INVESTIGATIONS OF SINGLE-CRYSTAL SILICIDE/SILICON STRUCTURES

Thesis by

Robert Joseph Hauenstein

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1987

(Submitted December 10, 1986)

Acknowledgements

I gratefully acknowledge the following people, who, during my years as a graduate student at Caltech, made significant contributions, either directly or indirectly, to this thesis, to my professional development, and to my personal enrichment. First and foremost, I would like to thank my thesis advisor, Prof. T. C. McGill, for his thorough and critical reading of this thesis, and for providing me with unique opportunities for professional development: excellent research facilities; relevant research topics; astute guidance and advice; and a range of research challenges, from theoretical to experimental, from immediate to extended, from well-defined to many-faceted, that most students never have the good fortune to encounter. I have profited greatly from his insights, on matters both scientific and otherwise, and he has, more than anyone else, helped me to understand myself. In addition, I am indebted to Drs. L. J. Schowalter, B. D. Hunt, E. L. Hall, N. Lewis, and S. Hashimoto, for growing and characterizing the silicide samples which were studied in this thesis. I would particularly like to express my gratitude to Mr. O. J. Marsh, whose interest and enthusiasm for Si MBE I have found to be contagious.

There are many others who have each contributed in their own way to this thesis. I particularly wish to acknowledge and thank Vere Snell, our group's administrative assistant, whose talents are largely responsible for making our research group run ever-smoothly, and for making the group the most professional operation I have had the privilege of being involved with. I also thank Carol McCollum for her excellent secretarial help, and for making sure that our supplies of paper, copier cartridges, magnetic tapes, and a host of other materials (of great significance in thesis writing), were always well-stocked.

The fabrication of the Si/silicide system discussed in Chapter 4 would have been impossible were it not for the many contributions of Caltech's Central En-

gineering Services (CES) staff. In particular, I would like to thank Mr. Norman Keidel, manager of CES, for his organization's high standard of professionalism, commitment to customer service, and willingness to take on things that most machine shops would quickly dismiss as "impossible" or "impractical." I would also like to acknowledge Mr. Louis Johnson, for his meticulous and constructively critical examination of every aspect of my designs, almost always resulting in their improvement, and also, for his skillful and thorough leak-checking of the many vacuum components at various stages of their fabrication. I am particularly grateful to Mr. Leif Larsson, first, for his several design contributions, most notably, to the shutter mechanism, and secondly, for his skillful and well-thought-out machining of some extremely difficult refractory metal and stainless steel components. Critically important contributions to be acknowledged (and admired) were also made by Mr. Ralph Ortega for his superlative welding of the many elaborate cryoshrouds and fittings used in our system.

Lastly, I would like to thank my friends and fellow students, past and present, who, along with me, are developing professionally and personally, to a large extent, through our mutual interaction. I acknowledge the following former students for their influence: Drs. R. T. Collins, R. M. Feenstra, S. R. Hetzler, C. Mailhiot, and A. Prabhakar. I particularly enjoyed and now miss many interesting and often lively discussions of physics over dinner with Drs. T. E. Schlesinger and A. Zur. I have further profited from useful discussions with the following people: A. R. Bonnefoi, D. H. Chow, M. K. Jackson, T. K. Woodward, G. Y. Wu, and Y. Rajakarunanayake. I acknowledge J.-M. Langlois for assistance with the UHV apparatus and the initial silicide growths. Finally, I particularly value and appreciate the scientific discussions, exchange of advice, and friendship that I have shared with Wesley Boudville, Hong-Man Chan, Matthew Johnson, John Klemic, and Richard Miles, and would like to thank Richard in particular for his

adroit translation of Chapter 1 of this thesis from "American" into "English."

Abstract

This thesis concerns the topic of molecular beam epitaxial (MBE) growth on Si, and presents work related to different aspects of this central topic. We have experimentally investigated the *electronic properties* of structurally near-perfect, metal-semiconductor interfaces involving metallic, transition-metal silicide films, grown epitaxially on Si by MBE. In particular, measurements of the Schottky barrier height of single-crystal, metal-semiconductor structures has been performed. In addition, inelastic electron tunneling spectroscopy has been used to look for elementary excitations of single-crystal, metal-semiconductor interfaces. Other work presented here concerns the building of a "home-made" Si MBE growth apparatus. Lastly, we present theoretical work on a topic unrelated to Si MBE; namely, the calculation of vibrational modes of point-defects in semiconductors.

In Chapter 2, the Schottky barrier heights of high quality, type-A and type-B $NiSi_2/Si$ structures are measured, with the use of photoresponse and forward *I-V* techniques. This work is motivated by a recent controversy over the barrier heights of type-A and type-B structures. Our principal finding is that a *substantial difference* (greater than 0.1 eV) is observed in Schottky barrier height between the type-A and type-B oriented films. For type-A epitaxy, both measurement techniques consistently yield a Schottky barrier height ϕ_{B_n} of 0.62 ± 0.01 eV. For the case of type-B, the *I-V* measurement yields a barrier height of 0.69 eV, in apparent disagreement with the photoresponse value (0.77 eV). Interpretation of the type-B photoresponse data is complicated by an unusual curvature in the plot of $\sqrt{\text{photoresponse}}$ vs. photon energy. It is shown that both the detailed shape of the type-B photoresponse data, and the apparent inconsistency between photoresponse- and *I-V*-determined Schottky barrier height values can be *quantitatively* explained in terms of a phenomenological model, in which we represent the type-B interface as a mixture of regions of high and low Schottky barrier

height (ϕ_{hi} and ϕ_{lo}), electrically in parallel with each other. In terms of this model, the type-B interface is explained as a mixed-barrier structure, having 91% areal coverage of $\phi_{hi} = 0.81\,eV$, and 9% of $\phi_{lo} = 0.64\,eV$.

In Chapter 3, the first electron tunneling spectroscopy experiments on single-crystal, type-A and type-B $NiSi_2$ films are described. For comparison, single-crystal $CoSi_2$ films, and polycrystalline Au films, all grown on similar, degenerate, n-type Si substrates, are also studied. The purpose of this study is to see if crystalline perfection of the *interface* results in additional or enhanced structure in the electron tunnel spectrum, due to rigorous $\mathbf{k}_{\|}$-conservation. This is not observed. Instead, we observe direct evidence of bulk, Si, $\mathbf{k}$-conserving phonons assisting the transport of electrons through the Schottky barrier. Our electron tunneling spectra appear to to be *independent* of the identity, crystalline perfection, and crystallographic orientation of the metal layer. No enhancements or new features in our tunneling spectra due to the single-crystallinity of the silicide films are evident, suggesting that, the observed inelastic transport is dominated by bulklike processes excited within the tunneling barrier rather than by processes intrinsically related to the metal-semiconductor interface.

In Chapter 4, the design and construction of a Si/silicide MBE growth system is presented. The original contribution in this work is the design of the apparatus: both of the overall system, and of its many detailed components. In particular, the design of this system reflects a number of issues which are of serious concern in state-of-the-art, Si MBE: the problem of unintentional p-type doping, the concern over the flaking of Si inside the growth chamber, and the accurate monitoring and control of substrate temperature. Most of the key assemblies, including the substrate heater, load-lock mechanism, sample holder, cryoshrouds, and shutter mechanism were designed and fabricated at Caltech. Our main result is the successful operation of this system. Preliminary silicide

growth attempts demonstrate the ability of the system to perform closed-loop, rate-controlled depositions onto a heated Si substrate.

Finally, Chapter 5 involves the theoretical calculation of vibrational modes of point-defects in semiconductors. An understanding of such modes is useful in providing structural information about point-defects. Often, vibrational modes are observable by optical techniques. In our work, three distinct defect systems are investigated: substitutional oxygen on a phosphorus site in GaP ($GaP:O_P$), substitutional and interstitial iron in Si (Si:Fe), and substitutional carbon on an arsenic site in GaAs ($GaAs:C_{As}$). Our theoretical method involves *modeling* the defect, *e.g.*, in terms of masses and springs. In the calculation, the masses are known, but the spring constants near the defect are unknown, and varied as parameters. This requires a comparison with experimental data. For the $GaP:O_P$ defect, we predict and fit two main defect modes through the adjustment of parameters, and in the process, find *detailed* agreement between theory and experiment for several additional modes, also predicted by our calculation, but not involved in the fit. Our calculations provide the first coherent framework to account for details from many different optical measurements on the $GaP:O_P$ defect, within a single, theoretical description. We have also calculated the low-energy defect modes associated with Fe substitutional and interstitial defects in Si, to determine whether a frequency shift is expected upon the isotopic substitution, $^{54}Fe \rightarrow {}^{56}Fe$. The calculation predicts the existence of a single, low-energy mode, and that a frequency shift should occur, in *disagreement* with an experimental result. Our calculation demonstrates that the observed mode cannot therefore be due to a defect *phonon*. Finally, the effect of the natural isotopic variation of Ga in GaAs on the infrared-active defect modes of the $GaAs:C_{As}$ system is investigated. The calculated results show detailed agreement with infrared absorption data with respect to defect-mode splittings and relative intensities,

and establish conclusively the identity of the impurity substitutional site as an As site.

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

Chapter 2:

Schottky Barrier Height Measurements of Epitaxial $NiSi_2$ on Si,
R. J. Hauenstein, T. E. Schlesinger, T. C. McGill, B. D. Hunt, and L. J. Schowalter, Appl. Phys. Lett. **47**, 853 (1985).

Schottky Barrier Height Measurements of Type-A and Type-B $NiSi_2$ on Si,
R. J. Hauenstein, T. E. Schlesinger, T. C. McGill, B. D. Hunt, and L. J. Schowalter, J. Vac. Sci. Technol. **B 4**, 649 (1986).

Schottky Barrier Height Measurements of Type-A and Type-B $NiSi_2$ Epilayers on Si,
R. J. Hauenstein, T. E. Schlesinger, T. C. McGill, B. D. Hunt, and L. J. Schowalter, J. Vac. Sci. Technol. A **4**, 860 (1986).

Structural Characterization and Schottky Barrier Height Measurements of Epitaxial $NiSi_2$ on Si,
B. D. Hunt, L. J. Schowalter, N. Lewis, E. L. Hall, R. J. Hauenstein, T. E. Schlesinger, and T. C. McGill, *Proc. of the Materials Research Society,* Boston, 1985, **54**, 479 (1986).

Chapter 3:

Tunneling Spectroscopy of Single-Crystal $CoSi_2$ and $NiSi_2$ Epilayers on n-type Si,

R. J. Hauenstein, L. J. Schowalter, B. D. Hunt, O. J. Marsh, and T. C. McGill, *Proc. of the Materials Research Society,* Palo Alto, April 15–18, 1986 (to be published).

Chapter 5:

Local Phonon Mode Calculations for Point Defects in Si,

R. J. Hauenstein, T. C. McGill, and R. M. Feenstra, in *Proceedings of the 163rd Meeting of the Electrochemical Society,* San Francisco, May 8–13, 1983.

Isotope Shifts for the P,Q,R Lines in Indium-doped Silicon,

T. E. Schlesinger, R. J. Hauenstein, R. M. Feenstra, and T. C. McGill, Solid State Commun. **46**, 321 (1983).

Vibrational Modes of Oxygen in GaP Including Nearest-Neighbor Interactions,

R. M. Feenstra, R. J. Hauenstein, and T. C. McGill, Phys. Rev. **B28**, 5793 (1983).

Vibrational Modes of Oxygen in GaP Including Second-Nearest-Neighbor Interactions,

R. J. Hauenstein, T. C. McGill, and R. M. Feenstra, Phys. Rev. **B29**, 1858 (1984).

Vibrational Modes of Oxygen in Gallium Phosphide,

R. M. Feenstra, R. J. Hauenstein, and T. C. McGill, *Proc. of the 13th International Conference on Defects in Semiconductors,* San Diego, August 12–17, 1984.

Contents

Acknowledgements ii

Abstract v

List of Publications ix

1 Introduction 1

1.1 Introduction to Thesis 1

1.2 Overview of Silicon MBE 2

1.2.1 Background 2

1.2.2 Novel Device Prospects with Si MBE 4

1.2.3 Recent Progress 15

1.3 Single-Crystal Silicides on Si 19

1.3.1 An Ideal Metal-Semiconductor Interface 23

1.3.2 Schottky Barrier Height Measurements 28

1.3.3 Electron Tunneling Spectroscopy 36

1.4 Vibrational Modes of Defects 44

1.4.1 Localized and Resonant Modes 45

1.4.2 Relation to Optical Spectra 48

1.5 Outline of Thesis 51

2 Schottky Barrier Height Measurements at Single-Crystal $NiSi_2$/Si Interfaces **59**

2.1 Introduction 59

2.1.1 Background 59

2.1.2 Results of this work 60

2.1.3 Outline of Chapter 61

2.2 Theory of Photoresponse 62

2.3 Experimental 66

2.3.1 Silicide Fabrication and Structural Characterization . . . 66

2.3.2 Electrical characterization 67

2.4 Results 71

2.4.1 Photoresponse Measurement 71

2.4.2 *I-V* Measurement 73

2.5 Double Schottky Contact Model 75

2.5.1 Formalism 75

2.5.2 Fit to Photoresponse Data 77

2.5.3 Prediction for *I-V* Measurement 80

2.6 Discussion 82

2.6.1 Interpretation of Results 82

2.6.2 Possible Physical Origins of Model 83

2.7 Conclusions 84

3 Electron Tunneling Spectroscopy Experiments on Single-Crystal $CoSi_2$ and $NiSi_2$ Epitaxial Films on *n*-type Si **87**

3.1 Introduction 87

3.1.1 Background 87

3.1.2 Results of this work 89

3.1.3 Outline of Chapter 90

3.2 Tunneling in Schottky Barriers 91
3.2.1 Elastic Tunneling: WKB Theory 91
3.2.2 Inelastic Tunneling 92
3.3 Experimental 93
3.3.1 Silicide Epitaxial Growth and Structural Characterization 93
3.3.2 Sample Preparation 96
3.3.3 Derivative Measurements 100
3.4 Results 103
3.4.1 Second Derivative Spectra 103
3.4.2 First Derivative Spectra 108
3.4.3 Lower Substrate Dopings 110
3.5 Discussion 113
3.5.1 Identification of Peaks 113
3.5.2 Mechanism of Inelastic Tunneling 114
3.5.3 Comparison to Au/Si Structures 115
3.5.4 Comparison to WKB Theory 117
3.5.5 Doping Enhancements and Band Tailing 120
3.6 Conclusions 123

4 Design and Construction of a Si MBE Growth Chamber 128
4.1 Introduction 128
4.1.1 Objectives 129
4.1.2 Results of this work 130
4.1.3 Outline of Chapter 131
4.2 Requirements for High Quality Epitaxy 131
4.2.1 Growth Chamber, Substrate, and Sources 131
4.2.2 Growth Conditions 136
4.3 Design Considerations 140

4.3.1 Overall Configuration 140
4.3.2 Substrate Heater 144
4.3.3 Evaporation Sources and Shutter Assembly 150
4.3.4 Load-Lock 151
4.3.5 Cryopanels: LN_2 vs. Water Cooling 155
4.4 Problems Addressed 157
4.4.1 B-doping of Epitaxial Si 157
4.4.2 Si Flaking 158
4.5 Choice and Preparation of Materials for UHV 159
4.5.1 Materials for UHV 159
4.5.2 Precleaning 162
4.5.3 Chemical Processing 163
4.5.4 *In Situ* Lubrication 165
4.6 Preliminary Results 166
4.6.1 System Performance 167
4.6.2 Sample Characterization 169
4.7 Problems and Suggested Remedies 173
4.8 Summary 174

5 Calculation of Vibrational Modes of Point-Defects in Semiconductors **177**
5.1 Introduction 177
5.1.1 Background 177
5.1.2 Results of this work 179
5.1.3 Outline of Chapter 181
5.2 Theoretical Techniques 181
5.2.1 Green's Function Approach 181
5.2.2 Group Theory 184

5.3 Vibrational Modes of Oxygen in GaP 187

5.3.1 Bulk Lattice Dynamics 190

5.3.2 Symmetry Considerations 192

5.3.3 Defect Perturbation 193

5.3.4 Results 195

5.3.5 Discussion 206

5.3.6 Summary of GaP:O_P Results 216

5.4 Vibrational Modes of Fe isotopes in Si 218

5.4.1 Bulk Lattice Dynamics 219

5.4.2 Symmetry Considerations 220

5.4.3 Substitutional vs. Interstitial Defects 221

5.4.4 Defect Matrix 222

5.4.5 Computed Results 223

5.4.6 Comparison to Experimental Spectra 226

5.4.7 Summary of Si:Fe Results 228

5.5 Vibrational Modes of C in GaAs 231

5.5.1 Defect Mode Splitting due to Ga Isotopes 231

5.5.2 Symmetry Considerations 232

5.5.3 Computed Results 233

5.5.4 Comparison to Experimental Spectra 235

5.5.5 Summary of GaAs:C_{As} Results 237

List of Figures

1.1 Metal-Base Transistor 7

1.2 Infrared Schottky Detector 9

1.3 Si_xGe_{1-x} Strained-Layer Epitaxy 13

1.4 Si_xGe_{1-x}/Si MODFET 14

1.5 High Throughput Si MBE 20

1.6 Type-A and Type-B Structures 22

1.7 Schottky Barrier Formation 24

1.8 Metal-Semiconductor Contact With Interface States 26

1.9 Transport Mechanisms in Schottky Barriers 30

1.10 Bias Dependence of Schottky Barrier Lowering 31

1.11 Determination of Barrier Height (I-V Method) 33

1.12 Photoresponse Measurement 34

1.13 Inelastic Tunneling 41

1.14 IETS for an Esaki Diode 43

1.15 Localized vs. Resonant Modes 47

1.16 Coupling Mechanisms to Local Phonons 50

2.1 Geometric Interpretation of Photoresponse Threshold 65

2.2 High-Resolution Lattice Image of Type-A $NiSi_2/Si$ Interface . . . 68

2.3 Experimental Setup for Photoresponse Measurement 70

2.4 Photoresponse of Type-A and Type-B $NiSi_2$/Si Schottky Barriers 72

2.5 Forward I-V Characteristics of Type-A and Type-B $NiSi_2$/Si Schottky Barriers 74

2.6 Double Schottky Contact Model Applied to Type-B Photoresponse 78

2.7 Derivatives of Compound Barrier Photoresponse 79

3.1 Irregular Surface Morphology of $CoSi_2$ Samples 97

3.2 Tunneling Device Structure 98

3.3 Zero-bias Resistance vs. Device Size 101

3.4 Three-point Contact Method 104

3.5 Tunnel Spectra of Single-Crystal Films on Si 105

3.6 Comparison of $CoSi_2$ Tunnel Spectra 107

3.7 dV/dI vs. V for $CoSi_2$ Epilayers on Degenerate n-type Si 109

3.8 Tunnel Spectra of $CoSi_2$ on Si:Sb 111

3.9 Tunnel Spectra of $NiSi_2$ on Si:Sb 112

3.10 Tunnel Spectra of Au Films on Si 116

3.11 Theoretical and Experimental Tunneling Characteristics for $CoSi_2$ Films on Si:As 118

3.12 dV/dI Spectra for Au Films on Si:As 122

4.1 Generic Si MBE System 135

4.2 Instrumentation and Control Electronics 138

4.3 Our Si/Silicide MBE System 141

4.4 Schematic Cross-Section of Our Si/Silicide Growth Chamber . . 143

4.5 Substrate Heater and Sample Carrier—Exploded View 146

4.6 Substrate Heater Attached to Sample Carrier 147

4.7 Top View of Shutter Assembly 152

4.8 Actual Sample Loading Sequence 154

4.9 Closed-Loop Control of Evaporation Rate 168

4.10 Surface Morphology of RDE-Grown $NiSi_2$ Film 171

4.11 Surface Morphology of SPE-Grown $NiSi_2$ Film 172

5.1 Bulk Lattice Dynamics of GaP in the DBA Model 197

5.2 Bulk Green's Functions for GaP 199

5.3 Local Vibrational Modes of the GaP:O_P Defect 202

5.4 Decomposition of A_1 LDOS for the GaP:O_P Defect 203

5.5 LDOS of T_2 Modes of the GaP:O_P Defect 205

5.6 Atomic Displacements in the Low Energy T_2 Mode 213

5.7 Calculated Bulk Phonon Dispersion Curves of Si 224

5.8 Calculated Bulk Phonon DOS of Si 225

5.9 Local Phonon Modes of the Si:Fe Defect 227

5.10 Low Energy Local Modes for Impurities in Si 229

5.11 Local DOS of Low Energy Resonant Modes 230

5.12 Experimental vs. Calculated LVM Band of the GaAs:C_{As} Defect 234

List of Tables

2.1 Schottky barrier height measurements of type-A $NiSi_2$ films on Si 75

2.2 Schottky barrier height measurements of type-B $NiSi_2$ films on Si 76

2.3 Comparison of Schottky barrier heights ϕ_{B0} of single-crystal $NiSi_2$ epilayers on n-type, (111)-Si 81

3.1 Samples grown for the electron tunneling study 95

4.1 Procedure used for degreasing vacuum parts 164

4.2 Procedure for chemical processing of vacuum parts 165

5.1 Energies (meV) of the phonons associated with the O^0 and O^+ defect in GaP 209

5.2 Experimental and theoretical LVM peak frequencies (ν) and relative intensities (I) for the GaAs:C_{As} defect 237

Chapter 1

Introduction

1.1 Introduction to Thesis

Much of the work presented in this thesis is related to Si molecular beam epitaxy (MBE)—the results reported in Chapters 2 and 3 were obtained on MBE-grown silicide films on Si, and Chapter 4 describes the building of a Si MBE growth chamber. The purpose of Chapter 1 is to provide introduction and background material for the topics covered in the later chapters of this thesis. To this end, Chapter 1 has been organized as follows: Section 1.2 presents an overview of Si MBE, with an emphasis on its potential as a device technology. Particular attention is devoted to the prospects for novel device structures, attainable in principle by MBE growth techniques, which may represent the future course of Si-based technology. Section 1.2 ends with a survey of recent progress in Si MBE processing, epitaxial growth, and throughput. Next, in Section 1.3, the nature of the metal-semiconductor interface is discussed. The silicides, $CoSi_2$, and particularly, $NiSi_2$, both of which can be fabricated as single-crystal epitaxial films on Si, provide the most structurally ideal and well characterized metal-semiconductor interfaces that can be fabricated at present. The contents of Section 1.3 provide

an introduction to the investigations into the electrical behavior of single-crystal silicide/Si structures reported in Chapters 2 and 3. Then, Section 1.4 follows with a description of the nature of vibrational modes due to point-defects in solids; this is the subject matter of Chapter 5, and is really a topic of research distinct from the chapters related to Si MBE. Finally, Chapter 1 concludes with Section 1.5 which contains an outline of the work to be reported in Chapters 2 thru 5 of this thesis.

1.2 Overview of Silicon MBE

1.2.1 Background

Molecular beam epitaxy (MBE) is a very clean vacuum deposition technique in which thin films are built up one monolayer at a time by directing one or more beams of atoms (or molecules) from appropriate sources onto an atomically clean, heated substrate. In this way, epitaxial* layers of precisely controlled thickness, composition, and doping can be fabricated with monolayer abruptness. In addition, multilayer epitaxial structures such as superlattices, in which the optical and electronic properties can be tailored through the direct structural control attainable with MBE, are possible.

The semiconductor industry today is based heavily on a technology in which devices are fabricated on high quality epitaxial layers of Si residing on a Si substrate. Within the past several years, the technique of MBE has been successfully applied to the problem of fabricating high quality, epitaxial thin films on Si. Owing largely to the dominant role of Si-based technology in commercial microelectronics, there is ever-increasing interest in the use of MBE for the fabrication of electronic and optical device structures on Si. At present, most existing Si de-

*An *epitaxial* film is one which is crystallographically related to the underlying substrate.

vice technology is based on another epitaxial growth technique, chemical vapor deposition (CVD). For the case of conventional bipolar Si technology, MBE represents a rather costly alternative to CVD. However, because of the desirability in general of reducing characteristic device sizes (to improve performance characteristics such as operating speed, to increase device density, to reduce power dissipation, etc.), MBE holds an advantage, at least in principle, over CVD which may lead to its emergence as a commercially viable technology. This advantage derives from the significant difference in growth temperatures between the two processes: for CVD, growth takes place at a high temperature ($\approx 1050\,^{\circ}$C or more) whereas for MBE, the substrate temperature is substantially lower (600 to $900\,^{\circ}$C).[1] The higher growth temperature of CVD limits the abruptness of doping profiles and interfaces because of diffusion,[1] which places a lower limit of $\sim 1\,\mu$m on the size of CVD-grown device structures. Consequently, MBE may become important commercially when it becomes desirable to fabricate smaller device structures having active regions less than $\sim 1\,\mu$m in depth.

In addition to the possibility of using MBE in the fabrication of scaled-down conventional bipolar devices, MBE also offers the exciting prospect of fabricating a new class of novel device structures. Such novel devices are based on the ability to reliably fabricate thin (as little as a few monolayers), hyperabrupt, heteroepitaxial and multilayered films, relatively free from impurities and structural defects. The physics governing the operation of these novel devices would also be quite different from that of conventional devices. Semiclassical physics, in which a semiconductor electron is viewed as a classical particle with an effective mass, combined with Boltzmann transport theory and the Maxwell equations, is adequate to describe most of the relevant processes in conventional devices, such as thermal excitation over potential barriers, and charge transport by drift and diffusion. However, as device sizes approach $\sim 100\,$Å or less, two important

changes occur which render this semiclassical description inappropriate: (1) the characteristic device size becomes comparable to the De Broglie wavelength of electrons (still assuming an *effective* electron mass); and (2) the characteristic device size becomes comparable to or less than the electron mean-free path. It then becomes necessary to consider the states and transport of electrons quantum mechanically. In this regime, effects such as tunneling and ballistic transport, quantization of spatially confined electronic energy levels, hot electron effects, quantum reflection, and resonant transmission, for example, are now the physics relevant to device operation. It is the prospects for novel device structures grown on Si, governed by just such quantum and ballistic effects, which are the subject of the following section.

1.2.2 Novel Device Prospects with Si MBE

A number of different materials have been fabricated epitaxially on Si by MBE. To date, this list includes metals, semiconductors, and insulators, and there is active interest in each case for potential novel device applications. The epitaxial metals fabricated on Si have been mainly silicide compounds. Some silicides, by virtue of their similar crystal structure and close lattice-constant match to Si, have actually been fabricated as single-crystal epitaxial films, and have been successfully overgrown with crystalline Si.[2,3,4] There also has been much interest recently in the fabrication of Si_xGe_{1-x} strained-layer films on Si for both optical[5] and microwave[6] device applications. In addition, the large-band-gap insulator CaF_2 has been grown as an epitaxial, single-crystal film on Si[7] and is potentially of interest for applications in silicon-on-insulator (SOI) technology, and in three-dimensional, epitaxial, very large scale integrated circuit (VLSI) architectures. Let us now consider some of these possibilities in more detail.

Silicides

Silicides are stoichiometric compounds of Si and a transition metal. (Silicide compounds with alkali and alkaline earth metals are also possible, but tend to be quite reactive, and hence, uninteresting for Si device applications.) The number of known transition-metal silicide compounds is large (over 100), partly because there are many possible stoichiometric phases of the form M_xSi_y, for any given transition metal M.[8] For some time now, transition-metal silicides have been of interest in VLSI applications mainly as *metallizations*, *i.e.*, as gates and interconnects. This is because of their desirable electrical properties such as high conductivity, their ability to form reliable Schottky barriers on Si, their chemical stability (particularly their ability to withstand oxidizing ambients at typical processing temperatures), and their relatively low temperatures of formation by solid phase epitaxial (SPE) techniques.[8] However, silicides are interesting materials in their own right, and with the capabilities of MBE, may be exploited to bring about the next generation of Si-based device technology.

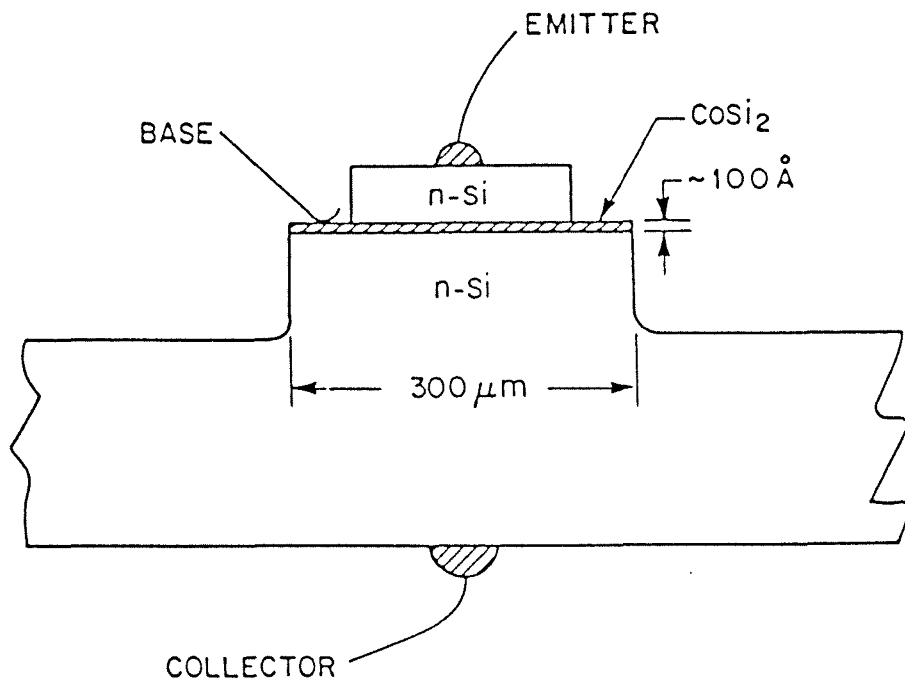
Transition-metal silicides display a range of physical behavior. Electrically, most silicides are metallic; some of these undergo a superconducting phase transition at low temperature.[8] Although most of the known transition temperatures T_c for silicides tend to be below 3 K, that of V_3Si is significantly higher[8] ($T_c \approx 17\,K$). Some silicides, such as $CrSi_2$, are thought to be narrow-band-gap semiconductors ($E_g \sim 0.3\,eV$). In addition, certain silicides, *e.g.*, Gd_5Si_3, exhibit ferromagnetism at room temperature though in general the Curie temperatures of silicides tend to be much lower.[9]

Structurally, transition-metal silicides crystallize in most of the seven crystallographic systems, including cubic. Zur, McGill, and Nicolet[10] have systematically examined the possible lattice matches of over 100 transition-metal silicides to the (100), (110), and (111) surfaces of Si. (Their definition of a lattice match[10]

requires the existence of some two-dimensional lattice common to the Bravais lattices of both materials, but this lattice may be specified by such a *large* unit cell that only a *small fraction* of the total number of interfacial atoms may actually constitute the match.) The results of Zur *et al.* show that there are a large number of transition-metal silicides having lattice mismatches to Si of less than 2%. Silicides well matched structurally to the Si lattice may be good candidates for heteroepitaxial growth.* If the lattice mismatch is small enough and the crystal structures are similar, the silicide may grow as a single-crystal film on Si, with the possibility of heteroepitaxial overgrowth of Si on the silicide. This is the case for $NiSi_2$ (lattice mismatch of 0.4%) and $CoSi_2$ (lattice mismatch of 1.2%), both of which crystallize in the CaF_2 structure, an fcc structure much like that of Si but with a three-atom basis (Ca at [000] and F at $\pm\frac{1}{4}[111]$ within the conventional unit cell). The capability of multilayer heteroepitaxy is important to the fabrication of novel device structures.

One novel device structure actively being pursued is the metal-base transistor (MBT).[3,2] The MBT structure is shown in Fig. 1.1(a). The structure of the MBT is similar to that of an *n-p-n* bipolar transistor, with the base region being replaced by a thin metal layer about 100 Å thick. Figure 1.1(b) shows the band diagram of the MBT under bias conditions for normal operation. Operation of this device is dependent on having a very thin, high-mobility region which acts as the base. Electrons, thermally excited over the emitter-base Schottky barrier, are injected "hot" into the thin base, where they quickly traverse the distance to the collector, pass over the reverse-biased collector-base Schottky barrier, and are collected. The hot electrons (of energy several kT above the Fermi energy) are meant to traverse the base region ballistically. Hence, the base material and thickness must be chosen so that the electron mean-free path,

*It must be stressed, however, that the interfacial *chemistry* as well as the geometry must be considered in the accurate prediction of heteroepitaxy.



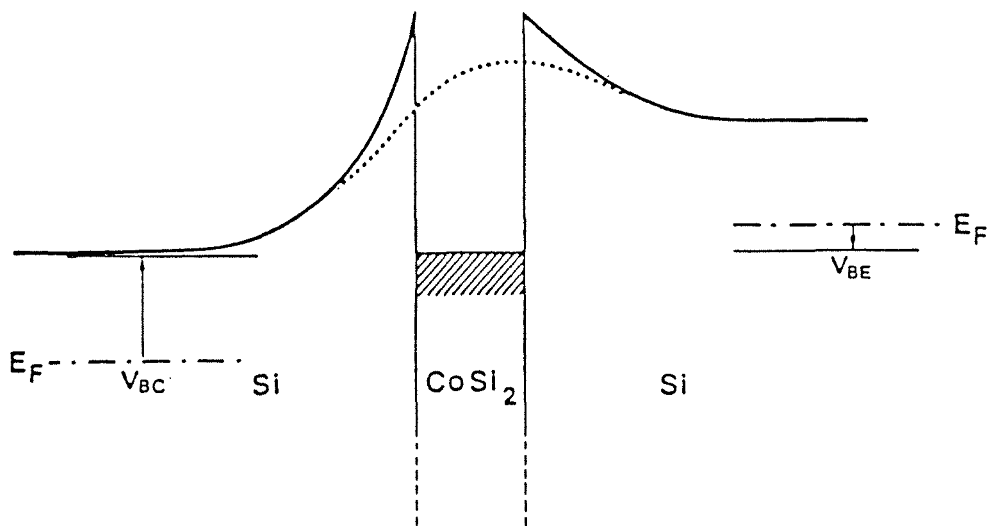


Figure 1.1: (a) Structure of the metal-base transistor. (b) Band diagram of device under bias. (From Ref. 3.)

or scattering length, is larger than the base thickness. Thus far, $CoSi_2$ has been used because of its comparatively large scattering length[11] ($\approx 200\,Å$), and the fact that this material can be grown as a single-crystal film on a Si (111) surface and subsequently overgrown with single-crystal Si. Because the electrons travel across the thin base ballistically, the MBT is in principle capable of very high speed, high frequency operation. An order of magnitude estimate, taking an electron velocity of $10^7\,cm/sec$, gives 0.1 psec for the transit time. (In practice, other characteristic times, such as the RC time constant associated with the emitter-base Schottky barrier, may be slower and thereby dominate the transient response of the device.)

Integrated structures of silicides and Si are also of interest for optical device applications. Preliminary efforts to fabricate an epitaxial $NiSi_2$/Si infrared Schottky detector by MBE codeposition of Ni and Si onto n-type Si substrates have been reported.[12] The Schottky barrier height (SBH) of $NiSi_2$ on n-type Si is about 0.7 eV, putting the detection threshold in the wavelength range of interest for optical communications (1.3–1.6 μm). The device structure is shown in Fig. 1.2 and has a quantum efficiency of 0.22%. With an optimized structure, it should be possible to improve the quantum efficiency by at least a factor of two.[12] It has also been suggested that a Si capping layer might be used to enhance the quantum efficiency.[13] Spatial quantization of electronic energy levels by fabricating very thin $NiSi_2$ layers in this case might be used to shift and sharpen the detection threshold. Also, it is possible that electron-induced avalanche multiplication may be used to significantly enhance the quantum efficiency.[12] The principal interest in an epitaxial $NiSi_2$/Si infrared (IR) detector is that, in principle, it can be vertically integrated with other devices such as amplifiers on a Si substrate.

Given the variety of electrical properties exhibited by the transition-metal

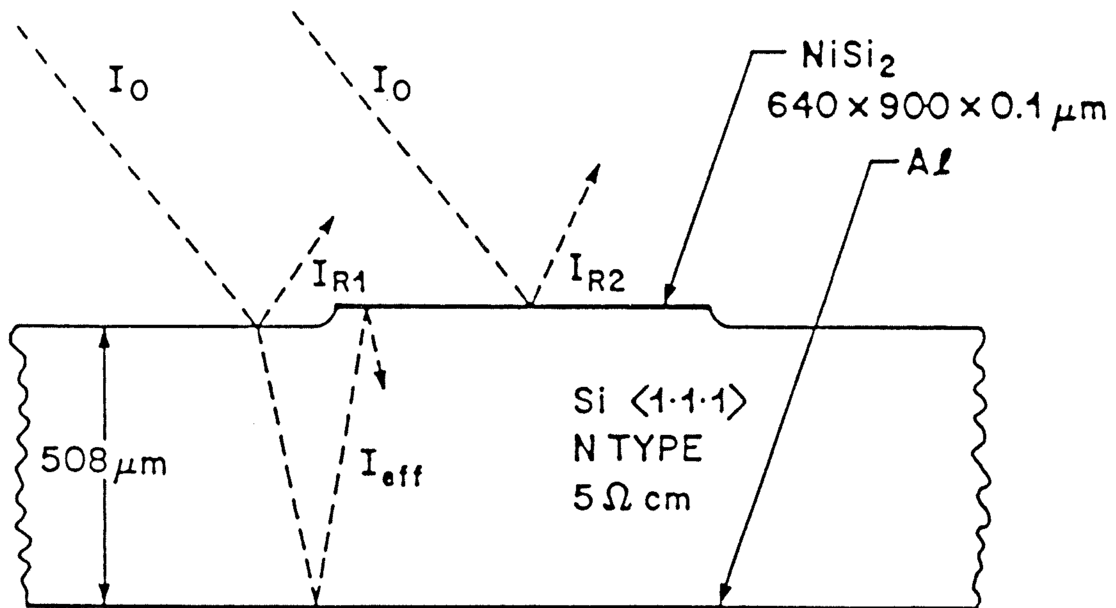


Figure 1.2: $NiSi_2$/Si infrared Schottky detector structure. Illumination of the metal-semiconductor interface is effected by reflection of below-band-gap incident light off the Al metallization (back contact). (From Ref. 12.)

silicides, it is possible to speculate on further novel device possibilities. On the basis of temperature-dependent resistivity or Hall coefficient studies, it is inferred that many of these materials are semiconductors; in particular, $FeSi_2$, $CrSi_2$, and $ReSi_2$. Recent optical transmittance and reflectance studies on orthorhombic, β-$FeSi_2$ thin films on Si have demonstrated that β-$FeSi_2$ is a direct-gap semiconductor with a band gap of 0.87 eV.[14] This is near midrange in the optical communication window described above. In addition, Zur *et al.* have calculated the lattice mismatch with respect to (100)-Si to be 1.0% although no information is yet available about whether β-$FeSi_2$ can be grown single-crystal on Si.[10] Similarly, the lattice mismatches for $CrSi_2$ ($E_g \sim 0.3\,\mathrm{eV}$)[15] and $ReSi_2$ ($E_g \sim 0.12\,\mathrm{eV}$)[16] on (111)-Si have been determined to be 0.1% and 0.2%, respectively.[10] Should it prove possible to fabricate high quality heteroepitaxial structures using these semiconducting silicides, one can imagine fabrication of light emitting and detecting structures analogous to those fabricated with the type III-V, and more recently, the type II-VI semiconductors. Finally, the possibility of using epitaxial superconducting silicides in high speed mixing and switching devices, such as Josephson Junction structures, should be mentioned. The main advantage to the fabrication of the silicide/Si-based devices discussed above is that, because of their compatibility with already established Si device technology, they could be integrated into production more readily than the more exotic III-V and II-VI materials.

Si/Ge Structures

Si and Ge are both indirect-gap semiconductors. The lattice mismatch between these two elemental semiconductors is substantial (4%). Nevertheless, with the use of MBE growth techniques,[5] it is possible to fabricate high-quality, coherently strained layers of Si_xGe_{1-x} alloys epitaxially on Si. Further, the com-

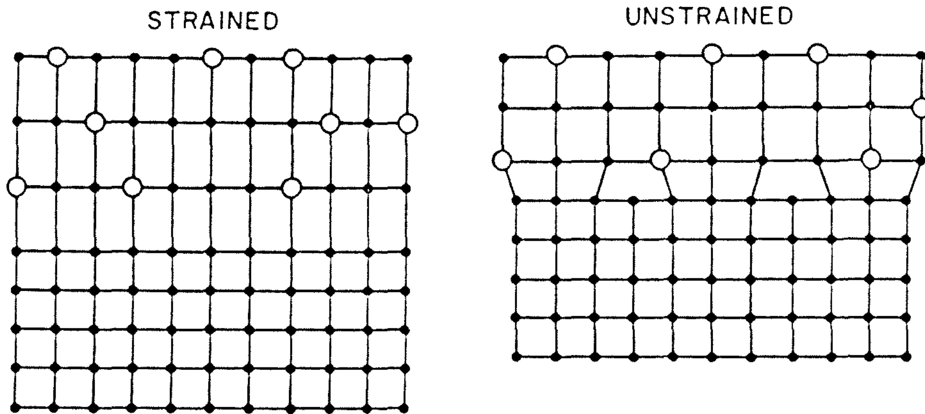
position of the alloy can be precisely controlled during growth, permitting the fabrication either of compositionally graded alloy layers, or abrupt heterojunctions. Moreover, the technique of secondary implantation[17] results in Sb doping profiles as sharp as 2.5 orders of magnitude over less than 100 Å. Strained-layer superlattices (SLS) of Si_xGe_{1-x}/Si have been grown and characterized.[5] As with any superlattice, the optical properties are in general quite different from those of an alloy of the same average composition, and can be tailored artificially by varying the individual layer thicknesses and the alloy composition x. Strained layers of Si_xGe_{1-x} on Si have also been studied.[18] Strain effects play a significant role in determining the band offsets of this poorly lattice-matched system. Theoretical and experimental attempts to understand the effects of strain in determining the band gap of coherently strained Si_xGe_{1-x} films have recently been undertaken.[18,19,20]

The technological interest in strained-layer Si_xGe_{1-x}/Si heterostructures lies in two applications: light detectors for optical communications, and microwave electronic devices. First, let us consider the optical properties of Si_xGe_{1-x} films. Figure 1.3(a) shows schematically the structural difference between an unstrained and a coherently strained Si_xGe_{1-x} epilayer on a substrate of Si. Since the unstrained film has a larger lattice constant than the substrate, misfit dislocations at the interface must occur periodically to accommodate the two incommensurate structures. However, if the Si_xGe_{1-x} film is thin enough, the energy required to uniformly strain the film compressionally along the interface is less than the energy required to form misfit dislocations. The resultant compressionally strained film, shown in Fig. 1.3(a), free from misfit dislocations, also undergoes a dilatational strain normal to the interface according to a factor analogous to Poisson's ratio. Since the electronic band structure is affected by strain, the difference in optical properties of strained and unstrained films can be significant. Fig-

ure 1.3(b) shows the fundamental band gap for the two cases. Also indicated is the spectral window of interest for optical communications. Note that, in the range of interest, the fractional Si content is much greater and the variation with composition much smoother for the case of strained-layer epitaxy. This is important because it means that optical devices using strained Si_xGe_{1-x} layers will be structurally more compatible with Si, and the optical properties of the devices will be more easily controlled.

MBE-grown Si_xGe_{1-x} structures are also of interest in microwave and high frequency device applications. An example of such a device is shown in Fig. 1.4. This device is a type of modulation-doped field effect transistor (MODFET). The structure is fabricated epitaxially on a Si substrate. First, an unstrained buffer layer of Si_xGe_{1-x} is deposited. The buffer layer is chosen to have roughly the average composition of the layers to be fabricated above it since the closer match to the net strained-layer lattice constant leads to better structural quality in the overlayers. The actual device structure begins with a 200 Å undoped Si layer. On top of this goes a 50%/50% alloy layer 100 Å thick, in which a very sharp doping spike (Sb) is centered. A 100 Å-thick compositionally graded layer ($x = 0.5$ to 0) is then grown, and, finally, an undoped Si film of thickness 100 Å is deposited. All layers above the buffer layer are coherently strained.

The resultant band gap lineup for the MODFET structure is shown in Fig. 1.4(b). The active region of this device involves a two-dimensional electronic band of energy levels at the heterojunction between the intrinsic Si and 50%/50% alloy which occurs due to spatial confinement of these energy states in the direction normal to the interface. Population of this two-dimensional band comes from electrons donated by the nearby Sb donors. Since the resultant two-dimensional electron gas is spatially separated from the donor atoms, charge transport normal to the page in Fig. 1.4 can occur without the mobility reduction associated



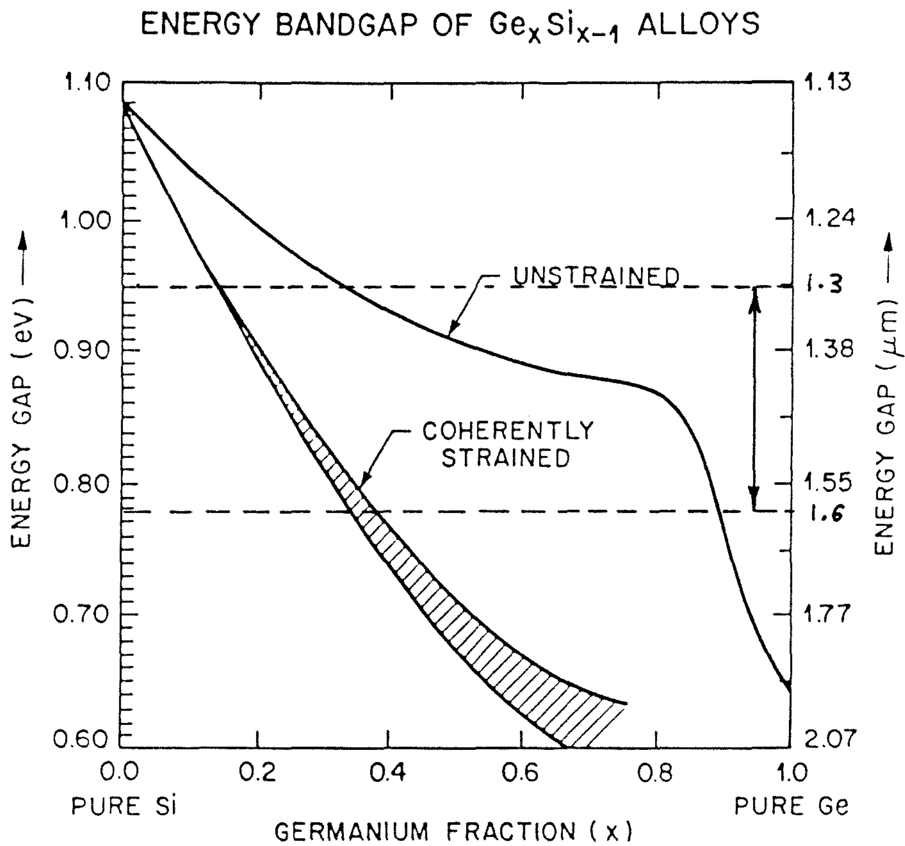


Figure 1.3: (a) Epitaxial structure of unstrained and coherently strained Si_xGe_{1-x} alloy layer on Si. (b) Energy band gap of Si_xGe_{1-x} strained and unstrained films as a function of alloy composition x. (From Ref. 5.)

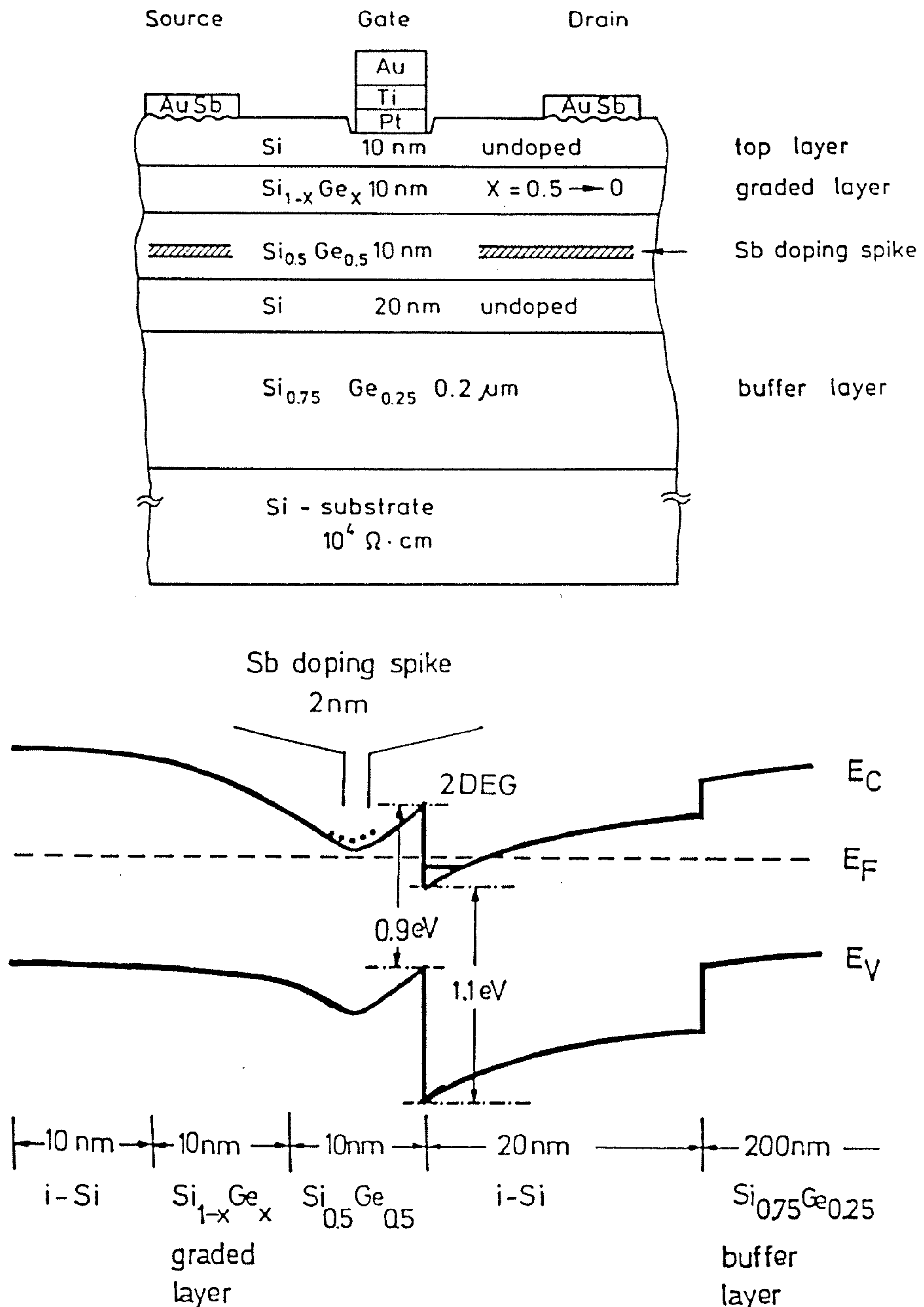


Figure 1.4: (a) Structure of the epitaxial MODFET grown by MBE. (b) Band structure lineup showing band offsets and band bending of the MODFET structure in thermodynamic equilibrium. (From Ref. 6.)

with scattering off ionized donor impurities. Such a two-dimensional conduction channel is the basis of the so-called high electron mobility transistor (HEMT). For the device shown in Fig. 1.4, transistor action is achieved through the electric field action of the Schottky gate contact on the HEMT channel. Current is conducted into and out of the channel through the (Ohmic) source and drain contacts, respectively. Finally, we mention here that another kind of microwave device, the IMPATT (standing for "impact ionization avalanche transit time") diode, has also been successfully fabricated by Si MBE, and in fact, is being produced commercially.[21]

1.2.3 Recent Progress

Processing

Much progress has been made over the past few years in the development of Si MBE processing, homo- and heteroepitaxial growth techniques, and in the development of schemes for improving throughput for commercial viability. One of the prerequisites to high quality MBE growth on Si is the preparation of a good substrate surface. An atomically clean surface free from carbon or oxygen is essential. Carbon and oxygen bond strongly with Si, tending to nucleate stacking faults in the epitaxial film due to three-dimensional island growth rather than the desired planar, two-dimensional growth. Surface cleaning of Si is generally done in one of two ways. In one procedure, the substrate is sputtered with keV-range Ar ions followed by a heat treatment around 850 °C to anneal out the sputter-induced damage.[22] The other method, the so-called Shiraki technique, involves the growth of a thin (< 50 Å) oxide film by chemical processing outside the MBE system, which is then volatilized *in situ* by heating the substrate under ultrahigh vacuum (UHV) conditions.[23] The Shiraki technique is viewed by most workers as the procedure which yields the superior surface although there is not

complete agreement on this point. Originally, the Shiraki approach involved a volatilization temperature which was fairly high (1200 °C).[1] More recently, it has been shown that such high-temperature treatment may lead to unintentional *p*-type doping effects.[24] Refinements of the Shiraki cleaning technique have led to a procedure which now uses a much lower temperature (850 °C). At this reduced temperature, the undesired effects of high-temperature annealing are not observed,[24] and carbon- and oxygen-free reconstructed Si surfaces suitable for high quality epitaxial growth are obtained routinely.

Progress also has been reported on doping of epitaxial Si layers during growth. While in principle it is always possible to use the technique of ion implantation to achieve good control over doping in epitaxial films, in practice, this is a very expensive and complicated addition to a Si MBE machine. It is preferred to be able to use conventional sources to codeposit dopants with good control during growth of the epitaxial films. Initially, it was believed that only the dopants Al, Ga, and Sb were suitable for coevaporation from conventional Knudsen cells because of the stringent requirements of UHV.[25] Unfortunately, these dopants have low sticking coefficients, and show a complex incorporation behavior very dependent on substrate temperature, rendering good control over doping profile difficult. Then, B was discovered to have near unity sticking coefficient and simple incorporation behavior,[25] and it was found that excellent control over doping profiles (sharpness better than 100 Å/decade) and a wide range of concentrations (up to the solid solubility limit) were possible. For *n*-type doping, several related enhancement techniques were discovered which dramatically increased Sb incorporation and profile control.[25,17] Thus, significant progress has been achieved in the problem of doping, but it is still an area of vital concern, with real problems left to solve before Si MBE could become a commercial technology.

Epitaxial Growth

Much progress has been reported recently on the perfecting of homo- and heteroepitaxial growth techniques. In 1983, Tung, Gibson, and Poate developed a new technique for growing ultrathin metallic films (thickness less than 50 Å) of $CoSi_2$ and $NiSi_2$ on (111)-Si which were continuous, single-crystal, and defect-free.[26] These and other workers have also shown that, once such a "template" layer has been established, it is possible to thicken the ultrathin film (while preserving crystalline perfection) by standard SPE, MBE, or reactive deposition epitaxial (RDE) techniques.[27,7,22] Prior to the development of the template growth technique, continuous, singly-oriented silicide films ($CoSi_2$ and $NiSi_2$) thinner than about 500 Å were not possible. In addition, it has been shown[26] that for the case of $NiSi_2$ on (111)-Si, it is possible to control the crystallographic orientation of the template layer with respect to the Si substrate, resulting in the so-called type-A and type-B orientations. The measurement of the electrical properties of type-A and type-B $NiSi_2$/Si interfaces is the subject of Chapters 2 and 3 of this thesis.

In connection with the MBT structure shown on page 7, it is of interest to achieve good quality epitaxial overgrowth of Si on thin silicide layers. Up to now, the MBT has been structurally less than ideal. A significant problem has been the formation of pinholes in the thin $CoSi_2$ base layer.[3] Such defects in the base significantly affect the electrical behavior of the MBT, causing it to behave instead like a permeable-base transistor (PBT).[29,30] Another problem has been the electrical quality of the epitaxial Si/$CoSi_2$ junction.[31] Structurally, $NiSi_2$ may prove to be a better choice for the base because of its closer lattice match to Si. By using a template technique for the Si overlayer, Tung, Gibson, and Levi have succeeded in making very high quality epitaxial, single-crystal, strained semiconductor-metal-semiconductor heterostructures.[4] This result may

prove significant in resolving some of the unsolved problems with the MBT.

Other advances in heteroepitaxy have been recently reported. Two groups[32,33] have successfully grown double heterostructures consisting of epitaxial CaF_2 on top of epitaxial $CoSi_2$ and $NiSi_2$ films on Si. The CaF_2 surface morphology and crystallinity appears to be better for growth on $NiSi_2$ than on $CoSi_2$. To date, however, there is no report of successful growth of a metallic silicide on top of a CaF_2 film on Si, which would constitute a more interesting, metal-oxide-semiconductor (MOS) device structure. Other progress includes reported electron mobility enhancements up to a factor of five in nearly defect-free, Sb-modulation-doped SLS structures with alternating layers of $Si_{0.5}Ge_{0.5}$ and Si.[34] This result is consistent with a negative conduction band offset of Si with respect to the smaller band-gap $Si_{0.5}Ge_{0.5}$ strained layers. Finally, homoepitaxial doping superlattices have been investigated.[35] Results obtained on *n-p-n-p* doping superlattice structures ($N_{Ga} \approx 10^{18}/\mathrm{cm}^{-3}$ and $N_{Sb} \approx 10^{16}/\mathrm{cm}^{-3}$) reveal *p*-type conductivity, with room temperature hole mobilities approximately double those of *pure* bulk Si crystals.[35] The origin of this mobility enhancement is not clear at present.

High throughput MBE

We conclude this overview of Si MBE with a discussion of the problem of throughput. Typical present day MBE machines are complicated UHV systems consisting of a growth chamber, an analysis chamber for Auger and low-energy electron diffraction (LEED) characterizations (both chambers normally maintained under UHV conditions), and a loading chamber which is vented to atmosphere. These MBE systems are designed to process 3-in diameter wafers one at a time although most allow for the initial loading of multiple wafers into some sort of cassette mechanism. The problem with the present scheme is that, because

of the complexity and expense of existing hardware, problems with equipment reliability, and the need for highly trained operating personnel, the cost per wafer is not competitive with other technologies except for specialized, exotic device structures.[36] One cost reduction approach would be simply to reduce the number of parts in a typical laboratory system. However, some of the appeal of Si MBE dervies from its versatility, making this option undesirable. The other approach is to *significantly* increase throughput. One scheme for doing this is presently being developed by Bean and is shown in Fig. 1.5.[36] The figure shows a multi-chamber, production-line style apparatus optimized for high throughput through the use of platens which are loaded with many wafers. In this scheme, the loaded platens are inserted into a loading chamber, are passed sequentially from chamber to chamber (under UHV conditions), and are finally removed at the other end. Each chamber in the line performs a specific processing step. Because these processing steps can be performed concurrently, throughput is limited by the longest individual step rather than by the time required for the entire process.[36] In addition to greatly improved throughput, the design of this apparatus promises also to resolve problems of flaking and film uniformity encountered in conventional systems.[36]

1.3 Single-Crystal Silicides on Si

The discussion up to this point has been in the form of an overview of Si MBE, with an emphasis on its *technological* prospects. This discussion constitutes the background to Chapters 2, 3, and 4 of this thesis. Now, we wish to turn our attention back to science and introduce the object of the research reported in Chapters 2 and 3; namely, the study of the electrical nature of single-crystal silicide/Si interfaces. With the use of MBE, the silicides, $NiSi_2$ and $CoSi_2$, because

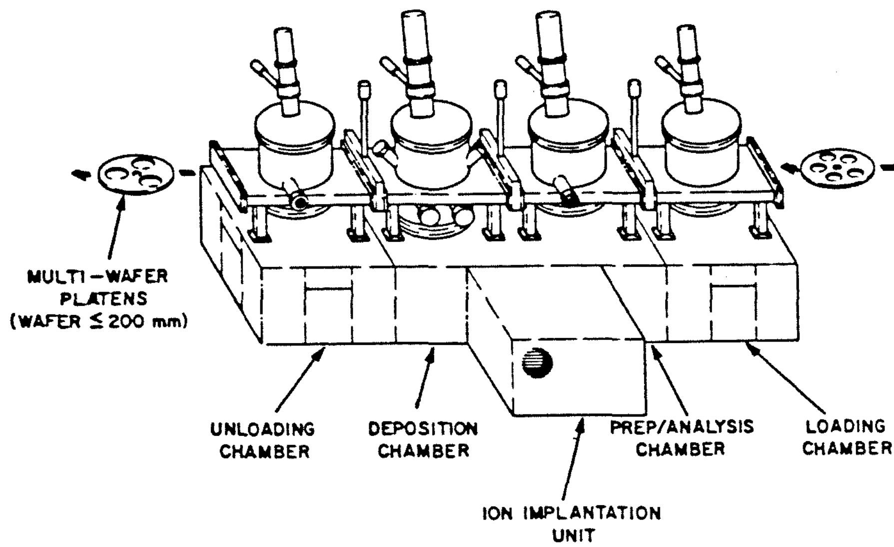
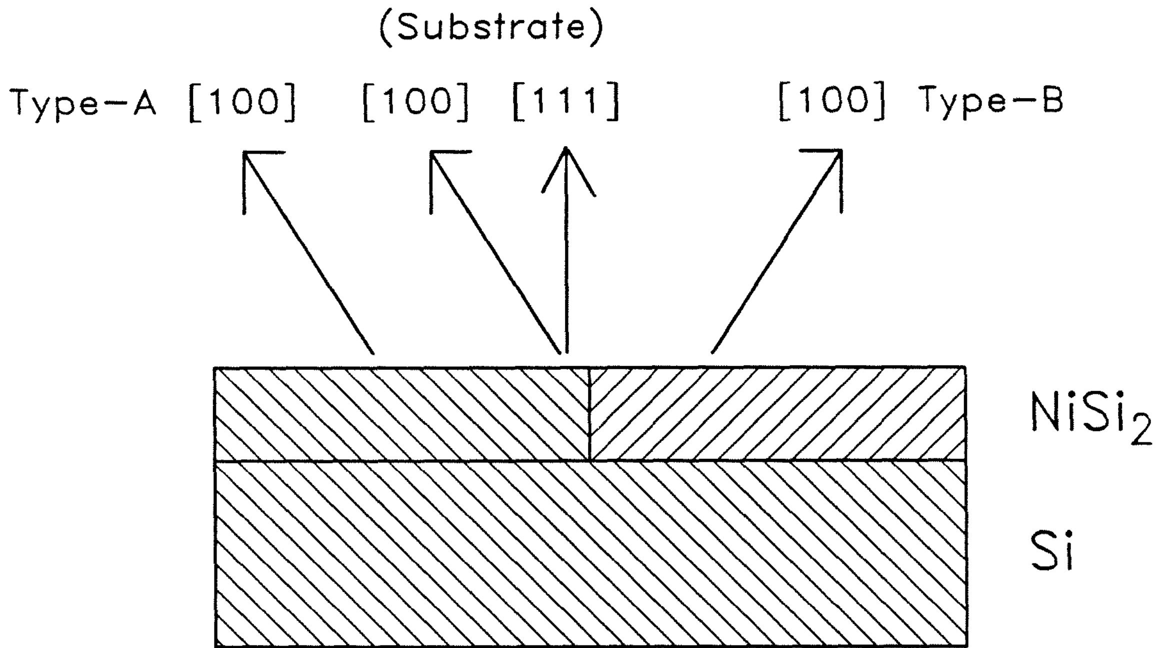


Figure 1.5: Apparatus designed to achieve high throughput MBE processing of Si wafers. (From Ref. 36.)

of a close lattice match and similar crystal structure, can be grown to a very high degree of structural perfection as single-crystal films on the (111) surface of Si.[27] It is also possible to grow $NiSi_2$ single-crystal films on (100)-Si.[26] In addition, for the case of $NiSi_2$ grown on (111)-Si, it is possible to control the orientation of the epitaxial growth with respect to the substrate so that one of two possible orientations, designated type-A or type-B, is obtained.[26] The type-A and type-B structures are shown in Fig. 1.6. The type-A film is aligned crystallographically with the substrate such that the epilayer and substrate [100] directions are parallel while the type-B film differs only by a 180° rotation about the substrate normal. Structurally, both orientations match to the Si (111) surface equally well. Despite the success with $NiSi_2$, it is not yet possible to stabilize growth of type-A $CoSi_2$ interfaces; in single-crystal form, only type-B films of $CoSi_2$ can be grown at present.[26]

Despite the ubiquity and technological importance of metal-semiconductor interfaces, a fundamental understanding of their nature is still lacking. It is easy to account for the existence of the electrostatic (Schottky) barriers which are observed at such interfaces, using little more than simple electrostatics and the band picture of solids. However, it has proven considerably more difficult to understand, even phenomenologically, the formation of Schottky barriers well enough to consistently explain *all* of the vast body of data accumulated over the past 30 years.[37] Part of the difficulty is that an *ab initio* treatment of this problem is extremely formidable. The rest of the difficulty is that good experimental control and characterization of metal-semiconductor interfaces has not been possible. As we shall see below, the interface plays an important role in the formation of Schottky barriers. Almost all of the experimental data gathered from metal-semiconductor interfaces were obtained on polycrystalline metal samples, often poorly characterized chemically (with respect to the presence of



Type–A aligns with substrate

Type–B is rotated by 180 deg

Figure 1.6: Type-A and type-B orientations of $NiSi_2$ on (111)-Si. The type-A film is in exact crystallographic alignment with the Si substrate while the type-B structure is rotated by 180° about the substrate normal. Both orientations match to the (111)-Si surface with equal perfection.

interfacial impurities) and structurally.

Now, thanks to MBE growth techniques and recent advances in physical characterization, it appears that some of the experimental difficulties may be coming under control. The distinct epitaxial orientations and high degree of structural perfection of the type-A and type-B $NiSi_2$/Si interfaces provide an ideal system in which to investigate fundamental metal-semiconductor interface behavior. These silicide-Si systems are also technologically interesting. We have already noted earlier in Sect. 1.2.2 the successful demonstration of Si overgrowth on both $CoSi_2$ and $NiSi_2$ films. Before continuing with a description of the investigations on silicide/Si interfaces presented in Chapters 2 and 3, however, let us briefly review the relevant aspects of metal-semiconductor interfaces.

