PROPERTIES AND APPLICATIONS OF THE MIS TUNNEL DIODE

PROPERTIES AND APPLICATIONS OF THE METAL-INSULATOR-SEMICONDUCTOR (MIS) TUNNEL DIODE

By

MARTIN ANDREW GREEN, B.E., M.Eng.Sc.

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfilment of the Requirements

for the Degree

Doctor of Philosophy

McMaster University

June 1974

(c) MARTIN ANDREW GREEN 1975

DOCTOR OF PHILOSOPHY (1974)
(Electrical Engineering)

MCMASTER UNIVERSITY
Hamilton, Ontario.

TITLE: Properties and Applications of the Metal-Insulator-Semiconductor (MIS) Tunnel Diode

AUTHOR: Martin Andrew Green, B.E. (University of Queensland)
M.Eng.Sc. (University of Queensland)

SUPERVISOR: Professor J. Shewchun

NUMBER OF PAGES: xix, 271

ABSTRACT

This work describes a theoretical and experimental investigation into the electrical properties and possible device applications of the three layer Metal-Insulator-Semiconductor (MIS) structure when the insulating layer is very thin (< 60Å). Under these conditions, current can flow between the metal and the semiconductor by quantum mechanical tunneling processes. This work is limited to the case where the semiconductor is not degenerately doped.

To aid in the discussion of the properties of this structure, its electrical behaviour is simulated by describing the transport processes through the semiconductor and the insulator regions mathematically and solving the resulting system of mathematical equations numerically. Solutions for both the AC and DC properties of the structure are obtained in this way. Using these solutions as a guideline, a general theoretical framework is established which allows the device properties to be predicted when parameters such as the work function of the metal contact, the insulator thickness, the properties of surface states at the insulator-semiconductor (IS) interface, the device temperature, and the dopant type in the semiconductor are varied. Particular importance is placed upon the charge conditions in the semiconductor at the IS interface under zero bias conditions. Depending upon whether this region is strongly inverted, depleted, or accumulated, the dominant component of the diode current tends to flow between the metal and the minority carrier energy band, surface state levels, and the majority

carrier energy band in the semiconductor respectively. Correspondingly, the devices are classified as minority carrier, surface state, and majority carrier diodes. In addition to this classification, a distinction between "equilibrium" and "non-equilibrium" devices is made. In the former, the insulator layer is thick and the device currents small while in the latter, this layer is thinner and the enhanced current flows cause the semiconductor to depart significantly from thermal equilibrium. Experimental measurements upon the metal-silicon dioxide-silicon system are used to demonstrate the validity of this theory and also demonstrate the utility of the classification scheme described.

Non-equilibrium diodes are of particular interest because of their possible device applications. Majority carrier non-equilibrium devices exhibit current multiplication properties which are discussed theoretically and measured experimentally. These make the devices of interest in transistor structures and in such applications as photo-diodes with internal multiplication properties. Minority carrier non-equilibrium devices are shown both theoretically and experimentally to possess properties similar to p-n junction diodes under reverse and small to moderate forward bias. Since they are considerably simpler to fabricate, they have possible application as replacing junction devices in a number of applications such as those as injecting contacts, photo-diodes, or members of photo-diode arrays. One particularly promising application is to direct energy conversion using the electron- or photo-voltaic effect. For an energy conversion situation using the radioisotope Pm^{147} as the primary energy source, it is demonstrated that the predicted higher conversion efficiency of the MIS tunnel diode compared to the p-n junction diode can be experimentally realized.

ACKNOWLEDGEMENTS

I wish to express my appreciation to my supervisor, Dr. J. Shewchun, for the guidance provided during this thesis work and to Dr. D.A. Thompson for assuming the responsibilities of this role during its final stages. I also gratefully acknowledge the contributions made by R.A. Clarke, F.D. King, and V.A.K. Temple to various aspects of this work. I would also like to thank B. Erskine for her accurate and rapid typing of this manuscript.

I gratefully acknowledge the receipt of a Canadian Award under the Commonwealth Scholarship and Fellowship Plan and the funding provided by the National Research Council of Canada and Atomic Energy of Canada Limited.

Finally, I would like to thank Judy Green for her companionship and encouragement during this work and to her I dedicate this thesis.

TABLE OF CONTENTS

PAGE

1. INTRODUCTION 1

2. CHARACTERIZATION OF THE MIS TUNNEL SYSTEM 6

2.1 General Discussion 6

2.2 Semiconductor Region 12

2.2.1 Transport Equations 12

2.2.2 Small Signal Transmission Line 16

2.3 Band Tunnel Currents 19

2.4 Surface States 22

2.5 Application of the Small Signal Transmission Line to the AC, DC, and Transient Analysis of Semiconductor Devices 28

2.5.1 Introduction 28

2.5.2 Small Signal AC Analysis 34

2.5.3 DC Solutions 36

2.5.4 Large Signal Transient Analysis 40

2.5.5 Performance of Algorithm 43

2.6 Additional Topics 44

2.6.1 Multi-Dimensional Effects 44

2.6.2 Parameter Values 47

3. THEORY OF THE MIS TUNNEL DIODE 51

3.1 Preliminary Discussion 51

3.2 Equilibrium Diodes 54

3.2.1 Capacitance-Voltage Characteristics 54

PAGE

3.2.2 Current-Voltage Characteristics 56

3.2.3 Duality of n and p Type Diodes 63

3.3 Non-Equilibrium Diodes 65

3.3.1 Effect of Insulator Thickness 65

3.3.2 Metal Work Function Effects 73

3.3.3 Complementary n and p Diodes 78

3.4 Surface State Effects 78

3.4.1 Introduction 78

3.4.2 Effects Upon Equilibrium Diodes 82

3.4.3 Effects Upon Non-Equilibrium Diodes 85

3.5 Temperature Effects 89

3.6 Summary 93

4. EXPERIMENTAL PROPERTIES OF THE MIS TUNNEL DIODE 97

4.1 Fabrication and Measurement 97

4.2 Interface Concentration at Zero Bias 100

4.3 Comparison of n and p Type Devices 104

4.4 Effect of Metal Work Function 107

4.5 The Effect of Temperature 127

4.6 Comparison with Previous Experimental Work 138

4.7 Summary 144

5. CURRENT MULTIPLICATION IN MAJORITY CARRIER DIODES 146

5.1 Introduction 146

5.2 Theory of the Multiplication Process 147

5.2.1 Qualitative Description 147

5.2.2 I-V Characteristics of a Reference Diode 150

PAGE

5.2.3 An Expression for the Multiplication Factor M 154

5.2.4 Effects of the Substrate Resistivity and Orientation 160

5.2.5 The Effects of Metal Work Function and Injected Current Density 161

5.2.6 The Effects of Oxide Charge and Oxide Thickness 164

5.2.7 Effects of Temperature and Frequency Response 166

5.2.8 The Effects of Surface States 169

5.3 Experimental Measurement of the Multiplication Process 175

5.3.1 Experimental Method 175

5.3.2 Experimental Results 179

5.4 Applications of the Multiplication Effect 190

5.5 Summary 193

6. MINORITY CARRIER DIODES AND THEIR APPLICATION TO DIRECT ENERGY CONVERSION 196

6.1 Introduction 196

6.2 Theory 197

6.2.1 I-V Characteristics 197

6.2.2 Clamping of the Inversion Layer 207

6.2.3 The Effect of Surface States 210

6.3 Experiment 215

6.3.1 Minority and Majority Carrier Diodes 215

6.3.2 I-V and C-V Characteristics of Minority Carrier Diodes 218

6.3.3 Comparison of Minority Carrier Diodes 226

6.3.4 The Effects of Substrate Resistivity 228

6.4 Application to Direct Energy Conversion 230

PAGE

6.5 Similarities to p-n Junction and Schottky Diodes 236

6.6 Summary 238

7. CONCLUSIONS 240

APPENDIX A Expressions for the Tunnel Currents Between the Metal and the Semiconductor Bands 245

APPENDIX B Efficient Numerical Solution of the Transmission Line Equivalent Circuit Model of a Semiconductor 252

APPENDIX C Expressions for the Capacitance of MIS Tunnel Diodes 258

C.1 Minority Carrier Diodes 258

C.2 Majority Carrier Diodes 261

APPENDIX D Expression for the Electric Field at the IS Interface 263

REFERENCES 266

LIST OF TABLES

TABLE		PAGE
2.1	Definition of Semiconductor Equivalent Circuit Elements	18
2.2	Element Values for Figure 2.4	24
2.3	Element Values for Figure 2.5	30
2.4	Numerical Values Assigned to the Parameters of the Metal-Silicon Dioxide-Silicon System	49
4.1	Metal to Insulator Barrier Heights for Ti, Al, Ni, Au, and Pt	103
4.2	Summary of Reverse C-V Characteristics for p Type Devices with Mg, Ti, Al, Ni, Au, and Pt Contacts	126
4.3	Summary of Previously Published Measurements Upon the Non-Degenerate Metal-Silicon Dioxide-Silicon System	139
5.1	Details of Experimental Devices Used to Study the Multiplication Process	182
5.2	Frequency Response and Current Gain of Device N0515	188
6.1	The Effect of Substrate Resistivity Upon the Initial Beta Cell Output	234

LIST OF FIGURES

FIGURE		PAGE
2.1	Schematic of the Metal-Insulator-Semiconductor (MIS) diode and schematic energy band diagram showing the major tunnel currents in the interfacial region of the device.	7
2.2	Small-signal transmission line equivalent circuit model of a semiconductor device.	17
2.3	A typical section of the semiconductor small-signal transmission line equivalent circuit as used in AC analysis.	17
2.4	Small-signal equivalent circuit describing the insulator and the band tunneling currents.	23
2.5	Small-signal equivalent circuit describing surface states and tunneling between the metal and these states.	29
2.6	A section of the semiconductor small-signal transmission line including modifications necessary to employ the circuit for DC analysis.	39
2.7	A section of the semiconductor small-signal transmission line in the form suitable for the transient analysis of semiconductor devices.	39

FIGURE		PAGE
3.7	Computed C-V curves corresponding to the I-V curves of Figure 3.6.	69
3.8	Demonstration of frequency dispersion effects in both the C-V and G-V curves for minority carrier "non-equilibrium" devices even in the absence of surface states.	74
3.9	Computed I-V curves for minority carrier non-equilibrium diodes with the metal to insulator barrier height as the variable parameter. Also shown is the dependence of the intercept, V_c, of the linear $1/C^2$ versus V portion of the C-V characteristics with the voltage axis upon this parameter.	75
3.10	Comparison of the computed I-V characteristics for non-equilibrium devices with the same metal contact upon n and p type substrates.	79
3.11	Assumed distribution of surface states across the energy band gap of silicon.	81
3.12	Computed effects of surface states upon the I-V characteristics of equilibrium diodes.	83
3.13	Computed effects of surface states upon the I-V characteristics of "non-equilibrium" diodes as the metal to insulator barrier height varies.	87

FIGURE		PAGE
3.14	Computed effects of temperature upon the I-V characteristics of equilibrium and non-equilibrium diodes.	91
3.15	Schematic showing the division of the MIS tunnel diodes into minority carrier, surface state, and majority carrier diodes as a function of the metal to insulator barrier height for both n and p type substrates.	95
4.1	Experimental C-V curves for n and p type devices with Ti, Al, Ni, Au, and Pt contacts.	102
4.2	Experimental comparison of the I-V characteristics of an n type and a p type device with Al contacts in both the equilibrium and non-equilibrium cases.	105
4.3	Experimental I-V and C-V characteristics of Al n and p type devices for oxide thickness in the range 23-39Å.	108
4.4	Experimental I-V and C-V characteristics of n and p type Au devices for oxide thickness in the range 21-35Å.	116
4.5	Experimental I-V curves for Ti, Al, Au, and Pt devices with both n and p type substrates.	122

FIGURE		PAGE
4.6	Experimental demonstration of the linear $1/C^2$ versus V region in the C-V characteristics of Pt, Au, and Ni non-equilibrium devices.	125
4.7	Experimental measurement of the effect of temperature upon the I-V characteristics of a p type Al minority carrier diode.	129
4.8	Experimental measurement of the effect of temperature upon the I-V characteristics of both a p type Au and an n type Al majority carrier diode.	131
4.9	Effect of temperature upon the I-V characteristics of both a p type and an n type Ni diode.	135
4.10	C-V, G-V, and I-V characteristics of a p type Mg device as measured by Kar and Dahlke [17].	141
5.1	Schematic energy band diagram for an n type majority carrier diode.	148
5.2	Calculated I-V characteristics of a reference Al-SiO_2-n type Si tunnel diode for varying supply rates of minority carriers. Both optimal and non-optimal methods of supplying minority carriers are demonstrated.	152
5.3	Computed I-V characteristics of a MIS tunnel diode similar to the reference diode of Figure 5.2 but with a higher density of impurities in the semiconductor substrate.	162

FIGURE		PAGE
5.4	Calculated dependence of the small signal multiplication factor, M, upon the metal to insulator barrier height and the supply rate of minority carriers.	163
5.5	Calculated effect of oxide charge and oxide thickness upon the I-V characteristics and multiplication properties of an Al-SiO_2-Si tunnel diode.	165
5.6	The effect of temperature upon the computed I-V characteristics of the reference diode of Figure 5.2.	167
5.7	The calculated frequency dependence of the small signal multiplication factor, M.	168
5.8	Assumed distribution of surface states for the calculations of Chapter 5.	170
5.9	Calculated effects of surface states upon the I-V and the multiplication properties of the reference diode.	171
5.10	Calculated frequency dependence of the small signal multiplication factor, M, in the presence of surface states.	174
5.11	Experimental method for investigating the multiplication properties of majority carrier diodes.	176

FIGURE		PAGE
5.12	Measured values of majority (J_{maj}) and minority (J_{min}) carrier device current flow and the resulting small signal current gain, M, for four Al-SiO_2-Si devices.	180
5.13	Observed response of the minority and majority carrier devices to a small chopped light signal.	185
5.14	Observed response of both minority and majority carrier devices to a low intensity flash of light from a stroboscope.	187
5.15	Comparison of the low frequency values of the small signal multiplication factor, M, determined experimentally by both the incremental and the light impulse methods.	189
5.16	"Ring-dot" structure and the measured results using the ring to inject minority carriers.	191
5.17	Schematic of the Surface Oxide Transistor (SOT).	192
6.1	Computed I-V characteristics demonstrating the "equilibrium" to "non-equilibrium" transition for a minority carrier p type diode as the thickness of the insulating layer is reduced.	198
6.2	Schematic energy band diagram of the minority carrier MIS tunnel diode showing the electron and hole quasi-fermi levels.	200

FIGURE | PAGE

6.3 Schematic showing the changes in the energy band diagram for a non-equilibrium minority carrier MIS tunnel diode as the applied bias is varied from large reverse bias to large forward bias. 205

6.4 The computed value of the ratio of the band tunnel currents at 0.3 V bias as a function of the insulator thickness and the metal to insulator barrier height. 206

6.5 Calculated value of the concentration of minority carriers at the IS interface as a function of the diode bias and the thickness of the insulating layer. 208

6.6 The computed effect of surface states upon the I-V characteristics of a minority carrier MIS tunnel diode. Also shown is the effect upon the minority carrier injection ratio in the bulk semiconductor regions of the device. 211

6.7 Experimental comparison of the I-V and C-V characteristics of diodes of similar oxide thickness but with gold and aluminum top contacts. 216

6.8 Experimentally observed I-V and C-V characteristics for an Al-SiO_2-p-100-Si diode and their comparison with theory. 219

FIGURE		PAGE
6.9	Experimentally observed dependence of the I-V characteristics upon the thickness of the insulating layer.	223
6.10	Experimental comparison of the I-V and C-V characteristics of two minority carrier diodes formed on the same substrate but with magnesium and aluminum top contacts.	227
6.11	Experimentally observed dependence of the I-V and C-V properties of Al-SiO_2-p type <100> Si tunnel devices upon the resistivity of the silicon substrate.	229
6.12	Experimental comparison of the energy conversion properties of a diffused p-n junction diode and a MIS tunnel diode using the radioisotope Pm^{147} as the primary energy source.	235
A1	Plot of the Fermi-Dirac Integrals for small positive and negative arguments.	248
A2	Schematic showing the "shadows" of the constant energy surfaces for both the metal and the semiconductor (silicon) at energies near that of the conduction band edge of silicon.	250
B1	A section of the semiconductor small signal transmission line after the elimination of the trapping node, v_t.	253

FIGURE		PAGE
B2	Conversion of the $[Y_k]$ matrix to the $[M_k]$ matrix in three steps of Gaussian elimination.	253
C1	The assumed space charge density and the corresponding electric field and electrostatic potential as a function of distance into the semiconductor for the calculations of Appendix C.	259

CHAPTER 1

INTRODUCTION

The Metal-Insulator-Semiconductor (MIS) structure has established itself as one of the most important in electronic device technology and current developmental work indicates that it will maintain this position for some time in the future. In its technologically important form, the insulator layer is of the order of thousands of angstroms thick and DC current flow through it is extremely small except when the electric field in the insulator attains very large values [1a].

This thesis is concerned with the properties of the MIS structure when the insulator layer is so thin (< 60Å) that measurable current can flow between the metal and the semiconductor by quantum-mechanical tunneling processes without requiring the presence of large insulator fields. Tunneling in this structure was first investigated experimentally by Gray [2] and subsequently by Esaki and Stiles [3] and Shewchun, Waxman, and Warfield [14] as a natural outcome of the interest rearoused in tunneling by the work of Esaki [4], Giaever [5], and Josephson [6]. While Gray [2] used non-degenerately doped semiconductor, Esaki and Stiles [3] employed degenerate semiconductor and observed a negative resistance region. Such degenerate MIS tunnel diodes will not be considered in this thesis since they can be regarded as forming a class of diode distinct from the non-degenerate diodes which are treated. The comparison is similar to that of degenerate and non-degenerate p-n junction devices. The degenerate MIS tunnel diode possesses several

properties not observed in the non-degenerate case. In addition to exhibiting a negative resistance feature in some systems [3,7,8,9], degenerate MIS tunnel diodes appear suitable for the low temperature observation of superconducting band gaps in the degenerate semiconductor [10] and the metal [10,11], and of the interactions of the tunneling current with phonons and impurities in the semiconductor and insulator [12]. Even at room temperature, the presence of dopant impurity bands and band tailing effects in the semiconductor and the importance of tunneling through the combined barrier presented by the insulator and the space charge region of the semiconductor, cause the characteristics of degenerate MIS tunnel diodes to differ significantly from the non-degenerate case [9,11].

Over the last decade, tunneling in the non-degenerate case has been studied by several groups [2,13-20]. The earliest work was with relatively thick (> 40Å) insulating layers [2,14] and direct current flow in these devices was very small. The tendency in subsequent work using thinner insulating layers has been to interpret DC current flows in terms of Schottky diode theory modified to include non-idealities introduced by the insulating layer [13,17,18]. However, the recent experimental work of Clarke and Shewchun [15] established that these thin insulator diodes constitute a unique class of their own. The aim of this thesis is to investigate the properties of this new class of diode both theoretically and experimentally with particular attention paid to properties that make the structure suitable for device application.

The first portion of this thesis, Chapter 2, describes the manner in which the MIS tunnel diode is characterized to study its theoretical

AC and DC properties. Calculations of these properties are carried out by solving the set of differential equations describing carrier transport in the semiconductor region of the diode in conjunction with the expressions describing tunneling between the metal and the semiconductor and charge transfer between surface states at the insulator-semiconductor (IS) interface and the semiconductor bands. A conceptually appealing method developed during this thesis work for solving these equations is described. It is based on the small signal transmission line equivalent circuit developed by Sah [21-23] as an exact analogue of the small signal part of the semiconductor transport equations. Although the use of this circuit previously has been restricted to computing the small signal AC properties of semiconductor devices [22,24-26] it is shown that it can be applied in a simple fashion to compute not only the AC properties but also the DC and the large signal transient properties [27].

The next portion of the thesis, Chapters 3 and 4, is devoted to a study of the current-voltage and capacitance-voltage properties of the non-degenerate MIS tunnel diode. Chapter 3 presents the first comprehensive discussion of the theoretical characteristics of this device. A broad theoretical framework is established to provide a basis for the interpretation of experimental results. It is shown that, depending upon the metal work function, the insulator thickness, the dopant species in the semiconductor and the surface state properties, a given diode falls naturally into one of three classes. A diode is classified as a minority carrier, surface-state, or majority carrier diode depending upon whether the dominant component of the diode tunnel current flows between the metal

and the minority carrier band, surface-state levels, or the majority carrier band in the semiconductor. Each of these classes is further divided into two subclasses, labelled "equilibrium" and "non-equilibrium", depending upon whether or not current flows are small enough to leave the semiconductor undisturbed from thermal equilibrium conditions over the bias range of interest. In Chapter 4, the results of experimental measurements upon the Metal-SiO_2-Si system are described. These substantiate the predictions of the theoretical treatment including the utility of the concepts of minority, surface-state, and majority carrier diodes.

The final part of the thesis, Chapters 5 and 6, deals with those properties of the MIS structure which have device application. Clarke and Shewchun [15] predicted that devices which on the basis of this work would be described as "majority carrier non-equilibrium" might exhibit current multiplication properties. This subsequently was verified in a qualitative fashion in their Surface-Oxide-Transistor structure [28]. In Chapter 5, the multiplication process is subject to a quantitative analysis. Gains of the order of 100 to 1,000 are predicted [29]. Also described are the results of the first experimental measurements of these gains and their dependence upon operating conditions. Gains of the order predicted are observed. The results of an experimental investigation of the frequency response of the process are described and compared to the theoretical predictions. In Chapter 6, the properties of another class of diode, the "minority carrier non-equilibrium" diode, are discussed in more detail. It is shown that, at reverse and small forward bias,

they exhibit current-voltage and capacitance-voltage characteristics similar to those of p-n junction devices [30,31]. Since the MIS tunnel devices are far more simply fabricated, they are attractive alternatives for p-n junctions in many applications including those as injecting contacts, photo-diodes or members of photo-diode arrays. They are ideally suited for direct energy conversion applications employing the electron- or photo-voltaic effects. An investigation using the radioisotope Pm^{147} as the primary energy source is described and this establishes that the predicted higher conversion efficiencies compared to p-n junction devices can be experimentally realized [31]. In fact, when compared to p-n junction diodes fabricated by a variety of methods including gaseous and contact diffusion, single and multiple layer chemical epitaxy and vacuum evaporation epitaxy, the MIS tunnel diode proved the most efficient [32].

CHAPTER 2

CHARACTERIZATION OF THE MIS TUNNEL SYSTEM

2.1 General Discussion

In contrast to the relative ease with which the MIS tunnel structure (Figure 2.1(a)) can be fabricated, the self consistent analysis of the structure from a basic quantum-mechanical formulation presents a formidable problem whose solution appears far off. The difficulties become apparent when we consider that questions remain to be resolved about simple heterojunctions [33a] and the present structure consists of two different heterojunctions separated by a small number of atomic layers from each other. In order to reduce the problem to a tractable form, a simpler approach to the analysis of the system must be employed. The approach used during this work is outlined below.

Each of the metal, insulator, and semiconductor is assumed to have its bulk electronic band structure and electronic properties right up to the planes defining the interfaces where abrupt changes are considered to occur. At the interfaces, the energy levels in each material are matched by defining semiconductor-insulator and metal-insulator barrier heights which may be regarded as quantities to be determined experimentally. The approximations are introduced that all of the applied voltage is absorbed across the insulator and semiconductor and none across the metal and that the effect of interface states at the metal-insulator interface is negligible. Interface states at the

Figure 2.1(a): Schematic of the Metal-Insulator-Semiconductor (MIS) diode. d is the thickness of the insulator layer and L is the thickness of the semiconductor substrate.

(b): Schematic energy band diagram of the MIS tunnel diode system with a p type semiconductor region. Also shown are the major current components in the surface region of the device. The semiconductor is biased positively with respect to the top metal contact by a voltage V_a.

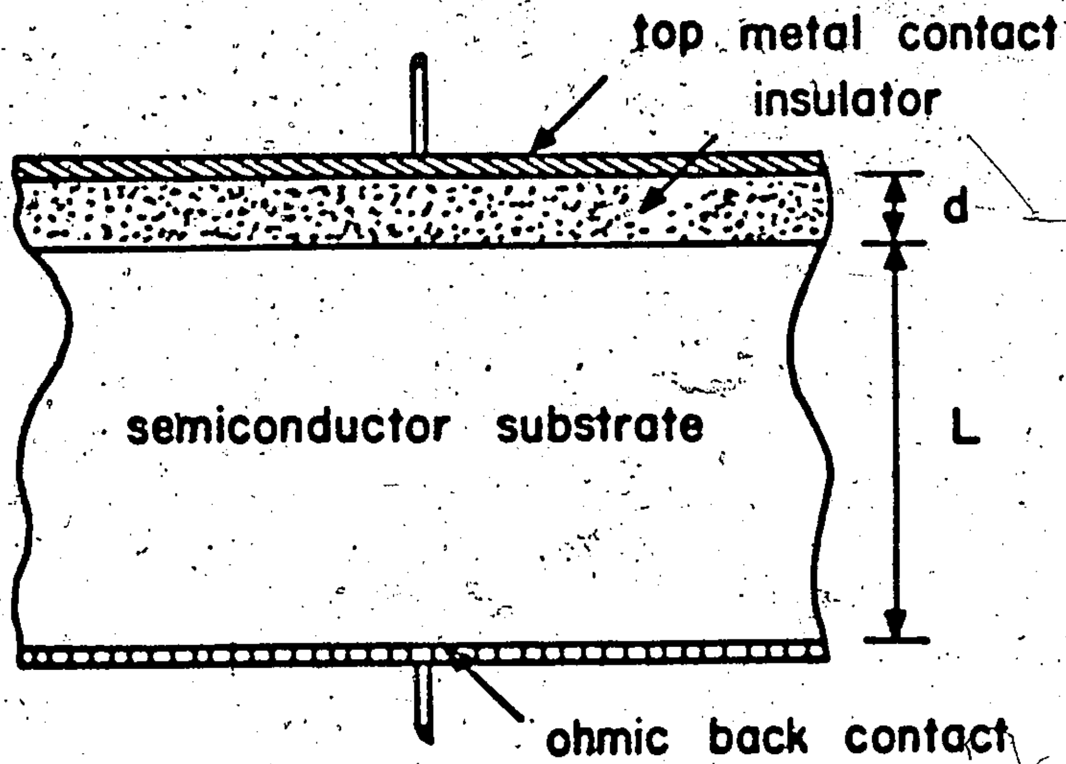


FIGURE 2.1(a)

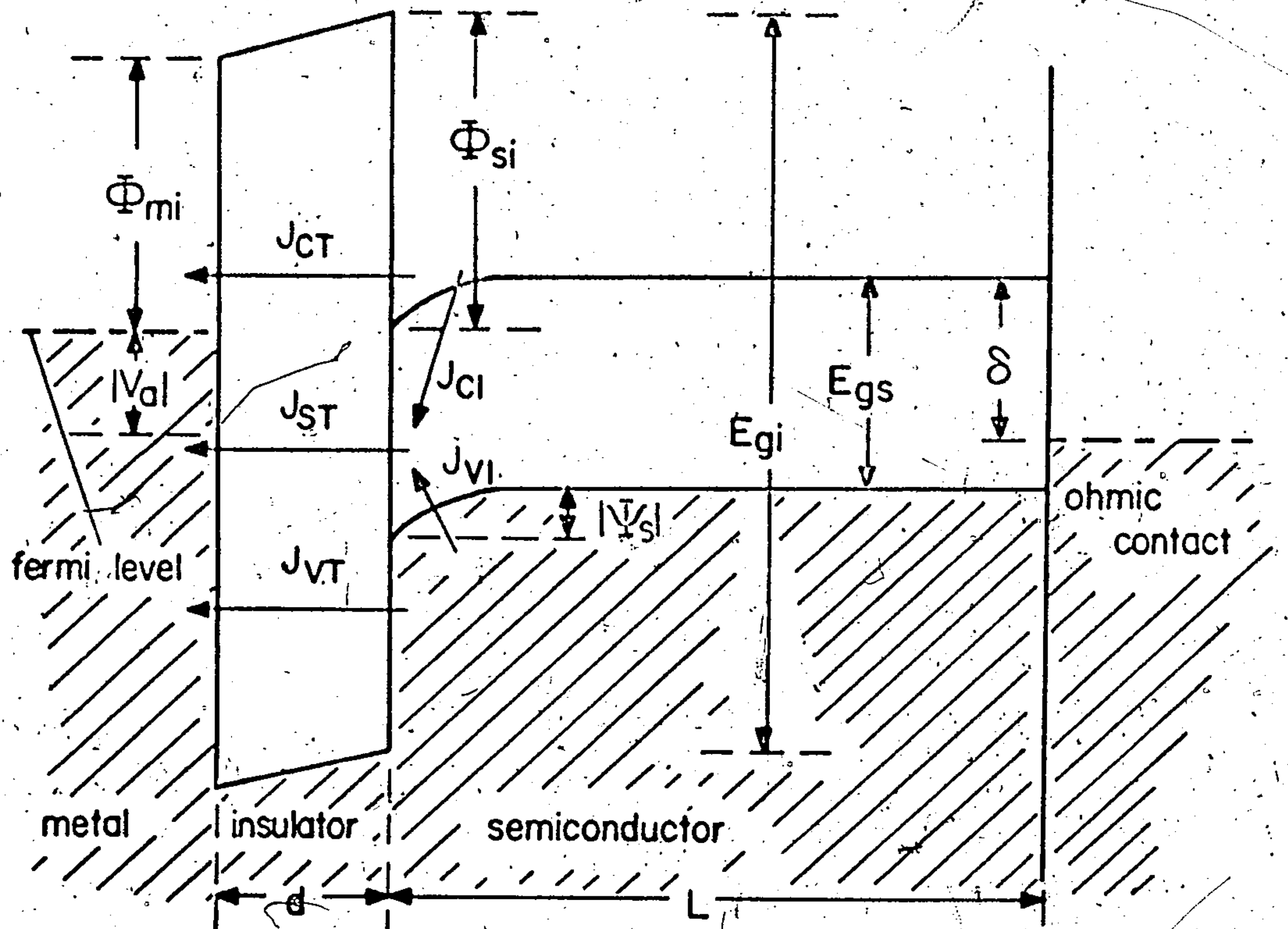


FIGURE 2.1(b)

insulator-semiconductor interface will be shown to play an important role in determining the device characteristics. On this basis, the energy band diagram of the system takes the form shown in Figure 2.1(b).

The semiconductor region of the device is characterized in the manner usual for semiconductor devices [1b]. The properties of this region are described by Maxwell's equations and the carrier continuity equations in conjunction with the Boltzman transport equation developed to its usual stage of approximation to describe electron and hole carrier flows. Shockley-Read-Hall (SRH) statistics were used to characterize recombination-generation processes assumed to occur via single level trapping centres.

Current transport between the metal and the semiconductor through the thin insulator region was modelled as occurring exclusively by tunneling processes. There exists considerable experimental evidence that suggests this is not unrealistic for the M-SiO_2-Si devices fabricated in this laboratory. This evidence has been discussed in some detail by Clarke [11] and only a brief summary will be given here. Firstly, the devices exhibit current-voltage characteristics that are relatively temperature insensitive over predictable bias regions. While this does not prove tunneling is the dominant process, it certainly makes other transport processes seem unlikely. In addition, at a given bias, the current flow has been found to depend exponentially on the average insulator thickness. This dependence would not be expected to be so strong for other transport mechanisms. However, perhaps that most conclusive evidence is that if the devices are made with a degenerately doped substrate and a metal contact which is superconducting at low

temperatures, a pronounced dip in the low temperature conductance-voltage characteristics of the device can be observed. This is interpreted as being due to the superconducting band gap in the metal [11] and establishes tunneling as the dominant process in the bias region where the gap is observed, i.e. within a few millivolts of zero bias.

The tunneling process between the metal and the conduction and valence bands in the semiconductor was characterized using the independent electron approach as described by Harrison [34] assuming specular, elastic transmission. This general approach can be criticized validly on a number of theoretical grounds [33]. Nonetheless, it remains the most widely applied theory for calculating and interpreting tunnel current flow in such barrier dominated structures as the MIS tunnel diode. This is because the formulation, besides producing expressions which can be evaluated relatively simply, is known to be capable of predicting the properties of most experimental devices in at least a qualitative fashion. For the purposes of the investigation described in this thesis, a more elegant and formalistically correct theory is not warranted. Surface-state occupations and recombination-generation characteristics were modelled using SRH kinetics modified to include the effect of tunneling between these states and the metal. Distributions of surface states across the semiconductor band gap were modelled as a large number of independent surface state levels. The current tunneling between the metal and the surface states was characterized simply in terms of "effective" tunnel capture cross sections [35].

The precise form of the equations used to describe the diodes

is given in the following sections. From a mathematical viewpoint, the complete system of equations can be reduced to a set of four coupled non-linear partial differential equations describing the semiconductor regime with fixed boundary conditions at the ideal ohmic back contact. At the IS interface, complicated boundary conditions which depend non-linearly upon the solutions within the semiconductor regime are imposed to describe tunneling and surface state properties. As might be expected, general analytical solutions are impossible to obtain so the problem was tackled numerically. A simple, conceptually appealing method for obtaining numerical solutions is described.

Suppose a solution for the equations is known under DC conditions. If a very small disturbance is applied, the equations describing the resulting small variations in the solution are linear. Sah [22] has shown that those for the semiconductor regime can be represented by a small signal equivalent circuit consisting of linear circuit elements (resistors, capacitors, and voltage-controlled current sources only). The metal, insulator, and surface states can be simply incorporated into this small signal equivalent circuit. It is shown that the use of this small signal equivalent circuit is not restricted to obtaining small signal properties of such semiconductor devices, but that it can be used also for obtaining DC and large signal transient solutions. Consequently, in the following sections, along with the equations describing each region of the device, the corresponding small signal equivalent circuit will be described.

The general theory of the equivalent circuit approach is given in

Section 2.5. Details of a computer program implementing this approach are given in reference [102]. Two additional computer programs were used to generate portions of the data described in the theoretical sections of the following chapters. These are also described in more detail in reference [102]. One of these merely uses a different, more cumbersome algorithm to solve the complex system of equations used to characterize the device. Its use was restricted to DC conditions. Since both this program and the "equivalent circuit" program solved the same system of equations to the same high degree of accuracy (usually 1/2%), both gave virtually identical DC results [102]. The use of the third program was more restricted than either of the above. It could be used only to analyze thick insulator "equilibrium" diodes where the semiconductor region of the device remains undisturbed from thermal equilibrium. This program takes advantage of the simplifications in calculating the semiconductor properties possible under such conditions [102]. All three programs gave nearly identical solutions under these conditions although the third was at least an order of magnitude more efficient.

2.2 Semiconductor Region

2.2.1 Transport Equations

The basic set of equations describing carrier transport in semiconductors consists of Poisson's equation and continuity equations for carriers in the conduction band, valence band, and trapping centres. For the "one-dimensional" geometry of Figure 2.1, these take the following form using Sah's notation [22].

$$\frac{\partial^2 v_I}{\partial x^2} - \frac{q}{\varepsilon}\left(n - p + \sum_i n_{Ti} - \sum_j N_{TTj}\right) = 0 \qquad (2.1)$$

$$\frac{\partial n}{\partial t} - \frac{1}{q}\frac{\partial j_N}{\partial x} + r_{CN} = 0 \qquad (2.2)$$

$$\frac{\partial p}{\partial t} + \frac{1}{q}\frac{\partial j_P}{\partial x} + r_{CP} = 0 \qquad (2.3)$$

$$\sum_i \frac{\partial n_{Ti}}{\partial t} - r_{CN} + r_{CP} = 0 \qquad (2.4)$$

where v_I is the electrostatic potential, n and p are the electron and hole carrier concentrations, n_{Ti} is the density of electrons occupying trapping centre i, and N_{TTj} is the total density of trapping centre j. The subscript i refers to all trapping and impurity levels, both donor and acceptor, while j refers to donor levels only. j_N and j_P are particle current densities while r_{CN} and r_{CP} are the net capture rates of electrons and holes by trapping processes. Additional relationships between n, p, n_{Ti}, j_N, j_P, r_{CN} and r_{CP} must be defined before any solution of (2.1 - 2.4) is possible. These additional relationships can be expressed entirely in terms of v_I, the electrostatic potential, and v_N, v_P, and v_{Ti}, the electron, hole, and trapping level quasi-fermi potentials.

The electron and hole concentrations, n and p, can be expressed as

$$n = \int G_c(E) f_c(E) dE \qquad (2.5a)$$

$$p = \int G_v(E') f_v(E') dE' \qquad (2.5b)$$

where G_c and G_v are the conduction and valence band densities of states, the integrals are over the relevant bands, E and E' are the energies above the conduction band edge and below the valence band edge, and f_c and f_v are the occupation factors for states in each band by electrons and holes respectively, i.e.

$$f_c(E) = 1/(1 + \exp[(E - b)/kT]) \tag{2.6a}$$

$$f_v(E') = 1/(1 + \exp[(E' - b')/kT]) \tag{2.6b}$$

where

$$b = q(v_I - v_N - V_{CI}) \tag{2.7a}$$

$$b' = q(v_P - v_I - V_{IV}) \tag{2.7b}$$

V_{CI} and V_{IV} are the potential differences between the intrinsic fermi level and the semiconductor band edges, while v_N and v_P are quasi-fermi potentials. Similarly

$$n_{Ti} = N_{TTi}/(1 + \exp[(v_{Ti} - v_I + v_{TI})q/kT]) \tag{2.7c}$$

The above information is valid regardless of the energy dependence of the density of states in the semiconductor and whether the carrier concentrations are degenerate or not [36]. Under the same conditions, the carrier current density equations can be derived from the Boltzman transport equation assuming the semiconductor is isothermal [37]. These equations can be conveniently expressed in terms of quasi-fermi gradients to give

$$j_N = -q\mu_N n \frac{\partial v_N}{\partial x} \tag{2.8}$$

$$j_P = -q\mu_P p \frac{\partial v_P}{\partial x} \tag{2.9}$$

where μ_N and μ_P are the carrier mobilities which are bias dependent in the general case [36].

Considering only a single level trapping centre and using a SRH model of recombination-generation processes [38], the expressions for the net capture rates r_{CN} and r_{CP} in the presence of either degenerate or non-degenerate carrier concentrations become [39,40]

$$r_{CN} = c_N[nN_{TT} - (n + n_1)n_T] \tag{2.10}$$

$$r_{CP} = c_P[(p + p_1)n_T - p_1 N_{TT}] \tag{2.11}$$

where N_{TT} is the total density of recombination centres of which n_T are occupied by electrons. n_1 and p_1 are defined as [40]

$$n_1 = n \exp[\frac{q}{kT}(v_N - v_I - V_{TI})] \tag{2.12}$$

$$p_1 = p \exp[\frac{q}{kT}(v_I - v_P - V_{TI})] \tag{2.13}$$

where V_{TI} is the potential difference of the trapping level above the semiconductor intrinsic fermi level. In the general case c_N and c_P are not constant but may be expressed in terms of more fundamental quantities [40]:

$$c_N = \frac{1}{n}\int A_{CT}(E)G_c(E)f_c dE \tag{2.14}$$

$$c_p = \frac{1}{p} \int A_{VT}(E) G_v(E) f_v dE \qquad (2.15)$$

where the integrals are over the respective bands and A_{CT} and A_{VT} are rate constants connecting states in the conduction and valence band respectively and the trapping level [39,40].

If these subsidiary relations (2.5 - 2.15) are inserted into the transport equations (2.1 - 2.4), the resulting equations contain only the potentials v_I, v_N, v_P, and v_T as the dependent variables [102].

2.2.2 Small Signal Transmission Line

Sah [21-24] has shown that the small signal version of the above equations can be described exactly by a small signal transmission line equivalent circuit. The exact small signal equivalent circuit of a typical semiconductor device has the form shown in Figure 2.2. The elements comprising a basic section of the transmission line are shown in Figure 2.3 and defined in Table 2.1. This form of the equivalent circuit is valid only if the carrier concentrations are non-degenerate throughout the device. During the course of this thesis work, the modifications required to incorporate the effect of carrier degeneracies were investigated [36]. It was found that the circuit retained the same basic topology with additional current sources introduced only in the trapping branch associated with node v_T. The values of some of the elements were modified by degeneracy factors whose values approach unity as the carrier concentrations become non-degenerate [36].

As this thesis is restricted to non-degenerately doped semiconductors, the only times where carrier degeneracies are important are when the region in the vicinity of the IS interface is very heavily

Figure 2.2: Small signal transmission line equivalent circuit model of a semiconductor device.

Figure 2.3: A typical section of the semiconductor small-signal transmission line equivalent circuit as used in AC analysis. Element values are given in Table 2.1.

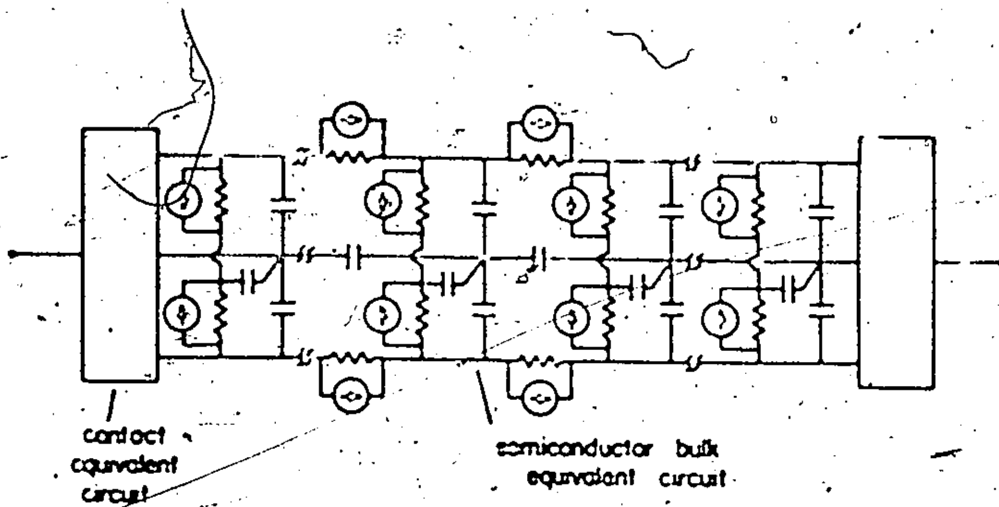


Figure 2.2

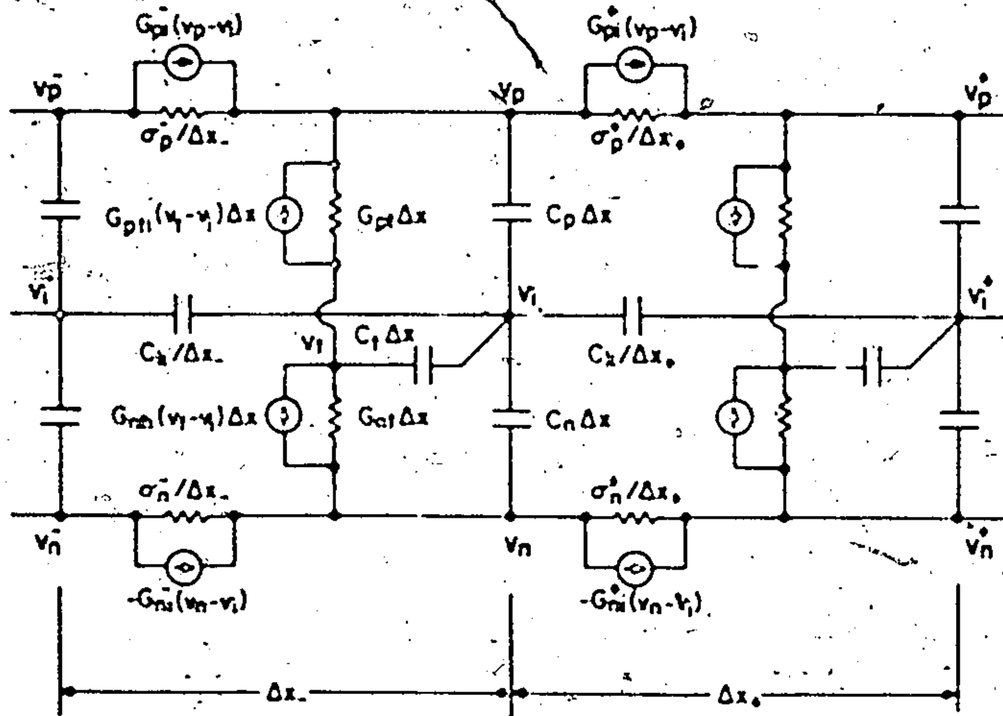


Figure 2.3

TABLE 2.1

Definition of Semiconductor Equivalent Circuit Elements

n, p, n_T, p_T electron, hole densities in bands and SRH level*

N_{TT} total trap density

J_N, J_P electron, hole current densities*

r_{CN}, r_{CP} net capture rates for electrons, holes*

V_N, V_P, V_T, V_I electron, hole, and trapping level quasi-fermi potentials and electrostatic potential*

$C_K = \epsilon$

$C_n = q^2N/kT$ $\quad$ $C_p = q^2P/kT$

$G_{ni} = qJ_N/kT$ $\quad$ $G_{pi} = qJ_P/kT$

$\sigma_n = 1/\rho_n = q\mu_N N$ $\quad$ $\sigma_p = 1/\rho_p = q\mu_p P$

$C_t = q^2N_TP_T/kTN_{TT}$

$G_{nt} = q^2c_nNP_T/kT$ $\quad$ $G_{pt} = q^2c_pPN_T/kT$

$G_{nti} = q^2R_{CN}P_T/kTN_{TT}$ $\quad$ $G_{pti} = q^2R_{CP}N_T/kTN_{TT}$

$R_{CN} = c_n[NN_{TT} - (N + n_1)N_T]$ $\quad$ $R_{CP} = c_p[(P + p_1)N_T - p_1N_{TT}]$

$G_n = C_n/\alpha\Delta t$ $\quad$ $G_p = C_p/\alpha\Delta t$

$G_t = C_t/\alpha\Delta t$ $\quad$ $G_K = C_K/\alpha\Delta t$

$Q_{ei} = C_K(V_I^+ - V_I)/\Delta x_+ - C_K(V_I - V_I^-)/\Delta x_- - q(N - P - N_{IMP})\Delta x$

$N_{IMP} = N_{TT} - N_T$ (donors) $\quad$ $N_{IMP} = -N_T$ (acceptors)

$I_{en} = \sigma_n^+(V_N^+ - V_N)/\Delta x_+ - \sigma_n^-(V_N - V_N^-)/\Delta x_- + qR_{CN}\Delta x$

$I_{ep} = \sigma_p^+(V_P^+ - V_P)/\Delta x_+ - \sigma_p^-(V_P - V_P^-)/\Delta x_- - qR_{CP}\Delta x$

$I_{et} = -q(R_{CN} - R_{CP})\Delta x$ $\quad$ $I'_{ei} = Q_{ei}/\alpha\Delta t$

$I'_{en} = I_{en}(\underline{V}_K) + I_{en}(\underline{S}_K)(1 - \alpha)/\alpha + q[N(t + \Delta t) - n(t)]\Delta x/\alpha\Delta t$

$I'_{ep} = I_{ep}(\underline{V}_K) + I_{ep}(\underline{S}_K)(1 - \alpha)/\alpha - q[P(t + \Delta t) - p(t)]\Delta x/\alpha\Delta t$

$I'_{et} = I_{et}(\underline{V}_K) + I_{et}(\underline{S}_K)(1 - \alpha)/\alpha + q[N_T(t + \Delta t) - n_T(t)]\Delta x/\alpha\Delta t$

---

* These symbols written as capitals represent the DC component of the variable (Fig. 2.3) or the trial value of the variable (Figs. 2.6 and 2.7).

inverted or accumulated. Since these conditions usually occur outside the bias range of interest, Sah's form of the equivalent circuit was used exclusively in this thesis (Figure 2.3). An outline of how this circuit is derived is given in Section 2.5.

2.3 Band Tunnel Currents

The tunneling process between the metal and the semiconductor bands was analyzed using the independent electron approach as outlined by Harrison [34]. Introducing the assumptions that both energy and transverse momentum are conserved during the tunneling transitions between the metal and the semiconductor (elastic, specular tunneling) and employing a WKB approximation results in the following expression for the current flow between the metal and the semiconductor:

$$J = \frac{q}{2\pi^2\hbar}\int dE(f_m - f_s)\int dS e^{-\eta} \qquad (2.16)$$

f_m and f_s are the occupation factors for states in the metal and semiconductor respectively. Defining the "shadow" of a constant energy surface to be its projection in wave number (k) space on a plane parallel to the barrier, the integral over S is over the overlap of the shadows from the metal and the semiconductor for the energy E [34]. This is discussed in more detail for the case of silicon in Appendix A. η is given by

$$\eta = \frac{2}{\hbar}\int_{x_a}^{x_b}(P_{T1}^2 - k_1^2)^{1/2}dx \qquad (2.17)$$

where x is the direction perpendicular to the barrier, x_a and x_b are classical turning points, and P_{Ti} and P_i are the transverse and the total momentum of the tunneling particle in the insulator region. Assuming there exists a parabolic relationship between the transverse energy, E_T, and the transverse momentum, P_T, in the metal, insulator, and semiconductor regions of the device, the integral over S can be replaced by one over the transverse energy E_T (a more general case is discussed in Appendix A).

$$E_{Tm,i,s} = P^2_{Tm,i,s}/2m_{Tm,i,s} \tag{2.18}$$

A two band model of the insulator was employed. In a one band model, the height of the barrier presented by the forbidden energy band in the insulator to a particle of given energy is roughly the energy to the conduction band edge in the insulator. In a two band model, the effect of the valence band in the insulator is taken into account. The necessity of the two band approach has been very simply and convincingly demonstrated by the theoretical work of Schnupp [41-43] employing a simple Kronig-Penney model of the insulator region.

To incorporate two band effects, the energy-momentum relationship in the forbidden band of the insulator must be specified. As this is not known for silicon dioxide, the following empirical relationship originally proposed by Franz [44] and subsequently employed by Stratton [45] in this context is used.

$$P_i^2/2m_i = [E - \phi_{CB}(x)](1 + [E - \phi_{CB}(x)]/E_{gi}) \tag{2.19}$$

where $\phi_{CB}(x)$ represents the energy of the edge of the insulator conduction band. Note that this expression reduces to parabolic relationships for energies near both the valence and the conduction bands of the insulator.

Combining Equations (2.16 - 2.19) and assuming m_{Ti} and m_i are equal gives the following expression for the electron current flow between the metal and the semiconductor conduction band

$$J_{CT} = 4\pi q \frac{m_{Ti}}{h^3} \int_0^{E_{max}} dE(f_m - f_s) \int_0^{\frac{m_{Ts}}{m_{Ti}} E} dE_T e^{-\eta} , \qquad (2.20)$$

where

$$\eta = \frac{2}{\hbar} (2m_{Ti})^{1/2} \int_{x_a}^{x_b} \{E_T - (E - \phi_{CB})[1 + (E - \phi_{CB})/E_{gi}]\}^{1/2} dx, \quad (2.21)$$

and the semiconductor conduction band edge is taken as the energy reference. A similar relationship holds for the current flow density between the metal and the semiconductor valence band.

The assumption of a parabolic relationship between the transverse energy and momentum (2.18) restricts the applicability of the above formulation. For example it is applicable to tunneling into <100> orientated silicon but not <111> silicon (see Appendix A). The integrals involved in evaluating Equations (2.20), (2.21) and related equations were performed numerically [102].

In addition to these tunneling equations, there is an additional expression related to the insulator region describing the applied voltage,

V_a, in terms of the voltage across the insulator, V_{ins}, the semiconductor surface potential, ψ_s, the barrier heights ϕ_{mi} and ϕ_{si}, and the potential separating the bulk fermi level from the conduction band edge in the semiconductor (δ). From Figure 2.1, the following relationship holds

$$V_a = \psi_s + V_{ins} + \phi_{mi} - \phi_{si} - \delta \tag{2.22}$$

The small signal equivalent circuit of the equations described in this section is shown in Figure 2.4 where the element values are defined in Table 2.2. This circuit, with additions to include surface states as described in the next section, is used to terminate the transmission line semiconductor equivalent circuit in the manner suggested in Figure 2.2.

2.4 Surface States

Surface states associated with the IS interface were characterized using SRH kinetics modified to include the effects of tunneling between these states and the metal. This formulation has been described previously by Freeman and Dahlke [35]. For a single level surface state, the differential equation for the occupancy of the state, f_{ss}, is

$$N_{ss} \frac{df_{ss}}{dt} = U_{cn} - U_{cp} + \frac{j_{ss}}{q} \tag{2.23}$$

where N_{ss} is the density of surface states and j_{ss} is the current density tunneling from the metal into the state. U_{cn} is the net rate of capture of electrons in the conduction band. U_{cp} is the corresponding term for holes in the valence band. Equation (2.23) is merely a continuity

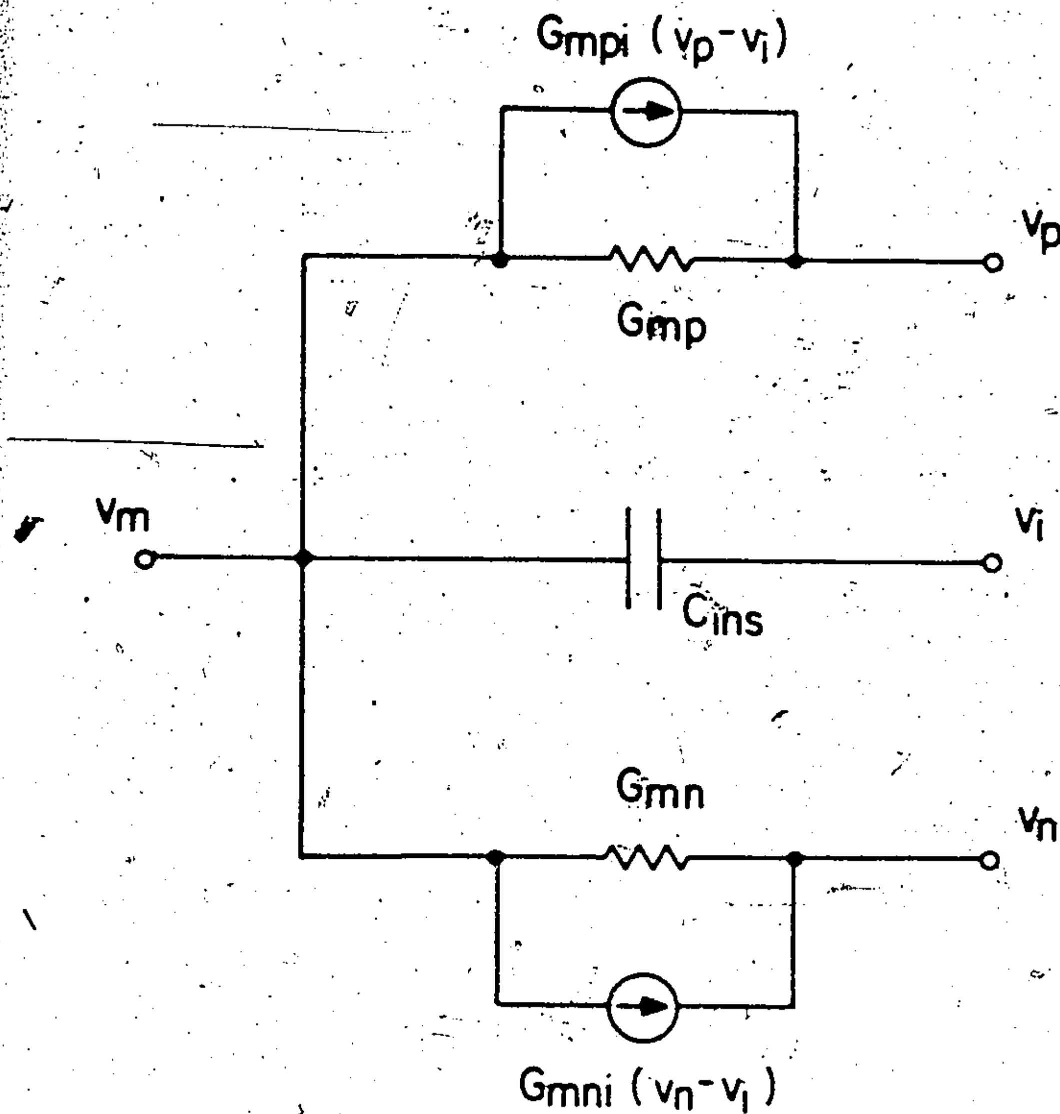


Figure 2.4: Small-signal equivalent circuit describing the insulator region and the band tunneling currents. Element values are given in Table 2.2.

TABLE 2.2: Element Values of Figure 2.4

$$C_{ins} = \epsilon_i / d_{ins}$$

$$G_{mp} = -\partial J_{VT} / \partial V_a$$

$$G_{mpi} = \partial J_{VT} / \partial \psi_s$$

$$G_{mn} = -\partial J_{CT} / \partial V_a$$

$$G_{mni} = \partial J_{CT} / \partial \psi_s$$

expression for the charge entering, leaving, and being stored in the surface state level. U_{cn} and U_{cp} can be expressed in terms of capture and emission terms as in the usual SRH approach [38] to give

$$U_{cn} = S_n[(1 - f_{ss})n_s - f_{ss}n_1] \quad (2.24)$$

$$U_{cp} = S_p[f_{ss}p_s - (1 - f_{ss})p_1] \quad (2.25)$$

where S_n and S_p are surface recombination velocities given by

$$S_{n,p} = \sigma_{n,p}v_{th}N_{ss} \quad (2.26)$$

σ_n and σ_p are capture cross sections [35] and v_{th} is the thermal velocity of electrons. n_1 and p_1 are familiar SRH parameters [38]. n_s and p_s are the electron and hole concentrations in the semiconductor at the IS interface. By using a simple potential well model of a surface state level and the same general independent electron tunneling approach as for the computation of the band tunnel currents (Sections 2.1 and 2.3), Freeman and Dahlke [35] have derived an expression for the surface state tunnel current, j_{ss}. This formulation, although relatively crude, has several features which make it suitable for gaining insight into the effects of surface states in the present work. Besides producing a relatively simple expression for j_{ss}, it is compatible with the band tunnel current formulation having the same general dependence upon the barrier properties. Moreover, the formulation requires the introduction of an effective tunnel capture cross section, σ_T, which in the theoretical sections of the following chapters is varied over wide ranges. In this

way, broad qualitative statements can be made about the effects of surface state tunneling. The expression for J_{ss} as derived by Freeman and Dahlke [35] takes the form

$$J_{ss} = \sigma_T q N_{ss} A_{ss} \exp(-\eta_{ss})(f_m - f_{ss}) \tag{2.27}$$

where η_{ss} is given by Equation (2.21) with E equal to E_{ss} and E_T equal to zero. Including two band effects in the insulator neglected in [35], A_{ss} is given by

$$A_{ss} = \frac{2\hbar}{\pi m} \frac{[\eta'(x_b)]^2 \eta_s'}{[\eta'(x_b) + \eta_s']} \Bigg/ \int_{x_a}^{x_b} \frac{1}{\eta'(x)} dx \tag{2.28}$$

where

$$[\eta'(x)]^2 = \frac{2m_i}{\hbar^2} [\phi_{CB}(x) - E_{ss}](1 + [E_{ss} - \phi_{CB}(x)]/E_{gi}) \tag{2.29}$$

$$[\eta_s']^2 = \frac{2m_s}{\hbar^2} [(E_c - E_{ss})(E_{ss} - E_v)/E_{gs}] \tag{2.30}$$

where E_c and E_v are the energies of the conduction and valence band edges in the semiconductor. Other terms have been introduced in the previous sections. The above equations (2.23 - 2.30) completely specify the static and dynamic behaviour of a single level state. However, some auxiliary expressions which can be derived under DC conditions will prove to be useful in discussing the role of surface states in subsequent chapters. The first is the expression for the occupation probability, f_{ss}, which can be derived in a straightforward manner from Equations (2.23 - 2.25)

under DC conditions;

$$f_{ss} = \frac{n_s S_n + p_1 S_p + j_{ss}/q}{(n_s + n_1)S_n + (p_s + p_1)S_p} \quad (2.31)$$

The second is the expression for the recombination rate U_{cn}, which is obtained simply by substituting (2.31) into (2.24);

$$U_{cn} = \frac{S_n S_p(n_s p_s - n_i^2) - (n_s + n_1)S_n j_{ss}/q}{(n_s + n_1)S_n + (p_s + p_1)S_p} \quad (2.32)$$

with a similar expression for U_{cp}. For small values of j_{ss} this reverts to the conventional surface recombination expression. For larger values of j_{ss}, in addition to its recombination role, the surface state diverts a fraction of the current tunneling into it to either band.

A continuous distribution of surface states was modelled as a large number (20 in the programs of [102]) of distinct surface state levels. Each of these states was assumed to be independent, i.e. it was assumed that each level exchanged carriers more rapidly with the semiconductor bands than with another level. The above theory can then be applied in turn to each level.

An additional relationship associated with surface states is derived by the application of Gauss's law across the IS interface.

$$\epsilon_i F_{ins} - \epsilon_s F_s + q(Q_{ss} + Q_i) = 0 \quad (2.33)$$

F_{ins} and F_s are the fields in the two regions, Q_{ss} is the positive charge contribution from surface states and Q_i is the effective charge at the IS interface representing the fixed charge in the insulator region.

The small signal equivalent circuit corresponding to the equations above takes the form shown in Figure 2.5(a) where the element values are given in Table 2.3. Each independent level introduces an additional node, v_{ss}, into the device equivalent circuit. However, as has been described for the related case of bulk trapping levels [46,47], the equivalent circuit for each level can be transformed into the form shown in Figure 2.5(b) where element values are also given in Table 2.3. The complete equivalent circuit for all the surface state levels can then be obtained simply by summing parallel admittances [47]. By terminating the semiconductor transmission line equivalent circuit with the combination of the surface state and band tunneling equivalent circuits, the resultant small signal equivalent circuit for the MIS tunnel diode is obtained. It will be shown in the following section how such circuits can be applied to the AC, DC, and transient analysis of semiconductor devices.

2.5 Application of the Small Signal Transmission Line Equivalent Circuit to the AC, DC, and Transient Analysis of Semiconductor Devices

2.5.1 Introduction

One important approach to the analysis of semiconductor devices that has been developed in recent years is based on the direct computer solution of the set of differential equations governing carrier transport in semiconductors subject to boundary conditions imposed at the device contacts [48,49]. This approach was first applied to the DC analysis of semiconductor devices by Gummel [48] and it subsequently has been extended to both the large signal transient analysis [49,50,51] and to the small signal AC analysis [49,52]. As noted by several of the above workers,

Figure 2.5(a): Small signal equivalent circuit describing surface states and the tunneling process between the metal and these states. Element values are given in Table 2.3.

(b): Circuit of Figure 2.5(a) transformed to eliminate the node v_{ss}. The values of the transformed elements are also given in Table 2.3.

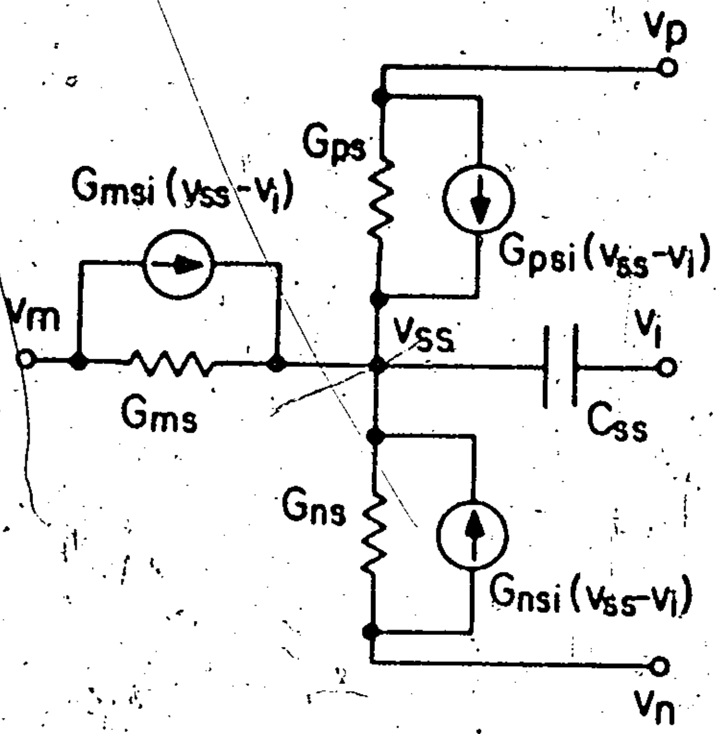


Figure 2.5 (a)

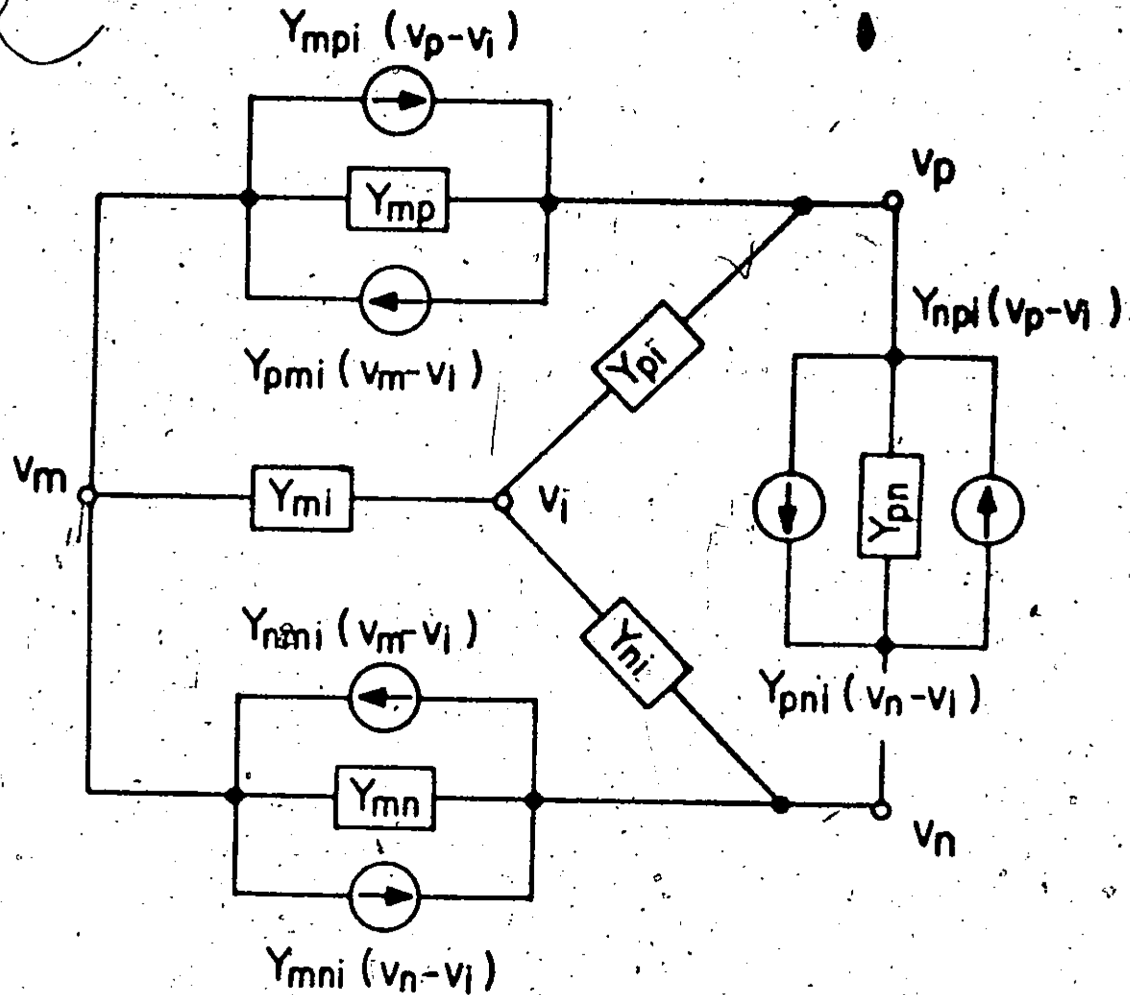


Figure 2.5 (b)

TABLE 2.3: Element Values for Figure 2.5

(a)

$$G_{ms} = -\partial J_{ST}/\partial V_a$$

$$G_{msi} = \partial J_{ST}/\partial \psi_s$$

$$C_{ss} = q^2 N_{ss} f_{ss}(1 - f_{ss})/kT$$

$$G_{ps} = q^2 S_p P_s f_{ss}/kT$$

$$G_{psi} = q^2 U_{CP} f_{ss}/kT$$

$$G_{ns} = q^2 S_n N_s (1 - f_{ss})/kT$$

$$G_{nsi} = q^2 U_{CN}(1 - f_{ss})/kT$$

$$U_{CN} = S_n(N_s - (N_s + n_1)f_{ss})$$

$$U_{CP} = S_p((P_s + p_1)f_{ss} - p_1)$$

(b) $\Sigma G = G_{ms} + G_{ps} + G_{ns} + sC_{ss} - G_{msi} - G_{psi} - G_{nsi}$

$$Y_{mi} = sC_{ss}G_{ms}/\Sigma G \qquad Y_{mp} = G_{ms}G_{ps}/\Sigma G$$

$$Y_{pi} = sC_{ss}G_{ps}/\Sigma G \qquad Y_{mpi} = G_{msi}G_{ps}/\Sigma G$$

$$Y_{ni} = sC_{ss}G_{ns}/\Sigma G \qquad Y_{pmi} = G_{ms}G_{psi}/\Sigma G$$

$$Y_{pn} = G_{ps}G_{ns}/\Sigma G \qquad Y_{mn} = G_{ms}G_{ns}/\Sigma G$$

$$Y_{pni} = G_{psi}G_{ns}/\Sigma G \qquad Y_{mni} = G_{msi}G_{ns}/\Sigma G$$

$$Y_{npi} = G_{ps}G_{nsi}/\Sigma G \qquad Y_{nmi} = G_{ms}G_{nsi}/\Sigma G$$

the solution of the coupled set of non-linear partial differential equations presents a complex problem involving several preliminary stages of manipulation of the equations before they reach a form which can be treated numerically.

A reduction in the complexity of the problem was achieved by the work of Ohtsuki and Kani [53] and Sah [23] who showed that the set of governing differential equations can be modelled by a large signal transmission line equivalent circuit involving non-linear resistive and capacitative elements. Although transmission line analogues had been applied to the analysis of semiconductor devices for some time prior to this development [54], this non-linear transmission line was unique in that it corresponded exactly to the complete set of governing transport equations. Effects such as those due to charge storage, carrier recombination and generation, carrier drift and diffusion, and displacement current flow are all included into this compact model. This large signal equivalent model can be solved using network analysis methods appropriate to non-linear circuits. Kani and Yokota [55] have solved it under DC conditions using state variable techniques. Transient solutions are also possible using this method.

A transmission line equivalent circuit model for the semiconductor which has found wider use than the above non-linear circuit is the small signal circuit developed by Sah [21-23]. This equivalent circuit (Figure 2.2) involves only linear resistors, capacitors, and voltage controlled current sources but, in the form derived by Sah, is valid only for small signal variations about a DC solution. It is ideally suited to the

analysis of the small signal AC properties of semiconductor devices. Not only can it be used in conjunction with a computer program to obtain precise solutions for these AC properties [25] but, under a number of circumstances, simplifications in the circuit topology can be made allowing analytical solutions to be obtained [22]. It represents a very significant contribution to the techniques available for the small signal AC analysis of semiconductor devices. However, since a DC solution is required at the bias point where the AC properties are being computed, its use has been restricted largely to thermal equilibrium or quasi-thermal equilibrium situations where the DC solution is relatively easy to obtain by other methods [24-26].

The following sections show that the use of the small signal transmission line equivalent circuit is not restricted solely to the small signal AC analysis of semiconductor devices but that it can be used for DC analysis and for large signal transient analysis. Used with a digital computer, it can provide solutions as accurately and as efficiently as any of the methods previously mentioned. Moreover, instead of requiring elaborate numerical methods, the solution procedure at its most complex reduces to the successive solution of the potential distribution along a linear network. The technique is also well suited for the computation of the properties of semiconductor devices when embedded in external circuitry.

The equivalent circuit for the MIS tunnel diode has the form shown in Figure 2.2. The network labelled "contact equivalent circuit" corresponds to the band tunneling and surface state equivalent circuits

(Figures 2.4 and 2.5). The unlabelled network corresponds to the back contact which was modelled as an "ideal ohmic" contact. This meant that this network was simply a short circuit across the branches of the transmission line [26,56]. The tunneling and surface state network is considerably more complex. Reduced to its simplest form by combining the band tunneling and surface state equivalent circuits, it will have the same topology as that of Figure 2.5(b). The tunneling process asserts its control over the device properties by the values of the elements of this network. For the formulation described in the previous sections, more computational effort is required to calculate these values than to calculate those of the hundreds of elements associated with the semiconductor equivalent circuit.

From a topological viewpoint, the band tunneling and surface state network serves to terminate the semiconductor transmission line. The equivalent circuits of other semiconductor devices such as p-n junction diodes [26,27] and thick insulator MIS diodes [24,25] also have this same general form. The algorithms to be described will apply to the typical internal regions of this semiconductor equivalent circuit. The small modifications required in the region adjacent to the terminating network are straightforward but will vary for different semiconductor devices [102].

As mentioned in Section 2.2, the governing differential equations in the semiconductor region, (2.1) - (2.4), can be expressed entirely in terms of the electrostatic and quasi-fermi potentials, v_I, v_N, v_P, and v_T, as dependent variables. The spatial derivatives can be eliminated from these equations by dividing the semiconductor region under analysis

into a large number of lumps or sections. This allows finite difference approximations to be made in the normal way [57]. At each lump or section, there will be a set of ordinary differential equations corresponding to the original set of partial differential equations (2.1 - 2.4). The former will be expressed in the symbolic form

$$F_V(\underline{v}_K) = 0 \qquad (2.34)$$

$$F_N(\underline{v}_K) = \dot{G}_N(\underline{v}_K) \qquad (2.35)$$

$$F_P(\underline{v}_K) = \dot{G}_P(\underline{v}_K) \qquad (2.36)$$

$$F_T(\underline{v}_K) = \dot{G}_T(\underline{v}_K) \qquad (2.37)$$

where $\underline{v}_K$ represents the values of the potentials at the lump under consideration and at the adjacent lumps. The functions $F_{V,N,P,T}$ and $G_{N,P,T}$ are non-linear functions of these potentials. The dot represents differentiation with respect to time. Having expressed the lumped equations in this symbolic form, it is now possible to show quite simply the relationships between the AC, DC, and transient analyses.

