problems and was never used to fabricate devices for electrical measurements. However, the problems manifested by the process partly lead to the development of the third process investigated. The process, investigation work and difficulties are discussed in detail in section 4.3.



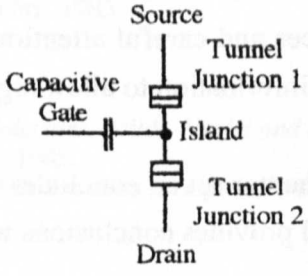
a. Suspended mask process device structure.



b. Crossed track process device structure.



c. Patterned granular thin film process device structure.



d. Circuit diagram for all three process technologies.

Figure 1.8 Schematics of device structures under investigation.

The third process under investigation used patterned granular thin films. This process attempts to form tunnel junctions in the connection between grains in the aluminium film. Figure 1.8c depicts the structure of devices formed with this process. The aim with this process was to produce the smallest possible tunnel junctions using aluminium films, thereby reducing the junction capacitance to raise the transition temperature for observing single electron effects. Both the previous processes exhibited unreliable processing when the tunnel junction size was significantly reduced therefore, the patterned granular thin film process was seen as the best possibility for high temperature single electron effects. As this technique requires the patterning process to define the grain positions and to some extent size, a large amount of film and pattern characterization work was required. This work is documented in section 4.4, along with the technique itself. The tunnel junction in this process is formed by creating an oxide film on the grain surface. This should result in an oxide barrier between the grains

forming the junction. Whether this barrier is formed during or after the evaporation is a question which is addressed in chapter seven.

Chapter five begins the discussion of the electrical characterization work. The use of patterning to control single grain placement in the film and device structure raises questions over the types of film used and the dominant conduction mechanisms. For this reason, some brief background is given to conduction effects in thin films along with characterization work to determine the film type.

Chapter six details measurement made on the devices fabricated and provides some detail on the measurement system developed and used in this work. The need for a low noise fempto-amp range current measurement system necessitated the design and construction of this electronic measurement system. Initial problems with fabricated devices and electrostatic discharge damage (ESD) resulted in a low number of successful device measurements. However, concentration on the patterned granular thin film devices and careful attention to and elimination of the ESD problems resulted in sufficient information to allow significant conclusions about the device character.

Finally, chapter seven concludes the thesis by discussion of the device and fabrication results and provides conclusions with suggestions for further work.

1.5 REFERENCES

1. R. Amatruda, *Semicond. Int.* **19**(10),63 (1996).

2. J. Bardeen & W.H. Brattain, *Phys. Rev.* **74**,230 (1948).

3. W. Shockley, *Bell. Syst. Tech. J.* **28**,435 (1949).

4. D.Halliday & R.Resnick, *Physics*, 3rd Ed., Ch. 26, p565, J Wiley & Sons, Toronto, 1978.

5. H.M. Leicester, *The Historical Background to Chemistry*, Ch.12. Dover Publications Inc., New York, 1971.

6. B. Bowers, *IEE Review*. **37**(9),291 (1991).

7. J.C. Maxwell, *Proc. Royal Soc. London* **13**,513 (1864).

8. B.H. Bransden & C.J. Joachain, *Introduction to Quantum Mechanics*, Ch.s. 1 & 2., Longman Scientific and Technical, Essex, 1989.
and R. Eisberg & R. Resnick, *Quantum Physics of Atoms, Molecules, Solids, Nuclei and Particles*, Chs. 1 to 5., 2nd Ed., John Wiley and Sons, New York, 1985.

9. P. Drude,. *Annalen der Physik* **1**,566 and **3**,369 (1900).

10. R. Eisberg & R. Resnick, *Quantum Physics of Atoms, Molecules, Solids, Nuclei and Particles*, Ch. 13., 2nd Ed., John Wiley and Sons, New York, 1985.
and N.W. Ashcroft and N.D. Mermin, *Solid State Physics*, W.B. Saunders Company, Philadelphia, 1976.

11. H.S. Black, *Bell System Tech. J.* **14**(1),1 (1934).

12. H.W. Bode, *Network Analysis and Feedback Amplifier Design*, D. Van Nostrand Company, Princeton, New Jersey, 1945.

13. See G.W. Drummer, *Electronic Inventions and Discoveries*, 2nd Ed., Pergammon Press, 1978.

14. J. Mavor, M.A. Jack & P.B. Denyer, *Introduction to MOS LSI*, Ch. 3., Addison-Wesley Publishers Ltd, 1983.

15. S.M. Sze, *High Speed Semiconductor Devices*, Ch. 1, John Wiley & Sons Inc., New York, 1990.

16. N. Cameron, H. McLelland, M.C. Holland, M.R.S. Taylor, S.P. Beaumont., *IEE Colloquium on "Modelling, design and applications of MMICs*, IEE Digest No: 1994/092, Professional group E12, (1994).

17. C.M. Snowden, *IEE Electronics & Communication Engineering Journal*. **8**(2),51 (1996).

18. R.M. Burger, J.A. Glaze, T. Seidel & O. Williams, *Solid State Technol.* **38**(1),38 (1995).

19. M.D. Levenson, *Solid State Technol*. **38**(2),57 (1995).

20. M.D. Levenson, *Microlithography World*. **1**(1),6 (1992).
and G. Cognié. *Microlithography World*. **1**(1),13 (1992).

21. H.G. Craighead, Proc., *38th Scottish Universities Summer School in Physics*, eds., J.H. Davies & A.R. Long, IOP Publishing, Bristol, England, 1992.

22. S.M. Sze, *Semiconductor Devices - Physics and Technology*, p435, John Wiley & Sons, New York, 1985.

23. J.B. Khare, W. Maly, S. Greip & D. Schmitt-Landsiedel, *IEEE. Trans. Semicond. Manufacturing* **8**(2),195 (1995).

24. C. Mead & L. Conway., *Introduction to VLSI systems*, Ch. 9., Addison-Wesley Publising Co., Reading, Massachusetts, 1980.

25. E.H. Li & B.L. Weiss, *IEE Electronics and Communication Eng. J.* **3**(2),63 (1991).

26. M.A. Reed, W.R. Frensley, R.J. Matyi, J.N. Randall & A.C. Seabaugh, *Appl. Phys. Lett.* **54**,1034 (1989).

27. D.A. Wharam, T.J. Thornton, R. Newbury, M. Pepper, H. Ahamed, J.E.F. Frost, D.G. Hasko, D.C. Peacock, D.A. Ritchie & G.A.C. Jones, *J. Phys. C.* **21**,L209 (1988).

28. G. Bergmann, *Physics Reports.* **107**(1),1 (1984).

29. H. Haucke, S. Washburn, A.D. Benoit, C.P. Umbach & R.A. Webb, *Phys. Rev. B.* **41**(18),12454 (1990).

30. M. Büttiker, *Phys. Rev. Lett.* **57**,1761 (1986).

31. C.P. Umbach, P. Santhanam, C. van Haesendonck & R.A. Webb, *Appl. Phys. Lett.* **50**,1289 (1987).

32. C.J.B. Ford, T.J. Thornton, R. Newbury, M. Pepper, H. Ahmed, C.T. Foxon, J.J. Harris & C. Roberts, *J. Phys. C (Solid State Physics).* **21**,L325 (1989).

33. C. Weisbuch & B. Vinter, *Quantum Semiconductor Structures*, Ch. 6., Academic Press Ltd. London, 1991.

34. S.M. Sze, *Physics of Semiconductor Devices*, Ch. 9., John Wiley & Sons. New York, 1981.

35. F. Capasso & S. Datta, *Physics Today.* **2**,74 (1990).

36. M. Rahman, *Electron Transport in Dry-Etched n^+-GaAs Nanostructures*, Chs. 3 & 4., Ph.D. Thesis, University of Glasgow, 1992.

CHAPTER TWO

Tunnelling and Single Electron Devices

2.1 INTRODUCTION

Devices which operate on the basis of single electron transport are possible through the process of quantum mechanical tunnelling and its impact on the charging energy of a tunnel barrier. In this chapter the tunnelling process is introduced in terms of the wave mechanics formulation of quantum theory. After discussing the tunnelling effect and results relevant to aluminium tunnel junctions, the subject of single electron tunnelling is developed. Most of the references cited develop the theory in a more general sense, including single electron effects observable in superconducting systems. Studies of single electron effects in semiconducting systems also exist but these are not discussed in detail within this work. Reviews by Beenakker[1], van Houten[2] and Kastner[3] with papers by Scott-Thomas[4], Meirav[5] and Kouwenhoven[6] provide a good introduction to the semiconductor field. Effects observed through scanning tunnelling microscope probing of granular films are mentioned, where relevant, as these fit easily into the theory for normal metal systems.

The theory discussed in this chapter follows the development by Averin and Likharev[7] in 1985/86, which has become termed the orthodox theory of single electron tunnelling. Various modifications have been added to the theory to account for other effects, such as noise (both inherent and coupled), leakage, environmental effects, effects in more

complex junction systems and effects in 2-3 dimensional systems. As such, the field has become subject to an increasingly large number of publications over the past ten years. Discussion of these additional effects is limited to their main points and conclusions.

The limitation of the discussion to normal metal systems is due to the aims of this thesis, as set out in section 1.3. Use of the available lithographic capabilities to reduce device dimensions to the extent that higher temperature operation of these devices is possible precludes the use of superconducting systems in this work. This work solely employs aluminium based devices having a superconducting transition temperature of $T_C \approx 1.178$K for bulk Al, with a T_C as high as 3.5K reported in granular Al films[8]. As we are aiming for greater than 4.2K operation (the temperature of liquid ^{4}He at 1atm. pressure), superconductive effects are not relevant in these devices.

This chapter divides into two sections and serves as a fuller introduction to the background and work of this thesis. Section 2.2, covers the theory of tunnelling and some relevant tunnelling experiments. Section 2.3 discusses single electron tunnelling by a brief introduction to the theory and discussion of the effects observed. Where relevant the discussion points out the potential device applications and circuit structures with a description of their operation. At the end of section 2.3, some discussion of the applications of single electron devices and technological requirements for useful single electron devices is given.

2.2 TUNNELLING

Development of the theory of quantum mechanics led very quickly to an explanation for the phenomenon of tunnelling. Application of quantum mechanical ideas to the calculation of particle behaviour in a region with a potential barrier provided contradictory results to that of the classical mechanics description. The theory of tunnelling demonstrates the ability to penetrate barriers which are, classically, impenetrable. In the classical description, a particle with insufficient energy to overcome a barrier will be reflected. Also, particles with sufficient energy are able to freely cross the barrier with no impact on their state. In this instance, the particle traverses the barrier with unity probability. Conversely, in a quantum mechanical description a particle with insufficient energy retains a finite probability of traversing the barrier. Where the particle has sufficient energy to cross the barrier, quantum mechanics does not guarantee unity probability of transmission. Quantum mechanically, the probability of transmission depends upon exponential functions of the barrier height and width. Barriers which are wide or high may yield a vanishingly

small, but non-zero, probability of transmission.

During the late 1920s and early 1930s, quantum theory was applied to a number of physical problems. In 1928, field ionization of atomic hydrogen was described by Oppenheimer. Again in 1928, Fowler and Nordheim described field emission from a free electron metal and Gamow, Condon and Gurney independently described alpha particle decay using the theory of tunnelling. Of more relevance to this work, the 1930's saw Frenkel[9] describe electron transport between disconnected metals through metal-vacuum-metal tunnelling and subsequently, in 1933, Sommerfeld and Bethe described metal-insulator-metal (MIM) tunnelling.

In 1958, Esaki[10] discovered the tunnel diode and the theory of tunnelling found its way into the realms of electronic devices. A great deal of work followed this, both in technological and theoretical[11] improvements. These resulted in the description of ohmic contacts with high doping levels in semiconductors and a number of tunnel devices suitable for microwave applications[12]. Of particular interest to this work, theoretical studies of MIM junctions were carried out by Simmons[13,14] around 1963, with corresponding experimental verification by Pollack and Morris[15] around 1965. Further suggestions relating to tunnel effect transistors[16] have been made but as yet remain unexploited. Towards the late 1970's, investigations into light emission from tunnel junctions became an active area of research. Emission from metal-oxide-metal[17,18] junctions was found to exhibit bias-voltage dependent broadband light emission at frequencies up to $h\upsilon_0 = e|V_b|$. This effect was explained by inelastic tunnelling causing the excitation of plasma modes at the surface of the metal electrodes, leading to radiative emission of energy.

Returning to the 1960s, tunnelling was used to sense the electronic energy structure of superconductors. The energy dependence of the density of states was measured by Giaever[19,20] and Nicol et al.[21] in normal metal-insulator-superconductor and superconductor-insulator-superconductor junctions. Later, Josephson predicted a number of effects leading to a rapid development of the field of superconductive tunnelling. The technological development of superconducting quantum interference devices (SQUIDs) for accurate measurement of voltages and magnetic fields soon followed these theoretical discoveries. Around this time, tunnelling was also used to describe the behaviour of thin granular films, especially at low temperatures where a high resistance phase can exist. The description followed that of an activated tunnelling process, dependent to some degree on the charging energy of individual grains in the film[22].

More recently, tunnelling has been an active area of interest in the development of new electronic devices, especially those operating on the principles of resonant tunnelling[23]. In this case, a double tunnel barrier either side of a quantum well leads to resonances in the particle transmission through this structure. When the energies of the incoming electrons match one of the discrete energy levels of the quantum well, a low resistance resonant state exists in the device. Increasing the electron energy (by increasing the voltage bias across the device) reduces the resonant effects and increases the device resistance. This effect provides a negative differential resistance region in the device I(V) characteristic. There are existing applications for Esaki's tunnel diode and for resonant tunnelling devices in oscillator and mixing systems, particularly at microwave frequencies.

The subject of single electron tunnelling results from the interaction of a junction's charging energy with the tunnelling process. This subject is relatively new and the main theories appeared only in 1985[7] although knowledge of charging effects existed long before[22,24] through studies of discontinuous granular films. This subject is returned to in Section 2.3.

2.2.1 TUNNELLING PROBABILITY

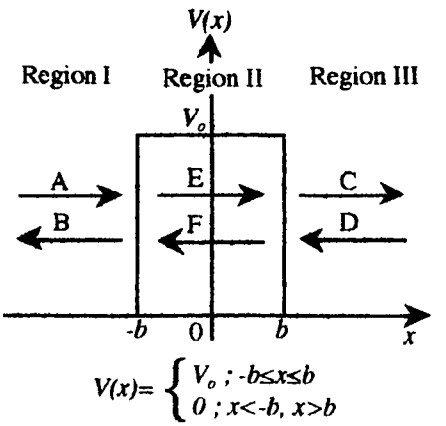


Figure 2.1. Potential landscape for barrier problem.

The tunnelling problem can be analysed through a one dimensional model, as shown in figure 2.1. Here we see a potential barrier of width $2b$ centred on $x=0$ and a problem which splits into three regions. A solution to this problem can be found using the one-dimensional, time independent Schrödinger equation as given in equation 2.1. In this equation, we have $\psi(x)$ as the particle wave-function with m as the particle mass and $\hbar$ being Plank's constant ($\hbar=h/2\pi$). $V(x)$ describes the potential barrier structure and E is an Eigenvalue solution to equation 2.1.

$$-\frac{\hbar^2}{2m}\frac{d^2\psi(x)}{dx^2} + V(x)\psi(x) = E\psi(x) \qquad [2.1]$$

In solving this problem, we are interested in calculating a measure of the probability of particle transmission through the barrier. Each of the regions contain wave-function solutions to the Schrödinger equation, having probability amplitudes A to F. However, we are interested in the transmitted flux through the barrier and therefore a solution to the problem can be found in terms of the ratio of transmitted, C, to incident, A, probability amplitudes. From this ratio a transmission coefficient, the ratio of transmitted to incident probability current densities, can be defined. The probability current density of a plane wave, $\psi(x) = A\,exp(ikx)$, is given by[25];

$$\begin{aligned} j &= \frac{\hbar}{2im}\left(\psi^*\frac{d\psi}{dx} - \frac{d\psi^*}{dx}\psi\right) \\ &= \frac{\hbar}{2im}\left(A^*e^{-ikx}A\,ik\,e^{ikx} - Ae^{ikx}A^*(-ik)e^{-ikx}\right) \\ &= \frac{\hbar k}{m}|A|^2 = \frac{p}{m}|A|^2 = v|A|^2, \end{aligned}$$

If the probability current densities for regions I and III are considered as j_1 and j_3 respectively, then the transmission coefficient is defined as:

$$T = \frac{j_3}{j_1} = \frac{v_3|C|^2}{v_1|A|^2} = \frac{|C|^2}{|A|^2}\ , \qquad [2.2]$$

where we have used $k_1=k_3=k$, which requires that $v_3 = v_1$.

In order to find the transmission coefficient, we need to determine the general form of the wave function in each region. When we have these solutions to the Schrödinger equation, for each of the three regions, it is necessary to match the solutions across the discontinuities in $V(x)$ to ensure continuity of $\psi(x)$ and $\partial\psi(x)/\partial x$.

In regions 1 and 3, we find that $V(x)=0$. As the particle energy is greater than the background potential, $E>V(x)$, we find the solutions to the Schrödinger equation take the form of travelling waves. Equations 2.3 and 2.4 describe these regions.

$$\psi_1(x) = Ae^{ikx} + Be^{-ikx} \ ; \ x < -b \qquad [2.3]$$

$$\psi_3(x) = Ce^{ikx} + De^{-ikx} \ ; \ x > b \qquad [2.4]$$

$$\text{where, } k = \sqrt{\frac{2mE}{\hbar^2}} \qquad [2.5]$$

In equation 2.4, D=0 since on the right hand side of the potential there is no mechanism for reflection, therefore no contribution to the solution can be allowed from waves travelling right to left in region III.

In region II, when the particles have lower energy than the barrier height, tunnelling is the only process whereby particles can traverse the barrier. The solution to the Schrödinger equation in this situation, where $E<V(x)=V_o$, takes the form of a decaying wave,

$$\psi_2(x) = Fe^{\kappa x} + Ge^{-\kappa x} \; ; \; -b > x > b. \qquad [2.6]$$

$$\text{where, } \kappa = \sqrt{\frac{2m(V_o - E)}{\hbar^2}}. \qquad [2.7]$$

Applying the boundary conditions $\psi_1(x)=\psi_2(x)$ and $\partial\psi_1(x)/\partial x=\partial\psi_2(x)/\partial x$ at x=-b, and $\psi_2(x)=\psi_3(x)$ and $\partial\psi_2(x)/\partial x=\partial\psi_3(x)/\partial x$ at $x=b$, yields the following set of equations:

$$\psi_1(x) = \psi_2(x) \; ; \; x = -b \Rightarrow \; Ae^{-ikb} + Be^{ikb} = Fe^{-\kappa b} + Ge^{\kappa b} \qquad [2.8]$$

$$\frac{\partial\psi_1(x)}{\partial x} = \frac{\partial\psi_2(x)}{\partial x} \; ; \; x = -b$$
$$\Rightarrow ik\left(Ae^{-ikb} - Be^{ikb}\right) = \kappa\left(Fe^{-\kappa b} - Ge^{\kappa b}\right) \qquad [2.9]$$

$$\psi_2(x) = \psi_3(x) \; ; \; x = b \Rightarrow \; Ce^{ikb} = Fe^{\kappa b} + Ge^{-\kappa b} \qquad [2.10]$$

$$\frac{\partial\psi_2(x)}{\partial x} = \frac{\partial\psi_3(x)}{\partial x} \; ; \; x = b \Rightarrow \; ik\,Ce^{ikb} = \kappa\left(Fe^{\kappa b} - Ge^{-\kappa b}\right) \qquad [2.11]$$

In this set, we have only four equations for five variables. However we are interested in the transmitted flux through the barrier therefore we can find the solution to the problem by using the definition of the transmission coefficient, equation 2.2. In this solution, we need to find an expression for the ratio of C/A to use in equation 2.2. After some algebraic manipulation, we find:

$$\frac{C}{A} = \frac{2ik\kappa\exp(-2i\kappa b)}{(k^2-\kappa^2)^2\mathrm{Sinh}2\kappa b + 2ik\kappa\mathrm{Cosh}2\kappa b}. \qquad [2.12]$$

Using the ratio of C/A in the equation 2.2 yields an expression for the transmission

coefficient in terms of the particle wave vector, k;

$$T(k) = \left[1 + \frac{\left(k^2 + \kappa^2\right)\mathrm{Sinh}^2 2\kappa b}{4k^2\kappa^2}\right]^{-1}, \quad \kappa = \sqrt{\frac{2m(V_0 - E)}{\hbar^2}} \text{ for } E < V_0. \qquad [2.13]$$

Taking $w=2b$ to be the barrier width and substituting for k and κ, $T(k)$ may be expressed, more conveniently, as $T(E)$;

$$T(E) = \left[1 + \frac{V_0 \mathrm{Sinh}^2 \kappa w}{4E(V_0 - E)}\right]^{-1}, \text{ or, } T(E) = \left[1 + \frac{\left(e^{2\kappa w} - e^{-2\kappa w}\right)^2}{16\frac{E}{V_0}\left(1 - \frac{E}{V_0}\right)}\right]^{-1}. \qquad [2.14]$$

This function is graphed in figure 2.2. In equation 2.14, if $\kappa w > 1$ the $e^{2\kappa w}$ term dominates the expression and the transmission coefficient can be approximated by,

$$T(E) = 16\frac{E}{V_0}\left(1 - \frac{E}{V_0}\right)e^{-2\kappa w}. \qquad [2.15]$$

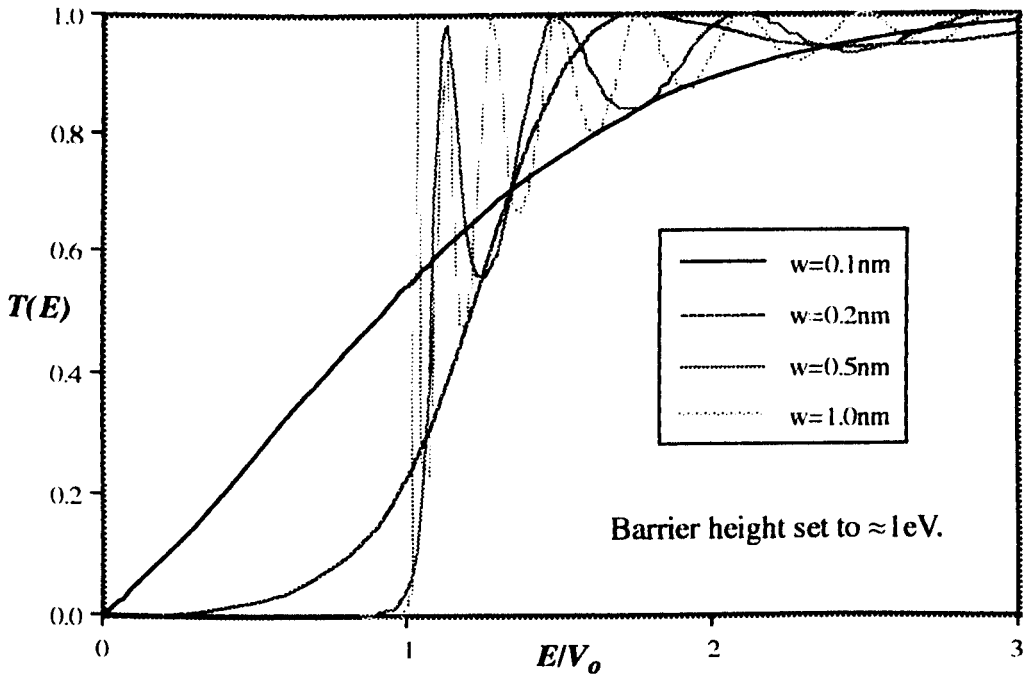


Figure 2.2. Transmission probability variation with energy.

In figure 2.2 the transmission probability in the quantum mechanical case is plotted against the energy ratio E/V_0 for a number of barrier widths. This displays two important properties of the transmission probability which demonstrate the departure from the classical situation. Classically, we have a situation where the particle must have enough energy to surmount the barrier for it to be transmitted. This can be

described by noting that $T(E) = 0$ for $E/V_O \leq 1$ and T(E) = 1 for $E/V_O > 1$. In other words, $T(E)$ is a step function going from 0 to 1 at $E/V_O = 1$ in figure 2.2. Therefore, classically, no transmission takes place where the particle energy is less than the barrier energy. Also, where the particle energy is greater than the barrier energy, transmission takes place with unity probability and the presence of the barrier has no impact on the particle behaviour.

From figure 2.2, we see that the quantum mechanical picture is significantly different from the classical picture. The results given by the quantum mechanical model clearly show that in both regions, $E/V_O \leq 1$ and $E/V_O > 1$, severe modifications to the transmission probability exist. In the case where $E/V_O \leq 1$, we find that the value of $T(E)$ is not zero. $T(E)$ may be vanishingly small, but it is not zero. This means that a particle with energy less than the barrier height can traverse the barrier through the quantum mechanical process of tunnelling. Where $E/V_O > 1$, we see that $T(E)$ is not necessarily unity. In this case there is a probability of reflecting a particle from the barrier, $R(E) = 1 - T(E)$, when it has enough energy to classically traverse it. When $E/V_O > 1$, travelling wave solutions exist in region 2 and the transmission coefficient is described by equation 2.16[26], which results from equation 2.14 where the wave vector, κ, has become complex and describes the travelling waves in region 2. Again, figure 2.2 shows that the behaviour of $T(E)$ in this region has an oscillatory component which easily extends to $E/V_O > 5$, implying that for $T(E)$ to equal unity large particle energies of at least six to ten times the barrier height are required. The oscillations in $T(E)$ are sometimes referred to as "over the barrier resonances".

$$T(E) = \left[1 + \frac{Sin^2 k_2 w}{4\frac{E}{V_o}\left(\frac{E}{V_o} - 1\right)}\right]^{-1}, \quad k_2 = \sqrt{\frac{2m(E - V_o)}{\hbar^2}}, \text{ for, } E > V_o. \qquad [2.16]$$

From figure 2.2, we also see the effect of variation of the barrier width. The strength of a barrier is determined by the product of its height and width, giving an indication of the barrier transparency. In figure 2.2, the barrier height has been kept constant while the width has been varied from 0.1nm to 1nm to increase the barrier strength. The lower barrier widths result in more transparent or weaker barriers which consequently have higher transmission probabilities when $E/V_O \leq 1$. Also, with stronger barriers, such as the 1nm width, the frequency of the over the barrier resonances of $T(E)$ is higher and $T(E)$ approaches unity faster than with the other barriers. It can be seen from figure 2.2 that as the barrier strength increases the tunnelling probability falls off and the barrier resonances are compressed until the model produces the classical situation.

2.2.2 TUNNEL CURRENT

The transmission probability through a barrier was discussed in the previous section. However, we need to calculate the current in a system containing such a barrier to determine the electrical behaviour. Again, a one-dimensional approximation is employed, and figure 2.3 displays the model used for the calculation. Treatment of the three dimensional situation requires the model to be broken down into a one-dimensional (tunnelling) part and a two dimensional solution. The two dimensional solution is, in effect, a calculation of part of the density of states in the electrodes. The electrons in the reservoirs on either side of the barrier occupy free electron states with an occupation probability given by the Fermi function,

$$f(E,\mu) = \left[1 + exp\left(\frac{E-\mu}{kT}\right)\right]^{-1}. \qquad [2.17]$$

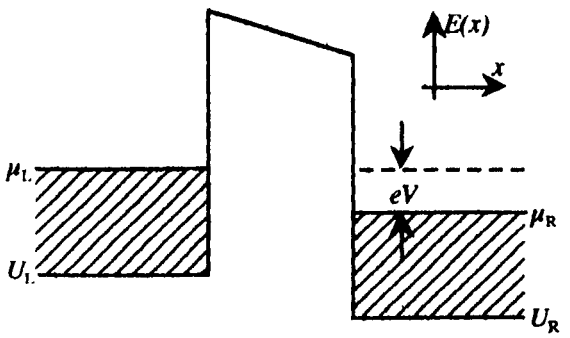


Figure 2.3. Model for calculation of tunnel current.

The current density in a bulk material containing no barriers is defined as,

$$j = 2q\int_{-\infty}^{\infty}\frac{dk}{2\pi}f(E,\mu)v(k), \qquad [2.18]$$

where $v(k)$ is the velocity of an electron state with wave-vector k, and q is the particle charge. We can see from equation 2.18 that we are considering the density of electron states along with their probability of occupation and the velocity of those states which carry current. Firstly, we consider transport from left to right across the barrier, the current density for this direction is,

$$J_{L\to R} = 2q\int_{0}^{\infty}\frac{dk}{2\pi}\left\{f(E_L(k),\mu_L)\left[1 - f(E_R(k),\mu_L)\right]\right\}v(k)\,T_{L\to R}(k). \qquad [2.19]$$

The rules governing tunnelling are reflected in equation 2.19 through the presence of the Fermi functions. They are arranged, in this case, to require an occupied state on the left hand side of the barrier and a corresponding unoccupied state on the right hand side for electron transfer to take place. When these conditions are met, $T(k)$, provides the proportion of current passing through the barrier. To express the current density in terms of energy, we note first that;

$$dk = \frac{dk}{dE}dE = \frac{1}{\hbar v}dE, \qquad [2.20]$$

which gives the equation,

$$\begin{aligned} J_{L\to R} &= 2q\int_{U_L}^{\infty}\frac{dE}{2\pi\hbar v}\Big\{f(E_L,\mu_L)\big[1-f(E_R,\mu_R)\big]\Big\}vT_{L\to R}(E) \\ &= \frac{2q}{h}\int_{U_L}^{\infty}f(E_L,\mu_L)\big[1-f(E_R,\mu_R)\big]T_{L\to R}(E)dE. \end{aligned} \qquad [2.21]$$

The velocity term has cancelled, leaving the current density dependent only on the transmission coefficient and the Fermi functions. These higher energy states have higher velocities and therefore potentially greater current carrying capacity, but their lower density negates this contribution. Current due to electrons tunnelling from right to left can similarly be calculated, giving

$$J_{R\to L} = -\frac{2q}{h}\int_{U_R}^{\infty}f(E_R,\mu_R)\big[1-f(E_L,\mu_L)\big]T_{R\to L}(E)dE. \qquad [2.22]$$

The sum of these current densities provides the overall current density through the barrier. Where we are dealing with small voltage differences, it is reasonable to assume that $T_{L\to R}(E)=T_{R\to L}(E)=T(E)$. Also, we can note that a tunnelling electron requires an empty state on the other side of the barrier, which can only exist above the greater of U_L or U_R. The net current is therefore,

$$\begin{aligned} J &= J_{L\to R}(E) - J_{R\to L}(E) \\ &= \frac{2q}{h}\int_{U_L}^{\infty}\big[f(E_L,\mu_L)-f(E_R,\mu_R)\big]T(E)\,dE. \end{aligned} \qquad [2.23]$$

If the electron densities in the two electrodes are the same, they will have the same Fermi energies at equilibrium. Also, if the electrodes are at the same temperature we

find that the bias voltage provides the difference in the chemical potentials between the right and left hand leads and therefore equation 2.23 becomes,

$$J = \frac{2q}{h}\int_{U_L}^{\infty}[f(E) - f(E+eV)]T(E)dE. \qquad [2.24]$$

Equation 2.24 displays the fact that the current density is not a simple function of the bias. The properties of $T(E)$ are energy dependent and the electron distributions are temperature dependent, through the Fermi functions. Therefore, the current density is dependent on the total energy of the incident electrons of which the applied bias only forms a part. A number of important notes can be made however and are discussed below.

2.2.2.1 LARGE BIAS

At large bias, none of the states on the right hand side will have available states on the left hand side into which they may tunnel, where the system is biased as in figure 2.3. Therefore, the contribution from $f(E_R,\mu_R)$ can be neglected. Also, large bias can distort the shape of the barrier, reducing its strength and resulting in a higher $T(E)$ and consequently higher current densities.

2.2.2.2 LOW TEMPERATURE

At low temperatures, the occupation of states near the chemical potential becomes degenerate and the Fermi function approximates to a step function and the integral of 2.22 reduces to:

$$J(E) = \frac{2q}{h}\int_{\mu_R}^{\mu_L}T(E)dE. \qquad [2.25]$$

2.2.2.3 SMALL BIAS

Writing $\mu_L = \mu + 1/2qV$ and $\mu_R = \mu - 1/2qV$ (where μ is the Fermi energy at equilibrium and V is the voltage bias), the sum of the Fermi functions in Equation 2.23 can be approximated by a Taylor expansion, to first order, as,

$$f(E,\mu_L) - f(E,\mu_R) \approx -eV\frac{df(E,\mu)}{d\mu}, \qquad [2.26]$$

where $q = -e$. Therefore, the current density becomes,

$$J \approx \frac{2e^2}{h} V \int_{U_L}^{\infty} \left(-\frac{df(E,\mu)}{d\mu} \right) T(E) dE. \qquad [2.27]$$

The small bias has little effect on $T(E)$ and, thus, the current varies linearly with the applied voltage giving rise to a tunnel conductance which obeys Ohm's law;

$$G = \frac{2e^2}{h} \int_{U_L}^{\infty} \left(-\frac{df(E,\mu)}{d\mu} \right) T(E) dE. \qquad [2.28]$$

It is important to realize the difference between this tunnel conductance and a standard drift-diffusion type conductance. In this tunnel conductance, the electron transfer is quantized in units of one electron. The process of tunnelling is, however, stochastic; that is electrons only have probabilities of tunnelling, they are not constrained to tunnel within a given time. Therefore, the current through the barrier is subject to variations in the timing of individual tunnel events and subsequently the current through the barrier displays a shot noise spectrum[27,28], characterized by the Schottky formula for noise spectral density of $S_I(f)=2eI_{DC}$ A^2Hz^{-1}.

2.2.2.4 LOW TEMPERATURE AND SMALL BIAS

In the combined limits of low temperature and small bias, we find that $-df/d\mu \rightarrow \delta(E-\mu)$. This results in the integral of equation 2.28 disappearing and the current depending only on the transmission coefficient, giving a simple form for the tunnel conductance which is;

$$G = \frac{2e^2}{h} T(\mu). \qquad [2.29]$$

2.2.3 WKB APPROXIMATION

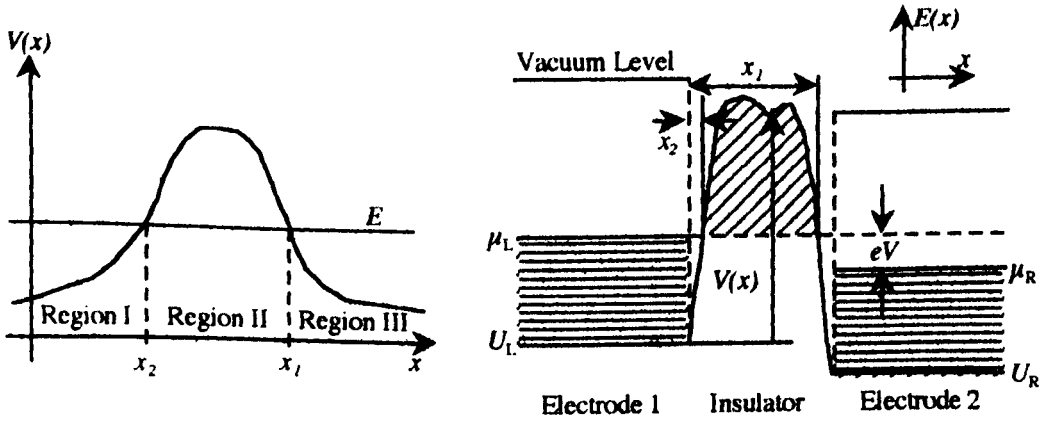
The difficulty with the rectangular barrier model is that it is an idealization. Few barriers in real devices will be strictly rectangular and where biases are applied, no barriers are rectangular. Changes in barrier shape and rounding of the potential steps occur due to depletion effects and image potentials, resulting in complex barrier shapes. In metals the situation is somewhat simpler since the large carrier density significantly reduces depletion effects. However image potentials still exist and affect the barrier shape. For arbitrary or complex barrier shapes, the tunnelling probability or

transmission coefficient becomes difficult to calculate. Matrix methods[23] can be applied, for example scattering, **S**, matrices or transmission, **T**, matrices. These matrix techniques allow complex barrier shapes to be constructed with basic barrier models and numerical methods to provide near exact solutions for the modelled problem.

The most widely applied analytical method is the WKB approximation. This was introduced to quantum mechanics in 1926 independently by G. Wentzel, H.A. Kramers and L. Brillouin. This approximation requires slowly varying potentials and is generally applied to one dimensional problems. A slowly varying potential is such that $V(x)$ changes only slightly over the de Broglie wavelength of the particle. The de Broglie wavelength is calculated as:

$$\lambda(x) = \frac{h}{p(x)}, \quad p(x) = \sqrt{2mE}. \qquad [2.30]$$

At an energy of 11.7eV, E_f for Al, $\lambda(x)$=3.6Å. For naturally grown AlO_x in Al based MIM junctions, the oxide thickness is generally quoted to be 2 to 3nm. This is roughly 5 to 9 times the de Broglie wavelength in Al and therefore variations in $V(x)$ should be of the order of $\lambda(x)$ or larger. Therefore, the WKB approximation applies to this case.



a. Example barrier potential for WKB calculation.

b. "Arbitrary" barrier used for Simmons' calculations.

Figure 2.4. Arbitrary barrier potentials for WKB approximation.

The WKB approximation assumes that the total tunnelling probability for an arbitrary barrier can be formed by the product of the transmission probabilities of small "sub-barriers". These sub barriers are arranged to simulate the shape of the barrier in question. The transmission probability for each sub barrier can be calculated from equation 2.15, however, where the particle energies are small, $E \ll V_o$, the transmission probability result is dominated by the exponential term of equation 2.15. Therefore,

provided the assumption of small particle energy holds for all points in the barrier, we can examine the limit where the sub-barrier widths tend to zero. Here, the WKB[25] approximation finds the transmission coefficient of an arbitrary barrier potential such as figure 2.4a to be,

$$T \approx exp\left(-2\int_{x_2}^{x_1} \kappa\, dx\right), \quad \kappa = \sqrt{\frac{2m(V(x)-E)}{\hbar^2}}. \qquad [2.31]$$

Simmons[13,14] applied this approximation to the barrier shown in Figure 2.4b. This analysis included the effects of image forces on the barrier shape. Image forces tend to lower and round the barrier profile, thereby reducing the barrier area and strength. In analysing the problem Simmons looked at variations in the barrier height and width, with each barrier including the modifications resulting from image potential created by the bias and electrodes. In each case curves for the tunnel resistivity, ρ_T (= V/J), were presented and showed good agreement with theories using rectangular barriers where large barrier potentials were used. However, at lower barrier potentials a marked deviation from this theory was observed. This deviation is marked by a lower tunnel resistivity where image potentials affect the barrier strength. Pollack and Morris[15] compared experimental measurements on oxidized Al electrodes having varying metal counter electrodes with Simmons' model and found excellent agreement over nine orders of magnitude in tunnel current. Finally, a study of the potential barrier structure within thin film Al-Al_2O_3-metal diodes[29] was carried out. In this case, the aluminium oxide was formed by anodization, resulting in a measured value of ≈2.0eV for the barrier height in the Al MIM system.

2.3 SINGLE ELECTRON TUNNELLING

Tunnelling is an inherently single particle process. That is, individual particles transfer through the barrier and do not exist simultaneously on both sides of the barrier. A tunnel junction can be characterized by its resistance and capacitance, although these may be significantly different from geometrical estimates[30,31]. The fact that the junction has a capacitance is the root of single electron tunnelling (SET) behaviour. The capacitance of a junction depends on its size and for a junction which can be approximated by a parallel plate model, the capacitance is given by the relation,

$$C = \frac{\varepsilon_r \varepsilon_0 A}{d}, \qquad [2.32]$$

where the junction has an overlap area, A, a plate separation, d, and a dielectric between

the electrodes of relative permitivity ε_r. The presence of this capacitance requires that the junction has an energy associated with the charging process. This electrostatic charging energy is given by:

$$E_c = \frac{Q^2}{2C}, \qquad [2.33]$$

where Q is the charge on the plates. If charge transport is governed by tunnelling then an energy cost of $e^2/2C$ is associated with each electron transported through the barrier. If the contact area of the junction is relatively large, capacitances with only a small associated charging energy result. Table 2.1 gives an indication of the magnitude of this charging energy, assuming an Al_2O_3 dielectric with ε_r=10, and a square, parallel plate, junction of side l and d = 2nm, giving $\varepsilon_r\varepsilon_o/d$=66.41x10^{-3}Fm^{-2}.

l	A /m^2	C /F	E_c /eV	Transition temperature T_o /K
1mm	1x10^{-6}	66.41x10^{-9}	1.205x10^{-12}	13.97x10^{-9}
1μm	1x10^{-12}	66.41x10^{-15}	1.205x10^{-6}	13.97x10^{-3}
0.1μm	1x10^{-14}	664.1x10^{-18}	120.5x10^{-6}	1.397
10nm	1x10^{-16}	6.641x10^{-18}	12.05x10^{-3}	139.7
1nm	1x10^{-18}	66.41x10^{-21}	1.205	13.97x10^{3}

Table 2.1. Capacitance and charging energy scales for Al MIM tunnel junctions.

To see the significance of the charging energy for a given junction size, it must be compared with the energy scale of thermal fluctuations, i.e. $E \approx k_B T$. Electrons will tunnel through the barrier if they are provided with enough energy to make the tunnelling process probable. This energy can be derived from the applied bias or from the environment, through temperature or other energy inputs. If the charging energy is less than or comparable to the available thermal energy, thermal fluctuations will mask any effects due to the charging energy. This results in the charging energy having little or no effect on the tunnelling process through the junction. This leads to the definition of the transition temperature, T_o, for the onset of SET effects;

$$T_o = \frac{E_c}{k_B} = \frac{e^2}{2k_B C} \qquad [2.34]$$

From table 2.1, we can see that for junction areas above 100 x 100nm^2 and temperatures greater than 1K, the charging energy is small and as a result plays little part in the tunnelling process. If, however, the junction area is reduced below the 100nm x 100nm scale we see some significance in the charging energy, rising above the value of around

100μeV. In this instance, cooling a sample below 1K would result in reduction of thermal fluctuations to such an extent that any effect of the charging energy could be observed. This leads to an inequality relating the SET effect observation temperature, T, to the charging energy through the transition temperature;

$$T << \frac{e^2}{2k_B C}. \qquad [2.35]$$

By similar arguments, SET effects will not be observable if other relevant energy scales are not less than the charging energy. These energy scales include environmental noise (from measurement circuits, etc.), electromagnetic noise (from circuits, electromagnetic radiation and radio frequency interference) and quantum fluctuations. In the case of quantum fluctuations, the energy time expression of the Heisenberg uncertainty principle leads to a further constraint on the junction characteristics. The lifetime, τ, of a given energy state, E, is constrained by the relation $\Delta E \Delta \tau \geq h$. In a tunnel junction a characteristic time, τ_T, can be defined as $\tau_T = R_T C$, where R_T is the tunnel resistance. Using the charging energy and this characteristic time, an order of magnitude inequality results which governs the value of the tunnel resistance for which SET effects are observable:

$$R_T > R_q, \quad R_q = \frac{h}{e^2} = 25.8\text{k}\Omega \qquad [2.36]$$

When the constraints on charging energy scale, temperature and junction resistance are met, SET effects can become observable. Single electron tunnelling effects result from the fact that the energy cost involved in transferring one electron across the junction can prevent the transfer of further electrons until such time as the charge, $\Delta Q = e$ is removed from the junction by the electron current. Effects caused by this mechanism were observed as an increase in the low electric field resistance of granular thin film structures some years ago[32-35]. The new effects of concern here arise from the realization that charging effects can cause correlations in tunnelling events even though the tunnelling process is itself stochastic.

2.3.1 COULOMB BLOCKADE OF TUNNELLING

The orthodox theory[36] of single electron tunnelling emerges from a Hamiltonian based description of a tunnel junction connected to a current source, as shown in Figure 2.5. In this description, the charging energy, the electronic states of the electrodes, the tunnelling process and coupling to a shunt resistance are considered. Development of the theory in this framework readily describes both normal metal and superconducting

systems alike. To consider low dimensional semiconductor structures, modifications which account for discrete energy levels in the system are required but the framework still applies.

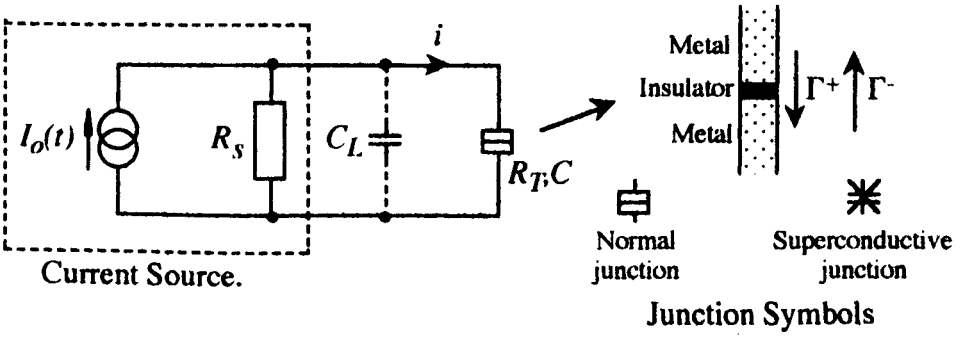


Figure 2.5. Circuit diagram for single junction model.

To calculate the tunnelling rates, the tunnelling Hamiltonian is treated as a perturbation of the system. Fermi's "Golden rule" is then applied to give the rate of scattering from an initial to a final state. It is this scattering which provides the tunnel current in this model. This treatment requires the tunnelling resistance to be large compared with the resistance quantum, as in equation 2.36 and that equilibrium is established in between successive tunnelling events. This ensures that the electron states in either electrode inter-mix only very weakly, allowing the assumption of a perturbation and that charges are redistributed before each tunnelling event.

For the circuit of Figure 2.5, we require $R_S >> R_q$ and $R_T \geq R_q$ and a relatively small current ($i < 0.1e/\tau_T$, $\tau_T = R_T C$). In this case the tunnelling rate is found to depend on the energy change in the circuit and is, for a normal metallic system given by;

$$\Gamma^{\pm}(Q) = \frac{\Delta E^{\pm}}{e^2 R_T}\left[1 - exp\left(-\frac{\Delta E^{\pm}}{kT}\right)\right]^{-1}. \qquad [2.37]$$

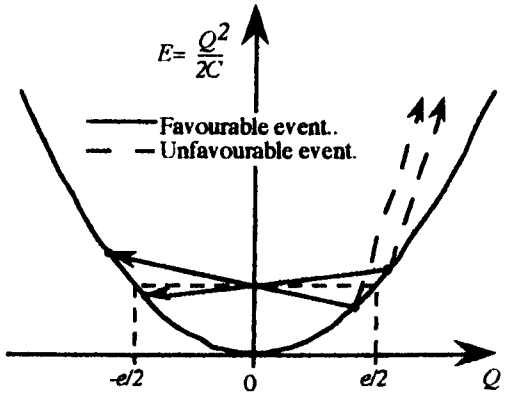
Here, $\Delta E^{\pm}$ is the change in electrostatic energy caused by the tunnelling event. Tunnelling of an electron causes a change in the junction charge of $Q \rightarrow Q \pm e$. For tunnelling to occur $\Delta E^{\pm}$ must be positive. Therefore;

$$\begin{aligned} \Delta E^{\pm} &= E(Q) - E(Q \pm e), \\ &= \frac{Q^2}{2C} - \frac{(Q \pm e)^2}{2C}, \\ &= \mp \frac{e}{C}\left(Q \pm \frac{e}{2}\right). \end{aligned} \qquad [2.38]$$

and,

$$\Delta E^{\pm} > 0 \ , \quad \text{for } |Q| > \frac{e}{2}. \qquad [2.39]$$

Therefore, a region of "energy - charge" space exists for the junction where tunnelling is energetically unfavourable. This leads to a blocking of the current flow through the junction which is termed the "Coulomb blockade of tunnelling". Figure 2.6 shows the energy changes required by the quadratic dependence of ΔE and illustrates the energetically favourable events.



Change in energy of a small tunnel junction due to tunnelling of a single electron. For two values of initial charge (Q).

Figure 2.6. Energy-charge space diagram for a single junction.

2.3.1.1 DC CURRENT-VOLTAGE CHARACTERISTIC

Where we have the circuit of figure 2.5, we have three requirements for the observation of Coulomb blockade effects. Firstly, the shunt resistance is large, $R_S \to \infty$. Secondly, the tunnelling resistance obeys the inequality $R_T \geq h/e^2$. Thirdly, we have low temperatures, ideally $T=0$. When these conditions are met, we find that for $|Q|<e/2$ tunnelling is blocked. The effect of this behaviour on the dc-current-voltage characteristic is to split the ohmic ($I=V/R_T$) characteristic about $V=0$, as shown in figure 2.7. This produces a non-conducting state if the junction voltage is less than the critical or threshold voltage, $V_t=e/2C$. As this threshold is exceeded, the junction enters a conducting state where the current approaches the asymptotic value of $I=V/R_T$ but contains a voltage offset from zero of $V_{off}=\pm e/2C$. The $I(V)$-curve has the functional dependence given by equation 2.40.

$$I(V) = \frac{1}{R_T}\left[V - V_{off}(sign.I)\right], \text{ where, } V_{off} = \frac{e}{2C}. \qquad [2.40]$$

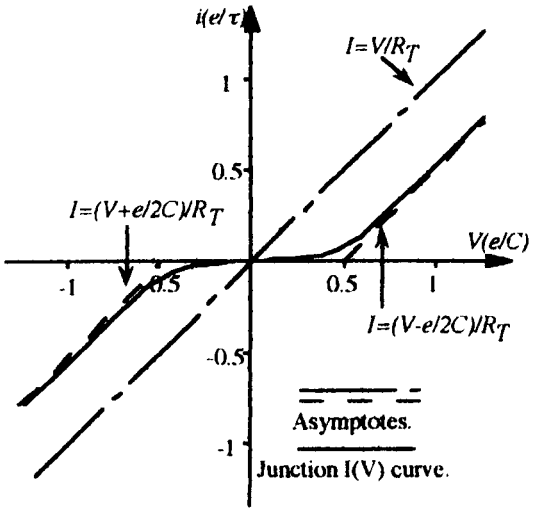


Figure 2.7. dc I(V) characteristic for a single junction.

2.3.1.2 TIME CORRELATION OF TUNNELLING

If we consider a small tunnel junction in the non-conducting state, the bias current of figure 2.5 controls the charge on the junction. Here we assume metallic, low resistance electrodes which allow for rapid charge redistribution. A low bias current, $i<0.1e/\tau_T$, will provide a continuous polarization of the junction charge with rate $dQ/dt=i$. As $|Q|$ approaches $e/2$, under the action of this current, the tunnelling probability increases until an electron tunnels and the junction charge is set to the opposite end of the region,

$$-\frac{e}{2}<Q<\frac{e}{2}. \qquad [2.41]$$

This process leads to a time correlation of the tunnelling electrons producing a voltage oscillation with amplitude $\approx V_t$, the threshold voltage, and frequency f_{SET} plus harmonics[36]. The oscillation frequency is given by equation 2.42, where $\bar{V}$ is the average voltage across the junction.

$$\begin{aligned} f_{SET} &= \frac{1}{e}\left(i-\frac{\bar{V}}{R_s}\right) \\ &= \frac{i}{e}, \quad \text{for, } R_s >> R_T \ \&\ i << \frac{e}{R_T C}. \end{aligned} \qquad [2.42]$$

Superposition of a high frequency current onto the dc bias current should result in resonances creating steps in the $I(V)$-curve at currents $i=(n/m)ef$. n and m describe phase locking of the m-th harmonic of the SET oscillation with the n-th harmonic of the external frequency. This effect has not yet been observed in single junction circuits.

Increasing i above $\approx 0.1e/\tau_T$, causes the amplitude of the oscillations to fall and the average junction voltage and background noise level to rise. As the tunnelling process is stochastic, a shot noise frequency characteristic is superimposed on the spectral components of the oscillations. This reduction in the oscillation amplitude at higher current causes the SET oscillations to become masked by the increased shot noise level. At these low voltages, the *I(V)*-curve has a quadratic dependence and the average voltage across the junction is given by[36],

$$\bar{V} = \frac{e}{C}\left(\frac{\pi i \tau_T}{2e}\right)^{1/2} = \sqrt{\frac{e\pi i R}{2C}} \qquad [2.43]$$

Increasing the current washes out the SET oscillations and produces the quadratic dependence of the *I(V)*-curve at low voltages. The *I(V)*-curve approaches the offset asymptotes at higher junction voltages.

2.3.1.3 EFFECT OF THE ENVIRONMENT AND TEMPERATURE

The theory discussed above is an idealized case of the single junction circuit. While the junction itself and the processes involved are adequately detailed, no account is taken of a real system. The environment of a single junction includes the physical connections to it and therefore, any series (leads etc.) or parallel (leakage) resistances, and any stray (usually parallel) capacitances. The shunt resistance cannot be infinite, especially as any real impedance has frequency dependent components, but if $R_S \geq 10R_T$, the theory is a good approximation[37]. The resistance in question must be taken as $R=min[R_T,R_S]$.

As R approaches R_T, the current range over which the SET oscillations have a narrow frequency spectral linewidth and large amplitude reduces. Continued reduction of R will eventually result in the SET oscillations vanishing. As R_S becomes less than R_T, time correlations between tunnelling events are lost. The SET oscillations will vanish and the gap in the *I(V)*-curve begins to exhibit offsets and asymptotic behaviour only at large voltages.

The charging energy of a single junction is described through its capacitance. However, the total capacitance in the single junction circuit is that seen by the bias current, $I_O(t)$. This capacitance necessarily involves at least three capacitances; any stray capacitance of the leads (C_L), the self capacitance or capacitance to ground of the junction structure (C_S), and the junction capacitance itself (C_T). These components combine to give a Thévenin equivalent capacitance of C_Σ, for a current biased single junction, as depicted

in figure 2.5, of,

$$C_\Sigma = C_T + C_S + C_L \qquad [2.44]$$

with the charging energy becoming,

$$E_c = \frac{Q^2}{2C_\Sigma} \qquad [2.45]$$

Generally, $C_S << C_T$. However, $C_L >> C_T$, and the charging energy becomes very much reduced. This lowers T_o of the single junction and masks the charging effects by thermal fluctuations. Resistances, with $R > R_q$, may be added close to the junction to decouple this effect[38]. For temperatures, $T > T_o$, SET oscillations become washed out and the Coulomb gap in the $I(V)$-curve disappears. However, for $0 < T < T_o$ random tunnelling events are induced at $V < V_t$ and the width of the spectral components of the oscillations are increased, with a corresponding decrease in the voltage amplitude[37,39]. The environment to which a junction is connected is therefore extremely important in determining the effects that can be observed[40]. The discussion above shows that to observe the full range of SET effects, the junction must be decoupled from the measurement environment. It is for these reasons that, to date, SET oscillations have not been observed in a single junction and why most experiments employ at least double junction structures which provide an island decoupled from its environment.

2.3.1.4 OBSERVED SINGLE JUNCTION CHARACTERISTICS

Single junctions isolated from their environment through arrays of tunnel junctions[41,42] or through high impedance leads[38] have been measured. In all cases, the decoupling impedance has been fabricated on chip. Each of these methods have produced clear $I(V)$-curves with well defined offset voltages. SET oscillations from a single junction have not yet been indisputably directly observed. Some authors have reported phase locking of Bloch oscillations in granular systems[43,44], and in superconducting single junction systems[45,46] resulting in plateaux in the I(V)-curve. These "Bloch" oscillations are the superconducting equivalent of the SET oscillation, having the same physical root and characterized by $f = I/2e$, where the $2e$ term arises from the transfer of Cooper pairs across the junctions. The measurements of Kuzmin et al[45,46] provide clearer observation of the single junction effect, as they used a single junction fabricated with high impedance leads. Addition of a high frequency component to the bias voltage produced resonances in the curve at currents of $I = 2nef$, where ($n = \pm 1, \pm 2, \ldots$). In the experiments by Yoshihiro[43,44], however, it is doubtful if the resonances observed result from a single junction.

2.3.2 DOUBLE TUNNEL JUNCTIONS

The double tunnel junction comprises two junctions in series, forming an island in the middle. Transport into and out of the island is controlled by tunnelling and the number of excess electrons on the island is controlled by the island electrostatic energy. In this system, the requirements placed on R_T and T are the same as for the single junction case. This ensures that the charge state of the island is well defined. In effect, this decouples the island from its environment. For example, if the single junction in the circuit of Figure 2.5 is replaced by two junctions in series, the island is decoupled from C_L and the action of the bias current on C_L is to create a variable voltage bias across the junctions. The capacitance of the island then becomes the sum of the two junction capacitances and any stray capacitance of the island.

2.3.2.1 DESCRIPTION OF THE DOUBLE JUNCTION CIRCUIT

Figure 2.8 shows the equivalent circuit of a double tunnel junction single electron device. This figure shows the device being operated in voltage bias mode. Any external circuit connected to this device will see an equivalent capacitance, C, and equivalent voltage, V, given by equations 2.46 and 2.47 respectively.

$$C = \frac{C_1 C_2}{C_1 + C_2}. \qquad [2.46]$$

$$V = \frac{Q_1}{C_1} + \frac{Q_2}{C_2}. \qquad [2.47]$$

As such, there exists a continuous variable, Q, which is the total charge transferred through the circuit exists and is given by equation 2.48.

$$Q = CV = \frac{C_2 Q_1 + C_1 Q_2}{C_1 + C_2}. \qquad [2.48]$$

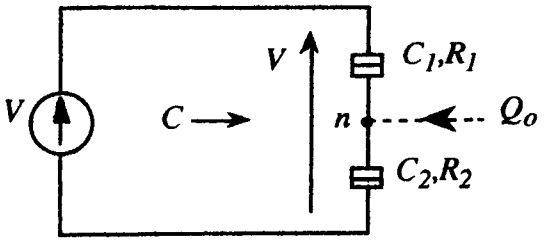


Figure 2.8. Equivalent circuit of a double junction device.

The metallic island formed between the tunnel junctions contains a charge of

$Q_1 - Q_2 = ne$, where n is the number of excess electrons on the island. $n = n_1 - n_2$, where n_1 and n_2 are the numbers of electrons transferred through junctions 1 and 2 respectively. This charge state is well defined if the tunnel resistances satisfy $R_1, R_2 >> R_q$. The island charge may be modified by a background charge Q_o. This charge can arise from effects of charge distributions in the vicinity of the junction having an electrostatic effect on the island charge. Similarly, the work functions of the electrode materials may induce a particular charge state on the island, leading to an initial state where the island contains excess electrons. In any case, Q_o is effectively continuous on the scale of the electron charge and therefore has a continuous effect on the island charge. The island charge, Q_i, can then be expressed as,

$$Q_i = ne + Q_o \qquad [2.49]$$

Where we have a system at low temperatures connected to a low impedance voltage bias, we find the tunnelling rates are given by equation 2.50. In this equation, $\Gamma_j^{\pm}(n)$ is the rate of tunnelling through junction j for ($\Delta n_j = \pm 1$). For tunnelling off the island, $n_j = -1$ and for tunnelling onto the island, $n_j = +1$.

$$\Gamma_j^{\pm}(n) = \frac{\Delta E_j^{\pm}}{e^2 R_{Tj}} \left[1 - exp\left(- \frac{\Delta E_j^{\pm}}{kT} \right) \right]^{-1} \quad , \text{ where, } \quad j = 1,2. \qquad [2.50]$$

Now,

$$\begin{aligned} \Delta E_1^{\pm} &= \frac{Q^2}{2C_\Sigma} - \frac{(Q \pm e)^2}{2C_\Sigma} \mp \frac{eC_2}{C_\Sigma} V \\ &= \frac{e}{C_\Sigma} \left(-\frac{e}{2} \mp (ne + Q_o) \mp C_2 V \right) \end{aligned} \qquad [2.51]$$

$$\begin{aligned} \Delta E_2^{\pm} &= \frac{Q^2}{2C_\Sigma} - \frac{(Q \pm e)^2}{2C_\Sigma} \mp \frac{eC_1}{C_\Sigma} V \\ &= \frac{e}{C_\Sigma} \left(-\frac{e}{2} \pm (ne + Q_o) \mp C_1 V \right) \end{aligned} \qquad [2.52]$$

From equations 2.51 and 2.52, we see that the energy change on tunnelling is governed by the island charge ($ne + Q_o$) and the bias voltage. If the island charge and bias voltage are such that ΔE_1 & $\Delta E_2 < 0$, the system exists in a stable state where electrons cannot transfer through the island and the tunnelling process is blocked. Provided the temperature is low enough, it is impossible for current to flow through the island while the system remains in this state. In this case the island charge reflects a near integer number of electrons, n, and addition or removal of an electron incurs an energy cost given by the charging energy of the island. Tunnelling is therefore energetically

unfavourable. That is, the initial energy state, $E(n)$, is much lower than the energy state where the island charge differs by one electron, $E(n+1)$ or, $E(n)<<E(n+1)$. The difference between $E(n)$ and $E(n+1)$ in this case is the island charging energy and this phenomenon provides Coulomb blockade of tunnelling in a double junction system.

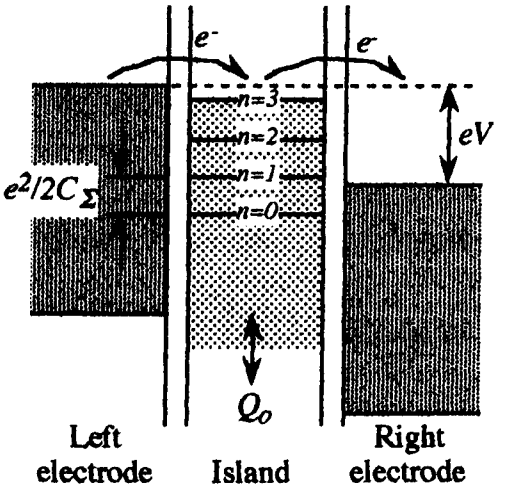
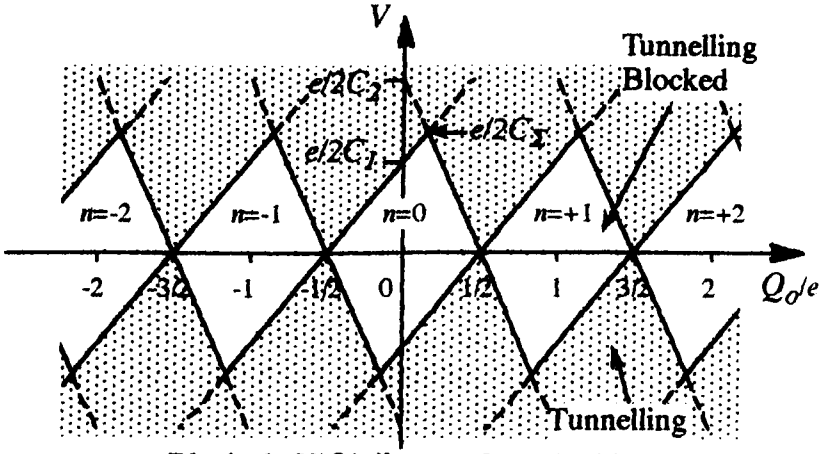


Figure 2.9. Quantization of electron numbers in the island.

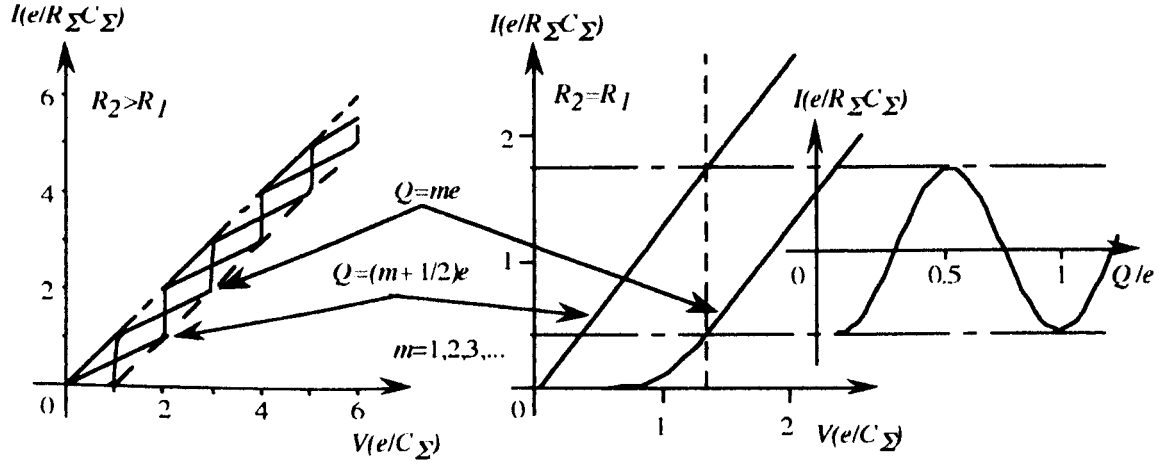


Blockade $V(Q)$ diagram for a double junction circuit at zero temperature with $C_1<C_2$

Figure 2.10. Double junction Q_o-V diagram.

Increase of the bias voltage can overcome this situation, as can modification of the island charge. However, without extra energy input this system will remain in a non-conducting state. Therefore, if we increase of the bias voltage to a threshold of $V_t \approx min[|e/2C_1|,|e/2C_2|]$ leads to tunnelling of an electron through one of the junctions. This event changes the electrostatic energy of the island. Correspondingly, the island potential changes by $\Delta V=e/C_\Sigma$, which leads to a tunnel event in the other junction, returning the island potential to the initial value. As we have a low impedance voltage bias, the potential across both junctions is fixed and thereby, the total charge in the

circuit is fixed, compared with the current bias described with the single junction circuit where the charge across the junction is continuously varied. As the second tunnelling event returns the island potential to the value which caused the initial tunnelling event, the complete charge transfer can be repeated, allowing current to flow through the island.



Effect of background charge on *I(V)* behaviour

a. "Coulomb staircase" in a double junction due to unequal tunnelling parameters.

b. Double junction with equal junction parameters.

Figure 2.11. Schematic *I(V)*-curves for $C_1 \neq C_2$ and $R_1 \neq R_2$ and variation in Q_o. (for $R_1, C_1 = R_2, C_2$)

Recall that the island potential is also modified by the charge Q_o, which is a continuous variable. This charge has the effect of shifting the island charge state and therefore the threshold required to transfer electrons through the island. Therefore, for any given bias voltage, variation of Q_o allows switching of the device between conducting and non-conducting states. This is illustrated in the Q_o-V diagram of figure 2.10, and is the effect which allows use of this system as a single electron transistor. In figure 2.10, the conducting and non-conducting states of the system are mapped out against the bias voltage and the background island charge. The shaded areas represent regions where the system is in a conducting state. This shows clearly that increasing the bias voltage above the appropriate threshold switches the double junction system into a conducting state. However, changing the background charge of the system, for a bias voltage $V<V_t$, results in the system being switched between conducting and non-conducting states periodically in Q_o. This state switching arises because the background charge can force an island charge of $Q=(n\pm 1/2)e$, as easily as a charge of $Q=ne$. Where $Q=ne$ we have Coulomb blockade of tunnelling as already discussed. However, where $Q=(n\pm 1/2)e$, we have a situation where the initial state $E(n)$ has energy roughly equal to that of $E(n+1)$ or, $E(n) \approx E(n+1)$. There is therefore little energy cost in the tunnelling

process and the blockade of tunnelling is lifted.

The dc $I(V)$-curve of this system depends on the characteristic parameters of the junctions (C_1,R_1,C_2 & R_2) that is, the tunnelling rates and charging energies. For unequal junction parameters, $C_1 \neq C_2$ and $R_1 \neq R_2$, the two junctions have different thresholds and different tunnelling rates resulting in steps in the $I(V)$-curve. Two distinct periods ($\Delta V_1 = e/C_1$ and $\Delta V_2 = e/C_2$) are observed at low temperature which are sometimes termed the Coulomb staircase, shown in Figure 2.11a. Any variation in the background charge Q_o will cause a change in the phase of these oscillations, $\Delta\theta = (2\pi/e)Q_o$. The effect of Q_o is periodic in electronic charge because a change of Q_o by e leads to an extra charge state being absorbed into the island charge state, ne. Effectively this is the same as adding an extra electron to the island. The effect on the $I(V)$-curve is shown in Figure 2.11b.

Increasing the bias voltage well above the threshold levels leads to suppression of the oscillations, and the $I(V)$ behaviour approaches the asymptotes of,

$$I(V) = \frac{1}{R_\Sigma}\left[V - V_{off}(sign.V)\right] \text{ , for, } |V| >> \frac{R_{min}}{R_{max}}\frac{e}{C_\Sigma}. \qquad [2.53]$$

$$\text{where, } R_\Sigma = R_1 + R_2, \quad V_{off} = \frac{e}{C_\Sigma}$$

If we have a strongly asymmetric structure, with for example $R_1 << R_2$, only one step will be visible. In this case, $\Delta V = e/C_2$ and the low current behaviour of the I(V)-curve is described by equation 2.54. A study of the conditions for observation of the Coulomb staircase was carried out by Hanna and Tinkham[47].

$$I(V) = \frac{1}{R_2 C_\Sigma}\left[Q_o + C_1 V_1 + en_o - \frac{e}{2}(sign.V)\right], \quad n_o = Int\left[C_2 V - Q_0 + \frac{e}{2}\right] \qquad [2.54]$$

$$\text{for, } R_1 << R_2$$

In the double junction circuit, increased temperature leads to a smoothing of the oscillations in the I(V)-curve until, eventually, the Coulomb blockade is lifted at $T \approx e^2/2C_\Sigma$.

2.3.2.2 LEAKAGE OR CO-TUNNELLING

As tunnelling is a quantum mechanical effect, the process is governed by quantum mechanical rules. The key issue here is that unless a process is strictly forbidden, through some fundamental property or symmetry of the system, quantum mechanics

allows it to occur. The probability and therefore rate of such processes may be small, but it still exists. In a system with two or more tunnel junctions, a probability exists for electrons to tunnel coherently through the island, even when the system is in the Coulomb blockade state. This effect was first considered by Averin and Odintsov[48,49] and results in a leakage current[50] through the junction system, even at zero temperature. The general effect is similar to a raised system temperature. In a single junction system, it is possible for the charge packet to tunnel through the $E=e^2/2C$ barrier, thus changing the junction charge. However, in this system it is unlikely that this process will lead to a change in the system state. In the case of double or multi-junction systems this is not necessarily the case. In a double junction system this type of process leads to a leakage current rather than state changes. The following considers this in more detail.

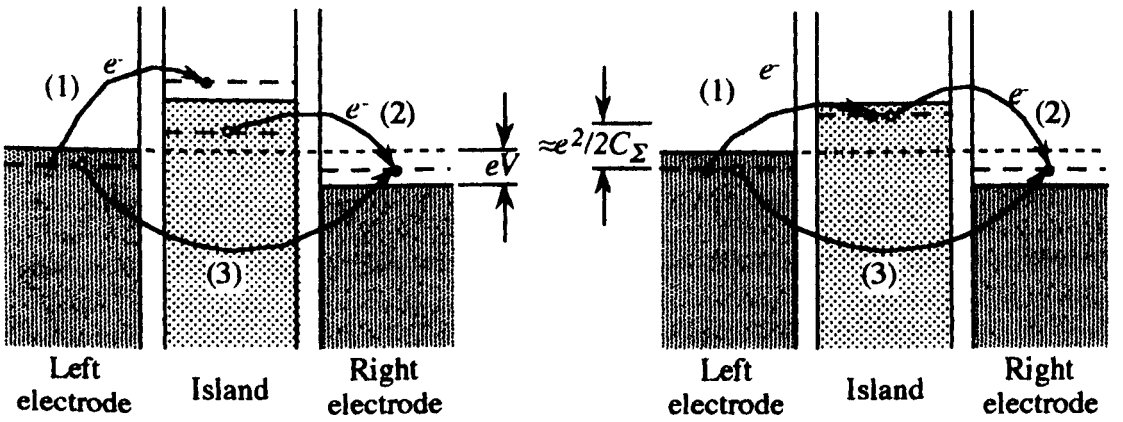


Figure 2.12. Illustration of co-tunnelling in a double junction.

Figure 2.12 illustrates possible processes for a double junction system biased at $eV<e^2/2C_\Sigma$. The process can be described as tunnelling through the Coulomb barrier via a virtual state, or as macroscopic quantum tunnelling of charge, q-MQT. The two processes represented by figures 2.12 a and b are those of inelastic q-MQT and elastic q-MQT respectively. In the case of elastic q-MQT, the virtual state is the same for the incoming electron as for the outgoing electron. That is, tunnelling through both barriers employs the same particle state. This necessarily requires more stringent selection rules within the quantum mechanical model. Elastic q-MQT is likely to be significant where the density of states in the island is low, such as in low dimensional semiconductor devices. In the metallic systems studied here, the island density of states will be large and therefore inelastic q-MQT will be the most dominant of the two processes.

In figure 2.12, an electron tunnels through the first barrier, by process 1 and, by the uncertainty principle, is allowed to exist there for a time $\Delta t=\hbar/\Delta E$ where $\Delta E\approx e^2/2C_\Sigma$. If

the electron state decays during this time by tunnelling of a second electron through the other barrier, process 2, an electron has effectively been transferred through the island via process 3. Provided the time Δt is short, ΔE can be very large, (through $\Delta E \Delta t \leq \hbar/2$) providing enough energy for an electron to tunnel through the entire energy barrier, allowing the vitual but truely quantum mechanical transfer of an electron via process 3. The rate of co-tunnelling is given by equation 2.55[57],

$$\Gamma_c(N,eV) = \frac{2\pi E_c N^{2N}}{\hbar(2N-1)!(N-1)!^2}\left(\frac{R_q}{4\pi^2 R_T}\right)^N \left(\frac{eV}{E_c}\right)^{(2N-1)} \qquad [2.55]$$

$$\text{for,} \quad eV << E_c \quad , \quad E_c = \frac{e^2}{2C_\Sigma}$$

where N is the number of tunnelling events required to transfer one electron charge through the structure. Since this tunnelling rate reduces as R_T rises above R_q and as N increases, the co-tunnelling effect is small but is nevertheless able to modify the device current[50]. The co-tunnelling rate is also dependent on the bias voltage. This effect arises from the $(eV/E_c)^{(2N-1)}$ term and reflects the increase in the number of electronic states available for the co-tunnelling process. The main difficulty caused by co-tunnelling arises in attempts to use SET systems for metrological work and as memory elements. In both cases, the leakage effect results in unreliable operation of the system. In metrological applications the accuracy of the measurement device can be severely reduced by this leakage. Similarly in memory systems, leakage of the island charge state would result in loss of the stored information.

2.3.2.3 THE SINGLE ELECTRON TRANSISTOR AND ELECTROMETER

The ability to alter the charge state of the island allows the development of SET transistors and electrometer devices. Figure 2.13 depicts the circuit set-up for such devices. In the case of the transistor, the gate voltage is treated as the input signal, while in the electrometer, V_g is a measured variable which senses the island charge to sub-e accuracy. The bias system represented in Figure 2.13 depicts a differential voltage bias of V volts for the double tunnel junction system. Addition of V_g and C_g, a normal capacitance, to this circuit allows the island charge to be controllably modified through $Q_g = C_g V_g$.

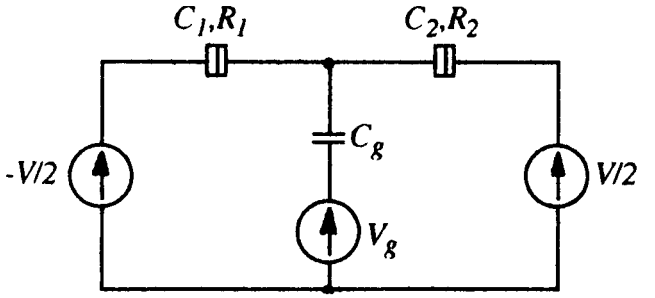


Figure 2.13. SET transistor / electrometer circuit.

Using Kirchoff's laws and superposition to analyse this circuit, we find that the charges are,

$$\begin{aligned} Q_1 &= \frac{C_1}{C_\Sigma}\left[\left(C_2+\frac{C_g}{2}\right)V+C_gV_g+ne\right], \\ Q_2 &= -\frac{C_2}{C_\Sigma}\left[-\left(C_1+\frac{C_g}{2}\right)V+C_gV_g+ne\right], \\ Q_g &= -\frac{C_g}{C_\Sigma}\left[\frac{1}{2}(C_2-C_1)V-(C_1+C_2)V_g+ne\right]. \end{aligned} \qquad [2.56]$$

The tunnelling rates are again given by equation 2.50, provided the supply voltage impedance is low. Also, the energy changes for $n \rightarrow n+1$ are given by,

$$\begin{aligned} E_1 &= \frac{e}{C_\Sigma}\left[e\left(n+\frac{1}{2}\right)-\left(C_2+\frac{C_g}{2}\right)V+C_gV_g\right], \\ E_2 &= \frac{e}{C_\Sigma}\left[-e\left(n+\frac{1}{2}\right)-\left(C_1+\frac{C_g}{2}\right)V+C_gV_g\right], \end{aligned} \qquad [2.57]$$

noting that $E_1(n)+E_2(n+1)=E_2(n)+E_1(n+1)=-eV$, is the work done by the voltage source.

The behaviour of this system is exactly as for the double junction in Figure 2.10, only Q_o is replaced by $Q_g=C_gV_g$. Thus tunnelling is suppressed by Coulomb blockade for voltages satisfying the inequalities,

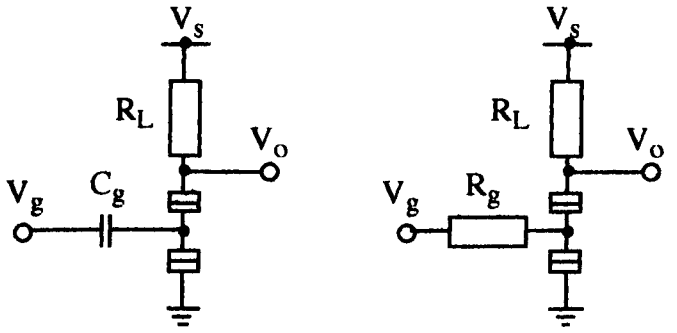
$$\begin{aligned} &\text{Junction 1;} \quad e\left(n-\frac{1}{2}\right) < C_gV_g+\left(C_2+\frac{C_g}{2}\right)V < e\left(n+\frac{1}{2}\right), \\ &\text{Junction 2;} \quad e\left(n-\frac{1}{2}\right) < C_gV_g-\left(C_1+\frac{C_g}{2}\right)V < e\left(n+\frac{1}{2}\right). \end{aligned} \qquad [2.58]$$

Therefore, at a constant voltage V the Coulomb gap in the dc I(V)-curve can be suppressed at gate voltages,

$$V_g = \frac{e\left(k+\frac{1}{2}\right)}{2C_g} \quad , \quad \text{for,} \quad k = 0, \pm 1, \pm 2, ... \qquad [2.59]$$

Sweeping of the gate voltage in a double junction system produces an oscillation in the tunnel current which is periodic in V_g, with period eC_g as shown in Figure 2.11b. This effect allows the use of the double junction as an electrometer. As current flows through the device, the island potential is continually modified by the presence of the electrons passing through it. Using the gate electrode to capacitively couple to a region where the potential is to be measured gives the ability to sense electric charge to sub-e accuracy.

Transistor behaviour can be achieved in a capacitively coupled SET transistor, figure 2.14a, if the device parameters are adequately controlled. The voltage gain, $A_V=|dV/dV_g|$ can be larger than 1 if $C_1=C_2=C<<C_g$ since $A_V \approx C_g/C$, whilst the current gain is $A_I(\omega) \approx 1/\omega R_T C_g$ (with $R_1=R_2=R_T$). Another device which has been proposed is the resistively coupled SET transistor. This device could exhibit large gain around $V_g \approx e/2C_\Sigma$ at $T=0$, due to the very large change in resistance when the device is switched between the conducting and non-conducting states. However to maintain island decoupling it is required that $R_g >> R_q$. The input resistance to this device is by necessity large and results in a limited current gain and low operational bandwidth through frequency limitations in the input circuit.



C-SET Transistor R-SET Transistor

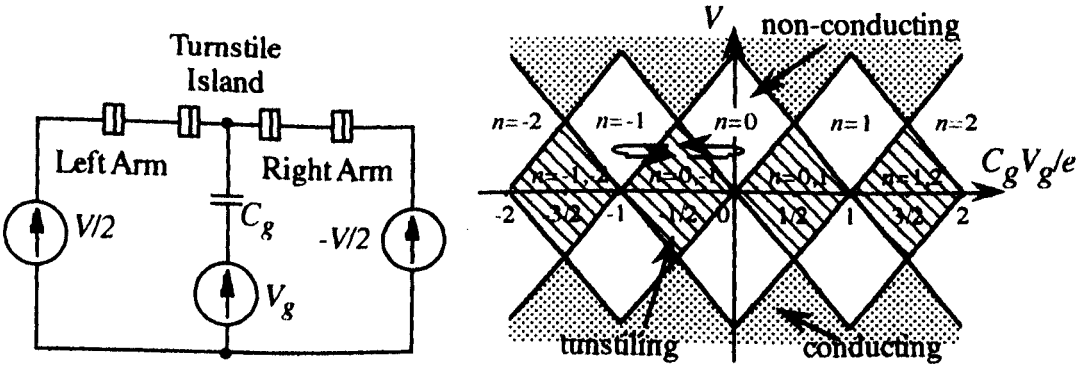
Figure 2.14. Single electron transistor configurations.

2.3.2.4 SINGLE ELECTRON TURNSTILE AND CHARGE PUMP DEVICES

The bias voltage in a double junction system can be used to put the circuit into either a conducting or non-conducting state. The gate voltage adds a further degree of freedom

to this system, allowing modulation or sensing of the island potential. As we have seen, this can be used either to sense the number of electrons flowing through the island, modulate the device current or switch the device between conducting and non-conducting states. However, the double junction configuration is unstable against the storage of an extra electron on the island. Adding an electron to the island, leads to an increase in the island potential which then results in an electron tunnelling off the island. As we have a system operating on the basis of single electron transfers, it should be possible to control individual tunnelling events and thereby provide absolute control of the current through the structure. The ability to trap an electron on an island is essential to controlling the charge transferred per unit time. This ability is not provided by the double junction configuration as this device is either conducting or non-conducting. Two solutions to this limitation have been developed, the "single electron turnstile"[52,53] and "charge pump"[54,55].

The single electron turnstile, shown in Figure 2.15a, includes an extra tunnel junction in each arm of the SET transistor. If the junction capacitances are equal and the gate capacitance is arranged to be $C_g=C/2$, all junctions have the same critical charge for tunnelling, $Q_c=\pm e/3$. In this system appropriate bias and gate voltages allow stable states to exist where excess electrons are trapped on the central island. Use can be made of the gate voltage to modify the charge distribution in the system, producing a controlled current flow through the device in the direction of the applied bias.



a. Single electron turnstile circuit diagram.

b. Single electron turnstile charge stability diagram, at zero temperature with all junction parameters equal.

Figure 2.15. Circuit and V-V_g diagram for a turnstile.

The V-V_g diagram of figure 2.15 illustrates the presence of locally stable states. The overlap of these states, the hatched areas of figure 2.15b, allow transfer of electrons between the islands without the junction array entering the conducting regime represented by the shaded area. In these stable states, the number of electrons on the island may cycle, for example, from $n=0$ to $n=1$, to $n=0$, ..., simply by controlled application of V and V_g. It is possible to arrange the voltages such that Q_c is exceeded

for the left hand arm but not for the right arm of the turnstile. After one electron has tunnelled through the left arm of the array, it causes a polarization of the gate capacitor and all the junctions in the device. The charge on all the junctions is then less than Q_c and the electron is trapped until the bias configuration is changed. The bias voltage, V, in this system provides an asymmetry in the charge configuration, allowing control over the direction of tunnelling through the device. In this case the bias is allowing electron transport from left to right. The trapped electron can be ejected from the array through the right hand arm by changing the gate voltage such that the critical charge of the right arm is exceeded. Due to the asymmetry caused by the applied bias voltage, this condition will be satisfied first in right hand arm. The electron then tunnels out through the right electrode, transferring one electron through the entire structure.