1.3.1 An Ideal Metal-Semiconductor Interface

When a metal and a semiconductor are brought into intimate contact, an electrostatic barrier known as a Schottky barrier (SB) is formed. The manner in which the SB is formed is shown conceptually in Figs. 1.7(a)–(d). For simplicity, we consider here an "ideal" interface, *i.e.*, one in which the interface is simply a planar, abrupt transition between metal and semiconductor layers, each of which retains all of its bulk properties right up to the interface. Let us consider an *n*-type semiconductor; the *p*-type case is analogous. Figure 1.7(a) shows the isolated metal and semiconductor layers, their band diagrams each referred to a common energy "zero" known as the vacuum level. The vacuum level is the energy possessed by an electron with zero kinetic energy outside the material and infinitely far away from the surface.* Note that in drawing Fig. 1.7(a) we have made another assumption; namely, that the metal and semiconductor work func-

*The usual definition puts the electron just outside the surface, but because of surface charge or dipole layers, this definition can be ambiguous.

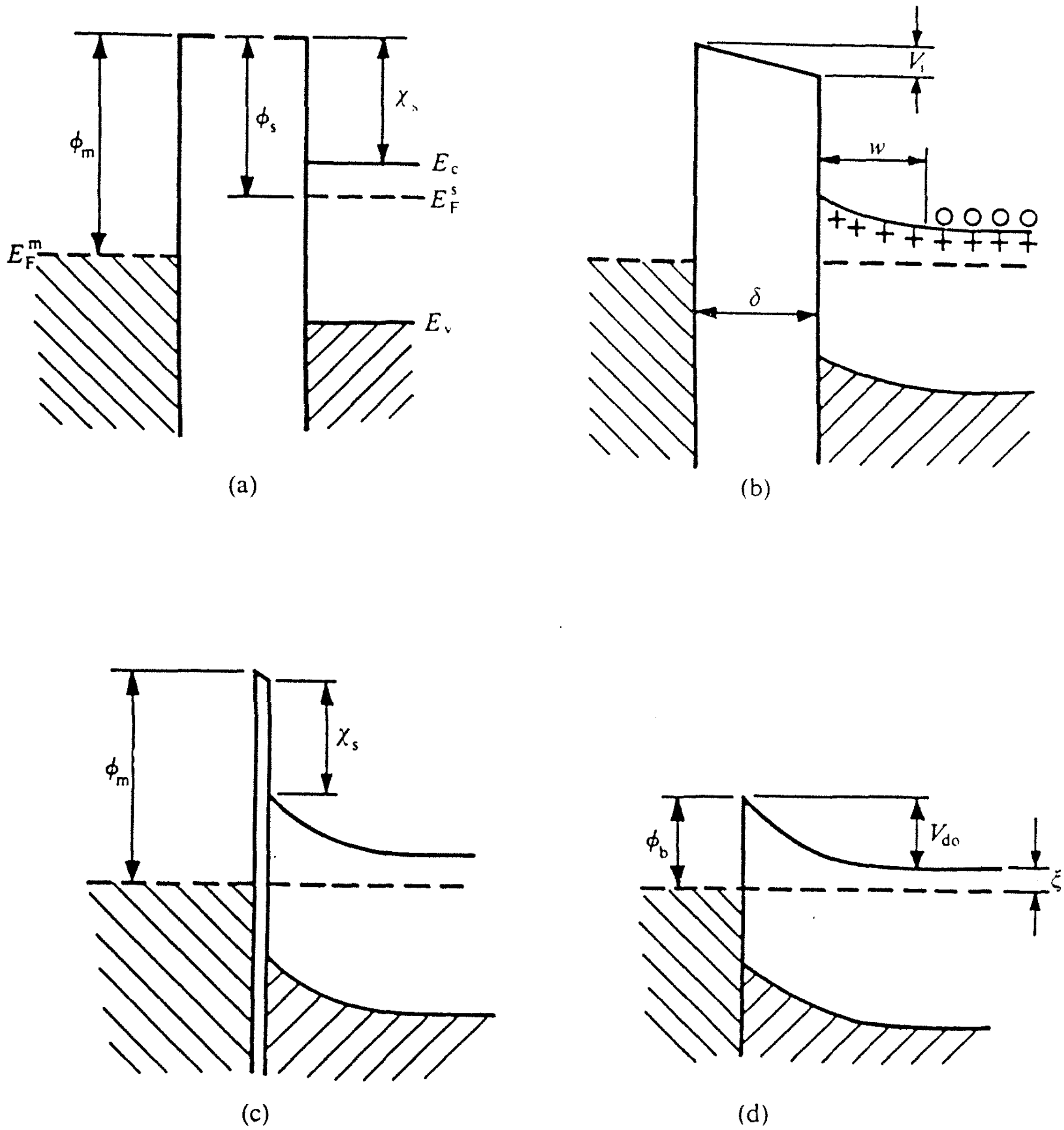


Figure 1.7: Sequence showing formation of a Schottky barrier. Metal and semiconductor are (a) neutral and isolated, (b) allowed to exchange electrons, (c) separated by a narrow gap, and (d) in intimate contact. (From Ref. 41.)

tions ϕ obey the relation $\phi_m > \phi_s$. This is the most important case in practice. In Fig. 1.7(b), the metal and semiconductor are allowed to exchange electrons. The electrons move from the semiconductor (high chemical potential) to the metal (low chemical potential) until thermodynamic equilibrium is established and the Fermi levels coincide. The removal of electrons from the semiconductor results in the formation of a positive space charge due to carrier depletion, with an equal but opposite charge occurring within a screening length (≈ 0.5 Å) of the surface in the metal. The situation is somewhat like that in a parallel plate capacitor where a voltage V_i is dropped uniformly in the gap δ between the layers. A diffusion or built-in voltage, V_{d0}, is also dropped in the semiconductor due to carrier depletion, and from Figs. 1.7(a) and (b) we have $\phi_m - \phi_s = V_i + V_{d0}$. As the two layers are brought closer together, as in Fig. 1.7(c), the voltage dropped across the shrinking gap becomes negligible until, in Fig. 1.7(d), we have intimate contact, with the SBH given by

$$\phi_{B0} = \phi_m - \chi_s. \tag{1.1}$$

This is the original model proposed by Schottky[38] and also by Mott[39] in 1938. The basic physical idea is simple: electrostatic barriers between dissimilar materials arise from a net transfer of electron charge, which in turn occurs because at finite temperature the electrons can escape from one material more readily than the other.

Unfortunately, Eq. 1.1 is not obeyed very well in practice. A number of assumptions have been made, perhaps the most serious of which is the neglect of the interface in determining SB formation. A very different view of SB formation was proposed by Bardeen[40] in 1947 to include the effect of interface electronic states. Consider the situation shown in Fig. 1.8. The metal and semiconductor are assumed to be separated by a thin ($\approx$ 10–20 Å) oxide such as that commonly found on chemically etched metal-semiconductor contacts. In this model, the ox-

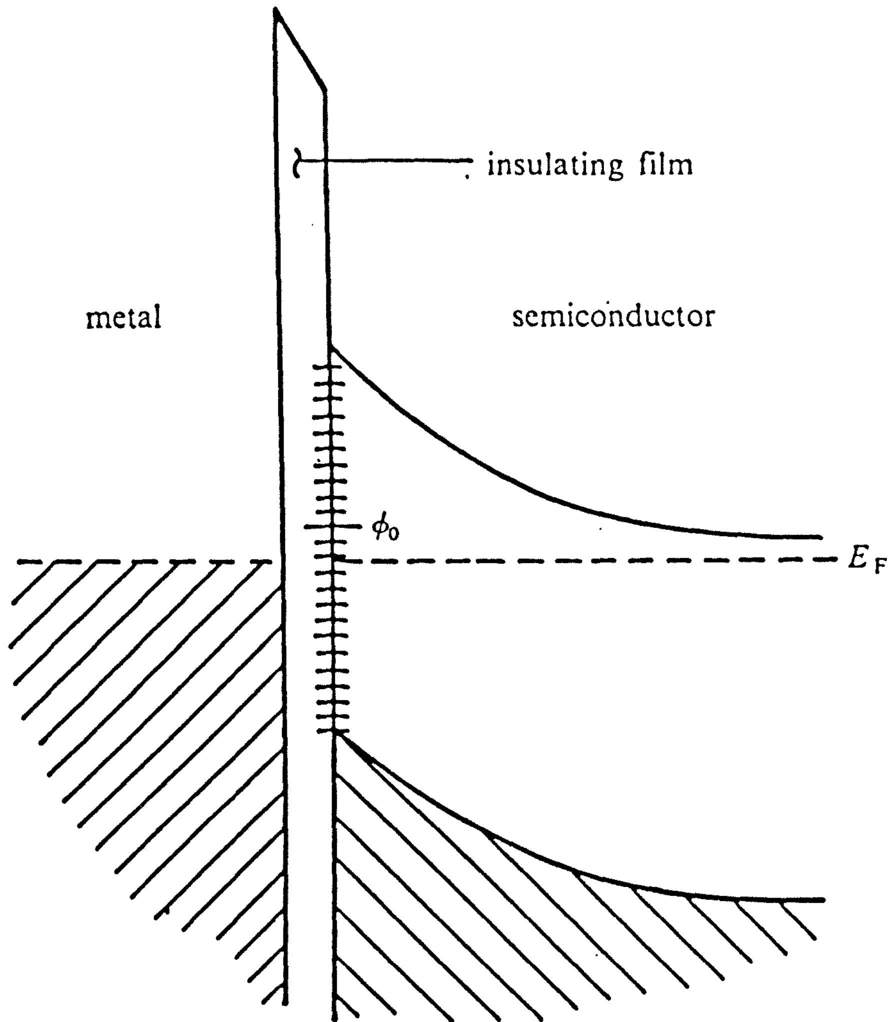


Figure 1.8: Schottky barrier formation including the effect of interface states. (From Ref. 41.)

ide separates the metal and semiconductor enough to allow the surface properties of each to be considered independently of the other while at the same time being thin enough so that electrons can easily tunnel through.[41] The other key feature of the model is the existence of a distribution of surface electronic states of surface density D_{ss} in the semiconductor. These states lie at energies within the forbidden gap of the semiconductor and are spatially localized near its surface. Surface states may be *extrinsic* due to impurities and defects at the interface, or may be *intrinsic* since even a perfect surface breaks the lattice translational symmetry in the perpendicular direction. In any event, there will be a "neutral level" ϕ_0 such that the surface is electrically neutral when all surface states below this energy are filled and those above are empty. Thus, the semiconductor surface charge is $Q_{ss} = D_{ss}(\phi_0 - \varepsilon_F)$. To demonstrate the importance of the surface states in SB formation, consider the case when D_{ss} is large. In this case, since the insulator width δ is fixed and insulator voltage V_i remains finite, so also must the metal charge Q_m remain finite. Further, since $-Q_m$ must equal the *total* charge in the semiconductor (bulk plus surface), Q_{ss} must remain finite as $D_{ss} \to \infty$. This requires that $\varepsilon_F \to \phi_0$. Since at the surface, the SBH is given by $(E_g - \varepsilon_F)$, the SBH in this limit is then given by

$$\phi_{\mathrm{B0}} \approx E_g - \phi_0. \tag{1.2}$$

Equation 1.2 is known as the Bardeen limit. In the Bardeen limit, the SBH is determined entirely by the electronic states of the interface, and the Fermi level is said to be "pinned" at the value ϕ_0.

Equations 1.1 and 1.2 represent two extreme points of view: in one case, the bulk properties alone determine the SBH, and in the other, the interface determines the SBH. For real Schottky barriers, the truth lies somewhere in between. The charge in the semiconductor will in general be divided between the space charge of the depletion layer Q_d and the charge at the surface Q_{ss}. The

SBH is then obtained from elementary electrostatics, subject to overall charge neutrality ($Q_d + Q_{ss} - Q_m = 0$) and the condition that the Fermi level be constant at zero applied bias. For the model shown in Fig. 1.8, the SBH can be shown to be

$$\phi_{B0} = \gamma(\phi_m - \chi_s) + (1-\gamma)(E_g - \phi_0), \quad (1.3)$$

where

$$\gamma = \frac{\epsilon_i}{\epsilon_i + q\delta D_{ss}}. \quad (1.4)$$

Here, ϵ_i is the dielectric constant of the insulating layer and $-q$ the electron charge. Hence, Eq. 1.3 contains Eqs. 1.1 and 1.2 as limiting cases, according to the relative importance of surface states (given by D_{ss}). There are additional complications which affect SB formation: modification of the surface contributions of ϕ_m and χ_s and field penetration into the metal,[41] interface quantum mechanical dipole layers,[42] and surface reconstruction at the interface. The tremendous computational complexity in treating these effects has prevented a quantitative prediction of ϕ_{B0} for actual metal-semiconductor combinations. However, a self-consistent many-electron calculation considering many of these effects using a jellium approximation for the metal and semiconductor has been undertaken recently. This calculation validates Eq. 1.1 in the absence of surface reconstruction between the free-surface and interface-surface situations.[43]

1.3.2 Schottky Barrier Height Measurements

We now wish to consider the SBH measurement. Assuming that a SB of barrier height ϕ_{B0} has formed, how is it measured? There are four standard techniques for SBH measurement, three of which are based on electron transport over the barrier while the fourth is based on the depleted space charge in the semiconductor.[44] We shall now discuss the techniques used in Chapter 2 to determine the SBH of $NiSi_2$/Si structures; namely, the *I-V* and photoresponse

methods. Both methods are based on electron transport. The mechanisms responsible for electron transport across the SB are shown in Fig. 1.9. The processes, numbered 1 thru 4, correspond to thermionic emission, quantum mechanical tunneling, recombination in the depletion region, and hole injection, respectively. For high quality Schottky diodes on nondegenerate semiconductors, thermionic emission is the dominant transport mechanism and so our attention will be focused on this process.

As an electron in the semiconductor nears the barrier, its charge $-q$ induces a response in the metal. The interaction of the electron with the metal can be represented as an electrostatic interaction between the electron and an image force in the metal, and is given by

$$F = \frac{-q^2}{4\pi(2x)^2\epsilon_s} = \frac{-q^2}{16\pi\epsilon_s x^2}, \tag{1.5}$$

where x is the perpendicular distance to the interface and ϵ_s is the semiconductor dielectric constant. The effect of the attractive interaction given by Eq. 1.5 is to *lower* the SB by an amount[44]

$$\Delta\phi = \sqrt{\frac{q\mathcal{E}_{\mathrm{max}}}{4\pi\epsilon_s}}, \tag{1.6}$$

where $\mathcal{E}_{\mathrm{max}}$ is the maximum electric field (near the interface) in the depletion layer. The barrier height measured by *I-V* and photoresponse methods is the lowered barrier height $\phi_{\mathrm{B}n} = \phi_{\mathrm{B}0} - \Delta\phi$. Because of the bias dependence of $\mathcal{E}_{\mathrm{max}}$, the precise value of $\phi_{\mathrm{B}n}$ depends on the voltage V applied across the SB as shown in Fig. 1.10. The thermionic-emission–diffusion theory of Crowell and Sze[45] gives the rectifying *I-V* characteristic of a nondegenerate Schottky diode at temperature T as

$$\begin{aligned} J &= J_s(e^{qV/nkT} - 1) \\ J_s &= A^{**}T^2 \exp\left(\frac{-q\phi_{\mathrm{B}n}}{kT}\right), \end{aligned} \tag{1.7}$$

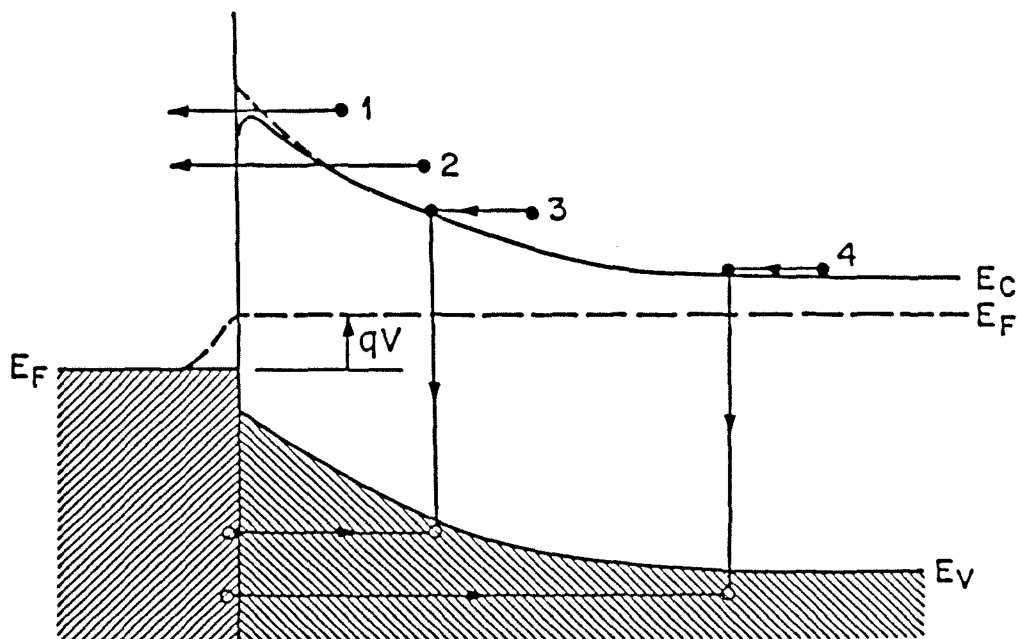


Figure 1.9: Possible mechanisms of electron transport in Schottky barriers. These are (1) thermionic emission over the barrier, (2) quantum mechanical tunneling through the barrier, (3) electron-hole recombination in the space charge region, and (4) recombination in the neutral region (hole injection). (From Ref. 41.)

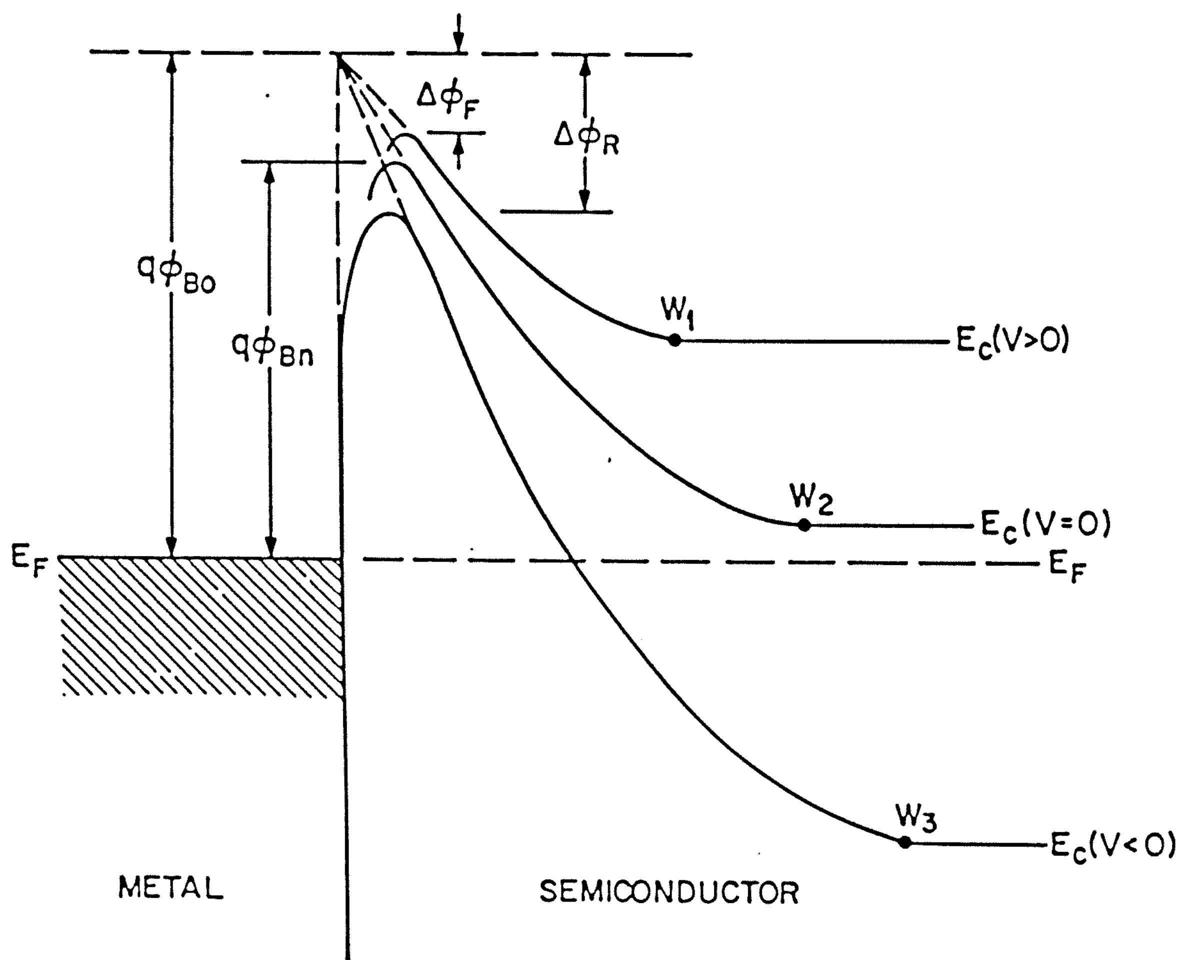


Figure 1.10: Dependence of Schottky barrier height on applied bias. For forward bias the barrier is raised while for reverse bias it is reduced for an n-type semiconductor-metal junction. (From Ref. 44.)

where A^{**} is the effective Richardson constant.[44] The quantity n in Eq. 1.7 is known as the ideality factor and is given by[44]

$$n \equiv \frac{q}{kT}\frac{\partial V}{\partial(\ln J)} = \left[1 - \frac{\partial \phi_{\mathrm{B}n}}{\partial V} + \frac{kT}{q}\frac{\partial(\ln A^{**})}{\partial V}\right]^{-1} . \qquad (1.8)$$

The ideality factor would be precisely unity for thermionic emission over a *fixed* SBH, but is actually slightly in excess of unity because of the applied-voltage dependence of the SBH, shown in Fig. 1.10. Other processes such as generation-recombination via deep impurity levels (process 3 in Fig. 1.9) correspond to a value $n = 2$. A high quality Schottky diode with a negligible fraction of recombination current in forward bias is expected to have $n \approx 1$.

The forward *I-V* method of SBH determination is a straightforward application of Eq. 1.7. For sufficient forward bias $V > 3kT/q$, the exponential term in the first equation dominates, leading to the result

$$\phi_{\mathrm{B}n} = \frac{kT}{q} \ln\left(\frac{A^{**}T^2}{J_s}\right) . \qquad (1.9)$$

The value J_s is obtained from the forward *I-V* characteristic by plotting $\ln J$ vs. V and extrapolating to the ordinate as in Fig. 1.11. Series resistance becomes important as V increases, limiting the useful range of data from which to extrapolate J_s. Fortunately, it is usually possible to apply numerical series resistance corrections to get around this difficulty.[14]

Schottky barrier heights can also be measured by the photoresponse technique. The photoresponse measurement is shown in Fig. 1.12. In this method, monochromatic photons of energy $\hbar\omega < E_g$ are used to illuminate the metal-semiconductor interface. Since the incident light used in this measurement is below band-gap, almost all of the light is absorbed by the free carriers in the metal and almost none by the semiconductor, which is assumed to be nondegenerate. Illumination may take place from the metal (front illumination) or, more

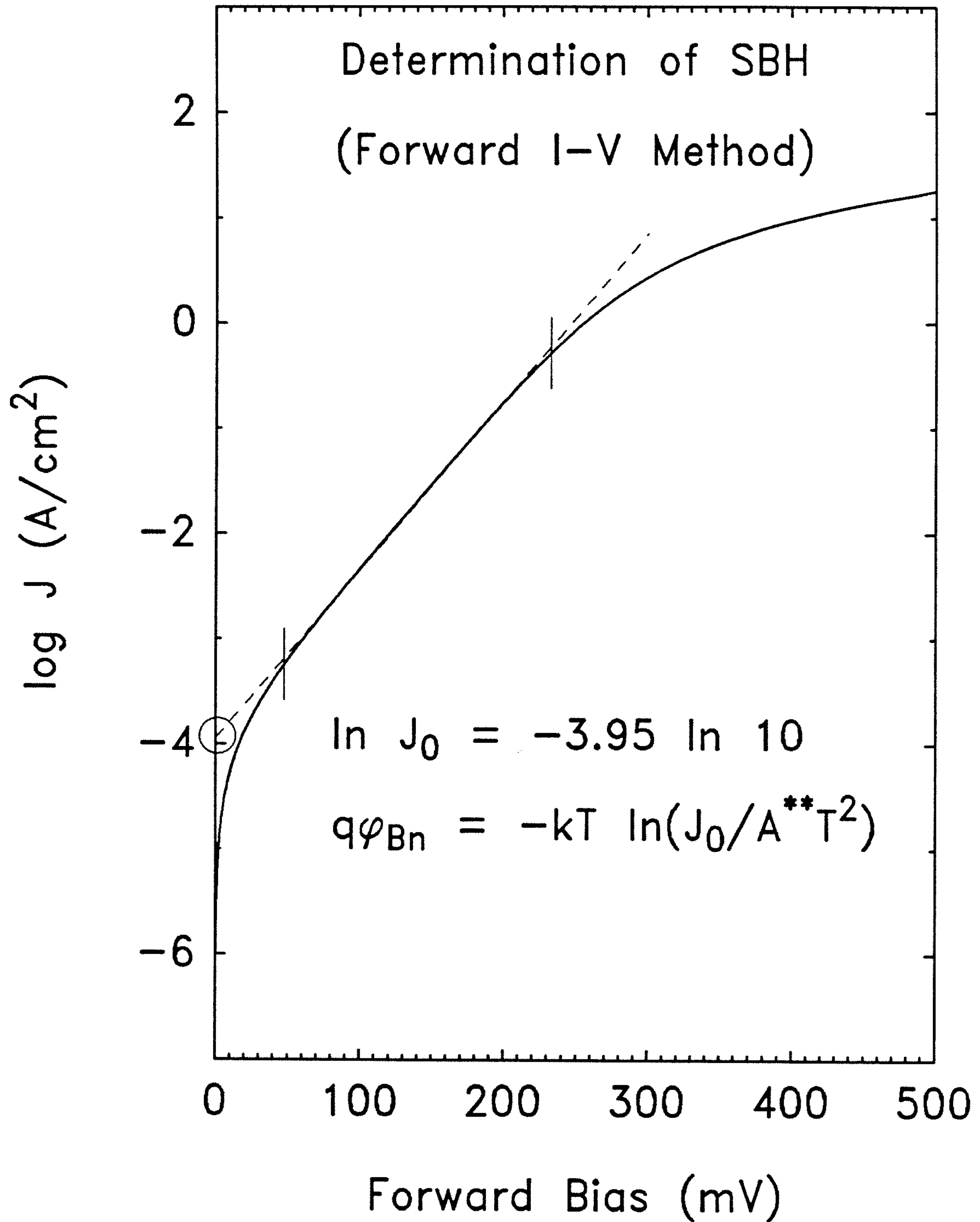
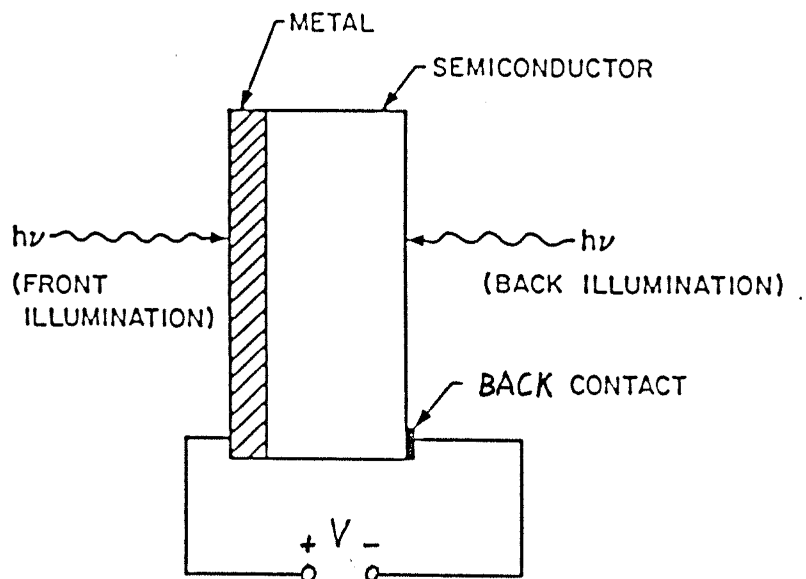


Figure 1.11: Determination of Schottky barrier height ϕ_{B_n} from forward I-V characteristics. Series resistance effects become important as V increases. The linear region should span at least two decades in current.



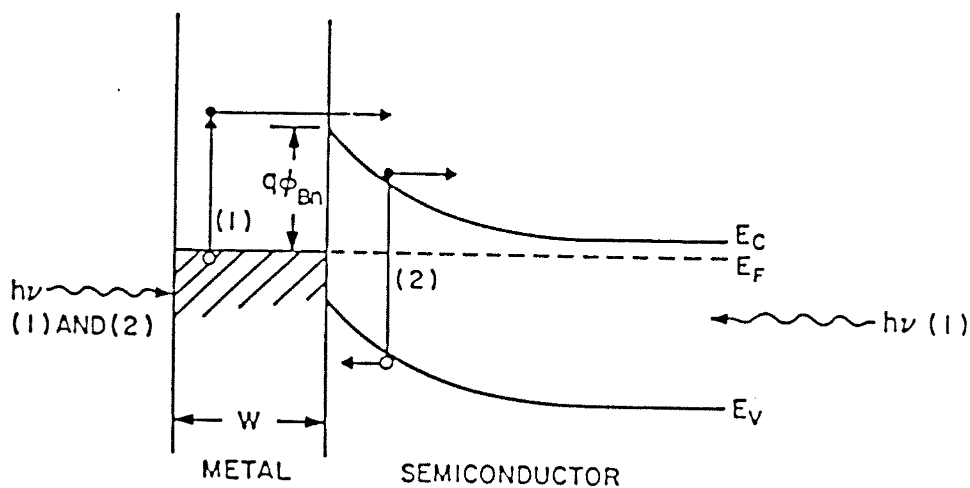


Figure 1.12: Experimental configuration for Schottky barrier height determination by the photoresponse method. (Adapted from Ref. 44.)

commonly, from the semiconductor (back illumination) side of the interface as indicated in Fig. 1.12. Referring to the figure, the physical situation is the following: an incident photon is absorbed by an electron (and **k**-conserving phonon) in the metal; if the electron is within a mean-free path ($\approx 100\,\text{Å}$) of the interface and has a perpendicular energy component in excess of $(\phi_{\mathrm{B}n} - \hbar\omega + \varepsilon_F{}^{\mathrm{metal}})$, the photoexcited carrier will surmount the barrier, giving rise to a photocurrent or photovoltage, depending on whether the measurement is made under closed or open circuit conditions, respectively. Since the photoresponse is proportional to the light intensity near the interface, the back illumination method is preferred when the metal layer thickness exceeds its optical attenuation length ($\approx 200\,\text{Å}$). The above analysis is a bit oversimplified. It would be more correct to speak of an energy-dependent "escape depth" of photoexcited carriers rather than the mean-free path,[47] but these lengths are probably of the same order of magnitude in the energy range of interest here.

Measurement of the SBH is accomplished by analyzing the photoresponse per incident photon, *i.e.*, the photoyield F, as a function of incident photon energy $\hbar\omega$. A soft threshold, turning on at $T = 0$ like $(\hbar\omega - \phi_{\mathrm{B}n})^2$, is found for Schottky barriers. A derivation of this result and a simple geometric interpretation will be given in Chapter 2. For now, we simply state the photoyield result originally worked out by Fowler[10] for the case of non-zero temperature

$$F = C\frac{T^2}{\sqrt{\varepsilon_F{}^{\mathrm{metal}} + \phi_{\mathrm{B}n} - \hbar\omega}}\left[\frac{x^2}{2} + \frac{\pi^2}{6} - \left(e^{-x} - \frac{e^{-2x}}{4} + \frac{e^{-3x}}{9} - \cdots\right)\right]; \qquad x \geq 0 \tag{1.10}$$

where $x \equiv (\hbar\omega - \phi_{\mathrm{B}n})/kT$, and C depends in general on the metal-layer reflectivity, absorptivity, electron mean-free path, and sample geometry,[49] but for our purposes may be taken to be constant. Since the terms $\hbar\omega$ and $\phi_{\mathrm{B}n}$ are typically an order of magnitude smaller than $\varepsilon_F{}^{\mathrm{metal}}$, the photoresponse behavior

approximately $3kT$ or more above the threshold is given by

$$F \approx C'(\hbar\omega - \phi_{Bn})^2 \qquad \text{for } x \geq 3. \tag{1.11}$$

Hence, extrapolation of a plot of $\sqrt{F}$ vs. $\hbar\omega$ to the abscissa should yield ϕ_{Bn} directly. This type of analysis is referred to as a Fowler plot. Photoresponse is a more direct method of SBH determination than the *I-V* method, and, as we shall see in Chapter 2, can give information about the presence of compound barriers that the *I-V* measurement cannot directly provide. Generally, for high quality Schottky diodes, the precision of both the photoresponse and *I-V* methods is about 0.01 eV.

1.3.3 Electron Tunneling Spectroscopy

In Chapter 3 we shall consider the transport of electrons through SB structures consisting of single-crystal silicide films on degenerate Si. When the semiconductor becomes sufficiently heavily doped, the barrier thickness decreases to ~ 100 Å and quantum mechanical tunneling effects begin to dominate the electron transport. The tunneling of electrons through the SB may take place either elastically or inelastically. Elastic tunneling is by far the dominant tunneling mechanism in SB structures, but for sufficiently heavily doped semiconductors, the effects of inelastic tunneling may be observed. In principle, analysis of *I-V* curves in terms of elastic tunneling theories can yield information about the dispersion relations of evanescent electron "states" within the band gap (which may be real states near a surface or interface), and also about the electron density-of-states (DOS) in the metal and the semiconductor. In practice, order of magnitude agreement in *I* between calculated and actual *I-V* curves is about all that can be expected. Fortunately, qualitative agreement is usually much better. Even so, DOS effects are usually not observed unless the DOS contains singular

features such as gaps or discrete levels. In contrast, inelastic tunneling effects yield quantitative information and have been exploited as a spectroscopic tool in a technique known as inelastic electron tunneling spectroscopy (IETS).[52] IETS spectra are used to identify quantized elementary excitations of the tunneling barrier such as phonons and plasmons, and to obtain their energy values. Both IETS and analysis in terms of elastic tunneling have been used in our studies of $CoSi_2$/Si and $NiSi_2$/Si structures. This work is presented in Chapter 3. In preparation for the results to be presented, let us consider now the important ideas associated with elastic and inelastic tunneling.

Elastic Tunneling

The elastic tunneling process was indicated earlier (process 2) in Fig. 1.9. In contrast to thermionic transport, tunneling transport will persist down to $T = 0\,\mathrm{K}$, and in the absence of impurity-induced or defect-induced leakage currents, will be the dominant process at low temperatures. In order for tunneling currents to be observed at low temperature, the semiconductor must be doped sufficiently so that carrier freezeout does not occur; *i.e.*, the semiconductor must remain *metallic* at $T = 0\,\mathrm{K}$. For Si, this requires an n-type doping level of $\approx 5 \times 10^{18}\,\mathrm{cm}^{-3}$ or greater. A general expression for the elastic tunneling through a potential barrier at arbitrary temperature has been given by Duke[50] as

$$J_{rl} = -\frac{2q}{(2\pi)^3} \int d^3k_r\, v_{xr} f_r(\varepsilon_r)[1 - f_l(\varepsilon_l)] D_{rl}. \tag{1.12}$$

This equation gives the current density component flowing from right to left through the barrier, taken to be normal to the x-direction. In the elastic tunneling process, it is assumed that: (1) the total electron energy, E, and (2) the electron wave vector parallel to the interface, $\mathbf{k}_{\parallel}$, are conserved quantities. For a bias qV applied across the barrier, the energies on either side are related by

$$\varepsilon_r = \varepsilon_l - qV \equiv E. \tag{1.13}$$

The integral in Eq. 1.12 runs over all initial-state wave vectors on the right-hand side of the barrier consistent with energy and $\mathbf{k}_{\parallel}$ conservation conditions, and with a perpendicular velocity component v_x directed *toward* the barrier. D_{rl} is the transmission probability or tunneling matrix element which couples an initial state ψ_r to a final state ψ_l. Tunneling electrons originate in occupied states to the right of the barrier (occupation probability given by the Fermi-Dirac function f_r) and tunnel into unoccupied states on the other side [vacancy probability given by $(1 - f_l)$]. The "2" in the prefactor takes care of electron spin degeneracy. There is an expression analogous to Eq. 1.12 to describe the electron tunneling current from left to right, J_{lr}.

Following Duke, the *net* current to the right is given by[50]

$$J = J_{lr} - J_{rl} = \frac{2q}{\hbar} \int dE \; [f(E) - f(E + qV)] \int \frac{d^2 k_{\parallel}}{(2\pi)^2} \, D(E, \mathbf{k}_{\parallel}), \tag{1.14}$$

where we have invoked energy and parallel wave vector conservation, and have made use of the relations

$$v_{x\{l,r\}} \, dk_{x\{l,r\}} = (1/\hbar) \, d\varepsilon_{\{l,r\}}, \tag{1.15}$$

$$d\mathbf{k}_{\parallel_l} = d\mathbf{k}_{\parallel_r} \equiv d\mathbf{k}_{\parallel}, \tag{1.16}$$

$$d\varepsilon_r = d\varepsilon_l = dE, \tag{1.17}$$

and

$$D_{rl} = D_{lr} = D(E, \mathbf{k}_{\parallel}; V). \tag{1.18}$$

In Eq. 1.18, the first equality follows from time-reversal symmetry, and the last term indicates a possible dependence of the transmission probability arising out of a modification of the shape of the tunneling barrier with applied bias.

Equation 1.14 is general and is the starting point for calculations of elastic tunneling through a potential barrier, including the SB structure. However, the theory that led to this equation is not exact. We have assumed an effective

mass picture for the electrons on either side of the barrier. The wave-function-matching conditions which lead to the transmission probability are imposed on the envelope functions rather than on the Bloch wave functions themselves. We have ignored internal degrees of freedom of the system (inelastic tunneling). Further, no satisfactory account has been taken of the details of the interface. Hence, we must be careful to recognize that any calculations based on Eq. 1.14 will not in general be *quantitatively* correct. With these cautions noted, let us proceed.

In order to calculate the I-V characteristics of a tunnel junction with the use of Eq. 1.14, it is necessary to first calculate the transmission probability $D(E,\mathbf{k}_{\parallel};V)$. In principle, D can be calculated by solving the effective mass Schrödinger equation $H_{\text{eff}}\psi = E\psi$ exactly for the case of interest. D is then obtained by imposing the boundary conditions on the effective mass wave function ψ. Unfortunately, analytical solutions are possible only in a few very simple cases. The more common approach is to approximate D by the WKB method as

$$D(E,\mathbf{k}_{\parallel};V) = \left|\exp\left\{2i\int_{x_l}^{x_r} k_x\left[E,\mathbf{k}_{\parallel},\phi_V(x)\right]\,dx\right\}\right|, \tag{1.19}$$

where k_x is the wave vector for a specified total energy E and parallel wave vector $\mathbf{k}_{\parallel}$ within the potential barrier $\phi_V(x)$, whose shape depends on the applied bias V. The limits of integration x_l and x_r are the classical turning points of the electron within the barrier, and are specified by the condition, $\phi_V(x_{\{l,r\}}) = E$. Within the tunneling barrier, k_x is either complex or purely imaginary. Hence, Eq. 1.19 gives a transmission probability characteristic of single-barrier tunneling which goes like $e^{-\overline{\kappa}w}$, where w is the barrier width and $\overline{\kappa}$ is a mean attenuation constant, determined by the barrier shape, and by E and $\mathbf{k}_{\parallel}$. The value of k_x is calculated by analytic continuation of the bulk band structure, usually given in terms of some model calculation, subject to the condition $E = \varepsilon(k_x,\mathbf{k}_{\parallel}) + \phi_V(x)$ within the barrier, where $\phi > E$. The advantage of the WKB approach is

that the case of arbitrary models for the bulk band structure, and modifications of the potential barrier to reflect realistic effects such as screening or image forces, are handled relatively easily. In general, the transmission probability given by Eq. 1.19 is correct apart from a prefactor reflecting a DOS, an error which becomes important for states near band edges (where the DOS approaches zero).[51]

Inelastic Tunneling

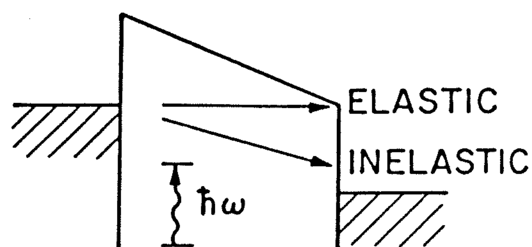
Inelastic electron tunneling occurs when the initial-state and final-state energies of an electron tunneling through a barrier are different. Since total energy must be conserved, the energy difference is taken up by a phonon or other elementary excitation which is emitted or absorbed in the tunneling process. For concreteness, let us restrict our attention to phonon processes. At low temperatures, the thermal background of phonons is so small that only the phonon emission process is significant. Figure 1.13(a) shows an inelastic electron tunneling process assisted by the emission of a phonon. At $T = 0\,\mathrm{K}$, a phonon can be emitted only when the bias across the tunnel junction qV exceeds the phonon energy $\hbar\Omega$ because only then is it possible for an electron to tunnel into a state of energy $\hbar\Omega$ below its initial energy level. Thus, when the bias voltage threshold $\hbar\Omega/q$ is exceeded, an inelastic tunneling current J_{in} is effectively switched on in parallel with the elastic current J_{el}. For non-zero temperature, this turn-on becomes thermally smeared on the order of kT, with J_{in} near the threshold being given by the expression[53]

$$J_{\mathrm{in}} = CkT\frac{x}{1-e^{-x}}, \tag{1.20}$$

where

$$x \equiv (qV - \hbar\Omega)/kT,$$

and C is a constant. The total current density J is the sum of the parallel



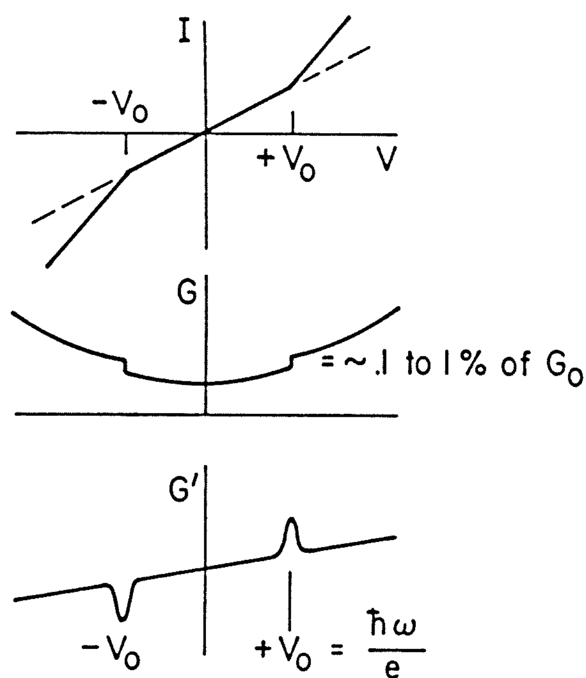


Figure 1.13: (a) Inelastic tunneling of an electron, assisted by the emission of a phonon of energy $\hbar\Omega$. (b) Effect of inelastic tunneling process on the I-V characteristic and its first and second derivatives. (From Ref. 52.)

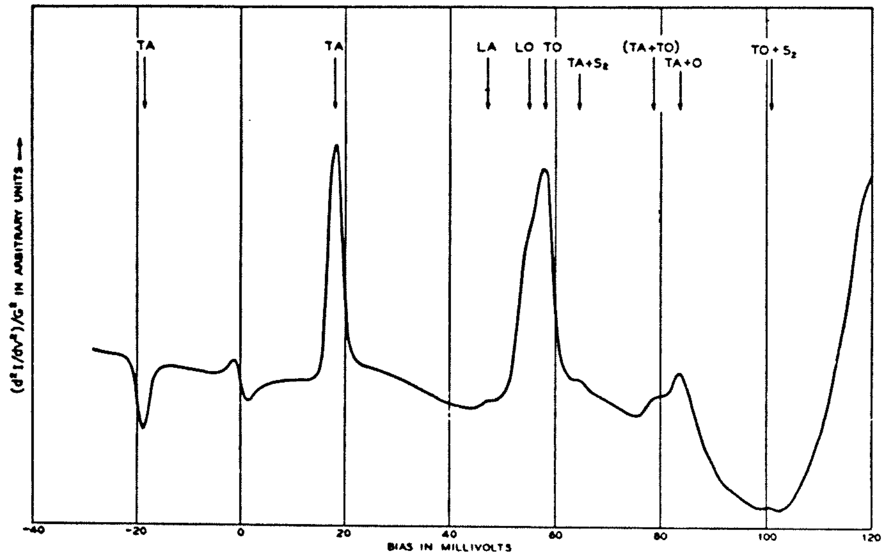
contributions

$$J = J_{\text{el}} + J_{\text{in}}. \tag{1.21}$$

The resultant change in slope of the *I-V* curve near a threshold $V = \hbar\Omega/q$ is indicated in Fig. 1.13(b). The presence of this (usually imperceptible) change in slope is made clearer by considering higher derivatives of the *I-V* curve: in the first derivative the threshold appears as a step, and in the second derivative as a peak. Hence, a plot of d^2I/dV^2 vs. V results in a spectrum of peaks whose positions allow for identification of vibrational and other excitations within or at the surface of the tunneling barrier. The direct measurement of d^2I/dV^2 vs. V forms the basis of IETS.

Phonon-assisted electron tunneling through an ideal SB is subject to the conservation of total energy. Further selection rules are dependent upon the nature of the inelastic process. Before considering phonon-assisted tunneling in SB structures, let us first examine the inelastic tunneling in a Si degenerate *p-n* (Esaki) diode. An actual tunneling spectrum is shown in Fig. 1.14(a).[54] Figure 1.14(b) shows the bulk phonon and electron band structures of Si.[55] In the Esaki diode, electrons near the indirect conduction band minimum ($\mathbf{k} = \mathbf{k}_0 \neq \vec{0}$) tunnel across the band gap in the depleted region into available hole states near $\mathbf{k} = \vec{0}$. The difference in electron wave vector states $\mathbf{k}_0$ is taken up by the emission of bulk phonons. As seen in the figure, at $\mathbf{k}_0$ there are four possible phonon frequencies corresponding to the TA, LA, LO, and TO phonon modes, all of which are observed experimentally. In fact, for the Esaki diode, the emission of wave-vector-conserving phonons is necessary for electronic tunneling transport.

In contrast to the Esaki diode, inelastic tunneling through a SB is a much less significant process, and hence, much harder to observe. This is because the final electron states (within qV of the Fermi surface of the metal) span a large range in $\mathbf{k}$-space rather than being confined near $\mathbf{k} = \vec{0}$. Hence, it is possible for the



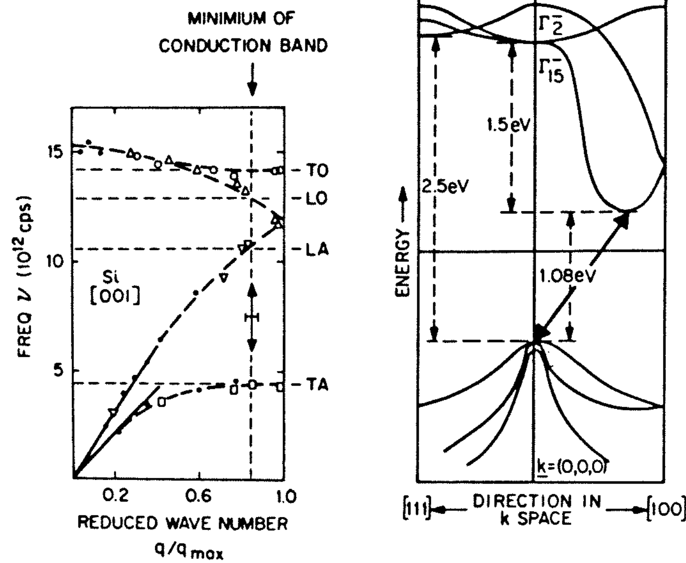


Figure 1.14: (a) Inelastic electron tunneling spectrum for a Si Esaki diode at 4.2 K (from Ref. 54). (b) Phonon dispersion and electronic band structure for bulk Si (from Ref. 55).

indirect-valley electrons to tunnel elastically without the need to emit a wave-vector-conserving phonon. The elastic tunneling process completely dominates the electron transport in degenerate SB structures. We have just seen that, for the Esaki diode, it was *total* $\mathbf{k}$ rather than total $\mathbf{k}_{\parallel}$ which was conserved in the tunneling process. While total $\mathbf{k}_{\parallel}$ should be rigorously conserved across a planar tunnel junction, evidently, the electric field within the Esaki diode depletion region represents a perturbation strong enough to cause interband tunneling, but not strong enough to break the three-dimensional lattice translational symmetry completely. Similarly, for the SB, inelastic processes excited *within* the depletion layer should be essentially bulklike and thereby conserve total wave vector $\mathbf{k}$. On the other hand, processes which are a fundamental excitation of a perfect interface *e.g.*, local interface phonons, should be expected to conserve overall $\mathbf{k}_{\parallel}$ whereas those associated with a disordered interface should not. Evidence for $\mathbf{k}$-conservation has already been observed in SB tunneling for the case of polycrystalline metal films on a semiconductor.[56] It is the purpose of the IETS measurements reported in Chapter 3 on *single-crystal* $CoSi_2$/Si and $NiSi_2$/Si SB diodes to look for the first evidence of $\mathbf{k}_{\parallel}$-conservation in assisted metal-semiconductor electron tunneling.

1.4 Vibrational Modes of Defects

Chapter 5 of this thesis presents theoretical work on the calculation of vibrational modes of point-defects in semiconductors. The point-defects considered explicitly in this work are impurity atoms, both substitutional and interstitial, within an otherwise perfect host lattice. The presence of a substitutional impurity atom can modify the vibrational spectrum of the lattice in two ways. First, the mass of the impurity in general will differ from that of the original atom.

Secondly, the interatomic interactions near the substituent site in general will be modified due to the different bonding chemistry of the impurity. Both of these effects break the lattice translational symmetry and give rise to vibrational modes which are spatially confined to the vicinity of the impurity. These "local" vibrational modes (LVM) are also obtained for interstitial impurities. Optical measurements on semiconductors containing electronic deep-level impurities provide one means for observing LVMs. Characterization of an impurity in terms of its LVM spectra provide information about the atomic motion near the impurity. Such information is helpful in telling us how strongly an impurity is bound to the lattice, and what the local lattice relaxation might look like. In the remainder of this section, we shall discuss the nature of LVMs, and their relation to optical spectra. The discussion of how such modes are calculated is deferred until Chapter 5.

1.4.1 Localized and Resonant Modes

It is well known that normal vibrational modes of a perfect crystal are delocalized plane waves. Let us consider the modes of a lattice containing a substitutional impurity, ignoring for the moment any modification of the interaction with neighboring atoms. First of all, many of the bulk modes will remain unaffected by the substitution of the impurity (*e.g.*, any mode which has a node at the impurity site). However, there will be certain frequencies for which most of the energy of vibration (displacement amplitude) occurs within the immediate neighborhood of the impurity. For example, let us suppose that the impurity atom is less massive than any bulk atom. In this case, there exists a particular frequency above the bulk cutoff frequency at which the impurity may oscillate. Since the bulk crystal cannot sustain oscillation above its cutoff frequency, the amplitude of oscillation is localized, the motion being damped roughly like $e^{-\alpha R}/R$ as one

moves a distance R from the impurity. Such a mode is known as a *localized* mode.[57] This situation is illustrated for a linear diatomic chain in Fig. 1.15(a). Localized modes occur at forbidden frequencies in the bulk phonon spectrum, either within gaps or above the cutoff frequency. On the other hand, it is possible that modes of enhanced impurity atom motion may lie within the allowed bulk phonon spectrum. In this case, illustrated in Fig. 1.15(b), the atomic motion is largest at the impurity but does not vanish as one moves far from the impurity site.[57] This situation is known as a *resonant* "mode."* Localized modes occur at discrete frequencies whereas resonant modes have a finite width because of the possibility of energy loss to the lattice. Both localized and resonant modes involve substantial vibrational motion in the region near the point-defect. Henceforth, the term LVM shall be used to include both localized and resonant modes.

The physical origin of LVMs is readily understood. In the case of the simple mass substitution just considered, it is clear that if we immobilized the entire lattice except for the impurity atom, the impurity would have some natural frequency of oscillation ω_0. The actual imperfect lattice may be viewed as a coupling between this mode ω_0 and the bulk phonon modes. In the limit of weak coupling between these two "subsystems," the case of ω_0 occurring at forbidden bulk frequencies leads to a localized mode of the total system. Similarly, a resonant mode can be viewed as bulk phonon modes driving the natural impurity motion near ω_0 in the case that ω_0 is an allowed frequency of the lattice. A somewhat different view of the interaction between bulk and impurity is adopted in the Green's function approach to be discussed in Chapter 5.

For concreteness, we have discussed localized and resonant modes in terms of a particular kind of point-defect; namely, a simple mass substitution. However,

*Strictly speaking, a resonant mode is *not* a normal, *i.e.*, stationary mode of the system since any vibrational energy initially localized near the defect can radiate away. If the resonance is sharp, however, the lifetime may be long and the mode *quasi*stationary.

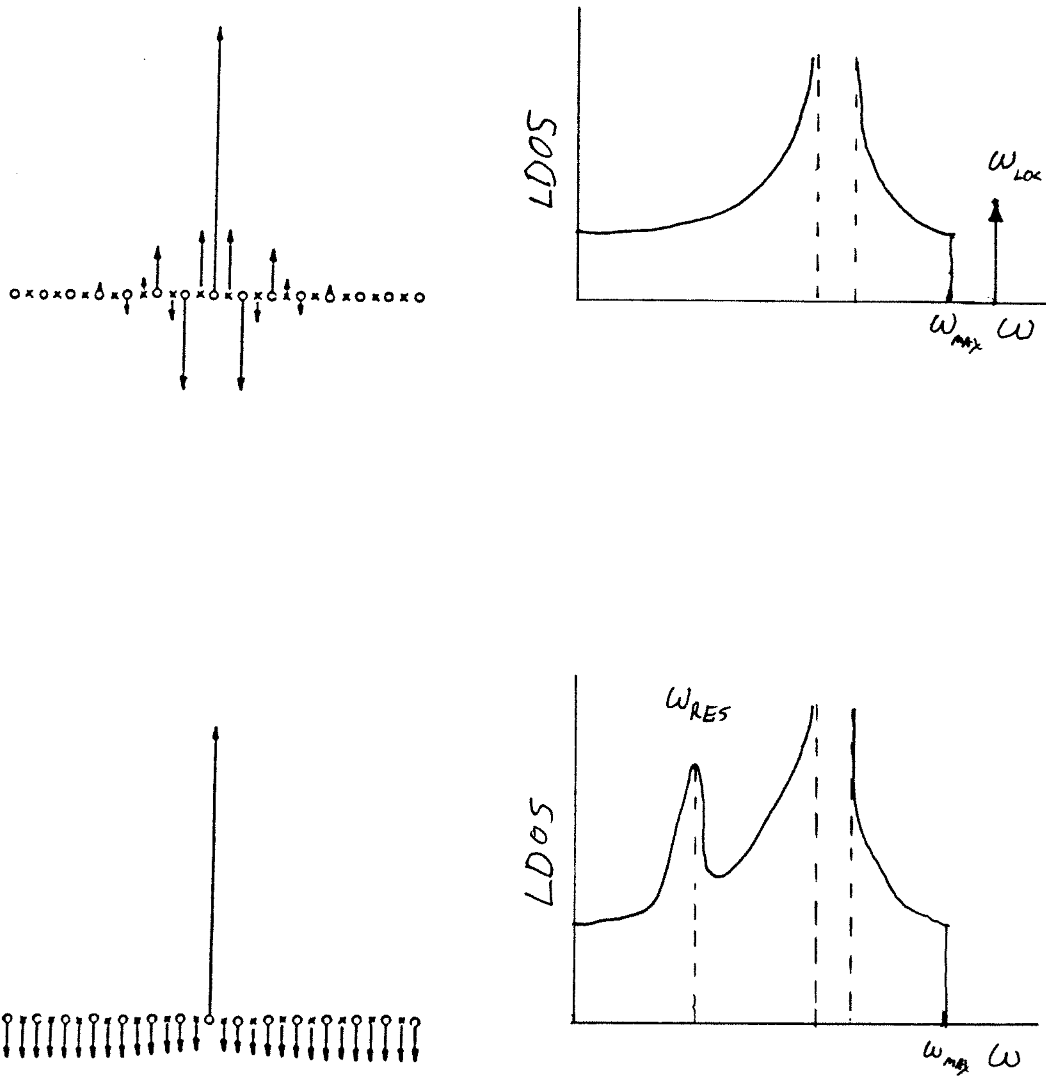


Figure 1.15: (a) Localized vibrational mode spatial (upper left) and frequency (upper right) distribution for a diatomic linear chain above the cutoff frequency. (b) Acoustic branch resonant mode spatial (lower left) and frequency (lower right) distributions. (From Ref. 57.)

it can be shown that LVM behavior can occur in general for other point-defects, including vacancies, interstitials, and defect complexes with internal vibrational degrees of freedom.[58] In the harmonic approximation, all interatomic interactions are described in terms of "force constants." In general, the introduction of impurities modifies the interaction with and among neighboring atomic sites. Hence, a general point-defect is modeled as a mass perturbation plus a perturbation of certain force constants. It can be shown quite generally that LVMs can arise out of perturbation of a combination of mass and force constants.[58]

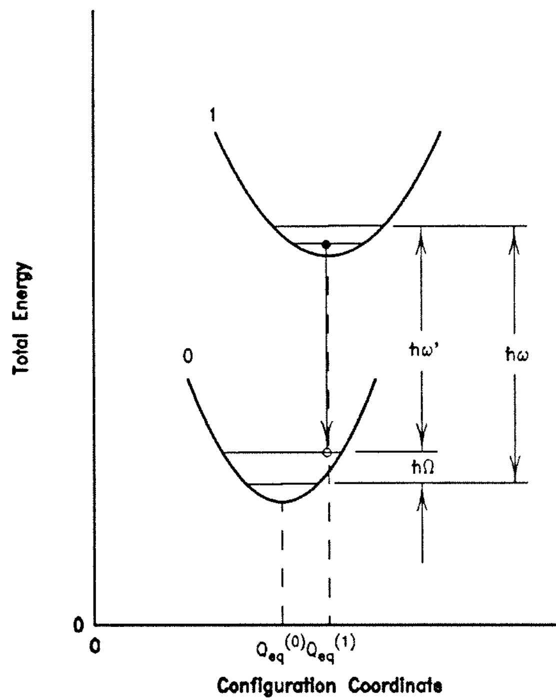
1.4.2 Relation to Optical Spectra

Optical spectroscopy provides a useful characterization tool in the very important and difficult theoretical problem[59] of electronic deep-level defects in semiconductors. Deep levels by definition are defects which present electrically active states deep within the band gap of a semiconductor, and often arise due to the presence of size-mismatched impurities such as transition-metal atoms. The atomic size mismatch leads to impurity electronic and vibrational states which are spatially localized near the impurity. The LVMs may couple strongly to the deep electronic levels, or to incident photons. Photoluminescence spectra obtained from deep levels are often informatively structured with peaks corresponding to principal and local-phonon-assisted electronic transitions. IR absorption spectroscopy on deep levels may also reveal the presence of IR-active LVMs. As mentioned earlier, a quantitative understanding of LVMs of a particular deep level impurity can be used to gain a further understanding of how that impurity is bound to the lattice.

One way in which LVMs are observed is through the photoluminescence from deep level impurities. The processes which lead to this photoluminescence have been described in detail by Feenstra.[60] Briefly, the situation for a deep donor level

might be the following: incident photons excite a deep-level ground-state electron up into the conduction band; the electron relaxes nonradiatively to the band edge through bulk phonon emission; the electron becomes trapped in a defect excited state, relaxes nonradiatively through successive excited levels; and finally, when the energy jump becomes large enough to make the radiative process competitive, the electron makes a transition to the ground state and emits a characteristic photon $\hbar\omega$. The minimum-energy spatial configuration of the impurity atom and its immediate neighbors depends on the electronic state of the defect as illustrated in Fig. 1.16(a). If there is a difference in configuration between excited and ground electronic states, the radiative transition to the ground state may leave the defect in a *vibrationally* excited state. In other words, the last step may take place with the emission of a photon $\hbar\omega'$ and a phonon $\hbar\Omega$, where $\hbar\omega' + \hbar\Omega = \hbar\omega$ is the energy difference between electronic excited and ground states. The phonon-assisted process is shown in Fig. 1.16(a). If the wave functions of the electronic levels are sufficiently localized, the phonon will likely be a LVM of the defect.[60] Hence, in photoluminescence spectra, LVMs are observed as subsidiary peaks shifted in energy by an amount $\hbar\Omega$ (or multiples thereof) below a principal no-phonon peak.

LVMs may also be observed by IR absorption spectroscopy. This process is shown in Fig. 1.16(b). In this case, the LVM $\hbar\Omega$ is coupled directly to the incident photon $\hbar\omega_i$ through the electric dipole moment formed by the effective charge centered on the impurity and neighboring atomic sites. The electronic state does not change in this process; only the vibrational state changes. Resonant absorption of the incident photon by a LVM phonon takes place when $\hbar\omega_i = \hbar\Omega$. Only those LVMs whose motion involves a change in dipole moment are IR-active and can resonantly absorb incident electromagnetic radiation.[61] In Chapter 5, we shall compare the results of calculated LVM spectra with photoluminescence and



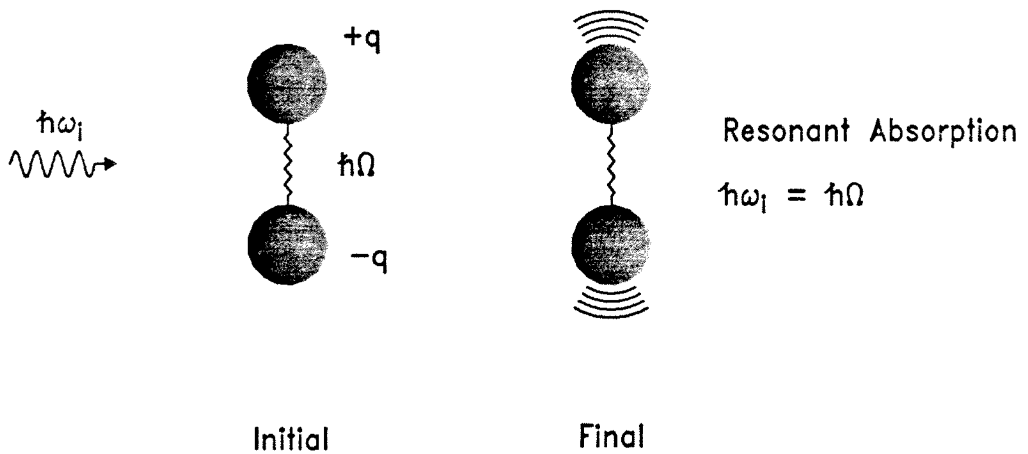


Figure 1.16: (a) Electronic transition ($1 \rightarrow 0$) between localized deep level states in which a LVM is excited. The configuration coordinate Q specifies the position of the impurity and surrounding bulk atoms. (b) Resonant absorption of an incident photon by electric dipole coupling to a LVM.

IR absorption measurements for selected impurities in Si, GaP, and GaAs.

1.5 Outline of Thesis

In Chapter 2, the results of SBH measurements on type-A and type-B $NiSi_2$/Si single-crystal structures are presented. The photoresponse and forward *I-V* methods are used in the determination of the SBH. The principal finding is that there is a substantial difference (greater than 0.1 eV) in SBH between the two cases. In addition, it is observed that the photoresponse and *I-V* measurements for the type-B structures yield an apparently inconsistent value for the SBH, and that the type-B photoresponse curves exhibit an unusual shape. It is shown that both the detailed shape of the photoresponse data and the apparent inconsistency with *I-V* results can be *quantitatively* accounted for in terms of a double Schottky contact model consisting of high and low SB regions electrically in parallel. Possible physical origins of this phenomenological model are discussed.

Chapter 3 presents the results of electron tunneling spectroscopy measurements performed on single-crystal films of $CoSi_2$ and type-A and type-B $NiSi_2$ on degenerate, *n*-type Si. The principle motivation for this work was to look for evidence of $\mathbf{k}_{\|}$-conservation and other enhancements of the IETS spectra, relative to that obtained for polycrystalline metal films. Our results do *not* show any evidence of total-$\mathbf{k}_{\|}$-conservation, but do show clear evidence of total-$\mathbf{k}$-conservation in assisted electron tunneling. In all cases we observe evidence of bulk Si $\mathbf{k}$-conserving phonons assisting the electron transport from the semiconductor into the silicide. No particular silicide-dependent or orientation-dependent features are observed, and the results obtained on nominally single-crystal films are essentially identical to those obtained for Au films thermally evaporated onto

similar substrates. As expected, the tunneling current is completely dominated by elastic tunneling. Comparison of our $CoSi_2$ tunneling I-V data with calculated WKB curves show reasonable quantitative agreement, and comparison of measured and calculated dV/dI spectra suggest the presence of impurity banding of the As donors.

In Chapter 4, the design and construction of a Si MBE growth chamber is presented. An attempt has been made in the design of this chamber to address many of the problems currently unresolved in the existing design of commercial Si MBE machines: in particular, the problem of unintentional B-doping of the epilayer, Si flaking in the chamber, and the temperature uniformity and measurement of the substrate. Many of the key aspects of this system, including the sample heater, substrate holder, load-lock mechanism, LN_2 shrouds, water-cooled shutters, etc., represent original conceptual and design work and were fabricated at Caltech. A discussion of practical considerations for the growth chamber and what should (and should not) go in it, and preliminary results, are also presented.