2.5.2 Small Signal AC Analysis

If a small signal disturbance is applied to a semiconductor device under DC bias, the potentials can be expressed in the form

$$\underline{v}_K = \underline{V}_K + \underline{v}_K' \qquad (2.38)$$

where $\underline{V}_K$ represents the DC values of the potentials and is therefore time

independent. $\underline{v}_k$ is the small signal variation about $\underline{V}_K$. Moreover we have

$$F_V(\underline{V}_K) = F_N(\underline{V}_K) = F_P(\underline{V}_K) = F_T(\underline{V}_K) = 0 \tag{2.39}$$

and

$$\dot{G}_N(\underline{V}_K) = \dot{G}_P(\underline{V}_K) = \dot{G}_T(\underline{V}_K) \equiv 0 \tag{2.40}$$

Substituting (2.38) into (2.34) - (2.37) and neglecting terms higher than first order in $\underline{v}_k$ (small signal approximation) gives

$$F_V(\underline{V}_K) + f_V(\underline{V}_K, \underline{v}_k) = 0 \tag{2.41}$$

$$F_N(\underline{V}_K) + f_N(\underline{V}_K, \underline{v}_k) = \dot{G}_N(\underline{V}_K) + \dot{g}_N(\underline{V}_K, \underline{v}_k) \tag{2.42}$$

$$F_P(\underline{V}_K) + f_P(\underline{V}_K, \underline{v}_k) = \dot{G}_P(\underline{V}_K) + \dot{g}_P(\underline{V}_K, \underline{v}_k) \tag{2.43}$$

$$F_T(\underline{V}_K) + f_T(\underline{V}_K, \underline{v}_k) = \dot{G}_T(\underline{V}_K) + \dot{g}_T(\underline{V}_K, \underline{v}_k) \tag{2.44}$$

where the $f_{V,N,P,T}$ and the $g_{N,P,T}$ are linear functions of $\underline{v}_k$. Because of (2.39 - 2.40) and the fact that the $g_{N,P,T}$ are linear in $\underline{v}_k$, Equations (2.41 - 2.44) reduce to

$$f_V(\underline{V}_K, \underline{v}_k) = 0 \tag{2.45}$$

$$f_N(\underline{V}_K, \underline{v}_k) = g_N(\underline{V}_K, \dot{\underline{v}}_k) \tag{2.46}$$

$$f_P(\underline{V}_K, \underline{v}_k) = g_P(\underline{V}_K, \dot{\underline{v}}_k) \tag{2.47}$$

$$f_T(\underline{V}_K, \underline{v}_k) = g_T(\underline{V}_K, \dot{\underline{v}}_k) \tag{2.48}$$

If the DC potentials V_K are known, the above are linear ordinary differential equations for the small signal potentials v_k. The equations are at the stage of development where they can be represented by a linear equivalent circuit. A more detailed derivation of the above equations has been given by Sah [22] who showed that, for non-degenerate material with a Shockley-Read-Hall model of recombination processes, the basic section of the transmission line equivalent circuit takes the form shown in Figure 2.3. Expressions for the element values are given in Table 2.1. More general forms have also been derived for degenerate semiconductor [36] and for different recombination and generation mechanisms [21,23].

As has been mentioned, the small signal transmission line equivalent circuit previously has been applied to the accurate computation of the small signal AC properties of semiconductor devices [25,26]. Once a DC solution is known at a given bias point (computed as in the next section or by any other method), the element values in the circuit can be calculated (Table 2.1). The device AC properties can then be obtained by computing the potential distribution along the linear equivalent circuit. An efficient method [56] for calculating this distribution taking full account of the specialized form of the circuit has been developed during this thesis work (Appendix B).

2.5.3 DC Solutions

DC solutions for the properties of semiconductor devices can be obtained using the small signal transmission line equivalent circuit. The method to be described is based on the standard mathematical technique for solving non-linear equations, Newton-Raphson iteration [52]. Suppose

there exists a trial DC solution which is close to the correct solution of the DC equations. A method for generating these trial solutions is described later. Again we will write the equation

$$\underline{v}_K = \underline{V}_K^{\delta} + \underline{v}_k \tag{2.38}$$

but will identify $\underline{V}_K$ as the trial solution, $\underline{v}_K$ as the correct solution, and $\underline{v}_k$ as the difference between them which is reasonably small. All time derivatives in equations (2.34 - 2.37) are zero. $\underline{v}_K$ therefore satisfies the relationships

$$F_V(\underline{v}_K) = F_N(\underline{v}_K) = F_P(\underline{v}_K) = F_T(\underline{v}_K) \equiv 0 \tag{2.49}$$

Substituting (2.38) into (2.49) and making a small signal expansion gives equations of the form

$$F_V(\underline{V}_K) + f_V(\underline{V}_K,\underline{v}_k) = 0 \tag{2.50}$$

i.e.

$$f_V(\underline{V}_K,\underline{v}_k) = -F_V(\underline{V}_K) \tag{2.51}$$

$$f_N(\underline{V}_K,\underline{v}_k) = -F_N(\underline{V}_K) \tag{2.52}$$

$$f_P(\underline{V}_K,\underline{v}_k) = -F_P(\underline{V}_K) \tag{2.53}$$

$$f_T(\underline{V}_K,\underline{v}_k) = -F_T(\underline{V}_K) \tag{2.54}$$

The terms on the left hand sides of (2.51 - 2.54) are exactly the same as those of the LHS of the small signal equations (2.45 - 2.48) which gave the equivalent circuit of Figure 2.3. The terms on the RHS of equations

Figure 2.6: A section of the semiconductor small-signal equivalent circuit including modifications necessary to employ the circuit for DC analysis. Element values are given in Table 2.1.

Figure 2.7: A section of the semiconductor small-signal transmission line in the form suitable for the transient analysis of semiconductor devices. Element values are given in Table 2.1.

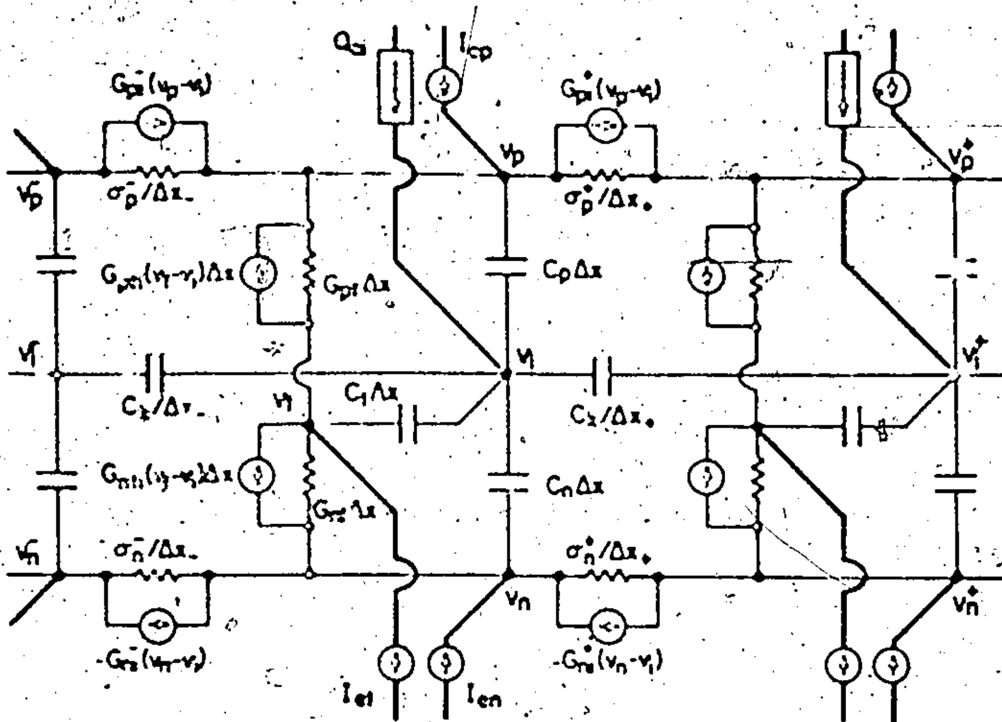


Figure 2.6

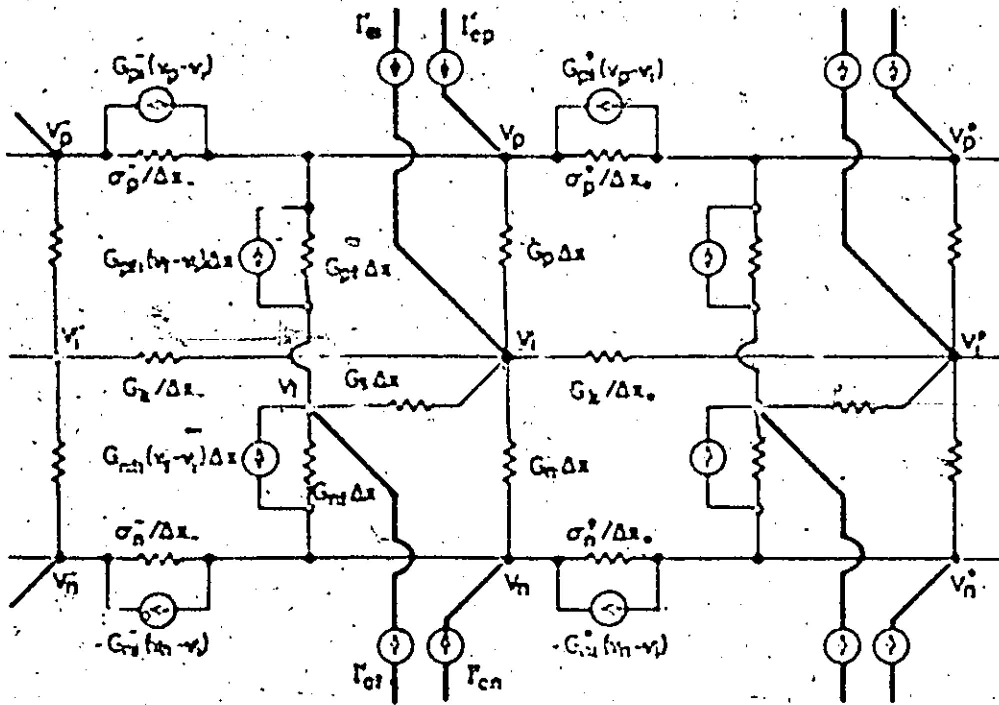


Figure 2.7

(2.51 - 2.54) are measures of the error in the trial solution, $\underline{V}_K$, since they would be zero if $\underline{V}_K$ was the correct solution. The equivalent circuit corresponding to Equations (2.51 - 2.54) is shown in Figure 2.6. Not surprisingly in view of the above, it is identical to the AC equivalent circuit but has an additional current or charge source at each node. These sources will be termed "error" sources for the reason indicated. Expressions for them are also given in Table 2.1. Suppose a trial DC solution at a given bias point is available. This is used to compute the values of the equivalent circuit elements, including the error sources, which represent a measure of the amount the trial solution deviates from the correct DC solution. The potential distribution along the transmission line under DC conditions is then calculated (using current conservation in the current branches, and charge conservation in the capacitative branch). A straightforward extension of the method of [56] described in Appendix B has proved very efficient for this purpose. This potential distribution is then added to the trial solution.

Provided the trial solution is not too poor, the new solution based on it should be closer to the correct DC solution. If the new solution is still not sufficiently close to the correct solution (as measured by the values of the error sources or by the potential changes between it and the original trial solution), it, in turn, becomes the trial solution and the above procedure is repeated until a solution as close to the correct solution as desired is obtained. When this solution is achieved, the error sources drop out and the equivalent circuit of Figure 2.6 reduces to exactly the same form as the small signal AC

equivalent circuit of Figure 2.3.

There exists a simple method for obtaining "trial" solutions as employed by Gokhale [52]. The solution under zero bias (thermal equilibrium) conditions is obtained relatively simply. This is a good trial solution for a bias point near zero, allowing a solution to be obtained at such a bias point. This solution is then a good trial solution for a bias point further removed from zero bias. Proceeding in this way, the solution at any desired bias point can be found.

2.5.4 Large Signal Transient Analysis

The transient analysis is performed using a method similar to that used for the DC analysis. In addition to using the spatial finite difference approximations included into Sah's circuit, similar approximations will be made in the time domain. The time derivative in Equation (2.35), for example, is replaced by a finite difference expression according to the general integration scheme [57]

$$\alpha F_N[\underline{v}_K(t+\Delta t)] + (1-\alpha)F_N[\underline{v}_K(t)] = \{G_N[\underline{v}_K(t+\Delta t)] - G_N[\underline{v}_K(t)]\}/\Delta t \tag{2.55}$$

where α has a value between 0 and 1. For α equal to 0, 1/2, and 1, this general scheme reduces to the explicit, Crank-Nicholson, and implicit schemes respectively [57]. The stability of the integration scheme increases as α increases from 0 to 1, while maximum accuracy is achieved for α equal to 1/2. Values of α of 1/2 and 1 are therefore the most useful.

Again we write the equation

$$\underline{V}_K = \underline{V}_K + \underline{v}_k \tag{2.38}$$

This time $\underline{v}_K$ is identified as the solution for the potentials at time $t + \Delta t$, $\underline{V}_K$ as the trial solution for these potentials at this time, and $\underline{v}_k$ as the reasonably small difference between them. Also the potentials $\underline{s}_K$ are defined as being the values of the potentials at time t. Substituting (2.38) into (2.55) and making small signal approximations gives

$$\begin{aligned} \alpha F_N(\underline{V}_K) + \alpha f_N(\underline{V}_K, \underline{v}_k) + (1 - \alpha) F_N'(\underline{s}_K) \\ = [G_N(\underline{V}_K) + g_N(\underline{V}_K, \underline{v}_k) - G_N(\underline{s}_K)]/\Delta t \end{aligned} \tag{2.56}$$

Thus Equations (2.34 - 2.37) take the form

$$f_V(\underline{V}_K, \underline{v}_k) = -F_V(\underline{V}_K) \tag{2.57}$$

$$\begin{aligned} f_N(\underline{V}_K, \underline{v}_k) = g_N(\underline{V}_K, \underline{v}_k)/\alpha\Delta t + \{-F_N(\underline{V}_K) - F_N(\underline{s}_K)(1 - \alpha)/\alpha \\ + [G_N(\underline{V}_K) - G_N(\underline{s}_K)]/\alpha\Delta t\} \end{aligned} \tag{2.58}$$

$$\begin{aligned} f_P(\underline{V}_K, \underline{v}_k) = g_P(\underline{V}_K, \underline{v}_k)/\alpha\Delta t + \{-F_P(\underline{V}_K) - F_P(\underline{s}_K)(1 - \alpha)/\alpha \\ + [G_P(\underline{V}_K) - G_P(\underline{s}_K)]/\alpha\Delta t\} \end{aligned} \tag{2.59}$$

$$\begin{aligned} f_T(\underline{V}_K, \underline{v}_k) = g_T(\underline{V}_K, \underline{v}_k)/\alpha\Delta t + \{-F_T(\underline{V}_K) - F_T(\underline{s}_K)(1 - \alpha)/\alpha \\ + [G_T(\underline{V}_K) - G_T(\underline{s}_K)]/\alpha\Delta t\} \end{aligned} \tag{2.60}$$

The terms on the LHS are again identical to those appearing on the LHS of the Equations (2.45 - 2.48) which gave the equivalent circuit

of Figure 2.3. The terms in $g_{N,P,T}$ are also closely related for both sets of equations. The additional terms on the RHS of (2.57 - 2.60) again represent a measure of the error in the trial solution $\underline{V}_K$.

The circuit corresponding to Equations (2.57 - 2.60) takes the form shown in Figure 2.7. This is identical to the circuit of Figure 2.3 except that capacitors appearing in the former circuit have been replaced by resistors and "error" current sources have been introduced at each node. As can be seen from Table 2.1, the value of the resistor is closely related to the value of the original capacitor. Knowing a solution for the device properties at time t, this circuit allows the solution at time $t + \Delta t$ to be found.

The technique employed is similar to that of the previous section. Using a trial value for the desired solution, all the element values including "error" sources are calculated. The potential distribution along the transmission line is then computed and added to the trial value of the potentials. This procedure is then iterated as in the previous section until solutions at time $t + \Delta t$ are known to the desired accuracy. The solution at time t provides an excellent starting value for the trial solution at time $t + \Delta t$.

The accuracy, stability, and efficiency of the transient algorithm depend critically upon the time step, Δt. For greatest efficiency, this step size should be changed during the evolution of the transient solution. The magnitude of the potential changes during previous time steps are convenient parameters for automatically adjusting Δt. The transient algorithm can also be employed solely to compute DC solutions at arbitrary bias

points provided a DC solution is known at one bias point such as zero bias. In this context, it is of interest to note that if Δt is set to a large number and α is given the value 1, the equations (2.57 - 2.60) approach those derived for the DC analysis (2.51 - 2.54) and the transient algorithm becomes identical to the Newton-Raphson iteration method described in Section 2.5.3.

2.5.5 Performance of Algorithm

To demonstrate the previous techniques, a silicon N^+P junction diode was analyzed under AC, DC, and transient conditions. The results are described elsewhere [27]. The computation times for the MIS tunnel diodes currently of interest were larger than those required for the junction diode because of the extra time required to evaluate the tunneling integrals of Sections 2.3 and 2.4. To obtain a DC solution at a bias point 0.1 V distant from a known DC solution required about 12 seconds computation time using a CDC 6400 computer, while an AC solution required about 4 seconds. No transient solutions were computed for the MIS diode. To ensure that negligible error was introduced by replacing the partial differential equations of Section 2.2.1 by the lumped equations of Section 2.5.1, solutions obtained with a given number of lumps were compared to those obtained with exactly twice the number of lumps. In this way, it was estimated that about 100 lumps were required to give solutions for the terminal properties which deviated by less than 0.5% from the exact solutions to the governing system of equations described in the preceding sections.

As mentioned in Section 2.1, an alternate method for solving these

equations under DC conditions was also developed during this thesis work [102]. The method lacked the conceptual simplicity of the transmission line approach and was not as well suited for calculating AC properties. However, it was ideal for obtaining DC solutions at isolated bias points, typically requiring between 5 and 40 seconds computation time depending upon the familiarity of the user with the MIS diode behaviour. Since both programs solved the same system of equations to a similar high degree of accuracy using widely different techniques, it was comforting to observe that they gave near identical solutions for a given set of operating conditions [102].

2.6 Additional Topics

2.6.1 Multi-Dimensional Effects

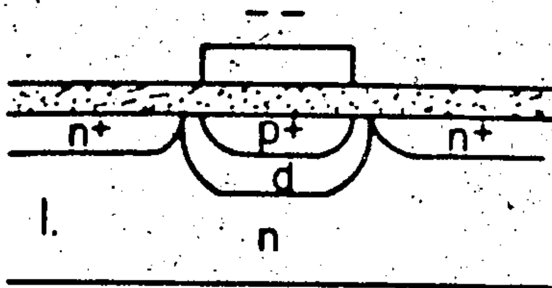
In the above characterization, the device properties have been assumed uniform in the y and z directions to enable the equations to be formulated in terms of one spatial variable, x. However, experimental devices can only approach this idealized one-dimensional situation. In this section, some likely effects of this multi-dimensional nature of experimental devices will be described.

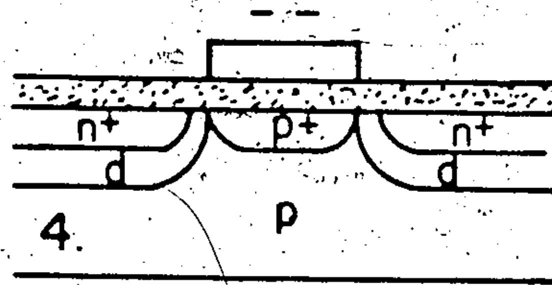
Perhaps the most striking effect for the present devices will be due to insulator non-uniformities. In addition to the surface potential variations due to spatial non-uniformities in insulator charge and surface state distributions as in conventional MIS devices [59], more important effects will be caused by inevitable non-uniformities in the insulator thickness. Since tunnel currents depend exponentially upon this thickness, these non-uniformities will be greatly magnified in the distribution of

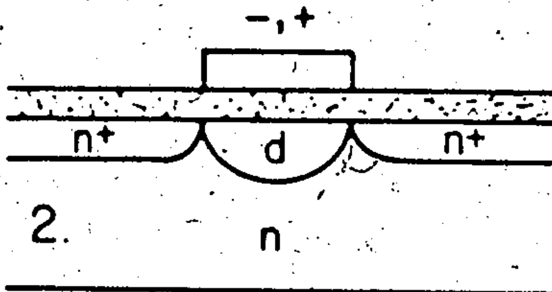
the current flow between the metal and the semiconductor. One is forced to accept the picture of most of the device current occurring at small areas distributed randomly across the device cross section. From a tunneling point of view, the device will be equivalent to an idealized one with a thinner insulator thickness than the average thickness of the actual device and of smaller cross sectional area [60]. This effect is incorporated into the tunnel algorithm as described in the next section. For the thinner insulator values where transport through the semiconductor becomes important in determining the device properties, these current density non-uniformities will tend to be washed out. The insulator non-uniformities may also cause the failure of the assumption of specular transmission employed in deriving the band tunnel equations [33b,34,61]. The consequences of such a failure have been discussed by Stratton [61].

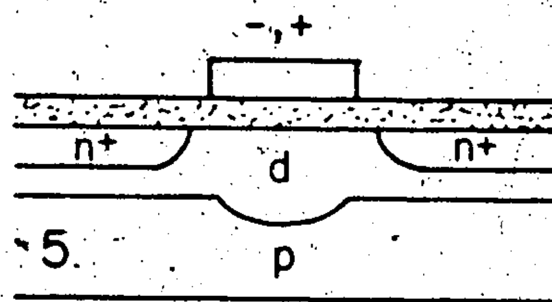
The other multi-dimensional effect to be discussed will be referred to as the "lateral" effect. It will be discussed specifically for the M-SiO_2-Si system and the fabrication procedures employed in this work (Section 4.1). After the growth of the oxide layer, it is known that electrons are attracted to the IS interface, presumably due to the presence of positive charges in the oxide. A metal contact deposited on top of this layer allows charge conditions directly under the contact to be varied. This results in the six possible lateral configurations shown in Figure 2.8 of which "2" and "6" are known to influence the properties of Schottky [62] and conventional MIS [63] diodes, respectively. These effects can be controlled using a guard ring [62], but this involves additional processing and complicates measurements. Whenever lateral

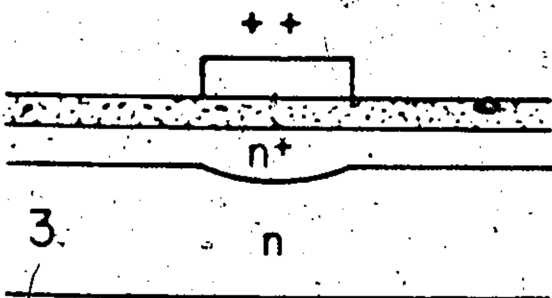
Figure 2.8: Schematic showing the six possible lateral distributions of charge in the semiconductor along the IS interface for both n and p type devices.

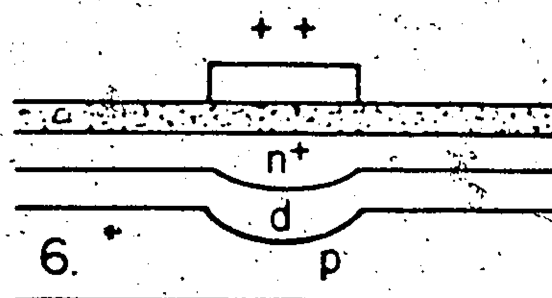












N type substrates   P type substrates

d — depleted

n+ — large electron density

p+ — large hole density

effects were found to be hampering the interpretation of experimental measurements upon the MIS tunnel diode, the device was subject to one of the simple chemical treatments known to reduce the tendency for electrons to congregate at the IS interface [64].

2.6.2 Parameter Values

The parameters appearing in the equations characterizing the semiconductor can be assigned values with some confidence. The opposite is true of those appearing in the characterization of surface states. Attempts to measure them experimentally have resulted in values varying over several orders of magnitude [17]. For this reason, surface-state effects will be discussed only qualitatively in the following chapters. In the band current tunneling equations, the barrier heights are assigned the same values as have been measured upon thick insulator devices [65]. The parameters remaining to be assigned values are the effective mass parameters and the barrier thickness, d. For simplicity, all effective masses were assumed equal. d was assigned the value most accessible experimentally, i.e. the value computed from the accumulation capacitance of the device. The insulator effective mass was then determined on the basis of accumulated experimental results. For the M-SiO_2-Si system, a rule of thumb is that the current flow at a given bias changes by an order of magnitude for a 2Å change in d [16]. Assigning the insulator effective mass a value equal to 0.65 times the free electron mass simulated this behaviour. The right was then reserved to adjust the magnitudes of the band tunnel currents using an effective area parameter. However, it was not found necessary to use this parameter to give the desired

qualitative fit between theoretical and experimental current magnitudes for a given insulator thickness.

Values assigned to the semiconductor, surface state, and tunneling parameters throughout this thesis are listed in Table 2.4.

TABLE 2.4: Numerical Values Assigned to the Parameters of the Metal-Silicon Dioxide-Silicon System.

$\phi_{si} = 3.2$ eV

$E_{gi} = 8.0$ eV

$E_{gs} = 1.1$ eV

$$n_i = \begin{cases} 1.5 \times 10^{16} m^{-3} (300°K) \\ 2.0 \times 10^{11} m^{-3} (200°K) \end{cases}$$

$N_c/N_v = 1.73$

$\epsilon_i = 3.82\ \epsilon_o$

$\epsilon_s = 11.7\ \epsilon_o$

$$\mu_n = \begin{cases} 0.13\ m^2 V^{-1} s^{-1} (300°K) \\ 0.40\ m^2 V^{-1} s^{-1} (200°K) \\ 0.07\ m^2 V^{-1} s^{-1} (0.1\ \Omega cm\ Si) \end{cases}$$

$$\mu_p = \begin{cases} 0.045\ m^2 V^{-1} s^{-1} (300°K) \\ 0.150\ m^2 V^{-1} s^{-1} (200°K) \\ 0.035\ m^2 V^{-1} s^{-1} (0.1\ \Omega cm\ Si) \end{cases}$$

$$t_{no}, t_{po} = \begin{cases} 10^{-5} s \\ 10^{-6} s\ (0.1\ \Omega cm\ Si) \end{cases}$$

$$n_1 = \begin{cases} n_i \\ 21\ n_i\ (\text{Figure 6.8}) \end{cases}$$

$$p_1 = \begin{cases} n_i \\ n_i/21\ (\text{Figure 6.8}) \end{cases}$$

TABLE 2.4 (continued)

$m_i, m_{Ti}, m_{Ts} = 0.65\ m_o$

$\sigma_{co} = 10^{-19} m^2$

$\sigma_{To} = 4 \times 10^{-21} m^2$

$v_{th} = \begin{cases} 10^5 (300°K) \\ 8 \times 10^4 (200°K) \end{cases}$

$L = \begin{cases} 2.5 \times 10^{-4} m \text{ (Chapters 3,6)} \\ 1.0 \times 10^{-4} m \text{ (Chapter 5)} \end{cases}$

CHAPTER 3

THEORY OF THE MIS TUNNEL DIODE

3.1 Preliminary Discussion

A schematic energy band diagram of the MIS tunnel diode has been given in Figure 2.1. The insulator layer is thin (< 60Å) so that the currents shown can flow between the metal and the semiconductor by tunneling processes. Shown is the case of a p-type semiconductor biased positively with respect to the top metal contact by a voltage V_a. Under this bias condition, there are net current flows J_{CT} and J_{VT} from the conduction and valence bands of the semiconductor to the metal contact due to electron tunneling transitions between states in the metal and these bands. Similarly a net current J_{ST} flows to the metal from surface states at the IS interface. J_{CI} and J_{VI} are effective current flows due to the interchange of charge between the conduction and valence bands of the semiconductor and these states by recombination-generation processes.

In this thesis, a diode is classified as a majority carrier, surface-state, or minority carrier diode depending on whether the dominant component of the diode current near zero bias flows between the metal and the majority carrier energy band in the semiconductor, between the metal and surface state levels, or between the metal and the minority carrier band, i.e. depending upon which of J_{VT}, J_{ST}, and J_{CT} is dominant near zero bias. In addition, a distinction is made

between equilibrium and non-equilibrium diodes. Equilibrium diodes are those with a relatively thick insulator layer ($d_{cr} < d < 60\text{Å}$) where the current flows are so small in the bias range of interest (usually ± 2.0 volts for the Metal-SiO_2-Si system) that they can easily pass through the semiconductor. The semiconductor remains essentially in thermal equilibrium and the electron and hole concentrations can be described by a single fermi level which, to a high degree of approximation, is constant throughout the semiconductor region. Non-equilibrium diodes have thinner insulator layers ($d < d_{cr}$). Consequently current flows disturb the semiconductor from thermal equilibrium and electron and hole quasi-fermi levels which vary throughout the semiconductor region are required to describe the carrier distributions. Non-equilibrium behaviour can be obtained in conventional thick insulator MIS diodes if they are subject to very high bias [66], or under transient conditions as exploited in charge transfer devices [67]. However, in this thesis, the term non-equilibrium will be used solely to describe effects resulting from DC current flow within approximately one to two volts of zero bias. The utility of subjecting MIS tunnel diodes to the above classification system will readily become apparent in this and the subsequent chapters.

In the following sections, the device characteristics will be evaluated initially neglecting surface states. Their effects upon this idealized description will then be evaluated in Section 3.4. This approach was chosen for two reasons. Firstly, surface states represent defects and the defect free situation is of obvious intrinsic interest. Continued work with these thin insulator diodes is likely to result

in drastic reductions in surface state densities as has been observed in conventional MIS devices over the last decade. Experimental devices should approach the idealized surface state free case more and more closely. Secondly, surface state distributions and capture cross sections are known to vary widely in experimental devices [17]. This thesis therefore is concerned with a qualitative discussion of their effects using the idealized surface state free case as a frame of reference.

Attention will be focussed upon the Metal-SiO_2-Si system since this is the system which is presently the most suitable for experimental investigation. Its theoretical properties will be discussed using the results of the simulation technique described in Chapter 2 and implemented into the computer programs described in reference [102]. For this system, the equilibrium to non-equilibrium transition occurs at a value of insulator thickness, d_{cr}, approximately equal to 30Å at room temperature. For values of insulator thickness larger than this, the semiconductor region remains essentially in thermal equilibrium. The theoretical value of d_{cr} can be determined from the numerical simulation by observing the behaviour of the semiconductor electron and hole quasi-fermi potentials. If, over the bias range of interest, these are virtually constant and equal to each other throughout the entire semiconductor region, the device is an equilibrium diode. Once variations in these potentials become comparable to kT/q (26 mV at 300°K), the device is a non-equilibrium diode. Initially the properties of the simpler equilibrium devices will be discussed to gain some insight into the MIS tunnel diode behaviour. The insulator thickness is then reduced

and the properties of the resulting non-equilibrium devices are then discussed in terms of those established for the equilibrium case.

The numerical simulation allows both the small signal AC and the DC properties of the MIS tunnel diode to be obtained [102]. The terminal properties of most interest in this chapter are the DC current-voltage (I-V) characteristics and the capacitance-voltage characteristics (C-V). The device capacitance is evaluated from the imaginary part of the diode admittance [102]. The diode conductance is also available at the same time since it is merely the real part of the diode admittance. Although conductance measurements are important in determining the surface state properties of experimental devices [17], the conductance characteristics of the devices discussed in the following sections generally will not be described. The current and capacitance characteristics are of far more use in accomplishing the task of the present chapter which is to establish a broad theoretical framework for interpreting the properties of MIS tunnel diodes.

3.2 Equilibrium Diodes

3.2.1 Capacitance-Voltage Characteristics

The energy band diagrams at zero bias are shown in Figure 3.1 for two diodes differing only in the value of metal-insulator barrier height, ϕ_{mi}. The diodes have 34Å of SiO_2 separating the metal contact from a p-type silicon substrate of dopant density $7 \times 10^{21} m^{-3}$. The major difference between the two cases is that the IS interface is inverted for the lower value of ϕ_{mi} (3.2 eV) and accumulated for the higher (4.2 eV). This difference will be shown to be of fundamental

Figure 3.1: Schematic energy band diagrams at zero bias for MIS tunnel diodes with a 34Å layer of silicon dioxide separating the metal from a 2 Ωcm p type silicon substrate ($N_A = 7 \times 10^{21}\ m^{-3}$) of <100> orientation for two values of the metal to insulator barrier height, ϕ_{mi}.

(a): $\phi_{mi} = 3.2$ eV.

(b): $\phi_{mi} = 4.2$ eV

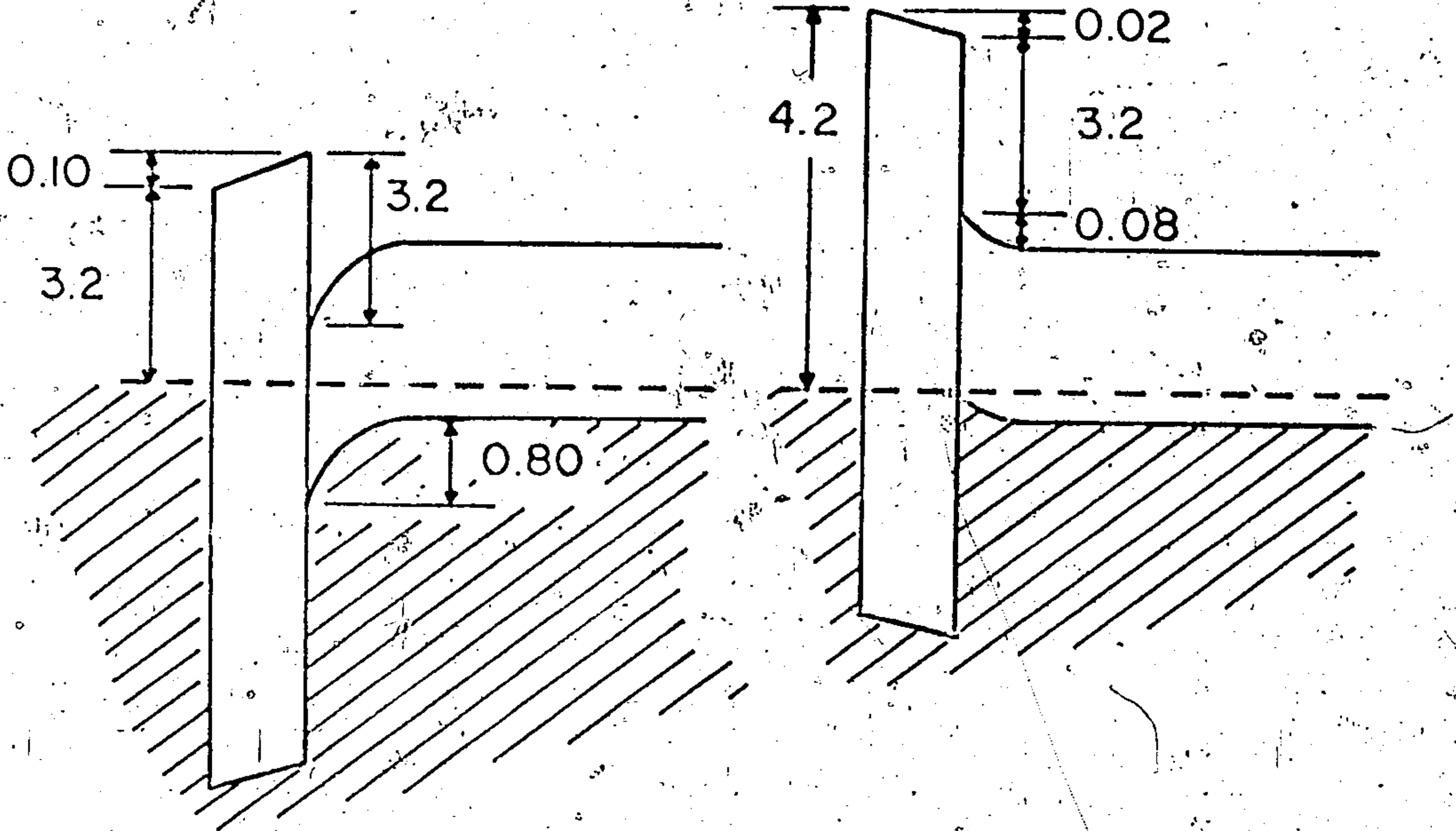


FIGURE 3.1

importance.

As might be expected, the capacitance-voltage characteristics of these equilibrium diodes do not differ significantly from those of conventional MIS capacitors [1c]. They consist of frequency independent regions (neglecting surface states) corresponding to accumulation and depletion at the IS interface, and a frequency dependent region corresponding to inversion. The frequency independent portion of these curves are shown in Figure 3.2 for the p-type diodes of Figure 3.1 and a range of ϕ_{mi} values. The transition from inversion at zero bias to accumulation as ϕ_{mi} increases is evident.

3.2.2 Current-Voltage Characteristics

The band tunnel currents, J_{CT} and J_{VT}, are expected to be described at least qualitatively by the independent electron approach described in Section 2.3. As is shown in Appendix A, this approach can be developed further to give the following approximate expressions for these currents;

$$J_{CT} = \frac{A_n T^2}{N_C} P_n (n_m - n_s) \tag{3.1}$$

$$J_{VT} = \frac{A_p T^2}{N_V} P_p (p_s - p_m) \tag{3.2}$$

n_s is the electron concentration at the semiconductor surface and n_m is a quasi concentration which is equal to the electron concentration which would be calculated at the semiconductor surface if the metal fermi level was used instead of the electron quasi fermi level. Values of n_s

Figure 3.2: Computed C-V characteristics for p type "equilibrium" MIS tunnel diodes with 34Å of SiO_2 separating the metal contact from a 2 Ωcm <100> silicon substrate ($N_A = 7 \times 10^{21}\ m^{-3}$). Other semiconductor and tunneling parameters are given in Table 2.4. The variable parameter is the metal to insulator barrier height, ϕ_{mi} (Surface states assumed negligible).

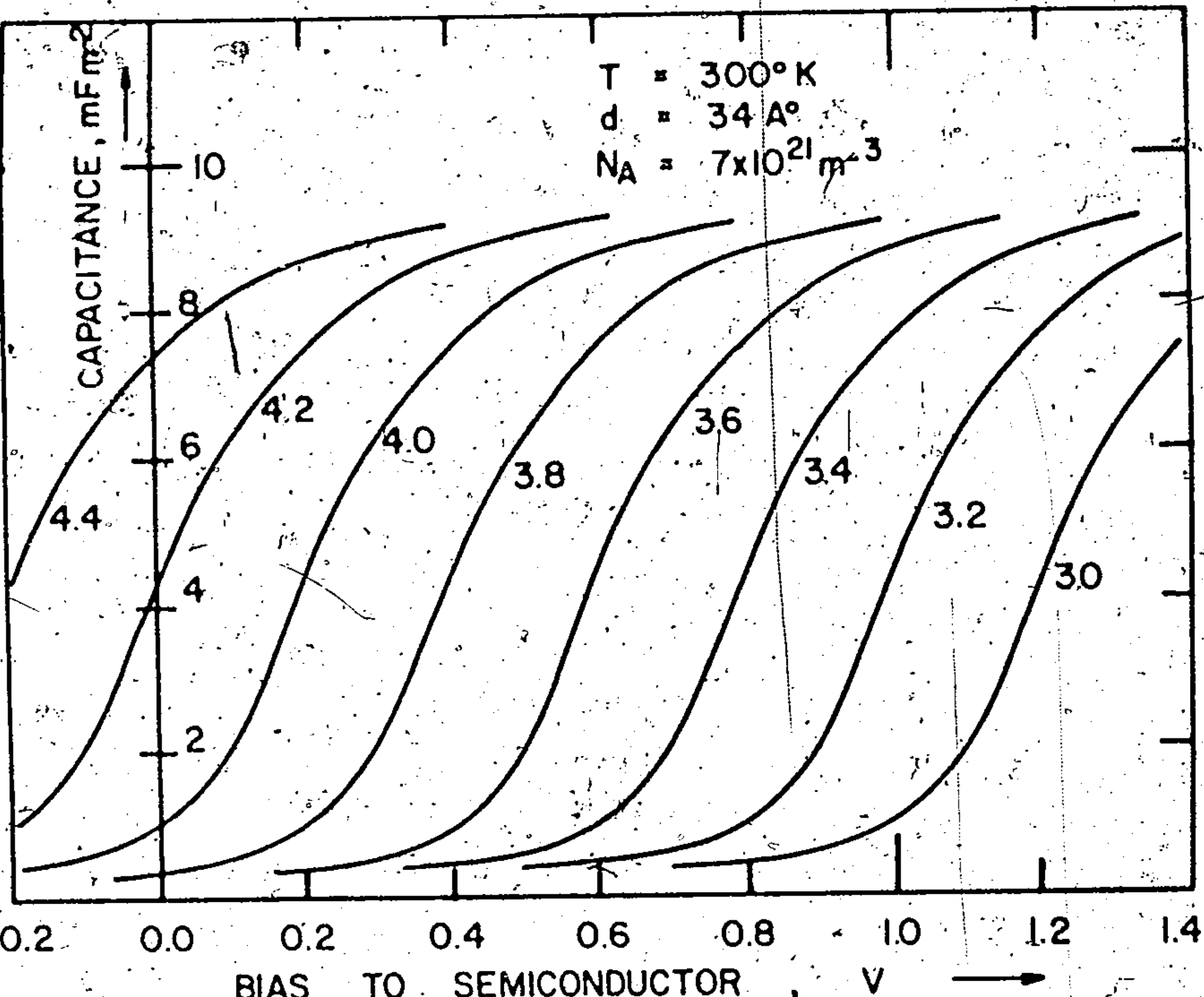


FIGURE 3.2

and n_m (as well as p_s and p_m) are compared in Figure 3.3 for a p-type diode under zero, forward and reverse bias. A_n is a constant and P_n is the probability of an electron moving perpendicularly to the barrier plane with an energy near that of the semiconductor band edge making a tunneling transition (Appendix A). The terms in (3.2) are defined similarly. The expressions are good approximations to results computed using the complete formulation described in Section 2.3 provided

$$n_s, n_m \leq N_C \quad \text{and} \quad p_s, p_m \leq N_V \tag{3.3}$$

Consequently, they will be used to describe qualitatively the results of the numerical simulation. One striking feature of these equations is their similarity to the equations used in the combined thermionic emission-diffusion (TED) theory of Schottky diodes [68-71]. The major differences are that current flows are reduced by the P terms and that the quasi concentrations, n_m and p_m, can be strongly bias dependent in the present case whereas they are relatively bias insensitive for Schottky diodes [70].

At zero bias, n_s and n_m are equal as are p_s and p_m (Figure 3.3(a)). In the present case of equilibrium MIS tunnel diodes, a positive bias increases n_m and p_s, and decreases n_s and p_m (Figure 3.3(b)). The ratio of J_{CT} to J_{VT} near zero bias is given by

$$J_{CT}/J_{VT} = P_{no}n_{so}/P_{po}p_{so} \tag{3.4}$$

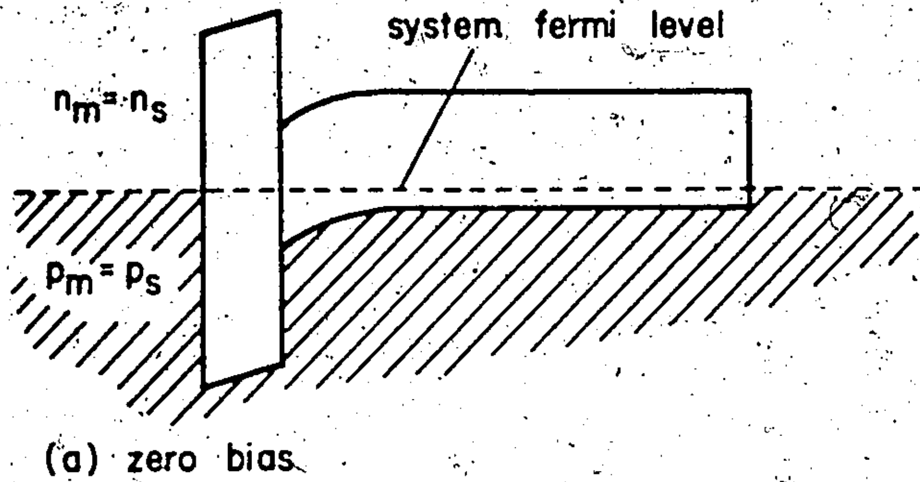
For a two band formulation, the ratio of P_{no} to P_{po} approaches unity [14]

Figure 3.3: Schematic energy band diagrams of the MIS tunnel diode (p type) under different bias conditions. Also shown are the device fermi levels and the relative magnitudes of the quasi concentrations, n_m and p_m.

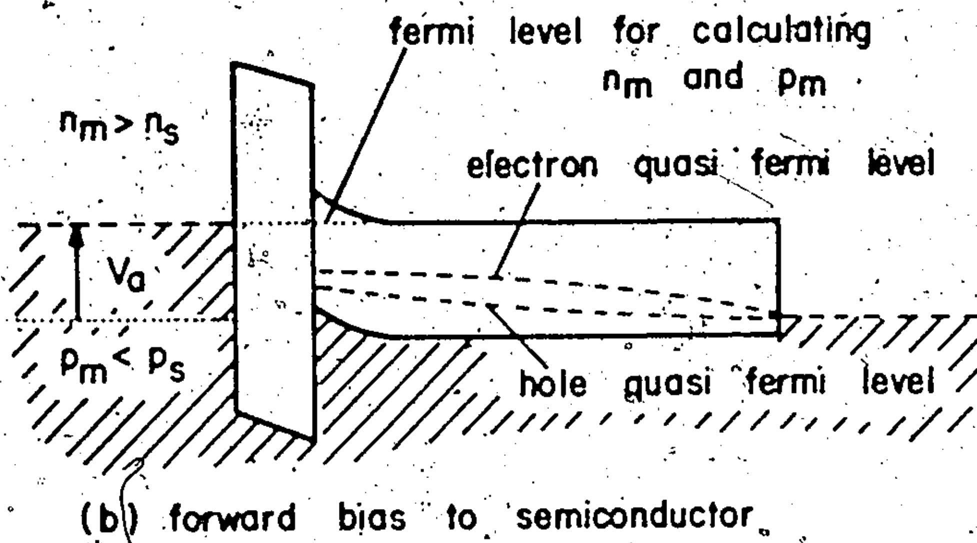
(a): zero bias

(b): forward bias to the semiconductor

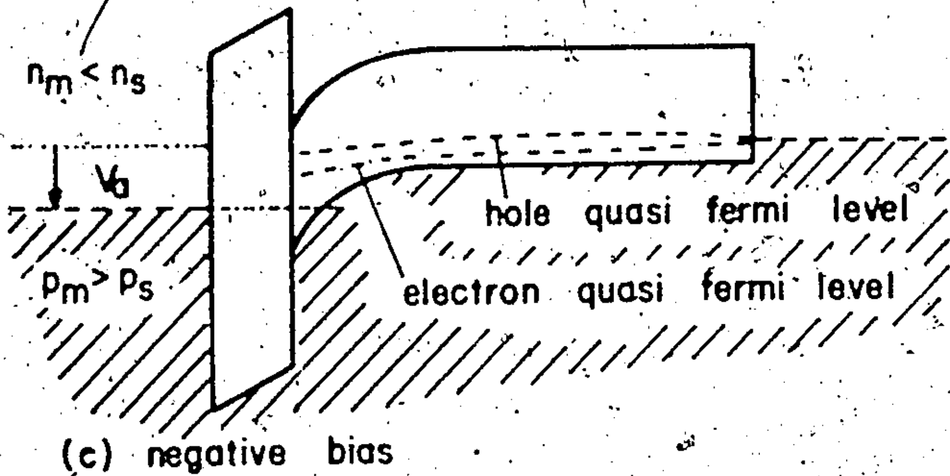
(c): negative bias to the semiconductor



(a) zero bias



(b) forward bias to semiconductor



(c) negative bias

Figure 3.3

and even with a one band formulation, this ratio will be overwhelmed by the extremely large variation which can be obtained in the ratio of n_{so} to p_{so} by varying the metal work function (Figure 3.1). The conclusion is reached that J_{CT} will be dominant for low work function contacts and J_{VT} for high work function contacts. Thus low work function metals should form minority carrier diodes for p-type substrates and majority carrier diodes for n-type substrates, with the opposite holding for high work function metals.

This conclusion is confirmed by a numerical evaluation of the current-voltage characteristics of the p-type diode of Figures 3.1 and 3.2 [102]. The oxide thickness is 34Å and the doping density in the semiconductor is 7×10^{21} m^{-3}. Other parameters are given in Table 2.4. The resulting curves are shown in Figures 3.4(a) and (b). As expected, for the low values of ϕ_{mi} (< 3.6 eV) shown in (a), the dominant component of current at small forward and reverse bias is minority carrier current. For the high values of ϕ_{mi} shown in (b), the majority carrier current is dominant. The value of ϕ_{mi} dividing the two regimes, ϕ'_{mi}, is approximately equal to

$$\phi'_{mi} = \phi_{si} + E_{gs}/2 + \frac{kT}{q} \ln(P_{no}/P_{po}) \tag{3.5}$$

The features of the "equilibrium" characteristics of Figure 3.4 can be described by referring to Figure 3.3 and Equations (3.1) and (3.2). For the equilibrium devices under discussion the electron and hole quasi-fermi potentials would not vary throughout the semiconductor and would be equal to each other. Under forward bias to the semiconductor, n_m

Figure 3.4: Computed tunnel current flows for p type "equilibrium" diodes with the same semiconductor and insulator parameters as in Figure 3.2. The variable parameter is the metal to insulator barrier height, ϕ_{mi}.

(a): I-V characteristics for low values of ϕ_{mi} (< 3.6 eV), where minority carrier current dominates at small applied bias.

(b): I-V characteristics for higher values of ϕ_{mi} (> 3.8 eV) where majority carrier flow is dominant at small applied bias.

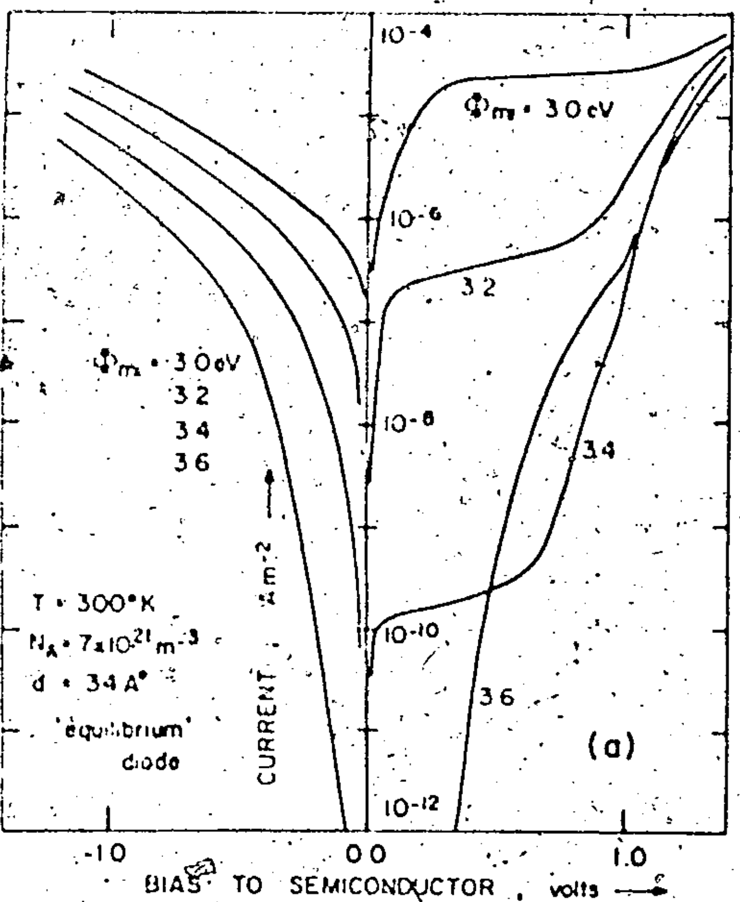


Figure 3.4(a)

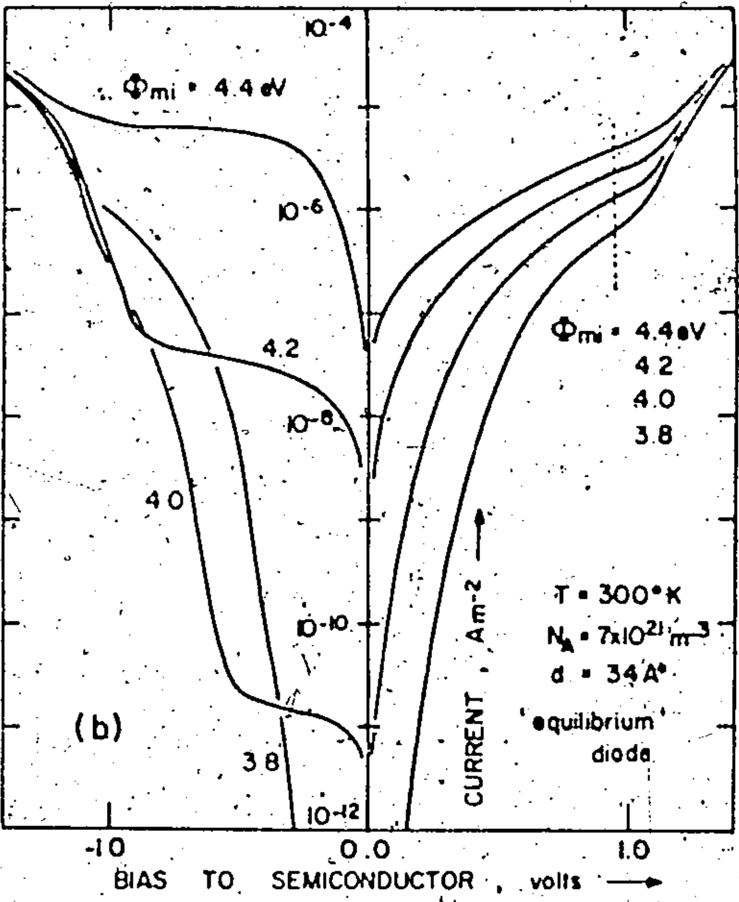


Figure 3.4(b)

and p_s are the dominant concentrations. J_{CT}, therefore, depends to a large extent upon the position of the metal fermi level relative to the conduction band edge which in turn depends upon the voltage across the insulator (Figure 3.3(b)). J_{VT} depends upon the hole concentration at the IS interface. Under reverse bias, the opposite conditions hold. J_{TC} (the current in the opposite direction to J_{CT}) depends upon the electron concentration at this interface, while J_{TV} depends upon the voltage across the insulator (Figure 3.3(c)). The variation of P with bias is much smaller than that of the above concentrations, at least when Equation (3.3) is satisfied.

For the lower ϕ_{mi} values of Figure 3.4(a), the IS interface tends to be inverted at zero bias. The dominant component of current near zero bias is therefore J_{CT} (or J_{TC}) which is minority carrier band current. Increasing the reverse bias to the semiconductor increases the degree of inversion causing J_{TC} to increase rapidly. Most of the reverse voltage is absorbed across the insulator, so J_{TV} also increases rapidly. J_{TC}, dominant near zero bias, maintains its dominance over the entire reverse bias region for the lower ϕ_{mi} values. As the bias is increased in the forward direction, the semiconductor goes from inversion, to depletion, to accumulation. The distinctive feature in this region is the plateau in each curve near zero bias. This corresponds to the IS interface being depleted. In this region, changes in applied bias are absorbed mostly across the depletion region rather than across the insulator so J_{CT} varies only slowly with bias. As the surface goes from depletion to accumulation, changes in bias are absorbed across the insulator resulting in a pronounced hump caused by the resultant increase

in J_{CT}. Since holes are accumulating at the IS interface, J_{VT} increases rapidly over the forward bias region. For the lower ϕ_{mi} values, J_{CT} dominates over the entire bias range, but, for higher ϕ_{mi} (3.4, 3.6 eV), additional structure is added to the curves in the bias region 0.5 - 1.0 V where J_{VT} is larger.

For the majority carrier diodes of Figure 3.4(b), the characteristics are similar but mirror reflected in the current axis. This is due to the basic symmetry between the conduction and valence bands as well as between accumulation and inversion. J_{VT} dominates to higher than 1 V forward bias, the depletion plateau occurs under reverse bias as does a hump due to the IS interface becoming inverted, and, for the lower ϕ_{mi} values (3.8, 4.0 eV), additional structure occurs under reverse bias due to J_{TC} becoming larger than J_{TV} for a small range of bias. As mentioned in Section 3.1, the above description is influenced by charge in the insulator and by the presence of surface states and these effects will be treated in Section 3.4.

3.2.3 Duality of n and p Type Diodes

Due to the similarity between accumulation and inversion, the current-voltage characteristics of n and p type equilibrium diodes are closely related. These characteristics are shown in Figure 3.5 for ϕ_{mi} equal to 3.2 eV, d equal to 34Å, and for a range of substrate doping densities. For lightly doped substrates, the current-voltage curves are nearly identical. Since the same band tunnel current dominates in each case, the following observation can be made. If a p-type diode is a minority (majority) carrier device, the n-type diode with the same metal contact will be a majority (minority) carrier device. This statement

Figure 3.5: Computed I-V characteristics for both n and p type "equilibrium" MIS tunnel diodes as the dopant density in the semiconductor varies over the range 10^{20} to 10^{23} m^{-3}. The curves are calculated with ϕ_{mi} equal to 3.2 eV and other device parameters as in Figure 3.2.

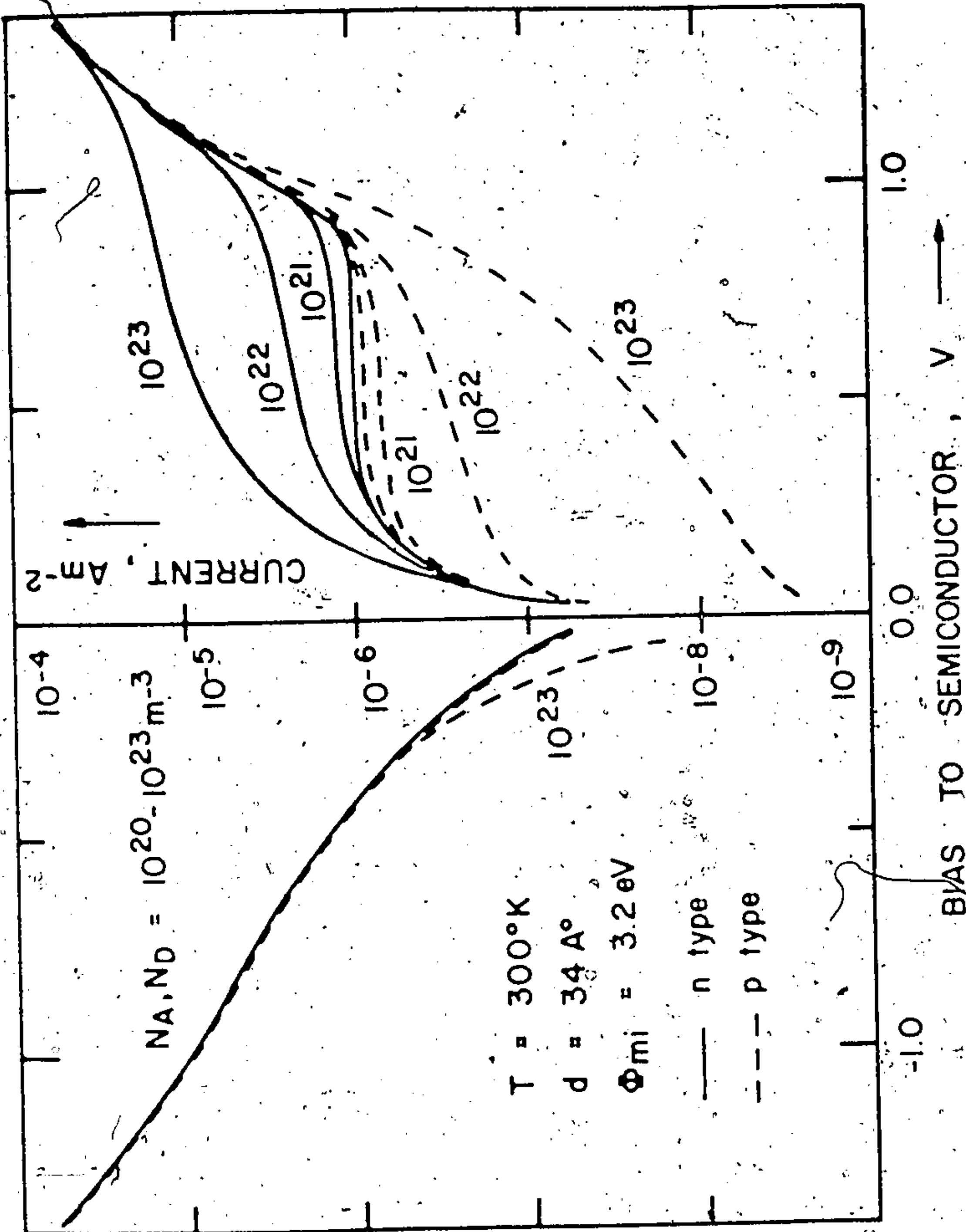


FIGURE 3.5

may be affected if large surface state densities are present (Section 3.4).

As the insulator thickness is reduced, the magnitudes of the tunnel currents increase rapidly. They eventually become so large that they disturb the semiconductor significantly from thermal equilibrium. Different changes are produced in majority and minority carrier devices, causing the characteristics of n and p type non-equilibrium diodes with the same metal contact to differ. These will now be discussed.

3.3 Non-Equilibrium Diodes

3.3.1 Effect of Insulator Thickness

The results of reducing the insulator thickness for typical minority and majority carrier diodes are shown in Figures 3.6(a) and (b) respectively where I-V characteristics for p-type diodes ($N_A = 7 \times 10^{21}$ m^{-3}) are shown for a range of d values. The behaviour of the p-type majority carrier diode with ϕ_{mi} equal to 4.2 eV (Figure 3.6(b)) will be described initially. With positive bias to the semiconductor, the current flow at a given bias merely increases exponentially with decreasing insulator thickness. This is the type of behaviour expected when the diode current is limited by the rate at which carriers can tunnel between the metal and the semiconductor. The same effect is observed at reverse bias up to about -0.8 V. Then the distinctive majority carrier non-equilibrium effect occurs. The hump present in the thicker device curves disappears. As described in Section 3.2.2, this hump is caused by the IS interface becoming inverted so its absence indicates the lack of such an inversion regime. The transition from equilibrium to non-equilibrium behaviour occurs at a value of d approximately equal

Figure 3.6: Computed I-V characteristics of the MIS tunnel diode demonstrating "non-equilibrium" effects. The substrate is 2 Ωcm p type silicon of $\langle 100 \rangle$ orientation. The variable parameter is d, the thickness of the insulating layer. Other device parameters are given in Table 2.4. (Surface states assumed negligible).

(a): For the case where ϕ_{mi} equals 3.2 eV demonstrating minority carrier MIS tunnel diode behaviour.

(b): For ϕ_{mi} equal to 4.2 eV demonstrating majority carrier diode behaviour.

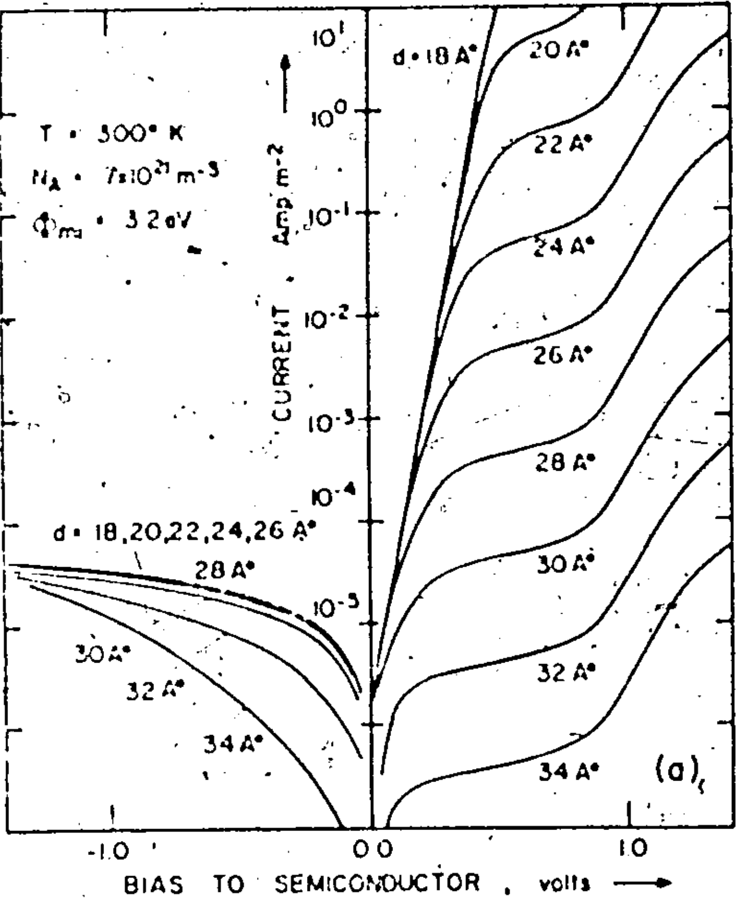


Figure 3.6(a)

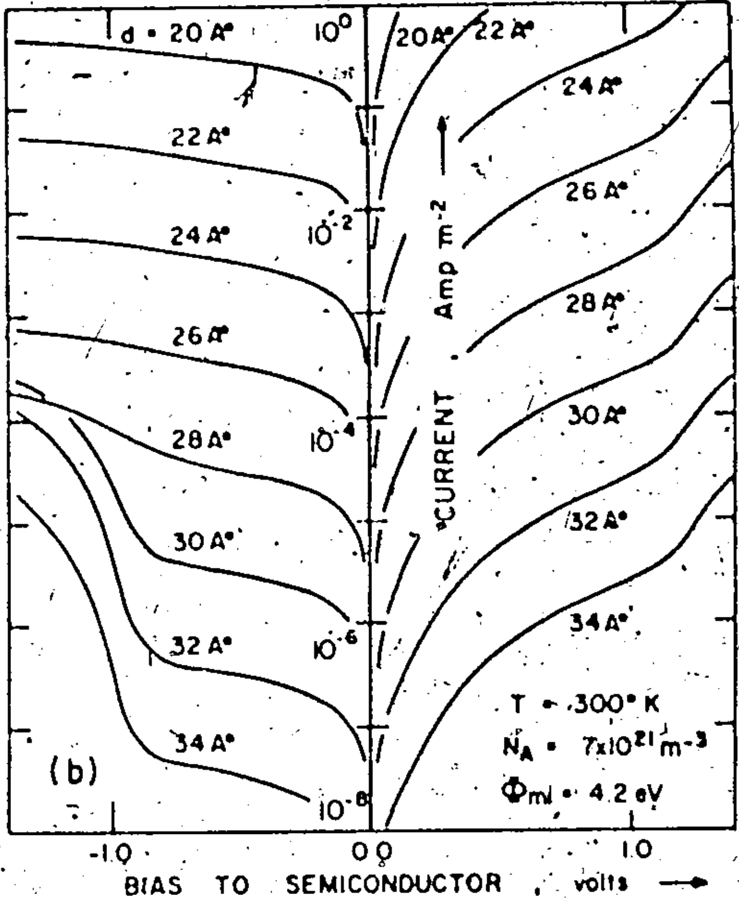


Figure 3.6(b)

to 28Å.

Although the devices of Figure 3.6(b) are majority carrier diodes, the reason for the disappearance of the inversion layer becomes clear if minority carrier flows are considered. In this bias region, the minority carrier tunnel current (J_{CT}) depends upon the electron concentration at the IS interface (Equation (3.1)). As the thickness decreases, the magnitude of this current increases exponentially. However, the rate at which the minority carriers can be supplied from the semiconductor is limited to the rate they can be generated in the semiconductor and this is restricted under the reverse bias conditions at the IS interface. Since the two currents must match, J_{CT} is also restricted which, from (3.1), means that the electron concentration at the IS interface is limited. The following description of the thinner insulator devices applies. At zero bias, the IS interface tends to be accumulated. As reverse bias is applied, the interface goes into depletion and the minority carrier concentration builds up to its limiting value. Until this stage, non-equilibrium devices are the same as equilibrium devices. Past this point, any additional voltage applied is absorbed mainly across the depletion region which consequently expands with bias. Little voltage is absorbed across the insulator which explains the near saturation of J_{VT}. However, if additional minority carriers are supplied by external processes such as illumination by light or injection from a third contact, the hump in J_{VT} can be restored. This forms the basis of the current multiplication process which will be described in Chapter 6.

The features of this majority carrier non-equilibrium behaviour can be seen to influence the high-frequency (100 kHz) capacitance-voltage

characteristics shown in Figure 3.7(b) for the thickness range of Figure 3.6(b). For the thicker equilibrium devices, the characteristics exhibit the normal MIS features including nearly constant capacitance in the inversion bias regime. For the thinner insulator devices, this behaviour is replaced by a depletion mode characteristic ($1/C^2 \propto V$). For the entire thickness range, the only frequency dispersion of the curves at frequencies less than 100 kHz occurs at reverse bias for the thicker devices where the inversion layer response becomes apparent at very low frequencies as in normal MIS diodes.

Minority carrier devices exhibit far different behaviour as the insulator thickness is reduced (Figure 3.6(a)). At large forward bias (> 0.4 V), the diode current increases rapidly with decreasing insulator thickness as expected of a tunneling limited process. However, at reverse and small forward bias (< 0.4 V), there is a distinct upper limit to the current flow at a given bias point regardless of how thin the insulator is made. As might be expected, the diode current in this bias region is not limited by the rate carriers can tunnel through the insulator, but by their transport properties in the semiconductor.

As indicated in Figure 3.1, the minority carrier diodes under discussion have the IS interface inverted near zero bias. This inversion region is followed by a depletion region and finally a space charge neutral region as the bulk semiconductor is approached. Under reverse bias, the rate that minority carriers can be supplied to the IS interface is limited by the rate they can be generated in the depletion region for all but small bandgap semiconductors. This rate is represented by the current flow in Figure 3.6(a) for the thinner insultator devices under

Figure 3.7: Computed C-V curves corresponding to the I-V curves of the p type MIS tunnel diode of Figure 3.6. The curves are calculated at the relatively high frequency of 100 kHz. The variable parameter is d, the thickness of the insulating layer.

(a): For the case where ϕ_{mi} equals 3.2 eV demonstrating "non-equilibrium" effects for minority carrier devices.

(b): For ϕ_{mi} equal to 4.2 eV demonstrating "non-equilibrium" effects for majority carrier devices.