Application of a high frequency modulation to the gate voltage results in this process being repeated for every cycle of V_g if the amplitude and bias voltage are appropriately fixed. The resulting current is then given by $I=ef$, with f being the ac gate voltage frequency. The effect is observed in the low voltage portion of the $I(V)$-curve as a plateau at the frequency controlled current value until the bias voltage becomes large enough to switch the array into a conducting state.

The charge pump system consists of three series junctions, producing two islands, with each of the islands gated as shown in Figure 2.16.

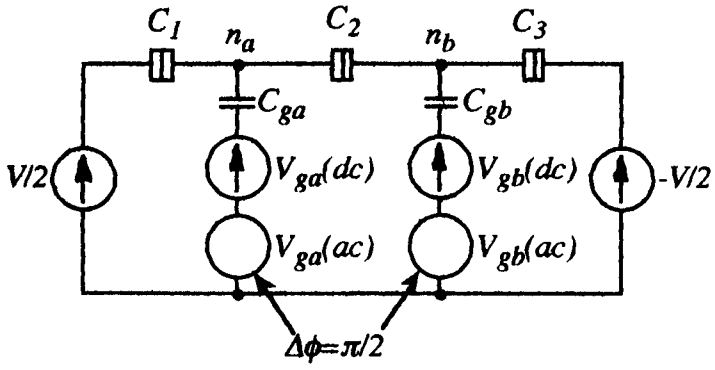


Figure 2.16. Charge pump system.

In this arrangement, the overall system energy relaxes to a minimum for any configuration of V, V_{ga} and V_{gb}. It is therefore possible to control electron tunnelling through any of the three junctions by altering any one of the voltages. If these voltages are properly controlled, it is possible to use the Coulomb blockade of tunnelling to ensure that only one electron tunnels at a time.

An example of a cycle to transfer one electron through this three junction array is to

increase V_{ga} until the voltage across junction 1 exceeds its threshold and an electron tunnels into island a. This tunnel event minimizes the system energy for this voltage configuration and creates a new stable state. Reduction of V_{ga} while at the same time increasing V_{gb} will result in a tunnel event taking place between island a and b when the threshold voltage of junction 2 is exceeded. Again a new minimum energy state is created and the system is stable. Reduction of V_{gb} will eventually result in an electron tunnelling out of island b into the right hand electrode, completing the transfer of one electron through the charge pump. If ac signals are applied to the gates a and b, and the phases are arranged to differ by $\Delta\phi=\pi/2$, a current $I=ef$ can be generated through the device, provided the ac amplitudes are properly controlled. Change of the phase by π results in a reversal of the current. This effect can be observed as a plateau in the $I(V)$-curve as in the turnstile circuit.

Comparing the operation of the turnstile and charge pump, it appears that the charge pump device would provide the most stable device structure. This is due to the energy relaxation effect within this system which allows it to find a stable state for any bias voltage condition. However, some Monte-Carlo simulation work undertaken by Roy[56] has shown that this advantage may be lost through variations in component values. This work has shown that the variations in component values tend to perturb the area of turnstiling in the V-V_g diagram. This perturbation leads to a greater area loss in the charge pump type devices than in the simple turnstile device. This in effect renders the charge pump device more difficult to operate than the turnstile.

2.3.2.5 EXPERIMENTS WITH DOUBLE JUNCTIONS AND TURNSTILES

Due to the simpler operation of the double junction device, a large amount of early experimental work was carried out using such systems. The first reported observations of conductance oscillations in the $I(V)$-curve appeared at the same time. Fabricated three junction systems were used by Fulton and Dolan[57] to investigate single electron charging effects. These systems demonstrated the scaling of E_C with junction area and variation of the Coulomb gap with gate voltage. Kuzmin and Likharev[58] observed conductance oscillations in the $I(V)$-curve through metal particles insulated from and sandwiched between metal electrodes, as shown in figure 2.18 of section 2.3.4. Barner and Ruggiero[59] used a similar technique to observe the Coulomb staircase in asymmetric junctions. Scanning tunnelling microscope probing of granular films is a widely used technique for observing double junction behaviour because of its flexibility. The ability to control the capacitance through small sized particles and the tunnel resistance through tip-particle spacing makes this technique useful for observing effects due to variations in the R_T and C parameters in double junctions[47]. However,

the island cannot easily be gated. These experiments are discussed more fully in section 2.3.4.

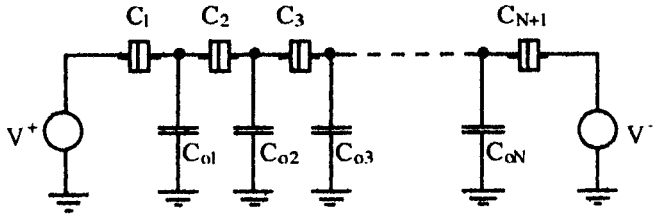
To develop devices operating on single electron effects, fabrication of gated islands is required. A large amount of work on multi-junction circuits has been carried out[42,50,57,60-62] demonstrating the variation of island state with gate voltage[57], the presence of co-tunnelling as a leakage phenomenon[50], the ability to use the double junction as a transistor[52], and multiple junctions as a turnstile[53] and charge pumps[54]. Use of the double junction system as an electrometer was thoroughly explored by Lafarge et al[63]. Here a double junction island was capacitively coupled to an electrometer and the island charge was measured through the effect on the electrometer current. This technique was subsequently used to measure the energy difference between even and odd electron number states in a superconducting island[64]. The charge sensitivity of these electrometers is of the order $10^{-4}eHz^{-1/2}$ at 10Hz. Further use of the electrometer has been made in the development of a precision capacitance bridge[65].

Experiments with turnstiles and charge pump devices have shown a reasonable degree of accuracy in the generated current. However, the limiting factor in most of the experiments has been the accuracy of the measurement system. A large amount of work is on-going in this area in an attempt to improve the accuracy of these systems. To allow these devices to be used as current standards[55,66,67], an accuracy of around 1 part in 10^7 or 10^8 is required.

2.3.3 ONE DIMENSIONAL ARRAYS

Long one dimensional arrays of tunnel junctions, as shown in Figure 2.17, exhibit both time correlations and spatial correlations of the voltage oscillations within the junction array. Time correlations are similar to the single junction SET oscillations. These correlations are a result of the frequency of oscillation being strictly dependent on the value of the array current. Spatial correlations arise from the spacing of charge polarizations within the array. Injection of an electron into the array results in polarization of the array by that electron's charge. This polarization behaves in a similar manner to a solitary wave, or soliton. This soliton can propagate through the array by application of a voltage greater than some threshold, similar to the threshold voltage for single and double junction systems. Applied voltages of less than this threshold result in a non-conducting state in the array at low temperatures. The behaviour of these solitons throughout the array is similar to that of charged particles in that like charged solitons will repel each other while unlike charged solitons will attract

each other. However solitons of opposite charge will neutralize the polarization on contact. Anti-solitons are those caused by the absence of an electron, therefore carrying positive charge. This type of polarization behaves in exactly the same manner as single electron solitons.



a. Circuit diagram of a ID array of tunnel junctions.

b. Array of metal islands, with central island charged

Figure 2.17. Schematic of a long 1D array of tunnel junctions[68].

The array depicted in figure 2.17 is assumed to be homogeneous and, as usual, $T<<T_o$ and $R_T>h/e^2$. Neglecting capacitive coupling to next nearest neighbours and higher order terms, a capacitive matrix can be defined[68,69],

$$C_{ij} = \begin{cases} C_o & , \text{ for } i = j \\ C & , \text{ for } i = j \pm 1 \\ 0 & , \text{ for } i \neq j \, , \, j \pm 1 \end{cases} , \qquad [2.60]$$

where i and j are number indices to identify the junctions in the array. Placing an electron on electrode k in an infinitely long array, a potential φ_k is generated on the island. Using an effective capacitance, $C_{eff} = \sqrt{C_o^2 + 4C_oC}$, this potential is $\varphi_k = -e/C_{eff}$. Generally $C_o<<C$ therefore, $C_{eff} \approx (4C_oC)^{1/2}$, assuming the island is far enough away from the edges of the array. The potential on an electrode i as a function of the distance from the charged electrode k is given by (where k is an integer index),

$$\varphi_i = -\frac{e}{C_{eff}} e^{-\lambda |i-k|} . \qquad [2.61]$$

λ in equation 2.59 is the decay length of the potential which is given by,

$$\lambda = \cosh^{-1}\left(1 + \frac{C_o}{2C}\right) . \qquad [2.62]$$

The potential distribution across the electrodes decays exponentially from the charged

electrode, having a maximum value of $|e/C_{eff}|$ at $i=k$. If the charge is moved from $k \rightarrow k+1$, the entire potential shifts the same way. The energy associated with such a soliton, or anti-soliton is,

$$E_s = \frac{e^2}{2C_{eff}} \quad \text{, with, } \quad C_{eff} = \begin{cases} C_o & \text{for } C_o >> C \\ \sqrt{4C_oC} & \text{for } C_o << C \end{cases}. \qquad [2.63]$$

Applying a bias voltage polarizes the array, to a distance of $M=1/\lambda$ junctions. Solitons within the array interact with this edge polarization and are either attracted or repelled by it. Therefore a one-dimensional array must be very much larger than this polarization length. In other words we must have $N>>1/\lambda$, where N is the number of junctions in the array. Application of a large enough voltage will cause an electron to tunnel into the first electrode in the array, setting up a soliton polarization. The array has a threshold voltage as in the single and double junction case which is given by;

$$V_t = \frac{E}{e}\left(1+e^{-\lambda}\right) = \frac{e}{C_{eff}}\left(1+e^{-\lambda}\right). \qquad [2.64]$$

Biasing symmetrically at $V=V_t$, injects solitons into the array at one end, with anti-solitons at the other end. This leads to annihilation of the polarization at the centre of the array. This effect creates a stationary state in the array and current only flows through the array at a voltage,

$$V_s \geq 2V_t. \qquad [2.65]$$

Shorter arrays cause the bias voltage to tail into the middle of the array, lowering the value for V_S of Equation 2.64.

Due to the repulsion of like solitons these polarizations become equally spaced throughout the array, resulting in spatial correlations of the solitons and the SET events. Movement of a soliton to an adjacent island causes a change in the polarization of the array, affecting all other solitons in the array and causing then to shift through the electrostatic interactions. Thus the tunnelling events also become time correlated, in a similar manner to the SET oscillations in a single junction. Likharev[68] found by numerical simulation that the frequency spectrum of an island clearly displays a peak at $f=I/e$ for currents $I<e/R_TC$. Increasing I above e/R_TC destroys these oscillations through the same processes as in the single junction. Application of high frequency modulation of the current phase locks the oscillations and results in resonances appearing as steps in the $I(V)$-curve at $I=nef$ ($n=\pm1, \pm2, \pm3, \ldots$). At high voltages, however, the $I(V)$-curve reaches the asymptotes

$$I(V) = \frac{1}{NR_T}\left[V - V_{off}(sign.I)\right] \text{ , where , } V_{off} \approx N\frac{e}{2C_{eff}}. \qquad [2.66]$$

Most of the effects relating to transport and charging in one dimensional arrays have been observed, with most of the work being published between 1988 and 1992[42,67,70-72]. The major observation from 1D arrays involves the phase locking of SET events with an external high frequency signal[71], producing the predicted plateaus in the *I(V)*-curve.

2.3.4 TWO DIMENSIONAL ARRAYS AND GRANULAR FILMS

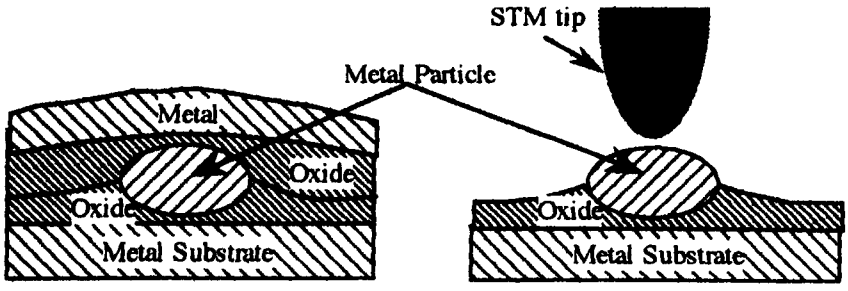
The subject of single electron tunnelling in two dimensional arrays overlaps with the field of conduction in granular films. Charging effects similar to those observed in high sheet resistance granular thin films have been observed in regular 2D arrays of tunnel junctions. The main difference between the two classes of structure come from the characteristics of the individual tunnel barriers in the array. Granular films exhibit a higher degree of disorder and variation than exists with fabricated 2D tunnel junction arrays. The effects of charging energy were first discovered in granular films, where massive resistance increases were observed in the low bias regime, at low temperature. This effect was described in 1951 by C.J. Gorter[24] as the result of the action of a mesoscopic particle's charging energy on the transport process between the particles. A number of experiments on sandwiches of granular films between electrodes and discontinuous films then verified this charging effect on the transport properties of a film[32-35].

In these films, a low voltage conductance behaviour varying as $V^{1/2}$ was observed[32,73]. Although the tunnel rate equations take the same form and the general shape of the *I(V)*-curves were predicted[74] the effect of the charging energy on the correlation of tunnel events was not appreciated. Also, large variations in "array" parameters and the two dimensional nature of the films prevented exploration of effects with controlled numbers of junctions. A review of the field until 1974 is given by Abeles et al[22]. Further developments in weak localization[75] and conduction in disordered systems[76] tended to draw attention away from single electron charging effects. This subject is returned to, in part, in section 5.2.

The discovery of conductance oscillations in the dc *I(V)*-curve came first in a granular film insulated from the contact electrodes[58] and the observation of incremental charging states of metal particles in an oxide between electrodes[59,77]. Both discoveries occurred as a direct result of the presentation of Averin and Likharev's

orthodox theory. Following these experiments, bismuth films deposited under varied oxidation conditions demonstrated that charging effects become dominant when the average barrier resistance exceeded h/e^2[78]. Further, single electron charging effects were observed in insulating wires[79] at temperatures as high as 10K. This illustrated the potential for using granular or discontinuous media to raise the transition temperature. In 1991, Ruggiero and Barner[80] published work on the effect on the capacitance of the average particle size within a discontinuous granular film. A good fit was found between experimental data and the model of a spherical capacitor for particle sizes from 5 to 20nm diameter. At the sub 5nm diameter, they found little difference between the model for an isolated sphere and a spherical capacitor in relation to the experimental data.

These experiments lead to attempts to use the disordered nature of such films to fabricate higher temperature SET devices by the use of granular or polycrystalline films. Use has been made of polycrystalline silicon films[81] and discontinuous platinum films with a liquid crystal coating[82] to observe offset *I(V)*-curve behaviour. However, this work is limited by the irreproducibility of the system's characteristics and the inability to controllably gate the structure, although promising attempts have been made[83].



a. Oxidised film and contact. b. Deposited metal grain for STM observation.

Figure 2.18 Granular film sandwiches and STM probes.

The ability to use a scanning tunnelling microscope (STM) tip to probe a metallic island has led to the highest reported operating temperatures so far recorded[82,84-86], being up to and beyond room temperature. Some of the clearest investigations of the double junction characteristics[47,87] have also utilized STM probing of films. Again, the device operation is limited because of the difficulty in gating such a structure. The small size of the particles, being less than 10nm in diameter, clearly demonstrates the requirements on fabrication technologies for room temperature SET devices.

Conduction in these arrays is described in a similar form to that of 1D arrays in the low

temperature, low bias limit. A soliton like potential distribution exists, polarizing the array in two dimensions. At low temperatures, no free charges exist and a threshold voltage must be overcome to inject charge into the array, and allow conduction to take place. In regular arrays, where junction parameters are better controlled and ordered, a number of physical effects are expected to be observable[88,89], especially in the superconducting state. Some of these have been observed. However, in the case of normal metallic systems, the effects of importance are the transition from a non-conducting to a conducting state at high temperatures ($T>T_o$) and the development of a threshold voltage for the onset of conduction at low temperatures.

2.3.5 DEVELOPMENT OF SINGLE ELECTRON DEVICES

Most of the predicted single electron effects have been observed, resulting in a fairly well characterized theory for device operation. Exploitation of these devices has been promised to yield significant benefits in future microelectronic systems. As the voltage scale for operation is roughly $V_t=e/2C$, the power scale is therefore roughly $P=V_t^2/R_T$ Watts, and the energy scale for operation would be around E_C J/bit. Therefore, reduction in energy requirements and power dissipation along with a massive increase in packing densities for logic and storage elements are among the predictions. In terms of metrological applications, the production of a current standard and the ability to perform supersensitive electrometry are the main areas open to exploitation.

The technological goals for the fabrication of such devices are not simply the reduction of structure dimensions, although this is probably the most pressing requirement if operational temperatures are to be raised. Repeatability in structure sizes and junction R_T and C parameters are equally important. Currently, shadow evaporation techniques, producing around 60x60nm^2 junctions have a yield of 50% for circuits up to 10 junctions. The junction parameters have variations of around 20% in R_T and 10% in C[90].

For current standard applications, the reduction of device dimensions is not critical, as the system can be operated at the very low temperatures required. In this application, the repeatability of device parameters is more important. However, use of devices in electrometry applications requires both higher temperature operation (smaller devices) and well characterized, repeatable device parameters. Finally, in digital electronics, true room temperature operation is essential while variability in the junction parameters is less critical.

The size/temperature problem is compounded by the presence of thermally activated

transitions where $T \neq 0$ but $T < T_O$. The rates of any such transitions must be small to prevent errors in digital logic and storage. A requirement of $\Gamma_T < 10^{-50} s^{-1}$ has been suggested, leading to the requirement that T should be around 100th of T_O for correct error free operation. This requires a T_O of around 3×10^4K for a device suitable for room temperature circuits. Therefore, current technologies are limited to helium dilution refrigerators.

It is encouraging however that structures capable of operating as logic circuits have been demonstrated[91] and that device operational temperatures are rising[92,93].

2.4 REFERENCES

1. C.W.J. Beenakker, *Phys. Rev. B.* **44**(4),1646 (1991).
2. H. van Houten, C.W.J. Beenakker & A.A.M. Staring, in *Single Charge Tunnelling* eds. H. Grabert & M.H. Devoret, Plenum Press, New York, 1992.
3. M.A. Kastner, *Rev. Mod. Phys.* **64**(3),849 (1992).
4. J.H.F. Scott-Thomas, S.B. Field, M.A. Kastner, H.I. Smith & D.A. Antoniadas, *Phys. Rev. Lett.* **62**(5),583 (1989).
5. U. Meirav, M.A. Kastner, M. Heiblum & S.J. Wind, *Phys. Rev. B.* **40**(8),5871 (1989).
6. L.P. Kouwenhoven, N.C. van der Vaart, A.T. Johnson, W. Kool, C.J.P.M. Harmans, J.G. Williamson, A.A.M. Staring & C.T. Foxon, *Z. Phys. B.* **85**(3),367 (1992).
7. D.V. Averin & K.K. Likharev, *J. Low Temp. Phys.* **62**,345 (1985).
8. J.J. Hauser, *Phys. Rev. B.* **3**(5),1611 (1971).
9. J. Frenkel, *Phys. Rev.* **36**,1604 (1930).
10. L. Esaki, *Phys. Rev.* **109**,603 (1958).
11. E.O. Kane, *J. Appl. Phys.* **32**(1),83 (1961).
12. S.M. Sze, Ch. 9,*Physics of semiconductor devices*, 2nd ed. John Wiley & Sons. Inc., Singapore, 1981.
13. J.G. Simmons, *J. Appl. Phys.* **34**(6),1793 (1963).
14. J.G. Simmons, *J. Appl. Phys.* **34**(9),2581 (1963).
15. S.R. Pollack & C.E. Morris, *J.Appl. Phys.* **35**(5),1503 (1964).
16. L.L. Chang & L. Esaki, *Appl. Phys. Lett.* **31**,687 (1977).
17. R.K. Jain, S. Wagner & D.H. Olsen, *Appl. Phys. Lett.* **32**(1),62 (1978).
18. S.L. McCarthy & J. Lambe, *Appl. Phys. Phys.* **33**(10),859 (1978).
19. I. Giaever, *Phys. Rev. Lett.* **5**(4),147 (1960).
20. I. Giaever, *Phys. Rev. Lett.* **5**(10),465 (1960).
21. J. Nicol, *Phys. Rev. Lett.* **5**(10),461 (1960).
22. B. Abeles, in *Applied Solid State Science, Vol. 6.* ed. R. Wolfe, Academic Press, London, 1976.
23. C. Weisbuch & B. Vinter, *Quantum semiconductor structures*, Academic Press, London., 1991
24. C.J. Gorter, *Physica* **17**,777 (1951).
25. B.H. Bransden & C.J. Joachain, *Introduction to quantum mechanics*, Longman Scientific and Technical, Malaysia, 1989, pp84-88, 102-103, 132-135 and pp 388-406.
26. R. Eisberg & R. Resnick, *Quantum physics*, 2nd Ed., John Wiley & Sons. Inc., New York, 1985, p201.
27. M.J. Buckingham, *Noise in electronic devices and systems*, Ellis Horwood Ltd., Chichester,

1983.

28. E. Ben-Jacob, E. Mottola & G. Schön, *Phys. Rev. Lett.* **51**(22),2604 (1983).

29. O.L. Nelson & D.E. Anderson, *J. Appl. Phys.* **37**(1):77 (1966).

30. Z. Hurych, *Solid State Electron.* **9**,967 (1967).

31. J. Pochobradsky, *Solid State Electron.* **10**,973 (1967).

32. C.A. Neugebauer & M.B. Webb, *J. Appl. Phys.* **33**,74 (1962).

33. I. Giaever & H.R. Zeller, *Phys. Rev. Lett.* **20**,1504 (1968).

34. H.R.Zeller & I. Giaever, *Phys. Rev.* **181**(2),789 (1969).

35. J. Lambe & R.C. Jaklevic, *Phys. Rev. Lett.* **22**,1371 (1969).

36. D.V. Averin & K.K. Likharev, in *Quantum Effects in Small Disordered Systems*, eds. B.L. Altshuler, P.A. Lee & R.A. Webb, Elsevier, Amsterdam, 1991.

37. K.K. Likharev, *IBM. J. Res. Develop.* **32**(1),144 (1988).

38. A.N. Cleland, J.M. Schmidt & J. Clarke, *Phys. Rev. B.* **45**(6),2950 (1992).

39. M. Ueda, *Phys. Rev. B.* **42**(5),3087 (1990).

40. G.-L. Inglold & Yu.V. Nazarov, in *Single Charge Tunnelling*, eds. H. Grabert & M.H. Devoret, Plenum Press, New York, 1992.

41. P. Delsing, K.K. Likharev, L.S. Kuzmin & T. Claeson, *Phys. Rev. Lett.* **63**(11),1180 (1989).

42. L.J. Geerligs, V.F. Anderegg, C.A. van der Jeugd, J. Romijn & J.E. Mooij, *Europhys. Lett.* **10**(1),79 (1989).

43. N. Yoshihiro, J. Kinoshita, K. Inagaki, C. Yamanouchi, S. Kobayashi & T. Karasawa, *Jpn. J. Appl. Phys.* **26**(suppl. 3),1379 (1987).

44. N. Yoshihiro, T. Akeyoshi, M. Kojima & S. Sugahara, *Jpn. J. Appl. Phys.* **26**(suppl. 3),1625 (1987).

45. L.S. Kuzmin, Yu. V. Nazarov, D.B. Haviland, P. Delsing & T. Claeson, *Phys. Rev. Lett.* **67**(9),1161 (1991).

46. L.S. Kuzmin & D.B. Haviland, *Phys. Rev. Lett.* **67**(20),2890 (1991).

47. A.E. Hanna & M. Tinkam, *Phys. Rev. B.* **44**(11),5919 (1991).

48. D.V. Averin & A.A. Odintsov, *Phys. Lett A.* **140**:251 (1989).

49. D.V. Averin & A.A. Odintsov, *Sov. Phys. JETP.* **69**(4),766 (1989).

50. L.J. Geerligs, D.V. Averin & J.E. Mooij, *Phys. Rev. Lett.* **65**(24),3037 (1990).

51. D. Esteve, in *Single Charge Tunnelling*, eds. H. Grabert & M.H. Devoret, Plenum Press, New York, 1992.

52. L.J. Geerligs & J.E. Mooij, in *Granular Nanoelectronics*, ed: D.K. Ferry, Plenum Press, New York, 1991.

53. L.J. Geerligs, V.F. Anderegg, P.A.M. Holweg, H. Pothier, D. Esteve, C. Urbina, M.H. Devoret & J.E. Mooij, *Phys. Rev. Lett.* **64**(22),2691 (1990).

54. L.J. Geerligs, S.M. Verbrugh, P. Hadley, J.E. Mooij, H. Pothier, P. Lafarge, C. Urbine, D. Esteve & M.H. Devoret, *Z. Phys. B.* **85**(3),349 (1991).

55. J.M. Martinis, H. Nahum & H. Dalsgaard Jensen, *Physica B.* **194-196**,1045 (1994).

56. S.A. Roy, Ph.D Thesis, No. 10390, University of Glasgow, 1994.

57. T.A. Fulton & G.J. Dolan, *Phys. Rev. Lett.* **59**,109 (1987).

58. L.S Kuzmin & K.K. Likharev, *Jpn. J. Appl. Phys.* **26**(suppl 3),1387 (1987).

59. J.B. Barner & S.T. Ruggiero, *Phys. Rev. Lett.* **59**(7),807 (1987).

60. T.A. Fulton, P.L. Gammel & L.N. Dunkleberger, *Phys. Rev. Lett.* **67**(22),3148 (1991).

61. T.A. Fulton, P.L. Gammel, D.J. Bishop, L.N. Dunkleberger & G.J. Dolan, *Phys. Rev. Lett.* **63**(12),1307 (1989).

62. P. Delsing, T. Claeson, K.K. Likharev & L.S. Kuzmin, *Phys. Rev. B.* **42**(12),7439 (1990).

63. P. Lafarge, H. Pothier, E.R. Williams, D. Esteve, C. Urbina & M.H. Devoret, *Z. Phys. B.* **85**(3),327 (1991).

64. P. Lafarge, P. Joyez, D. Esteve, C. Urbina & M.H. Devoret, *Phys. Rev. Lett.* **70**(7),994 (1993).

65. R.N. Ghosh, E.R. Williams, A.E. Clark & R.J. Soulen, *Physica B.* **194-196**,1007 (1994).

66. S.M. Verbrugh, S.J. Jans, P. de la Court, C.J.P.M. Harmans & J.E. Mooij, *Physica B.* **194-196**,1041 (1994).

67. P. Delsing & T. Claeson, *Physica Scripta.* **T42**,177 (1992).

68. K.K. Likharev, N.S. Bakhvalov, G.S. Kazacha, & S.I. Serdyukova, *IEEE. Trans. Magn.* **25**(2):1436 (1989).

69. M. Amman, E. Ben-Jacob & K. Mullen, *Phys. Lett A.* **142**(6,7):431 (1989).

70. L.S. Kuzmin, P. Delsing, T. Claeson & K.K. Likharev, *Phys. Rev. Lett.* **62**(21),2539 (1989).

71. P. Delsing, K.K. Likharev, L.S. Kuzmin & T. Claeson, *Phys. Rev. Lett.* **63**(17),1861 (1989).

72. P. Delsing, in *Single Charge Tunnelling*, eds., H. Grabert & M.H. Devoret, Plenum Press, New York, 1992.

73. R.C. Dynes & J.P. Garno, *Phys. Rev. Lett.* **46**(2),137 (1981).

74. I.O. Kulik & R.I. Shekhter, *Sov. Phys. JETP.* **41**(2),308 (1994).

75. G. Bergmann, *Physics Reports.* **107**(1),1 (1984).

76. P.A. Lee & T.V. Ramakrishman, *Rev. Mod. Phys.* **57**(1),287 (1985).

77. J.B. Barner, K. Mullen, M.J. Honkanen, S.T. Ruggerio, E. Ben-Jacob & A. Pelton, *IEEE. Trans. Magn.* **25**(2),2542 (1989).

78. J.L. Cohn, E. Ben-Jacob & C. Uher, *Physics Letters A.* **148**(1,2),110 (1990).

79. V. Chandrasekhar, Z. Ovadyahn & R.A. Webb, *Phys. Rev. Lett.* **67**(20),2862 (1991).

80. S.T. Ruggiero & J.B. Barner, *Z. Phys. B.* **85**(3),333 (1991).

81. K. Yano, T. Ishii, T. Hashimoto, T. Kobayashi, F. Murai & K. Seki, in Proc. International

Electron Devices Meeting Washington DC. Dec. 5-8th, 1993.

82. H. Nejoh & M. Aono, *Jpn. J. Appl. Phys.* **32**(1B),532 (1993).

83. H. Nejoh & M. Aono, *Appl. Phys. Lett.* **64**(21),2803 (1994).

84. C. Schonenberger, H. van Houten & H.C. Donkersloot, *Euro. Phys. Lett.* **20**(3),249 (1992).

85. S. Gregory, *Phys. Rev. B.* **44**(23),12868 (1991).

86. J.C. Wan, K.A. McGreer, N. Anand, E. Nowak & A.M. Goldman, *Phys. Rev. B.* **42**(9),5609 (1990).

87. R. Wilkins, E. Ben-Jacob & R.C. Jaklevic, *Phys. Rev. Lett.* **63**(7),801 (1989).

88. J.E. Mooij & G. Schön, in *Single Charge Tunnelling*, eds. H. Grabert & M.H. Devoret. Plenum Press, New York, 1992.

89. C.D. Chen, P. Delsing, D.B. Haviland & T. Claeson, *Physica Scripta.* **T42**,182 (1992).

90. SETTRON Single Electronics, *ESPRIT Basic Research Project 9005*, Periodic Progress Report PPR1, 1 Dec. 1993 to 1 Dec. 1994.

91. K. Nakazato, R.J. Blaikie, J.R.A. Cleaver & H. Ahmed, *Electron. Lett.* **29**(4),384 (1993).
and K. Nakazato, R.J. Blaikie & H. Ahmed, *J. Appl. Phys.* **75**(10),5123 (1994).

92. D. Ali & H. Ahmed, *Appl. Phys. Lett.* **64**(16),2119 (1994).

93. Z. Borsosfoldi, M. Rahman, I.A. Larkin, A.R. Long, J.H. Davies, J.M.R. Weaver, M.C. Holland & J.G. Williamson, *Appl. Phys. Lett.* **66**(26):3666 (1994).

CHAPTER THREE

Fabrication Techniques

3.1 INTRODUCTION

The fabrication techniques used in this work fall into three categories; preparation, lithography and pattern transfer. Preparation includes cleaning, oxidation and deoxidation steps. These process were required before the lithography step to prevent contamination or other defects and damage to the electrical characteristics of the finished sample. Lithography is used to place a pattern into a material which can then provided a means for other processes to effect pattern transfer. This pattern transfer step results in a physical layer or structure being defined which has a pattern geometry related to that used at the lithography stage. These general processes may be repeated in order to build up a number of layers of the device structure. The exact details of the processing steps depend on the structures being fabricated. Figure 3.1 depicts the general flow of the process for various structure types. An important point to note here is that a lithography step is necessarily followed by a pattern transfer step. The lithography step, although essential and central to this work, does not provide the patterns required in itself. The pattern transfer step uses lithography to produce a patterned layer thus imparting functionality to the structure.

Detailed discussion of fabrication techniques is split between this chapter and chapter four. This chapter provides the background and some detail of the main fabrication processes used in this work. Details of the process characterization and optimization for

the production of single electron devices in aluminium metal-insulator-metal tunnel junctions are discussed in chapter four, along with issues relating to the formation of metallic structures at nanometre dimensions. This chapter firstly discusses each of the three generic processing steps; preparation, lithography and metallization. This is followed by discussion of the metallization systems used and the suspended mask fabrication process. After this discussion the sample patterning and process flows are discussed.

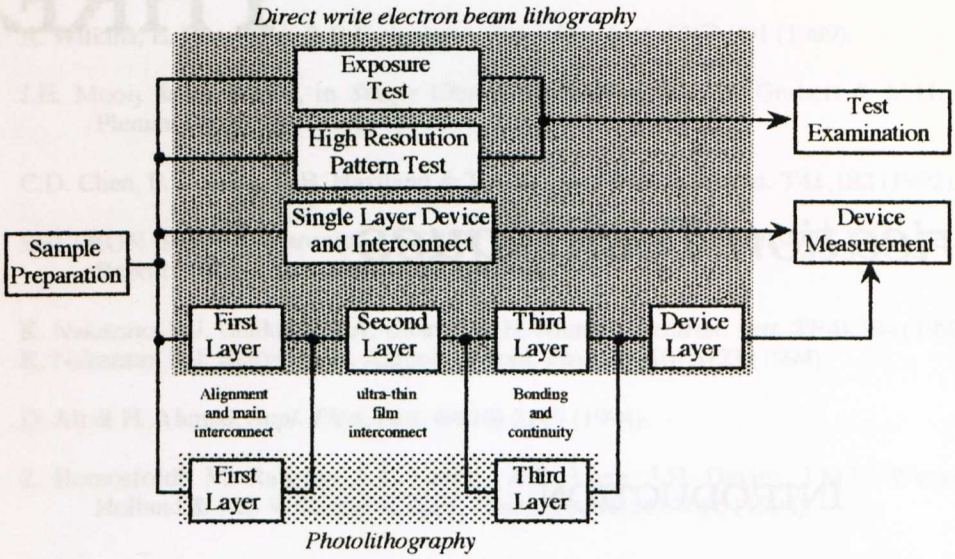


Figure 3.1 Process flow.

3.2 SAMPLE PREPARATION

Samples were processed as whole 3 inch wafers or 24, 12 or 10mm squares of lightly doped (100) 390μm thick p or n-type silicon. Resistivity was 10 to 50Ωcm. Some samples were oxidised to provide insulating substrates. Bulk wafers were used to provide flat substrates for the very thin films being deposited in this work, typically around 25nm thick in device metallizations. Processing large samples yielded several 2 or 5mm square test structures from one sample.

The first stage in sample preparation was cleaning. The aim here was to treat the wafer surface such that dirt, grease or inorganic contaminants were removed. This was required to help prevent failures, such as lack of adhesion, in subsequent processing steps. Cleaning was particularly critical before oxidation of the wafer, to help ensure high quality oxide growth free from defects and pin-holes. Two cleaning steps were carried out. The first clean removed organic contamination while the second clean was designed to remove inorganic contaminants. The cleaning steps and oxidation steps are described below.

3.2.1 ORGANIC CONTAMINATION CLEAN

The first stage of cleaning was a solvent based clean, designed to remove organic contaminants. This involved ultrasonic agitation of the sample in a series of organic solvents. The duration of agitation was 5 minutes in each solvent. After agitation in the solvent, the sample and container were rinsed with the next solvent before being placed in the ultrasonic bath again. The order of solvent cleaning was as follows: trichloroethane or trichloroethylene, methanol, acetone, iso-propyl alcohol (propan-2-ol). After this cleaning sequence the sample was blown dry of iso-propyl alcohol using a filtered nozzle nitrogen gun, fed from the laboratory nitrogen supply.

3.2.2 INORGANIC CONTAMINATION CLEAN

This second stage clean was used to remove inorganic contaminants. This cleaning process could be either a single or two step clean. The single step clean used a 1:4 v/v. solution of sulphuric acid (H_2SO_4) in hydrogen peroxide (H_2O_2). On mixing, this solution heats to around 70°C and evolves oxygen due to H_2O_2 decomposition. For this reason, the solution must be used fresh (ie. within ten minutes). Ultrasonic agitation was used to aid the cleaning process.

The two step clean was as detailed by Kern and Poutinen[1] and involved solutions of either ammonium hydroxide (NH_4OH) or hydrochloric acid (HCl) with H_2O_2 in de-ionized water (H_2O). The first step in this clean utilised the NH_4OH solution and was designed to remove any organic residue remaining on the sample. The second step, using the solution containing HCl removed inorganic contaminants from the sample. Both solutions were used at 70°C for around 10 minutes. Ultrasonic agitation was applied. The ratios of these solutions were not critical, and were typically used at 1:1:5 v/v. NH_4OH:H_2O_2:H_2O and 1:1:6 v/v. HCl:H_2O_2:H_2O. This two stage clean was used where samples were to be oxidized, helping ensure uniform void free oxide growth.

3.2.3 OXIDATION

Cleaned samples were oxidised in the department's wet oxidation furnaces. 24x24mm^2 samples were placed in the furnace and heated to around 1200°C. Wet oxidation of the samples took place using nitrogen as the carrier gas. After two hours exposure to these conditions an oxide of around 0.6μm thick was achieved, which was more than adequate for electrical isolation.

3.2.4 DEOXIDATION OR OXIDE ETCH

Silicon dioxide was etched using hydroflouric (HF) acid solutions. The solution used here was a buffered solution of 1:4 v/v of HF:H_2O. 1:4 HF:H_2O solution etches silicon dioxide (SiO_2) at roughly 0.1μm.min^{-1}, allowing estimation of the etch depth. Where deoxidation was required, the sample was left in the solution until the surface de-wet in H_2O. HF:H_2O solutions do not etch silicon[2] and a bare silicon surface is completely hydrophobic whereas the SiO_2 surface is hydrophilic. Therefore, the de-wet point was a clear indication of the removal of the surface oxide from a sample.

3.3 LITHOGRAPHY

Lithography is a general term describing patterning of a material. In this work, the process involved the formation of a mask on a substrate to allow pattern transfer through either etching or metal deposition. Removal of the mask left a substrate with a pattern which defined a layer of the final structure. Two lithography techniques were employed here, electron beam and photo-lithography. Electron beam lithography was the most heavily used of the two processes, being employed for both low and high resolution pattern transfer. Photolithography, however, was used only for low resolution pattern transfer, that is for feature sizes of greater than 2μm. The patterns employed were designed using CAD systems and the design process is described in section 3.3.1.2 while the patterns are detailed in section 3.7.

3.3.1 ELECTRON BEAM LITHOGRAPHY

To define structures with dimensions less than 0.1μm, electron beam lithography was the only method immediately available. Indeed, electron beam lithography was used in this work where structures below 2μm in size were required and in the production of the masks for the photo-lithography process. Electron beam lithography is capable of defining structures of around 10nm[3] or less[4], dependent on the system used. It is the only method currently capable of reducing device dimensions to the sizes required for high temperature single electron device operation. Electron beam lithography is, however, a relatively slow process. The need to scan the electron beam during the exposure process produces write times proportional to the pattern area. The need for small beam currents and high electron doses for small features further compounds this problem. The end result is very long write times where large pattern areas of high resolution features are required.

3.3.1.1 ELECTRON BEAM LITHOGRAPHY SYSTEM

The electron beam lithography system used exclusively during this work was a Leica-Cambridge EBPG5-HR electron Beamwriter. This Beamwriter employed a vector scan writing system with beam blanking. Beam currents between ≈100pA to ≈200nA were possible, producing spot sizes between 12 to 400nm at 20, 50 and 100keV accelerating voltages. Manual selection of the final aperture size (with 200, 400 or 800μm settings) provided fine control of small spot diameters or large currents for large spots. Resolution and beam step size were controlled via 15bit x and y DAC's, allowing 32000 pixels in both x and y. The resolution step could be continuously varied from 5 to 312.5nm, giving a minimum single writing frame of 160x160μm^2 with a maximum frame size of 800x800μm^2. Writing frequencies could be varied from 0.5kHz to 10M Hz.

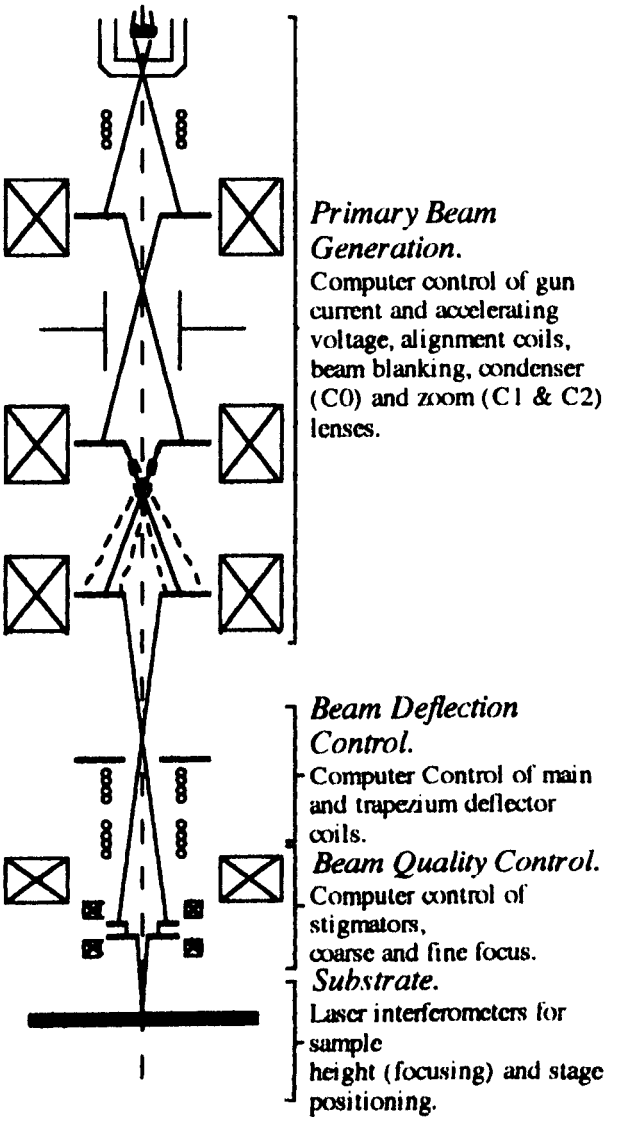


Figure 3.2
Beamwriter System Schematic.

Where butting of writing frames was required to cover larger area patterns, a process called stitching was used. Stage positioning was measured by laser interferometry to a resolution of 5nm (λ/120). This position was servo-looped to two systems, the stage positioning motors and the frame "pull-in" system. Frame pull-in adjusted the main beam deflection system to compensate for stage positioning errors to a maximum of ±20μm. Using these systems, average stitching errors of around 40nm with 800x800μm^2 frames were achievable. These systems allowed the Beamwriter to perform high resolution direct writing on relatively large area substrates (127x127mm^2). Alignment to previous layers was achieved to an average error of 50nm. Focus was dynamically adjusted using a laser height meter to measure the

substrate height and adjust beam focus.

3.3.1.2 DATA PROCESSING FOR ELECTRON BEAM LITHOGRAPHY

Pattern files were created with the aid of a standard computer aided design package. In this case MICAD (produced by EESOF) was used. The pattern files were created from drawings produced using this package, with each pattern in the drawing being formed from a series of rectangles and polygons. When the pattern was designed, two further data processing steps were required before the patterns were in a form usable for electron beam writing. Transcription Enterprise's Computer Aided Transcription System (CATS) was used to produce fractured pattern files for use with the Beamwriter. The data files produced by the CAD system required to be translated to allow CATS to understand the data. Translation produced a CATS library file which was then be used, by the CATS program, to produce a properly fractionated pattern file. Fractionation is the process of dividing a pattern into rectangles and trapezia which can be written by the Beamwriter control system. This process took into account frame sizes to allow stitching. A final step of adding relative exposure data, if required, and then producing the Beamwriter workfile completed the processing of the pattern data. This workfile contained all the pattern and exposure data required to write the pattern onto the sample.

Control of pattern writing was exercised through the Beamwriter control software, produced by Leica-Cambridge. However, to simplify the process, a program was written in house by Dr. S. Thoms to allow control of the Beamwriter layout process. The program, named BWL, allowed simple input of information related to job type, exposure dose, beam accelerating voltage, substrate positioning, registration data, repeat step size data and number of repeats. BWL took this information and translated it into control commands for the Beamwriter control system. Together, the control data and pattern data formed the complete information for a job.

3.3.1.3 POSITIVE ELECTRON BEAM RESISTS

Poly-methylmethacrylate (PMMA) was used in this work as a positive electron beam resist in two forms, one high sensitivity and the other lower sensitivity. This difference is derived from the polymer's molecular weight. Polymeric molecular weight can be defined by a number average, $\overline{M}_n$, and a weight average, $\overline{M}_w$, molecular weight, with a dispersivity index defined as $H=\overline{M}_w/\overline{M}_n$. Figure 3.3 illustrates the difference between the number and weight average molecular weights. A distribution of molecular weights results from the random nature of termination in addition polymerization. The number average represents the most probable molecular weight while the weight average

molecular weight is derived from the average molecular size[5 & 6].

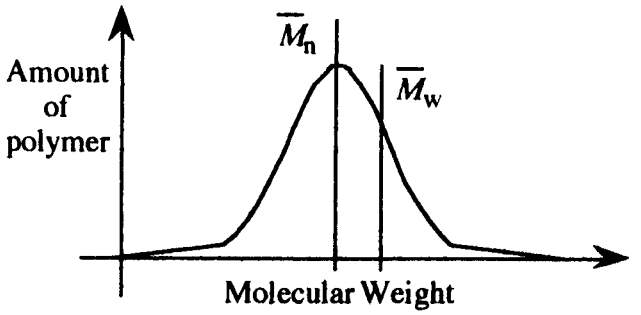


Figure 3.3 Polymer molecular weight distribution.

The high sensitivity resist is termed BDH, (BDH Chemicals Ltd.) while the lower sensitivity resist is Elvacite-2041 (DuPont). These molecular weights and distribution have been previously characterised[7] using gel permeation chromatography and are presented in table 3.1. Henceforth, these polymers will be referred to as "BDH" and "Elvacite" or "Elv." respectively.

	BDH.	Elvacite 2041
$\overline{M}_w$ / gmol^{-1}	85000	361000
$\overline{M}_n$ / gmol^{-1}	51000	187000
H	1.65	1.93

Table 3.1 Molecular weight characteristics of BDH and Elv. PMMA[7].

Generally, the exposure sensitivity of a positive resist polymer is determined by its chain length and therefore its average molecular weight. The exposure process results in breakdown of the main polymer chain, reducing its molecular weight. This affects the dissolution rate of the polymer in a given solvent. A critical molecular weight, M_c, exists for a given polymer/solvent combination[8]. The greater the molecular weight, the greater the number of chain scissions required to reach M_c and therefore the lower the sensitivity to electron beam exposure. The distribution of molecular weights has been shown to have some bearing on the resist contrast[9], with narrower distributions yielding higher exposure contrasts. Resist contrast is defined from the dose sensitivity to a given developer and can be measured by the ratio of the dose to clear the resist from an exposed area to the dose to begin resist develop in that area. Figure 3.4 illustrates this definition, while equation 3.1 defines the contrast. Table 3.2 details some of the exposure characteristics for the electron beam resists used here, but it is important to note that the exact dose and contrast numbers depend on the beam accelerating voltage, beam shape and developer conditions.

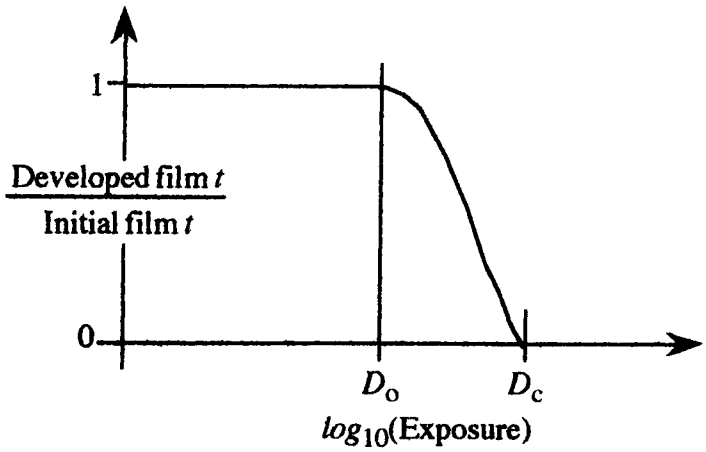


Figure 3.4 Resist contrast determination.

$$\gamma = \left[log_{10}\left(\frac{D_c}{D_o}\right) \right]^{-1} \qquad [3.1]$$

	BDH.	Elvacite 2041
D_o / μCcm^{-2}	100 - 120	150 - 200
D_c / μCcm^2	220	280
γ	2.9 - 4.1	3.7 - 6.8

Typical exposure conditions[7].
50keV exposure. 45s at 23°C in 3:1 developer.

Table 3.2. Exposure characteristics at 50keV for BDH and Elv. resists.

To allow spinning of these polymers onto samples, the polymer beads (as supplied) were dissolved into a casting solvent as a weight percentage. Spinning allowed creation of a smooth polymer/solvent layer which was then baked to drive off the solvent, leaving behind the polymer as the electron beam resist. The baking temperature was above the glass transition temperature, T_g=120°C, and below the melting point, T_m≈225°C, allowing the polymer to flow and the solvent to easily evaporate[9]. The solvent used was typically ortho-xylene but for higher weight percentage resists chlorobenzene was used. The greater solubility of PMMA in chlorobenzene was useful for these high weight percentage solutions and also produced a lower viscosity solution than an ortho-xylene counterpart. This lower viscosity produced a more uniform film which was easier to dispense and spin. Higher weight percentages and/or molecular weights produce more viscous solutions and therefore thicker resist films. Typical resists are tabulated in table 3.3 along with the casting solvent used and the thickness at 5krpm spin speed with a bake of 180°C. The resist thicknesses for these conditions were characterised as a matter of routine.

Resist	Casting Solvent	Thickness / nm
2.5% BDH	ortho-xylene	35±5
4% BDH	ortho-xylene	72±5
8% BDH	ortho-xylene	208±10
15% BDH	chlorobenzene	1250±100
2.5% Elv.	ortho-xylene	40±5
4% Elv.	ortho-xylene	100±10

Table 3.3. Table of common resist thicknesses, spun at 5krpm for 60s.

Resist	Spin Speed /rpm	Thickness / nm
2.5% BDH	3000	50±5
2.5% BDH	5000	35±5
2.5% Elv.	5000	40±5
2.5% Elv.	8000	28±5
2.5%BDH/2.5%Elv.*	5000/5000	70±5
2.5%BDH/2.5%Elv.*	3000/8000	75±5

*-layer 1 spun at quoted speed for 60s, baked for 1 hour at 180°C, layer two spun at quoted speed for 60s and baked for ≥3 hours at 180°C.

Table 3.4. Table of resist thicknesses for high resolution resists cast in ortho-xylene.

3.3.1.4 BILAYER RESIST SYSTEMS

Bilayer resists are necessary for lift-off processing, as described in section 3.4. In this type of pattern transfer process, it is necessary to develop an undercut profile in the resist to guarantee a discontinuity in the deposited material. This discontinuity allows the lift-off solvent (acetone in the case of PMMA) free access to dissolve the resist mask. Dissolving this mask allows the unwanted deposit to float away in the solvent leaving the defined pattern behind. Use of two different sensitivity resists allowed development of such a profile. Therefore spinning a bilayer resist, with the higher sensitivity BDH resist as the first layer, with the lower sensitivity, high molecular weight Elvacite as the top layer provided this undercut structure. When written, pattern size was determined by the top layer resist while the bottom layer provided support and an undercut in the resist cross-section, as shown in figure 3.5.

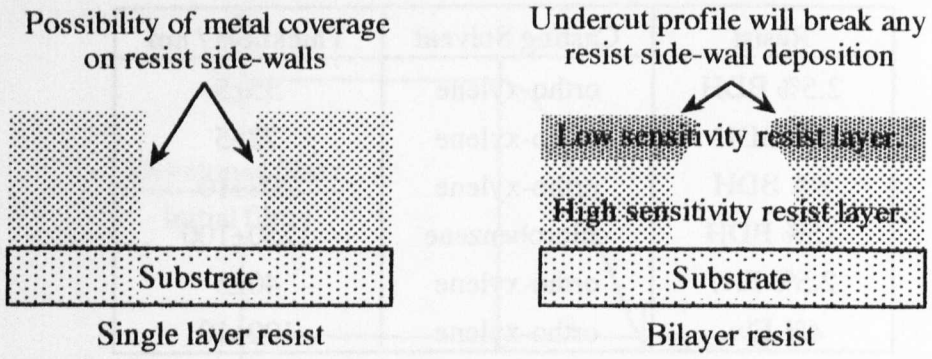


Figure 3.5 Schematic exposed and developed profiles for single layer and bilayer resist systems.

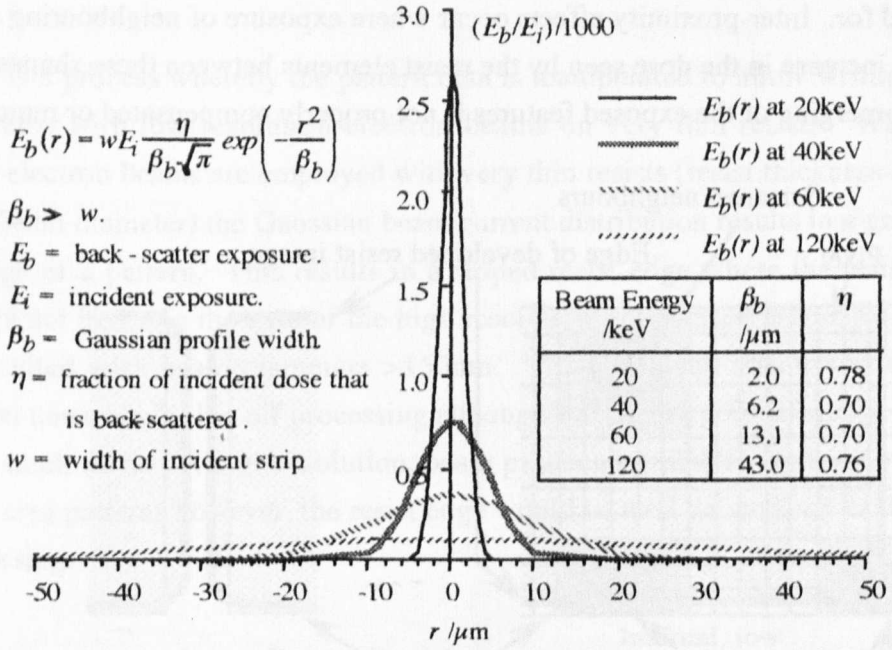
Bilayer resist systems have also been used to create a suspended mask structure by using a very high sensitivity bottom layer[12]. The use of an extremely high sensitivity bottom layer with a mechanically stable top layer produces an extreme version of the undercut profile shown in figure 3.5. This type of structure can be used for shadow mask processing with angled evaporations in the same way as the suspended mask process used in this work. Section 3.6 details the processing steps for this germanium suspended mask process used here. This process contained five distinct processing steps; support deposition, mask deposition, lithography, mask etch and support etch. A bilayer resist process therefore greatly simplifies the traditional suspended mask process.

3.3.1.5 ELECTRON BEAM RESIST EXPOSURE

For an electron beam resist to be useful, exposure to the electron beam must cause a localized chemical change which enhances the materials solubility in a given solvent, the developer. In the case of PMMA, the main mechanism is thought to be main chain scission[13], giving rise to reduced molecular weight in the exposed region and therefore increased solubility in a developer solution[14].

Four sources of electron exposure exist in electron beam lithography. These are primary, forward-scattered, back-scattered and secondary electron exposures. Primary electron exposure would result from an incident high energy electron interacting directly with nuclei in the main polymer chain to produce chain scission. However, it is thought that the interaction cross-section of these high energy primary electrons is too small for this to be the main exposure mechanism. Exposure from forward-scattered electrons takes place by incident electrons suffering elastic collisions with nuclei in the resist. These electrons are deflected from their incident trajectory, leading to exposure

by an apparently bigger primary electron beam. Back-scattered electrons have suffered elastic collisions in the substrate and found their way back to the surface to exit through the resist. This leads to resist exposure over a large region within the back-scatter radius. The back-scatter radius is proportional to the electron beam accelerating voltage[15], as illustrated in figure 3.6. Both forward and back scattering tend to be the main culprits in electron beam proximity effects as discussed in section 3.3.1.6.



Beam Energy /keV	β_b /μm	η
20	2.0	0.78
40	6.2	0.70
60	13.1	0.70
120	43.0	0.76

Figure 3.6 Diagram illustrating back-scatter radius dependence on kV.

Secondary electron exposure is thought to limit the ultimate resolution of electron beam resists[3, 16-18]. Secondary electrons are formed when primary electrons suffer inelastic collisions in the resist or substrate. The energy and range of secondary electrons is thought to be small. Experiments with UV exposure of PMMA[19] show that energies greater than ≈5 eV can cause main chain scission. Also, exposure simulations have shown that a secondary electron range of a few nanometres provides good agreement with experimental results[16]. The overall effect is an apparent lithographic limitation in PMMA of between 6 and 10 nm[3,4,20-24].

3.3.1.6 PROXIMITY EFFECTS

Proximity effects arise from the exposure conditions within neighbouring pixels in a pattern. These effects result in differences between the designed and exposed pattern shapes. Exposure by scattered electrons produces two types of proximity effect; intra-proximity and inter-proximity, see figure 3.7. Intra-proximity effects produce changes within the pattern area and can occur in isolated shapes, while inter-proximity effects

occur between neighbouring pattern areas. The two situations are represented in figure 3.7. These shape changes result from variations in the total electron dose seen by an element of resist, dependent on the surrounding exposure conditions and where the resist element lies within a pattern. For example in a square, intra-proximity effects arise from the fact that an element or pixel at the edge of a shape receives a reduced dose compared with a central pixel, while corner pixels receive a greatly reduced dose. This can lead to rounding of corners in the developed pattern if not properly compensated for. Inter-proximity effects occur where exposure of neighbouring shapes results in an increase in the dose seen by the resist elements between these shapes. This can result in merging of the exposed features if not properly compensated or removed.

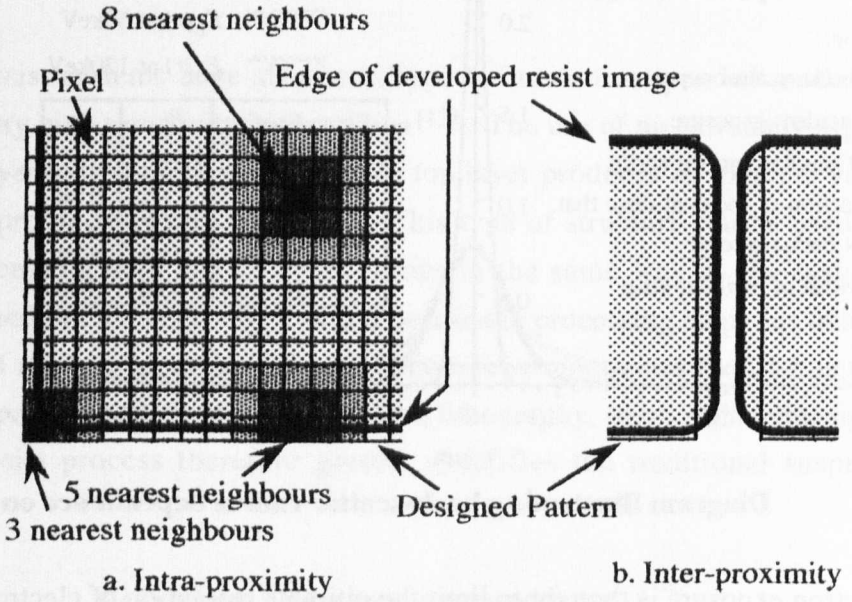


Figure 3.7 Intra and inter proximity effects in electron beam exposure.

Two forms of scattering combine to produce proximity effects, forward scattering and back scattering. Various algorithmic or computer aided modelling techniques can be applied to designs to help compensate for these proximity effects[25], however none were used within this work. Proximity exposure compensation was achieved here through the use of exposure tests. This method ensured that the defined pattern was correct before it was used to form a device structure. Scattering effects can, however, be further reduced through the choice of processes used. The effects of forward scattering can be almost eliminated by using thin (<0.1μm thick) resists. The effects of back scattering can be reduced through three main techniques, the first two of which were used in this work. Firstly, the use of high contrast resist/develop processes produces greater selectivity between the main exposure and the back scattered exposure. This subsequently requires a higher back scatter dose to affect the developed feature shape and resist profile. Secondly, the use of higher electron accelerating voltages

results in a larger back scatter radius and therefore a lower back scatter dose per unit area, as shown in figure 3.6. A third method of back scatter reduction is to use membranes instead of solid substrates[26]. In this case, the primary electrons exit through the bottom face of the sample before they can be scattered back, thus eliminating back scatter exposure.

3.3.1.7 SLEEVING

Sleeving is a process whereby the pattern data is manipulated to allow writing of large area patterns with low resolution electron beams on very thin resists. Where large diameter electron beams are employed with very thin resists (resist thickness much less than the beam diameter) the Gaussian beam current distribution results in a graded dose at the edge of a pattern. This results in a sloped resist edge where the resist process contrast is not high. In the case of the high resolution resists used in this work, sloped edges resulted with beam diameters >150nm. The end result was a written pattern which was unusable for lift-off processing although the pattern was defined in the resist. Use of a small diameter, high resolution, beam produced prohibitively long write times for large area patterns however, the resist edge would be well defined and usable for lift off processing.

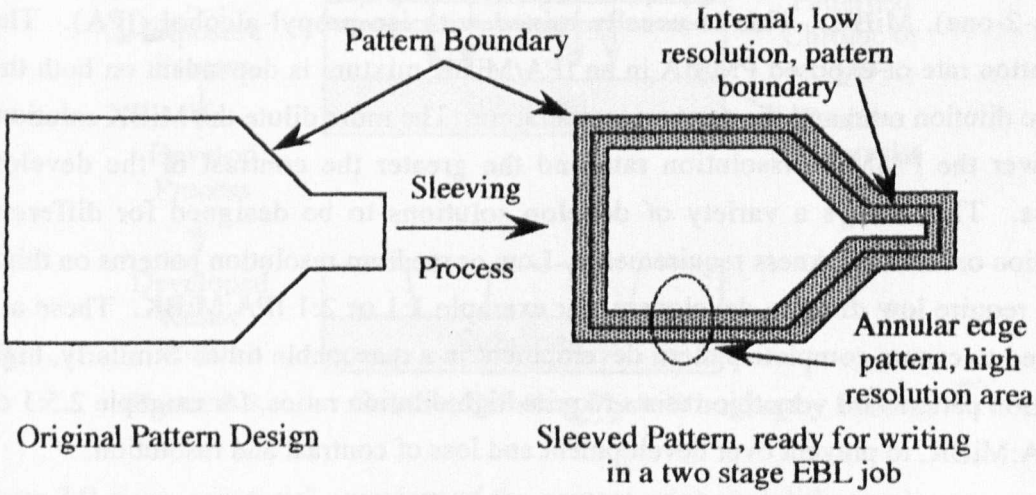


Figure 3.8 Schematic representation of sleeving.

To overcome the problem created by the need for large areas on very thin resists, the sleeving process is employed. The end result of the sleeving process is two pattern files which may be used together to define the pattern. The first pattern, the low resolution pattern, is used to define the bulk of the writing area and may use a large diameter, low resolution, high current electron beam. This provided a short write time for large areas. The second pattern defines the pattern resist edge. A high resolution small diameter

electron beam is used here, to preserve the resist edge profile. This pattern contains a much smaller write area than the original pattern and therefore eliminates excessively long write times.

The patterns used in this work were sleeved, when required, using the CATS software. To produce the low resolution pattern area, the designed pattern was simply sized down by a uniform dimension. For example 0.5μm. This pattern was then ready to use with the sample. To produce the high resolution edge pattern, the original pattern was again sized down (as with the low resolution pattern) but using a larger factor, say 1.0μm. This second reduced pattern was then subtracted from the original to produce an annular edge pattern. For the figures used here, a 1.0μm annulus would exist, similar to that shown in figure 3.8. The overlap between the low and high resolution patterns would be 0.5μm all round in this case. Therefore, in this example, a 0.5μm wide strip, 0.5μm from the pattern edge would exist which would be double exposed. However, this was not a problem as the doses used in large area patterns were relatively low.

3.3.1.8 ELECTRON BEAM RESIST DEVELOP

Development of electron beam written patterns defined in PMMA has been studied by Greeneich[14]. The most common developer used is methyl-iso-butyl-ketone (methyl-pentan-2-one), MiBK. This is usually mixed with iso-propyl-alcohol (IPA). The dissolution rate of exposed PMMA in an IPA/MiBK mixture is dependent on both the mixture dilution ratio and the mixture temperature. The more dilute the MiBK solution, the lower the PMMA dissolution rate and the greater the contrast of the develop process. This allows a variety of develop solutions to be designed for different resolution or resist thickness requirements. Low or medium resolution patterns on thick resists require low dilution developers, for example 1:1 or 2:1 IPA:MiBK. These are designed to ensure complete pattern development in a reasonable time. Similarly, high resolution patterns on very thin resists require high dilution ratios, for example 2.5:1 or 3:1 IPA:MiBK, to prevent over development and loss of contrast and resolution.

The PMMA dissolution rate increases logarithmically with temperature, but exhibits an activation temperature[14]. In order to control this temperature dependency a fixed temperature develop was used. These developers have been characterized in order to determine the appropriate develop times[7,16 & 27]. A standard develop time was then chosen for the resist and developer combination, typically around 30 to 45 seconds at a fixed temperature of 23°C. This time period was long enough to repeatably measure but short enough to limit the change in the developer temperature. The different developers were designed for use on specific resists to allow standardization of process conditions.

Exposure doses and resist thicknesses were then selected to give appropriate pattern development for these developer conditions.

3.3.2 PHOTOLITHOGRAPHY

While the devices fabricated in this work required very high resolution lithography, not all the layers needed to produce a measurable structure had the same dimensional requirements. It is possible to use the Beamwriter to write large area low resolution patterns however the time taken to process the resist system and to write the job can be prohibitive. Where pattern dimensions are not less than 2μm, photolithography can easily be used. This has the advantage of simple, high throughput processing, capable of defining structures at 5 to 2μm. The process resolution, in this case, is limited through the use of a contact printing system and layer registration is limited by this system to ± 1 to $\pm 2\mu$m. The following sections detail the photolithography process used while exact details of the process flow used are left to section 3.8.

3.3.2.1 CONTACT PRINT PHOTOLITHOGRAPHY PROCESS

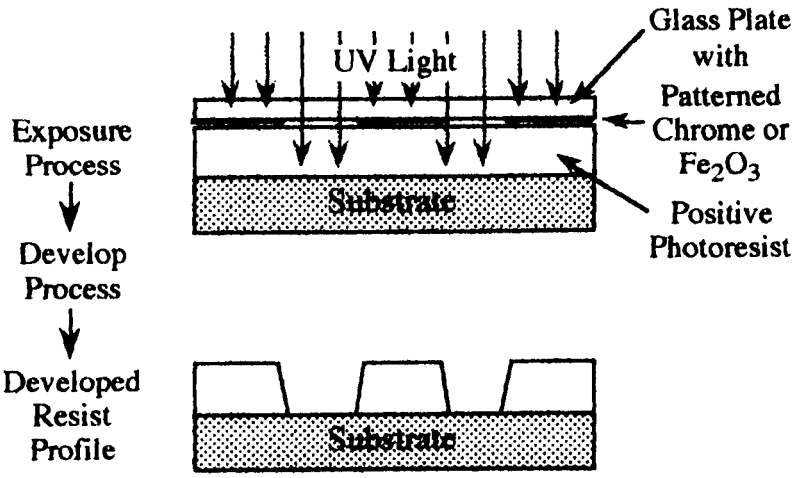


Figure 3.8 Contact Printing Photolithography Process

Figure 3.8 gives a pictorial overview of the contact print photolithography process. In this process the masking pattern is brought into contact with the resist surface. The resist itself is spun onto the sample in the same manner as the electron beam resists. Exposure of the resist to UV light through the mask results in chemical changes within the photoresist which make the exposed regions more soluble in the developer solution than the unexposed regions. When the exposed resist is developed, the pattern is transferred into the photoresist and the substrate plus resist is ready for further processing.

3.3.2.2 MASK GENERATION

Masks used in the photolithography process were fabricated from three inch square glass plates coated with either ferric oxide (Fe_2O_3) or chrome. Ferric oxide is partially transparent and can be easily used to align to previous layers. However, these masks cannot be used to define the smallest features possible with this process and are limited to minimum features of $\approx 3\mu m$. Chrome masks provide the photoresist with a higher contrast light image than ferric oxide masks and therefore can be used to print ≈ 1 to $2\mu m$ features. However, chrome masks cannot easily be used where alignment is required.

Both ferric oxide and chrome masks could be patterned by either electron beam lithography or by photolithography using a master mask plate. The resist pattern was then transferred to the mask material through wet chemical etching. Mask production by electron beam lithography used the Beamwriter and the same pattern generation process as described in section 3.3.1.2. Where photolithography was used to produce a mask plate the process was simply a contact print from a master plate to the working plate using a contact printer. The master plate was a chrome mask patterned by direct write electron beam lithography in the Beamwriter. The minimum feature sizes on the master plates were therefore determined by the mask write and etch process. Where a working plate was written using the Beamwriter, care had to be taken to flip the pattern through the vertical axis. This is because the contact print process flips the image though the vertical axis during the pattern transfer from the mask front face to the sample front face.

3.3.2.3 PHOTORESISTS

The photoresists used in this fabrication work were positive tone photoresists supplied by Shipley Europe Ltd. Either of S1400/31 or S1818 Microposit resists were used as both have nominally the same characteristics. Both resists respond to ultraviolet light of 350 to 420nm. The resist was spun onto clean dry samples using a spin speed of 7000rpm for 60s and baked for 30min. at 90°C to drive off the casting solvent. The resulting resist thickness was around $1\mu m$. Resists were applied using a syringe fitted with a Millipore "Fluoropore" $0.45\mu m$ filter.

3.3.2.4 RESIST TREATMENT FOR LIFT-OFF PROCESSING

To provide an undercut resist profile to aid lift-off processing, the spun and baked photoresist was soaked in chlorobenzene. Chlorobenzene (C_6H_5Cl) soaking reduces the

sensitivity of the resist surface to the developer, resulting in an undercut resist profile after development of the exposed image. Samples were soaked in chlorobenzene for 15min. and then baked for a further 15min. to drive off the solvent. This process increased the required exposure time for the resist as shown in table 3.4.

3.3.2.5 EXPOSURE

Two photolithography systems were used. Where no pattern alignment was required, a simple contact printer was used. This equipment consisted of a support for the mask, a manually operated shutter and a mercury vapour lamp. The mask was placed with the patterned side facing up and samples were placed face down on the mask and weighted down to ensure good contact. The samples were then exposed to the UV light for the required time, dependant on the contact printer and resist used.

The second system was an Hybrid Technology Group System 3 Mask Alignment and Exposure System. With this system, the mask was mounted onto a mask stage and the sample onto an x, y, z and theta stage. Both the mask and the sample were held in place by vacuum. When both the sample and the mask were in place the mask was brought into close proximity with the sample by the mask stage and use of the z motion of the sample stage. The sample was then aligned using the microscope and the x, y and theta controls on the sample stage. It is for this reason that the masks used for pattern alignment must be partially transparent, ie. ferric oxide. When the sample and mask were aligned, the sample was brought into hard contact with the mask and exposed. Exposure conditions are given in table 3.5. This system was capable of alignment to around ± 1 to 2μm accuracy.

Resist.	Mask Aligner. Exposure time /s.
S1400/31.	10
S1400/31 with C_6H_5Cl soak.	12
S1818.	11
S1818 with C_6H_5Cl soak.	13

Table 3.5 Mask aligner exposure data for photoresist processes.

3.3.2.6 DEVELOPMENT

Pattern development, for both resists, used Shipley S1400/31 developer, diluted at 1:1 developer:H_2O. The sample was immersed in developer for around 75s, or until the

pattern cleared, with some sight agitation applied to the sample to ensure the dissolved resist was removed from the sample surface. When complete, the sample was rinsed in de-ionized water and dried by blow drying with an N_2 supplied air-gun. Care was taken not to develop the sample for more than 150s as the developer contained either potassium or sodium hydroxide, both of which etch silicon.

3.3.2.7 POST BAKE

Samples requiring treatment in silicon dioxide etch (1:4 HF:H_2O, buffered) were given a post development bake of 30min at 120°C. This improved the resistance of the photo-resist to hydrofluoric acid and improved the resist adhesion to the substrate.

3.4 METALLIZATION AND LIFT-OFF

After the resist has been patterned, that pattern must be transferred into some material to complete a useful patterning stage in the processing of a structure. This pattern transfer can be achieved either through etching or deposition. All but one of the patterned layers in this work used metal to form the pattern and the lift-off process to transfer the pattern from the resist to the metal. The exception was the germanium masking layer formed in the suspended mask process, as discussed in section 3.6. All metallizations were carried out via vacuum evaporation of the metal onto the substrate. This particular physical vapour deposition technique provides a very directional deposition flux which works well with the lift-off process. The following three sections detail, in order, the lift off process and the metallization process used for the interconnect structures and the devices. The metallization systems used are detailed in section 3.5.

3.4.1 LIFT-OFF PROCESSING

Lift-off is the process of resist mask removal after the metal has been deposited. For ease of lift-off processing, an overhung or undercut resist profile is desirable. This is the main reason for using bilayer electron beam resists and chlorobenzene soaking of photoresists as described in sections 3.3.1.4 and 3.3.2.4. The process is illustrated in figure 3.9.

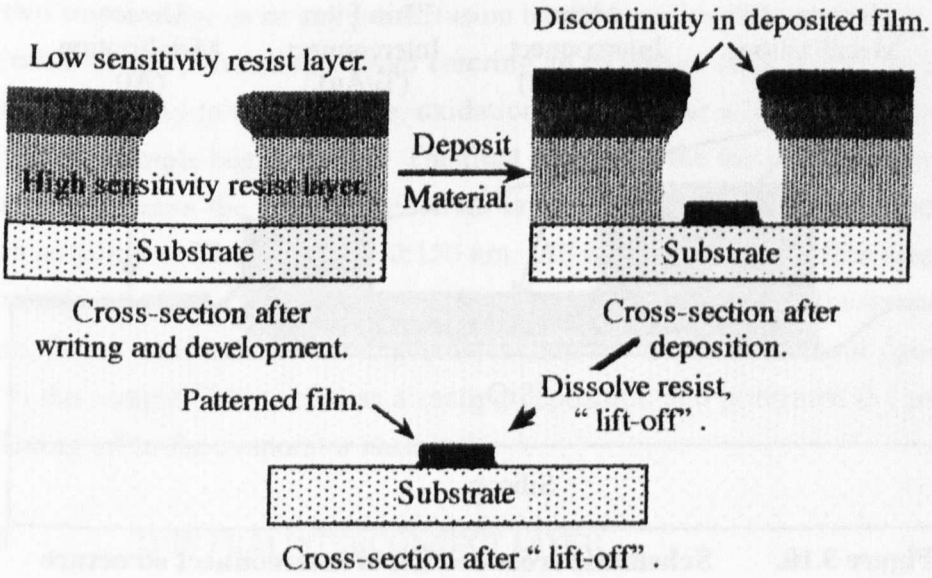


Figure 3.9. Lift-off processing.

In this process, metal is deposited after the resist pattern has been developed. Removal of the resist mask also releases and removes the unwanted metal pattern, thus transferring the resist pattern to the deposited metal layer. To remove the resist, the sample is immersed in a solvent which dissolves the resist polymer. For the resists used in this work, S1400/31 or S1818 and PMMA electron beam resists, soaking in acetone readily dissolved the resist and allowed removal of the unwanted material. Where difficulty was experienced with resist removal, the acetone solution was heated (to not more than 40°C) and a jet of acetone was applied to the sample using a syringe. In extreme cases, ultrasonic agitation could be applied to aid removal of the resist. Ultrasonic agitation was not used with aluminium metallizations as this damaged the patterns. In all cases care was taken to ensure the removed metal did not settle back on the sample surface as it is possible for such material to become permanently adhered to the sample.

3.4.2 INTERCONNECT METALLIZATION

The interconnect metallization system is a three stage metallization process that forms a complete interconnect structure. The interconnect structure provides contact to the device layer thin film metallization, an interconnect to communicate signals to large bondpad areas and bondpads, for gold wire bonding. A cross section of the structure is shown in figure 3.10 and each of the layers are discussed in the following three sections. The device layer metallization is discussed in section 3.4.3.

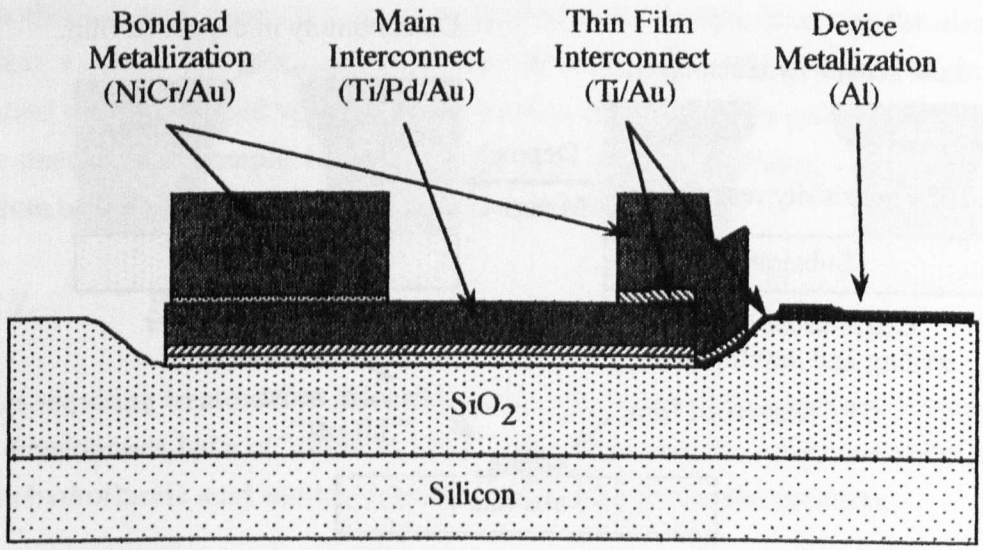


Figure 3.10. Schematic cross-section of interconnect structure

The interconnect structure was designed to be fabricated prior to the device metallization. This was to avoid either damage to or ageing of the device metallization through processing of the interconnect structure. As electron beam resists of thickness around 100nm were used for patterning of the device layer metallization, some of the silicon dioxide is removed in the region of the main interconnect. This action partially planarized the sample surface, allowing better resist coverage of the sample. SiO_2 etch was used for around 1 minute, as detailed in section 3.2.4, giving an oxide thickness reduction of roughly 100nm within the patterned area. The isotropic nature of the etch ensured that no metal step coverage problems occurred through abrupt etch steps.

3.4.2.1 MAIN INTERCONNECT LAYER

This layer serves three main functions. Firstly, it provides electrical continuity into the central test area of the sample. Secondly, as the first layer to be processed, it provides registration and alignment markers for electron beam and photolithographic processing. Thirdly, it ensures good adhesion of the metal film to the sample surface. This ensures that a good electrical contact can be made and maintained with the interconnect metal film with an ultrasonic wire bonding process. The patterning for this layer is described in section 3.7.

The metallization was formed from two thin layers and a thicker gold layer. Titanium was immediately in contact with the substrate, allowing a strong mechanical bond to be formed with either silicon or silicon dioxide[28]. Palladium was used as an interfacial layer between the titanium and gold layers. Palladium appears to prevent oxidation of the titanium layer. It is thought that palladium in a Pd/Ti, or Pd/W, metal couple acts in

one of two ways; either as an oxygen diffusion barrier or as an electrochemical modifier for the other metal, preventing it from entering an oxidation state. As titanium forms the mechanical bond to the substrate, oxidation of this layer will cause failure during bonding as the sample begins to age. The final reason for the use of palladium is to aid the adhesion between the gold and titanium layers[29-31]. The full metallization was therefore as follows: Ti:Pd:Au 33:30:150 nm. All three films in this film stack had to be deposited in the same vacuum cycle. Note that after the palladium evaporation was complete, 30nm of titanium was deposited within the chamber without opening the shutter to the sample. This acted as a sealing deposition and prevented the palladium layers flaking off in the evaporator chamber.

3.4.2.2 THIN FILM INTERCONNECT LAYER

This layer is required to electrically connect the main interconnect to the device. It must be thin enough such that the device metallization may overlay to it without step coverage problems. This is, however, incompatible with its electrical connection to the previous layer. The thickness of this layer had to be less than the device metallization which was usually around 15 to 20nm. A bilayer of titanium and gold was used, titanium being used to bond the metal to the substrate. Again, both films were deposited in the same vacuum cycle.

Two films types were examined for this layer, one of Ti:Au 2:8nm and one of Ti:Au 3:10nm. The electrical characteristics of both these films were examined and the details of these tests are discussed in chapter 5, section 5.5.3. The thicker of the two films was found to be the most mechanically and electrically stable.

3.4.2.3 BONDPAD LAYER

This metal layer provides bonding pads for connection of the main interconnect to a chip header and electrically connects the main interconnect to the thin film interconnect layer. The total bond pad metallization was some 190nm, made up from a eutectic alloy of nickel/chrome (NiCr) and gold. The thicknesses were: NiCr:Au 30:160nm. Again both films were deposited in the same vacuum cycle.

3.4.3 DEVICE METALLIZATION

The metal used for the device layer metallization was aluminium. This material was evaporated from aluminium wire of 99.999% purity. The purity of the evaporator system was important as aluminium readily reacts with nitrogen, oxygen and hydrogen,

all of which are common background gases in vacuum systems. The ability of aluminium to readily form an insulating oxide layer, when exposed to air or preferably oxygen, was the main property determining its usefulness in the production of tunnel junctions for this work. The granularity of the vacuum evaporated films was another property of this material which makes it of use in this work as further discussed in chapter 4. It was for the reasons of purity and controllability of the background gases during evaporation that the ultra-high vacuum, UHV, evaporator (see section 3.5.3) was used for device metallizations.

Aluminium metallization were usually between 15 and 40 nm thick, dependent on the type of structure to be formed. Oxidation of the film was effected by use of an oxygen partial pressure during the film evaporation and by post deposition oxidation through partially or fully pressurizing the vacuum chamber with oxygen. Aluminium readily forms a native oxide when exposed to oxygen containing atmospheres. The oxide thickness is typically in the range of 1.5 to 3nm for simple exposure oxidation[32,33]. This oxide can then be use to form the insulator between two aluminium layers, resulting in a tunnel junction. Evaporation of aluminium in an oxygen partial pressure also forms an oxidised aluminium film since the solubility of oxygen in solid aluminium is low[34]. In this case aluminium grains were formed with an oxide coating of the grain boundaries.

3.5 EVAPORATION SYSTEMS

3.5.1 PLASSYS MEB450

The Plassys MEB 450 is an electromagnetically controlled electron beam evaporator. This system is capable of evaporating up to six different materials in a single vacuum cycle. Each material is contained in an evaporation boat within a cooled copper hearth. The system is oil diffusion pumped, with a load lock fitted to allow quick sample exchange.

This system was used in this work for deposition of gold, titanium, palladium, nickel/chrome and germanium. Typical evaporation pressures were $\approx 2\times10^{-6}$mbar with 5×10^{-7}mbar achievable where titanium depositions were included. Films thicknesses were monitored using a quartz crystal oscillator deposition monitor. The oscillation frequency change was converted into a thickness by the controlling software. The same software controlled the operation of the evaporator and the deposition shutter, ensuring accurate control of system pressures and deposition thicknesses.

3.5.2 HIGH VACUUM EVAPORATOR

The high vacuum evaporator was used for some aluminium depositions, see chapter 4 and section 5.5.1, as it provided a quick method of depositing aluminium for test samples. However, it was not used for device layer depositions. This system was a simple oil diffusion pumped bell jar evaporation chamber. No load lock system was present, therefore the entire system was evacuated during a sample exchange. The system was vented using the laboratory nitrogen supply. No other gases were fed to the system. Materials were evaporated using resistance heating of tungsten wire baskets or tungsten boats. The samples were positioned directly above the evaporation boats at a distance of roughly 350mm. Thickness deposition was monitored through a quartz crystal oscillator. Thicknesses were measured through the change in oscillation frequency, and a calibration produced for each material used. Base pressures for depositions in this system were typically 2 to 3 $x10^{-6}$mbar, with run pressures between 1 and 5 x 10^{-5}mbar.