Finally, the thesis concludes with Chapter 5, which concerns the theoretical calculation of LVMs for selected point-defects in semiconductors. The following systems are considered: substitutional oxygen on a phosphorus site in GaP ($GaP:O_P$), substitutional and interstitial iron in Si (Si:Fe), and substitutional carbon on an arsenic site in GaAs ($GaAs:C_{As}$). In each case, comparison is made to available photoluminescence or IR absorption spectra. For the extensively studied $GaP:O_P$ defect we find detailed agreement between theory and various photoluminescence measurements with respect to LVM energies, symmetries, and isotope-shift behavior. Comparison of theory and experiment suggest that the oxygen atom is weakly bound in GaP, with an effective force constant which is 25% of the bulk value for a phosphorous site. Our theoretical investigations of

Si:Fe predict a low-energy LVM involving motion of the Fe impurity. In theory, this mode should shift in energy upon isotopic substitution of the impurity, in *disagreement* with photoluminescence spectra containing the expected low-energy mode. Lastly, the effect of the natural isotopic variation of Ga in GaAs on the IR-active LVM of the GaAs:C_{As} system is investigated. The calculated results show detailed agreement with IR absorption data with respect to LVM splittings and relative intensities, and establish conclusively the identity of the impurity substitutional site as an As site.

References

1. Y. Ota, Thin Solid Films **106**, 3 (1983).

2. E. Rosencher, S. Delage, Y. Campidelli, and F. Arnaud D'Avitaya, Electron. Lett. **20**, 762 (1984).

3. J. C. Hensel, A. F. J. Levi, R. T. Tung, and J. M. Gibson, Appl. Phys. Lett. **47**, 151 (1985).

4. R. T. Tung, J. M. Gibson, and A. F. J. Levi, Appl. Phys. Lett. **48**, 1264 (1986).

5. J. C. Bean, *Science* (to be published).

6. H. Daembkes, H.-J. Herzog, H. Jorke, H. Kibbel, and E. Kasper (to be (published).

7. R. W. Fathauer and L. J. Schowalter, Appl. Phys. Lett. **45**, 519 (1984).

8. M.-A. Nicolet and S. S. Lau, in *VLSI Electronics*, Vol. 6, edited by Norman G. Einspruch and Graydon B. Larrabee (Academic, New York, 1983), Chap. 6.

9. S. R. Hetzler, private communication.

10. A. Zur, T. C. McGill, and M.-A. Nicolet, J. Appl. Phys. **57**, 600 (1985).

11. J. C. Hensel, R. T. Tung, J. M. Poate, and F. C. Unterwald, Appl. Phys. Lett. **44**, 913 (1984).

12. T. R. Harrison, A. M. Johnson, P. K. Tien, and A. H. Dayem, Appl. Phys. Lett. **41**, 734 (1982).

13. J. Lambe, private communication.

14. M. C. Bost and J. E. Mahan, J. Appl. Phys. **58**, 2696 (1985).

15. D. Shinoda, S. Asanabe, and Y. Sasaki, J. Phys. Soc. Jpn. **19**, 269 (1964).

16. G. V. Samsonov, *No. 2 Properties Index* (Plenum, New York, 1964), p. 161.

17. H. Jorke, H.-J. Herzog, and H. Kibbel, Appl. Phys. Lett. **47**, 511 (1985).

18. R. People, Phys. Rev. B **32**, 1405 (1985).

19. R. People and J. C. Bean, Appl. Phys. Lett. **48**, 538 (1986).

20. D. V. Lang, R. People, J. C. Bean, and A. M. Sergent, Appl. Phys. Lett. **47**, 1333 (1985).

21. O. J. Marsh, private communication.

22. B. D. Hunt, L. J. Schowalter, N. Lewis, E. L. Hall, R. J. Hauenstein, T. E. Schlesinger, T. C. McGill, Masako Okamoto, and Shin Hashimoto, in *Proceedings of the Materials Research Society*, Boston, 1985 (to be published).

23. A. Ishizaka and Y. Shiraki, J. Electrochem. Soc. **133**, 666 (1986).

24. R. T. Tung, A. F. J. Levi, J. M. Gibson, K. K. Ng, and A. Chantre, in *Proceedings of the Materials Research Society*, Palo Alto, 1986 (to be published).

25. R. A. A. Kubiak, W. Y. Leong, and E. H. C. Parker, J. Vac. Sci. Technol. B **3**, 592 (1985).

26. R. T. Tung, J. M. Gibson, and J. M. Poate, Phys. Rev. Lett. **50**, 429 (1983).

27. R. T. Tung, J. M. Gibson, and J. M. Poate, Appl. Phys. Lett. **42**, 888 (1983).

28. M. Liehr, P. E. Schmid, F. K. LeGoues, and P. S. Ho, Phys. Rev. Lett. **54**, 2139 (1985).

29. E. Rosencher, S. Delage, F. Arnaud D'Avitaya, C. D'Anterroches, K. Belhaddad, and J. C. Pfister, Physica **134B**, 106 (1985).

30. R. T. Tung, A. F. J. Levi, and J. M. Gibson, Appl. Phys. Lett. **48**, 635 (1986).

31. S. Delage, P. A. Badoz, E. Rosencher, and F. Arnaud D'Avitaya, Electron. Lett. **22**, 207 (1986).

32. Julia M. Phillips and W. M. Augustyniak, Appl. Phys. Lett. **48**, 463 (1986).

33. R. W. Fathauer, B. D. Hunt, L. J. Schowalter, Masako Okamoto, and Shin Hashimoto, Appl. Phys. Lett. **49**, 64 (1986).

34. Th. Ricker and E. Kasper, J. Phys. (to be published).

35. K. Nakagawa and Y. Shiraki, in *Proceedings of the Second International Conference on Modulated Semiconductor Structures*, Kyoto, 1985, p. 724.

36. J. C. Bean and P. Butcher, in *Proceedings of the First International Symposium on Silicon Molecular Beam Epitaxy*, edited by J. C. Bean (The Electrochemical Society, Pennington, 1985), p. 429.

37. E. H. Rhoderick, *Metal-Semiconductor Contacts* (Clarendon, Oxford, 1978), Chap. 1.

38. W. Schottky, Naturwissenschaften **26**, 843 (1938).

39. N. F. Mott, Proc. Cambridge Philos. Soc. **34**, 568 (1938).

40. J. Bardeen, Phys. Rev. **71**, 717 (1947).

41. E. H. Rhoderick, *Metal-Semiconductor Contacts*, Ref. 37, Chap. 2.

42. A. J. Bennet and C. B. Duke, Phys. Rev. **162**, 578 (1967).

43. C. Mailhiot and C. B. Duke, Phys. Rev. Lett. (to be published); Phys. Rev. B (to be published).

44. S. M. Sze, *Physics of Semiconductor Devices*, 2nd ed. (Wiley, New York, 1981), Chap. 5.

45. C. R. Crowell and S. M. Sze, Solid State Electron. **9**, 1035 (1966).

46. H. Norde, J. Appl. Phys **50**, 5052 (1979).

47. J. I. Pankove, *Optical Processes in Semiconductors* (Dover, New York, 1971), pp. 292ff.

48. R. H. Fowler, Phys. Rev. **38**, 45 (1931).

49. T. Okumura and K. N. Tu, J. Appl. Phys. **54**, 922 (1983).

50. C. B. Duke, in *Tunneling Phenomena in Solids*, edited by Elias Burstein and Stig Lundqvist (Plenum, New York, 1969), Chap. 4.

51. R. T. Shuey, in *Tunneling Phenomena in Solids*, Ref. 50, Chap. 7.

52. R. C. Jaklevic, in *Inelastic Electron Tunneling Spectroscopy*, edited by T. Wolfram (Springer, Berlin, 1978), Chap. 1.

53. R. T. Collins, Ph.D. Thesis, California Institute of Technology, 1985.

54. A. G. Chynoweth, R. A. Logan, and D. E. Thomas, Phys. Rev. **125**, 877 (1962).

55. R. T. Logan, in *Tunneling Phenomena in Solids*, Ref. 50, Chap. 11.

56. F. Steinrisser, L. C. Davis, and C. B. Duke, Phys. Rev. **176**, 912 (1968).

57. O. Madelung, *Introduction to Solid State Theory* (Springer, Berlin, 1978), Chap. 9.

58. A. A. Maradudin, E. W. Montroll, G. H. Weiss, and I. P. Ipatova, in *Solid State Physics*, edited by Henry Ehrenreich, Frederick Seitz, and David Turnbull, Suppl. 3, 2nd ed. (Academic, New York, 1971), Chap. 7.

59. A. Prabhakar, Ph.D. Thesis, California Institute of Technology, 1985.

60. R. M. Feenstra, Ph.D. Thesis, California Institute of Technology, 1982.

61. G. Herzberg, *Molecular Spectra and Molecular Structure*, Vol. 2 (Van Nostrand, Princeton, 1945), Chap. 3.

Chapter 2

Schottky Barrier Height Measurements at Single-Crystal $NiSi_2$/Si Interfaces

2.1 Introduction

2.1.1 Background

Since 1983, there has been much interest in single-crystal $NiSi_2$ thin films grown epitaxially on Si by molecular beam epitaxy (MBE).[1,6,7,4] As mentioned in Sections 1.2.2 and 1.3, $NiSi_2$ is one of two metallic silicides which closely lattice-matches Si ($< 0.4\%$) and has been successfully grown on Si by MBE to form a nearly ideal metal-semiconductor interface structure. We have also mentioned that $NiSi_2$ is unique in that single-crystal silicide films can be be fabricated on (111)-Si in one of two distinct crystallographic orientations with respect to the substrate (type-A and type-B).[1] For these reasons, the $NiSi_2$/Si structures provide novel, well-characterized interfaces for the study of Schottky barrier (SB) formation, and have recently been the subject of active investigation.[6,7,4] Re-

cently, it was reported by Tung[6] that the SBHs of type-A and type-B $NiSi_2/Si$ structures differ substantially (greater than 0.1 eV). Later, this finding was disputed in the work of Liehr, Schmid, LeGoues, and Ho,[7] who reported that the SBH depends *not* on the *orientation* of the epitaxial silicide layer but rather on the *structural perfection* of the $NiSi_2/Si$ interface. Should the SBH prove indeed to depend on silicide orientation, the finding would have serious implications for basic theories of SB formation,[5,6,7] which at present do not predict such an orientation dependence.

At present, there are two distinct explanations[6,7] for the observed difference of barrier heights of $NiSi_2/Si$ systems. Here, we present the results of our independent study undertaken to help resolve this issue. Since the effects of processing steps, in particular, the substrate surface preparation, on the characteristics of these high-quality $NiSi_2/Si$ structures are not completely known, we believed it of value to perform SBH measurements on samples in which the processing details were varied from those of the previous two studies. There are two important differences between our work and that of Tung, and Liehr *et al.* First, in our work, the substrate surface precleaning is achieved with an Ar sputter-cleaning technique followed by an 850 °C anneal, rather than the Shiraki and *high temperature* thermal cleaning methods used by Tung,[6] and Liehr *et al.*,[7] respectively. The second difference involves the use of the photoresponse technique (see Section 1.3.2) in determining SBH, which as we shall see in Section 2.5, gives more information about the SB composition in the event that it is actually a *compound* barrier.[8] (In addition, a third difference, now also believed to be significant, concerns the different substrate doping levels employed in the three studies.[6,7,4])

2.1.2 Results of this work

The SBHs of high quality MBE-grown, type-A and type-B $NiSi_2$ films on

(111)-Si have been measured by photoresponse and forward I-V methods. Both photoresponse and I-V techniques demonstrate a clear difference in SBH between the two silicide orientations. Our SBH results are in good agreement with those of Tung,[6] but are in disagreement with the results of Liehr *et al.*[7] For type-A structures, both methods are found to consistently yield a SBH of $\phi_{Bn} = 0.62 \pm 0.01\,\text{eV}$. However, the SBH of our type-B structures are less clear: from I-V measurements we consistently obtain a value of $\phi_{Bn} = 0.69 \pm 0.01\,\text{eV}$, and from photoresponse, a value of $\phi_{Bn} = 0.77 \pm 0.05\,\text{eV}$. Interpretation of the photoresponse results is complicated by the fact that the type-B films consistently exhibit an unusual "bowing" in their Fowler plots.[4] Fortunately, both the unusual shape of the type-B photoresponse curves *and* the numerical result of the I-V measurement are quantitatively accounted for by adopting a phenomenological model of the type-B SB structure.[9] In this model, the actual barrier is represented as a mixture of local regions of low and high SBH. With the use of this phenomenological model, all of our type-B SBH data are consistently explained as a double Schottky contact consisting of 9% areal coverage of low SBH ($\phi_{lo} = 0.64 \pm 0.01\,\text{eV}$) and 91% coverage of high SBH ($\phi_{hi} = 0.81 \pm 0.01\,\text{eV}$).

2.1.3 Outline of Chapter

In Section 2.2, a derivation of the photoresponse of an ideal Schottky barrier at $T = 0\,\text{K}$, and a simple geometric interpretation of the quadratic threshold, which can be easily generalized to finite temperatures, are presented. Next, the silicide fabrication, device preparation, and experimental apparatus used in the photoresponse measurement are discussed in Section 2.3. The results of our photoresponse and I-V measurements of SBH are presented in Section 2.4. The unusual results obtained on our type-B $NiSi_2$ films leads us to formulate a phenomenological model to explain the data; this model is defined in Section 2.5

and is seen to account for the photoresponse and I-V measurements in detail. In the same section, we also compare the results of this study with previous work reported on single-crystal $NiSi_2$/Si structures. Then, Section 2.6 contains a discussion of the interpretation and possible physical origin of our results. Finally, the conclusions from this study are presented in Section 2.7.

2.2 Theory of Photoresponse

The general result for the near-threshold photoresponse* of a SB at finite temperature has already been given on page 35. Now, we wish to derive the corresponding result at $T = 0\,\mathrm{K}$, and to show that the photoresponse does indeed turn on quadratically in the quantity, $(\hbar\omega - \phi_{\mathrm{B}n})$. We discussed the physical process of internal photoemission of an electron over a SB in Section 1.3.2. There, we noted that, for an electron to surmount the barrier, it was required to move *toward* the barrier with an energy component ε_x which satisfies the condition

$$\varepsilon_x > \xi, \tag{2.1}$$

where

$$\xi \equiv \phi_{\mathrm{B}n} - \hbar\omega + \varepsilon_F. \tag{2.2}$$

The photoresponse will be proportional to the *number* of electrons N_{ph} which lie in initial states in the metal which satisfy condition 2.1. Hence, the problem is essentially to calculate

$$N_{\mathrm{ph}} = \frac{1}{2}\int_{\varepsilon_x > \xi} d\varepsilon\, \rho(\varepsilon); \qquad \xi \leq \varepsilon_F. \tag{2.3}$$

Here, ρ is the electronic density-of-states (DOS) and ε the electron energy in the metal, and the prefactor 1/2 takes account of the fact that we only want to

*By the term "photoresponse" we mean the photoinduced current or voltage *per* incident photon.

count electrons in initial states with group velocities directed toward the barrier. To accomplish the integration indicated in Eq. 2.3, we integrate separately over perpendicular and parallel energy components, obtaining

$$\begin{aligned} N_{\text{ph}} &= \frac{1}{2}\int_{\xi}^{\varepsilon_F} d\varepsilon_x\, \rho_x(\varepsilon_x) \int_0^{\varepsilon_F-\varepsilon_x} d\varepsilon_\parallel\, \rho_\parallel(\varepsilon_\parallel) \\ &= \frac{1}{2}\int_0^{\hbar\omega-\phi_{\mathrm{B}n}} dE_x\, \rho_x(\varepsilon_F - E_x) \int_0^{E_x} d\varepsilon_\parallel\, \rho_\parallel(\varepsilon_\parallel); \qquad \hbar\omega \geq \phi_{\mathrm{B}n}, \end{aligned} \tag{2.4}$$

where we have made use of Eq. 2.2, and have defined

$$E_x \equiv \varepsilon_F - \varepsilon_x.$$

Near the SB threshold, the condition $(\hbar\omega - \phi_{\mathrm{B}n}) \ll \varepsilon_F$ is well satisfied. Therefore, to a reasonable approximation, we can take the one-dimensional DOS ρ_x out of the integrand to get

$$N_{\text{ph}} \approx \frac{1}{2}\rho_x(\varepsilon_F) \int_0^{\hbar\omega-\phi_{\mathrm{B}n}} dE_x \int_0^{E_x} d\varepsilon_\parallel\, \rho_\parallel(\varepsilon_\parallel); \qquad \hbar\omega \geq \phi_{\mathrm{B}n}. \tag{2.5}$$

This last expression is about as far as we can go without making specific assumptions about the electronic structure of the metal. Following Fowler,[10] let us for simplicity assume that we are dealing with a free-electron metal. In this case, the functional forms of the DOS factors ρ_x and $\rho_\parallel$ are given by

$$\rho_x(\varepsilon) = \varepsilon^{-1/2} \times \text{const.}, \tag{2.6}$$

and

$$\rho_\parallel(\varepsilon) = \text{const.} \tag{2.7}$$

Substituting Eqs. 2.6 and 2.7 into Eq. 2.5 gives the desired result

$$N_{\text{ph}} = \frac{(\hbar\omega - \phi_{\mathrm{B}n})^2}{\sqrt{\varepsilon_F}} \times \text{const.}; \qquad \hbar\omega \geq \phi_{\mathrm{B}n}. \tag{2.8}$$

From Eq. 2.8 we see that N_{ph}, and hence the photoresponse, turn on quadratically above the SB threshold $\phi_{\mathrm{B}n}$. Below this threshold, the photoresponse in the low-temperature limit is zero.

A more complete analysis would have included additional factors in the integrals of the above equations corresponding to physical processes such as free-carrier absorption, and possibly, phonon scattering. However, the matrix element of, *e.g.*, the free-carrier absorption process, is at most a function of $\hbar\omega$, and is not expected to change rapidly in the neighborhood of $\hbar\omega \approx \phi_{\mathrm{B}n}$. Hence, just as we pulled the DOS factor ρ_x out of the integral in Eq. 2.5 above, to a good enough approximation we may regard such additional factors as constant in the range of interest here.

The quadratic behavior of the photoreponse threshold can be easily understood in terms of the simple picture shown in Fig. 2.1. Referring to the figure, the metal electrons at absolute zero lie in states within the Fermi sphere (radius k_F). The barrier to electrons is represented by a plane at the value of k_x corresponding to $\varepsilon_x = \varepsilon_F - (\hbar\omega - \phi_{\mathrm{B}n})$, where $k_x > 0$. As $\hbar\omega$ is varied, the plane moves with respect to the edge of the Fermi sphere by an amount δk, given by

$$\delta k \approx k_F \frac{(\hbar\omega - \phi_{\mathrm{B}n})}{2\varepsilon_F}; \qquad \delta k \ll k_F. \tag{2.9}$$

Positive displacement δk corresponds to movement of the plane toward the sphere, *i.e.*, in the $-k_x$-direction. For $\hbar\omega > \phi_{\mathrm{B}n}$, the plane intersects the sphere, and the electrons in states for which $k_x \geq (k_F - \delta k)$ may give rise to photoresponse. Now, the interpretation of the quadratic threshold is clear: the photoresponse is proportional to the *volume* of the Fermi surface which lies to the right of the plane at $k_x = k_F - \delta k$. For small enough δk that Eq. 2.9 is valid, this volume is proportional to δk^2, and, hence, from Eq. 2.9, to $(\hbar\omega - \phi_{\mathrm{B}n})^2$. For negative values of δk, *i.e.*, for $\hbar\omega < \phi_{\mathrm{B}n}$, *no* filled electron states lie to the right of the plane (at $T = 0\,\mathrm{K}$) and so there is no photoresponse in this case.

With this geometric interpretation it is also easy to understand the photoresponse behavior at finite temperatures. As the temperature is increased, the sharp edge of the Fermi surface becomes thermally smeared on the order of

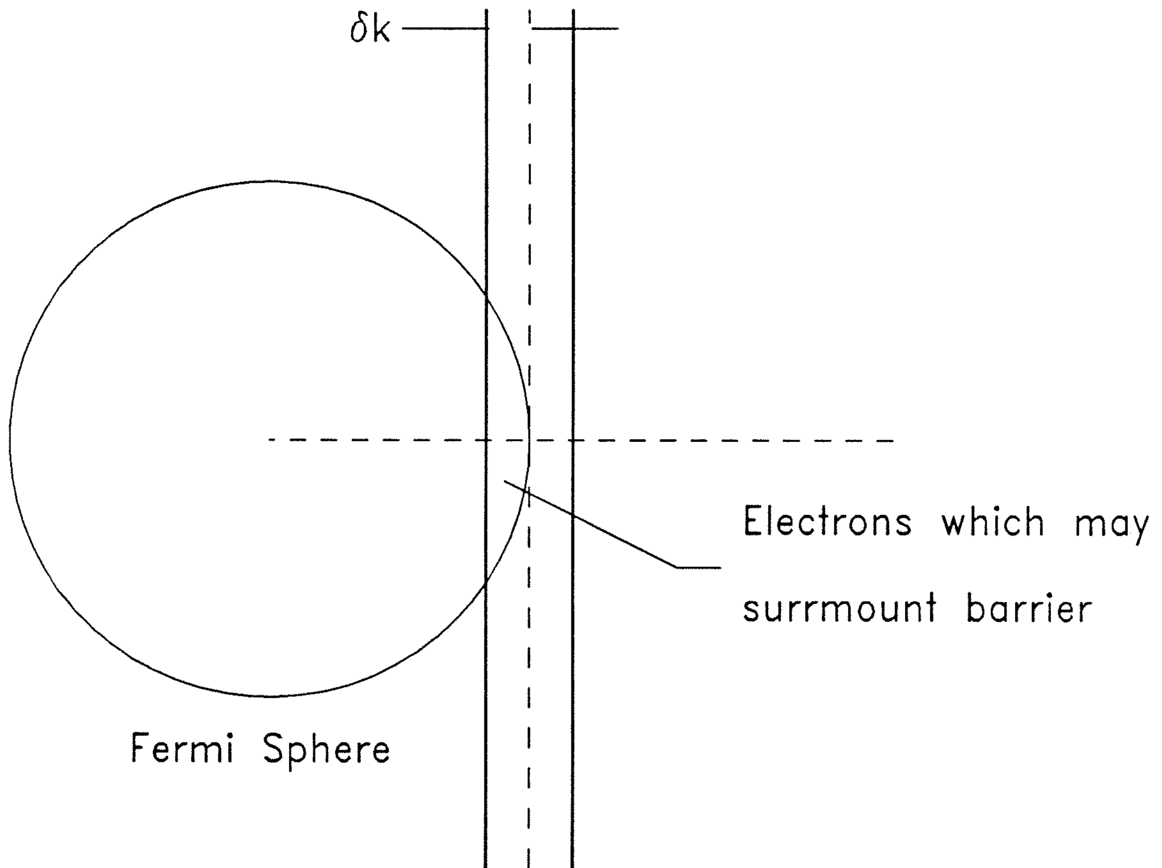


Figure 2.1: Geometric interpretation of photoresponse threshold of a Schottky barrier. The photoresponse is proportional to the number of electrons, *i.e.*, volume of the Fermi sphere, lying to the right of the vertical plane. The plane's displacement from the edge of the sphere δk is approximately proportional to the difference, $\hbar\omega - \phi_{\mathrm{B}_n}$ (see text).

kT about the Fermi energy. Hence, the threshold will also be smeared by this amount. From our picture we see that, for $\delta k = 0$, corresponding to $\hbar\omega = \phi_{B_n}$, the volume of the smeared region to the right of the plane is proportional to T^2. This is the origin of the T^2 prefactor occurring in Eq. 1.10, essentially due to the bulk DOS. Finally, as the plane is moved by approximately $3kT$ or more from the threshold, the zero-temperature behavior is recovered once again.

2.3 Experimental

2.3.1 Silicide Fabrication and Structural Characterization

The epitaxial $NiSi_2$ films used in this study were fabricated by *e*-beam-evaporation of Si and Ni in a VG Semicon, Ltd. Si MBE system.* The Si substrates used in our work were doped with phosphorous ($\approx 1.5 \times 10^{15}\,\mathrm{cm}^{-3}$). Prior to Ni deposition, the Si substrates were sputter-cleaned with Ar and annealed at 850 °C under ultrahigh vacuum (UHV) conditions. [The chemical cleaning technique (Ishizaka *et al.*[19]) followed by thermal evaporation of the oxide in the UHV system as used by Tung[1,6] was not used in our work.] After the anneal, all wafers showed a high quality 7×7 reflection high-energy electron diffraction (RHEED) pattern with no indication of SiC spots.[12] Template layers of type-A and type-B $NiSi_2$ were then formed by *e*-beam evaporation of Ni with subsequent annealing steps, following the procedure of Tung.[1] Type-A and type-B structures thicker than 70 Å were formed either by deposition of pure Ni or by codeposition

*The samples studied here were grown and structurally characterized at the General Electric Research and Development Center under the direction of Dr. Leo J. Schowalter, except for the ion channeling measurements, which were performed under Prof. W. M. Gibson at SUNY-Albany.

of Ni and Si onto the appropriate template layer at 650 °C. Both the cleaning and film deposition steps took place in a single UHV chamber (base pressure $\approx 2 \times 10^{-11}$ Torr). A total of three type-A and four type-B samples were grown with thicknesses ranging from 70 Å to 600 Å. Cross-sectional transmission electron microscopy (TEM) studies of a 250 Å-thick type-A sample and a 600 Å-thick type-B sample showed very uniform, high quality $NiSi_2$ films predominantly of one orientation. The structural perfection of our epitaxial films and the sharpness of our interfaces is evident in the high-resolution TEM lattice image shown in Fig. 2.2. In addition, channeling measurements on these two samples showed good channeling minimum yields ($\chi_{min} < 4\%$), and indicated that each sample was predominantly of a single orientation. Further details on the results of structural characterizations will be given in Section 2.6.2 below.

2.3.2 Electrical characterization

Photoresponse and forward *I-V* methods were used to measure the SBH of the type-A and type-B samples. The photoresponse measurements were performed by the back-illumination method (see Section 1.3.2) with the use of a calibrated spectrometer. The open-circuit photovoltage was synchronously detected at levels such that $V_{ph} \ll kT$ in order to insure linearity of the voltage response. The *I-V* measurements were made on chemically etched, black-wax-defined mesas[13] ranging between 360 and 1300 μm in diameter. A Au-wire probe directly against the silicide was used as the device contact in the *I-V* measurements presented here although it was observed that wire-bonding directly to even the thinnest (70 Å) $NiSi_2$ layers resulted in only a slight increase in leakage current. Because of native oxide (SiO_2) formation on the $NiSi_2$ films, it was necessary to "burn through" the thin oxide through the momentary application of a large forward bias. This procedure did not appear to damage the devices. Back

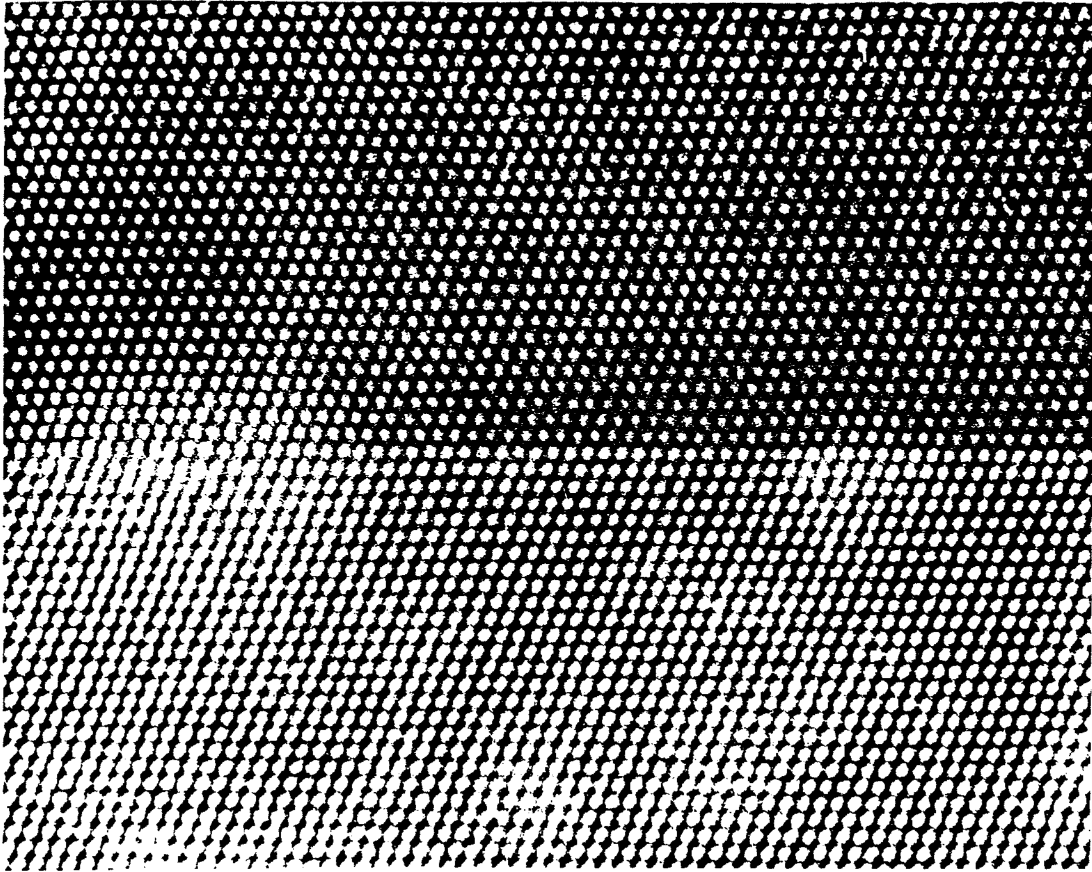


Figure 2.2: High-resolution lattice image of a type-A $NiSi_2$/Si interface, taken by transmission electron microscopy. The image shown was taken with the use of a Philips EM430 electron microscope at beam energy of 300 keV. Individual atomic planes are resolved. The bright spots are $[1\bar{1}0]$ channels in the lattice (normal to the page). Note the interface abruptness and crystallographic alignment of the heteroepitaxial structure. (Taken by E. L. Hall and N. Lewis at the G.E. Corporate Research and Development Center.)

contacts for the I-V measurements were formed by scratching the back Si surface with a diamond-tipped scribe, and then immediately painting on an In-Ga amalgam. Series resistance due to the high bulk substrate resistivities necessitated numerical correction to the raw I-V characteristics, following Norde.[14] All SBH measurements were performed at room temperature.

The photoresponse measurements were made on broad-area-coverage portions of the type-A and type-B wafers. The wafers were cleaved into $\sim$ 3-mm triangular specimens and then mounted with a drop of silver paint, silicide down, onto a transistor header. Care was taken not to puncture the silicide layers. The silver paint mount constituted the device front contact; the back contact was achieved by Al wire-bonding to the back side of the specimen at suitable points. (The back contact described is not a good Ohmic contact, but this fact did not affect the open-circuit measurement discussed below because the contact impedance was always much less than the $100\,M\Omega$ input impedance of the lock-in amplifier.) The I-V characteristics were checked using a curve tracer. No appreciable leakage currents aside from the normal reverse currents of the $NiSi_2$/Si diodes were observed in each instance, and all reverse characteristics showed strong light sensitivity. The open-circuit photovoltage was then measured as a function of photon energy.

The apparatus used to make the photoresponse measurements is shown schematically in Fig. 2.3. Light from a tungsten-filament lamp was directed through a SPEX model 1269 spectrometer, chopped at 270 Hz, and focused with the use of a microscope objective lens onto a small ($\approx$ 0.1mm-dia) spot on the Si side of the sample, avoiding illumination of the wire-bond contacts (though in practice this last precaution did not seem to affect the results). A long-wavelength-pass filter was used to prevent above-band-gap light from illuminating the sample. (A piece of Si would also have been suitable for this purpose.) The open-circuit photovolt-

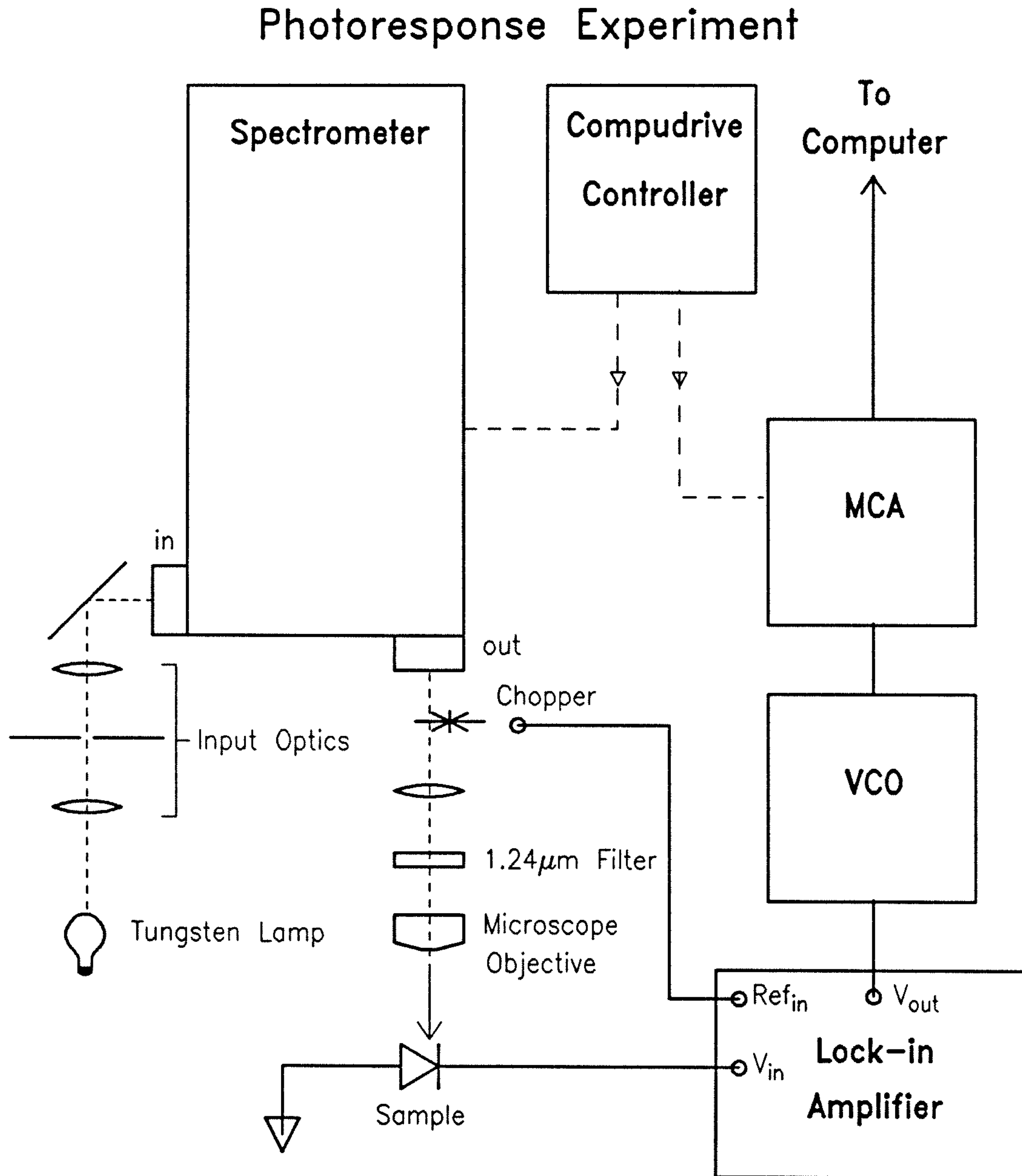


Figure 2.3: Experimental apparatus used for the photoresponse measurement. Monochromatic light exiting the spectrometer illuminates the $NiSi_2$/Si interface through the Si. The filter is used to remove higher order (above-band-gap) light. Sample photovoltage is synchronously detected under open-circuit conditions with the lock-in, and digitized and stored as a function of photon wavelength with the voltage-controlled oscillator and multichannel analyzer.

age signal (typcially $\sim 100\,\mu$V) was synchronously detected with a P.A.R. model 124A lock-in amplifier. The output voltage from the lock-in was then digitized with the use of a voltage-controlled oscillator (dynamic range of 1000:1) and a Nicolet model 1170 multichannel analyzer. This data acquisition scheme allowed for summing of successive scans in order to improve the signal-to-noise ratio, if necessary. Completed scans could be downloaded to a VAX-11/785 computer for analysis and permanent storage. Finally, the optical system response was measured in a separate run with the aid of a Molectron Model P4-42 pyroelectric detector, typically, immediately following the acquisition of photoresponse spectra.

2.4 Results

2.4.1 Photoresponse Measurement

The results of the photoresponse measurements of SBH are shown in Fig. 2.4. In this figure we present response curves taken at $T = 300\,$K for typical type-A and type-B samples plotted according to the conventional Fowler analysis.[10,15] In the conventional analysis the photoresponse data are expected to contain a linear region for incident photons $\hbar\omega$ which satisfy the condition, $(\hbar\omega - \phi_{\mathrm{B}n}) \geq 3kT$. This linear region, when extrapolated to the abscissa, gives the *actual* SBH $\phi_{\mathrm{B}n}$ directly, apart from an "apparent barrier height" correction.[8] As shown in Fig. 2.4, the type-A data can be extrapolated quite unambigously to an intercept value of 0.61 eV. Applying the apparent barrier height correction (difference between $\phi_{\mathrm{B}n}$ and the extrapolated value) as discussed by Okumura and Tu,[8] we finally obtain for the type-A SBH the value $\phi_{\mathrm{B}n} = 0.62\,$eV. This value is reproduced within 0.01 eV on all of our type-A samples. Turning now to the type-B photoresponse, we see from the figure that the data do not contain a clearly

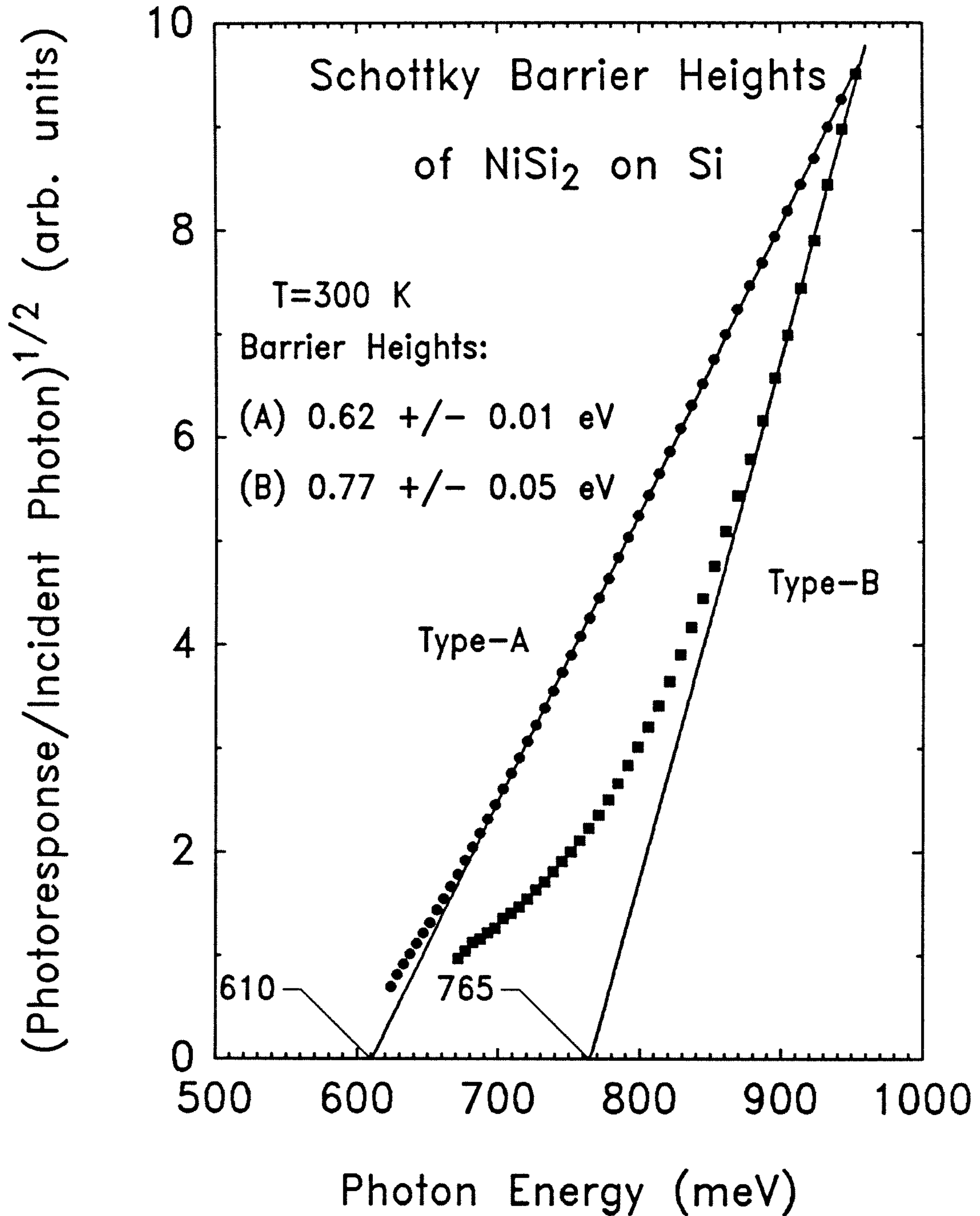


Figure 2.4: Photoresponse data for $NiSi_2/Si$ samples, for both type-A and type-B orientations. For clarity, the spectra are displayed only once every 28 data points. The barrier heights for type-A and type-B samples are (with apparent barrier height correction) $\phi_{B_n} = 0.62 \pm 0.01$ and $0.77 \pm 0.05\,\text{eV}$, respectively.

linear region, making conventional SBH interpretation difficult. Despite their unusual shape, our type-B photoresponse curves are quite *reproducible* over the full range of $NiSi_2$ layer thicknesses from 600 Å down to 70 Å, where one expects the best pseudomorphic growth. We note here that plots of the *cube* root of the photoresponse vs. $\hbar\omega$ also fail to yield a linear plot. If we are to interpret the type-B photoresponse results in the conventional manner, we are led to extrapolate from above the knee in the type-B curve in Fig. 2.4 as this is the most linear region of the data. This procedure results in a SBH of $\phi_{B_n} = 0.77 \pm 0.05\,\text{eV}$, where the estimated uncertainty is associated mainly with the *interpretation* rather than the *reproducibility* of the result. It is clear from the results shown in Fig. 2.4 that the type-A and type-B barriers differ substantially. What remains to be resolved is the nature of the type-B barrier.

2.4.2 *I-V* Measurement

In an attempt to further characterize the type-A and particularly the type-B barriers, the forward *I-V* characteristics of several diodes have been analyzed according to the method described in Section 1.3.2. In most cases, the results (after applying a series resistance correction) are reproducible from device to device and from sample to sample within 0.01 eV, the only complication arising from samples with nonlinear back contacts, which are readily identified and excluded from consideration. The excellent reproducibility of SBH and linearity over two or more decades of $\log J$ vs. V demonstrates the adequacy of our procedures. Representative curves are given in Fig. 2.5. As shown in the figure, we obtain once again a clear difference between type-A and type-B epitaxy. The type-A SBH is found to be $\phi_{B_n} = 0.63 \pm 0.01\,\text{eV}$, in excellent agreement with the photoresponse result. In addition, the ideality factors for type-A devices are in general quite good ($n \approx 1.02$). However, for type-B devices we consistently find

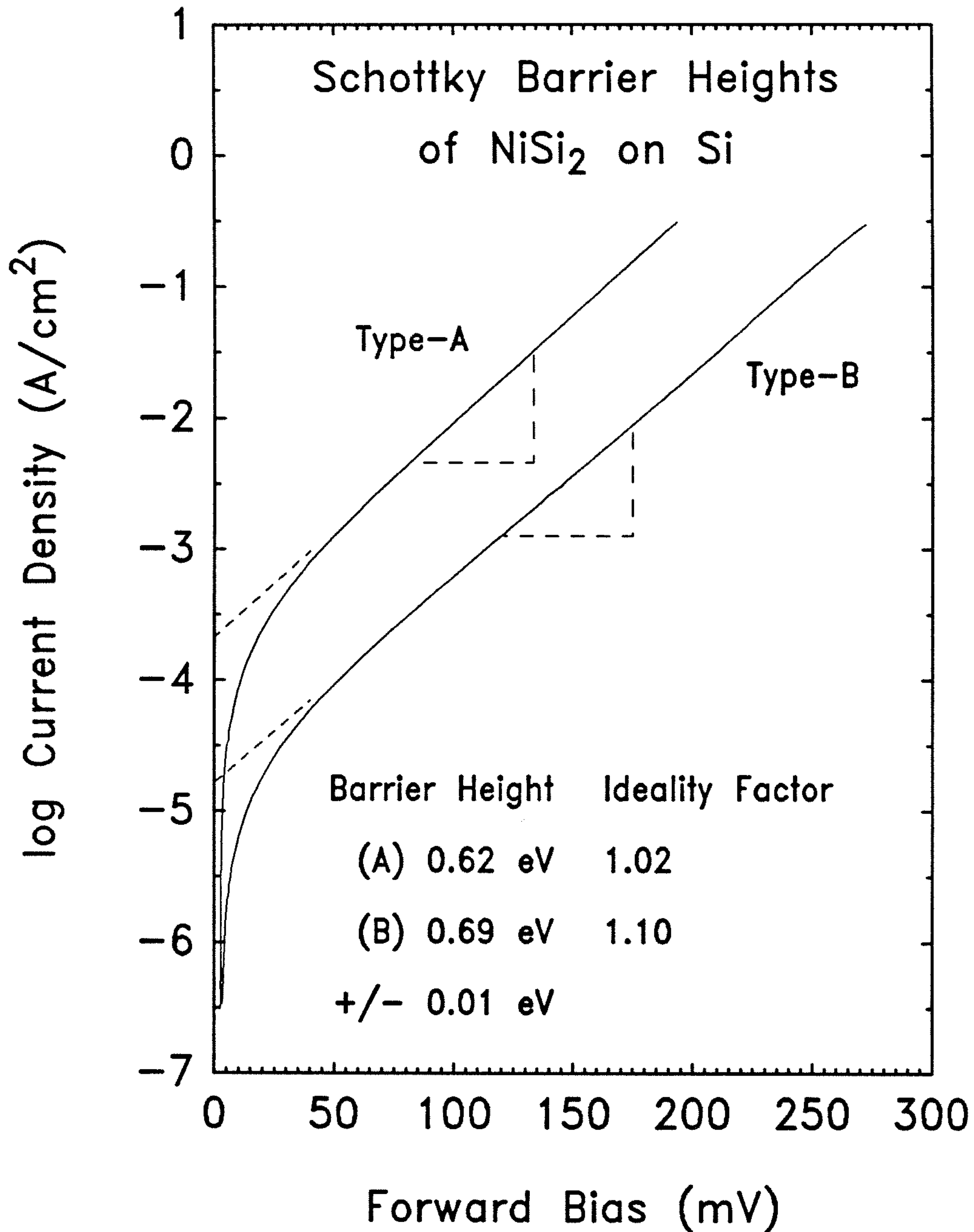


Figure 2.5: Forward *I-V* measurements, corrected for series resistance, of Schottky barrier height for type-A and type-B $NiSi_2$/Si samples, taken at $T = 300\,K$. The type-A and type-B barrier heights are 0.62 and $0.69 \pm 0.01\,eV$ with ideality factors of 1.02 and 1.10, respectively. A value of $110\,A/cm^2 \cdot K^2$ was used for the effective Richardson constant.

Table 2.1: Schottky barrier height measurements of type-A $NiSi_2$ films on Si. The results of *I*-*V* and photoresponse measurements are shown. An apparent barrier height correction (see text) of 0.01 eV has been added to the photoresponse values.

Sample	Thickness (Å)	${\phi_B}_n^{IV}$ (eV)	(n)	${\phi_B}_n^{\mathrm{PR}}$ (eV)
2C-67	250	0.63	(1.01)	0.62
2C-149	250	0.62	(1.05)	0.61
2C-151	600	0.63	···	0.62

a SBH of 0.69 ± 0.01 eV, which is different from the photoresponse value, and we also observe a larger ideality factor ($n \approx 1.10$). The results of our photoresponse and *I*-*V* measurements of SBH on type-A and type-B structures are summarized in Tables 2.1 and 2.2.

2.5 Double Schottky Contact Model

2.5.1 Formalism

We have seen from Figs. 2.4 and 2.5 that the type-A SBH measured by photoresponse and *I*-*V* methods are in excellent agreement, but that there is a significant discrepancy (≈ 0.08 eV) for the case of type-B. We now show that we can account for both the detailed shape of the type-B photoresponse and the numerical value of the SBH obtained from the *I*-*V* measurement in terms of a model of the type-B interface consisting of an electrically parallel mixture of regions of high and low SBH, ϕ_{hi} and ϕ_{lo}. Okumura and Tu[8] have considered the response from compound barrier structures in detail. Their analysis is based on

Table 2.2: Schottky barrier height measurements of type-B $NiSi_2$ films on Si. *I-V* and photoresponse results are presented. The values $\phi_{\rm lo}$, $\phi_{\rm hi}$, and α are obtained by fitting the photoresponse data in terms of the double Schottky contact model described in the text.

Sample	Thickness (Å)	${\phi_{B_n}}^{IV}$ (eV)	(n)	$\phi_{\rm lo}$ (eV)	$\phi_{\rm hi}$ (eV)	α (%)
2C-88	600	0.69	(1.10)	0.63	0.81	8
2C-93	70	...	...	0.65	0.81	11
2C-150	250	0.68	(1.10)	0.64	0.81	7
2C-152	600	0.68	...	0.65	0.81	14

the combined photoresponse Y of a parallel set of SBH's $\{\phi_i\}$ being given by

$$Y(\hbar\omega, T; \{\phi_j\}) = \sigma \sum_i \alpha_i C_i F(\hbar\omega, T; \phi_i), \tag{2.10}$$

where $F(\hbar\omega, T; \phi_i)$ is a normalized Fowler photoyield function of the type discussed earlier (Eq. 1.10 in Section 1.3.2), α_i corresponds to the fractional area coverage of barrier ϕ_i, σ is the active device area, and C_i is a factor which depends on the thickness, reflectivity, absorptivity, and electron attenuation length of the appropriate metal layer.[8] Apart from variation in optical and photoelectric parameters, Eq. 2.10 gives the photoresponse of a compound SB as an area-weighted linear superposition of the photoresponse due to each single-barrier component. We now specialize Eq. 2.10 to the case of our double Schottky contact model. In this high/low barrier model we take $C_i = C =$ const. and define a new normalization constant $\sigma' \equiv C\sigma$. Eq. 2.10 then becomes

$$Y(\hbar\omega, T; \phi_{\rm lo}, \phi_{\rm hi}) = \sigma' \left[\alpha F(\hbar\omega, T; \phi_{\rm lo}) + (1-\alpha) F(\hbar\omega, T; \phi_{\rm hi})\right], \tag{2.11}$$

where α refers to the effective areal fraction covered by ϕ_{lo}.

2.5.2 Fit to Photoresponse Data

With the use of the above model we are now in a position to analyze the type-B photoresponse data. Okumura and Tu[8] have developed a general method to analyze mixed Schottky barrier structures which requires differentiation of the photoresponse curve with respect to $\hbar\omega$, which is then analyzed in terms of derivatives of Eqs. 2.10 or 2.11. However, in our work we find it both adequate and much simpler to use Eq. 2.11 directly. Treating the quantities, ϕ_{hi}, ϕ_{lo}, α, and σ' as adjustable parameters, we perform a nonlinear least-squares fit to our experimental type-B photoresponse data using Eq. 2.11 evaluated at T = 300 K. The result is shown in Fig. 2.6. We see that good agreement with experiment is obtained for the values, $\phi_{\mathrm{hi}} = 0.81\,\mathrm{eV}$, $\phi_{\mathrm{lo}} = 0.64\,\mathrm{eV}$, and $\alpha = 0.09$. We mention here that applying this fit procedure to all of our type-B samples, from 70 to 600 Å film thicknesses, yield consistency in barrier heights of $\pm 0.01\,\mathrm{eV}$ and in α of ± 0.02, as can be seen from Table 2.2. In principle, a double Schottky contact should be directly observable as a "kink" in the first derivative[8] or as a double plateau in the second derivative[16] of the photoresponse curve. However, it is clear from Fig. 2.7 that for the present case, the first and second derivative spectra do not themselves yield useful information about the double contact structure at room temperature without again resorting to the type of fit described above, due to the close spacing of ϕ_{hi} and ϕ_{lo} compared to kT. Direct numerical differentiation of our room temperature photoresponse data, also shown in the figure, confirms this fact. On the other hand, it may be possible to see direct evidence of the double Schottky contact by performing the photoresponse measurement at or below $T = 77\,\mathrm{K}$, and taking the first derivative numerically, provided that artefacts of the numerical procedures can

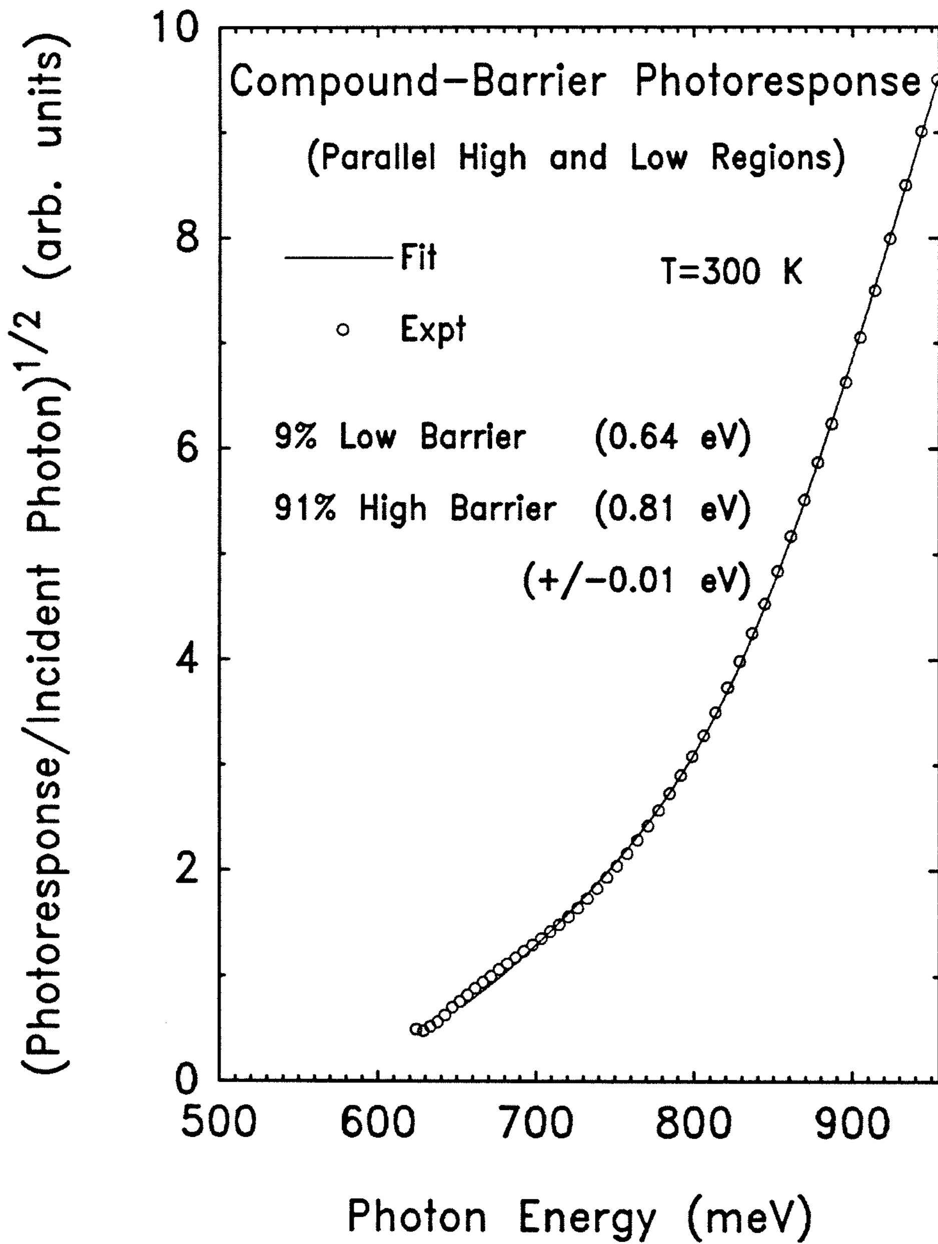


Figure 2.6: Nonlinear least-squares fit of the double Schottky contact model to type-B photoresponse data, for theoretical response (Eq. 2.11) evaluated at $T = 300\,\mathrm{K}$. The fit ranges over 1200 data points, corresponding to energies covered by the solid line. The three physically meaningful fit parameters are the barrier heights $\phi_{\mathrm{hi}} = 0.81\,\mathrm{eV}$, $\phi_{\mathrm{lo}} = 0.64\,\mathrm{eV}$, and low-barrier areal coverage $\alpha = 9\%$.

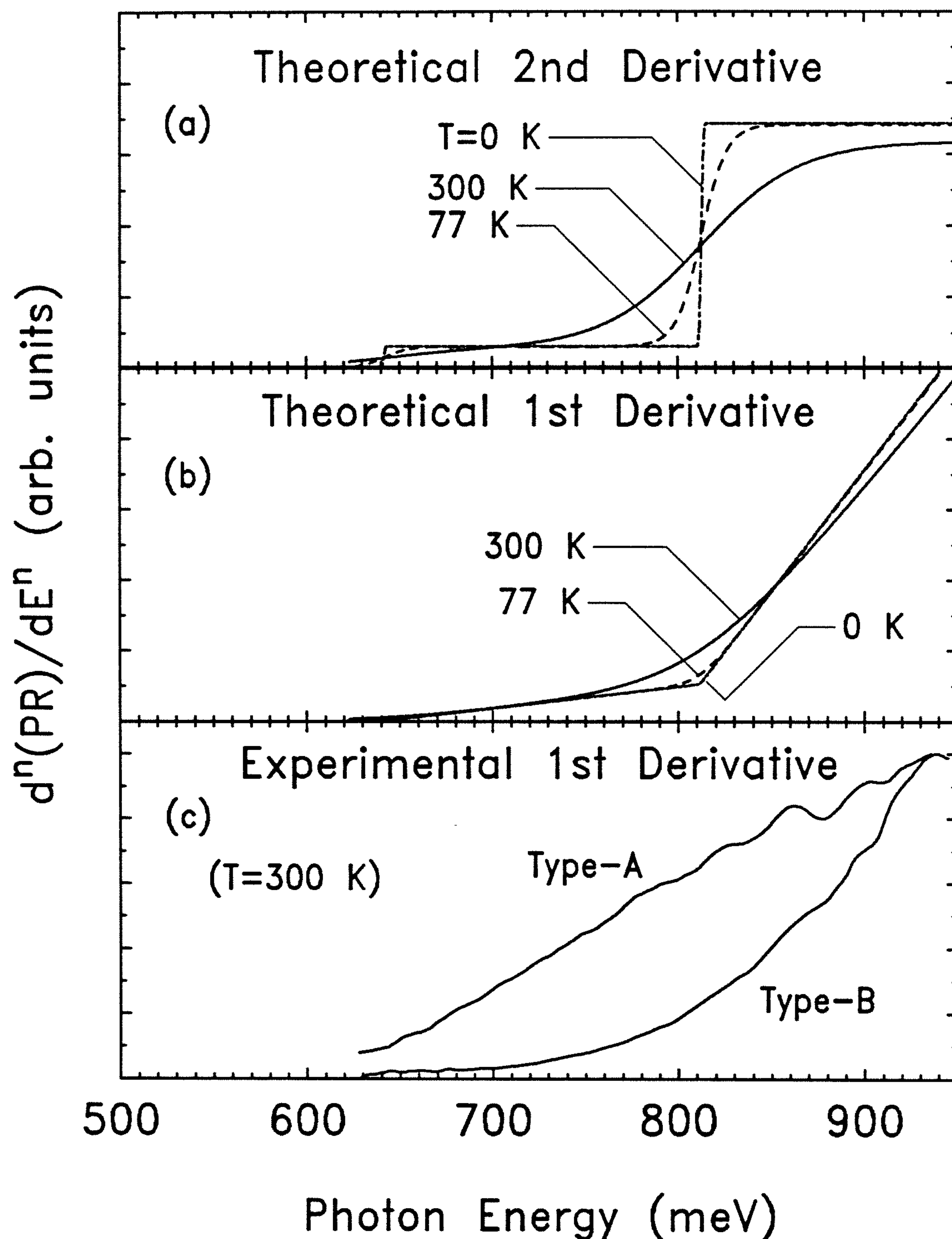


Figure 2.7: Second (a) and first (b) derivatives of theoretical compound-barrier photoresponse assuming the composition given in Fig. 2.6. (c) Numerically differentiated photoresponse of type-A and type-B spectra.

be minimized.

2.5.3 Prediction for *I-V* Measurement

We may use the results of the model fit discussed in Section 2.5.2 to account for the observed *I-V* SBH results of our type-B samples. In general, for the case of multiple Schottky barriers ϕ_i in *parallel*, the total current is simply the area-weighted superposition of the individual current density components J_i:

$$J = \sum_i \alpha_i J_i = \sum_i \alpha_i J_{si} \left(e^{qV/n_i kT} - 1 \right) . \tag{2.12}$$

Each J_i is given in turn by a Schottky diode expression in the form of Eq. 1.7 on page 29. An average barrier height $\overline{\phi}$ can be defined through the expression for the saturation current density (see Eq. 1.7) as

$$e^{-\overline{\phi}/kT} \equiv \sum_i \alpha_i e^{-\phi_i/kT} = \frac{1}{A^{**}T^2} \sum_i \alpha_i J_{si}. \tag{2.13}$$

Neglecting appreciable variation in the ideality factors n_i, we see from Eqs. 2.12 and 2.13 that it is the *single* SBH, $\overline{\phi}$, that is observed in the *I-V* measurement. It is straightforward to show that, for our double Schottky contact model, Eq. 2.13 leads to the expression,

$$\overline{\phi} \approx \phi_{\rm lo} - kT \ln \alpha, \tag{2.14}$$

valid whenever $\alpha \exp[(\phi_{\rm hi} - \phi_{\rm lo})/kT)] \gg 1$. Hence, the low-barrier SBH is preferentially weighted in the *I-V* measurement because of the nonlinear averaging given by Eq. 2.13. Using the values of $\phi_{\rm hi}$, $\phi_{\rm lo}$, and α obtained from the model fit above yields the result $\overline{\phi} = 0.70\,\text{eV}$, in excellent agreement with the type-B SBH directly measured by the *I-V* method (0.69 eV). It is easy to show that taking proper account of the ideality factors leads to a first order correction term in Eq. 2.14 of $-qV(\Delta n/\overline{n}^2)$, where qV is the forward bias (in eV), $\overline{n}$ is the mean ideality factor, and Δn the difference between the low-barrier ideality factor and

Table 2.3: Comparison of Schottky barrier heights ϕ_{B0} of single-crystal $NiSi_2$ epilayers on n-type, (111)-Si. The results of the present study are compared with the previous work of Tung, and Liehr *et al.*, respectively.

Orientation	PR[a]	IV[a]	Tung[b]	Liehr *et al.*[c]
	(±0.01 eV)	(±0.01 eV)	(±0.01 eV)	(±0.02 eV)
Type-A	0.63	0.64	0.65	0.66/0.78[d]
Type-B	0.65/0.82[e]	0.70	0.79	0.66/0.78[d]

a. This work.

b. Reference 6.

c. Reference 7.

d. The barrier heights are reported to depend on structural perfection rather than epitaxial orientation.

e. Values of ϕ_{lo} and ϕ_{hi} obtained from the model fit described in the text.

$\overline{n}$. From this we see that neglecting the ideality factors leads to an error of less than 0.01 eV for reasonable values of the ideality factors. Thus, with a simple model we have shown *quantitative* consistency between apparently disparate SBH results for the case of the type-B barrier heights as measured by two different methods.

Comparison of our SBH measurements with the results of the previous work of Tung, and Liehr *et al.*, respectively, is made in Table 2.3. For purposes of comparison, the numbers presented in the table are the *intrinsic* or flatband

Schottky barrier heights, ϕ_{B0}, related to ϕ_{Bn} through the addition of the image-force lowering term, *i.e.*, $\phi_{B0} = \phi_{Bn} + \Delta\phi$. From the table, we see, first of all, that the results of our work are in at least qualitative agreement with the results of Tung, but in disagreement with the work of Liehr *et al.* (The barrier heights observed by Liehr *et al.* are reported to depend on the structural quality of the interface and to be independent of the particular epitaxial orientation.) Further, we see that our type-A SBH value (0.63 eV), obtained both by photoresponse and *I-V* techniques, is in excellent agreement with that of Tung (0.65 eV). For the type-B barriers, the values of SBH obtained by the *I-V* technique differ between Tung's and our studies. However, from photoresponse studies we know that our type-B results can be explained by the presence of a double Schottky contact. Our value for ϕ_{hi} (associated with most of the effective device area) given in the table is in good agreement with the SBH obtained by Tung (0.79 eV) using *I-V* and *C-V* methods.