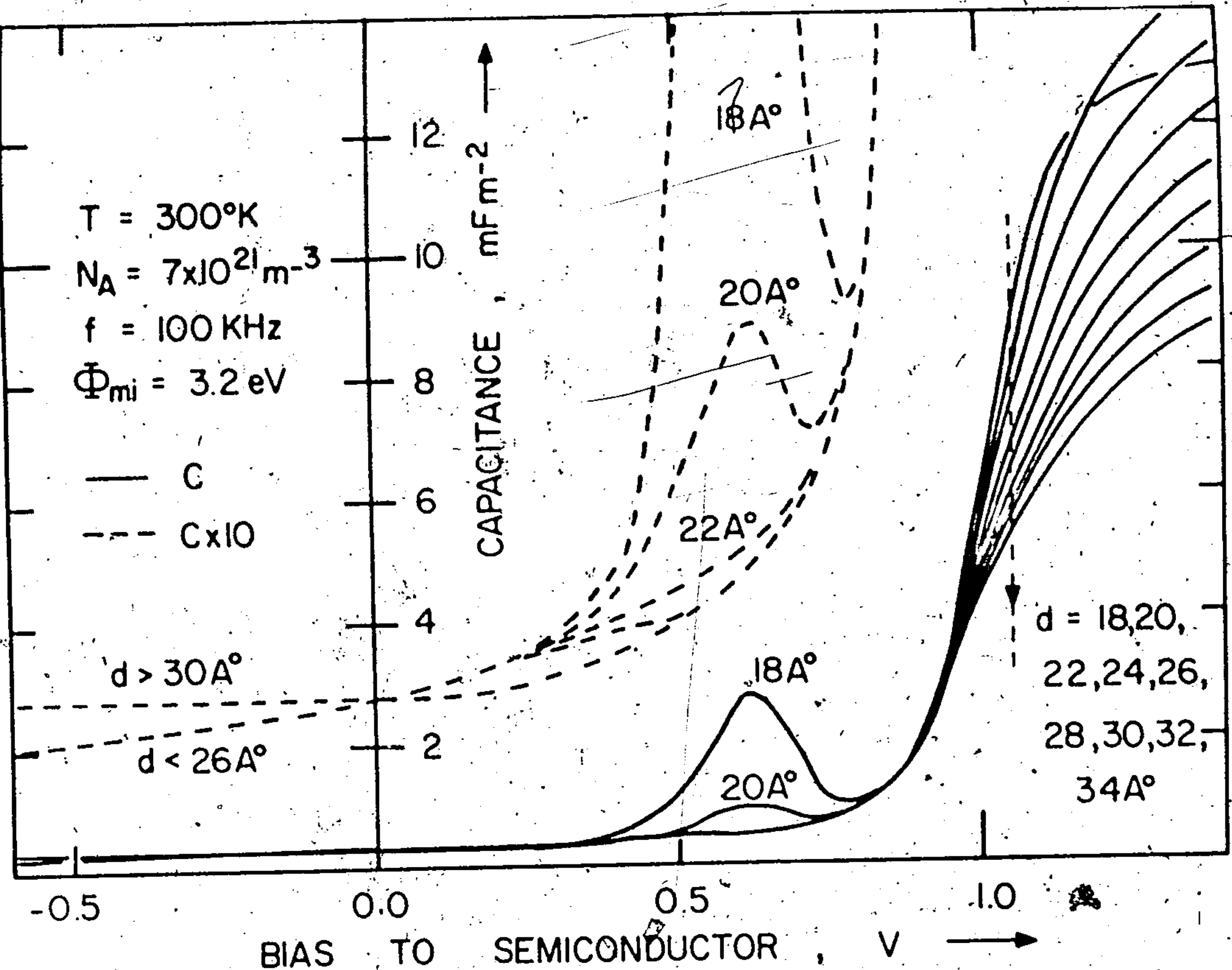


FIGURE 3.7(a)

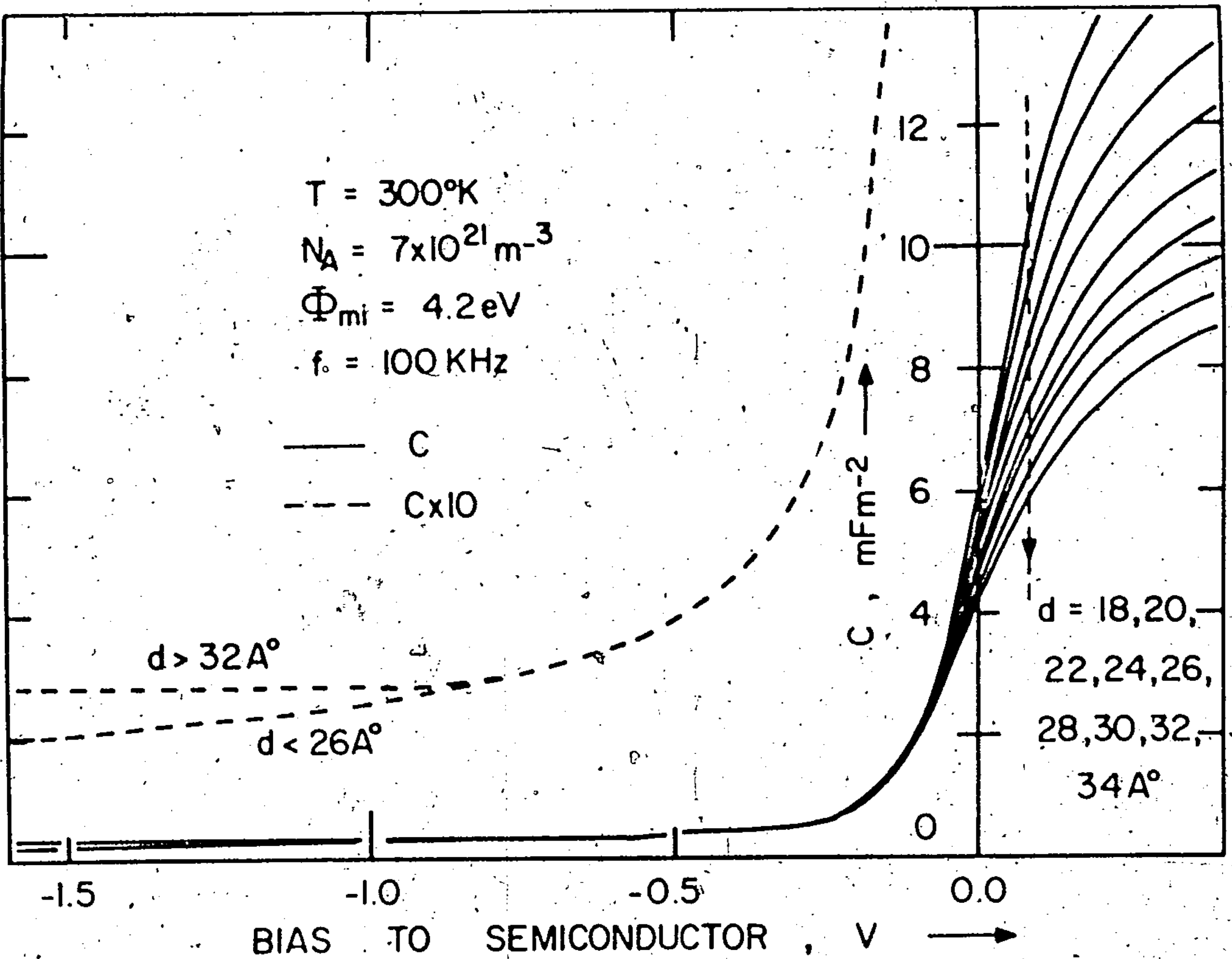


FIGURE .3.7(b)

reverse bias. Under forward bias, plenty of minority carriers can be supplied to the IS interface by the tunneling process, but these carriers must be able to flow through the semiconductor. The maximum rate for such a flow is given by the sum of the recombination rate in the depletion region and the diffusion rate of minority carriers through the semiconductor bulk. This total rate is represented by the asymptotic curve approached for thin insulator values at small forward bias in Figure 3.6(a).

A more detailed description of the minority carrier non-equilibrium diode is given in Chapter 6 where is is shown there that the asymptotic curve of Figure 3.6(a) is the same curve as would be computed for the current flow in an N^+P junction diode. This is not surprising in view of the above discussion. A more surprising result concerns the behaviour of the inversion layer present at zero bias. It is shown also in Chapter 6 that this layer remains at the IS interface throughout the entire bias region where the device current is semiconductor limited. Changes in applied voltage in such bias regions are absorbed across the depletion layer. At high forward bias where the current changes from semiconductor to tunnel-limited, the inversion layer eventually disappears.

The above behaviour is reflected in the high frequency (100 kHz) capacitance-voltage curves shown in Figure 3.7(a) for such minority carrier diodes. The curves for the thicker insulator diodes (>28Å) are the same as for conventional MIS diodes. At zero bias, the capacitance has nearly the same value independent of insulator thickness. However, the difference between equilibrium and non-equilibrium devices becomes

apparent under reverse bias. The width of the depletion range increases in the latter case, causing the device capacitance to decrease in a $1/C^2$ - V mode. Under forward bias, this width decreases for the thin insulator case. Note the distinctive humps under moderate forward bias (0.4 to 0.8 V) which increase in size as the insulator thickness is decreased. These humps are due to inversion layer response. As mentioned above, the inversion layer is clamped at the IS interface until the device current becomes tunnel limited when it is then able to disappear with increasing bias. Its frequency response is much enhanced due to the large quantities of minority carriers being injected by the tunnel contact. The situation is similar to that which arises in conventional MIS devices when large quantities of these carriers are supplied by illumination (e.g. see [25]). At large forward bias, the equilibrium and non-equilibrium devices exhibit similar characteristics differing only in the value of accumulation capacitance.

The value of accumulation capacitance provides an estimate of the insulator thickness for experimental devices. However, Figure 3.7(a) shows this approach is not suitable for thicknesses less than 20Å for these minority carrier diodes. In addition to experimental difficulties in measuring this capacitance due to the large current flows present, the value for the 18Å case would give incorrect results. The reduced value of accumulation capacitance in this case is attributed to an inductive contribution from the semiconductor bulk due to conductivity modulation effects at these large minority carrier injection levels [72].

The frequency dependence of the inversion layer capacitance hump is illustrated in Figure 3.8 for the 20Å p-type diode of Figure 3.7(a). Note the corresponding dispersion in the conductance-voltage characteristics. The range of frequencies over which the hump appears depends upon the insulator thickness and the metal work function. For lower work functions, the hump extends over a larger voltage range. For very thin insulator values or lower work functions, diffusion capacitance [73] due to minority carriers in the diode bulk can contribute to the initial portion of the hump.

3.3.2 Metal Work Function Effects

The effect of metal work function upon majority carrier non-equilibrium current-voltage curves is simply evaluated since the major non-equilibrium effect is to remove the inversion regime. A comparison of the I-V characteristics for non-equilibrium majority carriers would be similar to that for the equilibrium devices of Figure 3.4(b) except that in all cases the humps appearing in the latter under reverse bias would not be present. In the case of minority carrier diodes, the effect is more striking. As illustrated in Figure 3.9(a), the current flow at reverse and small forward bias is independent of the metal work function. Note that both minority and majority carrier non-equilibrium diodes have near saturating reverse characteristics. However, whereas for minority carrier diodes this characteristic depends entirely upon the semiconductor properties, for majority carrier diodes it depends upon the tunneling properties of the device and increases rapidly with decreasing insulator thickness (Figure 3.6(b)).

Figure 3.8: Demonstration of frequency dispersion effects in both the C-V and G-V curves for minority carrier "non-equilibrium" devices. Note that surface state effects have not been included into the analysis. The curves are for p type <100> silicon (N_A = 7 x 10^{21}, d = 20Å, ϕ_{mi} = 3.2 eV, other parameters as in Table 2.4).

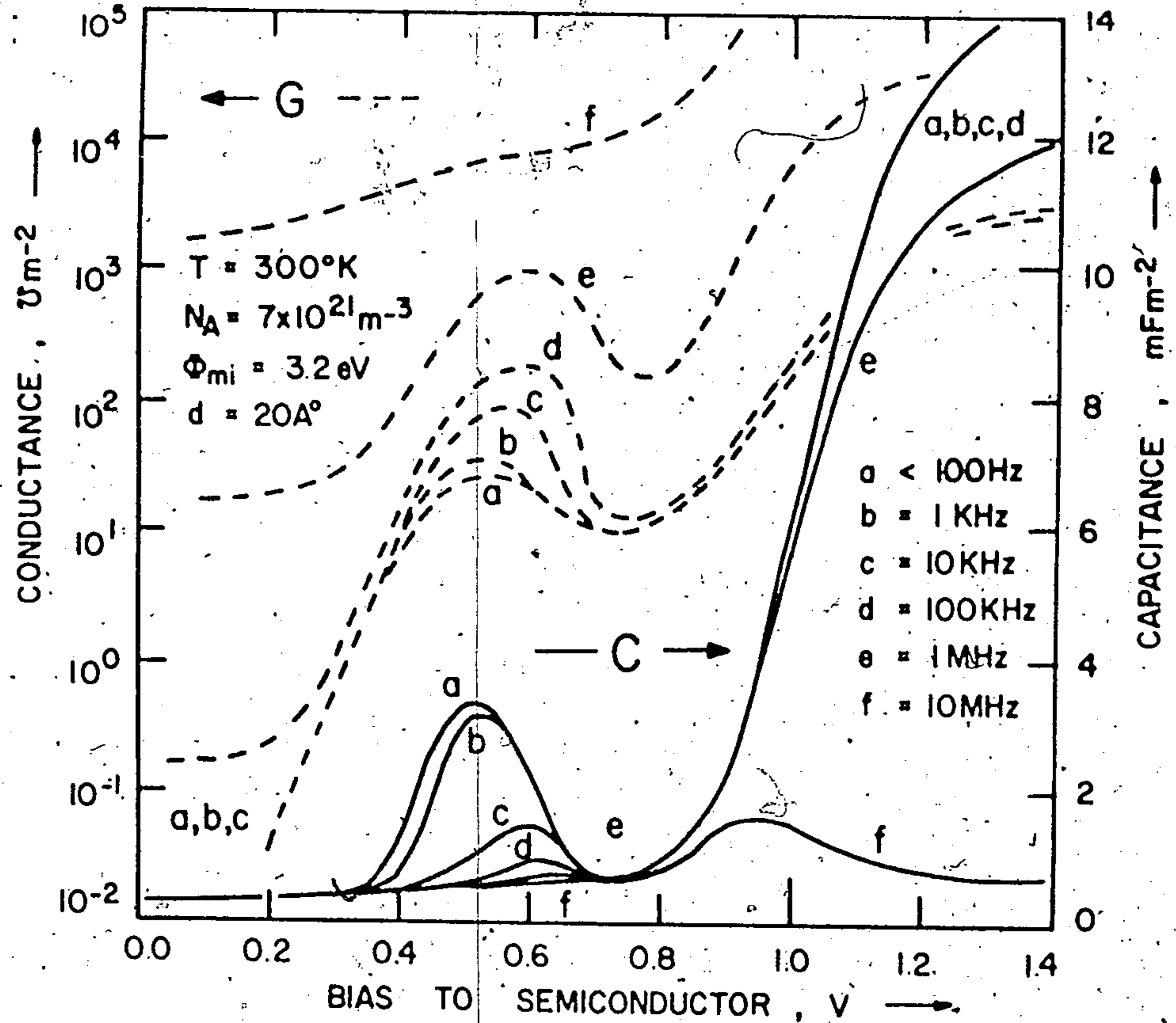


FIGURE 3.8

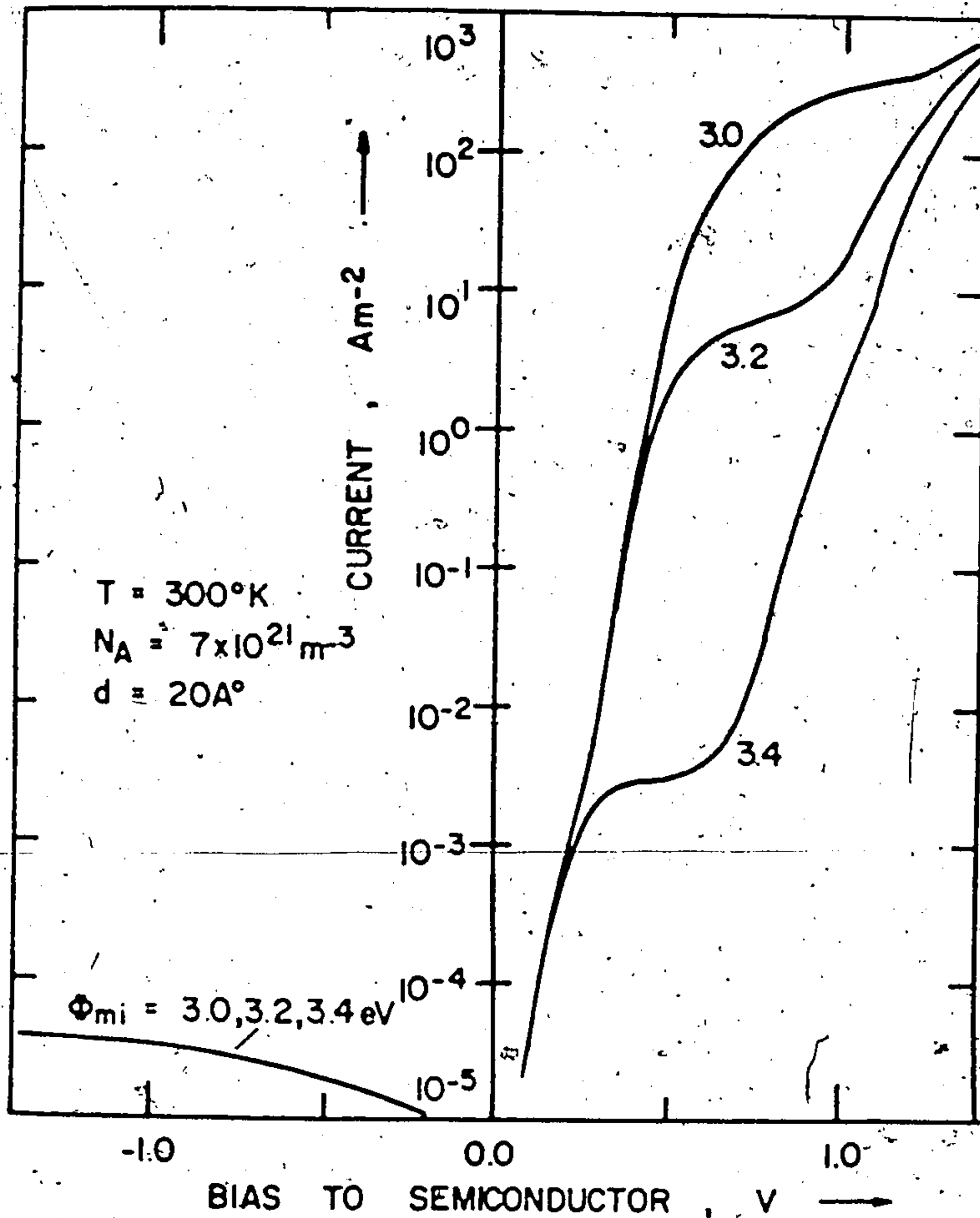


Figure 3.9(a): Computed I-V curves for minority carrier "non-equilibrium" p type MIS tunnel diodes with ϕ_{mi} as the variable parameter. Other parameters are the same as in Figure 3.8.

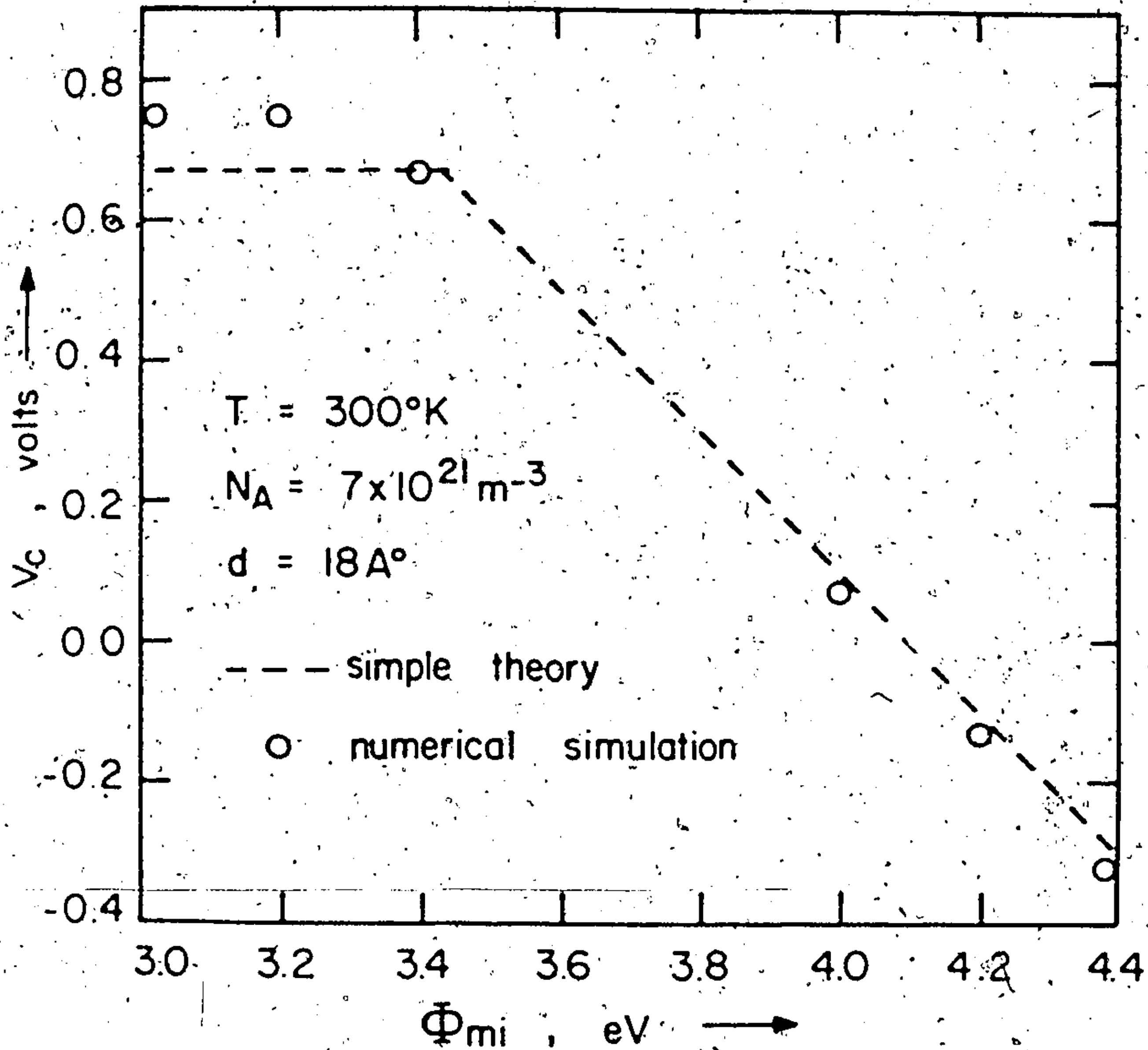


Figure 3.9(b): Dependence of V_C (intercept of the linear $1/C^2$ versus V portion of the C-V characteristics with the voltage axis) upon Φ_{mi}. The dashed curve shows the results of the simple theory of Appendix C while the circles represent the results of the more complete numerical simulation. Device parameters are as in Figure 3.9(a) except that d has the value 18Å.

Both minority and majority carrier diodes also exhibit capacitance-voltage characteristics which vary in the "depletion" mode ($1/C^2 \propto V$) under reverse bias. The intercept of the plot of $1/C^2$ versus V with the voltage axis (V_C) is a quantity which is of interest since it is readily accessible experimentally. Provided the insulating layer is sufficiently thin so that the IS interface never becomes inverted, V_C is proportional to the metal work function for majority carrier devices. However, for minority carrier devices, V_C tends to saturate with metal work function as is discusses in a simple manner in Appendix C. The results of this treatment are compared to the exact theoretical treatment in Figure 3.9(b). To a good approximation, V_C is given by (see Figure 2.1 for definition of symbols)

$$
\begin{aligned}
V_C &= \phi_{si} - \phi_{mi} + \delta \qquad & V_C < \psi_{sinv} \\
&= \psi_{sinv} \qquad & V_C \geq \psi_{sinv}
\end{aligned}
\tag{3.6}
$$

where ψ_{sinv} is the surface potential required to invert the semiconductor surface (Appendix C). In view of the above and the fact that the accumulation portion of the capacitance curves are all displaced according to the value of ϕ_{mi}, the extent of the low-frequency capacitance hump observed for minority carrier diodes under forward bias can be estimated as

$$\delta V_H = \phi_{si} - \phi_{mi} + \delta - \psi_{sinv} \tag{3.7}$$

$$= \phi_{si} - \phi_{mi} + \frac{kT}{q}\ln(N_C/N_A) \tag{3.8}$$

The hump appears as soon as the diode switches from the semiconductor limited to the tunnel limited mode. This voltage point depends upon the metal work function and the insulator thickness.

3.3.3 Complementary n and p Diodes

For a given metal, n and p type substrates will produce diodes belonging to different classes. The current-voltage characteristics in the non-equilibrium case will therefore be different. This is illustrated in Figure 3.10. Portions eliminated from the p type curve by non-equilibrium effects remain unchanged in the n type curve. The hump deleted from the n type curve by these effects remains in the p type curve. The form of the equilibrium characteristics could be estimated from such a pair of complementary non-equilibrium curves.

3.4 Surface State Effects

3.4.1 Introduction

Surface states associated with the IS interface can cause deviations from the above theory. These deviations arise due to the following properties of these states;

(a) they are charge storage centres, the amount of charge stored depending on the bias conditions;

(b) they provide a spatially concentrated region of recombination-generation centres;

(c) they provide additional tunneling paths between the metal and the semiconductor.

All of these properties were included self-consistently into the model used to characterize these states (Section 2.4). A standard

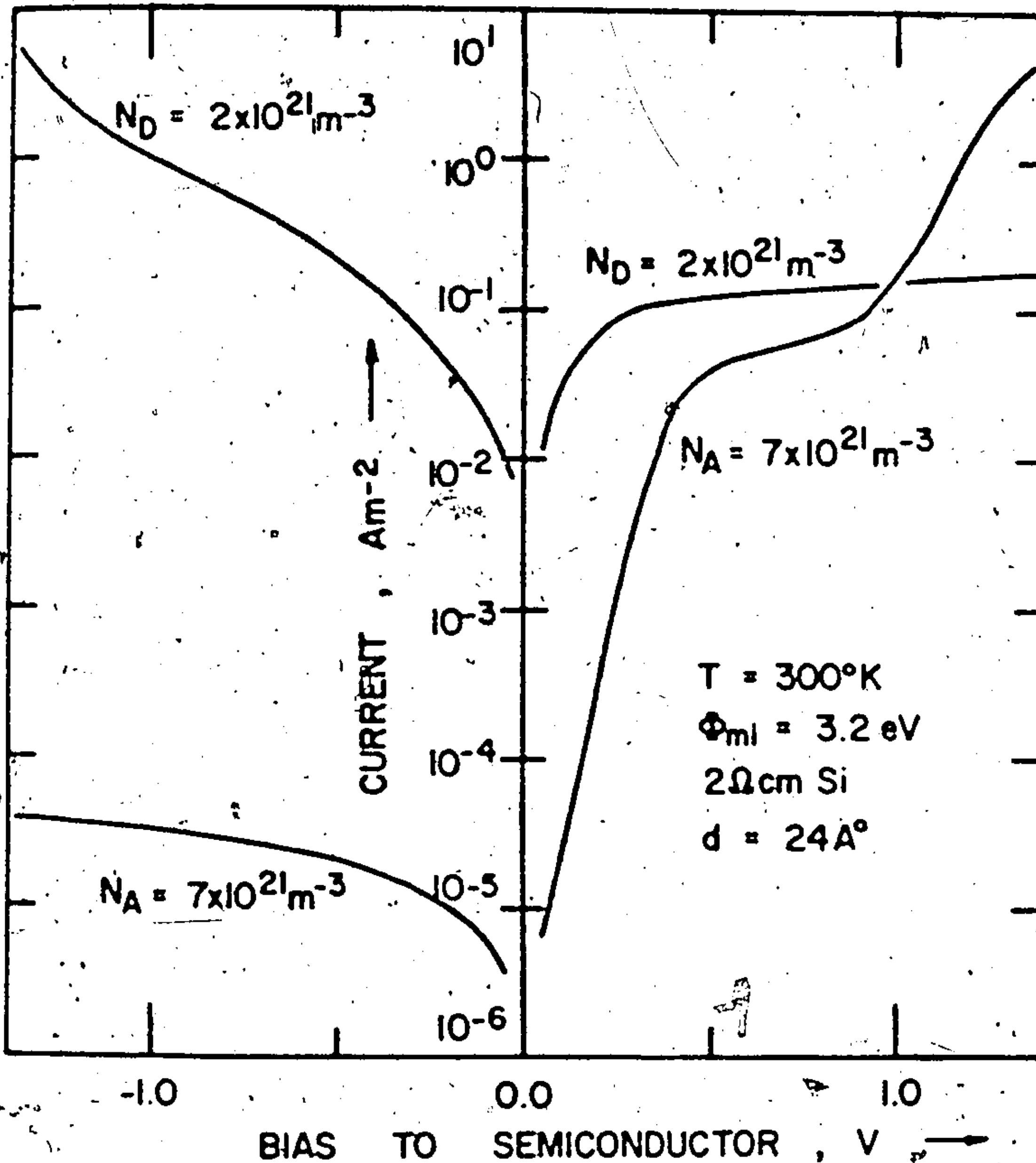


Figure 3.10: Comparison of the computed I-V characteristics for two "non-equilibrium" devices with the same metal contact (ϕ_{mi} = 3.2 eV) but with n and p type substrates. The substrate resistivity is about 2 Ωcm in each case ($N_A = 7 \times 10^{21}$ m^{-3}, $N_D = 2 \times 10^{21}$ m^{-3}) and the oxide thickness is 24Å.

distribution of surface states was used throughout this chapter. This distribution is illustrated in Figure 3.11 and is based upon experimental measurements upon thick insulator M-SiO_2-Si devices [74]. It represents a moderate to large surface state density for such devices.

The effect of fixed charge in the insulator region will also be treated in this section. It was found to have an effect virtually indistinguishable from a change in the metal-insulator barrier height ϕ_{mi}. In fact, the effect of insulator charge and the charge storage role of surface states can be qualitatively described by assuming an effective value of ϕ_{mi} given by

$$\phi''_{mi} = \phi'_{mi} - q(Q_{ss} + Q_i)d/\varepsilon_i \tag{3.9}$$

where Q_i is the effective charge at the IS interface and Q_{ss} is the charge in the surface states.

From Equation (2.31), surface states are heavily occupied by electrons if the electron concentration in the surface region is large. Under such conditions, the sum of the surface state and insulator charges has its most negative value which will be denoted by Q_{ss}^-. Conversely, when the hole concentration at the IS interface is large, it has its most positive value denoted by Q_{ss}^+. In view of this, one consequence of the charge storage properties of surface states can be described. High metal work functions attract holes to the IS interface under zero bias while low values attract electrons (Figure 3.1). This tends to reduce the difference between the effective ϕ_{mi} values. This effect is equivalent to the one usually invoked to explain the similarity in the

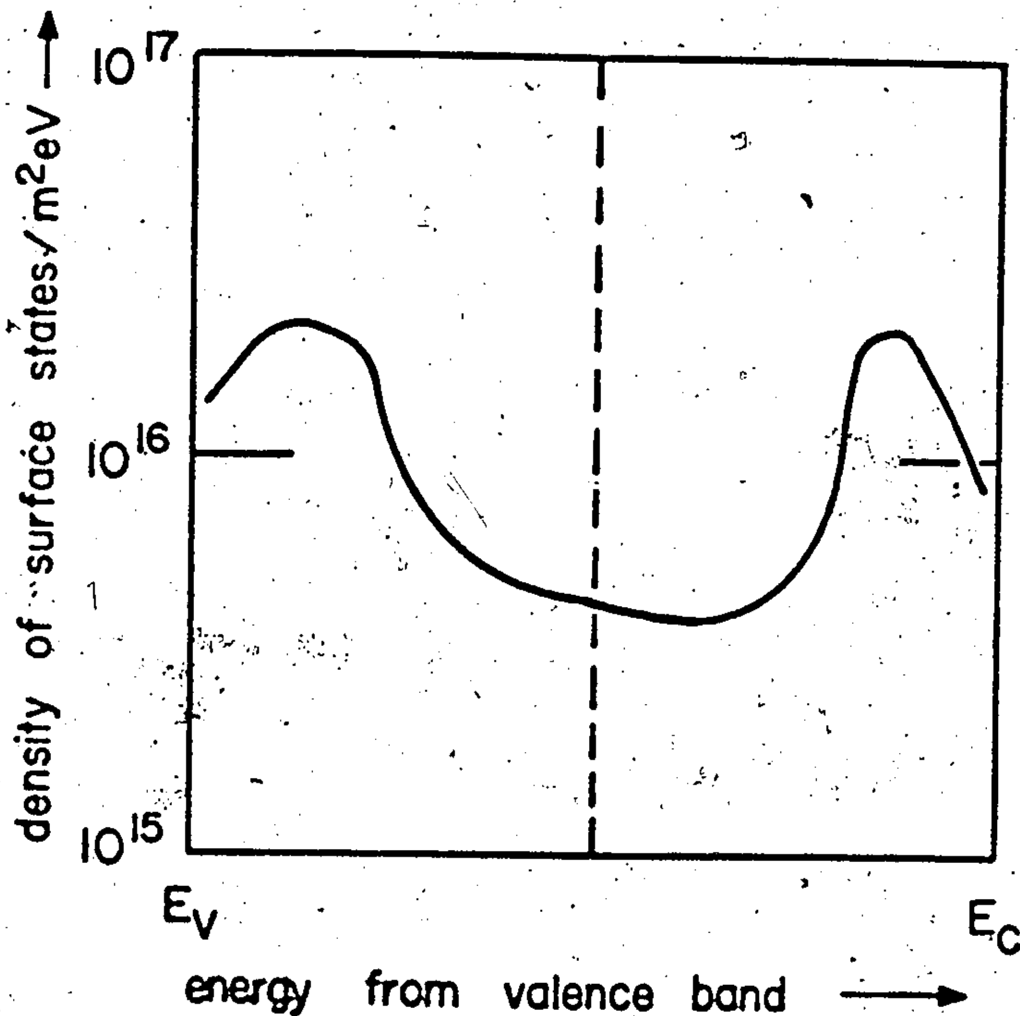


Figure 3.11: Assumed distribution of surface states across the energy band gap for the subsequent calculations in Chapter 3 (also Chapter 6). The states below midgap were treated as donor type levels, while those above midgap were modelled as having acceptor type properties.

properties of metal-semiconductor contacts [75]. However, at least for the diodes fabricated during this thesis work, this effect was limited to less than about 0.3 V.

3.4.2 Effects Upon Equilibrium Diodes

As might be expected on the basis of the curves of Figure 3.4, surface states have the most striking effects upon devices which have a value of ϕ_{mi} which lies between ϕ_{si} and $\phi_{si} + E_{gs}$ close to the value ϕ'_{mi} given by Equation (3.5). For such devices under zero bias, the metal fermi level lies opposite the semiconductor band gap and away from the band edges. This makes the band currents very small and enhances surface state tunnel current, allowing it to dominate. For devices with ϕ_{mi} values removed from the value ϕ'_{mi}, the appropriate band current should dominate unless surface state densities are very large.

Using the computational method of Chapter 2, the effect of surface states upon devices with values of ϕ_{mi} near that of ϕ'_{mi} (Equation (3.5)) is demonstrated in Figure 3.12(a) for several surface state treatments. Shown is the p type diode of Figure 3.4(a) with a value of ϕ_{mi} equal to 3.6 eV. The simplest case shown is that where surface states are neglected ($N_{ss} = 0$, $\sigma_T = 0$). The curve in this case is the same as that occurring in Figure 3.4(a). The next simplest case is where the surface state distribution of Figure 3.11 is assumed, but tunneling between the metal and surface states is neglected ($\sigma_T = 0$). The remaining two cases show the results of increasing the value of the tunneling capture cross section of these states. As a result of Figure

Figure 3.12: Computed effects of surface states upon the I-V characteristics of "equilibrium" p type MIS tunnel diodes. Several different treatments of surface states are shown. In the simplest case ($N_{SS} = 0$, $\sigma_T = 0$), surface states and oxide charge are assumed negligible and the curves are the same as those in Figure 3.4(a). In the other cases, the surface state distribution of Figure 3.11 is assumed in addition to an oxide charge, Q_{ox}, of $5 \times 10^{15}\ m^{-2}$. The capture cross section for electrons and holes (σ_n and σ_p) have the value σ_{co} (Table 2.4). The value of the tunneling capture cross section is then the variable parameter.

(a): I-V characteristics for a moderate value of ϕ_{mi} (3.6 eV).

(b): I-V characteristics for a low value of ϕ_{mi} (3.2 eV).

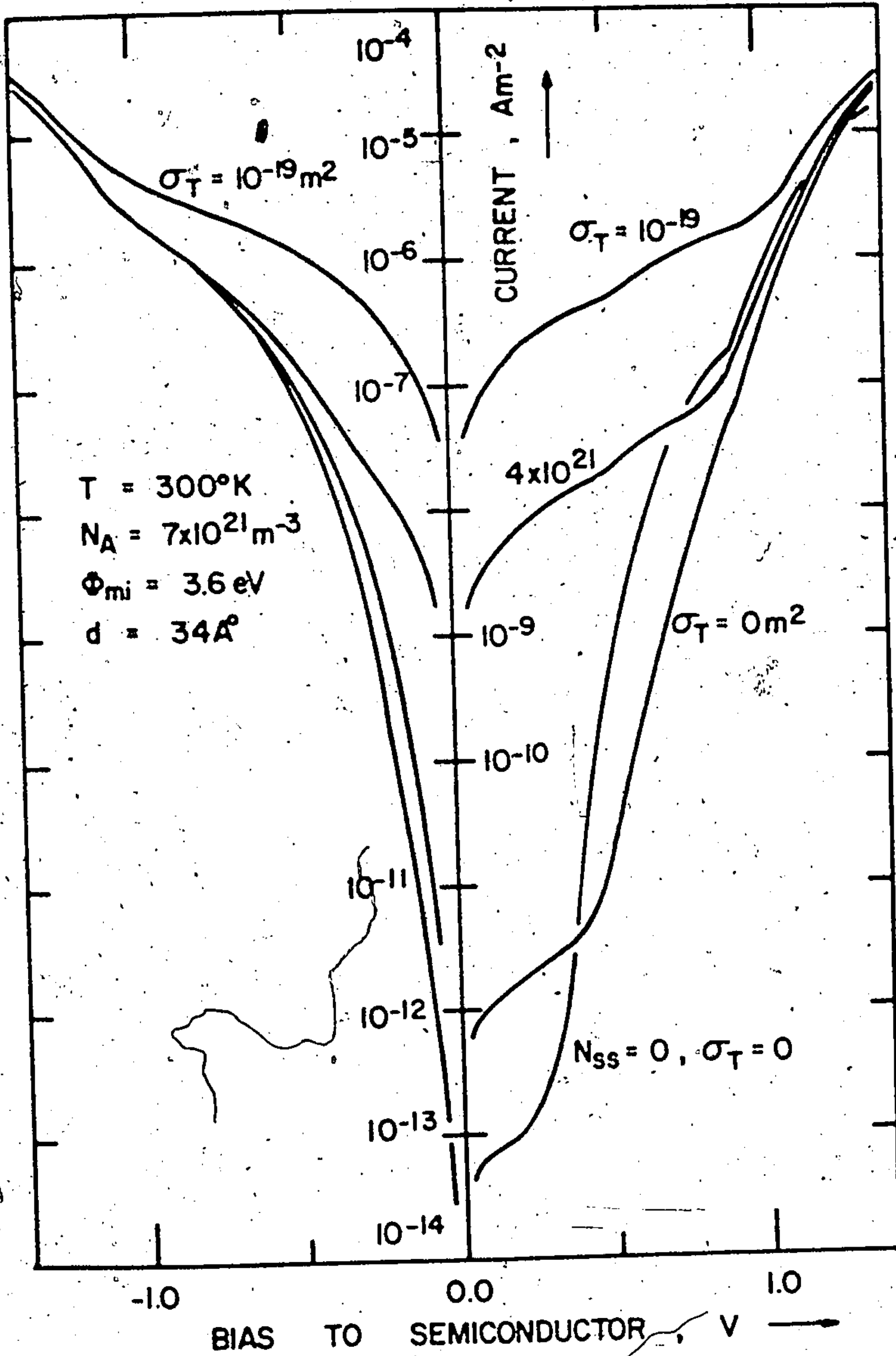


FIGURE 3.12(a)

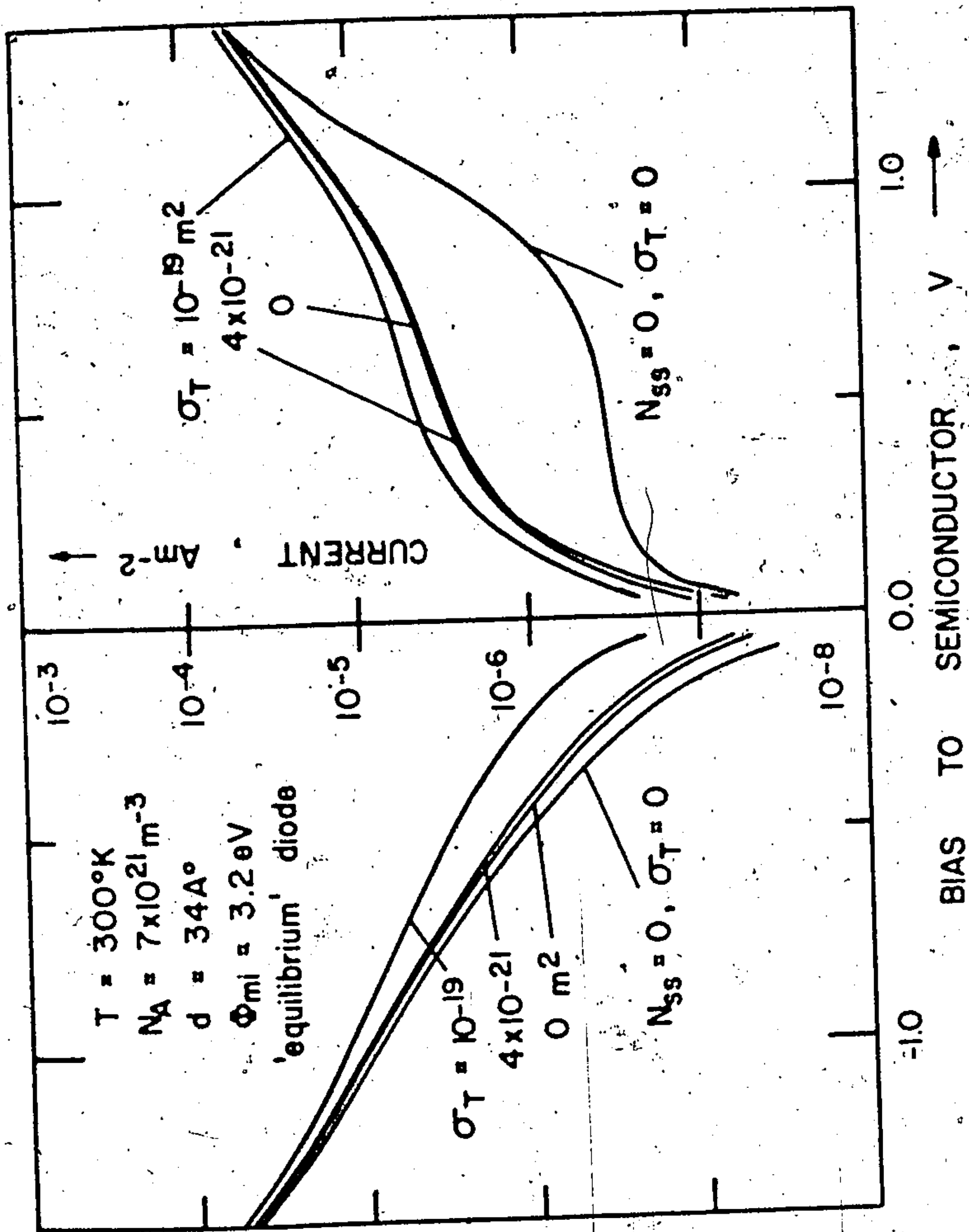


FIGURE 3.12(b)

3.12(a), the conclusion is reached that the dominant component of device current near zero bias is likely to be that due to the direct current flow between the metal and surface state levels for these ϕ_{mi} values. Such devices will be termed surface state diodes. The surface state tunneling current (J_{ST}) fills in the regions of small current flow near zero bias. Since this surface state current is not as bias dependent as the band tunnel currents once the IS interface becomes accumulated or inverted, it eventually loses its dominance as the diode bias becomes high.

For devices with ϕ_{mi} values removed further from the value ϕ'_{mi}, surface states have a far less direct effect. This is demonstrated in Figure 3.12(b) where the I-V characteristics are shown for a relatively low ϕ_{mi} value (3.2 eV) again with several surface state treatments. The major role of surface states in this case is as charge storage centres. By comparing the case where surface states are neglected (N_{SS} = 0, σ_T = 0) to that where only their charge storage role is considered (σ_T = 0), it is seen that their main effect is to increase the slope of the depletion plateau region of the I-V characteristic since they increase the share of any applied voltage change absorbed across the insulator. Not until the tunneling capture cross section attains very large values (σ_T = $10^{-19} m^2$) is there any appreciable contribution to the total device current from the current flowing directly between the metal and the surface state levels.

3.4.3 Effects Upon Non-Equilibrium Diodes

The current-voltage characteristics for a thin insulator (20Å)

p type device are illustrated in Figure 3.13(a) for the case where large currents can tunnel between the metal and surface state levels. For minority and majority carrier diodes, the major effect is an enhanced slope of the depletion plateau region of the curves. At reverse and small forward bias, minority carrier diodes retain their N^+P junction behaviour. A cursory analysis might suggest that, since surface states provide additional recombination-generation sites, they would add an additional component to the N^+P diode current. However, the IS interface is clamped in inversion in this semiconductor limited regime and the presence of an inversion layer is known to suppress such recombination-generation processes [76]. In the case of majority carrier diodes, surface states provide an additional path for minority carriers to flow out of the surface region under reverse bias and enhance the non-equilibrium effect previously described for these devices (Section 3.3.1). They also have a strong influence upon the multiplication process associated with this effect (Chapter 5).

There is no sharp transition in properties between surface state diodes and minority or majority carrier diodes. Surface state diodes have properties which range from those similar to minority carrier diodes to those similar to majority carrier diodes. This range in properties corresponds to the variation of the IS interface from near inversion to accumulation in the bias region close to zero volts. In near minority carrier diodes, tunneling via surface states provides an additional path between the metal and the majority carrier band in the semiconductor. Current flowing via this path dominates the restricted minority carrier current for the non-equilibrium devices of Figure 3.13(a).

Figure 3.13: Computed effects of surface states upon the I-V and C-V characteristics of "non-equilibrium" diodes with the metal to insulator barrier height as the variable parameter. The substrates are 2 Ωcm p type <100> Si ($N_A = 7 \times 10^{21}$ m^{-3}), the oxdie is 20Å thick, surface states have the distribution of Figure 3.11, and the oxide charge, Q_{ox}, has the value 5×10^{15} m^{-2}. All surface state capture cross sections have the value σ_{co} (Table 2.4).

(a): I-V characteristics.

(b): High frequency C-V characteristics (f = 100 kHz).

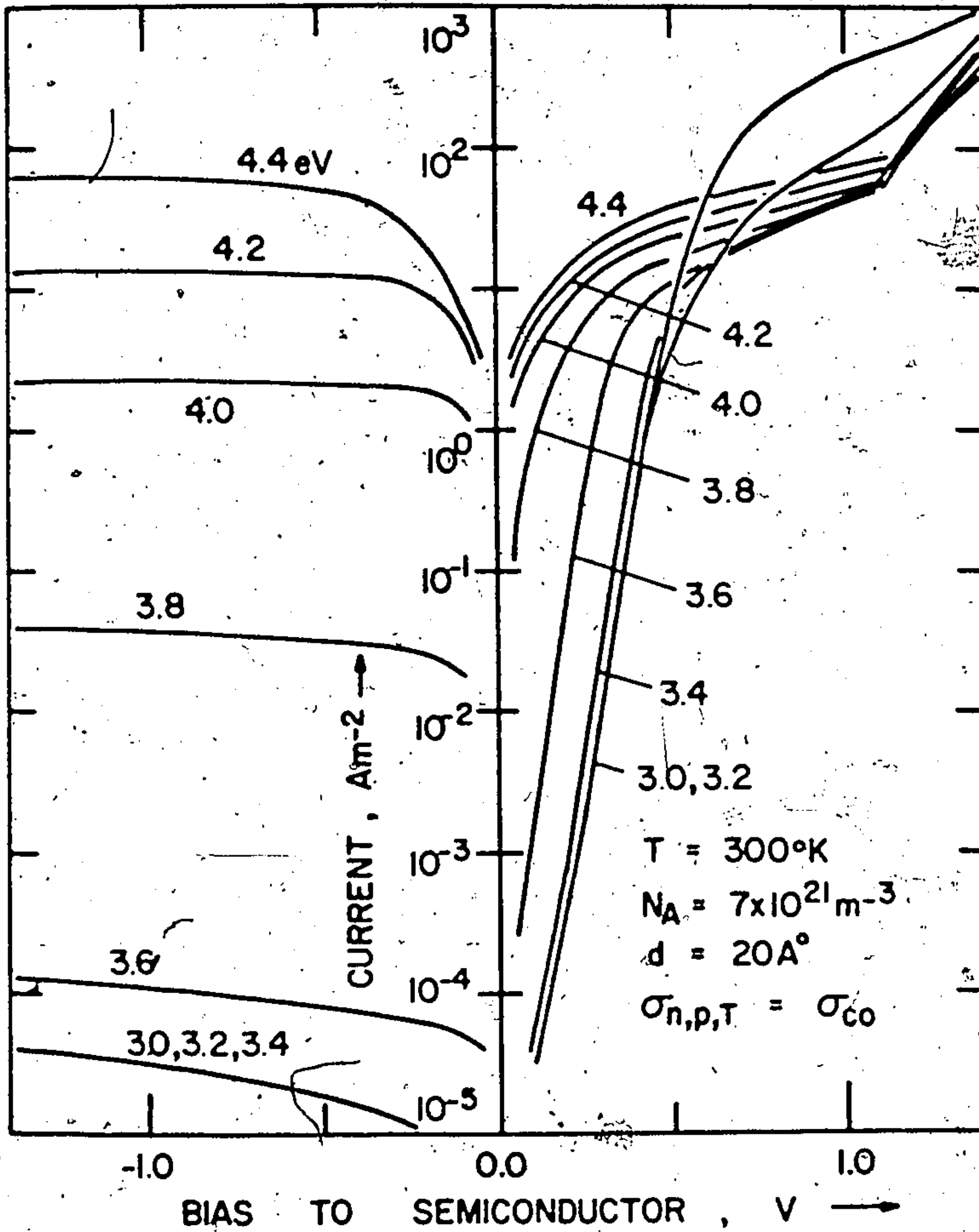


FIGURE 3.13(a)

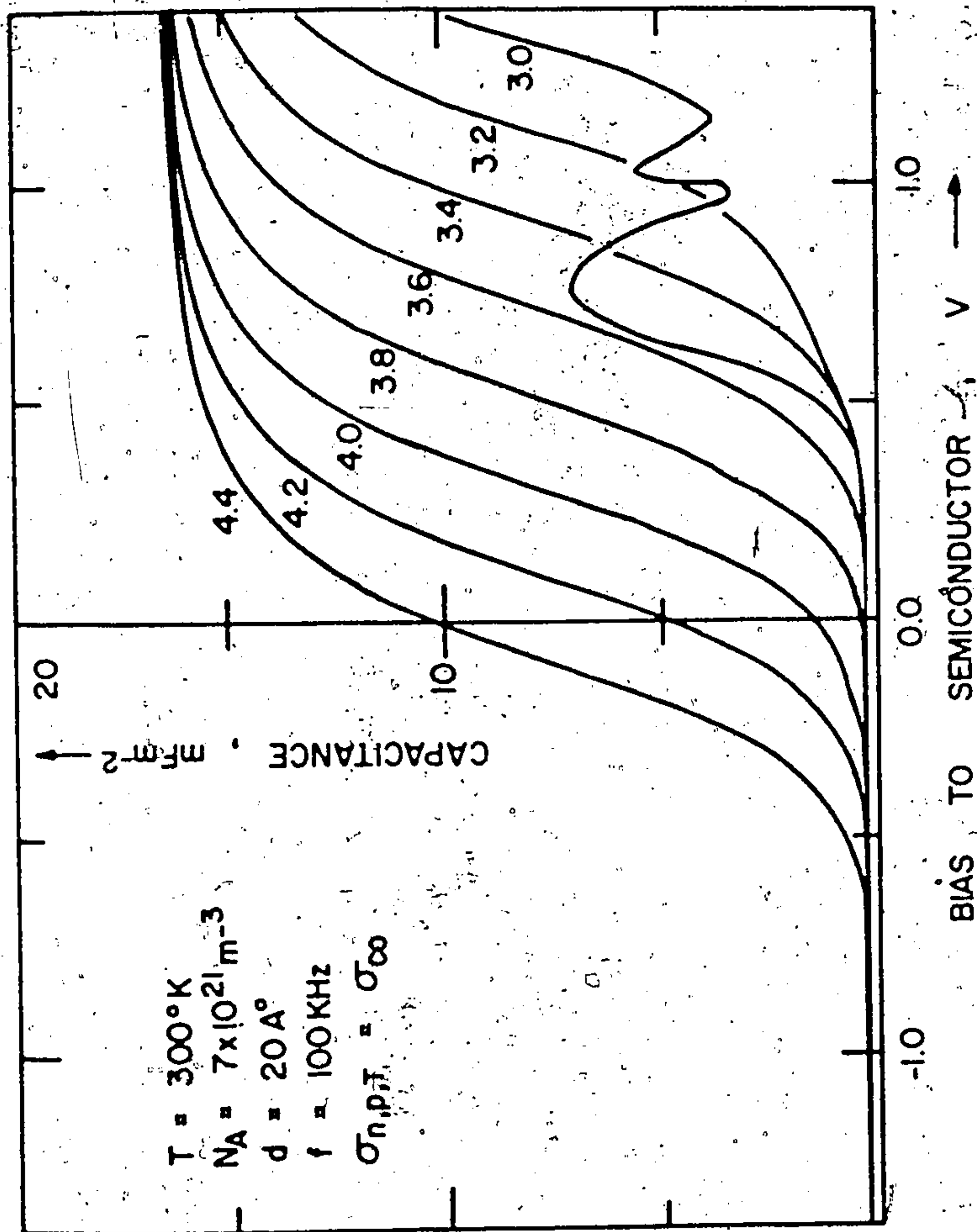


FIGURE 3.13(b)

The resulting current-voltage characteristics are demonstrated by the curves corresponding to ϕ_{mi} equal to 3.4 and 3.6 eV. In near majority carrier diodes, the inversion layer again does not form under reverse bias as demonstrated by the fact that all curves saturate under reverse bias to the semiconductor.

The capacitance-voltage characteristics corresponding to the I-V curves of Figure 3.13(a) are shown in 3.13(b) for a frequency of 100 kHz. Although this cannot be seen too clearly from Figure 3.13(b), all devices exhibit depletion mode characteristics under reverse bias. Additional frequency dependent structure has been introduced into these curves near the accumulation hump by the presence of surface states. For ϕ_{mi} greater than 3.4 eV, all curves have the same shape but are merely shifted along the voltage axis. For smaller values of ϕ_{mi} the curves have distinctly different shapes. In particular, the curve corresponding to ϕ_{mi} equal to 3.0 eV displays the distinct hump due to the delayed response of the diode inversion layer. The second peak is due to the surface state response. These curves demonstrate that the evaluation of surface state parameters from the C-V characteristics of non-equilibrium devices is not straightforward. The equilibrium device would appear far more suitable for this application.

3.5 Temperature Effects

Current flows in non-degenerate MIS tunnel diodes are generally not temperature insensitive. Temperature dependencies are evident in both equilibrium and non-equilibrium devices. In equilibrium devices, the band tunnel currents, J_{CT} and J_{VT}, are very temperature sensitive

when the semiconductor quasi fermi levels and the metal fermi level all lie at energies corresponding to the semiconductor band gap. This situation is demonstrated in Figure 3.3 where this condition is satisfied for the three bias conditions shown. It is most likely to occur if the value of ϕ_{mi} lies between ϕ_{si} and $\phi_{si} + E_{gs}$ when the IS interface is in depletion. Consequently, regions of the I-V characteristics corresponding to inversion and accumulation in equilibrium diodes are relatively temperature insensitive. The entire curves for high and low ϕ_{mi} values also have this property. J_{ST} is reasonably temperature insensitive under the conditions given for maximum sensitivity for the band tunnel currents. Consequently, surface state diodes have temperature insensitive characteristics.

The consequences of the above effects are illustrated in Figure 3.14(a) where current-voltage curves are compared for equilibrium diodes at 200 and 300°K with a range of ϕ_{mi} values. For low ϕ_{mi} (3.0 eV), the curve is temperature insensitive. Increasing ϕ_{mi} (3.0-3.4 eV) increases the temperature sensitivity in the depletion plateau until the device becomes surface state dominated. The device characteristics are then temperature insensitive (3.6 eV). The above temperature effects provide a possible method for distinguishing surface state dominated diodes experimentally.

In non-equilibrium diodes, the effect of temperature is to greatly influence minority carrier flow rates through the semiconductor. The current flow in minority carrier diodes in the semiconductor limited regime is very temperature dependent as in p-n junction devices. In

Figure 3.14(a): Computed effects of temperature upon the I-V characteristics of "equilibrium" MIS tunnel diodes. The semiconductor is p type ($N_A = 7 \times 10^{21}$ m^{-3}), the insulator thickness is 40Å, the surface state distribution is as in Figure 3.11, and Q_{ox} has the value 5×10^{15} m^{-2}. The surface state capture cross sections for electrons and holes habe the value σ_{co} while the tunneling capture cross section has the value σ_{TO} (Table 2.4). The variable parameter is ϕ_{mi} in eV. The solid curve corresponds to 300°K while the dashed curve is for a temperature of 200°K.

(b): Effects of temperature upon the I-V characteristics of an n type MIS tunnel diode ($N_D = 5 \times 10^{20}$ m^{-3}) with an oxide thickness of 34Å and ϕ_{mi} equal to 3.2 eV. The equilibrium to non-equilibrium transition is demonstrated as the temperature changes from 300°K to 200°K. The effects of oxide charge and surface states are neglected.

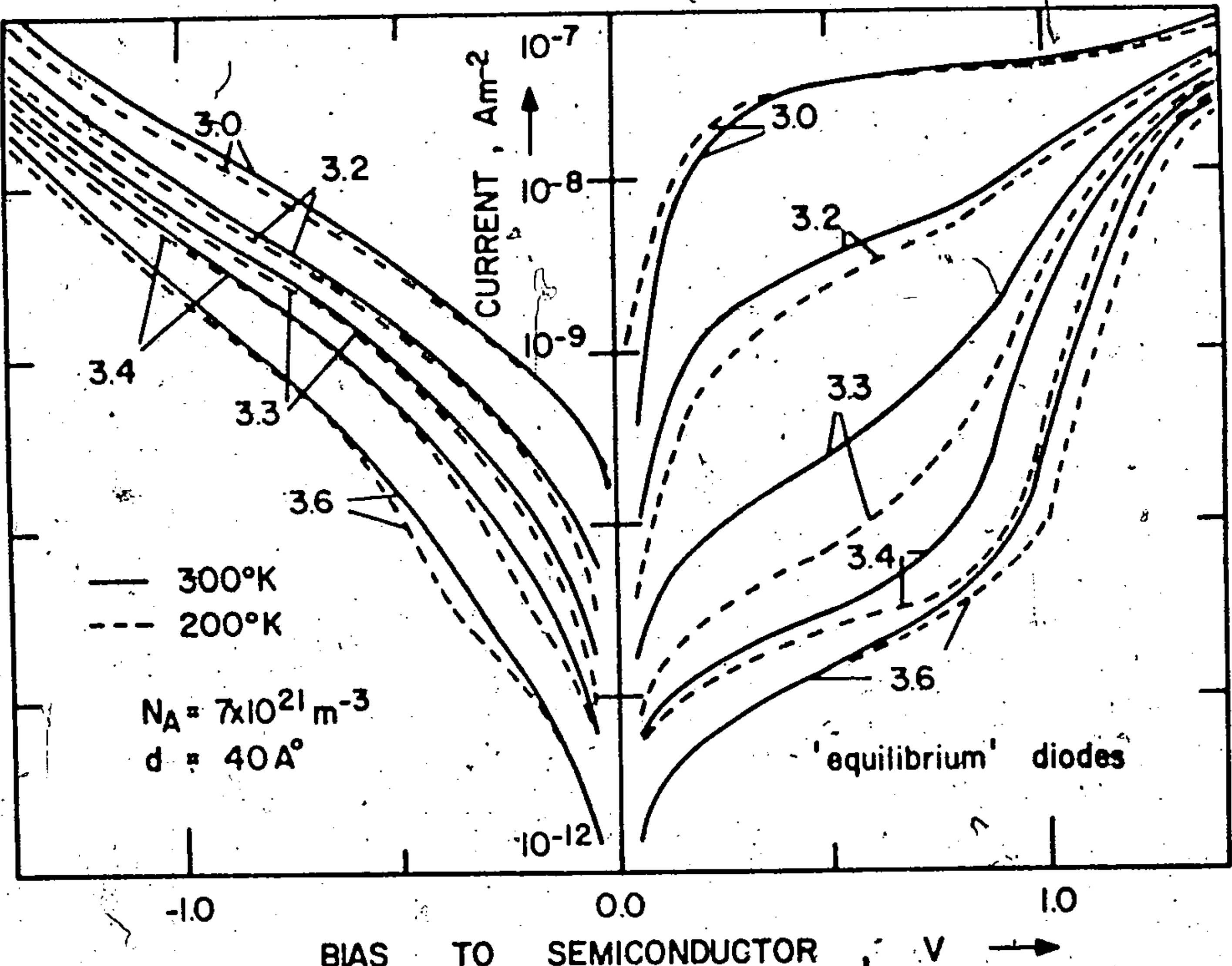


FIGURE 3.14(a)

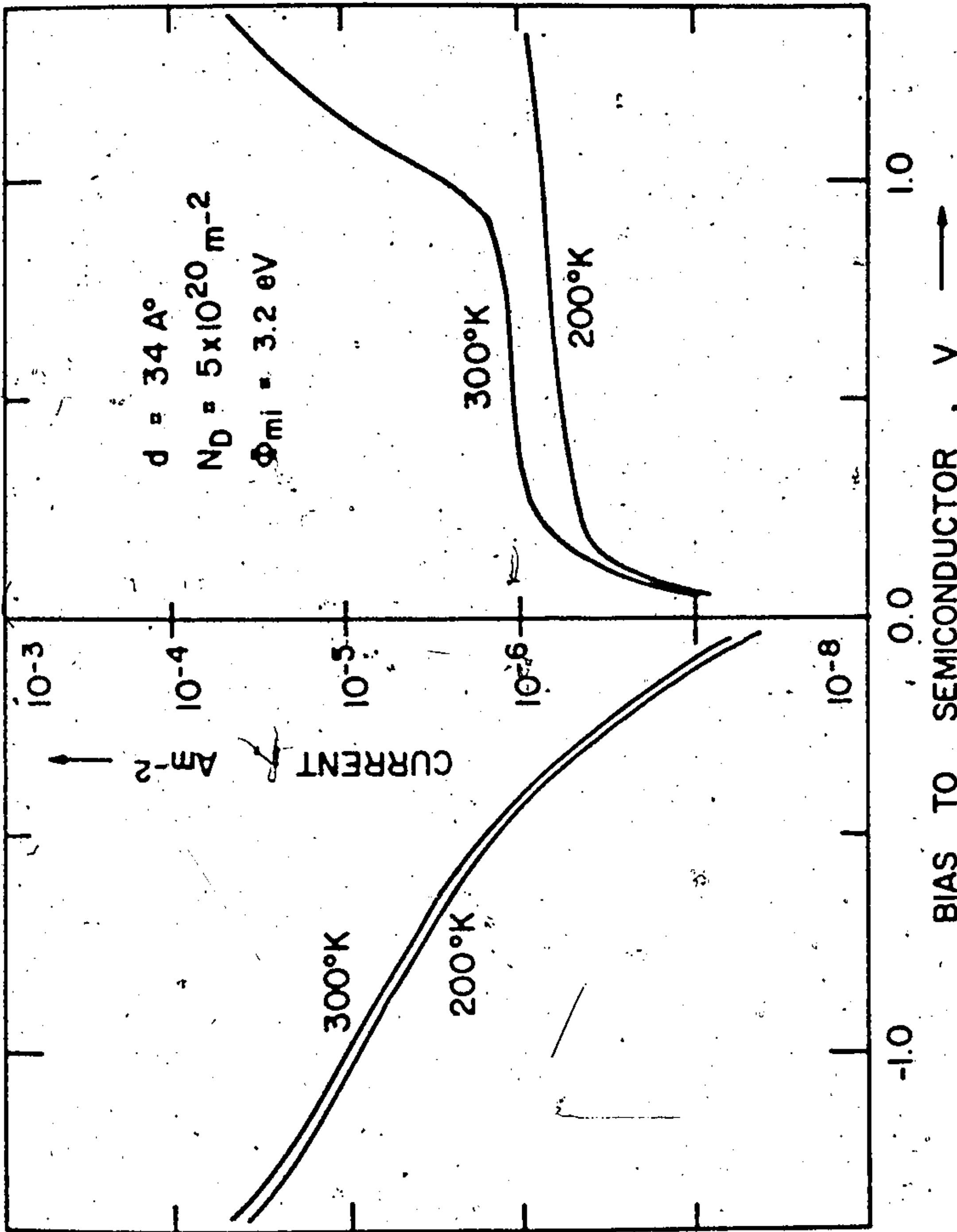


FIGURE 3.14(b)

tunnel limited regions, the temperature dependence is similar to that described for equilibrium devices. For majority carrier non-equilibrium devices, the major temperature dependence occurs in the region of the I-V characteristics corresponding to depletion at the IS interface as in equilibrium devices.

One additional effect of varying the device temperature arises from the associated large variation in the minority carrier flow rates in the semiconductor. It is possible to observe a diode undergoing an equilibrium to non-equilibrium transition merely by changing the device temperature. This is demonstrated in Figure 3.14(b) where the I-V characteristics of an n type diode of doping density $2 \times 10^{21} m^{-3}$ with oxide thickness 34Å are plotted at 200°K and 300°K. At 300°K, the minority carrier supply rate under inversion bias (positive to the n type semiconductor) is sufficiently high to maintain the inversion layer and the device displays equilibrium characteristics. At 200°K, the supply rate has been reduced considerably, the inversion layer cannot be maintained, and the device displays non-equilibrium characteristics.

In Section 4.5, it will be shown how the temperature dependence of the diode surrent in the depletion plateau region of the I-V characteristics can be used to evaluate the device parameters for band current dominated devices.

3.6 Summary

This chapter has established a broad theoretical framework for interpreting the properties of MIS tunnel diodes as the device parameters

are varied. Special emphasis has been placed upon the M-SiO_2-Si system although the general trends described and the conclusions reached should be applicable to other systems.

Central to the discussion is the division of diodes into majority carrier, surface state, and minority carrier devices. Although these classes are defined in terms of the energy distribution of the current flowing between the metal and the semiconductor, it is shown that devices fall into these categories depending upon whether the IS interface at zero bias is accumulated, depleted or in strong inversion, respectively. The metal to insulator barrier height (which is related to the metal work function) is the most important system parameter determining the state of this interface under zero bias. The division of the M-SiO_2-Si system into these classes is indicated in Figure 3.15 as a function of this parameter. The uncertainty in the boundary between adjacent classes is due to its dependence upon surface state parameters, the device temperature, and the insulator thickness in this order of importance.

Each class is subdivided into equilibrium and non-equilibrium diodes depending upon the level of current flow which depends primarily on the insulator thickness. Equilibrium diodes exhibit capacitance-voltage characteristics similar to conventional MIS capacitors. Consequently, the most striking feature is the accumulation hump. In the voltage range of interest, the most important feature of their current-voltage characteristics is the hump evident when the IS interface becomes accumulated for minority carrier diodes and inverted for majority carrier diodes. Non-equilibrium diodes are the devices of principal interest

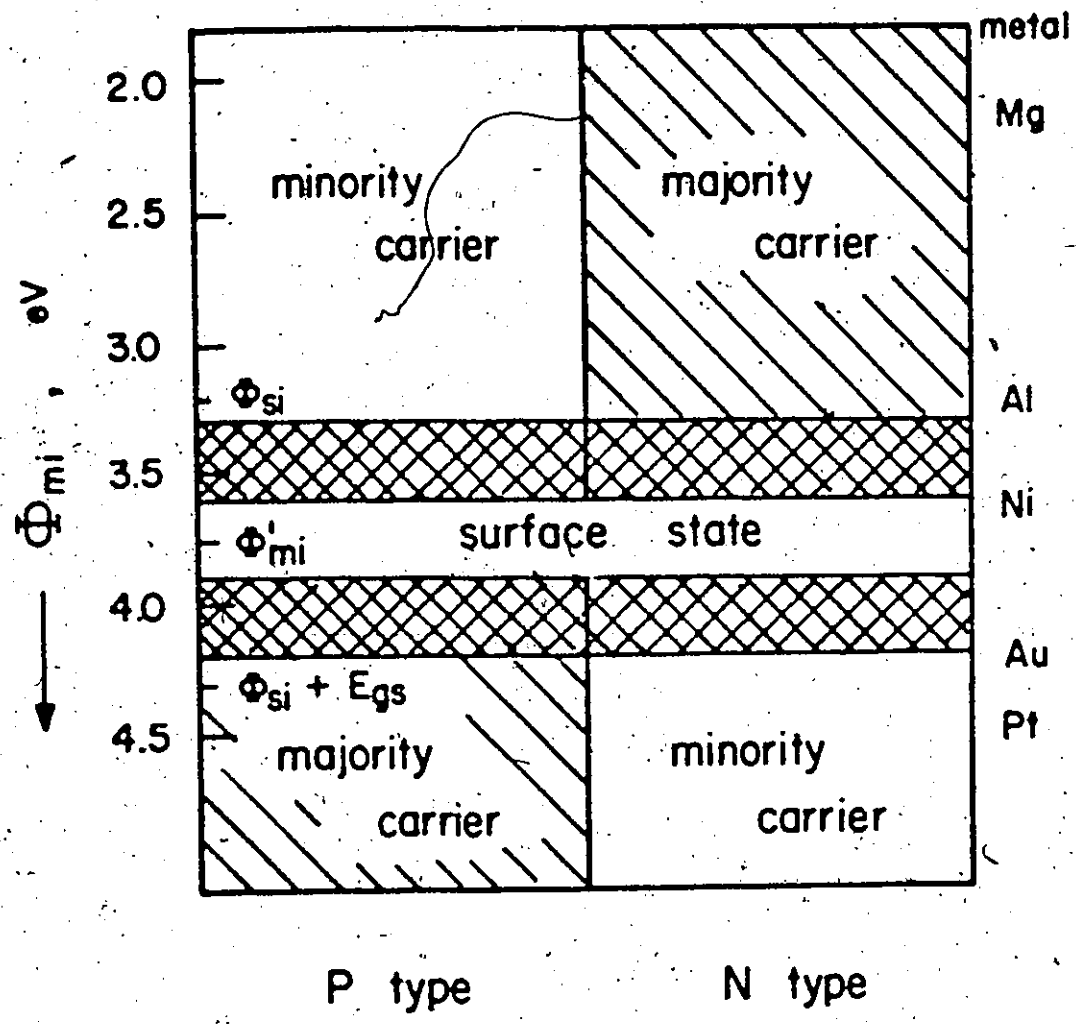


Figure 3.15: Schematic showing the division of the MIS tunnel diode into minority carrier, surface state and majority carrier diodes as a function of ϕ_{mi} (the metal to insulator barrier height) for both n and p type substrates. The heavily shaded regions represent the likely range of the transition region to surface state diodes. The value of ϕ_{mi} at which this transition occurs will depend largely upon the surface state densities.

in this thesis. In majority carrier non-equilibrium devices, the inversion layer at the IS interface never forms without some external stimulus due to limitations upon the rate that the semiconductor can supply minority carriers. This non-equilibrium effect produces the multiplication mechanism described in Chapter 5. Minority carrier non-equilibrium diodes have characteristics identical to p-n junction diodes under reverse and small forward bias. This arises because in each case the process limiting current flow is the transport of minority carriers through semiconductor regions. These diodes are described in more detail in Chapter 6 where they are shown to be well suited for direct energy conversion applications. Surface state non-equilibrium diodes have properties that vary from those similar to minority carrier diodes to those similar to majority carrier diodes. All non-equilibrium devices exhibit capacitance-voltage characteristics which vary as $1/C^2 \propto V$ under reverse bias. Minority carrier non-equilibrium diodes also exhibit a form of inversion layer response under moderate forward bias up to very high frequencies which produce large frequency dependent humps in the capacitance-voltage characteristics.

CHAPTER 4

EXPERIMENTAL PROPERTIES OF THE MIS TUNNEL DIODE

4.1 Fabrication and Measurement

There are three techniques available for fabricating the MIS thin insulator structure. In the first method, the insulating oxide is grown thermally or anodically on a semiconductor substrate and the metal is subsequently deposited onto this oxide layer. This method produces the best devices since, as well as producing a high quality insulating layer, the semiconductor substrate can be single crystalline. It has been demonstrated for Si [2,11-20] and InSb [77]. In the second method, the metal is deposited onto a suitable substrate, oxidized, and the semiconductor is then deposited on this metal oxide layer. Presently, the semiconductor layer of devices fabricated by this process have an amorphous or polycrystalline structure with properties inferior to the single crystalline layers possible with the first method. Even so, this method could be the most suitable for the production of inexpensive large area devices required for the application of the class of diode described in Chapter 6 to terrestial solar energy conversion. It is also the best method available for semiconductors on which a suitable oxide cannot be grown. It has been demonstrated experimentally for SnTe [3], GeTe [10], Ge [78], and CdS [79] in all cases using Al and its oxide as the metal and insulator layers. The third method produces a poorer quality insulating layer. It involves the deposition of the insulating layer rather than its

growth. This layer would normally be deposited on a single crystal semi-conductor substrate. Although it would be expected to produce the poorest quality tunneling device, the method is of interest due to its practical application in the production of electroluminescent diodes using CdS with a variety of insulators [80]. It has also been used for Si [20].

The first method mentioned has been used exclusively in this thesis since it produces the highest quality devices. Thin oxides (15-60Å) were grown in steam at temperatures in the range 700-900°C on chemically prepared [81] n and p type wafers of <111> and <100> orientated silicon obtained from Monsanto. Metal top contacts were evaporated onto this oxide through a metal mask in an NRC vacuum evaporation system equipped with a liquid nitrogen trapped oil diffusion pump of base pressure 5×10^{-7} Torr. After removal of the oxide from the unpolished bottom side of the Si wafer, aluminum was evaporated onto this region to form a large area ohmic back contact. Because of the irregular nature of the wafer surface on its unpolished side and the consequent large density of recombination centres in this region, such contacts exhibited satisfactorily ohmic characteristics to current densities usually as high as 300 A/m^2. This was adequate for most purposes. Improved contacts for higher current density operation of the p type energy conversion diodes of Chapter 6 were obtained by evaporating and sintering the Al back contact after the oxide growth but before the evaporation of the top metal contact.

Devices fabricated by such a procedure tended to attract electrons to areas of the IS interface remote from the top metal contact presumably due to positive charge in the oxide. This resulted in the device having

one of the six possible lateral distributions of surface charge described in Section 2.6.1. If desired, the effect could be substantially reduced by additional processing. The simplest method used was to expose the diode to the room atmosphere for several hours. This could be accelerated by heating for 10 minutes at 100°C in such an atmosphere. Immersion of the diode in weak sodium dichromate solution and heating at 60°C after removal from this solution was also effective [64].

The thickness of the oxide layer was determined either by ellipsometry [82] or, for the thicker oxide devices, by measuring the accumulation capacitance of the device. The two methods gave results which agreed to within 2Å. The semiconductor resistivity and the lower bound on the minority carrier lifetime were specified by the supplier. Measurements of the latter quantity after processing using the diodes of Chapter 6 and the switching method of Lederhandler and Giacoletto [83] gave values in the range of 2 to 10 μsecs.

The capacitance-voltage (C-V) characteristics of the devices were evaluated continuously over the voltage range of interest using a phase sensitive detection system described elsewhere [79]. The current-voltage (I-V) characteristics were evaluated point by point using a Keithley picoammeter (range ~ 10^{-13} to 10^{-2}A).