Aluminium wire of 99.999% purity was used in this system. The aluminium wire was loaded into a tungsten wire basket and a degas evaporation was used before any film deposition. The degas evaporation was usually equivalent to a deposition of 2 to 5nm. The tungsten wire baskets were used only once as these baskets were partially consumed during the evaporation and became brittle and unreliable for further use. The thickness calibration for aluminium in this evaporator is detailed in the next section.

3.5.2.1 HIGH VACUUM EVAPORATOR THICKNESS CALIBRATION

In order to calibrate the thickness measurements displayed on the crystal oscillator, a number of samples of different thicknesses of aluminium were deposited in this system. Thickness measurements from these samples were taken using a Rank-Taylor-Hobson Talystep instrument. The thickness measurements were then plotted against the change in oscillation frequency to give the system calibration for aluminium. Figure 3.11 shows this calibration and the line fit to the data. The calibration graph of figure 3.11 displays a line fit of $t(Al)$ /nm = $0.056\Delta f + 1.9$ or, a frequency to thickness calibration factor of 0.056nmHz^{-1} for films between 10 and 50nm thick.

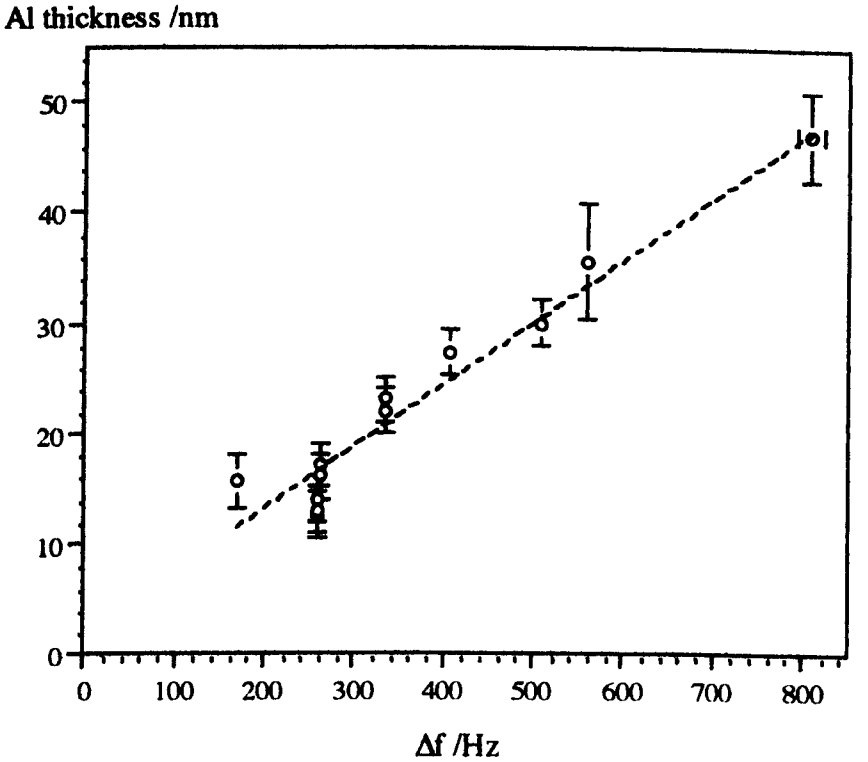


Figure 3.11. Thickness calibration for aluminium in the high vacuum evaporator.

3.5.3 ULTRA HIGH VACUUM EVAPORATOR

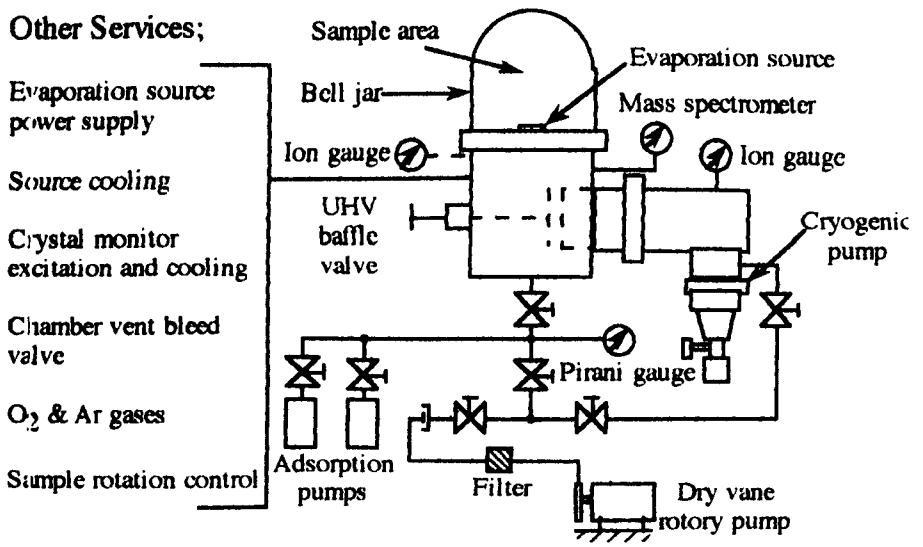


Figure 3.12. UHV evaporator system.

The basis of the device metallization system was a cryogenically pumped stainless steel ultra-high vacuum (UHV) system. This system was designed and supplied by Vacuum Generators (VG) Ltd. The upper section of the main chamber was replaced by a glass bell jar having a VITON™ rubber seal. This allowed access to the chamber without

requiring a load lock or copper gasket seals. The rubber seal, however, limited the ultimate vacuum to 5×10^{-9} mbar. A schematic of this system is shown in figure 3.12.

Two chambers exist in this system; the cryopump chamber and the main, or evaporation, chamber. These two stainless steel sections are isolated by means of a UHV gate valve. Each chamber can be individually rough pumped allowing the operating cryopump to remain under UHV conditions while the main chamber is vented or rough pumped. Change over from rough pumping to cryopumping of the main chamber must take place when the main chamber pressure is at or below 1×10^{-3} mbar.

3.5.3.1 PUMPING SYSTEMS

Rough pumping takes place in two stages. Initially a dry vane rotary pump is used to pump from atmospheric pressure to around 300 mbar. Two 40 litre capacity MS-200 liquid nitrogen cooled adsorption pumps, supplied by VG, then pump the system from 300 mbar to $\approx 10^{-4}$ mbar. Only one pump is required to pump the main chamber. This arrangement allows one pump to be purged of adsorbed gasses while the other pump is in use. After rough pumping the main chamber is pumped via an Air Products HV 202-6C cryogenic pump by opening the UHV gate valve.

3.5.3.2 PRESSURE SENSING

Pressure measurement in this system takes place via three gauges. Atmosphere to 100 mbar is measured via a Bourdon gauge in the rough pumping line. Also in the rough pump line is a Pirani unit, measuring pressure to 1×10^{-4} mbar. UHV pressure and compositional analysis in the main chamber are measured with a MassTorr FX residual gas analyser (RGA). Ion gauge ports are also present on both the main and cryopump chambers, to be driven from an IGC 21 ion gauge controller. All gauges are VG supplied.

3.5.3.3 DEPOSITION MONITORING

Deposition monitoring was carried out with an Intellemetrics IL-100 quartz crystal deposition controller. The crystals used were 6MHz Au coated quartz discs. Using data for the density and acoustic impedance of aluminium, the thickness and rate were directly read from the monitor. Calibration of this system, to account for crystal loading was not carried out as the crystal was replaced before becoming severely loaded. Use of constant density and acoustic impedance values for aluminium in the film thickness range of 10 to 50 nm may be invalid as experiments have shown that

aluminium film density is heavily thickness dependent in this range[33]. However, the readings supplied were a repeatable gauge of the film thickness although not necessarily accurate in an absolute sense.

3.5.3.4 SAMPLE HOLDER

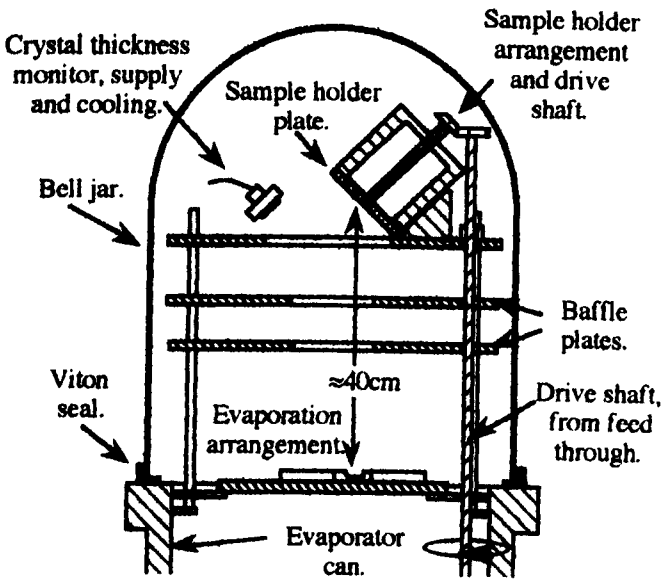
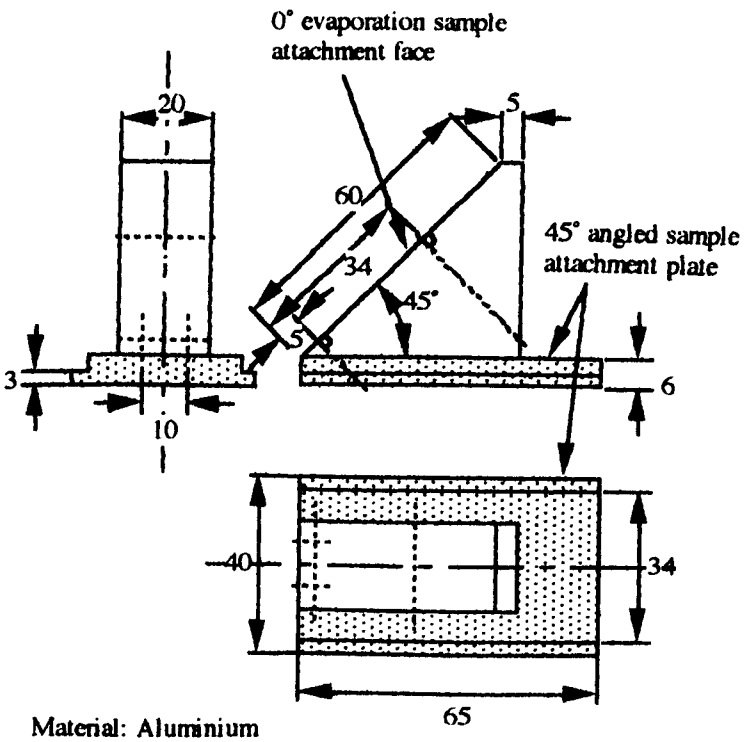


Figure 3.13. Schematic of bell jar arrangement.



Material: Aluminium
All dimensions in millimetres.
Screw holes drilled through block with 2.5mm diameter,
Screw threads tapped in from angled surface for M3 thread.

Figure 3.14. Sample holder attachment for 0° evaporations.

The sample holder was angled at 45° to the evaporant source as shown in figure 3.13. The holder face is capable of rotating through 180°, with indexing points at -90°, 0° and 90°. This design allows for fabrication of shadow evaporated junctions and crossed track junctions as detailed in chapter four. An attachment allowing the sample to be held normal to the source also fits this holder, as shown in figure 3.14. The standard sample holder, which allows 45° angled evaporations, was simply the base of the sample holder of figure 3.14. That is, there was no 0° evaporation sample attachment face present.

3.5.3.5 EVAPORATION SOURCE

The evaporant source was initially an electrostatically focused electron gun. However, evaporation of aluminium in this system was problematic. Molten aluminium, near the evaporation temperature at a given pressure, had virtually no contact angle on any of the crucible surfaces used[35,36]. Crucibles used in this system were produced from graphite, molybdenum and tungsten. This lack of contact angle resulted in the evaporant flowing over the entire crucible surface producing a very large evaporant surface. Another disadvantage of a larger surface area for an electrostatic gun arose from the beam focusing. The electrostatics focus the beam energy onto a static point and therefore require a true point source evaporation. A larger evaporant surface requires larger power input. This resulted in the electrostatic gun running too hot for the cooling system to be effective and aluminium was evaporated from all exposed surfaces of the crucible. Failure of the electron gun was then inevitable since the high voltage insulators became coated with aluminium. Also, the high temperature and greater surface area resulted in increased reaction with the crucible and reduce the evaporation time for a single aluminium charge. Aluminium is known to react, under certain conditions, with most refractory metals[37] and carbon[38], leading to deterioration of the crucible.

Change of the source to a resistance heated tungsten boat resulted in more reliable and reproducible evaporation. However, as aluminium reacts with tungsten[37], the boats deteriorated with time. Therefore, a single boat was used for only one full evaporation. Each vacuum cycle allowed for degassing one source, for the next evaporation, and evaporation from a degassed source. The presence of two positionally different evaporation sources has implications for the shadow evaporation processes used with the suspended mask and crossed track tunnel junction fabrication techniques. These implications are discussed in section 4.2.

3.5.3.4 SYSTEM GASES

Two gases were used with this system, for the purposes of venting and aluminium oxidation. Zero grade argon was used as the vent gas while zero grade oxygen was used for the oxidation process. Zero grade refers to zero water content. These gases were bled into the system through a fine needle valve in the evaporation chamber, the cryogenic pump chamber being isolated using the gate valve during the vent process. Where a partial pressure of oxygen was used during the evaporation, the cryopump was still used as the pump system; however, the pressure in the main chamber was maintained below 1 x 10^{-5}mbar. Argon was used as the vent gas as it is inert. Venting the chamber to nitrogen, as in the high vacuum evaporator, would result in direct reaction between aluminium and nitrogen to form AlN[38]. In air, aluminium is around 10 times more likely to react with oxygen than nitrogen, Al_2O_3 having a ΔG^o = -1526 kJmol^{-1} and AlN having a ΔG^o = -186.3 kJmol^{-1} at STP[39]. However since the nitridation reaction has a negative ΔG, AlN will be formed in a nitrogen vent. The oxygen in this system was used for oxidation of the aluminium either through a chamber vent to oxygen or through bleeding a partial pressure of oxygen into the chamber through the needle valve during an evaporation.

When the chamber was brought to atmospheric pressure using oxygen, care was taken during rough pumping. The adsorption pumps were liquid nitrogen cooled, therefore liquid oxygen collected in the pump traps, presenting an explosion hazard. Where an oxygen vent was used, the chamber was flushed with argon to dilute the pump contents and the pumps were allowed to warm slowly to room temperature before baking to drive off trapped gas.

3.6 SUSPENDED MASK FABRICATION

A suspended mask is a masking pattern formed in a material which is supported above the substrate by some other material. This type of masking system is used in situations where mask pattern must not contact the sample surface. This is especially the case where shadowed or angled evaporation is to be employed to deposit device features[40]. The mask pattern can be formed in two ways, either from a polymer supported thin film mask or from a bilayer of electron beam resist[12]. Both of these processes were investigated in this work; however, the former process gave more consistent results. This section describes the fabrication processes used to produce the suspended mask structure. However, the use of suspended masks for shadow evaporation in this and other work is discussed in chapter 4, section 4.2. Figure 3.15 summarizes the process flow in the fabrication of a suspended mask.

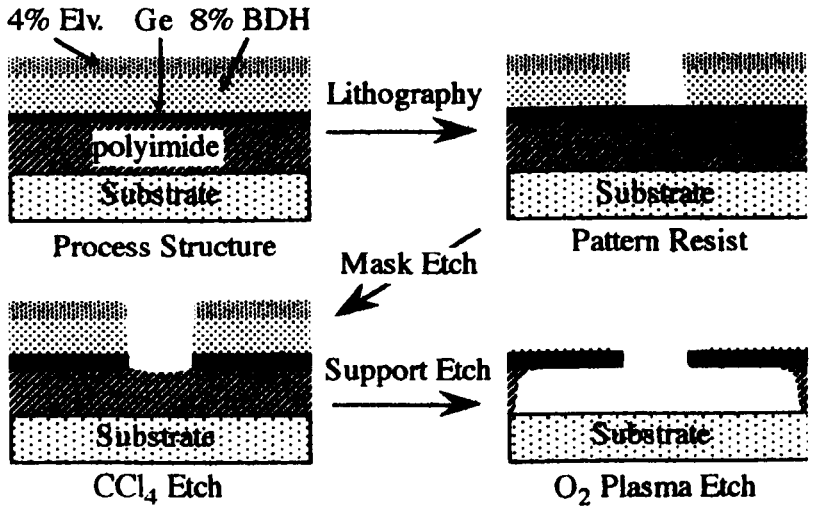


Figure 3.15. Process flow for Ge suspended mask fabrication.

3.6.1 MASK LAYERS

The mask consisted of an evaporated germanium thin film supported on a 200nm thick polyimide film. The polyimide layer was a 6% wt./wt. solution of polyimide granules dissolved in a 65:35 v/v acetophenone:ortho-xylene mixture. This was spun onto cleaned samples at 4.5krpm for 60 seconds. The samples are then baked for around 3 hours at 120°C giving polyimide thicknesses of 200nm ±20nm. The polyimide granules were Ciba-Geigy XU-218. After processing of the polyimide layer, the germanium layer was deposited using a Plassys MEB-450 electron beam evaporator. The germanium layer thickness was 65±2nm. This thickness of germanium provided a mechanically stable mask when the polymer immediately underneath was removed to prepare for the device metallization.

3.6.2 MASK PATTERNING

Patterning of the mask was carried out by reactive ion etching. Firstly, a pattern was defined by electron beam lithography in resist coated onto the germanium surface. The resist used was a bilayer of 8% BDH under 4% Elv., both spun at 5krpm. for 60s and baked at 120°C for at least 1hr and 3hrs. respectively. This relatively thick resist process was chosen to withstand the germanium etch process, which also attacks the PMMA resist. The resist pattern was developed using a 2.5:1 IPA:MiBK mixture at 23°C for 30s. Etching of the germanium layer took place in a Plasma Technology (Oxford Instruments) BP-80 dry etching system. The etch process used a carbon tetraflouride, CF_4, plasma. Etching conditions for the CF_4 etch were as follows: base pressure = 1.5mTorr, sample plate temperature = 19°C, CF_4 flow rate = 37scc min^{-1},

etch pressure = 23 to 24.7mTorr, power = 50W, reflected power = 1 to 2W, dc bias = -320 to -310V, etch time = 3min. A 7min. oxygen pre-clean was carried out before etching and the etch parameters were set and allowed to stabilize for 2 minutes before the etch commenced.

3.6.3 MASK SUSPENSION

To suspend the etched masking layer above the substrate, the underlying support material had to be removed from the regions around the pattern. The final step in suspended mask fabrication was, therefore, an oxygen reactive ion etch to remove the supporting polyimide from the etched mask areas. A PlasmaFab M500 Barrel Asher was used for this process. The process parameters were: oxygen pressure = 110mTorr, power = 80W, time = 50min. This oxygen plasma ash process removed the polyimide support in the area around the etched gaps in the germanium providing an undercut of roughly 0.5μm from the mask pattern edge. The remaining surface resist was also removed by this process, leaving the mask structure ready for use without further processing.

3.6.4 MASK REMOVAL

After the device processing steps, which are discussed in section 4.2, the mask and supporting polyimide required to be removed. This process was exactly the same as the lift off process, leaving a sample with the deposited metal pattern and no remainder of the mask materials used to define the pattern. Removal of the polyimide and the supported germanium mask was achieved by dissolution of the polyimide in a heated 65:35 v/v acetophenone:ortho-xylene mixture. The solution was heated to not more than 40°C and gentle agitation of the sample in the solution was applied.

3.7 PATTERN DETAILS

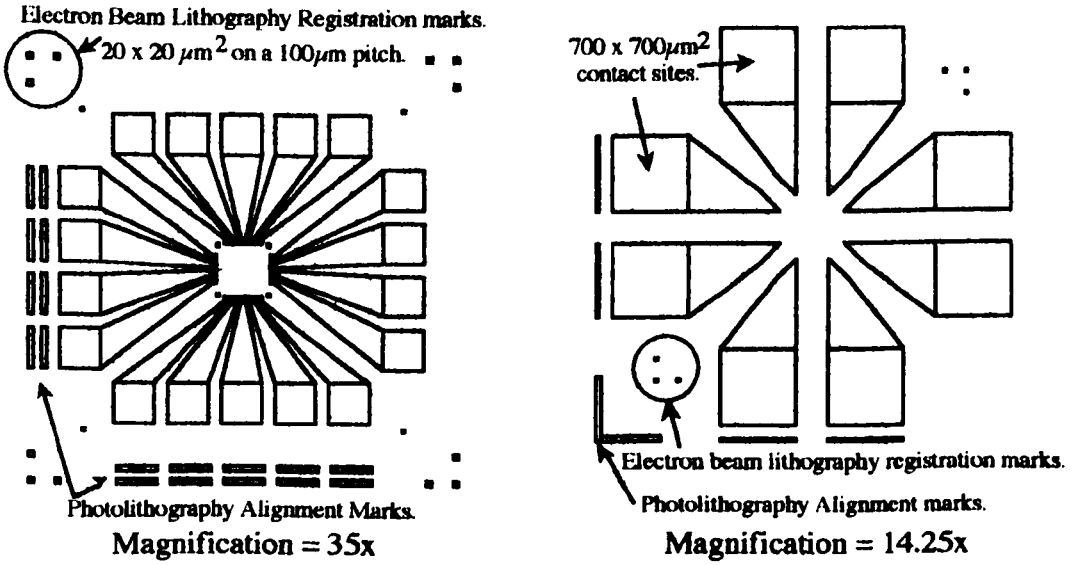
Excluding exposure tests and pattern development, two standard sets of patterns were designed for lithographic processing. These patterns follow the three level process of section 3.4.2, providing interconnect and bonding structures for ultrasonic wedge and manual wire bonding. The metallizations for these layers are described in section 3.4.

The first set of patterns defined a lead frame with 150x150μm^2 bonding pads and a 160x160μm^2 central area for placing device test structures. The overall chip size was 2x2mm^2 Eighteen leads were patterned to allow bonding to 18pin leadless ceramic chip headers. These headers were used on standard inserts for the available cryostats,

see section 5.4.

The second set of patterns provided interconnect with 700x700μm^2 bonding pads and a central device area of 46x46μm^2. The chip size in this case was 5x5mm^2. This was larger to allow bonding to be carried out by hand using pressed indium beads. Although these larger samples could still be bonded into a chip carrier, they were designed to be mounted on an insert designed for a 65 litre helium dewar. This insert is described in section 6.3.10. The 5x5mm^2 samples could also be fabricated in a single layer device and interconnect process by inclusion of the appropriate interconnect and device pattern files in the Beamwriter control job. The interconnect pattern requirements are detailed in the following sections.

3.7.1 MAIN INTERCONNECT LAYER



a. Pattern for 2x2mm^2 samples b. Pattern for 5x5mm^2 samples

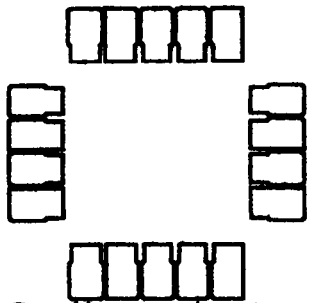
Figure 3.16. Diagrams of first layer lithography patterns.

Figure 3.16 shows the first layer layout for an individual sample of both types. This layer utilizes the metallization process of section 3.4.2. All alignment and registration marks are put down on this layout. This layer had to be included in the Beamwriter job file for single layer device and interconnect processing as it provides the bonding pads for this process. The pattern required to be sleeved as discussed in section 3.3.1.7.

3.7.2 THIN FILM INTERCONNECT LAYER

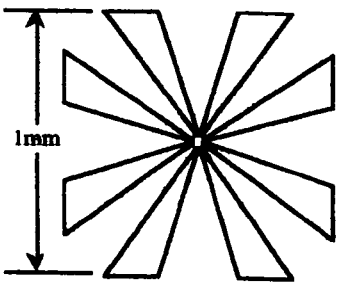
Figure 3.17 shows the second layer pattern layout for both sample types. The bulk of this pattern overlaps the first layer pattern and sits in the central area of the main

interconnect. These patterns use the metallization process of section 3.4.3, providing the connecting layer for the device structure. This layer was only processed using electron beam lithography as the minimum gaps between features are 1μm. Again, where single layer device and interconnect processing is employed, the 5x5mm^2 pattern is sleeved and included in the job file.



Central interconnetion pattern, allowing connection to device layer. Central area is 160 x 160μm^2.
Magnification = 130x

a. Pattern for 2x2mm^2 samples

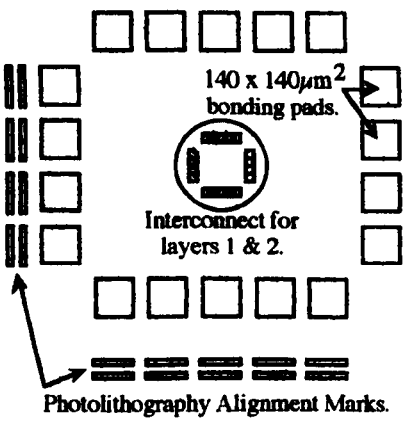


Central interconnection pattern, allowing connection to device layer. Central area is 46 x 46μm^2.
Magnification = 36x

b. Pattern for 5x5mm^2 samples

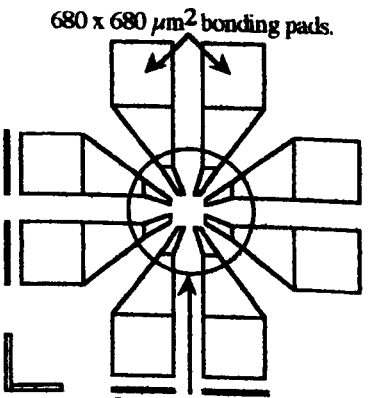
Figure 3.17. Diagrams of second layer lithography patterns.

3.7.3 BONDING LAYER



Magnification = 35x

a. Pattern for 2x2mm^2 samples



Magnification = 12x

b. Pattern for 5x5mm^2 samples

Figure 3.18. Diagrams of third layer lithography patterns.

This final interconnect layer is aligned to the first layer, either by photolithography or electron beam lithography. The metallization is as described in section 3.4.4. Where the three layer metallization process is used, this layer provides continuity between the layer 1 and 2 metallizations and the bondpads. The central areas of these patterns overlap the layer 1 and layer 2 patterns. The cross-section depicted in figure 3.10 is

appropriate to the 2x2mm^2 samples, with the central area bondpad metallization being shown as the metal stack connecting the main interconnect to the thin film interconnect. These bondpad patterns were not included in single layer device and interconnect processing.

3.7.4 PHOTOLITHOGRAPHY MASKS AND SAMPLE LAYOUT

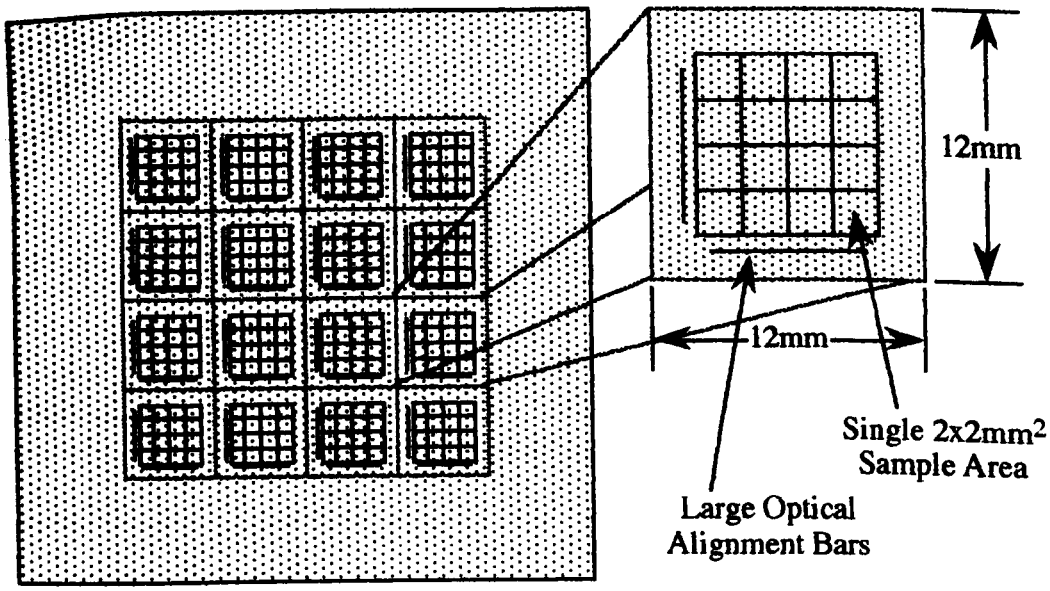
Photolithographic masks for the layer 1 and layer 3, main interconnect and bondpads, were produced for both sample types. Table 3.6 details the BWL file used to produce the mask along with the mask serial number and purpose.

BWL filename	Mask Serial Number	Mask Purpose
RLN0002	GU0049	Bonding: 2x2mm^2 sample, layer 3
RLN0004	GU0052	Bonding: 5x5mm^2 sample, layer 3
RLN0005F	GU0089	Main Interconnect & Alignment: 2x2mm^2 sample
RLN0006	GU0124	Main Interconnect & Alignment: 5x5mm^2 sample

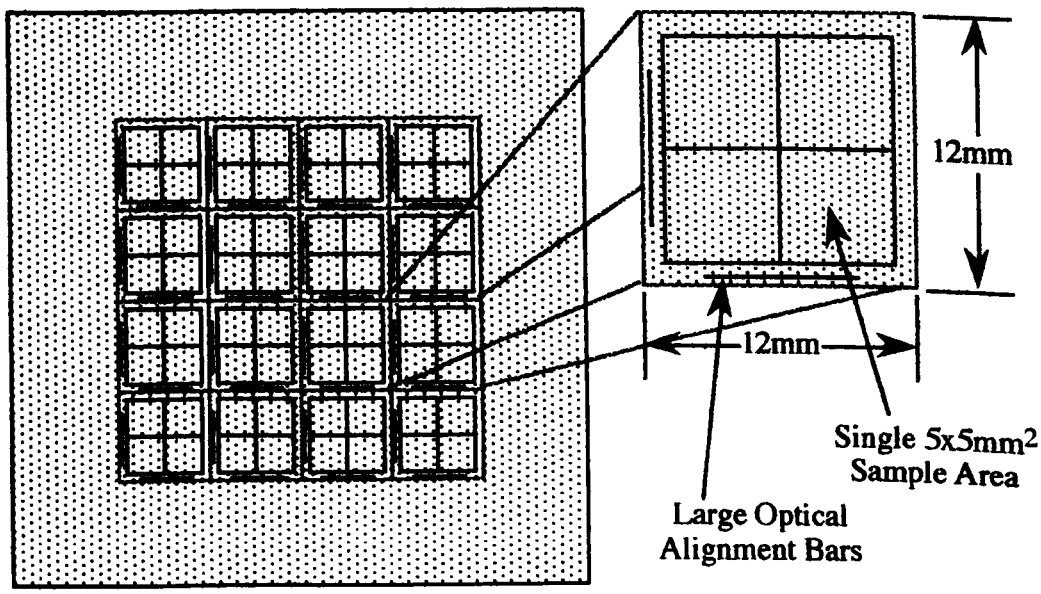
Table 3.5. Photolithographic Mask Details

The sample layouts depended on the type of sample being produced. The masking pattern and sample layout for 2x2mm^2 samples are depicted in figure 3.19a. It can be seen from this diagram that the 2x2mm^2 samples were fabricated on a 12x12mm^2 repeating area. Each 12x12mm^2 area contained a 4x4 array of samples. The corners of this array included test patterns for characterizing the film structures. The 12x12mm^2 sample could then be split into four to allow processing of 3 samples and a characterization sample. These process characterization samples provided van der Pauw sheet resistance squares for the main interconnect, the thin film interconnect, the bondpads and the device metallization. The centre of the characterization sample also included the device pattern for SEM inspection.

The 5x5mm^2 samples had a similar masking layout, as shown in figure 319b, although in this case the 12x12mm^2 pattern area only contained a 2x2 array, with no process characterization features.



a. Mask Layout for 2x2mm^2 Machine Bondable Samples



b. Mask Layout for 5x5mm^2 Manually Bondable Samples

Figure 3.19 Masking Layout for Photolithography Masks

3.8 PROCESS DETAILS

The following sections detail the exact processing sequence for the lead frame samples, high resolution resist processes and single layer device and interconnect samples. The processing of the suspended mask structures is not defined here as this is fully detailed in section 3.6. Also, details of the exact metallizations used to form devices or examine test patterns are not detailed here as these are covered fully where they are discussed in

chapters 4 to 6.

3.8.1 LEAD FRAME (5x5mm^2 or 2x2mm^2)

3.8.1.1 SAMPLE PREPARATION

1. Section wafer into 24x24mm^{2}sample.
2. Clean, following section 3.2.1 and 3.2.2.
3. Oxidize to ≈0.5μm of SiO_2 as in section 3.2.3.

3.8.1.2 LAYER 1 LITHOGRAPHY

This stage in the process could use either photo or electron beam lithography, the two processes, PL1 and EBL1 are covered in the two sub sections below.

A PHOTOLITHOGRAPHY PROCESS (REFER TO SECTION 3.3.2)

PL1:

1. Spin on S1400/31 or S1818 for 60s at 7krpm.
2. Bake at 90°C for 30min.
3. Soak in C_6H_5Cl for 15min.
4. Expose sample using mask
   GU0089 for 2x2mm^2 samples or
   GU0124 for 5x5mm^{2}samples.
   and exposure conditions as given in table 3.5.
5. Develop sample in 1:1 v/v de-ionized H_2O:Developer for 75s.
6. Rinse thoroughly in de-ionized H_2O.
7. Blow dry using N_2 air-gun and inspect.

B ELECTRON BEAM LITHOGRAPHY PROCESS (REFER TO SECTION 3.3.1).

EBL1:

1. Spin on 15% BDH in C_6H_5Cl for 60s at 5krpm.
2. Bake at 180°C for 1hr.
3. Spin on 4% Elv. in ortho-xylene for 60s at 5krpm.
4. Bake at 180°C for ≥3hr.
5. Expose to layer 1 pattern using standard exposure of 300μCcm^{-2}.
6. Develop sample in 1:1 v/v IPA:MiBK at 23°C for 30s.
7. Rinse thoroughly in IPA.
8. Blow dry using N_2 air-gun and inspect.

3.8.1.3 LAYER 1 METALLIZATION (REFER TO SECTION 3.4.2)

1. Oxide etch in 1:4 v/v buffered HF:H_2O for 1min.
2. Rinse in de-ionized H_2O and blow dry.
3. Deposit main interconnect metal (as section 3.4.2.1).
   recipe is 33:30:150nm Ti:Pd:Au.
4. Lift-off mask as in section 3.4.1.

3.8.1.4 LAYER 2 LITHOGRAPHY

In order to produce minimum feature spacings of 1μm and below, this layer requires the use of electron beam lithography.

EBL2:
1. Spin on 12% BDH in ortho-xylene for 60s at 5krpm.
2. Bake at 180°C for 1hr.
3. Spin on 4% Elv. in ortho-xylene for 60s at 5krpm.
4. Bake at 180°C for ≥3hr.
5. Expose to layer 2 pattern using standard exposure of 300μCcm^{-2}.
6. Develop sample in 1:1 v/v IPA:MiBK at 23°C for 30s.
7. Rinse thoroughly in IPA.
8. Blow dry using N_2 air-gun and inspect.

3.8.1.5 LAYER 2 METALLIZATION (REFER TO SECTION 3.4.3)

1. Deposit thin film interconnect metal (as section 3.4.2.2).
   recipe is 3:10nm Ti:Au. (in some cases, 2:8nm Ti:Au)
   Refer to chapter 5, section 5.5.3 for discussion of recipe.
2. Lift-off mask as in section 3.4.1.

3.8.1.6 LAYER 3 LITHOGRAPHY (EITHER PHOTO OR ELECTRON BEAM LITHOGRAPHY)

As with layer 1, this process stage could use either photo or electron beam lithography, the two processes, PL3 and EBL3 are detailed in the two sub sections below.

A PHOTOLITHOGRAPHY PROCESS (REFER TO SECTION 3.3.2)

PL3: Exactly as PL1 (section 3.8.1.2a) except:

4. Expose sample using mask
   GU0049 for 2x2mm^2 samples or

GU0052 for 5x5mm^{2}samples.
and exposure conditions as given in table 3.5.

B ELECTRON BEAM LITHOGRAPHY PROCESS (REFER TO SECTION 3.3.1)

EBL3: Exactly as EBL1 (section 3.8.1.2b) except:

5. Expose to layer 3 pattern using standard exposure of 320μCcm^{-2}.

3.8.1.7 LAYER 3 METALLIZATION (REFER TO SECTION 3.4.4)

1. Deposit bondpad metal (as section 3.4.2.3).
   recipe is 3:160nm NiCr:Au.
2. Lift-off mask as in section 3.4.1.

3.8.1.8 PREPARATION FOR DEVICE LAYERS

The samples produced through this route were then ready for device layer processing. This was the case regardless of the device layer process. It remained simply to further divide the 24x24mm^2 wafer sections into 12x12mm^2 sections for processing of the device layer.

3.8.2 HIGH RESOLUTION RESIST PROCESSES

The electron beam resist processes were used at various stages in the development of processes for the high resolution patterning stages. These resists were used on exposure and pattern tests and also for final device fabrication.

3.8.2.1 EBL-HR1

This resist process contained the standard high resolution resists however, the developer solution was slightly more aggressive than the standard high resolution EBL process. The processing order was as follows.

1. Spin on 2.5% BDH in ortho-xylene for 60s at 5krpm.
2. Bake at 180°C for 1hr.
3. Spin on 2.5% Elv. in ortho-xylene for 60s at 5krpm.
4. Bake at 180°C for ≥3hr.
5. Expose pattern using appropriate exposure.
6. Develop sample in 2.5:1 v/v IPA:MiBK at 23°C for 30s.
7. Rinse thoroughly in IPA.

8. Blow dry using N_2 air-gun and inspect.

3.8.2.2 EBL-HR2

This process was the standard high resolution PMMA EBL resist process used with the Beamwriter. The resist thickness was the same as for the EBL-HR1 process although the use of a high dilution ratio developer provided slightly more resolution and more process contrast. The process was:

As EBL-HR1 except:

6. Develop sample in 3:1 v/v IPA:MiBK at 23°C for 30s.

3.8.2.3 EBL-HR3

This resist process was developed for the pattern granular thin films. It was similar to the EBL-HR1 process but the resist thicknesses were different and the developer used was the same as that of EBL-HR2. The developer again provided resolution and contrast improvements while the thickness changes provide a different resist structure. In this case, a thicker bottom, BDH, layer was used with a thinner Elv. layer on top. This had the effect of allowing greater thickness of aluminium deposition while retaining the ability to perform lift off. The process contrast and resolution was thought to be the same as the EBL-HR2 process.

As EBL-HR1 except:

1. Spin on 2.5% BDH in ortho-xylene for 60s at 3krpm.
3. Spin on 2.5% Elv. in ortho-xylene for 60s at 7krpm.
6. Develop sample in 3:1 v/v IPA:MiBK at 23°C for 30s.

3.8.3 PROCESS FOR "SINGLE LAYER DEVICE AND INTERCONNECT"

Single layer device and interconnect processing allows patterning of the lead-frame and the device layer in one step. This works with any of the resist processes described in section 3.8.2 however, the EBL-HR3 process was used for fabricating patterned thin film devices. This process was used to prepare device samples where indium bonding would take place directly to the aluminium film. In this case, a Beamwriter layout and control job was written which would merge the lead frame (section 3.7.2), thin film interconnect (section 3.7.3) and the device layer pattern. Sleeving was used on the edge sections of the large (layer 1 and 2) patterns to allow proper definition of the resist

sidewall, as discussed in section 3.3.1.7. The following sections detail the processing steps.

3.8.3.1 SAMPLE PREPARATION

1. Section wafer into 24x24mm^{2}sample.
2. Clean, following section 3.2.1 and 3.2.2.
3. Oxidize to $\approx$0.5μm of SiO_2 as in section 3.2.3.

3.8.3.2 LITHOGRAPHY

Using processes EBL-HR1, 2 or 3, except:

5. Expose using sleeved lead-frame and thin film interconnect patterns with standard exposure of 350μCcm^{-2}.
   Also, device layer pattern with exposure as determined by exposure tests

3.9 CONCLUSION

This chapter detailed the general fabrication processes and systems used for both device and interconnect fabrication. The main emphasis of this chapter was to describe the fabrication processes by discussing the interconnect fabrication, where possible. This allowed the processes to be explained without the complication of discussing the techniques and process optimization for fabrication of single electron devices. The following chapter discusses, in turn, each of the three single electron device fabrication techniques studied. Heavy reliance is placed on the information given in this chapter in discussing the device fabrication techniques. In particular, with the suspended mask, shadow evaporated tunnel junctions no reference is made to the suspended mask fabrication process as this is detailed fully in this chapter in section 3.6. Similarly, the lithography systems or pattern generation have been fully discussed in this chapter.

3.10 REFERENCES

1. W. Kern & D.A. Puotinen, *RCA Review.* **31**(2),187 (1970).

2. S.M. Sze, *Semiconductor Devices - Physics and Technology*, John Wiley & Sons, Singapore, 1985, p455.

3. A.N. Broers, *IBM J. Res. Develop.* **32**(4),502 (1988).

4. D.R.S. Cumming, S. Thoms, S.P. Beaumont & J.M.R. Weaver, *Appl. Phys. Lett.* **68**(3),322 (1996).

5. F.W. Billmeyer, Jr., *Textbook of Polymer Science*, John Wiley & Sons, Singapore, 1985, pp16-18.

6. J.M.G. Cowie, *Polymers: Chemistry and Physics of Modern Materials*, International Textbook Company, Glasgow, 1973, pp146-168.

7. C. Binnie, *The fabrication of small geometry MOSFETs using electron beam lithography*, Ph.D. Thesis, No. 7150, University of Glasgow (1984).

8. V.K. Sharma, S. Affrossman & R.A. Pethrick, *British Polymer Journal.* **16**(2),73 (1984).

9. H.W. Deckman & J.H. Dunsmuir, *J. Vac. Sci. Technol. B.* **1**(4),1166 (1983).

10. M. Hatzakis, J. Paraszczak & J. Shaw, *Proc. Microcircuit Eng.* 1981, pp386-396.

11. S.P Beaumont, T. Tamamura, C.D.W. Wilkinson, *Proc. Microcircuit Eng.* 1981, pp 381-388.

12. P. Delsing, T. Claeson, K.K. Likharev & L.S. Kuzmin, *Phys. Rev. B.* **42**(12),7439 (1990).

13. W.M. Moreau, *Semiconductor lithography - principles practices and materials*, Plenum Press, New York, 1988, Chapter 3.

14. J.S. Greeneich, *J. Electrochem. Soc.* **122**,970 (1975).

15. L.D. Jackel, R.E. Howard, P.M. Mankiewich, H.G. Craighead & E.W. Epworth, *Appl. Phys. Lett.* **45**(6),698 (1984).

16. S.A. Rishton, *Resolution limits in electron beam lithography*, Ph.D. Thesis, No. 6907, University of Glasgow, 1984.

17. K. Murata, D.F. Kyser & C.H. Ting, *J. Appl. Phys.* **52**(7),4397 (1981).

18. D.F. Kyser, *J. Vac. Sci. Technol. B.* **1**(4),1391 (1983).

19. B.J. Lin, *J.Vac. Sci. Technol.* **12**(6),1317 (1975).

20. A. Muray, M. Issacson, I. Adesida & B. Whitehead, *J. Vac. Sci. Technol. B.* **1**(4),1091 (1983).

21. H.G. Craighead, *J. Appl. Phys.* **55**(12),4430 (1984).

22. P.B. Fischer & S.Y. Chou, *Appl. Phys. Lett.* **62**(12),1414 (1993).

23. W. Chen & H. Ahmed, *Appl. Phys. Lett.* **62**(13),1499 (1993).

24. P.B. Fischer & S.Y. Chou, *Appl. Phys. Lett.* **62**(23),2989 (1993).

25. G. Owen, *J. Vac. Sci. Technol. B.* **8**(6),1989 (1990).

26. K.Y. Lee, *The fabrication of GaAs membrane and GaAs/GaAlAs heterostructure field effect transistor devices*, Ph.D. Thesis, No. 7687, University of Glasgow, 1987.

27. W.S. MacKie, *Techniques and applications of very high resolution electron beam lithography*, Ph.D. Thesis, University of Glasgow, 1984.

28. M. Caulton, W.L. Sked & F.S. Wozniak, *RCA. Review.* **40**(2),115 (1979).

29. R. Buck, B.W. Sloope & H. Leideiser, Corrosion. **15**,566 (1959).

30. M. Stern & H. Wessenberg, *J. Electrochem. Soc.* **106**(9),759 (1959).

31. D. Mattox, *Thin Solid Films.* **18**,173 (1973).

32. G. Hass, *Z. anorg. Chem.* **254**,96 (1947).

33. T.E. Hartman, *J. Vac. Sci. Technol.* **2**,239 (1965).

34. J.P. Hirth & K.L. Moazed, in *Physics of Thin Films.* **4**:97, eds. G. Hass & R. E. Thun, Academic Press, 1967.

35. Y. Moriya, N. Okuma. & R. Sugiura, *Trans. 8th Int. Vacuum Symposium and 2nd Int. Congress of the American Vacuum Soc.* **2**:1055 (1961).

36. Y. Moriya, *Adv. Vac. Sci. & Technol.* **2**:744 (1960).

37. L. Holland, *Vacuum deposition of thin films*, 3rd printing, Chapter 6, pp169-176, Chapman & Hall, 1970.

38. F.A. Cotton & G. Wilkinson, *Advanced Inorganic Chemistry*, 5th ed., Chapter 7, p215, John Wiley & Sons, New York, 1988.

39. Calculated from thermophysical data in: J.G. Stark & H.G. Wallace, *Chemistry Data Book*, 3rd ed., J. Murray Publishers, 1982.

40. G.J. Dolan, *Appl. Phys. Lett.* **31**,377 (1977).

PAGE
MISSING
IN
ORIGINAL

CHAPTER FOUR

Single Electron Device Fabrication

4.1 INTRODUCTION

A large variety of processes and techniques exist for the fabrication of single electron devices. This variety is mainly due to the large array of materials systems in which tunnelling phenomena can be observed. Some of these systems were described in chapter two. In relation to single electron devices each system has its own merits but metallic based systems provide access to the smallest dimensions available with current fabrication technology. However, they still retain essentially classical device properties in that there is no quantization of particle states in the island between the tunnel junctions. The ability to make use of these small structure dimensions provides access to the very small island capacitances required for high temperature single electron devices.

This work is concerned solely with the Al-AlO_x-Al metal insulator metal (MIM) systems and the manipulation of these materials to provide nanometre scale structures for use as single electron effect devices. Three techniques were studied for the formation of tunnel junctions. These techniques are termed: 1. suspended mask, 2. crossed track and, 3. patterned granular thin film. The first two techniques employ multiple angled evaporation schemes while the last process is a strictly planar single level process. Study of these processes was carried out against the requirements of repeatability and reduction of device dimensions.

This chapter contains four main sections. The first, section 4.2, discusses the suspended mask structures in terms of the technique and the dimensional limits to which the technique can be applied for reliable ultrasmall tunnel junction fabrication. Section 4.3 discusses the crossed track technique in terms of the process itself, the process characterization work and the difficulties arising from the technique. Section 4.4 discusses the patterned granular thin film process, covering the basic technique and the process characterization work. Included in this section is a discussion of the underlying thin film formation properties which make this technique attractive for single electron devices. Finally, section 4.5 concludes this chapter and looks forward to the electrical measurement work documented in the following two chapters.

4.2 SUSPENDED MASK STRUCTURES

The technique of using suspended masks with angled evaporations was originally developed in 1977 by G.J. Dolan[1]. Dolan used this technique to reduce the linewidth of deposited metallic lines below the mask pattern linewidth. Used in conjunction with angled evaporation and lift-off processing, it has been used to produce features in the region of a few tens of nanometres from microlithographically defined patterns[2]. The use of suspended masks with angled evaporation produces two effects; linewidth reduction and pattern offset. Both these effects are shown schematically in figure 4.1. The degree of linewidth reduction is described by equation 4.1 and the offset by equation 4.2.

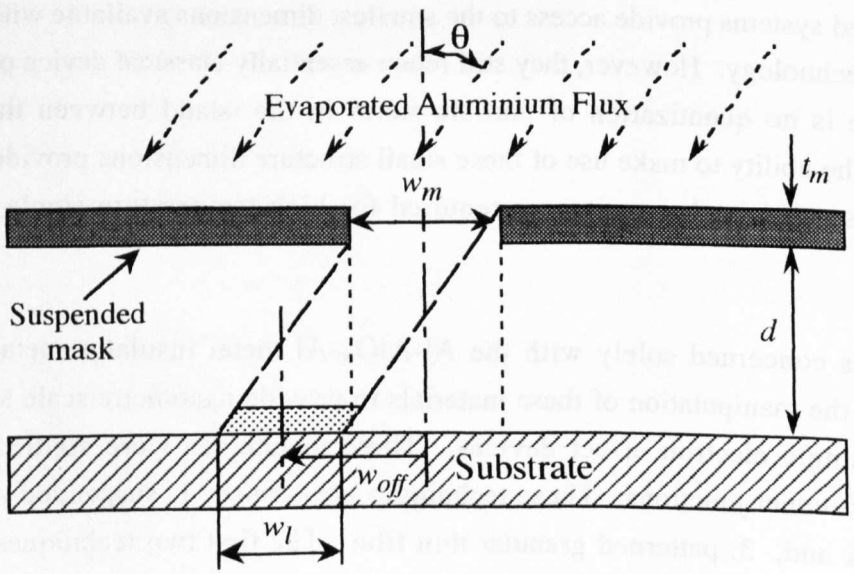


Figure 4.1. Effect of suspended mask interaction with angled evaporation.

$$w_l = w_m - t_m \tan\theta \qquad [4.1]$$

$$w_{off} = (t_m + d)\tan\theta - \left(\frac{w_m - w_l}{2}\right) \quad [4.2]$$

Since angled evaporation through a suspended mask produces an offset between the deposited pattern and the masking pattern, this process has been widely employed in tunnel junction fabrication[3-6]. By evaporating two individual metal layers from opposite angles, it is possible to arrange the masking pattern to produce overlap tunnel junctions as described in figure 4.2. However, it is important to note that the patterning process results in a restriction of the circuit layout to one dimension. There is very little flexibility in this type of patterning to allow for a two dimensional, or truly planar, circuit fabrication process.

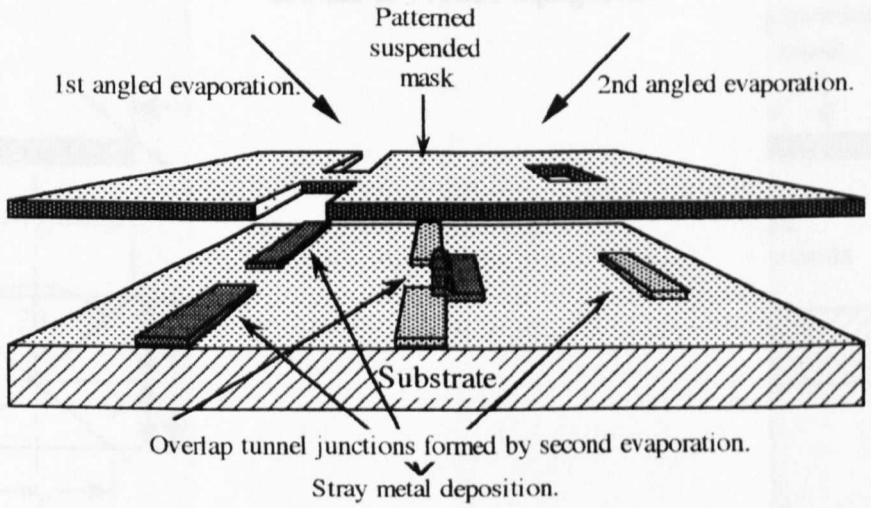


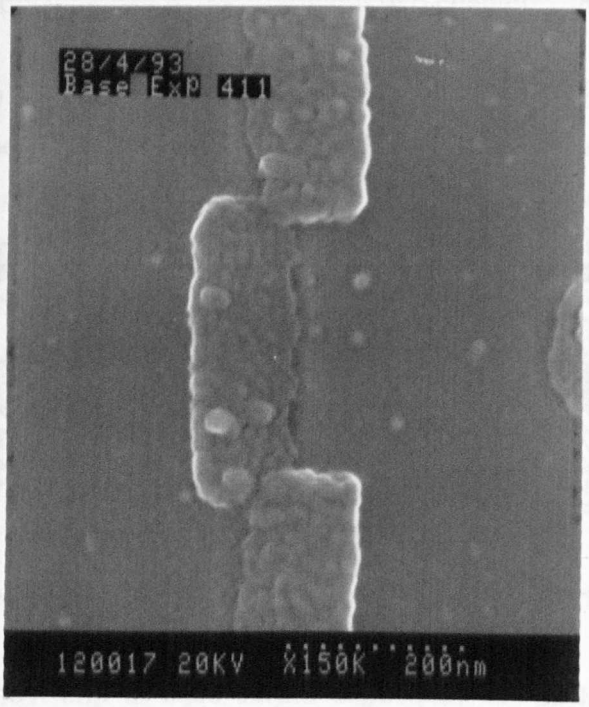
Figure 4.2. Overlap tunnel junction formation using Ge suspended mask.

The process for fabricating the suspended masks used in this work is described in section 3.6. In summary, this process utilizes a solid, germanium, mask layer supported above the substrate by a layer of polyimide. The mask layer is defined through lithography and reactive ion etching. Removal the support polyimide from the vicinity of the patterned areas of the mask allows the mask to be used for shadowing the substrate during metal deposition.

4.2.1 ANGLED EVAPORATION DEVICE FABRICATION

Figure 4.2 illustrates how overlap tunnel junctions can be formed using angled evaporation through a suspended mask. In this diagram, the mask is shown patterned and suspended above the substrate. In the areas to the left and right, off the diagram, support material will still exist, as depicted in figure 3.15. The angled evaporations are achieved by mounting the sample at the correct angle to the evaporation source. In this scheme, the first evaporation takes place and the sample is rotated through 180° to allow

the second evaporation. This is equivalent to flipping figure 4.2 through the vertical axis. The second evaporation arrives from the same source but appears to arrive at the sample from the opposite evaporation direction. Overlap tunnel junctions can be formed by this method provided the first layer is oxidized before the second layer evaporation. This oxidation step will form the barrier of the tunnel junction. It is therefore essential that this oxidation can be performed in the same vacuum cycle as the evaporations. A major difficulty with this fabrication technique is the presence of large areas of unwanted metal deposition as shown in figure 4.2. This effect requires careful design of the pattern and limits the structures, in most cases, to a predominantly one dimensional array.



Micrograph 4.1. Suspended mask double tunnel junction structure.

Micrograph 4.1 shows an overlap double tunnel junction structure formed by this technique, equivalent to the device structure shown in figure 1.8a. The effective overlap of this structure is approximately 40 x 40 nm^2. Section 3.5.3 describes the evaporator system used to deposit overlap tunnel junctions using this technique and section 3.5.3.4, in particular, describes the sample holder arrangement which allowed the angled evaporations. The angle between the sample and the evaporation source is fixed at 45°. Section 6.4.1 describes the circuit structure of the devices fabricated by this technique.

4.2.3 CONTROL OVER DEVICE FABRICATION

Good control of the process parameters is required for reliable junction fabrication with

this technique. The mask thickness and height above the substrate determine the size and position of the deposited pattern and mask design must account for offset effects. The following sections discuss the sources of variation and areas where these variations impact the ability to reliably form tunnel junctions with this technique.

4.2.3.1 PATTERN DIMENSIONS FOR OVERLAP JUNCTIONS

Simple analysis of the structure geometry allows calculation of the dimensions required to form an overlap tunnel junction. The overlap of the masking pattern deposition shadows is a major parameter in control of tunnel junction formation. Figure 4.3 illustrates the geometry of the masking and deposition.

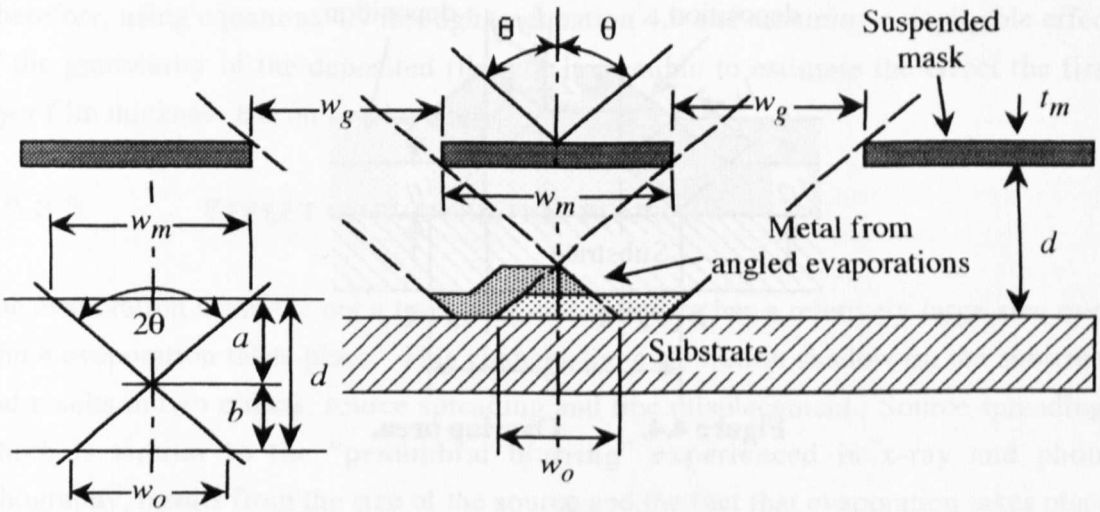


Figure 4.3. Pattern dimensions required for overlap.

From this inset of figure 4.3, we see how the shadow overlap is created from the deposition angle and the masking pattern. It can be seen that for a given shadow mask width, w_m, the support thickness contains two regions; the non-overlap region, a, and the overlap region, b. The relative sizes of these regions are controlled by the evaporation angle and the width of the shadow mask. Equation 4.3 defines each of these parameters.

$$d = a + b, \text{ where, } a = \frac{w_m}{2\tan\theta}, \text{ and, } b = \frac{w_o}{2\tan\theta}. \quad [4.3a, 4.3b \& 4.3c]$$

Therefore,

$$d = \frac{w_m}{2\tan\theta} + \frac{w_o}{2\tan\theta} \quad [4.4]$$

This equation for the mask support thickness is readily re-arranged to give equation 4.5

which evaluates the overlap in terms of the shadow mask width, support thickness and evaporation angle.

$$w_o = 2d\tan\theta - w_m \qquad [4.5]$$

This allows design of the structure to provide adequate overlap but also allows us to see the dependencies of the overlap on processing parameters, assuming no other effects come into play. We therefore have three sources of variation to control; support thickness, evaporation angle and shadow mask width.

4.2.3.2 EFFECTIVE JUNCTION OVERLAP

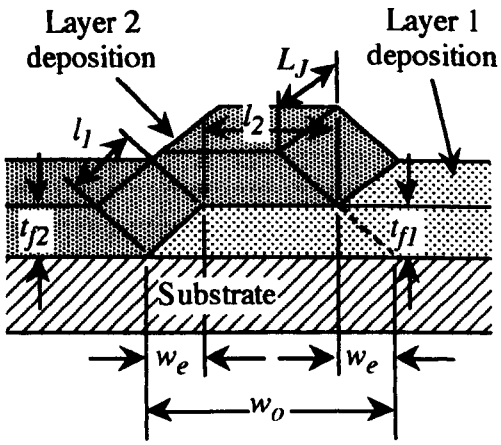


Figure 4.4. Overlap area.

If we assume that the junction overlap is defined as in equation 4.5, we are ignoring the effect of the film thicknesses in determining the dimensions of the junction. Given that, in this case, film thicknesses of 10 to 20% of the support thickness are being used, we must consider the effect of this on the overlap tunnel junction. In this scheme, the junction overlap area is affected by the overlap width and the length of overlap. If we examine the junction structure as in figure 4.4 we see that the film thickness has an effect on the junction overlap. Ignoring any effect of film granularity, the film growth tends to be at an angle to the substrate, due to the angular evaporation method. This leads to a junction area affected by the thickness of the first metal layer. Therefore, an estimate of the junction area can be made from the representation shown in figure 4.4. Equations 4.6 & 4.7 evaluate the important factors in the change of junction area, and equation 4.7 evaluates the reduction in the "parallel" overlap area due to the thickness of the first layer. The parameters of interest in this model are;

$$w_e = t_{f1}\tan\theta, \qquad [4.6]$$

$$l_1 = \frac{t_{f1}}{\cos\theta}, \qquad [4.7]$$

and the reduced parallel overlap, l_2, is given by:

$$\begin{aligned} l_2 &= w_o - 2w_e \\ &= w_o - 2t_{f1}\tan\theta. \end{aligned} \qquad [4.8]$$

The junction area can be estimated by,

$$\begin{aligned} \text{Junction area} &= \text{Junction length} \times \text{effective overlap} \\ \Rightarrow \quad A_J &= L_J(l_1 + l_2). \end{aligned} \qquad [4.9]$$

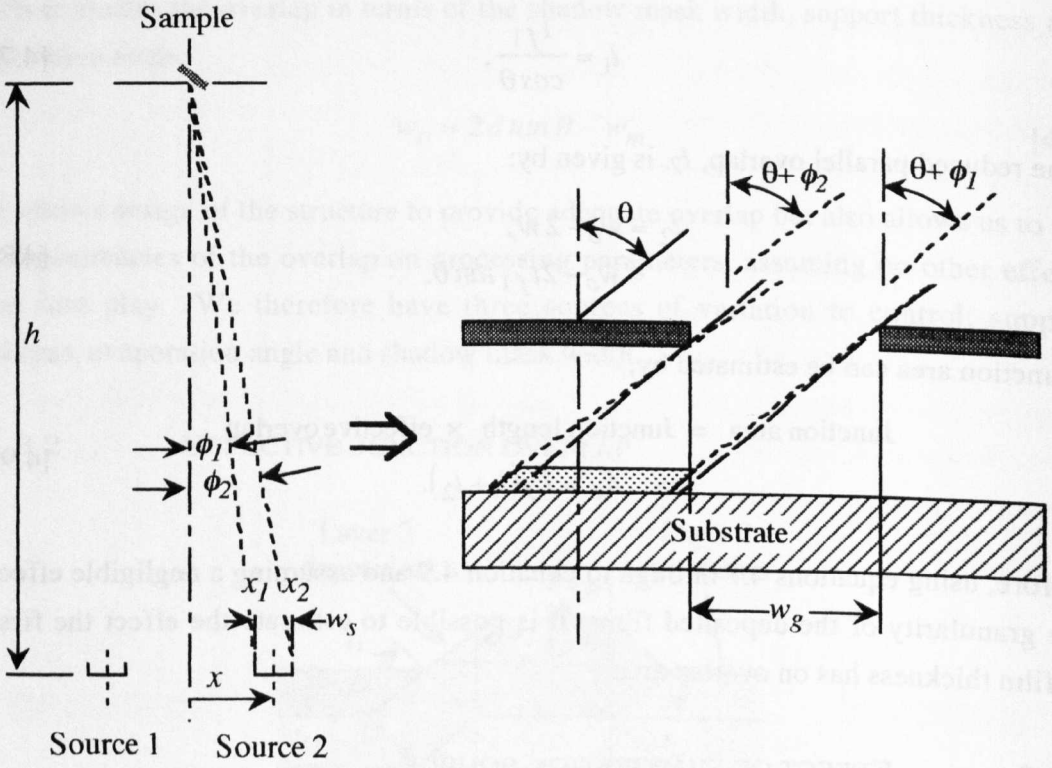
Therefore, using equations 4.7 through to equation 4.9 and assuming a negligible effect of the granularity of the deposited films, it is possible to estimate the effect the first layer film thickness has on overlap area.

4.2.3.3 EFFECT OF EVAPORATION SOURCE

The evaporation source is not a true point and therefore has a relatively large area over which evaporation takes place. This changes the evaporation shadow on the substrate and results in two effects; source spreading and line misplacement. Source spreading, which is similar to the "penumbral blurring" experienced in x-ray and photo lithography, results from the size of the source and the fact that evaporation takes place across the entire source surface. Although the spatial distribution of the evaporation rate may vary, there will be deposition on the substrate from the edges of the evaporant source. The models developed in the preceding sections assumed a point source at a distance large enough to result in very little angular distribution of the incident flux. If the source is no longer a point source, some degree of angular distribution will exist in the incident flux. The image of the mask on the substrate is therefore enlarged due to this distribution. Figure 4.5 depicts a source which is both misaligned from the sample and is large compared to the pattern dimensions.

Figure 4.5 examines the effect of both source misplacement and the physical size of the source. This model uses the change in evaporation angle to account for this, calculating it from the source misplacement, x, and the size of the source, w_s. Given the position of the source centre, the edges of the source are easily calculated,

$$x_1 = x - \frac{w_s}{2}, \quad \text{and,} \quad x_2 = x + \frac{w_s}{2}. \qquad [4.10a \ \& \ 4.10b]$$



a. Variation in evaporator source arrangement.

b. Effect of source variation on suspended mask deposition (from source 2).

Figure 4.5. Errors due to source dimensions and position.

The angles depicted in figure 4.5 can then be calculated, as follows:

$$\phi_1 = arctan\frac{x_1}{h} = arctan\left(\frac{2x - w_s}{2h}\right). \qquad [4.11]$$

$$\phi_2 = arctan\frac{x_2}{h} = arctan\left(\frac{2x + w_s}{2h}\right). \qquad [4.12]$$

This in effect changes the angle of deposition and alters the overlap of the metal films. Source-substrate misalignment also results in misplacement of the lines of the deposited pattern. The error here should be small, due to the relatively large source-sample separation. However, complete misplacement and therefore a lack of overlap can result if the source is too far off centre. When considering the effect of the source on the deposition shadow we see that the pattern linewidth is also altered. However, where we are forming the tunnel junctions by the overlap shown in figures 4.2 and 4.3, this change in linewidth is not relevant. The most important effect here is the deposition angles presented by the two different sources. The model shown above, in figure 4.5, considers deposition from source 2 and in this case the greatest deposition angle is the angle which determines where the film edge for the overlap is placed. This is exactly the same for deposition from source 1. The difference between the sources arises from

the spread of deposition angles, θ_d: For source 1 the spread is, $\theta-\phi_2 < \theta_d < \theta-\phi_1$ and for source 2 the angular spread is, $\theta+\phi_1 < \theta_d < \theta+\phi_2$. Therefore the angle to be considered in examining the change in overlap is the largest deposition angle for that particular source, ie. $\theta-\phi_1$ for source 1 and $\theta+\phi_2$ for source 2.

In the case of the system being used for this work, the source-sample separation is about 400mm, with the source widths being 5mm and offset from the centre by 15mm in each case. Using these values we find that ϕ_1 is 1.8° and ϕ_2 is 2.5°, from equations 4.11 and 4.12 respectively. This makes the evaporation angle (relevant to junction overlap) 47.5° if source 2 is used and 43.2° if source 1 is used. It is important, then, that the same source is used for both evaporations to ensure that the same deposition angle is maintained for both modes.

4.2.3.4 VARIATION IN PROCESS PARAMETERS

The models developed above all assume nominal process parameters. Some of the models include variations due to dimensional effects, such as the film thickness and the source position and size. However, the greatest problem posed to ultrasmall tunnel junction fabrication by this technique is control over variations in process parameters. Errors incurred in this system because of the physical size and position of the source affect the formation of tunnel junctions and are systematic. The scale of these errors is discussed in section 4.2.3.3.

Three main variations in the process come into play, firstly the thickness of the polyimide support, secondly the thickness of the layer 1 metal film and thirdly positional errors in the masking pattern. The thickness of the germanium masking will affect the linewidth of the deposited track, but it does not affect the overlap dimensions. In order of scale in the variations, the polyimide mask support results in ±20nm variation on 200nm. The layer 1 metal thickness is typically 17nm with an error of at worst ±0.4nm within a sample. Sample to sample, the thickness variation can be as much as ±2nm on the target thickness. Finally, although the masking pattern is written within one stitching field on the Beamwriter, it is possible for variations in exposure to shift the pattern edges, that is w_m in figure 4.3, by up to 5nm relative to each other. An rms. calculation of the effect of these variations is presented in table 4.1.

It is clear from the analysis presented in table 4.1 that the ability to form ultrasmall overlap tunnel junctions by this method is seriously affected by the control over polyimide thickness. In this fabrication process we are likely to have serious reproducibility problems on junctions fabricated around the 40nm overlap mark.

Improved control over the polyimide would improve the reliability of junction formation but it is clear that small overlap junctions will be irreproducible with this process. This conclusion is supported by inspection of the single electron device test structures fabricated by this method, as described in section 6.4.

Parameter	Equation	Error /nm
Polyimide thickness Δd	---	±20.0
Layer 1 film thickness, Δt_{f1}	---	±2.0
Masking tolerance, Δw_m	---	±5.0
Effect on shadow overlap, Δw_o	$\Delta w_o = \sqrt{(2\Delta d \tan\theta)^2 + (\Delta w_m)^2}$	±40.3
Effect on junction sloped edge, Δl_1	$\Delta l_1 = \sqrt{(\Delta t_{f1}/Cos\theta)^2}$	±2.8
Effect on junction flat edge, Δl_2	$\Delta l_2 = \sqrt{(\Delta w_o)^2 + (2\Delta t_{f1} \tan\theta)^2}$	±40.5
Effect on junction overlap, $\Delta(l_1+l_2)$	$\Delta(l_1+l_2) = \sqrt{(\Delta l_1)^2 + (\Delta l_2)^2}$	**±40.5**

For evaporation angle $\theta = 45°$.

Table 4.1. Errors produced by process tolerances.

4.3 CROSSED TRACK STRUCTURES

This technique was developed in an attempt to circumvent some of the problems associated with the suspended mask technique. The two main benefits sought were pattern formation in a more planar, two dimensional, manner and a reduction in the number of process steps required to form the deposition mask. The type of structure sought is depicted in figure 1.8b. All development work for this process involved electron beam lithography using the Beamwriter at 50keV accelerating voltage. The samples used were cleaned, deoxidised Si samples. No oxide growth was employed. Evaporation took place in the UHV evaporator described in section 3.5.3, using an electrostatically focused electron gun arrangement.

4.3.1 CROSSED TRACK TECHNIQUE

Angled evaporation is employed in this process to shadow regions of the substrate through a resist mask. The written pattern, the aspect ratio of the developed resist profile and angled evaporation are used to control the order and position of the deposited metal. Lines are written in the resist at right angles to each other, giving two modes for the evaporation. Tunnel junctions are therefore formed at the intersection of the lines and the junction dimensions are controlled by the wire width. Figure 4.6

illustrates this process.

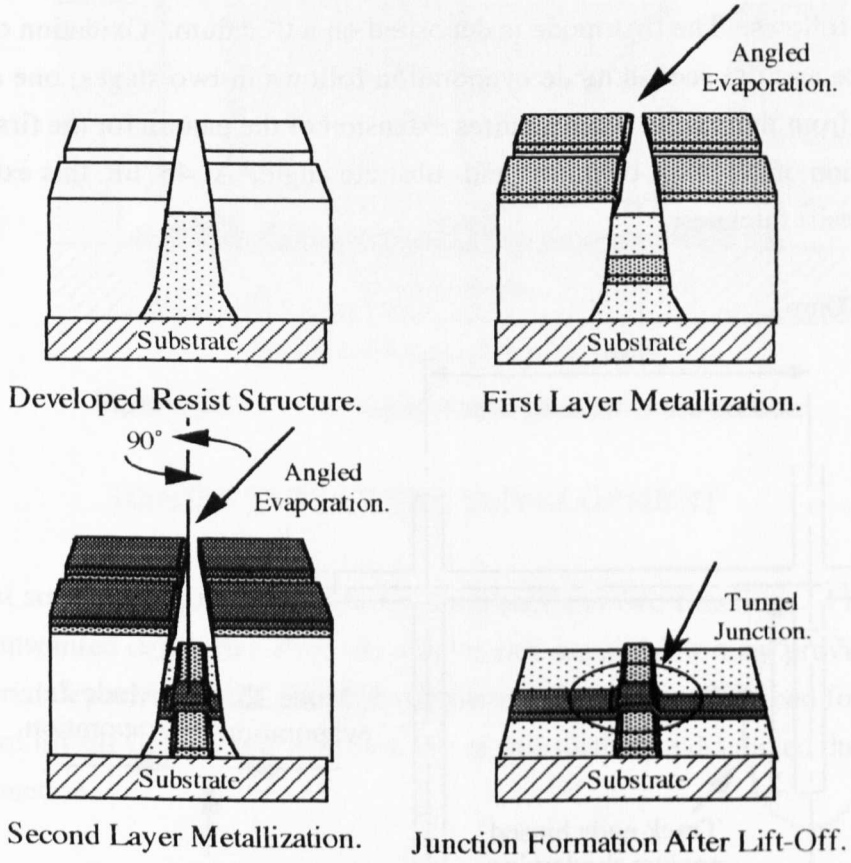


Figure 4.6. Crossed track evaporation technique.

The lines are presented to the source on an angled substrate in such a way that they are either parallel or perpendicular to the line of evaporant flux. Lines parallel to the flux are open for substrate deposition while the substrate is completely shadowed on those lines perpendicular to the flux direction. Rotation of the sample through 90° allows selection of the mode in which deposition will take place. The sample holder is described in section 3.5.3.

4.3.2 PATTERN DESIGN AND DEVELOPMENT

The 45° angle of the sample to the source in the UHV evaporator constrains the pattern design. As the resist is used to shadow the substrate from the source, shadowing will also take place at one end of the parallel (or open) tracks. The end closest to the source of any open track will be shadowed from film deposition by the resist. Without biasing of the track lengths against this effect, fore-shortening of the tracks and therefore incomplete pattern formation will occur. This biasing is shown on the lower end of the vertical tracks of figure 4.7. The need for biasing can be eliminated if the sample is rotated through 180° and evaporation repeated along the same mode. However, the

UHV sample holder arrangement has a limit of 200° rotation and as such a decision is required over the order of evaporation. The evaporation order for the pattern design of figure 4.7 is as follows. The first mode is deposited on a 0° datum. Oxidation of these lines takes place and the second mode evaporation follows in two stages; one at +90° and one at -90° from the datum. This requires extension of the pattern for the first mode lines by a function of the resist thickness and substrate angle. At 45° tilt, this extension is equal to the resist thickness.

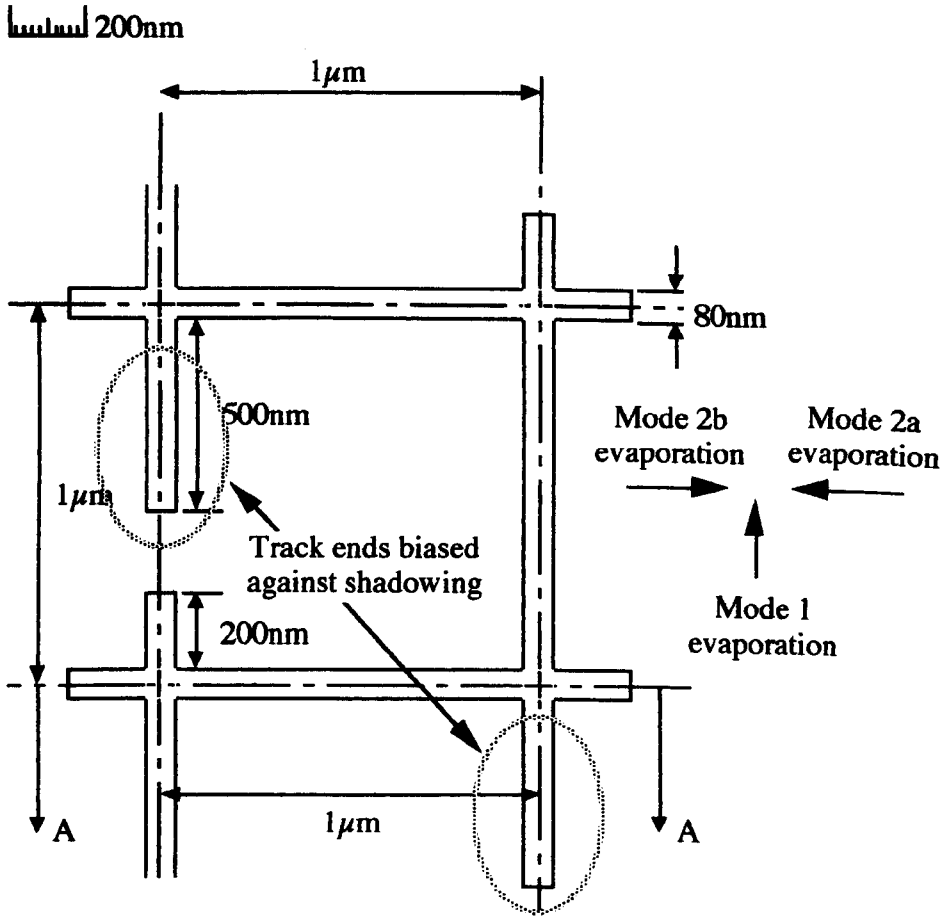
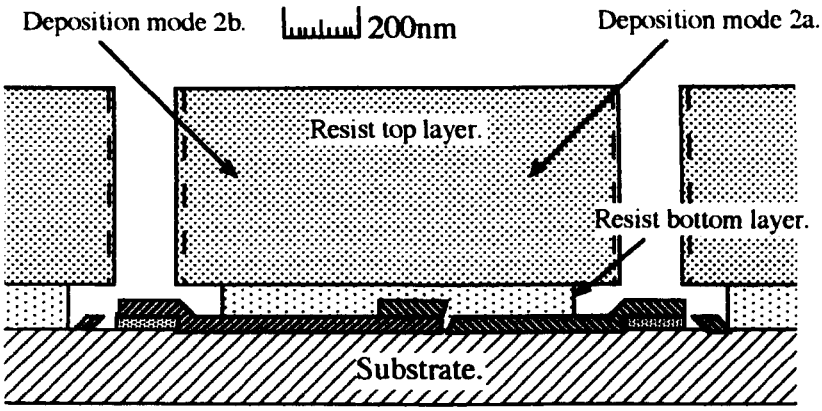


Figure 4.7. Cross track pattern design.

Figure 4.8 illustrates an idealized cross section through the pattern of figure 4.7 at AA. This cross section shows the relative arrangements of the metal layers after all depositions are complete. The mode 1 tracks go into the plane of the page.



Section through AA of figure 4.7.

Figure 4.8. Crossed track pattern cross section.

4.3.3 RESIST STRUCTURE DEVELOPMENT

The resist structure used in this process must perform two functions. Firstly, it must shadow unwanted deposition from the source and secondly, it must provide a break in any material deposited on the sidewalls of the resist, similar to the need for an undercut in standard lift-off processing. The first test of this technique displayed the need for the second function.

4.3.3.1 INITIAL RESIST STRUCTURE

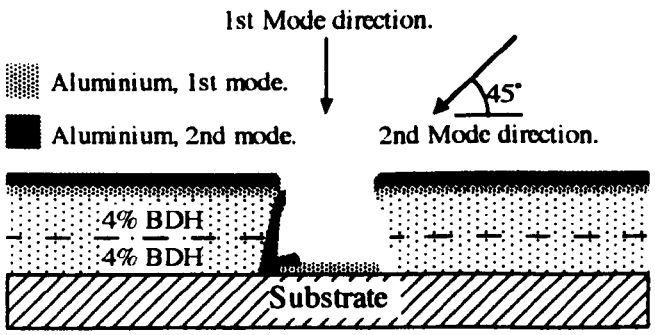


Figure 4.9. Schematic of connected sidewall and substrate deposit.

To test the process, a sample was prepared using a bilayer of 4% BDH resist spun at 5krpm with normal baking conditions. An 80nm linewidth crossed track pattern was written in the resist as part of an exposure test using 50keV accelerating voltage. Micrograph 4.2 and figure 4.9 display the problem encountered with this test where the resist sidewall deposition is not properly separated from the deposited metal track. In this instance the resist slot depth to width aspect ratio was around 1.5. For a 45° substrate angle this aspect ratio must be at least 2 to separate the sidewall and substrate deposited material. Values in excess of 4 are more useful since the sidewall metal will

be well separated from the substrate, resulting in more reliable lift-off processing.

Micrograph 4.2. Initial crossed track fabrication test.

4.3.3.2 REQUIRED RESIST PROFILE

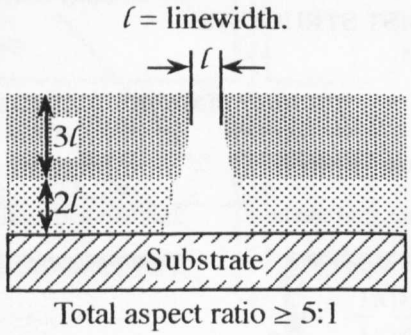


Figure 4.10. Illustration of resist profile dimensions.

For standard lift-off processing using the PMMA resist system the aspect ratio for isolated lines can approach 10:1 before resist stability problems begin to affect the process resolution. Therefore, it should be possible to find a resist combination to fit the profile shown in figure 4.10 for any given linewidth. In the profile shown in figure 4.10, sidewall deposition should only take place on the top resist layer. This results in a large enough gap between the sidewall deposition and the metal tracks. The undercut profile of the resist should result in the track end sidewall deposition being broken, allowing lift-off processing. This resist profile therefore requires use of a high sensitivity, low sensitivity (BDH, Elv.) bilayer resist system. To fabricate 60 to 100nm lines in this manner, using a bilayer resist, resist thicknesses of 120 to 200nm on the

bottom and 180 to 300nm on the top were required. None of the standard resists provide these thickness ranges.

4.3.3.3 CALIBRATION AND SELECTION OF RESIST THICKNESSES

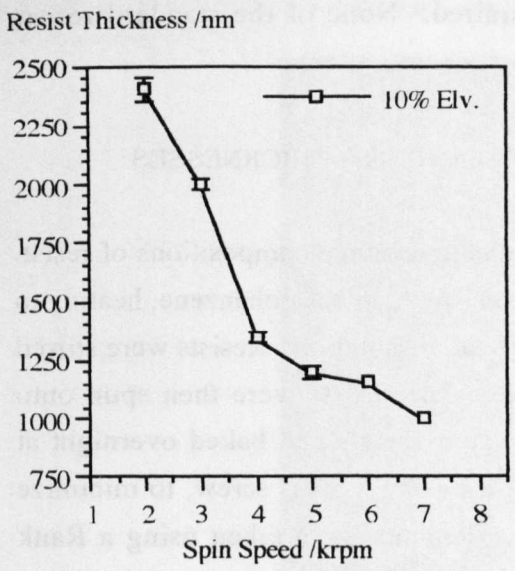
A series of calibrations were undertaken on various percentage compositions of resist. In each case, the weight percentage of resist was dissolved in chlorobenzene, heating to around 50°C and stirring with a magnetic stirrer to aid dissolution. Resists were stirred until no polymer residue remained in the bottle. The resists were then spun onto cleaned 10 x 10mm^2 silicon samples at various spin speeds and baked overnight at 180°C. Lines were scratched into the resist coating using a brass screw, to minimize damage to the silicon surface. Thickness measurements were taken using a Rank Taylor-Hobson Talystep surface profileometer, using a stylus force of around 10mg .

Resist compositions tested were as follows:- 10% Elv., 7% Elv., 6% Elv., 7% BDH, and 6% BDH. All resists were dissolved in chlorobenzene. The most suitable thickness ranges were found in the 6% BDH and 6% Elv. resists. Table 4.2 presents the thickness measurements for various spin speeds for these resists while figure 4.11 and 4.12 graph these results.