2.6 Discussion

2.6.1 Interpretation of Results

The interpretation of our photoresponse and *I-V* SBH measurements suggest that most (91%) of the interfacial area of our type-B $NiSi_2$/Si structures is associated with a SBH of $\phi_{Bn} = 0.81$ eV, but also present are localized regions, comprising 9% of the effective device area, that are associated with a reduced barrier height, 0.64 eV. There are two possibilities that immediately spring to mind to account for this phenomenon. Noting that the value of ϕ_{lo} is close to the type-A SBH, it may be that, present at the $NiSi_2$/Si interface are grains of type-A orientation mixed in among the type-B structure. This possibility is discounted by channeling and TEM observations, however, which show both our

type-A and type-B films to be single-crystal films of entirely one orientation.[2] A second possibility is that localized interfacial structural disorder gives rise to localized electronic states of sufficient density to locally pin the Fermi level. Misfit dislocations are observed[2] in plan-view TEM images of type-B $NiSi_2$ while none are observed in type-A films over the range of thicknesses reported here. However, since the density of misfits increases with increasing type-B film thickness, the consistent observation of bowing in the photoresponse for all type-B layers argues against misfit dislocations as the cause. Finally, the *consistent* absence of any non-ideality in the physical or electrical characterization of the type-A structures and the significant difference between type-A and type-B measured barrier heights would suggest the absence of some foreign surface contaminant which dominates the electrical behavior near the interface.

2.6.2 Possible Physical Origins of Model

It is possible that the low-SBH regions may be associated with localized *planar* defects seen in cross-sectional high-resolution TEM. These defects consist of an extra atomic plane, approximately 100 Å in extent, inserted at or near the interface. The average spacing of the defect regions is approximately 600–800 Å for both the type-A and type-B orientations. It has been suggested[2] that these localized planar defects are associated with the low SBH, and because of the proximity of the low-SBH to the type-A value, for the type-A structures, only *single*-SB behavior is observed. Because of the small linear extent of the planar defect compared with either the depletion length or the Debye length, the effective area of the defect might conceivably be different from the actual geometrical area, but would be the same effective area for both the photoresponse and *I*-*V* measurements.

2.7 Conclusions

In summary, we have made Schottky barrier height measurements on high-quality, MBE-grown $NiSi_2$/Si structures of type-A and type-B orientations with the use of photoresponse and forward *I-V* techniques. Comparison of the data obtained from both techniques clearly demonstrate a difference in barrier heights between type-A and type-B structures. This finding is in agreement with the results of Tung[6] but in disagreement with the results of Liehr *et al.*[7] The main contribution from this work is the side-by-side comparison of photoresponse and *I-V* measurements made on the *same* samples, and the unique comparison that these complementary measurements allow. For the case of type-A epitaxy, we consistently obtain a barrier height from both techniques of $\phi_{\mathrm{B}_n} = 0.62 \pm 0.01\,\mathrm{eV}$. For the case of type-B, a serious discrepancy in the barrier height between photoresponse and forward *I-V* methods is observed. Furthermore, the qualitative shape of the type-B Fowler plots are nonlinear but consistent from sample to sample, rendering determination of barrier height ambiguous. We have shown that both the detailed shape of the Fowler plot as well as the measured *I-V* value of type-B barrier height for our samples can be quantitatively accounted for in terms of a double Schottky contact model. In this model, the type-B structure consists of electrically parallel regions of high and low barrier height, where $\phi_{\mathrm{hi}} = 0.81 \pm 0.01\,\mathrm{eV}$, $\phi_{\mathrm{lo}} = 0.64 \pm 0.01\,\mathrm{eV}$, with the effective area of low- and high-barrier coverages being 9 and 91%, respectively. The possible physical origins of this model within the $NiSi_2$/Si system are discussed.

References

1. R. T. Tung, J. M Gibson, and J. M. Poate, Phys. Rev. Lett. **50**, 429 (1983).

2. R. T. Tung, Phys. Rev. Lett. **52**, 461 (1984); J. Vac. Sci. Technol. B **2**, 465 (1984).

3. M. Liehr, P. E. Schmid, F. K. LeGoues, And P. S. Ho, Phys. Rev. Lett. **54**, 2139 (1985).

4. R. J. Hauenstein, T. E. Schlesinger, T. C. McGill, B. D. Hunt, and L. J. Schowalter, Appl. Phys. Lett. **47**, 853 (1985).

5. M. Schlüter, Thin Solid films **93**, 3 (1982).

6. L. J. Brillson, Surf. Sci. Rep. **2**, 123 (1982).

7. E. H. Rhoderick, *Metal-Semiconductor Contacts* (Clarendon, Oxford, 1978).

8. T. Okumura and K. N. Tu, J. Appl. Phys. **54**, 922 (1983).

9. R. J. Hauenstein, T. E. Schlesinger, T. C. McGill, B. D. Hunt, and L. J. Schowalter, J. Vac. Sci. Technol. A **4**, 860 (1986).

10. R. H. Fowler, Phys. Rev. **38**, 45 (1931).

11. A. Ishizaka, K. Nakagawa, and Y. Shiraki, *Collected Papers of MBE-CST-2, 1982, Tokyo* (Japan. Soc. Appl. Phys., Tokyo, 1982), p. 183.

12. R. C. Henderson, R. B. Marcus, and W. J. Polito, J. Appl. Phys. **42**, 1208 (1971).

13. R. J. Hauenstein, T. E. Schlesinger, T. C. McGill, B. D. Hunt, and L. J. Schowalter, J. Vac. Sci. Technol. B **4**, 649 (1986).

14. H. Norde, J. Appl. Phys. **50**, 5052 (1979).

15. S. M. Sze, *Physics of Semiconductor Devices* (Wiley, New York, 1981), Chap. 5.

16. S. K. Krawczyk and T. Mrabeut, Appl. Phys. Lett. **46**, 837 (1985).

17. B. D. Hunt, L. J. Schowalter, N. Lewis, E. L. Hall, R. J. Hauenstein, T. E. Schlesinger, T. C. McGill, Masako Okamoto, and Shin Hashimoto, in *Proceedings of the Materials Research Society*, Boston, 1985 (to be published).

Chapter 3

Electron Tunneling Spectroscopy Experiments on Single-Crystal $CoSi_2$ and $NiSi_2$ Epitaxial Films on n-type Si

3.1 Introduction

3.1.1 Background

Through the development of molecular beam epitaxial (MBE) growth techniques, it has recently become possible to fabricate very high quality, single-crystal, epitaxial metal films on a Si substrate.[1,2] Thus far, the metallic silicides, $CoSi_2$ and $NiSi_2$, both of which have a very close lattice-constant match to Si, have each been fabricated with a high degree of crystalline perfection as single-crystal films on the (111) surface of Si.[1,2] In addition, we have already mentioned the fact (Sections 1.3 and 2.1) that $NiSi_2$ is unique in that single-crystal epilayers

can be fabricated on (111)-Si with comparable quality in either of two distinct orientations (type-A and type-B).[1] Single-crystal metal epilayers are of technological interest because of the possibility of overgrowth of other epitaxial films. A number of interesting device structures based on silicon-silicide heterostructures, in particular, the metal-base transistor (MBT), discussed earlier in Section 1.2.2 and shown in Fig. 1.1, are being pursued.[3,4,5] Silicide-on-Si structures are also ideally suited to *fundamental* investigations of important device issues; in particular, to our understanding of the metal-semiconductor interface. And, as we have discussed earlier in Chapter 2, silicide/Si structures have been the subject of recent controversy.[6,7,8,9]

In this chapter, we continue our study of the electronic transport across single-crystal metal-semiconductor interfaces. In Chapter 2, the results of Schottky barrier height measurements of epitaxial, single-crystal $NiSi_2$ films grown on non-degenerate Si were presented. Now, we wish to consider the transport in the case that the Si substrate is degenerate. We have already noted in Section 1.3.2 that, when the semiconductor is degenerately doped, the dominant electron transport mechanism in metal-semiconductor structures is elastic tunneling. In addition, there we saw that electron transport may also proceed by *inelastic* tunneling processes. In metal-semiconductor structures, the relative fraction of inelastic to total tunneling current is usually insignificant. Nevertheless, the inelastic tunneling current contains information, in the form of an energy spectrum, about the internal processes which can assist electron transport. The information contained in this small, inelastic component is most easily recovered by direct measurement of the *second derivative* of the I-V characteristic, *i.e.*, d^2I/dV^2 vs. V. In this chapter, we shall be primarily concerned with such second derivative tunnel spectra obtained from single-crystal $NiSi_2$ and $CoSi_2$ films on degenerate, n-type Si.

Electron tunneling in degenerate Schottky barrier (SB) structures have been

studied extensively in the 1960's.[10,11,12,13,14] However, these early studies always involved the study of *polycrystalline* metal films. As a consequence of the polycrystallinity, the extent to which $\mathbf{k}_{\parallel}$-conservation was obeyed in the electron tunneling process was never clear. In contrast to polycrystalline metal films, for the case of single-crystal silicides on Si, there is complete, two-dimensional lattice-translational symmetry in the plane of the interface. Hence, for the first time, $\mathbf{k}_{\parallel}$ must be conserved rigorously in metal-semiconductor tunneling processes. On this basis it has been suggested[15] that the rigorous $\mathbf{k}_{\parallel}$-conservation requirement might be manifested through the appearance of additional or enhanced features in silicide-on-Si tunneling spectra in comparison to tunnel spectra obtained with polycrystalline films. Initial tunneling spectra reported by Rosencher *et al.*,[15] which were obtained from $CoSi_2$ films grown epitaxially on Si, appeared to confirm this expectation. However, the reported tunneling spectra proved to be rather complicated,[15] making their interpretation unclear.

3.1.2 Results of this work

In this chapter, we report the first results of electron tunneling spectroscopy obtained from single-crystal, epitaxial films of $NiSi_2$ (type-A and type-B orientations), and also, of $CoSi_2$ films (type-B orientation), each grown on degenerate, n-type Si substrates. Two types of degenerate substrates were used in this work: a "heavily" doped ($\approx 2 \times 10^{19}\,\mathrm{cm}^{-3}$ As), and a "lightly" doped ($\approx 5 \times 10^{18}\,\mathrm{cm}^{-3}$ Sb) substrate. The tunneling spectra obtained for silicide films on the lightly doped substrate are similar to spectra obtained on $CoSi_2$ films by Rosencher *et al.*[15] in that they are complicated, and difficult to interpret. On the other hand, we obtain comparatively straightforward and reproducible tunneling spectra from all of our silicide films grown on the heavily doped substrate. The latter tunneling spectra contain peaks in forward bias at energies which correspond to

k-conserving Si phonons (see Section 1.3.3). In addition, for the $NiSi_2$ films we see a weak, *antisymmetric* peak in reverse bias corresponding to the **k**-conserving, Si TA phonon. The tunneling spectra obtained from single-crystal metal layers on the heavily doped substrate are essentially *identical* with those obtained from polycrystalline Au films on the heavily doped substrate. No new or enhanced spectral features due to the $\mathbf{k}_{\parallel}$-conservation condition are evident in our tunneling results. Instead, the results obtained here on the heavily doped substrate suggest that the inelastic transport of electrons is assisted through interaction with *bulk*, **k**-conserving, rather than *interfacial*, $\mathbf{k}_{\parallel}$-conserving processes.

3.1.3 Outline of Chapter

The theory of elastic and inelastic tunneling in Schottky barrier tunnel structures is presented in Section 3.2. Next, in Section 3.3, the experimental details of silicide growth, structural characterization, sample preparation, along with the tunneling and other electrical measurements, are given. In Section 3.4, our main results of electron tunneling spectroscopy (second derivative) and differential resistance (first derivative) measurements for single-crystal $CoSi_2$ and $NiSi_2$ films on As-doped Si are presented. The tunneling results from the Sb-doped substrates are also given there. Section 3.5 contains a detailed discussion of our tunneling results: the identification of spectral peaks, mechanism of tunneling, and comparison with the WKB theory of Section 3.2 are discussed. Also, in this section, our tunneling results of single-crystal films are compared with those obtained from polycrystalline Au films, and the significance of the peak seen in our first derivative (dV/dI vs. V) spectra is discussed. Finally, our conclusions from this work are given in Section 3.6.

3.2 Tunneling in Schottky Barriers

3.2.1 Elastic Tunneling: WKB Theory

The theory of elastic tunneling for degenerate Schottky barriers, in the WKB approximation, has been worked by several authors.[11,12,14] An "exact" calculation based on an effective-mass, one-band-model Hamiltonian, with the assumption of a one-dimensional, unscreened, parabolic potential in the semiconductor, has also been reported.[10] However, even for this oversimplified situation, the exact, analytical result is unnecessarily complicated for our purposes. Moreoever, a reasonable, *i.e.*, better than a factor-of-two agreement between the WKB and exact calculations has been demonstrated by Stratton *et al.*[16] So, we find it adequate and convenient to make use of the analytically simpler results of Stratton *et al.*[11,16] to aid in the analysis of our data.

We have previously discussed the WKB approach in Section 1.3.3. The results which we are now about to state are based on Eq. 1.14 in that section. Stratton[11] has derived simple analytic expressions to describe the forward-bias I-V characteristic of a tunneling Schottky diode at $T = 0\,\mathrm{K}$. For an n-type semiconductor, his results for the current density J as a function of voltage V are

$$J = \frac{4\pi q m_1}{h^3 c_{1F}^2} C e^{-b_{1F}} \left[1 - e^{-c_{1F} qV} - c_{1F} qV e^{-c_{1F}\mu_F}\right], \quad 0 \le qV \le \mu_F; \tag{3.1}$$

$$J = \frac{4\pi q m_1}{h^3 c_{1F}^2} C e^{-b_{1F}} \left[1 - e^{-c_{1F}\mu_F} - c_{1F}\mu_F e^{-c_{1F}\mu_F}\right], \quad qV > \mu_F; \tag{3.2}$$

where μ_F is the Fermi degeneracy, *i.e.*, the Fermi energy referred to the conduction band edge, and q is the magnitude of the electron charge. When the Fermi energy is sufficiently small in comparison with the barrier height ϕ_{B0} so that the condition, $\mu_F \ll \phi_{B0} - qV$ is satisfied, the quantities b_{1F} and c_{1F} appearing in

Eqs. 3.1 and 3.2 are given approximately by[11]

$$b_{1F} \approx \frac{\phi_{\mathrm{B}0} - qV}{E_{00}}, \tag{3.3}$$

$$c_{1F} \approx \frac{1}{2E_{00}} \ln \left[\frac{4(\phi_{\mathrm{B}0} - qV)}{\mu_F} \right], \tag{3.4}$$

where

$$E_{00} \equiv \frac{q\hbar}{2} \left(\frac{N_d}{m_1 \epsilon} \right)^{1/2}. \tag{3.5}$$

In this last expression, N_d is the donor concentration in the degenerate semiconductor, and ϵ is the semiconductor dielectric constant. The tunneling mass, m_1, which appears in Eqs. 3.1, 3.2, and 3.5, is an effective mass parameter which depends in general on the band structure of the semiconductor, and, for the case of anisotropic conduction band valleys, on the *direction* of electron tunneling. Also, the correction factor C in Eqs. 3.1 and 3.2 takes account of the valley anisotropy.[16] For the case of n-type Si, there are six indirect, anisotropic conduction band valleys along the equivalent $\langle 100 \rangle$ directions. For electron transport in the [111]-direction, each valley contributes equally. Equations 3.1 and 3.2 represent the contribution of just *one* valley; hence, these expressions must be multiplied by six to get the total current density. Corrections due to valley anisotropy have been worked out[16]; it is easy to show that, for tunneling in the [111]-direction in n-type Si, the tunneling mass is equal to the conductivity effective mass, $m_1 = m_c^*$, and $6C = 8.52$. It is with these parameters in Eqs. 3.1 thru 3.5 that we compare our experimental I-V and dV/dI vs. V results with theory.

3.2.2 Inelastic Tunneling

Inelastic tunneling effects are not apparent in the I-V characteristic of metal-semiconductor tunnel junctions. This is because the extent in **k**-space of the

metal Fermi surface is usually large compared to the extent of the semiconductor Fermi surface. As a consequence, it is usually possible to find initial states in the semiconductor (metal) and final states in the metal (semiconductor) for which electron energy *and* $\mathbf{k}_{\parallel}$ are simultaneously conserved (see the discussion in Section 1.3.3). Hence, the electron transport in metal-semiconductor tunnel structures is dominated by elastic tunneling. Inelastic effects must be observed in the *derivatives* of the I-V characteristic. In Section 1.3.3 we described the use of the d^2I/dV^2 vs. V curve in an experimental technique known as inelastic electron tunneling spectroscopy (IETS). IETS spectra have been obtained on vacuum-cleaved Si and Ge surfaces onto which polycrystalline metals have been evaporated.[13,14,17] IETS spectra have also been reported for metal-insulator-semiconductor (MIS) structures.[18] While phonon-assisted tunneling on metal–n-type-Ge structures has been observed,[14,17] to the best of our knowledge, similar observations on n-type Si have not been reported. However, in contrast to phonon-assisted inelastic tunneling effects, a *many-body*, self-energy (polaron) effect due to electron-phonon interaction with zone-center, LO phonons has been observed in p-type Si.[13,18] As we shall see below, our results on n-type Si are characteristic of phonon-assisted tunneling, much like the results obtained on Ge.[14,17]

3.3 Experimental

3.3.1 Silicide Epitaxial Growth and Structural Characterization

Single-crystal films of $CoSi_2$ (type-B orientation) and $NiSi_2$ (both type-A and type-B orientations) were fabricated* on As-doped and Sb-doped, degenerate

*The silicide samples were fabricated by Dr. L. J. Schowalter (General Electric Research and Development Center). Growth details are included here for completeness.

Si wafers. The Si substrates used here were all nominally (111)-oriented: the As-doped wafers were misaligned from the (111)-axis by 4° while the Sb-doped wafers were on-axis to within 0.25°. The substrate doping level was estimated from Hall and resistivity measurements* to be $N_d = (2.3 \pm 0.6) \times 10^{19}\,\mathrm{cm}^{-3}$ for the As-doped substrate, and $(5.4 \pm 1.0) \times 10^{18}\,\mathrm{cm}^{-3}$ for the Sb-doped substrate. Silicide growth took place in the same three-chamber Si MBE system which is mentioned in Section 2.3.1 above. Prior to growth, the Si substrates were cleaned either by giving them a modified RCA chemical clean[19] followed by thermal evaporation of the oxide in the UHV chamber, or by sputter cleaning with an Ar ion gun in the MBE system. These cleaning procedures are described in more detail elsewhere[2,9]; we note here, however, that no significant difference in the subsequent results was observed due to the different cleaning methods. The silicide samples which were grown for our tunneling study are summarized in Table 3.1.

The $NiSi_2$ films were grown by using the template techniques established by Tung *et al.*[1] to initiate single-crystal growth of type-A or type-B orientations. The final film thickness of 250 Å was achieved through codeposition of Si and Ni. $NiSi_2$ films grown by this technique appear mirror-smooth in Nomarski and have been previously shown to be high quality, single-crystal films.[9] $CoSi_2$ films were initiated by growing 70 Å of $CoSi_2$ using a pinhole-reduction technique,[2] and the final thickness of 250 Å was obtained through codeposition of Co and Si. In spite of the fact that thinner layers of $CoSi_2$ grown by this technique on *non*degenerate Si substrates have been of high quality with mirror-smooth surface morphologies,[2] the layers grown for this work on *degenerate* substrates showed some unusual surface features in Nomarski and scanning electron microscopy (SEM), and also in transmission electron microscopy (TEM) observations. These

*The Hall and resistivity measurements were made under the direction of O. J. Marsh at the Hughes Research Laboratories.

Table 3.1: Samples grown for the electron tunneling study. Single-crystal silicide films of thickness 250 Å were grown epitaxially on degenerate, (111)-oriented Si substrates. (The As-doped substrates used here are 4° off-axis.)

Sample	Silicide	Orientation	Dopant	N_d (cm^{-3})	Surface Prep.
2C-230	$NiSi_2$	A	As	2×10^{19}	Shiraki
2C-231	$NiSi_2$	B	As	2×10^{19}	Shiraki
2C-236	$NiSi_2$	A	As	2×10^{19}	Shiraki
2C-302	$NiSi_2$	A	Sb	5×10^{18}	Shiraki
2C-303	$NiSi_2$	B	Sb	5×10^{18}	Shiraki
2C-304[a]	$NiSi_2$	B	Sb	5×10^{18}	Shiraki
2C-370	$CoSi_2$	B	Sb	5×10^{18}	Shiraki
2C-371	$CoSi_2$	B	Sb	5×10^{18}	Sputter
2C-372	$CoSi_2$	B	As	2×10^{19}	Sputter
2C-373[b]	$CoSi_2$	B	As	2×10^{19}	Sputter

a. The silicide film thickness of sample 2C-304 was 72 Å.

b. A thin Si buffer layer was first grown. This buffer layer was of a thickness such that the template-forming solid-phase reaction [Ref. 2] completely consumed it, but that none of the substrate Si was consumed.

surface features correspond to highly irregular "craters" extending down into the Si substrate, as shown in Fig. 3.1. The cause of these surface features is not presently understood. However, the films still appear to be high quality, single-crystal films as judged by reflection high-energy electron diffraction (RHEED) and TEM dark-field imaging.

3.3.2 Sample Preparation

Following silicide layer formation and removal from the MBE growth chamber, individual Schottky diode devices were formed. The device structure is shown in Fig. 3.2. First, a small specimen (suitable for mounting on standard transistor headers) was cleaved from the central portion of the wafer and immersed in 2% HF for approximately 20 sec to remove the native oxide. The specimen was then immediately loaded into a metallization chamber and a Au overlayer ($\approx$ 1000 Å) evaporated. Subsequently, mesa devices ranging from 70 to 700 μm in diameter were defined by conventional photolithographic and wet etch techniques. A commercial etch was used to remove the Au, and then an etch consisting of 2% HF and 3% HNO_3 for approximately 30 sec was used to selectively remove the silicide.* The etch procedures all took place at room temperature under a mild mechanical agitation. Following the etch procedures, the back side of the specimen was scratched with a diamond scribe and an In-Ga amalgam immediately painted on for the Ohmic contact (typically 70–90 mΩ at T = 4.2 K). The specimen was then mounted with silver paint to a transistor header. Device contacts were made by ultrasonic wire-bonding of a single, 25 μm-diameter Al lead to the Au overlayer. The resulting package was suitable for repeated temperature cycling between room temperature and 4.2 K.

*The etch rates for $NiSi_2$ and $CoSi_2$ are approximately 25 Å/sec and 17 Å/sec, respectively [Ref. 2].

a)

b)

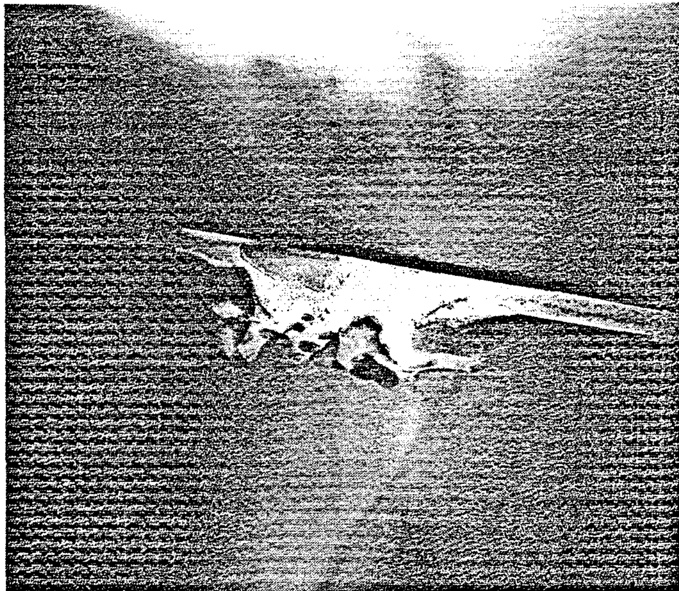


Figure 3.1: Unusual surface features of a 250 Å-thick $CoSi_2$ film grown epitaxially on degenerate Si, seen above in (a) SEM, and (b) cross-sectional TEM images. The surface features are highly irregular "craters" varying roughly between 0.1–1 μm across and extending a few thousand angstroms down into the substrate. The origin of these craters is unclear at present. (TEM photo taken by E. L. Hall, General Electric Research and Development Center.)

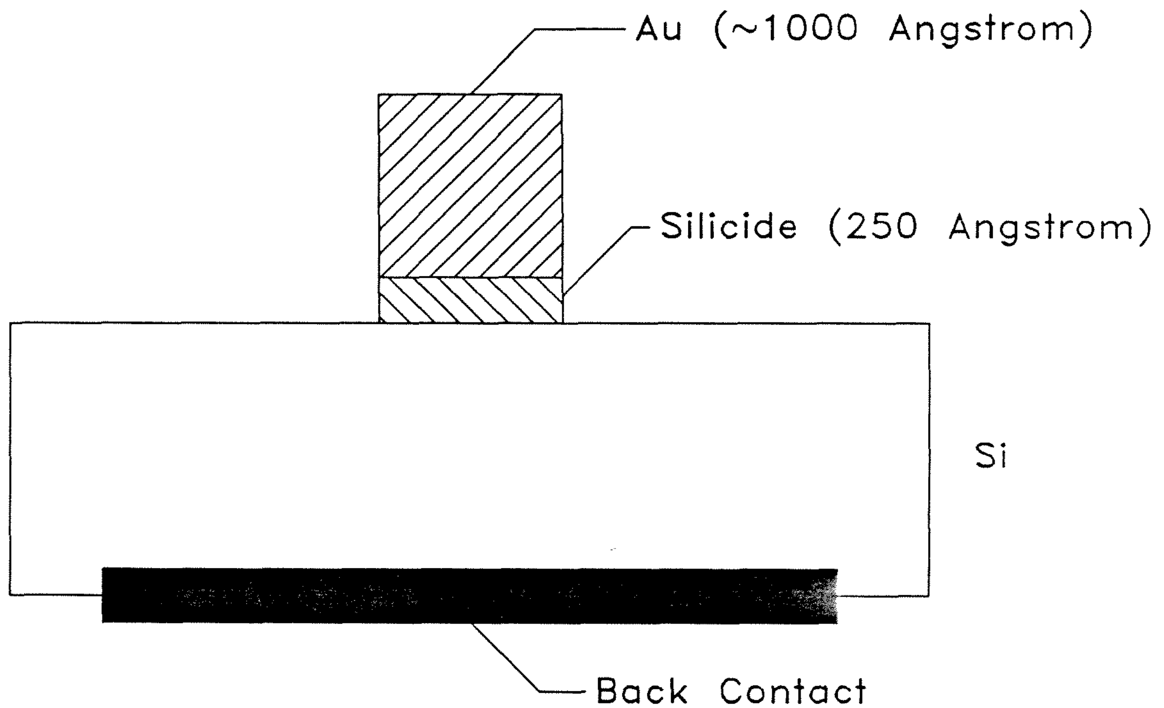


Figure 3.2: Device structure of tunneling Schottky diodes. An overlayer of Au ($\approx 1000\,\text{Å}$) is evaporated to facilitate the device contact (see text). Contact to the silicide is made by wire-bonding an Al lead to the Au film. The back contact is achieved by scratching an In-Ga amalgam into the Si substrate back surface.

There are two reasons for the use of the Au overlayer in our device structures: (1) to increase the effectvie contact area to the silicide, and (2) to reduce the leakage-current effects introduced by the ultrasonic wire-bonding to an acceptable level. Initially, the wire-bonding was made directly to the epitaxial silicide film. However, in this procedure it was observed that the zero-bias conductance of variously-sized devices scaled *sublinearly* with device area, rather than the required linear dependence (nor was the scaling proportional to device *diameter*). This is in contrast to a correct scaling with area observed earlier for silicide films on *nondegenerate* Si. The reason for the difference is that, in the nondegenerate case, the sheet resistance in the epilayer (and hence the voltage drop) is negligible in comparison with the resistance of the Schottky barrier itself whereas these two resistances are comparable in the degenerate case. In other words, with a small, single-lead contact to the silicide, the silicide–degenerate-Si interface is not an *equipotential* surface because of the non-negligible voltage drop along the film as one moves parallel to the interface. As we are about to see below, the use of a comparatively thick (1000 Å or more), low-resistivity Au overlayer solved this problem.

Minimum ultrasonic power and bonding time were found necessary to avoid damaging the silicide. When this precaution was not taken, the zero-bias conductance of a given device was observed to increase linearly with the *number* of wire-bond contacts made to the device, and, in fact, to dominate the observed conductance of the device. However, with the aid of the Au overlayer described above, along with minimal ultrasonic power and bonding time, a device conductance independent of the number of bonds was achieved routinely. As a check on our procedures, a thin, Au-wire probe, *i.e.*, a *nondestructive* contact, was used on a set of devices for comparison with results in which the *same* devices were later wire-bonded. As can be seen from Fig. 3.3, the validity of our

contacting technique is verified by comparison of zero-bias resistance vs. device diameter data for Al wire-bonded and Au wire-probed devices. When proper care was taken, zero-bias resistances were observed to scale with (diameter)$^{-2}$ at $T = 300\,\mathrm{K}$ and $4.2\,\mathrm{K}$, and the same specific contact resistance for both probed and bonded devices was obtained.

3.3.3 Derivative Measurements

Second Derivative

The tunneling spectra (d^2I/dV^2 vs. V) were taken at $T = 4.2\,\mathrm{K}$ with the sample immersed in liquid He. The second derivative signal was directly measured with the use of the second-harmonic synchronous detection technique.[20] In this technique, a small sinusoidal voltage $v = v_\omega \cos \omega t$ is superimposed on a dc voltage V_0. The current I through the sample is then given by

$$I(V) = I(V_0 + v) = I(V_0) + \frac{dI}{dV} v_\omega \cos \omega t + \frac{1}{2} \frac{d^2 I}{dV^2} v_\omega^2 \cos^2 \omega t + \cdots . \tag{3.6}$$

Using a simple trigonometric identity, we see that the above expression contains a *second harmonic* term of the form

$$i_{2\omega} = \frac{1}{4} \frac{d^2 I}{dV^2} v_\omega^2 \cos 2\omega t, \tag{3.7}$$

arising due to the nonlinearity of the *I-V* characteristic. It is the "response" of the device at frequency 2ω that is directly measured with a lock-in amplifier, which, for a fixed excitation voltage v_ω, is proportional to d^2I/dV^2. By measuring the second harmonic signal as a function of dc bias V_0, the tunneling spectrum is obtained. In practice, the second harmonic signal which is detected is not given simply by Eq. 3.7. Instead, the measured signal depends on the circuit design of the apparatus involved in the measurement, and on the impedance of the device under test. The actual apparatus used for second derivative measurements in

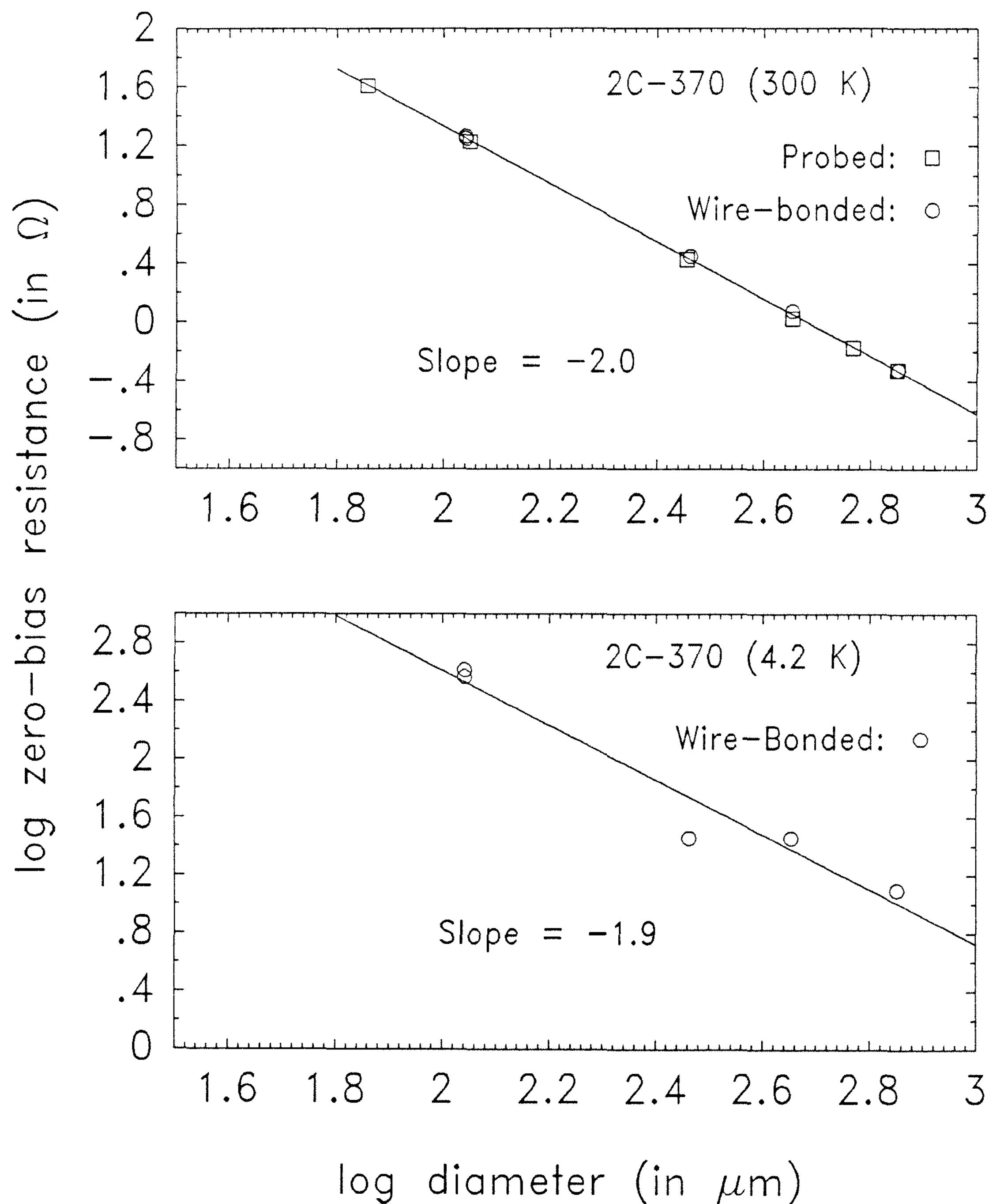


Figure 3.3: Zero-bias resistance vs. device diameter, plotted on a logarithmic scale, for several devices. The slope equals −2 as required. (a) Comparison of probed and wire-bonded device contacts at room temperature. (b) Wire-bonded device contacts at $T = 4.2\,\mathrm{K}$.

the present work is described in detail by Collins.[21] Here, the quantity measured turns out to be the second harmonic *voltage* $v_{2\omega}$ appearing across the sample (see the discussion in the Appendix of Ref. 21). This quantity is given by[21]

$$v_{2\omega} = \frac{1}{4}\frac{dV}{dI}\frac{d^2I}{dV^2}v_\omega^2. \tag{3.8}$$

Hence, our spectra are not strictly proportional to merely d^2I/dV^2, but have an additional voltage dependence due to the factor dV/dI. However, in this work we shall be concerned with the position of *peaks* in the second derivative spectrum, and not the detailed shape of the background. The factor dV/dI is a relatively smooth, slowly varying function compared with d^2I/dV^2. Therefore, in the results presented below, what is referred to as "d^2I/dV^2" is actually the second harmonic signal, *i.e.*, a quantity *proportional* to d^2I/dV^2 and a comparatively slowly varying function of bias; this is common practice in tunneling spectroscopy.

First Derivative

In addition to the second derivative measurements, direct measurements of dV/dI were made with the use of a similar ac method. In this case, a sine wave generator and a large (600 Ω) series resistance were used to supply an ac current i_ω through the device under test while the sample was biased at a nominal dc voltage V_0. The ac voltage at the fundamental frequency v_ω (proportional to dV/dI) was measured as a function of V_0. The measurement was calibrated with the aid of fixed resistors. The circuit used for the dV/dI vs. V measurement is similar to that shown in Fig. A.1 of Ref. 21. Oscillator frequencies of $\omega = 5\,\mathrm{kHz}$ and $50\,\mathrm{kHz}$ were used for the first and second derivative measurements, respectively.

Three-Point Method

Because of the very small average impedance of our tunnel devices ($\sim 1\,\Omega$), a

three-point measurement scheme[22] was employed to minimize the effects of the non-negligible series resistance introduced by the back contact. The main effect of a resistance in series with a nonlinear tunneling element is to shift features of the tunneling spectrum to higher values of bias amplitude. The three-point measurement is illustrated in Fig. 3.4. Current is supplied through the device under test through the device and back contacts. However, the voltage is sensed between the device contact and the contact on a neighboring device. Mikkor and Vassel[22] have discussed the conditions under which this scheme provides a valid measurement of the voltage across the active region, *i.e.*, tunnel junction, of the device under test, and these conditions are satisfied here. Our results show that the three-point measurement effectively removes the series resistance due to the substrate and back contacts. Still remaining, however, is the ($\approx 0.1\,\Omega$) resistance of the Al lead to the device. This lead resistance did become apparent in the spectra of some of our less resistive tunnel devices by shifting spectral features to higher bias amplitude. However, a simple numerical correction was sufficient to bring such spectra into coincidence with spectra obtained on more resistive devices.

3.4 Results

3.4.1 Second Derivative Spectra

The results of tunneling spectroscopy experiments performed on $CoSi_2$, and type-A and type-B $NiSi_2$ epitaxial layers on the Si:As substrate are summarized in Fig. 3.5. In this figure, we present d^2I/dV^2 (the second harmonic signal) as a function of dc bias applied across the tunnel structure. The upper two traces have been vertically displaced from the indicated zero level for clarity. We see that the tunneling spectra obtained from all three types of samples are quite

Three–Point Contact

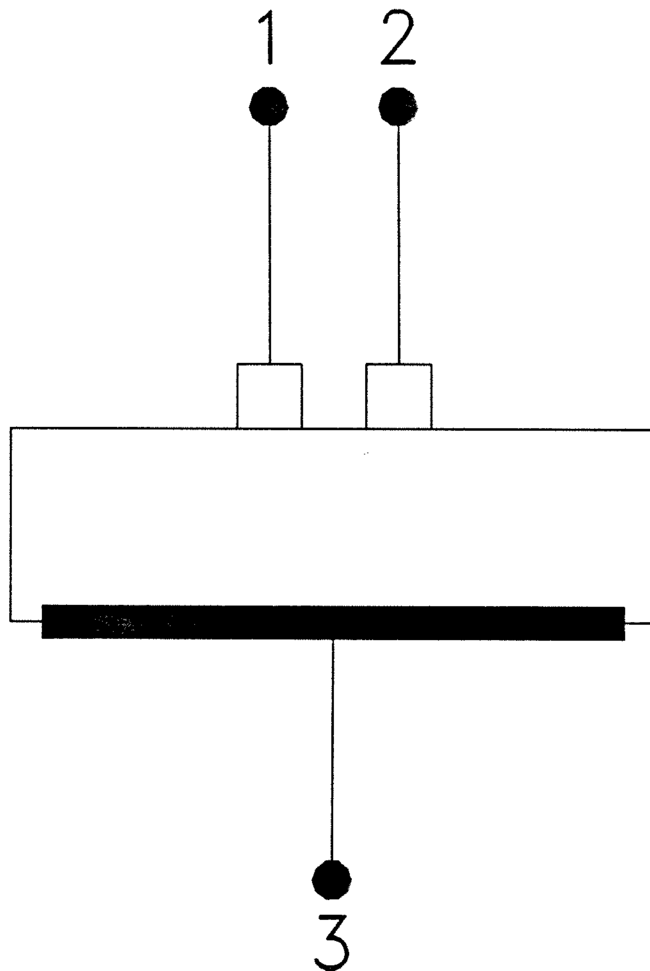


Figure 3.4: Three-point contact scheme employed in all electrical measurements. The ac and dc inputs are supplied through contacts 1 and 3, and the voltage components across the tunnel structure are measured across contacts 1 and 2.

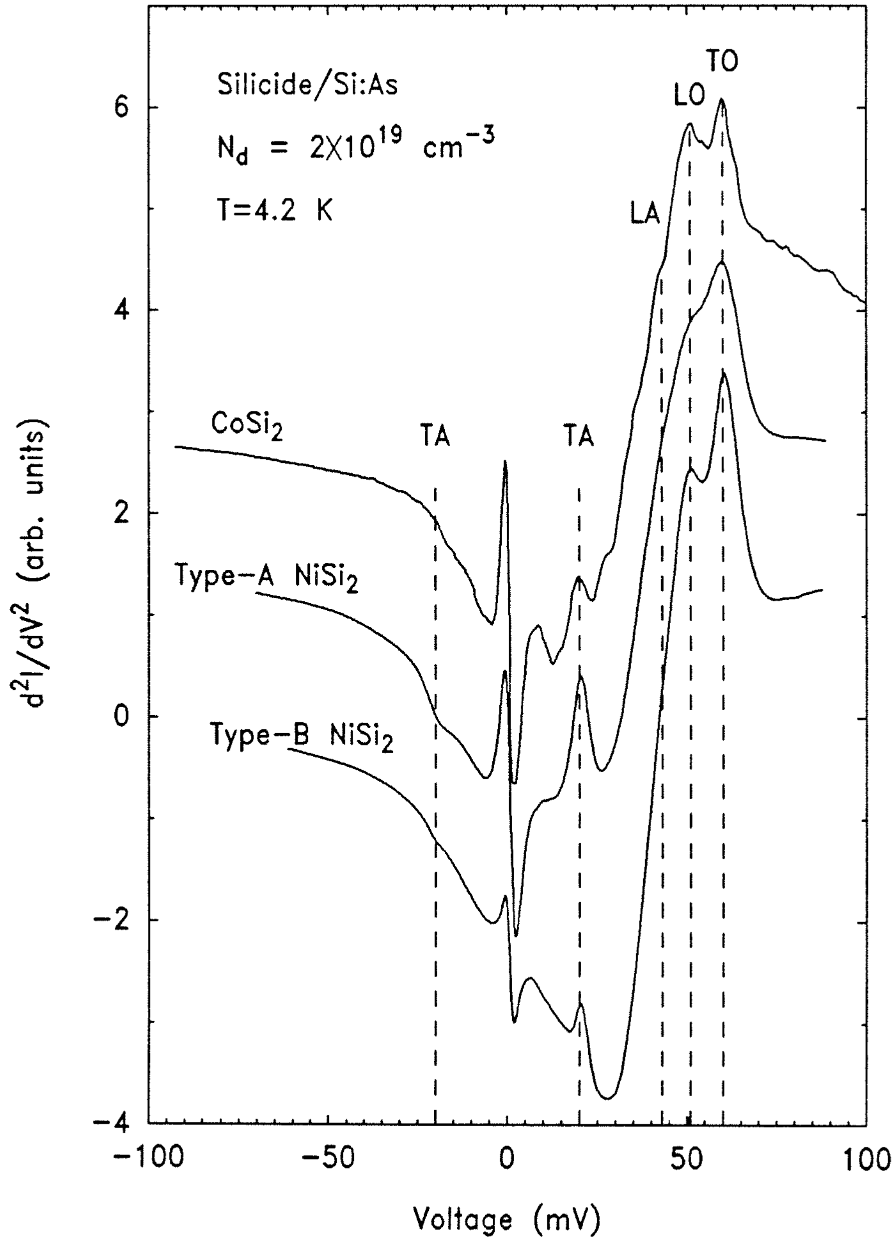


Figure 3.5: Second derivative tunneling spectra for $CoSi_2$ and $NiSi_2$ single-crystal films on Si. The top two spectra have been shifted vertically for clarity. The principal peaks seen in forward bias correspond to **k**-conserving bulk Si phonons at energies of 20,51, and 60 ± 1 meV for TA, LO, and TO phonons, respectively. The 43 meV LA phonon position is also indicated. Note the antisymmetric TA peak in reverse bias for $NiSi_2$.

similar in appearance. The strong, antisymmetric structure near $V = 0$ is the so-called zero-bias anomaly[23] and will not be of concern to us here. There are three strong peaks seen in forward bias (electrons tunneling from the Si to the silicide) corresponding to energies of 20, 51, and 60 meV, respectively. Following the discussion given in Section 1.3.3, we see that these three strong peaks appear to correspond to the TA, LO, and TO phonon branches of bulk Si. In addition, for the $NiSi_2$ samples there is a negative peak of smaller amplitude at a bias of -20 mV. Hence, let us tentatively identify the peaks at ± 20, 51, and 60 meV as TA, LO, and TO phonon peaks. (For now, this is merely a labeling scheme; the precise physical significance of these peaks shall become clear as we proceed.) We observe that the TA, LO, and TO peaks are highly reproducible from sample to sample although some temporal variation has been observed, which we believe to be related to the degradation of our back contacts with time.

The spectra obtained from $CoSi_2$ samples appear more structured than those from $NiSi_2$. This is shown in more detail in Fig. 3.6. There, we compare the forward bias portions of scans taken on several devices. We see that, while the major peaks (TA, LO, and TO) reproduce reasonably well from device to device, there appears to be additional, more complicated structure which does not reproduce well from device to device. This complicated structure appears to become more significant as the $CoSi_2$ devices age; or more precisely, as the *back contact* ages. We consistently observed an aging effect in our back contacts—they become more resistive with time. This degradation of the back contact appeared to correlate with a decreased reproducibility in our tunneling spectra. Hence, the results which we present here have all been taken on samples with *freshly* prepared back contacts. However, as Fig. 3.6 shows, even in this case we observe an additional, irreproducible structure on our $CoSi_2$ (but not our $NiSi_2$) films grown on the heavily doped substrate. At present, we do not understand

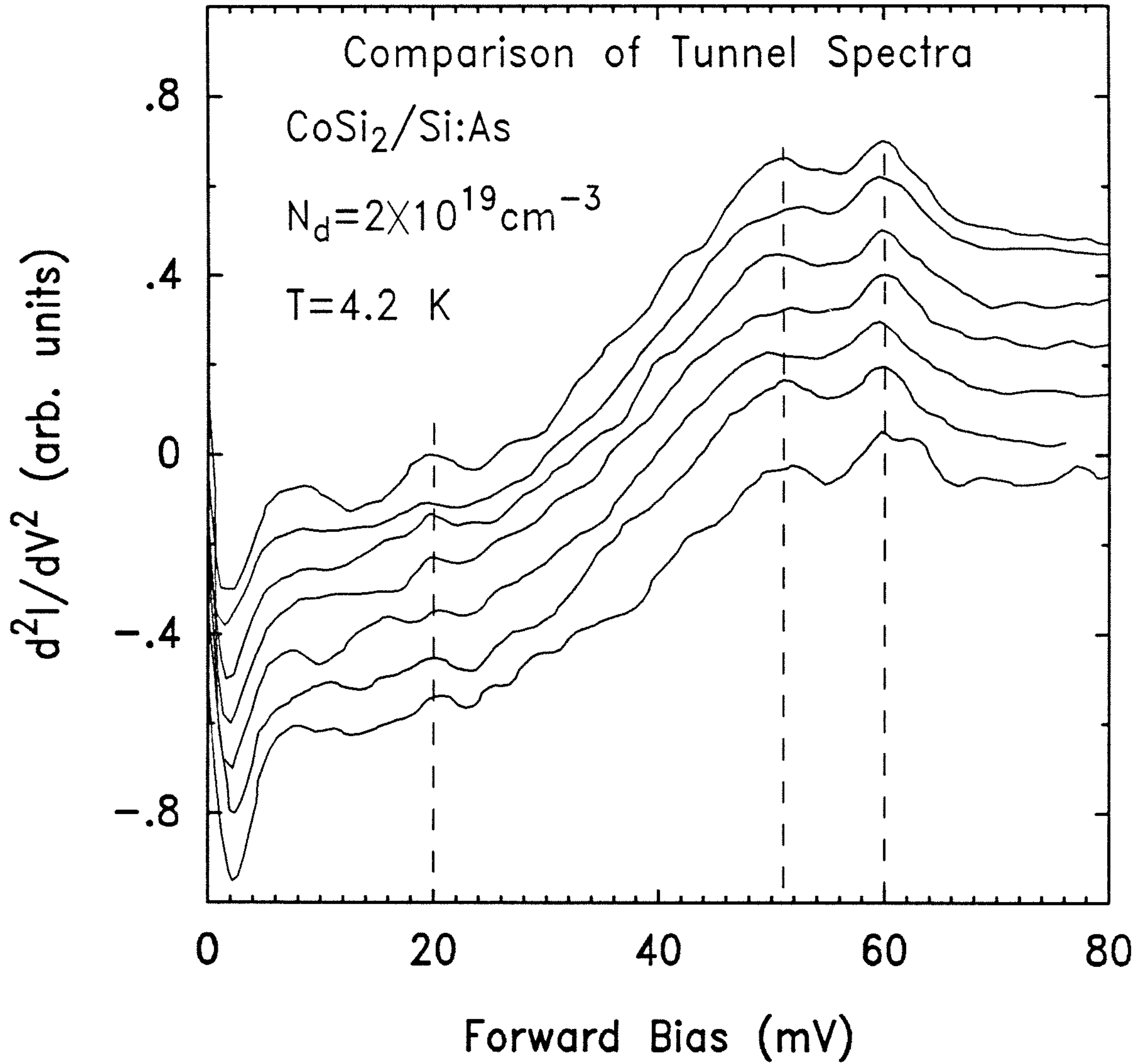


Figure 3.6: Tunnel spectra compared for several $CoSi_2$ devices. The three principal peaks (indicated) reproduce fairly well while the other features do not.

the origin of this additional structure though it may be related to the surface defects shown above in Fig. 3.1, which are also observed only on our $CoSi_2/Si$ structures.

3.4.2 First Derivative Spectra

Figure 3.7 shows the results of first derivative measurements, dV/dI vs. V. The data were taken on five devices ranging from 70–150 μm in diameter. We saw in Section 3.2.1 above that, for tunneling Schottky diodes, a plot of dV/dI vs. V is expected to contain a maximum in the forward bias direction. As can be seen from the figure, all of the devices shown exhibit the expected behavior in a consistent manner. (The variation shown in the figure is quite reasonable for tunnel structures.) The figure shows that the specific differential resistance is independent of device size. The zero-bias resistance is $R_0 = (5.7 \pm 0.3) \times 10^{-4}\,\Omega\text{-cm}^2$. For the case of a metal on n-type Si, the maximum in dV/dI occurs at a forward bias value corresponding to the Fermi degeneracy[24] μ_F. In the absence of carrier-freezeout and band-tailing effects, μ_F is related to the free-carrier concentration n according to

$$\mu_F = \frac{\hbar^2}{2m_d^*}\left(\frac{3\pi^2 n}{M_c}\right)^{2/3}, \tag{3.9}$$

where m_d^* is the density-of-states (DOS) effective mass of electrons, and M_c the number of degenerate conduction band minima, in Si. The position of the maximum corresponds to a Fermi degeneracy $\mu_F = 39\,\text{meV}$,[10] and hence, a free-carrier concentration $n = 3.9 \times 10^{19}\,\text{cm}^{-3}$. This value of n exceeds the carrier concentration $n \approx 2.3 \times 10^{19}\,\text{cm}^{-3}$ established nominally by Hall and resistivity measurements of our Si:As substrates. The significance of this discrepancy will be discussed further in Section 3.5.5 below.

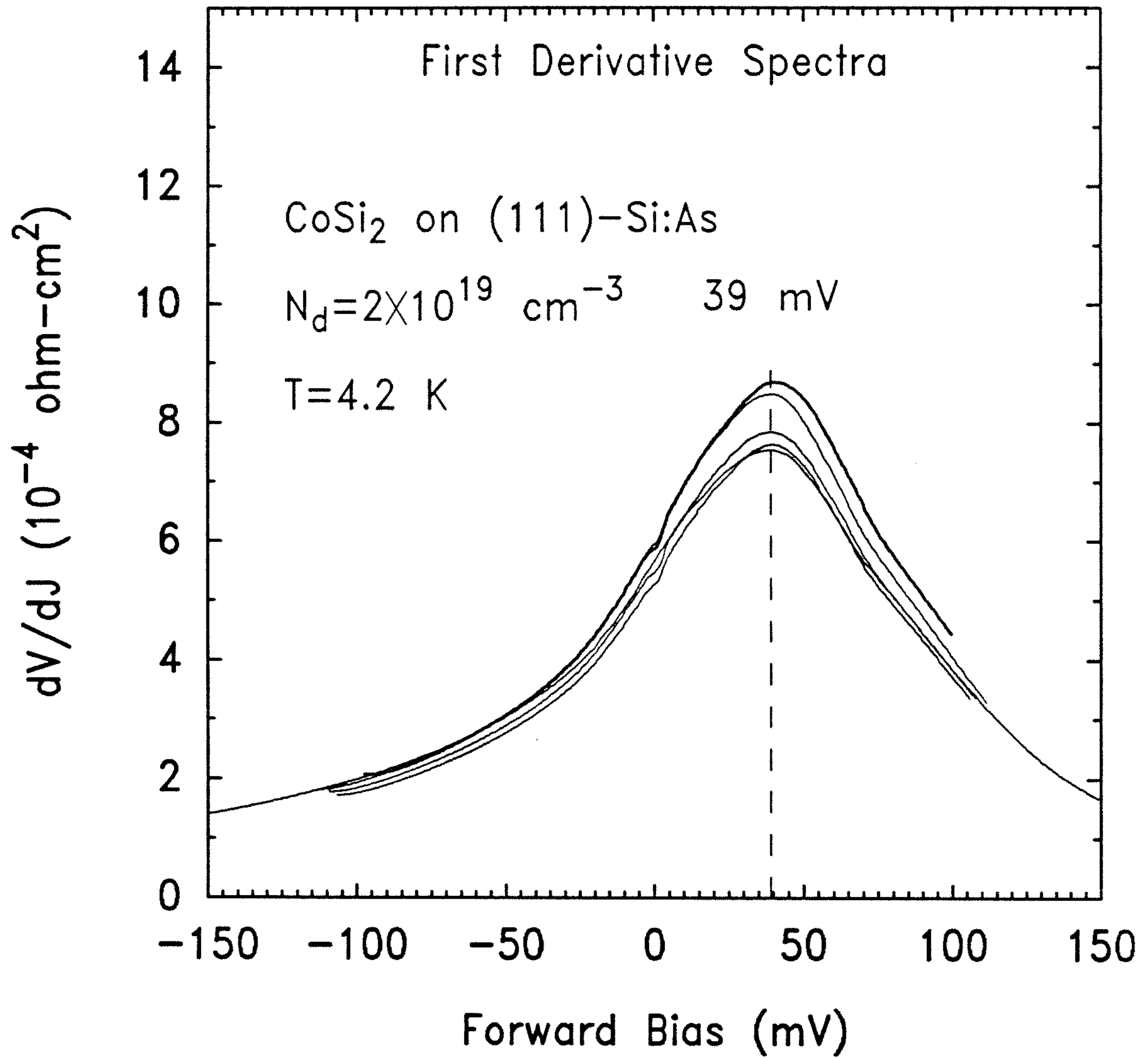


Figure 3.7: Small-signal specific resistance vs. bias for various device sizes. The peak corresponds to the Fermi degeneracy outside the Si depletion region.

3.4.3 Lower Substrate Dopings

Rosencher *et al.*[15] have reported tunneling spectra taken on $CoSi_2$ films grown on Si doped with P at a level of $N_d = 5 \times 10^{18}\,\mathrm{cm}^{-3}$. It is interesting that the tunneling spectra of Rosencher *et al.*[15] exhibit a number of significant differences from our tunneling spectra. Their spectra: (1) do not contain our prominent TA and TO peaks, (2) do contain relatively strong LA and LO peaks in forward bias whereas coupling to longitudinal phonons is much weaker than to transverse phonons for both our devices and for Esaki diodes,[25] and (3) in general appear much more complicated in both bias directions than our Si:As spectra. Points (1) and (3) are consistent with our own tunneling results on Si:Sb substrates of approximately the same doping concentration. Figure 3.8 shows representative tunnel spectra of $CoSi_2$/Si tunnel devices on the lower doped substrate. In contrast to the spectra obtained on the heavily doped substrate (Fig. 3.5), the reproducibility of these spectra is poor. The situation is even worse for $NiSi_2$ as seen in Fig. 3.9. With the exception of sample 2C-304 (70 Å-thick $NiSi_2$ layer), the tunneling spectra of $NiSi_2$ films on the lightly doped substrate consistently exhibit wild, oscillatory structure though the details of the spectra do not reproduce from sample to sample, or even among devices on the same specimen.

The origin of this extremely complicated structure is not understood. However, it is unlikely to be an artifact of the back contact for several reasons: the tunneling spectra from two devices of similar size and impedance on the *same* specimen are distinctly different; the spectra from two different specimens from the same sample (growth) are qualitatively similar, but qualitatively different from other *samples*, *e.g.*, compare Figs. 3.8, 3.9(a), (b), and (c); the results obtained by wire-bonding a lead directly to the Si substrate (without contacting silicide film) do not exhibit the detailed structure shown in Figs. 3.9(a) and (b), but merely show a large zero-bias anomaly. Nor is it likely that the complicated

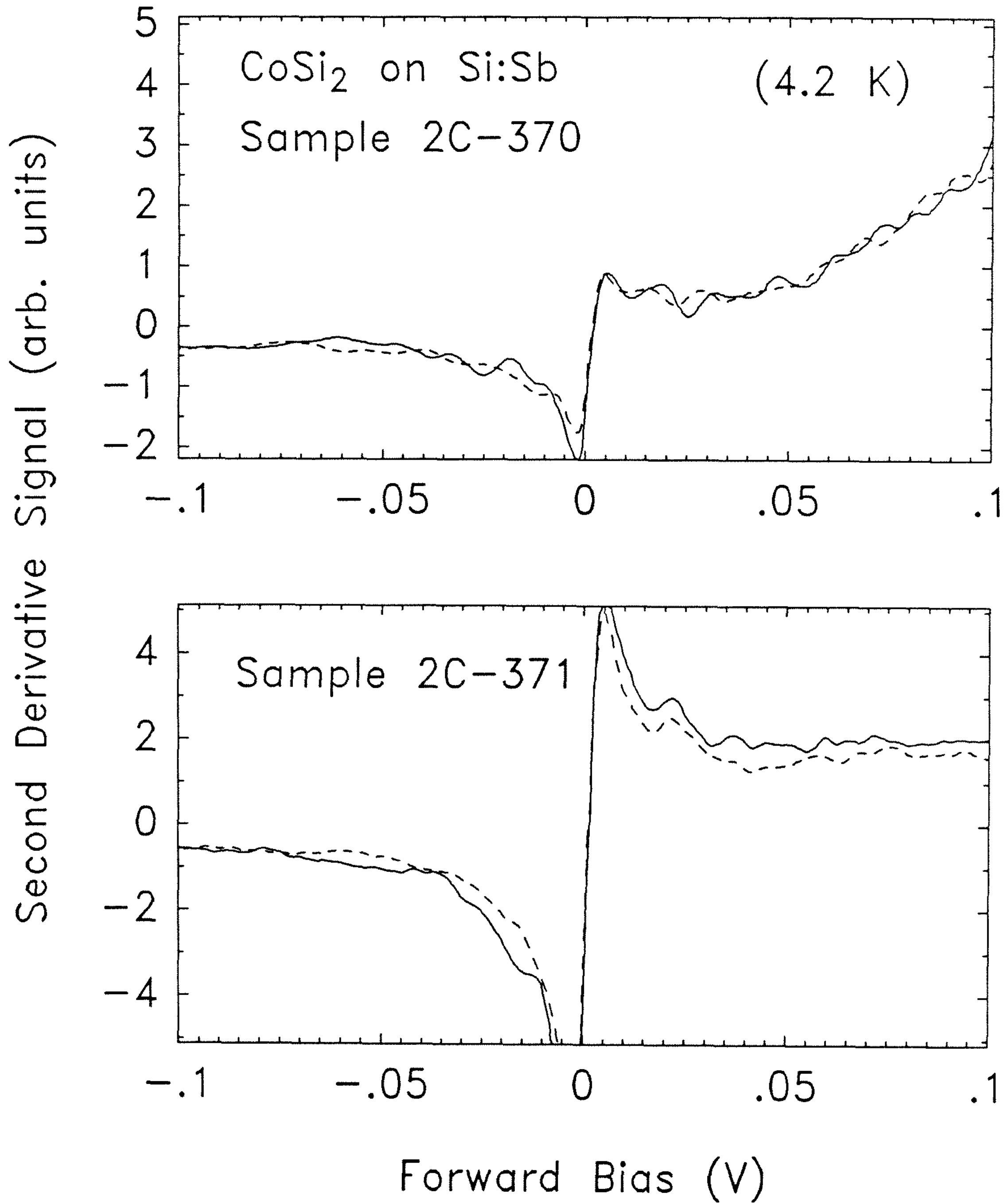


Figure 3.8: d^2I/dV^2 vs. V for 250 Å-thick $CoSi_2$ films on Si:Sb. Two devices from sample 2C-370, and two from 2C-371 are shown. Note the poor reproducibility of the spectra, and the absence of the major peaks which are seen on the Si:As substrate.

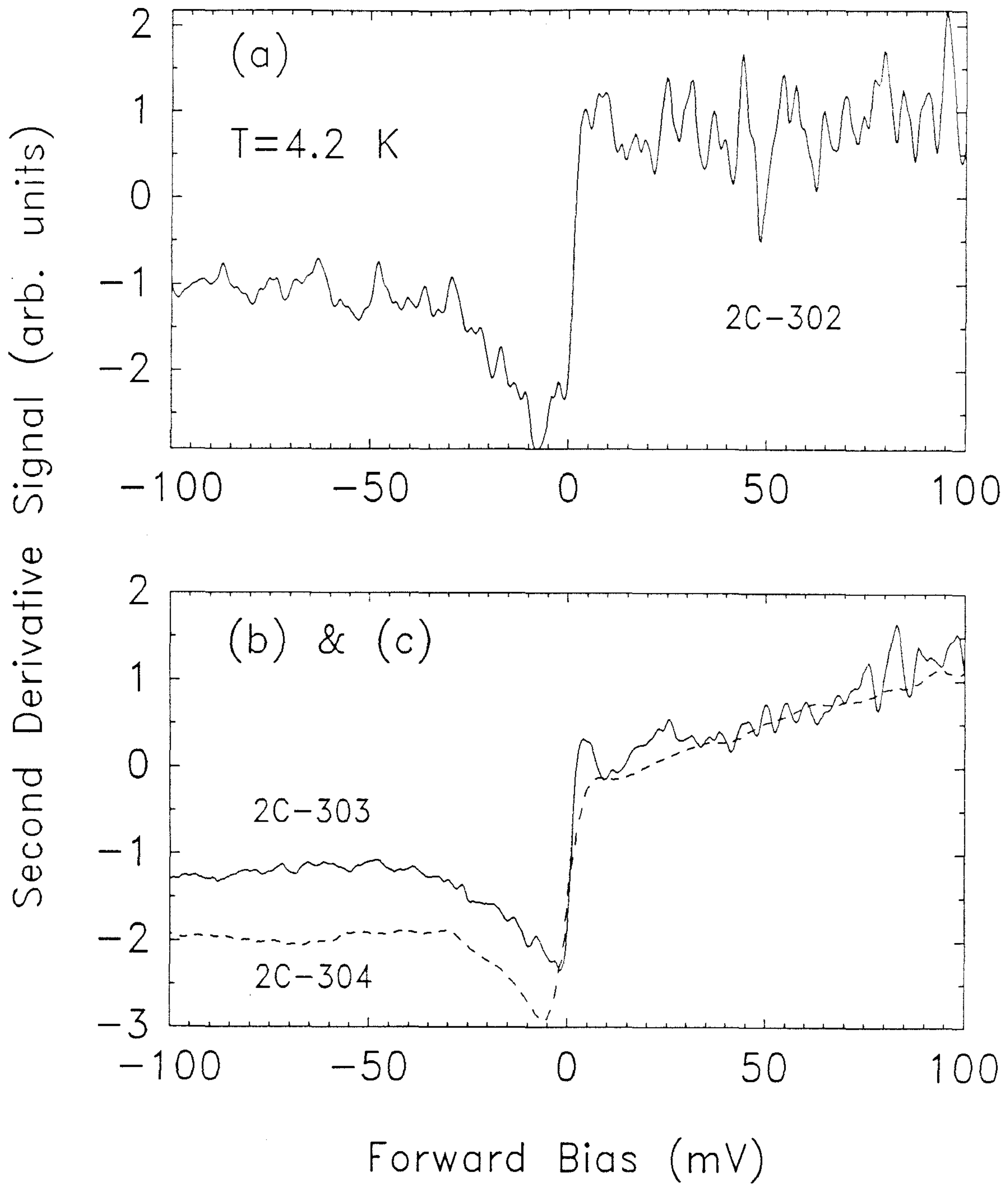


Figure 3.9: d^2I/dV^2 vs. V for type-A and type-B $NiSi_2$ films on Si:Sb. Typcial spectra are shown for (a) sample 2C-302, (b) 2C-303, and (c) 2C-304. The peak positions are not reproducibile. However, for a given sample, similar looking spectra are consistently obtained.

structure is an artifact of the second derivative measurement apparatus since the spectra on a given device are reproducible over time (each scan shown is the average of 25–100 scans), and a given structure persists reproducibly as scan rate and modulation voltage are varied. Hence, all we can say about the complicated structure is that it appears to be associated with the *tunnel structure* rather than the back contact.

3.5 Discussion

3.5.1 Identification of Peaks

Peaks in d^2I/dV^2 spectra similar to the ones shown in Fig. 3.5 have been reported much earlier on Si Esaki diodes by Chynoweth *et al.*[25] These peaks were identified as corresponding to $\mathbf{k}$-conserving Si bulk phonons. If we compare the tunnel spectra shown in Figs. 1.14 and 3.5, we see a number of similarities between the Esaki diode and our Schottky diode spectra: larger peak amplitudes in forward bias, an antisymmetric (TA) peak near $\pm 20\,\mathrm{mV}$, and a barely discernible (LA) peak near $43\,\mathrm{mV}$.[25] In Fig. 3.5 we saw a weak shoulder near the expected LA phonon energy in the $CoSi_2$ tunneling spectrum, but this structure is entirely absent from both of the $NiSi_2$ spectra, and, as we saw in Fig. 3.6, is not very reproducible from sample to sample as are the larger peaks.