Using the above procedure, between one hundred and two hundred different n and p type devices were fabricated and evaluated. The majority employed an Al top contact not only because of the features which make it the most important metal contact in conventional MIS technology, but also because its value of work function makes it optimal for studying the multiplication properties of the n type devices described in Chapter 5

and ideally suited for the energy conversion application of Chapter 6 using p type devices. Devices with Au, Ni, Pt, Ti and Mg contacts were made in smaller quantities to demonstrate features of the theory given in Chapter 3. In all cases the top contact had a circular geometry but the contact area varied from $1.1 \times 10^{-7} m^2$ (0.015" diameter), which was about the smallest size which could be conveniently probed without optical equipment, to $3 \times 10^{-4} m^2$ ($\sim$ 0.8" diameter), which approached the largest area device possible on a 1" silicon slice. The smaller contacts were deposited in an array pattern with up to one hundred contacts for a single wafer.

4.2 Interface Concentrations at Zero Bias

In view of the importance assigned in Chapter 3 to the zero bias concentrations of electrons and holes at the IS interface, it is appropriate to demonstrate how these vary with the choice of metal contact in experimental devices. It is well known [1c] that they show relatively little variation in the related Schottky diode structure. In fact, the control over these concentrations possible in the MIS tunnel diode is one of two factors contributing to the wide variety of properties which can be obtained compared to the Schottky diode (the other factor is the way in which the applied voltage is shared between the insulator and the semiconductor in the MIS device).

The simplest method for demonstrating this control is to measure the C-V characteristics for devices with different metal contacts. The location of the accumulation humps in these characteristics can then be used to determine whether the IS interface is accumulated, depleted, or

inverted at zero bias. Note that it has been shown (Figure 3.7) that the voltage at which the accumulation hump establishes itself is not strongly dependent upon the insulator thickness or whether the diode is an equilibrium or non-equilibrium device.

The accumulation region of the measured C-V characteristics of M-SiO_2-Si devices with both n and p type <100> substrates of resistivity 1-3 Ωcm and oxide thickness in the range 20-50Å are shown in Figure 4.1. From the location of the accumulation humps, it can be seen that, at zero bias, the IS interface of both n and p type devices can be varied over the entire range from accumulation to inversion by selecting an appropriate metal contact. Another feature of these curves is the absence of any gross deformities due to surface state effects. Throughout this thesis, surface state densities were observed to be at the lower end of the range of values reported elsewhere [17,84-86] for similar devices. This is attributed to the predominant use of <100> orientated silicon rather than <111>. Previous measurements on thick insulator devices indicate smaller densities for the former orientation [1d] as do measurements upon tunnel devices similar to those being reported here [15]. The relatively small densities of surface states allows the metal to insulator work functions to be estimated reasonably accurately from the position of the accumulation humps for both n and p type devices. Values estimated in this way are compared to the values calculated from the metal vacuum work function in Table 4.1. Similar evaluations for other metals have been performed recently by Kar and Dahlke [17].

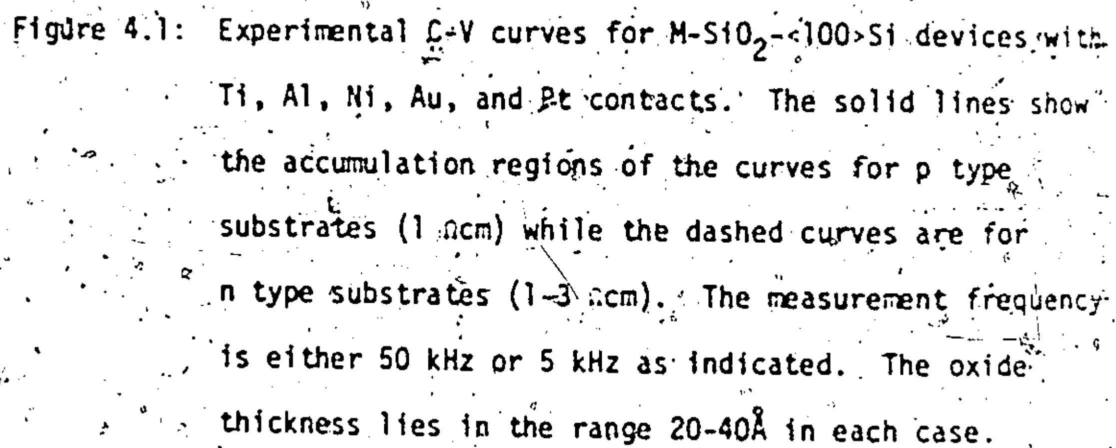


Figure 4.1: Experimental C-V curves for M-SiO_2-<100>Si devices with Ti, Al, Ni, Au, and Pt contacts. The solid lines show the accumulation regions of the curves for p type substrates (1 Ωcm) while the dashed curves are for n type substrates (1-3 Ωcm). The measurement frequency is either 50 kHz or 5 kHz as indicated. The oxide thickness lies in the range 20-40Å in each case.

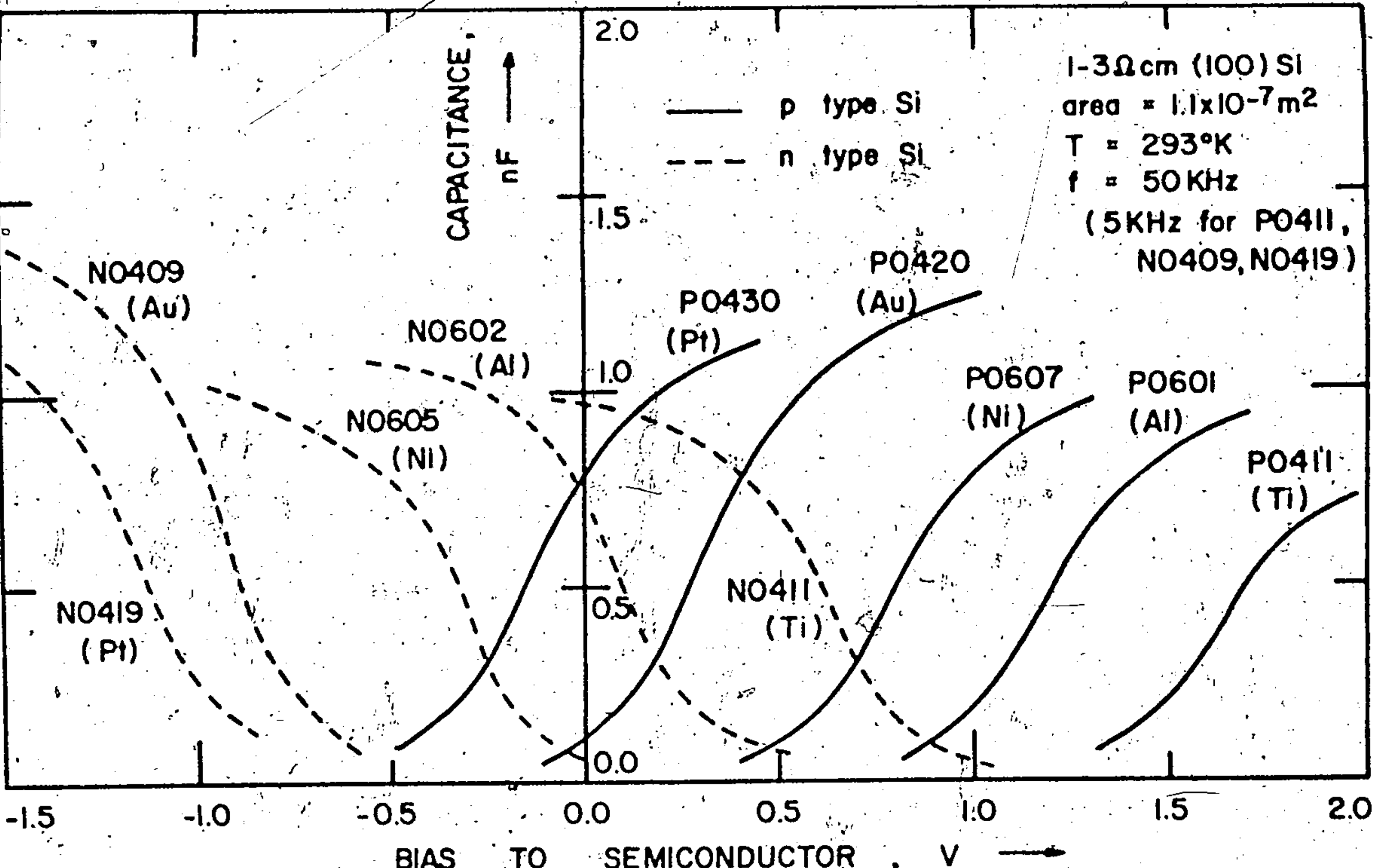


FIGURE 4.1

TABLE 4.1: Metal to Insulator Barrier Heights for Ti, Al, Ni, Au, and Pt.

Metal	Average Value From C-V Curves (eV)	Value Calculated From Vacuum Work Function*
Ti	2.7	3.0
Al	3.2	3.3
Ni	3.6	3.6
Au	4.2	3.9
Pt	4.5	4.4

* Obtained by subtracting the electron affinity of SiO_2 (0.9 eV) from the vacuum work function as given in [58].

4.3 Comparison of n and p Type Devices (Equilibrium and Non-Equilibrium Cases)

One prediction of Chapter 3 was that equilibrium devices with the same metal contact and similar insulator thicknesses should have qualitatively similar I-V characteristics independent of whether the substrate is n or p type since the same tunnel currents dominate in each case. This holds provided the substrate is not too heavily doped (Figure 3.5). The above feature was emphasized since, if a device on one type of substrate is a majority carrier device, it follows that one on the other substrate type must be a minority carrier device. This behaviour is demonstrated experimentally by the solid curves of Figure 4.2 which represent the measured I-V curves for a pair of Al-SiO_2-Si devices, one n type (N0511) and the other p type (P0511). The C-V curves for these devices can be found in Figure 4.3(c) which is to be discussed later. Since the devices have relatively thick oxides (37 and 39Å respectively), they display equilibrium characteristics. The most striking feature is the hump occurring near 1 V bias to the semiconductor where the surface of the p type device is becoming accumulated (Figure 4.3(c)) and conversely, that of the n type device is becoming inverted.

The dashed curves of Figure 4.2 show the characteristics for another pair of devices with Al contacts (N0602 and P0601) where the oxide thickness in each case has been reduced by about 6Å. This relatively small change has produced marked changes in the I-V characteristics. Under reverse bias to the semiconductor and forward bias to about 0.8 V, the curve for the n type device (N0602) has the same shape as for the previous equilibrium devices but has been displaced to current densities

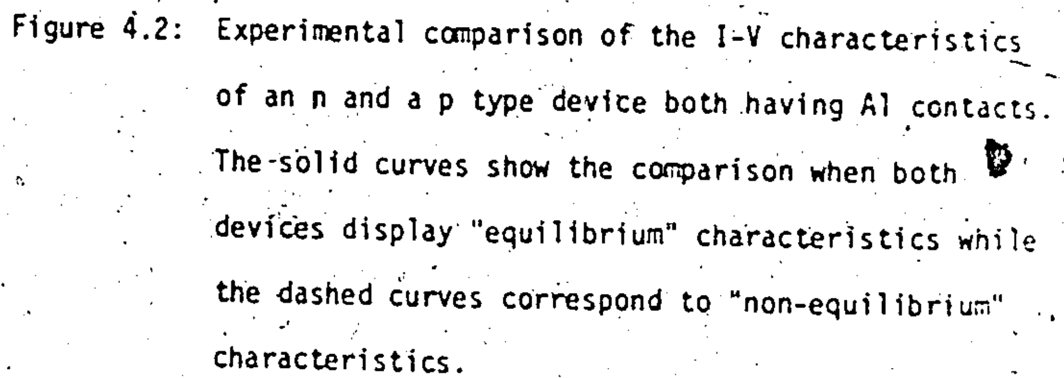


Figure 4.2: Experimental comparison of the I-V characteristics of an n and a p type device both having Al contacts. The solid curves show the comparison when both devices display "equilibrium" characteristics while the dashed curves correspond to "non-equilibrium" characteristics.

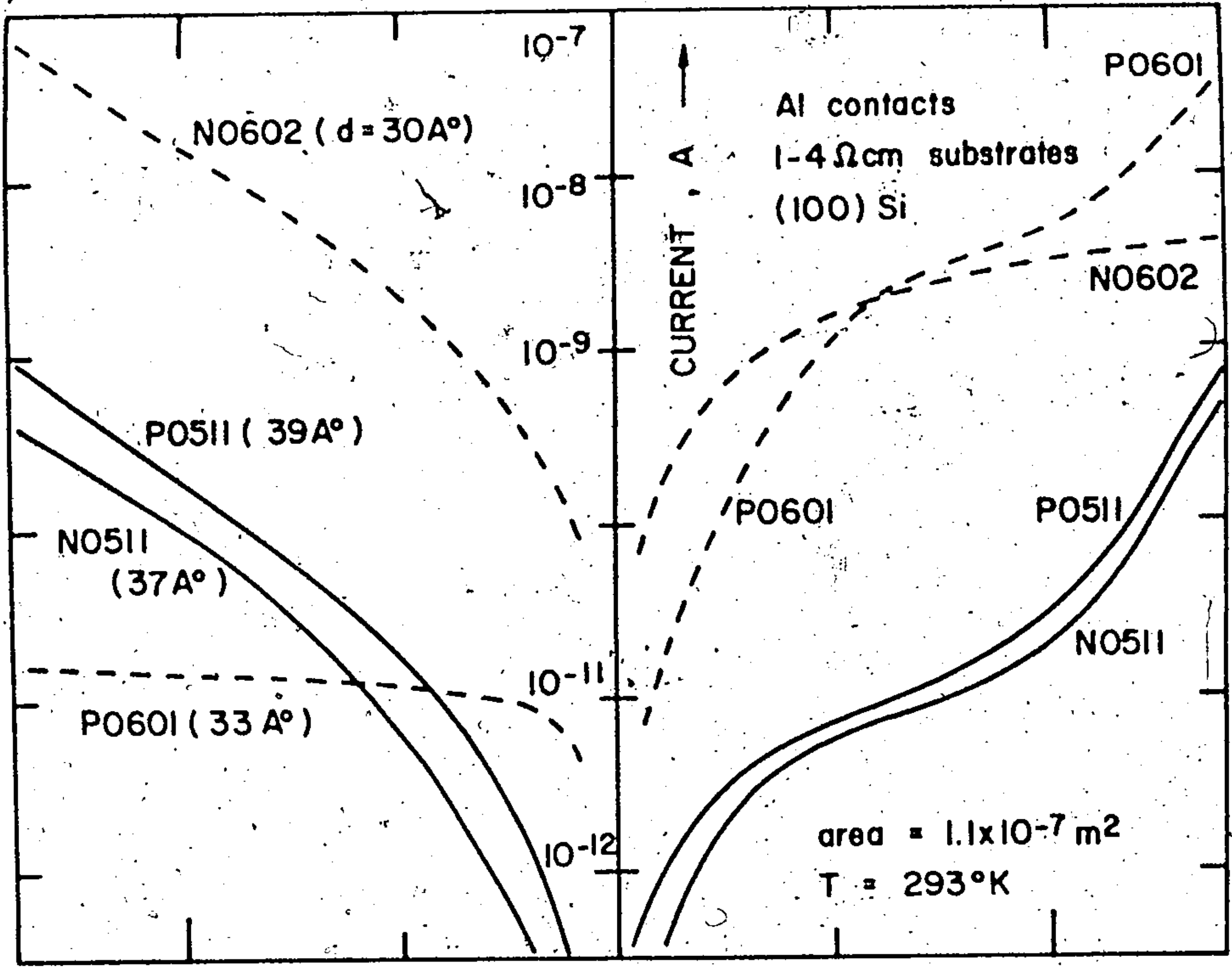


FIGURE 4.2

2-3 orders of magnitude larger. Above 0.8 V, the hump has completely disappeared. This is the predicted non-equilibrium effect for majority carrier devices (Figure 3.6(b)). The supply of holes to the IS interface is limited under this inversion bias direction for the n type device and holes trying to form an inversion layer quickly tunnel to the metal either directly or via surface states. The change in behaviour for the p type device (P0601) is even more striking. The current under reverse bias bears no relation to the previous curves. In fact, it merely reflects the current which can be generated in the depletion region of the device under such biases. It is determined entirely by the properties of the semiconductor. More importantly, a similar effect occurs under forward bias. At small forward bias, the device is also semiconductor limited in this case to the rate of recombination in the semiconductor depletion region (Chapter 6). As the diode bias increases, this rate increases rapidly. Consequently, by about 0.5 V the device current is again entirely tunnel limited as in the equilibrium diode. The distinct hump due to accumulation at the IS interface is therefore in evidence. These results are as predicted for minority carrier devices (Figure 3.6(a)). Some additional comments will be made about current flow under reverse bias for such p type Al devices in the following section.

In summary, the curves of Figure 4.2 demonstrate two important features of the theory of Chapter 3. Firstly, they demonstrate the similarity between the I-V characteristics of n and p type devices with the same metal contact in the equilibrium case. Secondly, the curves demonstrate the important non-equilibrium effects for both majority and minority carrier devices. These are the disappearance of the inversion

hump for majority carrier devices and the semiconductor limited nature of the device current at reverse and small forward bias for the minority carrier devices. In addition, the general shape of the characteristics are in good agreement with the theoretical expectations. This is demonstrated by the comparisons between theory and experiment contained in the next section where the major purpose will be to investigate the experimental consequences of varying the work function of the metal top contact.

4.4 Effect of Metal Work Function (Equilibrium and Non-Equilibrium Diodes)

The ability to produce minority and majority carrier devices by selecting the substrate as p type or n type has been demonstrated in the preceding section. In this section, it will be shown that a similar transformation in properties can be obtained by changing the metal contact from one with a relatively low work function metal (such as Al) to one with a high work function metal (such as Au). The approach employed will be to establish the properties of the Al-SiO_2-Si system by describing the measured characteristics of a number of these devices on both n and p type substrates with a range of oxide thicknesses. These will be compared with those established in a similar way for the Au-SiO_2-Si system. Finally the properties of other devices with either low (Mg,Ti) or high (Pt) metal contacts will be described.

In Figure 4.3, the measured characteristics of a number of devices with low work function Al contacts are displayed for values of the oxide thickness ranging from 23 to 39Å. Figure 4.3(a) shows the I-V characteristics for p type substrates, (b) for n type substrates, and (c) shows the C-V characteristics for the devices of both (a) and (b). The dashed

Figure 4.3: Experimental I-V and C-V characteristics for Al-SiO_2-<100>Si devices for oxide thickness in the range 23-39Å. The dashed curves are theoretical curves.

(a): Measured I-V curves for 1 Ωcm p type devices. The dashed theoretical curve is calculated with d equal to 23.5Å, N_A equal to $7 \times 10^{21} m^{-3}$, ϕ_{mi} equal to $3 \times 10^{15} m^{-2}$ and the surface state distribution of Figure 3.11. Other parameters are given in Table 2.4.

(b): Measured I-V curves for 1-3 Ωcm n type devices. The theoretical dashed curve is calculated with the same parameters as in (a) except that d has the value of 22Å and N_D is equal to $2 \times 10^{21} m^{-3}$.

(c): Measured C-V curves for the p and n type devices of (a) and (b). The measurement frequency is either 50 kHz or 5 kHz as indicated. The dashed curves are calculated at 50 kHz using the same parameters given in (a) and (b) for the p and n type device respectively.

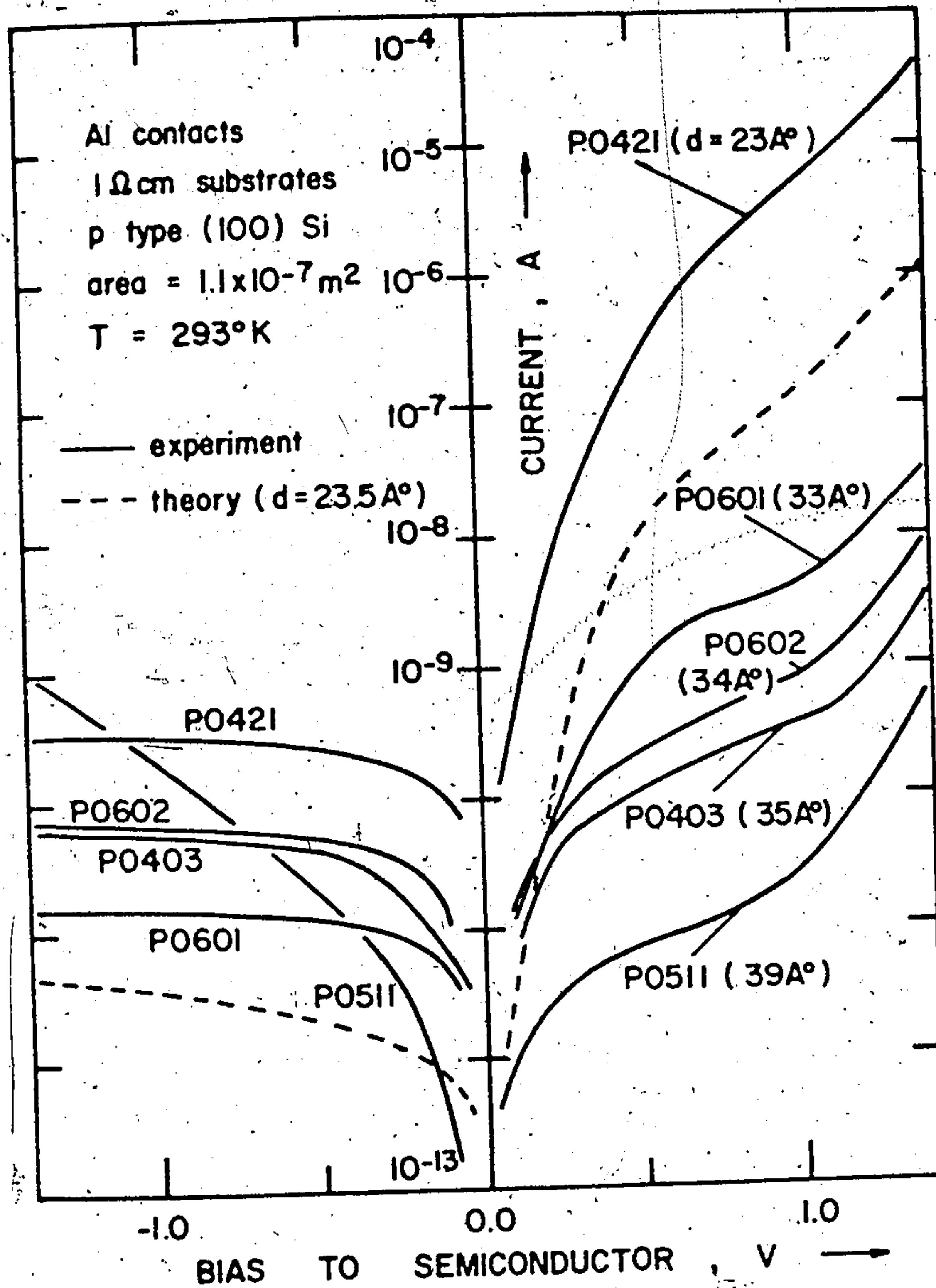


FIGURE 4.3(a)

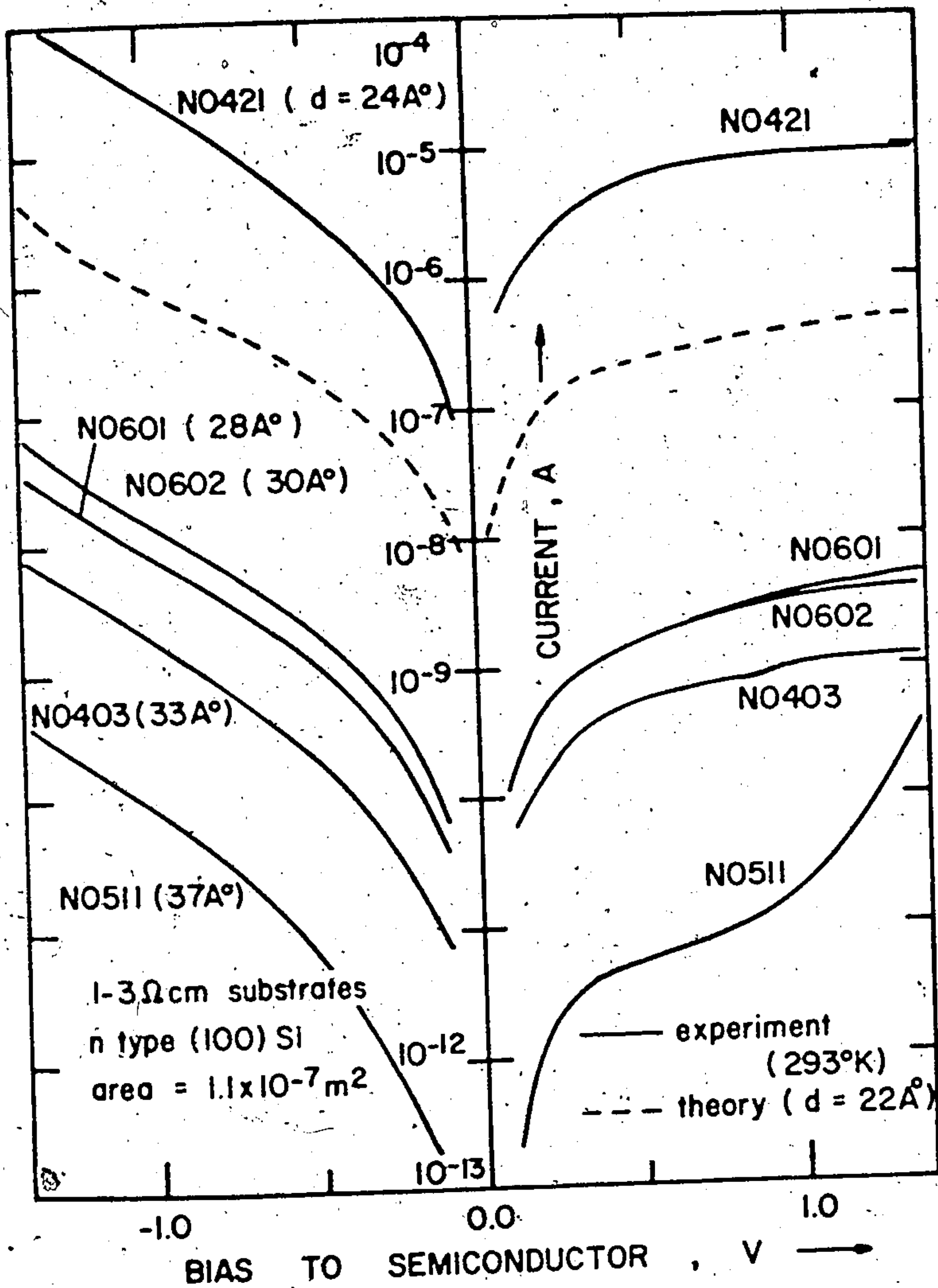


FIGURE 4.3(b)

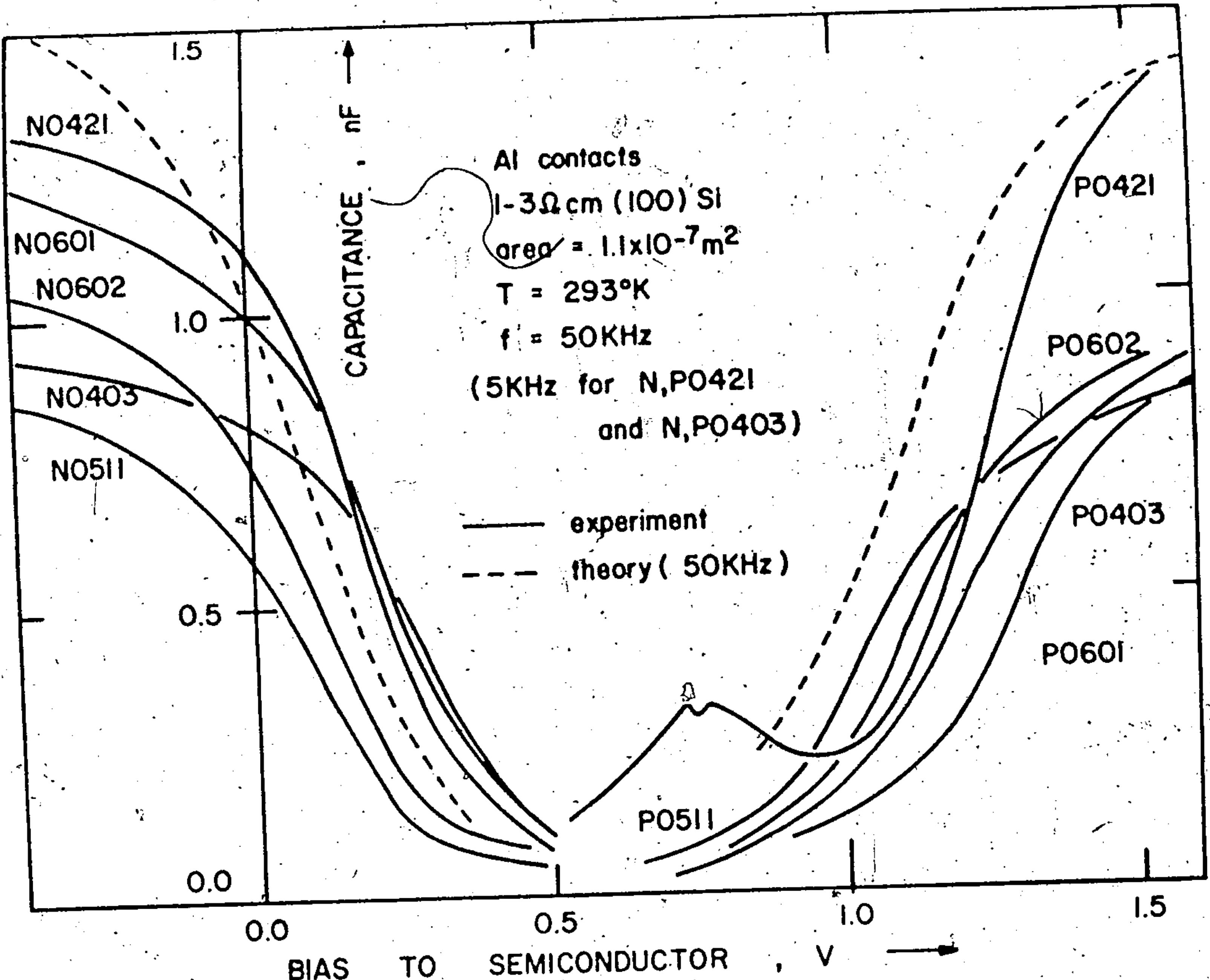


FIGURE 4.3(c)

curves in each case are theoretical curves calculated using the numerical simulation method of Chapter 2 and the typical values of the device parameters given in the figure caption.

The p-type devices of Figure 4.3(a) will be considered first. The set of theory curves indicating the expected behaviour of these devices most appropriately are those for the minority carrier devices of Figure 3.6(a). At moderate to large forward bias, the diode current should increase rapidly with decreasing insulator thickness as expected for a tunneling limited process. At reverse and small forward bias, semiconductor limited effects should become apparent for thinner insulator non-equilibrium devices. Non-equilibrium devices can be distinguished from equilibrium devices because they have small values of reverse current compared to current flows at forward bias.

On this basis, devices P0421 (23Å) and P0601 (33Å) as well as the theoretical curve (23.5Å) display non-equilibrium characteristics. Devices P0602 (34Å) and P0403 (35Å) are in transition between equilibrium and non-equilibrium behaviour while device P0511 (39Å) displays equilibrium characteristics. The most striking departure of the experimental curves from those expected occurs under reverse bias. Theory predicts that, under reverse bias, all these devices should have current flows less than or equal to the dashed curve which reflects the generation rate in the semiconductor depletion region. The observed deviation is attributed to the lateral effects described in Section 2.6.1 which were not included in the theoretical analysis. The appropriate configuration near zero bias will be either "5" or "6" of Figure 2.8. In each case, the semiconductor depletion region is considerably larger than that

modelled by the theoretical treatment allowing much larger generation rates. In addition, it was found that, in the lateral configuration "6", the surface inversion layer could provide a resistive DC path for minority carriers to the back contact. This explains why current flows in device P0511 can attain relatively large values under reverse bias withouth becoming semiconductor limited. The above conclusions regarding the importance of lateral effects was reached after it was observed experimentally that the reverse current flow in these p type Al devices could be varied by changing conditions in the "lateral" regions using the treatments mentioned in Section 4.1. Note that the lateral effects will also enhance recombination rates in the semiconductor under small forward bias.

Taking the above effect into account, the experimental curves are in good qualitative agreement with theory. The devices with the thinnest oxides display semiconductor limited effects at reverse and small forward bias and tunnel limited current flow at moderate and large forward bias. All devices display the depletion plateau and accumulation hump regions under forward bias. Note that with the standard set of tunneling parameters used for all the theoretical calculations in this thesis (Table 2.4) there is only qualitative agreement between the predicted and observed current densities for a given oxide thickness in the regions where the device current is tunnel limited. However, for devices where the lateral effects mentioned above are made small by additional processing (Section 4.1), it is possible to choose an individual experimental curve and obtain a very accurate fit to it by an appropriate selection of the theoretical device parameters. Such a

procedure is demonstrated in Section 6.3.2 where both the I-V and C-V properties of an experimental device are generated to a high degree of accuracy by the device simulation program [102].

The n type devices of Figure 4.3(b) also display characteristics ranging from equilibrium to non-equilibrium. Device N0421 (24Å) displays typical majority carrier non-equilibrium characteristics, device N0403 (33Å) represents a transition between equilibrium and non-equilibrium behaviour with the inversion hump just starting to appear, while device N0511 (37Å) displays equilibrium characteristics with the inversion hump present. With negative bias applied to the semiconductor, the shape of the I-V curve is virtually the same regardless of the insulator thickness. This shape is well described by the theoretical dashed curve as is the shape of the characteristics of the non-equilibrium devices under positive semiconductor bias. Note that device N0601 with an oxide on the average thinner than that of device N0602 as indicated by capacitance measurements has slightly smaller current flows over the negative bias region. This can be explained by postulating that its oxide is slightly more uniform, so that both devices have similar tunneling thicknesses.

The C-V curves of Figure 4.3(c) differ from the expected curves (Figure 3.7) in that there is some spread in the voltage at which the accumulation hump appears. This scattering is caused by variations in either the metal work function, the semiconductor to insulator barrier height, the amount of charge in the oxide, or the surface state densities. The experimental results indicate that a large component of the observed spread is due to the first effect. Under accumulation conditions, the combination of surface states and oxide charge makes its maximum and

minimum contribution (Q_{ss}^{+}, Q_{ss}^{-}) for p and n devices respectively. By defining

$$\phi_{ms} = \phi_{mi} - \phi_{si} \tag{4.1}$$

and fitting the location of the accumulation humps to theory, the following relationship is obtained for the p type-Al devices measured

$$\phi_{ms} - \frac{qd}{\epsilon_i} Q_{ss}^{+} = -0.15 \pm 0.10 \text{ eV} . \tag{4.2}$$

For n type Al devices, a similar relationship is found

$$\phi_{ms} - \frac{qd}{\epsilon_i} Q_{ss}^{-} = -0.10 \pm 0.10 \text{ eV} . \tag{4.3}$$

Changes in the position of the C-V accumulation humps correspond to variations in the I-V characteristics. For example, for the n type majority carrier devices of Figure 4.3(b), the value of the depletion plateau current (current at moderate positive bias) tends to increase with respect to the current at negative semiconductor bias as ϕ_{ms} becomes more negative. Another feature of the C-V curves is the large hump for the thinnest oxide device, P0421, centred at about +0.7 V. Prior to this work, this hump would have been interpreted in terms of surface states which had a peak in their distribution near mid-gap. While this hump has some contribution from surface states, on the basis of the theory of Chapter 3 it would be interpreted largely in terms of the delayed inversion layer response. A more striking experimental demonstration of this effect is described in Section 4.6.

As opposed to devices with relatively low work function contacts such as Al, it has been predicted that devices with high work function contacts should exhibit majority carrier properties on p type substrates and minority carrier properties on n type substrates. Gold is one of the relatively few metals with a work function high enough to demonstrate this effect. In a presentation similar to that of Figure 4.3, the measured characteristics of a number of Au devices are shown in Figure 4.4 for oxide thicknesses in the range 21 to 34Å.

The I-V characteristics for Au devices using a p type substrate are shown in Figure 4.4(a). For the range of oxide thicknesses shown, all the devices exhibit non-equilibrium behaviour. They display features similar to the n type majority carrier devices of Figure 4.3(b) but are mirror reflected in the current axis. With positive semiconductor bias, the curve shapes are similar to each other merely being displaced to higher current levels for the thinner oxides. Under reverse bias, all show saturating characteristics. All are too thin for the inversion hump to appear. However, by illuminating with light, minority carrier generation is enhanced and the inversion layer can be formed producing the characteristic hump shown by the chain line for device P0412. The variation of the effective value of the metal work function indicated by the C-V curves of Figure 4.4(c) is reflected in the I-V curves. The larger the effective work function as determined by the location of the accumulation hump in the C-V curves, the larger the depletion plateau current compared to the forward bias current. This effect is well demonstrated by comparing the I-V characteristics for the extreme cases of P0412 and P0420. Again the experimental curve shapes are well described

Figure 4.4: Experimental I-V and C-V curves for Au-SiO_2-<100>Si devices for oxide thickness in the range 21-35Å. The dashed curves are theoretical curves:

(a): Measured I-V curves for 1 Ωcm p type devices. The chain curve represents the measured characteristics of device P0412 when illuminated with light. The dashed theory curve is calculated with d equal to 20Å, ϕ_{mi} equal to 4.2 eV, and the other parameters exactly the same as in Figure 4.3(a).

(b): Measured I-V curves for 1-3 Ωcm n type devices. The theoretical curve (dashed) is calculated with d equal to 24Å, ϕ_{mi} equal to 4.35 eV, and the other parameters exactly the same as in Figure 4.3(b).

(c): Measured C-V curves for the p and n type devices of (a) and (b). The measurement frequency is either 5 kHz or 50 kHz as indicated. The dashed theory curves are calculated at 50 kHz using the same parameters given in (a) and (b) for the p and n type device respectively.

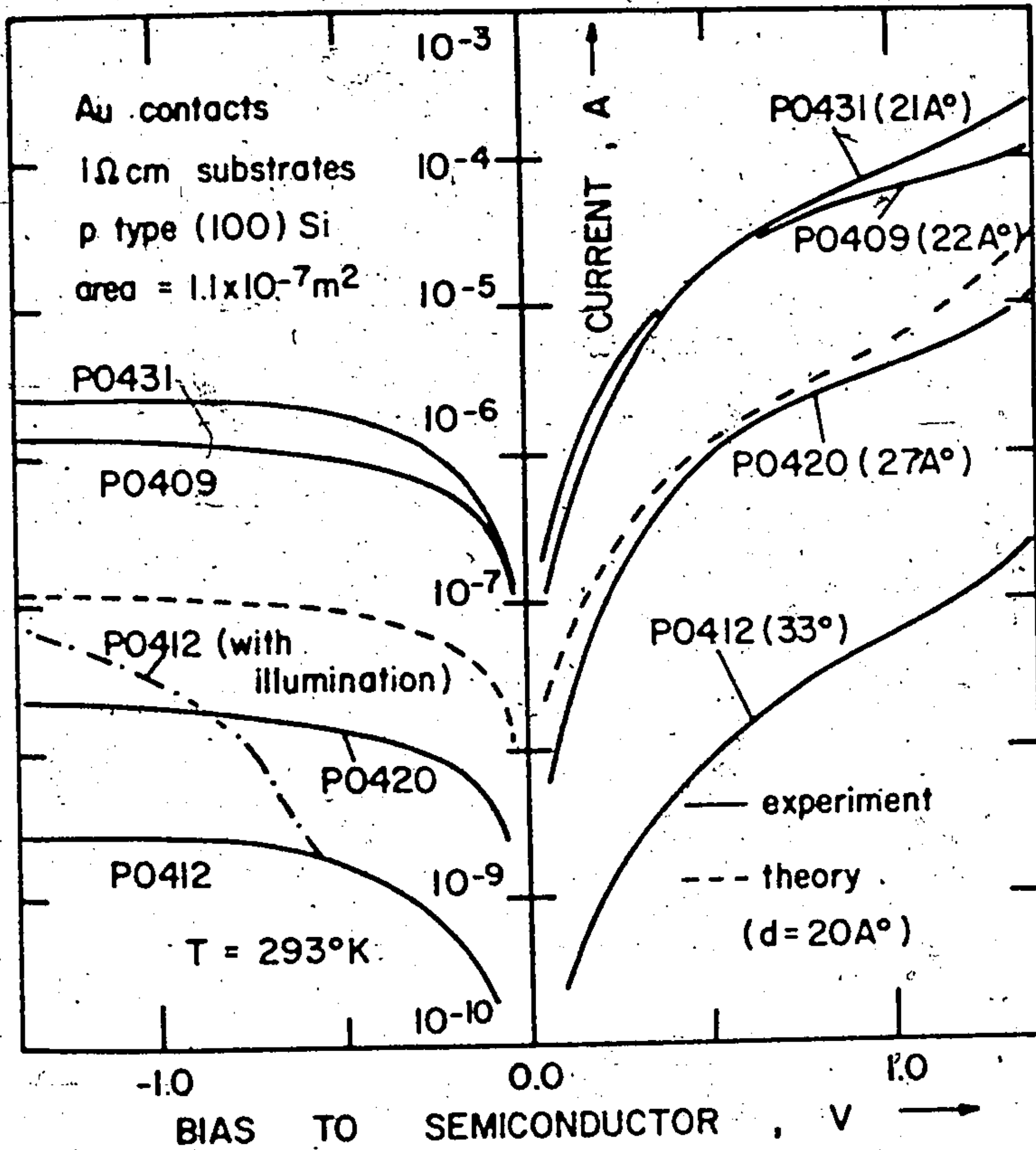


FIGURE 4.4(a)

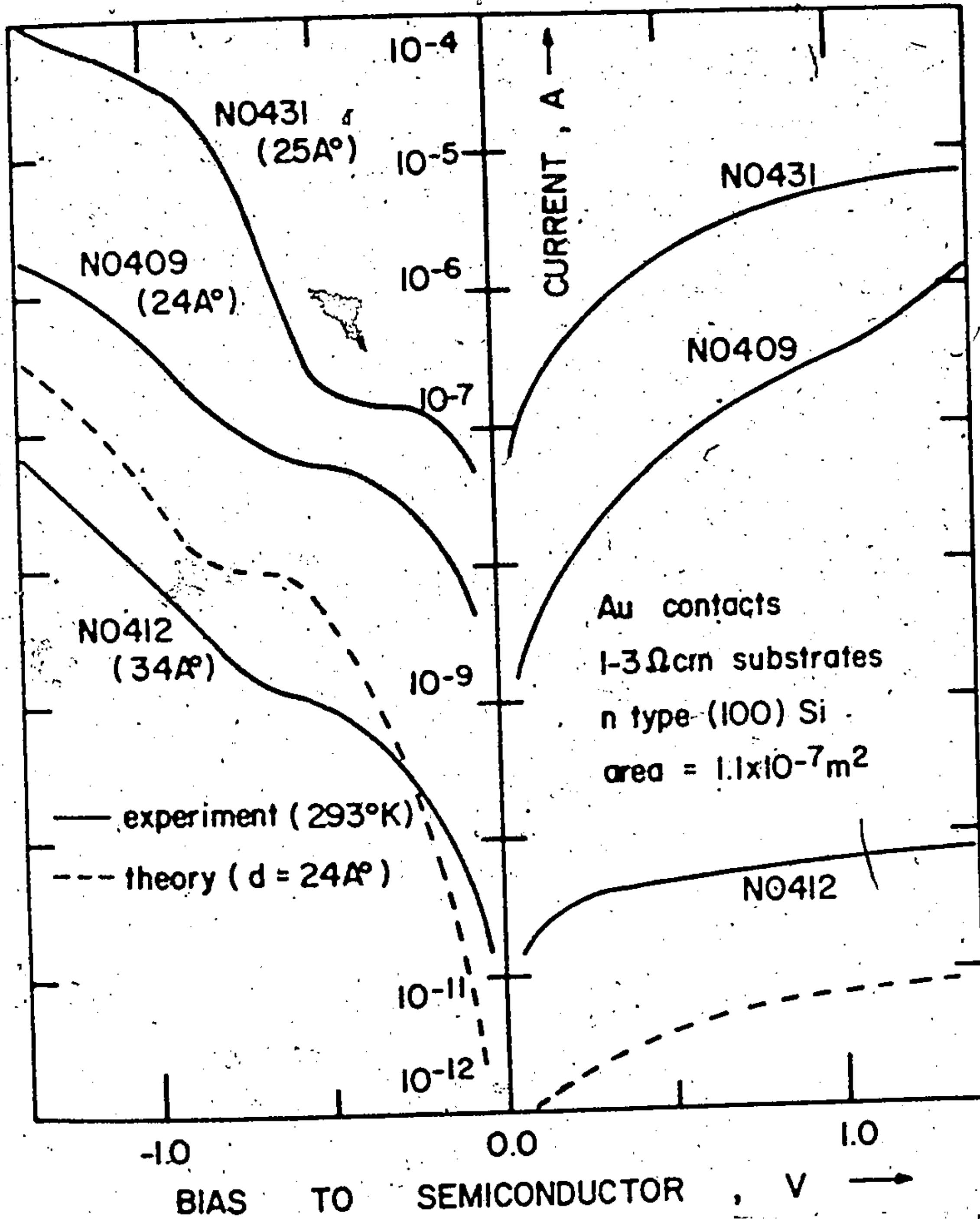


FIGURE 4.4(b)

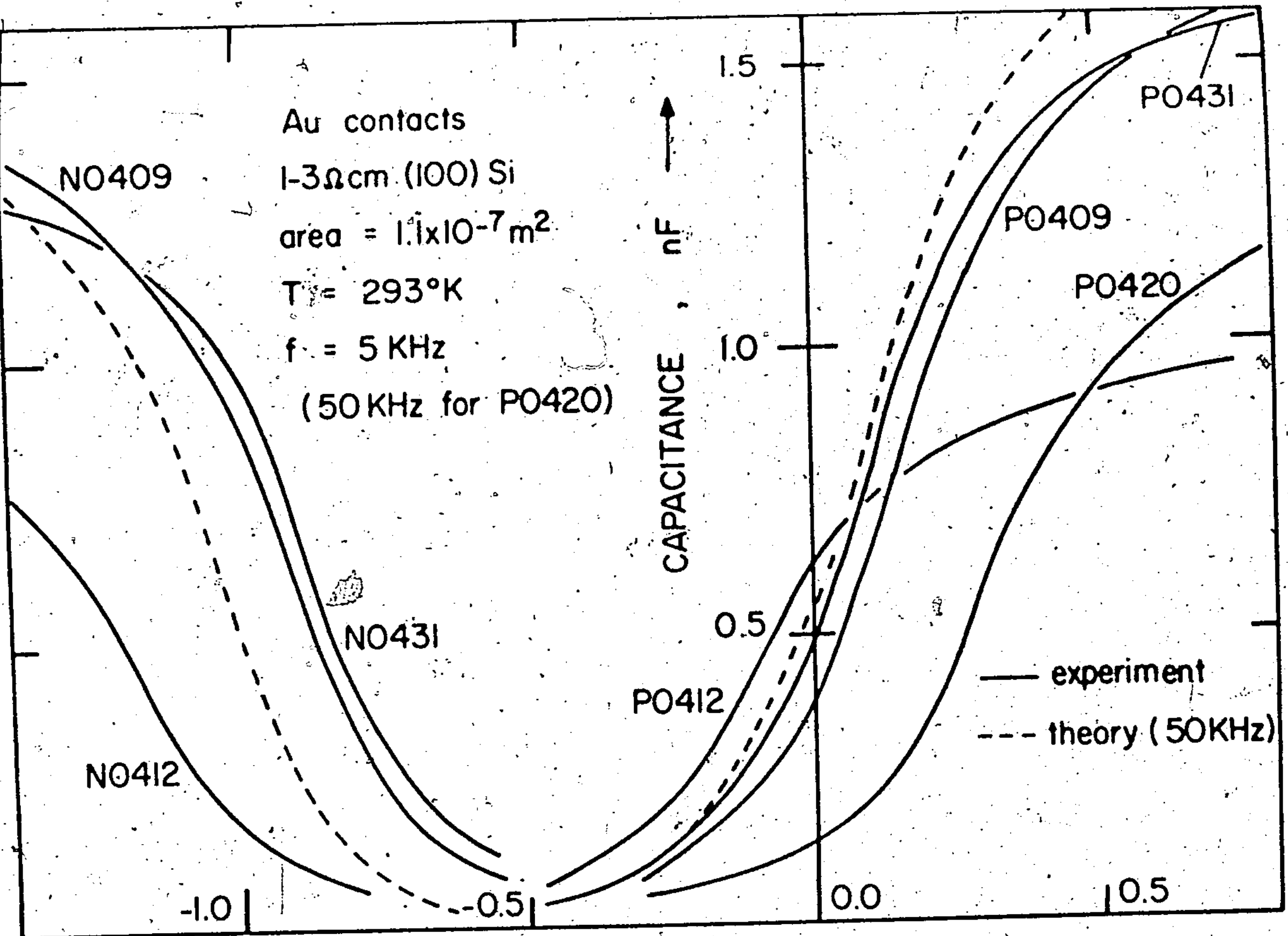


FIGURE 4.4(c)

by the dashed theoretical curve.

The n type I-V curves for these Au devices are shown in Figure 4.4(b). Their characteristics under negative semiconductor bias are similar to those of p type Al devices (Figure 4.3(a)) under the opposite bias direction. While N0413 is semiconductor limited under positive semiconductor bias, N0409 is tunnel limited over the entire bias range and N0431 is semiconductor limited only at the larger positive bias points. The relevant surface configurations are "2" and "3" of Figure 2.8. Yu and Snow [62] have studied the effects of these surface configurations for the related Schottky diode structure using a guard ring and they demonstrated enhanced minority current flow for positive or small negative semiconductor bias. The mechanisms responsible for these effects were also described and are not investigated further in this thesis. As observed for the p type devices, the spread in the effective work function evident from the C-V curves is reflected in the I-V characteristics. Under negative semiconductor bias, device N0431 with the lowest value of effective work function has additional structure above -0.5 V believed to be due to majority carrier tunnel currents as discussed with regard to the 3.8 and 4.0 eV curves of Figure 3.4(b). Under moderate to large negative semiconductor bias, the theoretical dashed curve adequately describes the depletion plateau and accumulation hump regions of the experimental I-V characteristics for the devices with the higher effective work function values (N0409 and N0412). Since lateral effects are not included into the theoretical simulation, the dashed theory curve displays typical minority carrier non-equilibrium effects, underestimating the experimentally observed curves at positive and small negative semiconductor

bias.

The C-V curves again display the spread observed for the Al devices. For the p type Au devices

$$\phi_{ms} - \frac{qd}{\epsilon_i} Q_{ss}^{+} = 0.95 \pm 0.15 \text{ eV} \tag{4.4}$$

and for the n type devices

$$\phi_{ms} - \frac{qd}{\epsilon_i} Q_{ss}^{-} = 1.05 \pm 0.20 \text{ eV} . \tag{4.5}$$

A large component of the variation in these quantities is attributed to work function differences in the metal film. For example, the high effective work functions for the devices N0412 and P0412 could be correlated with the fact that both had their contacts evaporated at the same time and these contacts were significantly thinner than those of the other devices of Figure 4.4(c).

Neglecting the reflection of the characteristics in the current axis, there is little doubt that the Au devices displaying properties most related to those of the Al p type devices of Figure 4.3(a) are the Au n type devices of Figure 4.4(b). Those displaying properties most related to the Al n type devices of Figure 4.3(b) are the Au p type devices of Figure 4.4(a). This provides experimental support for the theoretical prediction that devices can be changed from minority to majority carrier by an appropriate selection of a high or low work function metal for the top contact. Additional support is provided by the experimental properties of devices fabricated with other metals as top contacts.

Figure 4.5 compares the experimentally observed I-V characteristics of devices with Pt, Au, Al, and Ti contacts. The curves for p type devices are shown in Figure 4.5(a) while those for n type devices are in (b). The C-V characteristics for the devices shown can be found in Figures 4.1, 4.3(c), or 4.4(c), except those of device N0523 where current flows were too large to allow the capacitance to be measured over any appreciable bias region. The values of the metal to insulator work functions shown in Figure 4.5 are obtained from the location of the accumulation humps in the C-V curves for these particular devices. Pt has a high work function, this work giving a value on the average of about 0.3 eV higher than that of Au. As can be seen by comparing the characteristics of the non-equilibrium devices P0430 and P0420 in Figure 4.5(a), Pt devices on a p type substrate display majority carrier characteristics similar to those of gold but with enhanced depletion plateau current (current at moderate negative bias for these p type devices). This is in accord with theoretical expectations. Pt devices on n type substrates gave characteristics similar to those observed for the gold devices of Figure 4.4(b) with the higher values of effective work function. This is demonstrated in Figure 4.5(b) where devices N0419 (Pt) and N0409 (Au) are compared. Both display equilibrium characteristics despite the fact that the oxide layer in each case is relatively thin. This is attributed to the lateral effects described when discussing Figure 4.4(b). Ti has a lower work function than Al, and Ti devices on p type substrates give minority carrier devices. This is demonstrated in Figure 4.5(a) where the characteristics of the non-equilibrium devices P0411 (Ti) and P0601 (Al)

Figure 4.5: Experimental comparison of the I-V characteristics of M-SiO_2-<100>Si devices for Ti, Al, Au, and Pt devices.

(a): Measured I-V characteristics for 1 Ωcm p type substrates.

(b): Measured I-V curves for n type substrates with resistivity in the range 0.1-3 Ωcm.

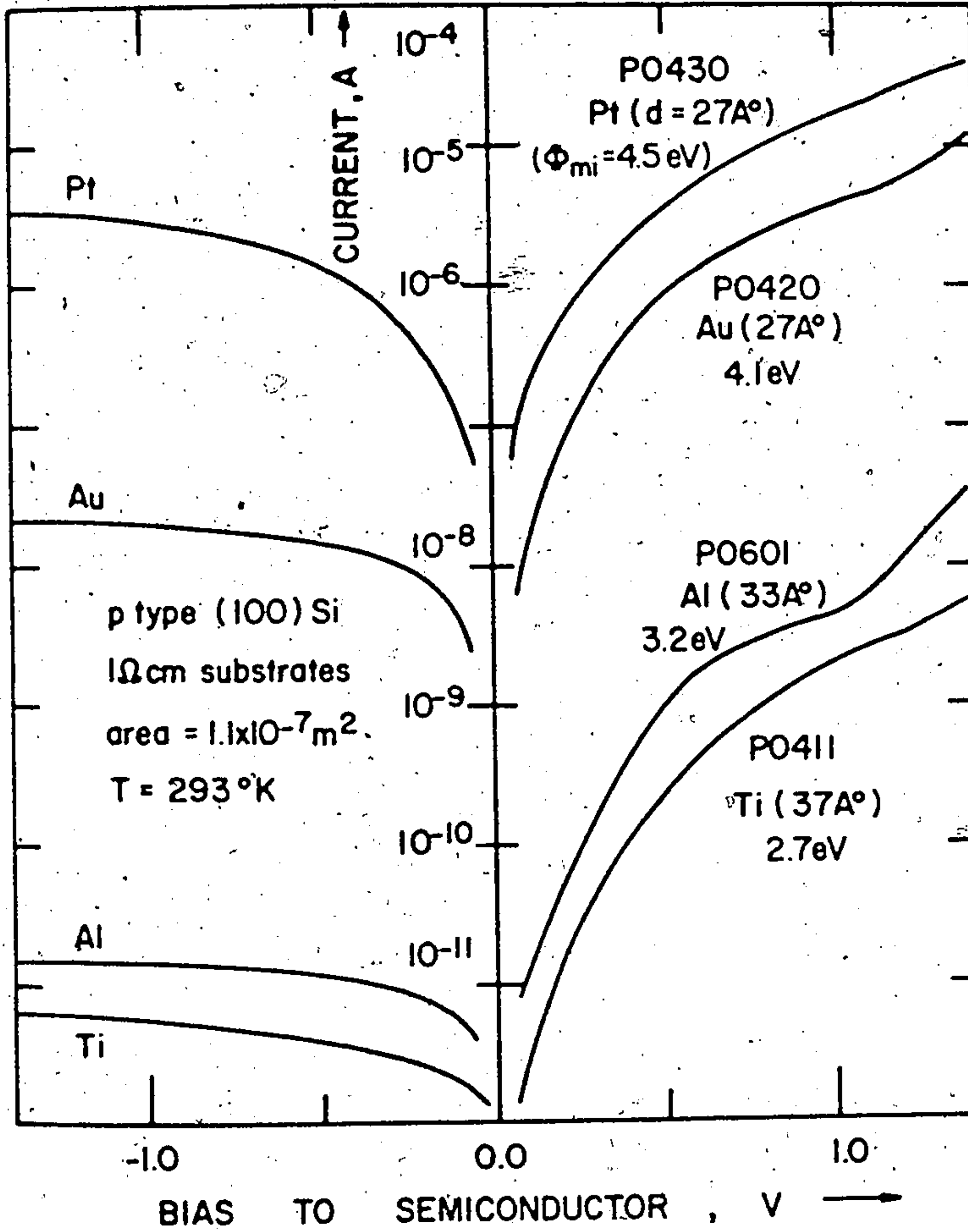


FIGURE 4.5(a)

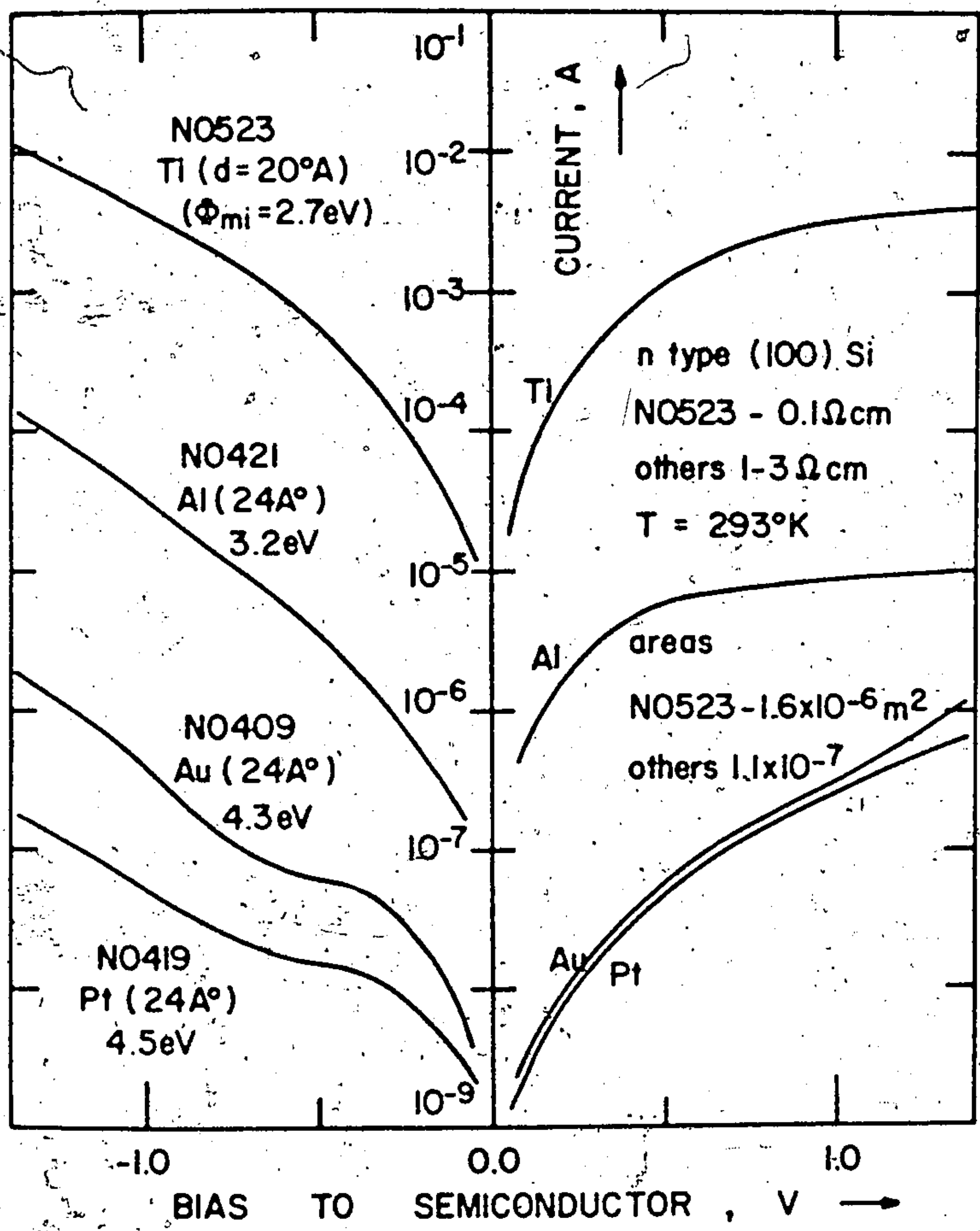


FIGURE 4.5(b)

are compared. Under reverse and small forward bias, both display semiconductor limited characteristics. Under forward bias above 1 V, the Al device (P0601) displays an accumulation hump. The Ti device (P0411) becomes accumulated at higher forward bias and the accumulation hump is just starting to become evident at 1.4 V. Ti devices on n type substrates give majority carrier devices. This is demonstrated in Figure 4.5(b) where devices N0523 (Ti) and N0421 (Al) are compared. The enhanced depletion plateau current (current at moderated forward bias for these n type devices) is evident for the Ti device. In Chapter 6, Mg and Al devices on p type substrates are compared. Mg has a very low work function and, as expected theoretically, Mg p type devices display minority carrier characteristics.

Although the C-V behaviour in the accumulation region has been described for several devices in Figures 4.1, 4.3(c), and 4.4(c), the behaviour of the device capacitance away from this region has not yet been discussed. As predicted theoretically, in the equilibrium case conventional MIS characteristics were displayed while in the non-equilibrium case linear $1/C^2$ versus V relationships were observed in the bias direction opposite that producing accumulation. Examples of this behaviour are given in Figure 4.6 for Pt, Au and Ni devices and additional examples are given in Chapter 6 for Mg, Al, and Au devices. A summary of measurements upon p type non-equilibrium devices is given in Table 4.2 for the complete range of metals used. The results for the intercept voltage, V_c, are in good agreement with the theoretical expectations.

Figure 4.6: Experimental demonstration of the linear $1/C^2$ versus V region in the C-V characteristics of Pt, Au, and Ni non-equilibrium M-SiO_2-p type <100> Si devices.

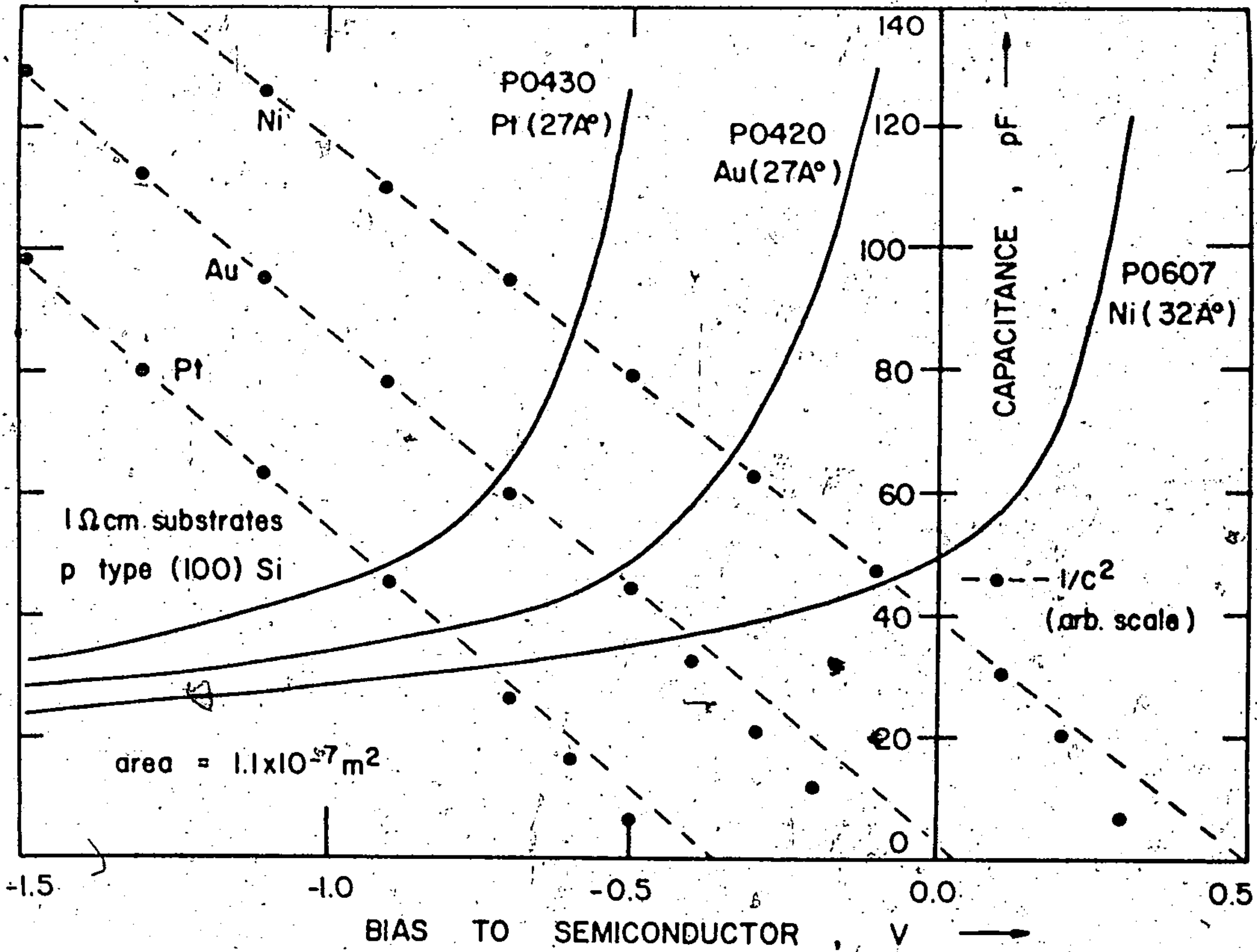


FIGURE 4.6

TABLE 4.2: Summary of Reverse C-V Characteristics for p Type Devices with Mg, Ti, Al, Au, Pt Contacts

Metal	Device No.	Orientation	Resistivity(Ωcm)	N_A(from slope m^{-3})	V_c
Mg	P1309M	<111>	5 - 10	1.3×10^{21}	0.84
Ti	P1043	<111>	5 - 10	2.5×10^{21}	0.73
Al	P1309A	<111>	5 - 10	1.4×10^{21}	0.74
	P1045	<111>	5 - 10	2.0×10^{21}	0.64
	P0224	<100>	2	6.0×10^{21}	0.70
	P0301	<100>	2	8.0×10^{21}	0.70
	P0302	<100>	5	3.0×10^{21}	0.74
Ni	P0607	<100>	1	1.3×10^{22}	0.50
Au	P1024	<111>	5 - 10	1.9×10^{21}	-0.07
	P0408	<100>	1	1.9×10^{22}	0.02
	P0420	<100>	.1	1.2×10^{22}	0.01
Pt	P0430	<100>	1	1.2×10^{22}	-0.36

4.5 The Effect of Temperature

As shown in the theory section, variation of the diode temperature is an invaluable aid in the investigation of MIS tunnel devices. In equilibrium devices, it gives an indication of the dominant tunneling path. It can also be used to demonstrate the equilibrium to non-equilibrium transition complementing the demonstration of Figure 4.2 where the insulator thickness is varied.

The I-V characteristics of both minority and majority carrier equilibrium devices are relatively temperature insensitive when the IS interface is accumulated or strongly inverted. When the interface is depleted, these devices display the depletion plateau region in the I-V curves. A simple analytical expression which follows from the analysis of Appendix A describes the temperature dependence of this depletion plateau current. Provided the argument of the Fermi-Dirac integral, F_1, is not a large positive number, the temperature dependence is given by

$$T^2 F_1(-qE_d/kT) \propto \begin{cases} \text{constant if } E_d < 0 \\ T^2 \quad , \quad \text{if } E_d = 0 \\ T^2 \exp(-qE_d/kT) \text{ if } E_d > 0 \end{cases} \tag{4.6}$$

where for low work function devices ($\phi_{mi} < \phi'_{mi}$ of Equation (3.5))

$$E_d = \phi_{ms} - \frac{qdQ^o_{ss}}{\epsilon_i} \text{ , where } (Q^-_{ss} < Q^o_{ss} < Q^+_{ss}) \tag{4.7}$$

and for high work function devices

$$E_d = E_{gs} - \phi_{ms} + \frac{qdQ^o_{ss}}{\epsilon_i} \tag{4.8}$$

As the device temperature is reduced, the devices eventually should make an equilibrium to non-equilibrium transition. Minority carrier non-equilibrium devices should display very strong temperature dependencies in the semiconductor limited bias regions similar to p-n junction devices. In tunnel limited region, the temperature dependence should be the same as in equilibrium devices. Majority carrier non-equilibrium devices have the same temperature dependence as in equilibrium devices except that, since the IS interface never becomes inverted, the depletion plateau type of temperature dependence (Equation (4.6)) extends over the entire inversion bias region.

The temperature effects experimentally observed for minority carrier devices will be described first. In Figure 4.7, the I-V characteristics for an Al p type device (P0602) are shown for temperatures in the range 200-350°K. At room temperatures, the device is in the transition region between the equilibrium and non-equilibrium modes of operation. At 350°K, minority carriers flow readily in the semiconductor in response to tunneling demands and the device is equilibrium over the whole bias range except at the largest negative semiconductor bias. The 350°K curve is similar to those presented earlier for equilibrium n and p devices (Figure 4.2). As the temperature is decreased to 300°K, non-equilibrium effects become more apparent under reverse bias. The device remains essentially in thermal equilibrium under small reverse bias and over the entire forward bias range. The temperature variation is smallest at large forward bias as expected since the IS interface is accumulated at this bias. At 250°K, non-equilibrium effects are evident over the entire reverse bias range and at small forward bias. By 200°K, the

Figure 4.7: Experimental measurement of the effects of temperature upon the I-V characteristics of the minority carrier device P0602 (Al-SiO_2(34Å)-p type<100>Si).

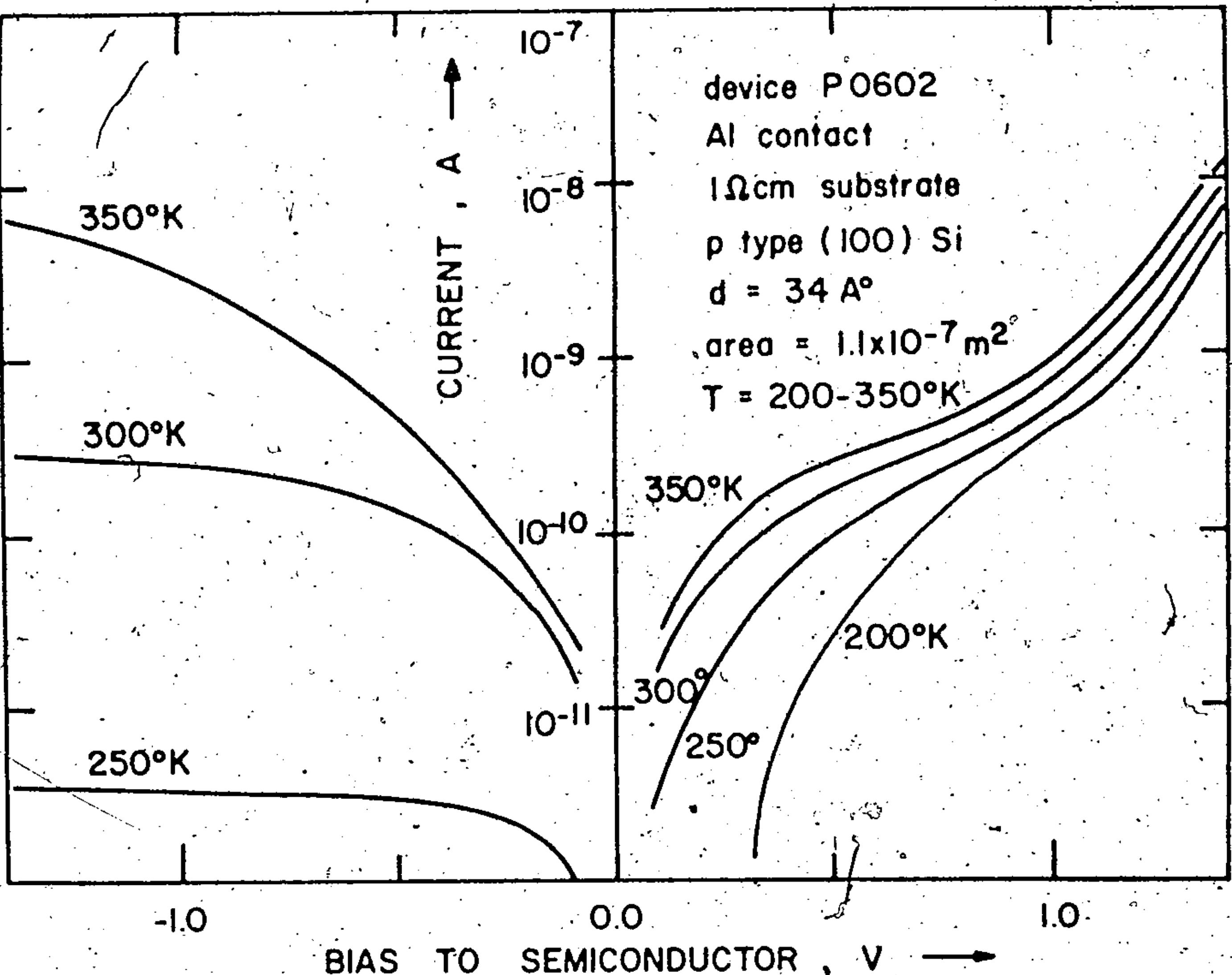


FIGURE 4.7

recombination-generation rate in the semiconductor has dropped so low that no portion of the curve appears below 0.3 V forward bias. Above this bias, the I-V curve blends into the expected equilibrium characteristic. Above 0.8 V, the current varies only slightly as the temperature changes from 350°K to 200°K. Under the largest value of reverse bias shown (-1.4 V), the current varies very rapidly with temperature in the manner $\exp(-E_a/kT)$. The value of the activation energy, E_a, is determined experimentally to be 0.57 eV. This is in excellent agreement with the value $E_{gs}/2$ expected since the generation current is proportional to n_i, the semiconductor intrinsic concentration (Section 6.2.1). The current in the depletion plateau region (moderate positive bias for the p type device) is relatively temperature insensitive. It varies approximately as T^2 at 0.8 V bias, giving a value of E_d (Equation (4.7)) of roughly zero.