Spin Speed /krpm.	Resist Thickness /nm				
	10% Elv.	7%Elv.	6% Elv.	7% BDH	6% BDH
2.0	2400±50	875±50	-----	-----	-----
3.0	2000±25	620±20	510±10	345±10	280±10
4.0	1350±25	640±20	420±10	280±10	270±10
5.0	1200±25	480±20	380±10	278±10	205±10
6.0	1160±10	480±30	350±10	238±10	185±10
7.0	1000±10	420±10	295±10	208±10	174±10
7.5	-----	420±10	-----	-----	-----
8.0	-----	420±10	263±10	203±10	163±10

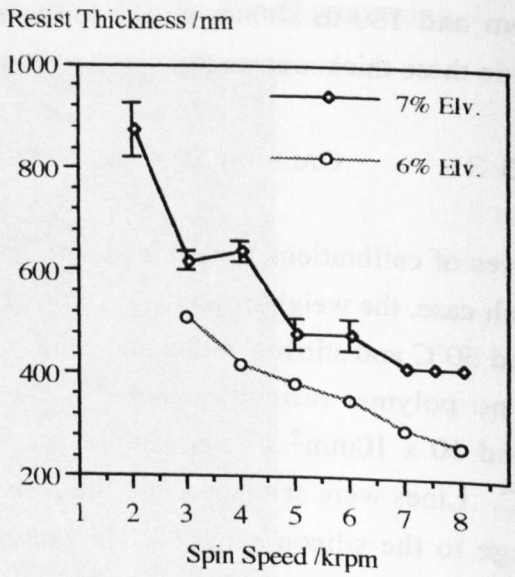
Table 4.2. Table of resist thicknesses at various spin speeds.

Figures 4.11 and 4.12 present the results of table 4.2 in graphical form. These show the standard variation of resist thickness with the inverse of spin speed as dictated by the resist viscosity. The lower weight percentage resists give the thinner films as these are the least viscose solutions. The most appropriate resists (cf. section 4.3.4.2) for this work are therefore the 6% BDH resist, spun between 5 and 8krpm for the bottom layer and 6% Elv. spun at 6 to 8krpm for the top layer.



Note; points without error bars have error too small to represent on the scale of the graph.

Figure 4.11a. Thickness vs spin speed graph for 10% Elv. resist.



Note; points without error bars have error too small to represent on the scale of the graph.

Figure 4.11b. Thickness vs spin speed graph for 6% and 7% Elv. resists.

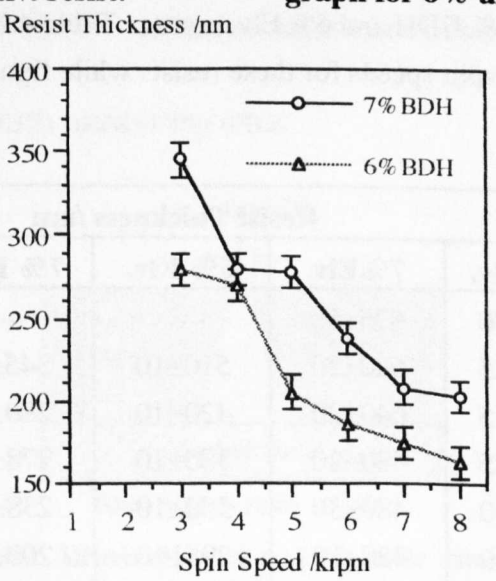


Figure 4.12. Thickness vs spin speed graph for 6% and 7% BDH resists.

4.3.4 FABRICATION PROBLEMS

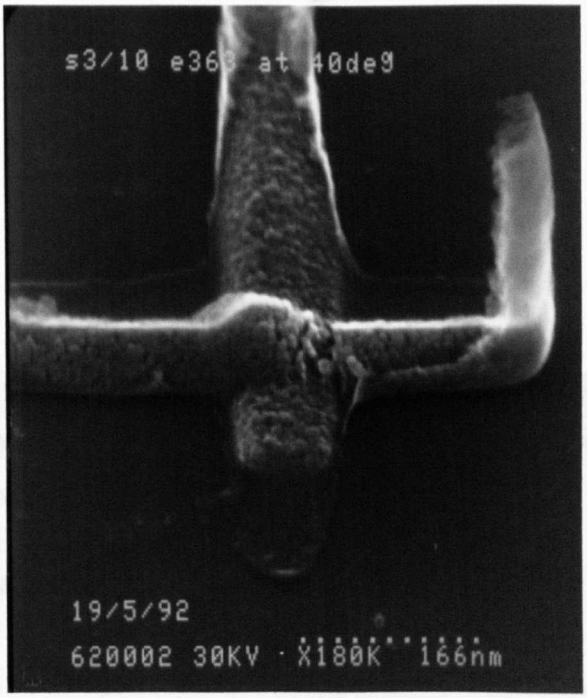
4.3.4.1 DEPOSITION ON RESIST SIDEWALLS

Two problems are associated with deposition on the resist sidewall. Firstly it reduces the resist slot width available for definition of the following tracks. This is discussed in section 4.3.4.4. Secondly, it leads to coating of the resist sidewall in such a way that the slot closes and forms a continuous metal film across resist top surface. When this occurs it significantly impedes lift-off processing. It also leads to metal bridges being

formed across the resist surface which do not detach from the sample during lift-off. Careful design of the linewidths and control of the deposition thickness are required to ensure the metallization does not bridge the track.

4.3.4.2 DEPOSIT AT TRACK ENDS

Since deposition occurs at an angle to the substrate and the resist sidewalls, the resist wall which is not shadowed at the track end will receive a full thickness metal film. The result of this is the creation of a large, unwanted, vertical metal track at one or both ends of the metal wire (depending on whether one or two evaporations were used to form the wire). Micrograph 4.3 displays this feature, which was common to most of the evaporations performed in this way.



Micrograph 4.3. Continuous vertical metal film formed at the track end.

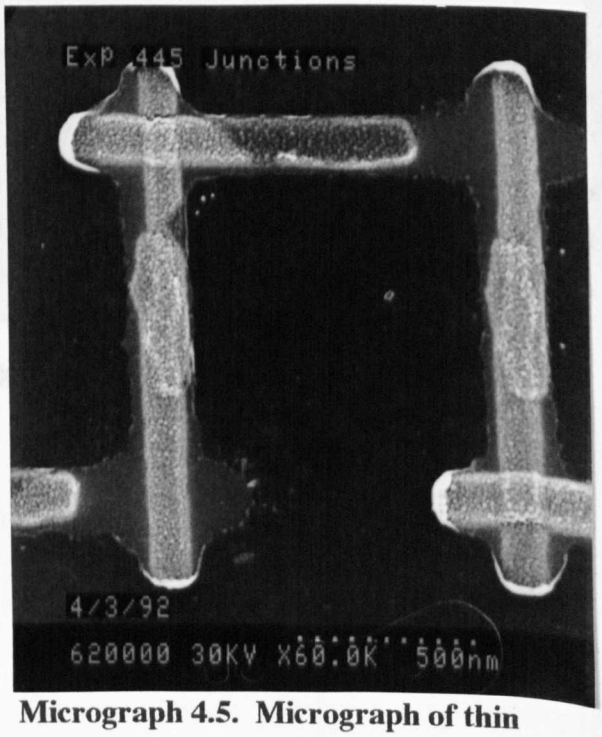
The solution to this problem lies in reducing the thickness of the bottom layer of resist and increasing the undercut to break the metal deposited on the resist. However, it is difficult, due to the small difference in sensitivity between BDH and Elv. (cf. table 3.2) resists to achieve the degree of undercut required.

4.3.4.3 ULTRA-THIN ALUMINIUM FILM DEPOSITS

It was quickly realized that the deposition of an ultra-thin film of aluminium on the substrate was causing problems with the fabrication process. This film was appearing in shadowed areas of the substrate and in some cases left behind after lift-off from a

deposit that was on the resist sidewalls. Two possible sources for this deposit were identified; 1. reflection of source flux from the evaporator bell jar, changing the effective substrate-source angle for some of the incident flux, and 2. surface migration of adsorbed atoms leading to coating of the resist sidewalls and areas which should remain shadowed. It is likely that a combination of these two processes exists to produce the observed effects.

Micrograph 4.4. Micrograph of thin film deposit in a junction area.

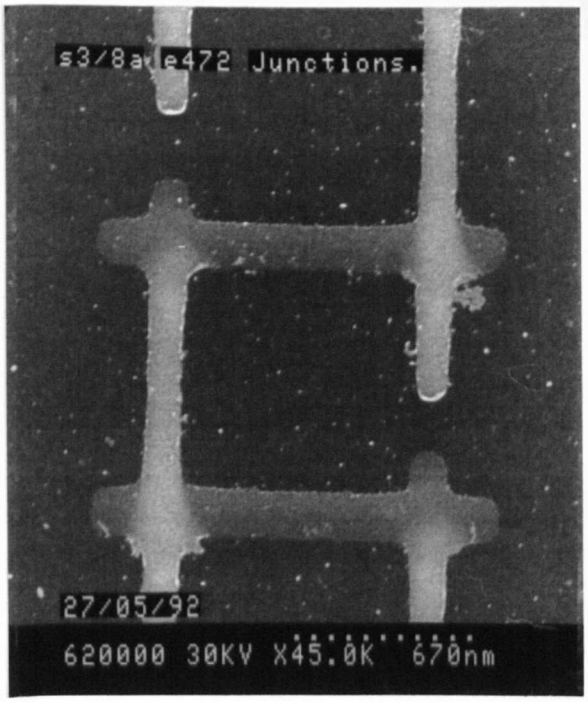


Micrograph 4.5. Micrograph of thin film deposit around crossed track pattern.

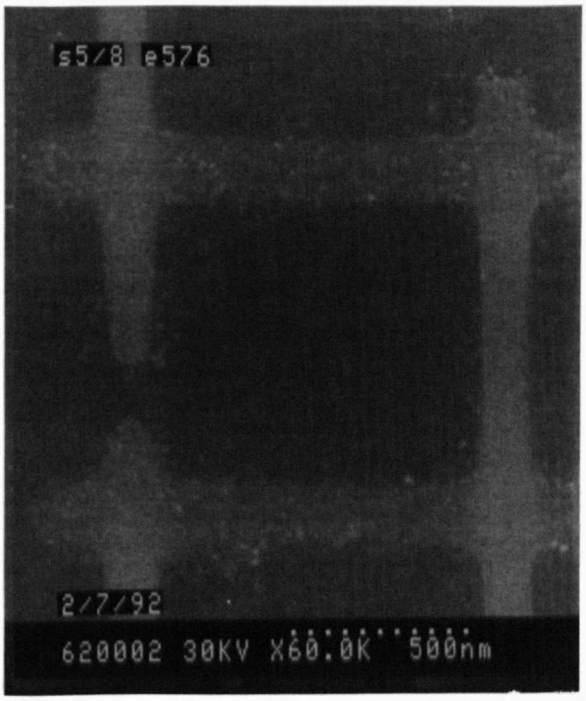
Micrograph 4.4 shows a single junction area covered with ultra-thin material which appears to have arisen from the resist sidewalls. Micrograph 4.5 shows that this material is present on the entire exposed sample surface suggesting some form of stray deposition.

4.3.4.3A BELL JAR REFLECTION

To test whether deposition from stray evaporant flux was causing this ultra-thin deposition, a sample was placed in the evaporator with a glass slide 50mm directly below the sample. This eliminated direct flux deposition while allowing deposition at angles roughly perpendicular to the substrate. A total direct flux thickness of 86nm was deposited at an average rate of $0.1 nms^{-1}$. The result of this test, after dissolving the resist, is shown in micrograph 4.6. The resist pattern has been reproduced exactly, indicating that deposition on the substrate is occurring from all directions within the evaporator system.



Micrograph 4.6. Micrograph of the crossed track pattern area. No direct deposition, no baffle plates.



Micrograph 4.7. Micrograph of the crossed track pattern area. No direct deposition, baffle plates present.

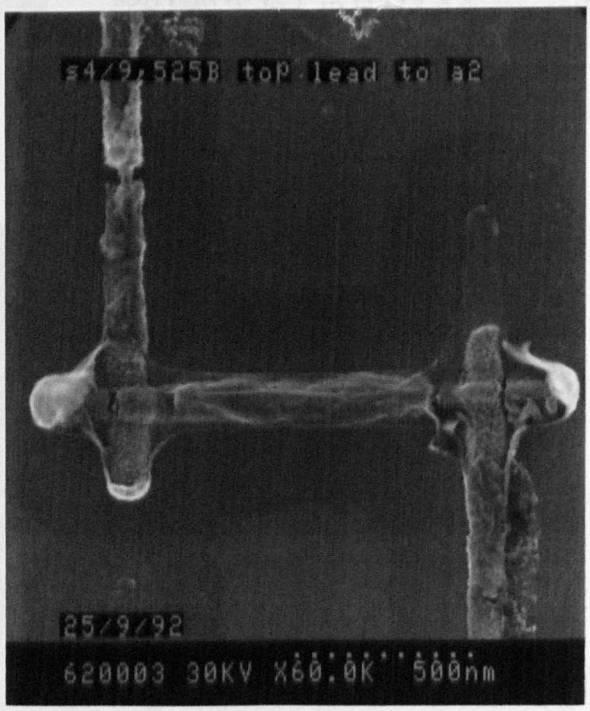
A component of this deposition is likely to arise from reflection off the evaporator bell jar, as there appears to be a directional component to the shading of the deposition in both micrographs 4.6 and 4.7. The shading of the deposited film suggests that the source of this reflection is roughly 5° above normal to the sample, since heavier deposition is experienced in the vertical (or first mode) lines of micrograph 4.7, while the horizontal second mode lines appear slightly shadowed.

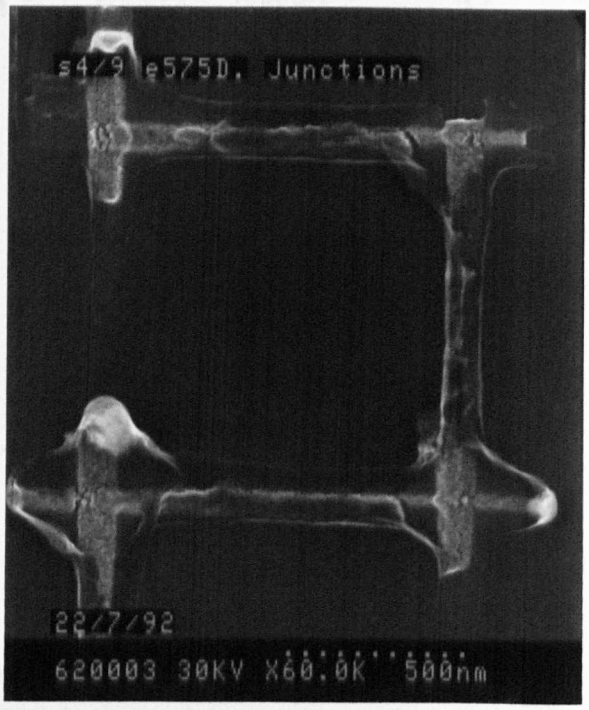
To reduce or eliminate this reflection, baffle plates were added to the evaporator set-up, as shown in place in figure 3.13. These were designed to remove this deposition. The above experiment was then repeated under the same conditions. The total direct Al deposition in this case was 70.6nm at a rate of 0.1nms^{-1}. The result after lift off is shown in micrograph 4.7. This micrograph shows, qualitatively, a thinner reflected deposit. A quantitative measurement was not possible due to the ultrathin nature of the film. Instead, the reduction in contrast between the film and the silicon substrate is taken as an indication of a reduction in the thickness of deposited material.

4.3.4.3B SURFACE MIGRATION

After the introduction of the baffle plates, this effect continued to be observed. Faithful reproductions of the resist pattern were observed after each evaporation. Micrographs

4.8 and 4.9 display the effect on samples where the baffle plates were in place.





Micrograph 4.8. & Micrograph 4.9.

Crossed track junction areas after lift off. Deposited with baffle plates present.

The processes of film nucleation, growth and surface migration are dealt with more thoroughly in section 4.4.2. Briefly, however, adatoms (in this case evaporated Al atoms) must interact with the surface upon which they are to nucleate and condense. In the process of condensing, the ad atoms can behave in a number of ways. They may re-evaporate, stick to the surface or migrate across the surface until a nucleation site is found or re-evaporation takes place. It was thought that a process of surface migration across the Si surface was leading to a greater thickness of unwanted deposition. Also, that re-evaporated adatoms were sticking to the resist and leading to a similar deposition on the resist side walls.

The solution to this problem involves cooling the substrate to ensure the probability of re-evaporation is low.

4.3.4.4 CLOSE OVER OF RESIST SLOT

During the deposition process, material is also deposited on the sidewalls of the shadowed mode. As the evaporation progresses through the deposition modes a reduction in linewidth takes place which may lead to the slot becoming completely blocked. This effect occurs even where the deposition is directly perpendicular to the substrate, see section 4.4.4.3. In this case the effect is accentuated by depositing

directly onto the sidewall of the shadowed mode. The slot is therefore narrowed by at least the same thickness as the deposited film.

The solution here is to grade the wire widths in the order of processing. The first wire deposited will, therefore, have a width equal to the wire width. The second wire requires the thickness of the previous deposition to be added to the resist slot width.

4.3.4.5 INCOMPLETE FORMATION OF TUNNEL JUNCTIONS

This problem relates to step coverage. As the first track deposited forms the bottom layer of the tunnel junction, the second track must properly cover the steps created by the underlying track. A problem exists here in that angled evaporation across a track causes shadowing by that track. This results in discontinuity of the overlapping wire, see figure 4.8.

4.3.4.6 DEPENDENCY ON ANGLE OF EVAPORATION

Changing the angle of evaporation in any angled evaporation scheme modifies the characteristics of the process. In this system the possibility of changing the angle is limited. Reduction of the evaporation angle has the benefits of reducing the degree of line extension required for correct pattern deposition, reducing the effect of reflected flux and easing the step coverage requirements. Reduced by too much, it is not possible to shadow the substrate when required. For example, a slot with a depth to width aspect ratio of 6:1 will receive substrate deposit for source angles less than ±9.5° off normal to the substrate.

Increasing the angle above 45° relaxes the aspect ratio requirements for complete shadowing of lines and reduces the depth to which the resist sidewalls are coated, simplifying lift-off processing. Increasing the angle may also reduce the effect from reflected deposition since to achieve such a high angle of incidence, a large number of collisions with the bell jar would be required. This increases the likelihood of the evaporant sticking to the bell jar. The difficulty with this situation arises from reduced film thickness on the substrate coupled with increased film thickness on the resist sidewalls, exacerbating the resist closure problem of section 4.3.4.4 and the step coverage problem of section 4.3.4.5.

4.3.5 REJECTION OF CROSSED TRACK PROCESS

Due to the difficulties experienced with this fabrication process, it was rejected as a

reliable method of forming MIM tunnel junctions in this work. The difficulties with the resist system made the patterning and the resist layer design specific to the dimensions of the tracks being deposited. Reduction of the track width used for tunnel junction formation would then require re-design and re-characterization of the process. Similarly, the development and characterization work undertaken to examine the process did not resolve the problems which arose with the process. The lack of availability of a sample holder with substrate cooling and variable substrate angle meant that some variants of the process could not be examined. Therefore, although not all the process optimizations were examined, the technique was rejected as none of the tests produced reliable defect free tunnel junctions.

4.4 PATTERNED GRANULAR FILM STRUCTURES

Granular thin films were one of the first systems in which single electron charging effects were discovered[7] and some of the highest recorded operating temperatures have involved such films[8,9]. The main reason for this is that the dominant transport mechanism in discontinuous granular films is tunnelling. In these systems small sized particles, with a very low capacitance, are coupled by this single particle transport process. Therefore the number of electrons on any particle can only be modified by one electron at a time. Two difficulties exist with the effects in these films. Firstly, the films are generally of a bulk two-dimensional nature. This, in effect, averages the film's electrical properties. Secondly, there is no way of controlling any individual current path as it is impossible to gate the structure on a scale which would significantly affect that region's electrical state. Experiments with granular films point the way to higher temperature metallic single electron effect devices but they do not provide for useful high temperature devices due to the bulk behaviour of these materials.

The scheme described here attempts to make use of granular aluminium films in a laterally patterned process to gain access to single electron devices with higher operating temperatures. The intention is that patterning a thin metal film on the scale of the natural grain size will lead to the grain growth being constrained in the narrowest regions of the pattern. It should therefore be possible to form a tunnel junction through the interface between the grain boundaries which meet in this region, as shown in figure 1.8c. Oxidation of the film, by appropriate means, should lead to the formation of a oxide barrier between these grains. Electron transport may then only take place through the grain boundaries and the islands formed by the metal grains. In this way, the film could be patterned into a controlled structure which could form a single electron device. Work carried out by MacLeod using a JEOL EBL system[10] demonstrated that evaporation of material onto a nanolithographically patterned resist/substrate system

coupled with lift off processing results in some constraint of the deposited material through the deposition kinetics. It is this effect which should be useful in the formation of single electron devices using the aluminium material system. This section of this chapter discusses this technique, giving an introduction to nucleation processes and describing the film and process characterization work.

4.4.1 PATTERNED THIN FILM STRUCTURE

The aim of this process is to form tunnel junctions by controlling the ordering and placement of grains in the thin film. Ideally we would have a system where the contact region or grain boundary between two grains would be fully oxidized. This would result in a tunnel barrier to charge transport between the grains. Control over grain growth and position is achieved by electron beam lithography. Patterning of the region of interest must take place on the scale of the natural grain size of the deposited film. Patterning in this manner leads to a degree of self organization in the film structure in the narrowest regions. Grain growth and the kinetics of film deposition have length scales associated with surface migration and the size of critical nuclei. These help to determine the organization of the structure whilst patterning controls the positioning of this essentially self organizing structure.

Initial tests aimed to form a tunnel junction by defining a pattern consisting of two opposing points written in resist. By this method, it was thought possible to utilize the effects present in electron beam exposure to vary the width of the contact and therefore the degree of control over grain growth and organization. The structure of the tunnel junction is shown schematically in figure 4.13. The diagram depicts a film patterned into opposing points by lithography, deposition and lift-off. The grain size is assumed to have some distribution around a given mean size and the patterning at the finest points is equal to that of the mean grain size.

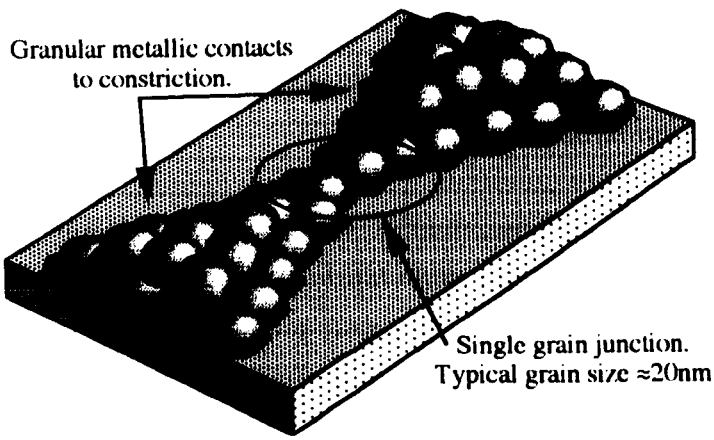


Figure 4.13. Schematic of patterned film structure.

4.4.2 NUCLEATION AND GROWTH OF THIN FILMS

Physical vapour deposition, PVD, involves presenting a target or substrate with vapour of the material to be deposited. The processes of adsorption, condensation and nucleation result in the formation of a solid phase of the deposit material on the substrate. Both evaporation and sputtering are forms of PVD. Evaporation is the mechanism used for deposition of the aluminium device layer and the interconnect layers, as discussed in sections 3.4 and 3.5. Both of these layer types are thin metallic films and as such their development is controlled by the processes of thin film nucleation and growth. Film microstructure and physical characteristics are determined by the growth mechanisms which can and do differ greatly depending on the materials involved. In this work, the greatest difference is observed in the microstructure of the aluminium films compared with the interconnect films. SEM investigations of these films display a definite granularity in the aluminium which is not present in the interconnect. It is these differences which are utilized in the technique developed here, and as such require some explanation of their roots.

Development of theories for the nucleation and growth of thin films has been hampered by the difficulty in measuring some or most of the parameters involved in the models. The theories themselves split into two main approaches, the Atomistic model and the Capillarity model. The atomistic approach discusses the formulation of nuclei in terms of adsorption and diffusion energies of deposited atoms. However, it is difficult to determine the parameters for this model. The Capillarity model treats the problem using macroscopic concepts such as surface energies and contact angles. The difficulty here lies in nuclei of small critical size, where only a small number of atoms constitute the nucleus. The application of macroscopic quantities to describe such a system is clearly dubious. A good introductory review of these theories is given by C.A. Neugebauer in the Handbook of Thin-Film Technology[11]. It is not the intention here to review the theories and applicability of the models but to describe the general processes involved in thin film formation and draw on that information to describe the formation of these patterned thin film structures.

Four clear stages exist in film formation[11], these are: 1. island stage, where initial adsorption and surface coverage takes place leading to nucleation, 2. coalescence stage, where nucleated material begins to join together producing larger islands, 3. channel stage, where surface coverage is almost complete but secondary nucleation and coalescence are filling the gaps in the film, and, 4. continuous film, where surface coverage by the film is complete and film growth continues. Outlines of these

processes are given in the following four sections.

4.4.2.1 ARRIVAL AND ADSORPTION

The initial step in film growth is arrival of the evaporant molecule at the substrate surface. Evaporation of a material at reduced pressures, $\leq 10^{-5}$mBar, produces long mean free paths, in excess of 50m, resulting in very little interaction between evaporated molecules. The arrival rate of nuclei from the source is calculated from the kinetic theory of gases and is given by equation 4.13. The pressures used are taken as the partial pressures of the component of interest.

$$R = (2\pi m k_B T)^{-1/2} P$$
$$= 2.6\text{x}10^{22}(MT)^{-1/2}\,\text{P} \quad (\text{m}^{-2}\text{s}^{-1}) \qquad [4.13]$$

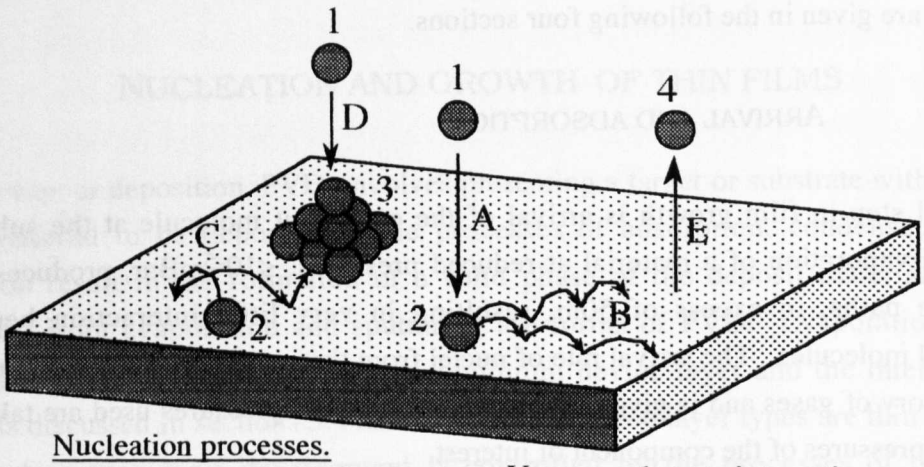
where, m = molecule mass, P = vapour pressure,
M = molecule mass /amu., P = vapour pressure /mBar,
T = vapour temperature /K.

The molecules generally arrive at the surface with energy distributions characteristic of their source temperature. This will normally be much higher than the energies associated with the substrate where metals are evaporated. The arriving molecule may either adsorb and permanently stick, adsorb and re-evaporate in a finite time or immediately reflect off the substrate. The impingent molecule must equilibrate with the substrate in order to remain adsorbed. Some of these processes are depicted in figure 4.14. Given these conditions a thermal accommodation coefficient and a sticking, or condensation coefficient can be defined for the arriving molecules as;

$$\text{Thermal accommodation coefficient} : \alpha_T = \frac{E_i - E_r}{E_i - E_s} = \frac{T_i - T_r}{T_i - T_s} \qquad [4.14]$$

E_i=Energy of incident molecule, E_r=Energy of re-evaporated molecule,
E_s=Molecule energy after equilibration with substrate,
T_i, T_r, T_s=Respective temperatures.

$$\text{Sticking coefficient} : S_C = \frac{\text{No. of molecules adhering to substrate}}{\text{No. of molecules arriving at substrate}} \qquad [4.15]$$



Nucleation processes.

A - Adsorption.
B - Surface diffusion.
C - Nuclei growth by surface diffusion.
D - Nuclei growth by evaporant arrival.
E - Desorption.

Vapour-substrate interactions.

1 - Evaporant arrival.
2 - Adsorbed evaporant.
3 - Embryonic nucleus.
4 - Desorbed evaporant.

Figure 4.14. Vapour substrate interaction[12].

α_T is a measure of the extent to which arriving molecules reach thermal equilibrium with the substrate and lose enough energy to prevent re-evaporation from the surface, while S_c is a measure of the sticking efficiency of the vapour - surface interaction. If all atoms are adsorbed and accommodated, α_T=1, but S_c will not necessarily be unity. An equilibrium process can exist between adsorbed molecules and the vapour whereby each adsorbed molecule remains on the surface for an average time, τ_a, defined by,

$$\tau_a = \frac{1}{\upsilon_o} exp\left(\frac{\Delta G_{des}}{k_B T}\right) \qquad [4.16]$$

where υ_o is the frequency of vibration of the desorbing molecule and ΔG_{des} is the free activation energy of the desorption process. If condensation of the adsorbed molecule does not occur, complete re-evaporation will eventually occur and S_c will be zero despite $\alpha_T \approx 1$. Measurements of the sticking coefficient have found that where clean surfaces are used S_c can be initially very low, but as surface coverage progresses the value of S_c rapidly approaches unity.

Some form of adsorption must take place at the substrate surface to allow the evaporant to condense. Adsorption takes two forms, physical and chemical adsorption. Physical adsorption or physisorption occurs between all solids and arises from weak interactions, such as van der Waals forces, electrostatic forces, etc. No alteration of the adsorbed molecule takes place. That is, no chemical bond is formed with the substrate and no change in the chemical state of the molecule takes place. The physisorption process does not depend on the chemical nature of the substrate and as such has no associated

activation energy. The process normally occurs at low temperatures, generally around the boiling point of the adsorbate. The enthalpy of adsorption (heat released when 1 mole of adsorbate changes from the gaseous to adsorbed states) is low[13], $-\Delta H^{\circ}_{ads} \leq 20 kJmol^{-1}$.

Chemisorption provides for the formation of new chemical bonds between the adsorbate and the substrate. The bond energies are similar to those found in molecules, giving $-\Delta H^{\circ}_{ads} \approx 200 - 500 kJmol^{-1}$ (of the order of chemical reaction energies). Chemisorption requires dissociation of the adsorbed molecule to form the bond, therefore O_2, for example, must split into its monomer state before it can be chemisorbed. The activation energy for this process is usually small, but not zero, whilst the temperature of occurrence depends on the activation energy but again is usually low. Due to the nature of the process, chemisorption can only allow for single monolayer coverage of a surface if this is the only active process[13].

Physisorption can provide a lower energy path to chemisorption by bringing a molecule close to the surface at low potential energy. The adsorbed molecule may then enter a transition state whereby the molecule transfers from the physically adsorbed state to the chemically adsorbed state.

4.4.2.2 NUCLEATION AND COALESCENCE

The model of nucleation suggests that across a surface adsorbed molecules must aggregate to prevent re-evaporation. This forms clusters of various sizes which must continue to grow to avoid decomposition by surface migration. Growth of these clusters deplete the surrounding area of adsorbed molecules in what is termed a capture zone. The radius of this zone is defined as in equation 4.17, where D is the coefficient of surface diffusion.

$$r_d = \sqrt{D\tau_a} \qquad [4.17]$$

The Capillarity model suggests that a critical size of cluster exists which is associated with the free energy of formation. This model describes a function for this energy which goes through a maximum at the critical size or radius, as shown in figure 4.15. Clusters or nuclei below this size must grow by incorporating more molecules or they will decompose through surface diffusion. Formation of critical sized nuclei is termed nucleation and it is at this point that the film structure will remain if the evaporant flux is removed. Growth of these nuclei will continue to super-critical dimensions allowing further decrease of their free energy and increase of their stability.

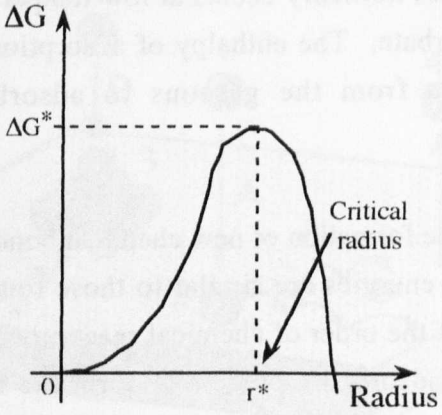


Figure 4.15. Nuclei free energy as a function of size.

As the nuclei grow, they will eventually touch and coalesce, forming a new island and reducing the nuclei surface area and free energy. This exposes some of the substrate surface, allowing secondary adsorption and nucleation to continue covering the surface. The process is shown schematically in figure 4.16.

Figure 4.16. Coalescence of nuclei to form a larger island.

4.4.2.3 CHANNEL STAGE

As coalescence continues at some point the increased film coverage results in small areas of exposed substrate. Coalescence, at this stage, does not produce the shape changes which form the rounded islands of the previous stage. Channels of exposed substrate therefore exist and are subject to secondary adsorption and nucleation. Gradually, the channels are filled and the entire surface becomes covered with a continuous film.

4.4.2.4 CONTINUOUS FILM

The film created in this manner is similar to an epitaxial film in the liquid like manner of its formation. However, it is polycrystalline islands which coalesce and these have a random distribution of crystallographic orientations. Some recrystallization can take place resulting in grain sizes larger than the separation of the initial nuclei. Defects are incorporated in these films by the presence of grain boundaries. As soon as islands

touch, a grain boundary is created which will include defects, unless complete recrystallization can take place to produce a single grain. Defects occur from a number of sources, which can include lattice orientation mismatch, differences in lattice parameters between the film and the substrate, stress producing dislocations at the edges of holes, dislocations propagated from the substrate into the film and contamination of the film grain boundaries.

The last defect source mentioned above is of interest in this work. As the aluminium films grows, contamination from the background gas will readily adsorb and react with the aluminium surface. Particularly if the contamination is oxygen, then an oxide should be formed at the grain boundary

4.4.2.5 IMPLICATIONS FOR GRANULAR DEVICES

The natural grain size of a film is governed by thermodynamic processes, therefore free energy of the structure must be minimised. The initial clusters may be much smaller than the final mean grain size, but as they are not stable any film will have a characteristic minimum grain size which cannot be smaller than the critical size. Forcing grain growth into a constrained patterned structure confines the thermodynamics of the growth process to limit the number of grains in the constriction or change their size, to minimize the grain free energy. This process should help form the single grain contact.

Another feature of constraining the film growth is forcing the presence of a grain boundary at the narrowest region. In this instance, the contact between two grains is constrained and coalescence of the grains either side of the constriction is energetically unfavourable. The grain boundary is a high stress region of the crystallite and therefore has a high free energy state. This should reduce the activation energy for oxidation and help relieve the stress on the boundary. Oxidation may be accomplished through a physisorption - chemisorption route, by inclusion of O_2 at the grain boundary as a contamination defect, or by inclusion of O_2 through its lack of solubility in Al[14]. Under this scheme, the result would be a naturally occurring and fully oxidized grain boundary.

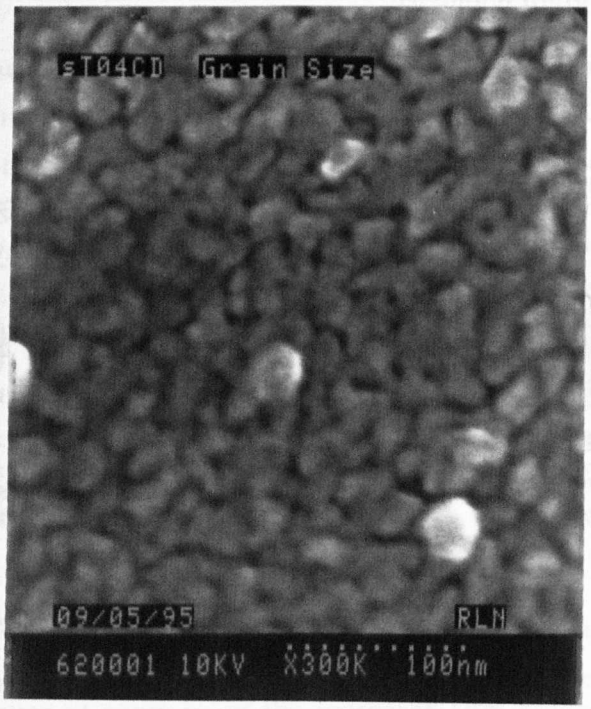
4.4.3 CHARACTERIZATION OF FILM STRUCTURE

In order to pattern the film on the scale of the grain size, knowledge of that size is required. Two means of gaining that information were available, scanning and transmission electron microscopy, SEM and TEM respectively. The three following

sections describe the work carried out using these techniques and present the results related to the film microstructure.

4.4.3.1 SEM INVESTIGATION OF BULK FILM STRUCTURE

SEM measurements of the film microstructure were the simplest to carry out, since micrographs of the films were take as a matter of routine during device fabrication. These measurements were taken using an Hitachi S900 field emission SEM. Micrographs of the film textures were taken at magnifications between 100k and 300k with accelerating voltages of 10keV.



Micrograph 4.10. Example of SEM grain measurement micrograph.

The accelerating voltage of this microscope can be easily varied between 6keV and 30keV. Lower accelerating voltages provide more information about the surface, since electron penetration of the sample and subsequent secondary electron yield from deep within the sample is lower. However, lowering the voltage too far results in a loss of resolution. The 10keV setting provides an optimum between these two effects. A typical example of a micrograph used for this analysis is shown in micrograph 4.10. Table 4.3 lists the micrographs used for the analysis and the errors incurred in the measurement.

Sample	Micrograph No.	Micrograph Date	Micrograph Magnification.	Errors incurred in measurement.
B3D	620012	16/1/95	150k	±1.1nm
A21	620001	19/1/95	300k	±0.6nm
B2AA	620000	16/2/95	300k	±1.6nm
B2AC	620000 & 620001	6/3/95	300k & 300k	±0.6nm
T04AJ	620000 & 620001	28/4/95	100k & 300k	±1.6 & ±0.6nm
T04CD	620000 & 620001	9/5/95	100k & 300k	±1.6 & ±0.6nm

Table 4.3. Micrographs used for grain size measurements.

From the micrographs obtained, as above, sample measurements of the grain size were taken as follows. The micrographs were scanned, using a flat-bed scanner, at 150 dots per inch horizontal resolution at appropriate contrast and brightness levels. The scanned image was printed and grain boundaries marked in by hand. A diameter was then measured and scaled to the micrograph's scale bar, as printed. This gave an effective magnification of 2.1 times the original image. This process is not accurate in any metrological sense as it ignores magnification errors in the SEM camera system of around 5%. However, the process gives a reasonable indication of the film grain sizes. Measurement of the grain diameter was taken to an accuracy of ±1/4mm. The assumptions made were that the grains were hemispherical and that the full grain diameter could be detected in the contrast variations within the micrograph.

As it was expected that the deposition conditions would have a bearing on the average grain size and film structure, the relevant deposition conditions are detailed in table 4.4. From the variations of the deposition conditions, it is possible to compare grain sizes obtained in a high vacuum evaporator (sample A21) and those obtained in the UHV evaporator. Comparison can also be made between those UHV films deposited using a partial pressure of oxygen and those where no oxygen partial pressure was used.

The grain size distribution may not be Gaussian as nucleation kinetics will force a lower limit on the grain size which may result in asymmetric distributions. However, in order to analyze and compare the distributions, mean and standard deviation figures have been calculated from the results. These are presented in table 4.5. The distribution for sample B2AC is shown in figure 4.18.

Sample	P_B /mbar	P_R /mbar	O_2 P_P /mbar	Rate /nms^{-1}	Film thickness /nm
B3D	2.3 x 10^{-8}	1 x 10^{-6}	---	0.2 - 0.3	34.5
A21†	2.1 x 10^{-6}	3.7 x 10^{-6}	---	≈0.2	≈20
B2AA	8.7 x 10^{-8}	1.6 x 10^{-6}	---	0.1 - 0.2	25
B2AC	8.7 x 10^{-8}	1.6 x 10^{-6}	---	0.1 - 0.2	25
T04AJ	3.3 x 10^{-8}	2.2 x 10^{-6}	2.2 x 10^{-6}	0.1 - 0.2	25
T04CD	2.3 x 10^{-8}	2.3 x 10^{-6}	2.3 x 10^{-6}	0.1 - 0.2	20

† - film deposited in the modified evaporator.

Table 4.4. Evaporation data for SEM grain measurement samples.

Sample	Micrograph ID.	Number of grains in sample.	Mean diameter. /nm	Standard deviation. /nm	Minimum diameter. /nm	Maximum diameter. /nm
B3D	620012	100	25.6	6.8	13.5	56.8
A21†	620001	101	15.8	3.3	10.1	28.6
B2AA	620000	67	14.6	2.3	10.1	21.4
B2AC	620000 620001	120	18.6	4.0	9.7	34.1
T04AJ	620000	100	28.8	7.9	12.6	60
	620001	60	19.3	4.8	12.2	33.7
T04CD	620000	100	24.4	7.2	13.9	47.4
	620001	81	17.1	4.6	9.3	29.5

† - film deposited in the modified evaporator.

Table 4.5. Summary statistics of grain size, SEM examination.

From the mean grain sizes presented in table 4.5, we see an average grain size of between 15 and 20nm. Those micrographs which yield larger mean grain sizes are also those with lower magnifications, cf. table 4.3, and therefore a greater measurement error. It should also be noted that on the basis of these measurements, there is no discernible difference between the high vacuum film and the UHV films. There is also little difference between the UHV films which did and did not have a O_2 partial pressure during the evaporation, suggesting similar deposition conditions at the pressures used here.

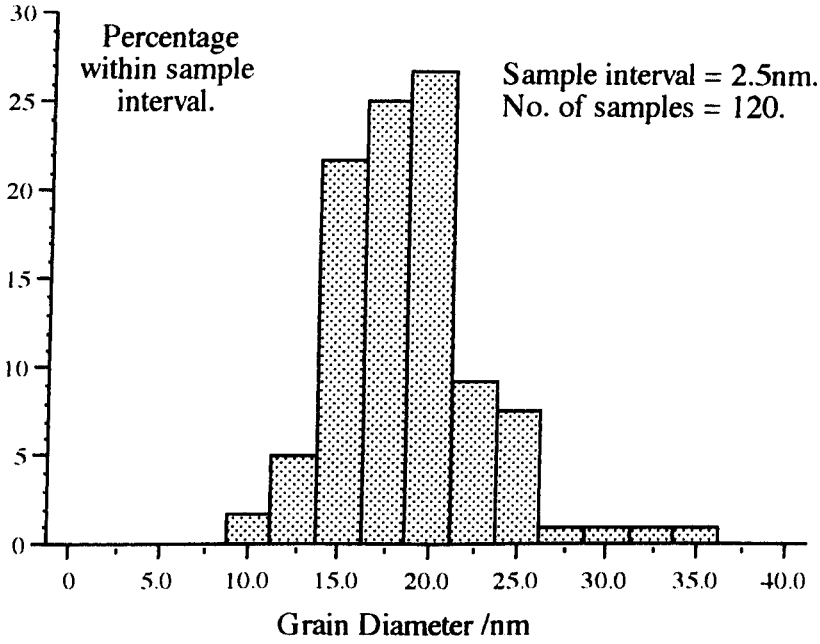


Figure 4.18. Histogram of grain diameter measurements for sample B2AC.

4.4.3.2 TEM INVESTIGATION OF BULK FILM STRUCTURE

A later set of measurements were taken to investigate the film structure, using a JEOL 1200EX TEM. These measurements were intended to yield more information about the microstructure of individual grains and also a clearer measurement of the film grain size. As transmission through the sample is required, the aluminium films were deposited onto specially prepared membrane samples. These samples contained a central 100μm x 100μm area membrane formed from 50nm thick amorphous silicon nitride, Si_3N_4, supported on etched Si substrates. The full sample dimensions were 2mm x 2mm, defined by etched cleave lines stopping around 50μm from the sample surface. Preparation of these samples and operation of the TEM were carried out by Dr. B. Khamsepour. The samples were mounted in the TEM holder with the membrane face down, using a small washer to help fix the sample in position.

Both bright field (normal image) and electron diffraction images were obtained from these samples using the TEM. The bright field images are discussed here, the electron diffraction patterns are discussed in section 4.4.3.3. Bright field images were taken using a 20μm objective aperture. The aperture was positioned in diffraction mode to ensure the main transmitted beam was central to the aperture. The microscope was then switched to bright field viewing mode for observation the film microstructure.

Table 4.6 details the evaporation data for the films deposited onto the membrane samples. All evaporations took place in the UHV evaporator, at normal angle of

incidence. Table 4.6 presents the base pressure, P_B, the run pressure, P_R, and the partial pressure of O_2 used, O_2 P_P, in the evaporator. The evaporation rate and the nominal film thickness, as measured directly on the crystal monitor, are also given.

Sample	P_B /mbar	P_R /mbar	O_2 P_P /mbar	Rate /nms^{-1}	Film thickness /nm
M08	2.3 x 10^{-8}	2.3 x 10^{-6}	2.2 x 10^{-6}	0.1	25
M09	2.3 x 10^{-8}	2.3 x 10^{-6}	2.2 x 10^{-6}	0.1	25
M10	2.5 x 10^{-8}	2.3 x 10^{-6}	2.3 x 10^{-6}	0.1 - 0.2	15
M11	2.5 x 10^{-8}	2.3 x 10^{-6}	2.3 x 10^{-6}	0.1 - 0.2	15
M12	2.3 x 10^{-8}	2.3 x 10^{-6}	2.3 x 10^{-6}	0.1 - 0.2	20
M13	2.3 x 10^{-8}	2.3 x 10^{-6}	2.3 x 10^{-6}	0.1 - 0.2	20

Table 4.6. Evaporation data for membrane samples.

The magnification used for examining each film was 120k times. Table 4.7 presents the data extracted from the TEM bright field micrographs of the microstructure. This data was extracted in a similar manner to the SEM micrograph data. The major difference between the TEM and SEM examination was that the grain structure was revealed to be more geometrical than spherical, as would appear to be the case from the SEM micrographs. Comparing the results from the TEM and SEM investigations it is possible to conclude that the grain size distribution is around 18 to 20nm in diameter with a typical 3 to 6nm standard deviation.

Sample	Mean diameter. /nm	Standard deviation. /nm	Minimum diameter. /nm	Maximum diameter. /nm
M08	18	6	9	36
M09	18	3	11	25
M10	18	3	13	25
M11	18	3	13	25
M12	17	3	11	25
M13	16	4	11	29

Sample size in each case was 50 grains.

Table 4.7. Summary statistics of grain size, TEM examination.

Figure 4.19 presents a histogram of the data from the TEM samples, pooled from all six samples.

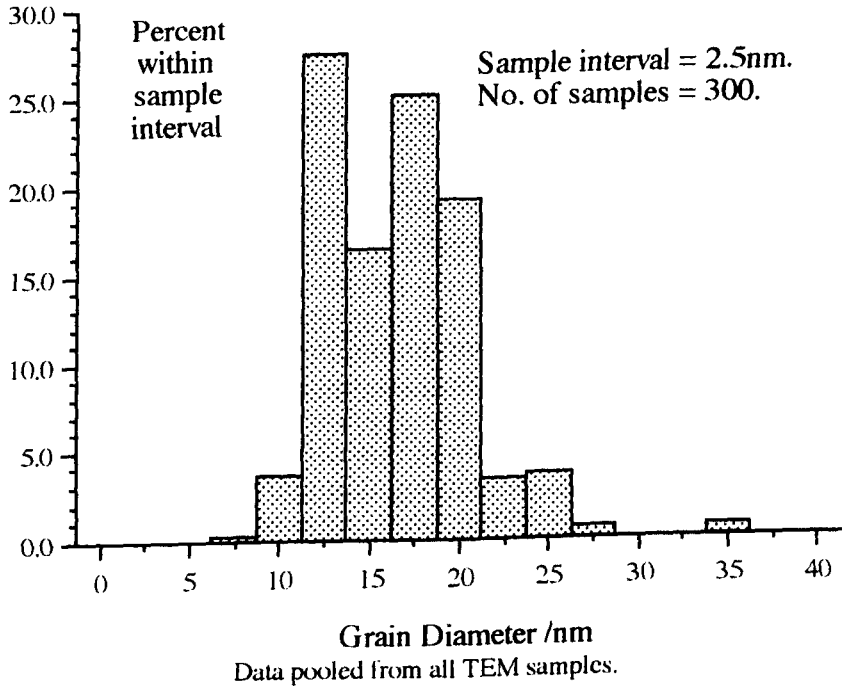


Figure 4.19. Histogram of grain diameters from TEM samples

4.4.3.3 TEM INVESTIGATION OF FILM CRYSTAL STRUCTURE

While carrying out bright field transmission measurements, it was a simple matter to view the film structure in electron diffraction mode. As the same samples were used, sample preparation and evaporation is the same as discussed in the previous section. The TEM was operated at the same accelerating voltage as for the bright field images. However, the 20μm objective aperture was removed to view the diffraction pattern. Before presenting the results, a brief introduction to electron diffraction is required.

4.4.3.3A INTRODUCTION TO ELECTRON DIFFRACTION

Any crystal structure can be described in terms of a co-ordinate system made up from an orthogonal set of basis vectors. This system is designed to form a repeating unit cell from which the entire lattice can be constructed. In the case of cubic, tetragonal and orthorhombic crystals a triplet set of basis vectors can be used: **a**, **b**, **c**. Planes in a crystal are normally defined in terms of Miller indices. These are formed by taking reciprocals of the plane's intercepts with the vectors **a**, **b**, **c** and reducing these numbers to the smallest set of integers h, k & l having the same ratio. Directions in a crystal are then defined by $|hkl|$ while planes are denoted (hkl). Families of planes eg. (100), (010), (001) etc., that is planes of the same order but not exactly the same orientation are denoted $\{hkl\}$. The interplanar distance, d, can be calculated as follows;

$$\frac{1}{d^2} = \frac{h^2}{a^2} + \frac{k^2}{b^2} + \frac{l^2}{c^2} \qquad [4.18]$$

or for cubic crystals where $c = b = a = a_o$, the lattice constant,

$$d = \frac{a_o}{\left(h^2 + k^2 + l^2\right)^{1/2}} \qquad [4.19]$$

Diffraction of coherent radiation in a three dimensional array of atoms is governed by Bragg's law, stating that an incident beam of monochromatic radiation is reflected by a set of lattice planes $\{hkl\}$ if,

$$2\,dSin\,\theta = l \qquad [4.20]$$

where θ is the angle between incident beam and the diffracting planes, $\{hkl\}$, and λ is the wavelength of the diffracted radiation. Where an electron beam is transmitted through a sample, portions of the beam are diffracted when Bragg's law is satisfied. If this diffraction is imaged on a plate a distance, L (the camera length), from the sample, the diffraction angle is given by $tan2\theta=r/L$, where r is the distance from the spot created by the directly transmitted beam to the spot or ring caused by the diffraction. Spots appear in a diffraction pattern when a single crystal is viewed. Polycrystalline materials produce a ring pattern, termed Deybe rings, due to the random relative orientations of each of the crystallites. Equating $tan2\theta=r/L$ with $2Sin\theta$ of Bragg's law gives the relation:

$$rd = \lambda L \qquad [4.21]$$

which with substitution of equation 4.19 results in:

$$r = \left(h^2 + k^2 + l^2\right)^{1/2} \frac{\lambda L}{a_o} \qquad [4.22]$$

For accelerated electrons , the wavelength is calculated from equation 4.23, where m is the electron rest mass, e is the electronic charge, V is the electron accelerating voltage, c is the speed of light and h is Plank's constant.

$$\lambda = \frac{\mathrm{h}}{\left[2meV\left(1 + \frac{eV}{2mc^2}\right)\right]^{1/2}} \qquad [4.23]$$

which yields $\lambda = 3.704$pm for 100kV accelerating voltage. If $(h^2+k^2+l^2)$ in equation 4.5 is replaced with, n, a plane index, we find that given crystal structures have a given sequence of numbers for n which cause diffraction. For a body centred cubic (BCC)

lattice, this sequence is[15] n = {2, 4, 6, 8, 10, 12, 14, 16, 18, 20,...} while for a face centred cubic (FCC) lattice[15] n = {3, 4, 8, 11, 12, 16, 19, 20, 21,...}. Hence, for an FCC structure, the families of planes which cause diffraction are {111}, {200}, {220}, {311}, {222}, {400}, {331}, {420}, {421},.... Therefore, with knowledge of the lattice constant, the crystalline structure of the film can be determined. Alternatively, knowledge of the structure of the film yields the lattice constant of the material.

4.4.3.3B ELECTRON DIFFRACTION RESULTS FOR ALUMINIUM FILMS

Viewing the deposited Al films in diffraction mode yielded a Deybe ring pattern, indicative of an FCC structure, confirming the creation of a polycrystalline FCC film from the Al evaporation. Spotting was also observed on these rings, due to the small number of grains viewed using the electron beam. Two effects cause this spotting. The first is due to the small spot size used, limiting the number of grains sampled by the spot. The second is the thickness of the film. These films are envisaged to be only one or two grain layers thick and therefore limiting of the number of grains sampled again. Examples of the electron diffraction patterns obtained are given in micrographs 4.11 and 4.12.

Sample	d_1 /mm	d_2 /mm	d_3 /mm	d_4 /mm	d_5 /mm	d_6 /mm	d_7 /mm
M08	22.75	26.25	37.00	43.75	45.00	57.50†	58.50†
M09	22.25	26.00	36.75	43.25	45.00	-----	-----
M10	22.50	26.00	37.00	43.00	-----	-----	-----
M11	22.50	26.25	36.75	43.25	-----	57.50†	-----
M12	22.50	26.00	37.00	43.00	44.75	-----	-----
M13	22.50	26.00	36.75	43.25	-----	-----	-----

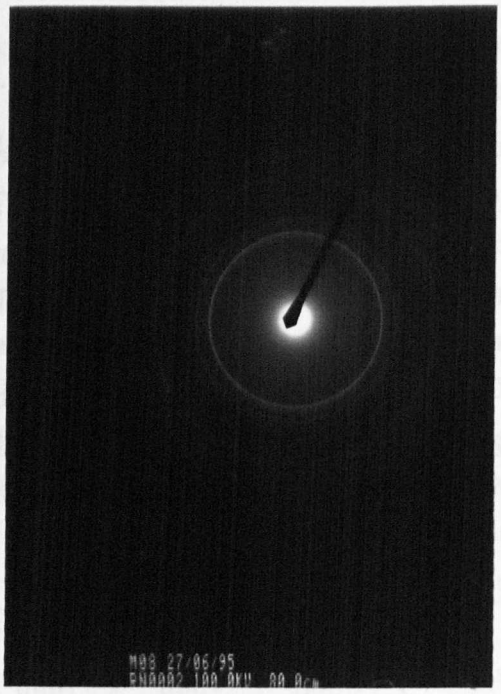
† - diameter estimate from a feint Deybe ring.

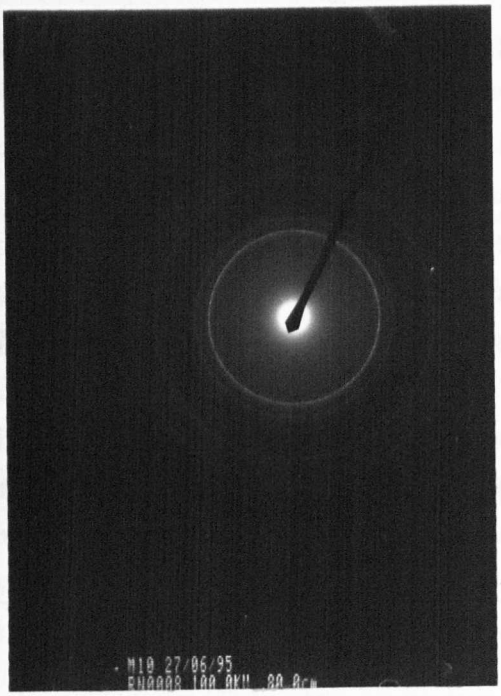
Table 4.8. Measurements of ring diameters from TEM diffraction patterns.

Direct observation of the patterns in the TEM revealed at least eight rings. However, as the plane index rises the ring intensity decreases and it becomes more difficult to measure the ring diameter. The accelerating voltage used for these measurements was 100keV and the camera length, L, was 80cm. The Deybe ring diameters were measured, from the micrographs, to the nearest 1/4mm and are given in table 4.8. No single crystal Al samples were used to calibrate the system against quoted lattice constants[16] due to the difficulty in preparing such samples. However, the manufacturers quoted absolute accuracy of the TEM & camera system is around 5%

with a worst case error of not more than 10%. A 5% error is assumed here.

As can be seen from table 4.8, all films have similar values for each of the ring diameters. This suggests similar lattice constants for the crystallites. Micrographs 4.11 and 4.12 display the diffraction patterns for two of these samples, M08 and M10 respectively.





Micrograph 4.11 & Micrograph 4.12.
Deybe rings diffraction patterns for samples M08 and M10 respectively.

Sample	$a_o(d_1)$ /Å	$a_o(d_2)$ /Å	$a_o(d_3)$ /Å	$a_o(d_4)$ /Å	$a_o(d_5)$ /Å	$a_o(d_6)$ /Å	$a_o(d_7)$ /Å
M08	4.51±0.28	4.52±0.27	4.53±0.26	4.49±0.25	4.56±0.25	4.12±0.22†	4.49±0.24†
M09	4.56±0.28	4.56±0.27	4.56±0.26	4.54±0.25	4.56±0.25	-----	-----
M10	4.65±0.28	4.56±0.27	4.53±0.26	4.57±0.26	-----	-----	-----
M11	4.65±0.28	4.52±0.27	4.56±0.26	4.54±0.25	-----	4.12±0.22†	-----
M12	4.65±0.28	4.56±0.27	4.53±0.26	4.57±0.26	4.58±0.25	-----	-----
M13	4.65±0.28	4.56±0.27	4.56±0.26	4.54±0.25	-----	-----	-----

† - calculated from estimated value.

Table 4.9. Lattice constants resulting from data of table 4.8.

The main differences in the films result from the thickness variations. The Deybe rings for M08 are more uniform than those of M10. This is a direct result of the difference in aluminium thickness for the two samples; nominally 25nm for M08 and 15nm for M10.

Table 4.9 presents the analysis of the ring data, determining the lattice constant, a_o. The values of a_o were calculated on the basis of an FCC structure, as indicated by the ring sequence and the match of the results.

The mean value for a_o=4.56±0.26Å, not counting the estimate values for the d_6 and d_7 rings. This value is larger than the lattice constant for bulk aluminium[16]. The most common naturally occurring phase of aluminium is αAl, having an FCC structure with lattice constant a_o=4.049Å at 25°C. If we take the mean values, these results are showing a 12.6±5% increase in the average lattice constant. Even if we assume manufacturers worst case error, the films still show and increase, albeit small.

A number of causes can be attributed to an increased lattice constant, stress in the films due to both film thickness and interaction with the substrate, incorporation of tungsten from the evaporation boat, incorporation (although unlikely for these films) of oxygen within the grains. O_2 incorporation is unlikely because there is no significant solid or liquid solubility of oxygen in Al[14] forcing incorporated oxygen to the grain surface during film growth where an oxide will form. It is possible, however, that such an oxide could contribute to an increased lattice constant. α-Al_2O_3 or corundum is the most common naturally occurring crystalline form of aluminium oxide, and has a rhombohedral structure (lattice constants a = 4.758Å & c = 12.991Å). However, any significant presence of such an oxide would be reflected in the diffraction pattern. It has been shown, however, that ultra-thin films of aluminium exhibit a volume increase the thinner the film becomes[17], which would explain the increase in lattice constant. Also, the fact that the Deybe rings are preserved in the FCC order shows that the main film material is aluminium in a polycrystalline state, with little effect on the microstructure by either incorporation of oxygen or other materials. It is therefore likely that the films are largely metallic in nature rather than discontinuous. However, this requires electrical measurements of the film transport properties to determine.

4.4.4 CHARACTERIZATION OF PATTERN AND STRUCTURE DEVELOPMENT

A number of experiments were carried out to determine the manner in which the opposing points structure develops, the limits to structure size and the possibilities for manipulating individual grains. All of the lithography work used the high resolution modes of the EBPG-5HR Beamwriter, at both 50 and 100keV accelerating voltages. The experiments and results are presented in the following sections.

4.4.4.1 PATTERNING OF OPPOSING POINTS

To determine the possibilities of patterning at or below the 20nm mean grain size, a series of exposure tests were undertaken. The tests were designed to vary the angle of the points and the gap between them. The aim of these variations was to observe the possibility of controlling the size of the connecting wire repeatably and achieve dimensions lower than the direct write dimensions by use of the proximity effect. Figure 4.20 illustrates the desired patterning effect. The results of these tests are described in Section 4.4.4.1a. The dependence of the grain structure in the constriction was also observed using Al metallization of the pattern and is discussed in section 4.4.4.1b.

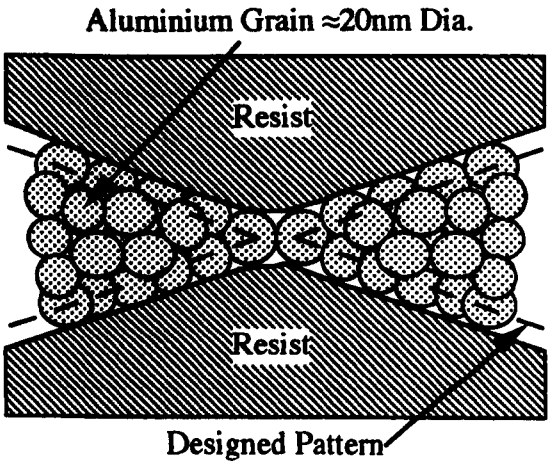


Figure 4.20. Opposing points patterning.

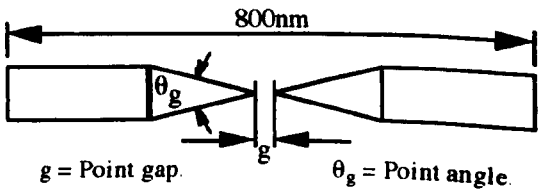


Figure 4.21. Opposing point variables.

4.4.4.1A LITHOGRAPHIC STRUCTURE DEVELOPMENT

Sample	Lithography Process†	Beam voltage	Metallization
B7	EBL-HR1	100kV	35-40nm Al
B8	EBL-HR1	100kV	35-40nm Al
B9	EBL-HR1	100kV	2x(10:10nm Ti:Au)
B10	EBL-HR1	100kV	2x(10:10nm Ti:Au)
B11	EBL-HR1	50kV	50nm Al
B12	EBL-HR1	50kV	2x(10:10nm Ti:Au)

† - see Section 3.8.

Table 4.10. Opposing point test sample data.

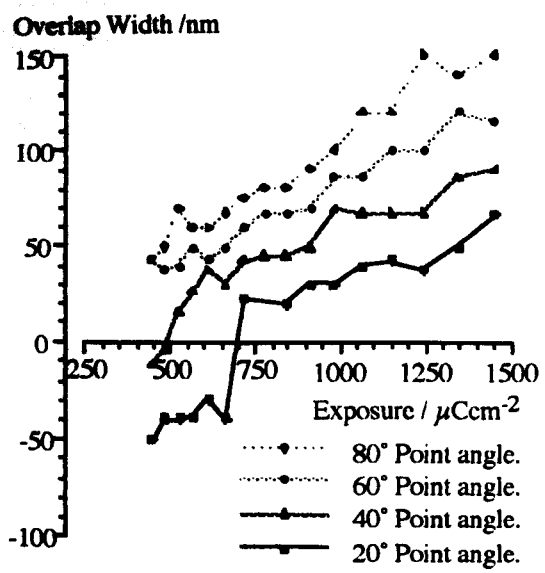
Using a beam spot size of 12nm diameter through a final aperture of 200μm, with a pattern resolution and step size of 5nm, the opposing point pattern of figure 4.21 was written as an exposure test. The gap variation was 5 to 25nm in 5nm steps and the angle variation was 20° to 110° in 10° steps. Doses from 450 to 1457μCcm^{-2} were used in a 16 step geometric series with 1.08148 increment. Accelerating voltages of 50 and 100kV were tested. Table 4.10 details the sample processing information. The terms point angle and point gap refer to the values assigned to the designed pattern and are not the same as those determined from the exposed and developed pattern.

Samples were initially metallized using aluminium, to observe the grain structure development but lack of contrast between Si and Al in SEM imaging of the pattern led to the use of Ti:Au metallization instead. Measurements were taken directly from the screen of the Hitachi S900 SEM. The measured dimension was compared with the scale bar presented on the screen, resulting in an error of 10 to 15% on any measurement. Measurements were taken at the highest magnification possible to allow good comparison of the structure with the scale bar.

It was realized that increasing the gap between the points does not aid reliable reduction of the width of the constriction, especially at 100kV beam voltage. This led to the conclusion that use of the proximity effect could not easily help reduce the critical pattern dimensions. Comparison of the results for a 5nm point gap, at 20°, 40°, 60° and 80° point angles for both 50 and 100kV accelerating voltages are shown in figures 4.22 and 4.23. These graphs are presented in terms of an overlap width vs. exposure. The overlap width is a measurement of the gap or wire dimensions produced by the exposure conditions. If the pattern receives too low an exposure, no connecting wire is formed in the resist constriction and a gap results. The width from point to point of this gap is measured and presented as a negative overlap width. Positive values of this width translate directly to the minimum wire width produced by a properly exposed and developed resist constriction.

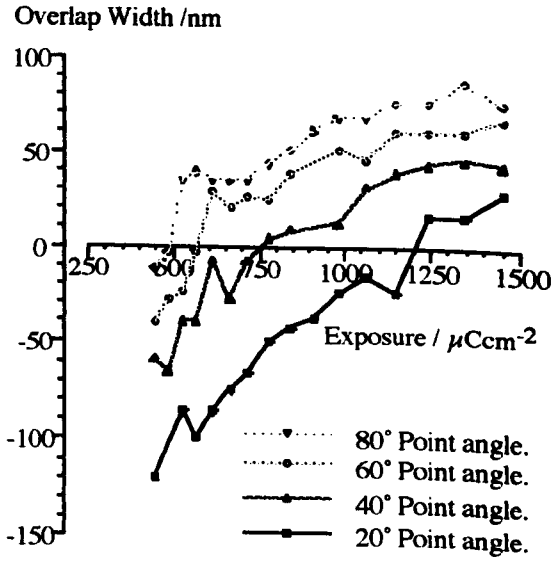
The samples represented in figures 4.22 & 4.23 were both prepared using the Ti:Au metallization to allow direct comparison. Two major differences occur between the 50 and 100kV tests. Firstly, the exposure required to produce an overlap in the 100kV test is roughly 1.8 times that required in the 50kV test. This feature is interpreted as a result of the reduction in exposure from backscattered electrons due to the greater backscatter radius in the 100kV tests. Secondly, the 100kV test has demonstrated a consistently smaller overlap width than those present in the 50kV test. Another feature of these graphs is the presence of a random variation in the overlap width. A trend in the behaviour with exposure occurs but the variation on this trend is sometimes significant,

particularly when the overlap width is in the ±20nm range.



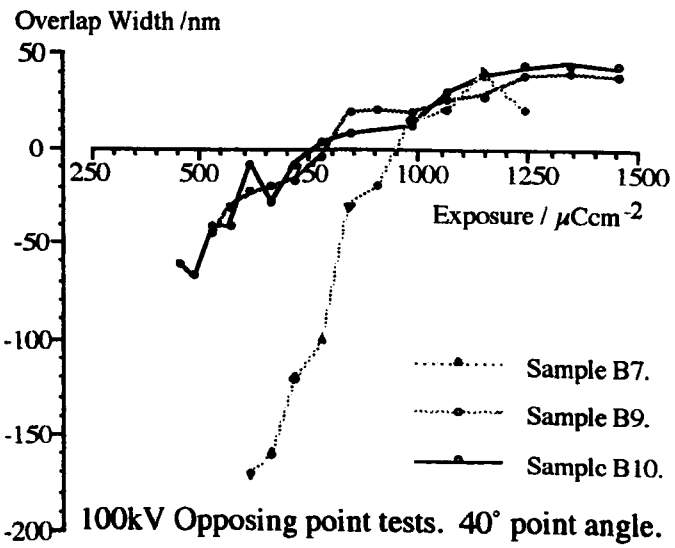
Sample B12. 50kV Opposing point test. 5nm point gap.

Figures 4.22. Overlap width vs exposure at 50kV. Point angles of 20°, 40°, 60° and 80° with 5nm point gap.



Sample B9. 100kV Opposing point test. 5nm point gap.

Figures 4.23. Overlap width vs exposure at 100kV. Point angles of 20°, 40°, 60° and 80° with 5nm point gap.



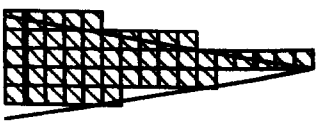
100kV Opposing point tests. 40° point angle.

Figure 4.24. Repeatability of exposure technique. Samples B7 to B10, 40° point 5nm gap.

Repeatability of the overlap width is important. Figure 4.24 presents the results from the 100kV samples, B7, B9 and B10, for a 40° point with a 5nm gap. In this graph, we see the results for samples B9 and B10 follow each other well, for the most part to within ±10nm. Sample B7, however, does not fit well. This sample was metallized using Al and it is thought that the larger granular structure is limiting the exact

reproduction of the resist image. Samples B9 and B10 used a Ti:Au metallization which maintains a much smaller grain size and can therefore reproduce the resist image more faithfully. Micrographs of similar regions on sample B7 are presented in the following section.

Although the tests proved that the appropriate lithographic structure could be developed for this form of patterning, two difficulties were discovered. Firstly, the EBL-HR1 process involves developing the high resolution bilayer resist (2.5% BDH, 2.5%Elv) with a 2.5:1 IPA:MiBK mixture. A more selective, higher contrast, 3:1 mixture is recommended for high resolution processing particularly for forming gratings and dense pattern structures. The advantage is probably limited in isolated features but processing of the more dense device patterns is likely to require the 3:1 developer mixtures of the EBL-HR2 & HR3 processes. Secondly, fractionation of the designed pattern produces an incorrectly pixelated structure for the pattern writing. Figure 4.25 compares the fractionated pattern with the designed pattern. This effect leads to asymmetry in the written pattern and slightly increases the dose seen by the constriction area.



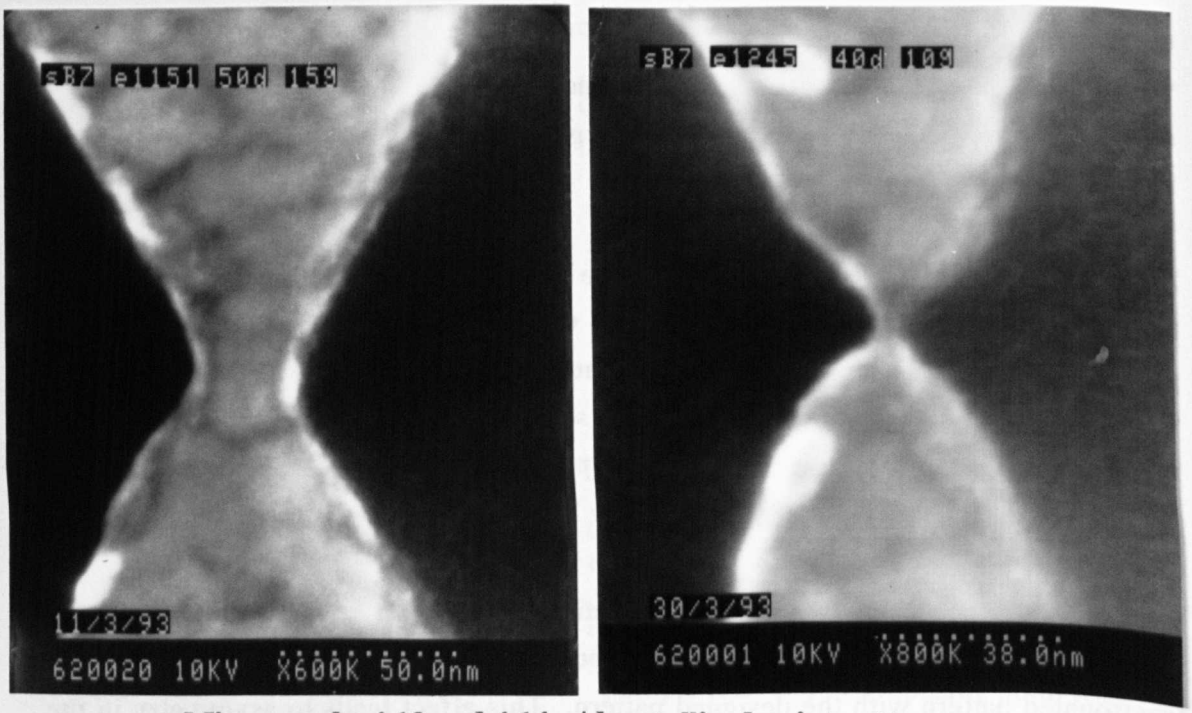
— Designed Pattern.
▨ Single beam pixel.

Figure 4.25. Designed and fractionated point patterns.

4.4.4.1B GRAIN STRUCTURE DEVELOPMENT

Observation of the tests metallized with Al revealed a lower limit on the useful constriction dimension. Micrographs 4.13 and 4.14 give images of the grain structure for constrictions which are 30nm and 7nm respectively, the lighter areas represent the Al film. These micrographs are representative of the detail found on these samples.

Where the constriction width is much smaller than the natural grain size, the formation of complete Al grains is found to be shifted back from the apex of the points and no observable grain structure is found by SEM inspection in the narrowest regions. Whether this effect occurs as a natural grain growth phenomenon, as discussed in section 4.4.2, or whether it is a result of the metallization closing the resist gap, as discussed in section 4.4.4.3, is uncertain.



Micrographs 4.13 and 4.14. Al metallized point patterns.

4.4.4.2 MODIFIED OPPOSING POINT TEST

The tests described in section 4.4.4.1 provided information on the repeatability of the structure formation but suffered from a number of difficulties. In this test the pattern was designed to fractionate explicitly into the shapes required, as shown by figure 4.26. The angle variation of the point was changed to a pixel gradient and an extra test, with no gaps between the points, was used to help lower the doses required for pattern development. The pixel gradients were varied through 1:1, 1:2, 1:3 and 1:5, equating to point angles of 90°, 53°, 37° and 23° respectively. The point gap was varied from 0 to 20nm in 5nm steps.

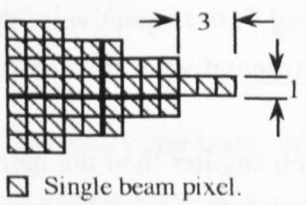
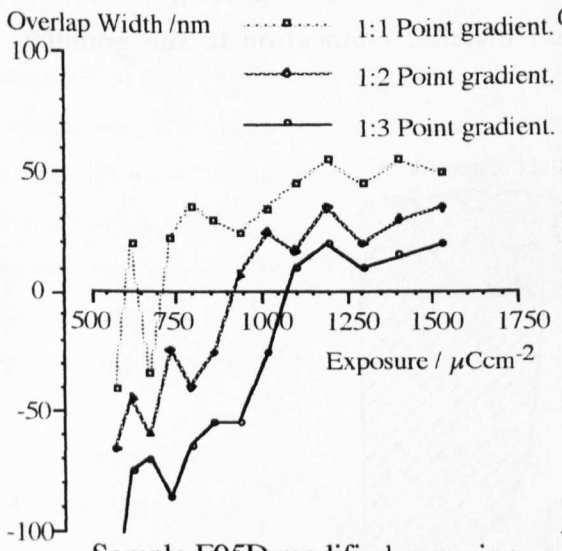


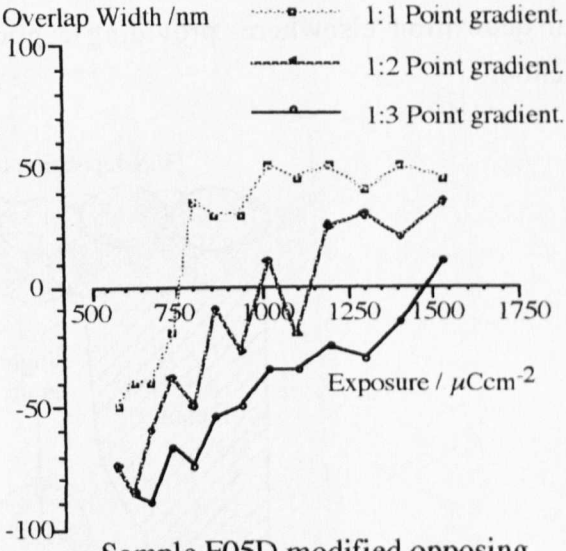
Figure 4.26. Redesigned point structure for 1:3 gradient.

The measurement technique was the same as discussed in section 4.4.4.1. The sample, F05D, was prepared using the EBL-HR2 process. An opposing point pattern was written as an exposure test, from 450 to 1530μCcm^{-2} in a 16 step geometric series with ratio 1.085. The resulting overlap width vs exposure plot for the zero and 5nm point gaps (with 1:1, 1:2 & 1:3 point gradients) are presented in figures 4.27 and 4.28.



Sample F05D modified opposing point test, zero point gap.

Figure 4.27. Overlap width vs. exposure for zero point gap.



Sample F05D modified opposing point test, 5nm point gap.

Figure 4.28. Overlap width vs. exposure for 5nm point gap.

The results presented in figure 4.28 can be compared with those in figure 4.23, where point gradients of 1:1, 1:2 and 1:3 compare roughly with the point angles of 80°, 60° and 40°, respectively. No results are shown for the 1:5 point gradient as this structure did not develop a positive overlap in this sample. Figure 4.27 displays the exposure dose reduction effected by removing the point gap from the designed pattern.

The change in developer mixture has caused the major differences between the results of figure 4.28 and figure 4.23. The dose required for similar features has roughly doubled. Other than this, the plots have roughly the same character, suggesting the pattern change has had little effect at these point angles. Comparison of figure 4.27 with 4.28 demonstrates that there is little advantage in using a designed point gap to reduce the overlap width.

4.4.4.3 TEST OF RESIST CLOSE-OVER MECHANISM

Similar to the problems found in section 4.3.4.4, it was observed that film thicknesses in narrow patterned regions self terminate. The cause of this is close-over of the resist slot due to deposition on the resist edges. This deposition reduces the slot width. If the gap is of the order of the film grain size, as in the case of Al with 20 to 30nm patterns, the slot can completely close. Closure of the slot prevents further deposition on the substrate underneath the bridge formed by the metal, as shown in figure 4.29. This property can be used to control the film thickness in the constriction. Being able to

control the film thickness to a single grain layer in the narrowest region allows thicker film deposition elsewhere, providing a good metallic connection to the granular junction.

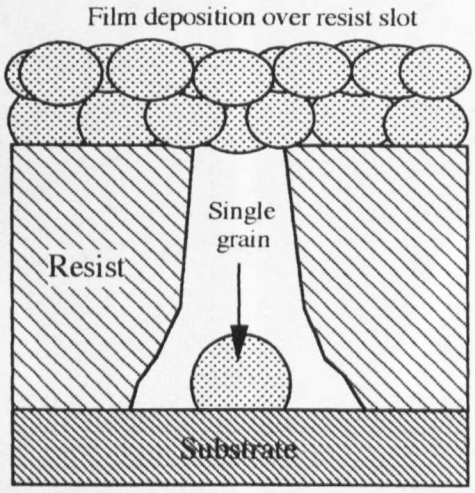
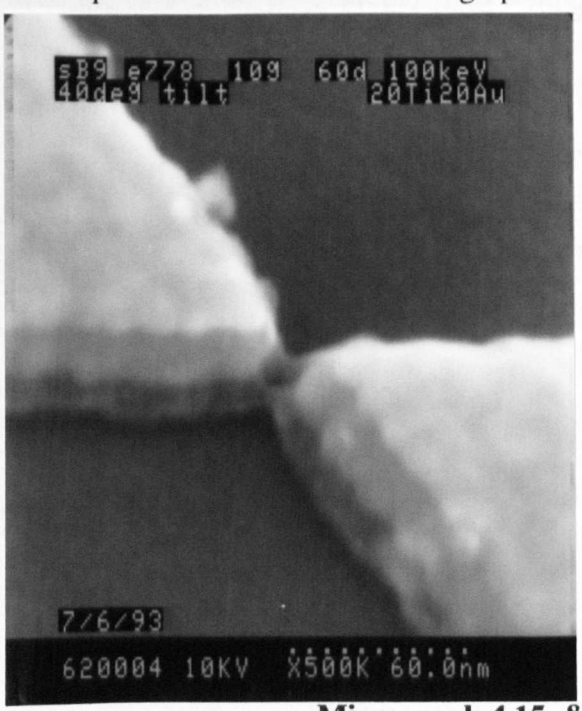


Figure 4.29 Schematic of resist closure mechanism.

To confirm this effect, a sample was prepared and metallized in a step-wise manner. The deposited film structure was 10nm each of Ti:Au:Ti:Au, giving a total thickness of 40nm. Films were deposited using the Plassys MEB-450 evaporator. The results of this experiment are shown in micrographs 4.15 and 4.16.



Micrograph 4.15. & Micrograph 4.16.

Constriction metallized with 2 x (10:10nm Ti:Au) showing broken top layer film.

These micrographs display a point structure viewed at an angle. The variation of

contrast is a result of the layered film structure. Care must be taken in interpreting these micrographs because of the differences in evaporation properties of Ti and Au. Au evaporates almost as a point source in the MEB-450 evaporator. Ti, however, will not evaporate until the whole surface of the boat reaches evaporation temperature and evaporation proceeds from the entire surface of the melt. The increased surface area increases the source footprint on the substrate. This effect is similar to the penumbral effect discussed in discussed in section 4.2.3.3. This is the reason for the apparent three layer structure in the micrographs, despite four layers being deposited. The second Ti layer covers the previous two layers.

Both micrographs clearly show a discontinuity in the top layer of the deposit. This is characteristic of other areas of the sample with similar constriction widths, and can be explained by the mechanism discussed above.

4.4.4.4 DEVELOPMENT OF DEVICE STRUCTURE

For the device structure, two sets of double tunnel junctions were chosen. This allows formation of two single electron electrometers which can be either individually gated or coupled through a common gate. Figure 4.24 shows the structure.

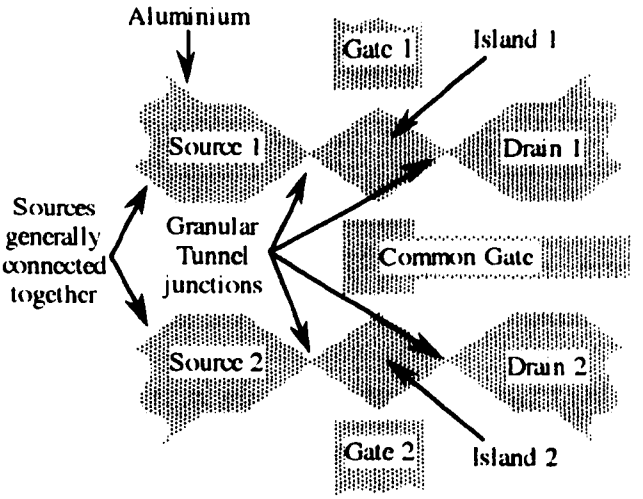
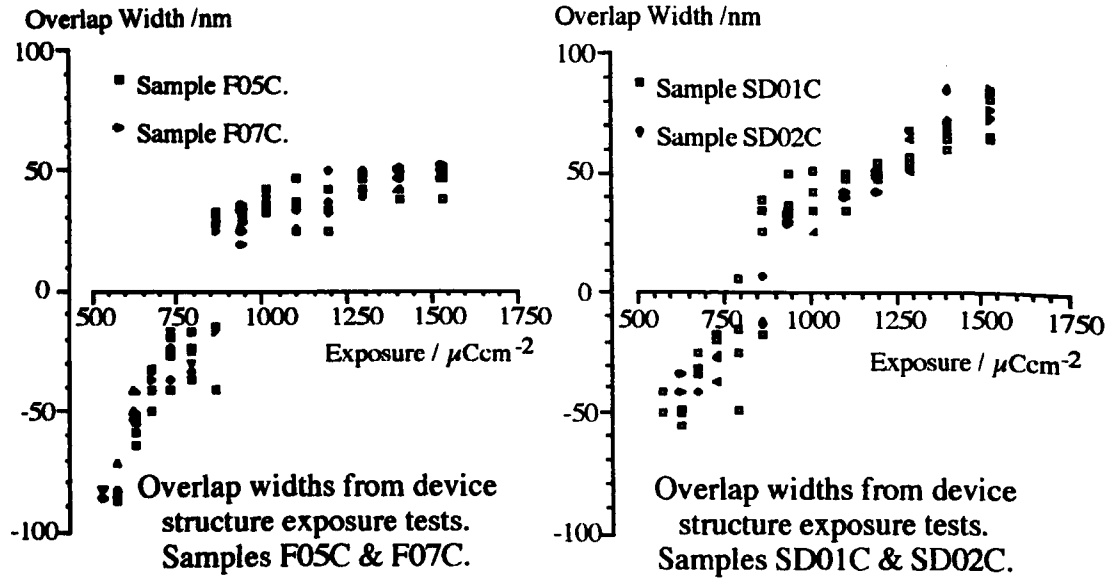


Figure 4.30. Schematic of device pattern.