For the Esaki diode, phonon-assisted transport is *necessary* to conserve crystal momentum as the electrons tunnel from the indirect valleys in the n-type region into zone-center hole states,[26] as we discussed in Section 1.3.3. However, for the case of metal-semiconductor tunneling, the Fermi surface of the metal is usually large enough in $\mathbf{k}$-space to permit unassisted electron tunneling transport to occur.[11] Therefore, it is generally much more difficult to observe $\mathbf{k}$-conserving phonons in metal-semiconductor transport. While $\mathbf{k}$-conserving phonons have

been reported for vacuum-cleaved n-type-Ge/metal systems,[17] we are unaware of any similar results for Si, particularly at the doping levels used in our study. Spectral peaks at the Si LO_Γ phonon frequency ($\approx 65\,\text{meV}$) have been reported for metal films on degenerate, p-type Si, but in that case, the peaks are *symmetric* with respect to bias direction, and are attributed to modification of the electronic band structure due to electron-phonon coupling.[13] In contrast, an *antisymmetric* structure (as observed at $\pm 20\,\text{mV}$ in Fig. 3.5) indicates *assisted, inelastic* transport.[13] Thus, on the basis of the observed peak energies and qualitative similarities between our results and the Esaki diode tunnel spectra, we identify the prominent peaks in Fig. 3.5 with inelastic electron transport assisted by **k**-conserving Si phonons.

3.5.2 Mechanism of Inelastic Tunneling

An evanescent gap state described by an imaginary wave vector connecting the zone-center band extrema has been used to theoretically explain the observed bulk phonon structure in n-type-Ge/metal tunneling spectra.[17] In that case, the tunneling electron, initially, *e.g.*, in one of the semiconductor indirect conduction band minima, first couples to a zone-center, evanescent gap state. In this process, a (bulk) **k**-conserving phonon is emitted, and an amount of electron energy $\hbar\Omega$ is lost. Then, the electron couples elastically from the gap state into a real metal state having $\mathbf{k}_\parallel = \vec{0}$ (in order to conserve energy and $\mathbf{k}_\parallel$ in tunneling across the interface). A similar process may account for our observation of **k**-conserving Si phonons in the electron tunneling spectra presented here. However, the direct energy gap in Si, relative to the indirect gap, is much larger[26] than it is in Ge. Hence, we should expect the fraction of inelastic tunneling current to be much weaker for the Si case* due to a larger imaginary wave vector. (The dependence of

*Compare, for example, Fig. 3.7 with Fig. 2 of Ref. 17.

tunneling probability on imaginary wave vector was discussed around page 39.) In fact, the inelastic features in our tunneling spectra do appear to be much weaker than those reported[17] for metal-Ge tunnel structures. Finally, it is worth mentioning here that we evidently do not observe any structures which may be *unambiguously* identified as silicide phonons or final state effects,[20] and the precise significance of the crystallographic perfection of the silicide layer in the observed tunneling spectra is unclear.

3.5.3 Comparison to Au/Si Structures

To better understand the significance of metal-layer single-crystallinity in our tunneling spectra, Au films were thermally evaporated onto chemically etched wafers which were nominally identical to the heavily and lightly doped substrates used for the silicides. The wafers were degreased and etched for 1 min in 20% HF, and then promptly loaded into a vacuum chamber ($\sim 10^{-6}$ Torr 1 min after loading), after which a 1000–2000 Å Au film was immediately evaporated. The resultant tunneling spectra are shown in Figs. 3.10(a) and (b). Figure 3.10(a) shows a typical spectrum for Au on the Si:As substrate. For the Si:As substrate, reproducibility was excellent. The TA, TO, and LO peaks previously identified in Fig. 3.5 are all readily apparent. In the scan shown here, there is, in addition, a very slight indication of the LA phonon peak in the forward bias direction. Further, in the reverse bias direction, the only significant feature is an antisymmetric peak at the TA phonon energy. Comparison of Figs. 3.10(a) and 3.5 shows that essentially the same result is obtained, independent of the identity or crystallographic orientation or perfection of the metal layer.

In contrast, the tunneling spectra of the Au film on the Si:Sb substrate appear complicated, and the features do not reproduce well from device to device. Figure 3.10(b) shows the d^2I/dV^2 vs. V spectra for two devices on the same

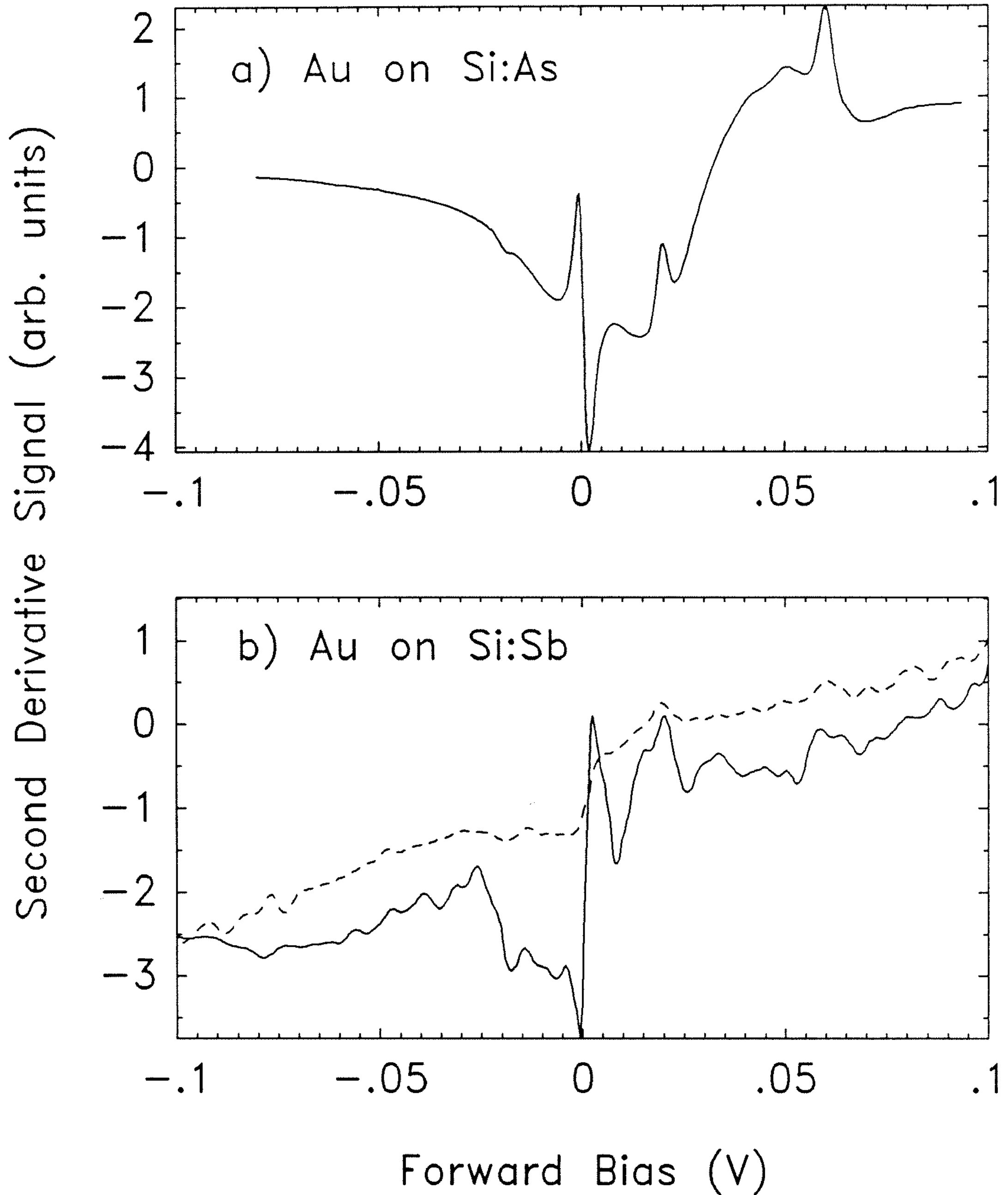


Figure 3.10: d^2I/dV^2 vs. V spectra for polycrystalline Au films evaporated onto (a) the Si:As, and (b) the Si:Sb substrate. Spectrum (a) is highly reproducible and is essentially identical to spectra obtained with single-crystal silicide films while spectra (b) contain many small features which do not reproduce well in detail.

specimen. While it is tempting to assign the forward bias peaks near 20 and 60 meV to TA and TO phonons, it must be stressed that the two scans shown do not provide sufficient evidence for this identification. The relatively detailed structure and lack of reproducibility are similar in character to the spectra shown in Figs. 3.8 and 3.9. Hence, the detailed structure, while associated with the tunnel structure rather than the back contact as argued in Section 3.4.3 above, is evidently *not* due to the single-crystallinity of the metal overlayer. In view of our results on Au/Si tunnel structures, there appears to be little if any significance to our results of the single-crystal nature of the metal films.

3.5.4 Comparison to WKB Theory

As mentioned above, the predominant electronic transport mechanism in metal-semiconductor tunnel structures is elastic tunneling. To characterize our silicide/Si tunnel structures, here, we compare measured $\log I - V$ and dV/dI vs. V characteristics with the WKB theory discussed above in Section 3.2.1. In Fig. 3.11(a), the $\log I - V$ forward characteristics of several devices are compared with theoretical WKB curves calculated with the use of Eqs. 3.1 and 3.2. The data shown in the upper figure are taken from seven $CoSi_2$/Si:As devices ranging in diameter from 70–150 μm. The theoretical curves are calculated assuming a Schottky barrier height (at T=4.2 K) of* $0.64 + 0.05 = 0.69$ eV. From Eq. 3.2, we see that, for sufficient forward bias, the tunneling current assumes the form

$$J \sim e^{qV/E_{00}}. \tag{3.10}$$

*The low-temperature Schottky barrier height of $CoSi_2$ on n-type Si is estimated by taking the room-temperature value, 0.64 eV [Ref. 27], and adding the band-gap increase, 0.05 eV [Ref. 28]. This assumes that the interfacial Fermi level position is pinned with respect to the valence band as discussed in Ref. 29.

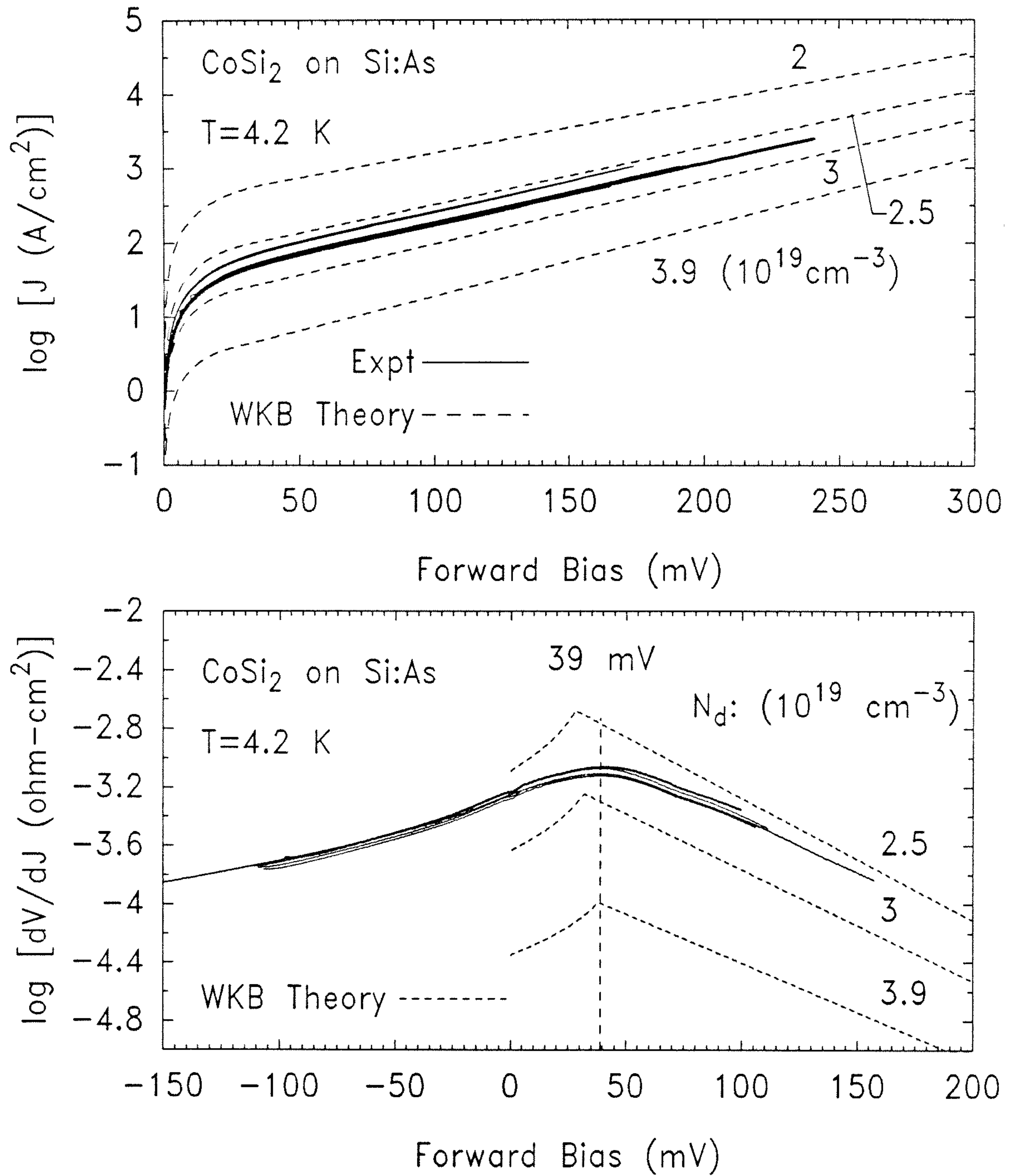


Figure 3.11: (a) Theoretical and experimental $\log I - V$ forward characteristics. A series resistance correction has been applied to the experimental curves. (b) Theoretical and directly measured dV/dI vs. V curves. The maximum is equal to the Fermi degeneracy μ_F.

As shown in the figure, our experimental data exhibit the exponential form* predicted by Eq. 3.10. From the mean slope of the experimental curves, we recover a value $E_{00} = (53 \pm 1)\,\mathrm{meV}$. Putting this value in Eq. 3.5 gives a value for the donor concentration of $N_d = (2.5 \pm 0.1)\,\mathrm{cm}^{-3}$, in good agreement with the results of Hall and resistivity measurements. We see that the experimental $\log I - V$ data agree well both in vertical position and in slope with the WKB theory. (However, the closeness of the agreement is partly fortuitous since WKB predictions of J are good to no more than the correct order of magnitude.)

In Fig. 3.11(b), directly measured $\log(dV/dI)$ vs. V curves for several devices are compared with WKB curves calculated for various dopant concentrations. The significant feature of this type of plot is the position of the maximum, which, as we noted earlier in Section 3.4.2, is equal to the Fermi degeneracy[10,11] μ_F. As shown in the figure, the maxima of the theoretical curves shift to increasing energy as N_d (assumed equal to n for our samples) is increased. The sharp cusp seen in the WKB curves is an artifact of the simple model, which uses two different functional forms (Eqs. 3.1 and 3.2) for forward bias values less than and greater than μ_F, respectively. However, the physical reason why a maximum should occur at $qV = \mu_F$ is straightforward: for forward bias less than μ_F, an increase in bias results in a nearly linear increase in current, primarily because of an increased *number* of electrons available for tunneling; as soon as the forward bias exceeds μ_F, no new electrons are available for tunneling, so the slope of the I-V curve locally flattens out, *i.e.*, the differential resistance (dV/dI) reaches a maximum. For further increases in forward bias, the current again increases, this time exponentially with voltage due to the increased quantum mechanical *transmission probability* of the Schottky barrier as it becomes *thinner* with increased bias (see the discussion on page 39). In this region, the I-V characteristic is

*The exponential dependence of the current on V for sufficient forward bias is a feature of single-barrier tunneling in general, not just Schottky-barrier tunneling.

given by Eq. 3.10. Hence, we see in the theoretical (and experimental) curves in Fig. 3.11(b) the linear dropoff past the maximum.

3.5.5 Doping Enhancements and Band Tailing

The position of the maximum in the first derivative spectrum provides a means of measuring μ_F. From this knowledge, and with the aid of Eq. 3.9, we have a means of determining the free-carrier concentration n. The first derivative maxima shown Fig. 3.11(b) occur at 39 meV for the $CoSi_2$/Si:As films shown (sample 2C-372). Similar results on sample 2C-373 give maxima around 38 meV. We have not made the first derivative measurement on our $NiSi_2$/Si:As samples, but numerically integrated second derivative spectra for these samples also consistently result in a minimum (dI/dV) around 38–40 meV. Using the value $\mu_F = 39\,\text{meV}$ in Eq. 3.9 gives a free-carrier concentration of $n = 3.9 \times 10^{19}\,\text{cm}^{-3}$. However, this value is greater than the free-carrier concentration determined by Hall measurements ($2.3 \pm 0.6 \times 10^{19}\,\text{cm}^{-3}$), and the concentration determined by resistivity measurements ($\approx 2.5 \times 10^{19}\,\text{cm}^{-3}$). We have also seen in Fig. 3.11 that, aside from the position of the first derivative maximum, the experimental data are in better agreement with theory for a concentration of $2.5 \times 10^{19}\,\text{cm}^{-3}$ than for $3.9 \times 10^{19}\,\text{cm}^{-3}$.

The Hall and resistivity determinations of carrier concentration are *bulk* measurements, sampling the entire thickness of the Si substrate. In contrast, the electrical properties of our tunnel structures are extremely sensitive to the properties of the substrate within only a few hundred angstroms at most of the metal-semiconductor interface. Therefore, to account for the discrepancy between the observed first derivative maxima and the bulk-determined carrier concentrations, the existence of a doping enhancement near the interface was postulated. Such an enhancement might arise due to the "snowplowing" of dopants during the

solid-phase reaction which forms the silicide; such effects have been observed in flash-annealed Pd_2Si films.[30] However, because of the manner in which our silicide films were prepared, *i.e.*, with the use of an ultrathin *template layer*, any region of enhanced doping, if it were to occur at all, would necessarily be exceedingly small (less than 70 Å). A direct measurement of the dopant profile with the use of secondary-ion mass spectroscopy* (SIMS) failed to provide unambiguous evidence (or lack thereof) of a doping enhancement near the interface. We mention here that the standard technique of C-V analysis[31] in determining doping profiles is not useful here due to the degeneracy of our samples.

To obtain clear and direct evidence of the supposed doping enhancement, the following experiment was performed: on one $CoSi_2$/Si:As sample, the silicide was chemically removed, and the wafer etched to a depth of $\approx 1\,\mu$m below the original silicide-Si interface; on a second Si:As wafer, on which no silicide had been fabricated, only the native oxide was removed. Au films were then thermally evaporated onto the "etched" and "unetched" specimens, respectively. Finally, first derivative spectra were taken in the same manner as for the silicide films. The results are shown in Fig. 3.12. There is no significant difference in peak position between the etched (39 meV) and unetched (37 meV) samples, nor do the peak positions differ from those obtained for the $CoSi_2$/Si:As samples shown earlier.

The consistent observation of a peak in the first derivative spectra near 37–39 meV on the etched, unetched, and silicide samples suggests the absence of a doping concentration enhancement, but the presence of an intrinsic property of the Si:As substrate that is responsible for the observed behavior. The occurrence of a first derivative maximum at precisely $qV = \mu_F$ assumes a bulk band structure; in particular, a sharp cutoff in the DOS at the conduction band edge

*SIMS measurements were done by Charles Evans and Associates. Oxygen ions were used because of their larger sputtering yield with respect to arsenic.

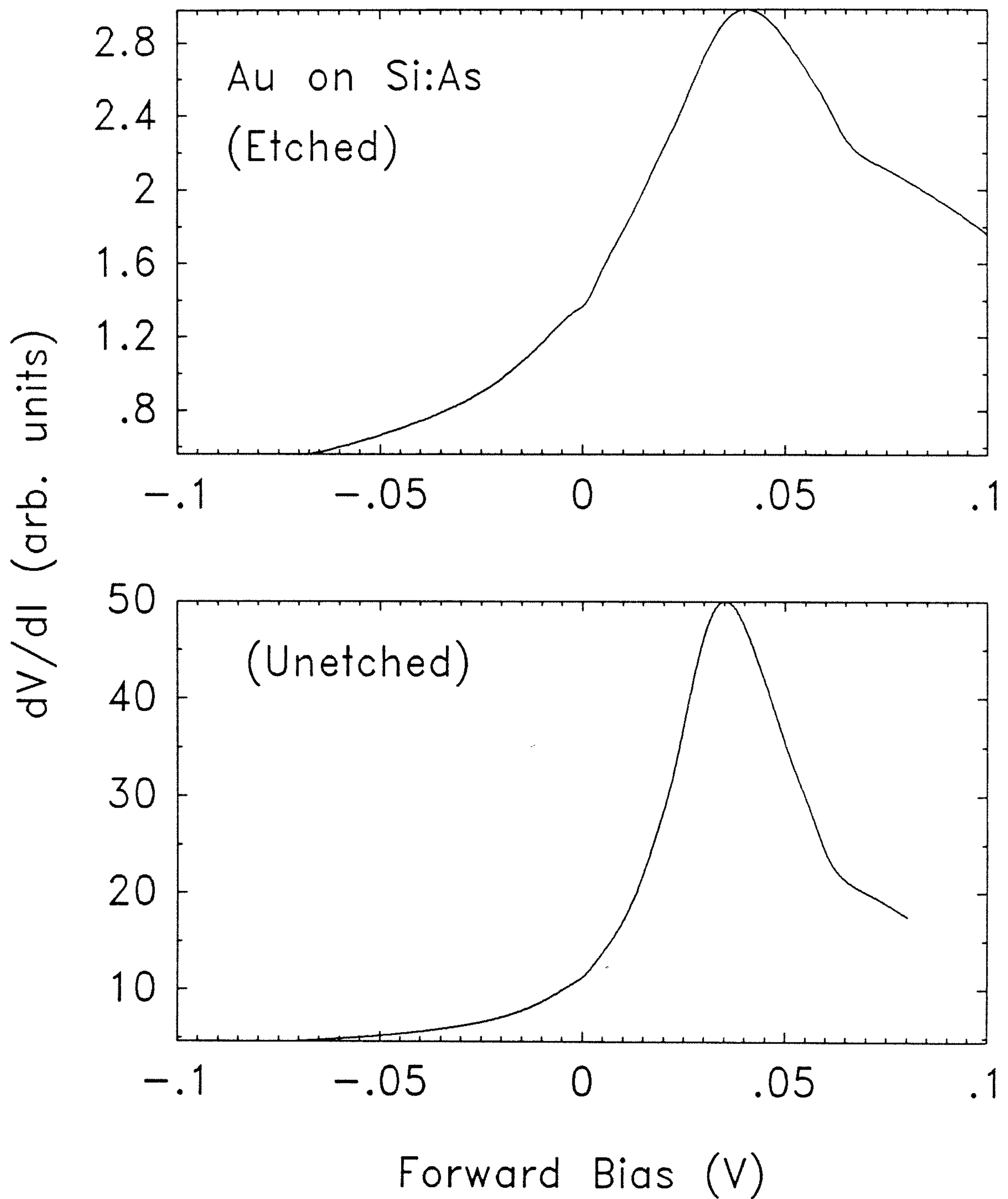


Figure 3.12: dV/dI vs. V for the etched and unetched Au/Si:As tunneling Schottky diodes described in the text. The maxima for the etched and unetched samples are at 39 and 37 meV, respectively.

is assumed. However, at high doping levels, impurity bands are known to form, which, along with local strain effects, smear the sharp edge of the band into the gap.[32] Impurity bands are also known to shift the position of the first derivative maximum to higher energy in *p*-type Si due to the relatively large heavy hole effective mass.[12] On the assumption of an actual doping level of $2.5 \times 10^{19}\,\mathrm{cm}^{-3}$, the upward energy shift due to impurity banding effects would be about 10 meV. This number is comparable to the donor binding energy of As in Si, 13 meV. The impurity band might also account for the rather broad maxima observed in our first derivative spectra.

3.6 Conclusions

We have examined the electron tunneling spectra obtained from single-crystal epitaxial films of $CoSi_2$, and, for the first time, of type-A and type-B $NiSi_2$ films on degenerate, As-doped Si. Our results show direct evidence of **k**-conserving phonons assisting the transport of electrons from the semiconductor into the silicide. Strong, reproducible peaks at the TA, LO, and TO phonon energies corresponding to the wave vector of an indirect valley in Si are observed in forward bias for both $NiSi_2$ and $CoSi_2$. The LA phonon peak is not observed in spectra obtained from $NiSi_2$ samples, but is sometimes weakly observed for $CoSi_2$ samples, along with other weak structure of unclear origin. In the reverse bias direction, the TA phonon is seen in $NiSi_2$ samples only, and is much weaker than and of opposite sign to the corresponding feature observed in forward bias.

The major features of our electron tunneling spectra appear to to be independent of the identity, crystalline perfection, and crystallographic orientation of the metal layer. In particular, essentially identical tunneling spectra are obtained for polycrystalline Au films evaporated onto like substrates. No enhancements

or new features in our tunneling spectra due to the single-crystallinity of the silicide films are evident. Instead, the observed *inelastic* processes appear to be associated with bulk wave vector conservation, suggesting that, at least, for the particular choice of dopant species and concentration considered here, that the observed inelastic processes are dominated by bulklike processes excited within the tunneling barrier rather than by processes related to the metal-semiconductor interface. In addition, the I-V and (dV/dI)-V characteristics of our $CoSi_2/Si$ tunnel structures are shown to be in good agreement with predictions of a simple WKB calculation. Finally, tunneling measurements are also made for single-crystal films grown on barely degenerate, Sb-doped Si substrates. However, these spectra, while usually highly structured, are not reproducible; hence, no useful information is obtained in this case.

References

1. R. T. Tung, J. M. Gibson, and J. M. Poate, Phys. Rev. Lett. **50**, 429 (1983); Appl. Phys. Lett. **42**, 888 (1983).

2. B. D. Hunt, N. Lewis, E. L. Hall, L. G. Turner, L. J. Schowalter, M. Okamoto, and S. Hashimoto, *Mat. Res. Soc. Symp. Proc.* **56** (to be published).

3. E. Rosencher, S. Delage, Y. Campidelli, and F. Arnaud D'Avitaya, Electronics Lett. **20**, 762 (1984).

4. J. C. Hensel, A. F. J. Levi, R. T. Tung, and J. M. Gibson, Appl. Phys. Lett. **47**, 151 (1985).

5. R. T. Tung, A. F. J. Levi, and J. M. Gibson, Appl. Phys. Lett. **48**, 635 (1986).

6. R. T. Tung, Phys. Rev. Lett. **52**, 461 (1984); J. Vac. Sci. Technol., **B2**, 465 (1984).

7. M. Liehr, P. E. Schmid, F. K. LeGoues, And P. S. Ho, Phys. Rev. Lett. **54**, 2139 (1985).

8. R. J. Hauenstein, T. E. Schlesinger, T. C. McGill, B. D. Hunt, and L. J. Schowalter, Appl. Phys. Lett. **47**, 853 (1985); J. Vac. Sci. Technol. **B4**, 649 (1986).

9. B. D. Hunt, L. J. Schowalter, N. Lewis, E. L. Hall, R. J. Hauenstein, T. E. Schlesinger, T. C. McGill, M. Okamoto, and S. Hashimoto, *Mat. Res. Soc. Proc.* **54**, 479 (1986).

10. J. W. Conley, C. B. Duke, G. D. Mahan, and J. J. Tiemann, Phys. Rev. **150**, 466 (1966).

11. R. Stratton, in *Tunneling Phenomena in Solids,* edited by Elias Burstein and Stig Lundqvist (Plenum, New York, 1969), Chap. 8.

12. J. W. Conley and G. D. Mahan, Phys. Rev. **161**, 681 (1967).

13. D. E. Cullen, E. L. Wolf, and W. Dale Compton, Phys. Rev. **B2**, 3157 (1970).

14. J. E. Christopher, H. M. Darley, G. W. Lehman, and S. N. Tripathi, Phys. Rev. **B11**, 754 (1974).

15. E. Rosencher, P. A. Badoz, A. Briggs, Y. Campidelli, and F. Arnaud D'Avitaya, in *Proc. of the First International Symposium on Silicon Molecular Beam Epitaxy,* edited by John C. Bean (The Electrochemical Society, Pennington, 1985), p. 268.

16. R. Stratton and F. A. Padovani, Phys. Rev. **175**, 1072 (1968).

17. F. Steinrisser, L. C. Davis, and C. B. Duke, Phys. Rev. **176**, 912 (1968).

18. L. B. Schein and W. Dale Compton, Phys. Rev. **B4**, 1128 (1971).

19. A. Ishizaka, K. Nakagawa, and Y. Shiraki, *Collected Papers of MBE-CST-2, 1982, Tokyo* (Japan. Soc. Appl. Phys., Tokyo, 1982), p. 183.

20. E. L. Wolf, *Principles of Electron Tunneling Spectroscopy* (Oxford, New York, 1985).

21. R. T. Collins, Ph.D. Thesis, California Institute of Technology, 1985.

22. M. Mikkor and W. C. Vassell, Phys. Rev. **B2**, 1875 (1970).

23. C. B. Duke, in *Tunneling Phenomena in Solids,* Ref. 11, Chap. 28.

24. R. Stratton and F. A. Padovani, Solid-State Electron. **10**, 813 (1967).

25. A. G. Chynoweth, R. A. Logan, and D. E. Thomas, Phys. Rev. **125**, 877 (1961).

26. R. A. Logan, in *Tunneling Phenomena in Solids,* Ref. 11, pp. 154–155.

27. S. M. Sze, *Physics of Semiconductor Devices* (Wiley, New York, 1981), p. 292.

28. In *Physics of Semiconductor Devices*, Ref. 27, p. 849.

29. In *Physics of Semiconductor Devices*, Ref. 27, pp. 270ff.

30. N. S. Alvi, D. L. Kwong, C. G. Hopkins, and S. G. Bauman, Appl. Phys. Lett. **48**, 1433 (1986).

31. In *Physics of Semiconductor Devices*, Ref. 27, pp. 286ff.

32. J. I. Pankove, *Optical Processes in Semiconductors* (Dover, New York, 1971), pp. 10ff.

Chapter 4

Design and Construction of a Si MBE Growth Chamber

4.1 Introduction

This chapter concerns the design and construction of an apparatus built for the purpose of growing epitaxial Si and silicide thin films on a wafer of Si. In this apparatus, "molecular"* beams are produced through *e*-beam evaporation of Si, and/or a selected transition metal. The molecular beams are directed onto a heated substrate under ultrahigh vacuum (UHV) conditions. Under proper conditions, this process, known as molecular beam epitaxy (MBE), results in the formation of a high quality, epitaxial thin film on a crystalline substrate. As we shall see below, the "proper conditions" for epitaxial growth of Si films on Si are actually quite stringent. It is the satisfaction of such stringent criteria that result in much of the cost and complexity of a Si MBE machine, and account for the delayed commercial appearance of Si MBE relative to GaAs MBE technology.

Transition metals can be deposited under vacuum on, and subsequently re-

*Actually, in Si MBE the beam fluxes are *atomic* rather than *molecular*, but the terminology "molecular beam epitaxy" persists.

acted with, Si, to form epitaxial silicides, by three different techniques: solid-phase epitaxy (SPE), reactive deposition epitaxy (RDE), and MBE. SPE, the simplest method, involves the deposition of a metal onto a room-temperature substrate, which is subsequently annealed, causing a silicide-forming, solid-phase reaction to take place. A silicide is formed, whose stoichiometry depends on the annealing conditions, and whose thickness, structure, and surface morphology depend on the initial thickness of the metal film. RDE differs in that the metal deposition takes place on a *heated* substrate. Thus, the silicide is formed as the metal film is being deposited. In MBE growth, the metal is stoichiometrically *codeposited* with Si onto a heated substrate, whose temperature is sufficient to provide good lateral mobility of the deposited species, but, also, low enough to limit diffusion perpendicular to the substrate surface. Of the three techniques, MBE offers structurally superior epitaxial films with the best abruptness, and the final silicide thickness, structure, and morphology (though interdependent) are most directly controlled.

4.1.1 Objectives

The purpose of this work is to develop a system capable of high-quality MBE growth of epitaxial silicide (and Si) films on Si. To accomplish this objective, a number of requirements must be met. First, we shall require two *e*-beam sources capable of concurrent, rate-controlled evaporation. It is also necessary to be able to heat the Si substrate *in situ*, with good uniformity and temperature control. Furthermore, as we shall see in Section 4.2, a UHV environment is required for high quality epitaxy. This UHV requirement seriously constrains our choice of materials, necessitates special processing of vacuum surfaces, and complicates the system design. Moreover, after exposure to atmosphere, even a properly built UHV chamber can achieve base pressures in the UHV regime *only* after

an extended period of bakeout (from one to several days). Consequently, to achieve reasonable throughput, a scheme is needed for loading and unloading samples *under vacuum*, so that the problem of exposing the growth chamber to atmosphere is avoided. Finally, more important than the base pressure of the system is the background pressure *during growth*. Hence, adequate provisions to handle the additional gas load due to outgassing from hot surfaces must also be made. One can quicky appreciate that a workable design must take a lot of things into account in a consistent way.

4.1.2 Results of this work

Our main accomplishment is the successful completion of a Si/silicide MBE growth system. Many of the components of this system, including the substrate heater, sample holder, load-lock mechanism, LN_2 shrouds, and water-cooled shutters, represent original conceptual or design work, and were fabricated at Caltech. Our design is unique in that we specifically attempt to address specific problems which are currently active issues in Si MBE: unintentional B-doping of the epitaxial Si layer, Si flaking, and substrate temperature control and uniformity. To verify the operational capabilities of our system, *preliminary* but successful attempts at epitaxial growth of $NiSi_2$ films have been made. While these first $NiSi_2$ films do not appear to be of high quality, as judged from observations of surface morphology and electrical characterizations, our attempts clearly demonstrate the ability of the system to reliably perform closed-loop, rate-controlled evaporations onto a heated substrate. To maximize system accessibility, all our initial growth attempts were performed in a minimal,* unbaked system configuration. Reasonable pressures during growth (10^{-8} Torr-range), comparable to the base

*A viton gasket was used on the Wheeler flange, and a Viton-sealed, quick-access door was used in place of the load-lock. With the use of these viton seals, the system was *not*, of course, in its UHV configuration.

pressure of the *unbaked* system, were achieved routinely, with the aid of specially designed, LN_2-filled shrouds and cryopanels. Finally, we mention that, with a proper bakeout, the growth chamber, loaded with most of its major system components (excluding only the shutters and the main cryopanel), has been able to achieve a base pressure of $\approx 5 \times 10^{-11}$ Torr.

4.1.3 Outline of Chapter

The rest of Chapter 4 is organized in the following way. Section 4.2 begins with a discussion of the stringent requirements of high quality epitaxial growth in Si MBE. These requirements strongly effect our system design. Next, the actual design of our system is discussed in Section 4.3. A focus on the particular problems of unintentional substrate doping, flaking, and substrate temperature uniformity, which we attempted to address in our design, is given in Section 4.4. In Section 4.5, we discuss our choice and preparation of materials for UHV, and in particular, for the present application. Results from our preliminary silicide growth attempts are presented in Section 4.6, and the main system-related problems observed during these initial attempts, along with suggestions to alleviate them, and additional suggestions for further improvement of the system, are given in Section 4.7. Finally, the work presented in this chapter is summarized in Section 4.8.

4.2 Requirements for High Quality Epitaxy

4.2.1 Growth Chamber, Substrate, and Sources

High quality epitaxial growth on Si imposes a number of cleanliness requirements on the growth chamber and its associated apparatus, the substrate itself, and the evaporation source materials. The main problem comes from sensitivity

to hydrocarbon contamination at the substrate surface during growth. In particular, Si forms strong chemical bonds with carbon and oxygen.[1] These contaminants give rise to the formation of stacking faults as their strong bonding tends to nucleate centers of three-dimensional, island growth.[1] The deliberate introduction of hydrocarbon contaminants has also been shown to affect the structural and electrical characteristics of silicide/Si interfaces.[7] Let us now consider some of the implications of this in more detail.

Growth Chamber

The growth chamber must be as hydrocarbon-free as possible. In practice, this means a bakeable, all-metal-sealed UHV chamber, equipped with oil-free pumps, for the purpose of both rough and high vacuum pumping. Typically, the chamber should have a base pressure at least in the low 10^{-10} Torr-range, and a pressure during Si growth in the range of 10^{-9} Torr. The requirements in the idle state are met with conventional UHV technology. For the purpose of roughing, sorption pumps are preferred because of their extremely high pumping speeds, and completely oil-less operation. In the high and ultrahigh vacuum regimes, some combination of ion pumps, titanium sublimation pumps (TSP), and closed-cycle-He cryogenic pumps (cryopumps) are the most common. The use of turbomolecular pumps (turbopumps) is occasionally found, but is generally discouraged because these pumps are more complicated and less fail-safe, because of the increased liklihood of oil contamination in normal operation, and because of the infrequent but *catastrophic* consequences of pump failure. The use of turbopumps should be limited to load-lock chambers, where their main advantage (high throughput) can be exploited, and where the consequences of a major failure are less serious.

Now, let us briefly digress, and discuss the nature of the gas load in ultra-

high vacuum. In a *leak-free* vacuum chamber in the high-vacuum regime, *i.e.*, free from *actual* and *virtual* leaks, the gas pressure in the chamber is entirely due to the desorption of gas molecules from the chamber internal surface area (outgassing). This is because any "volume gas" is removed very quicky after the high vacuum pumps are switched on, *e.g.*, the volume-gas pressure decreases exponentially, with a time constant of 0.5 sec, for a 200 liter/s pump, pumping a 100 liter chamber.[2] On the other hand, adsorbed gas molecules spend most of their time on the *internal surfaces*, and can be removed by the pumps *only* when they desorb. Consequently, the system pressure is dominated by outgassing from surfaces, and the pumpdown time dependence is limited by the removal rate of desorbed species from the walls and internal surfaces of the chamber. Empirically, a normal system pumpdown exhibits a pressure dependence *inversely* proportional to pumpdown time. Hence, a typical, unbaked UHV system with a pressure of 10^{-5} Torr one minute after the high vacuum pumps are switched on, would (without a bakeout) take over two months to reach a pressure of 10^{-10} Torr! The other point we wish to emphasize is that desorption is a *thermally activated* process with a strong temperature dependence, typically increasing an order of magnitude for a 20 °C temperature increase, near room temperature.[2] Consequently, the pressure within a properly designed UHV chamber is very strongly a function of the temperature of its internal surfaces.

Let us now continue with our original discussion. While conventional UHV chamber and pump technology are adequate to meet the requirements of Si MBE when the growth chamber is idle, some special provisions are needed to handle the additional gas load during growth. This additional gas load comes about because of the increased rate of desportion from heated surfaces. In a Si MBE growth chamber, the hot locations are situated near the (heated) substrate, and near the molten, evaporated source materials (including dopants if present). Un-

fortunately, the source and substrate areas are precisely the areas in which lower pressure is needed, particularly with respect to outgassing of hydrocarbons. The solution to this problem is the use of LN_2-filled shrouds to surround the hot locations in the system. These LN_2 shrouds serve to functions: first, they act as a cryopumping surface, and secondly, they act as a heat shield to prevent radiative heat transfer to surrounding parts and chamber walls. In our own system, we have found that the use of LN_2 shrouds significantly reduces the partial pressure of hydrocarbons during growth.

Substrate

The cleanliness requirement also applies to the sources and substrate as well as to the growth chamber. A schematic diagram of a Si MBE growth chamber is shown in Fig. 4.1. In the figure we see the *e*-gun sources and the heated substrate, both shielded by their respective LN_2 shrouds. An essential requirement in Si MBE is the capability of *in situ* cleaning of the Si substrate. As we mentioned already in Section 1.2.3, this is accomplished by one of two approaches. One approach consists of sputter-cleaning the substrate by bombardment with a broad beam of $\sim$ 1 keV, Ar ions, followed by a heat treatment to anneal out any sputter-induced damage. The sputter-cleaning technique is indicated in the figure by the presence of the ion-sputter gun. The second approach, the so-called Shiraki technique (see Section 1.2.3), involves the chemical removal of carbon and oxygen with the subsequent overgrowth of a temporary, protective, thin oxide skin which is easily volatilized under UHV conditions. The point to note is that some means of producing an atomically clean Si surface *in situ* is a prerequisite to epitaxial growth on Si.

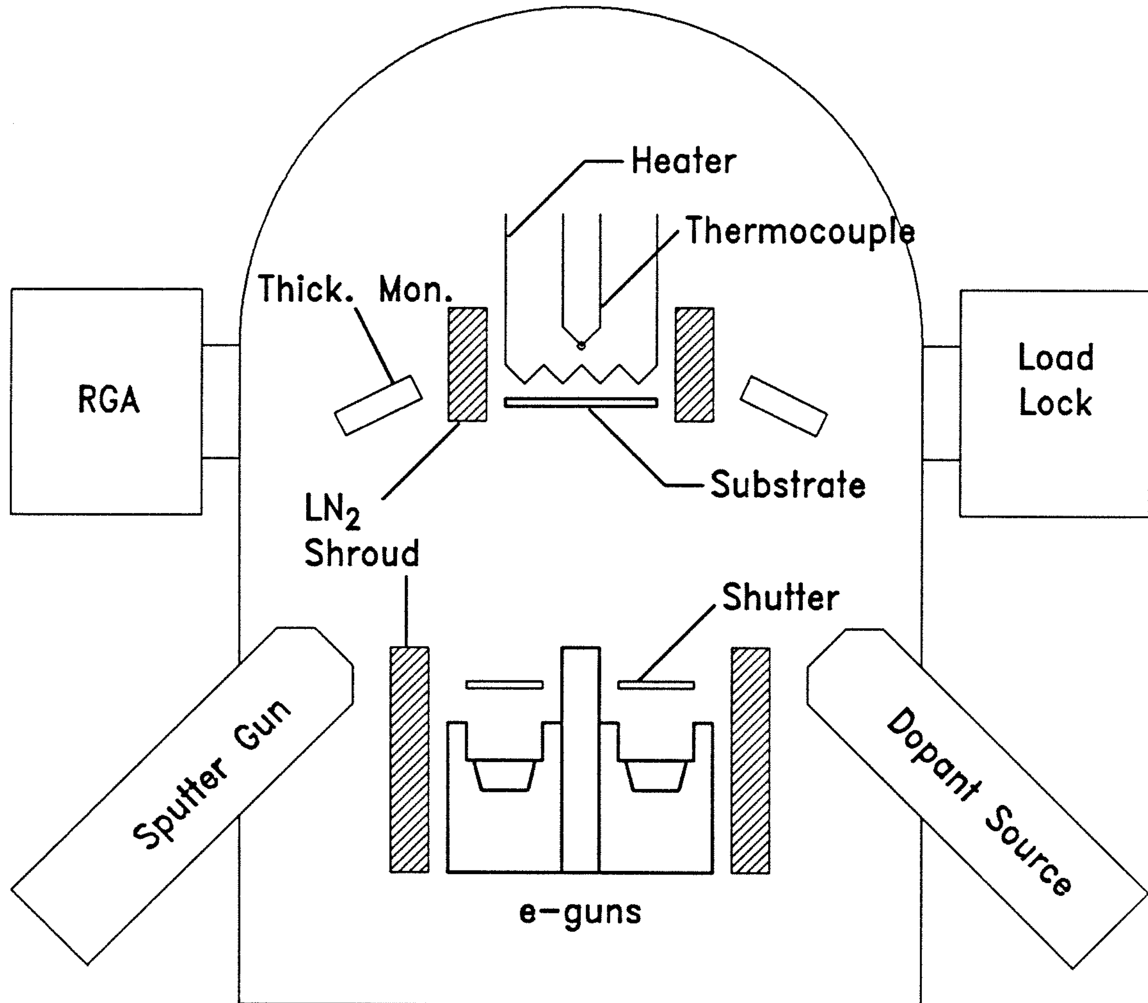


Figure 4.1: Schematic of a Si MBE system. Epitaxial growth takes place from *e*-beam evaporation sources onto a heated substrate in a UHV chamber. Hot regions are shrouded with LN_2-filled panels. Provisions for doping, sensing of substrate temperature and evaporation rate, characterization of the background gas, and *in situ* cleaning and loading of samples are also indicated.

Sources

As far as sources are concerned, a number of types of sources have been tried for vacuum deposition of Si, including resistive heating of Si, crucible heating of Si, sputtering, as well as *e*-beam evaporation.[1] However, only the *e*-beam evaporation technique offers the simultaneous advantages of the production of a "clean" molecular beam, along with high beam intensities, and good control over evaporation rate.[1] (By a clean molecular beam, we mean a flux consisting almost entirely of Si.) To achieve a Si flux of sufficient purity from *e*-gun sources, it is necessary to use specially shaped, auxiliary pieces of high purity Si, positioned around and adjacent to the Si source charge, due to the flood of stray and secondary electrons from the *e*-gun.[1] When this precaution is not taken, impurities of Cu and stainless steel from the body of the *e*-gun have been incorporated into epitaxial Si films, and have been correlated with the occurrence of structural defects.[1] Among the primary disadvantages of the use of *e*-guns as sources include filament burnout, a problem which is aggravated by Si flaking within the growth chamber, contamination due to secondary electron production, and lastly, the undirected, isotropic flux which is produced by *e*-beam evaporation. Nevertheless, at present, there is no satisfactory alternative to the use of *e*-guns as sources in Si MBE.

4.2.2 Growth Conditions

Control Electronics

The objectives in MBE are to fabricate epitaxial films of high structural quality, and controlled composition. These objectives necessitate *stable growth conditions*. Specifically, it is necessary to maintain precise control over such growth parameters as substrate temperature and deposition rate throughout the

growth process. This need is met with the use *closed-loop control* of the substrate temperature, and the deposition rates. A block diagram of the instrumentation and control electronics for our Si/silicide system is shown in Fig. 4.2. In our system, there are three separate control loops: one for the substrate heater, and one for each of the two *e*-guns. For example, the evaporation rate from an *e*-gun is monitored by a quartz-crystal thickness monitor, as indicated in Fig. 4.1. The actual rate is compared to the desired rate setpoint by the electronics in a controller (Inficon XTC), and an analog *error* signal is fed back to the *e*-gun power supply, thereby modifying the evaporation rate until it is in nominal agreement with the setpoint value. The control loop is said to be "closed" because the information about the output (evaporation rate) is fed back to the input (*e*-gun power) by means of a feedback function (the thickness monitor and its associated control electronics). In the case of our system, the characteristics of the feedback are a function of two parameters: "gain" and "overshoot." Ideally, these parameters are to be adjusted so that, for a step change in the rate setpoint, the response in actual rate is *critically damped*. Conceptually, the closed-loop control of the substrate temperature is identical, the only differences being in the sensor (a thermocouple) and control electronics (Eurotherm, Model 984 PID Controller). All three control loops are depicted schematically in Fig. 4.2.

In Situ Diagnostics

In addition to control electronics, it is important to be able to characterize both the UHV ambient, and the condition of the substrate or epitaxial film. A residual gas analyzer (RGA), depicted schematically in Fig. 4.1, is a commonly used instrument for the analysis of gas composition in a UHV chamber. This type of analysis provides a breakdown according to molecular mass* of the gaseous

*Actually, the RGA resolves ionized species as a function of *charge-to-mass ratio*. However, the largest contributions from a given gas species comes from its singly-ionized components.

Process Monitoring & Control

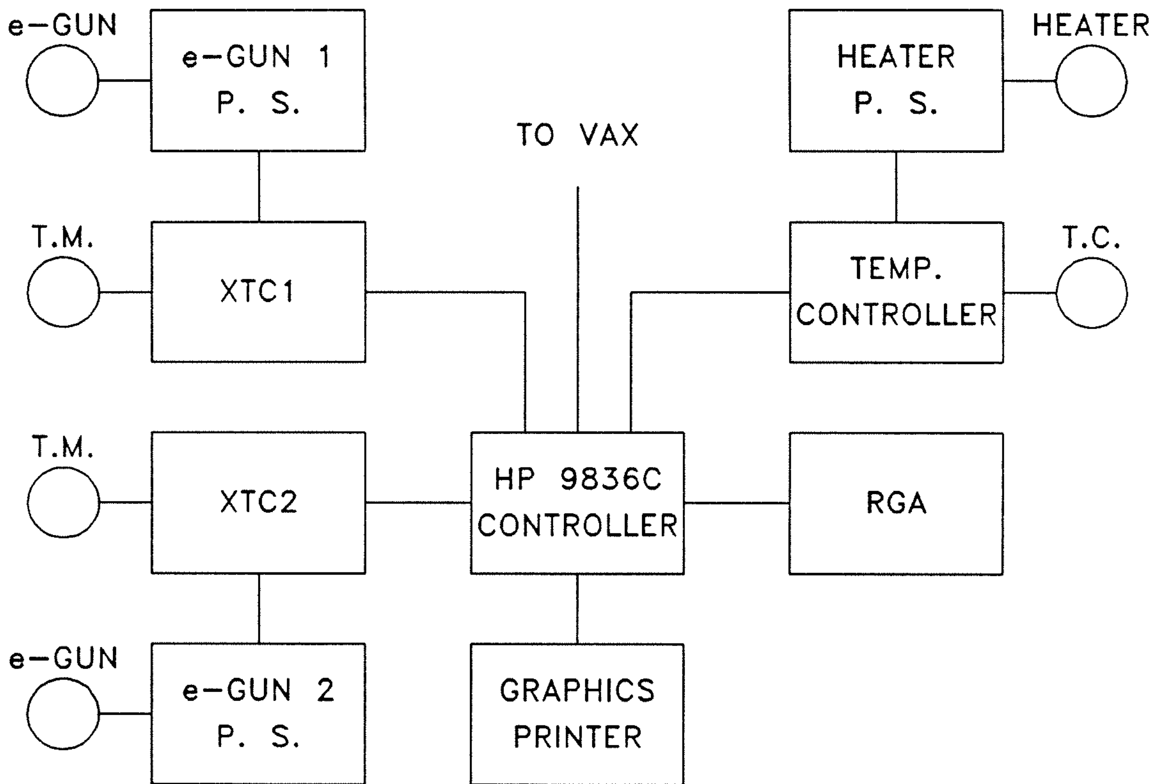


Figure 4.2: Instrumentation and control electronics for our Si/silicide system. Closed-loop control is provided for each of the two *e*-beam evaporators, and for the substrate temperature. Overall process monitoring and control is effected with the use of a microcomputer (HP 9836C).

molecular species in the chamber. Hence, a partial pressure decomposition, permitting the determination of identity and amount of specific hydrocarbons, for example, or of helium (for leak detection), are possible. In a properly sealed and baked UHV chamber, the dominant contribution to the total pressure in the UHV regime should be due to *hydrogen*. (In an unbaked system, *water vapor*, desorbing from the interior surfaces, is the dominant species.) Our system makes use of an RGA (UTI Model 100C) with a sensitivity to partial pressures as low as 10^{-14} Torr. Other useful *in situ* characterizations, commonly found, but not present in our system, are Auger electron spectroscopy (AES), and either low-energy electron diffraction (LEED) or reflection high-energy electron diffraction (RHEED). AES provides an analysis of the chemical composition near the substrate surface. This is particularly helpful in establishing the efficacy of the cleaning technique in the removal of carbon and oxygen from the surface. LEED and RHEED provide surface structural characterization of the substrate, or, alternatively, of the epitaxial film.

The instrumentation and control electronics for our system are summarized in the functional block diagram in Fig. 4.2. A microcomputer (Hewlitt-Packard Model 9836C) with a medium-resolution, color graphics display is used to control and monitor the overall process. The evaporation sequence (with closed-loop rate control) for each *e*-gun is controlled directly by a dedicated XTC. The XTC's are capable of being downloaded with program instructions from the main controller (microcomputer). Similarly, data and parameters can be read back to the microcomputer for processing and storage. Closed-loop control of the substrate temperature is effected with the aid of a temperature controller. The temperature setpoint can be remotely set from the microcomputer by an analog voltage, with the aid of a D/A converter. Similarly, analog voltage outputs from the temperature controller, RGA, and ion pump power supply (not shown) are read

with the use of a digital voltmeter (Keithley Model 195A) and scanner (Keithley Model 705), both of which are directly interfaced to the microcomputer. Immediate printout of text and screen graphics is possible on a local printer. Finally, data transfer to a VAX 11/785 computer is possible via a serial communications link.

4.3 Design Considerations

In this section, we shall concentrate our attention on the actual design of our Si/silicide MBE system. We shall begin with a discussion of the overall system configuration, and then discuss in further detail such individual items as the substrate heater, evaporation sources, shutter assembly, load-lock, and LN_2 shrouds.

4.3.1 Overall Configuration

Our Si/silicide MBE system configuration consists of a growth chamber, and a much smaller load-lock chamber, separated by a viton-sealed gate valve. The system is pictured in Fig. 4.3. The growth chamber is adapted from a standard (Varian Model FC-12E) UHV system, originally designed for surface analysis. The growth chamber consists of a 12-in. diameter, stainless steel bell jar, joined to a base chamber of the same diameter by a Wheeler flange. High vacuum pumping of the growth chamber is accomplished by means of five 40 liter/s ion pumps, a TSP, and a LN_2 cryopanel, all housed in the base chamber. An internal poppet valve, having a removable viton seal, isolates the pump section (below) from the rest of the growth chamber. The load-lock consists of a 4-in. O.D., stainless steel mitred elbow, to which three 2.75-in. Conflat ports have been added. One of these ports is used for rough pumping, another is connected to a 20 liter/s ion

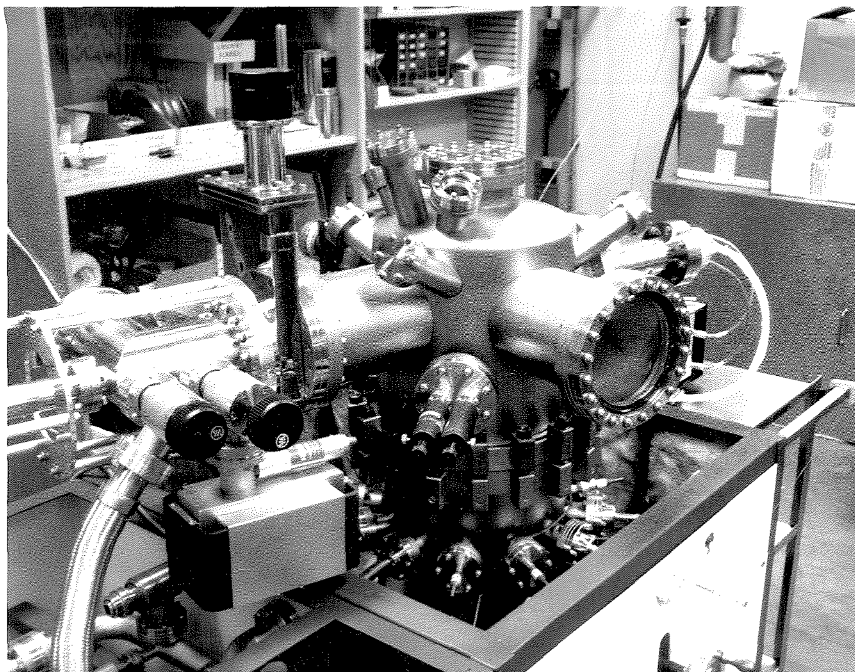


Figure 4.3: Photograph of our Si/silicide MBE growth system. The small chamber to the left of the gate valve is separately pumped (ion pump shown in foreground), and permits sample loading into the main (growth) chamber under vacuum.

pump (in the foreground of Fig. 4.3), and the third is used for a linear-rotary motion feedthru, by means of which a 2-in. diameter Si wafer can be introduced to the growth chamber under vacuum. Samples are introduced into the load-lock through a viton-sealed quick-access door (facing away from the reader in Fig. 4.3). Both the load-lock and growth chambers are rough-pumped from a single manifold connected to a pair of sorption pumps. The gate valve, pictured in Fig. 4.3, is a modified, 4-in., viton-sealed valve body (Huntington Model GVA-400), which has been fitted on one side with an 8-in. flange, and on the opposite side with a 6-in. flange, to minimize the extention of the loading mechanism. Also visible in Fig. 4.3 on the port opposite the load lock is the RGA probe, and, on the 4.5-in. flange in the foreground, a pair of linear motion actuators for the source shutters. Beneath the Wheeler flange, a ring of 2.75-in. Conflat flanges is shown, on which are mounted the electrical and liquid feedthrus for the e-beam evaporators, and their associated cryoshroud and shutter assembly. Maximum bakeout temperatures for the growth chamber are 250 °C with the gate valve removed, 200 °C with the gate valve in place *and cracked open*. For the load-lock chamber, the maximum bakeout temperature is 150 °C, limited by the quick-access door viton seal, and the viton-sealed, right-angle valves shown in the foreground of Fig. 4.3.

A cross-sectional view of the growth chamber is shown in Fig. 4.4. (The view is approximate because, for clarity, certain details have been omitted, and others have been rotated about the main vertical center line from the actual configuration.) The substrate heater, shown in the "growth position," is mounted on a horizontal shaft (not shown) which is capable of rotation, and also, of x-, y-, and z-axis linear motion. The substrate heater can be rotated into different positions: facing the e-guns (for growth), facing the load-lock (for sample loading/unloading), or facing presently unused flanges (for such possible uses as

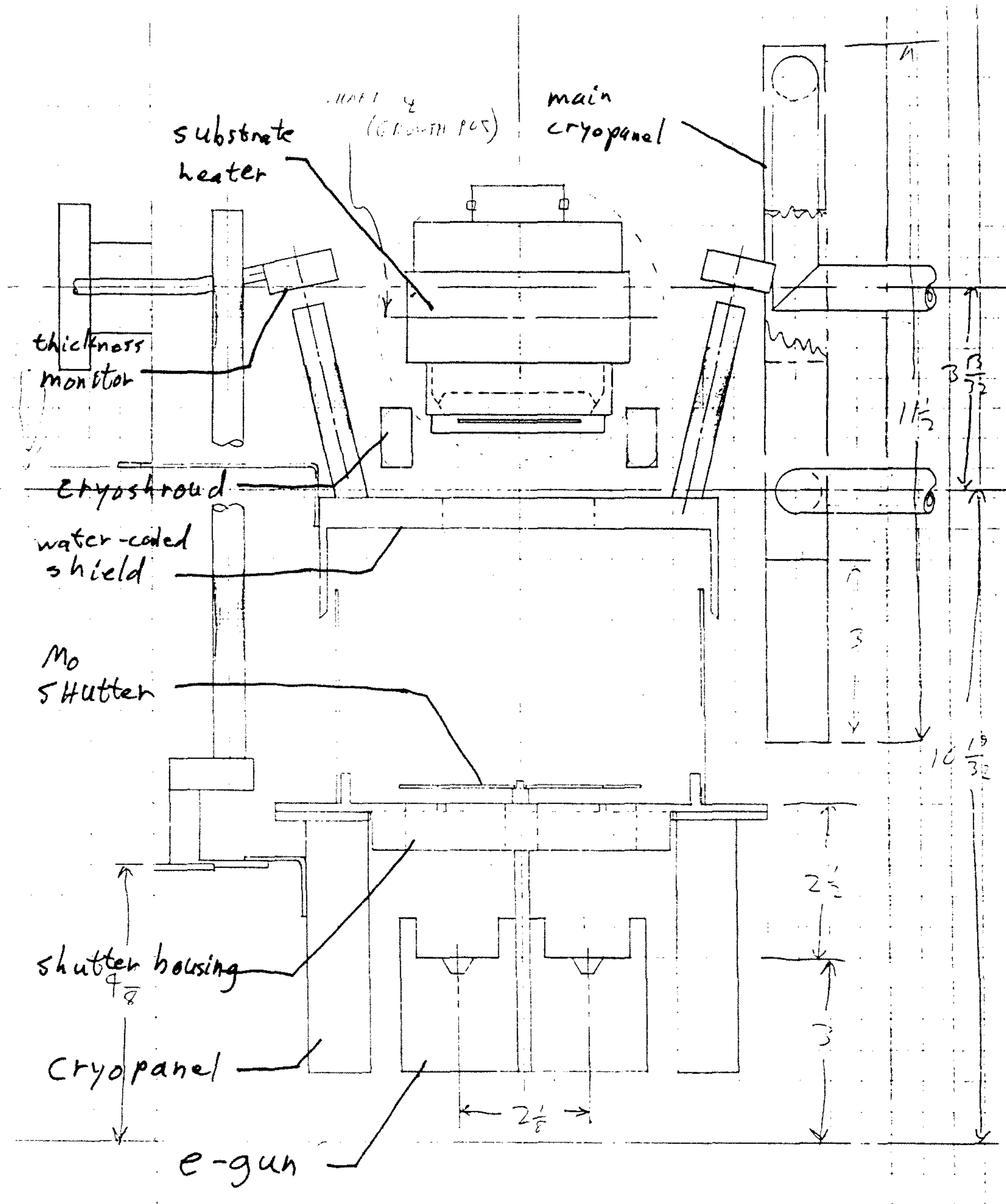


Figure 4.4: Interior configuration of our Si/silicide MBE growth chamber. The major assemblies are indicated. The substrate heater, sample holder, cryoshrouds, shutter assembly, and loading mechanism (not shown) were custom-designed and fabricated at Caltech.

sputter cleaning or surface analysis). A small mirror, mounted at the top of the substrate heater assembly, is used to view the *e*-beam sources during pre-melting or periodic inspection of the source charges. Two 3 kW, 1 cm^3-capacity *e*-beam evaporators (Thermionics) are positioned at a vertical distance of 23 cm below the substrate. With this configuration, approximately 1 μm of material can be safely deposited at the substrate from a fully loaded hearth. Surrounding the *e*-beam evaporators and the substrate heater are LN_2 shrouds, also seen in Fig. 4.4. A water-coolable shutter housing, containing two, three-position shutters, is mounted directly on top of the *e*-gun cryoshroud.* Deposited film thickness is monitored by two quartz thickness monitors. Each thickness monitor is directed at an *e*-gun hearth through a tube and a pair of collimating apertures which allow concurrent, closed-loop operation of both sources without cross-interference. Finally, a third, large LN_2 cryopanel is included in the growth chamber to assist in pumping the excess gas load during deposition. Now, let us consider the design aspects of some of the assemblies indicated in Fig. 4.4 in greater detail.

4.3.2 Substrate Heater

The substrate heater assembly is the most ambitious of the parts which were designed for this work. Here, our objectives were to design a heater capable of heating a standard 2-in. diameter Si wafer, as uniformly as possible, to a maximum temperature of 1300 °C, to be able to monitor and control the substrate temperature in a closed-loop fashion, to come up with a reliable and repeatable scheme for sample loading under vacuum, and to accomplish all these objectives in a manner consistent with UHV requirements. We have been successful in

*The shutter housing must *not*, of course, be water-cooled while LN_2 is being used to cool the cryoshroud.

achieving most of these objectives—the main problem being the monitoring and control of substrate temperature. The approach we have taken is *direct radiative heating* of the substrate. This differs from the "hot heater" approach most commonly found in GaAs MBE, in which the substrate is mounted with In onto a radiatively heated block of Mo. The hot heater approach is inappropriate in Si MBE because of the greater temperatures, and the particular sensitivity of Si epitaxy to outgassed or outdiffused contaminants from hot parts.[1] Accordingly, all commerically available Si MBE systems employ the direct radiative heating approach. One significant difference between our heater design and that of commercial heaters is the *absence* of boron-nitride ceramics in our heater. As we shall discuss in Section 4.4 below, this material has been suspected of giving rise to an unintended *p*-type doping in homoepitaxial Si films.

Our substrate heater assembly is shown diagramatically in Figs. 4.5 and 4.6. The first of these figures shows an exploded, cross-sectional view of the heater and carrier subassemblies, and the second shows the subassemblies connected (as they would be during growth). Before focusing on these subassemblies individually, let us review the overall concept. First, we want to heat the wafer *uniformly*. This means that we need the wafer to be *thermally isolated*, so as to avoid heat conduction, and hence, a *substantial* temperature gradient, along the *very thin* wafer. Also, this implies that we need the heater to be irradiated by a *planar* heat source, whose extent is at least as great as the wafer surface. Secondly, we want to heat the wafer *efficiently*. This means that the energy from our radiating element must, in the steady state, go out through the wafer, rather than into the body of the heater. Thirdly, we want a strain-free, reproducible, and reliable sample mounting scheme which will work *in vacuo*. This requires that the mechanism hold the wafer loosely yet precisely, and that the mechanism be rugged enough to withstand mechanical shocks and severe thermal stresses.