Measurements over the same temperature range upon a majority carrier Au p type Si device (P0420) are shown in Figure 4.8(a). The device is a non-equilibrium diode at room temperature. Immediately striking is the very small variation with temperature over the entire forward bias range. The temperature variation under reverse (inversion) bias is also much smaller than that observed for the previous device (P0602). Using Equation (4.6) to analyze the temperature variation at -0.5 V, an activation energy, E_d, of 0.11 eV is obtained. This corresponds to

$$\phi_{ms} - \frac{qdQ_{ss}^{o}}{c_i} = 1.0 \text{ eV} \tag{4.9}$$

in good agreement with the value of 0.9 eV obtained for this particular device from Table 4.2 using Equation (3.6). This value is also consistent

Figure 4.8: Experimental measurement of the effects of temperature upon the I-V characteristics of majority carrier devices.

(a): device P0420 (Au-SiO_2(27Å)-p type<100>Si).

(b): device N0602 (Al-SiO_2(30Å)-n type<100>Si).

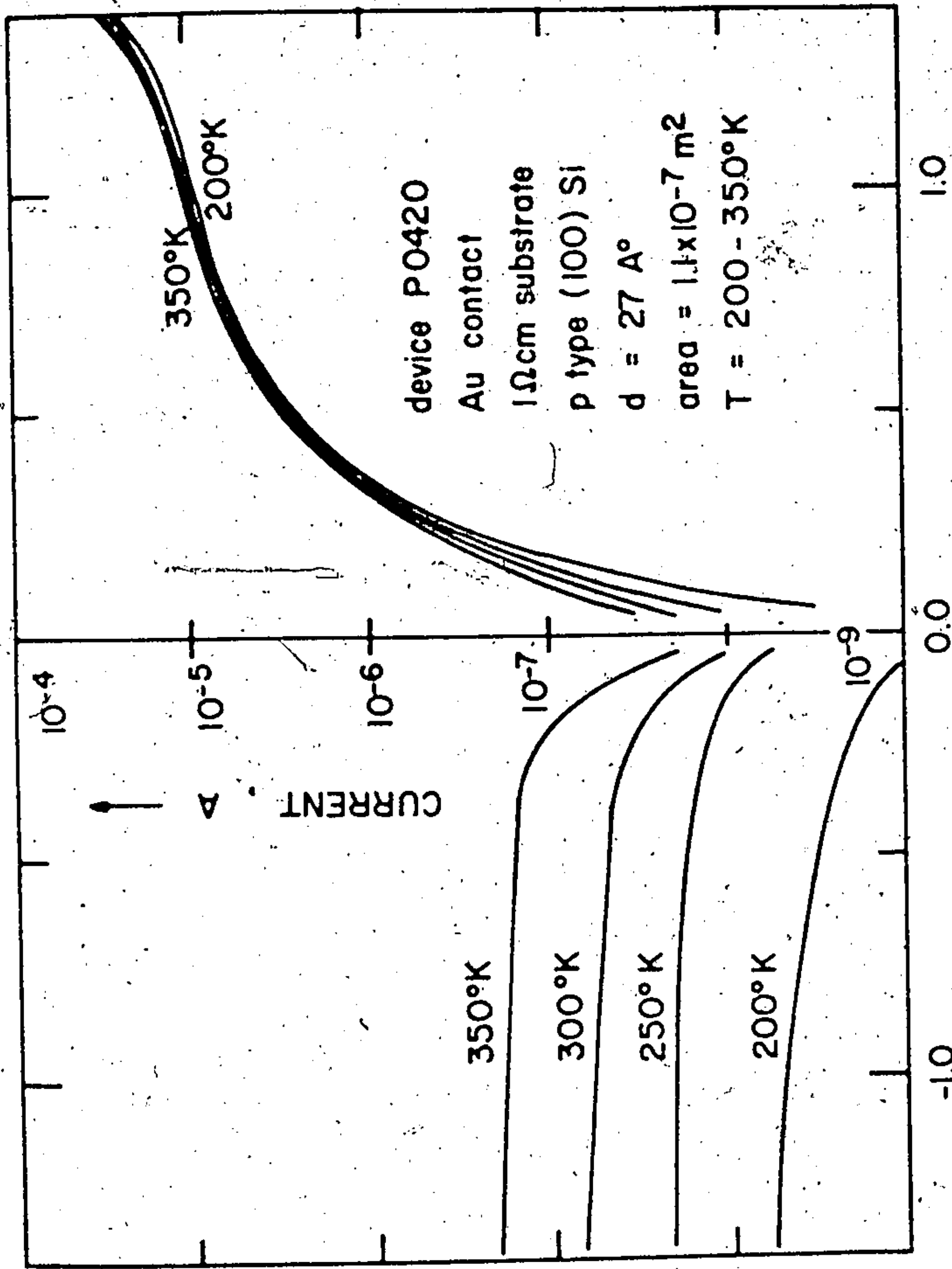


FIGURE 4.8(a)

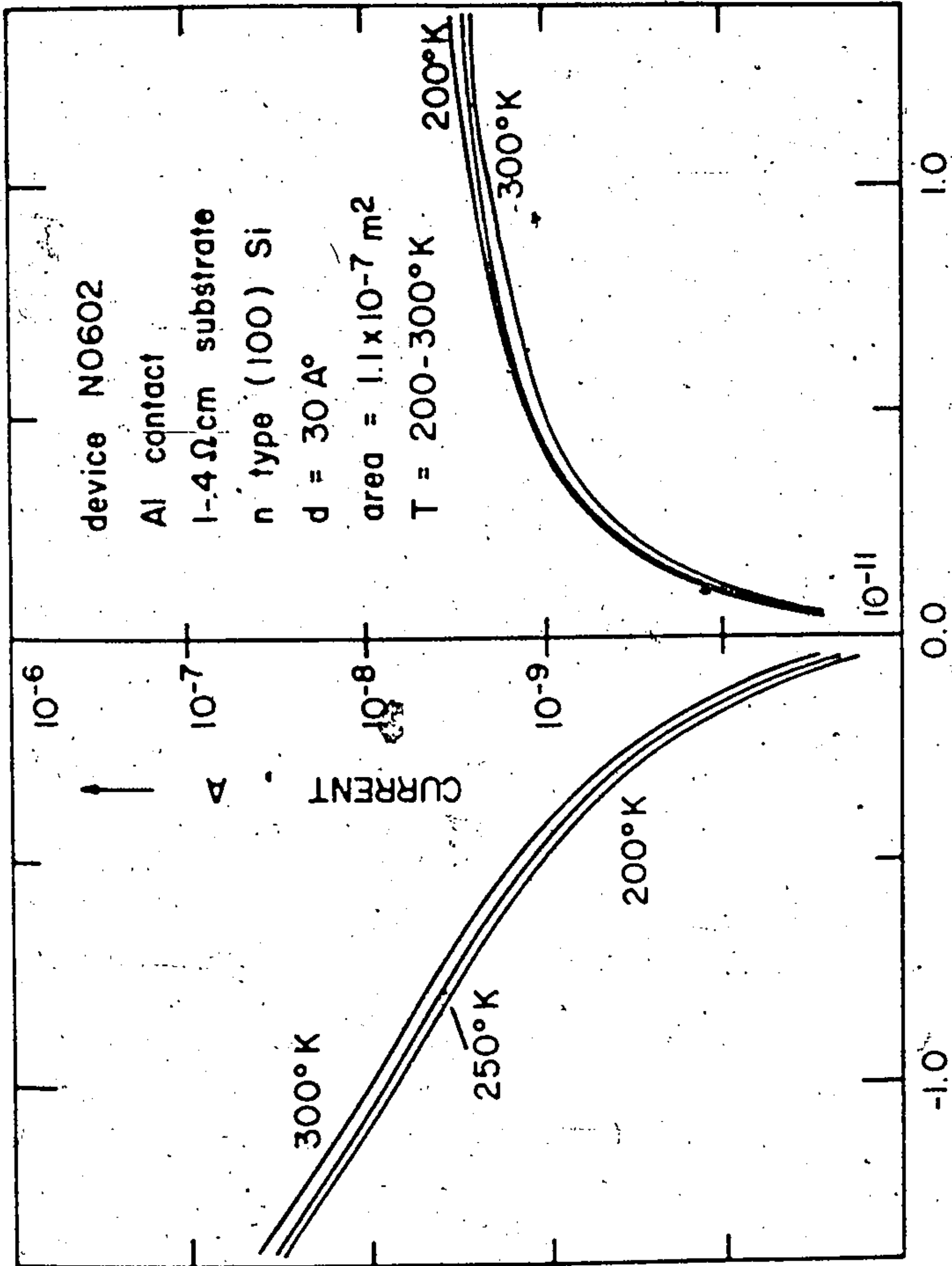


FIGURE 4.8(b)

with the relationship obtained from the location of the accumulation hump for this device (Figure 4.4(c)).

$$\phi_{ms} - \frac{qdQ_{ss}^{+}}{c_i} = 0.8 \text{ eV} . \tag{4.10}$$

The enhanced slope of the reverse bias I-V characteristics at low temperature is also in accord with theory (Equation (A9)). The inversion hump expected at large temperatures under negative bias has not appeared by 350°K. Similar curves are shown in Figure 4.8(b) for a different majority carrier system (Al on n type Si; device N0602). Again there is only a small temperature variation under accumulation bias (negative semiconductor bias in this case). Under the opposite bias, there is a slight increase in current as the temperature is decreased from 300°K to 250°K with an even smaller increase going from 250°K to 200°K. For analysis, these inversion bias characteristics will be regarded as temperature insensitive. From Equation (4.6), this gives a relationship of the form of Equation (4.9) where the right hand side (RHS) is more than a few kT negative, i.e. < -0.05 eV. This is consistent with the relationship obtained from the C-V curve shown in Figure 4.3(c)

$$\phi_{ms} - \frac{qd}{c_i} Q_{ss}^{-} = -0.05 \text{ eV}. \tag{4.11}$$

Nickel has a work function intermediate between Au and Al. Temperature measurements upon devices with Ni contacts are of interest because this is predicted to be an ideal situation for detecting surface state tunneling currents since the band currents are expected to be very small and also very temperature dependent in the bias region near zero volts.

For devices operating in the equilibrium mode, the device characteristics should be very temperature sensitive in this bias region if the band currents dominate. However, it was shown in Chapter 3 that surface state tunnel currents in this bias region are virtually temperature insensitive. Therefore, if surface state currents dominate, the characteristics of experimental devices operating in the equilibrium mode should not vary with temperature in the zero bias region.

The I-V characteristics for a p type device (P0607) over the temperature range 200-400°K are shown in Figure 4.9(a), those for an n type device (N0605) in (b), and the C-V characteristics in (c). Both devices are non-equilibrium at room temperature. The C-V curves confirm that the IS interface is depleted at zero bias. From the location of the accumulation humps, a relationship for the p type device (P0607) similar to Equation (4.10) is obtained with the RHS equal to 0.25 eV as is one for the n type device (N0605) similar to Equation (4.11) with the RHS equal to 0.4 eV. For the p type device, the I-V characteristics at 400°K indicate equilibrium mode operation over all the positive semiconductor bias range and most of the negative bias range. The 350°K curve indicates this mode over all the positive bias range and the initial portion of the negative range. The important feature is the small temperature dependence of the diode current in the bias region near zero bias where the device operates in the equilibrium mode. This is incompatible with the dominant component of diode current being band current. The features of the diode at lower temperatures are most related to those of the minority carrier device of Figure 4.7. On this basis, the device is interpreted as being a surface state dominated diode with minority carrier properties. Similarly,

Figure 4.9: Experimental measurement of the effects of temperature upon the I-V characteristics of Ni-SiO_2-<100>Si devices.

(a): device P0607 with a 1 Ωcm p type substrate (d = 32Å).

(b): device N0605 with a 1-4 Ωcm n type substrate (d = 31Å).

(c): C-V curves for devices P0607 and N0605 measured at 50 kHz at 293°K.

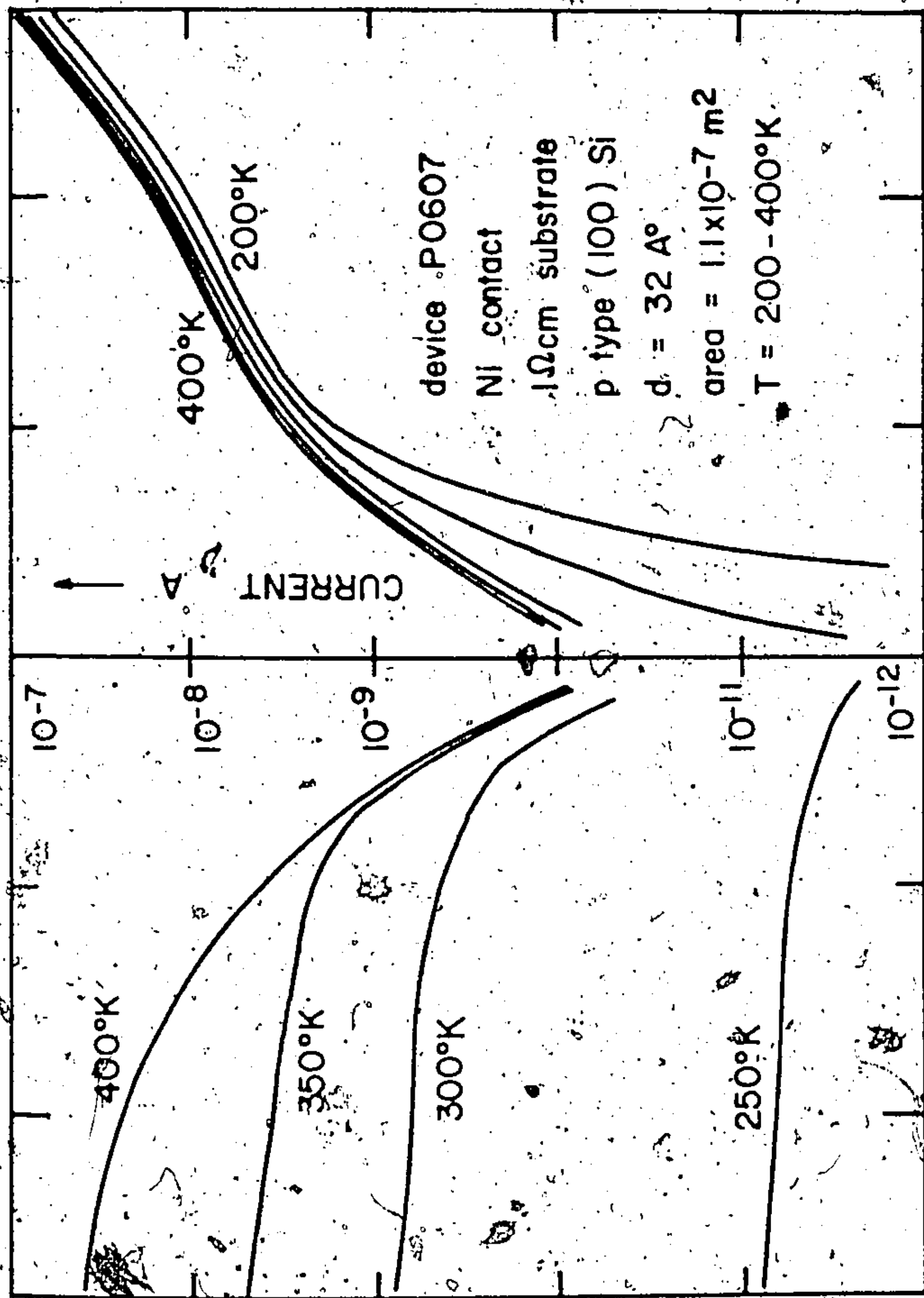


FIGURE 4.9(a)

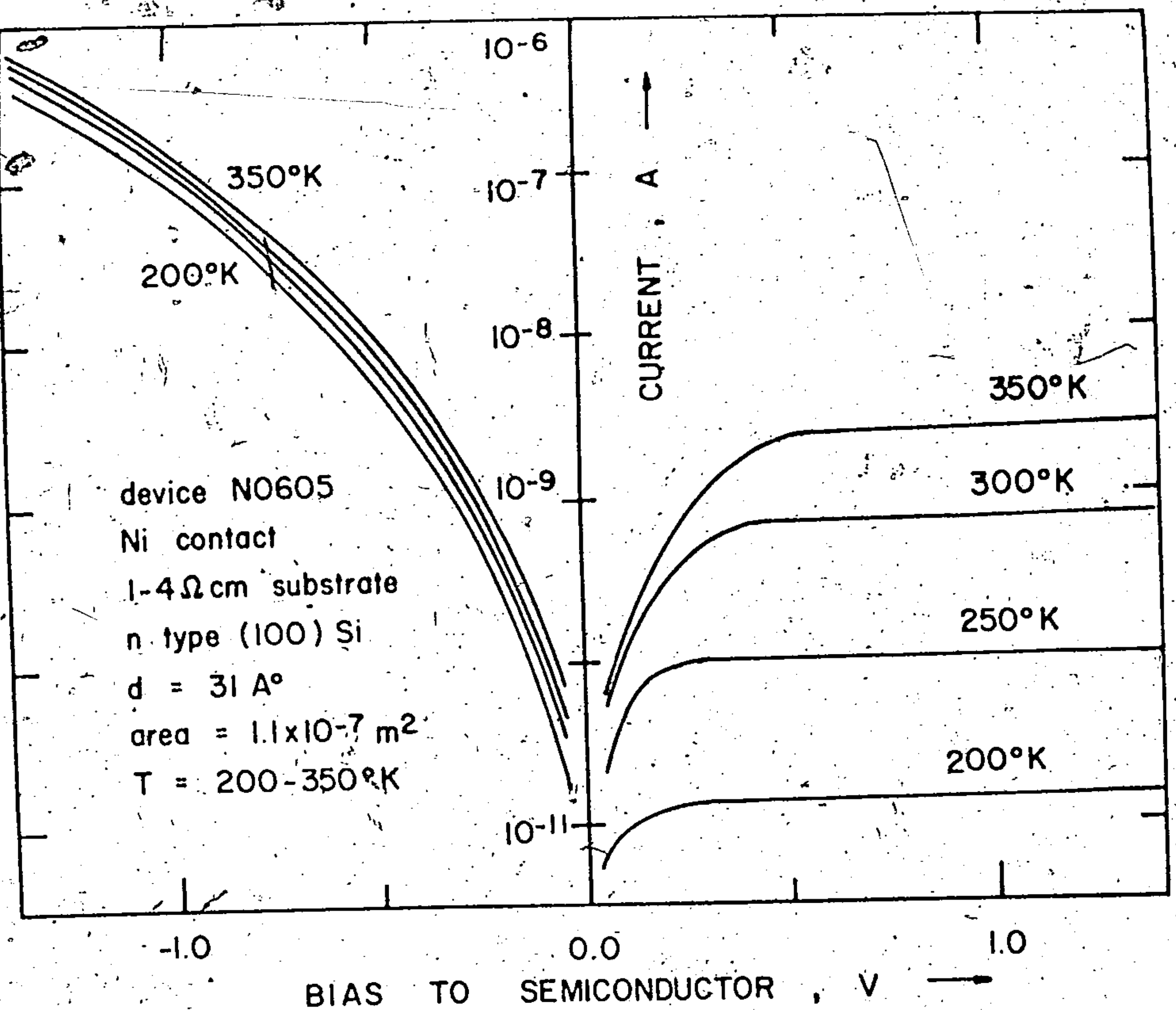


FIGURE 4.9(b)

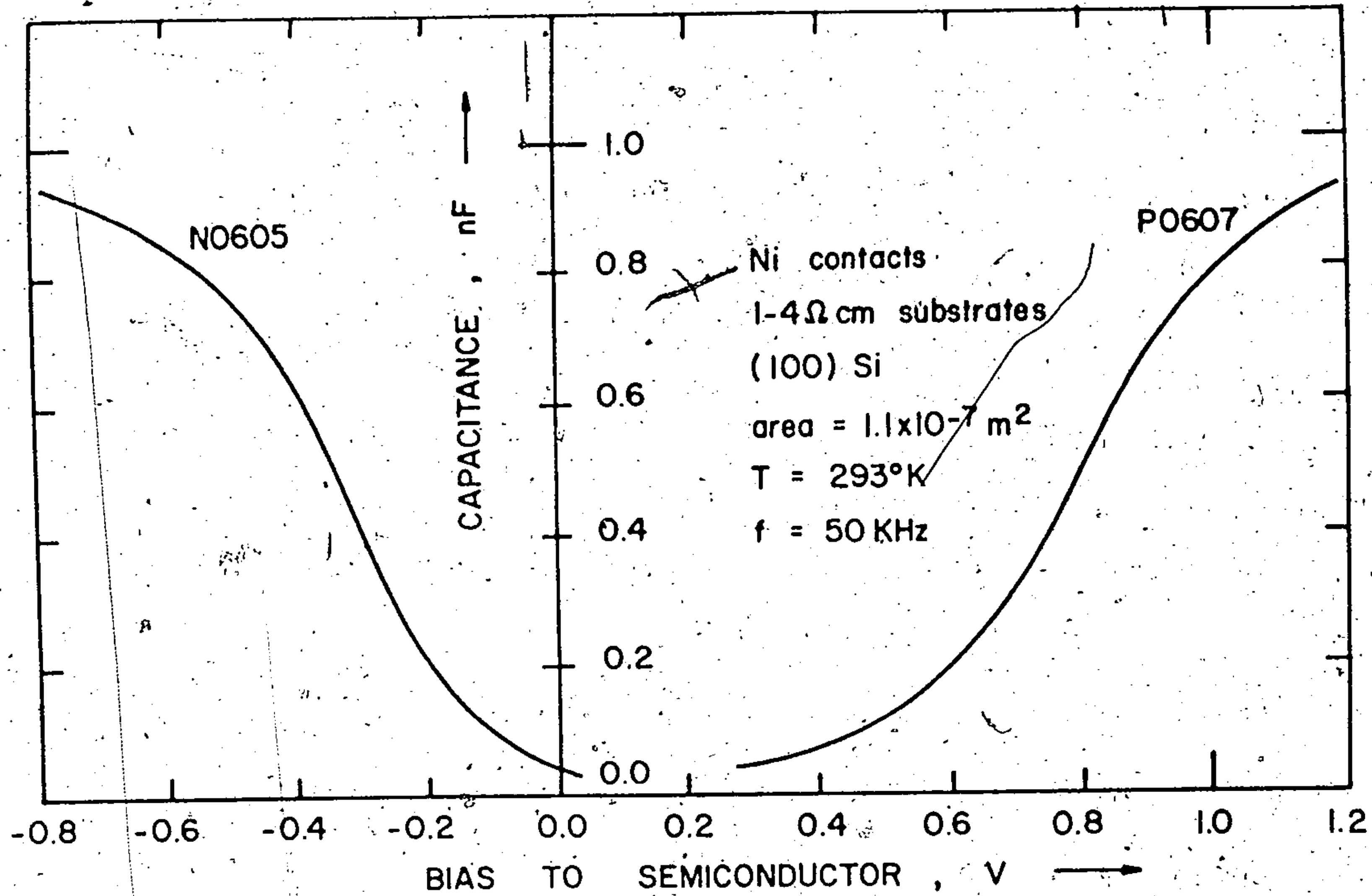


FIGURE 4.9(c)

for the n type device (N0605) at high temperatures, the current in the zero bias range is relatively temperature insensitive. Again, this is incompatible with the band currents dominating. The n type device is interpreted as being surface state dominated with properties most like majority carrier devices.

4.6 Comparison with Previous Experimental Work

A summary of published experiments made upon non-degenerate $M-SiO_2-Si$ tunnel diodes is given in Table 4.3. As shown, only a few groups have been concerned with the I-V characteristics of the device which are of principal interest in this thesis. As indicated by the work described in the previous section, temperature measurements of the device current in conjunction with C-V measurements are the most useful in determining the device physics. Only in the work of Clarke and Shewchun [15] has a temperature measurement of the diode current been mentioned. However, Kar and Dahlke [17] have described detailed room temperature I-V, C-V, and G-V measurements upon a $Mg-SiO_2$- p type <111> Si device. This section will be limited to a comparison of the results of these two groups with the theory of Chapter 3 since they are the only ones presenting sufficient data to allow a reasonably unambiguous interpretation.

The experimental work of [15] is restricted to the $Al-SiO_2$-n type Si system and results obtained are very similar to those described here for this system. Temperature measurements were made upon a device labelled No. 1048. Analysis of these results in the manner described for majority carrier devices in the previous section gives a relationship for ϕ_{ms}

TABLE 4.3: Summary of Previously Published Measurements Upon the Non-Degenerate Metal-Silicon Dioxide-Silicon MIS Tunnel Diode. († RT = Room Temperature.)

Metal	Insulator Thickness (Å)	Semiconductor Type	Temperature (°K)	Characteristics			Reference
				I-V	C-V	G-V	
Al	~ 55	~~p, n~~	77, 300		x	x	[2]
Al	45 - 70	p<100>, n<100>	77, 300		x	x	[14]
Cr/Au	~ 20	p<100>	300	x			[13]
Al	26 - 33	n<100>, n<111>	100, 300	x	x		[15]
Mg	28	p<111>	RT†	x	x	x	[17]
Au	24	n<100>	RT	x		x	
Mg, Al	≃ 10	p<111>	RT		x		
Cr, Cu	21 - 37	p<111>	RT		x	x	
Au	8 - 45	n<111>	RT	x	x		[18]
Au	10 - 100	n<111>	RT	x		x	[19], [20]
Cr/Au	35 - 60	p<111>	130, 297		x	x	[84]-[86]

similar to Equation (4.9) but with the RHS equal to -0.01 eV. This is consistent with the data given for describing the C-V characteristics of the device which gives a relationship similar to Equation (4.11) but with the RHS equal to 0.02 eV.

The measured characteristics of the Mg device described by Kar and Dahlke [17] are shown in Figure 4.10(a) and (b). One additional piece of data was that under negative semiconductor bias, the device had a linear $1/C^2$ versus V relationship with an intercept, V_C, of about 0.8 V. The C-V and G-V characteristics under forward bias are shown for a range of frequencies in Figure 4.10(a). Immediately striking is the extremely large hump in the C-V curves extending over a voltage range of 1.5 volts at low frequencies. In the absence of a suitable theory, this was quite naturally interpreted in terms of a large surface state density with a peak near mid-gap. However, by comparison with theoretical curves of Figure 3.8, it can be seen that the experimental C-V and G-V curves have the same shpae and display the same general frequency behaviour as those due to a delayed inversion layer response. The following is felt to be a more likely explanation of the capacitance hump than attributing it to surface states: The IS interface is without a doubt strongly inverted at zero bias. The inversion layer is clamped at the IS interface at reverse and small forward bias for the reasons explained in Chapter 6 associated with semiconductor limited current flow. At moderate forward bias, the current flow is no longer semiconductor limited and the inversion layer at the IS interface becomes unpinned, contributing to the device capacitance in the normal manner. Since the supply of minority carriers to the IS interface by tunneling through the oxide is very large by the

Figure 4.10: C-V, G-V, and I-V characteristics of an Mg-SiO_2-p type<111>Si device as measured by Kar and Dahlke [17].

(a): Measured C-V curves (solid line) and G-V curves (dashed) for frequencies ranging from DC to 1 MHz.

(b): Measured I-V curve.

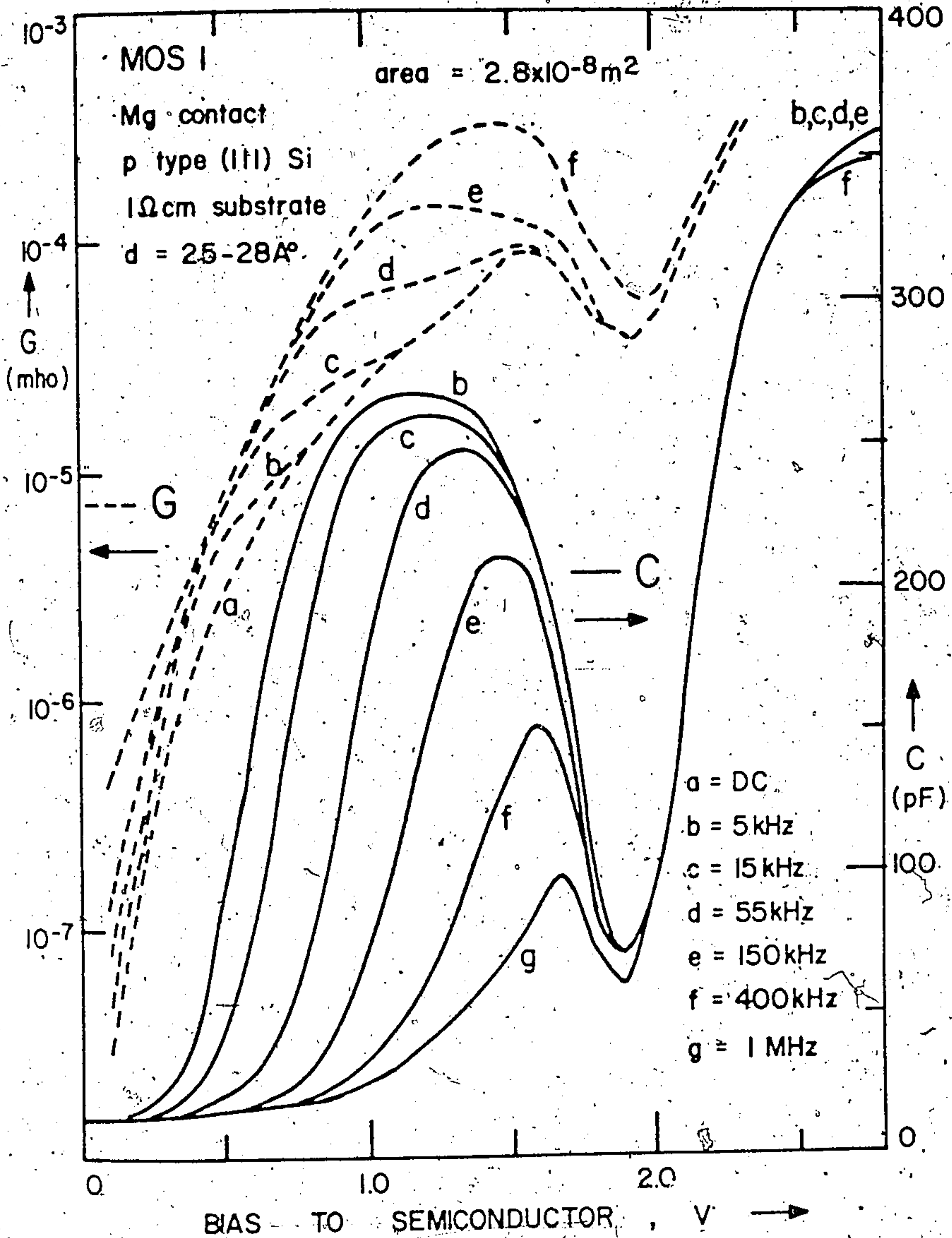


FIGURE 4.10(a)

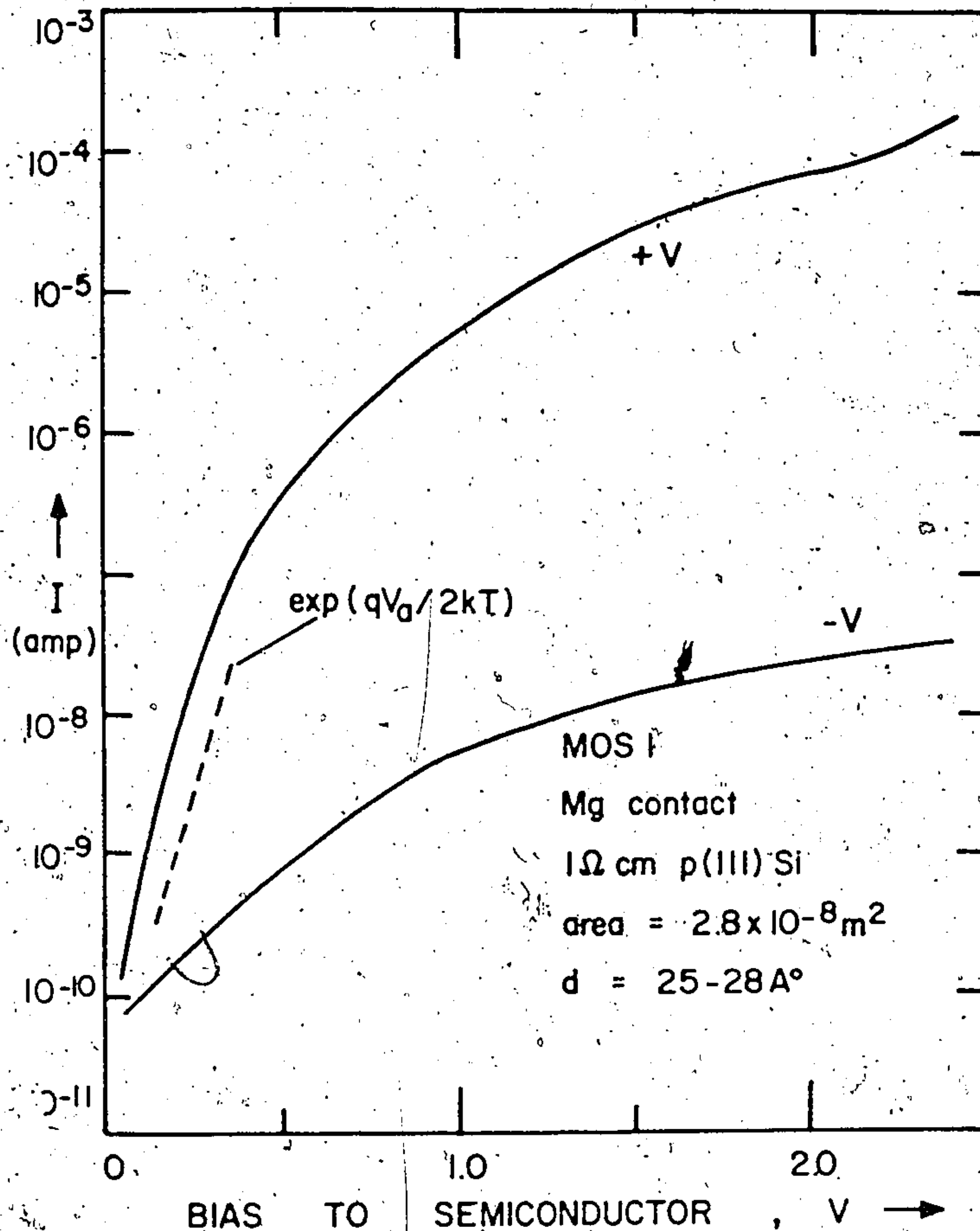


FIGURE 4.10(b)

time this bias region is attained, the frequency response of the layer is much higher than that normally associated with inversion layer response. For the Mg device, a relationship for ϕ_{ms} similar to Equation (4.8) can be obtained from the position of the accumulation hump with the RHS equal to -1.1 eV. The extent of the inversion layer response due to the preceding process is given approximately by Equation (3.8) which in the present case gives the value of 1.3 V in excellent agreement with the experimental value of 1.5 V. The small discrepancy is due largely to the fact that the theoretical estimate is based on the current flow when the hump first appears. In the present case, the flow when the hump disappears is significantly larger. The I-V characteristics (solid curve of Figure 4.10(b)) are consistent with this interpretation. In the forward direction, they display semiconductor limited effects, increasing approximately as $\exp(-qV_a/2kT)$. This type of behaviour is typical of current flow determined by recombination processes in the semiconductor depletion region. In the present instance, the extent of the depletion region is probably enhanced by the lateral effects described in Section 2.6.1. Above 0.3 V, they are envisioned as displaying tunnel limited effects. The hump at about 2.0 V bias corresponds to the IS interface becoming accumulated Under reverse bias, the current is again influenced by the lateral effects described in connection with Figure 4.3(a). Kar and Dahlke [17] also describe measured I-V characteristics for the case of Au on n type Si. They are not unrelated to the characteristics observed for device N0431 in Figure 4.4(b) having features similar to the theory curves of Figure 3.4(b) for ϕ_{mi} in the range 3.8-4.0 eV.

This section will conclude with some comments on surface state

densities in MIS tunnel devices. Although relatively large densities have been reported in [17], these correspond in all but one case to devices that have been annealed. This was found to degrade the diode properties. Of the five unannealed devices measured, only one has been shown to have a peak surface state density > $1.3 \times 10^{16} m^{-2} eV^{-1}$, a total surface state density > $3.9 \times 10^{15} m^{-2}$, and a capture cross section outside the range 10^{-20} to $10^{-19} m^2$. Other work [85,86] has given peak densities up to an order of magnitude larger and capture cross sections of the order of $10^{-18} m^2$. These measurements were made upon <111> orientated silicon. There exists some experimental support [15] for the view that <100> devices have lower surface state densities as in conventional MIS devices [1d]. Since, even with the relatively small number of tunnel devices that have been fabricated, densities with a peak value of $5 \times 10^{15} m^{-2} eV^{-1}$ and a total density of $1.2 \times 10^{15} m^{-2}$ have been reported [17], it is expected that further work will produce values for these quantities which approach even more closely the best values reported for conventional devices.

4.6 Summary

In this chapter, the major features of the theory of Chapter 3 have been investigated experimentally for the M-SiO_2-Si system. The existence of majority and minority carrier diodes has been demonstrated and it has been shown how the choice of metal and an n or p type substrate determines which type a given diode will be. The transition of minority and majority carrier devices from equilibrium to non-equilibrium behaviour has been demonstrated in two ways, by varying the insulator thickness and by varying the device temperature. The temperature variation of the current

flow in the MIS tunnel diode was shown to be consistent with its interpretation as tunnel current. Experimental evidence for the clamping of the inversion layer at the IS interface in minority carrier devices has been described. Temperature measurements upon Ni devices have produced evidence for surface state tunnel currents. The experimental results confirm the general validity of the theory of Chapter 3.

CHAPTER 5

CURRENT MULTIPLICATION IN MAJORITY CARRIER DIODES

5.1 Introduction

The recent experimental work of Clarke and Shewchun [15] established that the MIS tunnel diode possesses properties distinct from the related Schottky diode structure. It was predicted that devices which in this thesis are described as "majority carrier non-equilibrium" might provide the basis for a new transistor structure. This was subsequently verified in their Surface Oxide Transistor (SOT) device [28]. The aim of the present chapter is to provide the first quantitative analysis of the multiplication process which makes such an application of the MIS tunnel diode possible. The first experimental measurements of the gain available from this process and its frequency response are also described.

In Section 5.2, the multiplication process is analyzed theoretically using self-consistent numerical solutions for the complete MIS tunnel diode (Chapter 2). Analytical expressions also are derived to describe the process over certain ranges of the device parameters. In Section 5.3, a simple experimental technique is described for the measurement of this multiplication process using light as the primary excitation. Both DC and transient measurements are described. Gains in the range of 100 to 1,000 were observed. Finally in Section 5.4, the application of the multiplication process to a transistor structure and as a photodiode are described.

5.2 Theory of the Multiplication Process

5.2.1 Qualitative Description

A schematic energy band diagram for a MIS tunnel diode is given in Figure 5.1 for the case of an n type semiconductor and a negative voltage V_a applied to the top metal contact. The diagram is similar to Figure 2.1 except for the n type semiconductor region. J_{gen} is the total minority carrier current generated within the semiconductor under the bias conditions shown. The other parameters are defined in the text describing Figure 2.1. The diode properties of interest are obtained when the dominant tunnel current is the current tunneling between the metal and the majority carrier band in the semiconductor. Such conditions can be insured by an appropriate selection of the metal work function and the dopant species in the semiconductor as described in Chapter 3 and demonstrated experimentally in Chapter 4. For n type devices a low metal work function is required while p type devices require a high metal work function. A qualitative description of the multiplication mechanism present in such diodes is given below for the case when the semiconductor substrate is n type (Figure 5.1).

If the insulator layer in a majority carrier MIS tunnel diode is thicker than a critical value, d_{cr}, the semiconductor remains essentially in thermal equilibrium as the bias to the semiconductor is increased from zero. The experimentally determined value of d_{cr} for the Al-SiO_2-Si system is about 30Å at 300°K (Chapter 4). It follows that an inversion layer eventually forms at the semiconductor-insulator interface. However, as the insulator thickness is reduced below d_{cr}, the processes capable

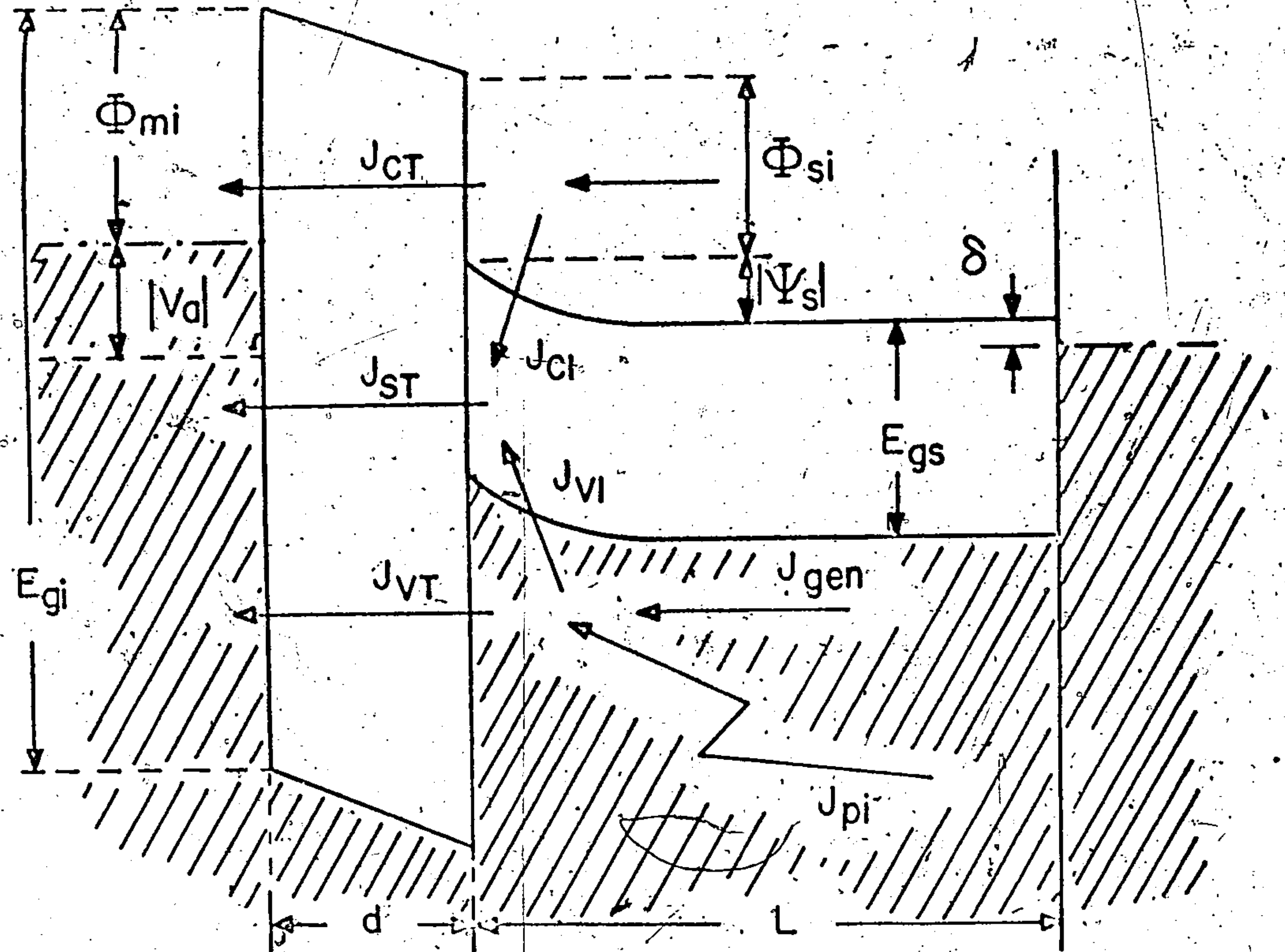


Figure 5.1: Schematic energy band diagram of the n type MIS tunnel diode system. Also shown are the major current components in the surface region of the device.

of removing holes from the surface region rapidly become more effective. The net hole current flow out of the surface region is given by the sum of J_{VT} and J_{VI} in Figure 5.1. The rate that the semiconductor can supply holes to the surface region (given by J_{gen}) is, on the contrary, restricted under the bias region of interest. As the thickness is reduced below d_{cr}, J_{gen} becomes too small to maintain the inversion layer at the surface region, causing the hole concentration in the surface region, p_s, to drop below its thermal equilibrium value. However, if additional holes are supplied to the surface region by external processes such as illumination by light or by injection from a third contact, p_s increases towards its thermal equilibrium value.

When the surface is in strong inversion ($p_s > N_D$), the field at the semiconductor surface is greatly enhanced over its value when the surface is merely in depletion and moreover is very dependent upon the degree of inversion in the surface region. In fact, for a wide range of values of p_s, the magnitude of the electric field, F_s, increases as the square root of p_s (Appendix D).

An increase in the surface field, F_s, increases the field in the insulator and hence the potential drop across it. This effectively causes the energy bands of the semiconductor to move to lower energies with respect to the metal fermi level, greatly enhancing the number of electrons in the metal of energy greater than that of the conduction band edge of the semiconductor. This greatly increases the tunnel current between the metal and the conduction band of the semiconductor (J_{CT}).

Thus, in summary, the multiplication process can be attributed to the following properties possessed by majority carrier MIS tunnel

diodes as the insulator layer becomes thin ($d < d_{cr}$);

(a) the degree of inversion at the semiconductor surface under inversion bias can be controlled by the supply rate of minority carriers to the surface.

(b) the electric field at the semiconductor surface is very dependent upon the degree of inversion at the semiconductor surface.

(c) the majority carrier tunnel current depends strongly upon the potential drop across the insulator and hence upon the surface field.

A change in the supply rate of minority carriers to the MIS diode therefore causes a change in the majority carrier current tunneling between the metal and the semiconductor under inversion bias conditions. With the diode parameters suitably chosen, this mechanism will be shown to result in the multiplication of any minority carrier current reaching the semiconductor-insulator interface by factors of 10^2 to 10^3.

5.2.2 I-V Characteristics of a Reference Diode

Initially consideration will be given to MIS tunnel diodes idealized in that the effects of surface states are neglected. The effect of these states upon the properties of the idealized diodes will be treated in Section 5.2.8.

The diode structure of Figure 5.1 was characterized using the equations described in Chapter 2. The semiconductor was modelled directly by the set of non-linear equations governing carrier transport

within it, and the tunneling processes were modelled using the independent electron approach as given by Harrison [34]. After formulating the problem in this way, numerical techniques were used to solve the system of nonlinear equations to a pre-specified accuracy (usually 1/2%) [102]. To study the multiplication process, a minority carrier current, J_{pi}, was injected into the semiconductor and the resulting current flow through the contact was calculated.

The calculated results depend on the manner in which J_{pi} is injected. If J_{pi} is injected by illuminating the device, the results depend upon the absorption properties of the semiconductors, whereas if it is injected from a third contact, the results depend upon the geometrical arrangement of the electrodes. Although any general case can be analyzed by the numerical technique, the approach chosen was that of the minority carriers being injected as close as possible to the semiconductor-insulator interface to ensure the maximum collection of these carriers by the MIS contact. Less ideal injected carrier distributions can then be discussed in terms of a collection factor, a_c, which will lie between 0 and 1.

A diode consisting of an aluminum contact separated by a 25Å layer of silicon dioxide from a 10 Ωcm n type silicon substrate of <100> orientation was selected as a reference diode. The results of a numerical analysis for the terminal properties of this device are given in Figure 5.2(a). The computed current-voltage characteristics of the diode are shown for various values of the injected minority current density, J_{pi}. At small inversion bias, the characteristics are relatively independent of the injected current density. At inversion voltages greater than 1

Figure 5.2(a): Calculated current-voltage characteristics for a reference Al-SiO_2-n type <100> Si MIS tunnel diode for varying supply rates of minority carriers to the surface region. The supply rate is expressed as an injected current density, J_{pi}. The diode temperature is 300°K, the oxide layer is 25Å thick, the aluminum to silicon dioxide barrier height is 3.2 eV, and the impurity density in the semiconductor is 5 x 10^{20} per cubic metre (10 Ωcm). Values assigned to other system parameters are given in Table 2.4.

(b): Computed current-voltage characteristics for the reference diode of (a) for two distributions of injected current density, J_{pi}. The solid curve represents the characteristics under optimal injection conditions while the dashed curve represents the case where the additional minority carriers are supplied by illuminating the semiconductor with light of 0.8 μm wavelength.

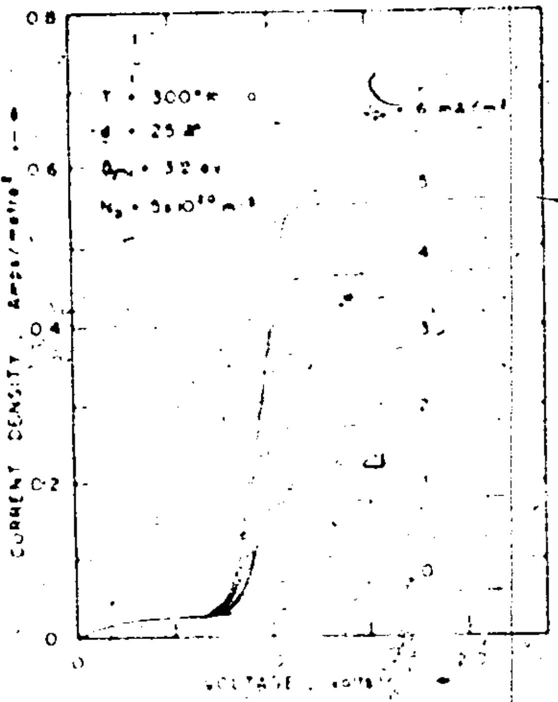


FIGURE 5.2(a)

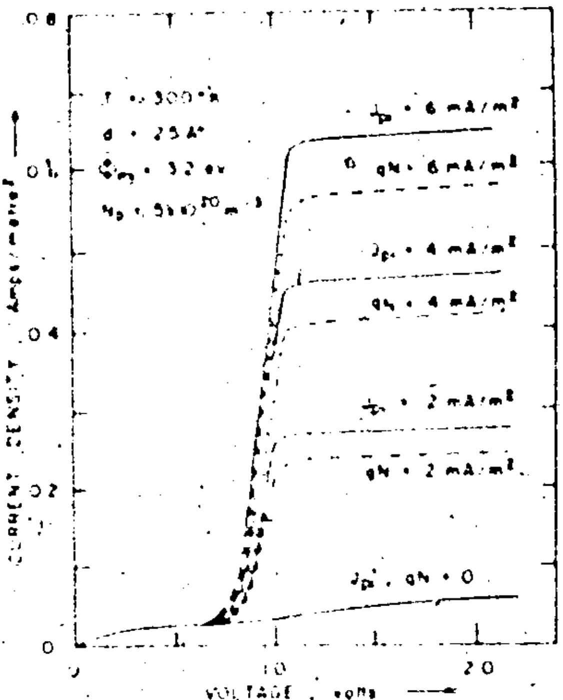


FIGURE 5.2(b)

volt, a small change in J_{pi} causes a large change in the total diode current. In the following section, an analytical expression for the small signal multiplication factor, M, will be derived for diodes biased in this region.

Figure 5.2(b) compares these results where minority carriers are injected as close as possible to the semiconductor-insulator interface with results obtained when light of wavelength 0.8 μm is incident on the device. In this case electron-hole pairs are generated by photons in the manner

$$G = \alpha N e^{-\alpha x} \tag{5.1}$$

where α is the absorption coefficient (10^5 m^{-1} in the present case) and N is the flux of photons entering the semiconductor per square metre per second. The total injected current density is obtained by integrating Equation (5.1) over the device. This gives

$$J_{pi} = qN \tag{5.2}$$

Because not all the generated minority carriers are collected since some diffuse to the back contact, Figure 5.2(b) shows that the effective value of J_{pi} is J'_{pi} where

$$J'_{pi} = \alpha_c J_{pi} \tag{5.3}$$

In the present case, α_c is not strongly bias dependent and can be seen to have a value of approximately 0.9.

5.2.3 An Expression for the Multiplication Factor M

As is shown in Appendix A, the following expression accurately describes the inversion bias dependence of the majority carrier electron current in regions where the argument of the Fermi-Dirac integral, F_1, is negative or a small positive number (see Equations (A5) and (A7)).

$$J_{CT} = A_n T^{1/2} N_C P_n F_1 \left[\frac{q}{kT} \left(|F_s| \frac{\epsilon_s}{\epsilon_i} d + \phi_{si} - \phi_{mi} + \frac{q}{\epsilon_i} Q_i d \right) \right] \qquad (5.4)$$

A_n is a constant, T is the diode temperature, N_C is the conduction band density of states in the semiconductor, ϵ_s and ϵ_i are permittivities in the semiconductor and insulator, Q_i is the effective charge density due to charges in the insulator, and d is the thickness of the insulator layer. P_n is the probability that an electron in the metal of energy approximately equal to the energy of the semiconductor conduction band edge and with the correct value of transverse momentum will make a tunneling transition between the metal and the semiconductor (Appendix A). For a <100> orientated silicon substrate, P_n is merely the transition probability of an electron of this energy moving perpendicularly to the barrier plane, while for a <111> substrate, the electron must have some of its total energy in a transverse direction to be able to make a transition (Appendix A). Consequently P_n is considerably smaller for <111> orientated silicon than for <100>.

P_n is slightly bias dependent, due to the changing shape of the tunnel barrier. However, the major bias dependence of J_{CT} originates from the F_1 term of Equation (5.4). This means that the electric field at the semiconductor-insulator interface, F_s, effectively determines J_{CT}.

In the region of interest, this field is given by (Appendix D)

$$|F_s| = \sqrt{2}\,[\frac{kT}{\epsilon_s} p_s + \frac{q}{\epsilon_s} N_d |\psi_s|]^{1/2} \quad (5.5)$$

for $p_s \leq N_v$, the valence band density of states. ψ_s is the electrostatic potential change across the semiconductor. For $N_d << p_s \leq N_v$, Equation (5.5) reduces to

$$|F_s| = (\frac{2kT}{\epsilon_s} p_s)^{1/2} \quad (5.6)$$

The current tunneling between the semiconductor valence band and the metal can be expressed as

$$J_{VT} = A_p T^{1/2} P_p p_s \quad (5.7)$$

for $p_s \leq N_v$. In the absence of surface states and at sufficient reverse bias to allow all the injected current, J_{pi}, to be collected by the contact, current balance at the semiconductor-insulator interface gives

$$J_{VT} = J_{pi} + J_{gen} \quad (5.8)$$

when $J_{pi} >> J_{gen}$, Equations (5.7) and (5.8) give

$$p_s = J_{pi}/A_p T^{1/2} P_p \quad (5.9)$$

This shows the direct proportionality between the injected current density and the hole concentration in the interface region.

With Equations (5.4), (5.6), and (5.9), there is sufficient

information to derive an expression for the small signal multiplication factor M.

$$M = \frac{d(J_{CT} + J_{VT})}{dJ_{pi}} \qquad (5.10)$$

i.e.

$$M = \frac{dJ_{CT}}{dJ_{pi}} + 1 \qquad (5.11)$$

Ignoring the bias dependence of P_n and P_p in comparison to more dominant terms gives the following relationships

$$\frac{dJ_{CT}}{dF_s} = A_n T^{1/2} N_c P_n \left(\frac{q}{kT} d \frac{\varepsilon_s}{\varepsilon_i}\right) F_0\left[\frac{q}{kT}\left(|F_s| \frac{\varepsilon_s}{\varepsilon_i} d + \phi_{si} - \phi_{mi} + \frac{q}{\varepsilon_i} dQ_i\right)\right] \qquad (5.12)$$

$$\frac{dJ_{pi}}{dF_s} = A_p T^{1/2} P_p \frac{\varepsilon_s}{kT} F_s \qquad (5.13)$$

Combining (5.11) - (5.13), and expressing F_0 in terms of more common functions gives

$$M - 1 = \frac{A_n}{A_p} \cdot \frac{P_n}{P_p} \cdot \frac{qdN_c}{\varepsilon_i} \cdot \frac{\ln\left(1 + \exp\left[\frac{q}{kT}\left(|F_s| \frac{\varepsilon_s}{\varepsilon_i} d + \phi_{si} - \phi_{mi} + \frac{q}{\varepsilon_i} Q_i d\right)\right]\right)}{|F_s|} \qquad (5.14)$$

Equation (5.14) is valid provided the argument of the exponential function appearing in it is not large and positive.

In the case where $\phi_{mi} = \phi_{si}$, F_s is large, Q_i is zero, and $A_n = A_p$, this can be reduced to

$$M - 1 > \frac{P_n}{P_p} \cdot \left(\frac{qdN_c}{\varepsilon_i}\right) \cdot \left(\frac{q}{kT} d \frac{\varepsilon_s}{\varepsilon_i}\right) \qquad (5.15)$$

With a two band tunneling model of the insulator (Chapter 2), the ratio of P_n to P_p is approximately equal to unity. Using this fact and substituting numerical values appropriate to the silicon-silicon dioxide system gives

$$M \sim 100 \tag{5.16}$$

This quantity should be compared to the results computed numerically in Figure 5.2(a). A change in the injected current density from 5 to 6 mA/m^2 changes the total diode current at 2 volts reverse bias from 5.6 to 6.5 Amps/m^2 giving a value of M under these bias conditions of 90.

To this point, the multiplication process has been discussed in the DC or low frequency case. The frequency response of the inversion layer depends on the same factors as in conventional MIS devices. Let Q_p be the total charge per unit area due to minority carriers in the IS interfacial region. Then, continuity of minority carriers gives

$$\frac{dQ_p}{dt} = J_{pi} + J_{gen} - J_{VT} . \tag{5.17}$$

Using the fact that holes are the dominant charge contributors to Poisson's Equation (Equation (D1)) in this interfacial region and that $J_{pi} >> J_{gen}$ gives

$$\epsilon_s \frac{dF_s}{dt} = J_{pi} - J_{VT} . \tag{5.18}$$

Applying Gauss's Law across the IS interface gives

$$\frac{\epsilon_i}{d}\frac{dV_{ins}}{dt} = J_{pi} - J_{VT} \qquad (5.19)$$

Imagine a small change, δJ_{pi}, suddenly made in the value of J_{pi} for a device originally under steady state conditions. This will eventually cause corresponding changes in V_{ins} and J_{VT}. The change in V_{ins} can be expressed as

$$\delta V_{ins} = \frac{dV_{ins}}{dJ_{pi}}\,\delta J_{pi} \qquad (5.20)$$

The maximum rate of change in V_{ins} occurs just after the change in J_{pi} and, from Equation (5.19), is given by

$$\left.\frac{\epsilon_i}{d}\frac{dV_{ins}}{dt}\right|_{max} = \delta J_{pi} \qquad (5.21)$$

This gives a lower bound on the time required for the change δV_{ins} to occur as

$$t_{min} = \delta V_{ins} \Big/ \left.\frac{dV_{ins}}{dt}\right|_{max} = \frac{dV_{ins}}{dJ_{pi}}\,\frac{\epsilon_i}{d} \qquad (5.22)$$

An upper bound upon the cutoff frequency of the multiplication process is therefore

$$\omega_c = \frac{d}{\epsilon_i} \Big/ \frac{dV_{ins}}{dJ_{pi}} \qquad (5.23)$$

$$\approx 2J_{pi}\,d/\epsilon_i V_{ins} \qquad (5.24)$$

where (5.24) is derived by evaluating the derivative using (5.6) and (5.7). Since J_{pi} is proportional to V_{ins}^2, it is seen that ω_c increases linearly with V_{ins} and as the square root of J_{pi}, the injected minority carrier current density.

This above analytical treatment predicts that the multiplication process is not restricted to the silicon system. Other MIS systems with thin insulating layers between the metal and the semiconductor should exhibit multiplication properties provided a metal of suitable work function is employed. Although the above treatment is for an n type semiconductor, a parallel development can be made for p type material. The major difference between the two cases is that the current multiplication factor M increases with decreasing metal work function for n type semiconductor but increases with increasing metal work function for p type.

The above derivation indicates that the quantities important in determining the multiplication properties of the MIS tunnel diode are ϕ_{mi} (Figure 5.1) which is related to the vacuum work function of the metal, d the insulator thickness, T the diode temperature, Q_i the fixed charge density in the insulator, N_D the dopant density in the semiconductor and the magnitude of the injected current density. The effects of these quantities will be determined for the metal-silicon dioxide-silicon system using the more general numerical approach, which is not subject to the restrictions it was found necessary to introduce to derive Equation (5.14).

5.2.4 Effects of the Substrate Resistivity and Orientation

As mentioned previously, the transmission probability P_n is considerably smaller for <111> orientated silicon than for <100> silicon. This should reduce the multiplication factor, M, for the <111> orientation. However, it is known experimentally that higher surface state densities are associated with the <111> orientation. As is shown later in Section 5.2, this will tend to reduce the differences between the two orientations, at least at large injected current densities.

To allow an evaluation of the effects of varying the resistivity of the substrate, the current-voltage characteristics of the reference diode (Figure 5.2(a)) will be discussed in more detail.

For each value of J_{pi}, the curve can be divided into three regions. The first region occurs at small inversion bias. In this region, the current-voltage characteristics are virtually independent of the injected current density. The MIS contact is unable to collect the injected current density, the injected carriers finding it easier to diffuse to the ohmic back contact. The current tunneling between the metal and the semiconductor is essentially all electron current. It is approximately constant because the semiconductor surface is in depletion and the surface field is small. The slope of the current-voltage characteristics in this region is determined by the doping density of the semiconductor.

In the second region, the injected carriers begin to be collected by the contact. The diode current increases approximately exponentially with increasing bias until the third region is reached. In the third

region, all the injected current density is collected by the MIS contact. The slope of the current-voltage characteristics is determined by a combination of two effects: the increase of current generated in the depletion region adjacent to the inversion layer; and the increasing field at the semiconductor surface due to the growth of the depletion region with increasing inversion bias (Equation (5.5)).

The current-voltage characteristics for a more heavily doped substrate (0.1 Ωcm) are shown in Figure 5.3. This illustrates primarily the enhanced slope in the first and the third regions described above.

5.2.5 The Effects of Metal Work Function and Injected Current Density

The metal chosen for the top contact to the diode plays an important role in determining the multiplication characteristics of the diode. This is demonstrated in Figure 5.4(a), where the small signal multiplication factor, M, is plotted as a function of the metal to insulator barrier energy (ϕ_{mi}) for two injected current densities. The curves obtained are in excellent qualitative agreement with the results predicted by Equation (5.14).

For large values of ϕ_{mi}, appreciable current multiplication occurs only at large values of injected current density. As ϕ_{mi} decreases, M increases rapidly until ϕ_{mi} is smaller than ϕ_{si}. Decreasing ϕ_{mi} past this point results in a much smaller rate of increase of M. In addition, for values of ϕ_{mi} less than ϕ_{si}, M decreases with increasing injected current density J_{pi}, whereas for large values of ϕ_{mi} the converse is true. This is demonstrated more clearly in Figure 5.4(b) where ϕ_{mi} is the variable parameter.

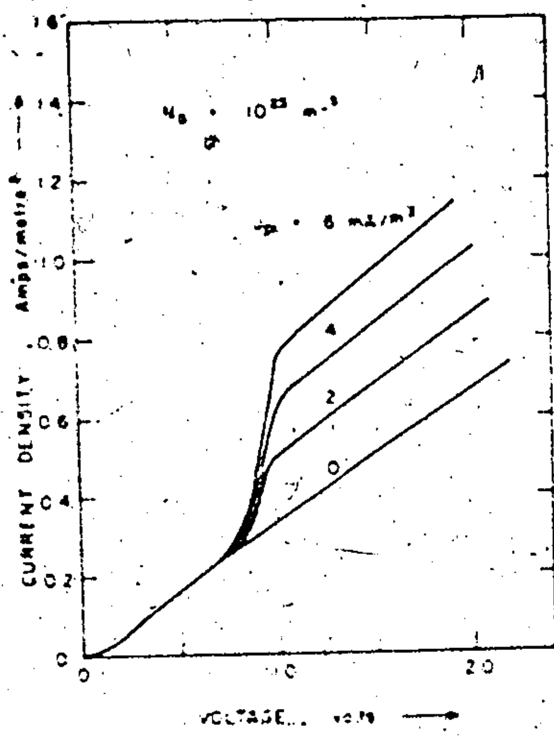


Figure 5.3: Computed current-voltage characteristics for a MIS tunnel diode similar to the reference diode of Figure 5.2(a) but with an impurity density of 10^{23} per cubic metre (0.1 Ωcm) in the silicon substrate.

Figure 5.4(a): The small signal multiplication factor, M, versus the metal to insulator barrier height, ϕ_{mi}, with the supply rate of holes to the semiconductor surface (J_{pi}) as the variable parameter. Other parameters as in Figure 5.2(a).

(b): The small signal multiplication factor, M, versus the supply rate of holes to the semiconductor surface, J_{pi}, with ϕ_{mi} as the variable parameter for the diode of (a).

(c): Calculated current-voltage characteristics of a MIS tunnel diode similar to the reference diode of Figure 5.2(a) but with the metal to silicon dioxide barrier energy decreased to 2.9 eV.

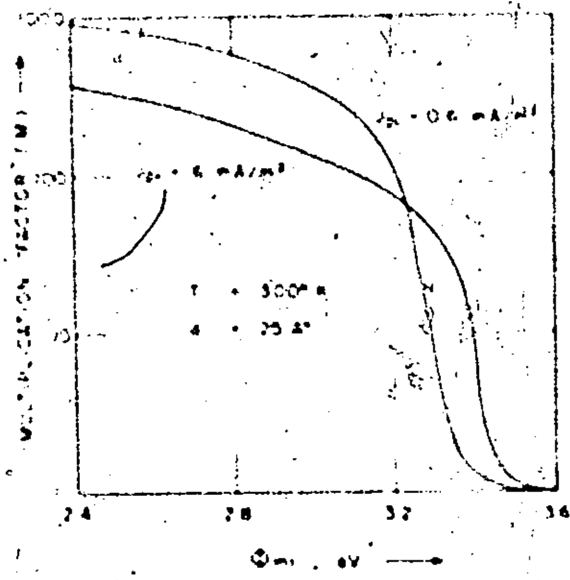


FIGURE 5.4(a)

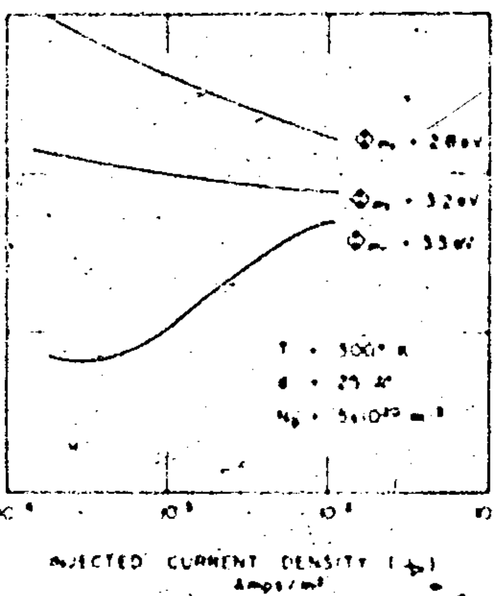


FIGURE 5.4(b)

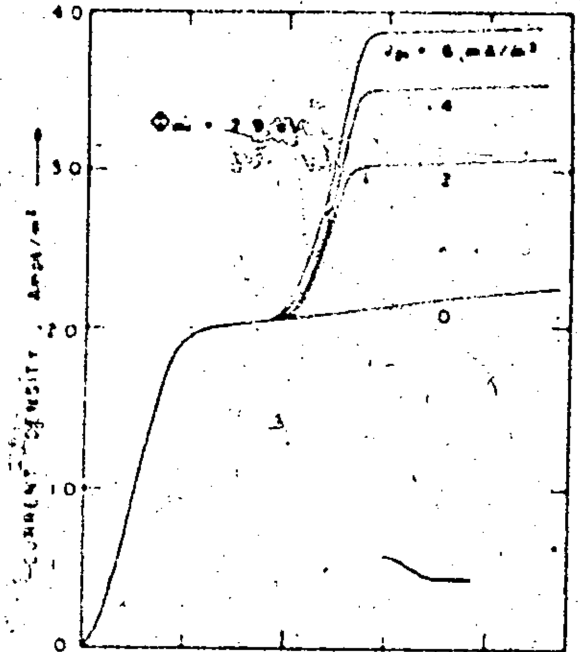


FIGURE 5.4(c)

It has been shown that decreasing ϕ_{mi} increases the multiplication factor, M. It also has the less desirable effect of increasing the current density when the injected current is zero. This is illustrated by the computed current-voltage characteristics for a value of ϕ_{mi} equal to 2.9 eV (Figure 5.4(c)). The diode chosen as a reference with ϕ_{mi} equal to 3.2 eV offers a reasonable compromise between large current multiplication and a small device current flow for zero injected current (Figure 5.2(a)).

5.2.6 The Effects of Oxide Charge and Oxide Thickness

As is illustrated by Equations (5.4) and (5.14), the effects of charge stored in the oxide are very similar to those caused by a change in the metal work function. This is demonstrated by the current-voltage characteristics shown in Figure 5.5(a) for the case of 10^{15} positive charges/m^2 in the oxide near the oxide-semiconductor interface. The current density for zero injection of minority carriers is increased as well as the small-signal multiplication factor M. Note that this charge density is nearly equivalent to a decrease of ϕ_{mi} equal to 0.012 eV.

The thickness of the oxide layer is an extremely important parameter of the MIS system. If the oxide layer is thicker than 30-35Å, the current generated in the surface region will be capable of supplying all the minority carrier current the contact can collect. This effectively eliminates the third region discussed in Section 5.2.4. Decreasing the oxide thickness from 25Å to 20Å allows the diode to operate at much higher injected current levels. Figure 5.5(b) shows the current-voltage characteristics of the reference diode with the oxide thickness reduced to

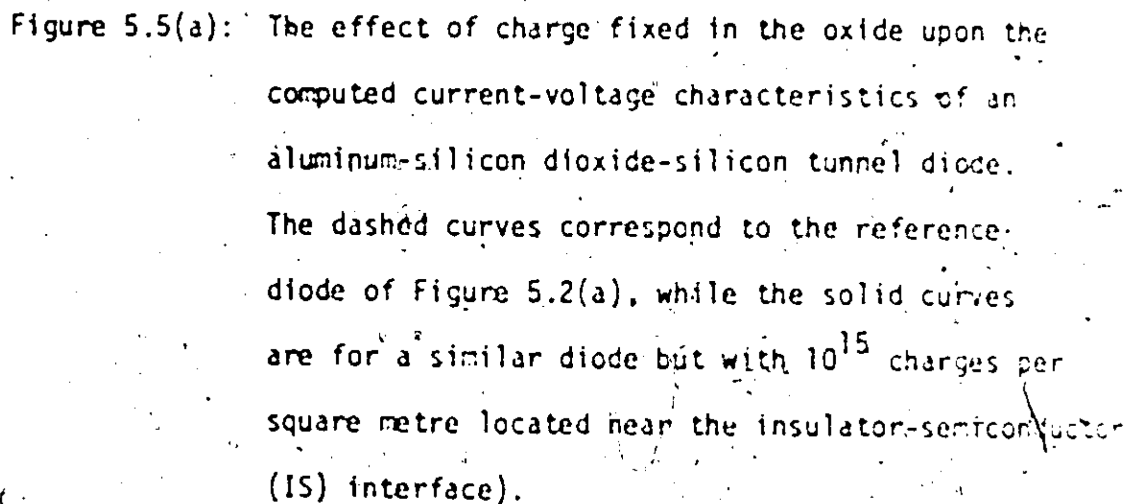


Figure 5.5(a): The effect of charge fixed in the oxide upon the computed current-voltage characteristics of an aluminum-silicon dioxide-silicon tunnel diode. The dashed curves correspond to the reference diode of Figure 5.2(a), while the solid curves are for a similar diode but with 10^{15} charges per square metre located near the insulator-semiconductor (IS) interface).

(b): Computed current-voltage characteristics of a MIS tunnel diode similar to the reference diode but with the layer of silicon dioxide reduced to 20Å thickness.

(c): The small signal multiplication factor, M, versus the supply rate of minority carriers (J_{pi}) for both the reference diode with the silicon dioxide layer 25Å thick and a similar diode with this layer reduced to 20Å.