A measure of the repeatability of the device formation can be gained by plotting all the overlap widths against exposure, from exposure tests on this structure. Figure 4.25 displays a plot of this for two samples, F05C and F07C.



a. Samples F05C and F07C. b. Samples SD01C and SD02C.

Figure 4.31. Device structure repeatability.

Results from four samples are presented in figure 4.25. Table 4.11 details the fabrication processes used.

Sample	Lithography Process†	Beam voltage	Metallization
F05C	EBL-HR2	100kV	40nm AuPd
F07C	EBL-HR2	100kV	40nm AuPd
SD01C	EBL-HR3	100kV	2x(5:5nm Ti:Au)
SD02C	EBL-HR3	100kV	2x(5:5nm Ti:Au)

† - see Section 3.8.

Table 4.11. Device structure tests.

The information provided by the graph of figure 4.31 provides a calibration for fabrication of the device structure and was used to determine the 910μCcm^{-2} exposure value for the devices fabricated for electrical measurement.

4.5 CONCLUSIONS

The work of this chapter sought to investigate three tunnel junction fabrication processes in terms of their ability to reliably and controllably fabricate ultrasmall Al-AlO_x-Al MIM tunnel junctions. The true test of this investigation is to be able to fabricate measurable device structures which yield working single electron devices at elevated temperatures. The work of chapter 6 discusses the measurements made on

such devices. The crossed track process was rejected from a purely fabrication point of view as this process proved too difficult to control and did not produce reliable tunnel junction structures. The remaining two processes were used to produce some single electron test structures and the conclusions from the fabrication work follow.

The suspended mask process is proven and reliable for the fabrication of overlap tunnel junctions. For the purposes of this work, the overlap must be reduced as much as possible. The analysis of section 4.2 has shown that for overlap widths of less than 40nm there are serious questions over the ability to form the junctions. However, as this is a random, sample to sample effect, it will be possible to use this process in electrical measurements if the overlap width is not too small. The overlap designs for devices fabricated using this process were intended to be around 30nm (producing 30 x $30nm^2$ junctions). As the expected overlap width error is about ±40nm, there is a reasonable probability of forming a junction with 30nm overlap although it is also likely that some structures will fail or produce larger overlap widths.

For high temperature operation of single electron devices in aluminium, the patterned granular film structures have the greatest potential of the two remaining techniques. The film characterization work has shown that the films are likely to be metallic in nature, with only a slight change in the lattice constant being experienced. This change is most likely to arise either from stress in the film or from the lower material density experienced in very thin films when compared with bulk materials. Therefore, it should be possible to form metallic contacts to the regions of interest, ie. the grain boundaries. As the mean grain diameter is about 20nm and patterning of the structure on this scale is certainly possible, it is likely that the position of the grain boundaries can be controlled through the lithography. It remains to be seen, however, if the grain boundaries can be formed such that useful single electron devices are produced.

With regards to lithographic pattern definition, it is clearly possible to create resist patterns which are below the direct write resolution limit of the electron beam system, as indicated in micrograph 4.14. However, it has been shown that it is difficult to controllably write such structures with the EBPG5-HR EBL Beamwriter system. Where metallization is used in a lift-off process to transfer the resist pattern into the metal film, it has been shown that the grain size of the material plays a significant role in the final structure size. Where aluminium metallization is used, minimum dimensions are limited to around 20nm for good control over the feature size. Given that the direct write resolution limit of the Beamwriter is roughly 12nm then alternative, higher resolution, lithography systems will be required to allow patterning which controls grain placement in smaller grain size materials.

4.6 REFERENCES

1. G.J. Dolan, *Appl. Phys. Lett.* **31**,377 (1977).

2. G.J. Dolan & J.H. Dunsmuir, *Physica B.* **152**(1),7 (1988).

3. R.H. Ono, J.E. Sanvageau, A.K. Jain, D.B. Schwartz, K.T. Springer & J.E. Lukens, *J. Vac. Sci. Technol. B.* **3**(1),282 (1985).

4. T.A. Fulton & G.J. Dolan, *Phys. Rev. Lett.* **59**,109 (1987).

5. P. Delsing, T. Claeson, K.K. Likharev & L.S. Kuzmin, *Phys. Rev. B.* **42**(12),7439 (1990).

6. L.J. Geerligs, V.F. Anderegg, P.A.M. Holweg, H. Pothier, D. Esteve, C. Urbina, M.H. Devoret & J.E. Mooij, *Phys. Rev. Lett.* **64**(22),2691 (1990).

7. L.S Kuzmin & K.K. Likharev, *Jpn. J. Appl. Phys.* **26**(suppl 3),1387 (1987).

8. H. Nejoh & M. Aono, *Appl. Phys. Lett.* **64**(21),2803 (1994).

9. C. Schonenberger, H. van Houten & H.C. Donkersloot, *Euro. Phys. Lett.* **20**(3),249 (1992).

10. R.W. MacLeod, *Fabrication and characterisation of GaAs-InP nanostructures*, Ph.D. Thesis, No. 7154, University of Glasgow, 1994, Chapter 3, p49, figure 3.15.

11. C.A. Neugebauer, in *Handbook of thin film technology*, eds. E.I. Maissel & R. Glang, McGraw Hill, New York, 1970, Chapter 8, Condensation, nucleation and growth of thin films.

12. J.P. Hirth & K.L. Moazed, *Nucleation processes in thin film formation*, Physics of Thin Films. **4**,97, eds., G. Hass & R. E. Thun, Academic Press, 1967.

13. G.C. Bond, *Heterogeneous Catalysis*, 2nd ed., Oxford University Press, 1987, Chapter 2.

14. Metals Handbook, Volume 8, *Metallography, structure and phase diagrams*, p343, *Phase diagrams of binary alloy systems*, Am. Soc. Metals, 1973.

15. S. Mader, in *Handbook of thin film technology*, eds., E.I. Maissel & R. Glang, McGraw Hill, New York, 1970, Chapter 9, Determination of structures in films.

16. D.R. Lide (ed.), in *Handbook of Chemistry & Physics*, Section **12**, p9, 72nd Edition, CRC Press, Boston, 1991.

17. T.E. Hartman, *J. Vac. Sci. Technol.* **2**,239 (1965).

CHAPTER FIVE

Thin Film Electrical Characterization

5.1 INTRODUCTION

The structures developed in this work are formed from a granular film of thickness comparable to the average grain diameter. This type of film structure poses a number of questions about the electrical characteristics of any device fabricated using this technique. Firstly, will the exhibited film behaviour be indicative of a metallic film or of a discontinuous granular film? Secondly, will the bulk or unpatterned regions of the film affect the charge transport properties of the device? Thirdly, will the device behaviour be swamped by the behaviour of immediately adjoining junctions or grains? Finally, and perhaps most importantly, will the film parameters and structure dimensions be of the correct scale to produce high temperature single electron effects?

Electrical characterization of both films and devices are required to answer these questions. In order to determine the behaviour and type of metallic film being used for these experiments, measurements were made of the resistivity temperature dependence. Measurements of the magneto-resistance and Hall effect characteristics were also taken. All measurements were made in order to gain information about the conduction processes within the films. However, the bulk of the information required was gained from the resistivity temperature dependence characteristics.

This chapter details the characterization work carried out to answer at least the first two

questions posed above. This chapter breaks down into five main sections. Firstly, some background is given to the subject of metallic and metallic film conduction, with emphasis on aluminium. Following this, section 5.3 provides the background to resistivity measurement techniques while section 5.4 details the experimental equipment used. The results are presented in section 5.5 with conclusions following in section 5.6. Chapter 6 is concerned with measurements on the fabricated devices and as such details the work carried out to answer the latter two of the four questions posed above.

5.2 METALLIC FILM RESISTIVITY

Where resistivities of metals are quoted in general reference texts[1,2], specially prepared samples of single crystal or polycrystalline material have been used. The materials measured in this section of the work are formed from thin oxidised films of polycrystalline, or granular, aluminium. These films, by definition, exhibit higher defect concentrations and their properties can differ significantly from those of "bulk" materials. The next section, 5.2.1, provides an introduction to metallic conduction while section 5.2.2 gives a brief discussion of the situation in thin metal films. Section 5.2.3 gives a brief discussion of conduction processes within granular thin metal films, with section 5.2.4 reviewing some experimental work with aluminium films.

5.2.1 METALLIC CONDUCTION

Metals are differentiated from non-metals by the nature of the bonding in the lattice. Valence electrons in the metal become delocalized from their original nucleus, resulting in free movement of the electrons throughout the material. With mono-valent alkali metals this results in a nearly uniform electron gas. It is this delocalization of the valence electrons which provides the good electrical conductivity of metals. In metals the resistivity, ρ, is generally in the range $10^{-8}<\rho<10^{-6}\Omega$m, whereas in semiconductors it is typically in the range $10^{-4}<\rho<10^{5}\Omega$m, while with insulators, $\rho>10^{15}\Omega$m is common. Other features which help with the classification of metals are obedience of Ohm's law and good thermal conductivity. The details of electrical and thermal behaviour of a metal, however, depend heavily on the composition and preparation of the material.

5.2.1.1 OHMS LAW AND RESISTIVITY

Ohm's law states that the current, I, flowing through a wire is proportional to the potential drop, V, along that wire, with the resistance being the proportionality constant:

$V=IR$. The wire resistance depends on its dimensions, but is independent of the magnitude of the current or potential drop. The dimensional dependency can be removed through the introduction of resistivity, allowing the current to be replaced by current density, **j**, and voltage by the electric field at a point in the wire, **E**. Current density is a vector, parallel to the direction of charge flow induced by the electric field with a magnitude equal to the amount of charge crossing a unit area perpendicular to the current flow per unit time. If we consider a wire of length L and cross-sectional area A, with current flowing along the length of the wire, we have a resistance (from the definition of Ohm's law) given by:

$$R = \frac{\rho L}{A}. \qquad [5.1]$$

This definition of resistance produces the dimensionally independent parameter ρ, the resistivity. This allows Ohm's law to be restated in terms of current density and electric field as:

$$\mathbf{E}=\rho\mathbf{j}. \qquad [5.2]$$

The resistivity can be expressed as a conductivity simply through the inverse relation of; $\mathbf{j} = \sigma\mathbf{E}$, where $\sigma=1/\rho$. Where a material is not uniform, or isotropic, ρ and σ may be represented by tensors.

In Drude's theory of electronic conduction, current density can be calculated by considering the number of electrons per unit volume, n, and their velocities, **v**. The resulting current density is simply the product of these factors times the charge carried by each electron or,

$$\mathbf{j}=-ne\mathbf{v}. \qquad [5.3]$$

Drude made the assumption that electrons were small light particles moving through a fixed background of heavy nuclei. This resulted in electron collisions with these nuclei randomizing the electron velocity. At thermal equilibrium and zero field conditions, the average electronic velocity will be zero as random or Brownian motion of electrons results from this assumption. Under the action of an electric field, each electron will develop a drift velocity of $\mathbf{v}_d=-e\mathbf{E}t/m$, where m is the electron mass and t is the time under action of the electric field. As each electron will travel a distance $l=v_d t$ between collisions, then they will, on average, travel a distance $\bar{l}$, with a characteristic time, τ, between collisions. This relaxation time is the average time between collisions for a conduction electron when under the action of field, while $\bar{l}$ is the mean free path of such a conduction electron. Gathering together the definitions for electric field, current density and drift velocity, we find that Ohm's law (equation 5.2) becomes;

$$\mathbf{E} = \left(\frac{m}{ne^2\tau}\right)\mathbf{j}, \qquad [5.4]$$

giving the resistivity as:

$$\rho = \frac{m}{ne^2\tau}, \qquad [5.5]$$

5.2.1.2 RELAXATION TIME APPROXIMATION AND MATTHIESSEN'S RULE

Drude's theory of metals visualises electrons freely moving within a fixed, background, lattice. The velocity distribution of the electrons are determined by Maxwell-Boltzmann statistics, as in the case of a ideal gas. The resulting theory led to a good description of electrical resistance but a poor description of the thermal conductivity. Sommerfeld modified this theory using quantum mechanical arguments, such as the Pauli exclusion principle. In this modification, he used Fermi-Dirac statistics to determine the velocity distribution of the electrons, providing better estimates of the thermal conductivities. The dependencies on electron energy and wave-vector were later added to this theory to produce the semi-classical or relaxation time approximation. In this semi-classical approximation, the m of equation 5.5 is replaced with m^* for the electron effective mass. This change accounts for the effects of the periodic lattice potential on the conduction electrons. Therefore within this theory, heavy reliance is given to the electronic band structure[3] and the shape of the Fermi surface in estimating electronic phenomena.

This relaxation time approximation leads to what is called Matthiessen's[4] rule. This rule states that where scattering mechanisms arise from independent sources, with their own characteristic relaxation times, the total behaviour of the metal can be characterized by the following expression:

$$\frac{1}{\tau} = \sum_{i=1}^{n} \frac{1}{\tau_i}, \qquad [5.6]$$

where n is the number of scattering mechanisms. If, for example, we have two independent scattering mechanisms, $\tau_{def.}$ for impurity or defect scattering and $\tau_{ph.}$ for phonon (or temperature dependent) scattering, this leads to the expression:

$$\frac{1}{\tau} = \frac{1}{\tau_{def}} + \frac{1}{\tau_{ph}}, \quad \text{or,} \quad \rho = \rho_{def} + \rho_{ph}. \qquad [5.7]$$

In other words, the resistivities from different scattering sources are additive, according to the relaxation time approximation. In this model, the temperature dependent scattering is described as it would be expected in a pure material, and is assumed to approach zero as temperature approaches zero. The impurity contribution ensures that the resistivity never reaches zero (while the material remains in the normal conducting state).

Where the relaxation time approximation is valid, it can be shown that Matthiessen's rule holds. However, we have seen that this rule requires the scattering mechanisms to be completely independent. When it is considered that the scattering rate of an electron depends on the configuration of other bodies in the system, it can be seen that these scattering events can be inter-dependent. In this case, Matthiessen's rule will break down.

5.2.1.3 MANY BODY INTERACTIONS

In order to describe the situation within metals more clearly, the "many body" effects touched on above need consideration. These many body effects take two main forms; electron - electron interactions and electron - phonon interactions. Electron - electron interactions provide little contribution to the bulk of the electronic and thermal behaviour of metals[5]. However at low temperatures ($T<<40K$) electron-electron interactions can contribute significantly to electronic behaviour, through superconductivity and electron - electron scattering.

Electron - phonon interactions are the effect of lattice vibrational modes on the electrons. These result in a number of clear electronic and thermal phenomena. Firstly, electron interaction with phonons results in electron scattering. This mechanism is the major contributor to the temperature dependence of metallic resistivity, as discussed in section 5.2.1.4. Secondly, as ions in the lattice are charged particles, there exists the possibility of effects due to attraction and repulsion of electrons and ions. With this type of interaction the valence electrons both distort the lattice and modify the electron energy to within $\pm\hbar\omega_n$ of the Fermi energy, where ω_n is the highest frequency vibrational mode of the system. This results in a major contribution to the thermal conductivity and heat capacity of a metal. Also related to electron interaction with lattice ions is superconductivity. Although not all metallic elements exhibit superconductivity (and some non-metallic elements superconduct under pressure), lattice distortions by an electron with energy around the Fermi energy can exert an attractive force on a second electron. This attraction is strongest when the electron states are configured with the same energy and wave-vectors of $+\mathbf{k}^{\uparrow}$ and $-\mathbf{k}^{\downarrow}$, ie. the

electrons have opposite k-space wave-vectors and are configured in opposite spin arrangements. At a temperature, $T<T_C$, (where T_C is the superconducting transition temperature) the electrons can lower their energy by forming a Cooper pair[6], resulting in a superconducting electronic ground state, characterized by vanishing resistivity and the Meissner effect. This is not the only interaction which can cause superconductivity, as periodic disturbances in the background potential landscape can aid the formation of a superconducting ground state.

5.2.1.4 TEMPERATURE DEPENDENCE OF RESISTIVITY

Temperature dependence of resistivity is exhibited by all materials, and can be defined as the temperature coefficient of resistivity (TCR):

$$\alpha = \frac{1}{\rho_0}\frac{d\rho(T)}{dT}, \qquad [5.8]$$

where ρ_o is the resistivity at a given temperature. This reference resistivity is generally defined as the resistivity at room temperature or, $\rho_{RT} = \rho(300\text{K})$. The value of the material TCR is generally quoted in parts per million resistivity change per degree of temperature change (or ppmK^{-1}). Figure 5.1 illustrates the resistivity temperature dependence of aluminium[2], with the inset plotting the TCR. The data given here is for a pure polycrystalline aluminium sample and as such the data below 40 to 50K relates purely to the sample from which the results were taken. The exact behaviour of a sample at low temperature is heavily dependent on the preparation technique and sample purity.

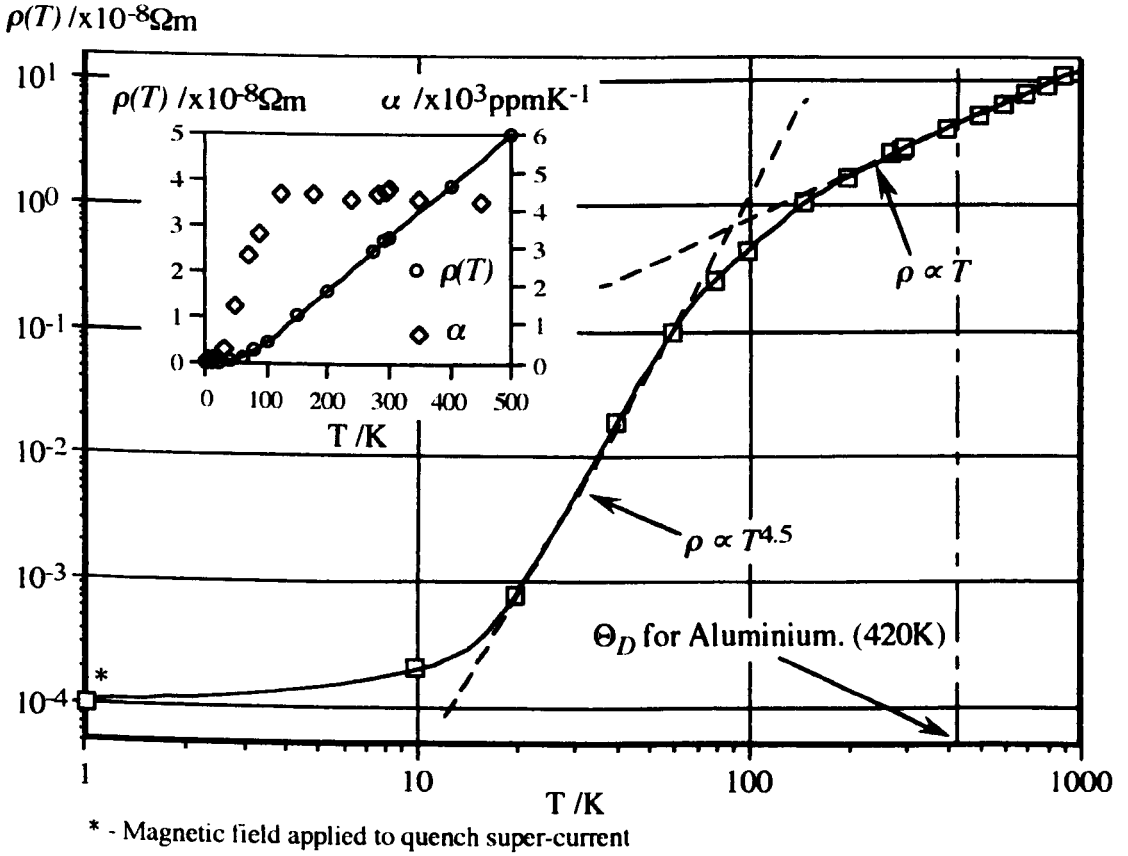
Around room temperature, most metals exhibit temperature independent behaviour of the electronic thermal conductivity while the electrical resistivity typically varies with temperature. That is $\rho \propto T$. This dependency results from the number of phonons which can take part in a scattering event increasing linearly with temperature. This is generally the case in metals where the temperature is near or above the Deybe temperature, Θ_D. The Deybe temperature for aluminium is 420K[7].

Increase in the thermal conductivity, and a faster decrease in the electrical resistivity, are observed when the material is cooled well below the Deybe temperature. Under these conditions $\rho \propto T^n$, where n = 3 to 5 in the steepest part of the range. The power factor of this behaviour arises from different forms of scattering event. Electron phonon interactions give rise to a dependency which varies as:

$$\frac{1}{\tau_{el-ph}} \propto T^3, \quad \text{for } T << \Theta_D. \qquad [5.9]$$

As temperatures lower, the increasing dominance of forward scattering (reduction of the ability of phonons to change the electron velocity vector by a large angle) produces a further dependency of:

$$\frac{1}{\tau_{fwd}} \propto T^2. \qquad [5.10]$$



* - Magnetic field applied to quench super-current

Figure 5.1. Resistivity temperature dependence for polycrystalline aluminium[2].

The combined effects of these two mechanisms (through Matthiessen's rule) lead to a theoretical $\rho \propto T^5$ law, which is sometimes referred to as the Bloch T^5 law. Modification of the T^5 law arises where complex Fermi surfaces are present and/or inter-band scattering is possible. These are termed "umklapp" or "U" processes, which tend to reduce the power of the T^n dependency through reduction of the effect of forward scattering events on the resistivity decrease. As can be seen from figure 5.1, for $T<<\Theta_D$, the steepest part of the $\rho(T)$ curve has a power law, $\rho \propto T^n$ with $n \approx 4.5$, indicating the presence of U-processes in the electron transport mechanisms for this material. As the Fermi surface of aluminium is relatively complicated[3,8] and the

material is generally polycrystalline, it is reasonable to expect modification to the T^5 law through inter-band scattering.

At low enough temperatures the resistivity temperature dependence, $\rho(T)$, plateaus out. The resistivity in this region is due to impurities and dislocations within the lattice. The resistivity which arises due to this disorder within the material is termed the residual resistivity. This residual resistivity is a function of the preparation of any given sample and can be measured simply as ρ_{RR}, the value of the plateau resistance, or as a residual resistivity ratio:

$$RRR = \frac{\rho_{300K}}{\rho_{4.2K}}. \qquad [5.11]$$

The higher the RRR, the purer the material. Also, the lower the temperature at which the plateau appears, the purer and more crystalline the material.

5.2.2 CONDUCTION PROCESSES IN THIN METAL FILMS

The definition of resistivity and conductivity removes the effect of sample dimensions on the electrical characteristics of a material. This assumes however that the resistivity is truly independent of the sample dimensions. The resistivity behaviour of a bulk material is affected by the preparation techniques, impurity and defect content, microstructure, chemical composition, bonding type, temperature and strain. Electrons within bulk materials suffer far fewer collisions with the sample boundaries than with defects within the material Therefore the resistivity may be defined as a dimensionally independent parameter. All of the above mentioned features affect conduction in thin films but we find that the film dimensions do affect the resistivity behaviour along with increased dominance of microstructure, defect content and strain effects. This determines not only the film structure but also the resistivity behaviour. The increased probability of a collision with the film boundaries leads to the introduction of a size effect into the film behaviour. The following sections briefly discuss classical and quantum effects in continuous metallic thin films which can significantly modify the resistivity behaviour.

5.2.2.1 SIZE EFFECT

The reason for the introduction of dimensional dependencies into the resistivity can be seen by considering the sources of scattering within a film. Normal impurity and defect scattering exist and indeed, increased defect densities due to interfacial strain, grain boundaries and surface roughness are all sources of increased scattering by defects.

However in the relaxation time approximation, we have defined an average carrier lifetime, or relaxation time, τ, which is the average time between collisions. Ignoring any drift velocity for the moment, the conduction electrons possess an instantaneous electronic velocity of magnitude v_F, the Fermi velocity. Therefore, an electron mean free path of $\bar{l} = v_F \tau$ can be defined as the average distance travelled between collisions. Reduction of the film thickness, d, can result in $\bar{l} >> d$, especially at very low temperatures where in relatively pure metals, $\bar{l}$ can extend several hundreds of microns.

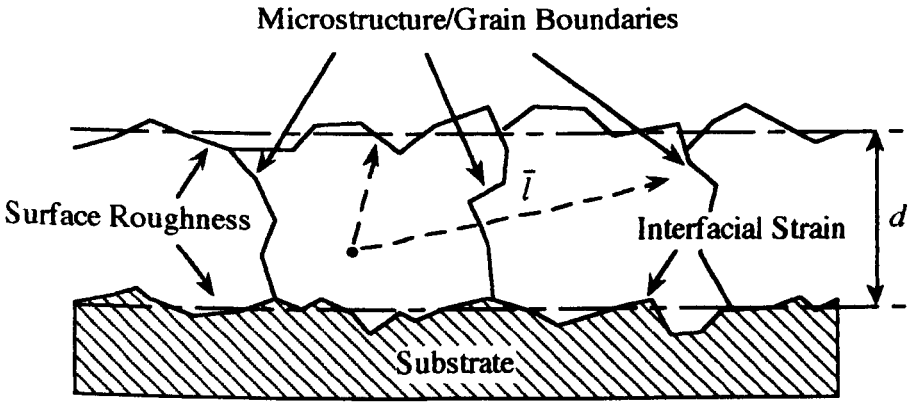


Figure 5.2 Picture of model for thin film conduction.

Using the relaxation time approximation[9], taking $n = 1.81 \times 10^{29} m^{-3}$, v_F to be $2.03 \times 10^6 ms^{-1}$ and m to be the electron rest mass, the resistivity data of figure 5.1 used with equation 5.5 yield mean free paths for free electrons of around 13nm at 300K and 400μm at 2K. This estimate of the mean free path provides an idea of scale for the term *thin* film. If it is possible to produce a film with the purity of the material whose characteristics are given in figure 5.1, then thicknesses of several hundred microns are required before the film can be described as a bulk material. Therefore, if $d < \bar{l}$, modifications to the resistivity can be expected due to the dimensions of the film. J.J. Thomson[10] was the first to consider this effect, using Drude's theory of conduction and a one electron model. The idea was taken further by Fuchs[11] who assumed quasi-elastic collisions with the film boundaries (ie: Matthiessen's rule holds for film boundary collisions).

To examine the problem, Fuchs defined the parameter $k = d/\bar{l}$. Solution of the problem using the Boltzmann transport equation ensured that all electron free paths (including those beginning at the film surfaces) were included in the calculation. This type of solution also accounted for the statistical distribution of the electron free paths. Fuchs' result reduces to very simple forms where the film is either very thick or very thin. Where the film is thick, $k >> 1$ and we have;

$$\frac{\rho_0}{\rho} = 1 - \frac{3}{8k}, \qquad [5.14]$$

where, ρ_0 is the resistivity for the "bulk" material. Where the film is thin, $k<<1$ and,

$$\frac{\rho_0}{\rho} = \frac{3k}{8} ln \frac{1}{k}. \qquad [5.15]$$

It is important to note that the term "bulk" here refers to an infinitely thick material with the same characteristics as the film in question. This defines an intrinsic resistivity, ρ_0, which is the resistivity without the effect of boundary collisions, instead of the resistivity value of a pure version of the material[12].

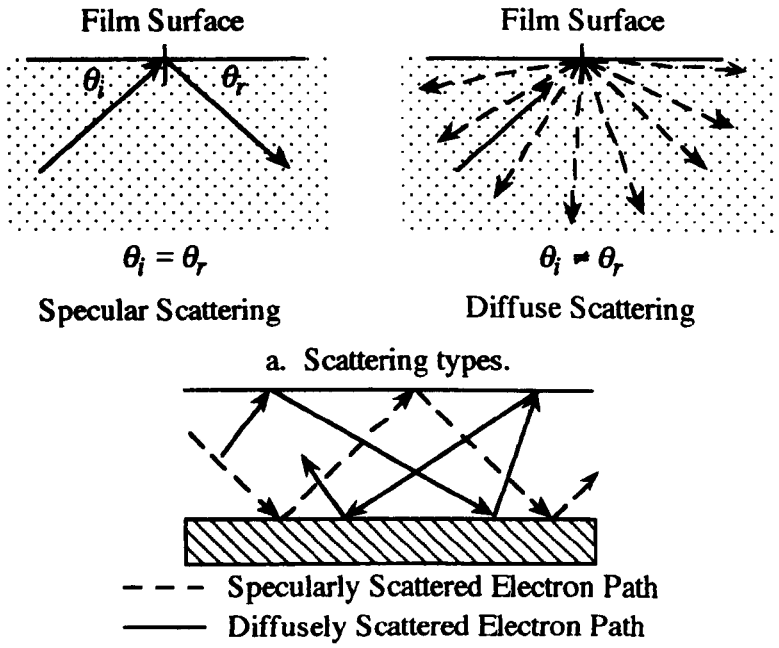


Figure 5.3 Diffuse and Specular Scattering from a film surface.

Figure 5.3 illustrates two types of scattering which can exist, specular and diffuse. Specular scattering is a simple reflection where the angle of incidence equals the angle of reflection whereas diffuse scattering randomizes the reflection angle. It is important to note that only diffuse scattering reduces the conductivity as this process modifies the forward component of the electron velocity. By making use of a parameter, p, for the ratio of specular to diffuse scattering, Fuchs also calculated a relationship which described the effects of scattering characteristics from one of the film surfaces. Where $k>1$, this relationship reduces to:

$$\frac{\rho_0}{\rho} = 1 - \frac{3}{8k}(1-p), \quad \text{where} \quad k>1 \qquad [5.16]$$

Further modifications to these classical theories have been made to account for reflections from both film surfaces, through the inclusion of further scattering parameters. However, most of these theories follow classical approaches, assuming both spherical Fermi surfaces and that the film approximates to a single crystal with parallel top and bottom surfaces.

In order to account for effects due to the quantum mechanical, or wave nature of the electron, further modifications have been made to size effect theories. Ziman[13] examined the scattering process when the surface or interface roughness was compared with the electron wavelength. In this view, the electron wave is seen as sampling the surface roughness, and the specularity or diffusivity of the scattering can be determined by whether the wavelength averaged out the surface roughness. If the wavelength is large compared with the surface roughness, the electron will, on average, be scattered specularly. Conversely, where the wavelength is short compared with the surface roughness, the electron is able to sense the surface topography during a collision and will, on average, scatter diffusely. When we consider that most metals have conduction electrons with de Broglie wavelengths of only a very few Angstroms, it is reasonable to expect most materials to produce diffuse scattering.

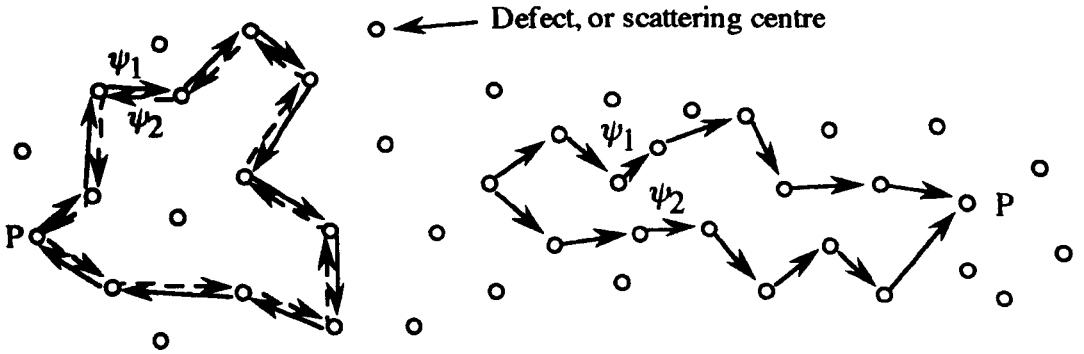
The size effect can also be considered in a quantum mechanical framework. The arguments behind this theory are similar to and included in the theories for weak localization, as briefly discussed in section 5.2.2.2. In this case, we have modifications to the weak localization theory through the effect of interference between closed electron paths which are scattered off the sample boundaries. A development of this theory has been given by Beenakker and van Houten[14].

5.2.2.2 ELECTRON COHERENCE EFFECTS AND WEAK LOCALIZATION

On the metallic side of the metal-insulator transition, we are considering samples in the limit of weak disorder. That is, scattering centres or impurities are randomly and sparsely distributed throughout the material. Where the scattering is elastic, the electron phase is preserved although a phase shift may be introduced into the electron wavefunction by the scattering. This gives rise to quantum effect corrections to the classical description of the conductivity through wavefunction interference effects. If we examine an electron path through a scattering medium, we find we can arrange pictures as shown in figure 5.4.

Here, two paths of an electron are shown, one described by the wavefunction ψ_1, and

the other by ψ_2. Either path may be taken by an electron, but if we consider the wave nature of the electron, we see the possibility for interference to affect the conduction. Both electron paths have the same source and therefore the same phase information. If these paths are scattered back onto each other, as shown in figure 5.4, the possibility of interference must be considered. This however requires that the phase information is not lost in the scattering process.



a. Coherent backscattering, leading to weak localization.

b. Coherent scattering, leading to non-linearities and universal conductance fluctuations.

Figure 5.4 Electron paths leading to interference effects.

In these systems, we can define two length scales, the elastic or phase coherence length, $l_\phi = \sqrt{D\tau_\phi}$, and inelastic or phase breaking length, $l_{in} = \sqrt{D\tau_{in}}$. In defining these length scales, we have used a diffusion constant, $D = v_F^2\,\tau/d$, (with d being the sample dimensionality), a phase coherence time, τ_ϕ, and a phase breaking time, τ_{in}. The inelastic lifetime generally depends on temperature, following a power law variation $\tau_{in} \propto T^{-p}$ (as discussed in section 5.2.1.4), but it is also dependent on the sample dimensionality. For electron coherence effects to modify the resistivity, the following condition on the electron mean free path exists, $\bar{l} < l_\phi < l_{in}$. Therefore, in figure 5.4, if $\bar{l} > l_{in}$, the scattering at point P is incoherent and the wavefunction *amplitudes* add to give the amplitude of the electron wavefunction at the point P of:

$$|\psi_P|^2 = |\psi_1|^2 + |\psi_2|^2 = 2|\psi|^2. \qquad [5.17]$$

If, on the other hand, $\bar{l} < l_\phi < l_{in}$ then phase coherence is preserved and the *wavefunctions* add, giving:

$$|\psi_P|^2 = |\psi_1 + \psi_2|^2 = 4|\psi|^2. \qquad [5.18]$$

Therefore, the probability of an electron being scattered to the point P is double where coherent, phase preserving scattering occurs compared with the case where incoherent,

phase breaking scattering occurs. Therefore, where sample dimensions are much greater than these coherence length scales the mean free path is dictated by the material properties and coherence effects are averaged out if $\bar{l} > l_{in}$. Reducing the sample dimensions, L, below the intrinsic mean free path limits the mean free path to the scale of the sample, that is $L < \bar{l}$. A situation can arise where interference effects occur because the sample dimensions meet the requirement of $L < l_\phi < l_{in}$, and the sample behaviour becomes size dependent

The two pictures of figure 5.4 relate to two types of interference effect. Figure 5.4a depicts coherent backscattering which is termed weak localization[15]. This effect leads to resistivity increases in the sample (or in a few cases where magnetic impurities exist, giving destructive interference, resistivity decreases). The case depicted in figure 5.4b represents a less specific case of coherent scattering which can lead to non-linearities in the $I(V)$ behaviour of small samples and also what is termed universal conductance fluctuations, UCFs.

This description of weak localization due to the sample dimensions can be counted as a modification to the classical (Drude) conductivity as[16]:

$$\sigma_{2D} = \sigma_o - \frac{e^2}{\hbar\pi^3} ln\left(\frac{L}{\bar{l}}\right), \qquad [5.19]$$

$$\sigma_{1D} = \sigma_o - \frac{e^2}{\hbar\pi}\left(L - \bar{l}\right), \qquad [5.20]$$

where L is the sample size and T=0. Therefore, the larger the sample (for the same degree of disorder, ie. same mean free path, $\bar{l}$), the greater the effect of coherent back scattering on the sample resistivity. Where the effect of temperature is concerned, assuming a power law temperature dependence as noted previously ($\tau_{in} \propto T^{-p}$) we find the following:

$$\sigma_{2D} = \sigma_o + \frac{pe^2}{2\hbar\pi^3} ln\left(\frac{T}{T_o}\right), \qquad [5.21]$$

$$\sigma_{1D} = \sigma_o - \frac{ae^2}{\hbar\pi} T^{-p/2}, \qquad [5.22]$$

where a is a proportionality factor and T_o is a constant. Here we see that in both the one and two dimensional cases, resistivity must rise with decreasing temperature. In the two dimensional case, the $ln(T)$ dependency agrees with experimental observations[17] however, the magnitude of the effect is wrong[18]. In two dimensions, disorder induced

electron-electron interaction effects lead to the same $ln(T)$ dependency[19]. However, application of a magnetic field clearly identifies the differences between the two mechanisms as follows. The two paths shown in figure 5.4a which lead to the interference effect of weak localisation can be considered to be the time reversal of each other. That is, $\psi_1(t)=-\psi_2(-t)=\psi_1^*(-t)$. This requires that the electron paths are time-reversal invariant if equation 5.18 is to hold. Application of a magnetic field breaks the time reversal invariance of the electron trajectories and therefore reduces the constructive interference at point P. Increase of this magnetic field reduces the effect of coherent backscattering and the resistance will begin to decrease, with low magnetic fields. This negative magneto-resistance in the sample continues until the magnetic length, $l_B=eB/m$, becomes smaller than the coherence length, l_ϕ[20]. The electron interaction effect is unaffected by **B** at low fields, but at high fields, where strong spin scattering of electrons exists, the electron-electron interaction is affected, allowing the different mechanisms to be clearly distinguished.

Universal conductance fluctuations, UCFs, are another manifestation of quantum interference at defects. UCFs results from individual electron paths being affected by the presence of magnetic fields. This produces repeatable fluctuations in the magneto-resistance of a sample[21] providing a sensitive measure of the impurity distribution within an individual sample. The study of interaction effects has shown that the phase information in an electron wave-function is not lost by diffusive transport. It is possible therefore, to experience effects which are due to the interference of electron wave-functions which lead to non-linear or apparently anomalous behaviour within small samples. In figure 5.4b, it can be seen that the presence of a magnetic field will move the electron paths off the trajectories shown. Therefore, the magnetic field changes the interference effect which in turn changes the sample conductance. This change is repeatable in magnetic field and is a function of the exact defect distribution within the sample.

The effects of UCFs, weak localization, interaction effects, etc. arise from the fundamental quantum nature of the electron. However, as with all quantum effects, the manifestation is smeared by energy input to inelastic scattering. This energy input is provided mostly by thermal energy but heating due to bias voltages and currents must also be considered. In examining samples for these effects it is most important that the energy input, whether temperature or electrical bias, is kept well below the critical or characteristic energy of the effects in question. That is, $E_c>k_BT$ and $E_c>eV_{bias}$.

5.2.3 CONDUCTION PROCESSES IN GRANULAR FILMS

Almost all thin films (as deposited) exhibit some degree of polycrystallinity in the structure. This polycrystallinity, or granularity, results from the nucleation and growth process as discussed in section 4.4.2. The resulting increase in the disorder of the material affects the conduction characteristics. Three main features can be identified which relate to this disorder. Firstly, grain boundaries are included in the conduction path. These produce an increased defect density within the material, resulting in greater scattering of conduction electrons. Secondly, where preferred crystallite orientations exist, anisotropic conductance behaviour can result. Thirdly, the chemical nature of the crystallite boundary can determine the bulk of the material properties. The chemical composition of this boundary can vary from a few dislocations, through a contaminated crystallite surface to a significant insulating or alloying component in the material matrix.

Work on granular metal films has identified three regimes within which individual films can be found: 1. the metallic or continuous regime, 2. the transition or semi-continuous regime and 3. the dielectric or discontinuous regime. The following sections briefly describe each of these regimes.

5.2.3.1 METALLIC OR CONTINUOUS REGIME

Granularity of the film introduces scattering into the conduction path and roughness into the top surface of the film. Anisotropy of electron conduction through the film can arise through preferred orientations of the film crystals. However, it is important to note that with deposited metal films, preferred orientation of crystallites is not common. To model the effect of grain boundaries, Mayadas & Shatzkes[22] produced a simplified model which requires elastic scattering. The model assumes columnar grains, resulting in grain boundaries oriented perpendicular to the conduction path. A scattering power parameter, α, is defined which is a function of the bulk mean free path, $\bar{l}$, the grain radius, r, and a reflection coefficient, R, of the boundary. Their definition is:-

$$\alpha = \frac{\bar{l}}{r}\left(\frac{R}{1-R}\right). \qquad [5.23]$$

So for large r (compared with $\bar{l}$) or small R, grain boundary scattering has little effect on the overall film resistivity. Therefore for small α the presence of grain boundaries has very little effect on electron transport. The ratio of background to grain boundary resistivity, ρ_g, is then found to be[22],

$$\frac{\rho_0}{\rho_g} = 3\left[\frac{1}{3} - \frac{1}{2}\alpha + \alpha^2 + \alpha^3 ln\left(1 + \frac{1}{\alpha}\right)\right]. \qquad [5.24]$$

An alternative method of viewing conduction in this region, as studied by Abeles et al[23], is given by examining the effect of the volume fraction of metal in an insulator, x, on the resistivity of the granular film. Using a percolative conduction model[24], a critical composition, x_c, will exist where crossover from metallic to tunnelling conduction will result. Metallic conduction near this crossover can be modelled as percolative, or random walk, electron transport. This is because any system with an insulating component forming thin barriers to conduction will have tunnelling conduction and metallic conduction processes existing in parallel. The crossover to tunnelling conduction will result when the metallic path is disconnected or there is no percolative path connecting both ends of the film. From the work of Abeles et al, x_c was found to be around 0.5 for unannealed granular films of co-sputtered metal and alumina or metal and silica. This result was higher than that predicted by theories[25], but attributed to the fact that the insulator tends to coat metallic particles in this type of film rather than producing a uniform composition material.

5.2.3.2 TRANSITION OR SEMI-CONTINUOUS REGIME

So far, the discussion of conduction processes in thin films has concentrated on films on the metallic side of the metal insulator transition. The metal insulator transition was first predicted by Mott[26]. He predicted that when $k_F\bar{l} = 1$, (k_F being the Fermi wavevector), an abrupt metal-insulator transition would occur, with the density of states at E_F going to zero at zero temperature. Consequently, the conductivity would become zero at zero temperature. However, due to other sources of scattering and electron coherence and interaction effects, such an abrupt transition is not observed as a rule in thin metal films.

We have already mentioned that thin films contain greater sources of scattering than bulk materials and that dependent on the length scale of the conduction process, quantum effects can introduce further sources of scattering through scale dependent effects. Increasing the disorder within a thin metal film moves the material closer to the metal-insulator transition through the introduction of more disorder. This disorder can be introduced into the film in a number of ways e.g., through incorporation of impurities or insulating material in the film structure or through successive reduction in the film thickness until the film becomes semi-continuous or discontinuous.

In this region, we are considering the condition where the metallic volume fraction is

approaching the critical value ($x \approx x_c$). As mentioned above, the material now exists with a labyrinth type conduction path. Electrons must follow a percolative or random walk type path to take part in the conduction process. For a discussion of this type of conduction see Coutts[27]. On the other hand, describing the effect using weak localization type ideas, the electron wave packet is becoming more and more spatially confined and the strength of the localization effect increasing. Eventually, through increased disorder, we are dealing with a material in the regime of strong localization. Such materials behave more like semiconductors or insulators rather than metals. It is in this region that the maximum metallic[28] resistivity is reached, with resistivities of the order of 1×10^{-4} to $1 \times 10^{-3}\Omega$m. As this resistivity is approached, from the metallic side, the TCR begins to approach zero and the resistivity begins to rise rapidly. Increasing the disorder, or decreasing the metallic volume fraction of the material pushes the TCR more negative and moves the film across the metal-insulator transition and into the next region.

An interesting feature of this transition region in granular metal films is the enhancement of the superconducting transition temperature (T_c) in superconducting materials[29]. Increases in the film sheet resistivity lead to corresponding increases in T_c. This increase continues to a maximum in T_c, with the width of the transition broadening and the variation in the T_c increasing, see section 5.2.4 and figure 5.5. At this maximum T_c, the sheet resistivity is near the maximum metallic resistivity. As the resistivity increases above this value, the T_c drops again until the degree of disorder, or discontinuity within the film, becomes too great to support the superconducting state.

5.2.3.3 DIELECTRIC OR DISCONTINUOUS REGIME

The structure of the film in this regime can be described as isolated metallic islands surrounded by a non-conducting medium. Electron transport within the film must therefore take place by either a tunnelling or an activated process (or possibly both). For this reason a large number of theories have been suggested for the conduction processes within these films[30]. Some of the theories are applicable at certain limits of electric field or frequency of field but by far the most successful have been those theories based on tunnelling transport between islands.

Gorter's[31] description of the importance of the electrostatic energy on the electron transport between two islands lead to the postulation from Neugebauer and Webb[32] that energy must be supplied when an electron is transferred between two initially neutral islands. After Gorter's suggestion, Neugebauer and Webb suggested that only electrons having energy e^2/r or more above the Fermi energy could be transferred. r in

this case is the island or particle radius. Assuming Boltzmann statistics, the total number of available charge carriers would then be:

$$n = N exp\left(-\frac{\Delta E}{k_B T}\right), \qquad [5.25]$$

with $\Delta E = e/4\pi\varepsilon r$.

The probability of electron transmission between islands was then calculated. The resulting conductivity displayed the following characteristics; inversely varying with island size, directly varying with the product of island separation and transmission coefficient, and exponentially varying with inverse temperature. The result was in broad agreement with experimental observations, but two main objections were raised. Firstly, the activation energy is part of the barrier which electrons must tunnel through and secondly, the transmission coefficient is a function of energy. Hill[33] removed these objections in his model by including these features in his theory. The resulting formula for current density reduces to Simmons'[34] tunnelling current density when ΔE tends to zero, as would be the case for very large metallic islands.

Objections to this theory relate to the following facts: Firstly, on the basis of the behaviour of three identical metallic particles, with identical barrier characteristics, the current density for a whole array has been calculated. This means that no account has been taken of the variation in particle size or separation. Secondly, no account has been taken of the effect of the substrate on the conduction process through the film. Thirdly effects due to the quantization of energy levels within the small particles in the structure have been ignored.

The first objection is by far the most serious when we consider we are dealing with materials which contain significant variations in island size and spacing. The effects of the substrate have been accounted for through theories relating to substrate assisted tunnelling[35] and effects due to discrete island energy levels have been addressed by Hartman[36] and others. Abeles et al[37] carried out significant work on addressing the problem of variations in the film island network. This work produced a theory based on tunnelling which provided a good fit between experimental data and the theoretical predictions for granular films[23].

The most important feature of this regime is the resistivity temperature dependence. Resistivities are much greater than $10^{-4}\Omega$m, and the TCR is negative as we are dealing with a thermally activated conduction process. This gives rise to a logarithmic

temperature dependency of $ln\rho \propto T^{-1}$.

5.2.4 CONDUCTION EFFECTS OBSERVED IN ALUMINIUM

Aluminium (having electronic configuration $[Ne]3s^23p^1$) is tri-valent. The valence electrons arising from the 3s and 3p orbitals produce the common oxidation states of Al^+ and Al^{3+}. In the solid state, the metal has an FCC structure with three valence electrons per primitive cell, or atom in this case. The free electron Fermi sphere for tri-valent metals in an FCC Bravais lattice extends into the fourth Brillouin zone. The fourth zone only contains small pockets of electrons, which disappear when the perturbation due to the weak periodic potential is considered. The first Brillouin zone is therefore full, with two valence electrons[8]. The remaining two zones are filled with the remaining valence electron. It is this property of the metal which leads to the high field Hall effect result indicating a carrier density of one hole per primitive cell[39]. That is, $n_o/n = -1/R_H ne = -1/3$. High field in this case is defined as $\omega_c\tau > 10$, where $\omega_c = eB/m$ is the electron cyclotron frequency.

Within thin aluminium films, the behaviour of the resistivity and structure has been studied and compared with Fuchs' size effect theory by Mayadas, Feder and Rosenberg[12]. This work examined film thicknesses from around 80nm to 1μm. The resistivity behaviour was found to change dramatically with thickness in the region of 450nm to 600nm thick. This feature was directly correlated with a change in the film structure which produced a preferred {111} orientation in the grains of the thicker films, as opposed to the random orientation with thinner films. The resistivity data collected for these films was fitted to Fuchs' theory. However, they required to use the parameters ρ_i and l_i, as fitting parameters. This is unsatisfactory, as these represent the intrinsic resistivity and mean free path, respectively, of a material with the same resistivity characteristics as the film being studied, but being infinitely thick. This suggests that the film characteristics are not well modelled by this classical theory. Also, other thickness dependent processes are responsible for changes in the conduction processes within the film.

Later studies of the resistivity of thin aluminium films looked at the effects of electron interaction and localization. Within two-dimensional films, the separate effects of weak localization and electron interactions were clearly identified[18]. In this case it was noted that localization effects dominate the magneto-resistance behaviour of aluminium at low temperatures, while electron interaction effects are required to understand the resistivity temperature dependence. The mean grain size of these films was around

3nm, with the thickness being around 1μm, resulting in three dimensional films. Further work[40] looking at the behaviour in one dimensional wires (0.2μm < W < 0.6μm, W being the wire width) in thin (15 to 25nm) films provided good agreement with 1D localization theory when modifications for super-conducting fluctuations were added to account for effects below 2K. This work extracted the inelastic diffusion length, l_i for these materials through the use of the "inelastic scaling field" as a fitting parameter. In the case of a sample of width $W = 0.2\mu$m, and thickness 25nm, $l_i \approx 1.5\mu$m at 4.5K, dropping to around 0.2μm by 20K, thereby showing the detrimental effect of inelastic scattering through temperature on these interference effects.

Enhancement of the superconducting transition temperature, T_C, has been studied with the resistivity of aluminium films. This T_C enhancement was found to be influenced by reductions in the metal grain size[41,42], and through reductions in the film thickness[43]. A maximum T_C of around 2.5K has been observed in Al/AlO_x granular films[44], along with the observation that the T_C peaks at $\rho(4.2K) \approx 10^{-5}\Omega m$. The effect of T_C enhancement has been described by Hauser[44] as resulting from reduction in the average phonon frequency. Reducing the film thickness and increasing the resistivity by oxidation both produce lower density thin films when compared with bulk material densities. This lower film density corresponds with a larger crystallographic volume, thus lowering the average phonon frequency. Alloys of Al and Ge where a corresponding metal grain volume increase results, have also exhibited this increase in T_C. The resistivity behaviour of these films with $\rho(4.2K) > 10^{-5}\Omega m$ was of the form $ln\rho \propto 1/T$. This T_C enhancement effect was studied, along with the resistivity required to induce a metal insulator transition[45]. In this work, the metal insulator transition was identified through superconductive tunnelling measurements of the electronic density of states. The film resistivity at which the density of states around the Fermi energy ($N(E_F)$) became zero was taken as the resistivity required to induce a metal-insulator transition. The resistivity at the metal insulator transition in Al was found to be around $\rho_{MI} \approx 1 \times 10^{-3}\Omega m$.

In the discontinuous regime, the $ln\rho \propto 1/T$ behaviour and various aspects of the field dependence of the film resistivity have been confirmed through work carried out by Abeles[23] and others. However, the exact process and the influence of various components within the material system are still a matter for debate. Aluminium CERMETs, or "ceramic metals" provide an example of materials which derive their usefulness from the high resistivity and stable TCR characteristics of this regime. These materials are formed from a discontinuous metal phase in an insulating (usually refractory oxide) material. They provide high sheet resistances, of up to 10MΩ per square, with a low and very stable TCR, making thin films of Al/Al_2O_3[46], for

example, useful for high value resistors in electronics applications.

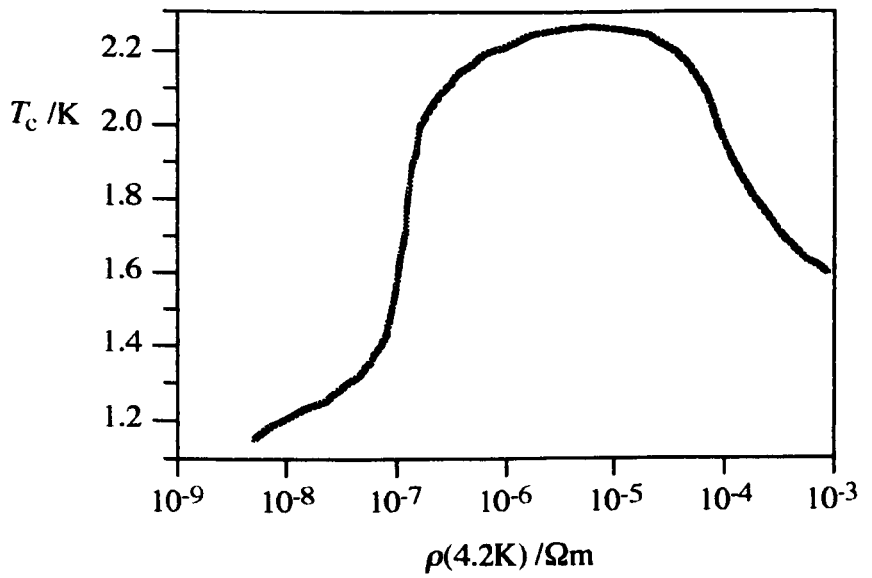


Figure 5.5 **Sketch of the T_c vs resistivity behaviour for Al/AlO_x films. After Dynes and Garno[45].**

A number of the films mentioned above (which are on or near the metallic side of the metal insulator transition) are similar in physical characteristics to those used in this work. Depositions were carried out under high vacuum conditions, usually in an oxygen partial pressure. The film thicknesses in a number of cases were comparable with those used in this work. It is therefore reasonable to use these results to predict and describe the effects observed in the films used here.

5.3 RESISTIVITY MEASUREMENT TECHNIQUES

Potentially the simplest method of determining resistivity is to prepare a sample of known dimensions, determine the sample resistance and apply equation 5.1, see for example figure 5.6. This method is subject to errors introduced from both the sample preparation and the measurement. When preparing a sample, the dimensional errors need to be known. Particularly with thin films, the error on the thickness measurement can be significant. This error can be somewhat offset by measuring the sheet resistance, which includes the film thickness in its definition as shown in equation 5.26.

$$R_s = \frac{\rho l}{A} = \frac{\rho l}{wt} = R_{sh}\frac{l}{w}, \qquad [5.26]$$

where $R_{sh}=\rho/t$ is the sheet resistance.

Similarly, when measuring the resistances of thin metallic samples, contact resistances can easily dominate the measurement. Whether the sample is being measured using a simple bridge circuit or by constant current application and measurement of the developed voltage, a contact resistance of a few ohms can easily dominate a measurement on a sample where $\rho \approx 1$ to $10\mu\Omega$cm.

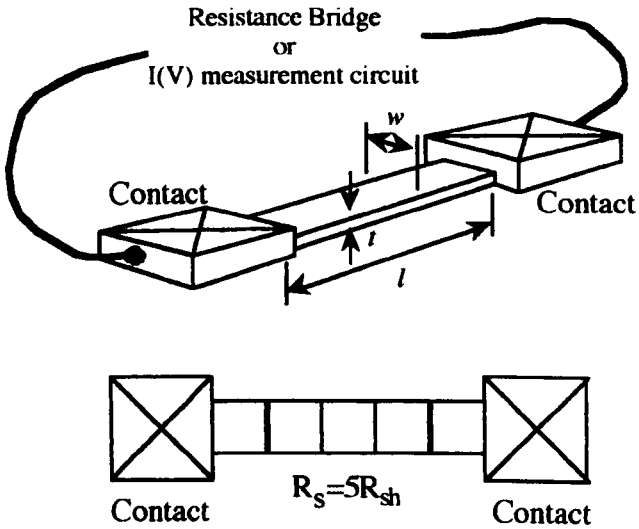
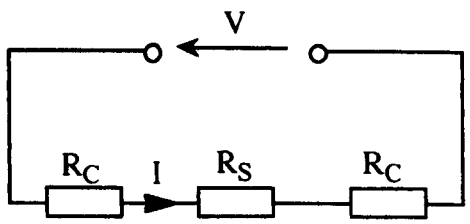
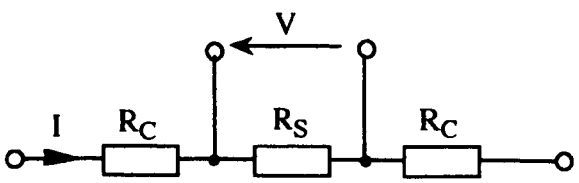


Figure 5.6. Resistivity and Sheet resistance sample.



$V=(2R_C+R_S)I$

a) Two Point Measurement



$V=IR_S$

b) Four Point Measurement

Figure 5.7. Comparison between 2 and 4 point probe measurements.

To overcome the difficulty of contact resistances, four point probe techniques need to be applied. Figure 5.7 demonstrates the application of this technique. The use of two probing contacts in addition to the current source/sink contacts eliminates the error in voltage measurement due to contact resistances. The voltage measurement is now independent of the current sourcing circuit and its associated resistances. Provided the

voltmeter input impedance is high enough, no significant current loading of the measurement will occur and the only current path will be that through the sample.

Finally, where sample resistivities are small, noise sources can generate significant errors in the measurement. In the case of thin film samples, there is a real constraint in the current levels which can be applied to a sample before the film is damaged by electromigration. Where we have small resistivities, significant currents are generally required to generate measurable voltage drops across the sample. It is therefore important to avoid coupling noise from the environment into the measurement circuit. It is still possible that limitations in the equipment used can cause the measured signal levels to be swamped in noise. In this case, frequency lock-in techniques can be applied, as described in section 5.3.2.

5.3.1 SHEET RESISTANCE

The most direct method of determining the sheet resistance is as shown in figure 5.6. The number of squares in the sample is simply counted as the length between the voltage probes divided by the sample width. Where four point probing is used, this situation can become slightly more complicated. In this case, two arrangements were employed, these are the Hall bar and the van der Pauw square.

5.3.1.1 THE HALL BAR

In the Hall bar arrangement, the sample is arranged with voltage probes as shown in figure 5.8. Provided the probe spacing, l, and the bar width, w, are large compared with the film thickness, the sheet resistance can be determined as in equation 5.26. If this is not the case, correction factors must be applied to measure the resistance, such as those for the in-line probe arrangement[47]. In the case of metallic films, the constraints on probe spacing related to film thickness arise from the electron mean free path in the film. If structure dimensions are many times the electron mean free path, conduction can be assumed to be diffusive and therefore the sample response is linear. In this case Ohm's law holds. If the mean free path is of the order of the structure dimensions, however, ballistic electron effects may determine the characteristics of the sample behaviour, leading to non-linearity in the $I(V)$ characteristics and non conformity to Ohm's law. In all the films examined here, the electron mean free path is much less than the structure dimensions.

The arrangement of the Hall bar provides for measurement of both the longitudinal, V_L, and transverse, V_T, voltages. This degree of flexibility allows measurement of the Hall

resistance of a sample under the action of a magnetic field. Measurement of the longitudinal voltage against applied current with varying magnetic field allows measurement of the sample's magneto-resistance.

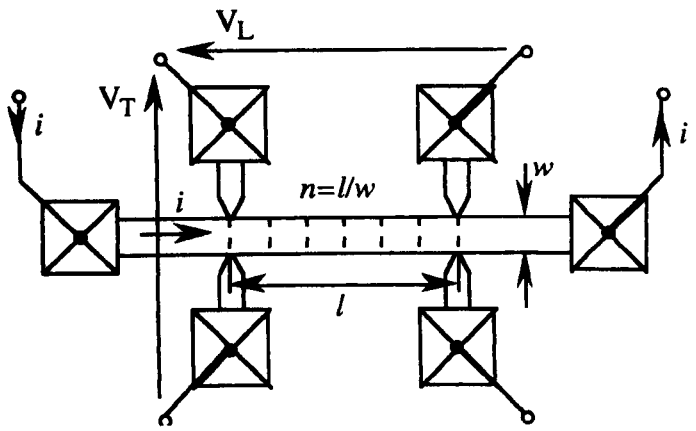


Figure 5.8 Hall bar sample arrangement.

5.3.1.2 THE "VAN DER PAUW SQUARE"

The "van der Pauw square" arises from a special case of a general arrangement devised by L.J. van der Pauw to allow sheet resistance determination on an arbitrary shape[48]. In this case a sample, depicted in figure 5.9, has an arbitrarily placed four point probe arrangement.

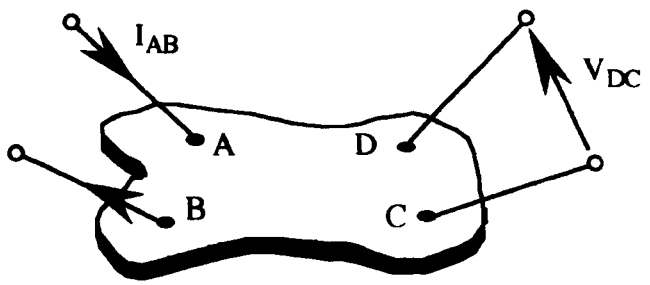


Figure 5.9. "van der Pauw" sample arrangement.

The current and voltage probes form a polygon. Current is passed between two terminals on one side of the polygon and the voltage generated on the opposite side of the polygon is measured. Rotating the contacts through 90° and repeating the measurements on the other sides of the polygon completes the measurement. The two sets of measurements are then related through equation 5.27, with reference to figure 5.9 for the senses of the currents and voltages.

$$exp\left(-\frac{\pi t}{\rho}\frac{V_{DC}}{I_{AB}}\right) + exp\left(-\frac{\pi t}{\rho}\frac{V_{AD}}{I_{BC}}\right) = 1. \qquad [5.27]$$

From this expression, the resistivity or sheet resistance ($R_{sh}=\rho/t$) can be determined. For the special case of an isotropic square film or, the "van der Pauw square", we find that opposing voltages and currents are symmetric, that is:

and
$$I = I_{AB} = I_{BC} = I_{CD} = I_{DA}. \quad [5.28]$$

$$V = V_{DC} = V_{AD} = V_{BA} = V_{CB}. \quad [5.29]$$

These voltages and currents being grouped in the order shown. Therefore, with these groupings and $R_{sh}=\rho/t$, equation 5.27 reduces to,

$$2exp\left(-\frac{\pi}{R_{sh}}\frac{V}{I}\right) = 1 \quad [5.30]$$

which re-arranges to give,

$$R_{sh} = \frac{\pi}{ln2}\left(\frac{V}{I}\right) \quad [5.31]$$

Therefore, comparing the voltages and currents noted provides an indication of the symmetry and isotropic nature of the sample, and the validity of the measurement.

5.3.2 FREQUENCY "LOCK-IN' TECHNIQUES

Where measurement of low level current and voltage signals is required in the presence of noise, it is possible to recover some of the signal from the noise by modulating the signal and looking for the response at the modulation frequency. Applying an ac. driving current to the sample while measuring the sample response in a phase locking circuit allows recovery of the response at the excitation frequency. The value returned is a root mean square of the ac. response along with the magnitude of any phase delay between the excitation and response. This assumes that the sample provides a linear response to the excitation signal or in other words, the response to the applied signal is not a distorted version of the input. Therefore, the applied excitation amplitude must be small enough so as not to average out any non-linearities in the sample behaviour. Figure 5.10 displays a schematic diagram of a lock-in amplifier circuit.

In this circuit a differential pre-amplifier is used for the first amplification stage. This signal will contain the ac. modulated response of the sample which is originally derived from the reference signal. Multiplying the output of the pre-amplifier and the reference signal picks out the signal at the reference frequency. It is possible to introduce phase delays to the reference signal to maximise the result of this multiplication, thus providing a measure of the phase delay between the applied signal and the measured

response. In this way the presence, and magnitude of any reactive elements in the sample response can be determined.

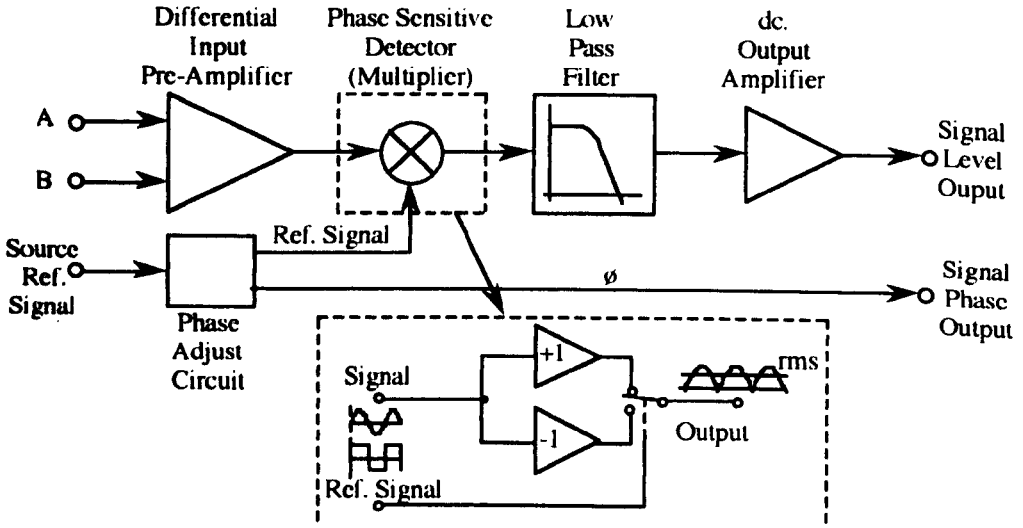


Figure 5.10 Schematic Diagram of Lock-in Amplifier Circuit.

Figure 5.11 displays an example of the type of circuit used for measuring a sample using this technique.

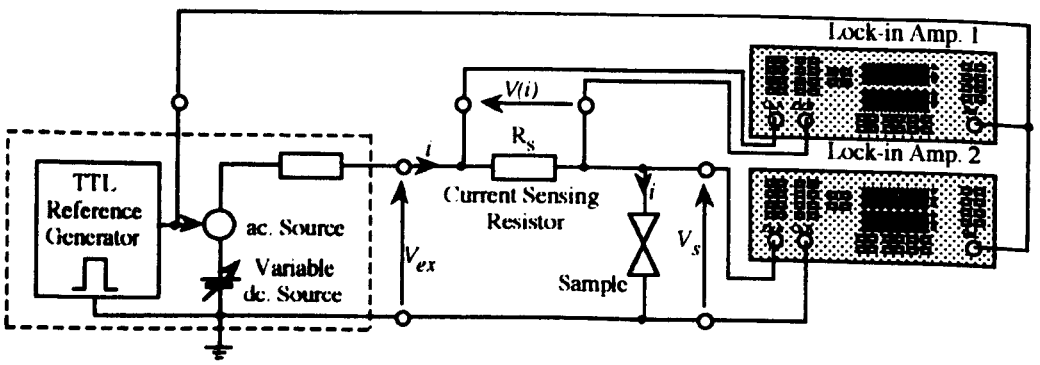


Figure 5.11 Schematic of Lock-in Four Point Probe Circuit.

In figure 5.11, we see that the circuit contains a resistive current sensing element, R_S, generating the measurable voltage, *V(i)*, which will be translated into a current. The current gain is therefore a combination of the gain setting of the amplifier used and the value of R_S. The sample voltage is directly measured by another lock-in amplifier.

It is important to note that this type of measurement examines the zero bias small signal response of the sample, ie. a measurement of *dV/dI* at *V*=0. Provided the response of the sample is linear around *V*=0, *dV/dI* represents the sample zero bias resistance. Equally, the sample conductance can be measured as *dI/dV*. As lock-in amplifiers reject frequency components away from the reference frequency, it is possible to sample the bias dependencies in the sample *I(V)* curve. Addition of a dc-bias to the ac. sample

excitation, allows the position of the *dV/dI* sample point to be moved within the *I(V)* curve. Taking, as an example, the *I(V)* curve of a current biased single electron transistor, we see it is possible to sample the bias dependencies of the sample response, as in figure 5.12

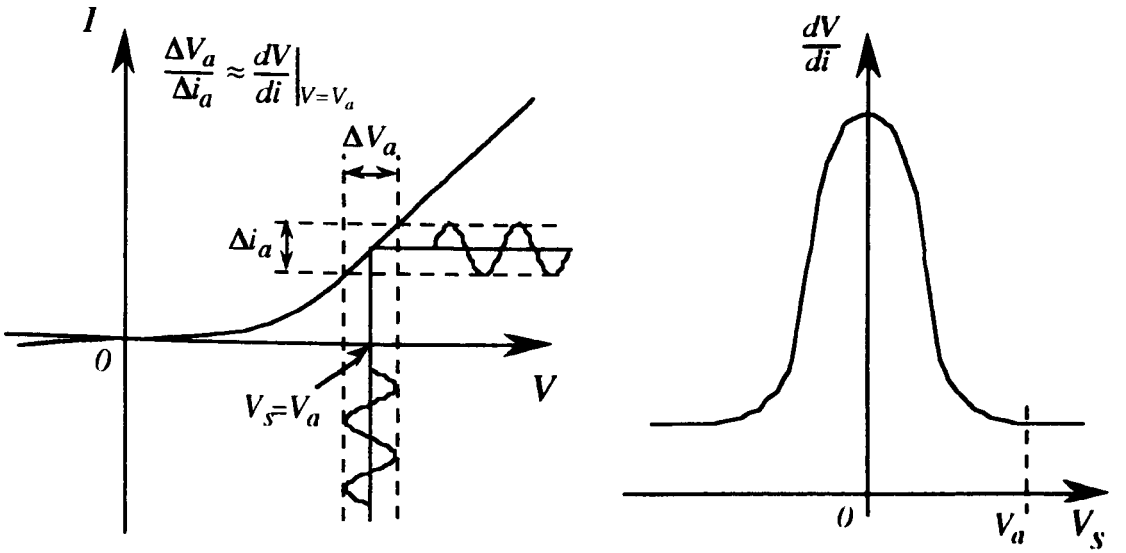


Figure 5.12 Use of ac. Techniques to Sample Differential Resistance.

It is important to note that while frequency lock-in techniques help recover signals from a noisy background, they do not reduce the level of noise in a system. Therefore, any effects which occur in the device due to the noise spectrum will still exist. In the simplest case, electron and sample heating can occur, raising the real temperature above the sample environmental temperature. On the other hand, where samples respond to particular harmonics in the frequency spectrum, it is possible for higher frequency signals to be down-mixed by the action of the device. This can result in measurement of effects due to harmonic mixing and not the device behaviour at the frequencies of interest.

5.4 RESISTIVITY MEASUREMENT SYSTEM

All temperature dependent resistivity measurements were performed in the same cryostat, with nominally the same circuit arrangement. Two main experimental set-ups were used. The first allowed for bonding directly to the aluminium films using small indium balls. This bonding was carried out by pressing the indium into the film and attaching wires to the indium balls. The second set-up utilised a standard ceramic chip carrier which allowed for gold wire bonding to the sample through the use of an ultrasonic wedge bonder. Bonding to this type of chip carrier required the use of the gold lead-frame described in chapter 3. This type of sample preparation allows the

samples to be used in other equipment. Initially, where manual bonding of samples was utilised, there was no facility available on the cryogenic insert for measuring temperature directly at the sample. This was later added for measurements on films deposited using the UHV system. Samples prepared using the ceramic chip carrier were mounted on a cryogenic insert with a temperature sensor fitted.

5.4.1 CRYOSTAT

The cryostat used was a Cryogenic Consultants ^{4}He Optical Cryostat. The system was equipped with a superconducting magnet capable of generating magnetic fields of up to 6 Tesla when driven from an Oxford Instruments Magnet Power Supply. Samples were inserted into the variable temperature insert (VTI) on an insert rod from the access port on the top of the cryostat. Samples were placed on the sample holder on the end of the insert, positioning the sample in the centre of the magnet poles. These magnets were arranged such that a horizontally mounted sample experienced magnetic fields perpendicular to the plane of the sample substrate.

Temperature in the VTI space could be varied from 1.2K to near room temperature. Heating of the sample space was effected by a heating element positioned around the sample space. Cooling to 4.2K took place by pumping cold ^{4}He gas from the helium reservoir through the sample space. A needle valve positioned between the helium reservoir and the VTI space controlled the gas flow rate. It was possible to fix the temperature through use of the heating element and control of the ^{4}He gas flow rate with the needle valve. To vary the temperature between near 1.2K and 4.2K, liquid ^{4}He was pumped into the VTI space. The needle valve was then closed and the pressure above the liquid ^{4}He reduced by pumping, thus dropping the liquid temperature. In practice temperatures less than 4.2K could be maintained for only a short time due to the heat loading from wiring on the electrical inserts, the insert rod itself and from the system in general.

5.4.1.1 TEMPERATURE MEASUREMENT

Temperature measurement could be effected by three methods and the method used depended on the VTI insert used and the equipment available. The first, and preferred method, used a rhodium-iron resistance thermometer mounted on the sample end of the insert. The sensor was driven by an Oxford Instruments ITC4 temperature controller which allowed direct measurement of the sample temperature after calibration. Calibration of this sensor/transducer combination took place via two set points, room temperature and liquid helium at atmospheric pressure (4.2K). The second method used

the temperature sensor on the heater plate at the bottom of the VTI. This method, however, did not accurately reflect the sample temperature as this sensor was not directly connected to the sample holder. The errors produced by this method were not linear. At liquid helium temperatures, the heater plate remained above the sample temperature and at high temperatures, the heater plate was able to reach temperatures much higher than the sample. Thirdly, where liquid helium is present in the sample space, the pumping line pressure can be measured. This pressure can then be related to the liquid temperature, or sample temperature provided the sample is immersed in liquid and the gas volumetric flow rate in the pump line is small enough to neglect line pressure losses. Where an insert was used without a rhodium-iron sensor, the low temperature (below 4.2K) measurements used the third method and all other measurements used method two.

5.4.1.2 SAMPLE INSERTS

During the thin film resistivity measurements, two inserts were used. The first insert could accept samples where the connections were made with pressed indium beads from which enamelled copper wire was soldered onto gold connector pins. This insert was initially used on films deposited in the HV, or "modified", evaporator as described in section 3.5. In this instance no temperature sensor was fitted to the insert. Temperature measurements were made through a combination of line pressure measurements and heating plate temperature readings. This resulted in some degree of hysteresis in the region from 20K to 200K. A rhodium-iron temperature sensor was fitted to this insert for measurements on films deposited under UHV conditions,

The second insert contained a sample holder which could accept 18pin ceramic chip headers. This chip header was a standard layout leadless carrier which can be used on various pieces of equipment. This insert was equipped with 18 signal leads, 2 leads connected to an LED, and leads to operate a rhodium-iron temperature sensor. To use this header with aluminium films, gold wires were ultrasonically bonded to the lead-frames fabricated on the sample. This insert was used for some aluminium film measurements but also for measurements on the characteristics of the lead frame thin film interconnect of section 3.4.2.2.

5.4.1.3 SAMPLE AND INSERT WIRING

Each insert terminated outside the cryostat in 2 x 10 pin connectors, which allowed connection to each of the electrical contacts with individual co-axial cables of type RG-316U. Both inserts used enamelled copper wire to connect between the sample contacts

and the connector on the top of the insert. The co-axial cable connected each contact to a numbered BNC connector on a wall mounted patch box. See section 5.4.2.3 for a fuller description of this box. The use of the patch box and co-axial cable was designed to help ensure shielding of the signal lines. The connection of the co-axial cable shielding was designed to prevent ground loops in the measurement circuit.

5.4.2 MEASUREMENT CIRCUIT

Figure 5.13 provides a system view of the circuit used to perform the resistivity measurements. The sample arrangement is generally as depicted in figure 5.11. The circuit elements are described in the following sections.

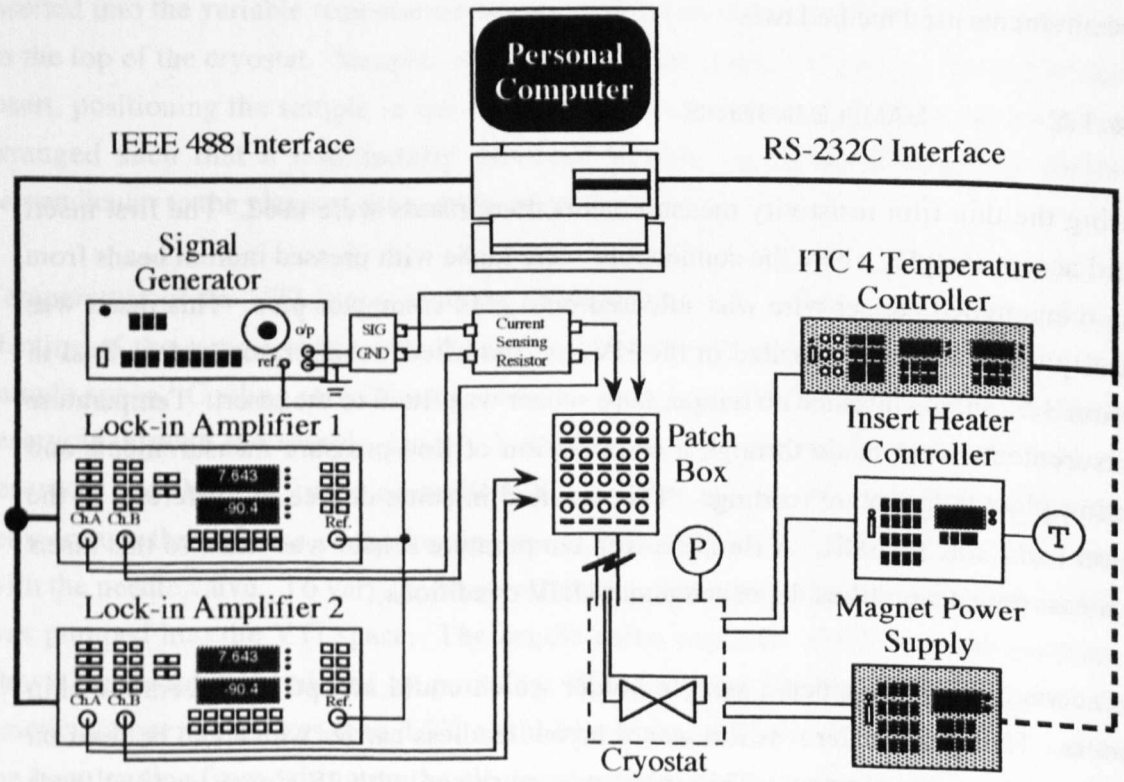


Figure 5.13 Schematic of Measurement System.

5.4.2.1 SIGNAL GENERATOR

The signal generator used was a Feedback Instruments FG601 0.2mHz to 10MHz sine / triangular / square - wave generator. This signal generator provides a 5V TTL reference output signal which has a 90° lead on the generator output signal. The function generator signal output, in sine-wave mode, provides the signal used as the sample excitation. The reference and output signals were applied to the circuit elements via coaxial cables.

5.4.2.2 CURRENT SENSING RESISTOR

The current sensing resistor was switchable through 1kΩ, 10kΩ, 100kΩ and 1MΩ. All resistors were of metal film type, having 1% tolerance with 100ppm°C^{-1} temperature coefficient. The resistors and switching mechanism were contained within a metal box, the signal lines were connected via co-axial cables with BNC connectors.

5.4.2.3 PATCH BOX

The patch box connects the chosen insert connection to the sample measuring circuit. The configuration of the circuit is therefore determined by making the appropriate connections from the measurement circuit to this patch box. Figure 5.14 provides a schematic of the box connections. The entire box is metallic, with the front panel BNC connectors being electrically isolated from the box connections to prevent ground loops. The entire circuit is wired to provide a single earthing point.

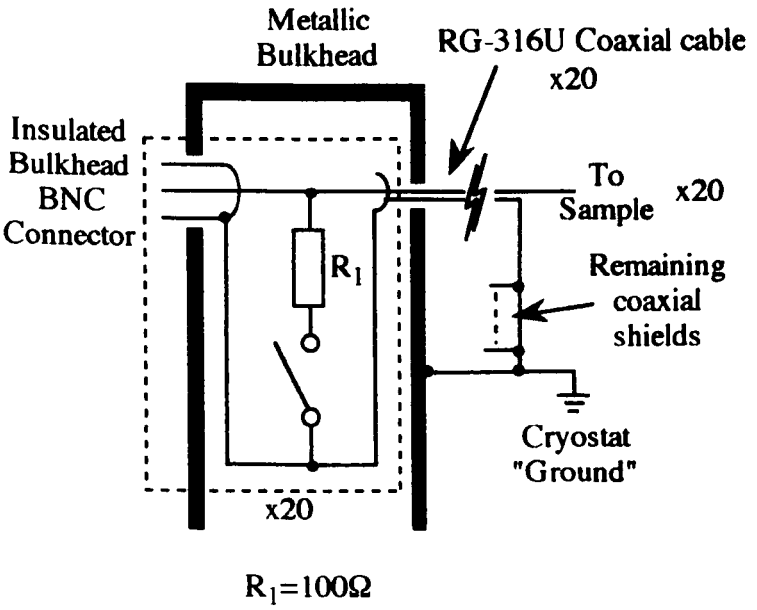


Figure 5.14 Patch Box Circuit Diagram.

The 100Ω resistor connected the shield to the coaxial inner connector allowing charge dissipation and prevention of ESD damage to the samples. This resistor was switched out of the circuit before measurements were taken.

5.4.2.4 LOCK-IN AMPLIFIERS

Two Stanford Research SR-530 lock-in amplifiers were used to perform the resistivity measurements. These amplifiers simultaneously measure the amplitude and phase of

the input signals. The presence of any reactive elements in the circuit will induce phase changes between the applied excitation and the measured response. At the low frequencies used here, these components should be minimal and no phase changes should be seen. Measurement of the phase allowed rapid diagnosis of open circuits and the possibility of determining the reactive components in the circuit if high frequency excitation was used. Both lock-in amplifiers were IEEE-488 addressable, allowing full computer control of the measurement. In the four terminal set-up, one amplifier was connected directly across the sample while the other was connected across the current sensing resistor, as in figure 5.11.

5.4.2.5 TEMPERATURE MEASUREMENT

The Oxford Instruments ITC4 temperature controller was controlled via an RS-232C link allowing computer control of the data acquisition on this instrument. Other temperature measurement equipment, such as the line pressure and the heater temperature controller were not computer controllable.

5.4.2.6 MAGNET POWER SUPPLY

The RS-232C link used with the ITC4 temperature controller could, alternatively, be used to control the magnet power supply. Measurement of the supply current allowed the magnetic field strength (**B**) to be calculated via a calibration factor stored in the control program.

5.4.2.7 CONTROLLING COMPUTER

An IBM PC clone was used to control the equipment connected via the IEEE-488 and RS-232C links. The software used for the control, IMS-3, was developed by Dr. M. Rahman using Turbo Pascal. The software collects the measured data and stores the results onto disk, allowing downloading and further manipulation.