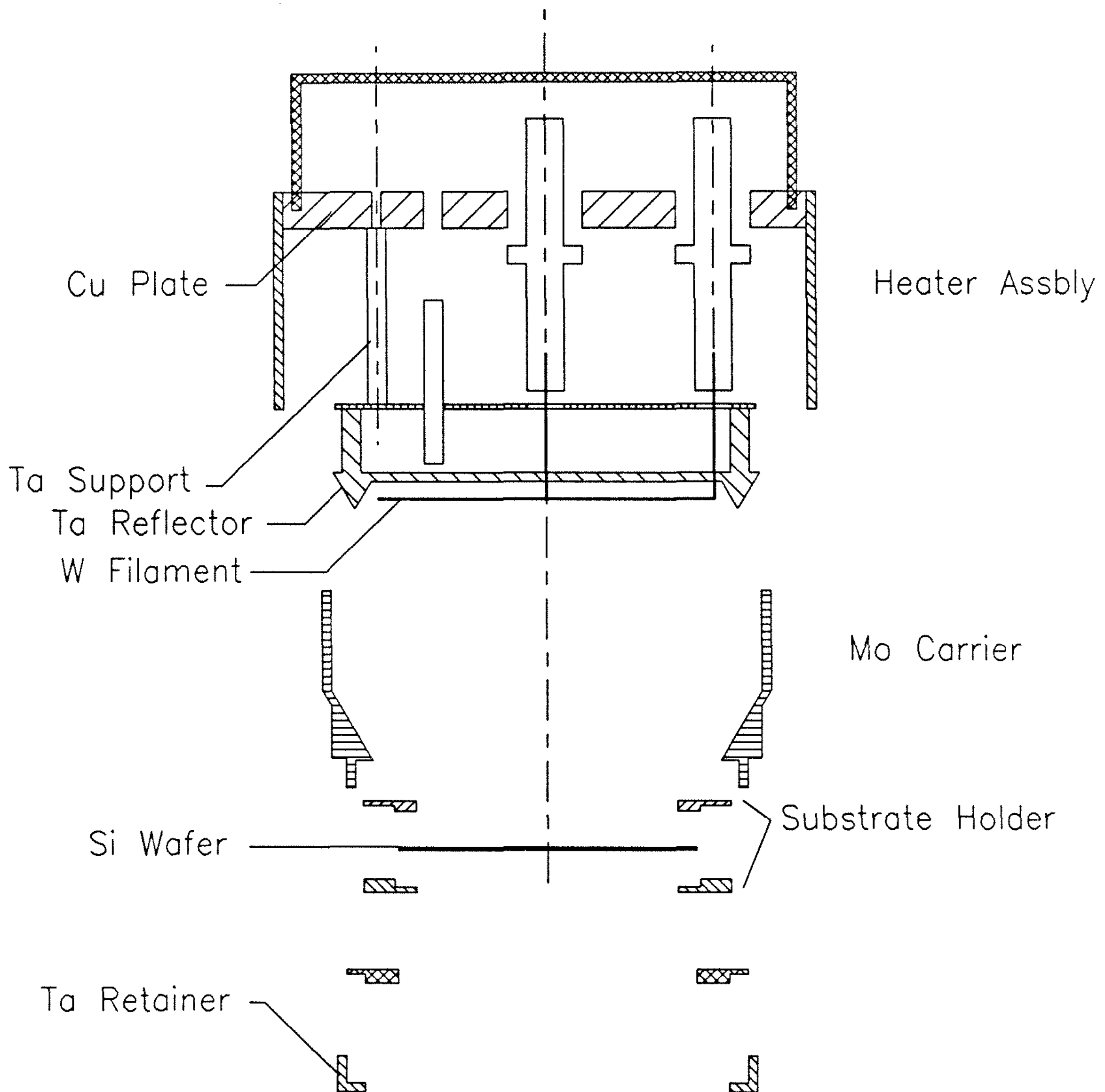


Figure 4.5: Si MBE substrate heater and sample carrier assemblies. A 2-in. diameter Si wafer is loosely held between two Ta rings. The ring-wafer assembly fits into the Mo carrier, which in turn mounts onto the substrate heater. The Si wafer is heated radiatively by a W filament, backed by a polished, Ta reflector assembly. The hot, refractory metal heater parts are stood-off from a cooled, Cu housing by means of three long, thin, Ta rods.

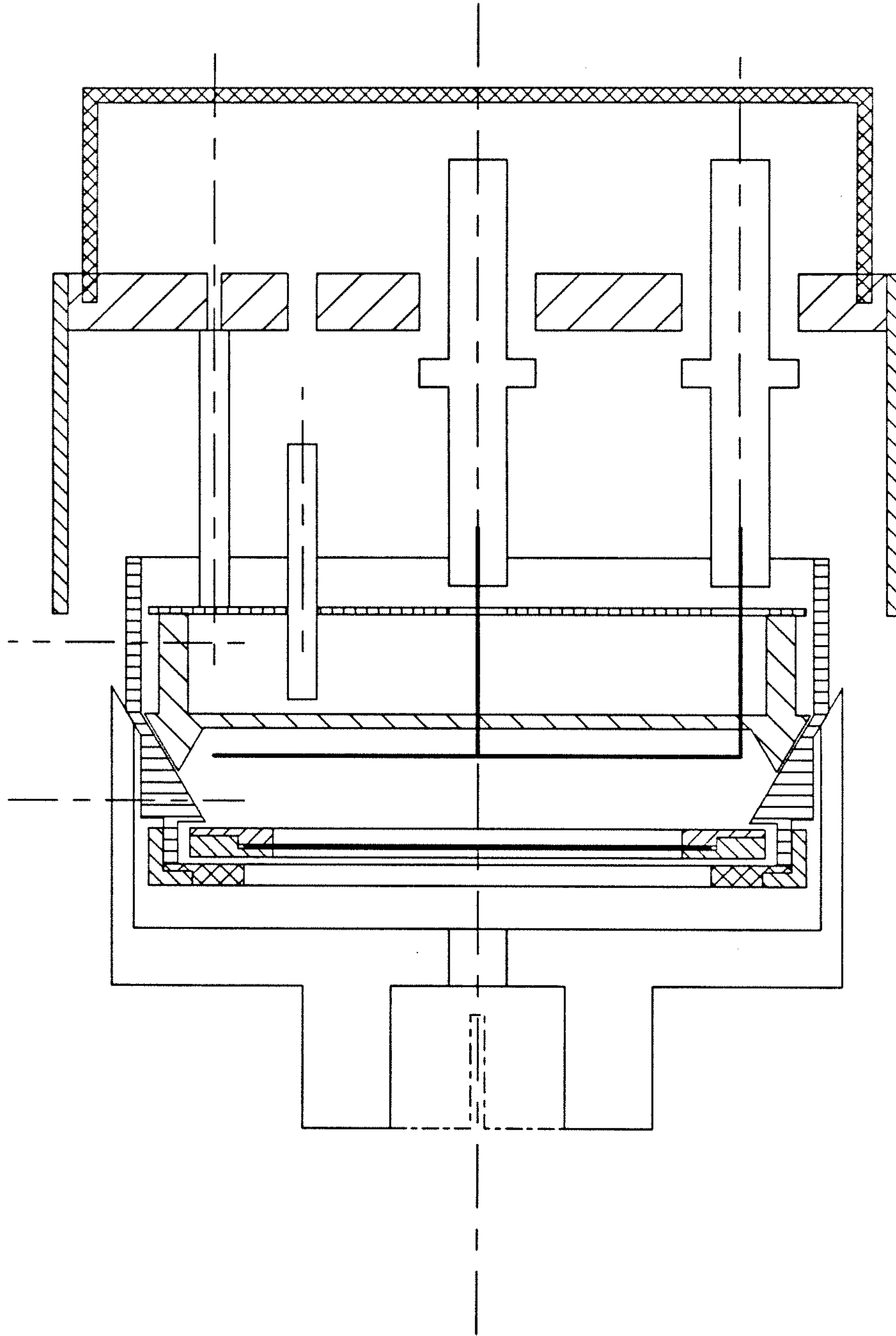


Figure 4.6: Sample carrier, docked with the substrate heater. The Si wafer, Ta reflector, and Mo carrier form a flat cavity which surrounds the radiative heater element. Also indicated is the transfer nosepiece, used for sample loading/unloading under vacuum.

Finally, we require a means to precisely monitor the substrate temperature—a notoriously difficult problem for radiatively heated objects.

The design shown in Figs. 4.5 and 4.6 accomplishes most of our objectives very nicely. All hot parts in the heater have been manufactured out of commercial, high-purity refractory metals (Mo Ta and W). Referring first to Fig. 4.5, the heater subassembly consists of a main, Cu housing, and a Ta reflector assembly. The Ta reflector, along with the substrate and Mo carrier assembly, form a flat cavity, which completely surrounds a Joule-heated tungsten-filament, as shown in Fig. 4.6. The Ta reflector is polished to a mirror-finish to minimize its emissivity relative to the Si wafer. The Ta reflector assembly is mounted to the Cu housing by a circuitous routing of three Ta struts, *e*-beam-welded to the back of the reflector, a Ta plate, and, three thin, Ta rods. These Ta parts are fastened together with six Mo screws, and the Ta rods are attached to the Cu housing with three stainless steel screws.* The purpose of this circuitous, high-thermal-resistance routing is to minimize conductive heat transfer from the (hot) refractory-metal parts to the (relatively cool) Cu housing. Although this is more typically achieved with the use of ceramic standoffs, we particularly wished to avoid the use of any ceramics in *direct contact* with the hot, refractory metal components. The ceramics which *are* are used are Macor, for electrical insulation of the Ta, filament electrodes from the housing (see Fig. 4.5), and alumina, to electrically insulate the thermocouple (type-K) from the metal components. To provide some measure of cooling, the housing is conductively cooled by means of a Cu braid, which connects the housing to a nearby LN_2 feed line. We chose Cu as the material for the housing because of its excellent thermal conductivity. Recalling that our substrate heater must articulate (see Section 4.3.1 above), the braid provides a simple alternative to a liquid-cooled design, which would have

*All screw threads, joints, and other potential gas entrapment volumes have been relieved, according to good vacuum practice.

necessitated the use of a flexible, and undesirably risky, fluid path such as formed bellows.

A Si wafer is attached to the heater by first mounting it in a cylindrical, Mo carrier, shown in Fig. 4.5. Then, the Mo carrier is bayonet-mounted onto the Ta reflector assembly with the aid of three, W pins. (An actual sample mounting sequence is pictured on page 154 below.) The manner in which the Si wafer is attached to the carrier is seen by examining Figs. 4.5 and 4.6. The wafer is first "sandwiched" between two Ta or Si rings,* which fit together precisely while leaving a gap slightly larger than the thickness and diameter of a commercial, 2-in. diameter Si wafer. In this way, a precise, yet strain-free mounting of the wafer is achieved. Under vacuum, thermal conduction between the rings and the wafer can only take place through the *three points* of physical contact, and is therefore poor. Next, the entire ring assembly is located in a slightly oversized recess at one end of the Mo carrier. A Ta retaining plate, with three, protruding, W pins, on which the ring assembly rests for thermal isolation (see Fig. 4.6), holds the ring assembly in place. The retaining plate itself is held in place against the carrier by a threaded, Ta collar.

The complete heater-carrier assembly is shown in Fig. 4.6. Also indicated in the figure is the transfer nosepiece (part of the load-lock mechanism), which is used to ferry the Mo carrier from the load-lock out to the substrate heater, and then to mount it, under vacuum. Matching tapers on the surfaces of the reflector, carrier, and nosepiece assemblies provide the critical self-alignment of the mechanism during sample loading/unloading. We mention here that the different thermal expansion characteristics between the Ta reflector and Mo carrier

*The purpose of using Si rings is to minimize slip line formation in the wafer when heated. This is because the transient thermal behavior of the rings, owing to their closely similar thermal properties, will more closely match that of the wafer, resulting in improved wafer temperature uniformity.

have also been taken into account in the sizing of these two parts.

The heater filament, consisting of a 100 cm length of 40-mil W wire, wound into a 5.7 cm O.D. flat coil, provides an approximately planar, radiative heat source for even heating of the wafer. The temperature is monitored at present with the use of a thermocouple which protrudes into the filament cavity through an opening in the Ta reflector. The tip of the thermocouple has a small piece of Si tied to it with very thin Ta wire. In princple, feedback from this thermocouple can be used to control the filament power supply in a closed-loop fashion. However, in practice, we find the thermocouple temperature to be lower, and transient response *much* slower, than that of the wafer; hence, this detection scheme does not appear to be useful for closed-loop control of the substrate temperature. (We note, though, that the accurate monitoring of the radiatively heated substrate is one of the important issues in Si MBE, and as yet has no good solution.) Thus, with the Mo carrier twist-locked onto the end of the heater, we have achieved most of our design objectives stated above.

4.3.3 Evaporation Sources and Shutter Assembly

Our system makes use of twin, 3 kW *e*-beam evaporators, mounted together as a unit on a single, 2.75-in., flanged, demountable feedthru. Each *e*-gun makes use of a permanent magnet to achieve *e*-beam deflection. The important consideration which we would like to emphasize here concerns the close placement of two such *e*-guns. Specifically, if the two *e*-guns are to be positioned side-by-side, as in our configuration, it is necessary to have one of the guns *rotated* (with respect to the other) about a vertical axis by 180°, in order that their respective magnetic fields *aid*, rather than *oppose* one another. Further, a "pole piece," *i.e.*, a bar of soft, ferromagnetic material such as mild steel (Ni-plated to prevent corrosion), placed between the guns will effectively isolate them magnetically, so

that the *e*-guns may then be placed directly against either side of the pole piece without mutual interference.

As we saw earlier in Fig. 4.4 on page 143, our *e*-guns are surrounded by a cryoshroud, directly on top of which sits a shutter assembly. A top-view of this shutter assembly (looking down from the substrate position) is indicated in Fig. 4.7. There, we see two, mitten-shaped, Mo shutter paddles, each of which can move linearly along the central slide rail by means of a guided, roller mechanism (not shown). Linear shutter motion is actuated by two direct-drive, linear motion feethrus (referred to earlier in Fig. 4.3). Since these motion feedthrus are mounted on a *bell-jar* flange, whereas the shutter housing itself is fixed to the *base chamber*, a means for readily decoupling the linkage between the two parts, *e.g.*, whenever the bell-jar needs to be lifted, had to be devised. For this purpose, we designed a magnetic linkage, which can be quickly decoupled by simply pulling the motion feethrus out past a point, beyond which, further shutter travel is restricted. Referring once again to Fig. 4.7, we see that there are *two* apertures for each *e*-beam evaporator. As shown in the figure, aside from a fully opened and a fully closed shutter position, there is an intermediate position, in which the evaporation source is not seen by the substrate, but *is* seen by its respective thickness monitor. The purpose of this intermediate position is to allow a steady-state, rate-controlled evaporation to be established *prior* to deposition on the substrate.

4.3.4 Load-Lock

An excellent view of our load-lock chamber has already been given in Fig. 4.3. As we mentioned before, our load-lock essentially consists of a standard UHV elbow, to which some additional ports have been added, to accomodate the pumping and motion transfer apparatus. The purpose of the load-lock is to improve

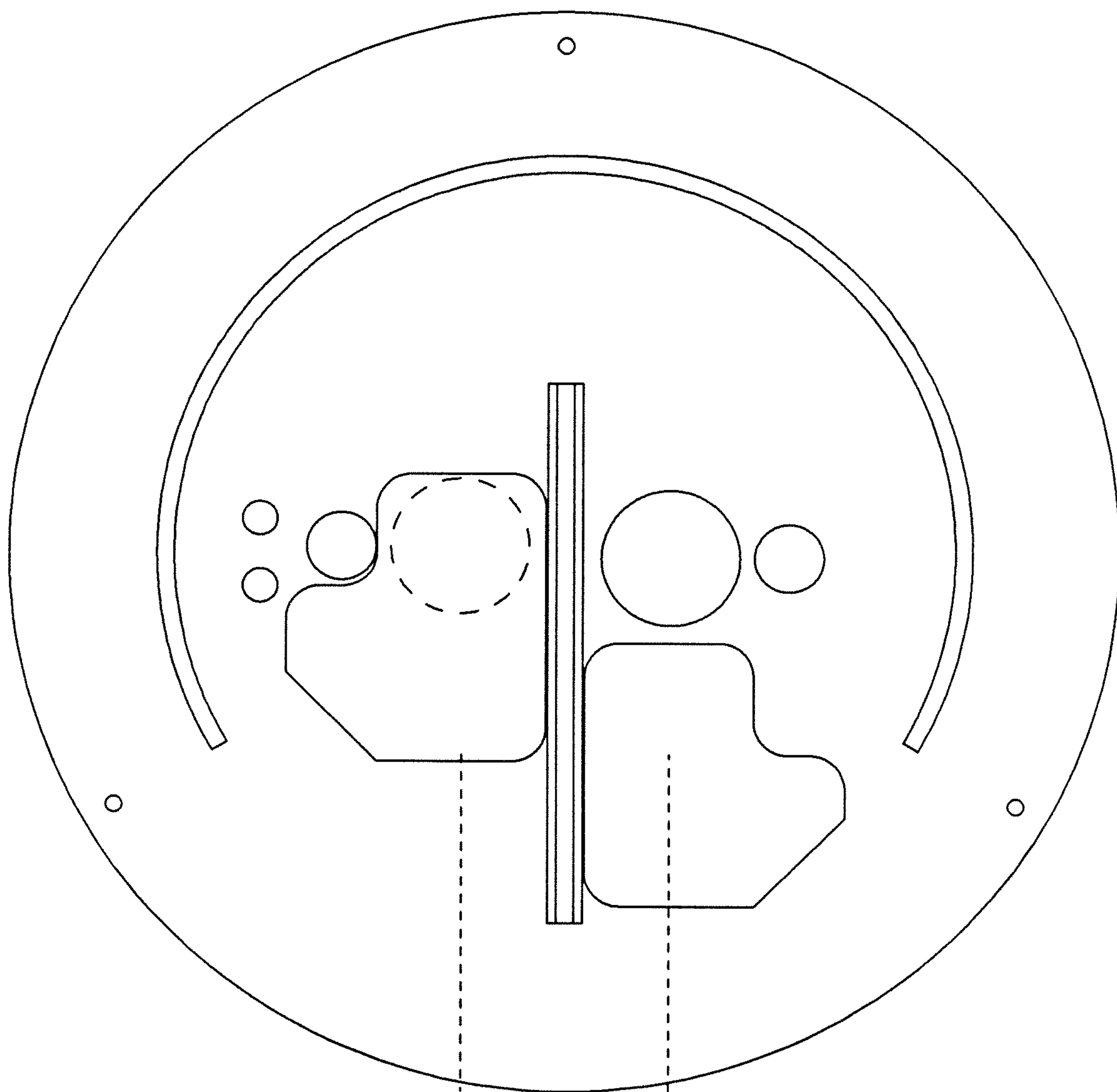


Figure 4.7: Top view of shutter assembly, as seen from the substrate heater. A pair of three-position, Mo shutters permit viewing of the two *e*-beam sources by the substrate heater (large apertures) and thickness monitors (small apertures). The shutter housing can optionally be water-cooled.

system *throughput* by providing a rapid means of sample introduction into the UHV ambient of the growth chamber. In terms of design requirements, this means that we want a chamber which can be quickly cycled ($\sim$ 10–30 min) from atmosphere to the transfer pressure ($\sim 10^{-7}$ Torr). As we discussed in Section 4.2.1, the desportion of *water vapor* from *internal surfaces* produces the pumpdown-limiting gas load in an unbaked, leak-free system. Therefore, to achieve a minimum of pumpdown time, we made our load-lock chamber as small as possible (to achieve the best ratio of *pumping speed* to *surface area*). Further, to minimize the adsorption of water vapor upon exposure to atmosphere, we have made provisions for backfilling the load-lock chamber with filtered, dry N_2. This provision is only effective, of course, if the system has been previously *baked* sufficiently to remove the water vapor from the chamber. In our load-lock, this bake/backfill approach has resulted in approximately an order of magnitude improvement in pumpdown time.

The operation of the loading mechanism is straightforward. A stainless steel nosepiece, having three angled, bayonet slots, which engage with corresponding pins in the Mo sample carrier (described in Section 4.3.2), is mounted onto the end of a shaft. This shaft is part of a magnetically coupled motion feedthru (partly shown in Fig. 4.3), and is capable of linear and rotary motion. To load a wafer in the growth chamber, the load-lock is backfilled and opened, and the Mo carrier, containing a wafer, is manually mounted into this nosepiece. Then, the load-lock chamber is pumped down to a pressure suitable for transfer ($\sim 10^{-7}$ Torr). The gate valve is opened (resulting in a temporary increase in pressure in the growth chamber), and the mechanism is extended through the valve, out to the sample heater. An actual loading sequence is seen in the four photographs shown in Fig. 4.8. In this figure, the upper left photo shows the substrate heater in the "load position," with the nosepiece and carrier approach-

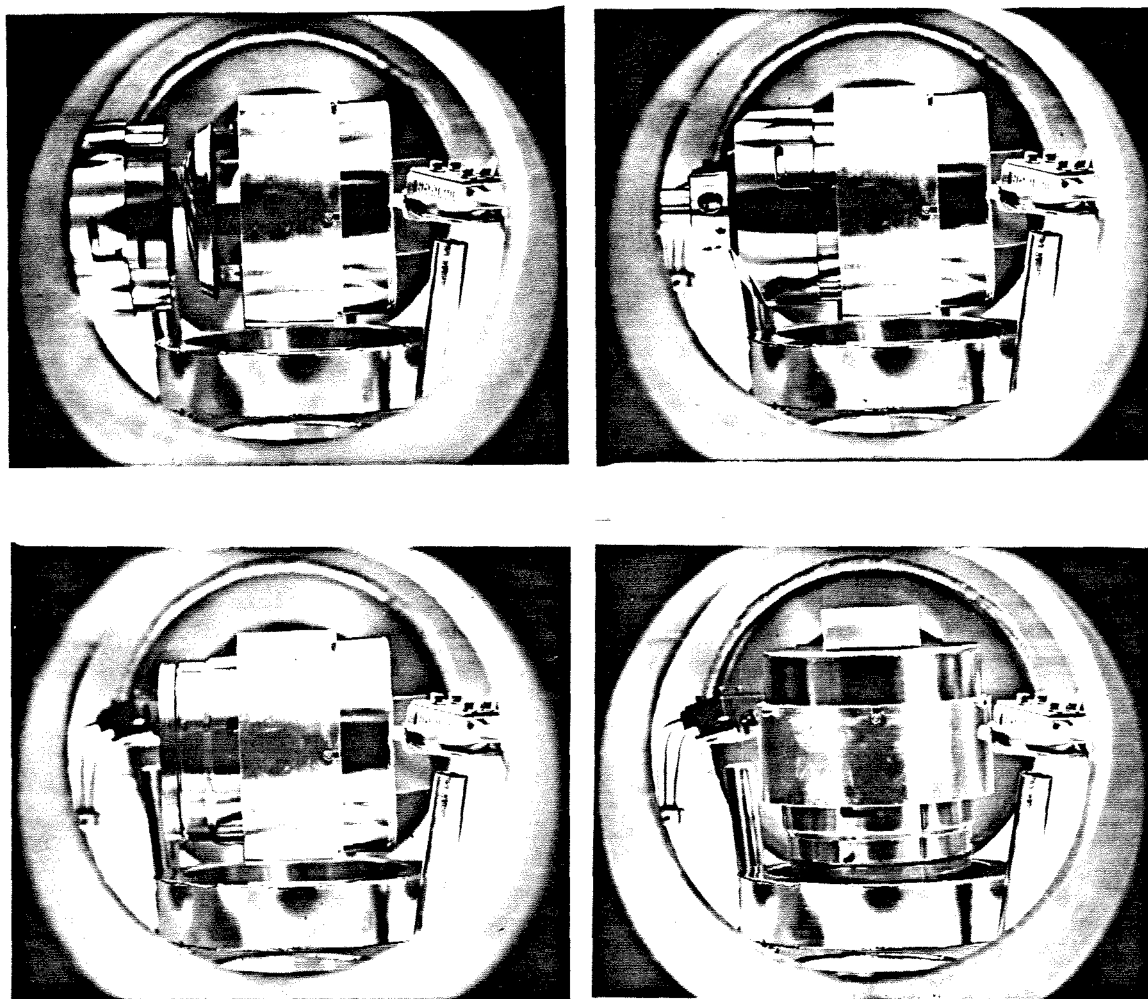


Figure 4.8: Sequence showing the substrate carrier being bayonet-mounted under vacuum onto the heater. From upper left to lower right: (a) Substrate heater in the load position with Mo carrier and transfer nosepiece approaching. (b) Carrier docked with heater. (c) Nosepiece withdrawn. (d) Heater rotated to the growth position.

ing. In the upper right, the carrier-nosepiece combination is docked with the sample heater. Note that the bayonet slots on the nosepiece are "right-handed" while those on the carrier are "left-handed"; hence, a single rotary motion will disengage one set mounting pins while engaging the other set. The lower left photo shows the substrate heater, now loaded with the wafer carrier, and the nosepiece has been withdrawn. Finally, the substrate heater is rotated into the growth position, as shown in the lower right photo. The sample unloading sequence is merely a reversal of the loading sequence. (Also visible in this sequence of photos is the heater cryoshroud, the water-cooled-shield, and the two thickness monitors.)

4.3.5 Cryopanels: LN_2 vs. Water Cooling

Our system design makes use of LN_2-filled cryogenic shrouds or panels in three places in the system: around the e-beam evaporators, around the heated substrate, and along the outer wall of the chamber. The e-beam shroud performs the dual function of cryopump and heat shield. The small cryoshroud around the heater serves to improve vacuum integrity in the immediate vicinity of the heated substrate, but because of the geometries involved, does not provide much heat shielding. For this latter purpose, an additional, water-cooled shield has been included, as seen in Fig. 4.4 above. (In principle, this shield could be LN_2-cooled although it has not been designed to be an efficient cryopanel.) The third, large cryopanel, contoured to fit along the inner wall of the growth chamber, provides a large surface area for the purpose of cryopumping excess gases which are produced during sample growth.

All of the LN_2 shrouds are made of thin, stainless-steel sheet metal (low thermal mass), which is cut, folded, and welded into hollow vessels with a high surface-area–to–volume ratio, to which inlet and outlet tubes are attached. The

processing of these shrouds, prior to their introduction into our UHV chamber, is described in Section 4.5 below. The shrouds are mounted by their plumbing, on either standard or custom LN_2 feedthrus. Because of the possibility of leakage due to thermal expansion and contraction, we tried to avoid the use of tube fittings and connectors with seals internal to the system. However, for the case of the *e*-gun shroud, it was necessary to use VCR fittings internally for both the LN_2 inlet and outlet lines. While these fittings are designed to be UHV-compatible, we found our VCR fittings to leak upon introduction of LN_2, presumably due to differential thermal expansion between gaskets and stainless-steel seal surfaces, when Cu gaskets were used. Therefore, it is important to *avoid* the use of Cu gaskets in VCR fittings when LN_2 is used. Fortunately, we found that the use of silver-plated, stainless steel gaskets adequately solves the leak problem although greater care must be taken when installing and removing these gaskets, so as not to scratch the seal surfaces.

At present, there is some disagreement over whether LN_2 or water is the most appropriate coolant. In particular, the well-known problem of flaking of accumulated evaporants appears to be aggravated by the constant thermal cycling, *i.e.*, thermal contraction and expansion, of cryogenic shrouds. Obviously, water-cooled shrouds have the advantage of being maintainable at a fixed temperature continuously. A further question over the appropriateness of LN_2 cooling concerns the suspicion that hydrocarbons, trapped by adsorption on the cryogenic surface of, say, an *e*-gun shroud, might be *desorbed*, either by *photons* (from the filament), or by *stray electrons*.[1] For this reason, at least one group has decided to use *water* rather than LN_2 to cool their *e*-gun cryoshroud.[3] Therefore, we have left open the possibility of water-cooling in the design of our own *e*-gun cryoshroud and shutter housing, by making the shutter housing *water-coolable*. Since the cryoshroud and shutter housing are in direct thermal contact (see

Fig. 4.4 above), cryogenic cooling of the shroud is sufficient to keep the shutter housing from getting too hot during growth.* However, *water-cooling* of the cryoshroud will *not* provide adequate cooling of the shutter housing. Therefore, we have designed our shutter housing so that it can be water-cooled directly. We warn, though, that in our design it is imperative that water-cooling of the shutter housing *never* take place concurrently with LN_2 cooling of the *e*-gun cryoshroud, *or else the cooling water will freeze*, resulting in a serious system failure.

4.4 Problems Addressed

4.4.1 B-doping of Epitaxial Si

In this section, we shall discuss our attempt to address two problems which are current issues of interest in Si MBE; namely, unintentional *p*-type doping of Si epilayers, and Si flaking from interior surfaces. First, let us describe the *p*-type doping problem. It has been found in certain cases[1] that, in attempting to fabricate nominally undoped epitaxial Si on Si, that there is an initial *p*-type spike in dopant profile, largest at the Si/Si interface, which corresponds to a concentration of $\sim 10^{17}\,\mathrm{cm}^{-3}$.[1] This doping spike, identified by secondary-ion mass spectroscopy (SIMS) analysis as being due to B impurities, disappears beyond a distance of $\sim 1\,\mu\mathrm{m}$ from the interface.[1] At one point, there was some speculation[1] that the doping might be due to B contamination as a result of the heating of boron-nitride ceramics.[1] Boron nitride, and its pyrolytic form (PBN), are a commonly used, UHV-compatible, high-temperature insulator; in particular, PBN is used for physical support of the heating element in commercial MBE substrate heaters (and sometimes as a crucible material in *e*-beam evaporation).

*The shutters themselves may get hot, however, by thermal radiation from the source melt, which is one reason why they have been made of Mo rather than stainless steel.

To remove the possibility of doping from this source, we have designed our substrate heater with a *free-standing* W filament. Filament support is achieved entirely by an all-refractory-metal mounting. The filament ends are inserted into holes at the tips of the two Ta electrodes (refer to Fig. 4.5 on page 146), and are secured with Mo set screws. The Ta electrodes themselves are mounted onto the Cu housing of the substrate heater, with Macor ceramics used for insulation from the housing. Since the ceramics are sufficiently far from the heating element, and are in direct contact with the relatively cool ($< 100\,^{\circ}$C) Cu housing, they should not get very hot during wafer heating. The main disadvantage of the free-standing design is the *sag* of the filament under its own weight when heated, as we will discuss further in Section 4.7.

4.4.2 Si Flaking

Another problem which we address in our design concerns the flaking of evaporated materials, particularly Si, which have collected on the interior surfaces. The flaking problem occurs because of stresses which build up with increasing film thickness, causing a reduction in the effective adhesion of the film.[4] The flaking problem is important in Si MBE for several reasons. First, it can result in particulate contamination of the sources, which in turn may cause morphological defects in an epitaxial film. In addition, flaking can introduce fluctuations in the evaporation rate, and has been known to result in filament burnout in *e*-beam evaporators.[5] Unfortunately for Si MBE systems, the adhesion of Si to stainless steel appears to be particularly poor, with the result that the floor of a Si growth chamber soon becomes covered with a "snow" of Si powder.[6]

In our system, we have taken steps to reduce the flaking problem. It has been found that the roughening of metal surfaces by *sand-blasting* results in greatly improved adhesion (up to $30\,\mu$m) of polycrystalline Si films.[1] Therefore, we have

designed our system so that deposition takes place only on surfaces which have been sand-blasted, and which can be removed for periodic cleaning. Specifically, the inner surface of the *e*-gun cryoshroud, the bottom of the shutter housing, and the bottom of the Mo shutters have all been sand-blasted. In addition, as we have already mentioned above, *thermal cycling*, particularly of smooth, electropolished LN_2 shroud surfaces, tends to promote the flaking of deposited material. Thus, to further minimize flaking problems aggravated by temperature cycling, our substrate cryoshroud has been deliberately shielded from the line-of-sight of the evaporation sources by a *water-cooled* shield, as shown in Fig. 4.4. Finally, this last point about thermal cycling is an additional reason why we have chosen to include water-cooling capability in our shutter housing design.

4.5 Choice and Preparation of Materials for UHV

In this section, we discuss our materials choices for the Si/silicide evaporation system. Also, we explicitly present the procedures that were used in the preparation of materials for use in our UHV chamber. These procedures include *precleaning*, and *chemical processing* such as electropolishing, and other processes such Au-plating and dicroniting. An excellent general discussion of the properties and suitability of materials with respect to UHV applications is given in Ref. 7. We do not wish to repeat that discussion here. Rather, the intent is to focus on the components of our particular system, and to account for our materials choices in terms of the function of these components.

4.5.1 Materials for UHV

Most of our homemade system components have been made out of type 304 stainless steel. The list of these assemblies includes: the cryoshrouds, shields,

and their associated plumbing; the shutter housing and mechanism; and, the load-lock transfer nosepiece. Type 304 stainless steel is the material of choice for the construction of UHV chambers because of its mechanical strength, bakeability to 450 °C, low vapor pressure content (unlike type 303), and particularly, its good machining and excellent welding characteristics (along with commercial availability). The low-carbon grade, type 304L, is in fact particularly popular in UHV technology, because of its superior weldability. The primary disadvantages of type 304(L) stainless steel are *poor* thermal conductivity, and high density (weight). Loosely speaking, type 304 or 304L stainless steel is the "default" UHV material, and other materials are chosen only when a particular requirement makes type 304(L) an unsuitable choice. For example, our cryoshrouds, which are thin-walled, intricately welded assemblies, require a material with excellent welding characteristics, but not excellent thermal conductivity. (Copper, which has the opposite properties, would *not* be appropriate in this case.) On the other hand, two parts in our shutter mechanism require attributes not possessed by type 304(L) stainless steel: we require a *ferromagnetic* material in the magnetically coupled linkage, refered to in Section 4.3.3 above; we also require a *hardenable* material for our roller bearings in the same mechanism. Since type 304 stainless steel is essentially nonmagnetic, and cannot be hardened by heat treatment, we chose (17-4PH) stainless steel, which is both ferromagnetic and heat treatable[8] while still retaining acceptable vacuum properties, for use in both our roller bearings and our magnetic linkage.

In our system, all metallic parts which get *hot* are made of high purity, refractory metals. Our most extensive use of refractory metals, specifically, Ta, Mo, and W, is made in the design of the substrate heater and sample holder assemblies. In addition, our shutter paddles are made of Mo since they are capable of getting fairly hot due to radiative heating from the molten *e*-beam

source materials. Refractory metals are expensive to obtain and comparatively difficult to machine—in fact, W is too brittle to be machined—so that their use is limited to those locations where they were deemed absolutely necessary. There are two reasons for the use of refractory metals instead of stainless steel at hot locations. First, the melting points of Mo, Ta, and W are very high (over 2500 °C) compared with that of stainless steel (1425 °C).[7] Consequently, the vapor pressures of refractories at high temperatures are much lower. Secondly, stainless steel is not a *pure* substance, but rather an alloy of Fe, Ni, and Cr, with a host of added impurities, which can outdiffuse at elevated temperature, and may adversely affect the electrical and structural properties of epitaxial Si films. The hottest metal parts in our system occur near the substrate heater and the *e*-guns: specifically, the tungsten filaments, and any adjacent surfaces which are thermally isolated, and not directly cooled.

In a few of our system parts, OFHC* Cu is used because of its excellent thermal and electrical conductivity. In particular, the electrical connections to the substrate heater and *e*-guns consist of Cu wire. The main housing of our substrate heater is also made of Cu because of the high thermal conductivity, as explained earlier in Section 4.3.2. Cu loses its structural strength at high temperatures, and in general should be avoided in UHV applications at temperatures in excess of 400-500 °C.[7] In normal operation of our sample heater, the Cu housing does not appear to be much getting hotter than 100 °C. If this were not the case, it would be necessary to redesign the heater housing, either by adding additional cooling, or by remaking the parts out of a refractory metal.

Macor and alumina ceramics are used for the purpose of electrical insulation, primarily in and around the cooler parts of the substrate heater. Macor

*Oxygen-free, high conductivity (OFHC) is a special grade of Cu which does not contain any dissolved oxygen. Ordinary Cu is not suitable for UHV service because it will act as a virtual gas source due to oxygen outdiffusion.

is the trade name for a new, *machinable* glass-ceramic with excellent outgassing characteristics (after bakeout), zero-porosity, and excellent dielectric breakdown strength. As a result of its excellent vacuum and electrical properties, combined with machinability, Macor has recently become one of the promising new ceramic materials for use in UHV applications, at temperatures below 1000 °C. We have used Macor insulators for bulkhead mounting of the substrate heater electrodes and thermocouple, in the Cu plate at the top of the substrate heater (Fig. 4.6). We also have a Macor cable guide on the substrate heater support shaft. For higher temperature service, we have used an alumina ceramic for support of the substrate heater thermocouple, near the tip. Finally, our substrate heater connecting wires are jacketed by a flexible, alumina sleeve.

4.5.2 Precleaning

At this point, we shall discuss the preparation of materials for UHV. The preparation of a suitable vacuum surface is essential, not only to allow the system to reach the UHV pressure range, but also, to minimize the release of contaminants during growth. The procedures which we have followed consist of two parts: (1) degreasing, *i.e.*, precleaning with detergents, and then organic solvents, for the purpose of removing oils and lubricants picked up during machining and handling; and (2) chemical treatment, which may include chemical etching or polishing, passivation, and electropolishing. It is important to note that it is never possible by these methods to produce an atomically clean surface, completely free of contamination. What is sought is a procedure by which (1) the microstructure is *smoothed* as much as possible to minimize microscopic surface area (recall our earlier discussion about desorption) and in particular to remove microscopic cracks and imbedded contaminants; and (2) all *loosely bound* chemical species are *removed* from vacuum surfaces (since these will be the ones

which are most troublesome in vacuum).[9]

The first step in preparing a newly manufactured part for UHV involves degreasing. For all of our parts, we found the procedure given in Table 4.1 to be adequate. The procedure begins with a detergent wash in *hot* water. Tap water is acceptable for this initial wash. Also, we found mechanical scrubbing to be important in removing stubborn grease and oils left over from machining. Next, a thorough rinse in deionized (D.I.) water ($10\,\mathrm{M\Omega \cdot cm}$) was performed. Then, the organic solvents, tricloroethylene (TCE), acetone, and methanol were used in sequence to finish degreasing the part. In particular, the initial TCE rinse was repeated with fresh solvent until the TCE bath remained clear (while ultrasonically agitating the part in the solvent). Next, D.I. water is used to remove any solvent residue from the surfaces. Finally, the part is thoroughly air-dried on brown paper in a relatively dust-free area, such as under a laminar flow hood. For storage, cleaned parts were wrapped in fresh Al foil, and then sealed in plastic bags.

4.5.3 Chemical Processing

Following the initial degreasing, a further chemical treatment is applied, depending on the material of which the part is made. In the case of stainless steel, the parts were sent out for commercial *electropolishing*.* For our Mo and W parts, chemical polishing, with the aid of a 20% (wt) KOH solution was performed. Our Ta components were also chemically polished by means of a saturated potassium dichromate solution in hot (110 °C), concentrated H_2SO_4. For the Cu assemblies, a "bright dip" chemical etch[9] consisting of 1/2 oz HCl mixed per gal of the solution, 60% (vol) H_2SO_4 and 15% (vol) HNO_3 in water,

*For stainless steel parts, we used the "Summa Process." (Huntington Mechanical Laboratories, Inc., Mountain View, CA)

Table 4.1: Procedure used for degreasing vacuum parts.

Step	Action	Duration (min)	Comments
1	Detergent Wash	10[a]	Scrub, Hot Water
2	D.I. Water Rinse	10[a]	...
3	TCE Rinse	10[a]	Repeat until clear
4	Acetone Rinse	10[a]	...
5	Methanol Rinse	10[a]	May substitute 2-propanol[b]
6	D.I. Water Rinse	10[a]	...
7	Air dry	...	Dust-free area

a. Use ultrasonic agitation if possible, or else pour fresh liquid over part after immersion.

b. Reference 9.

Table 4.2: Procedure for chemical processing of vacuum parts.

Material	Etchant	Duration	Comments
Cu	60% (vol) H_2SO_4 15% (vol) HNO_3 1/2 oz/gal HCl	5–45 sec	Room temp.
Mo,W	20% (wt) KOH soln.	5–15 min	Boiling
Ta	Saturated Potassium Dichromate soln. in hot (110 °C) conc. H_2SO_4	5–15 min	Keep soln. red

was used. The chemical processing of our metal system parts is summarized in Table 4.2. (The procedures shown in the table have been adapted from information contained in Refs. 9 and 10.) Following chemical processing, all parts were thoroughly rinsed in D.I. water, and allowed to dry completely in a dust-free location. In particular, we found it to be particularly important to allow the Mo and W parts to dry thoroughly to avoid discoloration, presumably due to oxide formation. For the preparation of Macor and alumina ceramics, only the degreasing procedure listed in Table 4.1 above was followed.

4.5.4 *In Situ* Lubrication

We close this section with a mention of a particular problem associated with moving parts of stainless steel against stainless steel. When stainless steel surfaces of the cleanliness required for UHV use are slid against one another, as in

internal screw threads, bearings, etc., there is a strong tendency of the extremely clean surfaces to gall, *i.e.*, to form a *cold weld*, resulting in seizure of the parts. The problem can occur outside the vacuum system as well as inside, since all of the flanges, nuts, bolts, and other fastening hardware are made of the same (type 304) stainless steel. Galling is avoided outside the system by means of lubrication; typically, a MoS_2-containing grease is applied. Internally, a dry form of lubrication, sufficient to prevent galling, and UHV compatible, is needed. Two approaches are commonly employed to prevent galling in the UHV environment. One is to *Au-plate* the stainless steel surface by means of conventional electroplating techniques. This is a particularly common practice for screw threads inside the system. The other means to prevent galling is to apply a thin, hard, WS_2 film over the stainless steel surface. This process is commerically available and is known as "dicroniting." Dicroniting results in a *hard, low-friction* surface with acceptable vacuum characteristics for temperatures below 1350 °C, and is commonly used on ball bearings in conventional sample manipulators. In our system, the rollers in the shutter mechanism have been dicronited as well.

4.6 Preliminary Results

In this section, we shall discuss the results of our *preliminary* growth attempts with the Si/silicide system. In these first trials, we attempted to fabricate epitaxial $NiSi_2$ films with the use of the SPE and RDE techniques. We did not attempt to fabricate any films by MBE (codeposition of Si and Ni) because of problems with one of our *e*-gun power supplies although we did, in an earlier trial, successfully evaporate Si from one of our *e*-beam sources. The primary purpose of our preliminary growth attempts was not to grow high quality epitaxial films, but rather to observe the characteristics our apparatus under actual

growth conditions. The results, which we shall now present, demonstrate the ability of our system to perform rate-controlled evaporations onto a heated substrate at a reasonable system pressure, which, after calibration and optimization of the critical growth parameters, should eventually lead to the fabrication of high quality epitaxial films.

Our initial growths all took place on phosphorous-doped, nondegenerate, (111)-oriented Si wafers (13–17 Ω·cm). Prior to growth, the wafers were degreased and etched in an HF solution, soaked in a solution containing H_2O_2 and NH_4OH, and finally etched again in an HF solution. The samples were then blown dry and loaded into the growth chamber. The system was pumped down until a pressure in the range of 10^{-8} Torr was reached. For the RDE samples, the substrates were first heated under vacuum to a temperature we believe to be in excess of 850 °C, for a period ranging from 15–30 min to degas and clean the substrates. (All of the temperatures quoted here are merely estimates as we do not have a reliable calibration of the substrate temperature at this time.) For the SPE samples, no substrate heating prior to deposition was performed. Substrate temperatures near 650 °C were sought for silicide formation; in practice, we now believe that the actual substrate temperatures were significantly higher.

4.6.1 System Performance

Deposition of Ni took place under closed-loop rate control, at evaporation rates ranging from 1 to 4 Å/sec. Film thicknesses ranged from 170 to 860 Å. (Once again, these values are estimated since we do not have a thickness calibration yet.) Figure 4.9 shows the evaporation rate vs. time for a nominal, 4 Å/sec, rate-controlled deposition. In this instance, the shutter was opened at a point beyond the initial rate overshoot seen in the figure. The small variations about the nominal rate are due to *fluctuations* in the physical process, combined with

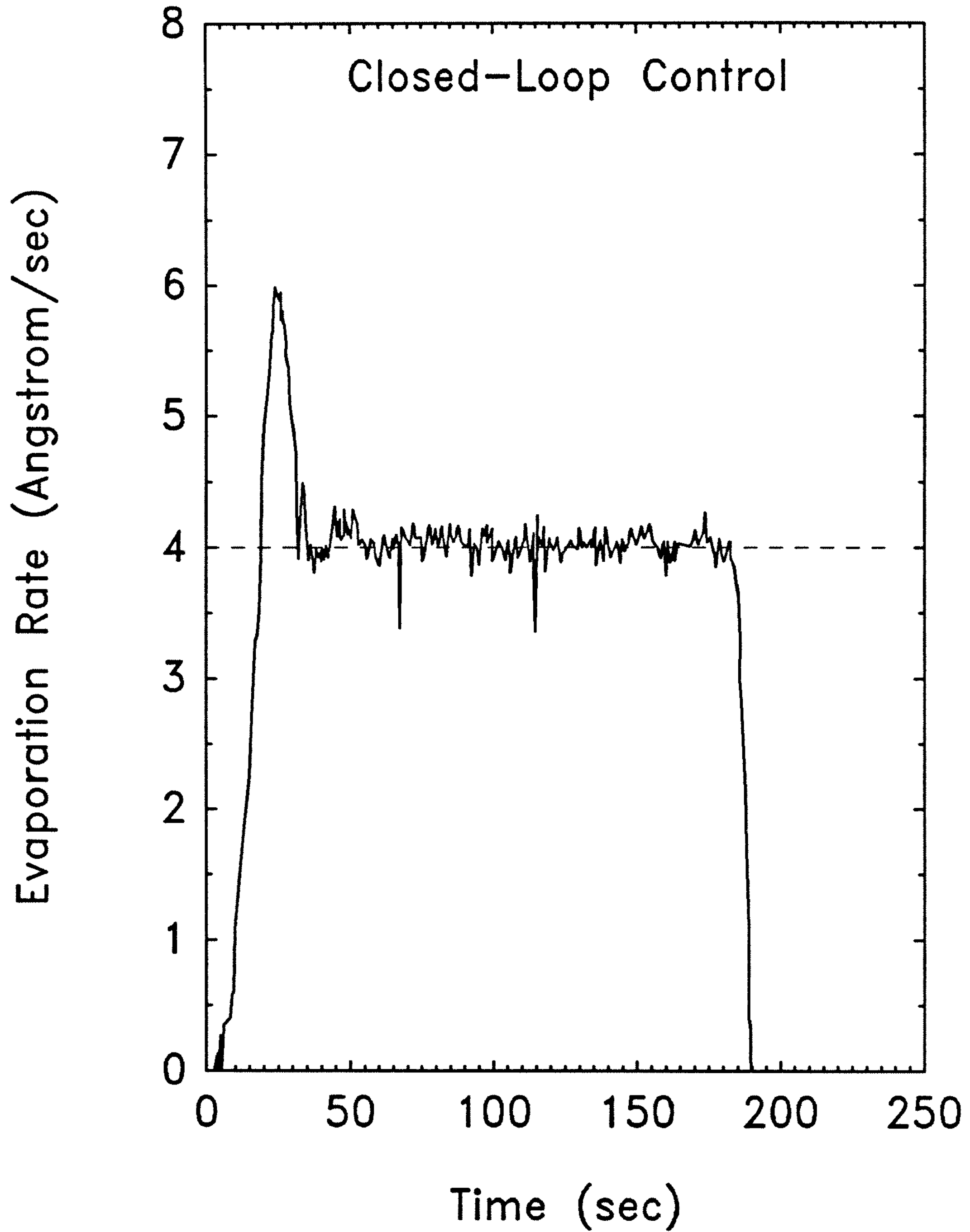


Figure 4.9: Actual evaporation rate of Ni from an e-beam source under closed-loop control. The nominal rate is 4 Å/sec. Aside from the initial overshoot, the (small) variations about the nominal rate are due to process flucutations and instrumental digitization noise.

digitization noise in our instrumentation, rather than *oscillations* arising out of an improper choice of control-loop feedback parameters. Typically, the background pressures during deposition were observed to be in the upper 10^{-8} Torr-range. Residual gas analysis shows most of the process gas to consist almost entirely of H_2, particularly when the *substrate heater* is in operation. Our *e*-guns appear to be producing some CO, particularly when first turned on after the system has been exposed to atmosphere. However, we must point out that our system has not been baked prior to the growths. Moreover, the contribution of H_2 to the total pressure exceeds all other components by more than an order of magnitude, under the growth conditions described, and with the shutter and substrate cryoshrouds filled with LN_2.

4.6.2 Sample Characterization

The four samples which we have grown have been characterized by optical microscopy and electrical measurements. While we cannot ascertain the stoichiometry of our films, we observe that they do not come off with adhesive tape, and that they resist an etch which is known to remove Ni but not $NiSi_2$. Therefore, we believe that the films are some silicide phase, Ni_xSi_y, and are probably epitaxial. Electrically, all of our silicide/Si structures appear to be "leaky" Schottky diodes, exhibiting a nearly linear *I-V* characteristic in reverse bias, but at the same time, the expected light-sensitivity. Attempts at *C-V* analysis have not yielded meaningful results. The forward *I-V* characteristics appear to be strongly dominated by series resistance, making Schottky barrier height determination from the forward characteristic unreliable. Hence, we were unable to obtain any quantitative information about our samples from electrical methods.

The surface morphologies of our four samples differ substantially. As an example, compare Fig. 4.10, which shows the surface of our RDE samples, with

Fig. 4.11, showing an SPE sample. In the case of the RDE sample, the silicide surface is seen to be essentially mirror-smooth, *except* for occasional "craters" in the substrate surface. These features are very similar in appearance to ones observed on the $CoSi_2$ samples studied in Chap. 3 (see Fig. 3.1 on page 97). In contrast, the SPE sample, shown in Fig. 4.11, is seen to have a much rougher, "plaster ceiling" appearance. However, we caution that *no* systematic variation of growth conditions was attempted in our preliminary growths, so that such comparisons as between Figs. 4.10 and 4.11 are illustrative, but not particularly meaningful; in particular, the substrate temperatures, evaporation rates, film thicknesses, and substrate thermal history were all different, quite apart from the fact that one sample had been fabricated by SPE and the other by RDE.

Our preliminary findings to this point suggest that our biggest problem at present is substrate temperature: both calibration and control. In particular, we believe that our substrate temperatures in all cases were significantly higher than the heater thermocouple reading would indicate. However, this is simply a matter of calibration. We cannot overemphasize that the preliminary growths that have taken place represent the *very first tests of a completely uncalibrated system*, and that an accurate choice of the growth parameters, particularly, the substrate temperature during growth,[6] is *crucial* to the attainment of good results. Moreover, the system itself was intentionally not used in its full, UHV configuration during these initial growths because this would have greatly restricted our degree of access to the growth chamber. Consequently, the poor silicide results obtained are due to our ignorance about the pertinent conditions of sample growth, and in no way represent an *inherent* limitation of the capability of our system. Nor is the ultimate system performance itself limited to that observed during these preliminary tests. Indeed, the performance of our system has lived up to or exceeded our expectations. The initial growth attempts were

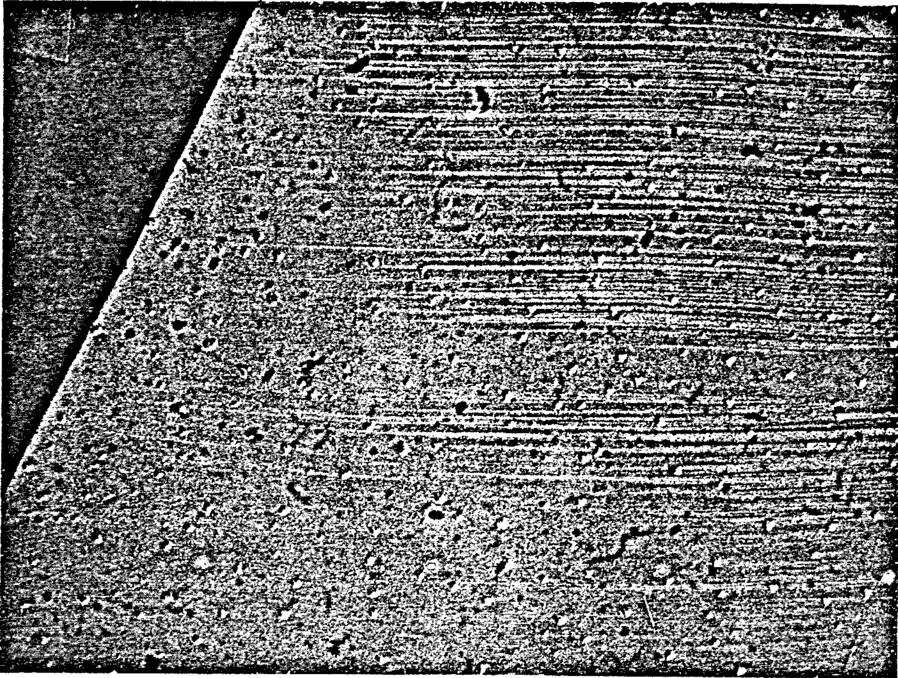


Figure 4.10: Surface morphology of a $NiSi_2$ film fabricated by RDE. Film thickness is believed to be $\approx 600\,\text{Å}$. The surface appears mirror-smooth, apart from occasional "craters" ($\sim 1\,\mu$m or less in size). We estimate the substrate growth temperature to be in the range from 700–950 °C.

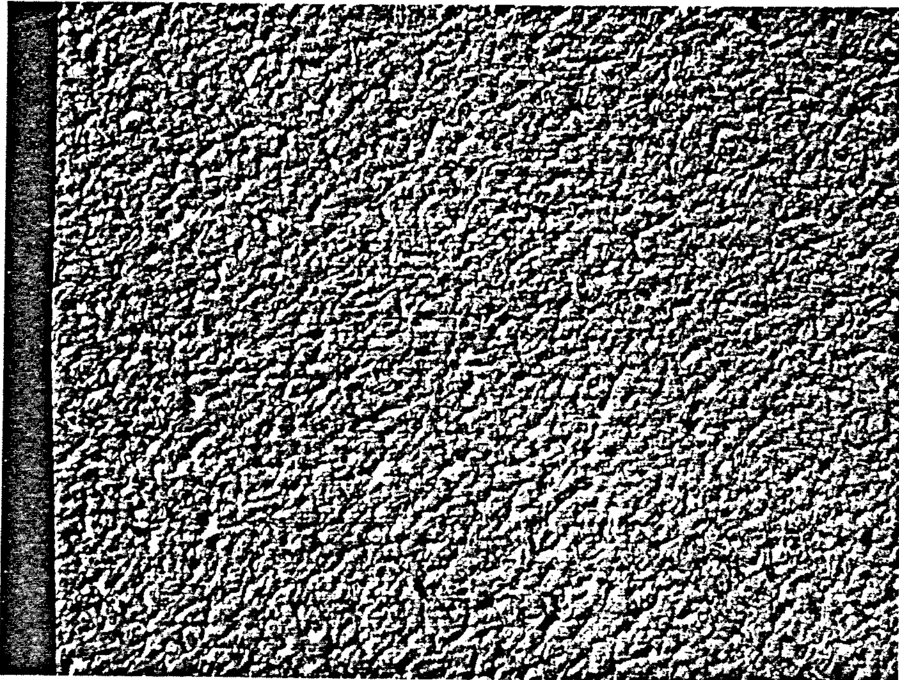


Figure 4.11: Surface morphology of a $NiSi_2$ film fabricated by SPE. We estimate the annealing temperature to be between 570 and 760 °C, and film thickness to be ≈ 3000 Å. The surface appears highly textured under Nomarski imaging.

designed merely to determine *if* our system would work, not *how well* it would work. In summary, we have tested our system by performing actual growths, and have found it to perform satisfactorily. Thus, we believe that, in principle, our system is capable of producing epitaxial silicide films of comparable quality to those studied in Chapters 2 and 3 of this thesis.

4.7 Problems and Suggested Remedies

From our preliminary growth attempts, we have observed a few problems, which, if addressed, would undoubtedly further improve the performance of our Si/silicide system. In this section, we discuss those problems, and make suggestions for minimizing or alleviating them. The first problem concerns the substrate heater filament. At present, we make use of a 100 cm-long, pancake-coil filament. The problem with this is that the filament, when heated, sags under its own weight, to the point of coming into physical contact with the wafer, and its holding rings. The result of this contact is a nonuniform substrate temperature distribution, which, in extreme cases, causes the wafer to crack. As a short-term solution, we have found that stacking *two* wafers in the ring assembly greatly improves the substrate temperature uniformity. The permanent solution is to modify the filament: either by changing its geometry to a *serpentine* pattern, and shortening its overall length (65 cm should be quite adequate), or by going to a rigid, SiC-coated *graphite* radiating element. The matter of substrate temperature control, refered to in the preceding section, is easily solved, too, by the rather trivial addition to the existing system of an *optical pyrometer* to monitor substrate temperature.

The next problem concerns the outdiffusion of H_2 gas from the Ta rings: At present, the filament touches these rings and heats them above 760 °C, resulting

in the outdiffusion. This is a well known problem with Ta, the solution to which is to bake the Ta rings in vacuum above 760 °C.

In our preliminary growth attempts, we noticed that alignments in the system appear to be very critical. In particular, there appears to be a larger-than-expected variation in film thickness across the wafer. A solution to this problem might be to modify the shutter housing (shown in Fig. 4.7) so that there is a single, larger aperture per *e*-gun, rather than the *pair* of apertures per *e*-gun that now exist. This modification would also simplify the loading of the *e*-guns, which with the present design, is somewhat awkward.

The next problem concerns flaking. While we have attempted to address this problem by roughening the collecting surfaces, we apparently have not gone far enough. It appears that the underside of the shutter housing and water-cooled shields (refer to Fig. 4.4) should be sand-blasted again, this time using a *course grit* of clean, alumina particles.

Finally, we and others[6] have observed that small Mo screws tend to seize in threaded holes, and become impossible to remove without drilling them out. One empirical solution that has been used elsewhere[6] is to rinse the Mo screws in methanol before assembly. Two possibilities that one might further suggest would be to use a dry lubricant, such as graphite powder, or the dicroniting process, on the Mo screws, to get around this difficulty.

4.8 Summary

We have described in detail the design, construction, and testing of an ultra-high vacuum, Si/silicide growth chamber, capable of molecular beam epitaxial growth of silicide films on a conventional Si wafer. The system described is capable of codeposition of Si, and a transition metal, under tightly controlled

conditions of evaporation rates, substrate temperature, and ambient pressure. Such controlled conditions are essential to the fabrication of extremely abrupt, defect-free, planar interfaces, and structurally ideal and stoichiometrically well controlled epitaxial films. The original contribution in this work lies in the *design*, both of the overall system configuration, and of the many, detailed system components, subject to the serious constraints imposed by the requirements of the MBE growth process. A number of currently active issues in Si MBE technology have been addressed, some uniquely, in the design of our system. Our main result is the demonstrated successful operation of the system. An extensive discussion of Si MBE process requirements, practical design and materials considerations, preliminary results of silicide growth, and observed problems and suggested remedies, has been given. Finally, based on our results and experiences thus far, we believe that our Si/silicide system is capable of fabrication of very high quality, epitaxial silicide structures, provided that the relevant growth parameters are properly calibrated and optimized.

References

1. O. J. Marsh, private communication.
2. W. M. Brubaker, in *Methods of Experimental Physics*, Vol. 14, ed. by G. L. Weissler and R. W. Carlson (Academic, New York, 1979), pp. 183–192.
3. Water-cooling around the e-beam evaporators is used in a system built at AEG-Telefunken, Forschungsinstitut, Ulm, Federal Republic of Germany because of the suspicion of stimulated desorption of hydrocarbons from cryogenic surfaces during growth, according to Ref. 1.
4. M. T. Thomas, in *Methods of Experimental Physics*, Ref. 2, Chap. 16.
5. T. J. Maloney, private communication.
6. L. J. Schowalter, private communication.
7. Y. Shapira and D. Lichtman, in *Methods of Experimental Physics*, Ref. 2, Chap. 8.
8. L. F. Johnson, private communication.
9. N. Milleron and R. C. Wolgast, in *Methods of Experimental Physics*, Ref. 2, Chap. 9.
10. G. Petzow, *Metallographic Etching* (American Society for Metals, Metals Park, 1978).

Chapter 5

Calculation of Vibrational Modes of Point-Defects in Semiconductors

5.1 Introduction

5.1.1 Background

This chapter concerns the calculation of local vibrational modes (LVM) of point-defects in semiconductors. The point-defects which are the subject of our calculations are single-atom impurities, both substitutional and interstitial. In contrast to the other chapters in this thesis, the work presented in this chapter is purely theoretical. (The particular systems studied here also differ from the rest of the thesis in that they have nothing specifically to do with Si MBE.) We have used the Green's function technique to calculate the LVM spectra of three distinct defect systems: substitutional oxygen on a phosphorus site in GaP (GaP:O_P), substitutional and interstitial iron in Si (Si:Fe), and substitutional carbon on an arsenic site in GaAs (GaAs:C_{As}). The first two of these represent

examples of a *deep-level* defect. We discussed deep levels earlier in Section 1.4. Because deep levels may be electrically active, *e.g.*, acting as electron (hole) traps, generation centers, or recombination centers, depending on the relative rates of carrier capture and emission,[1] they may significantly influence the electrical and optical properties of a semiconductor, and, therefore, the operational characteristics of electronic and optoelectronic devices. Clearly, then, the characterization and understanding of deep levels is of importance to device electronics.

Unfortunately, our present theoretical understanding of deep levels in semiconductors is not too good. An example is the Si:Fe defect, which, as we shall see in Section 5.4 below, is of technological importance, yet the *nature* of the defect, *e.g.*, *where* and *how strongly* the Fe defect is bonded in the lattice, whether the defect exists as a single-atom impurity, or forms defect *complexes* with its own or with other impurity species, is not completely understood.[2] In contrast, the GaP:O_P defect is a special case in that it has been extensively characterized, and so is an excellent "prototype" system on which our theoretical understanding of deep levels may be tested. We have already discussed (see Section 1.4) the usefulness of optical measurement techniques such as photoluminescence and IR-absorption spectroscopy in the characterization of deep levels. In Section 1.4, we also discussed some mechanisms by which defect LVM's may be observed in optical spectra, and how a detailed understanding of the vibrational properties near a defect can yield information about the *nature* of the defect. With the use of theoretical LVM calculations to aid in the interpretation of experimental spectra, it becomes possible to identify LVM energies, symmetries, and defect force constants, *i.e.*, "spring" constants, directly. Indirectly, assumptions such as defect site symmetry, local lattice relaxation, and the composition of the defect complex can be tested, and the bonding strength of the impurity to the lattice (related to but different from the force constants) can be qualitatively inferred. In order for

the comparison between theory and experiment to be meaningful, a fairly elaborate theoretical method, which takes into account the many vibrational degrees of freedom of the impurity *and* the bulk lattice, is required. Fortunately, such a theoretical approach, the Green's function method, exists, and will be discussed in Section 5.2.1 below.

5.1.2 Results of this work

Briefly, the major results of our LVM calculations for the GaP:O_P, Si:Fe, and GaAs:C_{As} defects are as follows. We have calculated the local vibrational modes of the GaP:O_P defect. A 15-parameter, lattice-dynamical model was used to describe the host GaP lattice, and the oxygen defect was represented as a mass perturbation, and a perturbation in the first- and second-nearest-neighbor interactions at a phosphorus site. Our calculated LVM's agree *in detail* with observed LVM's involving O^+ and O^0 defects. For the first time, the major LVM peaks observed experimentally are accurately predicted by a single calculation. In particular, with essentially only two adjustable parameters to specify the GaP:O_P defect, we accurately predict two modes of A_1 symmetry, and two modes of T_2 symmetry. Our calculations suggest that the O^+ and O^0 defects are *weakly* bound in GaP. We find detailed quantitative agreement between theory and experiment by weakening the O-Ga and O-P force constants to 38% of the bulk, P-site values, while simultaneously weakening the Ga-P back-bond force constants (interactions between the NN and 2NN shells) to 93% of the bulk value. Overall, taking long-range interactions into account, we find that the oxygen defects are bonded to the GaP lattice with an effective force constant of 3.5 eV/ Å^2, which is 25% of the bulk, P-site effective force constant. Finally, in order to determine the effective force constant with long-range interactions taken into account, we have derived a purely analytical result (Eq. 5.30), suitable for the computation of this quantity.

For the Si:Fe defect, we are interested in low-energy ($< 10\,\text{meV}$) LVM's. A photoluminescence feature, believed to be a 9 meV LVM of an Fe impurity, has been observed experimentally. This mode is thought to involve significant vibrational motion of the Fe impurity; consequently, it is expected to undergo a frequency shift under isotopic substitution of the Fe. Our calculations support this. We have calculated the LVM's for both substitutional and interstitial Fe single-atom impurities, and find that in both cases, a single, low-energy LVM involving significant Fe motion is predicted. We also verify by direct computation that the predicted LVM shifts upon the mass substitution, $^{54}\text{Fe} \rightarrow {}^{56}\text{Fe}$. In the process of treating the interstitial, we have extended the Green's function formalism quite generally to include an arbitrary defect *complex* (Eq. 5.36), which is here specialized to the case of a single impurity atom. Our major computed result is in *disagreement* with the experimentally observed absence of an isotope shift, however. From this, it is concluded that the experimental feature is not a LVM of an Fe impurity in Si.

Our last calculation concerns the experimentally observed, fivefold splitting of a localized mode (near 72 meV) of the carbon acceptor in GaAs. The splitting of this mode is significant since it has been attributed to the natural isotopic variation among the Ga NN shell, thereby establishing the substitutional site of the carbon as an As site. However, early theoretical work incorrectly predicted the splitting to be *fourfold*, and the original experimental measurements of this structure were unable to clearly resolve a fifth peak. Recent, high-resolution measurements by Hauenstein, Baukus, Feenstra, and McGill[3] have since clearly demonstrated a *fivefold* splitting. Hence, we have re-examined the problem theoretically, using both Green's function and finite-cluster calculations. We find that *nine* distinct modes are predicted rigorously from simple symmetry considerations, and these broaden into *five* bands. The peak positions and relative

amplitudes of our five bands are in good agreement with the experimentally observed peaks. Finally, the close agreement between our Green's function and finite-cluster calculations demonstrates that almost all of the mode energy is confined to the carbon atom and its Ga nearest neighbors.

5.1.3 Outline of Chapter

The remainder of Chapter 5 is organized in the following manner. In Section 5.2, we discuss the two main theoretical tools that are used in all our LVM calculations; namely, the Green's function technique, and the use of group theory to simplify our work by exploiting the symmetries in the problem. Next, in Section 5.3, the work on the GaP:O_P defect is presented. The calculation of interstitial and substitutional Fe defects in Si (Si:Fe) follows in Section 5.4. Finally, this chapter concludes with Section 5.5, in which our calculation of the fivefold-split localized mode of the GaAs:C_{As} defect is presented.

5.2 Theoretical Techniques

5.2.1 Green's Function Approach

The work presented in this chapter makes use of the Green's function technique for the calculation of vibrational modes of point-defects. The full vibrational problem of a point-defect in a lattice is a *many body* problem. For example, in a GaP lattice there are N unit cells, each cell containing two basis atoms, and each atom possessing three degrees of freedom, namely, x, y, and z motion. Consequently, there are $6N$ degrees of freedom in the problem of determining the $6N$ normal vibrational modes for the entire lattice. When a *substitutional* impurity is introduced, there remain $6N$ degrees of freedom, and the resultant *perturbed* system will also contain exactly $6N$ normal modes. Our problem is that we wish

to determine only those *relatively few* modes of the perturbed system which will have their energy spatially localized near the defect, without having to explicitly determine all $6N$ modes, particularly since, for a typical laboratory-sized specimen, $N \sim 10^{19}$–10^{22}.