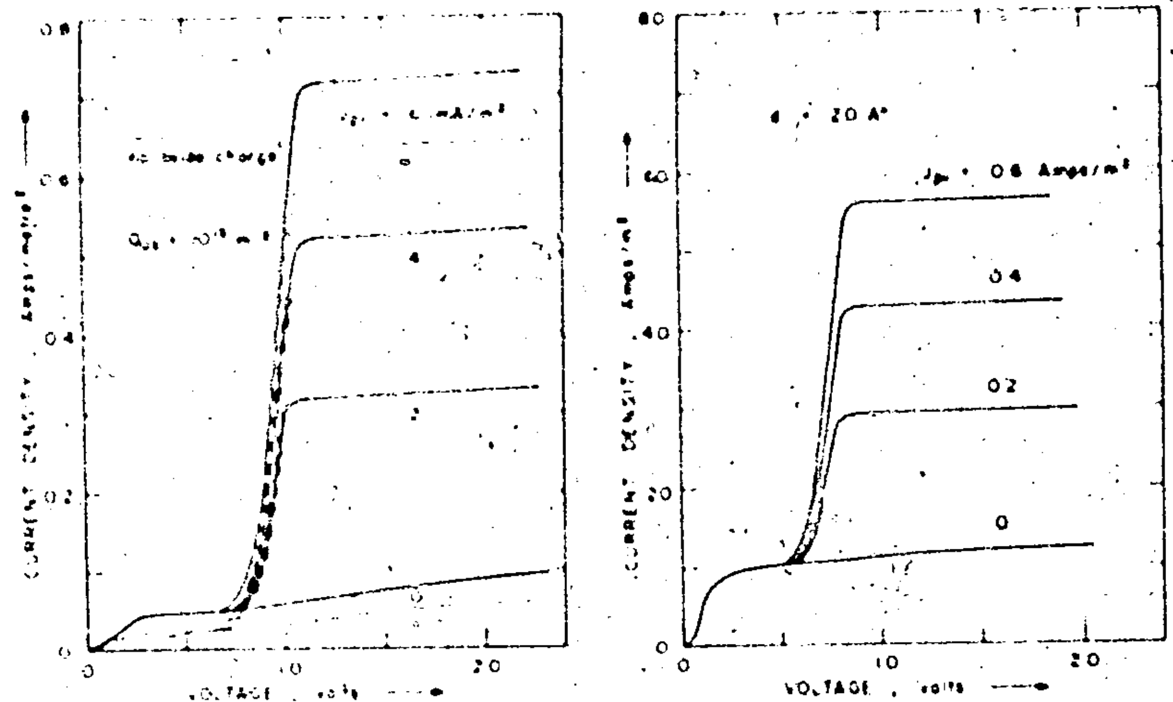


FIGURE 5.5(a)

FIGURE 5.5(b)

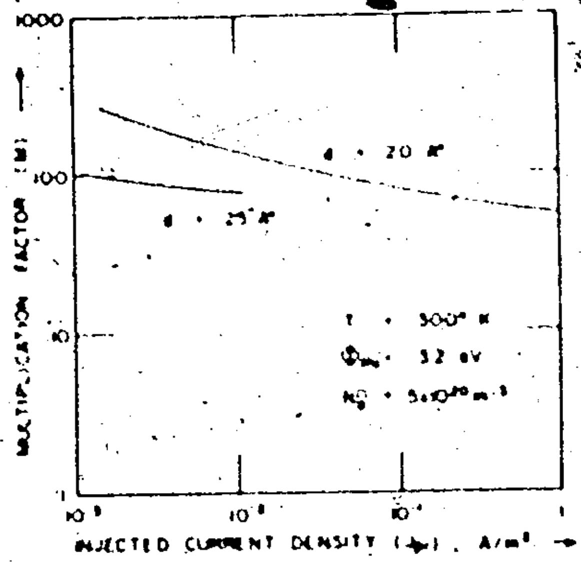


FIGURE 5.5(c)

20Å. All current levels have been increased by two to three orders of magnitude. Figure 5.5(c) compares the small signal multiplication factor, M, for diodes with 20Å and 25Å. Note that the current flow for J_{pi} equal to zero is considerably larger for the thin oxide diode.

5.2.7 Effects of Temperature and Frequency Response

The results of varying the diode temperature from 300°K to 200°K are given in Figure 5.6. It is seen that the characteristics of this diode are relatively insensitive to temperature variations. However, as indicated by Equation (5.14), the temperature performance of the diode is dependent upon the work function of the metal contact. Lower work function contacts are expected to exhibit larger temperature variations (Chapter 4). Experimental measurements of the temperature variations in Al and Au majority carrier devices have been described in Chapter 4 for the case when J_{pi} is zero.

The computed frequency response of the 20Å diode of Figure 5.5(b) is shown in Figure 5.7. The small signal gain, M, is shown as a function of the value of injected current bias, J_{pi}, with the frequency of the superimposed small signal disturbance as the variable parameter. For small values of J_{pi}, the frequency response of the process is very low. As J_{pi} increases, the frequency response also increases as predicted by Equation (5.24). Since the voltage across the oxide also increases with J_{pi}, the breakdown of the insulator will limit the maximum frequency response of the process. The relatively low frequency response of the device is expected to restrict its practical application. Results of experimental measurement of the frequency response of experimental devices

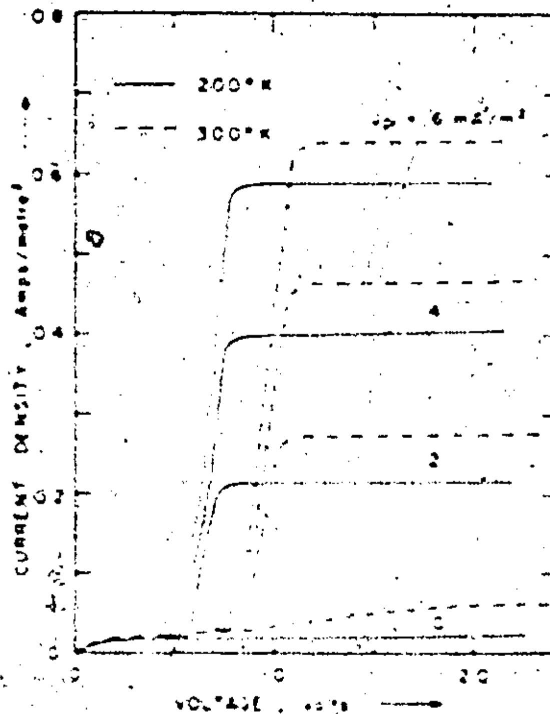


Figure 5.6: The effect of temperature upon the computed current-voltage characteristics of the reference aluminum-silicon dioxide-silicon tunnel diode. The dashed curve corresponds to a temperature of 300°K while the solid curve is for a temperature of 200°K.

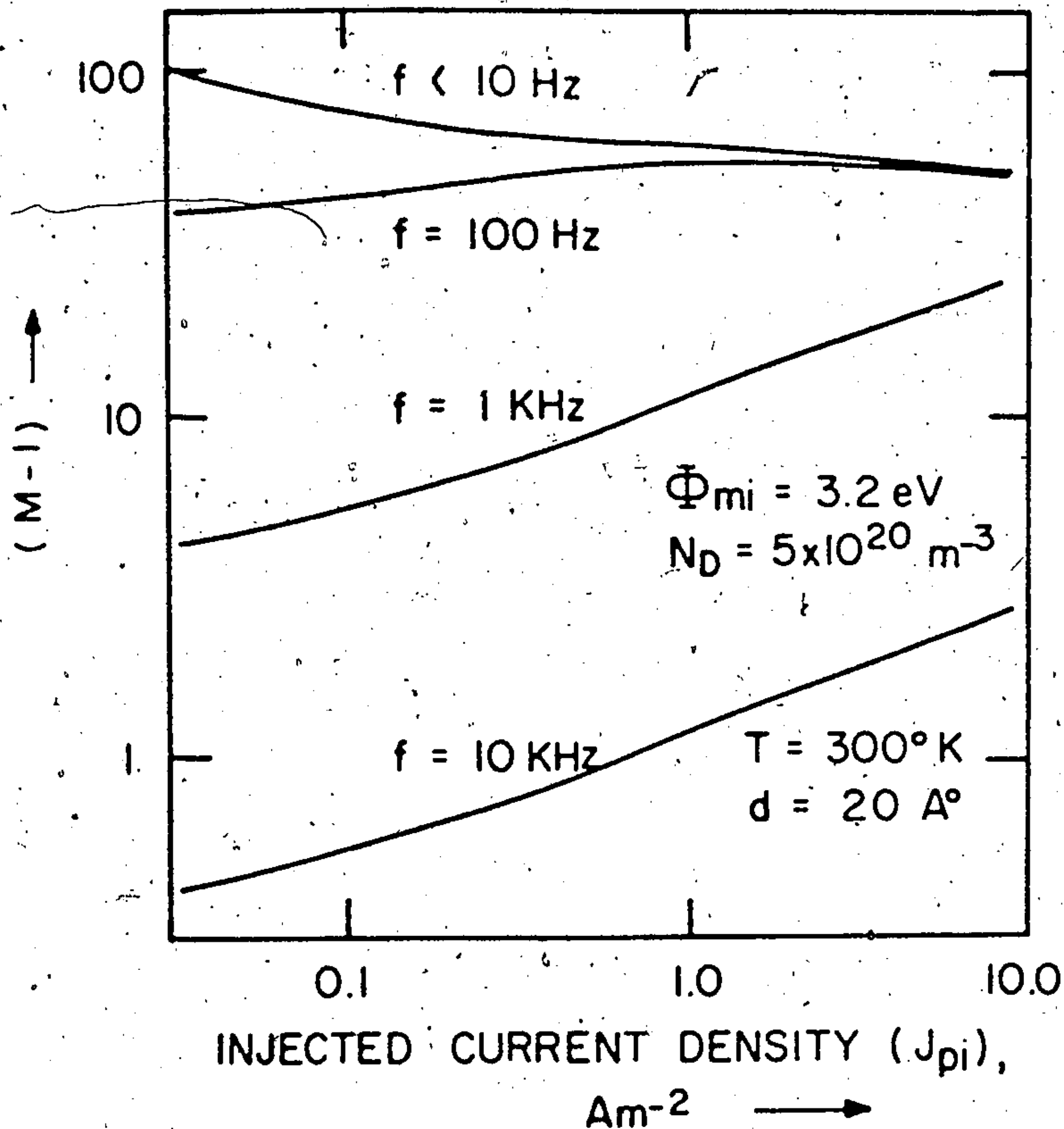


Figure 5.7: The calculated frequency dependence of the small signal multiplication factor, M, for a diode similar to the reference diode of Figure 5.2(a) but with the insulator thickness reduced to 20Å.

are described in Section 5.3.3.

5.2.8 The Effects of Surface States

Surface states associated with the semiconductor-insulator interface can cause deviations from the theory of the previous sections. These deviations arise due to the following properties of these states:

(a) they are charge storage centres, the amount of charge stored depending on the bias conditions;

(b) they provide a spatially concentrated region of recombination-generation centres;

(c) they can cause additional current flow betweem the metal and the semiconductor by tunneling transitions between the metal and the surface states.

All three of the above properties were included self-consistently into the model used to characterize surface states as described in Chapter 2.

To enable a simpler discussion of the effects of surface states upon the properties of the reference diode, the current tunneling between the metal and the surface states is initially treated as being very small; i.e. the effective tunnel capture cross sections (σ_T) of the states was set equal to zero. Assuming the surface state distribution of Figure 5.8 and a fixed charge in the oxide of 10^{15} charges/m^2, the computed current-voltage characteristics of the diode are shown in Figure 5.9(a). The effects of the surface states in this case are primarily as charge storage centres.

The current-voltage characteristics for zero injected current

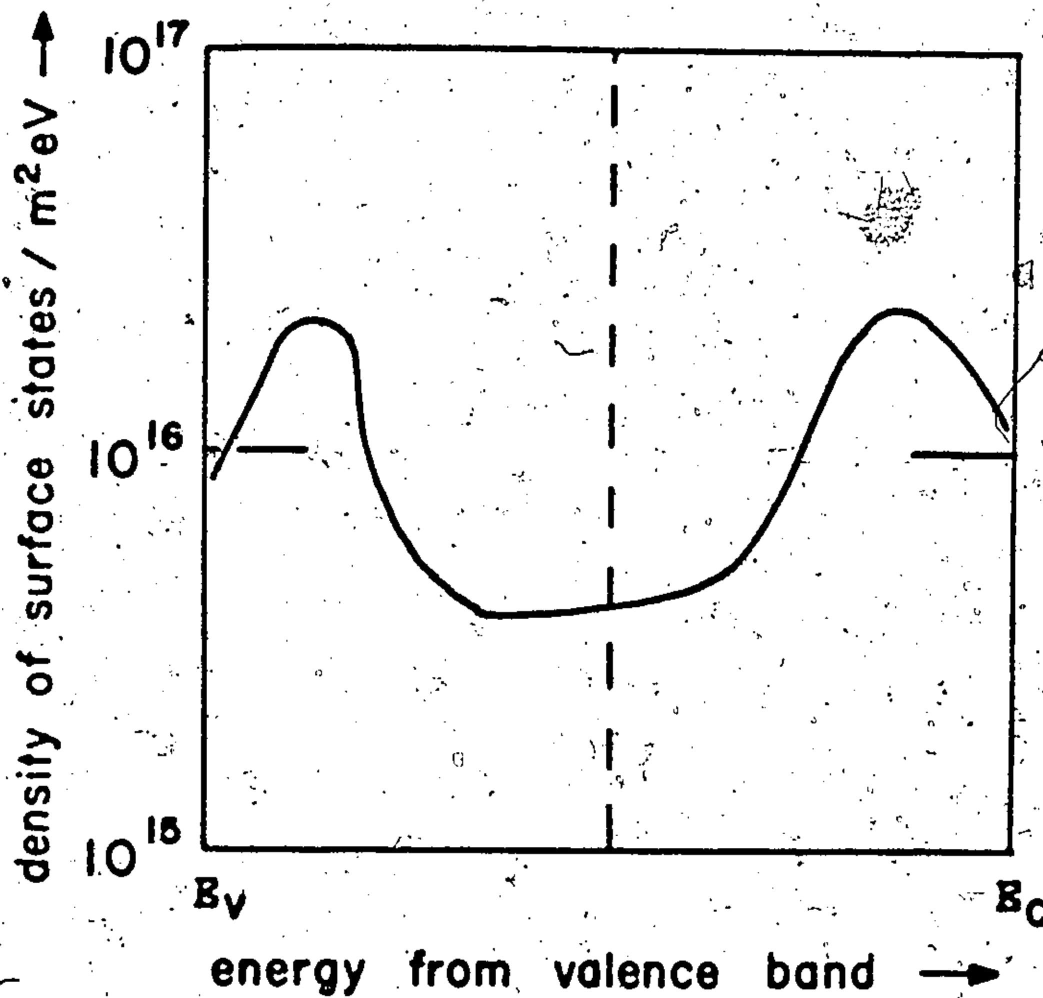


Figure 5.8: Assumed distribution of surface states across the energy gap of silicon for the subsequent calculations in Chapter 5. States below midgap were modelled as having donor type properties while those above midgap were modelled as acceptor type levels.

Figure 5.9(a): The effect of surface states upon the current-voltage characteristics of the reference diode of Figure 5.2(a) with currents tunneling between the metal and the surface-states negligible ($\sigma_T \to 0$) and these states acting only as charge storage and recombination centres. The assumed distribution of surface states is shown in Figure 5.8.

(b): The effect of surface states upon the current-voltage characteristics of the reference diode in the case where there is appreciable current tunneling between the metal and these states in addition to the role of these states as charge storage and recombination centres.

(c): The small signal current multiplication factor, M, versus the supply rate of minority carriers to the surface region (J_{pi}) for different surface state properties. The two cases of (a) and (b) are plotted as well as the case of negligible surface state density (Figure 5.2(a)).

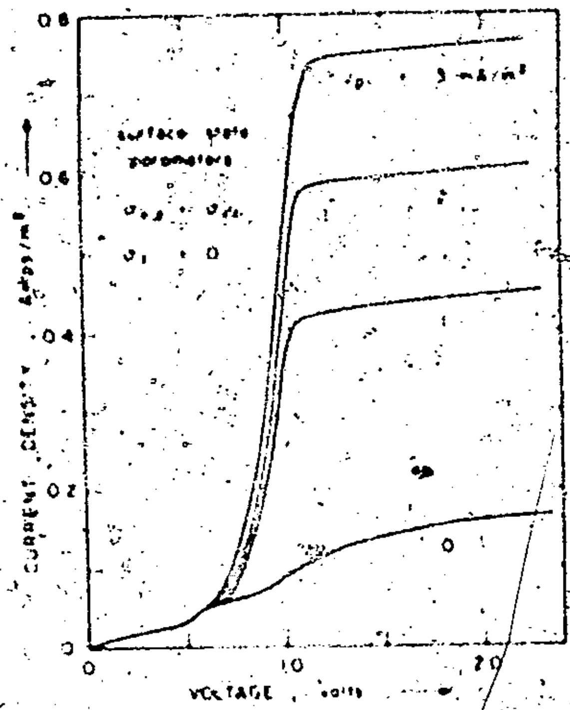


FIGURE 5.9(a)

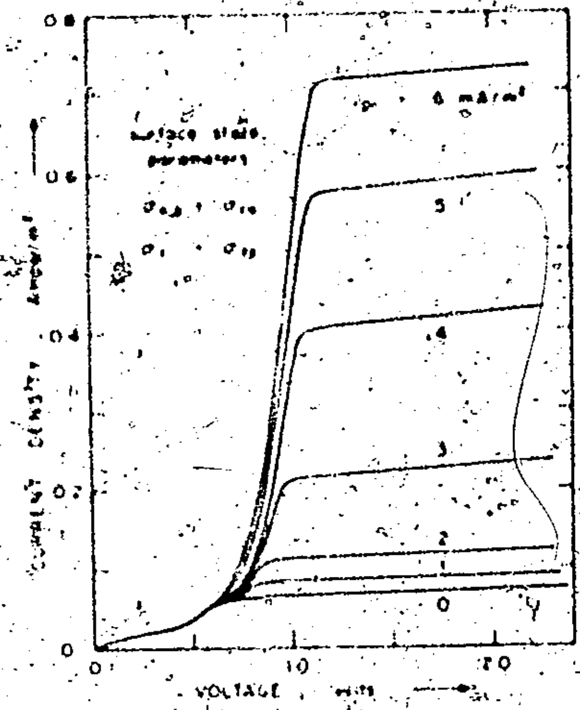


FIGURE 5.9(b)

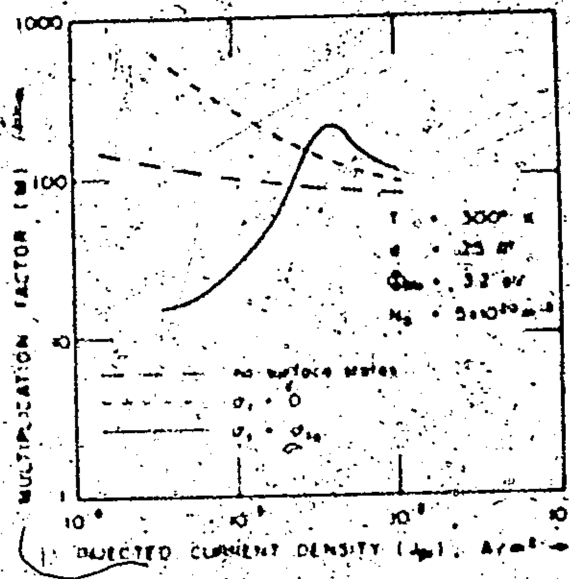


FIGURE 5.9(c)

density, J_{pi}, exhibit fluctuations caused by the increasing positive charge contribution of the surface states as the hole concentration in the surface region increases. The tendency for the positive charge stored in the surface states to increase with increasing hole concentration at the surface aids the multiplication process at least for the smaller values of J_{pi}. As the surface becomes more strongly inverted, the surface states are effectively all occupied by holes and contribute the maximum positive charge. Recombination-generation processes at these states are considerably reduced under these circumstances.

Thus at high injected current densities, the multiplication properties of the diode can be related to those of an ideal diode of Section 5.2.3 with a charge at the semiconductor-insulator interface equal to Q_{ss}^{+} where Q_{ss}^{+} is the maximum positive charge contribution from the combination of the surface states and charges in the oxide.

The computed current-voltage characteristics for the case where there is appreciable tunnel current flow between the metal and the surface states are shown in Figure 5.9(b). Although experimental values of σ_n and σ_p, the surface state capture cross sections for electrons and holes in the semiconductor conduction and valence bands, have been reported for these thin oxide devices [17,85,86], σ_T, the effective tunnel capture cross section, has not yet been evaluated experimentally. It was assigned the value of $4 \times 10^{-21} m^2$ which falls within the range of reported values of σ_n and σ_p. For the diode of Figure 5.9(b) under the moderate reverse bias conditions of interest, the computed value of the current tunnelling between the metal and the surface states, J_{ST}, is relatively bias independent and has a value of 3-4 mA/m^2 with this

choice of σ_T and the surface state distribution of Figure 5.8.

For large values of J_{pi}, the device operates as described above. For smaller values of J_{pi}, a new element is introduced into the operation of the diode. In this region, the dominant currents supplying and removing holes to and from the surface region are J_{pi} and J_{VI} (Figure 5.1). The hole concentration in the surface region adjusts itself to establish a balance between the two processes. If J_{pi} is much smaller than J_{ST}, the hole concentration at the surface is small. As J_{pi} increases, this concentration also increases so that multiplication can still be obtained in this region. The multiplication properties are dependent upon the surface state distribution and are generally inferior to those obtained for higher values of J_{pi}.

These results are illustrated in Figure 5.9(c). Shown are the multiplication factors for the reference diode with different surface state properties. It is seen that the gain at high injected current levels is slightly enhanced when surface states are present. At small current levels the multiplication is reduced, the current density for the transition between the two regions depending upon the density of surface states and their effective capture cross sections for particles tunneling between the metal and the surface states.

The frequency response of the multiplication process in the presence of surface states is demonstrated in Figure 5.10 where the surface state properties "c" of Figure 5.7 have been employed. Note how the peak apparent in the low frequency curves becomes smoothed out in the high frequency curves. This peak is caused by positive charges accumulating in surface states as the inversion layer is formed, a

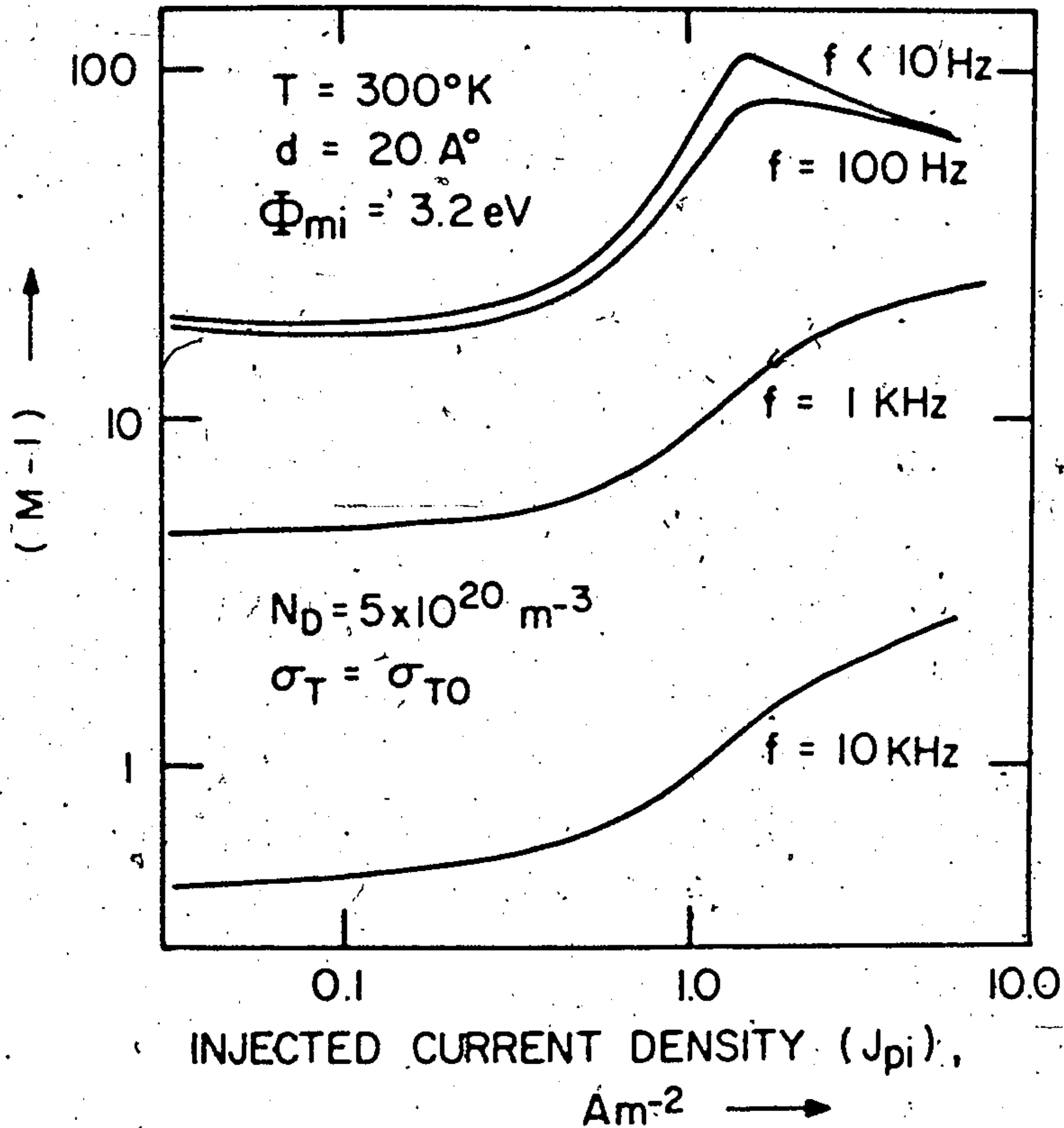


Figure 5.10: The calculated frequency dependence of the small signal multiplication factor, M, in the presence of surface states. The diode analyzed is similar to that of Figure 5.9(b) except the insulator thickness has been reduced to 20Å.

process with a restricted frequency response.

5.3 Experimental Measurement of the Multiplication Process

5.3.1 Experimental Method

The gains available from the multiplication process in majority carrier devices can be investigated using a simple experimental technique. The use of this technique depends upon two properties of the MIS tunnel diode established in this thesis. The first is demonstrated in Chapter 4 where it is shown that devices with the same metal contact form majority or minority carrier devices depending upon whether the substrate is n or p type. The second will be treated in Chapter 6, where it will be shown that minority carrier devices have properties similar to p-n junction devices when the insulating layer is very thin (< 25Å). In particular, these devices have the same photovoltaic properties so that minority carrier devices can be used as visible light detectors, one application of silicon p-n junctions.

Using this fact, the experimental arrangement employed to evaluate the multiplication properties of majority carrier devices is illustrated in Figure 5.11. Light is used to generate the injected current density (J_{pi}). White light from a source whose intensity could be adjusted impinges upon a pair of MIS tunnel diodes. These diodes have semiconducting regions of similar resistivity and minority carrier lifetime but one is n type and the other is p type. With the Al top contacts used exclusively in this chapter, the n type device is the majority carrier device whose properties are to be measured, while the p type minority carrier device operates essentially as a detector. The insulator

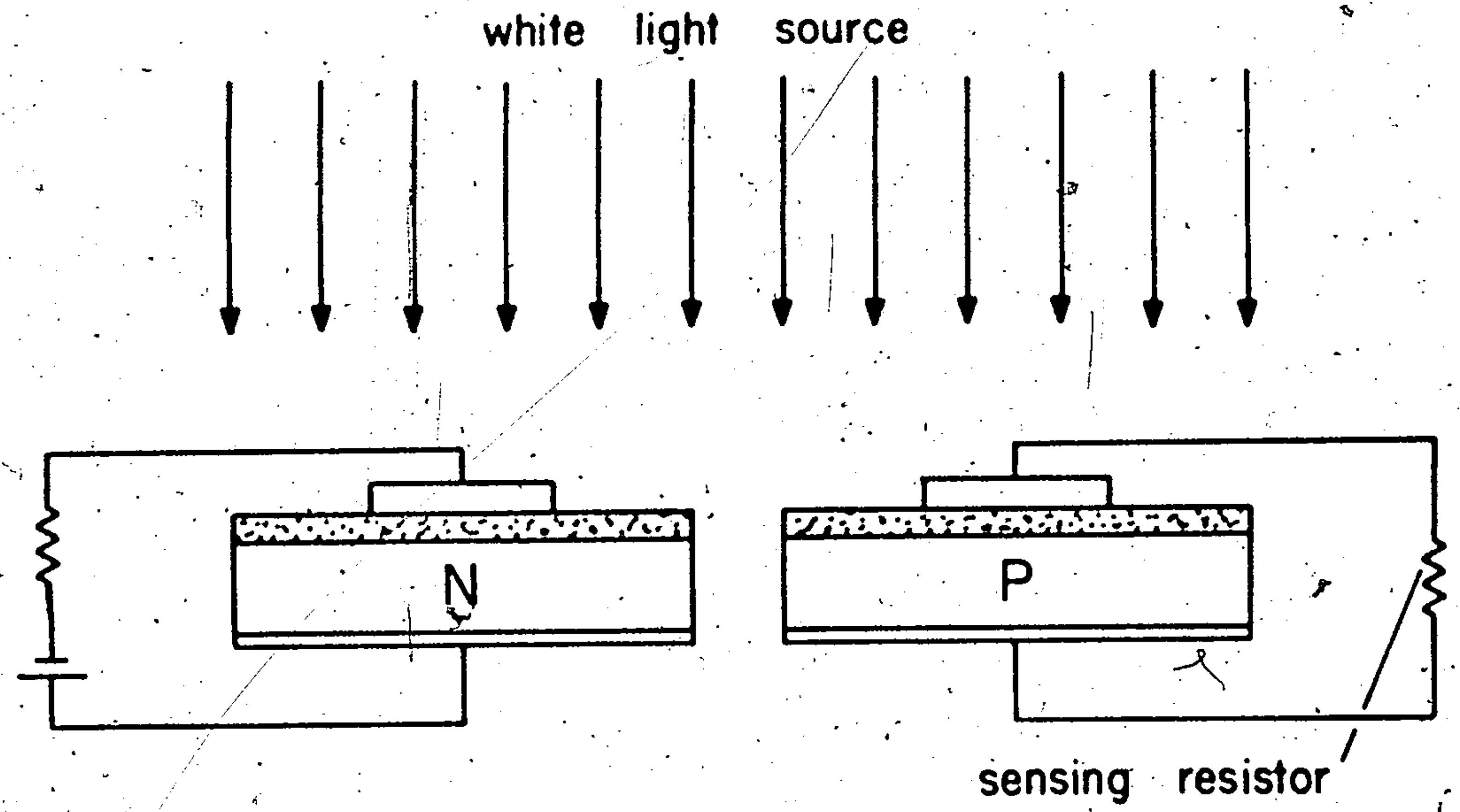


Figure 5.11: Experimental method for investigating the multiplication properties of the majority carrier MIS tunnel diode.

thickness in the minority carrier device is less than 20Å to ensure it operates in the desired detection mode over the intensity range of interest. The insulator thickness in the majority carrier diode is not subject to this restriction. The Al top contact is very thin (< 200Å) and has been evaporated at the same time for the n and p devices. It then has near identical optical properties for each device (in particular a similar transmission coefficient).

This experimental arrangement allows the multiplication properties of the majority carrier diodes to be measured simply. Since the n and p devices have similar bulk properties, the collection of the minority carriers generated in the semiconductor by the light passing through the metal contact will be nearly equal in each case (Section 6.4). This collection rate will be reflected in the current flowing in the minority carrier device, J_{min}, under short-circuit or inversion bias conditions. Thus J_{min} can be identified as the injected current density in the majority carrier device, J_{pi}.

The minority carrier device was shunted by a small resistance (effectively a short circuit) while the majority carrier device is shunted by a similar resistor in series with a DC power supply to bias the diode into the multiplication region (top contact typically 2-3 V negative with respect to semiconductor). DC measurements of the multiplication factor, M, are made simply by varying the intensity of the light source and recording the current flow in the majority carrier device, J_{maj}, as a function of the corresponding current in the minority carrier device, J_{min}. The factor, M, is then obtained simply by differentiating this curve. AC measurements of this quantity are made by

superimposing a small time dependent light signal onto the bias light signal. The resulting small signal response of J_{maj} and J_{min} are then recorded using an oscilloscope monitoring the voltages across the sensing resistors. One time dependent signal used is chopped light which approximates a square wave excitation. The other is a single flash from a stroboscope. Compared to the relatively slow majority carrier diode response, this can be regarded as an impulse excitation. Fourier transform techniques can then be used to extract frequency domain information from the resulting transient response.

The arrangement described above uses light which is transmitted through the metal contact as the source of the injected current, J_{pi}. A less efficient method of injecting current using light is possible even with thick, opaque metal contacts. Light striking the semiconductor surface in regions not covered by the metal contact is absorbed into the semiconductor in the usual way [87]. A small fraction of the minority carriers generated in this way can be collected by the MIS tunnel diode. Theoretically, because of their similar bulk properties, both the minority and majority carrier devices should collect the same quantity. However, it is known experimentally that the two configurations differ in that the surface of one tends to be accumulated while that of the other ranges from depletion to inversion (Section 2.6.1). While it can be said in the thin metal case that J_{min} is nearly identical to J_{pi} (expected difference less than 30%), the situation with the present injection method is more uncertain. While J_{pi} is expected to be proportional to J_{min}, the constant of proportionality is difficult to obtain theoretically.

Even without assigning a value to this constant, this method can be used to obtain quantitative frequency response data as well as the shape of the M versus J_{pi} curve.

A third method uses the "ring-dot" structure sketched in Figure 5.16(a). The dot forms the contact whose properties are being studied while the ring is used to inject minority carriers. Since both the ring and dot contacts were made with the same metal, both are majority carrier devices and consequently the ring is a poor injector of minority carriers. Moreover, the variation of injection ratio with the ring bias is unknown. These factors make this arrangement the least satisfactory for a quantitative study of the multiplication properties. The structure is of interest since the results obtained support those obtained by the other methods in a qualitative fashion and since it provides an introduction to the Surface Oxide Transistor device described in Section 5.4.

5.3.2 Experimental Results

The measured J_{maj} versus J_{min} curves for four sets of devices are illustrated in Figure 5.12(a). Details of the device parameters are given in Table 5.1. Two sets (N-P0516, N-P0517) had thin metal contacts (100Å to 200Å) and used the first method of injecting current. An evaluation of the p type devices indicated that the transmission coefficients of the metal films were in the range 0.20-0.25. This is in good agreement with earlier measurements upon such Al films [88]. Most of the incident light is reflected. The other two sets had thick contacts and used the second injection method. The factor determining the upper limit to the measured current densities was the majority diode

Figure 5.12(a): Experimental plots of J_{maj} versus J_{min} as the intensity of the light source in Figure 5.11 is varied. The parameters of the four pairs of devices shown are given in Table 5.1.

(b): Experimentally determined values of the low frequency value of the small signal gain, M', obtained by differentiating the curves of (a).

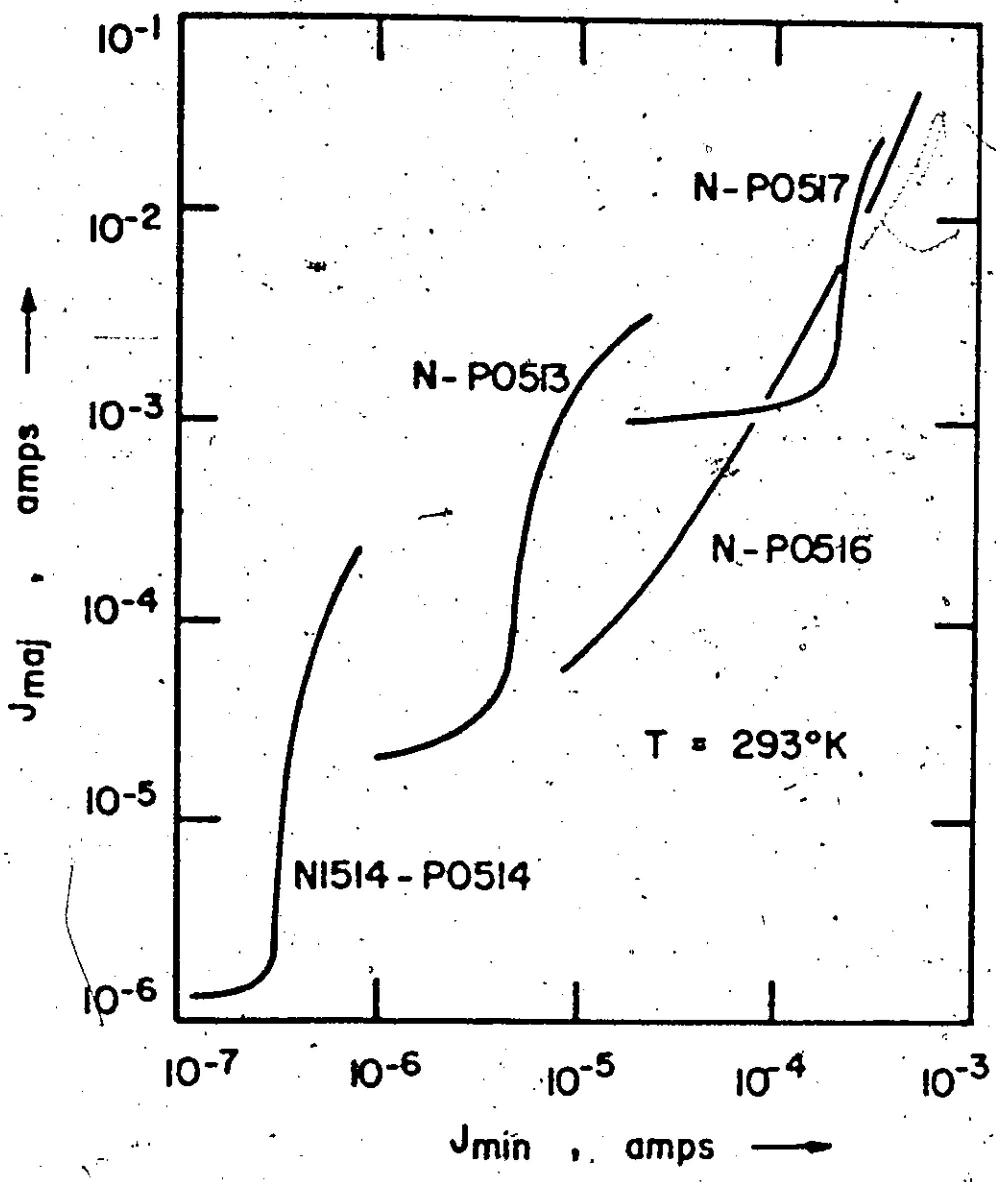


FIGURE 5.12(a)

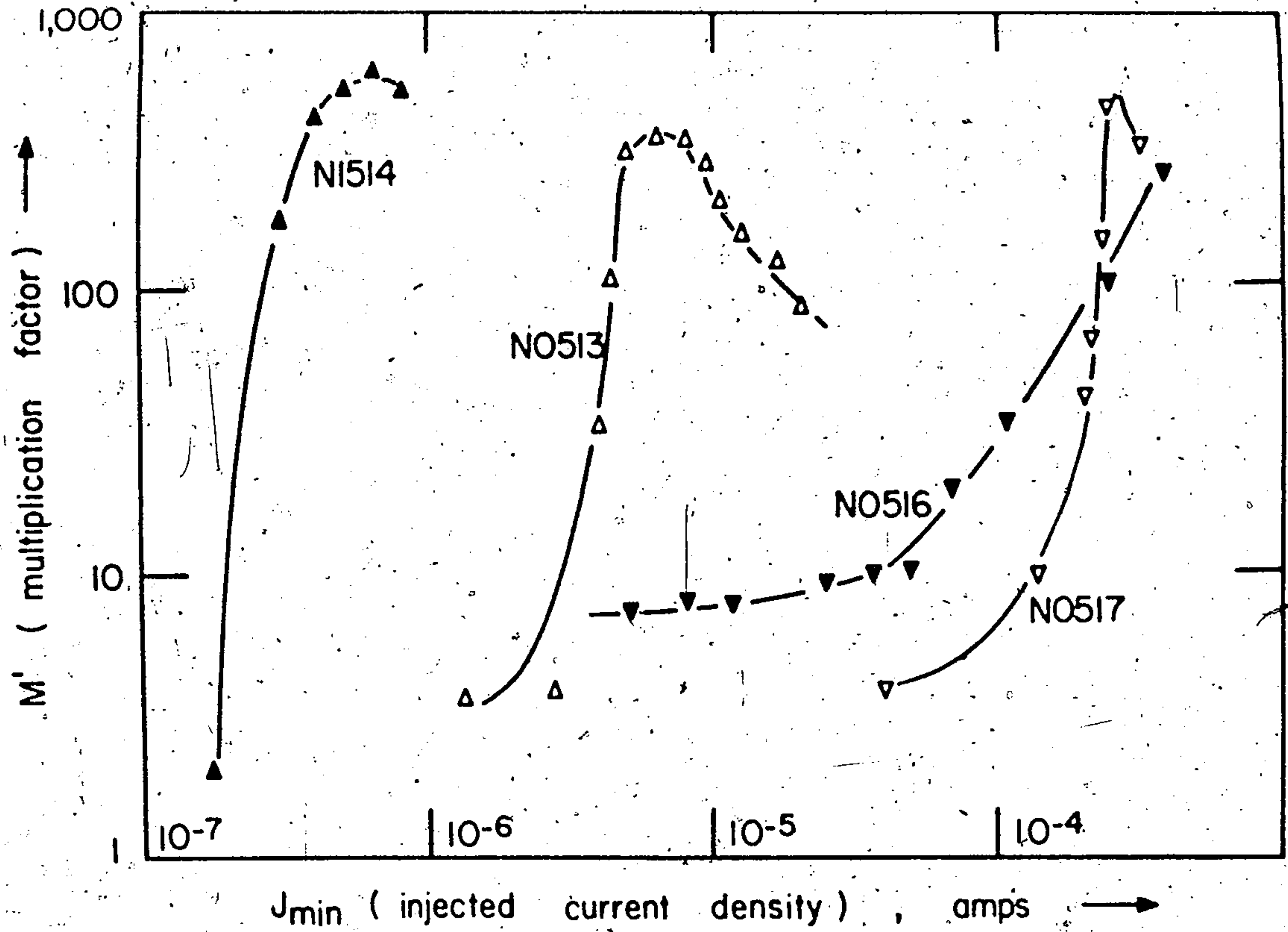


FIGURE 5.12(b)

TABLE 5.1: Details of Experimental Devices Used to Study the Multiplication Process

Device	Substrate	Area (mm^2)	d_{est} (Å)
N0513	0.1 Ωcm <100>	1.6	25
P0513	1 Ωcm <100>	1.6	18
N1514	1-3 Ωcm <111>	1.6	27
P0514	1 Ωcm <100>	1.6	18
N0515	1-3 Ωcm <100>	0.11	24
P0515	1 Ωcm <100>	0.11	18
N0516	1-3 Ωcm <100>	32.0	23
P0516	1 Ωcm <100>	32.0	18
N0517	0.1 Ωcm <100>	8.0	22
P0517	1 Ωcm <100>	8.0	18

bias voltage (usually 2 to 3 V). At the limiting current densities most of this voltage fell across the insulator eliminating the "third region" discussed in Section 5.2.4. Examination of the multiplication properties past this point required that this bias be increased further. Since the breakdown of the oxide occurs at a value given by 0.15 d (where d is the oxide thickness in Å) in thick insulator devices [89], this eventually led to a deterioration in the insulator properties and an irreversible change in the diode characteristics.

Curves of M' ($= dJ_{maj}/dJ_{min}$) versus J_{min} corresponding to Figure 5.12(a) are shown in 5.12(b). For the thin metal devices, M' and M are nearly identical. For the other two devices, this statement cannot be made with the same certainty but it is true that M' is proportional to M. Furthermore, it is known that the minimum value possible for M is unity. By equating the minimum value of M' measured to unity and dividing the rest of the curve by this value gives a conservative lower bound on M. The peak value of M obtained even with this conservative interpretation is over 100 in both cases. The following general features in the behaviour of M are evident. M is small at low injected current densities and increases rapidly after a certain threshold value of this current density is attained. M increases to values in the range 100 to 1000 for all the devices. In three (N1514, N0513, N0517), a peak gain was observed while in the other case (N0516), the gain was still increasing at the largest current density values.

The sudden increase in M after a threshold value of J_{pi} is reached can be accounted for in terms of the previous theory in one of two ways. If ϕ_{ms} (Equation 4.1) is small and positive, there will be a sudden

increase in gain when the voltage across the oxide becomes large enough to project the metal fermi level above the conduction band edge (Figure 5.1). A theoretical curve corresponding to this case is the 3.3 eV curve of Figure 5.4(b). Another possible interpretation is to attribute this feature to surface state effects. When the injected current density approaches the value of the current tunneling between the metal and surface states, there is also a sudden increase in gain as described in Section 5.2.7 and illustrated in Figure 5.9(c).

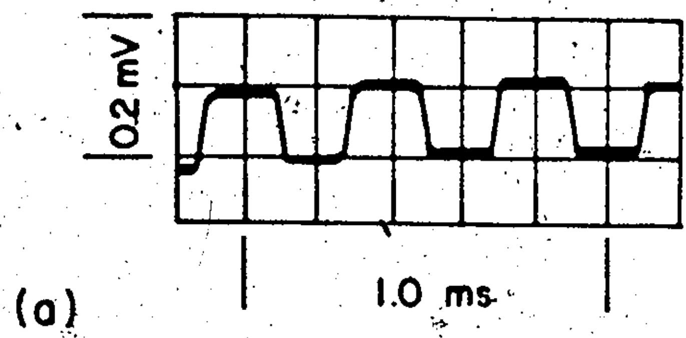
In addition to these DC measurements, AC measurements were made upon the device N0515 using P0515 as a detector. These devices had thick contacts and used the second mode of current injection. Not only was the n type device subject to a voltage bias, but both devices were subject to a light bias. A small chopped light signal was superimposed upon this bias light. The response of the p type device (P0515) was independent of the bias light and is shown in Figure 5.13(a). The response of the n type device (N0515) was very dependent upon the bias light level and is shown in Figure 5.13(b) for a low level and in (c) for a higher level. These curves show the expected features. The response of the minority carrier device is small but it has no difficulty in following the input light signal. At the small bias light level, the n type device had a response larger than the p type device, but had a restricted dynamic range. Note the opposite direction of the response. As the bias light level increases, not only does the magnitude of the response increase by a factor of 10, but the dynamic range has also improved significantly.

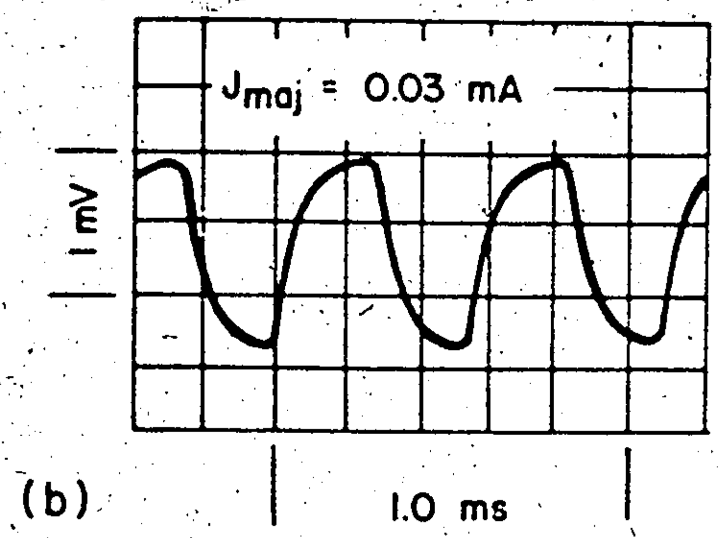
Figure 5.13: Experimentally observed response of two MIS tunnel diodes, P0515 and N0515, to a small chopped light signal superimposed upon the bias light. The diode parameters are given in Table 5.1.

(a): response of the minority carrier diode, P0515 (independent of bias light intensity).

(b): response of the majority carrier diode, N0515, when the bias light is of a relatively low intensity. The device is voltage biased at 2.5V.

(c): response of the majority carrier diode, N0515, when the bias light is of higher intensity. The device is voltage biased at 2.5V.





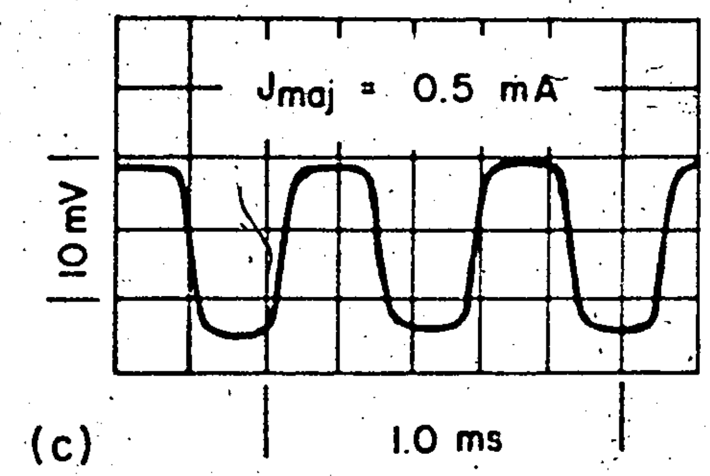


FIGURE 5.13

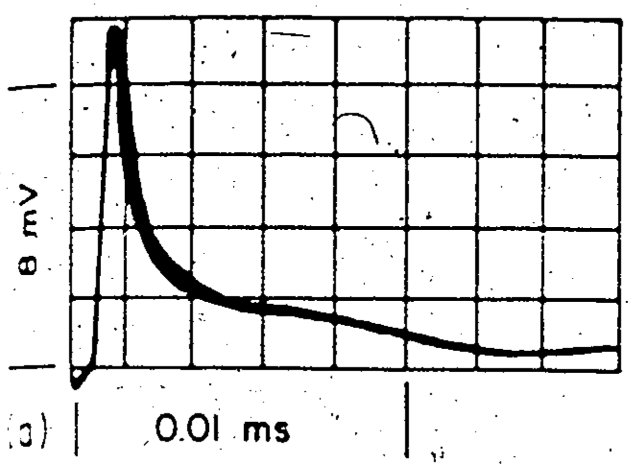
A method more suitable for the quantitative evaluation of these effects was to record the response of the devices to a small impulse of light superimposed upon the bias light level. The response of the p type device (P0515) is shown in Figure 5.14(a) and that of the n type device (N0515) under different light levels in Figures 5.14(b)-(f). The response of the p type device is shown on an expanded time scale and is much faster than the fastest response of the n type device. If the p type response is regarded as an impulse excitation, the responses of the n type device can be analyzed in the following way. The Fourier transform of each response gives the response of the multiplication process in the frequency domain [90] for that particular bias light level. The area under each response curve is therefore proportional to the low frequency gain of the process, while the characteristic time constant associated with the decay of the response is simply related to the cut-off frequency of the process. Table 5.2 contains a summary of the measurements made upon this pair of devices as well as the value of the low frequency gain, M', computed from the impulse response and a tabulation of the cut-off frequency at the different bias light levels. The DC and the transient methods give similar values form M' as is demonstrated in Figure 5.15. The measured cut-off frequencies are in order of magnitude agreement with the value of 2 kHz predicted by Equation 5.21 assuming J_{min} equals J_{pl}.

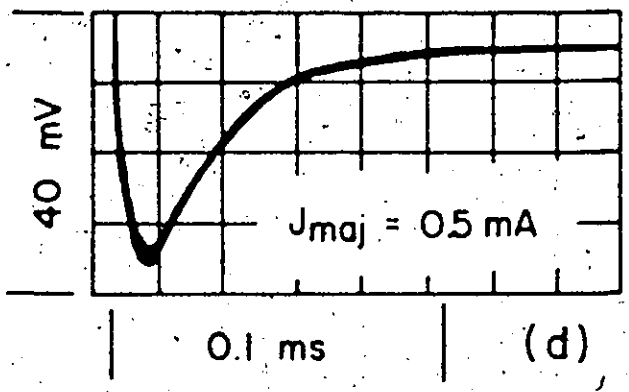
Injection of minority carriers from a nearby contact gave similar results for the DC properties of the multiplication process. Both the experimental arrangement and the measured results are indicated in

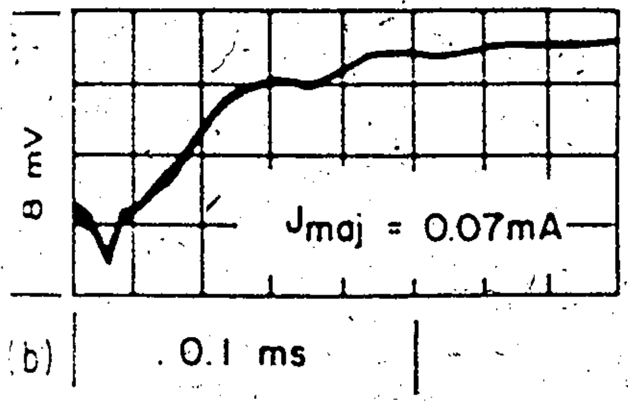
Figure 5.14: Experimentally observed response of two MIS tunnel diodes, P0515 and N0515, to a flash of light from a stroboscope superimposed upon the bias light. The diode parameters are given in Table 5.1.

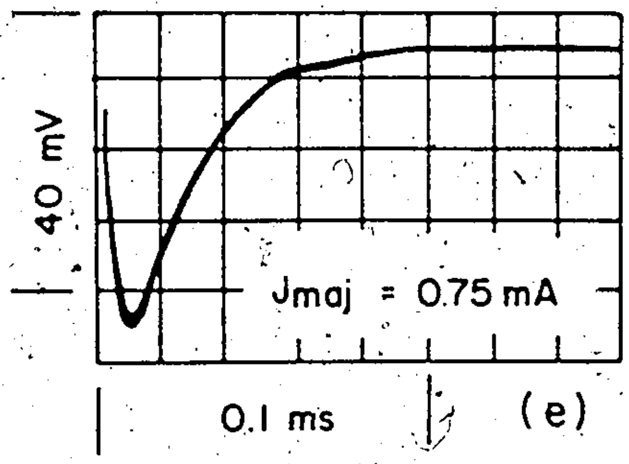
(a): response of the minority carrier diode, P0515 (independent of bias light intensity).

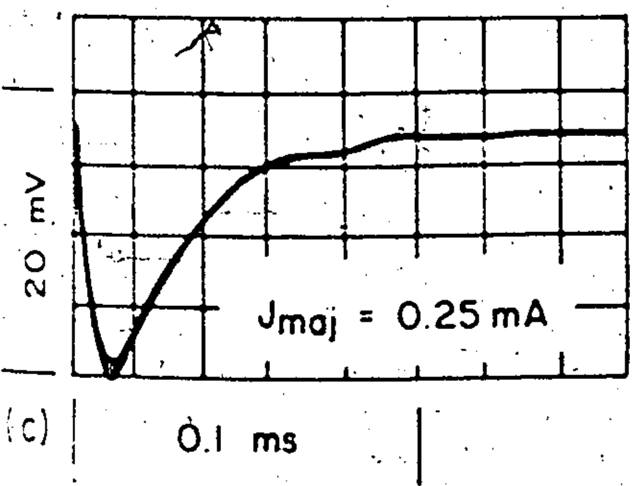
(b)-(f): response of the majority carrier diode, N0515, voltage biased at 2.5V as the intensity of the bias light is increased.











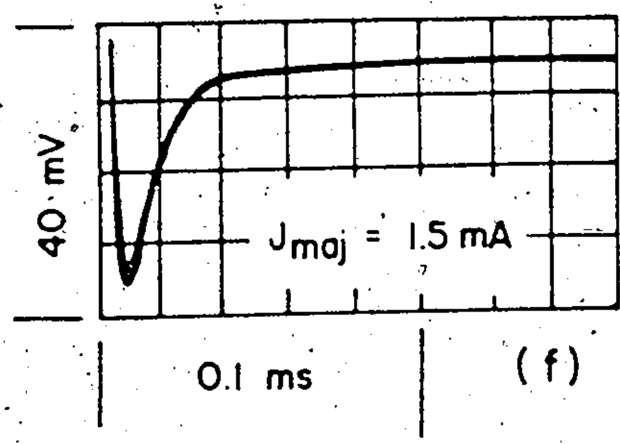


FIGURE 5.14

TABLE 5.2: Frequency Response and Current Gain of Device N0515

J_{maj} (mA)	DC Gain*	Cut-off Frequency (kHz)
0.014	7	2.9
0.03	15	3.2
0.07	27	4.0
0.25	58	5.1
0.50	122	5.9
0.75	148	6.9
1.0	92	8.0
1.5	68	9.9

* Ratio of area under transient response curve of N0515 to area under response curve of P0515.

Figure 5.15: Comparison of the low-frequency values of the small signal multiplication factor ,M', determined experimentally by two different methods for device N0515. One method was to evaluate dJ_{max}/dJ_{min} (incremental method), while the other method was to compare the areas under the impulse response curves (Figure 5.14) of N0515 and P0515 (impulse method).

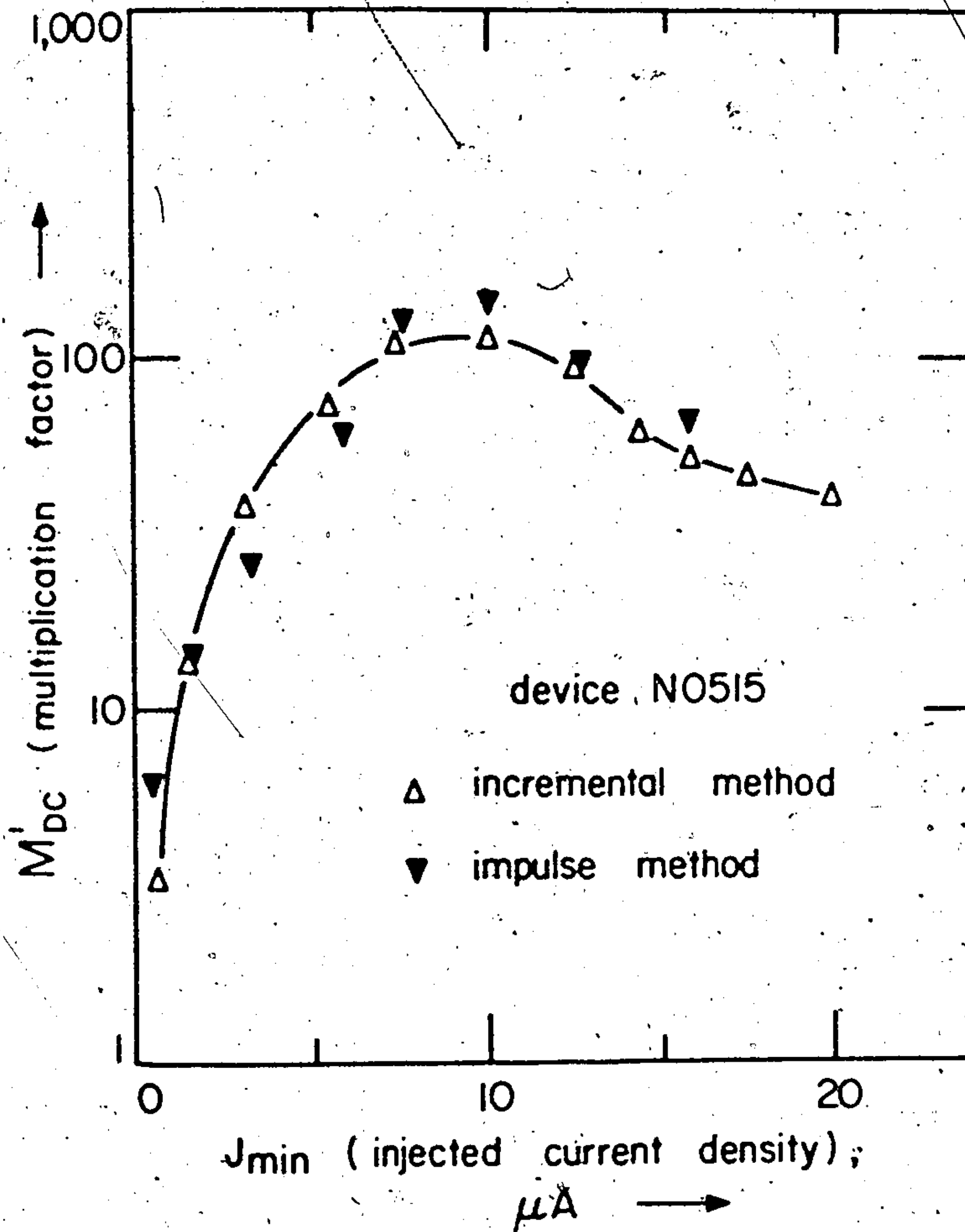


FIGURE 5.15

Figure 5.16. From Figure 5.16(b) and (c), it can be seen that the gain of the device is small until a certain threshold value of the ring current is attained (~50 mA) when it suddenly increases. This sudden increase in the gain of the structure is attributed primarily to the sudden increase in the multiplication factor, M, as observed earlier.

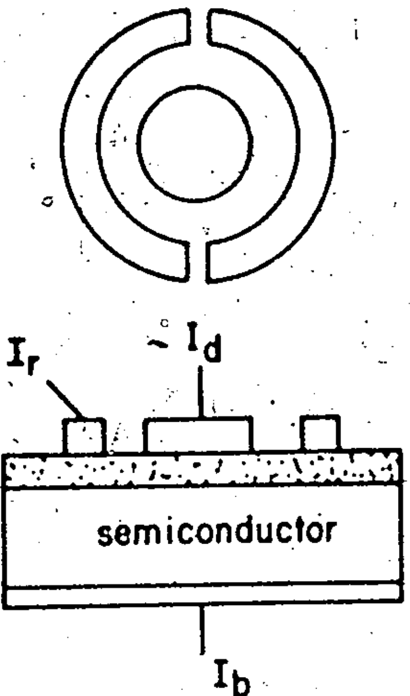
5.4 Applications of the Multiplication Effect

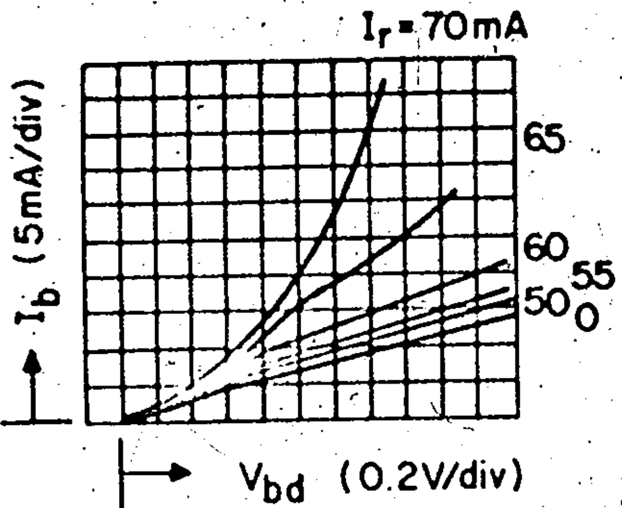
One application which has been investigated experimentally by Shewchun and Clarke [28] is to a new transistor structure called the Surface Oxide Transistor (SOT) and illustrated in Figure 5.17. The geometry is similar to that used in the measurement situation shown in Figure 5.16 except that all dimensions are optimized to give considerably larger device gains [28]. The transistor characteristics of the device were similar to those of Figure 5.16(b) exhibiting large gains after a threshold value of injected current was attained. The devices were also fragile in the sense that care had to be taken to ensure that the contact to substrate bias levels never exceeded a few volts. A more rugged structure was obtained by intentionally "forming" the oxide by subjecting it to voltage stress [28]. This reduced the gain of the device but shifted its operating characteristics to higher current levels and gave generally improved transistor characteristics. On the basis of this work, it is also expected that it would enhance the device frequency response. It is possible that the use of a different metal for the contacts will give more desirable multiplication properties. Even so, transistor structures employing the multiplication properties of unformed contacts will have the disadvantages of not being particularly rugged

Figure 5.16: (a) Schematic of the 'ring-dot' arrangement used to study the multiplication properties of majority carrier diodes.

(b) Transistor characteristics observed from the structure of (a). The device current, I_b, is shown as a function of the diode bias, V_{bd}, with the injected ring current, I_r, as the variable parameter.

(c) Small signal gain of the device at a bias of 1.6V as calculated from (b).





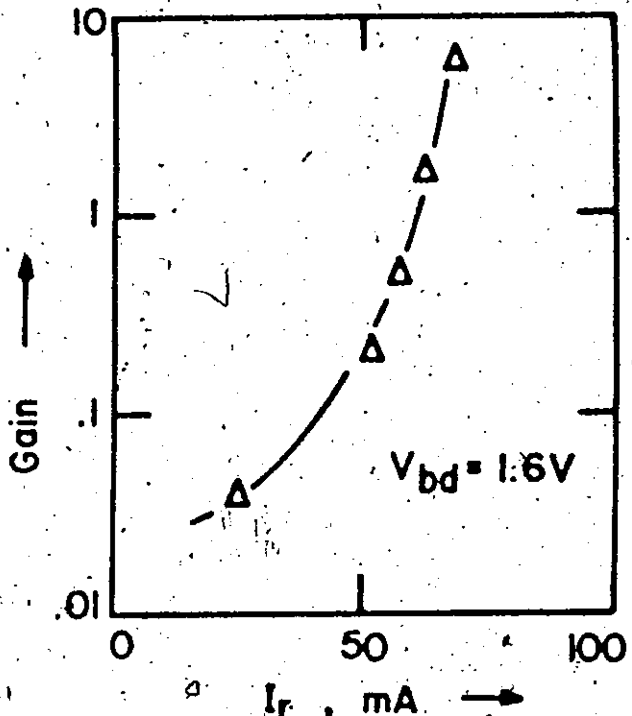


FIGURE 5.16

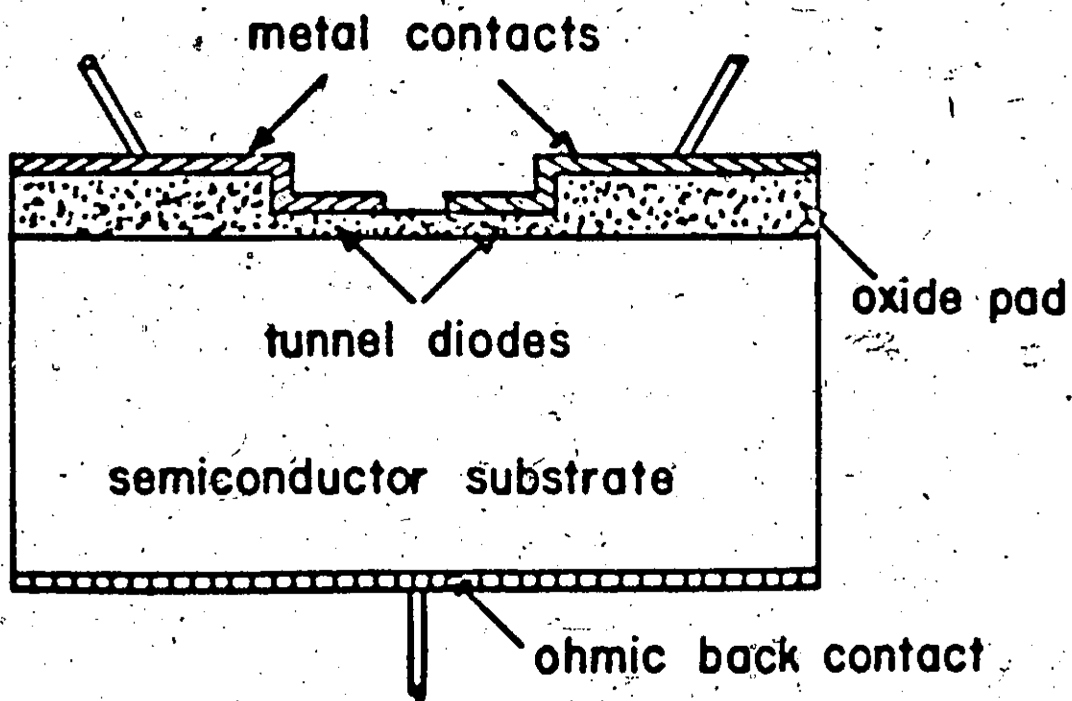


Figure 5.17: Schematic of the Surface Oxide Transistor (SOT).

and of possessing a relatively low frequency response.

Another possible application of the effect is to a solid-state photodiode with internal multiplication properties. The diode configuration would be similar to that of the thin metal contact (< 200Å) devices of the previous section. The addition of an anti-reflection coating [88] would ensure the maximum coupling of photons into the semiconductor region under the metal contact. The Al devices fabricated in this thesis would have the disadvantage that the large gain region is restricted to a relatively small range of light intensities. It is possible that the use of a different metal contact will improve this situation. The most successful solid-state photodiode with multiplication properties is the avalanche photodiode [le]. Compared to this device, the proposed device has a pronounced disadvantage in its restricted frequency response but has advantages in fabrication simplicity and in the fact that the multiplication occurs at small values of diode bias and is virtually independent of bias past this point.

5.5 Summary

This chapter describes the first quantitative theoretical and experimental analysis of the recently reported [15,28] multiplication properties of the MIS tunnel diode. Analytical expressions have been derived for a class of MIS tunnel diodes to describe the multiplication process and its dependence upon metal, insulator, and semiconductor parameters. In addition, results obtained by numerically solving the transport equations for the complete MIS system have been used to investigate the process under more general conditions.

The theoretical discussion centred about the metal-SiO_2-n type Si system. Theoretically, the most important parameters in determining the multiplication properties of this system are the metal work function, the oxide thickness, and the density of surface states associated with the IS interface. The optimal conditions for obtaining a high multiplication factor, M, are low surface state densities, small metal work functions, and a high resistivity <100> substrate. Surface states are expected to degrade the performance of the device at low current levels as are metal contacts with values of ϕ_{mi} greater than ϕ_{si}.

Experimental measurements were obtained using Al-SiO_2-n type Si majority carrier devices. The experimental technique used the photovoltaic properties of the minority carrier devices described in the following chapter. Multiplication factors, M, in the range 100 to 1000 were obtained but the gain was restricted to high current levels. The frequency response of the multiplication process has also been measured and has been found to be relatively poor as predicted theoretically. Physically, this is due to the fact that the charging of the oxide capacitance associated with the gain mechanism must be accomplished by the relatively small minority carrier currents.

Two possible applications of the multiplication process were discussed. The major factor limiting possible device application of the effect is felt to be its relatively poor frequency response. Since the value of M observed depends directly upon the ratio of the derivatives of the currents tunneling to the majority and minority carrier bands with additional structure at current densities determined by the value

of the surface state tunneling current, measurements of M contain a great deal of quantitative information about the tunneling process in a particular MIS device. One possible application suggested by this work is to a quantitative investigation of these tunneling processes.

CHAPTER 6

MINORITY CARRIER DIODES AND THEIR APPLICATION TO DIRECT ENERGY CONVERSION

6.1 Introduction

In the previous chapter, it has been shown that majority carrier MIS tunnel diodes exhibit current multiplication properties. Minority carrier tunnel diodes also have been found to possess properties which make them of interest for device application. In the non-equilibrium mode, the minority carrier MIS tunnel diode displays properties similar to p-n junction diodes under reverse and small forward bias to the semiconductor. In addition to exhibiting I-V characteristics which approach the "ideal diode" characteristics of classical junction diode theory [91], minority carrier diodes have similar C-V characteristics and electron- and photo-voltaic energy conversion properties.

In Section 6.2, the theory of Chapter 3 is developed further for the minority carrier diode with emphasis again placed upon the M-SiO_2-Si system. In addition to predicting terminal properties for these devices similar to p-n junctions, the theory shows that the inversion layer present under zero bias conditions remains clamped at the IS interface under reverse and small forward bias. Experimental support for this effect has been described in Section 4.6. In Section 6.3, experimental measurements complementing those of Chapter 4 are used to demonstrate aspects of the extended theory. Data is presented to show the behaviour of minority

carrier diodes as a function of the metal work function, insulator thickness, and the substrate resistivity and orientation.

Due to their fundamental simplicity and the consequent minimization of any degradation of the bulk semiconductor properties during processing, the devices are particularly suited to applications as direct energy conversion diodes employing the electron- or photo-voltaic effects. In Section 6.4, this application of minority carrier devices is treated. It is demonstrated experimentally that higher conversion efficiency is possible than with conventional diffused p-n junction devices. Finally in Section 6.5, the minority carrier devices are compared in a qualitative fashion with two other diodes, p-n junction diodes and Schottky diodes.

6.2 Theory

6.2.1 I-V Characteristics

The theory of the MIS tunnel diode has been described in Chapter 3. Portions relating to the operation of minority carrier devices are summarized below. Minority carrier diodes are formed on p type substrates by selecting metals for the top contact which have low values of the metal to insulator work function, ϕ_{mi}. Conversely, on n type substrates, metals with high values of ϕ_{mi} are required. These statements are well illustrated by Figure 3.15. Minority carrier diodes tend to have an inversion layer at the IS interface at zero bias. This favours the minority carrier band tunnel current over the majority carrier band and surface state tunnel currents. Changing the insulator thickness from high to low values produces the changes in the I-V characteristics shown in Figure 6.1 for the p type substrates which will be considered in this chapter. At high

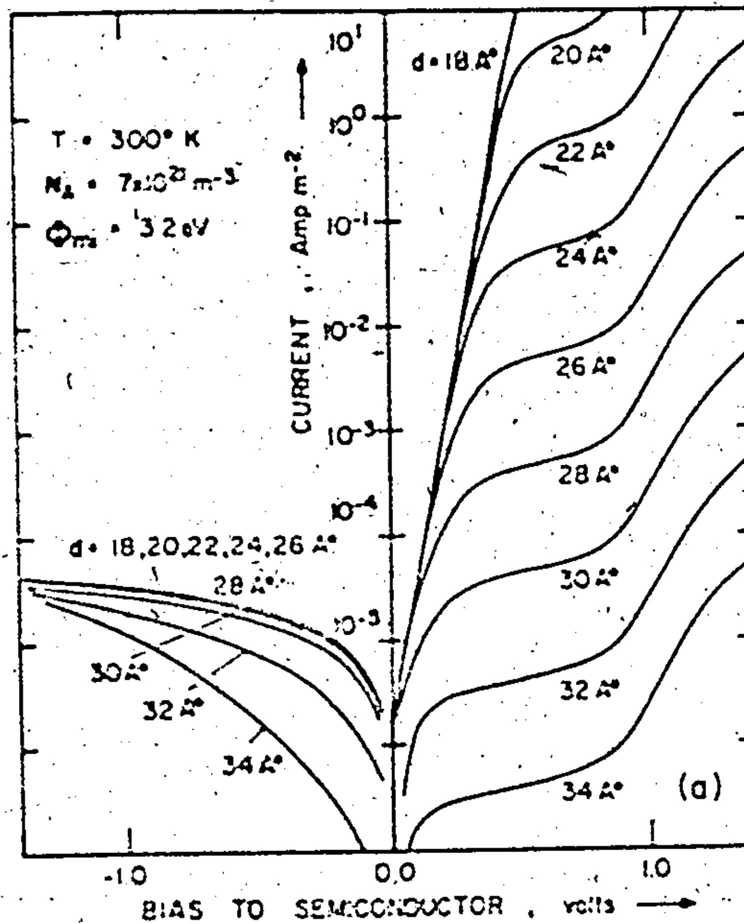


Figure 6.1: Computed I-V chacteristics of the MIS tunnel diode demonstrating 'non-equilibrium' effects for minority carrier devices. The substrate is 2Ωcm p type <100> silicon, ϕ_{mi} is equal to 3.2 eV, and the variable parameter is d, the thickness of the oxide layer. Other device parameters are given in Table 2.4.

forward bias (> 0.3 V), the current flow increases rapidly with decreasing insulator thickness. In this region, the diode current is limited by the rate at which particles can tunnel between the metal and the semiconductor, which varies approximately exponentially with the insulator thickness. However, at reverse and small forward bias there is a distinct upper limit to the current flow at a given bias point regardless of how thin the insulator is made. This difference arises because, under reverse and small forward bias, the diode current is limited by transport through the semiconductor rather than by tunneling through the insulator.