5.4.3 THE HP4145 SEMICONDUCTOR PARAMETER ANALYSER

When making measurements of sheet resistance at room temperature, a Hewlett Packard HP4145 semiconductor parameter analyser was used. This system allowed for probing of structures, using fine needle probes connected to stimulus measurement units (SMUs). Control over the currents and voltages supplied or measured via the SMUs was effected through the operation software of this system. This same software allowed

some analysis of the data collected, to extract the sheet resistance etc.

5.5 RESISTIVITY MEASUREMENTS

To determine the characteristics of the films being used, and their usefulness in the fabrication of granular single electron devices, a number of measurements were made. Firstly, measurements were made on aluminium films deposited in the high vacuum system; these measurements are discussed in section 5.5.1. Secondly, measurements were made on ultra-high vacuum (UHV) deposited aluminium films. The UHV aluminium films were co-fabricated with single electron device test structures, using an oxygen partial pressure atmosphere during deposition and oxidation after deposition (as discussed in section 3.4.3 and 3.5.3). Section 5.5.2 discusses the measurements taken from these films. Thirdly, some measurements were made with the thin film, Ti/Au, interconnect to diagnose difficulties with these films. These measurements are discussed in section 5.5.3.

5.5.1 HIGH VACUUM DEPOSITED FILMS.

The high vacuum system, described in section 3.5.2, was used to deposit aluminium films onto oxidized silicon substrates for measurement of the sheet resistance. These films were used to give some idea of the behaviour of aluminium films of this thickness. The source metal was of the same purity as used in the UHV depositions, with the same evaporation method. The main differences were the system purity, cleanliness and gases used within the system. The HV system was vented using nitrogen and depositions were carried out using base pressures of around 2×10^{-6}mbar with run pressures in the region of 4 to 8×10^{-6}mbar. The rate of deposition in all cases was kept below $0.3 nms^{-1}$.

Two sets of samples deposited in the HV evaporator system were measured. The first set ("SB" samples) were simple 5 x 5 mm van der Pauw squares. These were bonded with pressed indium beads and mounted onto an insert with no direct sample temperature measurement ability. In all, nine samples were measured at room temperature and three of these were characterized over the temperature range 1.6K to 300K. The samples had a range of thicknesses, as determined by a Rank-Taylor-Hobson Talystep profileometer. Figures 5.15 to 5.17 plot the measurements of sample sheet resistance and resistivity against temperature.

A second set of samples was produced, on lead-frame structures. These were, in general, difficult to measure and the majority of samples failed to yield any resistance

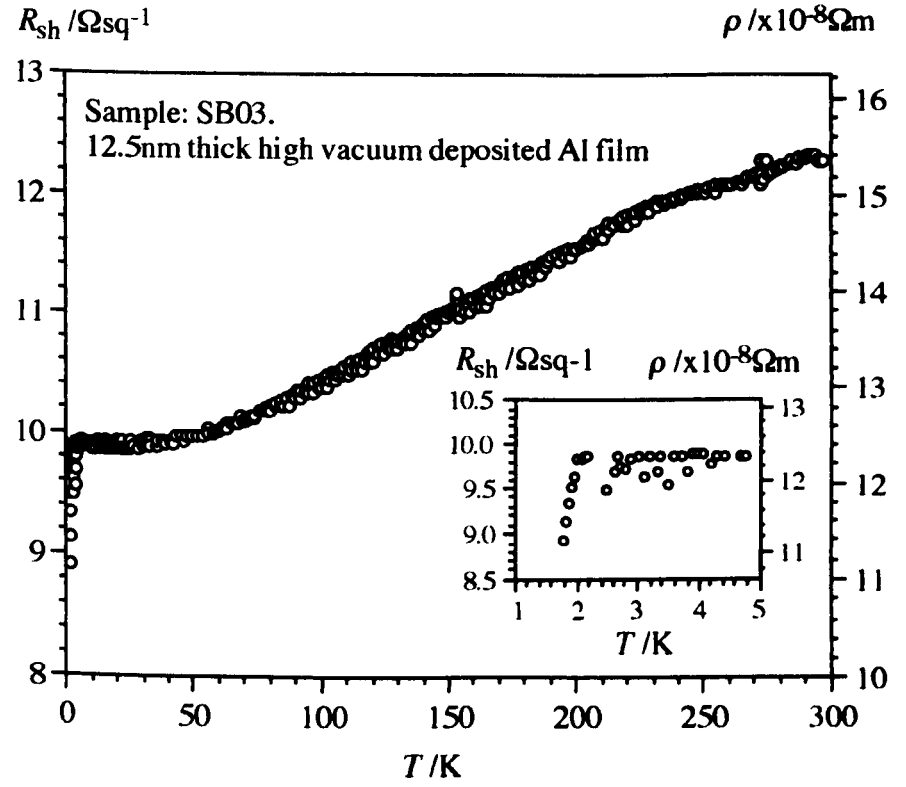
measurements. This was due to the thin film interconnect and is discussed in section 5.5.3. This second set of samples was arranged in a Hall bar structure of 4μm wide by 9μm between the voltage probe centres. The voltage probes were 1μm wide. The lead-frame structure allowed the use of ceramic chip headers and an insert with a sample temperature measurement sensor. The results from two 47nm thick aluminium films are presented in figures 5.18 and 5.19.

In all cases, the samples were measured using a 1kΩ current sensing resistor with root mean squared (rms) excitation voltages of 1V at 18Hz. This gives a maximum rms sample current of around 1mA, or a maximum current density of around 5 to 20GAm^{-2}. The high excitation voltage was used to overcome the noise present in this measuring system and to provide high sample currents for the measurement. The low resistance values of the samples meant that high currents were required in order to generate recoverable voltage signals from the sample. A dual lock-in circuit was employed using the arrangement shown in figure 5.11. The sample information is summarized in table 5.1.

Sample Identification	Base Pressure /x10^{-6}mbar	Run Pressure /x10^{-6}mbar	Measured Film Thickness /nm	Measurement Range
SB01†	5.2	6.2	14.0 ±2	300K - HP4145
SB03†	5.2	6.2	12.5 ±2	1.6K to near 300K
SB04†	5.2	6.2	13.0 ±2	300K - HP4145
SB05†	2.7	7.4	22.0 ±2	300K - HP4145
SB06†	2.7	7.4	22.0 ±2	300K - HP4145
SB07†	2.7	7.4	23.0 ±2	1.6K to near 300K
SB09†	1.2	4.4	30.0 ±2	300K - HP4145
SB10†	1.2	4.4	30.0 ±2	300K - HP4145
SB11†	1.2	4.4	30.0 ±2	1.6K to near 300K
S2ACA1‡	1.5	3.1	47.0 ±4	1.6K to near 300K
S2ACA3‡	1.5	3.1	47.0 ±4	1.6K to near 300K

† - van der Pauw sample. ‡ - Hall Bar sample.

Table 5.1 HV deposited sample and measurement details



Inset displays the detail of the low temperature portion of the measurement.

Figure 5.15 **Sample SB03 Sheet resistance and resistivity vs. temperature.**

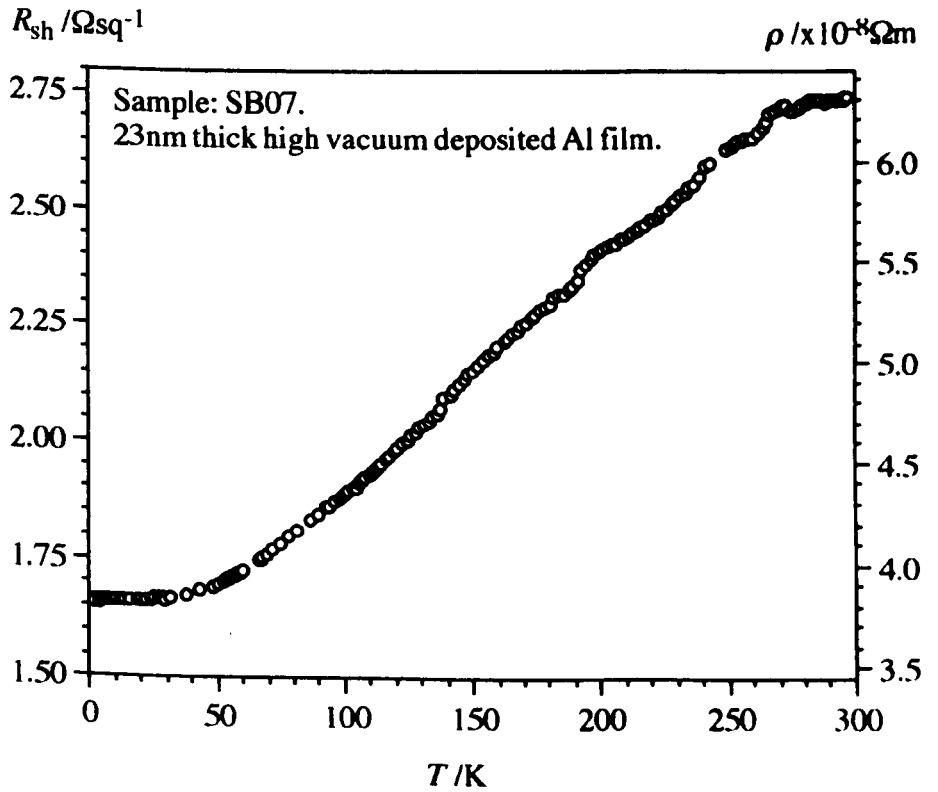


Figure 5.16 **Sample SB07 Sheet resistance and resistivity vs. temperature.**

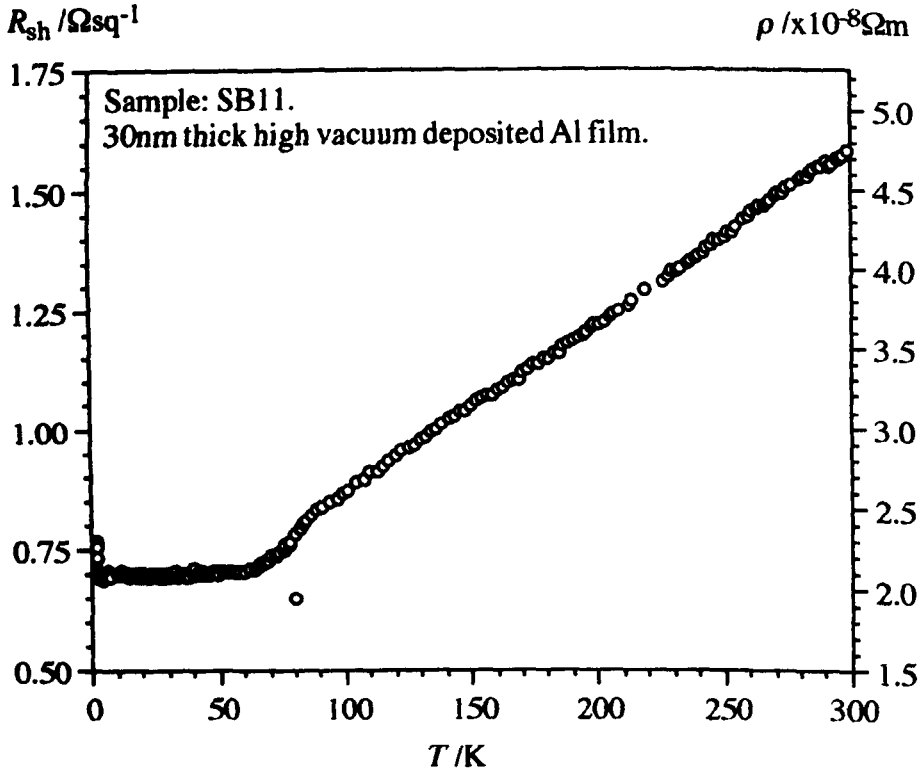


Figure 5.17 **Sample SB11 Sheet resistance and resistivity vs. temperature.**

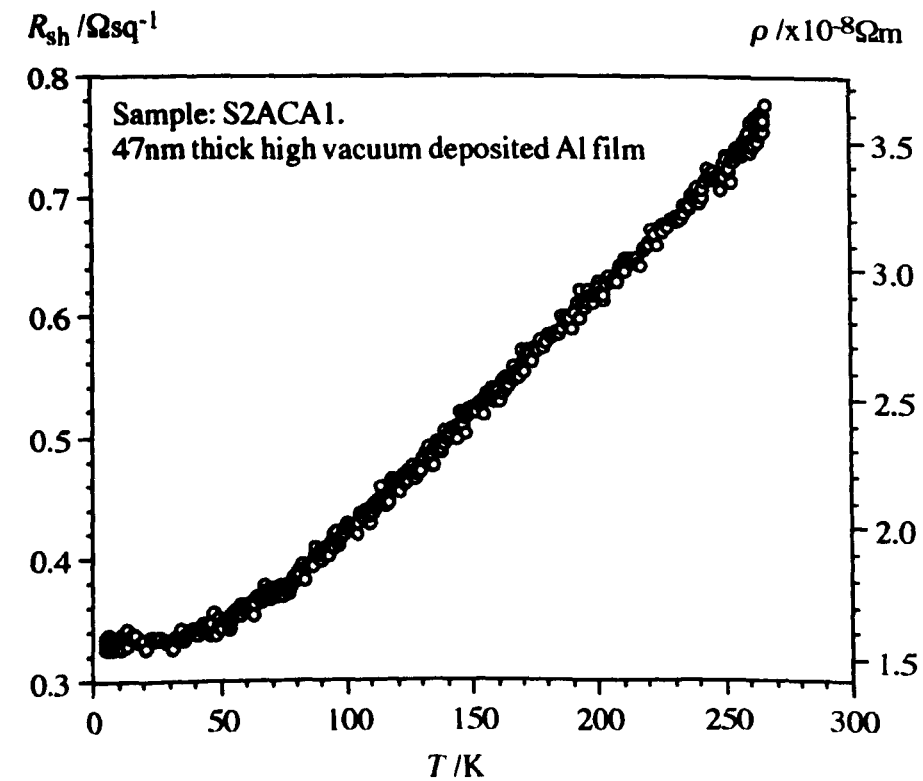


Figure 5.18 **Sample S2ACA1 Sheet resistance and resistivity vs. temperature.**

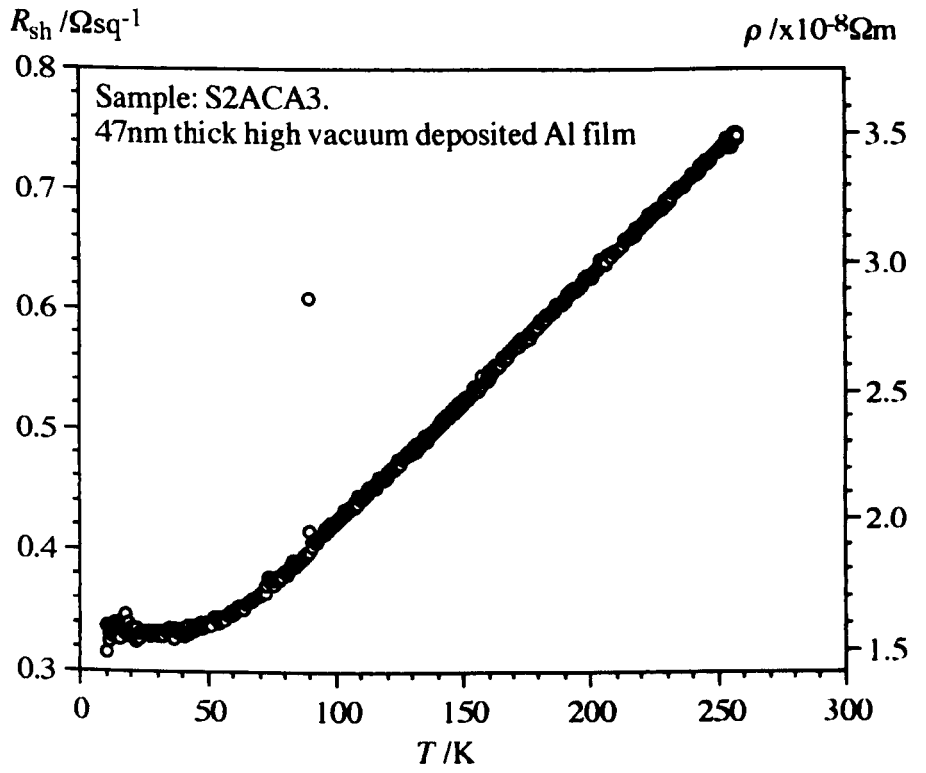


Figure 5.19 **Sample S2ACA3 Sheet resistance and resistivity vs. temperature.**

The following observations arise from the data presented above. Firstly, as the sample temperature was not directly measured, a number of experimental artefacts have been observed in the data of figures 5.15 to 5.17. These generally occur at the high temperature end of the measurement, but in the data of SB03, some hysteresis is present in the temperature sweep. The deviations from a straight line dependency are a result of the fact that the temperature measurement was taken from the heater plate temperature controller on the cryostat VTI. This results in a difference between sample temperature and measured temperature if the rate of change is high or the sample temperature does not follow the heater plate temperature. Care was taken in these measurements to keep the rate of change of temperature low, around 2 to 3°Kmin^{-1}. However, with this cryostat system, the sample space does not closely follow the heater plate temperature above 220 to 250K due to the cooling effect that the helium reservoir and liquid nitrogen cooling jacket have on the sample space temperature. These artefacts are not present in the two sets of data taken using temperature measurement directly at the sample, that is the data for samples S2ACA1 and S2ACA3.

Secondly, the resistivity is calculated from the sheet resistance. Measurement of the sheet resistance had an error of around 2% associated with it, however, the thickness measurement had an error of at best ±1.5nm on the stated thickness. The total error on the accuracy of measurement of the resistivity was therefore not less than 10%. The

scale of this error is too small to display on the graphs of resistivity temperature dependence.

Thirdly, with the 14nm film, an elevated transition to superconductivity is observed. The insert of figure 5.15 shows the low temperature portion of this data in detail. Temperature measurement below 4.2K for the "SB" samples was carried out by noting the helium pumping line pressure and converting that vapour pressure into liquid temperature[49]. Care was taken to ensure that the sample was immersed in liquid helium and that the gas flow rate was low when pressure measurements were taken. These actions ensured that the vapour pressure measurement of the liquid temperature could be safely taken as the sample temperature.

The transition to superconductivity occurred around 1.9K. This corresponds well with the data of Dynes and Garno[45], where films with resistivities of around $1x10^{-7}\Omega m$ at 4.2K were found to have transition temperatures in the range of 1.6 to 2.0K. In this case, the transition may have occurred because of the decreased film density which would lead to a decrease in the phonon frequencies, or through creation of a weak periodic or semi-periodic background potential, which aids the development of the superconducting state. The measurement shows, however, that heating effects due to the relatively high current used are not significant and the films are very conductive.

Fourthly, the resistivities observed and the highest resistivity of around $1.5x10^{-7}\Omega m$ for the 14nm thick film demonstrate that all films sampled are very metallic. The film resistivities lie well below the metal insulator transition resistivity of $\approx 1x10^{-3}\Omega m$, as identified by Dynes and Garno[45].

Table 5.2 summarizes the results from the samples measured in this section of the work. This table lists the sheet resistances and resistivities at both room temperature and 4.2K (where relevant). From these numbers, the residual resistivity ratio is calculated. The temperature coefficient of resistivity has also been calculated for the data presented in the graphs above. This was calculated by using a discretization of equation 5.8. The calculation was spread over five sample points, each point being separated by around 0.2K. The number presented in table 5.2 is then the average gradient of the temperature dependence of resistivity from 50 to 220K. This temperature range represents the linear portion of the temperature dependency on all samples.

Sample Identification	R_{sh}(300K) /Ωsq^{-1}	ρ (300K) /x10^{-8}Ωm	R_{sh}(4.2K) /Ωsq^{-1}	ρ (4.2K) /x10^{-8}Ωm	RRR	α (50 to 220K) /ppmK^{-1}
SB01	9.48	13.3 ±1.9	---	---	---	---
SB03	12.3	15.4 ±2.5	9.78	12.2 ±1.9	1.26	≈900
SB04	11.3	14.7 ±2.3	---	---	---	---
SB05	4.22	9.28 ±0.8	---	---	---	---
SB06	2.99	6.58 ±0.6	---	---	---	---
SB07	2.74	6.30 ±0.5	1.66	3.82 ±0.3	1.65	≈1600
SB09	1.81	5.43 ±0.4	---	---	---	---
SB10	1.50	4.50 ±0.3	---	---	---	---
SB11	1.58	4.74 ±0.3	0.70	2.10 ±0.1	2.26	≈2100
S2ACA1	0.82†	3.85 ±0.3	0.33	1.55 ±0.1	2.48	≈2170
S2ACA3	0.84†	3.95 ±0.3	0.32	1.50 ±0.1	2.63	≈2200

† - extrapolating to 300K from data for figures 5.18 & 5.19.

Table 5.2 Results for HV deposited samples.

The results summarized in table 5.2 show that increasing the film thickness reduces the residual resistivity of the material, as expected. The reduction in boundary scattering with thicker films is probably the greatest factor influencing this behaviour, although with thinner films a higher impurity content may exist due to a more open film structure. Figure 5.20 plots the thickness dependence of the resistivity for the measured samples. This graph displays a log(ρ) vs log(d) behaviour which is indicative of the more classical theories for conduction in this type of film. Recalling the equation for Fuchs' theory and that $k=d/\bar{l}$, equation 5.14 rearranges to give equation 5.32 (assuming that any boundary scattering is purely diffuse, i.e. $p = 0$):

$$1-\frac{\rho_o}{\rho}=\frac{3\bar{l}}{8d}. \qquad [5.32]$$

This allows us to take logs of both sides giving:

$$log\left(\frac{\rho-\rho_o}{\rho}\right)=log\frac{3\bar{l}}{8}-logd. \qquad [5.33]$$

From equation 5.33, it appears possible to determine the mean free path based on data for the thickness dependence of resistivity. Recall, however, that the parameter ρ_o is the resistivity of a material with the same physical characteristics as the films in question but that it is infinitely thick. Fabrication of such a film is almost impossible as some of the defects inherent in a thin film depend on the thickness. It may be possible using simultaneous depositions of different but known thickness films to determine the

mean free path in this manner. In the case of this data, where the log - log dependency is present in both the room temperature and 4.2K data, we may conclude that the greatest modification to the film behaviour results from classical boundary scattering. It is very likely, however, that any quantum mechanical modifications would be masked by lack of resolution in the measurement system ($\Delta R_{min} \approx 0.2\Omega$), inherent noise and the magnitude of the excitations used in these measurements.

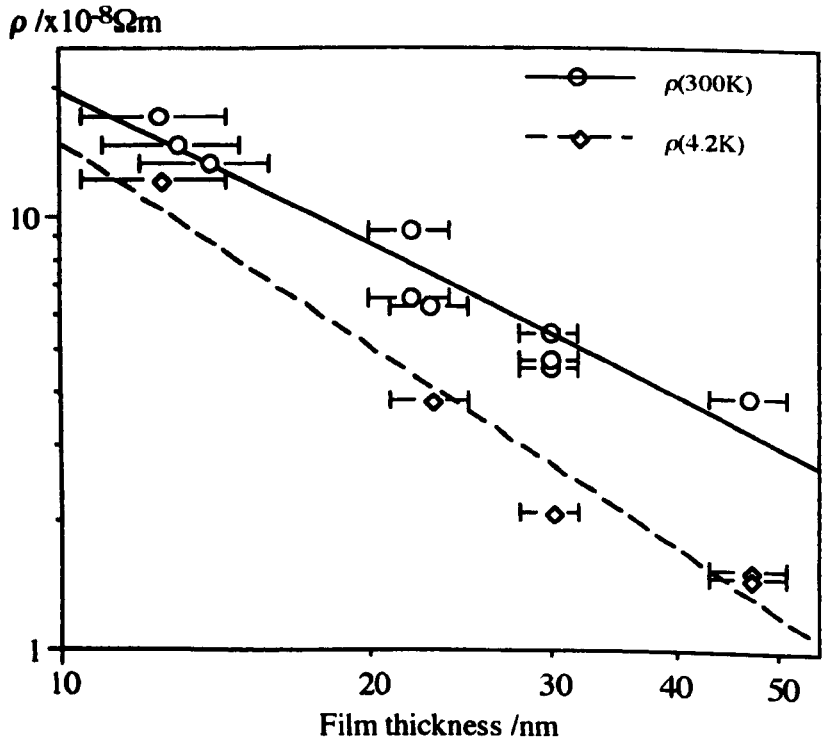


Figure 5.20 **Resistivity (for 300K and 4.2K) vs thickness for HV deposited films.**

The residual resistivity and temperature coefficient of resistivity also are dependent on the film thickness. This is to be expected as defectivity must increase with reduction in the film thickness. Also, thin metal films display a change in sign of the TCR as they cross from the continuous to dielectric regimes. This results from the film structures becoming more percolative and eventually discontinuous. Eventually, with a completely discontinuous film behaving in a dielectric manner, charge transport must occur by an activated process and therefore the TCR must be negative. Figure 5.21 plots the RRR and the TCR (α) against the log of thickness. It appears from this data that film thicknesses of around 5nm will produce a material with a negative TCR, indicating a discontinuous film. Similarly, for such a film, the RRR would be less than 1, which is likely to occur at a thickness of around 10nm.

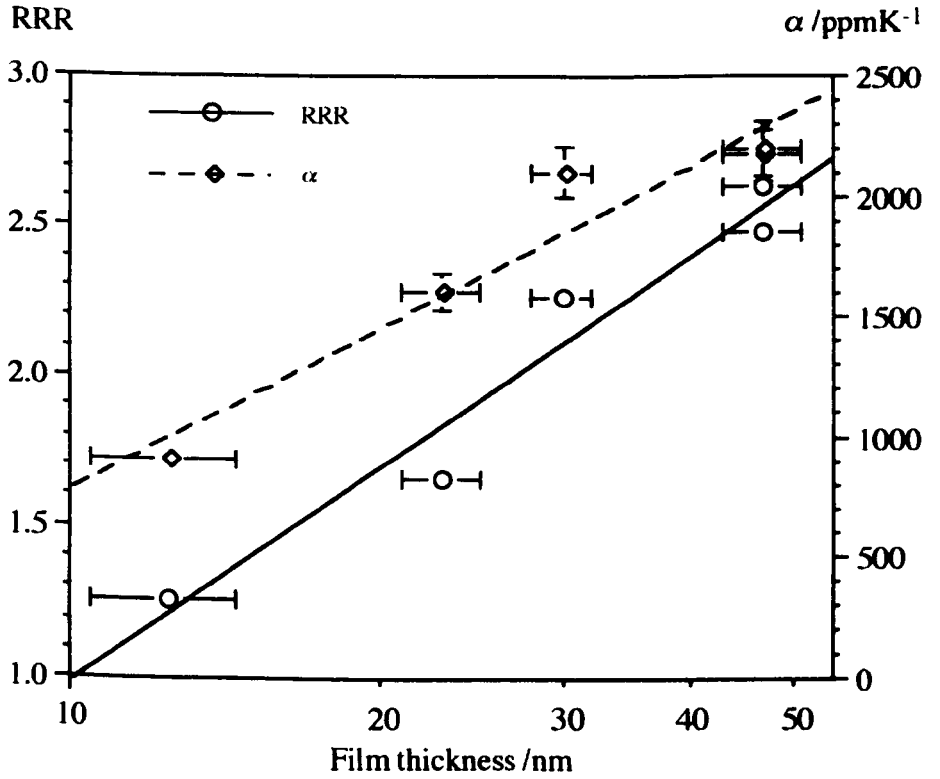


Figure 5.21 **Residual resistivity ratio (RRR) and Temperature coefficient of resistivity (α) vs thickness for HV deposited films.**

The magneto-resistance of the samples, SB03, SB07 and S2ACA1 was measured at 4.2K. With the Hall bar sample, S2ACA1, Hall measurements were also attempted, however, the high carrier density precluded the observation of the Hall resistance with this measurement circuit. All of the magneto-resistance measurements failed to show any change in sample resistance within the resolution of the measurement system. All of these measurements were carried out using the same measurement set-up as for the temperature dependence measurements.

5.5.2 ULTRA-HIGH VACUUM DEPOSITED FILMS.

Films measured in this section of the work were deposited using the UHV evaporator system, as described in section 3.5.3. These samples were deposited in the same vacuum cycle as some of the single electron device samples. This allowed characterization of the film used on an individual device. Measurements of these samples were carried out with a slightly modified measurement circuit and an insert fitted with a sample temperature sensor. All samples were of the Hall bar type, with the length between the voltage probes being 0.8mm and the width of the bar being 50μm. The width of the voltage probe contacts was 2μm. These samples were bonded by pressing indium beads into the film bond-pads.

The measurement circuit was modified by the addition of an EG&G model 113P instrumentation amplifier to the sample voltage (V_s) measurement circuit. This amplifier was connected between the sample voltage probes and the lock-in amplifier of figure 5.11. The voltage gain of the amplifier was set to 1000 in all cases. Inclusion of this amplifier allowed the use of much lower sample currents than in the previous measurements, reducing the possibility of sample damage through ESD or high sample currents. The excitation voltage used in these measurements was between 2 and $10mV_{rms}$, with a current sensing resistor of $10k\Omega \pm 1\%$. This gives a maximum rms sample current of $1\mu A$, limiting the maximum current density to between 10 and $100kAm^{-2}$ for these films.

In all, five samples were measured, two each of 15nm and 20nm thick films and one film of 25nm thick. The oxygen partial pressures used during the evaporation are given in table 5.3 along with the sample thicknesses. After the film deposition was complete, the samples were kept in the UHV chamber for over 1 hour in 1bar of oxygen, to ensure full oxidation of the film surface. The temperature sweeps are shown in figures 5.22 to 5.26.

Sample Identification	Deposited Film Thickness /nm	Oxygen Partial Pressure $P_p(O_2)$ /x10^{-6}mBar
T04DB	15 ±0.2	2.5 ±0.5
T04DK	15 ±0.2	2.5 ±0.5
T04CB	20 ±0.2	3.0 ±0.5
T04CK	20 ±0.2	3.0 ±0.5
T04AB	25 ±0.2	2.0 ±0.5

Table 5.3 UHV deposited sample details.

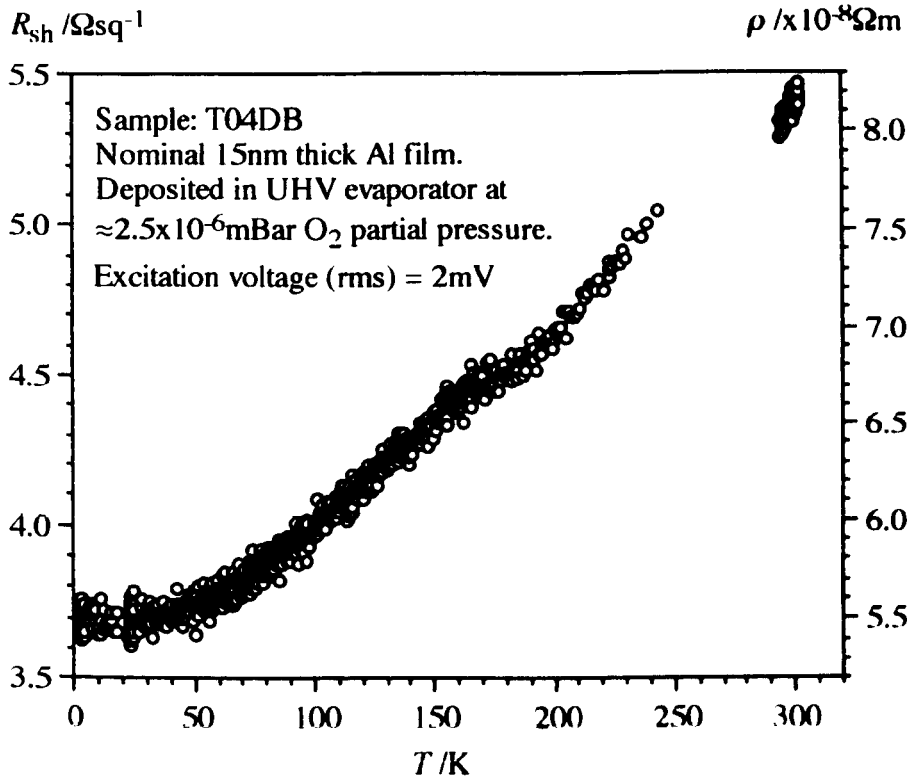


Figure 5.22 **Sample T04DB Sheet resistance and resistivity vs. temperature.**

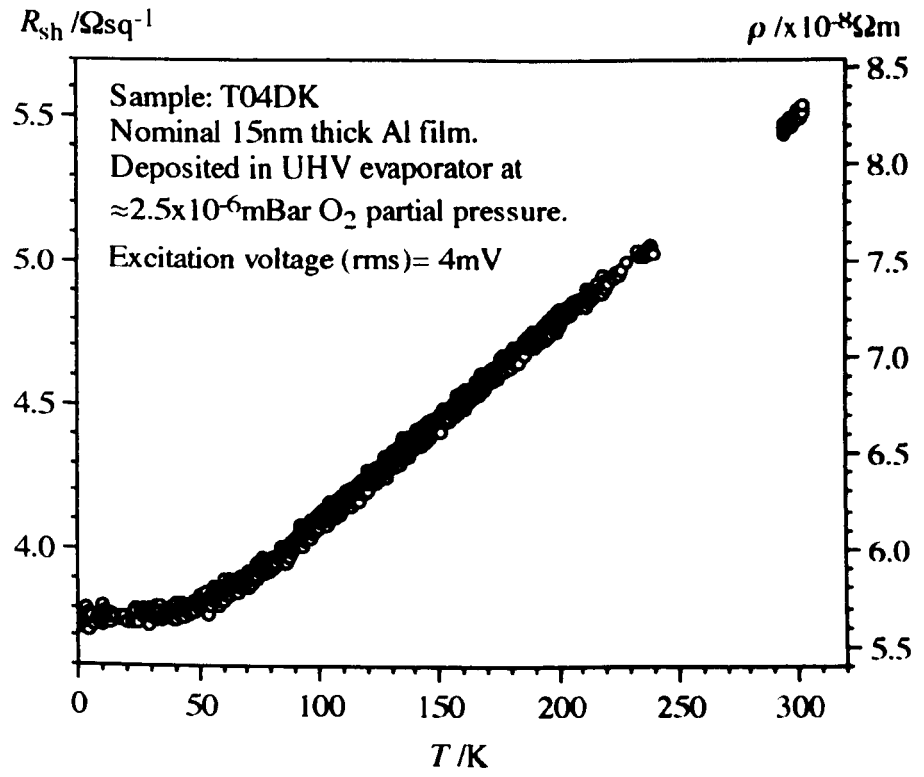


Figure 5.23 **Sample T04DK Sheet resistance and resistivity vs. temperature.**

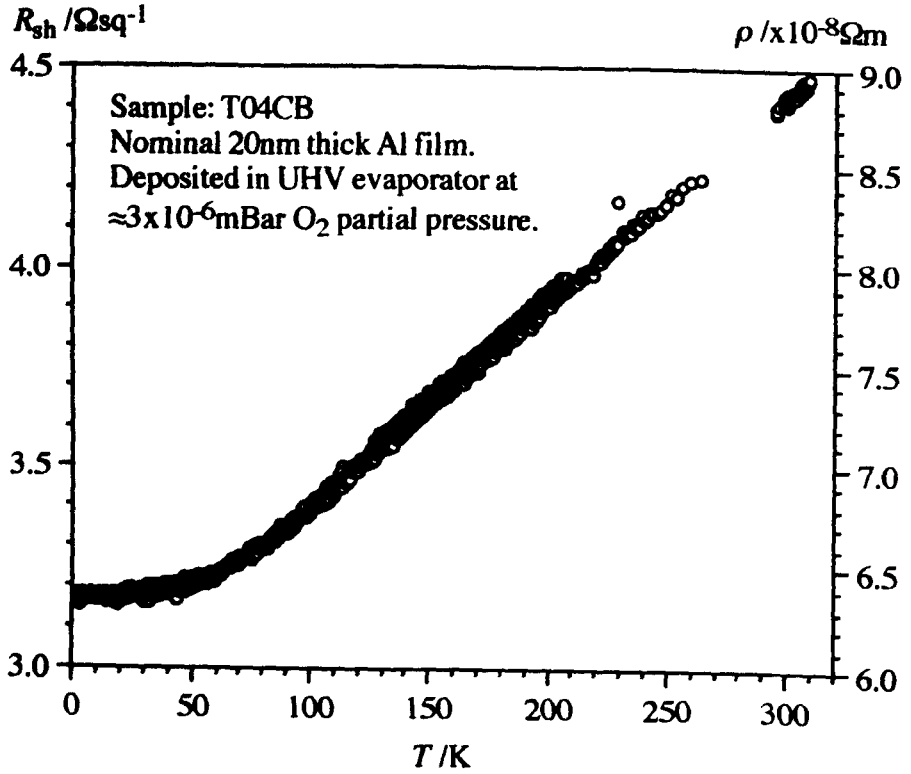


Figure 5.24 **Sample T04CB Sheet resistance and resistivity vs. temperature.**

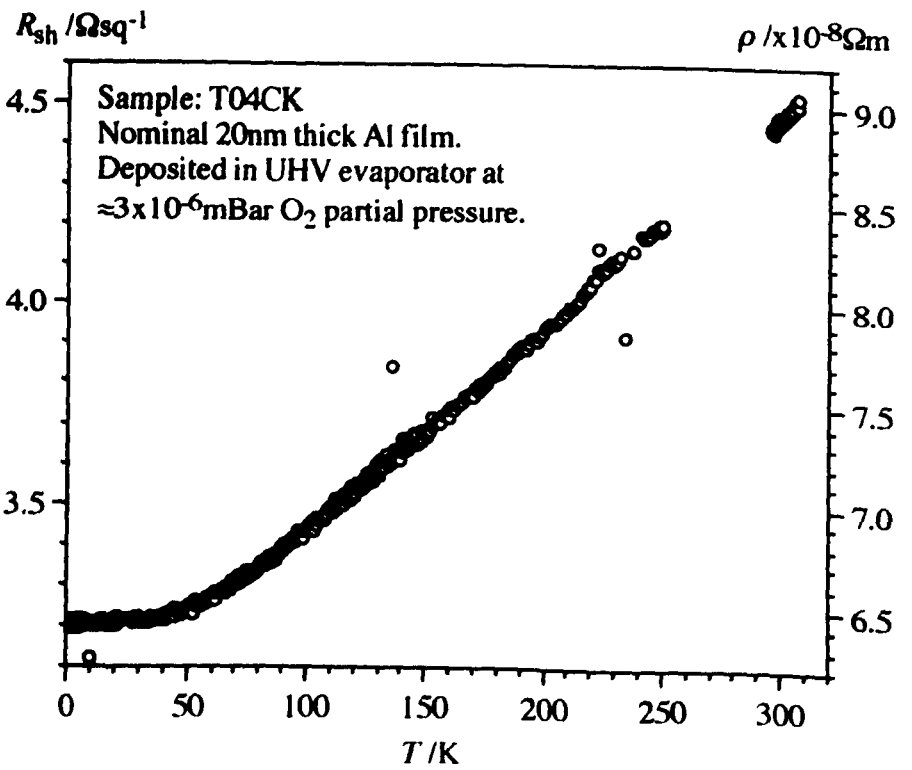


Figure 5.25 **Sample T04CK Sheet resistance and resistivity vs. temperature.**

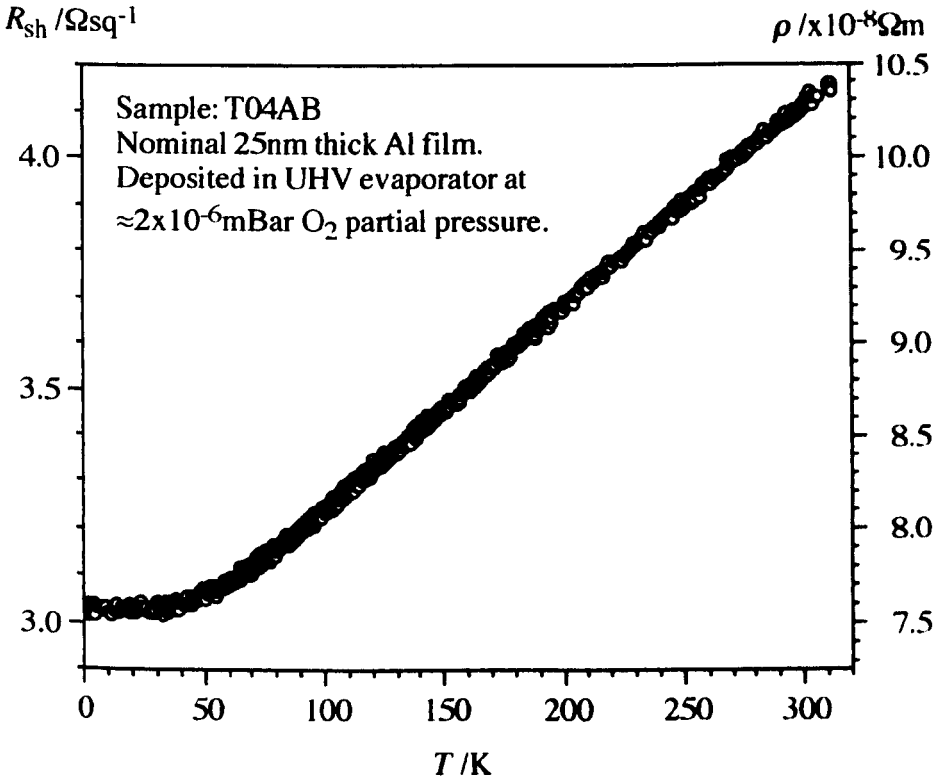


Figure 5.26 Sample T04AB Sheet resistance and resistivity vs. temperature.

The results shown in the figures above are summarized in table 5.4. All samples were measured in the same loading of the cryostat and room temperature measurements were obtained before cooling the samples. Measurements around room temperature were achieved by gently heating the sample end of the insert with a hot air gun. When the insert system had been cooled once, it was not possible to raise the sample space temperature above 250K due to cooling from the liquid helium reservoir and liquid nitrogen jacket. Therefore, T04AB is the only sample for which data over the full temperature range has been obtained as this was the first sample measured. The data from T04DB contains a higher degree of scatter than that of T04DK, due to the use of a lower excitation voltage. This indicates the degree of noise present in the measurement system, with this circuit arrangement.

Sample Identification	R_{sh}(300K) /Ωsq^{-1}	ρ (300K) /x10^{-8}Ωm	R_{sh}(4.2K) /Ωsq^{-1}	ρ (4.2K) /x10^{-8}Ωm	RRR	α (100 to 300K) /ppmK^{-1}
T04DB	5.38	8.07	3.65	5.48	1.47	1280
T04DK	5.51	8.27	3.75	5.63	1.47	1270
T04CB	4.42	8.84	3.16	6.32	1.40	1170
T04CK	4.47	8.94	3.20	6.40	1.40	1160
T04AB	4.11	10.3	3.02	7.55	2.03	1070

Table 5.4 Results for UHV deposited samples.

Table 5.4 shows that the films appear to become more resistive with increasing thickness. Two points should be noted here. Firstly, each film was deposited at a different partial pressure of oxygen and will therefore have a different resistivity contribution due to this type of impurity (see table 5.3), although, the spread of partial pressures is only $1x10^{-6}$mbar. Secondly, the quoted thickness is as read from the crystal monitor system in the UHV evaporator. The error in the absolute value of this measurement derives from the reduction in film density[50] as the thickness is reduced and from the presence of oxygen in the film[22]. No cross check of the thickness against profileometer measurements was made. Therefore, the absolute value of the resistivity is not accurately known but the magnitude of the resistivity is the important parameter here. Again, magneto-resistance measurements of these films were taken at 4.2K, however, no resistance changes were observed in the data.

The measurements above directly represent the films being used to fabricate single electron devices in this work and, therefore, the measurements of sheet resistance are valid. The absolute value of the film resistivity is, however, not accurately known but the magnitude of the resistivity places the films well on the metal side of the metal insulator transition. The films deposited in the UHV system are therefore, like the HV films, very conductive and show no signs of being affected by severe modifications due to percolative or activated conduction processes. The films are therefore useful as low resistance contacts to the tunnel junction regions of the single electron devices.

5.5.3 THIN FILM INTERCONNECT MEASUREMENTS.

The thin film interconnect, section 3.4.2.2, is used to allow connection of the device layer aluminium deposition to the bondpads on the lead frame structures. This interconnect must remain stable and not create step coverage problems for the aluminium films. However, a great deal of difficulty resulted from the use of the thinner of the two films. Measurements on a number of HV and UHV aluminium films deposited onto the 2:8nm Ti:Au interconnect failed due to open circuit behaviour. Similarly, measurements of aluminium films of <20nm for the HV evaporator and <25nm for the UHV evaporator failed as open circuits when the 3:10nm Ti:Au interconnect was used. SEM inspection of these failures showed that on the thicker interconnect the cause of the failure was step coverage problems, and in some cases ESD damage to the aluminium film. SEM inspection of the failures on the 2:8nm Ti:Au interconnect displayed no damage to the aluminium film or to the overlap between the interconnect and the aluminium. However, the SEM analysis showed that the interconnect film structure had changed.

Measurements of the thin film interconnect, of both thicknesses, were made to assess the suitability of the films for their purpose and to diagnose the problem with the thinner of the two films. One sample of each thickness of interconnect was measured (S2ACC and S3BAA), along with two samples of the thinner interconnect (T02BE and T02BB) to diagnose the problems discovered with this film type. The sample characteristics are detailed below in table 5.5.

Sample Identification	Deposited Ti Film Thickness /nm	Deposited Au Film Thickness /nm	Received 180°C bake.
S2ACC†	3	10	Yes
S3BAA†	2	8	Yes
T02BE‡	2	8	No
T02BB‡	2	8	Yes

† - van der Pauw sample. ‡ - Hall Bar sample.

Table 5.5 Interconnect characterisation sample details.

Measurements on samples S2ACC and S3BAA used the same set-up as the HV aluminium films where the lead-frame samples were bonded using an ultrasonic gold wire bonder. Measurements of T02BE and T02BB used the same set-up as the UHV aluminium films and both of these Ti:Au samples were indium bonded to the insert wiring.

When deposited, the gold film on top of the titanium film appeared smooth and continuous. However, after processing, the gold on all the 2:8nm Ti:Au films had agglomerated into a sparse island structure on top of the titanium. Fusing of this thin filamentary gold layer near the overlap with the bonding structure created an open circuit. This fusing was probably caused by the reduced cross-sectional area of the conducting film in this network or percolative material state. It was believed that the film structure was changed during processing of the final aluminium layer, in particular during the 180°C resist bake. In effect, the gold layer was being annealed into an island or network type structure on top of the titanium layer during this process. Two samples, T02BE and T02BB were used to determine whether or not this was the case.

Both the T02B samples were deposited in the same evaporation cycle. Both films were patterned into Hall bars and measured exactly as the UHV aluminium films described in section 5.5.2. One sample, T02BE, was baked overnight at 180°C, as would be the case for any film receiving e-beam lithography resist processing. The other sample, T02BB, was not baked. The samples were bonded to the insert wiring using pressed indium

beads. The sample excitation voltage was 10mV, rms, using the measurement circuit of section 5.5.2. However, at 10mV excitation, the 180°C baked sample would continually switch between a conducting state and a non-conducting or high resistance state every few minutes. The data collected (and presented in table 5.6) for sample T02BB was measured using 100mV excitation, at which the sample remained in a stable conducting state. The temperature dependence again took the form of an almost linear temperature dependence with a plateau at low temperatures. The results are summarized in table 5.6, below.

Sample Identification	R_{sh}(200K) /Ωsq^{-1}	ρ (200K) /x$10^{-8}\Omega m$	R_{sh}(4.2K) /Ωsq^{-1}	ρ (4.2K) /x$10^{-8}\Omega m$	RRR (ρ (200K) / ρ (4.2K))
S2ACC	2.41	3.13	1.74	2.26	1.38
S3BAA	16.0	16.0	12.3	12.3	1.30
T02BE	15.7	15.7	13.3	13.3	1.18
T02BB	18.8	18.8	16.6	16.6	1.13

Table 5.6 Results for Ti:Au interconnect samples.

The gold layer of the 2:8nm interconnect did change character through baking at 180°C. The island structure observed on S3BAA and other films deposited at the same time, was also evident on T02BB. However, T02BE displayed the structure of a smooth gold film on top of the titanium layer. The island film structure appears to be a direct result of the 180°C resist bake. This bake is part of the lithographic resist processing for the device layer on these structures. The change in character is also evident from the increased resistivity of the baked sample over the un-baked sample. The transition to a percolative structure is prohibitive to the use of these films as an interconnect to the aluminium film structures, as connection cannot be guaranteed, and the unstable current-voltage behaviour of this film produces unreliable contacts.

From the results collected from these samples, it is clear that the 3:10nm Ti:Au interconnect film (S2ACC) is the best choice for this type of processing. The film remains very conductive and retains its deposited form throughout the subsequent processing. However, it does limit the thickness of the aluminium film that can be used to form device structures with the lead-frame samples. No aluminium samples under 20nm thick were successfully measured on this interconnect film. Where the thinner, 2:8nm Ti:Au, interconnect was used with aluminium films the samples failed to yield any measurements. In all cases, ESD or high current damage was evident on the interconnect film. None of the aluminium films deposited on this type of showed any kind of step coverage or electron-migration problems suggesting that the interconnect

fused before the aluminium layer could be damaged.

5.6 CONCLUSIONS

This work was motivated by the need to answer four questions related to the conduction properties of the thin films being used for single electron device formation and in particular the granular thin film devices. The first two questions, to which answers have been gained, related to the need for metallic films which do not affect the conduction properties of the fabricated device structures. From the measurements carried out, it can be seen that the aluminium films used for this work are metallic. The resistivities in all cases are less than 2 x $10^{-7}\Omega$m and all films have a positive temperature coefficient of resistivity. The low values of the residual resistivity ratios for each of the films indicate the presence of significant background defectivity. This defectivity is probably due to grain and film boundary scattering. However, the films do not show any dielectric or discontinuous behaviour.

The high vacuum deposited films show a logarithmic dependency of resistivity on film thickness while the UHV films do not. The absence of this trend in the UHV films may be due to deficiencies in the measurement of the UHV films thickness and variation in the oxygen partial pressure during deposition, as noted in section 5.5.2. The logarithmic dependency is present in both Fuchs' classical theory and in the modifications through the quantum effects of weak localization and electron interactions. To determine the source of the scattering, magneto-resistance data would be required. However, none of the samples tested provided this data. The reasons for this may be three-fold.

Firstly, and most likely, the signal to noise ratio of the measurements was too low. To increase the accuracy and resolution of the measurement by reducing the noise contribution, the excitation voltage required to be increased. Increasing this voltage tends to mask any quantum effects. This is the case with the measurements on high vacuum deposited samples, where the circuit rms excitation was 1V with the voltage developed across the sample being around 10mV. Secondly, the sample temperature may have been higher than measured. This is unlikely as the presence of a superconductive transition in one of the thin HV deposited samples suggests that the error on the temperature measurement was not significant. However the effective electron temperature is much more significant. The excitation voltage imparts energy to the electrons, $E \approx eV$, and in almost all cases this is in excess of the effective temperature of the sample. In order to observe effects with characteristic temperatures of around 10K, electron energies of $<<900\mu$eV are required. In almost all measurements taken on

high vacuum deposited films, the excitation and coupled noise ensured this limit was exceeded. Thirdly, the samples are very conductive and the magnitude of any modifications due to quantum effects are likely to be very small on the scale of the resolution of the measurement circuit. Similarly, Hall effect measurements failed to yield any results. Again, the resolution of the measurement system was insufficient to measure the Hall coefficient where a large, metallic, number density of charge carriers was expected.

The question of the dominant conduction mechanism within the films therefore remains unanswered. However, if we return to the questions posed in the introduction to this chapter, we can conclude that the first two questions have been answered. We do indeed have metallic aluminium films which in all cases have resistivities less than $2 \times 10^{-7}\Omega m$. Also, provided there are no severe modifications to the conduction mechanisms through one or two-dimensional localization effects[40], the bulk of the film should not impact the behaviour of any device regions or affect the charge transport properties of these regions or constrictions. However, we are still without a clear answer to the third and fourth questions of the introduction. These questions relate to the microscopic behaviour of the metal films; will the film structure immediately adjacent to the device seriously affect the device properties and, will the device structures themselves provide high temperature single electron effects? Further work is required in order to answer these two remaining questions and this is discussed in chapter 6.

5.7 REFERENCES

1. E.M. Tennant (ed.), *Science Data Book*, p60, 1st Ed., Oliver & Boyd, 1971.

2. D.R. Lide (ed.), *Handbook of Chemistry & Physics*, Section **12**, p33, 72nd Edition, CRC Press, Boston, 1991.

3. B. Segall, *Phys. Rev.* **124**(6),1797 (1961).

4. A. Matthiessen & C. Vogt,*Pogg. Ann.* **122**,19 (1864).

5. R.G. Chambers, *Electrons in Metals and Semiconductors*, Chapman Hall, 1990.

6. L.N. Cooper, *Phys. Rev.* **104**,1189 (1956).

7. D.R. Lide (ed.), *Handbook of Chemistry & Physics*, Section **12**, p41, 72nd Edition, CRC Press, Boston, 1991.

8. N.W. Ashcroft & N.D. Mermin, *Solid State Physics*, Chapter 15 - Band structure of selected metals, The trivalent metals, pp299-302, W.B. Saunders Co., Philadelphia, 1976.

9. N.W. Ashcroft & N.D. Mermin, *Solid State Physics*, Chapter 13 - The semi-classical theory of conduction in metals. pp243 - 263, W.B. Saunders Co., Philadelphia, 1976.

10. J.J. Thomson, *Proc. Camb. Phil. Soc.* **11**,120 (1920).

11. K Fuchs, *Proc. Camb. Phil. Soc.* **34**,100 (1938).

12. A.F. Mayadas, R. Feder and R. Rosenberg, *J. Vac. Sci. Technol.* **6**(4),690 (1969)

13. J.M Ziman, *Electrons and Phonons*, Clarendon Press, Oxford, 1960.

14. C.W.J. Beenakker & J.H. van Houten, *Phys. Rev. B.* **38**,3232 (1988).

15. G. Bergmann, *Physics Reports*, **107**(1),1 (1984).

16. P.A. Lee & T.V. Ramakrishman, *Rev. Mod. Phys.* **57**(1),287 (1985).

17. G.J. Dolan & D.D. Osheroff, *Phys. Rev. Lett.* **43**,721 (1979).

18. T. Chui, P. Lindenfeld, W.L. McLean & K. Mui, *Phys. Rev. Lett.* **47**(22)1617 (1981).

19. B.L. Al'tshuler, A.G. Aronov & P.A. Lee, *Phys. Rev. Lett.* **44**,1288 (1980).

20. M. Rahman, *Electron Transport in Dry-etched n^+- GaAs Nanostructures*, Chapter 5 - Quantum Transport in Mesoscopic Systems, Ph. D Thesis, University of Glasgow, 1992.
and S. Washburn, in *Quantum Coherence in Mesoscopic Systems*, B. Kramer (ed.), Plenum Press, New York, 1991.

21. H. Haucke, S. Washburn, A.D. Benoit, C.P. Umbach & R.A. Webb, *Phys. Rev. B.* **41**(18),12454 (1990).

22. A.F. Mayadas & M. Shatzkes, *Phys. Rev. B.* **1**(4),1382 (1970).

23. B. Abeles, P. Sheng, M.D. Coutts & Y. Arie, *Adv. Phys.* **24**(3),407 (1975).

24. S. Kirkpatrick, *Rev. Mod Phys.* **45**,574 (1973).

25. R. Landauer, *J. Appl Phys.* **23**,779 (1952)

26. N.F. Mott, *Phil. Mag.* **B44**,265 (1981).

27. T.J. Coutts, *Electrical conduction in thin metal films*, Chapter 8 - Conduction in semi-continuous films, Elsevier, 1974.

28. D.J. Thouless, *Phys. Rev. Lett.* **39**(18),1167 (1977).

29. R.W. Cohen & B. Abeles, *Phys. Rev.* **168**(2),444 (1968).
and G. Deutscher, M. Gershenson, E. Grünbaum & Y. Imry, *J. Vac. Sci. Technol.* **10**(5),697 (1973).

30. J.C. Anderson, *J. Vac. Sci. Technol.* **9**(1),1 (1971).

31. C.J. Gorter, *Physica* **17**,777 (1951).

32. C.A. Neugebauer & M.B. Webb, *J. Appl. Phys.* **33**,74 (1962).

33. R.M. Hill., *Proc. Roy. Soc. (London)* **A309**,377 (1969).

34. J.G. Simmons, *J. Appl. Phys.* **34**(9),2581 (1963).

35. D.S. Herman & T.N. Rhodin, *J.Appl. Phys.* **37**,1594 (1966).

36. T.E. Hartman, *J. Appl. Phys.* **34**,943 (1963).

37. B. Abeles, in *Applied Solid State Science. Vol. 6.*, ed., R. Wolfe, Academic Press, London, 1976.
and P. Sheng & B. Abeles, *Phys. Rev. Lett.* **28**,34 (1972).

38. C.S. Barrett, *Metals Handbook*, 8th Ed., Vol. 8. - Metallography, structure and phase diagrams, Crystal Structure, p242, *American Soc. Metals*, 1973.

39. R. Lück, *Phys. Stat. Sol.* **18**,49 (1966).

40. P. Santhanam, S. Wind & D.E. Prober, *Phys. Rev. Lett.* **53**(12),1179 (1984).

41. W. Buckel & R. Hilsch, *Z. Physik.* **138**,109 (1954).

42. I.S. Khukharera, *Sov. Phys. JETP.* **16**,828 (1963).

43. M. Strongin, O.F. Kammerer & A. Paskin, *Phys. Rev. Lett.* **14**,949 (1965).

44. J.J. Hauser, *Phys. Rev. B.* **3**(5),1611 (1971).

45. R.C. Dynes & J.P. Garno, *Phys. Rev. Lett.* **46**(2),137 (1981).

46. H. Gurev, *Thin Solid Films.* **18**,275 (1973).

47. See, for example, S.M. Sze, *Physics of Semiconductor Devices*, 2nd Ed., pp31-32., John Wiley & Sons, Singapore, 1981.

48. L.J. van der Pauw, *Philips Res. Rep.* **13**,1 (1958), and *Philips Tech. Rev.* **22**,220 (1958).
and D.R Zrudsky, H.D. Bush & J.R. Fassett, *Rev. Sci Inst.* **37**(7),885 (1966).

49. D.R. Lide (ed.), *Handbook of Chemistry & Physics*, Section **6**, p88, 72nd Edition, CRC Press, Boston, 1991.

50. T.E. Hartman, *J. Vac. Sci. Technol.* **2**,239 (1965).

CHAPTER SIX

Single Electron Device Measurements

6.1 INTRODUCTION

Devices which exhibit Coulomb blockade effects at high temperatures require very small capacitance, of order of 10^{-18}F, and a correspondingly high, tunnelling, resistance. This was discussed in chapter 2. In the aluminium/aluminium oxide system, this necessarily requires ultra-small geometries. From the work documented in chapter 4, we see that the patterned granular films provide potentially the smallest controlled geometries possible using aluminium metallization. The characteristics of the films used in this work were investigated through the work documented in chapter 5. That chapter posed four questions related to the usefulness of these films in this application. The work detailed in chapter 5 answered the first two questions. We have seen that the films are indeed metallic and that the "bulk" material characteristics should not seriously affect the charge transport properties of a device. The latter two questions relate to the effect of patterning on the material and need to be answered through measurements on single electron device structures. Firstly, will the small geometry regions behave differently from the bulk material, either through granular, or other, effects? Secondly, can high temperature single electron devices be fabricated using these structures? The work documented in this chapter aims to answer these two questions.

This chapter is divided into four sections. First, section 6.2 provides a very brief discussion of the single electron devices measured in this work. Section 6.3 details the measurement system and section 6.4 documents the results gained from the devices measured. Finally section 6.5 concludes this chapter, reviewing the results.

6.2 DEVICE STRUCTURES

Chapter 2 dealt with the theory behind single electron device operation. Apart from small device size for low capacitance, good decoupling of the device from the electromagnetic environment and low current operation are needed. In some respects the last two of these requirements appear to be mutually exclusive since decoupling requires high resistances which generate high noise levels. This decoupling can, in effect, be accommodated through the device design, for example by using an electrometer[1,2] or one dimensional array structure[3]. Alternatively, high impedance leads can be employed at the cold end of the measurement system to minimise the Johnson noise generated by the decoupling impedances[4]. The following sections discuss the test circuit used for the device measurements and describe the implementations in the two fabrication processes used.

6.2.1 FABRICATION TEST CIRCUIT

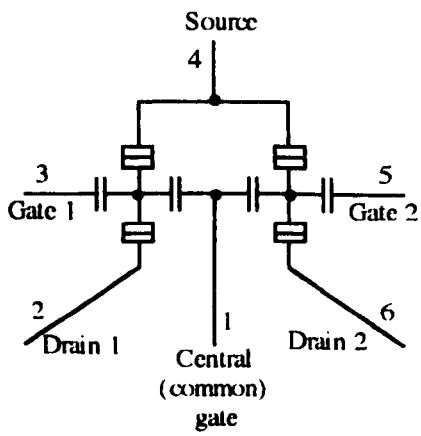


Figure 6.1 Circuit schematic of device test circuit.

The capacitively coupled double tunnel junction, or SET transistor was chosen as the basic structure for device measurements. It is the simplest structure in which single electron effects can be observed. More complicated arrangements, such as charge pump devices or one dimensional arrays, can provide progressively more decoupling and interesting effects, but become more difficult to set up and operate. The double junction circuit simply requires appropriate voltage bias across the source-drain contacts and the ability to measure the source-drain current. The only other variable required is

the gate voltage which can be easily applied and controlled.

The test circuit used is shown in figure 6.1. A number of features are built into this design as follows. Two SET transistor devices are fabricated per sample to double the chances of producing a working device per measurement and to allow for device state coupling experiments[5] if both devices operate. The devices have separate electron drains with a common electron source. Although in the granular implementation the source was patterned as two separate contacts, they were then connected together at the bonding stage. Each device island has a separate unique capacitive gate in order to independently bias the electron island and characterize each individual device. The common gate is provided to allow cross coupling of the islands in a state coupling experiment.

6.2.1.1 SUSPENDED MASK DEVICE STRUCTURE

These devices were fabricated using the techniques described in section 4.2. The pattern layout was designed to reflect the circuit of figure 6.1. Three devices of this type were fabricated and measured (see section 6.4). However, all failed through open circuits. On two devices the open circuits were caused by electrostatic discharge (ESD) damage, while on one device, the structure failed to overlap. Measurements on this type of device were discontinued in favour of the granular device structures. Section 6.4.1 discusses some of the measurements taken from these devices and provides micrographs of the device patterns and damage.

6.2.1.2 GRANULAR DOUBLE JUNCTION DEVICE STRUCTURE

Figure 6.2 depicts the pattern used for the fabrication of the circuit of figure 6.1 by the granular junction fabrication technique. The fabrication process for these devices is described in section 4.4. In figure 6.2, the individual SET junctions are formed from the constrictions in the pattern. The individual and common capacitive gates are shown, along with the separate source and drain connections to the electrometers. The total area of the pattern at this scale is some 0.64 μm^2. The pattern forming the constrictions used a 2 x 1 gradient of 5nm square pixels. The dose used for the bulk of the features was 510μCcm^{-2} while the regions for the constriction received 910μCcm^{-2}. The entire pattern (including bonding pads etc) was written using the EBPG-5HR Beamwriter. Patterns for each portion of the structure were combined into one job, as described in section 3.3.1.7, with differing exposures and spot sizes. The total job run time was around 1 hour.

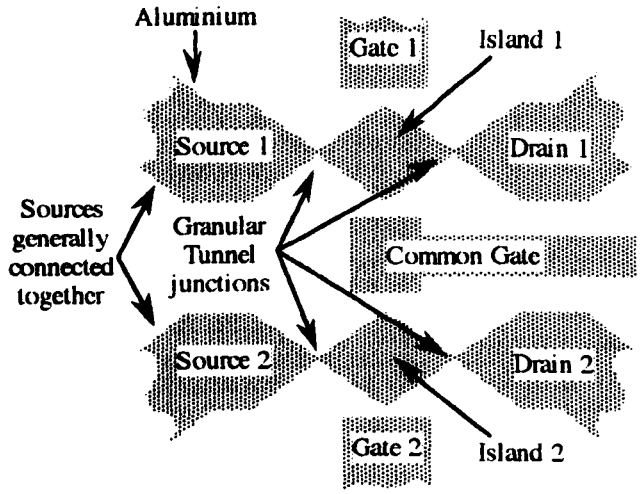


Figure 6.2 Layout of the granular double junction device.

In order to examine the resistance of a single junction, granular junction structures with a single constriction were also fabricated and measured. These structures had a simple single constriction design, with four terminals (two on each side of the tunnel junction) leading away from the constriction. This design allowed both two and four terminal measurements of the tunnel junction resistance.

6.3 MEASUREMENT SYSTEM DESIGN

The system used to measure the single electron devices was intended to measure the *I(V)* curve of the sample. This was to be achieved by applying a known voltage to the sample and measuring the current flowing through the sample. The voltage was varied and the sample current continuously measured. The resulting trace of the *I(V)* curve was plotted on an X-Y chart recorder. The system was, of necessity, a low noise system, capable of measuring at both room temperature and at liquid helium temperatures (4.2K). It must be able to controllably voltage bias the sample and measure the current flowing in the device circuit to an accuracy in the fempto-amp range. Noise coupling and transmission were to be kept to a minimum and ESD protection was required. The system was also to be capable of applying a controllable gate voltage to the sample and providing signals which were a mix of both ac and dc to allow curve gradient sampling.

6.3.1 SYSTEM OVERVIEW

The complete measurement system used for the SET device measurements is designed to exclude noise as far as possible and to avoid unnecessary noise generation[6,7]. The system also maintains high signal levels at all stages, except where application directly

to the device is required. This ensures a high signal to noise ratio at all signal processing stages. With these combined efforts, the noise transmitted to the device is kept as low as possible.

Figure 6.3 provides a schematic of the system layout, illustrating the main components. The system maintains a continuous metallic shield around the device and all signal processing elements. This is to reject any electrically coupled noise or interference. Magnetic field borne noise, however, is not rejected by this shield. To maintain the circuit shielding, signal transfer is via co-axial cables at all stages. Circuit elements are designed to ensure no ground loops are present, preventing common current path or "ground loop" noise coupling. In all cases, power supplies are capacitively decoupled from the active components to a separate "power supply" ground which connects to the "signal" ground for that circuit at only one point. All "ground" lines therefore become referenced to only one point, which is positioned as close to the device as possible.

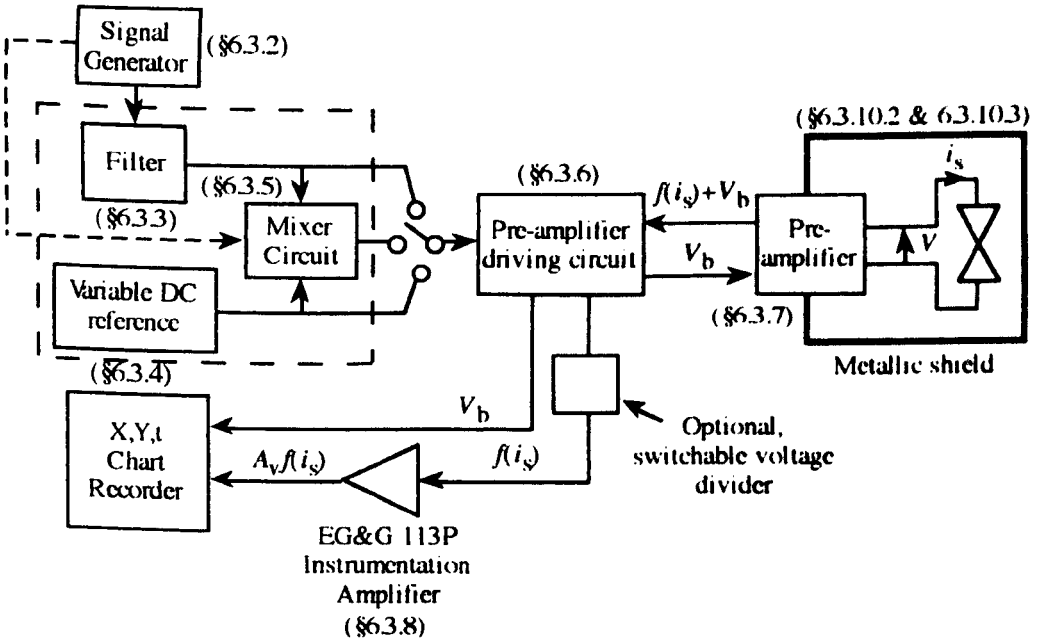


Figure 6.3. Schematic of measurement system.

Finally, device current and voltage signals are handled differentially, reducing common mode noise effects. The signals applied use a 'pseudo' dc. mode. The I-V curve of the device is sampled using low frequency triangular waveforms, producing a linear sweep of the I-V curve. The frequencies used for this sweep are in the range of 0.2mHz to around 50mHz. A complete sweep of the I-V curve can therefore take from around 20 seconds to 8 minutes.

The following sections detail the design and characterization of each of the system elements.

6.3.2 SIGNAL GENERATOR

The signal generator used was a Feedback Instruments FG-601 sine/square/triangle wave signal generator. Signal peak to peak voltage output was 0 to 20V at 0.2mHz to 10MHz.

Analysis of the noise output from this signal generator demonstrated components at 50Hz plus harmonics and major h.f. components at 50 and 100MHz, again with harmonics. The 50Hz components rated at around -55dBV (or 1.7mV), with higher order components falling off in magnitude and becoming undetectable at around 500Hz. The 50 and 100MHz components measured around -45dBm into 50Ω from a nominal 600Ω source. This represents roughly 16mV rms noise output into an open circuit at 50 and 100MHz. Higher order components, again, dropped off rapidly, reaching around -80dBm (or 0.3mV into an open circuit) by around 400MHz. The amplitude ratios and frequency component spacings of this noise were indicative of a square wave clock pulse coupling into the output signal.

6.3.3 FILTER STAGE

In order to supply a "clean" pseudo dc. signal to the pre-amplifier driving circuit, a suitable low pass "brick wall" filter was required. The cut-off frequency was chosen at 2Hz. To provide adequate filtering at all frequencies, the filter circuit was broken into two parts, a radio and high frequency section (rf & hf), and a low frequency (lf) active filter section. A final offset correction stage was also included. The relationship between output and input in this circuit need not be precise as the applied bias voltage can be measured before application to the pre-amplifier circuit. The major requirement of this circuit is the rejection of frequency components above 2Hz. The following sections detail the filter design while appendices 6.1 and 6.2 display the circuit diagrams.

6.3.3.1 RADIO AND HIGH FREQUENCY FILTER SECTION

Filtering at radio and high frequencies (>100kHz) is required as the components used in the active filter stages begin to pass signals with frequency above 100kHz to 500kHz. High frequency signals can easily be down-converted to lower frequencies by the non-linear circuitry of the active components, increasing the baseline noise level. The input stage of the filter was designed to provide hf and rf rejection, preventing these frequency components being passed onto the following active circuitry.

The connection of two bulkhead, π-section, rf-interference (rfi) filters in a differential arrangement within an isolated metallic box completely isolates this filter from the circuit shielding box. This arrangement effectively provides a short circuit to signals above 30MHz. The rfi filter stage feeds directly into a differential single pole RC low pass filter with a cut-off frequency of 128kHz and dc voltage division of 2. The RC filter stage feeds a differential buffer consisting of two unity-gain FET input operational amplifiers, each having low input bias current (10pA maximum). The high input impedance of the FET input operational amplifier allows the use of high R values in the filter without producing significant loading or offset errors in the input stage. The input RC filter also provides a return path for the buffer amplifier's input bias current. The low value of this bias current generates insignificant offset errors despite the relatively large impedance in the input circuit.

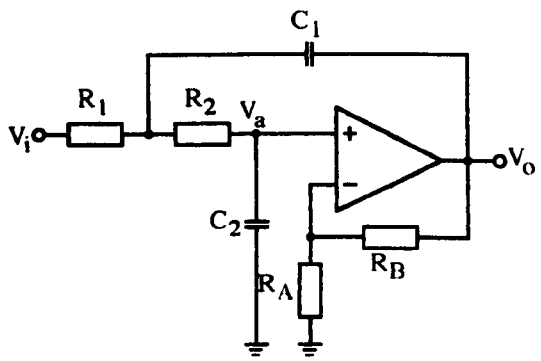
6.3.3.2 LOW FREQUENCY (2HZ TO 500KHZ) ACTIVE FILTER

This filter stage consists of a differential input buffer feeding an instrumentation amplifier in negative unity gain configuration. The instrumentation amplifier performs a differential to single ended signal conversion. The high common mode rejection ratio (CMRR) is useful in eliminating common mode ground borne or magnetically coupled noise. For an INA103KP, the CMRR is typically 86dB at dc. to 600Hz at gain = 1. This common mode rejection is in addition to that supplied by the differential unity gain input buffer stage which is capable of supplying CMR to above 1MHz. The input circuitry of the instrumentation amplifier consists of another differential single pole RC filter stage with cut-off frequency of 24kHz and dc. voltage division of 1.424. The RC filter provides a return path for the amplifier input bias current, at an appropriate impedance level to avoid significant offset errors. In conjunction with the first, or high frequency stage voltage divider, this second passive filter stage provides some compensation for the active filter gain through a total voltage division of 2.848 to about 2% maximum error on the division. With the gain compensation of these two input voltage dividers, an overall gain error of 0.5% is achieved on the nominal gain of 1 for the filter. The worst case error on voltage transfer through the input stage is 4% on the input voltage simply through resistor tolerances.

The active filter stage is a six pole Bessel-Thomson response 2Hz active filter. This filter is designed to give -120dB of attenuation at 20Hz, to ensure complete rejection of 50Hz components from the signal source, as described in section 6.3.2.

6.3.3.3 DESIGN OF THE BESSEL THOMSON FILTER.

The Bessel-Thomson response was chosen for this filter as this response eliminates oscillations at frequencies near the cut-off frequency and maintains the pass band phase response of the filter. Neither of these features are characteristic of Butterworth or Chebyshev filters. A Bessel-Thomson filter is a time domain implementation, not a frequency domain implementation as compared with Butterworth and Chebyshev filters. A full analysis of this filter type and comparison with Butterworth and Chebyshev responses is given in chapter 10 of van Valkenberg's book[8]. The filter implementation is that of three cascaded second order Sallen and Key[9] active filters. The Bessel-Thomson response is ensured by the appropriate choice of filter design values. The circuit provides a 6th order voltage controlled voltage source filter. The Sallen and Key circuit is shown in figure 6.4



Low Pass Sallen & Key filter section.

Figure 6.4. Sallen and Key active filter circuit.

Two transfer functions exist in this circuit. The first defines the low frequency pass-band gain, K, or the gain from V_a to V_o. K is given by equation 6.1. The second transfer function defines the transfer from V_i to V_o and the frequency response of the circuit. The general form of the Laplace transfer function for this circuit is given in equation 6.2

$$\frac{V_o}{V_a} = K = 1 + \frac{R_B}{R_A}. \qquad [6.1]$$

$$T(s) = \frac{V_o}{V_i} = \frac{K\omega_o^2}{s^2 + \left(\frac{\omega_o}{Q}\right)s + \omega_o^2}. \qquad [6.2]$$

where ω_o is the cut-off frequency, in radians per second, and Q is the filter quality factor. Now,

$$\frac{\omega_o}{Q} = \frac{1}{R_1C_1} + \frac{1}{R_2C_1} + \frac{1}{R_2C_2} - \frac{K}{R_2C_2}, \text{ and, } \omega_o^2 = \frac{1}{R_1R_2C_1C_2}. \quad [6.3]$$

If we normalize the resistance and capacitance values such that, $R_1=R_2=1$ and $C_1=C_2=1$, then equation 6.2 becomes:

$$T(s) = \frac{K}{s^2 + (3-K)s + 1} \quad [6.4]$$

This normalizes the cut-off frequency to 1 rads^{-1}. From equation 6.4 and the fact that normalised, $\omega_o=1$, we can see that

$$K = 3 - \frac{1}{Q}$$
$$\Rightarrow R_B = 2 - \frac{1}{Q}, \text{ for } R_A = 1 \quad [6.5]$$

Design of active filters utilizing the Sallen and Key circuit is normally done by the appropriate choice of filter section gains and cut-off frequencies. These parameters can easily be found in tables for a particular filter type to remove the need to calculate the parameters for each particular design[8,10]. Design of this filter was effected by deciding on a gain of -120dBV at 20Hz, and requiring a cut-off (-3dBV) frequency of 2Hz. This was to ensure maximum attenuation of the signal at 50Hz and above but still retaining the ability to pass signals with frequencies of less than 2Hz. This design demands a sixth order (or 6 pole) filter. The Bessel-Thomson response for a 6 pole transfer function has a combination of Q and ω_n as given in table 6.1[8]. n is the filter stage number.

Filter Stage, n.	Normalized Stage Cut-off frequency, ω_n, /rads^{-1}	Stage Q factor, Q_n.
1	4.336	0.510
2	4.566	0.611
3	5.149	1.023

Table 6.1. Design parameters for filter.

ω_m and Q are the pole pairs for the Bessel-Thomson response, $T_m(s)$, (with m = 6). As the Bessel-Thomson response is defined in the time domain, usual frequency normalization techniques do not apply. However, the normalized frequency at which a given loss exists can be determined. Again, this is available in design tables[11]. In this case we are specifying a -3dBV point which, from the use of tables, provides a

normalization frequency of $\omega_c = 2.7$ rads^{-1} for a 6 pole filter. This can then be used to normalize the stage cut-off frequencies through $\omega_{cn}=\omega_n/\omega_c$. Similarly, equation 6.5 allows calculation of the stage gain, K_n. This leads to the Sallen and Key *normalized* design table of:

Filter Stage, n.	Normalized Stage Cut-off frequency, ω_{cn}, /rads^{-1}	Stage Gain, K_n.
1	1.606	1.039
2	1.691	1.363
3	1.907	2.022
	Total Gain	2.863

Table 6.2. Normalized Sallen & Key Design parameters for filter.

From this table, we can find the actual cut-off frequencies based on our required 2Hz -3dB point, bearing in mind that $\omega_0 = 2\text{Hz} = 12.57\text{rads}^{-1}$. Denormalization is achieved through $\omega_{on}=\omega_0\omega_{cn}$. Using the stage gains of table 6.2 and choosing a value for the capacitance, C, we obtain the resistor values if we set $R_1 = R_2 = R_A = R$ and $R_B = (K-1)R$. We then have $\omega_0^2 = 1/R^2C^2$, or $\omega_0 = 1/RC$ in the circuit of figure 6.4. Table 6.3, below, details the full design for the filter. Appendix 6.2 details the full circuit design of this filter, including the component values and types used for design implementation.

Filter Stage, n.	Cut-off frequency, ω_{on}, /rads^{-1}	Stage Gain, K_n.	C /μF	R /kΩ	R_B /kΩ
1	20.19	1.039	1	49.53	1.932
2	21.26	1.363	1	47.04	17.08
3	23.98	2.022	1	41.70	42.62

Table 6.3. Design parameters for If filter.

6.3.3.4 OUTPUT OFFSET COMPENSATION

The final stage of the filter circuit contains an inverting amplifier (shown in appendix 6.1) in unity gain configuration having a single pole RC (f_c=9.5kHz) filter. Constant current diodes supply 5.1V Zener diodes, ensuring a reference voltage of 10.2V across the trim potentiometer. This trimming voltage is fed to the inverting amplifier in current injection compensation mode to provide ±160mV of trim range. This trim range more than covers the worst case projected offsets of this circuit.

6.3.3.5 FILTER FREQUENCY RESPONSE

Using the signal generator of section 6.3.2 to feed the filter circuit with a signal of known frequency and amplitude, together with the instrumentation amplifier of section 6.3.8 to amplify the output of the filter and an oscilloscope, the frequency response of the filter was obtained. Figure 6.5 displays a Bode and phase plot of the filter frequency response. High frequency characterization showed no significant addition of any frequency components, and also elimination of the 50 and 100MHz components noted in section 6.3.2.

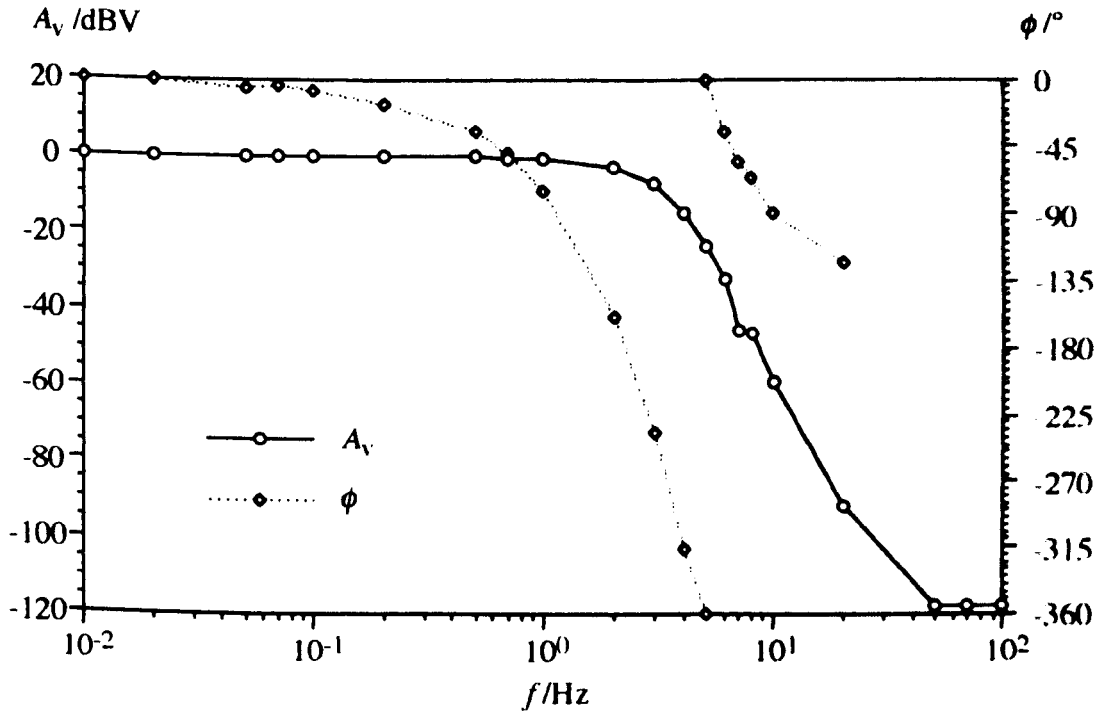


Figure 6.5 Bode plot and phase plot of input filter frequency response.

Figure 6.5 shows that the design criteria of the filter were realized. The total attenuation within one decade is almost 120dBV, while the total phase is indicative of a 6 pole filter, each pole in a filter providing -90° of phase shift. The attenuation is shown to level out at 120dBV, this is simply due to the noise level on the output. The signal on the oscilloscope became indistinct from the noise at this point. The peak to peak noise level as read from the oscilloscope was around 1μV.

6.3.3.6 FILTER NOISE OUTPUT

To measure the noise output from the filter circuit, the input connection was short circuited and the output noise signal measured. The output was fed to an EG&G Model

113P pre-amplifier. This provided enough gain to measure the circuit noise output, but also added its own noise. The output of the pre-amplifier was then fed to an HP3651 FFT spectrum analyzer. This arrangement is shown schematically in figure 6.6.