The Green's function technique is useful whenever the defect is *spatially localized*. Then, it turns out that we need only to work explicitly with the degrees of freedom *directly* affected by the perturbation. For example, a single-atom impurity is represented as a perturbation of mass and force constants at a particular lattice site. Let us consider a zinc blende lattice, and further consider that the lattice dynamics can be adequately modeled with only NN interactions (springs) acting between atoms. In this case, the substitutional defect involves only *five* atoms, and therefore $3 \times 5 = 15$ degrees of freedom, out of a possible $6N$.

Let us now outline the Green's function method more precisely. To determine the lattice dynamics of the entire problem, we begin by determining the lattice dynamics of the perfect lattice, with the use of Newton's second law:

$$(\underline{\mathrm{M}}\Omega^2 - \underline{\Phi})\underline{\mathrm{u}} \equiv \underline{\mathrm{L}}\,\underline{\mathrm{u}} = 0. \tag{5.1}$$

This is a matrix equation, of order $6N$ (continuing with our above example of a zinc blende lattice), whose rows and columns are indexed by the Cartesian coordinates of the $2N$ atoms. Note that in Eq. 5.1 and in all that follows, we have assumed a harmonic time-dependence of the displacements $\underline{\mathrm{u}}(t) = \underline{\mathrm{u}}e^{-i\Omega t}$. The quantity $\underline{\mathrm{M}}$ represents a diagonal, $6N \times 6N$ matrix whose elements are the atomic masses in the problem. The force constants describing interatomic interactions are completely contained in the $N \times N$ matrix, $\underline{\Phi}$. Finally, the atomic displacements are contained in the column vector $\underline{\mathrm{u}}$. Equation 5.1 is an eigenvalue problem, whose solution yields the normal mode frequencies and coordinates of the entire lattice.

The perfect-lattice Green's function is obtained through the formal definition,

$$\underline{G}^0 \equiv \underline{L}^{-1}, \tag{5.2}$$

where clearly, $\underline{G}^0$ is a function of Ω^2.* From Eq. 5.1, we see that $\underline{L}$, and therefore, $\underline{G}^0$, are diagonalized when expressed in terms of the bulk normal coordinates. In fact, the easiest way to calculate the matrix elements of $\underline{G}^0$ requires a knowledge of the bulk eigenfrequencies (normal modes) and eigenvectors (normal coordinates).[4]

The introduction of an impurity results in a mass and/or force constant perturbation which, in the present matrix formalism, is represented by the quantity $\underline{\delta L}$ as

$$\underline{\delta L} \equiv \underline{\delta \Phi} - \underline{\delta M}. \tag{5.3}$$

The equations of motion for the entire *perturbed* system can then be written as

$$(\underline{L} - \underline{\delta L})\underline{u} = 0. \tag{5.4}$$

By multiplying Eq. 5.4 from the left by $\underline{G}^0$ we get, using Eq. 5.2,

$$(\underline{1} - \underline{G}^0\underline{\delta L})\underline{u} = 0. \tag{5.5}$$

The matrix $\underline{\delta L}$ in general will have all elements equal to zero *except* those within the subspace directly affected by the perturbation. Now comes the big advantage of the Green's function method: because $\underline{\delta L}$ is nonzero only within a small subspace, Eq. 5.5 remains valid *entirely* within the subspace of the defect.[5] In other words, we have drastically reduced our computational task from a matrix equation of order $6N$ down to order n, where n is the total number of degrees of freedom affected by the perturbation (15 in our example above). The solvability

*Defined as in Eq. 5.2, $\underline{G}^0$ becomes singular at the bulk eigenfrequencies. This difficulty is avoided by the addition of a small imaginary component, *i.e.*, $\Omega^2 \rightarrow \Omega^2 + i\epsilon$.

condition for Eq. 5.5 requires that

$$\det\left|\underline{1} - \underline{G}^0(\Omega^2)\underline{\delta L}(\Omega^2)\right| = 0 \tag{5.6}$$

for nontrivial solutions. The frequencies Ω for which Eq. 5.6 is satisfied are the *localized* modes of the defect. Similarly, it can be shown[5] that a necessary (but not sufficient) condition for a *resonant* mode is given by

$$\mathrm{Re}\left\{\det\left|\underline{1} - \underline{G}^0(\Omega^2)\underline{\delta L}(\Omega^2)\right|\right\} = 0. \tag{5.7}$$

The reason that this condition is not sufficient is that the resonance may be too broad to be distinguishable from the background of bulk lattice vibrations, or it may be in the wrong direction (an *antiresonance*). As discussed by Feenstra,[4] it is usually necessary to examine the local density-of-states (LDOS) in order to establish the existence of a resonant mode. Hence, in order to solve for the LVM's, we need a means of computing the bulk Green's functions in the subspace of the defect, and also, a description of the defect. In practice, both $\underline{G}^0$ and $\underline{\delta L}$ are given in terms of some phenomenological model, in terms of which, the interactions between atoms are described.

5.2.2 Group Theory

We have just seen how the finite extent of the defect results in a significant reduction in dimension of the matrix equations that must be solved. Now, we shall describe the use of group theory for the purpose of further reducing the size of the matrices, and consequently, the number of independent matrix elements in the problem. This simplification is a consequence of the physical *symmetry* of the problem. As a familiar example, consider the calculation of bulk phonon modes for our zinc blende lattice. The $6N$ eigenvalues and eigenvectors are given in principle by solving Eq. 5.1, a $6N \times 6N$ matrix equation. However, because of

the *translational symmetry*, we know in advance that the eigenvectors must be *plane waves* of the form

$$\mathbf{u}_{l\kappa} = \mathbf{u}_{0\kappa} e^{i\mathbf{k}\cdot\mathbf{R}_l}, \tag{5.8}$$

where $\kappa = 1, 2$ identifies the basis atom, l labels the unit cell, $\mathbf{k}$ is a wave vector, and $\mathbf{R}_l$ is a lattice translation vector. This knowledge affords us a tremendous simplification. By considering displacements of the form of Eq. 5.8, Eq. 5.1 is reduced from a $6N \times 6N$ problem to N 6×6 problems. Said in another way, by making a (unitary) transformation from our original basis of Cartesian coordinates on separate atoms, to a new basis of "collective" coordinates which look like plane waves, Eq. 5.1 becomes *block-diagonalized* into a set of N 6×6 blocks. Each of these N blocks is *labeled* by an allowed value of $\mathbf{k}$, as are the displacement eigenvectors, and each block can be solved *independently*. (There are precisely N allowed values of $\mathbf{k}$ in the first Brillouin zone for an N-celled lattice.) Note that $\mathbf{k}$ also labels the eigenvalues. Hence, for a given $\mathbf{k}$, the solution of the 6×6 eigenvalue problem yields six eigenvalues, resulting in a total of $6N$ normal modes, as required.

In the language of group theory, $\mathbf{k}$ is said to label an *irreducible representation* of the lattice-translation group. The collective coordinates, *i.e.*, our plane waves in this example, are said to *transform* (under a lattice-translation operation) according to the irreducible representations of the group. For example, under any lattice translation $\mathbf{R}_l$, any plane wave coordinate labeled by $\mathbf{k}$ simply changes phase by an amount $e^{i\mathbf{k}\cdot\mathbf{R}_l}$. In other words, the irreducible representations, which in general are matrices, are all one-dimensional (phase factors) in the present example. Note that in choosing the collective-coordinate basis, we have not "guessed" the final answer; for a given $\mathbf{k}$, we must still solve an eigenvalue problem of order 6. What we *have* done is to make use of the symmetry of the problem in order to: break the original problem into smaller parts, which

can be handled independently; and, label and classify the normal modes by the irreducible representation (**k**) according to which their associated eigenvectors transform under a symmetry operation.

Now, let us introduce a substitutional impurity atom into the zinc blende lattice. The introduction of the impurity has broken the translational symmetry of the lattice, but the *point-group* symmetry remains. In the case of the zinc blende structure, which is what we consider in all of our calculations, the point-group is T_d. The use of group theory to simplify our LVM calculation proceeds in an entirely analogous manner to our above example. Within the subspace of the defect, spanned by the Cartesian coordinates, $x_1, y_1, z_1, \ldots, x_n, y_n, z_n$, we must *deduce* a set of collective coordinates which transform according to the irreducible representations of T_d. The mathematical procedure for accomplishing the decomposition of the Cartesian-coordinate (reducible) representation into irreducible representations is straightforward.[6,4] For example, the 17-atom cluster (considered in Section 5.3 below) consisting of a substitutional defect, and its first- and second-nearest neighbors, has a 51-dimensional *reducible* representation Γ which reduces to

$$\Gamma = 3A_1 \oplus A_2 \oplus 4E \oplus 5T_1 \oplus 8T_2, \tag{5.9}$$

where the labels, $A_1, \ldots, T_2$, refer to the irreducible representations of the T_d point-group. Some of these irreducible representations have a simple physical significance. For instance, A_1 refers to a fully symmetric, "breathing" motion, and T_2 denotes a vectorlike displacement. While the decomposition (Eq. 5.9) is readily performed, the determination of collective coordinates is somewhat more difficult. In practice, the collective coordinates used in our work were determined by inspection. Having deduced the collective coordinates, $\underline{G}^0$, $\delta\underline{L}$, and any other matrices exhibiting the T_d symmetry become block diagonalized[4] in the collective-coordinate representation. Then, Eq. 5.7 can be solved separately for the blocks,

$A_1, A_2, \ldots$, for the roots Ω^2 which correspond to a LVM of energy $\hbar\Omega$, labeled by the appropriate representation.

Essentially, the procedure for determining LVM's is as follows: (1) the bulk lattice dynamics are specified (with the use of a phenomenological model); (2) the defect is defined in terms of the same model, and the collective coordinates determined; (3) from this, the independent matrix elements of the bulk Green's function $\underline{G}^0$ are computed in the collective-coordinate representation; (4) the perturbation $\underline{\delta L}$ is constructed (analytically or numerically) in the collective-coordinate representation; (5) Eq. 5.7 is solved for the roots Ω^2, *separately* for each irreducible representation. The computational procedure for determining the matrix elements of $\underline{G}^0$ is given by Feenstra.[4] The major difference between the work of Feenstra and the present work concerns the location and extent of the defect, the collective coordinates needed, and the lattice-dynamical model used. What has been attempted in Sections 5.2.1 and 5.2.2 is a *description* of a fairly elaborate calculation. The reader is refered to Ref. 4 for further computational or mathematical details not covered in these two sections.

5.3 Vibrational Modes of Oxygen in GaP

In this section, we use the Green's function method to calculate the local vibrational modes (LVM) of a substitutional oxygen defect on a phosphorous site in GaP. Much theoretical and experimental effort has been devoted to the understanding of the GaP:O_P defect system. Our primary interest in the GaP:O_P defect is as a prototype for understanding the nature of deep levels in semiconductors (see also the discussion in Section 1.4 above). This work represents an extension of an earlier Green's function calculation performed by Feenstra, Hauenstein, and McGill.[7] In our earlier calculation, the phenomenological model

on which the bulk lattice dynamics, and, ultimately, the LVM spectra are based, was restricted to nearest-neighbor (NN) interactions between atoms. This limitation restricts the possible local vibrational modes to involving only the oxygen defect itself and its four nearest-neighbor Ga atoms.* Also, the simplicity of the NN model results in calculated local modes which are only *qualitatively* correct, *i.e.*, correct in symmetry, but not precisely correct in energy $\hbar\Omega$. Hence, it is desirable to consider a more realistic lattice dynamical model of the host lattice in order to more accurately predict the LVM's of an oxygen defect within it.

Here, we use the deformable-bond approximation (DBA) of Kunc *et al.*[8] as the phenomenological description of bulk GaP. In contrast to our previous work, the 15-parameter, DBA model includes both short-range interactions [out to second-nearest-neighbors (2NN)] *and* long-range, Coulomb effects. The long-range effects arise in the present model from the electrostatic interaction between charges, and electric dipoles associated with the various lattice sites.[8] The "defect" considered in this work consists of a *mass perturbation* at a *substitutional* site, and a perturbation in the *short-range interactions* among the 17-atom cluster consisting of the O, the four neighboring Ga, and the 12 2NN P atoms. We do not explicitly consider perturbations in the Coulomb interactions since their long range would render the Green's function approach intractable. As we shall see, however, the Coulomb portion of the total interaction is relatively small, and can be effectively contained in the short-range defect perturbation, for sufficiently localized modes.

The DBA, and the specifics of this Green's function computation are briefly described in Sections 5.3.1 and 5.3.2, respectively. In Section 5.3.3, we explicitly specify the defect perturbation considered in this work. Next, our results

*Actually, the restriction comes about because only the interaction between the defect (oxygen) and its nearest neighbors were modified in the previous calculation.[7] Thus, no mechanism for the existence of LVM's involving further neighbors was present.

our presented in Section 5.3.4. There, we first derive an *exact* analytical expression for the long-range contribution to the restoring force resulting from the displacement of a single atom in an infinite lattice. We shall need this result later in order to determine the bonding strength of the GaP:O_P defect. Next, in the same section, we present our computed dispersion curves, density-of-states (DOS), and selected Green's functions for bulk GaP. We conclude Section 5.3.4 with the results of LVM calculations for the GaP:O_P defect. (Here, we shall use the collective term "local vibrational mode" to refer to both strictly localized and resonant modes; the distinction between localized and resonant modes was explained earlier in Section 1.4.) Our interaction perturbation consists of varying the *short-range* O-Ga, O-P, and Ga-P (back-bond) force constants (in a manner consistent with fundamental constraints) from their bulk, unperturbed values. For convenience, we take the values of the O-Ga, O-P, and Ga-P defect force constants to be the respective fractions, η, η, and 0.12η, of the corresponding bulk, short-range values. Energies for local modes of a given symmetry are then presented as a function of the *single* parameter, η. The local density-of-states (LDOS) and "eigenvectors" are also presented for selected local modes, and a discussion of the major features, and the reason for the above choice of defect parameters is given.

An extensive comparison to available photoluminescence data is made in Section 5.3.5. There, we see that the electron capture luminescence data for the O^0 defect[9] is well described by our calculations for the value $\eta = -0.62$. In particular, we identify the phonon replicas at 24.7 and 28.4 meV with the peaks in our T_2 LDOS at 25.0 and 28.2 meV, respectively. Both peaks in the LDOS shift under isotopic substitution of ^{18}O for ^{16}O as expected. Our computed A_1 modes at 20.4 and 48.2 meV are observed for the charge states O^+ and O^0.[10,11] In all cases we find substantial improvement in the agreement, both qualitative

and quantitative, between calculated and observed values over that predicted by our earlier model.[7] Finally, the conclusions that can be drawn from this work are summarized in Section 5.3.6.

5.3.1 Bulk Lattice Dynamics

To describe the lattice dynamics of bulk GaP we have used the deformable-bond approximation (DBA). This model is a 15-parameter version of the deformable-dipole model of Kunc *et al.* and has been shown to provide a good description of several III-V compounds.[8] The quantities of interest for our calculations are the force constant matrix $\underline{\Phi}$ and the dynamical matrix $\underline{D}$, which is related to $\underline{\Phi}$ by the equation, $\underline{D} = \underline{M}^{-1/2}\underline{\Phi}\,\underline{M}^{-1/2}$. In this expression, $\underline{M}$ is a diagonal matrix whose elements are equal to the masses in the problem. In the DBA, the force constant matrix has the form

$$\begin{aligned} \underline{\Phi} &= \underline{\Phi}^{(sr)} + \underline{\Phi}^{(el)} \\ -\underline{\Phi}^{(el)} &= (\underline{Z} + \underline{N}^{\dagger})(\underline{1} - \underline{B}\,\underline{a})^{-1}\underline{B}(\underline{Z} + \underline{N}), \end{aligned} \tag{5.10}$$

where $\underline{\Phi}^{(sr)}$ and $\underline{\Phi}^{(el)}$ correspond, respectively, to short-range and electrostatic components. The short-range force constant matrix, $\underline{\Phi}^{(sr)}$, contains the non-Coulombic interactions which in the DBA extend out to second-nearest-neighbors. As a consequence of the T_d^2 space-group symmetry of GaP, the off-diagonal Cartesian blocks can be shown[8] to have the form

$$\underline{\Phi}^{(sr)}(\text{Ga;P}) = \underline{\Phi}^{(sr)}(\text{P;Ga}) = \begin{pmatrix} A & B & B \\ B & A & B \\ B & B & A \end{pmatrix}, \tag{5.11}$$

$$\underline{\Phi}^{(sr)}(\text{Ga;Ga}) = \begin{pmatrix} C_1 & D_1 & E_1 \\ D_1 & C_1 & E_1 \\ -E_1 & -E_1 & F_1 \end{pmatrix}, \tag{5.12}$$

and

$$\underline{\Phi}^{(sr)}(\mathrm{P;P}) = \begin{pmatrix} C_2 & D_2 & -E_2 \\ D_2 & C_2 & -E_2 \\ E_2 & E_2 & F_2 \end{pmatrix} \tag{5.13}$$

for interactions between nearest neighbors (Eq. 5.11) and second-nearest-neighbors (Eqs. 5.12) and 5.13), respectively. These Cartesian blocks must obey the transformation law[12]

$$\Phi_{\alpha\beta}(LK;L'K') = \sum_{\mu\nu} S_{\alpha\mu} S_{\beta\nu} \Phi_{\mu\nu}(l\kappa;l'\kappa'), \tag{5.14}$$

where $\underline{S}$ is the *point-group* operation associated with the space-group transformation which takes the site $(l\kappa)$ into the site $(l'\kappa')$. In this expression, l, l', L, and L' label the N unit cells in the lattice, κ, κ', K, and K' specify the basis atom, and the remaining indices refer to Cartesian coordinates. Hence, from Eqs. 5.11–5.14, and the translation-invariance condition[12]

$$\underline{\Phi}(l\kappa;l\kappa) = -\sum_{l'\kappa' \neq l\kappa}{}' \underline{\Phi}(l\kappa;l'\kappa'), \tag{5.15}$$

$\underline{\Phi}^{(sr)}$ for the entire lattice is determined. The description of the bulk, short-range interaction in terms of these 10 parameters (A–F_2) is the most general allowed by symmetry, and physically corresponds to both central and non-central pair interactions.[8]

In addition to purely short-range forces, the DBA force constant matrix contains a contribution from classical electrostatic interactions. Associated with each lattice site is an effective charge, $\underline{Z}$. Displacement of a charge from its equilibrium site effectively results in an electric dipole moment at the site. Dipoles may also be *induced*, either by the motion of neighboring atoms (mechanical polarizability $\underline{N}$) or by the electric fields produced by other dipoles (electronic polarizability $\underline{a}$). The dipoles, however generated, produce electric fields. These fields exert forces on the charges and mechanically induced dipoles, giving rise to

the Coulomb term $\underline{\Phi}^{(el)}$ of Eq. 5.10. Five parameters are used to determine $\underline{\Phi}^{(el)}$, giving a total of 15 for the DBA.[8] These 15 parameters have been determined for GaP by Kunc *et al.*[8]

5.3.2 Symmetry Considerations

The bulk Green's functions are numerically obtained from the eigenvalues and eigenvectors of the Fourier-transformed dynamical matrix in precisely the manner described in Ref. 7. From the bulk Green's function matrix $\underline{G}^0$, combined with the defect perturbation $\delta\underline{L}$, the standard formalism[5] is used to obtain the local modes as in our earlier calculation.[7] The major difference in the present work lies in the spatial extent of the defect considered. Here, the defect considered affects 17 atoms (the oxygen atom and its NN and 2NN), and, hence, 51 Cartesian coordinates. Thus, we are required to evaluate a 51×51 block of the bulk Green's function matrix.[5]

We shall consider our substitutional defect to retain the full point-group symmetry (T_d) of the lattice. Accordingly, it is natural and convenient to express the Green's function matrix in a basis of "collective coordinates" which transform according to the irreducible representations of T_d. It can be readily shown that the 51-dimensional, *reducible* representation (Γ) of T_d, given in terms of Cartesian coordinates of the 17 atoms, can be decomposed into irreducible representations according to Eq. 5.9. The collective coordinates, Q, transforming according to these irreducible representations (A_1–T_2), are related to the original Cartesian coordinates by an orthogonal transformation. The collective coordinates have already been determined for the five-atom cluster.[7] The additional collective coordinates needed here for the 12-atom, 2NN shell are given elsewhere.[13]

The advantage of the collective-coordinate transformation is that matrices

reflecting the T_d symmetry are block-diagonalized in the form[7,12]

$$G^{0(s)(r)}_{\mu\nu;\sigma\rho} = \delta_{(s)(r)}\delta_{\mu\nu}G^{0(s)}_{\sigma\rho}, \tag{5.16}$$

where (s) and (r) label the irreducible representation. The block-diagonalization allows us to consider each irreducible representation separately, and provides a classification scheme for the defect vibrational motion according to symmetry. For a given irreducible representation, μ and ν are indices labeling its dimension, and σ and ρ its occurrence, *e.g.*, $\mu = 1,2,3$ and $\sigma = 1,\ldots,8$ for T_2, a three-dimensional, irreducible representation which occurs eight times in Γ. The form of the block-diagonalization also restricts the number of independent matrix elements on the basis of symmetry. From Eq. 5.16, we see that the number of independent Green's function matrix elements is reduced to 68.

5.3.3 Defect Perturbation

The defect perturbation considered in this work consists of the mass substitution $m_P \rightarrow m_O$ at the defect site and a perturbation of the short-range force constants among the substituent and its first- and second-nearest neighbors. We wish to consider the most general force constant perturbation $\underline{\delta\Phi}$ having the block form of Eqs. 5.11–5.13 and consistent with T_d symmetry. Let us label the O-Ga perturbation parameters δA, δB; the O-P parameters $\delta C,\ldots,\delta F$; the Ga-P NN back bonds $\delta A'$, $\delta B'$; the intra-Ga-shell 2NN parameters $\delta C',\ldots,\delta F'$; and the intra-P-shell 2NN parameters $\delta C'',\ldots,\delta F''$. Now, not all these force constant parameters may be chosen independently. Applying the infinitesimal-rotation-invariance condition,[12] we find that

$$\delta E = \delta E' = \delta E'' = 0 \tag{5.17}$$

$$(\delta A' - \delta B') = 0 \tag{5.18}$$

$$(\delta A - \delta B) = -2\left[(\delta C' - \delta D') + \delta F'\right] \tag{5.19}$$

$$\delta F = -2(\delta C'' - \delta D'') \tag{5.20}$$

$$2\delta F'' = -(\delta C - \delta D). \tag{5.21}$$

The linear combinations of the form $(\delta A - \delta B)$ represent NN bond-bending spring constants.[7] Similarly, $(\delta C - \delta D)$, δF, and the corresponding primed expressions may be viewed as *anisotropic* bond-bending springs between 2NN pairs while the parameters δE correspond to a more complicated pair interaction. Essentially, Eqs. 5.17–5.21 impose constraints on the non-central pair interactions possible within an *isolated* cluster having the geometry of our 17-atom defect. The results to this point (given our above assumptions) are general, leaving nine free quantities to specify $\underline{\underline{\delta\Phi}}$.

We are now ready to parametrize the interaction perturbation. To reduce the number of free parameters, we assume the same fractional variation among *independent* O-Ga, O-P, and Ga-P parameters, giving

$$\frac{\delta A}{A} = \frac{\delta B}{B} = \eta_{\mathrm{O,Ga}} \tag{5.22}$$

$$\frac{\delta C}{C_2} = \frac{\delta D}{D_2} = \frac{\delta F}{F_2} = \eta_{\mathrm{O,P}} \tag{5.23}$$

$$\frac{\delta A'}{A} = \eta_{\mathrm{Ga,P}}. \tag{5.24}$$

If we make the analagous assumptions, $\delta C'/C_1 = \delta D'/D_1 = \delta F'/F_1$ and $\delta C''/C_2 = \delta D''/D_2$, we see from Eqs. 5.17–5.21 that the interaction perturbation $\underline{\underline{\delta\Phi}}$ is completely specified in terms of the η's. In this way, we specify directly the NN $(\eta_{\mathrm{O,Ga}})$, the adjustable part of the 2NN $(\eta_{\mathrm{O,P}})$, and the bond-stretching part, $(\delta A' + 2\delta B') = 3\delta A'$, of the back-bond interactions $(\eta_{\mathrm{Ga,P}})$. Other interactions within the cluster are then modified automatically according to rotation invariance.

The defect perturbation matrix is now $\underline{\delta L} = \underline{M}^{-1/2}\underline{\delta\Phi}\,\underline{M}^{-1/2} - \Omega^2\underline{M}^{-1}\underline{\delta M}$, where $\underline{\delta M}$ contains the on-site mass perturbation. (Note that we have put the explicit mass dependence in $\underline{\delta L}$ rather than $\underline{G}^0$, as in Ref. 7.) We numerically construct $\underline{\delta L}$ in Cartesian form, using Eqs. 5.11–5.15 and 5.17–5.24, and then we apply the transformation to collective coordinates. The local modes are given by Eq. 5.7. We solve such a determinantal equation *separately* for *each* irreducible representation. A useful quantity for characterizing the relative amount of motion of a certain symmetry type and at a particular frequency is the LDOS. The LDOS of the perturbed crystal is given by the diagonal elements of $-\mathrm{Im}(\underline{G}/\pi)$, where the perturbed crystal Green's function is given by[5]

$$\underline{G} = (\underline{1} - \underline{G}^0\underline{\delta L})^{-1}\underline{G}^0. \tag{5.25}$$

Strong resonant modes show up ideally as Lorentzian peaks in the LDOS while strictly localized modes appear as delta functions.[7] For representations which have multiple occurrences in the space of the defect, the LDOS for each occurrence provides a measure of the fraction of total kinetic energy in that particular type of motion, and this quantity is used in the following sections to describe the composition of defect modes.

5.3.4 Results

Restoring Force for One-Atom Displacements

In the discussion to follow, it will be useful to know the effective spring constant acting on the O atom under a unit displacement from equilibrium. This spring constant is given by $k^{(eff)} = k^{(sr)\prime} + k^{(el)}$, where

$$k^{(sr)\prime} = -4\left[(A+\delta A) + 2(C_2+\delta C) + (F_2+\delta F)\right] \tag{5.26}$$

denotes the modified short range part, and $k^{(el)}$ is the bulk, P-site, Coulomb component. We wish to compute $k^{(el)}$. In the notation of Maradudin,[12] we have,

for the unit displacement $u_\beta(l'\kappa') = \delta_{\alpha\beta}\delta_{ll'}\delta_{\kappa\kappa'}$

$$\begin{aligned} F_\alpha(l\kappa) &= -\sum_{\beta l'\kappa'} \Phi_{\alpha\beta}(l\kappa;l'\kappa')\delta_{\alpha\beta}\delta_{ll'}\delta_{\kappa\kappa'} \\ &= -\Phi_{\alpha\alpha}(l\kappa;l\kappa), \end{aligned} \tag{5.27}$$

where $\Phi_{\alpha\alpha}(l\kappa;l\kappa) = k^{(el)}$ for $\underline{\Phi} = \underline{\Phi}^{(el)}$. Now, using the identity

$$\delta_{ll'} = \frac{1}{N}\sum_{\mathbf{k}} \exp\{-i\mathbf{k}\cdot[\mathbf{x}(l\kappa) - \mathbf{x}(l'\kappa)]\} \tag{5.28}$$

in Eq. 5.27 gives

$$\begin{aligned} F_\alpha(l\kappa) &= -\frac{1}{N}\sum_{\mathbf{k}}\left(\sum_{l'}\Phi_{\alpha\alpha}(l\kappa;l'\kappa)\times\exp\{-i\mathbf{k}\cdot[\mathbf{x}(l\kappa)-\mathbf{x}(l'\kappa)]\}\right) \\ &= -\frac{1}{N}\sum_{\mathbf{k}}\hat{\Phi}_{\alpha\alpha}(\kappa\kappa|\mathbf{k}), \end{aligned} \tag{5.29}$$

where the inner sum has been identified as the Fourier-transformed force constant matrix. Recognizing that our $\hat{\underline{\Phi}}(\kappa\kappa|\mathbf{k})$ transforms as $T_2 \otimes T_2$, we can restrict the summation over $\mathbf{k}$ in Eq. 5.29 to the irreducible wedge $\mathcal{W}$ as in Ref. 7, obtaining finally

$$k^{(el)} = \frac{1}{3N}\sum_{\mathbf{k}\in\mathcal{W}} h_{\mathbf{k}} \sum_{\beta=1}^{3} \mathrm{Re}\left[\hat{\Phi}^{(el)}_{\beta\beta}(\kappa\kappa|\mathbf{k})\right]. \tag{5.30}$$

This sum was evaluated numerically by a method similar to the numerical computation of $\mathrm{Im}\,\underline{G}^0$ in this and in our previous[7] work. The result is $k^{(el)} = -0.177k^{(sr)}$ for the bulk, P-site.

Bulk Phonons

In Fig. 5.1, we show the bulk dispersion curves and bulk DOS for GaP as computed from the DBA. The DBA results constitute a substantial quantitative improvement over those of the two-parameter model used previously.[7] The calculated LO_Γ and TO_Γ energies, 50.0 meV and 45.3 meV, respectively, are split due to Coulomb effects and are in agreement with measured values.[8] The calculated TA_X energy is 12.8 meV compared with the actual value, 13.1 meV.[14] The bulk DOS

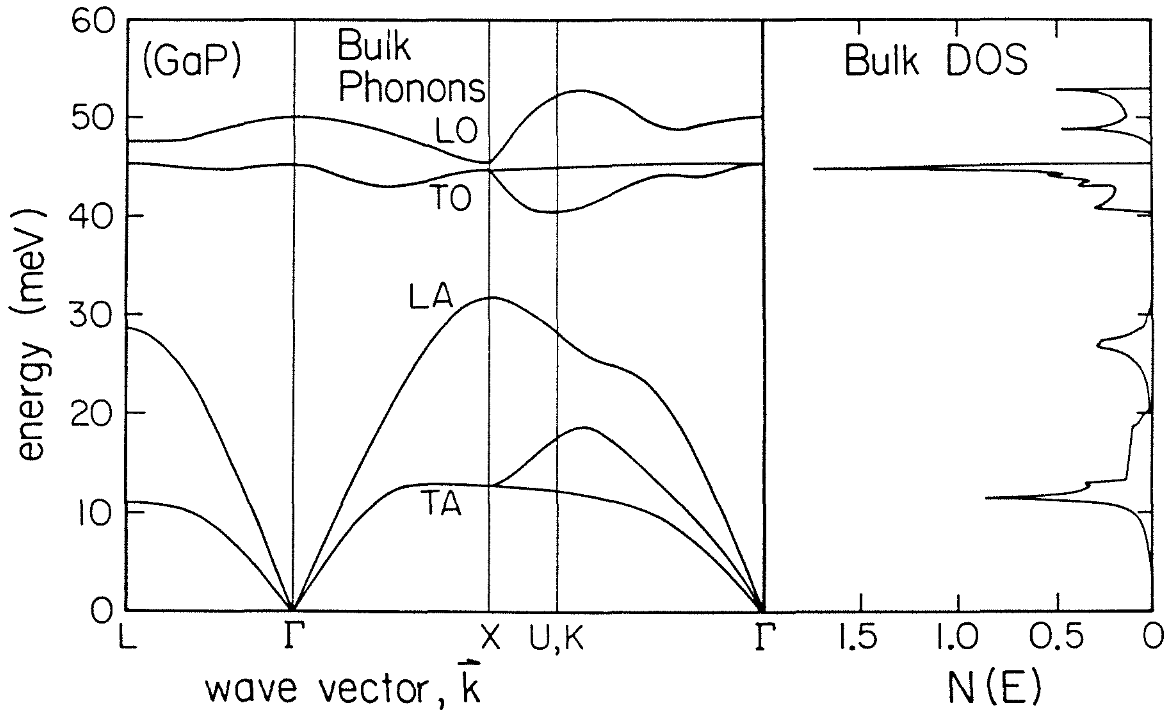


Figure 5.1: The bulk phonon dispersion curves and density-of-states for GaP calculated from the deformable-bond approximation. The calculated LO_{Γ}, TO_{Γ}, and TA_{X} energies are 50, 45.3, and 12.8 meV, respectively. The peaks in the bulk DOS lie at 11.3, 27, 44.8, 48.8, and 52.9 meV. In the units shown, the DOS integrates to 6.0004, compared with the exact value, 6.

presented here is highly structured compared with that of the two-parameter model.[7] As we shall see, this structure gives rise to local mode behavior not seen in the simpler model. In the units shown in the figure, the computed area under the bulk DOS curve is 6.0004, compared to the exact theoretical value, 6.

Bulk Green's Functions

We have computed the bulk Green's function matrix, $\underline{G}^0$, for both a P-centered and a Ga-centered 17-atom cluster. The method of computation is described in Ref. 7. The Green's functions were calculated as functions of $(\hbar\Omega)^2$ at intervals of $2.5\,\mathrm{meV}^2$ and then converted to functions of $E = \hbar\Omega$ (see Eq. 9 of Ref. 7). This corresponds to a resolution in E of about $0.5\,\mathrm{meV}$ near $E = 0$ and $0.025\,\mathrm{meV}$ at $E = 50\,\mathrm{meV}$. There are 68 bulk Green's functions each for both the P- and Ga-centered clusters—too many to display individually. Instead, for each irreducible representation (s), we sum the diagonal matrix elements and present $\mathcal{D}^{(s)}(E)$ vs. E in Fig. 5.2, where

$$\mathcal{D}^{(s)}(E) \equiv -\frac{1}{\pi}\,\mathrm{Tr}\,\mathrm{Im}\left[\underline{G}^{0^{(s)}}(E)\right]. \tag{5.31}$$

The quantity $\mathcal{D}^{(s)}(E)$ is the partial DOS[7] for the representation (s) and is proportional to the mean-squared amplitude of modes associated with (s)-type motion in the 17-atom cluster at energy E. In particular, the A_2 mode consists of 2NN motion only. From Fig. 5.2 we see that the P-centered, A_2 partial DOS has a large amplitude in the TO and small amplitude in the TA phonon branches while the opposite is true for the Ga-centered DOS. This is because the P atoms undergo larger displacements in optical modes than do the heavier Ga atoms. For the discussion to follow, the important symmetries are A_1 and T_2. A_1 modes consist of some combination of the following: "breathing" motion of the NN shell, 2NN breathing, and a third, more complicated 2NN motion. T_2 modes consist of defect-site motion along with collective motions within the NN and 2NN shells.

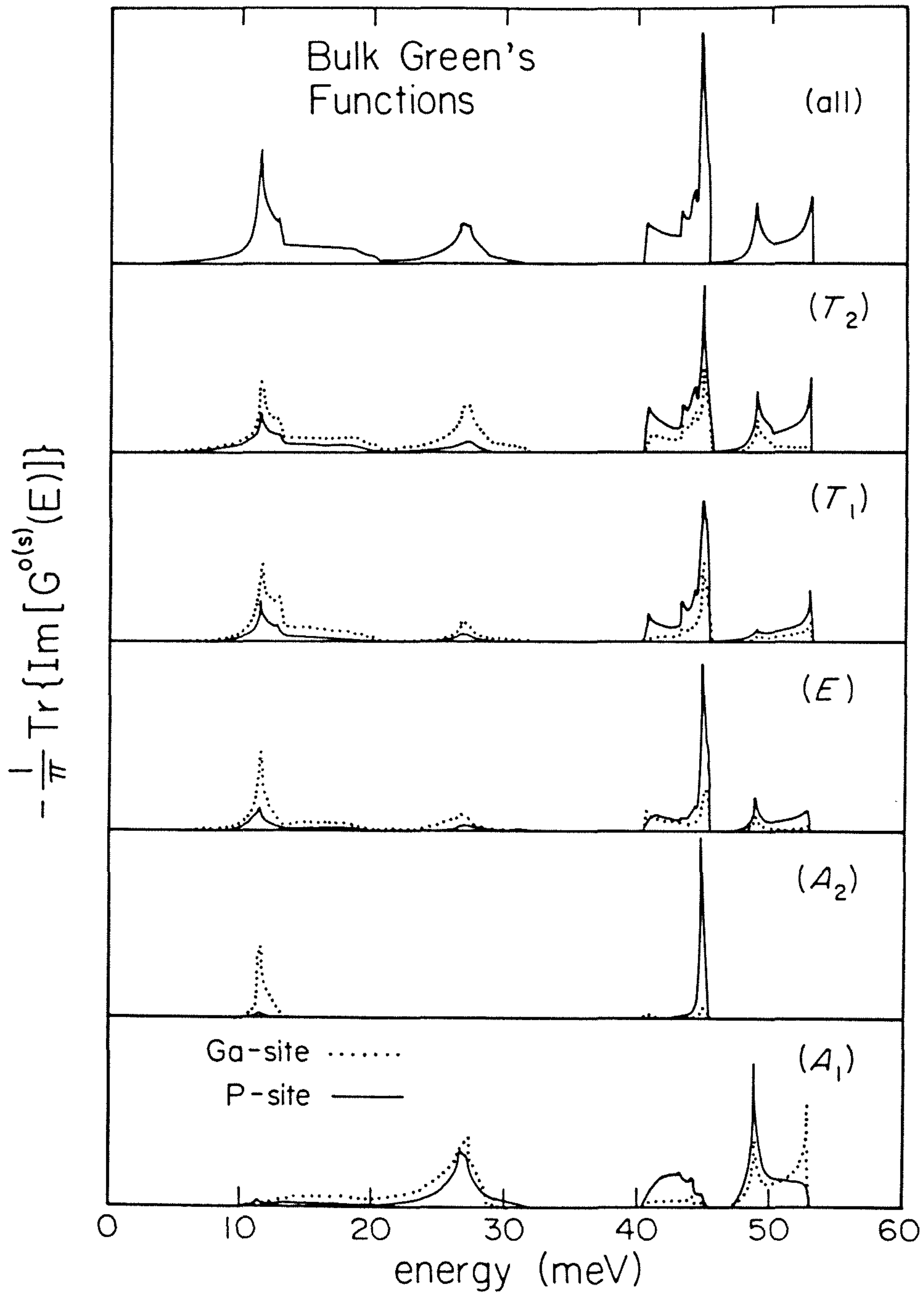


Figure 5.2: The partial density-of-states of bulk GaP for the representations, A_1–T_2. Green's functions for P- and Ga-centered clusters are shown. The A_2 partial DOS involves 2NN motion only and for the P-centered (Ga-centered) cluster is larger at optical (acoustic) frequencies.

Only T_2 modes can involve defect-site motion. (See the Appendices in Refs. 7 and 13 for a complete listing of the collective coordinates of the 17-atom cluster.)

Figure 5.2 gives the decomposition of the bulk DOS into the various representations of T_d appropriate for a point-defect. The bulk DOS can be recovered from the various partial densities by performing the sum

$$\sum_{(s)=A_1}^{T_2} d_{(s)} \left[\mathcal{D}_{\mathrm{P}}^{(s)}(E) + \mathcal{D}_{\mathrm{Ga}}^{(s)}(E) \right] = 17N(E), \tag{5.32}$$

where $d_{(s)}$ is the dimensionality of the representation (s). The renormalization factor of 17 on the right side of Eq. 5.32 occurs because, in summing over both P- and Ga-centered clusters, the two-atom basis of GaP is included 17 times.

Local Modes

Let us examine the local phonon modes of GaP:O_P as a function of the interaction perturbation. In Section 5.3.3, we have parametrized the interaction perturbation for GaP:O_P in terms of $\eta_{\mathrm{O,Ga}}$, $\eta_{\mathrm{O,P}}$, and $\eta_{\mathrm{Ga,P}}$. Now, we feel physically motivated to require a further constraint between $\eta_{\mathrm{O,Ga}}$ and $\eta_{\mathrm{O,P}}$. It is natural to expect the O-P interaction to vanish when the O-Ga interaction vanishes and the O-P interaction to have the bulk (P-P) value when the O-Ga interaction has the bulk (P-Ga) value. In this way we are led to consider defects for which $\eta_{\mathrm{O,Ga}} = \eta_{\mathrm{O,P}} = \eta$. In contrast, we expect the relation between the back bonds ($\eta_{\mathrm{Ga,P}}$) and our newly defined η to depend on details of the defect chemistry well beyond the scope of this paper. In practice, we have found that the effect of the parameter $\eta_{\mathrm{Ga,P}}$ is to "turn on" certain 2NN-shell local modes while not appreciably affecting those low-shell modes which have a strong dependence on η. Therefore, for the purpose of presentation, we *choose* (see Section 5.3.5) $\eta_{\mathrm{Ga,P}} = 0.12\eta$ and present our results as a function of the single parameter, η.

The local modes of GaP:O_P as a function of the perturbing force constant

parameter η are shown in Fig. 5.3. Only A_1 and T_2 modes satisfying Eq. 5.7 are shown here. Modes of A_2, E, and T_1 symmetry also occur but are omitted for clarity. η ranges from -1, corresponding to removal of the *short-range* O-Ga and O-P interactions, through 0, where all interactions have their bulk values, to $+1$. The effective O-site restoring force, $k^{(\mathrm{eff})} = (1+\eta-0.177)/(1-0.177)$, normalized to the bulk P-site value, is indicated along the top of the figure while the bulk DOS adjoins on the left. Note that the shaded region below the value $k^{(\mathrm{eff})} = 0$ is unphysical as this corresponds to a dynamically unstable configuration.

The important features of Fig. 5.3 are the steep T_2 branch in the energy range near 12–40 meV, the lowest-lying T_2 branches (around 10 meV), the low-energy A_1 branch (20–35 meV), and the high-energy A_1 branch (45–55 meV). The steep T_2 branch was seen in our earlier model[7] and consists mainly of O-atom motion. Also seen in our earlier work was the low-energy A_1 branch. In the present work, this A_1 branch, consisting over 95% of NN breathing, has been pulled up to higher energies ($\eta < 0$) in better agreement with experiment, as we shall see below. Two significant features appear in the present work not obtained previously. One is the high-energy A_1 branch, which appears only upon perturbation of the Ga-P back bonds and involves mostly 2NN breathing motion. The other is the low-energy T_2 branch. This T_2 branch is primarily composed of Ga-shell, bond-stretching motion ($E \approx 6$ meV) and seems to appear for $\eta < 0$ only when the bulk lattice dynamical model includes 2NN interactions.

In Section 5.3.5, we shall see that the O^0 defect is well described by the perturbation, $\eta = -0.62$. This corresponds to a $k^{(\mathrm{eff})}$ of 25% of the bulk value. The local mode energies at $\eta = -0.62$ are indicated in the caption of Fig. 5.3. Next, in Fig. 5.4, we compare the A_1 bulk and defect partial DOS for $\eta = -0.62$. These are decomposed into their Ga- and P-shell parts, showing the relative energy of motion associated with each shell at a given mode "energy," $E = \hbar\Omega$.

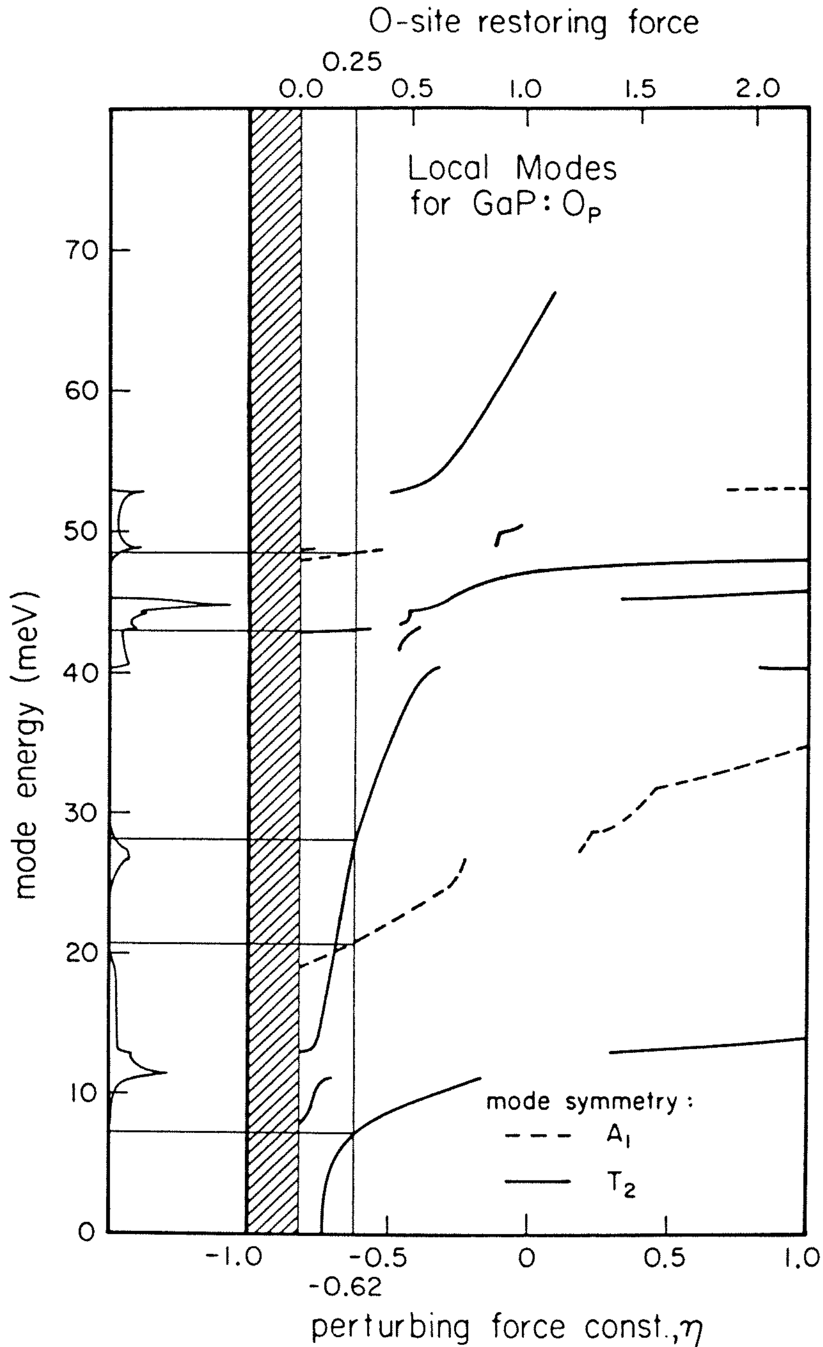


Figure 5.3: The A_1 and T_2 local mode energies for GaP:O_P as a function of the short-range force constant perturbation, η, discussed in the text. For $\eta = -0.62$ the A_1-mode energies are 20.4 and 48.2 and the T_2-mode energies 7.3, 28.2, and 43.1 meV. The shaded region is unphysical. The density-of-states for bulk GaP adjoins on the left.

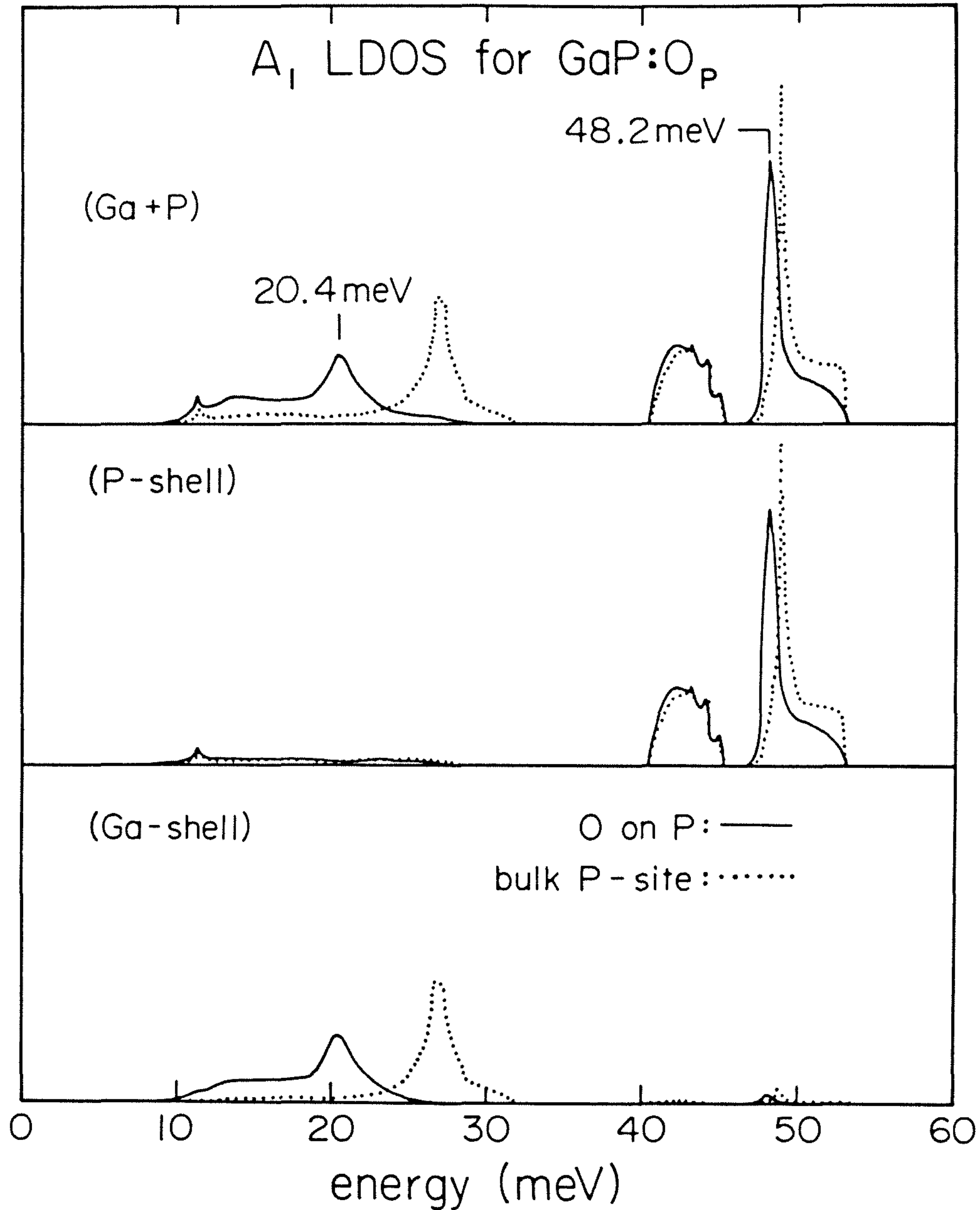


Figure 5.4: The A_1 local density-of-states for GaP:O_P decomposed into Ga- and P-shell components. The corresponding bulk partial densities-of-states are included for comparison. The 20.4 meV resonance involves mostly Ga-shell motion while the resonance at 48.2 meV consists mainly of P-shell motion. All curves have the same scale.

From the top curve in Fig. 5.4 we see resonant modes at 20.4 and 48.2 meV. The 20.4 meV mode is a rather broad resonance of width $\approx$ 5 meV and involves mainly NN Ga motion (96%) as mentioned above. The resonant mode at 48.2 meV is sharper ($\approx$ 0.5 meV) and consists 97% of motion of the less massive P atoms. It is evident from the figure that both resonant modes occur at frequencies appreciably different from the bulk peaks and hence are characteristic of the defect. In contrast, local modes confined near singular features in the bulk partial density of states may also occur, in general. Such modes evidently couple strongly to bulk modes and are thereby more "bulklike," *i.e.*, more spatially delocalized in character.

We show the T_2 LDOS for GaP:O_P at $\eta = -0.62$ in Fig. 5.5, once again decomposed into contributions from O-, Ga-, and P-shells. The mode energies which solve Eq. 5.7 are indicated on the top curve. Twin resonance peaks at 25.0 and 28.2 meV are seen in the O-site LDOS, occurring just above and just below the 27 meV LA peak in the bulk DOS. The 28.2 meV mode consists of 86% O-, 8% Ga-, and 6% P-shell motion and the 25.0 meV mode consists of 70% O-, 20% Ga-, and 10% P-shell motion, within the 17-atom cluster. We note that at $\eta = -0.62$ only the mode at 28.2 meV is a solution of Eq. 5.7. The twin-peak structure appears to be the result of coupling between isolated-defect and nearby bulk-LA phonon modes. The 27 meV LA peak in the bulk DOS involves very little P-sublattice motion. Thus, a mode involving large O motion cannot exist at 27 meV, and when $\eta \to -0.62$ the O mode splits into two modes. Both O-site peaks occur at energies different from bulk peaks and are well localized near the O atom, and upon the substitution of ^{18}O for ^{16}O, both peak positions undergo a shift in energy. The 25.0 and 28.2 meV resonances are, therefore, local rather than bulklike in character. In contrast, the weak mode at 43.1 meV is pinned to a van Hove singularity in the bulk DOS (see Fig. 5.3) and is evidently

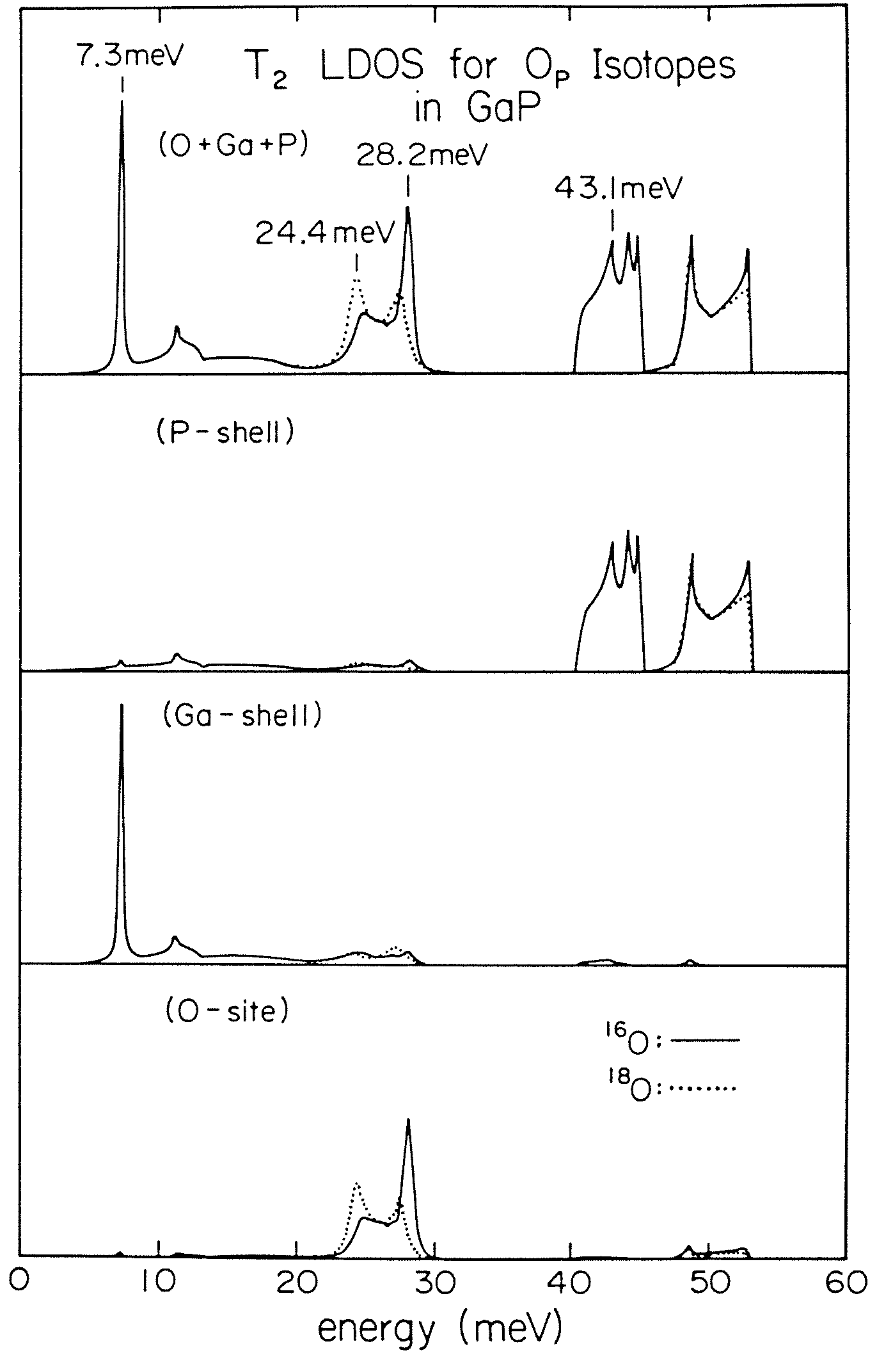


Figure 5.5: The T_2 local density-of-states of GaP:O_P for isotopes, ^{16}O and ^{18}O, decomposed into O-atom, NN-shell, and 2NN-shell parts. The twin resonances centered near 25 meV primarily involve O-atom motion and undergo isotope shifts. Other resonant modes occur at 7.3 meV (Ga-shell) and 43.1 meV (bulklike P-shell). All plots are on the same scale.

bulklike in character. Finally, the sharp resonance at 7.3 meV primarily involves Ga-shell motion (94%) along with small P (5%) and O (1%) contributions.

5.3.5 Discussion

Intensity of phonon replicas

The intensity of phonon-assisted optical transitions at defects have been thoroughly discussed by Rebane.[15] Selection rules for determining which transitions are allowed are similar to those used for molecules, as discussed by Herzberg.[16] Allowed transitions are those for which the dipole matrix element

$$\langle\, \Psi_m(r,R)\Phi_{mj}(R) \,|\, D_e(r) \,|\, \Psi_l(r,R)\Phi_{li}(R) \,\rangle$$

is nonzero. Here, Ψ is the electronic wave function, Φ is the vibrational wave function, l and m label electronic states, i and j label vibrational states, r is the set of all electron coordinates, R is the set of all nuclear coordinates, and $D_e(r) = -e\sum_i r_i$ is the electronic dipole moment operator. In place of the nuclear coordinates R we introduce new coordinates Q_{l_s} which are the normal coordinates of the vibrational problem. These normal coordinates may be chosen to form bases for irreducible representations $\Gamma(Q_{l_s})$ of the symmetry group of the imperfect crystal, where $\Gamma(x)$ refers to the representation according to which x transforms. We assume that the same set of normal coordinates may be used for both the initial and final electronic states (though the equilibrium values of these coordinates will depend on the electronic state). The total vibrational wave function can now be written as

$$\Phi_{l\{i_s\}} = \prod_s \phi_{i_s}(Q_s), \tag{5.33}$$

where i_s labels the occupation of the sth normal coordinate. The harmonic oscillator functions ϕ are even with respect to their arguments for even occupation,

and odd for odd occupation. By inspection, in the vibrational ground state ($i_s = 0$ for all s) the wave function $\Phi_{l\{i_s\}}$ is totally symmetric. For multiple one-phonon occupation, Φ transforms as $\prod_{\{i_s=1\}} \Gamma(Q_s)$, where the product runs over all s such that $i_s = 1$.

As an example of the above discussion, consider a defect with T_d symmetry, in which case the dipole moment transforms as T_2. An $A_1 \rightarrow A_1$ electronic transition is forbidden, but if one T_2 phonon is involved, the dipole matrix element transforms as $A_1 \otimes A_1 \otimes T_2 \otimes A_1 \otimes T_2 = A_1 \oplus E \oplus T_1 \oplus T_2$, which is allowed. Similarly, an $A_1 \rightarrow T_2$ electronic transition is allowed, and participation of a single A_1 phonon or a single T_2 phonon is also allowed since $A_1 \otimes A_1 \otimes T_2 \otimes T_2 \otimes A_1$ and $A_1 \otimes A_1 \otimes T_2 \otimes T_2 \otimes T_2$ both contain A_1. The intensity of phonon replicas is governed by the Franck-Condon principle.[15,16] Essentially, this principle states that the phonon replicas which will be observed in a transition are those which describe the change in equilibrium position of the atoms between the initial and final states. In other words, if the atoms move along some coordinate Q between their initial and final states, normal modes of the type Q will be strongly excited in the transition.

We can use this principle to demonstrate that "antisymmetric," *i.e.*, not totally symmetric, phonons cannot be strongly excited. If some antisymmetric vibrational mode is strongly excited, this necessarily leads to a change in the symmetry of the system, and in the new symmetry group the vibrational mode is totally symmetric. However, for weak excitation of the antisymmetric vibrations, this change in symmetry can be neglected. Thus, in electronically allowed transitions, totally symmetric vibrations will predominate, and in electronically forbidden transitions the occupation of the dominant antisymmetric modes changes by one.

Comparison of theory and experiment

In Table 5.1 we compare our theoretical results with the phonon energies observed in the optical spectra of the O^+ and O^0 defects. The apparent agreement between theory and experiment is achieved partly by matching unknown theoretical parameters with experiment, and partly by the predictive powers of the theory itself. If we view the bulk phonon dispersion curves as input to the theory, then the only unknown is the interaction perturbation within the 17-atom cluster. We have parametrized the interaction perturbation (Eqs. 5.22–5.24 in Section 5.3.3) and have invoked the further restriction, $\eta_{O,Ga} = \eta_{O,P} = \eta$ in Section 5.3.4, so that the total number of unknown parameters is two. We determine values for these parameters by matching the energies of theoretical modes with those observed. The theory then predicts a number of other vibrational modes not involved in the matching process with which we identify observed phonons. In this way we determine the symmetry and character of almost all of the phonon modes associated with the O^+ and O^0 defects. Strictly speaking, all of the phonons observed in the optical spectra involve some component of defect motion (*i.e.*, O-atom, nearest Ga-shell, or nearest P-shell motion). However, certain phonons involve much more defect motion than others, and it is useful to separate the modes into those which are associated largely with the defect and those which are bulklike. In Table 5.1, we have classified the modes as such. The theoretical defect mode energies are taken from peaks in the LDOS, and the theoretical bulk phonon energies are taken from peaks in the bulk DOS or critical points in the dispersion curves.

A set of very interesting phonon sidebands associated with the neutral oxygen defect, O^0, occur in the capture luminescence of Dean and Henry.[9] Here, intense phonon resonances near 25 meV are observed to involve significant O motion since they shift under the isotopic substitution, $^{16}O \rightarrow ^{18}O$, and must therefore be T_2

Table 5.1: Energies (meV) of the phonons associated with the O^0 and O^+ defect in GaP. $^{16}O \rightarrow {}^{18}O$ isotope shifts (meV) are given in parentheses following the energies. Tentative identifications are indicated by "?".

Label	Expt.	Theory	Identification
Zn–O	6.0[a]	7.3	defect (T_2, Ga shell)
TA	13.1[b]	11.5	bulk (TA)
a_1(Ga)	19.5,[c] 19[d]	20.4	defect (A_1, Ga shell)
loc	24.7 (−1.6)[b]	25.0 (−0.6)	defect (T_2, O atom)
loc′	28.4 (−0.5)[b]	28.2 (−0.7)	defect (T_2, O atom)
A	43.0[b]	43.1	defect (T_2, P shell) ?
TO	44.8[b]	44.8	bulk (TO)
...	...	45.3	defect (E, P shell)
...	...	45.4	defect (T_1, P shell)
...	...	45.8	defect (A_2, P shell)
B	46.1,[b] 46.4,[c] 46.5[e]	48.8	defect (T_2, P shell) ?
a_1(P)	47.5,[b] 47,[f] 48[d]	48.2	defect (A_1, P shell)
C	48.7,[b] 48.8[c]	48.8	bulk (LO)
LO^{Γ}	49.8[b]	50.0	bulk (LO_{Γ})

a. Reference 17.

b. Reference 9, capture luminescence spectrum.

c. Reference 9, luminescence excitation spectrum.

d. Reference 11.

e. Reference 18.

f. Reference 10.

modes. Since the O mass is less than that of P, these low energy resonances imply reduced O-neighbor force constants relative to the bulk force constants. This force constant reduction was first estimated by Feenstra and McGill[18] to be 30%, and later estimated by Baraff, Kane, and Schlüter[19] to be 20%, both estimates based on simple molecular-type vibrational models. Subsequent NN Green's function results[7] yield an estimate of 15%. In the present model, if we assume the same fractional reduction for O-Ga and O-P force constants, the best value for the purely short range force constant reduction is $\eta = -0.62$. After including Coulomb effects between the O atom and the rest of the lattice, this corresponds to a total effective force constant ($k^{(\mathrm{eff})}$) of $3.5\,\mathrm{eV}/\,\mathrm{\AA}^2$, which is 25% of the bulk, P-site value. We emphasize, though, that $k^{(\mathrm{eff})}$ includes long-range Coulomb effects from the entire lattice and must *not* be viewed simply as an equivalent short range force constant. Our LDOS for T_2 modes, summed over the O-atom, Ga-shell, and P-shell, is shown in Fig. 5.5. The theoretical peak energies and isotope shifts, listed in Table 5.1, compare reasonably well with experiment.

Having determined the O-neighbor force constants, the theory predicts an A_1 mode at 20.4 meV. We identify this mode with the observed 19 meV phonon listed in Table 5.1. As discussed by Morgan[20] and also in Section IV A, the observed intensities of this 19 meV mode and the prominent 47 meV mode indicate that they have A_1 symmetry. To produce a defect mode of A_1 symmetry near 47 meV, it is necessary to reduce the back-bond Ga-P force constants from their bulk values. We arbitrarily choose to reduce the Ga-P force constants by 12% of the O-neighbor reduction, resulting in $\eta_{\mathrm{Ga,P}} = -0.0744$ for the Ga-P interactions and yielding an A_1 mode at 48.2 meV. Further reductions in the Ga-P force constants reduce the energy of this mode, bringing it closer to the observed value. However, considering the uncertainties associated with our calculated optical phonons,

we feel that an accurate determination of these back-bond Ga-P interactions is outside the range of our theory. Suffice it to say that our computations do yield an A_1 mode in the optical branch and that the back-bond Ga-P interactions are reduced from their bulk values.