As previously mentioned and as indicated in Figure 6.2, the minority carrier diodes under discussion have the IS interface inverted near zero bias. This inversion region is followed by a depletion region and finally a space charge neutral region moving from the IS interface to the bulk semiconductor. Minority carrier current flowing in the semiconductor is the sum of the diffusion current in the space charge neutral region of the device and the generation-recombination current associated with the depletion region.

In the following discussion, an expression for the limiting value approached at reverse and small forward bias for these minority carrier devices will be obtained. For simplicity, surface states will be assumed to be negligible until Section 6.2.3. The net minority carrier current tunneling through the insulator (J_{CT}) can be expressed in terms of its components, J_{CM} and J_{MC}, the currents tunneling from the conduction band to the metal and from the metal to the conduction band respectively, as follows (Chapter 2).

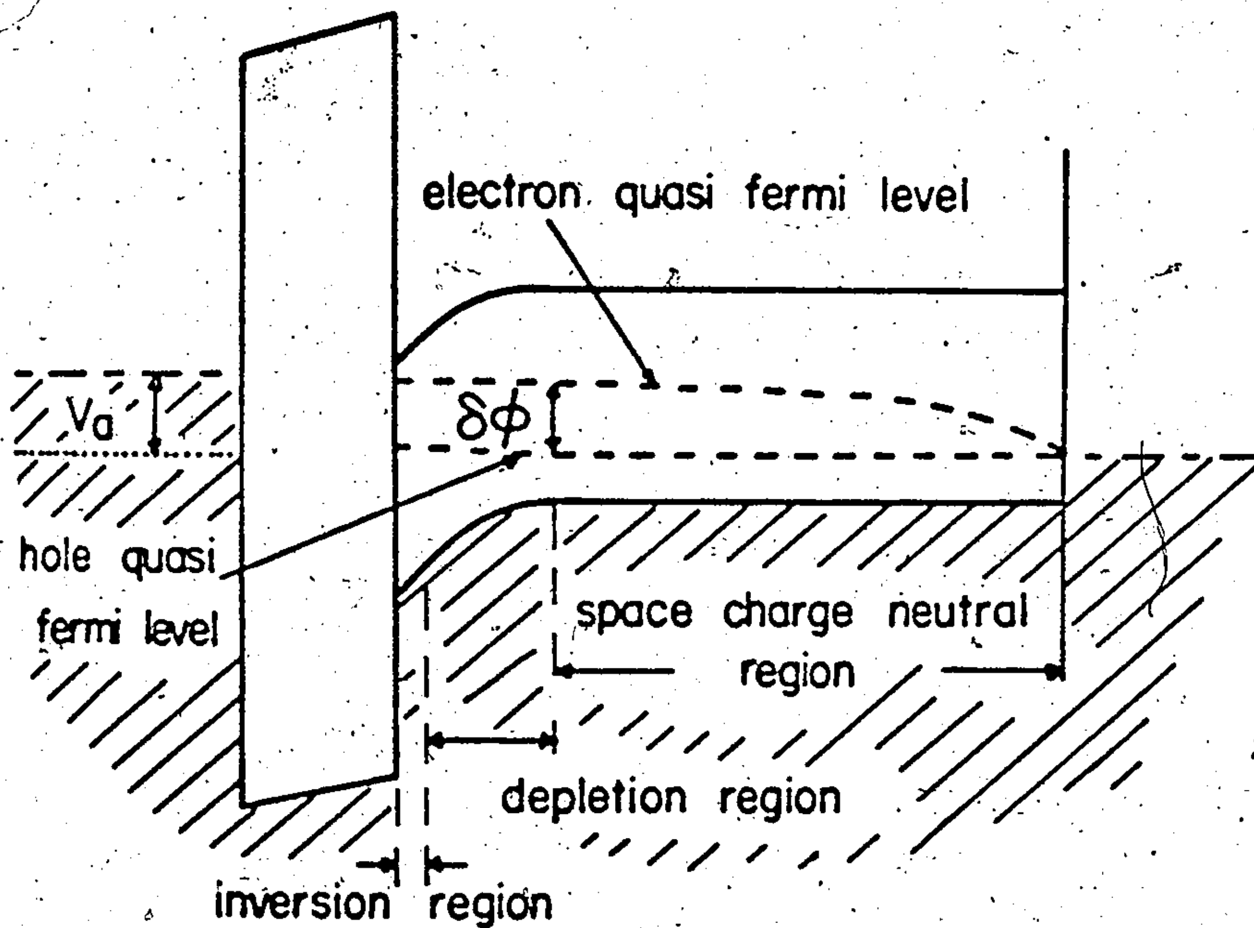


Figure 6.2: Schematic energy band diagram of the minority carrier MIS tunnel diode under "non-equilibrium" conditions. Shown in the diagram are the electron and hole quasi fermi levels in the semiconductor region.

$$J_{CT} = J_{CM} - J_{MC} \tag{6.1}$$

where

$$J_{CM} = \frac{q}{2\pi^2 h} \int_{\substack{\text{conduction}\\\text{band}}} dEf_m \int dSe^{-n} \tag{6.2}$$

$$J_{MC} = \frac{q}{2\pi^2 h} \int_{\substack{\text{conduction}\\\text{band}}} dEf_s \int dSe^{-n} \tag{6.3}$$

Employing the conventional assumptions of junction diode theory (i.e. that carrier quasi-fermi levels do not vary significantly across the depletion region and that low injection conditions apply) gives the following expression for the diffusion current flow [92]

$$J_d = \frac{qD_n n_i^2}{L_n N_A} \coth\left(\frac{L}{L_n}\right)[\exp(q\phi/kT) - 1] \tag{6.4}$$

where ϕ is the difference between the electron and hole quasi-fermi levels in the semiconductor depletion region (Figure 6.2), D_n and L_n are the electron diffusion coefficient and diffusion length in the p type semiconductor doped with N_A acceptors per unit volume.

Applying current conservation conditions at the semiconductor-insulator interface gives

$$J_{CM} - J_{MC} = J_d + J_{rg} \tag{6.5}$$

where J_{rg} refers to the recombination-generation current associated with the semiconductor region (e.g. see Equation (6.12)).

As the oxide becomes thin, both J_{CM} and J_{MC} increase rapidly (approximately exponentially) and eventually each becomes much larger than the terms on the right hand side of Equation (6.5), at least at reverse and small forward bias. In this bias region, the difference between J_{CM} and J_{MC} is much smaller than either J_{CM} or J_{MC}, i.e.

$$J_{CM} \approx J_{MC} \tag{6.6}$$

One consequence of Equations (6.2), (6.3) and (6.6) is

$$f_m \approx f_s \tag{6.7}$$

where f_m and f_s are the energy distributions of electrons in the metal and the semiconductor conduction band respectively. Under thermal equilibrium conditions they are identical and are given by the Fermi-Dirac distribution function

$$f_m = f_s = \{1 + \exp[(E - E_F)/kT]\}^{-1} \tag{6.8}$$

where E_F is the system fermi level. At least for moderate departures from thermal equilibrium, the electron energy distributions can be described in terms of this distribution law by using quasi-fermi levels. For the metal, E_F is replaced by E_{Fm}, the metal fermi level, while for the semiconductor, E_F in replaced by E_{Fsn}, the electron quasi-fermi level. Therefore, Equation (6.7) is equivalent to

$$E_{Fm} = E_{Fsn} \quad (6.9)$$

showing that, for thin insulators, the semiconductor electron quasi-fermi level is effectively pinned to the metal fermi level under reverse and small forward bias.[9]

Since the majority carrier flow (J_{VT}) is small, there is little variation of the hole quasi-fermi level across the bulk regions of the semiconductor (Figure 6.2). Thus in Equation (6.4), $\Delta\phi$ can be replaced by V_a, the applied voltage, to give

$$J_{df} = \frac{qD_n n_i^2}{L_n N_A} \coth\left(\frac{L}{L_n}\right)\left[\exp\left(\frac{qV_a}{kT}\right) - 1\right] \quad (6.10)$$

The diode current-voltage characteristics at reverse and small forward bias can then be described as

$$J_{CT} = J_{df} + J_{rg} \quad (6.11)$$

where [1f]

$$J_{rg} = \frac{qn_i W}{\tau_e}\left[\exp\left(\frac{qV_a}{2kT}\right) - 1\right] \quad (6.12)$$

τ_e is an effective lifetime parameter and W is the width of the semiconductor depletion region. Equation (6.11) is identical with the formula for the characteristics of one sided abrupt p-n junction diode.

As the diode bias is increased in the forward direction, the diode current given by Equation (6.11) increases rapidly. Eventually the difference between J_{CM} and J_{MC} (Equation (6.5)) is no longer small compared

to either J_{CM} or J_{MC}. The fermi levels become unpinned and the current flow becomes tunnel limited as opposed to the semiconductor limited situation described above. Schematic energy band diagrams demonstrating the features of the above discussion are shown in Figure 6.3 for bias conditions ranging from large reverse bias to large forward bias. The electron quasi-fermi level, the minority carrier fermi level for the p type device illustrated, remains pinned to the metal fermi level at reverse and small forward bias, eventually becoming unpinned at moderate forward bias. The relative magnitudes of the current components J_{CM} and J_{MC} are also indicated.

To this stage, it has been assumed that, if the current tunneling to the majority carrier band, J_{VT}, is less than J_{CT} in the thicker insulator "equilibrium" case, J_{VT} will remain less than J_{CT} as the insulator thickness is reduced. However, even when J_{CT} is semiconductor limited, J_{VT} remains tunnel limited at small applied bias since it is a majority carrier flow which can easily pass through the semiconductor. It therefore increases with respect to J_{CT} when the insulator thickness is reduced through the semiconductor limited regime. This is demonstrated in Figure 6.4 where the ratio J_{VT}/J_{CT}, at 0.3 V forward bias is plotted for the thickness range of interest. When J_{CT} is tunnel limited the ratio varies little with thickness. Once J_{CT} becomes semiconductor limited at this bias point, the ratio increase rapidly with decreasing thickness. Figure 6.4 shows that the desired semiconductor limited behaviour occurs in the thickness range of interest only if $\phi_{mi} < 3.4$ eV for the particular diode configuration under analysis. For such values

Figure 6.3: Schematic showing the changes in the energy band diagram for a non-equilibrium MIS tunnel diode as the applied bias is varied from large reverse bias to large forward bias.

(a) large reverse bias

(b) moderate reverse bias

(c) zero bias

(d) small forward bias

(e) moderate forward bias

(f) large forward bias

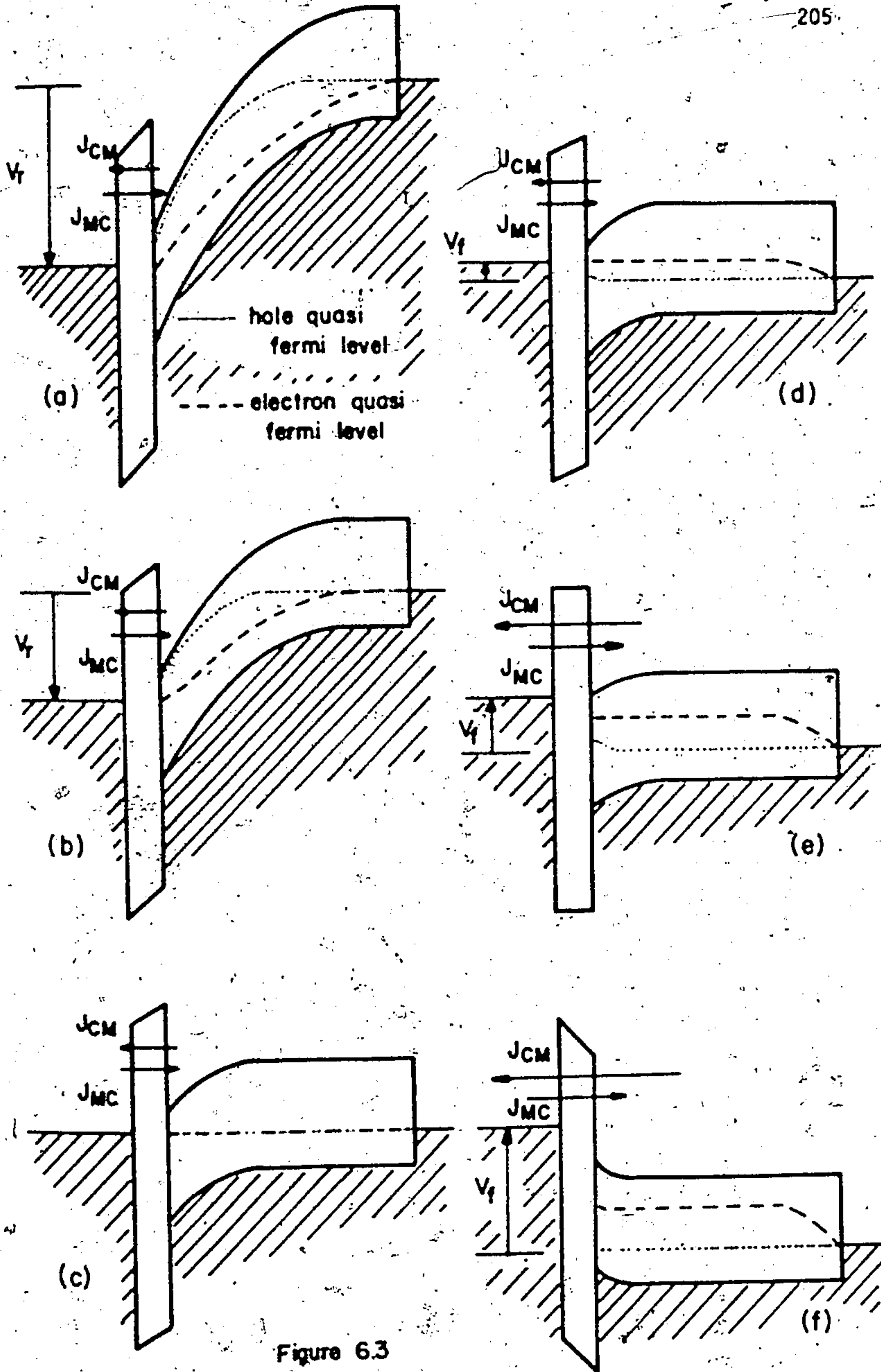


Figure 6.3

Figure 6.4: The computed value of the ratio of the band tunnel currents J_{VT} and J_{CT} versus the thickness of the insulating layer, d, for the monority carrier diode of Figure 6.1. The variable parameter is ϕ_{mI}, the metal to insulator barrier height.

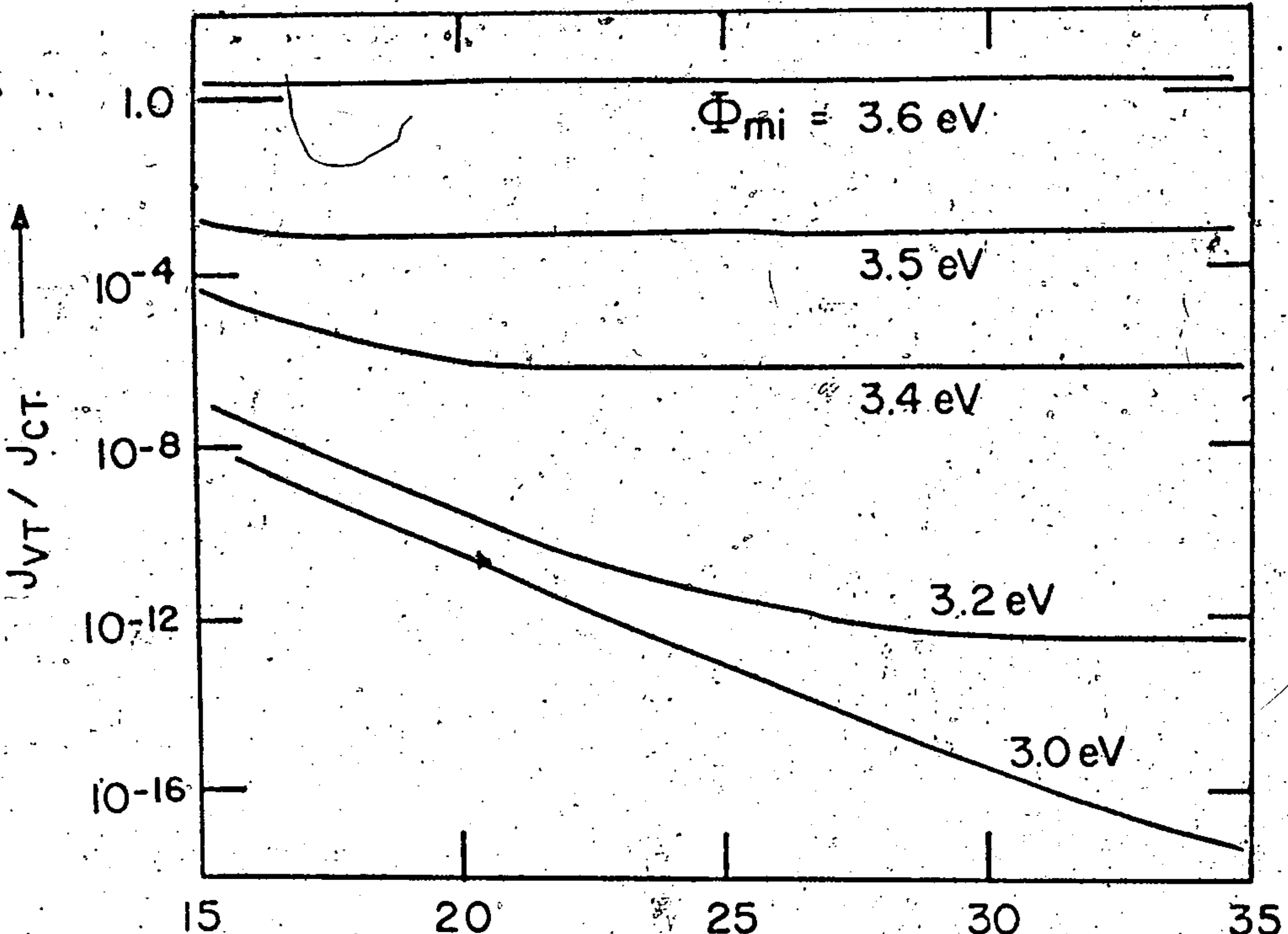


FIGURE 6.4

of ϕ_{mi}, the ratio of J_{VT} to J_{CT} is always less than 10^{-4}. This justifies the neglect of J_{VT} when computing the total diode current.

6.2.2 Clamping of the Inversion Layer

Another property of the diodes as the insulator layer becomes thin is illustrated in Figure 6.5 where the electron carrier concentration at the IS interface is plotted as a function of bias for several insulator thicknesses. In regions where the current is semiconductor limited, this concentration is essentially constant, increasing only slightly with increasing forward bias.

This behaviour can be explained in the following way. A voltage, V_a, applied to the semiconductor is absorbed as a change in the voltage drop across the insulator and a change in ψ_s, the semiconductor surface potential. If the insulator is thick and the pinning of the fermi level as described does not occur, the situation is identical to that of a conventional MIS capacitor. A positive voltage, V_a, would cause ψ_s to change by an amount less than V_a, reducing the electron concentration in the surface region which in turn reduces the electric field at the semiconductor surface since this field is strongly dependent upon the amount of inversion change (e.g. see Appendix D). V_a would therefore be absorbed partly as a reduction in the voltage across the insulator and partly as a change in surface potential (see Equation (2.22)).

However, if the electron quasi-fermi is pinned to the metal fermi level as has been established for thin insulator diodes, the electron quasi-fermi level would change V_a due to the application of this bias. If the electron concentration is to decrease, the potential of the

Figure 6.5: Calculated value of the minority carrier concentration at the IS interface for the diode of Figure 6.1. The variable parameter is d, the insulator thickness.

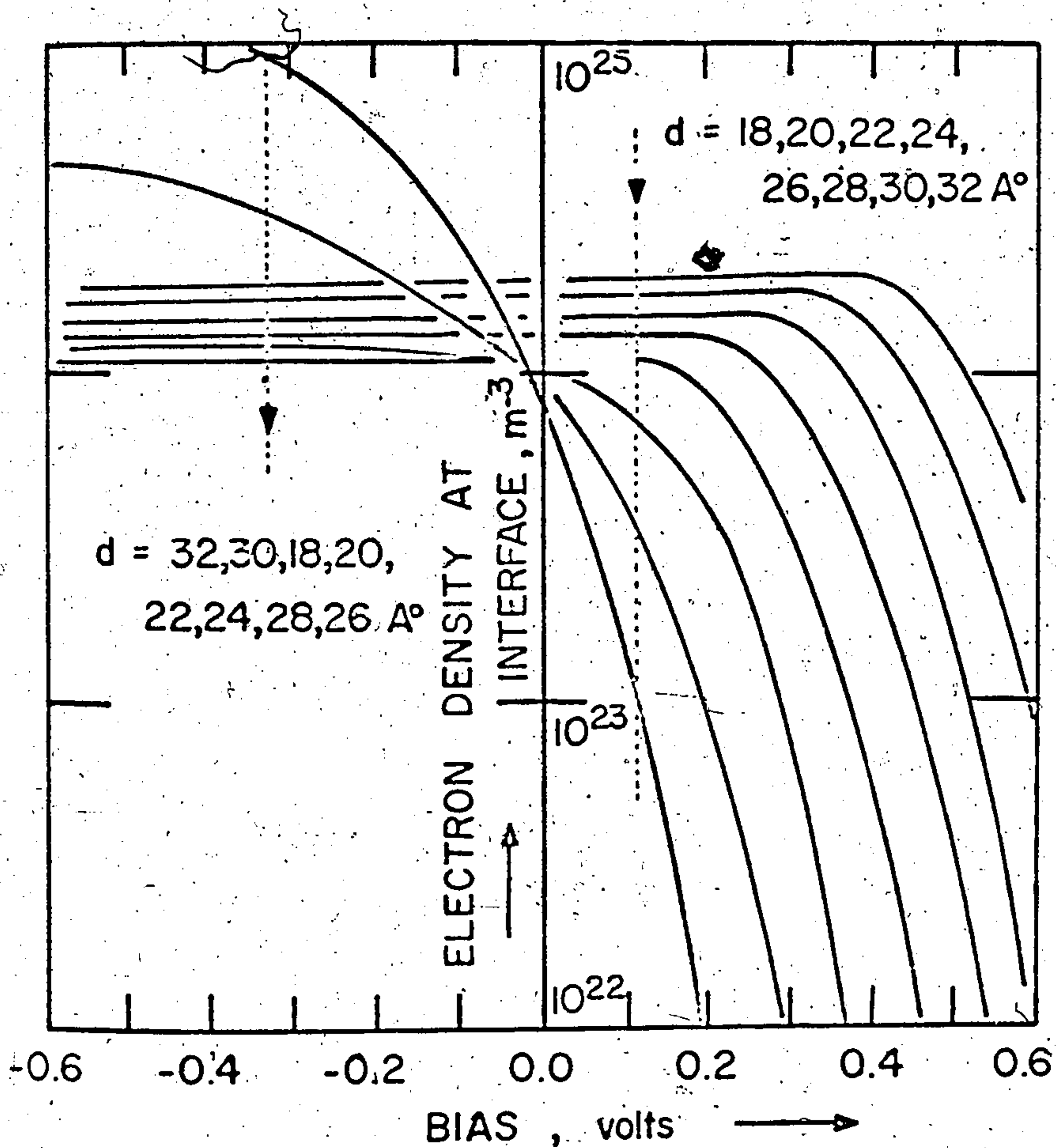


FIGURE 6.5

conduction band edge and consequently ψ_s must change by more than V_a since the electron concentration depends on the energy difference between the conduction band edge and the electron quasi-fermi level. A decrease in the electron concentration at the surface again implies a reduction in the surface field. Thus, some of V_a would be absorbed as a reduction in the voltage across the insulator. This is not possible since a voltage greater than V_a has already been absorbed in changing ψ_s. It is concluded that when the electron quasi-fermi level is pinned, the electron concentration in the surface region does not decrease with bias as in a conventional MIS capacitor. In fact, by continuing the above argument, it can be shown to increase slightly in agreement with Figure 6.5. It follows that the conduction band edge is also effectively pinned with respect to the metal fermi level in the semiconductor limited regime. Figure 6.3 illustrates the features of the above discussion.

To summarize, the following qualitative picture of the operation of minority carrier MIS tunnel diodes has been developed. These diodes tend to have an inversion layer at the semiconductor-insulator interface at thermal equilibrium (i.e. zero bias). When the insulator is so thin that current flow is semiconductor limited near zero bias, the electron quasi-fermi level in the semiconductor is pinned to the metal fermi level. The inversion layer is forced to remain at the surface for all biases where these levels are pinned. Further into the semiconductor is a depletion region. Since the inversion region does not alter significantly with voltage, changes in the applied voltage are absorbed mainly by variations in the width of this depletion region.

6.2.3 The Effect of Surface States

The effect of fixed charge in the insulator region of the device is virtually indistinguishable from a change in the metal to insulator barrier height, ϕ_{mi}. For the silicon system, a positive charge of density 10^{16} $metre^{-2}$ located near the IS interface causes an effective reduction in ϕ_{mi} of about 0.1 eV. For this reason, insulator charge assists in the formation of minority carrier MIS tunnel diodes using p type silicon and majority carrier diodes using n type silicon.

The effects of surface states are more complex. Surface states have the following properties:

(a) they are charge storage centres;

(b) they are recombination-generation centres;

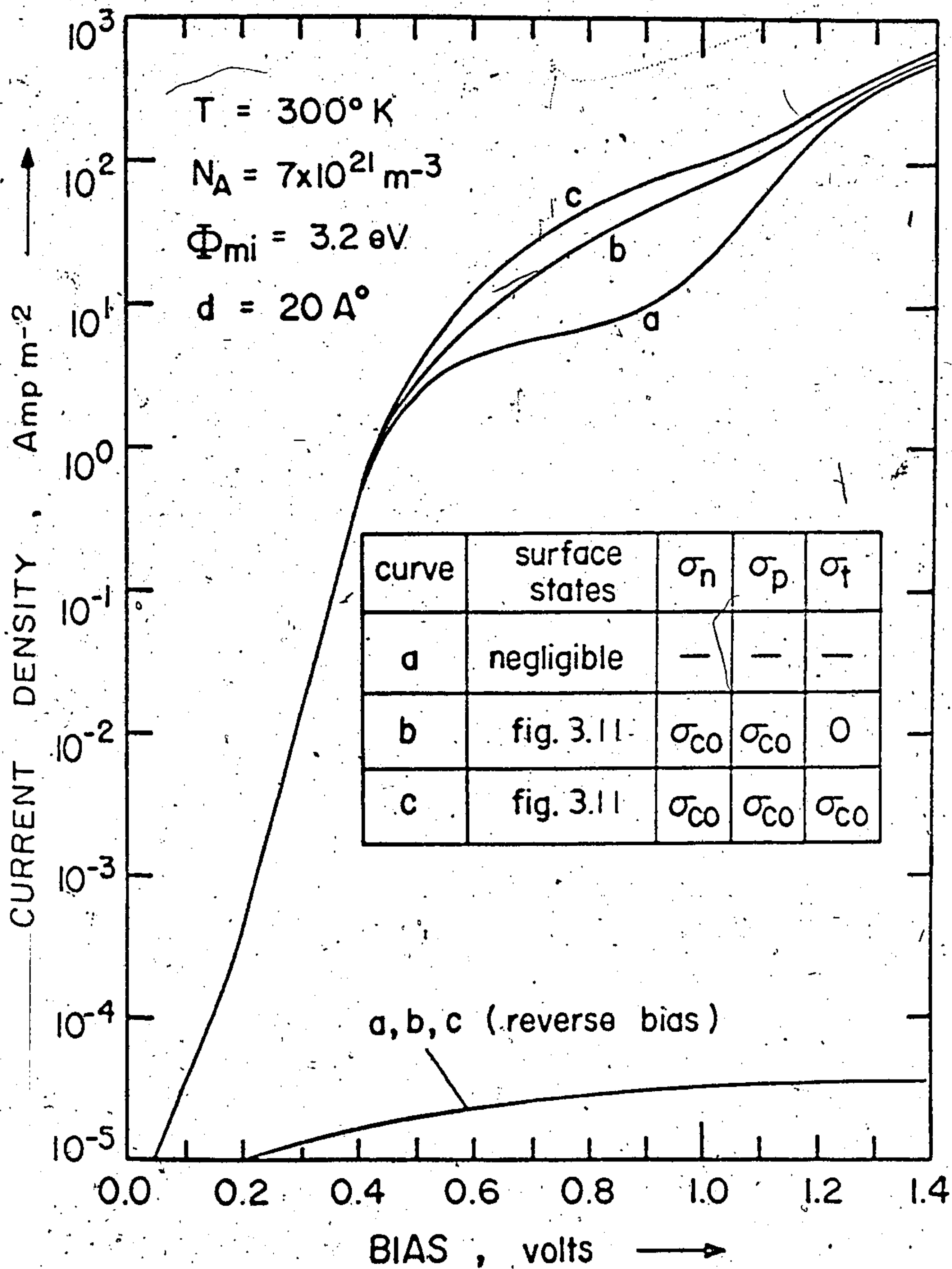
(c) they provide additional tunneling paths between the metal and the semiconductor.

All of these properties were included into the model used to characterize these states assumed to be located at the IS interface (Chapter 2). Their occupation and recombination-generation properties were characterized using SRH recombination statistics modified to include the effect of tunneling between the metal and these states. The latter process was formulated in terms of effective tunneling capture cross sections.

Current-voltage characteristics computed by the numerical method described in reference [103] for three surface state treatments are shown in Figure 6.6(a). Shown are the cases when surface states are negligible, when they are distributed across the silicon band gap as in Figure 3.11 and act as charge storage and recombination-generation centres

Figure 6.6: (a) The computed effect of surface states upon the I-V characteristics of a minority carrier MIS tunnel diode. The value of ϕ_{mi} is 3.2 eV, the silicon dioxide layer is 20Å thick, and the silicon substrate is p type of <100> orientation and resistivity 2Ωcm ($N_A=7\times10^{21}m^{-3}$). Shown are three cases; when surface states are neglected, when they have distribution of Figure 3.11 but act only as charge storage and recombination centres, and when there is also appreciable current tunneling between these states and the metal.

(b) The computed effect of surface states upon the minority carrier injection ratio, γ, at a point 3μm from the IS interface. Shown are the three cases of Figure 6.6 (a) labelled in the manner indicated in this figure.



T = 300° K
N_A = 7x10^21 m^-3
Φmi = 3.2 eV
d = 20 A°
c
b
a
curve
surface states
σn
σp
σt
a
negligible
b
fig. 3.11
σco
0
c
fig. 3.11
σco
a, b, c (reverse bias)
CURRENT DENSITY, Amp m^-2
BIAS, volts
10^3
10^2
10^1
10^0
10^-1
10^-2
10^-3
10^-4
10^-5
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4

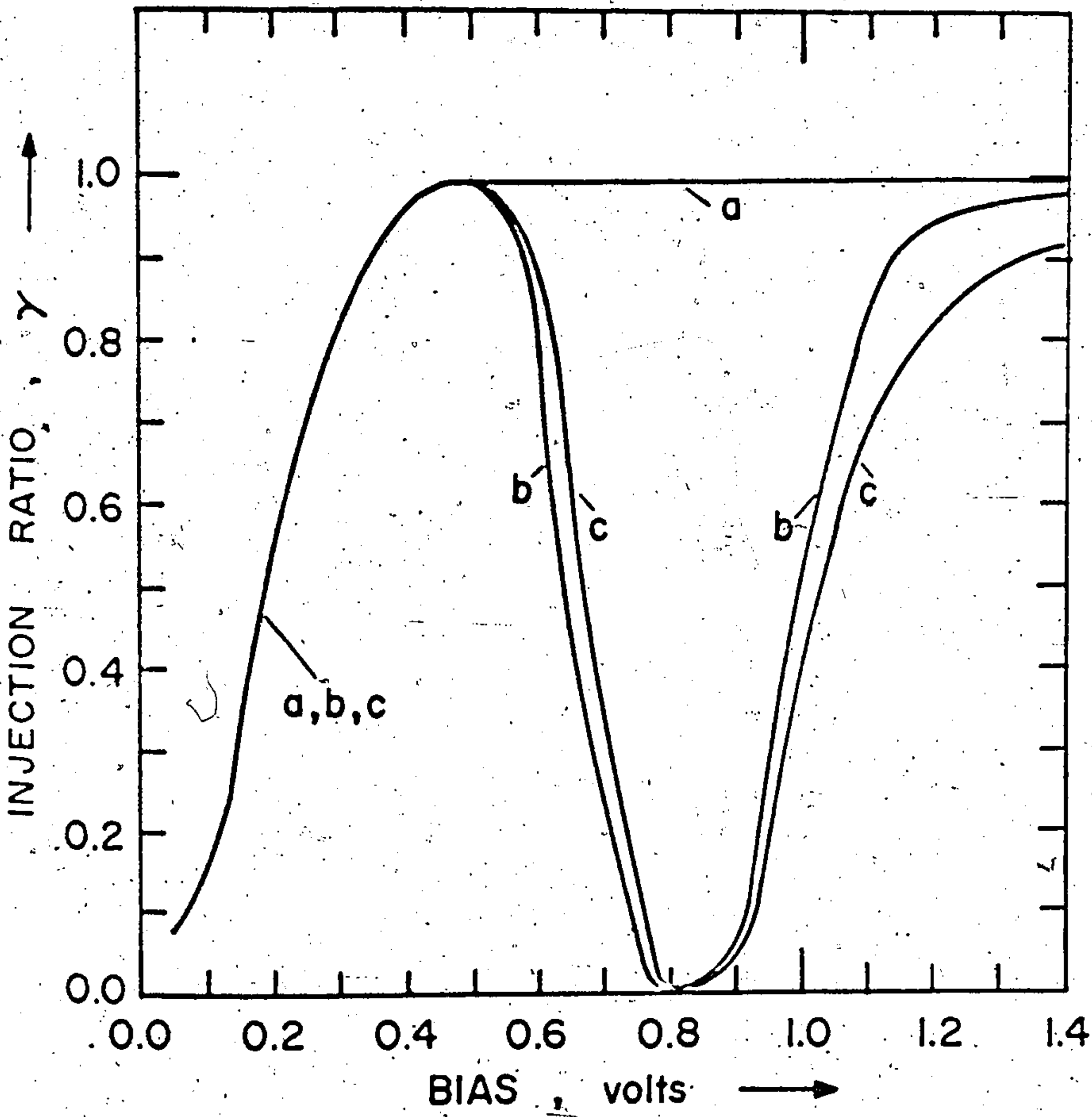
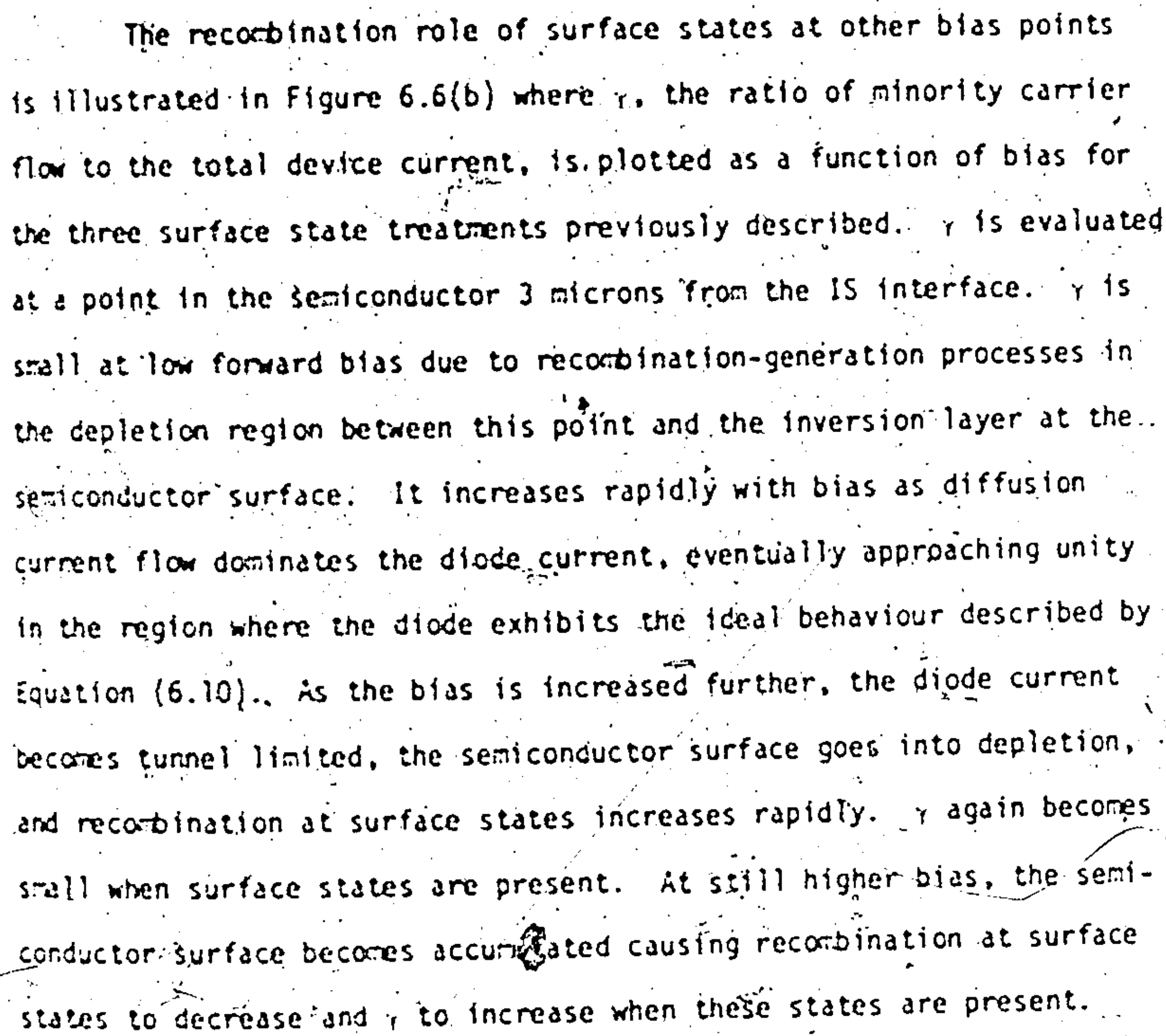


FIGURE 6.6(b)

only, and when surface states have the preceding properties and also provide additional tunneling paths between the metal and the semiconductor. The current flow at reverse and small forward bias is unaffected by the presence of surface states, while it is significantly altered in the tunnel limited bias range.

The occupation of the surface states depends strongly on the carrier concentration in the surface region. For minority carrier MIS diodes, this region is inverted when the current flow is semiconductor limited. For p type diodes, it follows that the surface states are almost entirely occupied by electrons. Thus, when the current is semiconductor limited, surface states contribute their smallest possible positive charge and this contribution is virtually bias independent. When the current becomes tunnel limited, the occupation of the surface states changes rapidly with increasing bias as the inversion layer disappears. The increasing positive charge in surface states determines the slope of the current-voltage characteristics between 0.6 to 1.2 volts.

Recombination-generation rates at surface states are greatest when the surface is in depletion and are considerably suppressed when the surface is accumulated or inverted since in these conditions the states are either too heavily or too lightly occupied by electrons for efficient recombination-generation. Since the surface is clamped in inversion in the semiconductor limited regimes, surface states do not contribute significantly to the recombination-generation current flow when reverse and small forward bias is applied to the minority carrier diodes under discussion.



The recombination role of surface states at other bias points is illustrated in Figure 6.6(b) where γ, the ratio of minority carrier flow to the total device current, is plotted as a function of bias for the three surface state treatments previously described. γ is evaluated at a point in the semiconductor 3 microns from the IS interface. γ is small at low forward bias due to recombination-generation processes in the depletion region between this point and the inversion layer at the semiconductor surface. It increases rapidly with bias as diffusion current flow dominates the diode current, eventually approaching unity in the region where the diode exhibits the ideal behaviour described by Equation (6.10). As the bias is increased further, the diode current becomes tunnel limited, the semiconductor surface goes into depletion, and recombination at surface states increases rapidly. γ again becomes small when surface states are present. At still higher bias, the semiconductor surface becomes accumulated causing recombination at surface states to decrease and γ to increase when these states are present.

The differences between the curves of Figure 6.6(a) occur in the tunnel limited regime. Curves "a" and "b" differ due to the charge contribution from surface states in "b". Curves "b" and "c" differ because of two effects. The first is due to the change in occupancy of the surface states with bias caused by the current tunneling between these states and the metal (Equation (2.31)). There is also a smaller contribution due to the direct addition of the surface state tunnel current (J_{ST}) to the total device current in this region.

Thus, in summary, surface state effects for minority carrier

diodes are confined to the bias regions where the current is tunnel limited. In the semiconductor limited regimes, surface states have little effect upon the device characteristics and act primarily as a fixed charge at the IS interface.

6.3 Experiment

6.3.1 Minority and Majority Carrier Diodes

The differences between minority and majority carrier diodes have been demonstrated in several sections of Chapter 4. Gold has been found to have a value of ϕ_{mi} equal to approximately 4.2 eV and is therefore expected to produce majority carrier MIS tunnel diodes using p type substrates, whereas aluminum with a value of 3.2 eV is expected to produce minority carrier diodes. In reverse and small forward bias range of interest in the current chapter, the differences are well illustrated by Figure 6.7(a) which shows the I-V characteristics of diodes with gold and aluminum top contacts compared on a linear scale. Both diodes were fabricated in an identical fashion using a p type <111> oriented substrate of resistivity 5-10 Ωcm with an oxide thickness of 18Å. The general behaviour is as described by the theory of Chapter 3. The gold diode exhibits majority carrier behaviour whereas the aluminum diode displays minority carrier behaviour. For all bias voltages, the gold diode allows current flows that are several orders of magnitude greater than the largest possible value of minority carrier flow that can pass through a semiconductor of this particular resistivity. The saturation of the gold diode current at reverse bias (semiconductor negative with respect to the top metal contact) is a result of the non-equilibrium

Figure 6.7: Experimental comparison of the characteristics of diodes of similar oxide thickness but with gold and aluminum top contacts. The substrate in each case was p type silicon of <111> orientation with resistivity in the range of 5-10 Ωcm.

(a) comparison of the I-V characteristics.

(b) comparison of the C-V characteristics.

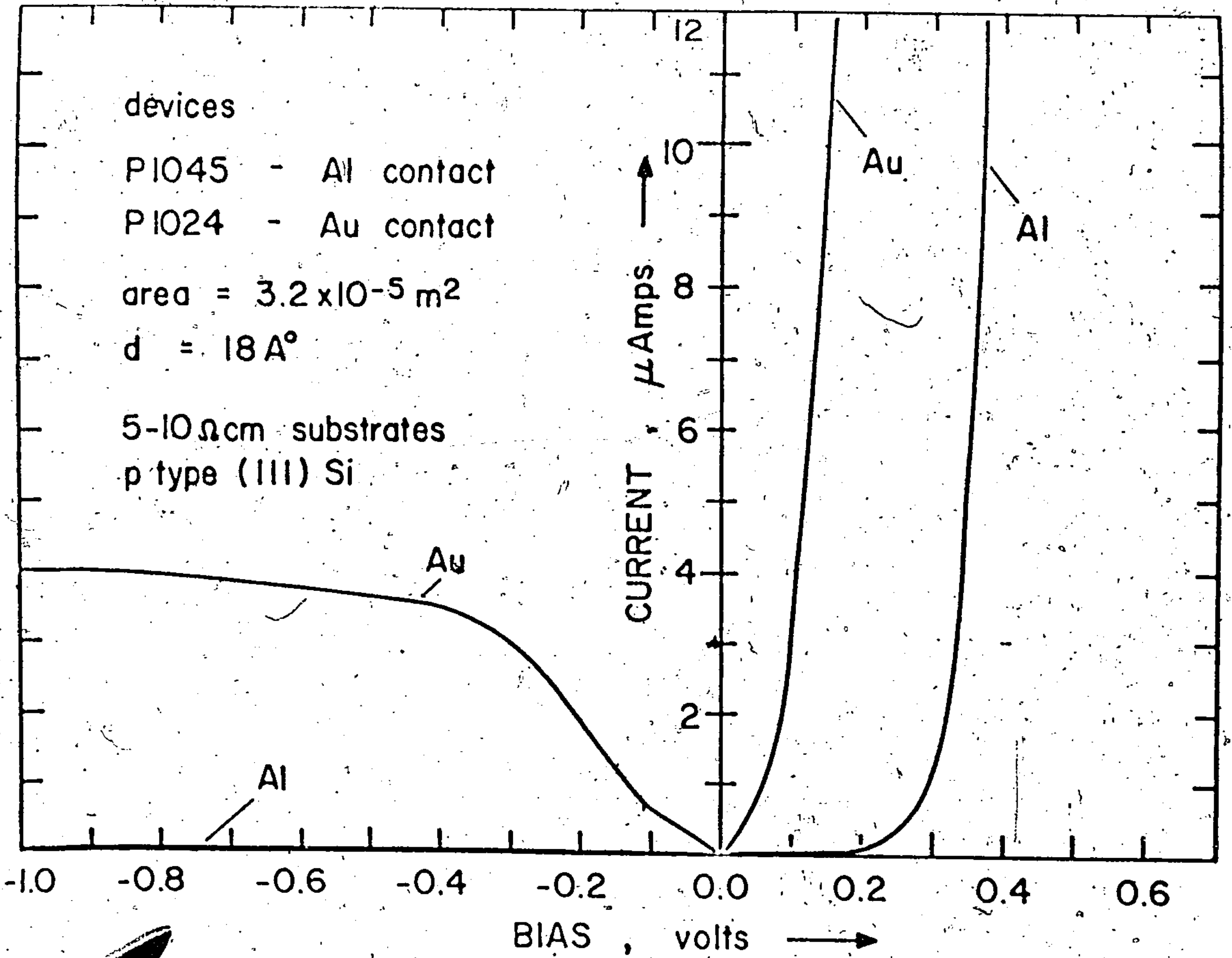


FIGURE 6.7(a)

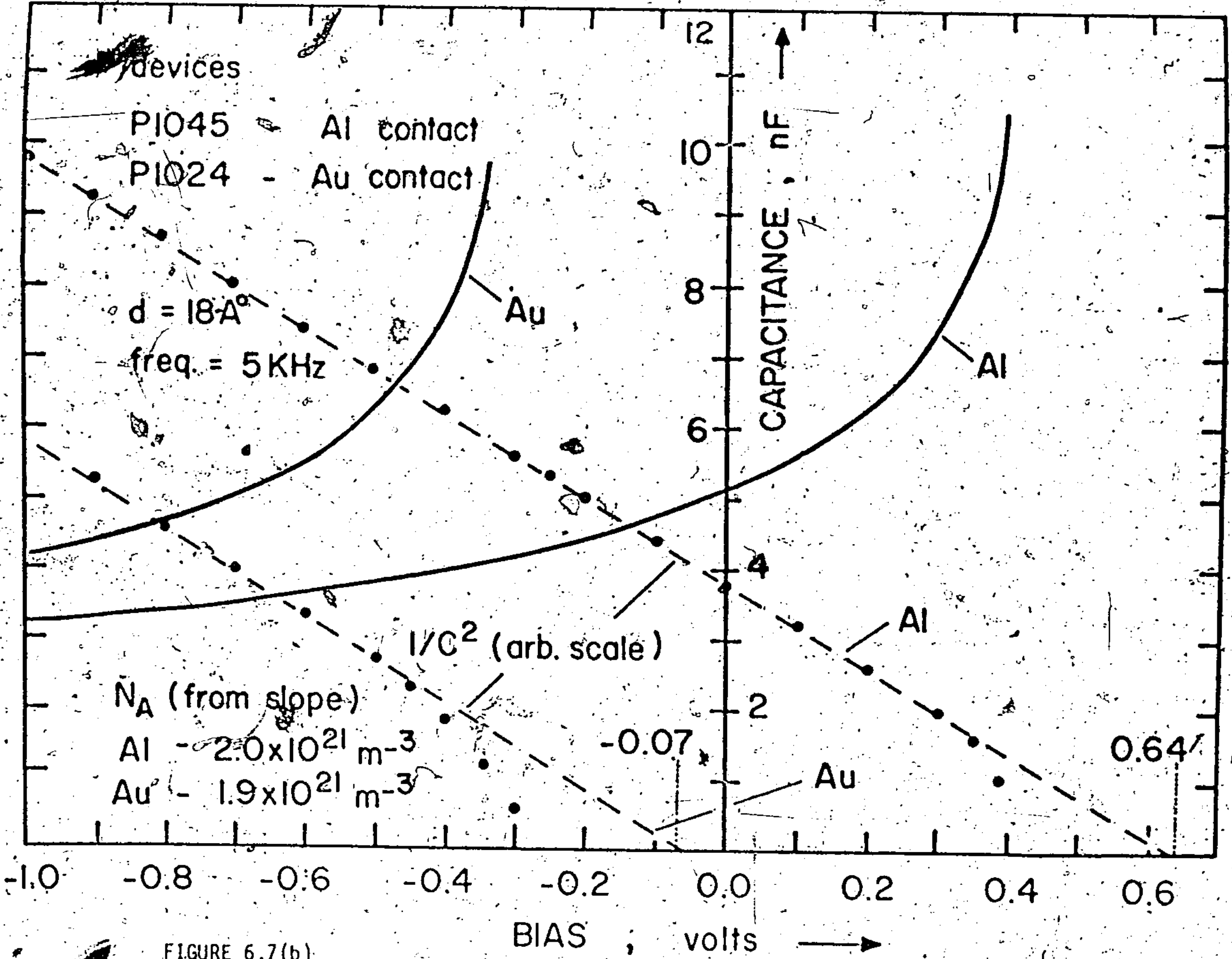


FIGURE 6.7(b)

condition in the semiconductor due to the limited generation rate of minority carriers as discussed in Chapter 3, 4 and 5. The reverse current of the gold diode is predominantly majority carrier flow being several orders of magnitude greater than the current due to the generation of minority carriers which is reflected in the aluminum diode characteristic.

The C-V characteristics shown in Figure 6.7(b) also agree with the theory of Chapter 3. Under reverse bias, both diodes display a linear $1/C^2$ versus V relationship with a slope determined by the dopant density in the substrate. The gold diode is displaced to the left by approximately 0.7 V. These results agree with those obtained for a range of metal contacts tabulated in Table 4.2.

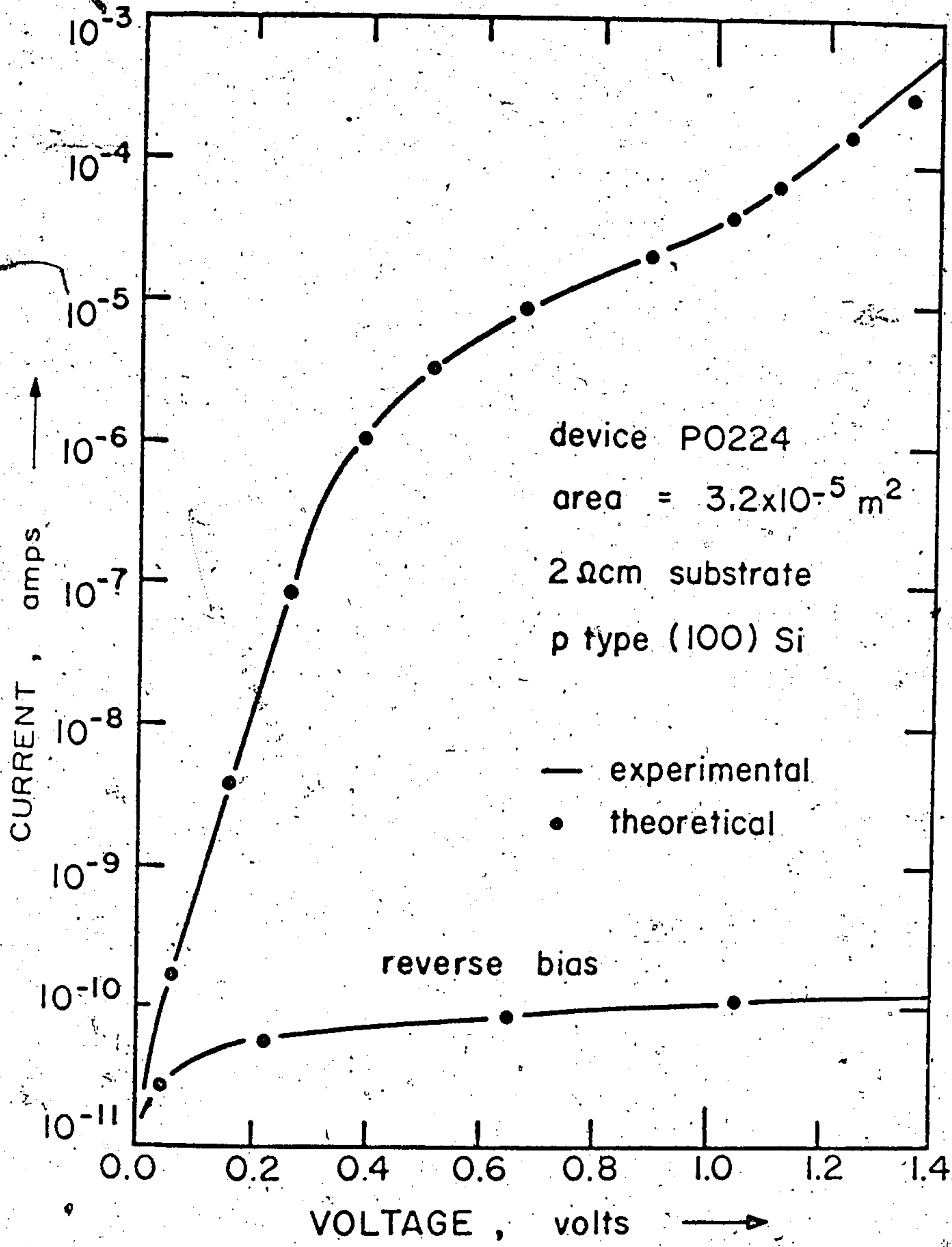
6.3.2 I-V and C-V Characteristics of Minority Carrier Diodes

The experimentally observed current-voltage characteristics of an aluminum minority carrier diode formed on a 2 Ωcm p type silicon substrate of <100> orientation with $d \sim 24$ Å are shown in the semi-logarithmic plot of Figure 6.8(a). With the semiconductor biased positively with respect to the aluminum contact, the current increases approximately exponentially with voltage until about 0.3 V, increases less rapidly to 1 V bias, and then it again starts to increase rapidly with bias above 1 V. In the reverse direction, the diode current is very small (~ 0.4 nA/cm^2) and a slowly increasing function of reverse bias.

The experimentally observed characteristics can be quantitatively described to a high degree of accuracy using the computer simulation program [102] assuming a surface state distribution as in Figure 3.11 and

Figure 6.8(a): Experimentally observed I-V characteristics for an aluminum-silicon dioxide-p type silicon diode. The silicon substrate was of <100> orientation and 2 Ωcm resistivity. The theoretical points were calculated assuming the surface state distribution of Figure 3.11 in addition to an oxide charge density of $3 \times 10^{15} m^{-2}$. d was assigned the value of 23.5Å and ϕ_{mi} was taken as 3.2 eV. Other parameters are given in Table 2.4.

(b): Experimentally observed C-V characteristics of the diode of Figure 6.8(a). Also shown are the characteristics of a diode formed on the same substrate but with the area reduced by a factor of 16. The theoretical points are calculated using exactly the same parameters as in (a) at a frequency of 50 kHz.



CURRENT , amps
VOLTAGE , volts
10-3
10-4
10-5
10-6
10-7
10-8
10-9
10-10
10-11
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
device P0224
area = 3.2x10-5 m2
2 Ωcm substrate
p type (100) Si
experimental
theoretical
reverse bias

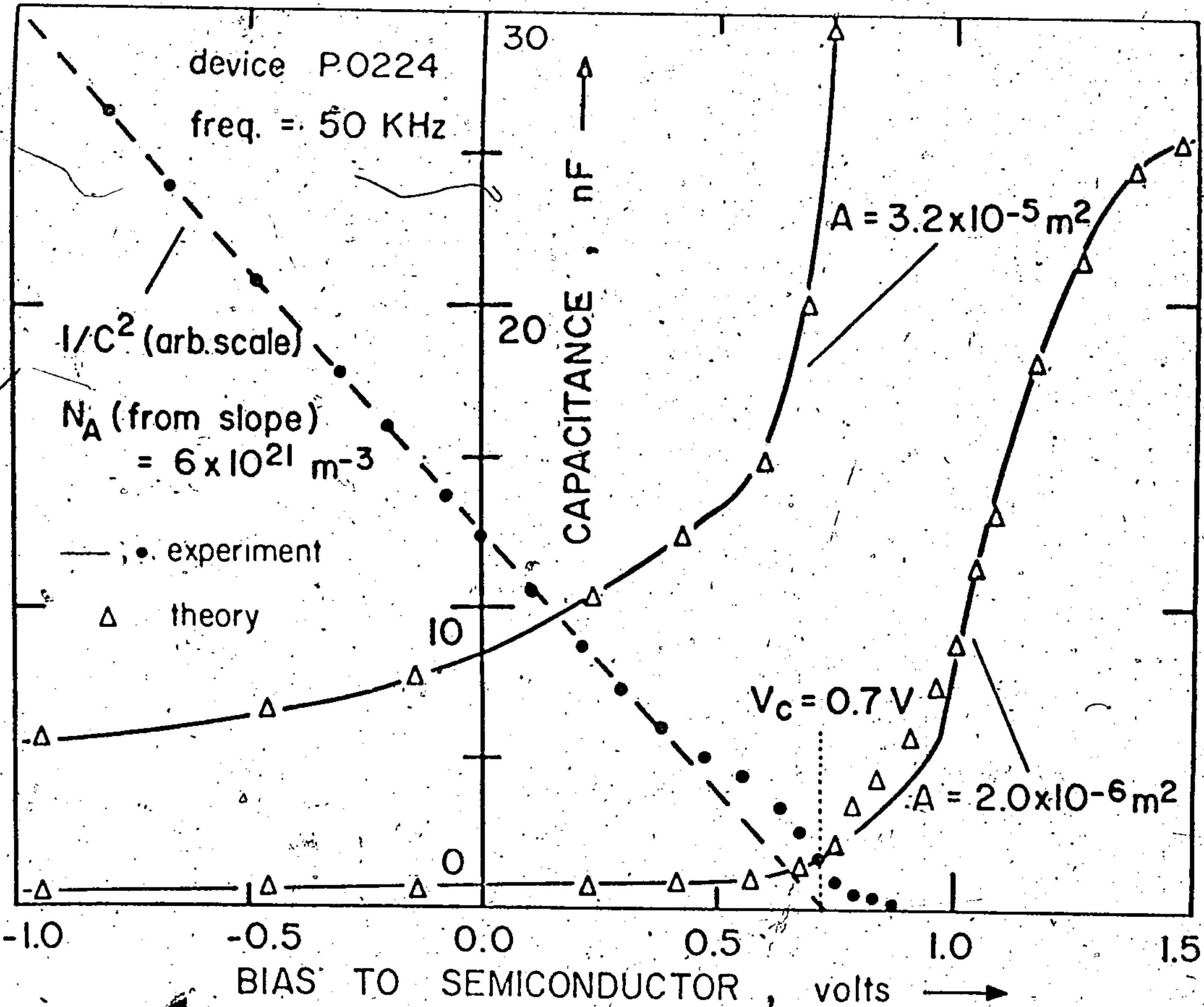


FIGURE 6.8(b)

using values for the diode parameters as given in Table 2.4. The current flow in the bias region below 0.3 V is semiconductor limited and is tunnel limited above 0.3 V. The increase in current flow as the bias is changed from 0.3 to 1 V is due to the increase in the positive charge contribution of surface states. The more rapid increase in current flow above 1 V bias is due to the surface going into accumulation. The experimental-theory fit is very accurate.

The corresponding C-V characteristics are shown in Figure 6.8(b). The diode capacitance is seen to increase rapidly with increasing forward bias. It should eventually saturate when the surface of the semiconductor becomes accumulated, as in the conventional thick oxide MOS devices. Due to the relatively large current flows (Figure 6.8(a)) at these bias points, it was found that the bridge used to measure the device capacitance went out of balance before this region was attained. However, a diode of 1/16 times the area was also formed at the same time on the same substrate. This diode had both its capacitance and the current flow through it reduced correspondingly. It was found possible to retain the bridge in balance in this case and the resulting C-V characteristics are also shown in Figure 6.8(b). Again the experimental results can be described to a high degree of accuracy by the simulation program [102] using exactly the same parameters as used for the theory points of Figure 6.8(a).

Also shown in Figure 6.8(b) is a plot of $1/C^2$ versus bias for the larger diode. It is shown in Appendix C that the C-V characteristics of minority carrier diodes when semiconductor limited can be expressed in the form

$$\frac{C}{A} = [\frac{2(V_c - V_a)}{q\epsilon_s N_A}]^{-1/2} \quad (6.13)$$

where A is the diode area and N_A is the impurity density. In Figure 6.8(b), the plot of $1/C^2$ versus V_a is linear under reverse bias conditions with deviations occurring under forward bias. The intersection of the linear portion with the voltage axis is V_c. This voltage is related to the value of the surface potential at the semiconductor-insulator interface under zero bias conditions. However, since minority carrier diodes tend to have an inversion layer at this interface under zero bias, the relationship is complex. As is shown in Appendix C, in such cases

$$V_c = \psi_{inv} \quad (6.14)$$

where ψ_{inv} is the value of the surface potential necessary to drive the surface into strong inversion. The calculated value for the diode in Figure 6.8(b) is 0.67 V. Since the measured V_c of 0.7 V is larger than ψ_{inv}, the capacitance curve indicates that the semiconductor surface is inverted at zero bias.

The density of dopant impurities can be compared from the slope of the linear portion of the $1/C^2$ curve of Figure 6.8(b). The computed value of N_A is $6 \times 10^{21} m^{-3}$ which is in good agreement with the value of $7 \times 10^{21} m^{-3}$ obtained from the semiconductor resistivity.

The current-voltage characteristics of Figure 6.8(a) are for a relatively thick insulating layer. Decreasing the oxide thickness will displace the tunnel limited regime to higher current levels. This behaviour is illustrated in Figure 6.9 where the current-voltage

Figure 6.9: The experimentally observed dependence of the I-V characteristics upon the thickness of the silicon dioxide layer. The top contact was aluminum and the silicon substrate was of 2 Ωcm resistivity and 100 orientation in all cases. The dashed line represents the "ideal diode" law of Equation (6.15).

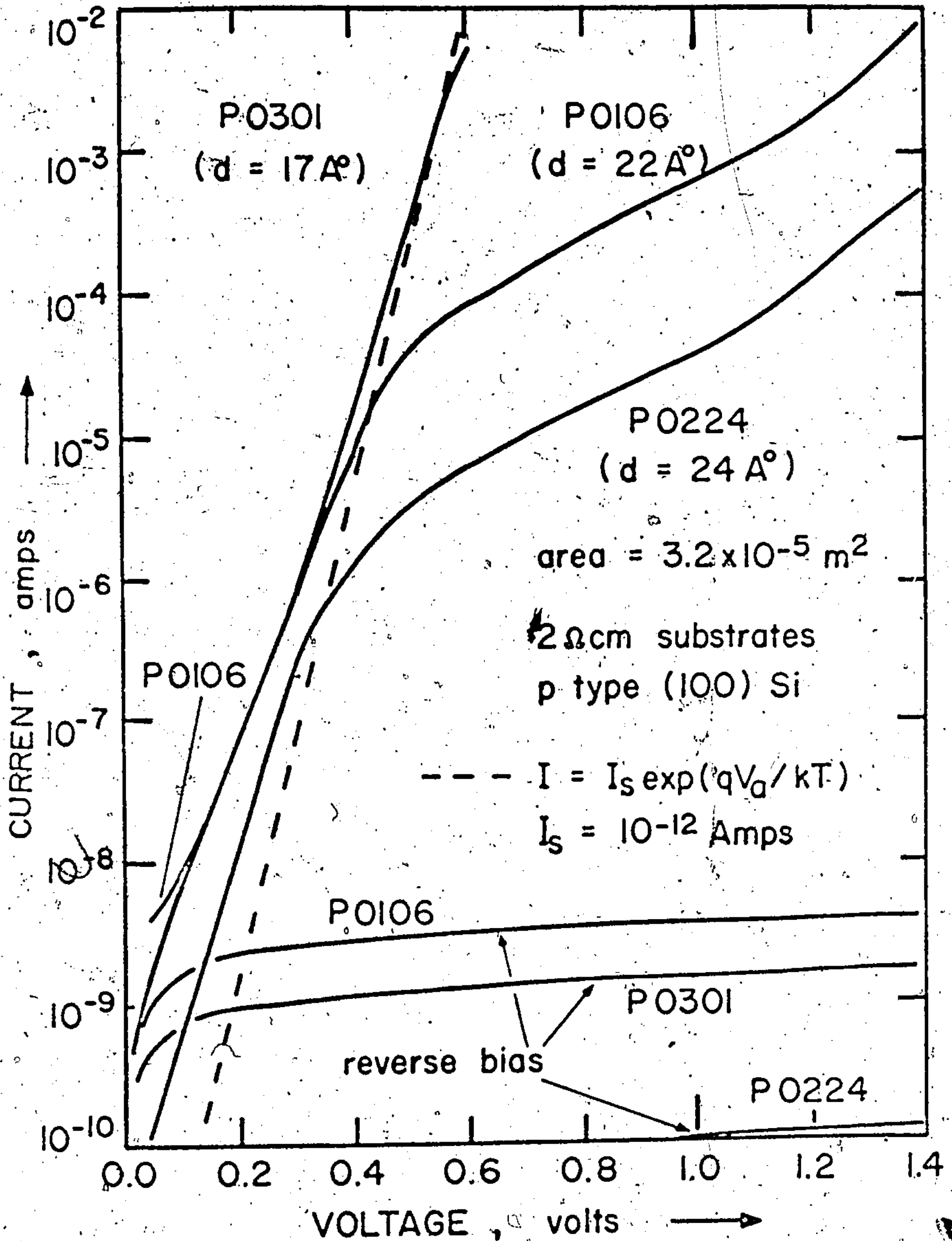


FIGURE 6.9

characteristics are shown for three oxide thicknesses in the range 17Å to 24Å. The substrate resistivity is the same in each case. The current-voltage characteristics of the 17Å and 22Å diodes are nearly coincident up to 0.4 V forward bias indicating that the diode current in this range is more likely to be semiconductor limited rather than tunnel limited.

Also shown on the same graph is the ideal diode law under forward bias

$$I = I_s \exp(qV_a/nkT) \tag{6.15}$$

with n equal to 1. The experimental curves never increase as rapidly with bias as this curve. This deviation from ideality can be described by a value of n, the ideality factor, greater than 1. The general behaviour of the diodes as expected on the basis of the theory of Section 6.2 is that at low forward bias ($V_a < 0.2$ V) the diode current is limited by recombination processes in the semiconductor depletion region giving an n value of 2. As the bias increases ($0.2 \text{ V} < V_a < 0.5$ V), the diffusion component of the diode current increases and n approaches 1. At high forward bias, the diode current can become either tunnel limited, limited by the resistance of the "ohmic" back contact, or limited by high injection effects. These current limiting processes cause the value of n to increase again.

The results shown in Figure 6.9 are in qualitative agreement with this description. The 17Å and 22Å diodes have n values of 1.6 and 1.8 respectively at 0.2 V decreasing to 1.2 at 0.4 V for the 17Å diode and to 1.5 at 0.35 V for the 22Å diode. The n values then begin to increase

due to back contact limitations for the 17Å diode, and due to entering the tunnel limited regime for the 22Å diode. The 24Å diode differs in that it has a much smaller recombination-generation component of current as discussed below. n decreases from 1.3 at 0.15 V to 1.2 at 0.2 V before increasing again as the current flow becomes tunnel limited.

The reverse currents, although very small and relatively bias insensitive in each case, differ from diode to diode. Theory predicts that this current should consist entirely of current generated in the semiconductor depletion region beneath the metal contact and should be equal for the different oxide thicknesses shown. The 24Å diode has an extremely small value of reverse current (< 0.4 nA/cm^2 at 1 V reverse bias) which is probably composed mainly of this generation current. The 17Å and 22Å diodes have reverse current flows 1 - 1½ orders of magnitude larger. Some variation in reverse current can be attributed to differences in the semiconductor parameters determining the generation rate in this depletion region. However, the major cause of the additional reverse current for these diodes is attributed to generation processes occurring in the vicinity of the semiconductor surface but in the regions that are not directly underneath the top metal contact. Such lateral effects were not included into the device analysis which was for a "one-dimensional" approximation to the actual experimental situation (Section 2.6.1). This conclusion was reached after it was observed experimentally that the reverse current depended upon the surface potential in these "remote" regions which was varied by the treatments mentioned in Section 4.1.

6.3.3 Comparison of Minority Carrier Diodes

Magnesium has a very low value of ϕ_{mi} which is estimated as about 2.2 eV (Section 4.6) and it is of interest to compare the properties of diodes made with this metal as the top contact with those of diodes made using aluminum. A thin oxide layer (16Å) was grown onto a p type <111> orientated substrate of resistivity 8 Ωcm and aluminum and magnesium contacts were evaporated onto different areas of this substrate.

The I-V characteristics of the aluminum and magnesium diodes are compared on a semilogarithmic plot in Figure 6.10(a). They are seen to be virtually identical over the current range of 10^{-7} to 10^{-3} Amps indicating that the current flow is semiconductor limited and that both contacts form minority carrier diodes. Theoretically, the only differences between the two diodes should be apparent at high forward bias when diodes current becomes tunnel limited. However, for both the diodes, the oxide layer was so thin that the diode current was limited by the "ohmic" back contact before this region was attained. The differences observed between the reverse currents are attributed to the lateral effects described above.

A comparison of the C-V curves for these diodes is shown in Figure 6.10(b). Under reverse bias, both conform to the behaviour given by Equation (6.13). The value of V_c for the Al and Mg diode is 0.74 V and 0.84 V respectively. Since ψ_{inv} for this doping density is 0.60 V, it can be deduced from the treatment given in Appendix C that both diodes are inverted at zero bias, with the Mg diode more strongly inverted. It should be noted that the difference between the V_c values

Figure 6.10: Experimental comparison of the properties of minority carrier MIS tunnel diodes with magnesium and aluminum contacts formed on the same p type silicon substrate of <111> orientation and of resistivity 8 Ωcm with a 16Å layer of silicon dioxide.

(a) comparison of I-V characteristics.

(b) comparison of C-V characteristics.

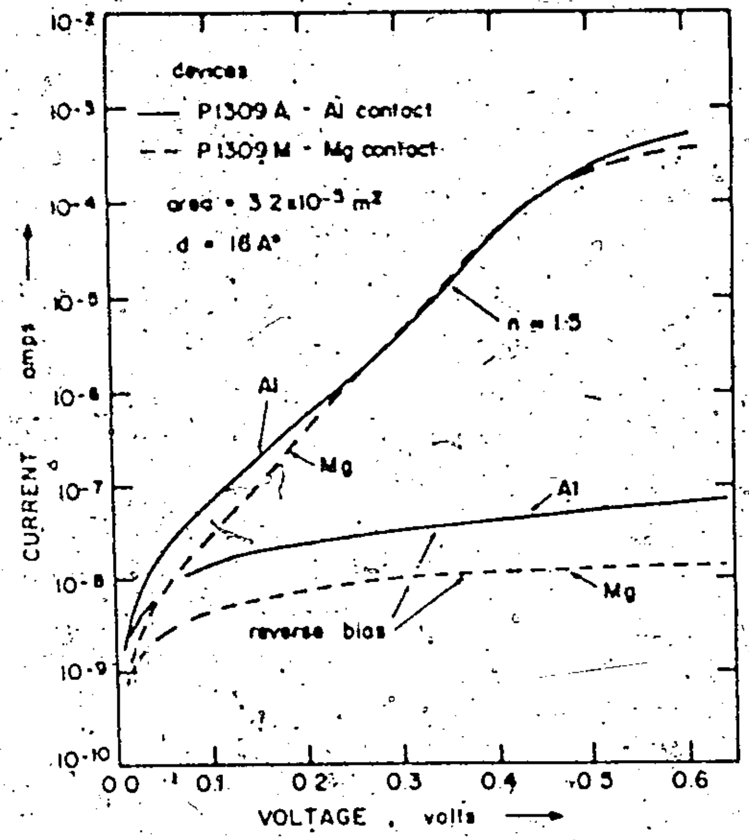


Figure 6.10(a)

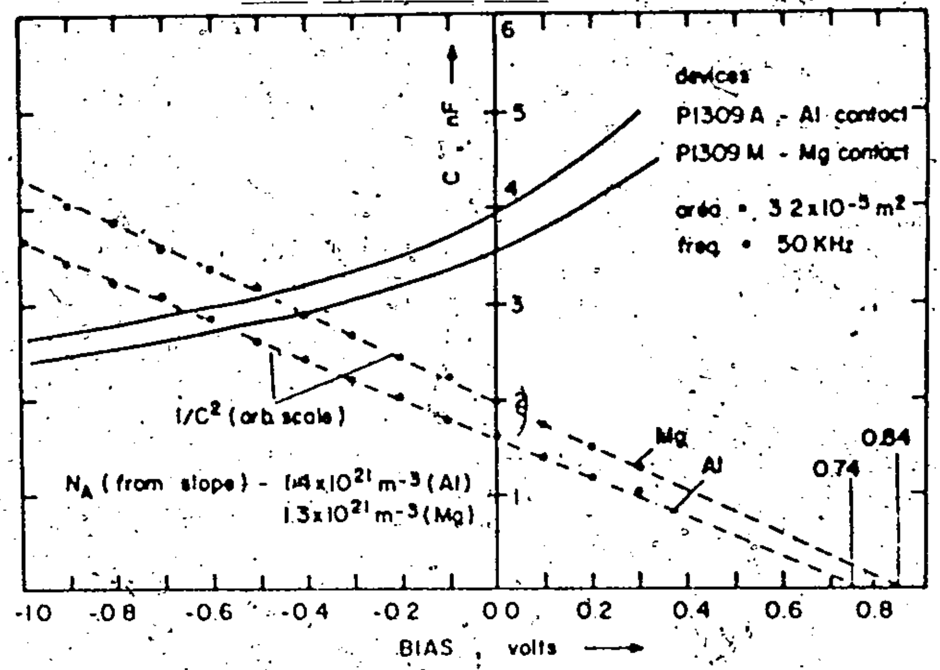


Figure 6.10(b)

cannot be equated to the difference between the values of ϕ_{mi} under such circumstances.