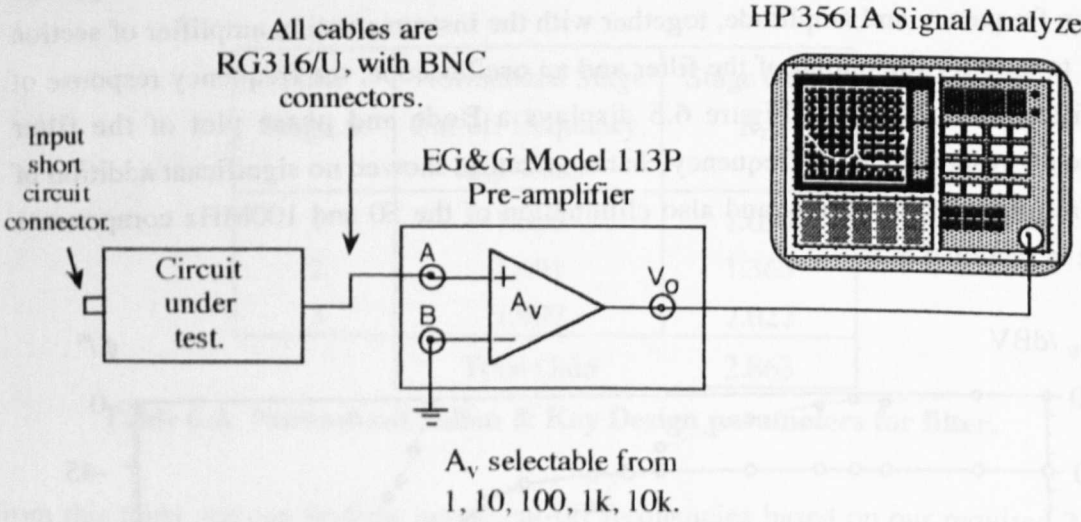


Figure 6.6. Circuit noise test arrangement.

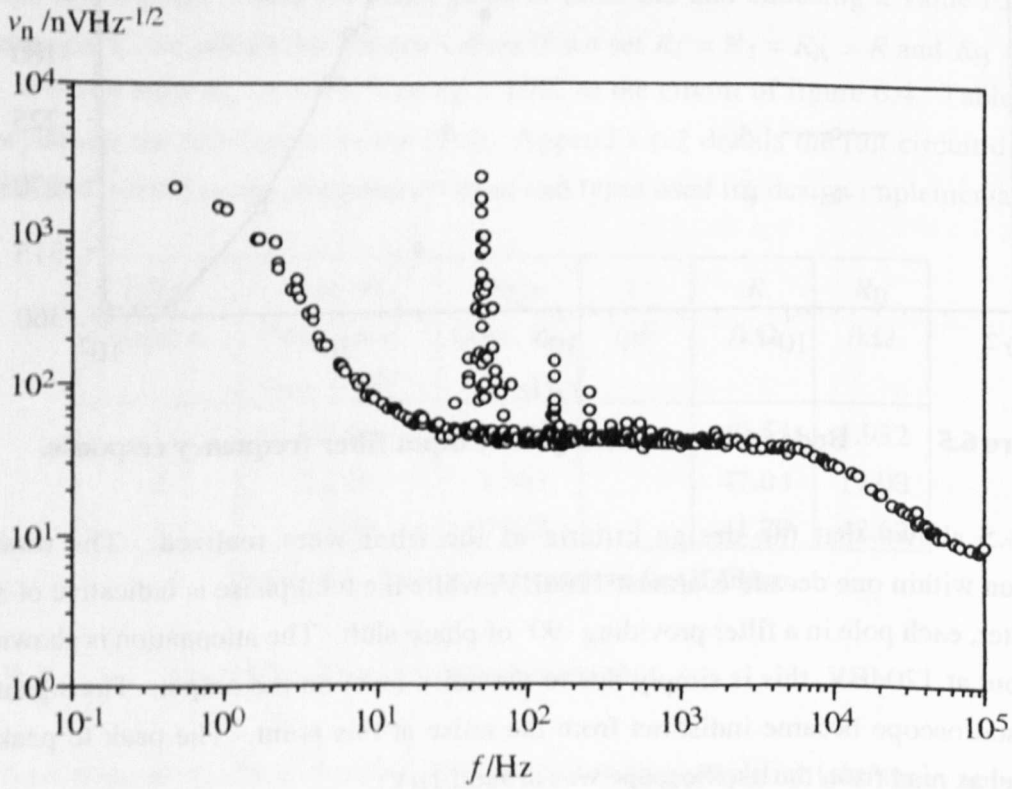


Figure 6.7 Filter circuit input referred voltage noise.

The test circuit displayed in figure 6.6 was used for all the noise measurements. Voltage noise measurements, V_m, were taken from the HP3651 spectrum analyzer. Each sample had a particular bandwidth, B, with the measured noise being expressed in

dBV. With the EG&G amplifier having a gain, A_V, and the circuit having a gain, A_b, the input referred voltage noise spectral density, $v_n(RTI)$, for the circuit is calculated from:

$$v_n(RTI) = \frac{1}{\sqrt{B}}\left[\frac{10^{(V_m/20)}}{A_V A_b}\right]. \qquad [6.6]$$

This spectral density is expressed in $VHz^{-1/2}$. The results for the filter circuit are presented in figure 6.7.

Figure 6.7 shows that the output from the filter circuit contains the most significant noise output at low frequency. This output is tending to level off, at around 2 to 4 $\mu VHz^{-1/2}$ for frequencies under 1Hz. The 50, 150 and 250Hz noise peaks arise from mains pickup through the coaxial cables used for the measurement, and were present in all noise measurements. Movement of the cables varied the noise peak amplitude and measurement of a short circuited co-axial cable in place of the circuit under test of figure 6.6 produced similar mains noise characteristics to those shown above. It is reasonable to assume that the mains noise pickup is due to pickup in the signal cabling. This however does mean that this noise will be present in the circuit and the signals supplied to a device.

6.3.4 VARIABLE DC. REFERENCE VOLTAGE

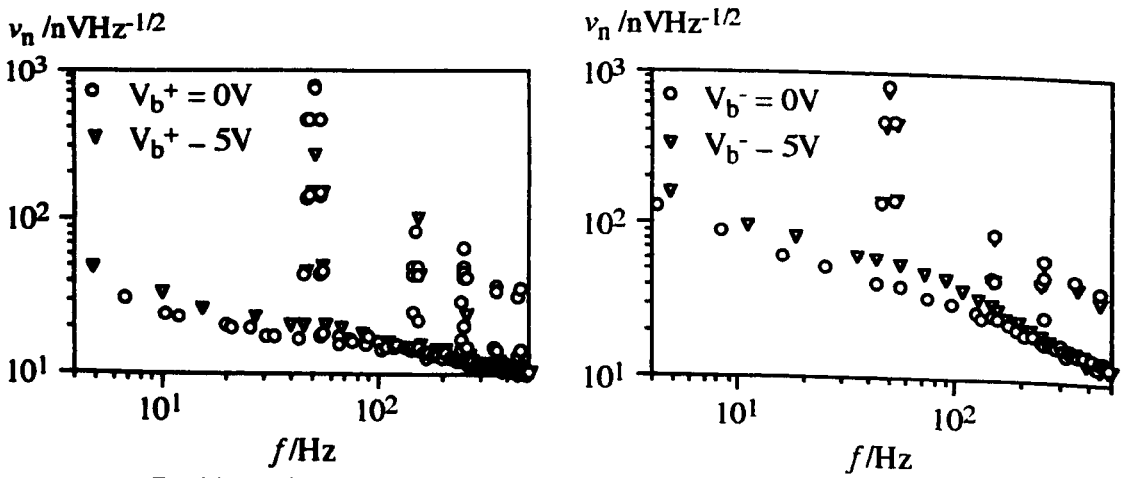
To supply dc. voltages for addition to the input signal, application to a device gate or constant voltage biasing of a device, a stable low noise voltage source is required. Appendix 6.3 gives the circuit diagram for this dc. voltage source. The source provides a positive and negative voltage on two separate output channels. A low noise 10V reference IC. is used for the reference voltage. This voltage is divided by 2 and use of a 10 turn potentiometer provides a variable 0 to 5V supply. The division error of the reference voltage is 2% maximum. A single pole low pass RC filter with a cut-off frequency of 640Hz is present on the output.

The variable voltage output is supplied to both a buffer amplifier and an inverting amplifier. The inverting amplifier provides the buffered negative voltage output. Each of the positive and negative circuits contain single pole RC filters with cut-off frequencies of 160Hz. The inverting stage output has a maximum gain error of 2%. Input loading effects are removed by feeding the low input impedance inverting amplifier direct from the output of the voltage buffer circuit. Each channel carries a worst case maximum offset error of $\pm 500\mu V$, arising from the amplifier input offset

voltage. Input bias current contributions to this offset voltage are negligible due to the low (worst case ±2pA) input bias current of the OPA111AM operational amplifiers.

6.3.4.1 REFERENCE VOLTAGE OUTPUT NOISE

A noise test of this circuit was carried out. This test was the same as that described in section 6.3.3.6. The results are detailed in figure 6.8.



a. Positive reference voltage noise spectral density.

b. Negative reference voltage noise spectral density.

Figure 6.8 Reference voltage noise spectral density.

As in all the noise tests, 50, 150, 250 etc. noise peaks are present in the noise spectral density. These are generated through coupling, as discussed in section 6.3.3.6. The two graphs, figures 6.8a & b, represent the noise output for the positive and negative reference voltages respectively. In both cases, data is presented for zero reference voltage output and full reference voltage output. This shows clearly the increase in noise due to dc. bias voltage supply. The higher level of noise present in the negative reference circuitry is due to the extra operational amplifier in the negative voltage output section.

6.3.5 MIXER CIRCUIT

To allow mixing of signals the circuit of appendix 6.4 was designed. This is a simple four channel unweighted summing amplifier with filtering. Mixing is provided through the four channel summing circuit, each channel having an input impedance of 20kΩ. The circuit is arranged to provide single pole RC filtering at 12kHz. Again, a low input bias current operational amplifier is used to negate bias current based offset voltage errors. The main error arises from the worst case max. ±500μV input offset voltage.

Worst case maximum gain errors of 2% exist on all channels.

The mixer output stage is a single unity gain buffer utilising a low noise bipolar input operational amplifier with low input bias current and input offset voltage. The behaviour of this circuit was verified through a frequency response Bode plot, confirming a 12kHz single pole low pass filter action, with unity gain in the pass band.

6.3.5.1 MIXER OUTPUT NOISE

Again, noise measurements were taken with this circuit as with the previous circuits. In this case, the input signals were grounded using the switching, rather than connecting a zero Ohm BNC connector. The results are presented below in figure 6.9.

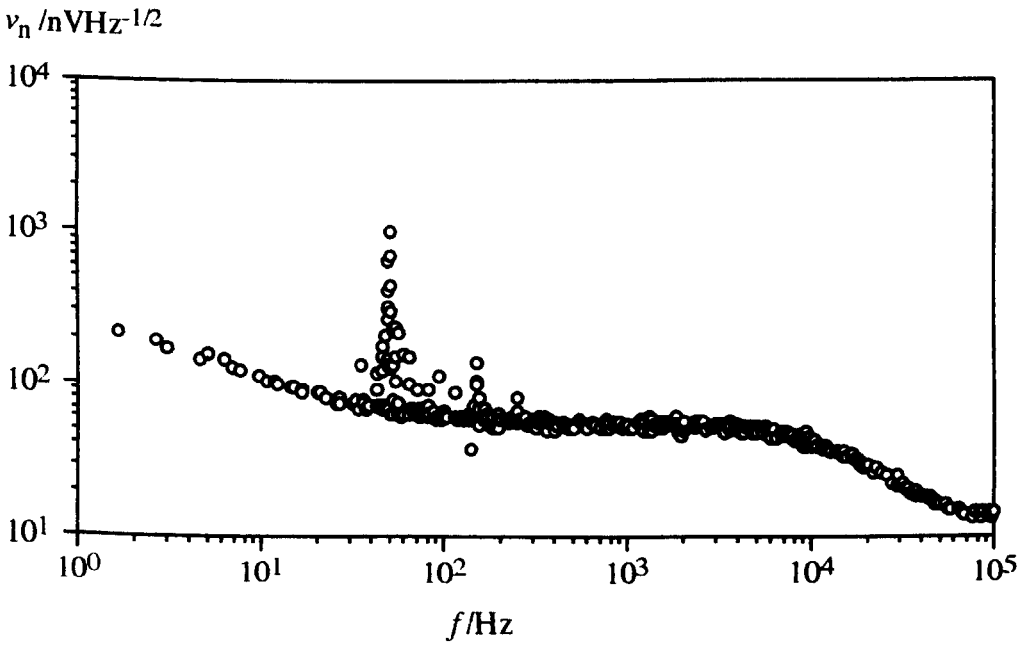


Figure 6.9 Mixer circuit input referred voltage noise.

Again, coupling of 50, 150, 250Hz etc mains noise is present. The bulk of the noise output arises from the OPA111AM operational amplifier used in the circuit, with the output buffer and the resistor circuitry adding little to the noise from this amplifier.

6.3.6 DRIVING CIRCUIT

This circuit takes the input signal and supplies a differential voltage bias for the measurement pre-amplifier. The pre-amplifier returns a differential voltage to the driving circuit which is a function of the sample circuit current plus the bias voltage, as applied to the sample. The driving circuit therefore removes this bias voltage to provide

a differential measurement of the sample current. Figure 6.10 shows a schematic of the driving circuit while appendix 6.5 details the circuit design.

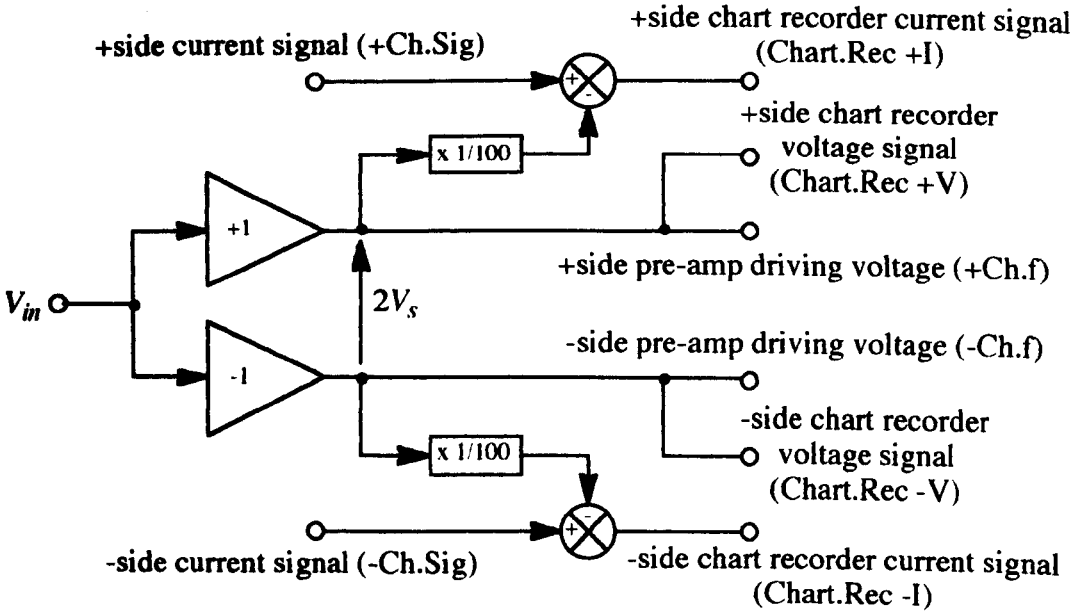


Figure 6.10 Driving circuit function schematic.

The circuit (as seen in appendix 6.5) provides two sets of trimming potentiometers. These are used during the circuit set-up to ensure correct operation and measurement. As the pre-amplifier circuit provides no facility to adjust offset voltages, this adjustment is provided by the driving circuit. Two offset compensations are required; pre-amplifier input offset adjustment and compensation for errors on the bias voltage removal. Input offset compensation is achieved through two offset trim circuits feeding the inverting inputs of the unity gain voltage amplifiers in the driving circuit. Where the supply voltage is zero, the output of both channels of the pre-amplifier should be zero (i.e. no current should be flowing). By shorting the device connections of the pre-amplifiers to ground, the pre-amplifier is placed into high gain voltage follower mode. The amplifier output is then set to zero using the trim circuits, thus removing any offset voltage errors. This results in the circuit measuring zero current output at zero volts input with the device circuit shorted.

The second stage in setting up the circuit is removal of bias voltage subtraction errors, which is achieved by the trim potentiometers on the non-inverting inputs of the driving circuit summing amplifiers. In this case, if no sample is connected to the pre-amplifier (or the device circuit is open) no current can flow. Therefore, regardless of the sample bias voltage supplied the driving circuit current signal should be zero. If this signal is not zero some of the bias voltage is being passed into the current measurement. To correct for this error, the sample bias voltage is set to a high value, around 5V, and any

resulting differential current signal is trimmed to zero.

Again, a requirement for low noise operation is in force, to maintain good signal to noise ratios. As with the filter and signal generating circuits, this constraint is less important than with the pre-amplifier circuit. To help achieve this the signals supplied to the pre-amplifier are 100 times greater than required. The pre-amplifier has a divide by 100 circuit on its input. This circuit is positioned close to the inputs of the operational amplifiers and appropriately shielded to prevent further noise pickup or generation. Therefore any noise generated or passed by the driver circuit will be divided by the same ratio as the signal. The signal to noise ratio will therefore remain constant but the total noise contribution is reduced by 100.

6.3.6.1 NOISE OUTPUT OF DRIVER CIRCUIT

Calculation of the noise contribution of the driving circuit concentrates on two sources. Firstly, the noise contribution from the two unity gain amplifier arrangements which provide the differential bias voltage. Secondly, the noise contribution to the current measurement from the two amplifiers which remove the bias voltage offset from the current measurement.

Dealing with the first source of noise, the amplifiers used to supply the bias voltage to the pre-amplifier are of type LF411. The circuit arrangement has a bandwidth limit of 234kHz, through the RC feedback on the amplifier feedback circuitry. The bias voltage input to the pre-amplifier has an RC filter with 47Ω and 10nF, giving 338kHz bandwidth. Therefore, the bandwidth limitation on the LF411 amplifier is the most significant here. Table 6.4 details the noise spectral density of the LF411 amplifier.

Frequency range	Noise spectral density
2 to 20Hz	$45nVHz^{-1/2}$
20Hz to 100Hz	$30nVHz^{-1/2}$
100Hz to 100kHz	$25nVHz^{-1/2}$

Table 6.4 LF411 voltage noise spectral densities.

From these numbers, it is possible to estimate the total output noise value. The components of this noise are detailed in table 6.5

Source	Bandwidth	Contribution (rms)
10kΩ resistor Johnson noise	234kHz	6.22μV
Amplifier input voltage noise	234kHz	12.1μV
Amplifier input voltage noise	80Hz	0.03μV
Amplifier input voltage noise	20Hz	0.05μV
Total noise output		14.3μV
Differential noise output (√2 times)		20.2μV

Table 6.5 Voltage noise calculation for driving circuit bias voltage supply.

Given that the pre-amplifier input circuit has a differential divide by one hundred input, the actual rms voltage noise delivered to the pre-amplifier circuit is 202nV, assuming a 234kHz bandwidth. In order for this to appear across the device, it must be passed by the pre-amplifier feedback circuitry. This is bandwidth limited to 21Hz by the RC filter in the pre-amplifier, as discussed in section 6.3.7. Therefore, the total noise delivered to the device from the bias voltage supply is insignificant.

Turning now to the second source, the noise added to the current signal from the bias voltage subtraction circuitry. The amplifiers used in this section of the circuit are of type OP-27GN. The noise characteristics of this amplifier are detailed in table 6.6. This circuit has a 142kHz bandwidth arising from the use of 5.1kΩ resistors with 220pF capacitors in the amplifier feedback circuitry. However, we have unity gain amplifiers, subsequent bandwidth limitations and a significant noise input to compare this noise with. The bandwidth limitation to noise on the current signal is set by the instrumentation amplifier. This amplifier combines the two resulting current signals to provide the overall differential current from the circuit. The amplifier bandwidth was normally set to at least 1kHz. Therefore the worst case noise addition from this amplifier is to be considered in a 1kHz bandwidth. Table 6.7 calculates the contributions for this.

Frequency range	Noise spectral density
0.1 Hz to 10Hz	0.09μV (rms)
f_o = 10Hz	3.8nVHz$^{-1/2}$
30Hz	3.3nVHz$^{-1/2}$
1kHz	3.2nVHz$^{-1/2}$
input current noise	1.70pAHz$^{-1/2}$
input current noise in 5.1kΩ	8.67nVHz$^{-1/2}$

Table 6.6 OP-27GN Noise contributions.

Source	Bandwidth	Contribution (rms)
5.1kΩ resistor Johnson noise	1kHz	290nV
Amplifier input voltage noise	1kHz	101nV
Amplifier input voltage noise	30Hz	19nV
Amplifier input voltage noise	10Hz	12nV
Amplifier input current noise in 5.1kΩ	1kHz	274nV
Total noise output		412nV
Differential noise output (√2 times)		582nV

Table 6.7 Voltage noise calculation for driving circuit current correction circuit.

Comparing the noise output from this circuit, with the noise on the current signal (section 6.3.7.1 and table 6.7), we see that this contribution is roughly three orders of magnitude less that the noise generated by the pre-amplifier. The contribution to the total noise from this circuit is therefore insignificant provided the bandwidth limitation on the instrumentation amplifier is applied.

6.3.6.2 ERROR IN DRIVER CIRCUIT DESIGN

There is one error in the design of the driver circuit. The driver circuit supplies the bias voltage to the pre-amplifier and also supplies that voltage directly to a chart recorder. Figure 6.10 and appendix 6.5 show this clearly. Initial measurements on single electron devices using these circuits resulted in a 100% ESD failure rate. The source of this problem was found to be the feed of the bias voltage to the chart recorder. Any noise or voltage/current spikes were fed back to the driver circuit and from there, directly into the pre-amplifier bias voltage. To overcome this problem, measurement of the bias voltage was taken directly from the input to the driver circuit, not the differential output of that circuit. The voltage measured on the chart recorder for the device was therefore:

$$V_b = \frac{2\,V_{in}}{100} \qquad [6.7]$$

The error in the bias voltage measurement is higher due to the increased errors in the circuit between the application of the voltage to the device and the measurement of that voltage. This error is around 2% on the voltage measurement

6.3.7 PRE-AMPLIFIER

The pre-amplifier circuit is a differential current to voltage converter, or trans-impedance amplifier. Figure 6.11 provides a schematic of the circuit. The current

which flows in the sample circuit is the same current that flows through the feedback resistors connected to the operational amplifiers. A normal trans-impedance configuration would have the non-inverting input tied to ground potential. The feedback resistor and the large open loop voltage gain of the operational amplifier produce zero voltage difference between the operational amplifier inputs, resulting in a virtual ground at the inverting input. Application of a voltage directly onto the non-inverting input allows the sample bias voltage to be controlled. This is because the shunt voltage feedback circuit and the amplifier gain combine to produce zero voltage difference between the operational amplifier inputs. The voltage of the inverting input therefore tracks that of the non-inverting input. The voltage applied to the pre-amplifier circuit, V_b in figure 6.11, becomes the sample bias voltage or, V_{xy}.

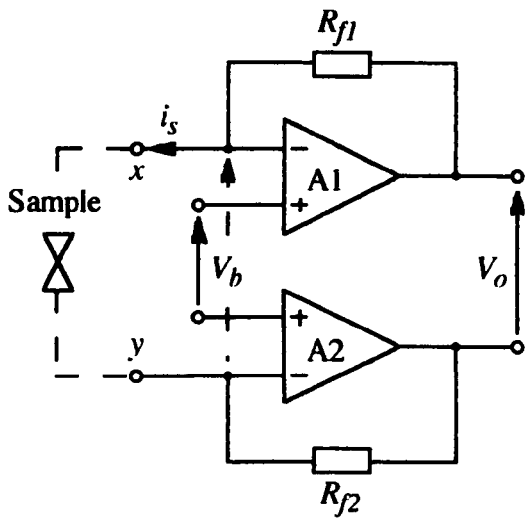


Figure 6.11 Current measurement pre-amplifier schematic.

As each channel of the amplifier contains an offset or bias voltage, the single ended output is given by:

$$V_O(A_n) = i_S R_{fn} + v^+_n \quad (n = 1 \text{ or } 2). \qquad [6.8]$$

For each amplifier we have, nominally, $v^+_n = V_b/2$. If the bias voltage is removed from the output voltage, we have the simple current to voltage gain relation of:

$$V_O(A_n)/i_S = R_{fn} \quad (n = 1 \text{ or } 2). \qquad [6.9]$$

Setting $R_{f1} = R_{f2} = R_f$ and looking at the total differential output, we have the following relation:

$$V_O = 2i_S R_f + V_b. \qquad [6.10]$$

So, subtracting V_b from the output (which is one function of the driving circuit) gives the relation between output voltage and current in the sample circuit as:

$$i_s = V_o/2R_f. \qquad [6.11]$$

With reference to figure 6.11, we see that the sample resistance is a function of the resistance between nodes x and y. That is,

$$R_{xy} = \frac{V_b}{i_s}. \qquad [6.12]$$

From this, the resistance of the sample can be easily calculated, but it is limited by the uncertainty of the resistors in the circuit. The measured sample resistance, R_{sm}, is given by equation 6.13. See figure 6.12 for the full circuit layout.

$$R_{sm} = \frac{V_b}{i_s} - 2R_1. \qquad [6.13]$$

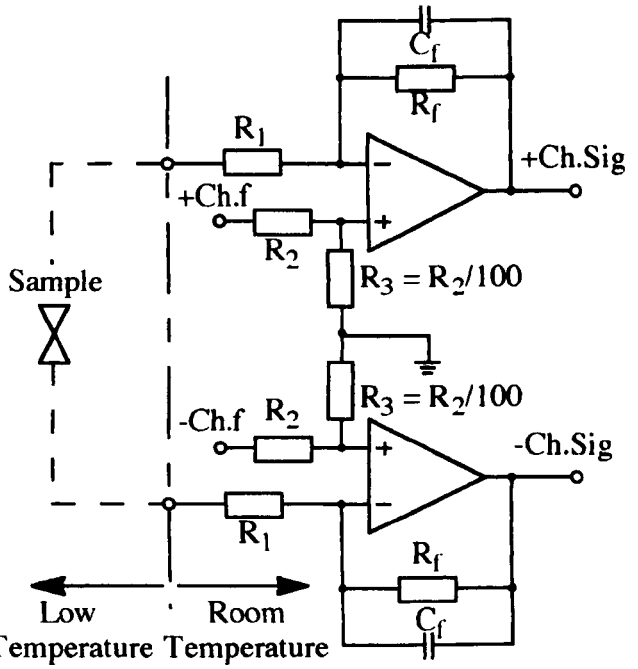


Figure 6.12 General circuit for pre-amplifier.

Appendix 6.6 details the final circuit design for this amplifier and the circuit architecture is shown schematically in figure 6.12. In the bias voltage input channel, we have a low source impedance, divide by 100, voltage divider. This reduces the bias voltage signal to safe operating levels and adds as little noise as possible to the bias signal. The feedback circuit contains a capacitive element to bandwidth limit the Johnson noise of the feedback resistor, the noise output and the noise appearing across the device. A resistor is included, in series with the sample, on both channels to help limit the rate of change of current and provide some ESD protection. This resistor also helps prevent gain peaking in the inverter amplifier circuit caused by the high

capacitance of the input co-axial cables (approx. 100pFm^{-1}).

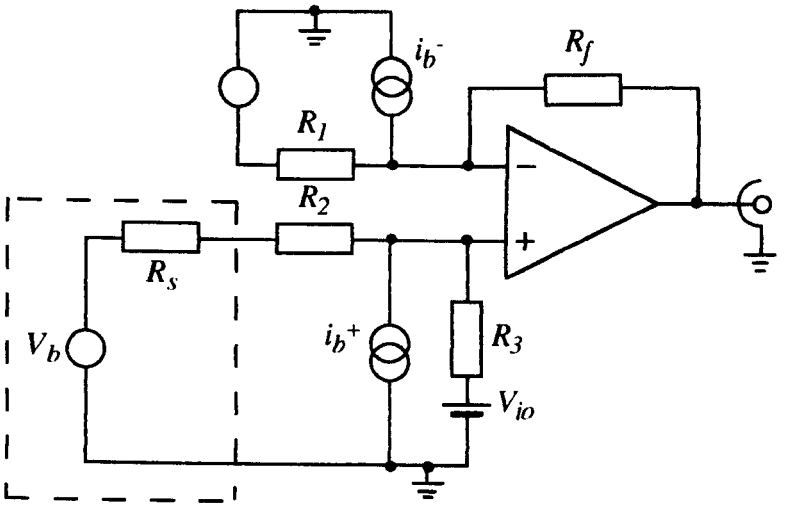
The circuit layout is important in the pre-amplifier to avoid unnecessary noise generation or coupling. As with all circuits, the power supply and signal grounds are kept separate. However, to prevent cross coupling of noise through the circuit board, all signal carrying components are mounted off-board using PTFE stand-offs. Where possible, all these components are mounted directly above a ground plane copper sheet defined on the circuit board.

6.3.7.1 ERRORS AND NOISE IN THE PRE-AMPLIFIER

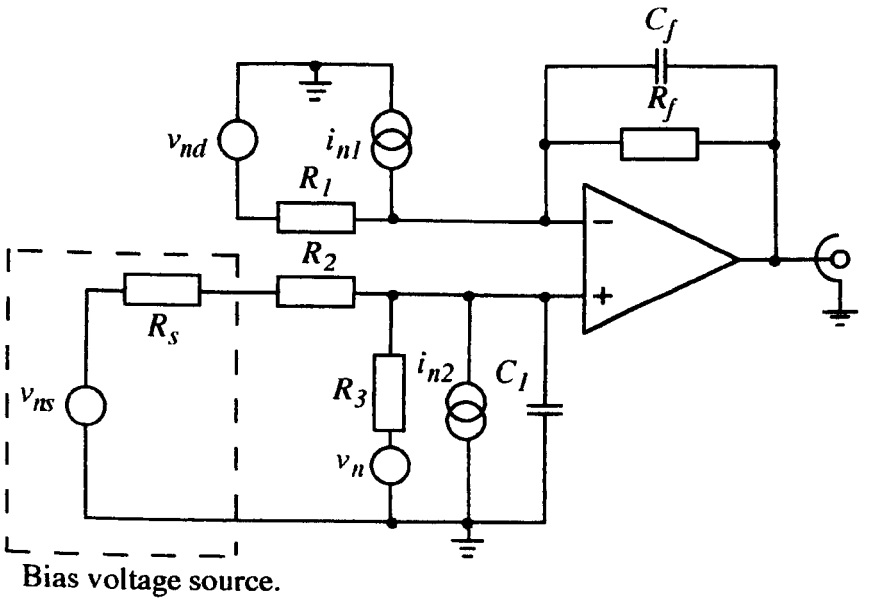
Operation of the pre-amplifier, in terms of accuracy and precision is defined by errors and noise in the circuit. As this circuit provides the main amplification of the sample current, knowledge of the noise behaviour and magnitude of errors is essential to be able to predict the measurement conditions of the sample under test and to properly analyze the results. Figure 6.13 presents two models of the amplifier circuit, figure 6.13a models the dc offsets, while figure 6.13b models the noise sources. Other errors in this circuit arise from the resistor tolerances, which are fixed, systematic, errors.

In the pre-amplifier circuit, there are a variety of sources of error and noise. In terms of systematic errors, or offsets, i_b^+ and i_b^- model the input bias currents of the amplifier while v_{io} models the input offset voltage. Noise sources are modelled by, i_{n1}, i_{n2}, v_{nd}, v_{ns} and v_n. These are the amplifier input noise currents, device and voltage source noise and amplifier input referred voltage noise respectively.

Examining the systematic circuit errors, we have the following parameters to note. The resistor tolerance is 1% in all cases, with 100ppm°C^{-1} temperature coefficient of resistance (TCR). For the amplifier used, the input offset current and voltage, and input bias current are given in table 6.8.



a) DC bias equivalent circuit.



b) Noise equivalent circuit.

Figure 6.13. Model of pre-amplifier circuit. (Single ended)

Parameter		Typ. Value	Max. Value	Units
I_b		±40	±75	fA
I_{io}		30		fA
V_{io}		±140	±500	μV
V_{io} drift			±5	μV°C^{-1}
Noise				
V_n	f_o = 10Hz	92		nVHz$^{-1/2}$
	f_o = 100Hz	78		nVHz$^{-1/2}$
	f_o = 1kHz	27		nVHz$^{-1/2}$
	f_o = 10kHz	15		nVHz$^{-1/2}$
	f_b = 10Hz to 10kHz (rms)	2.4		μV
	f_b = 0.1Hz to 10Hz (pk-pk)	4		μV
i_n	f_b = 0.1Hz to 10Hz (pk-pk)	2.3		fA
i_n	f_o = 0.1Hz throu'. 20kHz	0.12		fAHz$^{-1/2}$

Table 6.8. Noise and offset parameters for the Burr Brown OPA128LM†.
†Taken from the Burr Brown Linear I.C. data book, Vol. 33/33c.

These values for bias and input offset current lead to values for i_b^- and i_b^+ derived from, $i_b^+ = I_b + 1/2|I_{io}|$, and $i_b^- = I_b - 1/2|I_{io}|$ since, $I_{io} = i_b^+ - i_b^-$ and $I_b = (i_b^+ + i_b^-)/2$. Taking the model of figure 6.13a, we see that the total output offset voltage, V_{off}, is given by:

$$V_{off} = V_{io} + i_b^- R_f + i_b^+ R_3 \| (R_2 + R_s). \qquad [6.14]$$

Assuming we have $R_s << R_2$ (R_s=51Ω, R_2=4.7kΩ) and using other nominal component values (R_f=750MΩ, R_1=750kΩ and R_3=47Ω), we have, for the input referred offset voltage:

$$V_{off} = V_{io} + i_b^- R_f + i_b^+ \frac{R_3 R_2}{(R_3 + R_2)}. \qquad [6.15]$$

Using both typical and maximum errors, we find that the bias currents make no significant contribution to the offset voltage. Therefore, the voltage gain setting of the amplifier circuit does not significantly affect the offset voltage. With the definitions above, we find that the input referred offset voltage is typically ±181μV or a maximum of ±547μV. This offset voltage can be trimmed at circuit set-up, see section 6.3.6. Temperature drift in offsets and component values provides a more limiting error which cannot be trimmed. In this case, two significant sources exist which need to be accounted for. Firstly, the drift in V_{io} which is given in table 6.8. This translates to around 6.67fA°C^{-1} drift in the current measurement. The feedback resistor, R_f, is also a

significant source of drift error, at 100ppm°C^{-1}, or ΔR = 75kΩ°C^{-1}. This drift represents about 0.01% of the resistance value, per °C. This is small in relation to the input offset voltage drift error. However, the total error in measured current can be calculated from (for single ended operation):

$$\Delta i_m = \frac{\Delta V_{io} \Delta T}{R_f + \left(\frac{dR_f}{dT}\right) R_f \Delta T}, \text{ where, } \frac{dR_f}{dT} = \text{TCR}. \qquad [6.16]$$

Assuming a ±10°C change in amplifier operation temperature due to operation at extremes of the supply voltage, this provides an effective limit on the current measurement accuracy of ±67fA due to uncertainty in the operation temperature of the operational amplifier. Or, as previously, ±6.7fA°C^{-1} current measurement accuracy. The most significant drift is, therefore, that due to the drift in amplifier input offset voltage.

Finally, the resistor tolerances produce systematic errors in the circuit. The bias voltage divider on the non-inverting input contains a 51Ω resistor, from the driver circuit. This resistor is present to offset capacitive loading on the output of the driver circuit operational amplifiers when driving high capacitance coaxial leads. The 51Ω resistor in series with R_2 in figure 6.13a produces a 1.08% error in the division. The maximum error in the division is 2%. The total worst case error on the bias voltage is therefore 3%, or the likely error is around 2.3%. In the current measurement, the most significant errors occur from the 1% tolerances on R_f and R_1. With R_f, we have a direct 1% error on the circuit trans-impedance gain, providing a 1% error in the current measurement. The error resulting from the tolerance on R_1 is again 1%. This value is used in calculating the sample resistance, see equation 6.13. As we have 2 x R_1 in series with the sample in the differential configuration, then R_{sm} can be determined to 1% based on the current value. The total error for determining the sample resistance is likely to be 3.3%, or a worst case of 4%.

Noise sources in this circuit limit the measurement precision, through random errors. Figure 6.13b presents the model for the various noise sources. Table 6.9 lists the contributions to the output noise for various sources, while table 6.10 presents the calculated magnitude of these sources for three values of R_1. The initial circuit design (and the noise characterization) involved R_1=1MΩ. After the noise measurements were taken, this was changed to 750kΩ. Finally, after some measurements on granular devices, R_1 was further reduced to 75kΩ. Hence, the three sets of calculations. Although R_1 was decreased to improve the measurements, care was taken to prevent gain peaking, caused by the input capacitance from the coaxial leads. This would result

where the closed loop circuit had insufficient gain margin, resulting in the circuit oscillating. The OPA128 open loop gain has a maximum guaranteed value of 110dB, which rolls off at -20dBdec^{-1} around 1.5 to 2 Hz. With R_f = 750MΩ and R_1 = 75kΩ and the cut off frequency of 21Hz, adequate gain margin is maintained as the 80dB open loop gain is reached around 150 to 200 Hz.

Source Number, m.	Source	Noise Type	Voltage output contribution
m = 1	R_1	Johnson Noise	$\sqrt{4k_BTBR_1}\frac{R_f}{R_1}$
m = 2	$R_s = R_2 \| R_3$	Johnson Noise	$\sqrt{4k_BTBR_s}\left(\frac{R_f}{R_1}+1\right)$
m = 3	R_f	Johnson Noise	$\sqrt{4k_BTBR_f}$
m = 4	i_{n1}	Amplifier Current noise.	$i_{n1}R_f\sqrt{B}$
m = 5	i_{n2}	Amplifier Current noise.	$\left(i_{n2}R_s\sqrt{B}\right)\left(\frac{R_f}{R_1}+1\right)$
m = 6	v_n	Amplifier voltage noise.	$v_n\sqrt{B}\left(\frac{R_f}{R_1}+1\right)$

Table 6.9. Noise source contributions in the pre-amplifier†.
†In all cases, the bandwidth, B, is taken as 21Hz (f_c=1/2πR_fC_f).

The total root mean square noise output of the circuit can be estimated from the following equation:

$$V_o(noise) = \sqrt{\sum_{m=1}^{6} V_{nm}^2}\,. \qquad [6.17]$$

the results of these calculations are given in table 6.10.

As the circuit operates in a differential mode, signals are multiplied by a factor of two or, put another way, the trans-impedance gain doubles (to 1.5GΩ in the differential circuit from 750MΩ in the single ended circuit). However, the noise sources are independent and uncorrelated. The total output voltage noise is then given by √2 times $V_o(noise)$. For the final circuit design, the total noise output is some 6.4mV and, with the trans-impedance gain of 1.5GΩ, this translates to 4.3pA or 931fAHz$^{-1/2}$ in 21Hz bandwidth. Where R_1 = 750kΩ, the noise output would be 937μV, giving a

measurement current noise of 625pA or 136fAHz$^{-1/2}$.

	R_1 = 1MΩ	R_1 = 750kΩ	R_1 = 75kΩ
Inverter A_v	750	1000	10000
Follower A_v	751	1001	10001
Noise source			
v_n(m=1) /V	442 x 10^{-6}	510 x 10^{-6}	1.62 x 10^{-3}
v_n(m=2) /V	3.02 x 10^{-6}	4.03 x 10^{-6}	40.2 x 10^{-6}
v_n(m=3) /V	16.2 x 10^{-6}	16.2 x 10^{-6}	16.2 x 10^{-6}
v_n(m=4) /V	412 x 10^{-9}	412 x 10^{-9}	412 x 10^{-9}
v_n(m=5) /V	19.2 x 10^{-12}	25.6 x 10^{-12}	256 x 10^{-12}
v_n(m=6) /V	316 x 10^{-6}	422 x 10^{-6}	4.22 x 10^{-3}
Total rms output noise /V	*544 x 10^{-6}*	*663 x 10^{-6}*	*4.52 x 10^{-3}*
rms current noise /fA	*725*	*884*	*6030*
Voltage noise output in 21 Hz /VHz$^{-1/2}$	118 x 10^{-6}	144 x 10^{-6}	985 x 10^{-6}
Current noise in 21 Hz /fAHz$^{-1/2}$	157	192	1310
Voltage noise in 21 Hz (RTI) /VHz$^{-1/2}$	157 x 10^{-9}	144 x 10^{-9}	98.5 x 10^{-9}

Table 6.10 Noise contributions in pre-amplifier circuit (*single ended*), bandwidth = 21Hz.

These values represent the limit on the precision of the measurement for the pre-amplifier circuit. This is the best possible output from the system which also includes the driving circuit contributions and the contributions from the instrumentation amplifier. However, these values represent only estimated figures for the noise output and provide some understanding of the dominant noise sources. In order to determine the true noise limitation, the circuit requires to be tested for noise output. This was done as a separate noise measurement exercise (described in section 6.3.7.2) and during set-up of the measurement system, as will be discussed in section 6.3.11.

6.3.7.2 PRE-AMPLIFIER NOISE PERFORMANCE

Using the noise test circuit and technique described in section 6.3.3.6, the output voltage noise spectral density of the pre-amplifier was obtained. The measurement was taken in a single ended configuration with the device connections of the amplifier shorted to signal ground. From the measured output voltage noise, calculations were made for both the current noise (and therefore precision of the measured current) and the input referred voltage noise. The current noise was calculated using the trans-impedance gain of R_f = 750MΩ. The voltage noise was referred to the amplifier input

to provide comparison with other circuit designs. All three presentations of the data are given in figure 6.14 through adjusting the axis scales to match the curve for output referred voltage noise.

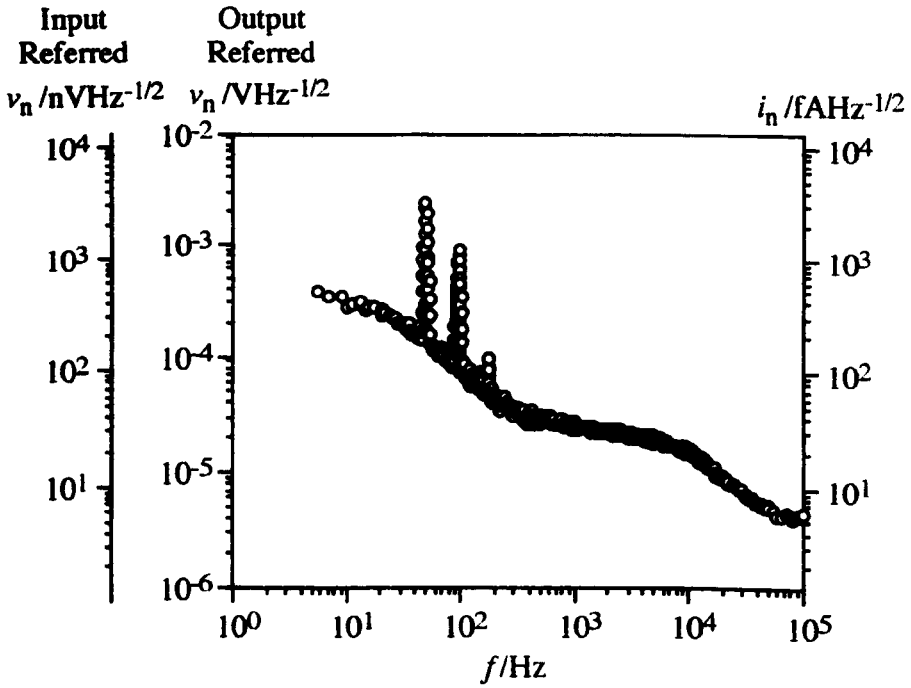


Figure 6.14 Current pre-amplifier noise spectral density (single ended).
(for the pre-amplifier of appendix 6.6 with R_f = 1MΩ)

From the data presented above, the amplifier implementation appears noisier than the calculation. From 5 to 21 Hz, the noise spectral density appears around 3 times higher than expected from calculations. However, the scale of the noise is not unexpected. The magnitude of the mains coupling suggests that the insert wiring is coupling some mains noise into the circuit and that this is being amplified through the voltage gain of the circuit. An interesting feature of the noise calculations (table 6.10) is that the voltage noise component of the total noise is reduced by reducing R_f, while the current noise increases. This is as expected. However, where the sample resistance is not extremely large, it is important not to lose sensitivity or precision of the voltage measurement across the sample in order to maintain an unnecessarily low current noise. This is the reason why R_f was eventually reduced to 75kΩ. The magnitude of sample resistances meant that the sample behaviour (particularly the room temperature sample behaviour) was being lost in noise arising from R_f.

This circuit is closest to the device and provides the amplification required to recover the device response to excitation. Therefore noise present at this stage will be coupled into the device affecting the measurement. Using the values of the noise, as referred to

input, we see the noise level presented to the device. All other noise sources are fed in through the input channel which had a divide by 100 potential divider. These noise sources are therefore insignificant on the scale of the noise measured here. The measured 50Hz noise peaks at 30μV. This noise is effectively common mode and is subtracted from the output when the differential configuration is used. It is, however present in the signal lines. This noise represents an effective temperature of around 350mK, which is well below the temperatures at which the devices are measured. Low frequency random noise of 500nVHz$^{-1/2}$ over 20Hz bandwidth in differential mode results in 4.5μV, rms noise or, a noise source temperature of around 50mK.

6.3.8 INSTRUMENTATION AMPLIFIER

The instrumentation amplifier used for measurement of the differential sample current signal was an EG&G Instruments Model 113P pre-amplifier. This amplifier takes the driver circuit differential output signal representing the sample current and produces an amplified single ended signal which drives a chart recorder Y channel. The amplifier uses dual channel (A & B) inputs in an A - B configuration. The differential gain can be selected from 10, 100, 1000 and 10000 times. Each input is selectable between ac. and dc. coupling. With ac. coupling, both cut-off frequencies in the amplifier pass band are selectable while in dc. coupling mode, only the upper cut-off frequency is selectable. The lower cut-off frequencies are selectable from 0.3, 1, 3, 10, 30, 100 and 300Hz while the upper cut-off frequency is selectable from 300, 1k, 3k, 10k, 30k, 100k and 300kHz. This amplifier was generally used in dc. coupling mode with 1kHz set as the upper cut-off frequency.

6.3.8.1 EG&G 113P NOISE OUTPUT

Again, noise output measurements were taken from this circuit, in order to characterize the contribution of this amplifier to any circuit noise output. The plot shown in figure 6.15 represented the input referred voltage noise of the EG&G 113P instrumentation amplifier. The circuit used for this measurement simply fed the amplifier output directly into to the spectrum analyzer, with the amplifier inputs grounded.

The plot in figure 6.15 was taken with the gain setting of the amplifier set to 1k. Similar results were achieved for 10k gain. The major features of this noise are again, the presence of 50, 150 and 250 noise peaks (also with an appearance of 100 and 200Hz peaks), and the generally low level of noise. The contribution of less than 10nVHz$^{-1/2}$, at most frequencies, rising to 30nVHz$^{-1/2}$ is insignificant when compared with the noise levels generated in the other measurement circuits.

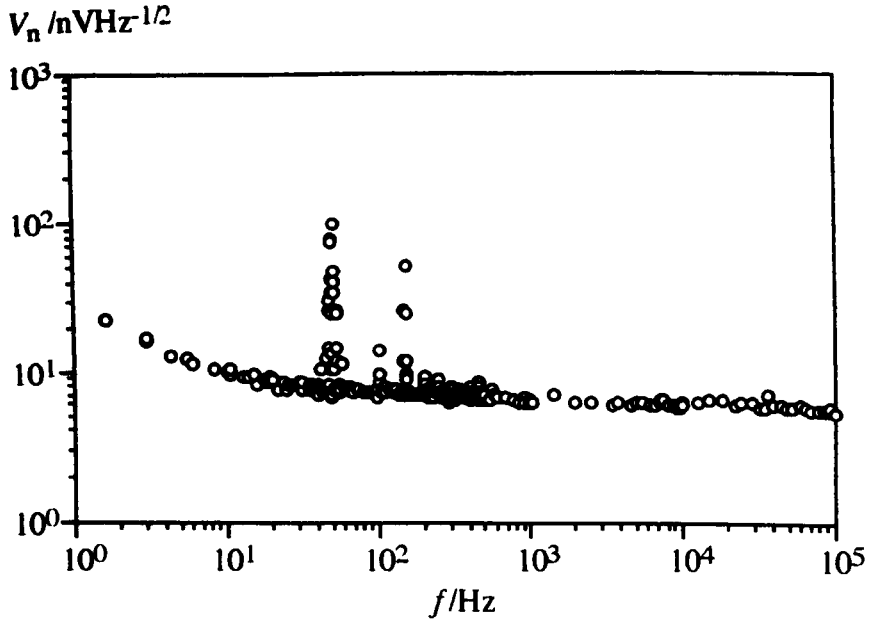


Figure 6.15 EG&G 113P input referred voltage noise.

6.3.9 POWER CIRCUITRY

The power source for these circuits is required to be stable and produce very little noise. The supply was split between two systems. A battery supply was used to power the driver circuit and pre-amplifier and a low noise power supply was used to power the remaining circuits. For the driver circuit and pre-amplifier, two 12V 6Ah lead-acid batteries were used to provide a ±12V power supply. Additional filtering was employed through the use of a circuit which also provided fuse protection and visual display of the battery supply output. The total supply was set to a minimum of ±11.5V, with 200mA fusing. LEDs were used to indicate sufficient battery supply voltage. These were arranged to extinguish when the supply voltage dropped below 11.5V. Appendix 6.7 provides a diagram of this circuit.

To supply the remaining circuits, a low noise ±15V power supply was designed and built by the electronics workshop. The circuit diagram for this is given in appendix 6.8. The circuit supplied ±15V at 3A rating from two supply channels. Again, the circuit of appendix 6.7 was applied before feeding power to the signal conditioning circuits.

As it is possible for coupled or generated high frequency noise to become mixed down to lower frequencies, the low noise power supply was tested for high frequency noise output. No significant output was found. The noise output spectrum for frequencies below 100kHz was tested as with the other circuits discussed here. In this case, the noise spectrum was analyzed both at open circuit output, and at 3A output current,

through a 10Ω power resistor. These tests demonstrated significant noise output at 50, 100, 150, 200Hz, etc. up to 2kHz. The dominant noise output arose from the 50 and 100Hz contributions, resulting in some 3mV peak to peak ripple on the 15V DC output, or -50dBV of ripple. When accepting the noise output of this circuit, consideration was taken of the 80 to 100dB (0 to around 500Hz) power supply rejection ratios (PSRR) of the active components used in these circuits. This PSRR, along with the decoupling arrangements in the circuit designs, ensures that feeding of supply noise to the signal lines should not be a significant problem. The 3mV ripple should therefore be reduced to around 30nV ripple on the signal lines due to the PSRR.

6.3.10 MECHANICAL CONSTRUCTION

While the electronic circuits were designed to eliminate, as far as possible, noise coupling or generation, it was important to maintain a good shield around the circuits and solid mechanical design. This is important if vibration induced noise is to be avoided. Equation 6.18 represents a major source of vibration induced noise, particularly for coaxial cables.

$$Q = CV, \quad \Rightarrow \quad \frac{dQ}{dt} = C\frac{dV}{dt} + V\frac{dC}{dt} \qquad [6.18]$$

In this equation, we see that current noise is induced capacitively through changing electric fields and through changing capacitance. Movement of leads or circuit components is the major source of the *dC/dt* term in equation 6.18 and, therefore, must be eliminated as far as possible.

To this end, each circuit was housed within an aluminium alloy die cast box. The cabling for the insert was supplied through a pipe and connected to a sample header, rigidly held in an aluminium can. The die cast boxes provided a continuous unbroken metallic shield connected to the supply ground. Where signals were passed through the box, insulated BNC bulkhead connectors were used. This required a ground, or shield, connection to be made explicitly. In this way ground loops in the signal lines were prevented.

6.3.10.1 PRINTED CIRCUIT BOARDS

All circuits, except the driving circuit, where constructed on printed circuit boards (PCBs). The PCBs were all formed from double sided resist coated copper boards. Patterning was carried out via masking patterns laser-printed onto acetates. Board layout followed standard 0.1 inch through hole component layout rules. The boards

were designed to have a continuous ground plane. In all cases, the ground plane was connected to the signal ground at only one point. The signals on the boards were separated, as far as possible, into power supply on one side and signal lines on the other. This was to provide maximum insulation and minimum capacitance between the power and signal lines, thereby reducing leakage, coupling, interference and crosstalk. Where PCBs were not used, circuit construction was on strip board (or Veroboard™). Again the layout and ground plane definitions given above were followed as far as possible.

The use of printed circuit boards maintained the physical placement of components and allowed the circuitry to be rigidly fixed into the circuit boxes. Also, the circuits could be easily replicated if required.

6.3.10.2 INSERT CONSTRUCTION

To perform measurements at liquid helium temperature, an insert was designed to fit a standard 65 or 100 litre liquid helium dewar. This insert provided a continuous metallic shield around the device and pre-amplifier circuit, communication of signals between the amplifier and device with as little cross coupling or generation of noise as possible, selectability of device connections, ESD protection to help prevent damage to the device and rigidity of design to reduce vibration induced noise. The following three diagrams show the general insert construction. Figure 6.16 depicts the relationship between the elements, Figure 6.17 provides the details of the insert shielding can, while figure 6.18 depicts the sample holder arrangement used.

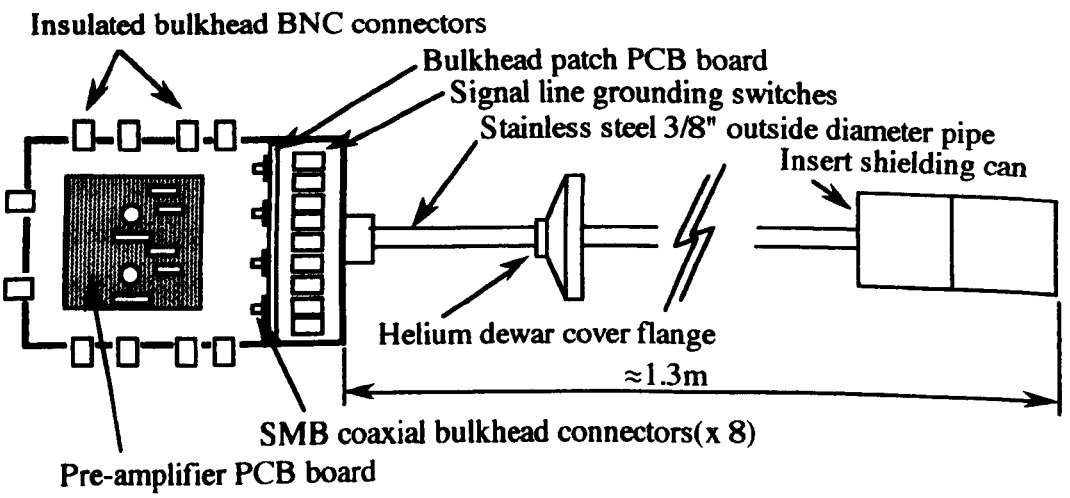


Figure 6.16 Schematic of measurement insert.

In this assembly, the pre-amplifier PCB is fixed into a die-cast aluminium box containing the switching and patch board. From this box, the cabling for the insert is fed down a stainless steel tube to the insert can. The can holds the sample holder

assembly (shown in figure 6.17 and figure 6.18). The bulkhead patch board performs two functions.

Firstly, it allows the appropriate signal line to be connected to the appropriate lead on the device. These signal lines are all co-axial, (RF-MCX or "microcoax" cabling) and are therefore isolated from each other. The cable shields are connected at the sample holder inside the insert shielding can. Secondly, the patch board ground plane provides a measure of shielding between the pre-amplifier board and the incoming signal lines and switching section of the assembly. This is to minimize noise coupled from high level signals to the unshielded switching arrangement to the right of the patch board in figure 6.16. This switching circuit provides ESD protection during sample handling, mounting and measurement and is the same arrangement as that shown in figure 5.14.

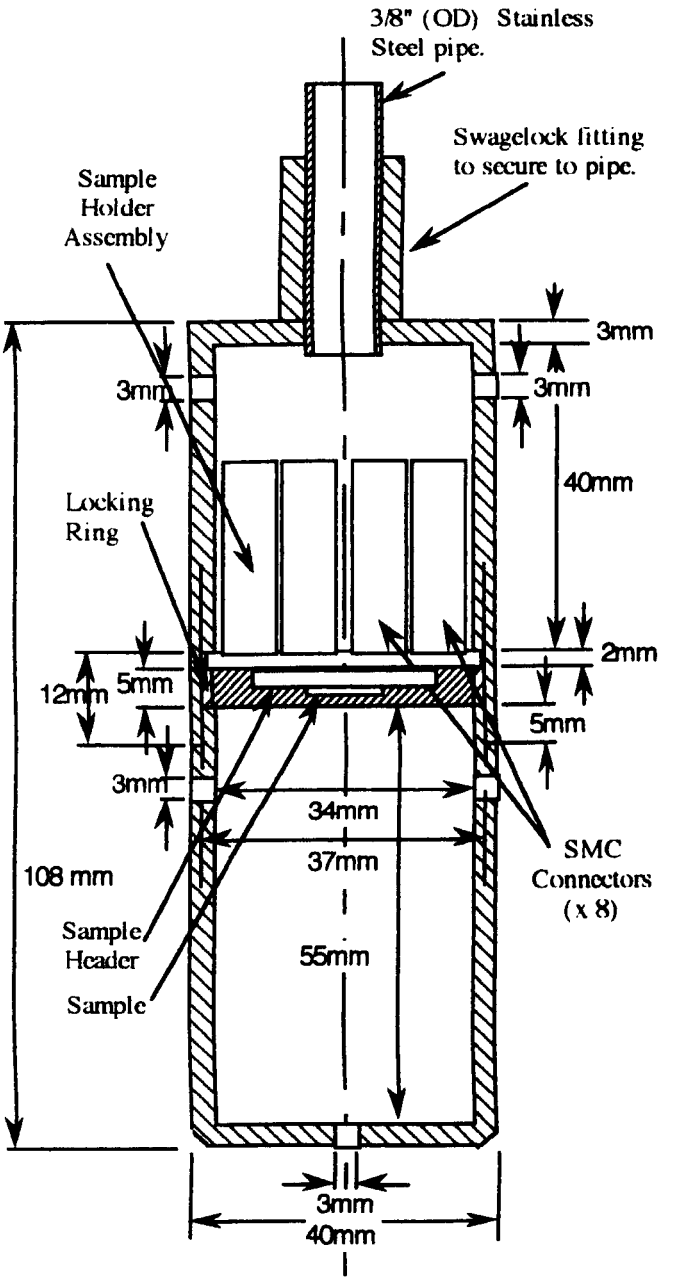


Figure 6.17 Sample holder shielding can construction.

The length of the insert (approx. 1.3m) allows access to the bottom of both 65 litre and 100 litre liquid helium dewars. The depth into the dewar is controlled through the position of the cover flange. This allows for most economical use of the available liquid helium. The sample sits 55mm from the bottom of the insert can (figure 6.17) which then requires a depth of at least 80mm of liquid helium within the dewar to cool the device to 4.2K.

The length of the insert (approx. 1.3m) allows access to the bottom of both 65 litre and 100 litre liquid helium dewars. The depth into the dewar is controlled through the position of the cover flange. This allows for most economical use of the available liquid helium. The sample sits 55mm from the bottom of the insert can (figure 6.17) which then requires a depth of at least 80mm of liquid helium within the dewar to cool the device to 4.2K.

The insert can is fabricated from three aluminium sections. First, the upper section, which provides space for the cabling to enter the can, and contains the sample holder assembly. 3mm diameter holes are drilled into this section, to allow free flow of liquid helium into the can and, more importantly, out of the can. This section is threaded at the bottom, to allow assembly. The second section is a locking ring. The sample holder fits into the upper section of the can, and rests on a 1.5mm wide lip at the end of the threaded portion. The locking ring screws into this thread and locks the sample holder in place. There is a break in the insert PCB ground plane to prevent this ring shorting the sample holder ground plane to the can. Finally, the third section is effectively the lid for the can. This lid screws into the remaining thread in the upper section, to close the can. Again holes are drilled in this section to allow free flow of helium liquid and, particularly, cold helium gas.

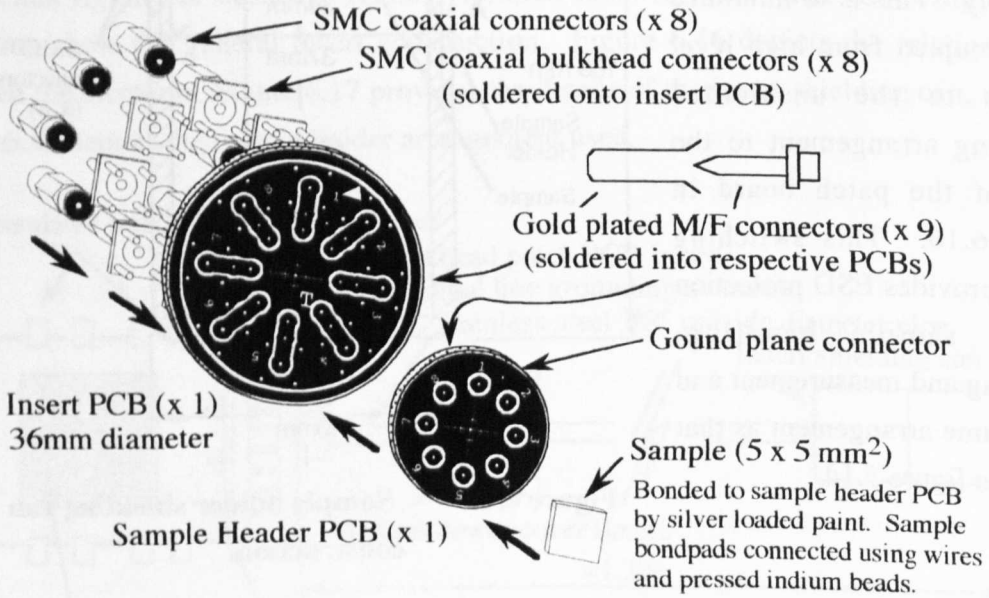


Figure 6.18 Sample header components

The sample holder arrangement is depicted in figure 6.18 (and assembled and in place in figure 6.17). This assembly is formed from two PCBs, eight PCB SMC connectors and nine sets of gold plated pin connectors. The SMC connectors are soldered into the insert PCB, to accept the SMC connectors fitted to the RF-MCX cabling of the insert.

The female sections of the gold plated connectors are fitted (and soldered) into the opposite face of the insert PCB. The sample header PCB has gold plated pins soldered into the opposite side to the sample area. Enamelled wires are soldered into these pins, to allow connection to the other side of the board and bonding to the sample mounted on the front face of sample header PCB. One extra pin (over and above the 8 signal lines) is fitted between connections 1 and 2, to ensure proper orientation of the sample header when inserted into the insert PCB and connection of the ground plane between the PCBs.

In order to mount samples, a duplicate of this sample header was produced. The SMC connectors used with this "bonding header" had the signal lines shorted to the ground plane through 8.2MΩ resistors. This allowed handling of the header and, coupled with ESD protection straps, prevented ESD damage to the sample during bonding, transportation and storage.

6.3.10.3 INSERT WIRING

As already mentioned, the insert is wired using coaxial cable. The cable used is of type RF-MCX (or "micro-coax"). This thin coaxial cable allows eight signal lines to be fed down the stainless steel pipe. The dielectric for this cable is PTFE, compatible with cryogenic applications. The connectors used in the insert cabling are of type SMB and SMC. The SMC connectors (providing screw connections) are used at the cold end of the insert. These connectors are permanently in position and do not require continual connecting and disconnecting. At the patch board, where signal line selection and repeated reconnection is required, SMB type connectors are used. These push fit connectors provide a good contact between the signal and shield lines while giving the flexibility of quick reconnection.

The low frequency characterization work has shown that the noise source temperatures are well below the measurement temperature for the devices. However, some work has been carried out examining the effect of coupled high frequency noise into the environment of a single electron device. This has shown the need for significant filtering at frequencies above a few GHz[12]. In order to assess the amount of high frequency noise received by the sample and the suitability of this insert arrangement for use with single electron devices, high frequency characterization measurements were carried out.

To examine the insert cabling attenuation in the high frequency and GHz regions, an HP8510B Network Analyzer was used in conjunction with an HP8515A two port S-

parameter test set. The S-parameter test set had a calibrated frequency range of 45 MHz to 26.5 GHz. The full frequency span was used in the attenuation measurement. Two measurements were taken. The first measurement was taken without a system calibration, while the second measurement was taken after a system and cabling calibration. To do this, the cabling used to connect to the SMB connectors on the insert was connected between the two ports on the test set. A calibration measurement was taken. Directly after this, the insert was connected to these cables and the frequency response measurement taken.

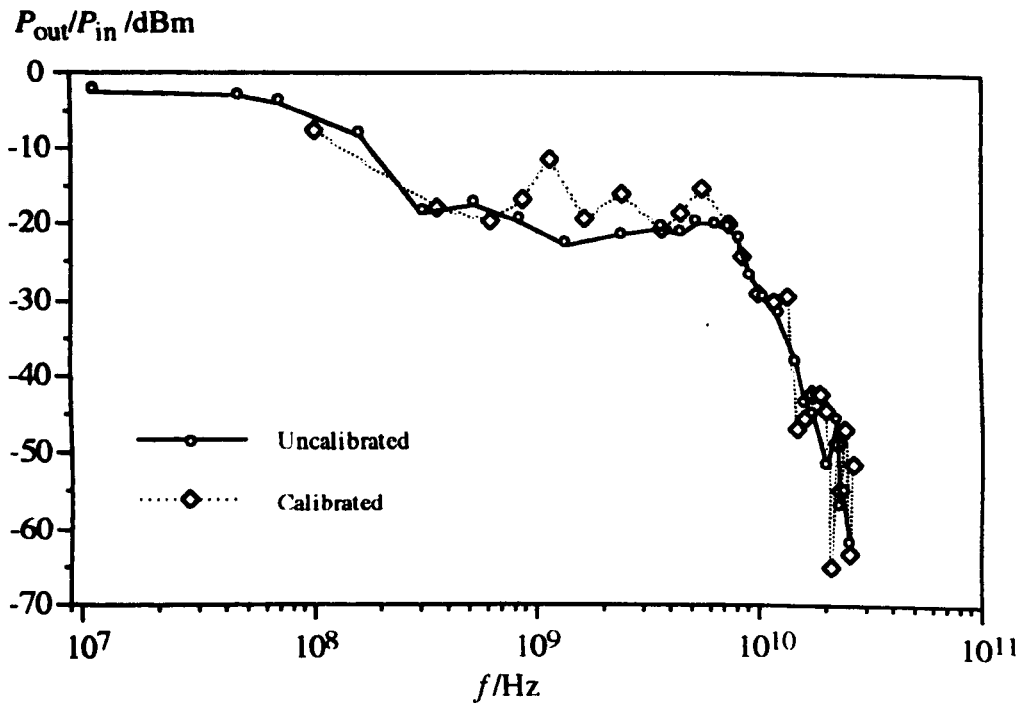


Figure 6.19 Insert cable attenuation.

In both cases, the measurement used a short circuit sample header connected into the insert PCB. Therefore the tests involved twice the length of cabling to the sample. The connections in the signal path consisted of an SMB connector to around two centimetres of wire soldered onto a switch with gold plated wiper and contacts. Further solder connections connected this switch and signal line to an RF-MCX cable. At the "cold" end of the insert, solder joins from the RF-MCX cable to an SMC connector, connect this cable onto the insert PCB signal line. From this signal line, the sample header is reached through gold plated pin connectors and another solder join. The return path to the second port in the test set is the reverse of this. It is clear therefore that a considerable number of sources of attenuation for microwave or very high frequency signals exist in this signal path.

The results of these measurements are shown in figure 6.19. These results are

measurements of the two port parameter S_{21}. The measurements are presented in a power ratio as decibels referred to 1 mW power. It is clear from the measurements that the frequency response of the cable collapses at around 7 GHz. The rate of collapse appears to be around 120dBm per decade. Provided this rate of collapse continues it is safe to say that no extra filtering for terahertz frequency noise is required in this circuit arrangement[12].

6.3.11 SYSTEM OPERATIONAL CHARACTERISTICS

During circuit set-up and trimming, a noise and leakage test was carried out. The leakage test provided information on the quality of the circuit set-up while the noise test provided information on the precision, arrangement and functionality of the circuit. For example, both broken cabling and/or shields or a damaged operational amplifier would increase the circuit output noise level. Similarly, poor set-up and trimming would result in feeding of the bias voltage into the current signal (as discussed in section 6.3.6). This would produce an apparent increase in the leakage current within the circuit. The noise tests were taken from the peak to peak signal at a given sample bias voltage as plotted on the chart recorder, with maximum system gain and an open device circuit. The leakage measurement was performed at the same time (again with the device connections open circuited).

The results of the noise tests (on the pre-amplifier R_f = 75kΩ configuration) showed that a good circuit set-up could achieve some 2.25mV peak to peak noise output, or a current noise of 1.5pA. Typical set-ups provided 4.5 to 6 mV peak to peak noise output (or 3 to 4pA). This translates into a typical current noise of around $0.87pAHz^{-1/2}$ in 21Hz bandwidth. These measurements, being taken at the same time as the leakage test, were taken as the worst case noise on the sweep of sample bias voltage. In terms of leakage, the maximum measurable resistance was used as a gauge of the quality of circuit set-up. This was generally in the 100s of GΩ range. Typical circuit set-ups provided a resistance trace of between 100GΩ and 300GΩ (leakage currents of $3.3pAV^{-1}$ to $10pAV^{-1}$) while the maximum achieved was 1178GΩ, or a leakage of $850fAV^{-1}$. As the sample bias voltages are generally in the millivolt range, expected current leakage due to set up errors is in the order of $3.3fAmV^{-1}$ to $10fAmV^{-1}$.

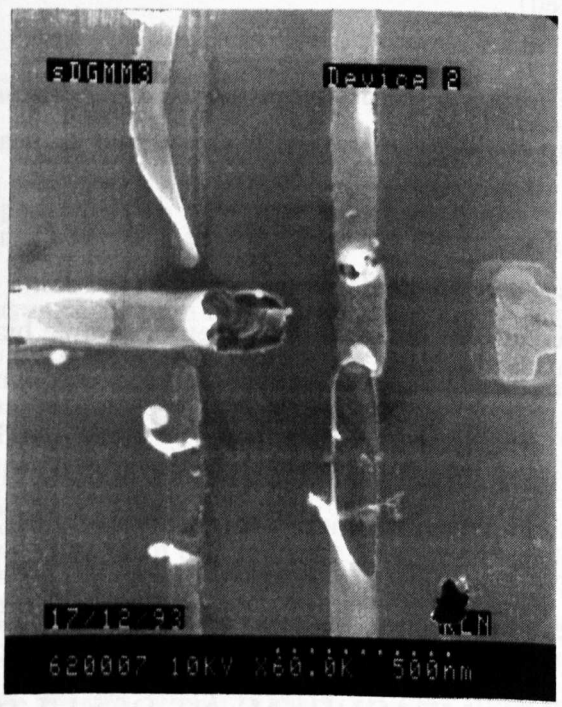
6.4 DEVICE MEASUREMENT RESULTS

In total, fifteen samples were fabricated and tested in this work. The samples were of three types; suspended mask samples, granular SET electrometer samples and granular single tunnel junction samples. Measurements were taken using the chart recorder to

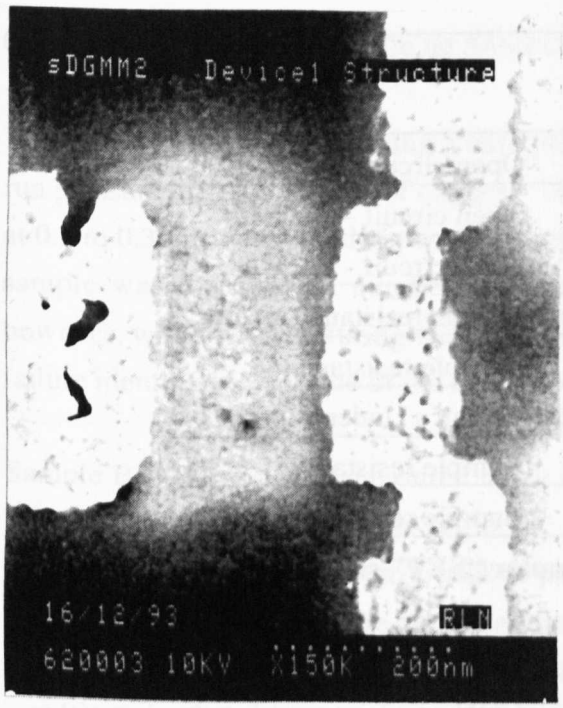
record the *I(V)* curve of the sample. Measurements of the sample current behaviour under action of a swept gate voltage were also taken (at 4.2K) but these did not yield visible single electron behaviour. Sample resistance calculations were taken directly from the *I(V)* graphs from the chart recorder. When the circuit parameters and chart recorder channel gains were known, sample currents and voltages could be calculated to the relevant accuracy. The following sections detail the results obtained from each set of samples in turn.

6.4.1 SUSPENDED MASK SAMPLES

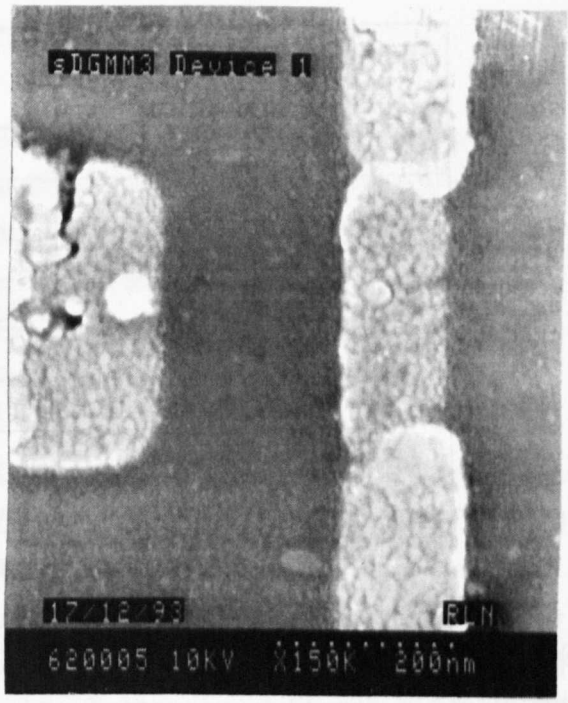
As discussed in section 6.2.2.1, suspended mask samples were prepared to form SET electrometers. Two devices were fabricated per sample according to the circuit of figure 6.1. All three of the samples prepared were found to be open circuit when tested. Subsequent analysis using an SEM showed that one sample did not form the overlap junctions while the other two samples failed due to ESD damage. Micrograph 6.1 illustrates the effect of ESD damage on these samples, while micrographs 6.2 and 6.3 illustrate the variation in overlap experienced between different samples. See also the discussion in chapter 4, section 4.2. Note that the two samples depicted in micrographs 6.2 and 6.3 were processed using the same process flow.



Micrograph 6.1 **Effect of ESD damage on suspended mask samples.**



Micrograph 6.2 Small overlap junction.



Micrograph 6.3 Large overlap junction.

Fabrication of this type of sample was discontinued in favour of the granular structures. This decision was taken due to the lack of control over the junction overlap area. The requirement to raise the operational temperature of the device, through reduction in the tunnel junction area and therefore the capacitance, meant that the granular structures provided a more promising option.

6.4.2 GRANULAR "SET" ELECTROMETERS

Granular tunnel junction structures were prepared to form the circuit depicted in figure 6.1. The pattern used was that of figure 6.2. In all, eight samples of this type were prepared, giving a total of 16 measurable SET electrometers. None of the samples displayed any single electron effects at 4.2K. The first three of these samples were all found to be open circuit. In each case, this was due to ESD damage of the sample. The fourth sample provided some resistance measurements before being damaged through ESD or electromigration. The fifth sample identified the source of the high current electromigration problem and provided useful resistance measurements. The sixth sample was again open circuit, due to under exposure of the pattern, while the seventh and eight samples yielded further resistance measurements. Table 6.11 summarizes the status of these samples.

Sample No.	Sample ID	Results
1	D03D	Open circuit - ESD
2	E01	Open circuit - ESD
3	F17C	Open circuit - ESD
4	F17D	Some sample resistance results
5	B3AB	Sample resistance
6	B2AC	Open circuit - Under exposed
7	T01AA	Sample resistance
8	T01AB	Sample resistance

Table 6.11 Summary of granular "SET" sample results.

All samples were patterned using the techniques described in chapters 3 and 4. The aluminium film deposition was carried out in the UHV evaporator, using the deposition conditions noted for each sample. After resist lift-off, the samples were bonded onto the sample header using silver loaded paint and electrical connections were made to the header through pressing indium beads into the contact pads and the header wiring. The following sections detail the samples and the results obtained.

6.4.2.1 RESULTS FROM SAMPLE F17D

This sample was deposited with a base pressure of 2.2×10^{-8} mbar, a run pressure of 6.6×10^{-6}mbar of oxygen. The evaporation rate was maintained at $0.2 nms^{-1}$, to a total film thickness of 20nm. Post evaporation oxidation was carried out at 1 bar pressure of oxygen for 1 hour. The sample was measured using R_J in the pre-amplifier circuit set to 750kΩ. The voltage bias used on this sample was from 100μV to 10mV peak to peak at frequencies of 1 and 2mHz.

At room temperature, the sample resistance was found to be 1.58MΩ ±44kΩ. When measured at 4.2K, the resistance was 1.60MΩ ±45kΩ. This sample showed no signs of offset *I(V)* behaviour in the *I(V)* trace. Table 6.12 summarizes this result.

Room Temperature resistance	4.2K resistance
1.58MΩ ±44kΩ	1.60MΩ ±45kΩ

Table 6.12 Sample F17D results summary.

6.4.2.2 RESULTS FROM SAMPLE B3AB

This sample was deposited using a base pressure of 2.3 x 10^{-8}mbar, and an evaporation run pressure of 1 x 10^{-6}mbar of oxygen. The aluminium deposition rate was maintained at 0.2 to 0.3nms^{-1} to a total deposition thickness of 27nm. Device 1 (B3AB1) from this sample was measured a number of times, under differing conditions. Device 2, however, was damaged during the initial room temperature measurement set-up. This failure identified the source of the ESD problems as discussed in section 6.3.6.2.

Sample B3AB1 was first measured using R_1 in the pre-amplifier circuit set to 750kΩ. The EG&G amplifier gain was set to 10. *I(V)* sweeps were taken with sample circuit bias voltages of 100μV to 10mV peak to peak at frequencies between 1mHz and 4mHz. Using this arrangement, the room temperature sample resistance was found to lie between 81kΩ and 88kΩ ±15kΩ, while the 4.2K sample resistance was between 72kΩ and 84kΩ ±15kΩ. Again, no signs of single electron behaviour were observed, either break-up of the *I(V)* curve or modulation of the source-drain current with gate voltage.

R_1	Room Temperature resistance			4.2K resistance		
	Upper	Lower	Accuracy	Upper	Lower	Accuracy
750kΩ	88kΩ	81kΩ	±15kΩ	84kΩ	72kΩ	±15kΩ
75kΩ	37.5kΩ	---	±1.5kΩ	1.02MΩ	---	±28kΩ
75kΩ	52.7kΩ	---	±1.5kΩ	---	---	---
75kΩ†	133kΩ†	---	±3.7kΩ	---	---	---
75kΩ†	207kΩ†	---	±5.8kΩ	---	---	---

† - these represent the most common resistance states.

Table 6.13 Sample B3AB (device 1) results summary.

Due to the low sample resistance observed, R_1 in the pre-amplifier circuit was changed to 75kΩ. This improved the resistance measurement resolution at these "low" resistances. However, as noted in section 6.3.7.1, the current noise was increased in this system. A differential divide by 21 voltage divider was added into the circuit between the EG&G instrumentation amplifier and the current signal outputs of the driver circuit. This allowed measurement of greater current ranges, as experienced on this sample. After the first measurement, the sample was removed from the helium dewar. Subsequent measurement of the device at room temperature found resistance state switching behaviour in the *I(V)* trace. This trace exhibited four clear resistance states of the device, these were: 37.5kΩ ±1.5kΩ, 52.7kΩ ±1.5kΩ, 133kΩ ±3.7kΩ and 207kΩ ±5.8kΩ. The two highest resistance value states were the most common and stable of

these states from the *I(V)* curve. Finally, the 4.2K measurement was repeated, yielding a sample resistance of 1.02MΩ ±28kΩ.

6.4.2.3 RESULTS FROM SAMPLE T01AA & T01AB

Both these samples were fabricated on the same substrate and the aluminium film was deposited in the same vacuum cycle. The evaporation conditions used were: base pressure = 1 x 10^{-8}mbar, with a run pressure of ≈ 1.4 x 10^{-6}mbar. Oxygen ambient was not used during this evaporation due to an oxygen line blockage. The oxygen partial pressure was therefore uncertain in this deposition. The evaporation rate was 0.2nms^{-1}, maintained to a total film thickness of 25nm. Post evaporation oxidation was carried out at 1 bar pressure of oxygen for 1 hour.

Both samples were measured using the R_I = 75kΩ pre-amplifier circuit, and in both cases, the EG&G amplifier was set to a gain of 10, with a differential divide by 21 voltage divider fitted to its input. Again, sample circuit bias voltages of 160μV_{p-p} to 2mV_{p-p} at 1mHz to 5mHz were used to sample the device *I(V)* curve. The results for each device on both these samples are given in table 6.14.

Sample	Room Temperature resistance			4.2K resistance		
	Upper	Lower	Accuracy	Upper	Lower	Accuracy
T01AA1	7.2kΩ	—	±1.5kΩ	8.2kΩ	—	±1.5kΩ
T01AA2	8.5kΩ	—	±1.5kΩ	7.5kΩ	6.6kΩ	±1.5kΩ
T01AB1	open	circuit	—	open	circuit	—
T01AB2	8.7kΩ	7.0kΩ	±1.5kΩ	9.2kΩ	7.6kΩ	±1.5kΩ

Table 6.14 Samples T01AA and T01AB results summary.

6.4.3 GRANULAR SINGLE TUNNEL JUNCTIONS

These samples were prepared in order to test the resistance of a single junction. In all, four samples were prepared, each containing one single junction pattern. All samples had the films deposited in the same vacuum cycle. The film deposition conditions were as follows; base pressure 2.3 x 10^{-8}mbar, with a run pressure of ≈ 2 to 2.3 x 10^{-6}mbar of oxygen. The evaporation rate was 0.1 to 0.2nms^{-1}, to a total film thickness of 20nm. Post evaporation oxidation was carried out at 1 bar pressure of oxygen for 1 hour.

Sample measurements used the standard circuit set-up with the current pre-amplifier R_I = 75kΩ. The EG&G amplifier gain was set to 10 and as before a divide by 21

circuit was used on the EG&G amplifier input. Again structure repeatability and ESD were problems with these measurements and only one sample yielded results. Table 6.15 summarises the results from these samples.

Sample ID	Room temperature Resistance	4.2K Resistance
T04CA	50GΩ - open circuit, ESD damage	
T04CC	534kΩ ±15kΩ	688kΩ ±19kΩ
T04DA	170GΩ - open circuit, pattern incorrectly formed	
T04DC	open circuit, cause unknown	

Table 6.15 Granular single tunnel junction results.

6.5 CONCLUSIONS

The conclusions to this chapter divide into two sections. Those relating to the circuits and characterization work and those relating to the device measurements. Finally, the questions posed at the beginning of this chapter will be addressed.

The circuits developed for this work represent a low noise measurement system for single electron devices. These circuits do not require the somewhat difficult gain matching of some other single electron measurement systems[1], however they do require set-up and offset trimming. The pre-amplifier circuit provides the majority of the gain in the system, allowing measurement of extremely low current signals. The flexibility of these circuits, however, is limited by the design of the pre-amplifier for measurement of high resistance samples. The low sample resistances experienced in this work resulted in reduced accuracy of the resistance measurement. The noise sources of these circuits have been well characterized and should not add significant noise into the device circuit, other than that generated by the pre-amplifier. The pre-amplifier is therefore the noise limiting portion of the circuit design. It appears, however, that the accuracy of measurement may have been insufficient to get the greatest extent of information out of the successful device measurements.