A number of phonon replicas in the optical branch are seen in the optical spectra of Dean and Henry.[9] We identify the symmetry and nature of these phonons as indicated in Table 5.1. Some of these optical phonon replicas are weak, and considering the uncertainty of our optical phonon computations, those identifications we feel should be regarded as tentative are marked by "?". The a_1(P) phonon has been discussed above. The phonon labeled B we take to be a defect mode since its energy falls well away from peaks in the bulk DOS. This phonon is observed in capture luminescence and luminescence excitation spectra,[9] and is also seen in the near-neighbor (m=4 and m=5) (Zn,O) pair spectrum[18] (labeled O in that spectrum). Considering the proximity of the donor-acceptor impurities in the latter case, it seems that this 46.5 meV mode corresponds to atomic motion which is directed along the donor-acceptor axis, *i.e.*, a T_2 mode with respect to the T_d symmetry of an isolated O impurity. The 46.1 meV B-phonon in the capture luminescence spectrum most likely has T_2 symmetry,[20] although the 46.4 meV mode seen in luminescence excitation could transform as either A_1 or T_2, as discussed in Section IV A. For now, we group all three of these modes together with the label B in Table 5.1, and we tentatively identify them with a computed T_2 resonance at 48.8 meV which is strongly tied to a LO peak in the bulk DOS (see Figs. 5.3 and 5.5).

Another T_2 defect mode occurs in the theory at 43.1 meV, and this mode we tentatively associate with the observed 43.0 meV phonon labeled A. The 49.8 meV phonon identified by Dean and Henry as LO_Γ is in agreement with our computed 50.0 meV LO_Γ phonon energy. The mode labeled C we associate with the promi-

nent peak in the bulk DOS occurring at 48.8 meV while we associate the observed 44.8 meV mode with the very prominent TO peak in the bulk DOS occurring at 44.8 meV. In both of these cases, the peaks in the bulk DOS do not correspond to phonons at or near critical points in the Brillouin zone but rather are associated with phonons in the interior of the zone. Finally, with regard to defect modes in the optical branch, we note that the present theory neglects defect-induced electrostatic interactions, and such effects are known to cause vibrational modes split off from the LO branch.[21]

Now, we come to the entry in Table 5.1 labeled Zn-O. This 6.0 meV mode is just barely resolved in the NN Zn-O pair spectrum of Henry, Dean, and Cuthbert[17] (a related mode[18] is seen at 9.2 meV for 2NN Zn-O pairs). Although the computational results presented here apply to the isolated O defect, simply replacing a NN Ga with a Zn atom changes the phonon energies by less than 0.1 meV since Ga and Zn have almost the same mass. Further force constant variations of the Zn-O, Zn-P, or Zn-Ga interactions could affect the mode energies, but we assume that these variations are not significant. The presence of the Zn atom will, however, drastically affect the selection rules for observing phonons. In particular, all modes which are T_2 for isolated O are now split to $A_1 \oplus E$, and strong phonon replicas of the allowed electronic transition transform as A_1. Thus, the low energy T_2 mode in Fig. 5.5 should appear in the Zn-O spectrum, and we identify it with the observed 6.0 meV mode. The theoretical "eigenvector" for this mode is shown in Fig. 5.6. We plot the displacement eigenvector, $\underline{u} = \underline{M}^{-1/2}\underline{\xi}$, satisfying the eigenvalue equation

$$\left[\underline{1} - \underline{G}^0(\Omega^2)\underline{\delta L}(\Omega^2)\right]\underline{\xi} = \lambda(\Omega^2)\underline{\xi} \tag{5.34}$$

at the resonant frequency, $\hbar\Omega = 7.3\,\text{meV}$, for which $\text{Re}\,\lambda = 0$.* Now, the actual

*The condition, $\text{Re}\,\lambda = 0$, is an alternative to Eq. 5.7 for identifying resonant modes. (See Ref. 5, pp. 386ff.)

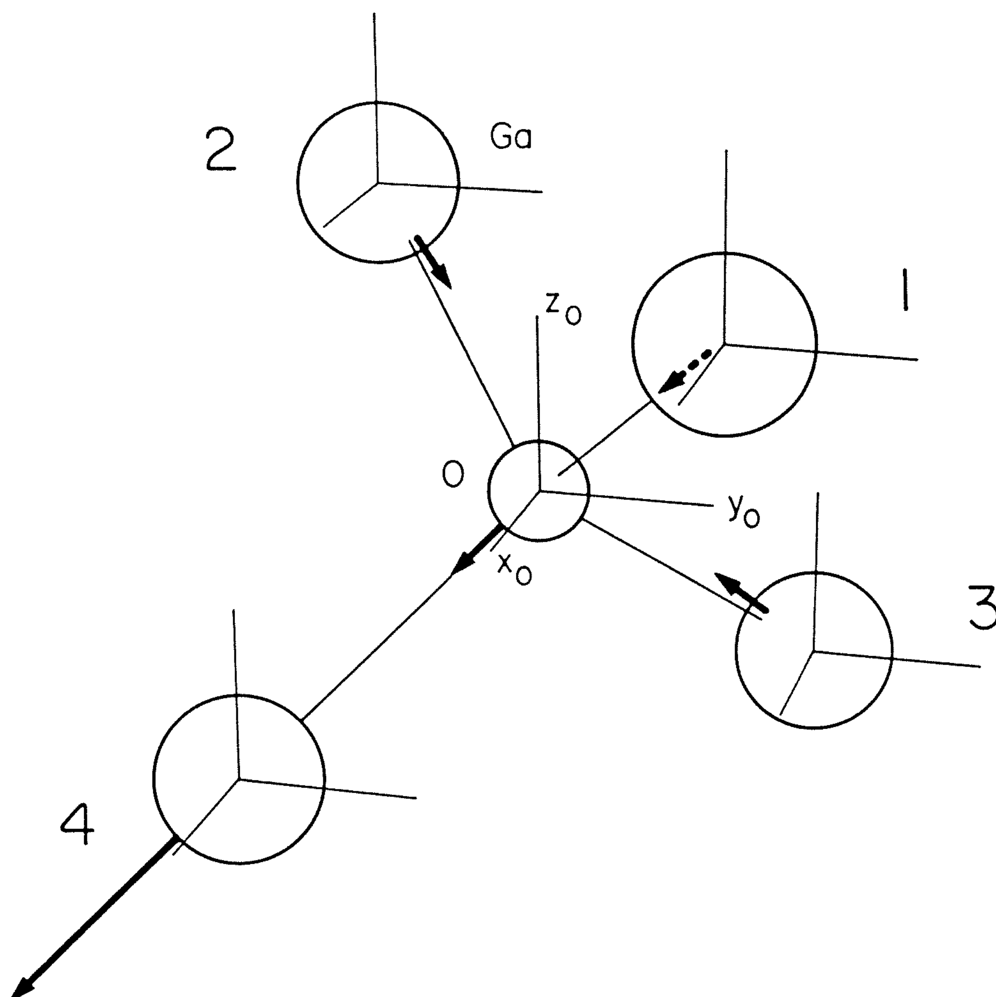


Figure 5.6: The "eigenvector" for the 7.3 meV T_2 mode. The atomic displacements are indicated for the O and its four Ga neighbors. The $Ga_{(4)}$ and O displacements (in the ratio 3.4 : 1) both lie along the bond direction. The eigenvector has C_{3v} symmetry relative to the O-$Ga_{(4)}$ bond. With Zn in place of the $Ga_{(4)}$ atom, this is the predominant motion involved in the dissociation of Zn-O pairs (see text).

defect normal mode, given by $(\underline{1} - \underline{G}^0 \underline{\delta L})\underline{\xi} = 0$, requires a complex value of Ω for its solution. However, since $\mathrm{Im}\,\lambda$ is proportional to the resonance width, we expect, for the very sharp 7.3 meV mode, that Ω is near the real axis so that Eq. 5.34 gives a reasonable representation of the atomic motion in this case. It has been argued that this particular resonant mode plays an essential role in the dissociation of Zn-O pairs.[18] In particular, the atomic motions which occur during the early stages of the dissociation may be as shown in Fig. 5.6. It is seen that the amplitude of Zn-atom motion is roughly three times that of oxygen. Thus, we conclude that in the pair dissociation it is the Zn atom which jumps to an interstitial site and diffuses away.

The vibrational modes of NN Cd-O pairs have been observed in optical spectra,[17,22,23] and can be qualitatively understood on the basis of our calculations. The Cd-O modes will be similar to the Zn-O modes, except that those vibrations involving significant Cd motion will be shifted down in energy. Replicas of the electronically allowed A line should have A_1 symmetry for the point-group C_{3v}, *i.e.*, the A_1 modes and the A_1 component of the T_2 modes for isolated oxygen.[22] In emission, Henry, Dean, Thomas, and Hopfield observe local modes at 7.0, 47.3, and 49.7 meV.[23] The 7.0 meV mode is similar to the 6.0 meV Zn-O mode discussed above. Thus, this mode involves mainly Cd motion in agreement with the observed isotope shift.[17] The optical resonances at 47.3 and 49.7 meV are probably associated with the defect B and a_1(P) modes, thus involving mainly P-shell motion. From the intensities of the observed phonon replicas we conclude that in the optical transition the major atomic motion is that of the Cd-atom moving relative to the almost stationary O-atom. A number of Γ_3 phonons are reported[23] which produce replicas of the forbidden B line. Theoretically, we expect these to arise from the E modes and the E components of the T_1 and T_2 modes for isolated oxygen. As shown in Table 5.1, we do compute various E

and T_1 modes. All these modes are very sharp (width $\approx 0.1\,\mathrm{meV}$) and are tied to singularities in the bulk DOS. These modes could account for some of the observed Γ_3 modes, and the E components of the T_2 mode at 43.1 meV may also contribute. Our optical phonon energies really are not accurate enough to make a definitive identification of these observed Γ_3 phonons.

Now, let us consider what our calculations tell us about the nature of the oxygen defect. In particular, two opposing models[24,25] presently exist for this defect, and it is desirable to try to choose between them based on our results. First, the defect we are considering in our calculations is only a model, without the same detailed properties of the actual oxygen defect although their general properties should be similar. The most important result from our calculations is the O-lattice effective force constant, $3.5\,\mathrm{eV}/\mathrm{\AA}^2$. This number is accurate to within, perhaps, 10% and so provides a test of the opposing models. Unfortunately, it is very difficult to accurately obtain the defect force constant from the existing theoretical models. To date, Baraff *et al.*[19] have attempted such a calculation for the classical model[26,25] and have obtained an O-Ga stretching constant of $4.4\,\mathrm{eV}/\mathrm{\AA}^2$, corresponding to an effective force constant of $4(4.4)/3 = 5.9\,\mathrm{eV}/\mathrm{\AA}^2$, a result which is somewhat too large. Nobody has yet attempted a similar computation for the weak-bonding model,[24] although the computed bond strength should be significantly different for the two models as they contain different numbers of bonding and antibonding electrons.[27] Morgan has recently argued that the observed structure near 25 meV in the O^0 phonon sidebands is due to a dynamic Jahn-Teller distortion of the O^0 defect.[27] Although such a distortion could occur, our calculations demonstrate that it is not necessary to invoke this mechanism in order to account for the observed structure. With one free parameter (η), we have obtained quantitative agreement between experiment and theory for those phonon energies, isotope shifts, and intensities.

One basis for the weak-bonding model[24] is the 19 meV a_1(Ga) phonon, present in donor-acceptor pair (DAP) spectra,[10,11] but apparently absent in photoluminescence excitation (PLE) spectra.[9] According to our calculations, this mode is rather broad as seen in the Ga-shell A_1 LDOS for O on a P-site, shown in Fig. 5.4. Clearly, this resonance is distributed throughout the acoustic branch, so let us take its "width" to be 5 meV. This width has no consequence in the DAP spectrum due to the more than 10 meV broadening caused by a distribution of pair separations. But in the PLE spectrum, the no-phonon line is relatively sharp, so that the ratio of peak-heights between an a_1(Ga)-phonon-assisted line and no-phonon line is significantly reduced from the DAP spectrum. Additional broadening for the phonon-assisted line is also possible; the 20.4 meV Ga-shell peak in the A_1 LDOS for O on a P-site shown in Fig. 5.4 represents mainly NN Ga motion. Motion of further lying Ga-shells will be concentrated at energies closer to 27 meV, and the actual phonon sideband probably involves some combination of these motions. We feel that a calculation of these broadening effects would be useful in order to quantify the apparent "anomaly."[24]

5.3.6 Summary of GaP:O_P Results

The major results of Section 5.3 are concerned with the existence of defect vibrational modes, including both strictly localized and resonant modes. It is impossible to determine the number of such modes by group theory alone since the existence of a mode is determined by a quantitative parameter: the degree of localization around the defect. Thus, it is necessary to undertake a computation of the vibrational modes of a defect. The theory we use for this purpose is a phenomenological one, in which the force constants of the perfect lattice and those of the defect are all chosen to match experiment. Bulk phonons are described by a 15-parameter model which includes first- and second-nearest-neighbor plus

electrostatic interactions. A defect is then introduced by perturbing the mass of an atom and perturbing its first- and second-nearest-neighbor force constants. A Green's function technique is used to evaluate the local-density-of-states (LDOS) near the defect, and thereby determine the presence of defect modes.

Our theoretical results are compared with experiment for the GaP:O_P defect. We use a two-parameter description of the short-range interaction perturbation near the O site. By a suitable choice of these two defect force constants we have obtained quantitative agreement with observed phonon modes involving O^+ and O^0 defects. For O-neighbor (including O-Ga and O-P) and Ga-P–back-bond short-range defect force constants at 38% and 93% of their respective bulk values, our calculations predict two A_1 modes and two T_2 modes which are reasonably localized at or near the defect, and also, other modes which are more bulklike in nature. We summarize the character of the four most localized modes as follows:

(1) O-atom T_2 mode—This mode involves mainly O-atom motion, and this motion is localized near the defect for practically any value of the defect force constants. To match the observed phonon energies near 25 meV, our choice of defect force constants correspond to an O-lattice effective force constant (taking all long-range and short-range interactions into account) of $3.5\,\mathrm{eV}/\mathring{A}^2$, which is 25% of the bulk, P-lattice value. This force constant may be compared with that obtained by assuming that the observed phonon is completely localized on the oxygen impurity, $(25\,\mathrm{meV}/\hbar)^2 m_\mathrm{O} = 2.4\,\mathrm{eV}/\mathring{A}^2$. Our calculations demonstrate that the observed two-peak structure in this resonant mode is due to coupling with LA bulk phonons.

(2) Ga-shell T_2 mode—This mode involves mainly vectorlike motion of the shell of Ga atoms surrounding the O impurity. The mode originates from the TA peak in bulk phonon DOS, and is significantly localized around the defect for values of the O-neighbor force constants which are about 80% or less of the bulk values.

Experimentally, this mode is not seen in spectra involving isolated O impurites, but is seen when a Zn or Cd atom is located near the O atom, thereby lowering the symmetry of the system.

(3) Ga-shell A_1 mode—This mode is a breathing motion of the Ga shell. The mode originates at the peak in the LA bulk phonon DOS and is localized for values of the O-neighbor force constants which are about 80% or less of the bulk values. This resonant mode is relatively broad with a width of about 5 meV (in agreement with experiment).

(4) P-shell A_1 mode—This mode mainly involves breathing motion of the P shell (second-nearest-neighbors) surrounding the O atom. The mode originates at the LO peak in the bulk DOS, and appears when the Ga-P back-bond interactions are reduced to 95% or less of the bulk values.

For all of the four defect modes discussed above, we feel that our theoretical description is fairly close to experiment in terms of energies and eigenvectors. We have demonstrated that the O^+ and O^0 defects are weakly bound in GaP, as expected, and that the Ga-P back bonds appear to be weakened. Our predicted two-peak, O atom T_2 mode shows that the experimental results for the O^0 defect are consistent with a T_d-symmetric geometry. Finally, our calculations identify the predominant atomic motion involved in the dissociation of Zn-O pairs in GaP.

5.4 Vibrational Modes of Fe isotopes in Si

In this section, we make use of the Green's function technique to calculate the LVM's due to substitutional and interstitial Fe defects in Si. In contrast to the GaP:O_P defect treated in the previous section, Fe deep level defects in Si are not well understood. In addition, these deep levels are of relative practical im-

portance. The presence of Fe deep levels provides an electron-hole recombination path which can shorten minority carrier lifetimes, adversely affecting the operating characteristics of certain electronic devices. Moreover, Fe contaminants may be readily introduced into Si during conventional device processing.[28] Hence, it is of considerable interest to better understand the nature of Fe defects in Si; in particular, the types of defect complexes that may form in Si, and how these may affect the bulk electrical and optical properties.

Photoluminescence has proven to be a useful characterization of deep levels in semiconductors. In the study of Fe deep levels in Si, a set of photoluminescence lines occurring near 1.7 μm have been observed, and are thought to be related to isolated Si:Fe centers.[2] This set of luminescence lines contain a subsidiary peak lying in energy 9 meV below the principal peak at 1.687 μm,[2] and this subsidiary peak is believed[29] to correspond to a phonon-assisted process involving a Si:Fe LVM. The main purpose of the present calculation is to predict the theoretical existence or absence of such a Si:Fe LVM, and if the 9 meV mode is predicted, to ascertain the nature of the defect vibration. In Section 5.4.5 below, we shall see that a 9 meV LVM is indeed predicted for weakly bound, isolated Fe defects, either substitutional or interstitial.[30] Further, the mode predicted is seen to *necessarily* involve motion of the Fe impurity, and hence, the LVM frequency is expected to shift upon isotopic substitution of the defect species. The theoretical prediction of a LVM frequency shift is not observed experimentally, however.[2] The observed lack of a shift, in view of our theoretical result, suggests that the experimentally observed subsidiary peak is *not* a vibrational excitation on an Fe defect complex in Si.

5.4.1 Bulk Lattice Dynamics

We consider a model of the Si lattice in which only nearest-neighbor (NN)

interactions are present. There are then only two parameters required to describe the lattice dynamics.[31] The appropriate linear combinations of these parameters represent "bond-stretching" and "bond-bending" spring constants, f_1 and f_2, respectively. Bond-stretching springs provide a restoring force for atomic displacements *along* the equilibrium bond direction whereas bond-bending springs provide a restoring force for displacements *perpendicular* to the bond direction. In this two-parameter model, the bond-bending interaction provides the **k**-dependence of the transverse-acoustic (TA) and transverse-optical (TO) branches of the bulk-phonon spectrum, which would be flat, *i.e.*, independent of **k**, in the absence of this interaction.[4]

5.4.2 Symmetry Considerations

All the point-defects treated in the present calculation are assumed to lie at positions having T_d symmetry. This assumption is clearly true for a substitutional site. It also turns out that the interstitial site lying at the center of the fcc conventional unit cell has precisely the same nearest-neighbor geometry as a lattice site, *i.e.*, coordinated by four Si atoms arranged tetrahedrally at the same interatomic spacings. For simplicity, we assume our interstitial defect to be located on this T_d-interstitial site. (The consequences of lifting this assumption will be discussed in Section 5.4.7 below.) The defect perturbation can then be chosen to preserve the T_d symmetry. Accordingly, we decompose the five-atom, 15-dimensional, *reducible* representation of T_d, consisting of the Fe impurity atom and its four Si nearest neighbors, into irreducible representations, according to

$$\Gamma = A_1 \oplus E \oplus T_1 \oplus 3T_2. \qquad (5.35)$$

The matrices $\underline{G}^0$ and $\delta\underline{L}$ are block-diagonalized when expressed in the basis of "collective coordinates" which transform according to these irreducible repre-

sentations. The form of the block-diagonalization and the collective coordinates are reported elsewhere.[7] Here, we simply observe that the independent, non-zero matrix elements are labeled by irreducible representation $(A_1, \ldots, T_2)$ and occurrence $(1, \ldots, 3)$. There are nine independent elements to evaluate for each matrix.

5.4.3 Substitutional vs. Interstitial Defects

The Green's function method is used to determine the LVM's for Fe and Cu impurities in Si. The substitutional defects considered in this paper consist of a mass defect at the substitutional site and a NN-interaction defect with the four surrounding sites. In this case, $\underline{G}^0$ and $\delta\underline{L}$ are 15×15 matrices determined in the space of coordinates, $x_1, y_1, z_1, \ldots, z_5$, of the five lattice sites directly affected by the defect. The local modes are determined with the use of Eq. 5.7.

For interstitial defects, we have formulated an extension of Eq. 5.7 to include the extra degrees of freedom due to the interstitial. We define

$$\underline{\hat{G}_0} \equiv \begin{pmatrix} \underline{G}^0 & \underline{0} \\ \underline{0} & \mathcal{G} \end{pmatrix} \qquad \delta\underline{\hat{L}} \equiv \begin{pmatrix} \delta\underline{L} & \mathbf{A} \\ \mathbf{A}^\dagger & \delta\mathcal{L} \end{pmatrix}, \tag{5.36}$$

where $\underline{G}^0$and $\mathcal{G}$ are Green's functions for the bulk and an interstitial defect complex *isolated* from each other. The bulk–defect coupling occurs through $\mathbf{A}$, and $\delta\underline{L}$ and $\delta\mathcal{L}$ contain the resultant perturbations on the respective isolated systems. Then, the interstitial defect modes are found by using $\underline{\hat{G}_0}$ and $\delta\underline{\hat{L}}$ in place of $\underline{G}^0$ and $\delta\underline{L}$ in Eq. 5.7. Equation 5.36 is a general result for an arbitrary defect *complex*, possibly containing many internal degrees of freedom. For the interstitial *point*-defect considered in this paper, three degrees of freedom are added to the problem, and four bulk Si sites are affected by the perturbation; hence, $\underline{G}^0$ and $\delta\underline{L}$ are 12×12, and $\underline{\hat{G}_0}$ and $\delta\underline{\hat{L}}$ 15×15 matrices, respectively.

5.4.4 Defect Matrix

The complete perturbation $\delta\underline{\mathrm{L}}$ for the substitutional defect consists of a mass perturbation, $\Delta m = m_{sub} - m_{Si}$, and a perturbation Δf_1 and Δf_2 in the defect-site force constants. The new bond-stretching and bond-bending interactions between the substituent and its NN's are then given by

$$f_i' = f_i + \Delta f_i, \qquad i = 1, 2. \tag{5.37}$$

In order to satisfy the infinitesimal rotation-invariance condition,[12] it is sufficient to connect 2NN bond-bending springs μ among the four NN Si atoms. It is then easily shown that we must have $\mu = -(\Delta f_2)/4$. In practice, the addition of the springs μ to the five-atom complex has negligible effect on local mode energies, but their inclusion allows variation of Δf_2 which would otherwise be forbidden by rotation-invariance. Because the effective spring constant $k^{(\mathrm{eff})}$ for the substitutional defect is $4(f_1' + 2f_2')/3$, the weak-bonding limit ($k^{(\mathrm{eff})} \to 0$) can now be considered in a reasonable way. $\delta\underline{\mathrm{L}}$ is given explicitly as

$$(\delta\underline{\mathrm{L}})_{A_1} = \frac{\Delta f_1}{m_{Si}} \tag{5.38}$$

$$(\delta\underline{\mathrm{L}})_{E} = \frac{\Delta f_2}{4m_{Si}} \tag{5.39}$$

$$(\delta\underline{\mathrm{L}})_{T_1} = 0 \tag{5.40}$$

$$(\delta\underline{\mathrm{L}})_{T_2}^{11} = \frac{5}{6}\left(\frac{\Delta f_2}{m_{Si}}\right) \tag{5.41}$$

$$(\delta\underline{\mathrm{L}})_{T_2}^{22} = \frac{1}{m_{Si}}\left(\Delta f_1 - \frac{1}{3}\Delta f_2\right) \tag{5.42}$$

$$(\delta\underline{\mathrm{L}})_{T_2}^{33} = \frac{4}{3m_x}(\Delta f_1 + 2\Delta f_2) - \frac{\Delta m}{m_{Si}}\Omega^2 \tag{5.43}$$

$$(\delta\underline{\mathrm{L}})_{T_2}^{12} = \frac{\sqrt{2}}{6}\left(\frac{\Delta f_2}{m_{Si}}\right) \tag{5.44}$$

$$(\delta\underline{\mathrm{L}})_{T_2}^{13} = \sqrt{\frac{8}{3m_{Si}m_x}}(\Delta f_2) \tag{5.45}$$

$$(\delta \underline{\mathrm{L}})^{23}_{T_2} = \sqrt{\frac{4}{3m_{\mathrm{Si}}m_x}}(\Delta f_1), \tag{5.46}$$

where $m_x = m_{\mathrm{Si}}$ for substitutional defects.

The perturbation $\delta\hat{\underline{\mathrm{L}}}$ for the interstitial case consists of connecting the interstitial atom of mass m_x to its four Si neighbors with defect bond-stretching springs, f_1'. We could have, in addition, included defect bond-bending springs, f_2', in analogy with the substitutional case just described. However, in this work we are really interested only in a low energy LVM; hence, we wish to consider the *weak-bonding limit* of our Fe point-defects. In the substitutional case it is necessary to *weaken* the bulk force constants to zero to reach this limit. In contrast, in the interstitial case we are *adding* springs. Further, we shall see that the weak-bonding limit involves motion of the defect atom only, so that the addition of a bond-bending interaction f_2 has little effect on the results, but adds an additional unknown to the problem. Therefore, we take $f_2' = 0$. Finally, $\delta\hat{\underline{\mathrm{L}}}$ can be obtained from Eqs. 5.38–5.46 with the substitutions,

$$\delta\underline{\mathrm{L}} \rightarrow \delta\hat{\underline{\mathrm{L}}} \qquad \Delta f_1 \rightarrow f_1' \qquad \Delta f_2 \rightarrow 0 \qquad \Delta m \rightarrow 0.$$

5.4.5 Computed Results

The lattice-dynamical results for bulk Si, as calculated from the two parameter model, are summarized in Figs. 5.7 and 5.8. In Fig. 5.7, we show the bulk dispersion curves and in Fig. 5.8 the bulk density-of-states (BDOS). The calculated X-point phonon energies are $\hbar\Omega_{\mathrm{LA,LO_X}} = 45.5\,\mathrm{meV}$ and $\hbar\Omega_{\mathrm{TO_X}} = 61.7\,\mathrm{meV}$, compared to the respective experimental values,[32] 49.2 meV and 58.7 meV. In the units shown in Fig. 5.8, the BDOS integrates to the value, 5.997, in comparison with the exact value of 6. Integrals over $\mathrm{Im}\left[\underline{\mathrm{G}}^0(\Omega^2)\right]$ show comparable relative numerical precision. Hence, we expect that our calculated local phonon energies are good (with respect to computational error) to at least three significant

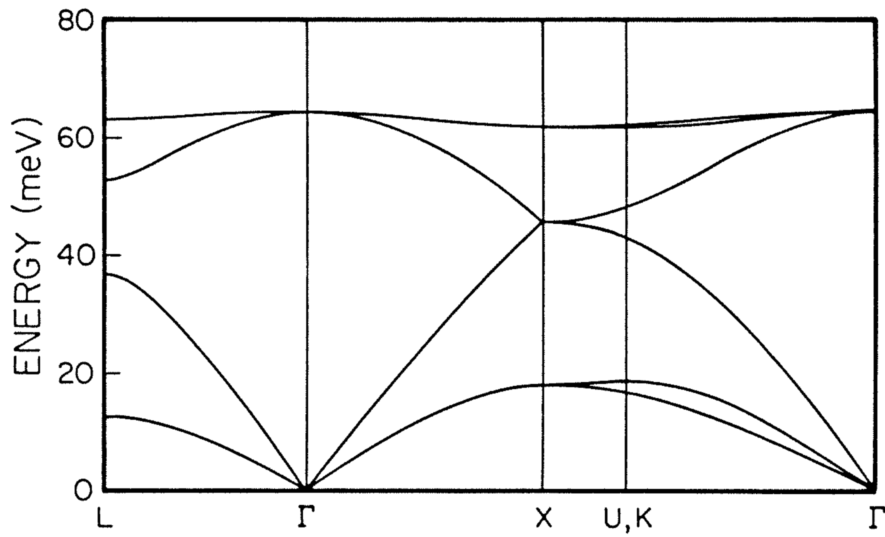


Figure 5.7: The phonon dispersion curves for bulk Si as calculated from the two-parameter model discussed in the text. The spring constants f_1 and f_2 were determined by fitting the points, LO_Γ and TA_X, exactly.

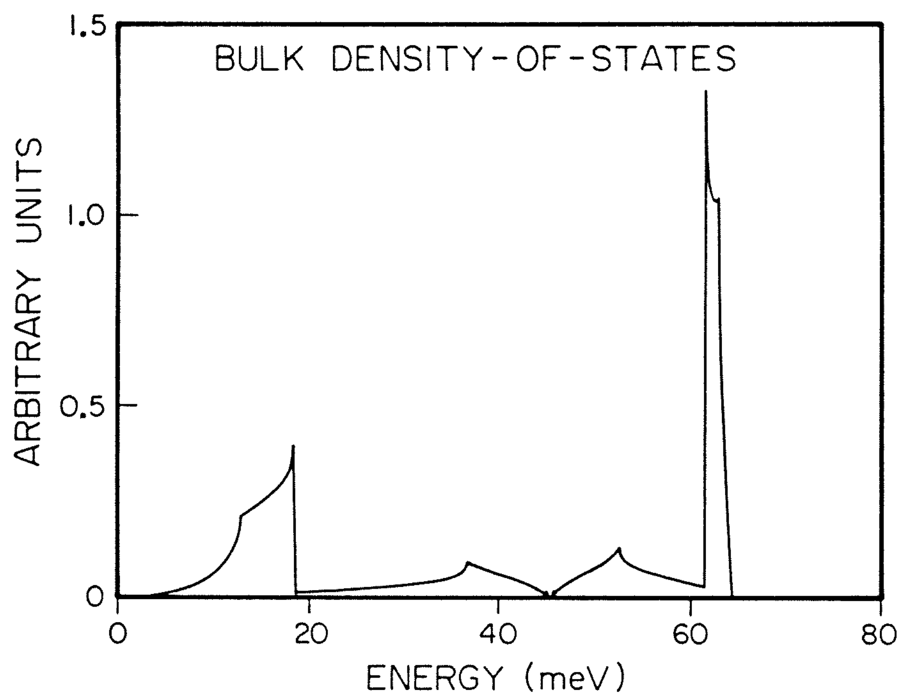


Figure 5.8: The bulk Si phonon DOS as calculated from the two-parameter model. The energy resolution is 0.03215 meV. In the units shown, the actual area under the curve is 5.997, compared to the exact theoretical value of 6.

figures.

In Fig. 5.9, we show the local modes calculated from Eq. 5.7 as a function of the parameter, $\alpha \equiv f_1'/f_1$, for interstitial ^{56}Fe. The local modes are labeled by the irreducible representations (A_1–T_2) of the symmetry group of the defect. In the figure, we see that the Green's function computation produces only A_1 and T_2 modes. A_1 modes consist of "breathing" motion of the four Si neighbors and do not involve motion of the defect. T_2 modes, on the other hand, are the only modes which couple defect-atom motion to Si-atom motion. Also, included in Fig. 5.9 are the results of a "molecular-defect-cluster" calculation in which the Fe and four Si atoms vibrate in a lattice which is immovable elsewhere. The molecular calculation is useful as a check on the Green's function computation for defect modes with strong spatial localization. This is the case in the lower left corner of Fig. 5.9 where, in the weak-bonding limit, the motion approaches that of Fe alone.

5.4.6 Comparison to Experimental Spectra

Photoluminescence spectra from certain Fe defect complexes in Si show a characteristic satellite about 9 meV below the principal line.[33,29,34] This satellite has been thought to be a local phonon mode involving Fe.[29,34] From Fig. 5.9, we see that a single T_2 mode is predicted for ^{56}Fe at a defect spring constant, $f_1' = (14.1\%)f_1$. The *calculated* 9 meV mode involves significant Fe motion and undergoes a shift in energy of $\hbar(\Delta\Omega) = 0.16$ meV upon the isotopic replacement, $^{56}\mathrm{Fe} \rightarrow {}^{54}\mathrm{Fe}$ (see Fig. 5.11). We have carried out similar calculations for substitutional Fe, in which case, we compute an isotope shift of 0.17 meV. However, recent experiments[2,36] demonstrate the lack of an isotope shift of the 9 meV feature.

To test our theoretical technique, we have made LVM calculations for sub-

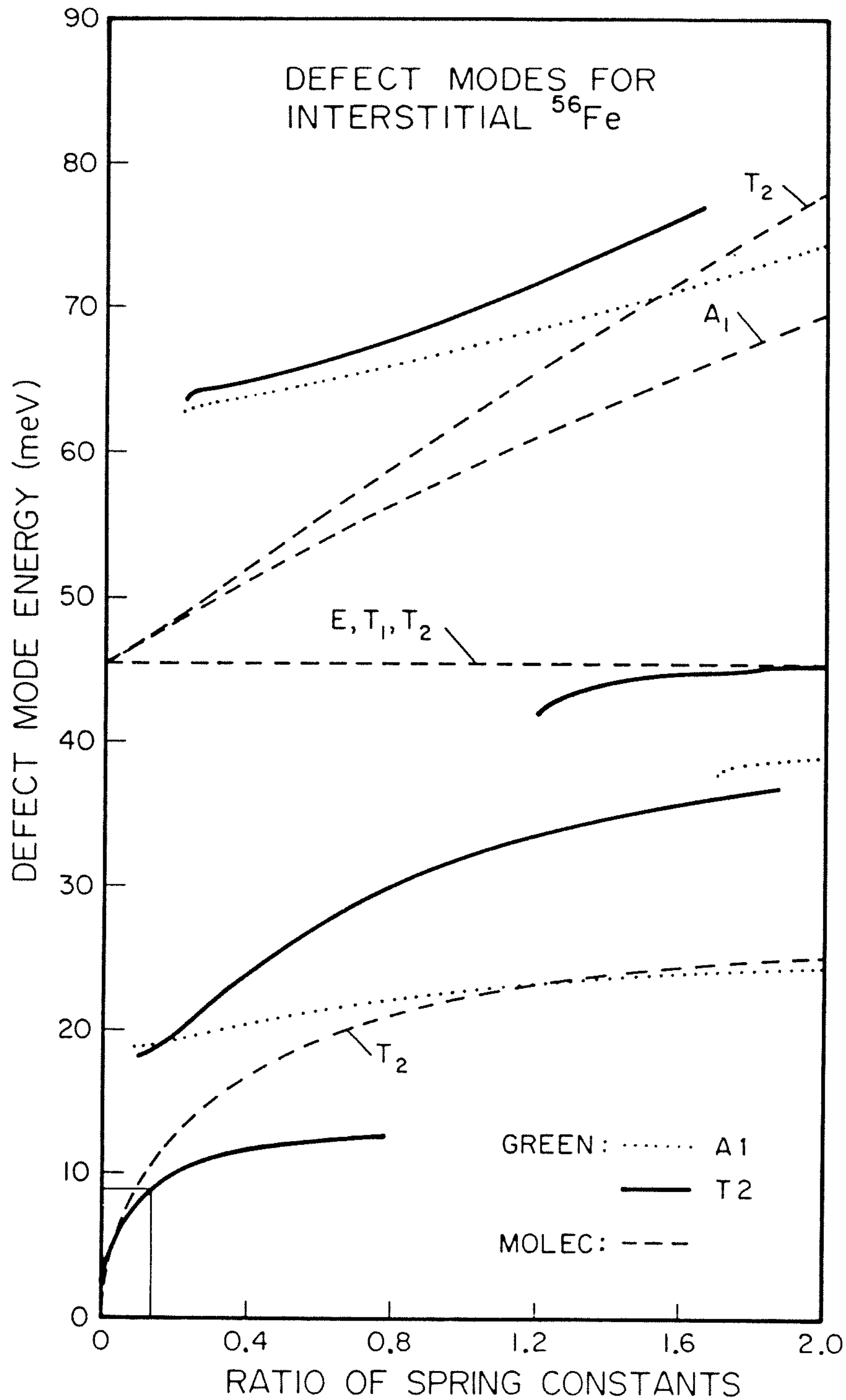


Figure 5.9: The energy and symmetry of calculated local modes of interstitial ^{56}Fe as a function of the parameter, α, defined in the text. The results of a five-atom-cluster, vibrational mode calculation are included for comparison with the Green's function results. A 9 meV mode requires $\alpha = 14.1\%$.

stitutional and interstitial Cu isotopes, for which an isotope shift *has* been observed.[35] Results similar to those shown in Fig. 5.9 were obtained for Cu; the weak spring results for all of our local mode calculations are presented in Fig. 5.10. There, we plot defect mode energy vs. α where $\alpha \equiv f_1'/f_1 \equiv f_2'/f_2$ for the substitutional defects. Experimentally, local phonon replicas at 6.96 meV and 7.03 meV, attributed to a Cu-pair complex, have been reported for ^{65}Cu and ^{63}Cu, respectively, by Weber *et al.*[35] We compute a shift of 0.11 meV, for both interstitial and substitutional Cu isotopes, in reasonable agreement with experiment, considering that our calculation was made for single-atom defects.

The Fe and Cu isotope shifts of the low energy T_2 resonant modes are apparent in the local density-of-states (LDOS) shown in Fig. 5.11. The LDOS is given by Eq. 5.31, where the trace sums over T_2 matrix elements in this case. Strong resonances appear as sharp Lorentzian peaks in the LDOS. From Fig. 5.11, we see that the width of the resonances is greater for Fe ($\approx$ 2 meV) than for the more weakly bound Cu ($\approx$ 0.7 meV). We also note that the 2 meV resonance width is of comparable magnitude to the spectral width of the R-line observed in the Fe-In luminescence seen by Schlesinger *et al.*[36,2] It is worth noting here that we have also obtained a 9 meV local T_2 mode of width $\approx$ 2 meV for substitutional In at a force constant ratio, $\alpha = 0.53$.

5.4.7 Summary of Si:Fe Results

Our calculations of Fe local modes in Si show that a 9 meV phonon should involve appreciable Fe motion and should consequently undergo a shift in energy upon isotopic substitution. Although our calculations treated the highest symmetry case with a particularly simple phenomenological model, we expect this statement to remain valid in more general cases. First, Fe is a heavy defect and will likely participate in any observed low energy local vibration. Secondly,

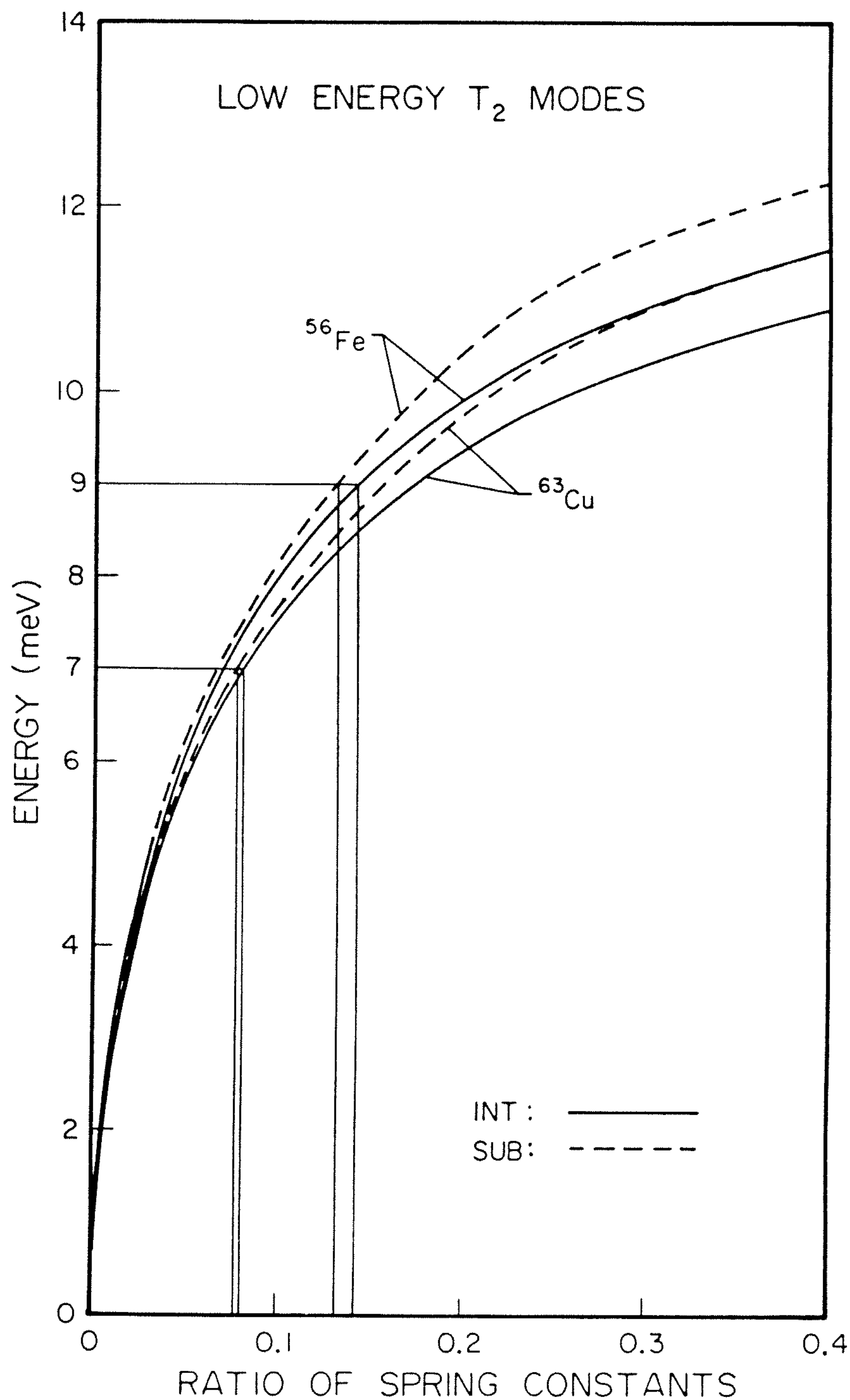


Figure 5.10: The low energy T_2 modes for T_d-interstitial and substitutional ^{56}Fe and ^{63}Cu, vs. the parameter, α. For the 9 meV Fe mode, $\alpha = 14.1\%$ and 13.0%, respectively. For the 7 meV Cu mode, $\alpha = 7.98\%$ and 7.64%.

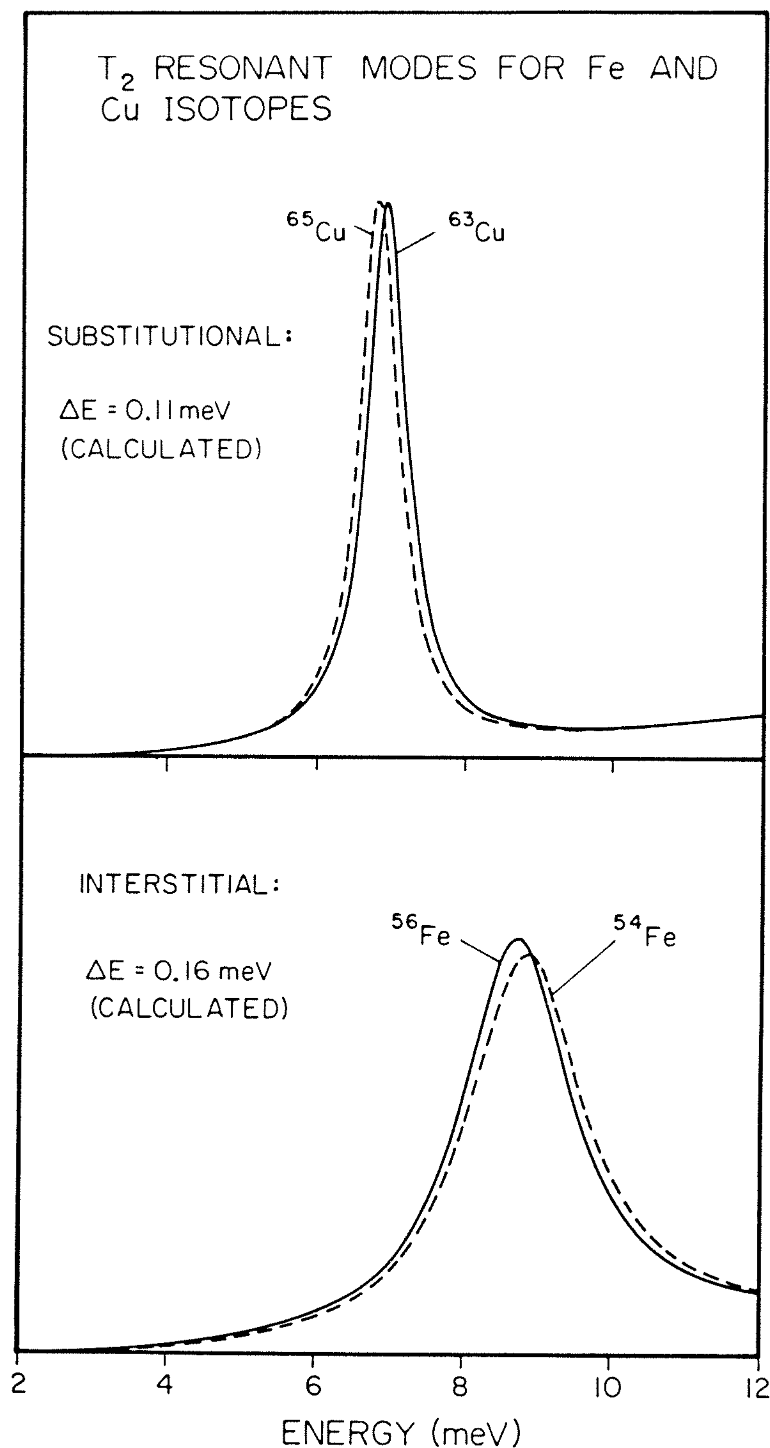


Figure 5.11: The T_2 LDOS showing the resonant character of the low energy Fe and Cu local modes. The local-mode energies calculated from Eq. 5.7 are 9.00 and 9.16 meV for ^{56}Fe and ^{54}Fe, and 6.89 and 7.00 meV for ^{65}Cu and ^{63}Cu, respectively.

the effect of reducing site symmetry will alter labels (such as T_2) but not the participation of Fe in low energy modes, which will, in general, be split by the symmetry-breaking. And finally, even in more complicated phenomenological interaction schemes, the strongest effects are short range, and these are the most important to the motion of a single atom (Fe).

Given this assertion, our calculations imply that if the 9 meV feature is an Fe *phonon*, it should undergo an isotope shift. However, measurements with different Fe isotopes show the expected shift not to occur (although certain experiments suggest that the 9 meV satellite might be an Fe *feature*). Also, no prominent phonon feature associated with bulk Si occurs near 9 meV. Thus, we conclude that the 9 meV feature is not a vibrational excitation of an Fe defect complex in Si.

5.5 Vibrational Modes of C in GaAs

5.5.1 Defect Mode Splitting due to Ga Isotopes

Substitutional carbon impurities in GaAs have recently been shown to reside on the As sublattice.[37] This site has been deduced from the observed splitting of the LVM of the carbon defect, the splitting being attributed to isotopic variation among the four nearest-neighbor Ga atoms. However, the *number* of peaks (four or five) into which the GaAs:C_{As} LVM is split is not clear from the original data, and, furthermore, the number of peaks predicted theoretically from a simple valence-force model has been incorrectly reported to be four.[37,38] In this section, we theoretically examine the GaAs:C_{As} LVM through the use of Green's-function and "molecular" defect-cluster calculations. Our theoretical analysis makes use of a two-parameter, nearest-neighbor-interaction model[39] to describe the bulk (As-Ga) and defect (C-Ga) interactions, and simple group-theoretic considerations to

predict the splittings. We find that group theory predicts a rigorous splitting of the GaAs:C_{As} LVM into nine distinct localized modes, which after broadening, merge into *five* bands, which are observed experimentally. Reasonably good agreement between theory and experiment for the peak frequencies and relative intensities is obtained, confirming the observation of Theis, Bajaj, Litton, and Spitzer[37] that substitutional carbon occupies the arsenic site in GaAs. Finally, the parameters required to fit the observed splittings indicate that the fractional variation between defect and bulk bond-bending interactions is much larger than for bond-stretching interactions for the GaAs:C_{As} defect.

5.5.2 Symmetry Considerations

The GaAs:C_{As} localized mode consists mostly of carbon atom motion, but because a small fraction of the motion ($\approx$ 6–7% of the total energy) involves the nearest-neighbor-Ga shell, the vibrational frequency of the mode has a slight dependence on these Ga masses. The splitting of the GaAs:C_{As} LVM band due to the mass variation of Ga isotopes in the NN shell can be readily predicted with the aid of group theory. There are two Ga isotopes, ^{69}Ga and ^{71}Ga, with natural abundances of 60.4% and 39.6%, and masses of 68.925574 and 70.924706 amu, respectively.[40] Hence, there are five possible nearest-neighbor-shell configurations of the Ga isotopes, which we label $l = 0, \ldots, 4$, where l denotes the number of light (^{69}Ga) NN atoms. Assuming that the GaAs:C_{As} defect is accurately substitutional on the As site, each NN configuration is associated with a point-group: T_d for the $l = 0$ and $l = 4$ configurations, C_{3v} for $l = 1$ and $l = 3$, and C_{2v} for $l = 2$. For T_d symmetry, the LVM is a triply degenerate, IR-active T_2 mode. Upon lowering the symmetry to C_{3v}, the T_d irreducible representation, T_2, decomposes according to

$$T_2 \rightarrow A_1 \oplus E,$$

splitting the GaAs:C_{As} LVM into two modes. Similarly, for the case of C_{2v} symmetry we have

$$T_2 \rightarrow A_1 \oplus B_1 \oplus B_2,$$

resulting in a threefold splitting of the LVM. Now, we simply enumerate the possible nearest-neighbor-shell configurations: for the cases, $l = 0, \ldots, 4$, we obtain 1, 2, 3, 2, and 1 peak(s) in the LVM spectrum, respectively, giving a total of nine peaks for natural Ga isotopic abundances, all of which are IR-active.[41]

5.5.3 Computed Results

The results of our Green's-function calculations are presented in Fig. 5.12. As expected, the calculations predict nine localized modes. The relative intensities of these lines have been computed from the nearest-neighbor-configuration probability, weighted by the degeneracy of each mode. Since the localized modes from the Green's-function calculations are δ-functions in Ω^2 [Ref. 3], we simulate the physical and instrumental broadening effects by using weighted Lorentzian curves[38] centered at the nine LVM frequencies, each identically broadened so that agreement with experimental linewidths is achieved. Upon broadening, we see from Fig. 5.12 that the nine lines become five peaks whose relative intensities are in reasonably good agreement with experiment.

The results of the Green's-function calculations are indistinguishable (on the scale shown in Fig. 5.12) from simpler, defect-cluster calculations that we have made, indicating that, to a very good approximation, the LVM motion can be considered completely restricted to the carbon atom and its Ga NN's. First, let us consider the results of our cluster calculation for the case, $l = 0$. We find that the LVM has energy $E_0 = E(^{71}\text{Ga}) = 72.170\,\text{meV}$, and that we can write the components of an eigenvector for this mode as

$$c_a = 0.0895$$

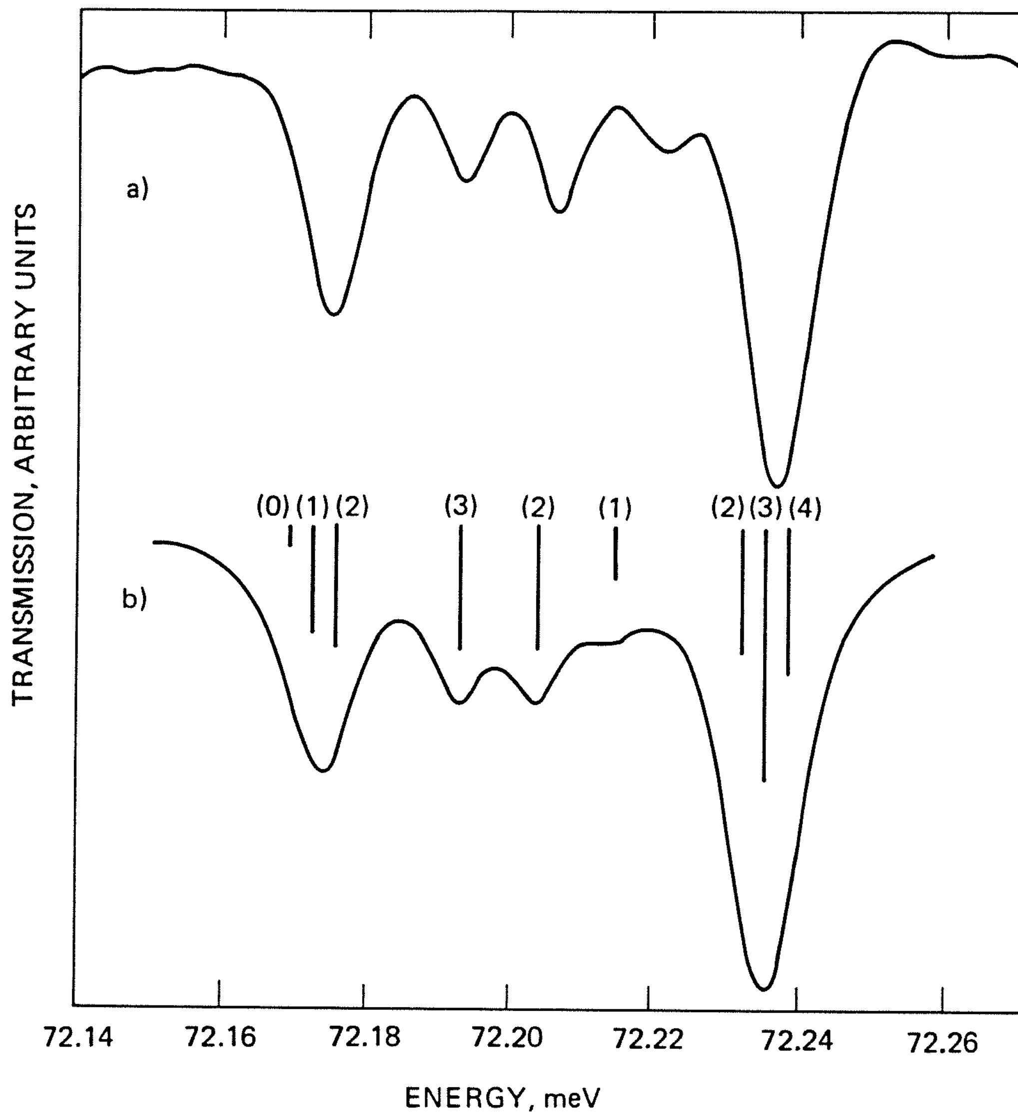


Figure 5.12: (a) High resolution, Fourier transform IR absorption measurement of the localized vibrational mode (LVM) band of $^{12}C_{As}$ in GaAs. At a resolution of $0.05\,cm^{-1}$, we see five peaks in the LVM spectrum (Courtesy of J. Baukus, Hughes Research Laboratories). (b) Calculated LVM spectrum, before and after broadening, based on a Green's-function computation. The values in parentheses indicate the number of light (^{69}Ga) atoms coordinating the defect for the particular mode. The baselines used in determining peak heights are also shown.

$$\begin{aligned} c_b &= 0.239 \\ c_c &= 0.967, \end{aligned} \tag{5.47}$$

where c_c corresponds to carbon motion, c_a to Ga-shell bond-bending motion, and c_b to Ga-shell bond-stretching motion.[39] From this result we can obtain approximate energies for the eight remaining LVM lines. To first order in the Ga mass perturbation, $\epsilon = 1 - (m_{69}/m_{71})$, the expressions for the energy shifts δE of the eight remaining LVM's are

$$(\delta E)_{A_1}^{l=1} = \frac{3}{8} c_b^2 \epsilon E_0 \tag{5.48}$$

$$(\delta E)_{E}^{l=1} = \frac{3}{16} c_a^2 \epsilon E_0 \tag{5.49}$$

$$(\delta E)_{A_1}^{l=2} = \frac{1}{4} (c_a^2 + c_b^2) \epsilon E_0 \tag{5.50}$$

$$(\delta E)_{B_1}^{l=2} = \frac{1}{8} (c_a^2 + 4c_b^2) \epsilon E_0 \tag{5.51}$$

$$(\delta E)_{B_2}^{l=2} = \frac{3}{8} c_a^2 \epsilon E_0 \tag{5.52}$$

$$(\delta E)_{A_1}^{l=3} = \frac{1}{8} (4c_a^2 + c_b^2) \epsilon E_0 \tag{5.53}$$

$$(\delta E)_{E}^{l=3} = \frac{1}{16} (5c_a^2 + 8c_b^2) \epsilon E_0 \tag{5.54}$$

$$(\delta E)_{T_2}^{l=4} = \frac{1}{2} (c_a^2 + c_b^2) \epsilon E_0. \tag{5.55}$$

The energy shifts predicted by Eqs. 5.48–5.55 differ from those computed directly (from both Green's-function and defect-cluster calculations) by less than 3%, showing that the LVM splittings are well described by first order perturbation theory.

5.5.4 Comparison to Experimental Spectra

In Fig. 5.12 we compare the results of the Green's-function calculations with the observed LVM spectrum. These results are also summarized in Table 5.2.

We see from Eqs. 5.48, 5.50, and 5.53 that the three center peaks in Fig. 5.12 should be equally spaced, and that the entire spectrum should be symmetric in frequency with respect to the center peak (Eq. 5.50). From Fig. 5.12, we see that this is indeed the case for the theoretical peaks, and also appears to be nearly true for the experimental spectrum, though all five theoretical peaks do not line up exactly with their experimental counterparts. For this work, we have arbitrarily chosen to fit our theory to the outer peaks. Alternatively, we could have fit the three middle peaks. In the latter case, the outer peaks would then be spread too far apart in our simple nearest-neighbor-interaction model. However, more realistic models, which include second-nearest-neighbor and long-range Coulomb effects, would probably tend to reduce the fraction of nearest-neighbor-Ga motion, *i.e.*, the magnitudes of c_a and c_b in Eq. 5.47, and thereby pull in the outer peaks. In deriving Eqs. 5.48–5.55 no special assumptions were made other than the on-site geometry of the carbon defect; in particular, these equations remain valid (as first order equations) in the presence of more complicated interactions. Hence, so long as values for c_a and c_b can be found that accurately, with the aid of Eqs. 5.48–5.55, predict all five peaks (which seems likely given the symmetry of the observed spectra), we may identify the Ga isotopic fluctuation as the source of the LVM splitting. The lack of exact quantitative agreement with experiment for all five peaks in our Green's-function and cluster calculations is then only a consequence of the simplicity of our model.

The calculated results presented here have made use of a two-parameter description of the C-Ga interactions. These two parameters (bond-stretching and bond-bending C-Ga force constants) were determined by fitting the two outer peaks. In our model we found it necessary to fractionally increase the C-Ga bond-bending much more than the bond-stretching force constants, relative to the bulk, to avoid significantly ($> 50\%$) overestimating the outer peak-to-peak splitting, as

Table 5.2: Experimental and theoretical LVM peak frequencies (ν) and relative intensities (I) for the GaAs:C_{As} defect.

Peak	ν_{exp} (cm^{-1})	ν_{calc} (cm^{-1})	I_{exp}	I_{calc}
1	582.12	582.12	0.59	0.50
2	582.27	582.28	0.27	0.28
3	582.37	582.36	0.35	0.28
4	582.48	582.45	0.14	0.12
5	582.61	582.61	1.00	1.00

Theis *et al.* have done using a monatomic, linear chain, defect model.[37] For our Green's-function calculations, we fit the outer peaks with the defect force constants, $f_1' = (1.0965)f_1$ and $f_2' = (4.25)f_2$. As argued in the preceding paragraph, the nearest-neighbor-interaction model assumed here is evidently too simple for this last result to be *quantitatively* meaningful, though it may remain qualitatively correct in the context of more elaborate models.

5.5.5 Summary of GaAs:C_{As} Results

We have made Green's-function and defect-cluster calculations on the local vibrational modes of the substitutional ^{12}GaAs:C_{As} defect in GaAs. We have shown from group theoretic considerations that, at the natural abundances of the Ga isotopes, the LVM band consists of nine discrete localized modes which, after suitable broadening, form *five* bands. The results of high resolution, Fourier-tranform infrared spectroscopy measurements[3] confirm the prediction of five LVM peaks. Reasonably good agreement between our theoretically

predicted peak frequencies and relative amplitudes, and experimental results,[3] are obtained. We have derived approximate expressions for the energy splittings which should remain valid for more complex models than the simple, nearest-neighbor-interaction model considered here. Qualitative predictions from these approximate expressions are consistent with experimentally observed features, further supporting the identification of Ga isotopic variation as the source of the LVM splitting. Thus, our results confirm in detail the observation of Theis *et al.*[37] that the carbon defect is substitutional for As in GaAs. Finally, the parameters required to fit the observed LVM spectra indicate that there is a much larger fractional variation in the bond-bending than the bond-stretching C-Ga force constants, relative to the values for bulk GaAs.

References

1. For a disscussion of the electrical properties of deep levels, see A. Prabhakar, Ph.D. Thesis, California Institute of Technology, 1985.

2. T. E. Schlesinger, R. J. Hauenstein, R. M. Feenstra, and T. C. McGill, Solid State Commun. **46**, 321 (1983).

3. R. J. Hauenstein, J. P. Baukus, R. M. Feenstra, and T. C. McGill (unpublished).

4. R. M. Feenstra, Ph.D, Thesis, California Institute of Technology, 1982.

5. A. A. Maradudin, E. W. Montroll, G. H. Weiss, and I. P. Ipatova, in *Solid State Physics*, edited by H. Ehrenreich *et al.*, Suppl. 3, 2nd ed. (Academic Press, New York, 1971), Chap. 8.

6. M. Tinkham, *Group Theory and Quantum Mechanics* (McGraw-Hill, New York, 1964).

7. R. M. Feenstra, R. J. Hauenstein, and T. C. McGill, Phys. Rev. B **28**, 5793 (1983).

8. K. Kunc and M. Balkanski, Phys. Rev. B **12**, 4346 (1975); K. Kunc, M. Balkanski, and M. A. Nusimovici, Phys. Status Solidi B **71**, 341 (1975); **72**, 229 (1975); **72**, 249 (1975).

9. P. J. Dean and C. H. Henry, Phys. Rev. **176**, 928 (1968).

10. P. J. Dean, C. H. Henry, and C. J. Frosch, Phys. Rev. **168**, 812 (1968).

11. B. Monemar and L. Samuelson, J. Lumin. **12/13**, 507 (1976); Phys. Rev. B **18**, 809 (1978).

12. A. A. Maradudin, in *Solid State Physics*, Ref. 5, Chap. 2.

13. R. J. Hauenstein, R. M Feenstra, and T. C. McGill, Phys. Rev. B **29**, 1858 (1984).

14. P. J. Dean, Phys. Rev. **157**, 655 (1967).

15. K. K. Rebane, *Impurity Spectra of Solids* (Plenum Press, New York, 1970).

16. G. Herzberg, *Molecular Spectra and Molecular Structure,* Vol. 3 (Van Nostrand and Reinhold, New York, 1966).

17. C. H. Henry, P. J. Dean, and J. D. Cuthbert, Phys. Rev. **166**, 754 (1968).

18. R. M. Feenstra and T. C. McGill, Phys. Rev. Lett. **47**, 925 (1981).

19. G. Baraff, E. O. Kane, and M. Schlüter, Phys. Rev. B **25**, 548 (1982).

20. T. N. Morgan, Phys. Rev. Lett. **40**, 190 (1968).

21. P. J. Dean, D. D. Manchon, Jr., and J. J. Hopfield, Phys. Rev. Lett. **25**, 1027 (1970).

22. T. N. Morgan, B. Welber, and R. N. Bhargava, Phys. Rev. **166**, 751 (1968).

23. C. H. Henry, P. J. Dean, D. G. Thomas, and J. J. Hopfield, in *Localized Excitations in Solids,* edited by R. F. Wallis (Plenum, New York, 1968), p. 267.

24. T. N. Morgan, Phys. Rev. Lett. **49**, 173 (1982); Physica **117B/118B**, 146 (1983).

25. P. J. Dean, M. S. Skolnick, Ch. Uihlein, and D. C. Herbert, J. Phys. C **16**, 2017 (1983).

26. P. J. Dean, Physica **117B/118B**, 140 (1983).

27. T. N. Morgan, Phys. Rev. (to be published).

28. K. Graff and H. Pieper, J. Electrochem. Soc. **3**, 669 (1981).

29. J. Weber and P. Wagner, J. Phys. Soc. Jap. **49**, 263 (1980).

30. R. J. Hauenstein, R. M. Feenstra, and T. C. McGill, in *Proceedings of the 163rd Meeting of the Electrochemical Society,* San Fransisco, 1983.

31. A. Grimm, A. A. Maradudin, I. P. Ipatova, and A. V. Subashiev , J. Phys. Chem. Solids **33**, 775 (1972).

32. B. N. Brockhouse, Phys. Rev. Lett. **2**, 256 (1959).

33. T. E. Schlesinger and T. C. McGill, Phys. Rev. **B25**, 7850 (1982).

34. J. Weber, R. Sauer, and P. Wagner, J. of Lumin. **24/25**, 155 (1981).

35. J. Weber, H. Bauch, and R. Sauer, Phys. Rev. **B25**, 7688 (1982).

36. T. E. Schlesinger and T. C. McGill, in *Proceedings of the 163rd Meeting of the Electrochemical Society,* San Fransisco, 1983.

37. W. M. Theis, K. K. Bajaj, C. W. Litton, and W. G. Spitzer, Appl. Phys. Lett. **41 (1)**, 70 (1982).

38. W. M. Theis, K. K. Bajaj, C. W. Litton, and W. G. Spitzer, in *Physics of Semiconductors,* edited by M. Averous (North-Holland, Amsterdam, 1983), p. 116.

39. For a description of the Green's-function and cluster calculations and the nearest-neighbor-interaction model used, see R. M. Feenstra, R. J. Hauenstein, and T. C. McGill, Phys. Rev. B **28**, 5793 (1983).

40. P. A. Tipler, *Foundations of Modern Physics* (Worth, New York, 1969), p. 496.

41. G. Herzberg, *Molecular Spectra and Molecular Structure* (Van Nostrand, Princeton, 1968), Vol. 2, ch. 3.