Diodes made with titanium and aluminum as the top metal contact were also compared and similar results were obtained.

6.3.4 The Effects of Substrate Resistivity

As the substrate resistivity is decreased, the diffusion current which flows in the semiconductor for a given applied bias also decreases. Thus, in the bias range 0.4 V to 0.6 V where diffusion current is dominant, there should be predictable differences between the I-V characteristics of diodes with different substrate resistivities.

Experimentally obserbed I-V characteristics for three diodes fabricated on different resistivity p type silicon substrates of <100> orientation are shown on a semilogarithmic plot in Figure 6.11(a). In the vicinity of 0.55 V, these are in reasonable agreement with the simple expression for the diffusion current,

$$I = I_s \exp(qV_a/kT) \tag{6.16}$$

where

$$I_s = \frac{qAn_i^2}{N_A}\sqrt{\frac{D_n}{\tau_n}} \tag{6.17}$$

D_n is the electron diffusion constant and τ_n is the electron lifetime. Equation (6.16) is also plotted in Figure 6.11(a) for each resistivity substrate. At smaller forward bias, the diffusion component of current is masked by the component due to recombination-generation processes within the semiconductor depletion region. The differences between the

Figure 6.11: Experimentally observed dependence of the properties of aluminum-silicon dioxide-p type silicon tunnel diodes upon the silicon resistivity. The substrate orientation in each case was <100> and the silicon dioxide layer was in the thickness range 15-20Å.

(a) comparison of I-V characteristics.

(b) comparison of C-V characteristics.

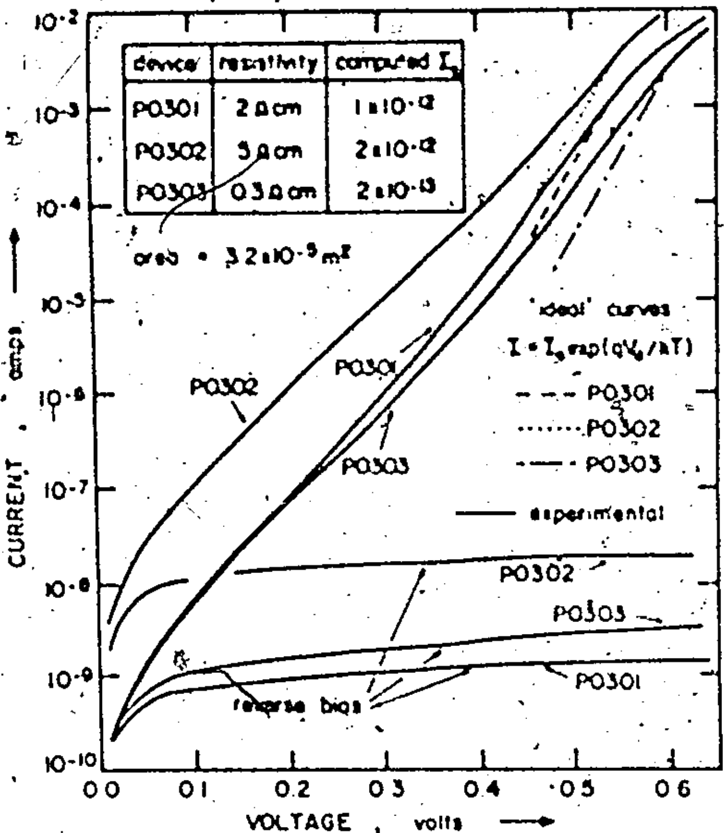


Figure 6.11(a)

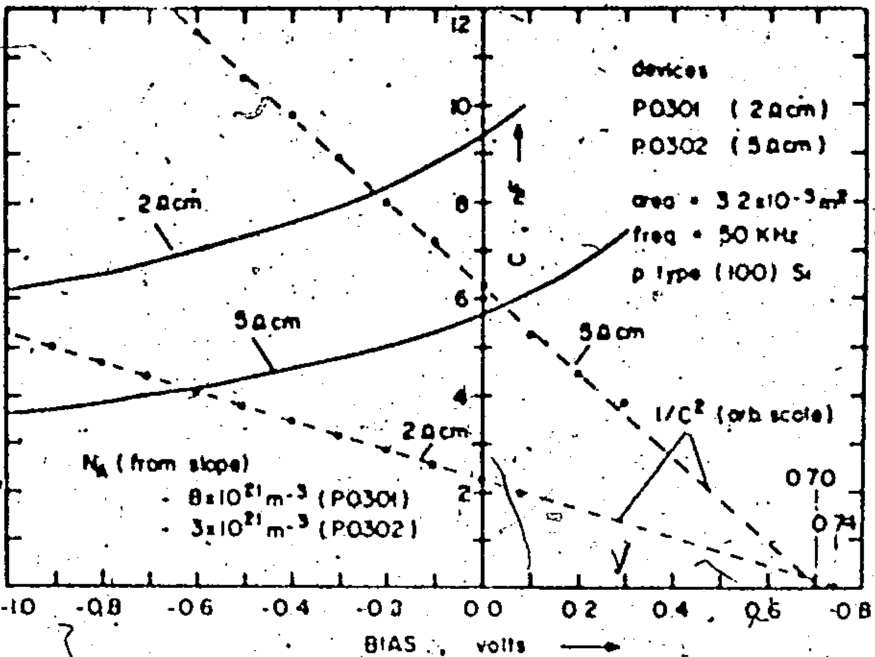


Figure 6.11(b)

reverse currents are again attributed to the lateral surface effects.

The C-V curves for two diodes with different substrate resistivities are compared in Figure 6.11(b). Again Equation (6.13) is satisfied. The more highly doped specimen gives a larger value of V_c due to the shift in the bulk semiconductor fermi level caused by the change in doping density. ψ_{inv} is computed as being 50 mV higher for the more highly doped specimen. The experimentally observed difference between the values of V_c is about 40 mV, in agreement with the above value.

6.4 Application to Direct Energy Conversion

One particularly promising application of the minority carrier MIS tunnel diodes is to energy conversion employing the electron- or photo-voltaic effects. These diodes have several advantages over the shallow diffused p-n junction diodes most commonly used in such applications [93]:

(a) the MIS diode is considerably simpler to fabricate;

(b) since high temperature steps can be avoided, the carrier lifetimes are not significantly degraded during fabrication;

(c) the collecting junction is located nearer the surface of the semiconductor for the MIS diode which results in improved collection of electron-hole pairs generated in the surface region. This is especially important for radiation such as short wavelength light which does not penetrate deeply into the semiconductor;

(d) shallow diffused p-n junction diodes commonly have an excess component of current at small forward bias [93] which degrades the diode performance at low power levels.

For photo-voltaic applications, the top metal contact of the MIS diode must be thin (~ 200Å) or have a grid-like structure so that it does not absorb an appreciable amount of the incident radiation. Assuming the radiation impingent on the device causes a generation rate of electron-hole pairs through the semiconductor region of the diode given by [94]

$$G = \alpha N e^{-\alpha x} \tag{6.18}$$

where α is the semiconductor absorption coefficient and N is the photon flux entering the semiconductor per square metre per second, the collected current density from electron-hole pairs generated in the diode bulk (assumed p type) is, under short circuit conditions [94]

$$J_{nb} = qN\alpha L_n\{\frac{\exp(-\alpha W)}{1 + \alpha L_n} - \frac{\exp(-\alpha L)}{(\alpha L_n)^2 - 1}[\frac{\exp(-(1 - \alpha L_n)(L - W)) - 1}{\sinh(L - W)}]\}. \tag{6.19}$$

W is the width of the semiconductor depletion region plus inversion region and L_n is the diffusion length in the semiconductor. This current is augmented by the contribution to the total collected current density made by electron-hole pairs generated in the depletion and inversion regions. The voltage generated under open-circuit conditions can be computed from this short-circuit current density as for the p-n junction diode [94].

For small values of the absorption coefficient, α, the MIS tunnel diode performance is similar to an N^+P junction diode with the N^+ region thin and identical lifetimes in the P regions. As α increases, the MIS diode performance is superior since the depletion region of the device is located nearer the semiconductor surface. The presence of the thin insulating layer can degrade the MIS diode performance at high radiation intensities when the short circuit current given by Equation (6.19) is larger than the current where the diode becomes tunnel limited under dark conditions. In such cases, the surface of the semiconductor must become even more inverted to allow the collected current to tunnel through the insulator. This causes an increase in the potential drop across the insulator which degrades the device performance in much the same way as a series resistance degrades the performance of the more conventional p-n junction devices [19]. However, this effect is simply eliminated by decreasing the insulating layer thickness which displaces the tunnel limited regime to higher current levels (Figure 6.9). For the p type silicon diodes, an oxide layer thinner than 16Å would be sufficiently thin for solar cell applications.

The use of the electron-voltaic effect to convert β radiation from a radioisotope to electrical energy has been investigated as a means of producing a small volume, low power, long lifetime atomic battery. Silicon N^+P junction diodes coupled with the use of the radioisotope Pm^{147} as the primary energy source previously have been found to be most satisfactory combination for this application [95]. The maximum theoretical conversion efficiency of this combination has been estimated as 4.5% [96]. This efficiency is low compared to the corresponding figure

for silicon solar cells (~20%) due to the facts that (a) about 4 eV of β radiation from the radioisotope is required for the creation of an electron-hole pair which can be accomplished with 1.1 eV by a completely efficient process, (b) incident power densities are much lower than that associated with sunlight (about three orders of magnitude smaller) resulting in small voltage and curve factors [94] and, (c) the β source tends to absorb some of the available radiation so there exists an optimum thickness of the radioisotope source where the ratio of the output power to the available power is a maximum [96].

The same theoretical considerations apply to the MIS diode as to the N^+P junction diode so that 4.5% also represents an upper value of the conversion efficiency in this case. The selection of the resistivity of the p type substrate depends on the same factors as in the design of junction solar cells [93]. A low resistivity substrate is required for large open circuit voltages, but a high resistivity substrate is necessary to obtain the high lifetimes necessary for large short circuit currents. The experimental results for different substrate resistivities and orientations for MIS diodes with an Al top contact are shown in Table 6.1. The optimum resistivity was found to be in the vicinity of 1 Ωcm. The performance of the diode was found to be independent of the oxide thickness provided it was less than 22Å. Diodes with titanium top contacts performed similarly.

The output characteristics of a MIS tunnel diode and a diffused silicon N^+P junction diode are compared in Figure 6.12. The absolute conversion efficiency of the MIS diode is 4% while that of the junction diode is 3%. These figures are 4-5 times larger than those reported for

TABLE 6.1: The Effect of Substrate Resistivity upon the Initial Beta Cell Output

Specimen No.	Resistivity(Ωcm)	Orientation	Voc(mV)	$I_{sc}(\mu A cm^{-2})$
P1303	50 - 150	<111>	322	32
P1323	5 - 10	<111>	378	31*
P1322	5 - 10	<111>	385	29*
P0304	1.6 - 2.4	<100>	402	33
P0312	1.6 - 2.4	<100>	424	34
P0321	0.7 - 1.3	<100>	448	30*

* These diodes had an active area of 3 cm^2. The other diodes had an active area of 1 cm^2.

Figure 6.12: Experimental comparison of the energy conversion properties of a diffused p-n junction diode and a MIS tunnel diode (device P0321). Shown are the output characteristics in each case using the radioisotope Pm^{147} as the primary energy source. The source contained 4 Curies.

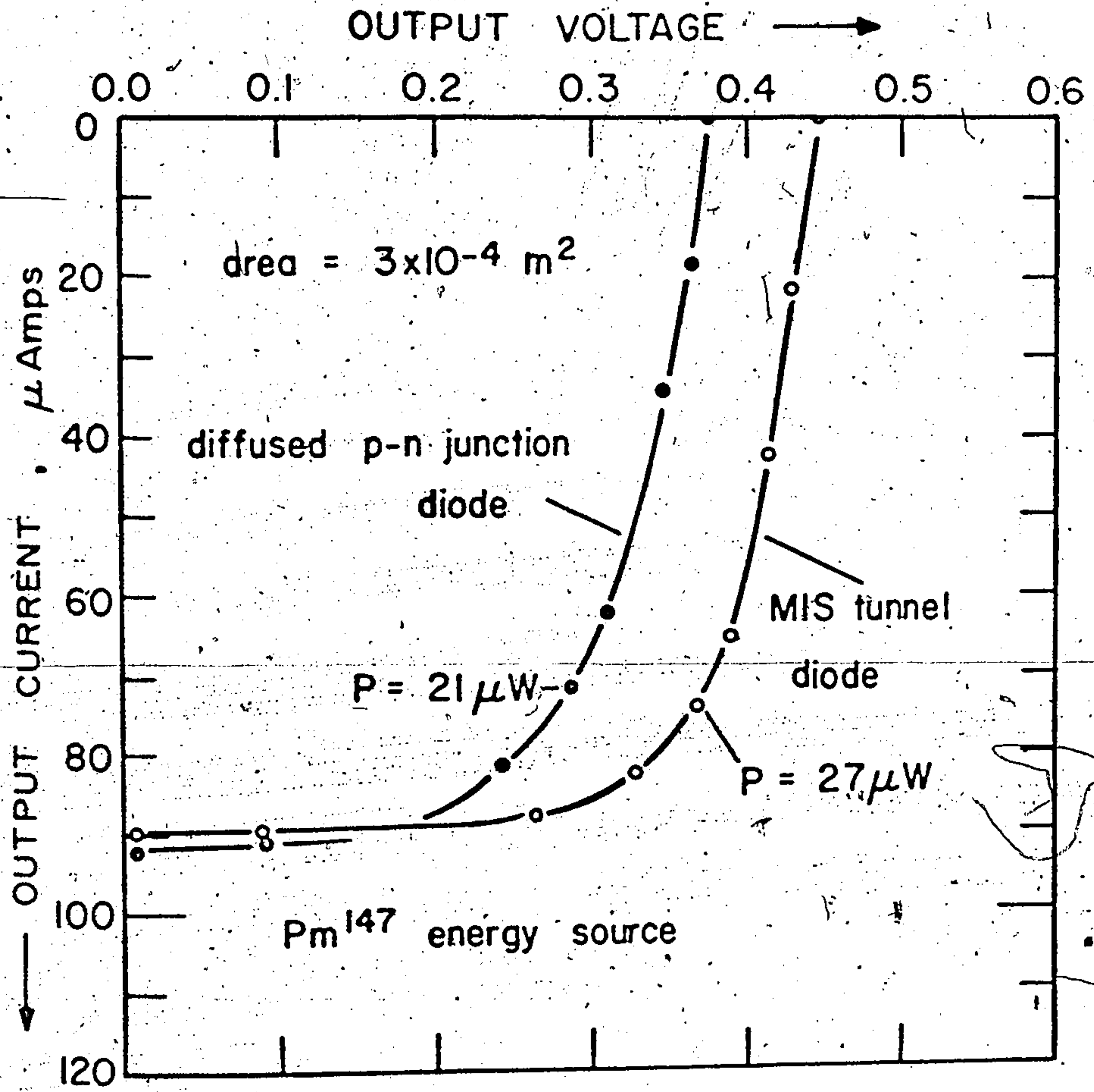


FIGURE 6.12

earlier attempts using Pm^{147} as the energy source [95]. In fact, the MIS tunnel diode gives results which approach the theoretical maximum. Its improved efficiency compared to the junction diode is caused by the larger open circuit voltage obtained without deterioration of the short circuit current. It is also aided by an improved curve factor caused by the more ideal (lower n values) current-voltage characteristics at these low current levels.

Subsequent to the above experimental work, conversion efficiencies for N^+P junction devices comparable to these results for the MIS tunnel diode have been reported [97]. Much of the improvement in performance of the junction device relative to the MIS device is attributed to the fact that, in reference [97], as well as optimizing the diode properties, the thickness of the radioisotope source was also optimized. This latter optimization step was not undertaken in the work described above.

6.5 Similarities to p-n Junction and Schottky Diodes

The minority carrier device when operating in the semiconductor limited bias regions has been shown to have properties similar to a p-n junction diode. In such regions, the device can be thought of as an N^+P junction diode, the N^+ region being the induced inversion layer at the semiconductor surface under the contact. The major difference is that the current flowing through the N^+ region is limited by tunneling processes through the insulator in one case while being limited by transport processes (drift and diffusion) through the N^+ region in the other case.

The device is also very similar to a Schottky diode both in the ease of fabrication and in the resulting energy band relationships. One

very important difference is that the dominant current is minority carrier current rather than majority carrier current. In a Schottky diode, majority carrier flow is limited by thermionic emission over the barrier while minority carrier flow is limited by diffusion through the semiconductor [98]. As the Schottky barrier height increases, the majority carrier flow decreases while the minority carrier flow is the same at a given small applied bias. Viewed from this aspect, the introduction of the insulating layer causes the minority carrier MIS tunnel diode behaviour to depart so significantly from Schottky diode behaviour due to the following two effects:

(a) it eliminates the pinning effects [75,99] which normally keep Schottky barriers low (about 2/3 of the semiconductor band gap);

(b) it provides an additional impediment to majority carrier flow (as well as to minority carrier flow).

These two effects are taken advantage of in the following manner. The metal contact is selected to produce large barriers at the semiconductor surface and the insulating layer is kept sufficiently thin so that minority carrier flow is still limited by diffusion through the semiconductor. Thus at small forward bias, the minority carrier flow is the same as in a Schottky diode while the majority carrier flow has been reduced several orders of magnitude by the increased barrier and additional orders of magnitude due to the fact that it must tunnel through the insulator.

Note that neither of the above comparisons attempts to explain

why the inversion (or N^+) region remains at the semiconductor surface when the semiconductor is forward biased. This is explained in Section 6.2.3.

6.6 Summary

Employing the general numerical analysis method of Chapter 2, it has been shown that minority carrier flow can dominate the current-voltage characteristics of MIS tunnel diodes with the metal contact suitably chosen. If the insulating layer is very thin, this current flow can significantly disturb the semiconductor from thermal equilibrium. The result is a diode with properties similar to those of an ideal p-n junction diode.

This surprising behaviour of the MIS tunnel contact is caused by the pinning of the metal fermi level and the semiconductor minority carrier quasi-fermi level under reverse and small forward bias conditions. It is shown that this effect tends to clamp the minority carrier concentration in the IS interface region to its value under zero bias conditions. Thus, for the minority carrier diodes under discussion, this means there is an inversion layer at the semiconductor-insulator interface under both reverse and moderate forward bias conditions. As the forward bias becomes large, the diode current is eventually limited by tunneling through the insulator rather than by diffusion through the semiconductor and the inversion layer disappears. Surface state effects are important only in this tunnel limited regime. Recombination-generation effects at these states are inhibited in the semiconductor limited regime by the presence of the inversion layer at the IS interface.

Experimental measurements confirming this analysis are reported

employing the metal-silicon dioxide-p type silicon system with the oxide layer in the thickness range 15-25Å. Several metals have been employed as the top contact to the device. Aluminum, magnesium, and titanium are found to produce the minority carrier diodes whereas gold is not. Also investigated were the effects of oxide thickness and substrate resistivity upon the diode properties. The experimental results are in good agreement with the theoretical description of the device operation.

Although having similar properties, the minority carrier MIS tunnel diode has a distinct advantage over p-n junction diodes in its ease of fabrication. In addition to having applications as an injecting contact, as a photodiode, or as a member of a photodiode array, the device is well suited to direct energy conversion using the electron- or photo-voltaic effect. An experimental investigation of its electron-voltaic energy conversion properties using the radioisotope Pm^{147} as the primary energy source shows that the device is capable of higher conversion efficiencies than those obtained using p-n junction didoes in this application. The experimental conversion efficiencies approached the theoretical maximum.

CHAPTER 7

CONCLUSION

The work presented in this thesis has been concerned with investigating the electrical transport properties and possible applications of the three layer metal-insulator-semiconductor (MIS) structure when the insulating layer is very thin (< 60Å). In this configuration, current can flow between the metal and the semiconducter by quantum mechanical tunneling processes. The work is restricted to the case where the semiconductor is not degenerately doped.

The first part of this thesis is concerned with characterizing the MIS tunnel device by appropriate mathematical expressions. The semiconductor region of the device is characterized directly by the basic differential equations describing carrier transport within it. These are Poisson's equation, the carrier continuity equations, and the carrier current density equations as derived from the Boltzman transport equation. Recombination-generation processes in the semiconductor are modelled as occurring through single level trapping centres. It was not found possible to characterize tunneling through the insulator in a manner as fundamental as this. An independent particle approach as formulated by Harrison [34] is used. Although this general approach can be criticized on a number of theoretical grounds as summarized by Duke [33], it is known to be capable of predicting most tunneling properties in at least a qualitative fashion. Surface states at the insulator-semiconductor

(IS) interface are important in MIS devices and so are included into the device characterization. In addition to their usual role as recombination-generation and charge storage centres, the effect of tunneling between the metal and these states is also included. At the IS interface an abrupt transition from insulator to semiconductor properties is assumed, and Gauss's Law and other continuity conditions are applied across the interface. The system of equations resulting from the above characterization are solved exactly with a digital computer.

Based on studies using this simulation technique, a broad theoretical framework for interpreting the properties of the MIS tunnel diode as device parameters vary is established. Special emphasis is placed upon the Metal-SiO_2-Si system although the general trends described are applicable to other systems. The most important device parameters for such a system are the work function of the metal contact, the thickness of the insulator layer, and whether the semiconductor is n or p type. Depending upon whether the dominant component of tunnel current near zero bias is between the metal and the minority carrier band, the majority band, or the surface-state levels in the semiconductor, devices are classified as minority carrier, majority carrier, or surface state diodes. This is shown to correspond closely to whether the IS interface is strongly inverted, accumulated, or depleted under zero bias conditions. Depending primarily upon the insulator thickness, two additional classifications are made, "equilibrium" and "non-equilibrium". In equilibrium devices, the insulator layer is relatively thick and the corresponding small values of tunnel current easily pass through the semiconductor

region. In non-equilibrium devices, the insulator layer is much thinner allowing much larger current flows and the transport properties in the semiconductor become important in determining the device properties. While the equilibrium devices are most important in analyzing the basic tunneling properties of the MIS contact, the non-equilibrium devices are the most important from an applications point of view. Detailed experimental results upon the Metal-SiO_2-Si system are reported. Data is presented showing the dependence of the properties of these devices upon the metal contact, the insulator thickness, whether the semiconductor is n or p-type, the semiconductor resistivity, and the device temperature. These results establish the validity of the previous theory and also the utility of the classification scheme described. Temperature measurements are shown to be of particular importance in interpreting the properties of experimental devices.

Non-equilibrium devices are treated further. The multiplication properties of majority carrier diodes are investigated both theoretically and experimentally. Gains of the order of 100 to 1000 are obtained. The frequency response of the multiplication process is measured and found to be relatively low as predicted in the theoretical treatment. Applications of the effect to a new type of transistor structure and as a photodiode with internal multiplication properties are briefly described. The relatively poor frequency response of the multiplication process is expected to be the major factor likely to limit its use in these applications. Minority carrier diodes are shown both theoretically and experimentally to possess properties the same as p-n junction diodes

under reverse and small forward bias conditions. The MIS tunnel diodes have a distinct advantage over p-n junction devices in their ease of fabrication. In addition to having applications as an injecting contact, as a photodiode or a member of a photodiode array, the device is well suited for direct energy conversion using the electron- or photo-voltaic effect. Using an electron-voltaic energy conversion situation, it is demonstrated that the predicted higher conversion efficiencies compared to p-n junction devices can be experimentally realized.

The major contributions of this thesis are in three areas. The first is the extension of the use of the conceptually appealing transmission line equivalent circuit model of a semiconductor from the AC analysis of semiconductor devices to the whole range of AC, DC, and transient analyses. This method of analysis is demonstrated in this thesis for the MIS tunnel device and elsewhere for the p-n junction device [27]. A second is the large leap in the understanding of the theoretical properties of the non-degenerate MIS tunnel diode and their dependence upon the device parameters represented by the theory sections of this work and the demonstration of these properties experimentally using the Metal-SiO_2-Si system including the demonstration of the multiplication properties of majority carrier diodes. The third is the theoretical prediction and the experimental demonstration of the semiconductor limited properties of minority carrier devices. This new class of diode with its p-n junction properties and structural similarity to Schottky diodes could prove important in energy conversion applications. It is felt that the most important goal for further research in this area

is to develop the understanding of the experimental properties of other non-degenerate MIS systems to the stage attained in this thesis for the silicon system. Of particular interest are systems of the type metal-metal oxide-deposited semiconductor (Section 4.1) which have possible large scale application in terrestial solar energy conversion.

APPENDIX A

EXPRESSIONS FOR THE TUNNEL CURRENTS BETWEEN THE METAL AND THE SEMICONDUCTOR BANDS

In this appendix, analytical expressions will be derived for the band tunnel currents from the tunneling formulation given in Chapter 2. These expressions are valid provided neither the electron nor hole concentrations are degenerate and the metal fermi level lies at energies corresponding to the band gap of the semiconductor. With the edge of the conduction band as the energy reference, the current flowing from the conduction band to the metal, J_{CT}, can be expressed as (Equation (2.20))

$$J_{CT} = 4\pi q \frac{m_{Ti}}{h^3} \int_0^{E_{max}} dE(f_m - f_s) \int_0^{E\frac{m_{Ts}}{m_{Ti}}} dE_T e^{-\eta} \qquad \text{(A1)}$$

The integral over E_T can be simplified by noting that it is sharply peaked at E_T equal to zero. Thus most of the contribution to the integral comes from values of E_T near zero where η can be approximated by a truncated Taylor series expansion. This gives to a good degree of approximation

$$J_{CT} = 4\pi q \frac{m_{Ti}}{h^3} \int_0^{E_{max}} dE(f_m - f_s)\exp(-\eta)\left\{\left[1 - \exp\left(-\frac{d\eta}{dE_T} E \frac{m_{Ts}}{m_{Ti}}\right)\right]\Big/\frac{d\eta}{dE_T}\right\}\Bigg|_{E_T=0} \qquad \text{(A2)}$$

If the metal fermi level lies at energies corresponding to the semiconductor band gap and the electron concentration is non-degenerate, both f_m and f_s are strongly peaked for values of E near zero. Hence the

maximum contribution to the integral comes from this region. Using this fact gives

$$J_{CT} = 4\pi q \frac{m_{Ts}}{h^3} \exp(-\eta|_{E,E_T=0}) \int_0^\infty dE(f_m - f_s)E \tag{A3}$$

and f_m and f_s are given by

$$f_{m,s} = 1/(1 + \exp[(E - b_{m,s})/kT]) \tag{A4}$$

where b_m and b_s represent the energy of the metal fermi level and semiconductor electron quasi-fermi level above the conduction band edge in the semiconductor. From the definition of the Fermi-Dirac integrals [100], F_j, Equation (A3) becomes

$$J_{CT} = 4\pi q \frac{m_{Ts}}{h^3} P_n(kT)^2[F_1(b_m/kT) - F_1(b_s/kT)] \tag{A5}$$

where

$$P_n = \exp(-\eta|_{E,E_T=0}) \tag{A6}$$

P_n is the probability that an electron in the metal of energy approximately equal to the energy of the conduction band edge in the semiconductor moving perpendicularly to the barrier will make a tunneling transition. From Figure 2.1 and Equation (2.22), b_m can be seen to have the value

$$b_m/q = F_s d \frac{\epsilon_s}{\epsilon_i} + \frac{q(Q_i + Q_{ss})d}{\epsilon_i} + \phi_{si} - \phi_{mi} \tag{A7}$$

while b_s is given by

$$b_s/q = v_I - v_N - V_{CI} \qquad \text{(A8)}$$

where these terms are defined in Section 2.2.1, and (A8) is evaluated at the semiconductor surface.

In the same manner, the valence band tunnel current can be derived. It takes the form

$$J_{VT} = 4\pi q \frac{m'_{Ts}}{h^3} P_p (kT)^2 [F_1(b'_s/kT) - F_1(b'_m/kT)] \qquad \text{(A9)}$$

where b'_s and b'_m represent the energy of the semiconductor hole quasi-fermi level and the metal fermi level below the valence band edge. They are given by

$$b'_s/q = v_p - v_I - V_{IV} \qquad \text{(A10)}$$

and

$$b'_m/q = -E_{gs} - b_m/q \qquad \text{(A11)}$$

Equations (A5) and (A9) can be expressed in the form used in Chapter 3 by noting a property of the Fermi-Dirac Integrals displayed in Figure A1. For small positive and all negative arguments

$$F_{1/2} = F_1 \qquad \text{(A12)}$$

Therefore, since the derivation was restricted to such values of the arguments, they become

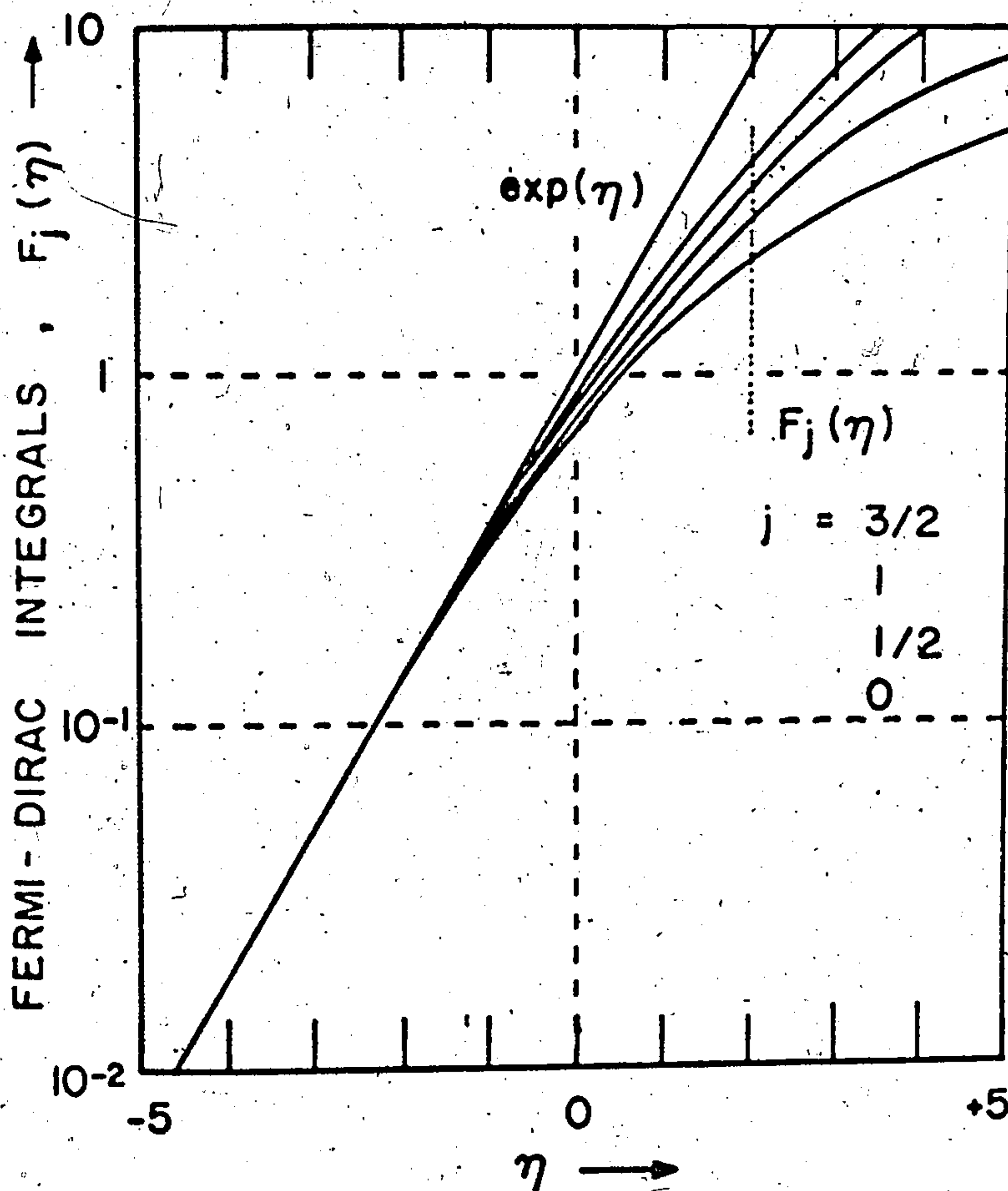


Figure A1: Plot of the Fermi-Dirac Integrals [100] for small positive and negative arguments.

$$J_{CT} = 4\pi q \frac{m'_{Ts}}{h^3} P_n (kT)^2 (n_m - n_s)/N_c \quad \text{(A13)}$$

$$J_{VT} = 4\pi q \frac{m'_{Ts}}{h^3} P_p (kT)^2 (p_s - p_m)/N_v \quad \text{(A14)}$$

n_s is the electron concentration at the IS interface and n_m is a quasi-concentration which would be calculated at the semiconductor surface if the metal fermi level was used instead of the electron quasi-fermi level. N_c is the effective density of states in the conduction band. The terms in Equation (A14) are defined similarly. The above formulation holds if

$$n_s, n_m \leq N_c \quad \text{and} \quad p_s, p_m \leq N_v \quad \text{(A15)}$$

The above derivation is valid for <100> silicon. By deriving Equation (A3) in a more basic fashion, it will be shown that similar results hold for <111> silicon. Figure A2 shows the overlapping of the "shadows" (Section 2.3) of the metal and semiconductor constant energy surfaces for energies near the conduction band edge of silicon. The situation for both <100> and <111> is shown.

For <100> silicon, it is possible to derive Equation (A3) directly from Equation (2.16) in the text. When the maximum contribution to the tunnel current comes from electrons with energies just above the conduction band edge of silicon, the following approximation can be made

$$\int dS e^{-\eta} \approx \text{area of shadow} \times e^{-\eta}\big|_{E,E_T=0} \quad \text{(A16)}$$

Figure A2: Schematic showing the "shadows" of the constant energy surfaces for energies near the conduction band edge of silicon. A "shadow" of a constant energy surface is defined as its projection in wavenumber space onto a plane parallel to the tunneling barrier. The dashed curve represents the metal "shadow", the solid curves represent the semiconductor "shadow", and the cross hatched regions represent their overlap. The case for two different substrate orientations are shown (a) <100> silicon, and (b) <111> silicon.

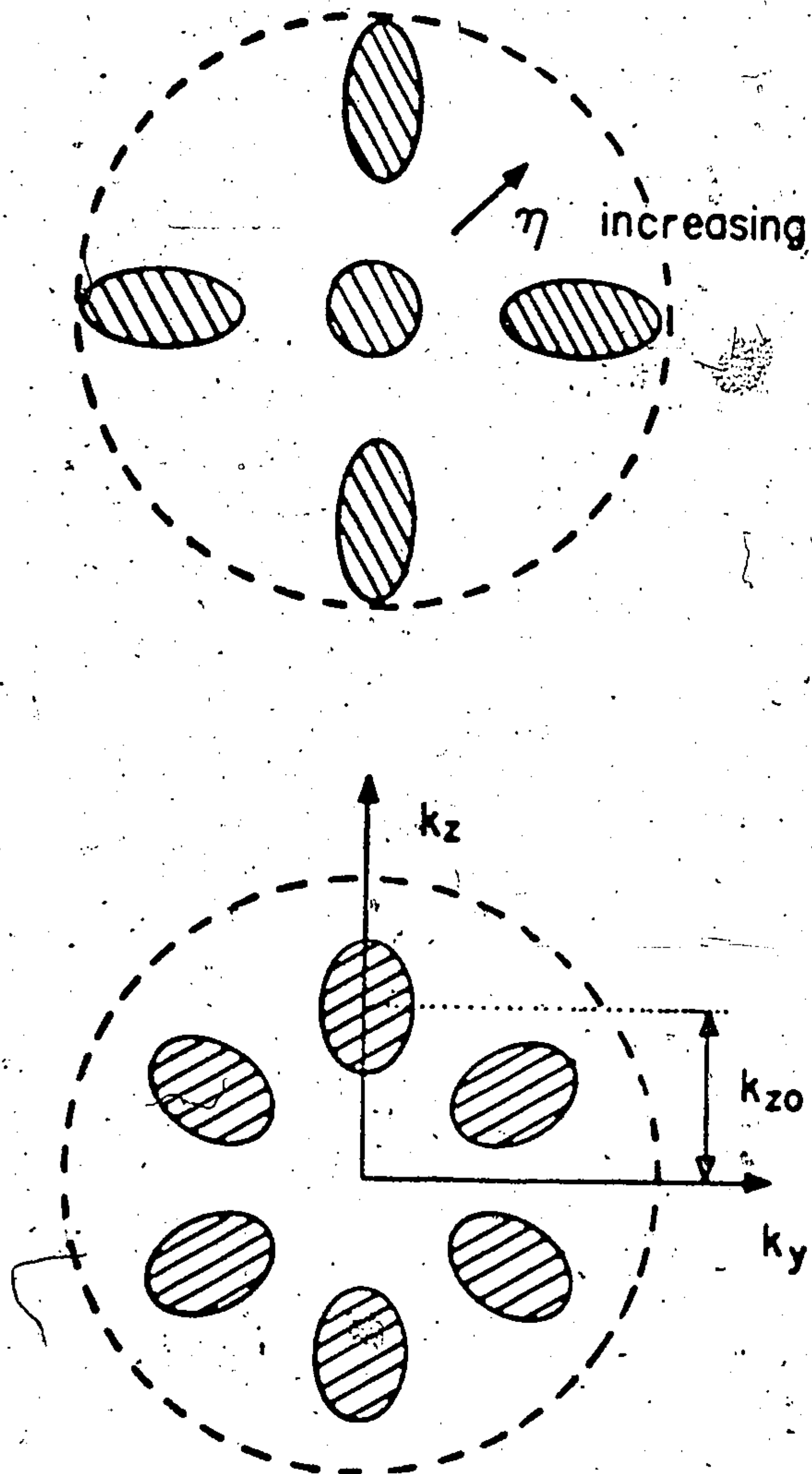


FIGURE A2

since the effective shadow area (near $k_T = 0$) is small and η is essentially constant across it.

Thus

$$\int dSe^{-\eta} = \pi k_{T_{max}}^2 e^{-\eta}|_{E,E_T=0} = \frac{8\pi^3 m_{Ts}}{h^3} Ee^{-\eta_o} \qquad (A17)$$

Substituting this into Equation (2.16) gives Equation (A3).

The same procedure can be followed for <111> silicon. In this case

$$\int dSe^{-\eta} = 6\pi k_{y_{max}} e^{-\eta}|_{E,k_y=0;k_z=k_{zo}} = \frac{48\pi^3 \cdot \sqrt{m_{Ty} m_{Tz}}}{h^3} Ee^{-\eta_1} \qquad (A18)$$

Therefore, equations similar to (A5), (A9), (A13), and (A14) can be derived under the same conditions for <111> silicon. The difference is that in the expressions for J_{CT}, m_{Ts} is replaced by $6 \cdot \sqrt{m_{Ty} m_{Tz}}$ and P_n is the transmission probability evaluated for an electron with the correct component of transverse momentum (Equation (A18)).

APPENDIX B

EFFICIENT NUMERICAL SOLUTION OF THE TRANSMISSION LINE EQUIVALENT CIRCUIT MODEL OF A SEMICONDUCTOR

In this appendix, a numerical algorithm developed during this work is described which allows the transmission line equivalent circuit model of a semiconductor to be solved very efficiently. Initially its application to obtaining small signal AC solutions will be described followed by a description of the small modification required for its application to the DC and transient analyses.

Figure 2.2 shows the small signal equivalent circuit for a semiconductor device with arbitrary doping profile. It has been assumed that the device can be modelled as one-dimensional and that its recombination-generation characteristics can be described by Shockley-Read-Hall (SRH) processes. The node, v_t, can be eliminated to give the small signal AC circuit shown in Figure B1 at the k'th section [46,47]. Applying Kirchoff's law in this section gives

$$[Y_k][v_n^+ \; v_i^+ \; v_p^+ \; v_n^o \; v_i^o \; v_p^o \; v_n^- \; v_i^- \; v_p^-]_k^T = [0 \; 0 \; 0]^T \tag{B1}$$

where the elements of $[Y_k]$ can be readily obtained from Figure B1. Assume there is also a relationship for the k'th section,

$$[M_{k-1}][v_n^o \; v_i^o \; v_p^o \; v_n^- \; v_i^- \; v_p^-]_k^T = [0 \; 0 \; 0]^T \tag{B2}$$

Combining (B1) and (B2) gives

Figure B1: A section of the semiconductor small signal transmission line after the elimination of the trapping node, v_t.

Figure B2: Conversion of the $[Y_k]$ matrix to the $[M_k]$ matrix in three stages of Gaussian elimination. Also formed in the process is the $[A_k]$ matrix.

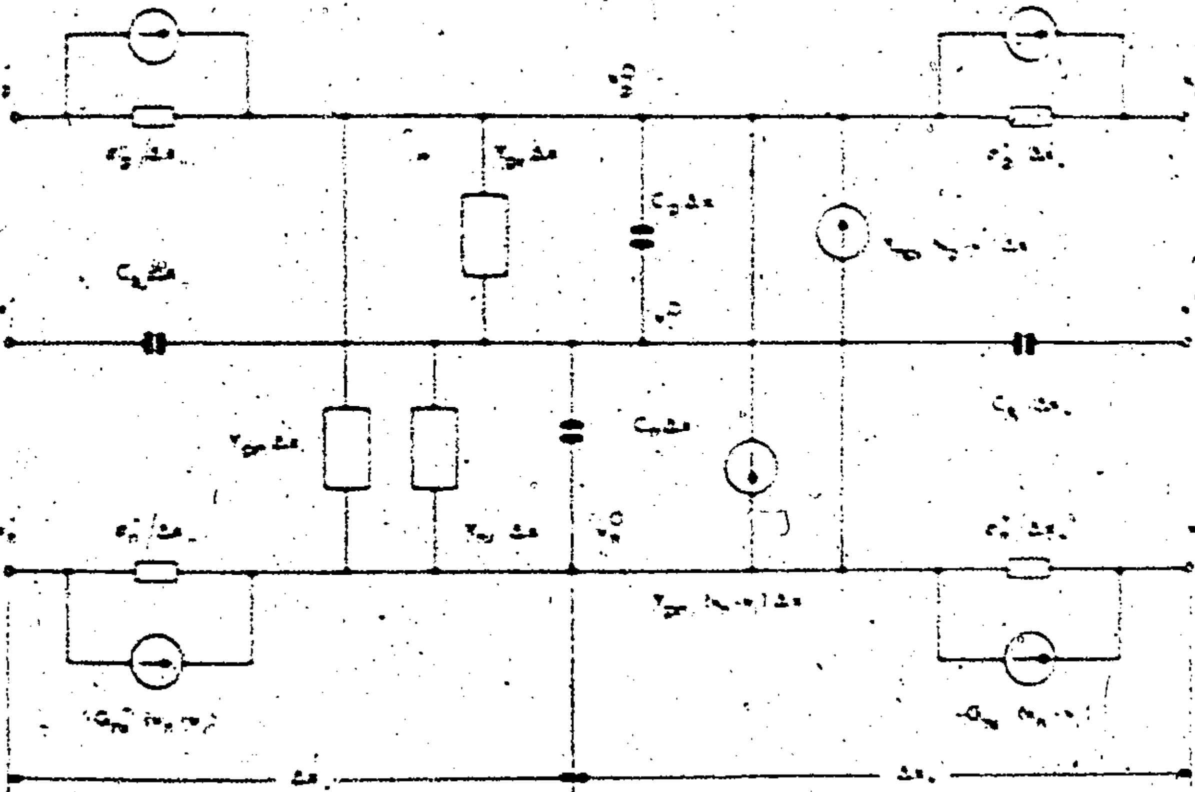
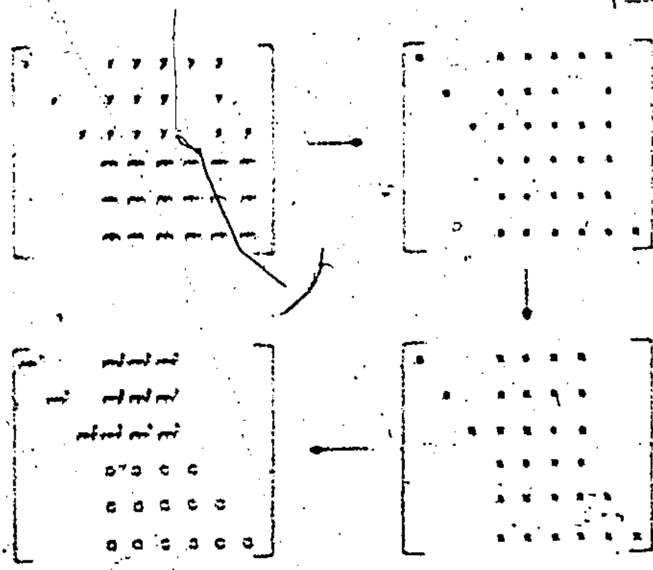


FIGURE B1



a FIGURE B2 b

$$[U][v_n^+ \; v_i^+ \; v_p^+ \; v_n^o \; v_i^o \; v_p^o \; v_n^- \; v_i^- \; v_p^-]_k^T = [0\;0\;0\;\;0\;0\;0]^T \tag{B3}$$

where the non-zero entries of [U] (denoted by a letter indicating whether they came from the "M" or "Y" matrix) are shown in Figure B2. After the three stages of Gaussian elimination indicated in this figure, the following expression is obtained

$$[U'][v_n^+ \; v_i^+ \; v_p^+ \; v_n^o \; v_i^o \; v_p^o \; v_n^- \; v_i^- \; v_p^-]_k^T = [0\;0\;0]^T \tag{B4}$$

where the non-zero entries of [U'] are also shown in Figure B2. From the form of [U'], it can be seen that the relationship

$$[M_k][v_n^+ \; v_i^+ \; v_p^+ \; v_n^o \; v_i^o \; v_p^o]_k^T = [0\;0\;0]^T \tag{B5}$$

can be extracted.

Thus the matrix $[M_k]$ can be derived from the matrix $[Y_k]$ if $[M_{k-1}]$ is known. (Note that during the process, the matrix $[A_{k-1}]$ indicated in Figure B2 is also formed. This can be discarded in the small signal AC case but is stored for the DC and the transient cases).

The first section which is in the vicinity of a contact will now be treated. It will be assumed that variations in the current flow across the contact and in the potential across the contact region can be expressed in terms of the conditions prevailing in the semiconductor immediately adjacent to the contact. This assumption is valid for "ideal ohmic" contacts, saturation velocity limited contacts, for Schottky barrier diodes (provided tunneling through the space charge region of the barrier is small), and for the non-degenerate MIS tunnel diodes currently of interest. Contacts not satisfying the above conditions can be incorporated but these complicate

the solution algorithm in the contact region.

The effective form of the $[Y_k]$ relationship with the above assumption is

$$\begin{bmatrix} & \vert y_{17} \\ Y_1' & \vert y_{27} \\ & \vert y_{37} \end{bmatrix} [v_n^+ \; v_i^+ \; v_p^+ \; v_n^o \; v_i^o \; v_p^o \; v_{c1}]_1^T = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad (B6)$$

i.e.

$$[Y_1'][v_n^+ \; v_i^+ \; v_p^+ \; v_n^o \; v_i^o \; v_p^o]_1^T = -v_{c1}[y_{17} y_{27} y_{37}]^T \quad (B7)$$

If v_{c1} is selected as the voltage reference (i.e. $v_{c1} = 0$), Equation (B7) shows that $[Y_1']$ can be identified as $[M_1]$. This allows $[M_2]$, $[M_3]$, ..., $[M_{N-1}]$ to be found.

An "ideal ohmic" contact can be modelled by shorting the transmission line [26] to v_{c1} (i.e. at section 1, $v_{c1} = v_n^o = v_i^o = v_p^o = 0$). This gives the following trival form to $[M_1]$ in this case

$$[M_1] = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (B8)$$

In the general case, at the other contact, the $[Y_k]$ relationship can be expressed in a form similar to that of Equation (B7).

$$[Y_N'][v_n^o \; v_i^o \; v_p^o \; v_n^- \; v_i^- \; v_p^-]_N^T = -v_{c2}[G] \quad (B9)$$

But, the following relationship is already available

$$[M_{N-1}][v_n^o \; v_i^o \; v_p^o \; v_n^- \; v_i^- \; v_p^-]_N^T = [0 \; 0 \; 0]^T \quad (B10)$$

These can be combined in a manner similar to that illustrated in Figure B2 and the combined matrix reduced in the way shown in this figure ([G] must also be included into the Gaussian reduction process since it is non-zero). After 5 stages of reduction all the potentials in the contact region can be calculated once v_{c2} is specified. The current flow can then be computed in this region and the device small-signal admittance can be found.

This method of solving the AC equivalent circuit takes more advantage of its specialized form than does the "transmission matrix" method [26,101] and the "cofactor" method [25] used for obtaining AC solutions in earlier work. It requires fewer than one third the number of multiplications [56]. As is shown in Chapter 2, the transmission line equivalent circuit can also be used for the DC and the small signal transient analyses. In these cases, the circuit has the same topology as Figure B1 but has "error" sources introduced at the nodes. Equation (B1) takes the form

$$[Y_k][v_n^+ \; v_i^+ \; v_p^+ \; v_n^o \; v_i^o \; v_p^o \; v_n^- \; v_i^- \; v_p^-]_k^T = [E] \qquad (B11)$$

where the column matrix [E] corrsponds to the error sources. The algorithm is the same as for the AC case except that the column matrix [E] must be operated on during the Gaussian reduction stage. Proceeding from one contact, the algorithm works through the semiconductor to the other. In this case the elements of the matrix $[A_k]$ (and the additional elements corresponding to [E]) are stored. After calculating the potentials in the second contact region, these stored elements are used to calculate

the potential distribution throughout the device working from this second contact back to the first.

The differences between the algorithms as described above are small. In actual implementation the differences are larger. The AC analysis requires the use of complex arithmetic or its equivalent while the DC and the transient do not. Also, AC analysis requires only small computer storage while the DC and the transient require at least 18 N storage locations, where N is the number of sections into which the semiconductor is subdivided.

APPENDIX C

EXPRESSIONS FOR THE CAPACITANCE OF MIS TUNNEL DIODES

C.1 Minority Carrier Diodes

In the semiconductor limited regime of minority carrier MIS tunnel diodes, it is shown in Chapter 6 that both the minority carrier fermi level and the semiconductor bands are clamped with respect to the metal fermi level. This means that changes with bias in the charge stored in the inversion region at the semiconductor-insulator interface are small. A change in applied bias is absorbed mainly as a change in the width of the depletion region behind the interfacial inversion layer. If the assumption is made that all the bias change is absorbed in changing the width of the depletion region, a simple analytical expression describing the device capacitance-voltage characteristics can be found.

The space charge, electric field, and electrostatic potential distributions are shown schematically in Figure C1 for thermal equilibrium conditions and when a negative voltage is applied to the p type semiconductor region. An abrupt transition has been assumed at the edge of the transition region, as in the depletion approximation. Integrating Poisson's Equation twice as indicated in this diagram gives

$$-V_a = \psi_s - \psi_{so} = \frac{q}{\epsilon} N_A W d_o + \frac{1}{2} \frac{q}{\epsilon} N_A W^2 \qquad \text{(C1)}$$

i.e.

Figure C1: The assumed space charge density (ρ), the corresponding electric field strength (F), and the electrostatic potential (ψ) as a function of distance into the semiconductor from the interface. ψ is obtained by integrating F, and F is obtained by integrating ρ. The distributions at zero bias and at a small reverse bias are shown.

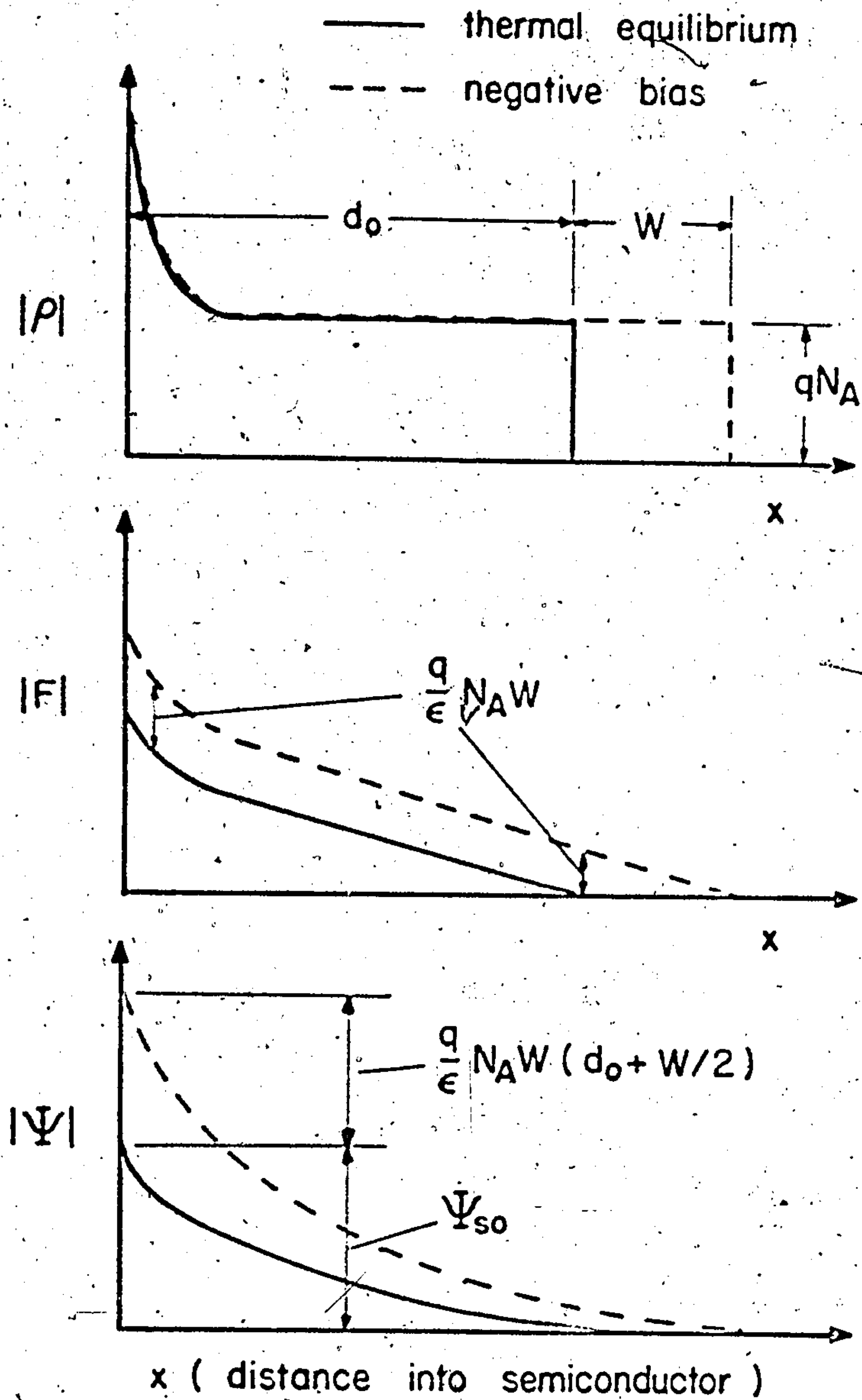


FIGURE C1

$$W = \sqrt{d_o^2 - 2\frac{\epsilon V_a}{N_A}} - d_o \qquad (C2)$$

$$C = \frac{dQ}{dV} = -qAN_A \frac{dW}{dV} \qquad (C3)$$

i.e.

$$\frac{C}{A} = \left[\frac{2}{\epsilon q N_A}(V_c - V_a)\right]^{-1/2} \qquad (C4)$$

where

$$V_c = q\frac{N_A}{2\epsilon} d_o^2 \qquad (C5)$$

d_o is the distance of edge of depletion region from the semiconductor-insulator interface at zero bias. If ψ_{so} is less than ψ_{inv} where ψ_{inv} is the surface potential which puts the surface into strong inversion ($n > N_A$) at thermal equilibrium, depletion theory holds and d_o can be expressed as

$$d_o = \sqrt{\frac{2\epsilon\psi_{so}}{qN_A}} \qquad (C6)$$

giving

$$V_c = \psi_{so} \quad \text{for} \quad \psi_{so} < \psi_{inv} \qquad (C7)$$

where

$$\psi_{inv} = \frac{2kT}{q}\ln(N_A/n_i) \qquad (C8)$$

However for $\psi_{so} > \psi_{inv}$, d_o increases much more slowly with increasing ψ_{so}, tending to saturate at a value approximated by

$$d_o = \sqrt{\frac{2\epsilon\psi_{inv}}{qN_A}} \qquad (C9)$$

giving

$$V_c = \psi_{inv} \quad \text{(C10)}$$

Therefore the intercept of the $1/C^2$ versus V curve with the voltage axis (V_c) is equal to ψ_{so} only when $\psi_{so} < \psi_{inv}$. For larger values of ψ_{so}, V_c increases monotonically with ψ_{so} but is smaller than this quantity.

The capacitance given by Equation (C4) is the total capacitance of the device in the semiconductor-limited regime since the insulator capacitance in series with it is much larger. Due to the clamping effects, the charge in surface-states is virtually constant in this regime (Chapter 6) and these states do not contribute significantly to the device capacitance at these biases.

C.2 Majority Carrier Diodes

For majority carrier diodes with very thin insulator regions, the inversion layer does not build up at the IS interface without external stimulation. Under reverse bias to a p type substrate, the interface changes from accumulation to depletion and remains in depletion. All voltage changes are absorbed primarily across this depletion region. As in the usual depletion approach, this gives

$$\frac{C}{A} = [\frac{2}{\epsilon Q N_A} (V_c - V_a)]^{-1/2} \quad \text{(C11)}$$

where V_c can be identified as the value of applied voltage where the depletion region has zero width, i.e. the flat band voltage. This gives (see Figure 2.1 for a definition of the symbols)

$$V_c = \phi_{si} - \phi_{mi} + \xi \qquad \text{(C12)}$$

In view of the above results, Equation (C7) will be developed further. For $\psi_{so} < \psi_{inv}$, the IS interface is not strongly inverted at zero bias, the surface fields and consequently V_{ins} are small. Equation (2.22) gives

$$\psi_{so} = \phi_{si} - \phi_{mi} + \xi \qquad \text{(C13)}$$

Thus Equation (C12) is the general form for V_c if the value of ψ_{so} is less than ψ_{inv}, i.e. if the surface is not strongly inverted at zero bias.

APPENDIX D

EXPRESSION FOR THE ELECTRIC FIELD AT THE IS INTERFACE

In this appendix, an expression is derived for the electric field strength, F ($= -dv_I/dx$), evaluated at the semiconductor surface as a function of the concentration of minority carriers at this interface. The derivation will be for n type semiconductor.

For non-degenerately doped n type semiconductor material in the bias range of interest, donor impurities are completely ionized. Poisson's Equation (2.1) can be written as

$$dF = \frac{q}{\epsilon}(p - n + N_D)dx \tag{D1}$$

Multiplying by F ($= -dv_I/dx$, where v_I is the electrostatic potential) gives

$$FdF = \frac{q}{\epsilon}(n - p - N_D)dv_I \tag{D2}$$

The current density equations (2.8) and (2.9) have the form

$$J_n = -q\mu_n n \frac{dv_N}{dx} \tag{D3}$$

$$J_p = -q\mu_p p \frac{dv_P}{dx} \tag{D4}$$

where $v_{N,P}$ are quasi-fermi potentials. Substituting these relationships into Equation (D1) gives

$$FdF = \frac{q}{\varepsilon}\left[n(dv_I - dv_N) - p(dv_I - dv_P) - N_D dv_I - \left(\frac{J_n}{q\mu_n} - \frac{J_p}{q\mu_p}\right)dx\right] \quad (D5)$$

n and p can be expressed in the following form (see Equations (2.5) - (2.7)) assuming parabolic densities of states in the conduction and valence bands

$$n = N_c F_{1/2}\left[\frac{q}{kT}(v_I - v_N - V_{CI})\right] \quad (D6)$$

$$p = N_v F_{1/2}\left[\frac{q}{kT}(v_P - v_I - V_{IV})\right] \quad (D7)$$

Integrating (D5) between the ohmic back contact (o) and the semiconductor insulator interface (s) gives

$$\frac{F^2}{2}\Big|_o^s = \frac{kT}{\varepsilon}\{N_c F_{3/2}\left[\frac{q}{kT}(v_I - v_N - V_{CI})\right]\Big|_o^s$$

$$+ N_v F_{3/2}\left[\frac{q}{kT}(v_P - v_I - V_{IV})\right]\Big|_o^s\} \quad (D8)$$

$$- \frac{q}{\varepsilon} N_D \psi_s - \int_{x_o}^{x_s}\left(\frac{J_n}{\varepsilon\mu_n} - \frac{J_p}{\varepsilon\mu_p}\right)dx$$

For $p_s \ll N_v$, the function $F_{3/2}$ can be replaced by $F_{1/2}$ to a very good approximation (see Figure A1)

$$\frac{F_s^2 - F_o^2}{2} = \frac{kT}{\varepsilon}[(n_s - n_o) + (p_s - p_o)] - \frac{q}{\varepsilon} N_D \psi_s$$

$$- \int_{x_o}^{x_s}\left(\frac{J_n}{\varepsilon\mu_n} - \frac{J_p}{\varepsilon\mu_p}\right)dx \quad (D9)$$

All but the most significant terms will now be eliminated. A quick calculation reveals that the current integral is negligible provided the diode current is less than 10^4 amps/m^2 with typical silicon parameters. Similarly F_o and p_o can be neglected.

$$|F_s| = \sqrt{2}\left[\frac{kT}{\epsilon} p_s + \frac{q}{\epsilon} N_D\left(|v_s| - \frac{kT}{q}\right)\right]^{1/2} \quad \text{(D10)}$$

i.e. for $N_V \gg p_s \gg N_D$

$$F_s = -\left(\frac{2kT}{\epsilon} p_s\right)^{1/2} \quad \text{(D11)}$$

REFERENCES

1. S. M. Sze, 'Physics of Semiconductor Devices', Wiley, New York (1969).

   a) Ibid, p.496; b) Ibid, p.65; c) Ibid, p.433; d) Ibid, p.471;

   e) Ibid, p.675; f) Ibid, p.102; g) Ibid, p.650.

2. P. V. Gray, Phys. Rev. 140, A179 (1965).
3. L. Esaki and P. J. Stiles, Phys. Rev. Letters 16, 1108 (1966).
4. L. Esaki, Phys. Rev. 109, 603 (1958).
5. I. Giaever, Phys. Rev. Letters 5, 147 (1960); 5, 464 (1960).
6. B. D. Josephson, Phys. Letters 1, 251 (1962).
7. L. L. Chang, P. J. Stiles, and L. Esaki, J. Appl. Phys. 38, 4440 (1967).
8. L. L. Chang, J. Appl. Phys. 39, 1455 (1968).
9. V. A. K. Temple and J. Shewchun, Solid-St. Electron., accepted for publication; to appear in 17 (1974).
10. P. J. Stiles, L. Esaki, and J. F. Schooley, Phys. Letters 23, 206 (1966).
11. R. A. Clarke, Ph. D. thesis, McMaster University (1972).
12. For example, E. L. Wolf, Phys. Rev. Letters 20, 204 (1968);

    R. B. Laibowitz, Appl. Phys. Letters 13, 221 (1968).

13. W. E. Dahlke and S. M. Sze, Solid-St. Electron. 10, 865 (1967).
14. J. Shewchun, A. Waxman, and G. Warfield, Solid-St. Electron. 10, 1165 (1967).

    A. Waxman, J. Shewchun, and G. W. Warfield, Solid-St. Electron. 10, 1187 (1967).

15. R. A. Clarke and J. Shewchun, Solid-St. Electron. 14, 957 (1971).

16. S. Kar and W. E. Dahlke, Appl. Phys. Letters 18, 401 (1971).

17. S. Kar and W. E. Dahlke, Solid-St. Electron. 15, 221 (1972); 15, 869 (1972).

18. H. C. Card and E. H. Rhoderick, J. Phys. D: Appl. Phys. 4, 1589 (1971).

19. H. C. Card and E. H. Rhoderick, Solid-St. Electron. 15, 993 (1972).

20. H. C. Card and E. H. Rhoderick, Solid-St. Electron. 16, 365 (1973).

21. C. T. Sah, Proc. IEEE. 55, 654 (1967).

22. C. T. Sah, Solid-St. Electron. 13, 1547 (1970).

23. C. T. Sah, Phys. Status Solidi (a) 7, 541 (1971).

24. C. T. Sah, R. F. Pierret, and A. B. Tole, Solid-St. Electron 12, 681 (1969).

25. V. A. K. Temple and J. Shewchun, Solid-St. Electron. 16, 93 (1973).

26. C. R. Smiley, L. D. Yau, and C. T. Sah, Solid-St. Electron. 16, 895 (1973).

27. M. A. Green and J. Shewchun, 'Application of the Small Signal Transmission Line Equivalent Circuit Model to the AC, DC, and Transient Analysis of Semiconductor Devices', Solid-St. Electron., accepted for publication.

28. J. Shewchun and P. A. Clarke, Solid-St. Electron. 16, 213 (1973).

29. M. A. Green and J. Shewchun, 'Current Multiplication in MIS Tunnel Diodes', Solid-St. Electron. 17, 349 (1974).

30. M. A. Green, F. D. King, and J. Shewchun, 'Minority Carrier MIS Tunnel Diodes and their Application to Electron - and Photo - Voltaic Energy Conversion; Part 1: Theory', Solid-St. Electron., accepted for publication; to appear in 17 (1974).

31. J. Shewchun, M. A. Green, and F. D. King, 'Minority Carrier MIS Tunnel Diodes and their Application to Electron - and Photo Voltaic Energy Conversion; Part II: Experiment', Solid-St. Electron., accepted for publication, to appear in 17, (1974).

32. J. Shewchun, R. Keefe, M. A. Green, F. D. King, W. Garland (unpublished); Atomic Energy of Canada, Report No. 5421 (1973).

33. C. B. Duke, 'Tunneling in Solids', Academic Press, New York (1969) a) Ibid, p. 28; b) Ibid, p. 53.

34. W. A. Harrison, Phys. Rev. 123, 85 (1961).

35. L. B. Freeman and W. E. Dahlke, Solid-St. Electron. 13, 1483 (1970).

36. M. A. Green and J. Shewchun, 'The General Transmission Line Equivalent Circuit Model for Degenerate and Non-Degenerate Carrier Concentrations in Semiconductors', Solid-St. Electron., accepted for publication, to appear in 17 (1974).

37. A. C. Smith, J. F. Janak, and R. B. Adler, "Electronic Conduction in Solids', Appendix K, McGraw-Hill, New York (1967).

38. W. Shockley and W. T. Read, Phys. Rev. 87, 835 (1952).

39. P. T. Landsberg and T. S. Moss, Proc. Phys. Soc. B69, 661 (1956).

40. E. L. Heasell, Solid-St. Electron. 12, 225 (1969).

41. P. Schnupp, Solid-St. Electron. 10, 785 (1967).

42. P. Schnupp, Phys. Status Solidi 21, 567 (1967).

43. P. Schnupp, Phys. Status Solidi 25, 455 (1968).

44. W. Franz, 'Handbuch der Physik' (edited by S. Flugge), Vol XVII, p. 155, Springer, Berlin (1956).

45. R. Stratton, J. Phys. Chem. Solids 23, 1177 (1962).

46. C. T. Sah, Electron. Letters 7, 713 (1971).

47. M. A. Green, V. A. K. Temple, and J. Shewchun, Solid-St. Electron. 15, 1027 (1972.)

48. H. K. Gummel, IEEE Trans. ED - 11, 355 (1964).

49. C. W. Gwyn, D. L. Scharfetter, and J. L. Wirth, IEEE Trans. NS - 14, 153 (1967).

50. V. A. J. van Lint, J. H. Alexander, D. K. Nichols, and P. R. Ward, IEEE Trans. NS - 14, 170 (1967).

51. A. de Mari, Solid-St. Electron. 11, 1021 (1968).

52. B. V. Gokhale, IEEE Trans. ED - 17, 594 (1970).

53. T. Ohtsuki and K. Kani, IEEE Trans. CT - 17, 26 (1970).

54. R. L. Pritchard, Proc. IRE. 42, 786 (1954).

55. K. Kani and A. Yokota, IEEE Trans. ED - 19, 1028 (1972).

56. M. A. Green and J. Shewchun, Electron. Letters 9, 474 (1973).

57. R. S. Varga, 'Matrix Iterative Analysis', Prentice-Hall, New Jersey (1962).

58. V. S. Fomenko, 'Handbook of Thermionic Properties', edited by G. V. Samsonov, Plenum Press, New York (1966).

59. E. H. Nicollian and A. Goetzberger, Bell Syst. Tech. J. 46, 1055 (1967).

60. Z. Hurych, Solid-St. Electron. 13, 683 (1970).

61. R. Stratton, Phys. Rev. 136, A 837 (1964).

62. A. Y. C. Yu and E. H. Snow, J. Appl. Phys. 39, 3008 (1968).

63. E. H. Nicollian and A. Goetzberger, IEEE Trans. ED - 12, 108 (1965).

64. T. M. Buck and F. S. McKim, J. Electrochem. Soc. 105, 709 (1958).

65. R. Williams, Phys. Rev. 140, A 569 (1965).

66. F. H. Hielscher and H. M. Preier, Solid-St. Electron. 12, 527 (1969).

67. W. S. Boyle and G. E. Smith, Bell Syst. Tech. J. 49, 587 (1970).

68. W. Schultz, Zeitschrift fur Physik 138, 598 (1954).

69. C. R. Crowell and S. M. Sze, Solid-St. Electron. 9, 1035 (1966)

70. M. A. Green and J. Shewchun, Solid-St. Electron 16, 1141 (1973).

71. R. A. Clarke, M. A. Green and J. Shewchun, J. Appl. Phys. 45, 1442 (1974).

72. T. Misawa, J. Phys. Soc. Japan 12, 883 (1957).

73. W. Shockley, 'Electrons and Holes in Semiconductors', Van Nostrand, New Jersey (1950), p. 316.

74. E. H. Nicollian and A. Goetzberger, Bell Syst. Tech. J. 46, 1055 (1967).

75. A. M. Cowley and S. M. Sze, J. Appl. Phys. 36, 3212 (1965).

76. A. S. Grove and D. J. Fitzgerald, Solid-St. Electron. 9, 783 (1966).

77. L. L. Chang, L. Esaki, and F. Jona, Appl. Phys. Letters 9, 21 (1966).

78. A. Nwachuku and M. Kahn, Appl. Phys. Letters 12, 163 (1968).

79. J. Shewchun and A. Waxman, Rev. Scient. Instr. 37, 1195 (1966).

80. R. C. Jaklevic, D. K. Donald, J. Lambe, and W. C. Vassell, Appl. Phys. Letters 2, 7 (1963).

81. W. Kern and D. A. Puotinen, RCA Rev. 31, 187 (1970).

82. K. H. Zaininger and A. G. Revesz, RCA Rev. 25, 85 (1964).

83. S. R. Lederhandler and L. J. Giacoletto, Proc. IRE 43, 477 (1955).

84. W. R. Hunter, D. H. Eaton, and C. T. Sah, Appl. Phys. Letters 17, 211 (1970).

85. H. Katto and C. T. Sah, Phys. Stat. Sol. (a) 13, 417 (1972).

86. T. P. Ma and R. C. Barker, J. Appl. Phys. 45, 317 (1974).

87. H. R. Phillip and E. A. Taft, Phys. Rev. Letters 8, 13 (1962).

88. O. S. Heavens, 'Optical Properties of Thin Solid Films', Dover, New York (1965).

89. M. Klein, IEEE Trans. ED - 13, 788 (1966).

90. J. C. Hancock, 'An Introduction to the Principles of Communication Theory', McGraw-Hill, New York (1961).

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

92. J. Lindmayer and C. Y. Wrigley, 'Fundamentals of Semiconductor Devices', Van Nostrand, New Jersey (1965).

93. M. Wolf, Energy Conversion 11, 63 (1971).

94. ~~M. Wolf, Proc. IRE. 48, 1246 (1960).~~

95. H. Flicker, J.J. Loferski, and T. S. Elleman, IEEE Trans. ED - 11, 2 (1964).

96. P. Rappaport, J.J. Loferski, and E. G. Linder, R.C.A. Rev. 17, 100 (1956).

97. L. C. Olsen, Energy Conversion 13, 117 (1973)

98. D. L. Scharfetter, Solid-St. Electron. 8, 299 (1965).

99. V. Heine, Phys. Rev. 138, A 1689 (1965).

100. J. S. Blakemore, 'Semiconductor Statistics', Pergamon, New York (1962).

101. F. Hennig and C. T. Sah, Solid-St. Electron. 16, (1973).

102 M. A. Green, 'Computer Simulation of the MIS Tunnel Diode', Tech. Report, Department of Engineering Physics, McMaster University, Canada.