Despite the low number of devices measured, a few conclusions can be reached as regards the devices fabricated in this work. Firstly, the resistance of the fabricated structures must be due to a tunnelling resistance. This is the case even in the low resistance samples, T01AA and T01AB. Given the results of the film characterization work in chapter 5, the high value of resistance through the structure is not explicable either through the film resistivity and structure dimensions or by any know mechanism

of resistance increase such as weak localization. Also, in all samples measured, the sample resistance is temperature invariant within the measurement resolution. This backs up the argument of a tunnelling resistance being formed in the structures. More conclusive proof of the tunnelling nature of the contacts may have been provided by high voltage sweeps of the *I(V)* curve. This would have shown a current exponentially dependent on bias voltage. However, the divide by 50 input stage to the sample pre-amplifier coupled with the high trans-impedance gain prohibited this test of the devices.

Secondly, the structure resistance is very dependent on feature size and deposition conditions. Within a single lithography/deposition cycle, the device resistance appears to be repeatable, as the results for T01AA and T01AB demonstrate. However, the resistance appears less repeatable between fabrication cycles. Compare, for example the single junction sample T04CC and the double junction sample B3AB1 which were deposited under almost identical deposition conditions. In these two samples, resistances differ by more than an order of magnitude. It is questionable, however, whether the second measurement of B3AB1 is valid as the sample had been immersed in liquid helium, then exposed to atmosphere. Thermal cycling and subsequent water condensation onto the sample must affect the material structure and as such the device resistance. Discounting this second measurement, the difference still stands. The differences are two fold, firstly T04CC was 20nm thick while B3AB1 was 27nm thick. This in itself cannot explain the difference. The differences in evaporation O_2 partial pressures is then the only other known difference in that T04CC was deposited in 2 to 2.3×10^{-6}mbar of oxygen while B3AB1 was deposited in 1×10^{-6}mbar of oxygen.

Thirdly, the spread of device resistances may well be prohibitive to single electron device fabrication if this cannot be controlled through some fabrication parameter. Whether this relates to the lithographic structure size, or to the deposition conditions is uncertain. However, the lithographic size and feature shape may have a bearing on the device formation as this may be another source of difference between samples T04CC and B3AB. Discussion of this subject is continued in section 7.2.

Returning to the questions posed at the start of this chapter, in section 6.1, we find we have answers to both questions. Firstly, the small geometry regions do indeed behave differently from the bulk material. This is clear from the high resistance of the structures. It is likely that the structures fabricated contain tunnel junctions as this would explain the high sample resistance and the temperature invariance of resistance. However, the answer to the second question is less clear. Clearly, the structures used in this work do not produce high temperature single electron devices. That is obvious from the absence of single electron effects in the devices measured. However, what is

not clear is the cause of the absence of these effects. It is unlikely that the effects would be "washed out" by noise coupled into or generated by the measurement system. The characterization work on the measurement circuits precludes this conclusion. The resulting conclusion is then that the structure capacitance is not sufficiently low to allow the manifestation of single electron effects at 4.2K. This conclusion will be discussed further in section 7.3.

6.6 REFERENCES

1. P. Delsing, T. Claeson, K.K. Likharev & L.S. Kuzmin, *Phys. Rev. B.*, **42**(12),7439 (1990).

2. L.J. Geerligs, V.F. Anderegg, P.A.M. Holweg, H. Pothier, D. Esteve, C. Urbina, M.H. Devoret & J.E. Mooij, *Phys. Rev. Lett.*, **64**(22),2691 (1990).

3. K.K. Likharev, N.S. Bakhvalov, G.S. Kazacha, & S.I. Serdyukova, *IEEE. Trans. Magn.* **25**(2),1436 (1989).

4. A.N. Cleland, J.M. Schmidt & J. Clarke, *Phys. Rev. B.* **45**(6),2950 (1992).

5. P. Lafarge, H. Pothier, E.R. Wiliams, D. Esteve, C. Urbina & M.H. Devoret, *Z. Phys. B.* **85**(3),327 (1991).

6. P. Brokaw, *An I.C. amplifier users' guide to decoupling, grounding and making things go right for a change*, Analog Devices - Linear I.C. Databook, Vol. 1., Application Notes, Ch. 20, pp13-20, 1984.
   and A. Rich, *Understanding interference type noise*, Analog Devices - Linear I.C. Databook, Vol. 1., Application Notes, Ch. 20, pp81-90, 1984.

7. P.J. Fish, *Electronic noise and low noise design*, MacMillan Press, London, 1993.

8. M.E. van Valkenberg, *Analog filter design*, Holt Saunders, 1982.

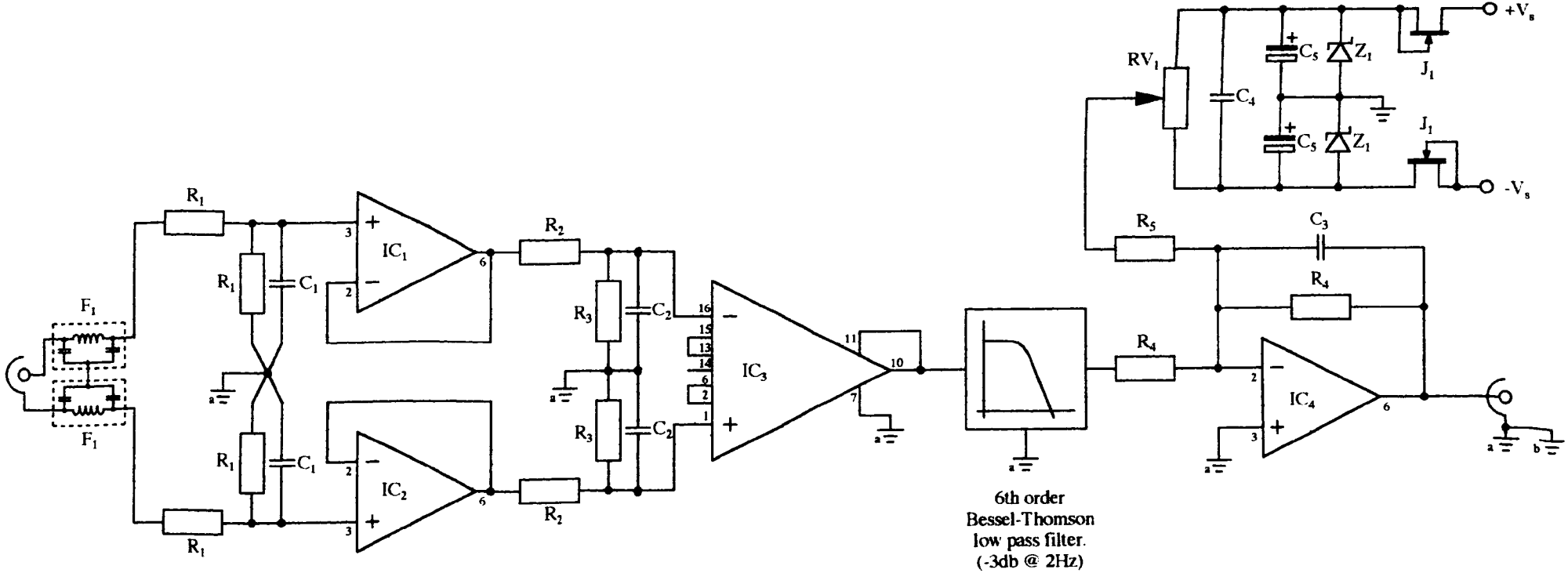
9. P.R. Sallen & E.L. Key, *IRE Trans. Circuit Theory*, **CT-2**(3),74 (1955).

10. P. Horowitz & W. Hill, *The art of electronics*, 2nd Ed., Cambridge University Press, 1990.

11. M.E. van Valkenberg, *Analog filter design*, Ch. 10, table 10.5, Holt Saunders, 1982.

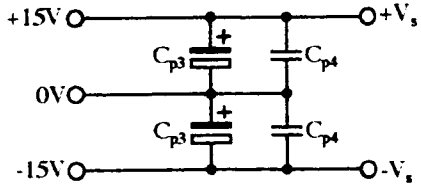
12. D. Vion, P.F. Orfila, P. Joyez, D. Esteve & M.H. Dévoret, *J. Appl. Phys.*, **77**,2519 (1995).

ORIGINAL COPY TIGHTLY BOUND

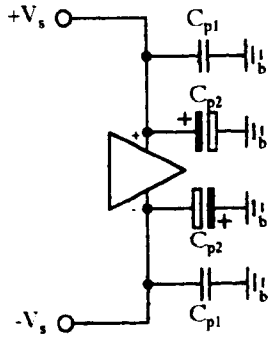


COMPONENTS

IC_1 & IC_2	OPA 627 AP	Burr Brown.	C_1	33pF	(33p ±2.5% polystyrene)
IC_3	INA 103 KP	Burr Brown.	C_2	4.7nF	(4n7 ±2.5% polystyrene)
IC_4	OPA 27 AJ	Burr Brown.	C_3	3.3nF	(3n3 ±2.5% polystyrene)
			C_4	100nF disc ceramic.	
F_1	5500pF π-section rfi. bulkhead filter.		C_5	47μF electrolytic.	
			C_{p1}	100nF sub-miniature ceramic.	
R_1	75kΩ	(75k ±1%)	C_{p2}	10μF sub-miniature electrolytic.	
R_2	2kΩ	(2k ±1%)	C_{p3}	100μF electrolytic.	
R_3	4.7kΩ	(4k7 ±1%)	C_{p4}	100nF disc ceramic.	
R_4	5.1kΩ	(5k1 ±1%)			
R_5	150kΩ	(150k ±1%)	J_1	J511	(4.7mA const. current diode)
			Z_1	5.1V Zener diode.	(BZX55C 5V1)
RV_1	10kΩ	(10k 10T 9236 64Y series trim pot.)			



Power supply arrangement.



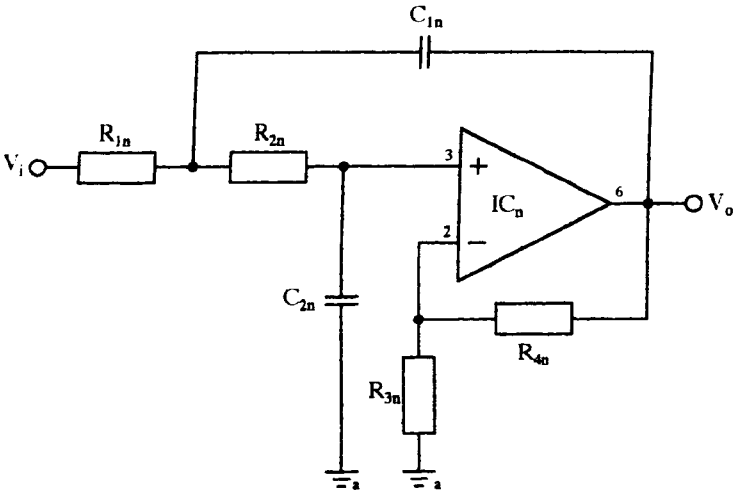
Supply connections to amplifiers.

Notes:
1. All power supplies are decoupled through ceramic and electrolytic capacitors, C_{p1} & C_{p2}.
2. Supply and signal grounds are separate and connected at only one point.
   - a signal ground.
   - b supply ground.
3. Resistors are type MRS25 metal film unless stated otherwise.

APPENDIX 6.1
Input Signal Filter Circuit

Active 6th order Bessel-Thomson filter.

Implementation:- 3 x cascaded 2nd order Sallen & Key circuits.
Design requirement: -3dB at 2Hz.



Low Pass Sallen & Key filter section.

COMPONENTS

IC_n	OPA 111 AM	Burr Brown.
C_{p1}	100nF sub-miniature ceramic.	
C_{p2}	10μF sub-miniature electrolytic.	
C_{p3}	100μF electrolytic.	
C_{p4}	100nF disc ceramic.	

Other resistor and capacitor values are detailed in the final table shown opposite.

Resistors: precision metal film. ±0.1% tol.
Capacitors: brass cased axial polycarbonate. ±5% tol.

Design parameters.

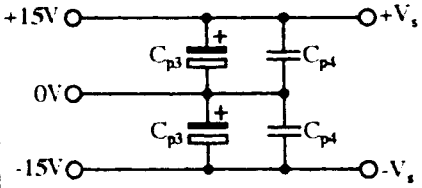
Filter stage n	ω_{cn} /rads^{-1}	Q_n	$K_n = (3-1/Q_n)$
1	20.19	0.510	1.039
2	21.26	0.611	1.363
3	23.98	1.023	2.022

Required circuit values.

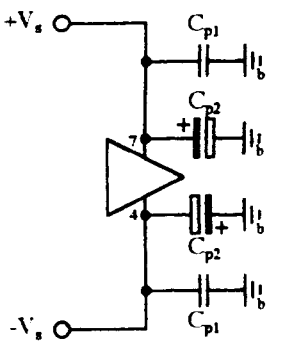
Filter stage n	$C_{1,2n}$ /μF	$R_{1,2,3n}$ /kΩ	$R_{4n}=(K_n-1)R_{3n}$ /kΩ	f_{cn} /Hz
1	1	49.53	1.932	3.21
2	1	47.04	17.08	3.38
3	1	41.70	42.62	3.81

Implementation.

Filter stage n	$C_{1,2n}$ /μF	$R_{1,2,3n}$ /Ω	R_{4n} /Ω
1	1	48k7 + 866R	1k96
2	1	46k4 + 649R	16k9 + 182R
3	1	41k2 + 523R	42k2 + 442R



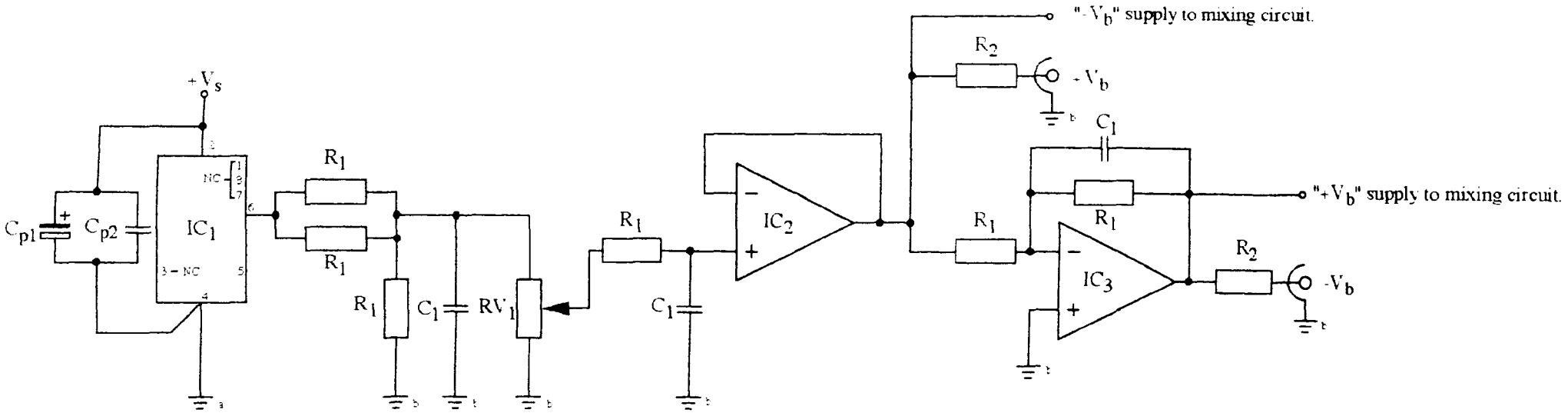
Power supply arrangement.



Supply connections to Op-amps.

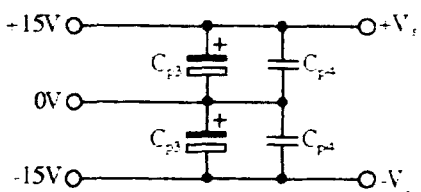
Notes:
1. All power supplies are decoupled through ceramic and electrolytic capacitors, C_{p1} & C_{p2}.
2. Supply and signal grounds are separate and connected at only one point.
   - a signal ground.
   - b supply ground.

APPENDIX 6.2
Bessel Thomson Response Filter Design

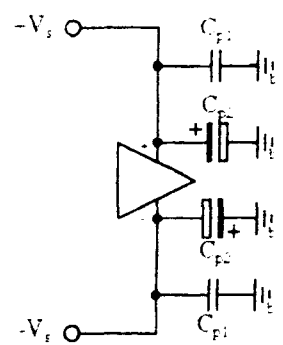


COMPONENTS

IC_1	REF 10 KM	Burr Brown
IC_2 & IC_3	OPA 111AM	Burr Brown
RV_1	3540s Series 10kΩ ± 5% 10T pot. Linearity ±0.25%	
R_1	10kΩ (10k ± 1%)	
R_2	50Ω (51R ± 1%)	
C_1	0.1μF (0.1μF ± 5% Polyester)	
C_{p1}	100nF sub-miniature ceramic.	
C_{p2}	10μF sub-miniature electrolytic.	
C_{p3}	100μF electrolytic.	
C_{p4}	100nF disc ceramic.	



Power supply arrangement.

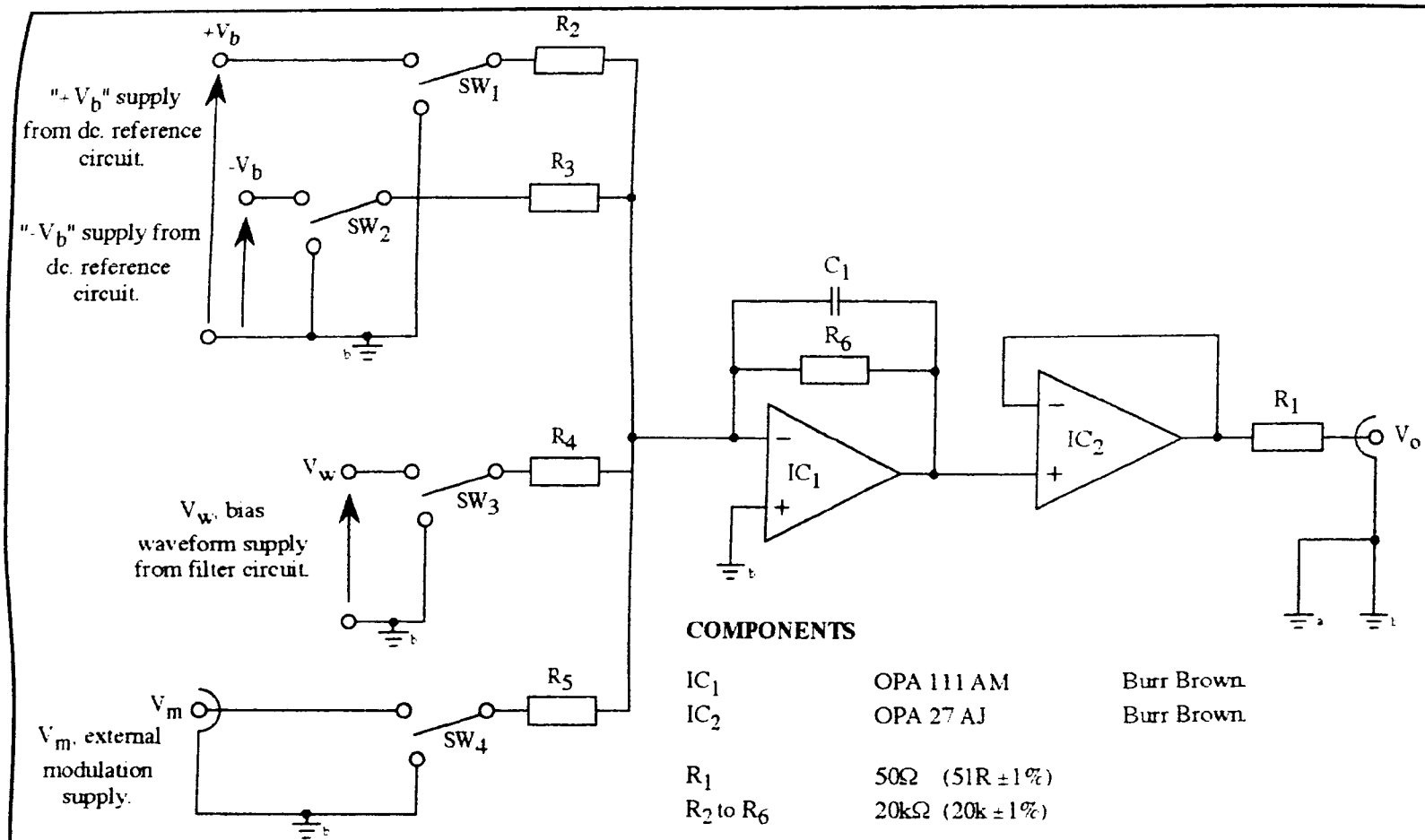


Supply connections to amplifiers.

Notes:
1. All power supplies are decoupled through ceramic and electrolytic capacitors, C_{p1} & C_{p2}.
2. Supply and signal grounds are separate and connected at only one point.
   a signal ground.
   b supply ground.
3. Resistors are type MRS25 metal film unless stated otherwise.

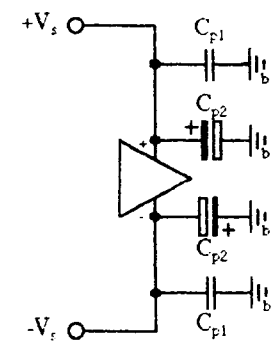
APPENDIX 6.3

Variable DC Reference Voltage Circuit

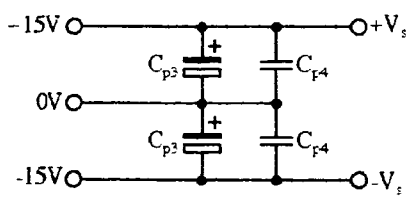


COMPONENTS

IC_1	OPA 111 AM	Burr Brown
IC_2	OPA 27 AJ	Burr Brown
R_1	50Ω (51R ±1%)	
R_2 to R_6	20kΩ (20k ±1%)	
C_1	680pF (680pF ±2.5% Polystyrene)	
C_{p1}	100nF sub-miniature ceramic.	
C_{p2}	10μF sub-miniature electrolyic.	
C_{p3}	100μF electrolytic.	
C_{p4}	100nF disc ceramic.	
SW_1 to SW_4	Single Pole, 2-Way Gold contact switches	



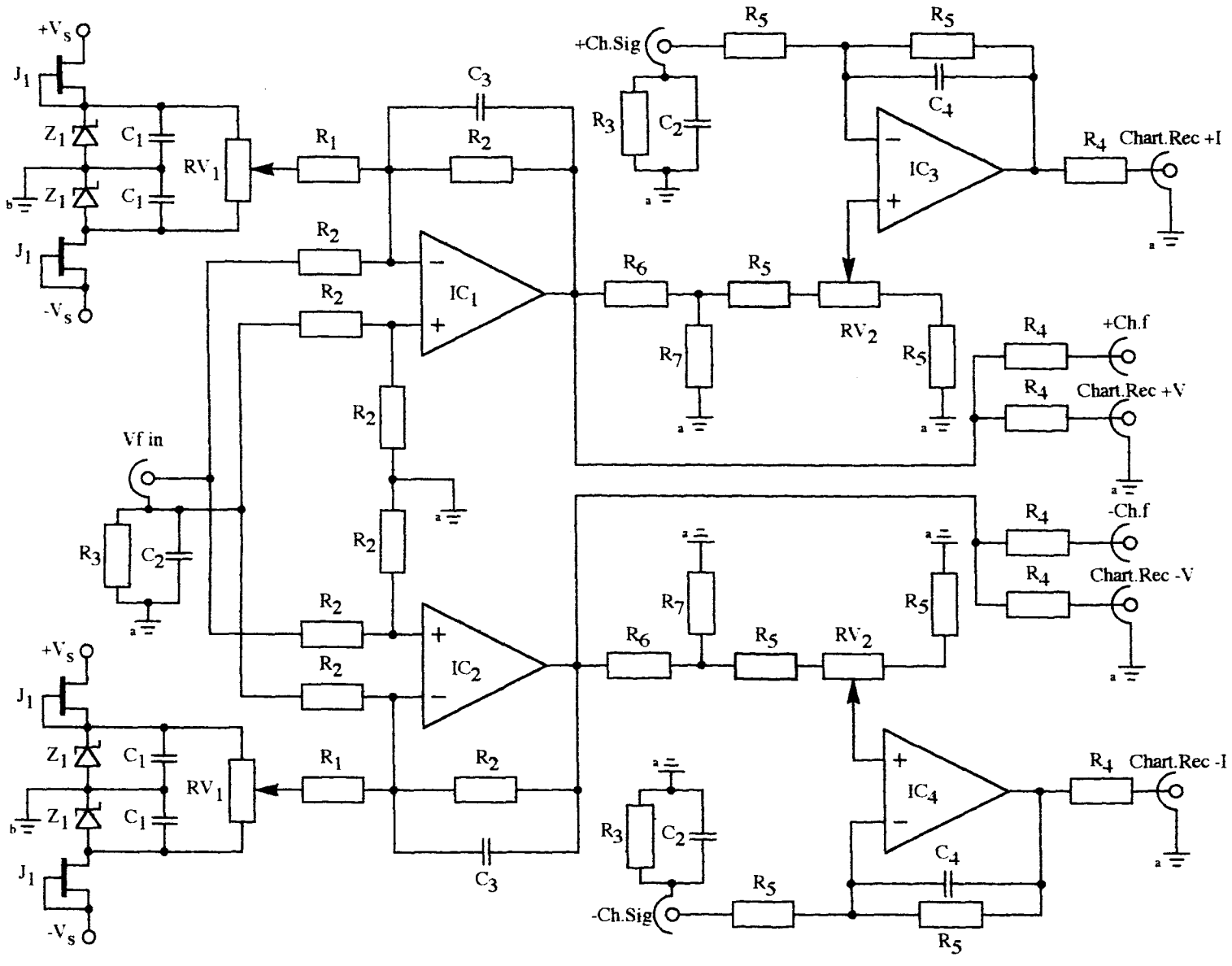
Supply connections to amplifiers.

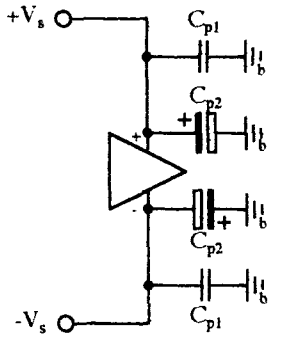


Power supply arrangement.

Notes: 1. All power supplies are decoupled through ceramic and electrolytic capacitors, C_{p1} & C_{p2}.
2. Supply and signal grounds are separate and connected at only one point.
   a signal ground.
   b supply ground.
3. Resistors are type MRS25 metal film unless stated otherwise.

APPENDIX 6.4
Output Mixing Circuit





Supply connections to amplifiers.

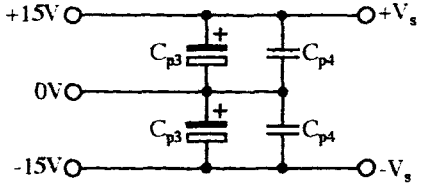
COMPONENTS

C_1 & C_2	100nF Ceramic.
C_3	68pF Polystyrene ±2.5%
C_4	220pF Polystyrene ±2.5%
J_1	J511 (4.7mA Const Current Diode)
IC_1 & IC_2	LF 411N
IC_3 & IC_4	OP 27 GN
R_1	1MΩ (1M ±1%)
R_2	10kΩ (10k ±1%)
R_3	100Ω (100R ±1%)
R_4	50Ω (51R ±1%)
R_5	5.1kΩ (5k1 ±1%)
R_6	10kΩ (10k ±1%)
R_7	100Ω (100R ±1%)
RV_1	9236 64Y Series 10kΩ 10T
RV_2	9236 64Y Series 10kΩ 10T
Z_1	BZX55 C 5V1

NOTES:
Power supply and signal grounds are separate, and connected at only one point.
- a Power supply ground.
- b Signal ground.

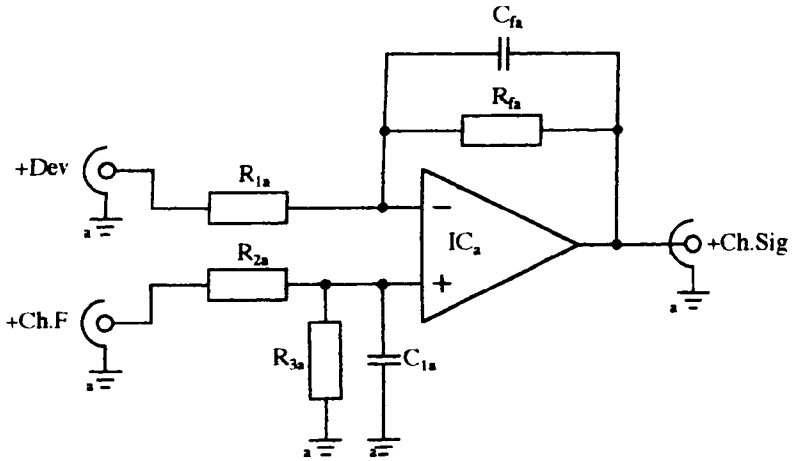
COMPONENTS

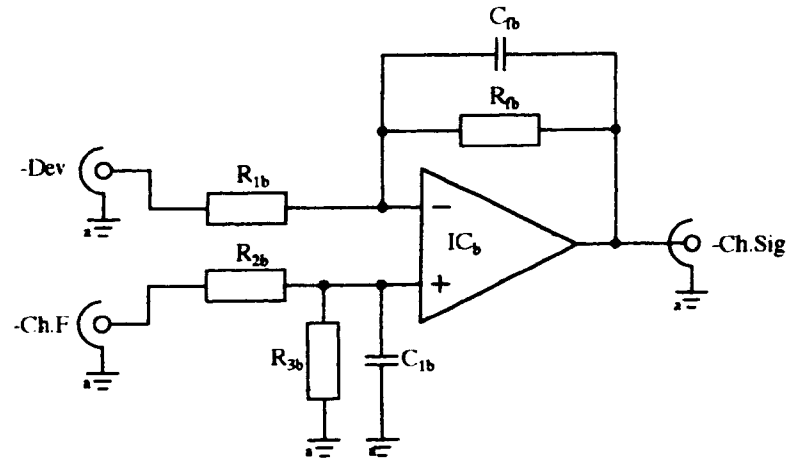
C_{p1}	100nF sub-miniature ceramic.
C_{p2}	10μF sub-miniature electrolyic.
C_{p3}	100μF electrolytic.
C_{p4}	100nF disc ceramic.

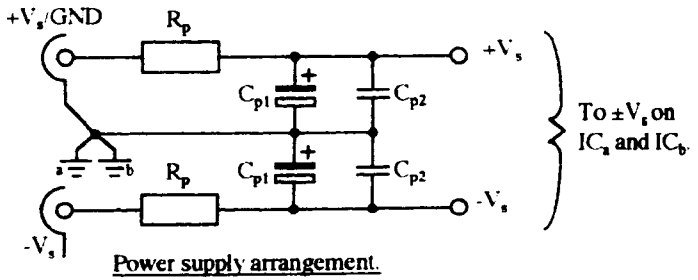


Power supply arrangement.

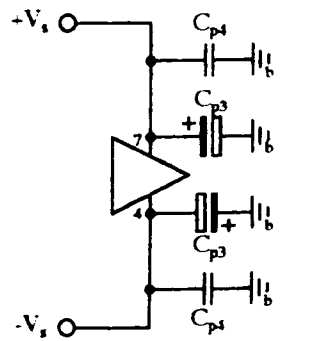
APPENDIX 6.5
Pre-Amplifier Driving Circuit







Power supply arrangement.



Supply connections to Op-amps.

COMPONENTS

IC_a & IC_b	OPA128LM Burr Brown	
R_{fa} & R_{fb}	750MΩ	(750M ±1%.)
R_{1a} & R_{1b}	75kΩ	(75k ±1%)
R_{2a} & R_{2b}	4.7kΩ	(4k7 ±1%)
R_{3a} & R_{3b}	47Ω	(47R ±1%)
R_p	51Ω	(51R ±1%)
C_{fa} & C_{fb}	10pF	(10p ±2.5% Polystyrene.)
C_{1a} & C_{1b}	10nF	(10n ±2.5% Polystyrene.)
plus paralleled	1μF	(1μF ±5% Polyester.)
C_{p1}	100μF	(100μ 63V Electrolytic.)
C_{p2}	100nF	(100n ±20% Disc Ceramic.)
C_{p3}	10μF	(10μ 25V Sub-miniature electrolytic.)
C_{p4}	10nF	(10n miniature disc ceramic.)

NOTES

Power supply and signal grounds are separate and are connected only at the insert case.

a Signal ground.

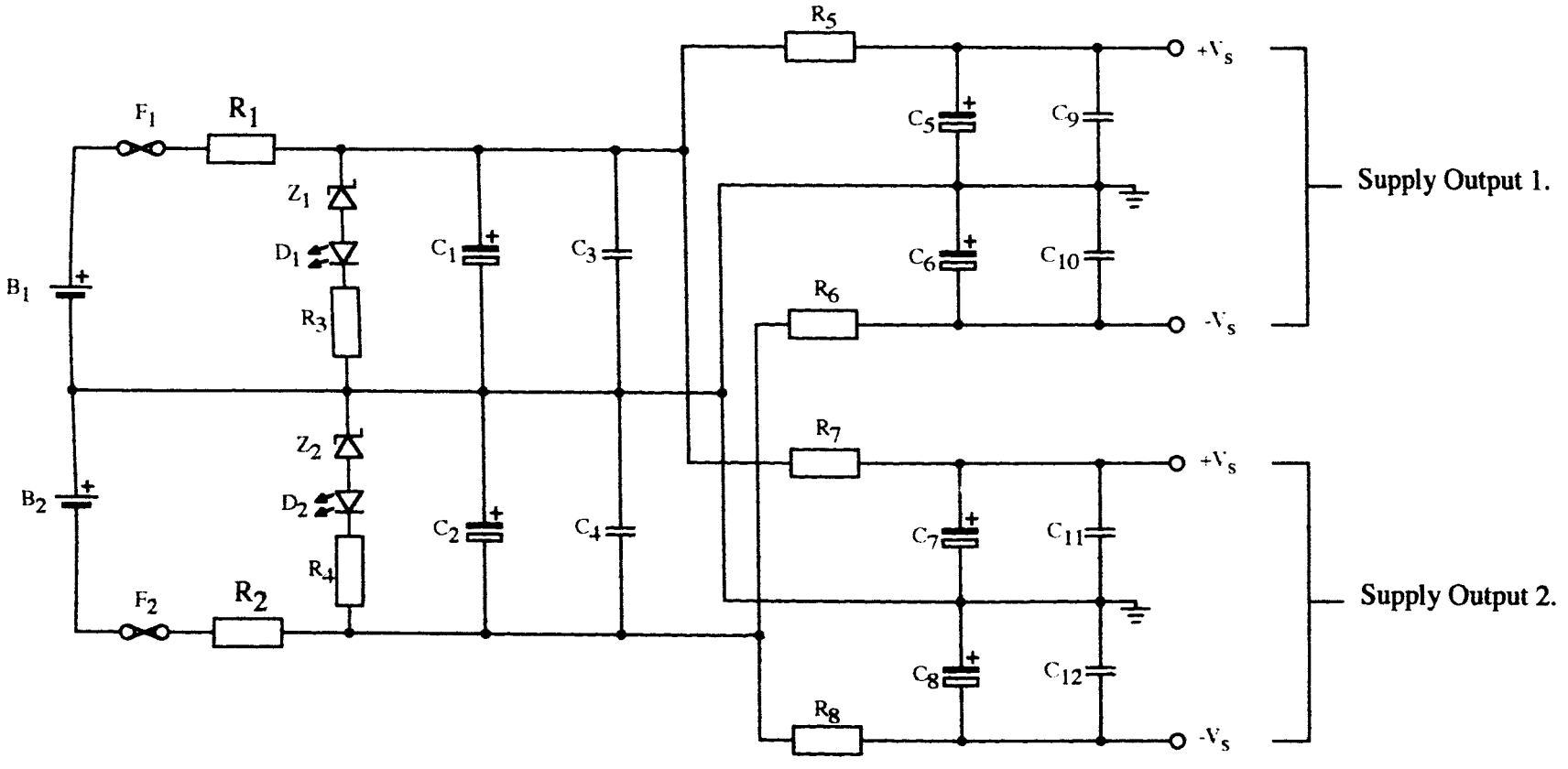
b Power supply ground.

Inputs to the OPA128 are ESD sensitive. Therefore, care should be exercised when handling.

The OPA128 input and output leads should be mounted directly onto PTFE stand-offs. The remaining components should also be mounted between stand-offs to ensure no leakage current paths. The PCB area directly beneath the input connections should be connected as a "guard plane". This plane should be connected to the non-inverting terminal.

APPENDIX 6.6

Current Measurement Pre-Amplifier

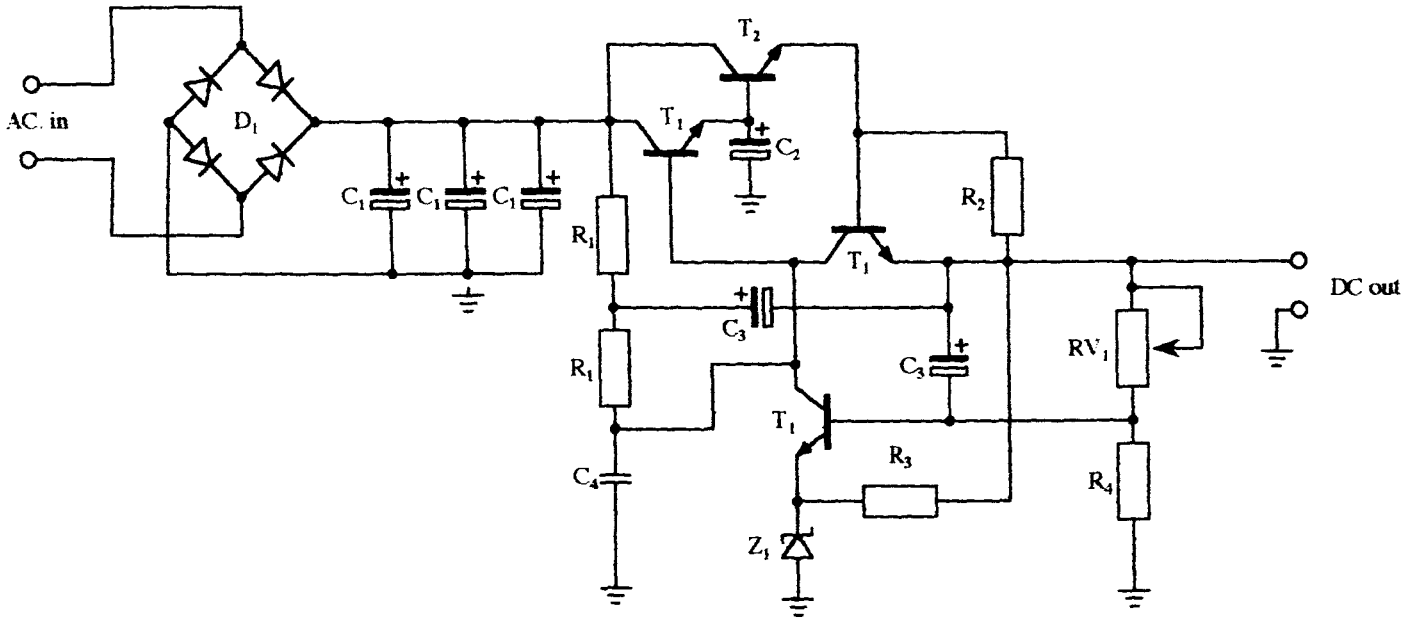


COMPONENTS

B_1 & B_2	12v Pb-Acid batteries (NP6 12v6. 6Ah)			
F_1 & F_2	200mA slow blow fuse.			
Z_1 & Z_2	BZX85C 9v1			
D_1 & D_2	CMD64530 (Green LED)			
			C_1 & C_2	100μF 63v Electrolytic.
R_1 & R_2	10Ω 6W		C_3 & C_4	100nF Disc Ceramic.
R_3 & R_4	160Ω nom.	(160R ±1%)	C_5 to C_8	10μF 63v Electrolytic.
R_5 to R_8	11Ω nom.	(11R ±1%)	C_9 to C_{11}	100nF Disc Ceramic.

APPENDIX 6.7

Battery Power Supply Conditioning Circuit



COMPONENTS

	Positive output	Negative Output
T_1	BC547	BC307
T_2	2N3055	2N3791
C_1	100μF electrolytic.	
C_2	10μF electrolytic.	
C_3	100μF electrolytic.	
C_4	100nF disc ceramic.	
R_1	1kΩ	
R_2	0.2Ω, 2W	
R_3	3.3kΩ	
R_4	4.7kΩ	
RV_1	10kΩ	
D_1	KBU4D Bridge rectifier	
Z_1	6.2V Zenner diode	

NOTES

Layout shown for positive DC output. For negative DC output, component changes are listed in the component list.

Transistor T_2 should be mounted with heatsink and off board.

Designed by: J. Hollander, S. Fairbairn.
Project No: ED1096.

APPENDIX 6.8

Low Noise Regulated Power Supply

CHAPTER SEVEN

Discussion

7.1 INTRODUCTION

Section 1.3 of Chapter 1 details the aims of the work presented in this thesis. These were to investigate the fabrication techniques for metal-insulator-metal tunnel junction based single electron systems. From this investigation, the development of a single electron tunnelling (SET) electrometer, operational at temperatures of 4.2K or above was the goal for demonstrating high temperature SET device operation. These aims were investigated predominantly through the technique of patterning granular thin films, as described in section 4.4.

A great deal of the initial work was spent investigating and characterizing the fabrication processes in order to achieve stable and reliable tunnel junction fabrication techniques. Of the three techniques investigated, the crossed track system failed to properly define tunnel junction structures. Of the other two techniques, the conventional suspended mask[1] did not adequately control junction overlap dimensions at small sizes. The devices fabricated also proved particularly sensitive to electro-static discharge (ESD). However, this was later discovered to be mainly an interconnect and measurement problem. The remaining technique of patterned granular thin films was therefore the most promising choice for raised temperature operation.

Although the project has failed with regard to achieving the main aim of a working SET

electrometer operational at 4.2K, a great deal of information has been gained with regard to the formation of the structures investigated. It is to these results that this chapter turns, in order to draw conclusions from the information available. To this end, this chapter divides into five parts. Firstly, the control over structure formation and device parameters is discussed in the light of the device measurement results. Secondly, the discussion is focused on the structure capacitance. Thirdly, possibilities for further development of the patterned granular film technique are explored. Fourthly, the possibilities for future single electron devices are discussed in the light of the preceding sections and current work. Finally, the chapter is concluded and recommendations are made for further work.

7.2 STRUCTURE FORMATION AND CONTROL

Control over fabricated device structures is essential in the development of any useful device technology. As the electron transport in single electron devices depends on tunnelling, the device operation is therefore heavily dependent on physical device parameters. We have seen in chapter 2 that a tunnel current is dependent on the transmission coefficient. This varies exponentially with barrier width, for any given barrier potential. If we consider a parallel plate model for the capacitance of a junction, we have a capacitance inversely dependent on plate separation or dielectric thickness, with a squared dependency on junction cross-sectional dimension. We therefore find that the junction parameters of tunnel resistance, R_T, and junction capacitance, C_J, are heavily dependent on physical dimensions of the device. To produce reliable and repeatable SET devices it is therefore essential to control the feature sizes used to form the junctions.

Chapter four describes the fabrication processes investigated in this work. In this chapter, the crossed track technique was rejected as the structure could not be reliably formed. This deficiency was apparent before any consideration was made of the control of junction dimensions. The suspended mask fabrication technique allows reliable formation of the junctions however, the junction overlap control was found to be very poor. As discussed in chapters four and six, the junction overlap was reasonably constant within a sample. However sample to sample variations were too great for a reliable SET device fabrication process. This then left the patterned granular thin film devices as the best candidate for repeatable high temperature devices.

The requirements of the patterning process for the patterned granular devices were therefore; to place grains forming the tunnel junction where they were required, to control the constriction dimensions to ensure repeatability of the junction R_T and C_J

values, and to provide repeatable device fabrication from sample to sample. The physical size of the patterned constriction is important in determining the tunnel junction capacitance and of secondary importance in determining the junction resistance. The exponential dependency of R_T on barrier width makes this parameter more sensitive to oxide barrier thickness variations and the physical separation of the grains rather than the cross-sectional area of the grain-grain contact. The patterned constriction dimension has an important part to play in determining any grain separation. If the constriction is very much smaller than the grain size, nucleation processes will result in large inter-grain spacings and, possibly, very small capacitances. However if the constriction is of the order of the mean grain size or slightly larger, nucleation kinetics will drive the grain formation and grain size, producing a single grain-grain contact in the constriction. The oxidation conditions will then determine the barrier width.

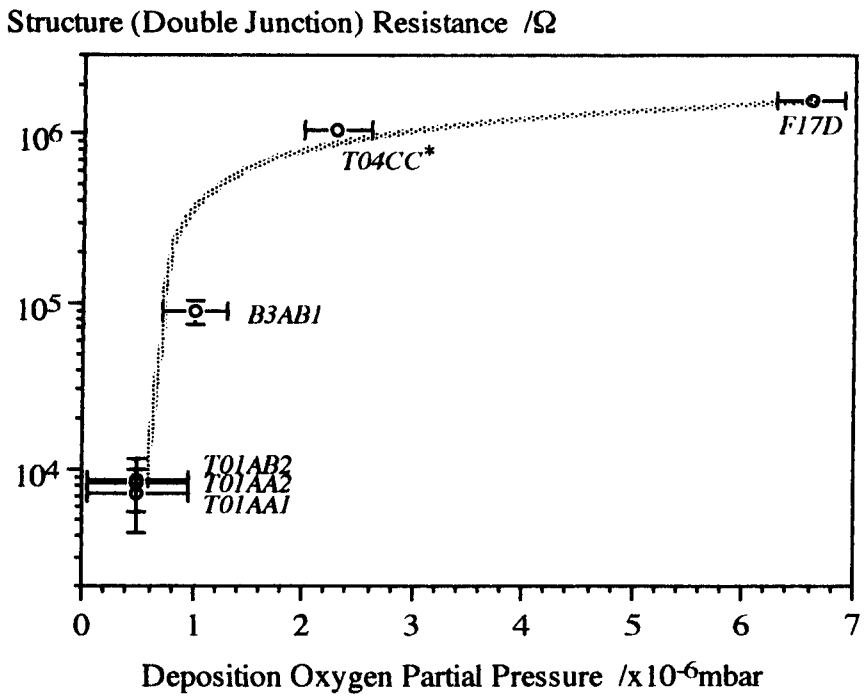
The results from chapter 4 on the characterization of the fabrication process appear promising for control of the constriction dimension. Figures 4.25 and 4.26 of section 4.4.3.4 show a ±5 to ±10nm spread on the mean constriction dimension of ≈30nm. This dimension is a little large compared with the mean grain size of around 20 to 25nm, but it provides single grains in the constriction region through grain nucleation effects and the resist close-over mechanism. In other words the deposition process is very unlikely to form more than one grain across the narrowest point of the pattern. From the pattern characterization results of chapter 4, the constriction regions maintain a width of one full grain until the pattern width is greater than around 50nm.

Control of the junction structure dimensions must be reflected in the electrical characteristics of the samples. The control would be exhibited both by variations within a sample and by sample to sample variations of the R_T and C_J values. No capacitance measurements were obtained from any of the devices as no single electron effects were observed. However, the resistance measurements demonstrate quite clearly the degree of process control. It appears that within a sample, or fabrication run, the repeatability of sample resistance is reasonable. Comparing the results from samples T01AA and T01AB we see consistent structure resistances. These samples were EBL written and developed together, with both of the samples having aluminium deposited in the same vacuum cycle. The mean of the structure resistances was 7.8kΩ with a range of ±1.3kΩ, a tolerance of 16%. The systematic error on the mean for these measurements was ±1.5kΩ.

Samples fabricated at different times show no such control over the structures, and therefore, junction, resistance. Review of the results of section 6.4.2 shows a sample

resistance spread over a range of 1.6MΩ through ≈100kΩ to ≈10kΩ. This is more than two orders of magnitude spread in what is believed to be a tunnel resistance. This demonstrates considerable variation in the junction characteristics where SEM investigations showed the junction region to be within the characterized process variation. Therefore, this variation must be through some change in the grain to grain spacing within the structure. This type of change will affect the sample resistance in a exponential manner by changing the tunnelling probability and also affect the junction capacitance through changes in the inter-grain capacitance. This capacitance is discussed more fully in section 7.3.

The fabrication scheme used here is intended to produce single grain-grain contacts. The oxidation conditions both during and after evaporation are intended to produce the tunnel barrier between the grains. This is expected to be predominantly through natural oxidation of the aluminium on the grain surface and at the grain boundaries, as discussed in section 4.4.3.3. This suggests that the oxidation conditions should be instrumental in the formation of the tunnel junction and in the control over the barrier width. Figure 7.1 graphs the results for the structure resistance versus the oxygen partial pressure used during the film deposition. The line included here is only a guide to a line of fit but it is displaying an $R_T \propto Log_{10}[P_p(O_2)]$ dependency.



*Sample T04CC was a measurement on a single junction The resistance has been multiplied by two to compare with the double junction samples. Sample identifications are given next to the data point.

Figure 7.1 Relationship between structure resistance and film deposition oxygen partial pressure.

Although the number of points presented in figure 7.1 is small, it can be tentatively suggested that the structure resistance depends logarithmically on the deposition oxygen partial pressure. The line shown in the graph is a logarithmic fit to the data, but is only presented as a guide to the dependency. From the results obtained from the devices, it is clear that the oxygen partial pressure during evaporation is essential in the formation of tunnel barriers. All the data presented in figure 7.1 are from samples which were exposed to 1 bar of oxygen after the deposition was complete. This may have an overall effect on the structure resistances, but the greatest effect clearly arises from the deposition oxygen pressure. We can conclude from these data that the barrier formation is dominated by oxidation of the grain surface during evaporation and not by post deposition oxidation of the grain boundaries.

From the results obtained, it is reasonable to assume that the structures are indeed tunnelling junctions. The basis for this derives from two sources. Firstly the temperature invariance of the sample resistance. In some samples a very small change in the resistance is observed between the room temperature and 4.2K measurements but this is mostly derived from set-up errors in the pre-amplifier. This temperature invariance is characteristic of a tunnel junction operating at low bias as discussed in section 2.2.2.4. The second basis is the magnitude of the resistance. We have seen from the characterization work of chapter 5 that the films used are very metallic. The resistivity has a worst case of $2 \times 10^{-7}\Omega m$. For structures of size approximately 30nm wide by 25nm thick, a resistance per unit length of $2.6 \times 10^{8}\Omega m^{-1}$ results from Ohm's law. Therefore, if we model the constriction as a wire of about 100nm long (a little longer than the actual structure length) we find we have a structure resistance of only 26Ω. Modifications through weak localization or interference effects would require some two orders of magnitude increase in this value to explain the observed resistance. This magnitude of correction is unlikely in these models. Also, as seen in section 5.2.2.2, equation 5.22, interference and weak localization effects are temperature dependent and the results here are temperature invariant.

Measurements of these structures below 1.2K, to examine the electronic density of states through superconducting tunnelling measurements would provide conclusive proof of tunnelling behaviour. These measurements were not feasible in the experimental set-up used in chapter 6 and use of other low temperature systems was limited through inability to contact to devices by methods other than indium bead bonding.

We can conclude from this discussion that the methods used to form the junction do not allow reliable control over the tunnel resistance, R_T, where low resistance junctions are

required. Recall that the junction resistance and capacitance define the junction time constant and frequency response. Also, the junction resistance must be greater than the resistance quantum to confine the electronic states. Keeping the resistance as low as possible is advantageous for the devices but does not meet with the process dependency on oxygen partial pressure during deposition. Where oxygen partial pressures are above $3x10^{-6}$mbar, the junction resistance appears to plateau out at around 600kΩ per junction. Therefore, provided junctions of this resistance are usable in single electron devices, the process is reasonably stable with these relatively high oxygen partial pressures. The variation in junction resistance, with reasonable control over the feature sizes indicates that the patterning dimensions may not be a major influence in the junction resistance. What has not been determined is the effect of dimensional and oxygen pressure variations on the structure capacitance. It is this parameter which will ultimately determines the useful temperatures of the devices.

It is clear form the degree of spread of the results that it is important to fully characterize the process window for these devices. However, the limited number of results obtained fail to allow firm conclusions to be made with regard to this process window.

7.3 STRUCTURE CAPACITANCE

We have so far discussed aspects of the process control relating to the measured resistance parameters of the structures; however, we still require to address the structure capacitance. In this situation, the junction capacitance should be determined by that of the inter-grain capacitance. As no single electron effects have been observed in any of the fabricated devices, we know that the characteristic island capacitance, C_Σ, is greater than 220aF (for $T_O = 4.2$K). In a double junction system, this characteristic capacitance is given by,

$$C_\Sigma = C_l + C_r + C_s = 2C_J + C_s \qquad [7.1]$$

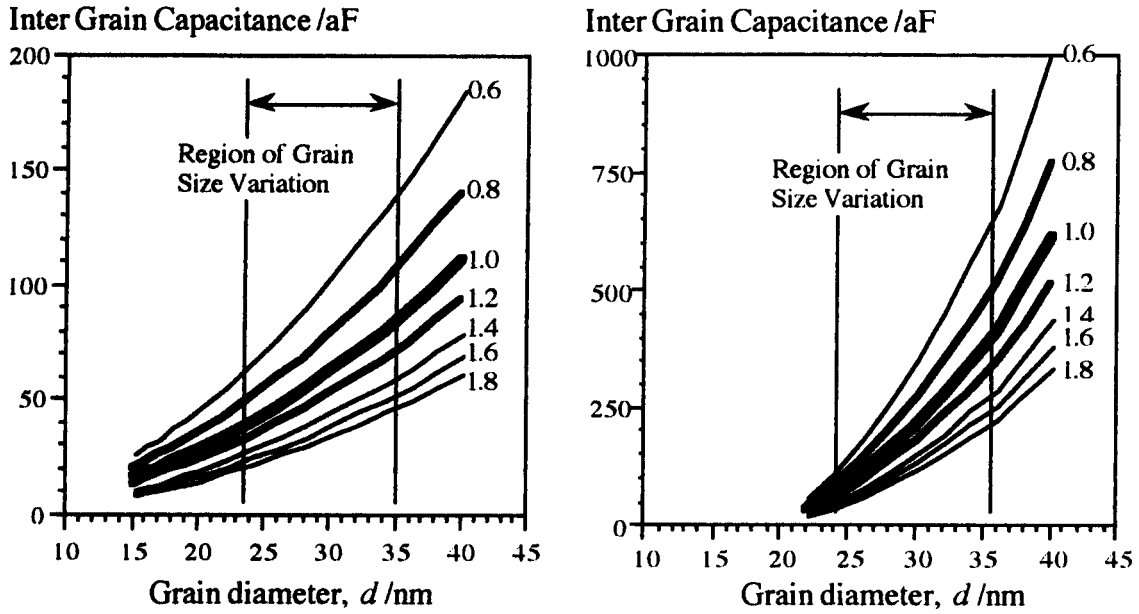
where C_l and C_r are the left and right hand junction capacitances respectively. In this case it is assumed that $C_l = C_r = C_J$. This gives a C_J of ≥110aF. In practice, C_J could be as low as 50 to 75aF for single electron effects to remain unobservable through washout by thermal fluctuations.

Some detailed work has been undertaken on analytical calculations of the structure capacitance of devices of this type[2]. This work compared a number of capacitance models to provide good estimates for the capacitances of these and similar single

electron device structures. The six models examined were; two types of parallel plate approximation, the capacitance of a conducting isolated sphere, the self capacitance of a conducting sphere with a dielectric shell, a sphere near an infinite conducting plane and two spheres represented in a bispherical coordinate system[3]. The later models attempt to account for the increasing importance of fringing fields as the feature sizes shrink.

The thin film work of Abeles et al[4] suggests a limiting case for the inter-grain capacitance in thin (discontinuous) metallic films. This limiting case was found to be the self capacitance of a dielectric sphere. This argument was arrived at through the charge screening action of neighbouring grains. However, all the self capacitance models suggest grain capacitances and therefore inter-grain capacitances of around 1 to 2 aF (for 25nm spherical grains with 1nm thick AlO_x having a dielectric constant of ε_r = 10). This would represent a SET T_O of around 460K! The model Abeles used examines the case of a grain fully surrounded by other grains. In this work, the structures are patterned into a point and the grains that are brought together have a significant amount of empty space around them. In a number of studies of single electron devices on granular thin films using STM probing, Abeles self capacitance model has found some credence[5]. The high temperatures at which single electron effects have been reported from many of these experiments tends to confirm the self capacitance value. However, the films are of the same type as those investigated by Abeles and they are not patterned as the films in this work. Where the films are patterned into opposing points, it is likely that a significant amount of stray capacitances are likely to result, increasing the overall capacitance.

Returning to the structures fabricated in this work and using a model similar to the parallel plate approximation, a value of around 44aF is obtained for 25nm diameter grains with an AlO_x barrier of 1nm thick, having ε_r = 10. From Roy's work[2] we see that the bispherical model increases the calculated capacitance by just over 2 times that given by the parallel plate model. This puts the capacitance of the junctions in the region of 100aF or produces a T_O of the order of 4.0K, excluding any stray capacitances. This T_O would not result in observable single electron effects at 4.2K in these structures.



a. Parallel plate approximation for structure capacitance

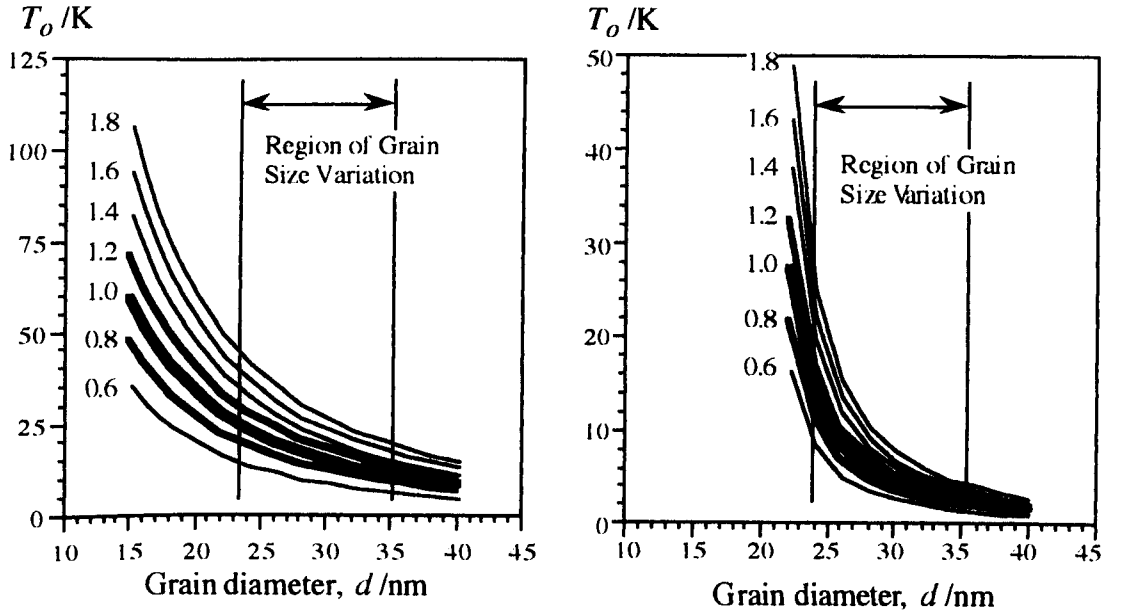
b. Estimate of fringing field effects, through use of bispherical coordinates[2]

Figure 7.2 Comparison of analytical capacitance models.

The two graphs in figure 7.2 estimate the effect of varying the barrier thickness and grain diameter on the modelled inter-grain capacitance. The barrier thickness used in the calculation, given in nanometres, is marked next to the corresponding line. The weight of the line used is an indication of the likelihood of the modelled barrier thickness being the actual barrier thickness, barriers of around 1nm thick being the most likely[6]. Figure 7.2a is a straight forward calculation from the parallel plate capacitance model where the plate area is taken as the maximum cross-sectional area of a sphere of diameter d and the plate separation, t, is taken as the estimated oxide thickness between the grains. The estimate of the effect of fringing fields on the capacitance, figure 7.2b, arises from Roy's comparison between this parallel plate model and the results from a bispherical coordinate based model. Roy calculated the ratio between the two approximations for varying grain separation. This ratio is used here with the parallel plate approximation result to estimate the capacitance, including fringing fields, between the grains. Roy used a 3nm barrier thickness with 20nm diameter grains to examine the effect of increasing the intergrain spacing. The graph presented here assumes that the change in barrier thickness does not significantly affect the results of the bispherical model.

It is important to note that the model used to include the fringing effects has some specific limiting criteria. The validity of the model depends on the ratio of grain size to barrier thickness. This is because reduction in the barrier thickness requires the number of terms in the series expansion of the analytical result to increase in order to maintain

reasonable calculation accuracy through series convergence.



a. T_0 for the Parallel plate approximation of structure capacitance

b. T_0 for the fringing field effect estimate.

Figure 7.3 Transition temperatures resulting from the capacitance models.

From figures 7.2 and 7.3, we can see that the parallel plate approximation is producing higher estimates of the junction T_0. At 30nm grain diameter, we have a likely junction capacitance of 62.5aF, giving a T_0 of 14.8K. This gives a structure capacitance of some 125aF with a T_0 of around 7.4K. This transition temperature should produce observable, albeit partially washed out, single electron effects in a system cooled to 4.2K. As single electron effects have not been observed in the devices fabricated here, we may agree with Roy's statements on the inaccuracies of the parallel plate model. That is, that inclusion of fringing field effects into the capacitance calculation tends to result in an increase in the capacitance of this type of structure.

Fringing fields arise from the spherical, or near spherical surface of the metal grains. This spherical surface increases the area of the equipotential surfaces in the electric field between the metal grains. As the curvature of the electric field increases, the field density decreases and therefore the contribution of these electric field components to the intergrain capacitance decreases. However, a parallel plate model assumes no fringing at the plate edges and in this case, we have two "spherical" surfaces with the majority of the electric field lines being made up of "fringing" fields. Given that this increases the effective surface area over which the capacitance should be calculated, it is natural to expect the intergrain capacitance to be greater than that depicted by a parallel plate model. Roy's work suggests this modification to be around a factor of 2 to 3 for the

grains used in this work, resulting in single junction transition temperatures either at or below 4.2K. The T_o for a double junction system will be effectively half the T_o for a single junction. Following this argument, we have a possible explanation for the absence of observable single electron effects in this work.

Stray capacitances can also play a significant role, as shown in equation 7.1. These stray capacitances are the capacitance of the island to ground and to the island gate voltage probes. The gate and island capacitances must be included in the island capacitance calculation if they are significant. As the island is supported on 1μm of thermally grown SiO_2 on 525μm of silicon (non-conducting at 4.2K), the island capacitance to ground can be neglected. Similarly the capacitance to the gates is small as the island is some 50 to 100nm away from the gate probes which should yield a capacitance of the order of 1 to 2 aF or less. The structure capacitance used in this work should therefore be dominated by the tunnel junction capacitances.

Turning this discussion to the parameters which can be modified in order to improve the device structure, we have only three; the material system, the barrier thickness and the grain size. Changing the material system allows the possibility of different, lower dielectric constant tunnel barriers to be used, for example chromium oxide. However, it is possible to introduce added complications through leaky oxides arising from varying oxidation states. Changing the barrier thickness can be accomplished through varying the oxidation conditions and also through physical separation of the grains in the contact. The greatest gain could be achieved through patterning to separate the grains but the control of this parameter is likely to be extremely poor using the lithography technique developed here, as witnessed by the work of section 4.4.4. In some respects this is the advantage given by the Step Edge Cut-Off (SECO) technique[7]. Control over the barrier thickness is limited due to the range in thicknesses over which a natural oxide can grow, that being around 1 to 3nm on aluminium. Oxidation through electrolysis is a possibility for increasing the oxide thickness, however corresponding increases in the device resistance will result. Reduction in the grain size is certainly possible. Within the natural variation of grain sizes, a lower limit of around 25nm on the mean grain size is likely. Size reduction below this can be achieved through a cooled substrate. However, limits are again imposed as aluminium evaporated onto a 77K substrate will form smaller grains than on a room temperature substrate, but these anneal to a larger size when brought to room temperature[7]. Reduction of the grain size would be a useful test of the results of the model in order to show whether or not the effects of fringing fields are as significant as they appear.

7.4 FURTHER DEVELOPMENT OF THE GRANULAR TECHNIQUE

One difficulty with the structure used to form the single electron devices in this work is environmental decoupling. The double junction system used in this work was detailed in section 2.3.2. This simple system is particularly prone to noise and fluctuations, including co-tunnelling effects, all of which tend to reduce the useful operational area of the device. Co-tunnelling and quantum fluctuation effects can be reduced through the use of environmental decoupling[8,9]. This decoupling is easily achieved through fabricated resistors or other tunnel junctions. Developing the technique used here to write narrow (20 to 30nm wide) lines would provide this decoupling. Metallization of a line which is as wide as the natural grain size should provide a single line of grains. When aluminium is deposited into this structure in the same manner as the films in this work, these lines should form a series of oxidised grains. This would result in a line of metallic islands formed by the aluminium grains with tunnel junctions formed at every grain boundary by the granular oxide layer. This is essentially a "single electron wire" or a short 1 dimensional array. Further development of this type of structure allows simple creation of one dimensional arrays of tunnel junctions which may be used as a simple vehicle for soliton and single electron shift register experiments.

A single electron system with such a small island size raises important questions over the ability to supply gate voltages to the metal islands. In such a system, the island to island spacing is 20 to 30nm. In order to provide a gate or island voltage probe for each island, alternate sides of the wire must be used, as depicted in figure 7.4a. However, the gate - gate spacing would be 40 to 60nm. This requires writing of a grating of 20 to 30nm lines with a 1:1 mark space ratio on either side of the "single electron wire". Using current technology, it is impossible to register this type of feature accurately enough to fabricate it using a second layer of a different metal type. Therefore, the wire pattern and the gate pattern must be fabricated in the same lithographic layer. This necessarily requires the gates to be formed from the same metallization. The gates will therefore have the same characteristics as the single electron wire resulting in gate potentials which are indeterminate due to the single electron nature of the gate wire. Also, with gate spacing so close to adjacent islands, the gating effect on adjacent islands will be significant due to the high stray capacitance components of such a structure. This problem could be solved by gating of every other island (as in the Geerligs turnstile[10]) allowing 80 to 120nm spacing between the gate electrodes on either side of the wire. This increase in the space available allows alteration of the gate electrode shape to ensure metallic electrodes and therefore, known electrode potentials on the

island gate.

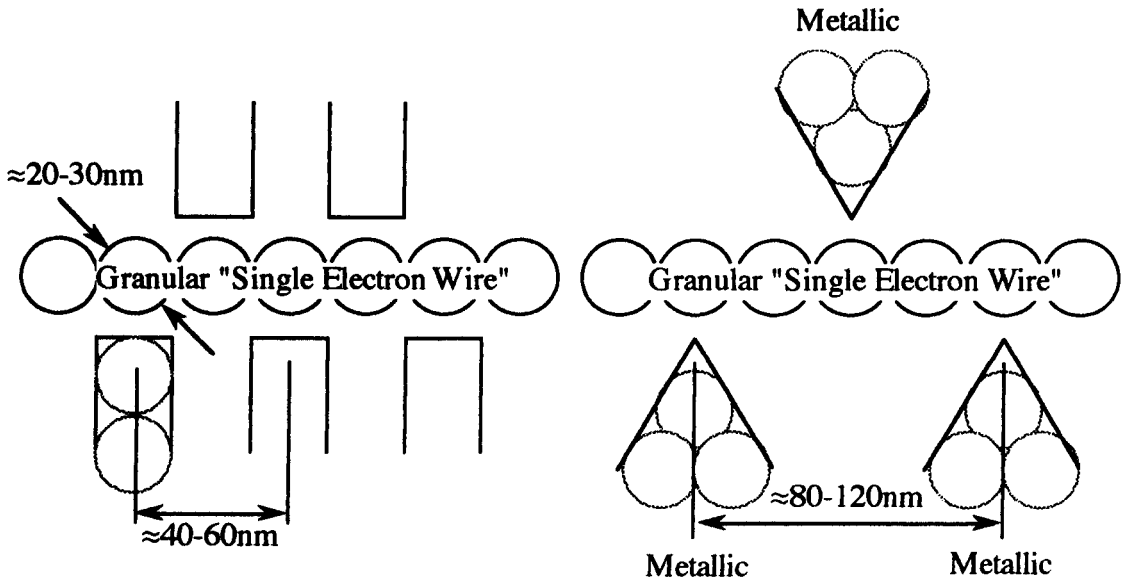


Figure 7.4 Gating schemes for granular "single electron wires".

The problem over structure dimensional control may be addressed by the use of higher resolution lithography techniques. The system used (described in section 3.3.1.1) provides minimum guaranteed Gaussian spot size of 15nm diameter, with the possibility of utilising a estimated 12nm diameter spot. Other electron beam lithography systems exist which are capable of providing spot sizes at least 4 times smaller than this. Such a system has been used to pattern 3 to 6nm wide wires using metallization and lift-off for pattern transfer[111]. Use of such a system would ultimately provide greater control over the structure dimension, as may be required through the discussion of section 7.2. The use of such a system is not necessarily in order to narrow the linewidths, but to provide greater resolution and therefore control of the 20 to 30nm lines or constrictions. This increase in resolution would provide more repeatable lithographic patterning and therefore more repeatable, controllable structures.

Reduction in the capacitance of the structure is required in order to raise the operating temperature of the device or single electron system. It appears that the possibilities for such reductions, within this material system, are limited as discussed in section 7.3. However, it is important to note that this relies heavily on the results of, albeit more accurate, capacitance models. It is important to determine whether these structures do indeed form single electron devices and at what temperature. It is only through these measurements that direct experimental estimates of the structure capacitance can be made. This provides an essential test of the models and also information about the likelihood of raising the temperature further in this material system.

Finally, fabrication of the films and device structures onto a conducting interconnection lead frame is essential in order to make use of very low temperature measurement equipment. However, any moves to reduce the grain size ultimately lead to a reduction in film thickness and therefore a reduction in the ability to connect to the films without using the pressed indium bead technique. As long as the films remain as the final layer in test structure fabrication, the interconnect layer must remain thinner than the aluminium film in order to prevent step coverage discontinuities.

7.5 POSSIBILITIES FOR HIGH TEMPERATURE SET DEVICES

Perhaps the most promising work from the point of view of high temperature single electron devices is that involving STM probing of thin films. This work was briefly discussed in section 2.3.4. This technique uses known processes of film deposition and tip fabrication to provide systems of extremely low capacitance tunnel junctions. The flexibility of the tip to sample spacing allows variations in the junction resistance and capacitance parameters in double and multi junction systems to be investigated. The major shortcoming of this type of system is the inability to controllably gate the structures and thereby independently control the device operation. Essentially, these techniques do not allow for three terminal device fabrication.

The work carried out on thin metallic films, however, extends the ideas of the STM work in order to gain the advantages of the STM approach while maintaining the ability to fabricate three terminal devices. This work relies on the small grain size of the film material coupled with the ability to pattern to small dimensions to provide tunnelling junctions with low capacitance. This feature has been investigated by the work of others and this thesis. Most notably, the work of Chen and Ahmed[12] and Langheinrich and Ahmed[13] has shown that the use of very small grain size materials and processing to deliberately introduce breaks in the material significantly reduces the tunnel junction capacitance. This correspondingly raises the operating temperature of the device. In both cases above, devices were operational at 77K, with traces of single electron effects present at room temperature. While the first technique[12] depends on the random spacing and placement of grains in a discontinuous film, the structures formed by the second technique[13] are directly controlled by the fabrication process. This technique has been further developed and termed the "step edge cut-off" technique[7].

Some work on metallic films has looked at changing the material system[14]. However, the benefits to be gained through such changes are at best second order improvements

compared with those achieved by pure size reduction. Perhaps an exception to this is where the change in material system moves the device technology away from metals altogether. Work on poly-silicon[15,16] and patterned silicon islands[17] along with work on other semiconductor[18] material systems has shown that it is possible to raise device operating temperatures in these systems by changes in fabrication techniques. However, all of these systems are operational at less than 130K, with the majority at less than 30K and a number produce either variable effects or effects uncontrollable by the fabrication technique[19].

There is little doubt, therefore, that the main method of increasing the operational temperature of single electron systems is to reduce the structure dimensions and thereby the capacitance of the tunnel junction. This then leads to the question of whether it is possible to use traditional fabrication techniques and material systems to gain room temperature single electron devices. The work of Chen and Ahmed[12] has employed some of the smallest dimension structures used so far and yet the devices are only just operational at room temperature. Similarly the control over the definition of these structures is poor and control of conduction through a single 2nm metallic island is almost impossible at present. Therefore, in order to have three terminal single electron devices, pattering on a true nanometre scale is required. This is the case whether or not some form of natural ordering or self-assembly can be utilized in the fabrication system.

Due to these extreme dimensional requirements, it is useful to consider the possibilities of using naturally nanometric scale materials, such as macro-molecules. It is clear that as the traditional fabrication systems approach the molecular and few atom limit, the importance of molecules for single electronics must increase. Studies into the electronic behaviour of molecular systems are already well underway[20] and some work has already been documented specifically for single electron behaviour[21,22]. From these results, where room temperature operation can be easily achieved, it appears that molecular electronics is likely to provide the structures for true high temperature single electron devices. Similarly from the more traditional or macro-scaled fabrication techniques, island sizes of 1nm or below are required for true room temperature operation. Clearly, a significant amount of work is still required before single electron devices are a viable, usable and operational technology.

7.6 CONCLUSIONS AND FURTHER WORK

This work set out to investigate the fabrication processes used in aluminium based single electron device technology. From this investigation the goal of a working single electron device was set. This goal was not met. However, the reasons for the failure to

meet this goal have been discussed in order to extract as much information about the situation as possible. From this discussion, a number of conclusions can be drawn and are discussed in the following paragraphs. Following these conclusions suggestions for further work in the development of single electron devices and this fabrication technology are made.

7.6.1 CONCLUSIONS

From the fact that the fabricated devices did not yield single electron effects at a temperature of 4.2K, we believe that we have structure capacitances which are too large. The reasons behind this have been discussed in section 7.3. The question of whether this is an effect of noise coupling into the measurement is eliminated by the noise transmission characterization work documented in chapter 6. This effectively leaves us with the conclusion that the material system being used here is either too large or inherently too capacitive to provide us with 4.2K operation. It is expected, however, that the structures are not too far away from being operational at 4.2K. The modelling work of S. Roy[2] tends to support this. Within the material and lithography system there is some room for improvement and reduction of the capacitance. This can be achieved through reduction in grain size, reduction in pattern size and changes in the pattern layout. However, as the modelling work suggests, it is unlikely that any of these approaches will give the required capacitance decrease to significantly raise the operational temperature of this type of single electron device.

While it is true that a useful single electron device technology will require to be operational at room temperature, raising of operational temperatures to 4.2K is a significant improvement. This allows simpler experimentation with the single electron structures and device types for evaluation and development purposes. This experimentation with working circuits and systems is essential, as so far there is little experimental work on fully controllable memory elements or logic circuits. Therefore, work related to the elevation of the operating temperature of this type of device is valid. While it may be possible to use this type of technology, or variants of it, to produce elevated operating temperature devices, it appears unlikely that the aluminium system will provide room temperature operation with current patterning techniques.

The tunnel junctions fabricated in this work have been characterized predominantly through the tunnelling resistance. This measurement, although containing some flaws as discussed in section 7.2, shows that there is a strong dependency on at least one process parameter, the oxygen partial pressure during film deposition. This leads to questions over the formation of the tunnel barrier and which parameters may be used to

control it. Within single electron devices, we have competing requirements of ultra-small structure size, to give small capacitances, and low structure resistance, to prevent bandwidth or speed limitations within the devices through *RC* time constants. For given capacitances and stray capacitances, high resistance structures will lead to lower device operational frequencies. Within this process, it appears that at "high" oxidation pressures we have a wide process latitude for high resistance structures. At the "low" pressure end of the scale, it is possible to produce low structure resistances, but with almost no process latitude.

Being able to connect to the aluminium thin films and structures used in this work is a problem, as mentioned in section 7.4, due to the very thin metal layers used. However, the work on the thin film interconnect has provided a solution. The thin film interconnect can easily result in discontinuities in the device layer due to step coverage problems. Use of a suitable interconnect would have made possible the measurement of these thin aluminium film devices in cryogenic electrical measurement systems other than that described in section 6.3. The development of the thin (3:10nm) Ti:Au film described in section 3.4 was aimed at solving this problem. This development work succeeded in so far as tests on the film resistivities were successful on the thicker of the two films used, see sections 5.5.1 & 5.5.3. This film may therefore be used with single electron devices to allow use of ceramic chip headers for bonding and mounting in other measurement systems.

In summary, therefore, we conclude that tunnel junctions have been formed by the patterned granular thin film technique. These tunnel junctions were formed into candidate single electron structures which did not yield single electron effects at 4.2K. The reason for this is believed to be a higher than expected capacitance from the fabricated system. This proposal is backed up through some analytical models which suggest that the operational point is not much below 4.2K. From this conclusion there appears to be some room for improvement within the patterning, processing and materials in order to raise the operational temperature. However, it is not expected that the techniques documented here will raise the operation temperature significantly above 4.2K.

From the discussions at the end of chapter 2 and the further discussion in section 7.4, it appears that the traditional technologies may not provide room temperature single electron devices. In all cases where room temperature single electron operation has been reported, the relevant structure sizes have been very much smaller than those used in this work. It is clear that to raise the operational temperature well above room temperature, structure patterning and control at the 1nm level is required. This remains

the major challenge for single electron device technology.

7.6.2 FURTHER WORK.

From the conclusions above, it is clear that the patterned granular thin film structures must be measured at a lower temperature in order to assess the true single electron operational temperature. This assessment will ultimately provide values for the junction capacitances and the gate capacitances. In determining these parameters, the capacitance models can be verified or dismissed. When this information is known, further development work for these structures can be undertaken. This development work is related to the pattern and film size reduction. In order to prove that tunnel junctions are indeed formed by this technique, the superconductive tunnelling characteristics of single junctions fabricated in this manner may be taken. These measurements would conclusively prove the tunnelling behaviour of the structures regardless of whether or not single electron effects were observed.

It is unfortunate that the project did not develop sufficiently to allow investigation of the process dependencies of the structure fabrication. A tentative suggestion of a logarithmic dependency of the structure resistance on the oxygen partial pressure during deposition has been made and it would be valuable to confirm this. If the dependency is proved to be logarithmic, then there is a real difficulty in producing low resistance single electron devices. The high resistance that we are then limited to would result in serious bandwidth limitations for the device. Another area for investigation relates to the dependency of the tunnel resistance on the structure size. Again if this is significant, greater control is required and therefore a change in the lithography system must result to provide this control.

In terms of developing the fabricated structure, there are some areas which require work. Firstly, there is a need to develop a method of separating the grains by patterning or processing. This will result in a capacitance reduction and therefore an increase of the T_O. Secondly investigation of the process dependencies is necessary. This is needed to see the effects of lithography pattern variation and variations in the deposition conditions. This investigation is require in parallel with measurement of the effect on device resistance. Thirdly, will changing the device geometry into a "single electron wire" as discussed in section 7.4 improve the T_O and the device behaviour? Finally, investigation of the scope and range over which the grain size can be reduced in aluminium is important in order to raise the device T_O. It is clear from the discussion and recent work that very small grain size systems yield the highest operational temperature single electron devices. It may therefore be worthwhile applying this

technique to other smaller grain size materials.

Within the electronics system we have the advantages of simplicity of set-up and ease of use. However, the fixed trans-impedance gain setting of the pre-amplifier is restrictive. The use of a T-network in the feedback circuit[23] to provide a wide range of switchable gains (without the use of costly high value resistors) would be a major benefit. Similarly reduction of the noise generated within this circuit will improve the precision of the measurement circuit. The fixed differential divide by 50 circuit on the input to the pre-amplifier should be modified to allow the division to be switchable. This change, coupled with switchable trans-impedance gain would allow measurement of large bias voltage and sample current ranges, providing a more flexible and versatile measurement system. A final improvement would be direct measurement of the voltage applied to the device. Currently this is not accurately measured and forms a significant portion of the circuit measurement tolerance. Improvement of the measurement accuracy would be gained through such a step and the usefulness of the system would be increased by allowing measurement of both the current flowing and the voltage applied to a device.

7.7 REFERENCES

1. T.A. Fulton & G.J. Dolan, *Phys. Rev. Lett.* **59**:109 (1987).

2. S.A. Roy, *Simulation tools for the analysis of single electronic systems*, Chapter 3: Analysis of device geometry and capacitance, Ph.D Thesis, No. 10390, University of Glasgow, 1994.

3. Y. Moon & D.E. Spencer, *Field Theory for Engineers*, 1st Ed., The Van Nostrand Series in Electronics and Communications, ed. H.J. Reich, D. Van Nostrand Company Inc. 530, Princeton, New Jersey, 1961.

4. B. Abeles, P. Sheng, M.D. Coutts & Y. Arie, *Adv. Phys.* **24**(3):407 (1975).
and B. Abeles, *Granular metal films*, In: *Applied Solid State Science. Vol. 6.*, ed. R. Wolfe, Academic Press, London, 1976.

5. C. Schonenberger, H. van Houten & H.C. Donkersloot, *Euro. Phys. Lett.* **20**(3):249 (1992).

6. T.E. Hartman, *J. Vac. Sci. Technol.* **2**,239 (1965).

7. S. Altmeyer, B. Spangenberg, F. Kuhnel & H. Kurz, *Microelectronic Eng.* **30**(1-4):399 (1996).

8. A.N. Cleland, J.M. Schmidt & J. Clarke, *Phys. Rev. B.* **45**(6):2950 (1992).

9. P. Delsing, T. Claeson, K.K. Likharev & L.S. Kuzmin, *Phys. Rev. B.* **42**(12):7439 (1990).

10. L.J. Geerligs, V.F. Anderegg, P.A.M. Holweg, H. Pothier, D. Esteve, C. Urbina, M.H. Devoret & J.E. Mooij, *Phys. Rev. Lett.* **64**(22):2691 (1990).

11. D.R.S. Cumming, S. Thoms, S.P. Beaumont & J.M.R. Weaver, *Appl. Phys. Lett.* **68**(3):322 (1996).

12. W. Chen & H. Ahmed, *J. Vac. Sci & Technol. B.* **13**(6):2883 (1995).

13. W. Langheinrich & H. Ahmed, *Jpn. J. Appl. Phys.* Prt. 1 **34**(12B):6956 (1995).

14. L.S. Kuzmin, Y.A. Paskin, A.N. Tarkhelidze, F.J. Ahlers, T. Weimann, D. Quenter & J. Niemeyer, *Appl. Phys. Lett.* **68**(20):2902 (1996).

15. K. Yano, T. Ishii, T. Hashimoto, T. Kobayashi, F. Murai & K. Seki, *Proc. International Electron Devices Meeting*, Washington DC., Dec. 5-8th, 1993.

16. A. Ohata, H. Niiyama, T. Shibata, K. Nakajima & A. Toriumi, *Jpn. J. Appl. Phys.* Prt. 1 **34**(8B):4485 (1995).

17. A. Fujiwara, Y. Takahashi, K. Murase & M. Tabe, *Appl. Phys. Lett.* **67**(20):2957 (1995).

18. M. Dilger, K. Eberl, R.J. Haug & K. von Klitzing, *Superlattices & Microstructures*, **21**(4):533 (1997).

19. K. Nakazato, R.J. Blaikie & H. Ahmed, *J. Appl. Phys.* **75**(10):5123 (1994).

20. M. Bryce, ed., *Molecular electronics*, Chemistry in Britain, August, 1991, pp707-723.

21. S.A. Iakovenko, A.S. Trifonov, F.S. Soldatov, V.V. Khanin, S.P. Gubin & G.B. Khomutov, *Thin Solid Films*, **284-285**:873 (1996).

22. M. Dorogi, J. Gomez, R. Osifchin, R.P. Andres & R. Reifenerger, *Phys. Rev. B.* **52**(12):9071 (1995).

23. G.B. Clayton, *Operational Amplifiers*, 2nd ed., Newnes-Butterworth, 1979.