Design of Quantum Well Infrared Photodetectors

by

Janet Lin Pan

Submitted to the
Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 1999

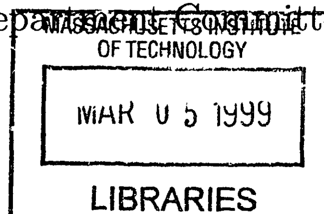
© Janet Lin Pan, MCMXCIX. All rights reserved.

The author hereby grants to MIT permission to reproduce and distribute publicly paper and electronic copies of this thesis document in whole or in part.

Author ....
Department of Electrical Engineering and Computer Science
30 October 1998

Certified by ....
Clifton G. Fonstad, Jr.
Professor of Electrical Engineering
Thesis Supervisor

Accepted by ....
Arthur C. Smith
Chairman, Department Committee on Graduate Students


MASSACHUSETTS INSTITUTE OF TECHNOLOGY
MAR 0 5 1999
LIBRARIES

ARCHIVES

Design of Quantum Well Infrared Photodetectors

by

Janet Lin Pan

Submitted to the
Department of Electrical Engineering and Computer Science
on 30 October 1998, in partial fulfillment of the requirements
for the degree of Doctor of Philosophy

Abstract

QWIPs which respond to normally incident radiation without the need for an optical grating are of particular interest because they can be fabricated with fewer process steps and increased expected yield. An important contribution of this work is the demonstration of the first n-type QWIP (n-QWIP) which showed a significant detectivity of $4\text{x}10^{10}\text{cm-}\sqrt{Hz}/\text{Watt}$ without the use of an optical grating. This detectivity corresponds to a conversion efficiency of 4 % or, equivalently, a responsivity of 270 mA/W. This detectivity is significant because it is large enough for focal plane array performance to be limited by the uniformity of processing rather than the size of the single pixel detectivity.

An important part of this work was the development of numerically accurate physical models yielding simple analytical expressions for the QWIP leakage current and photocurrent. This physical model yielded analytical expressions for the number of, and the distance over which, carriers are depleted from quantum wells whenever the photocurrent is larger than leakage current. This depletion capacitance is expected to be important at high frequencies, in situations where the photocurrent is much larger than the leakage current, and in QWIPs designed with a small number of quantum wells (as when the quantum efficiency is large or an optical cavity is used).

Studies of the microscopic physics of quantum wells are presented to elucidate the physical origin of the intersubband absorption of normally incident radiation. A key result of this work is the derivation within the framework of $\vec{k} \cdot \vec{p}$ theory of selection rules for the intersubband absorption of normally incident radiation by hole subbands in a p-QWIP (p-doped QWIP) in the absence of an optical grating. It is found that the absorption of normally incident radiation by holes in a p-QWIP in the absence of an optical grating is largest for heavy hole to light hole transitions. The intersubband absorption of normally incident radiation by electrons in an n-QWIP in the absence of an optical grating is found within $\vec{k} \cdot \vec{p}$ theory to be much smaller than that in a

p-QWIP. It is also found that $\vec{k}\cdot\vec{p}$ theory predicts that uniaxial strain does not have a large effect on the strength or the selection rules of intersubband absorption because the Hamiltonian describing uniaxial strain has the same (tetragonal) symmetry as that describing the confinement of carriers in the quantum wells along the growth direction.

Nonuniformity of device parameters across an array of QWIPs is an important issue. High Resolution X-ray Diffraction (HRXRD) was used to measure the layer width variations of QWIPs grown by molecular beam epitaxy. The spread of the measured full-width-half-maxima of superlattice diffraction peaks with the diffraction order was used with Bragg's Law to obtain the measured layer width variation in the growth direction. It was found that the fractional layer width variation was about 2% for three example growths. This layer width variation is consistent with an effusion cell temperature variation of 1°C during growth.

A theoretical study has been made of different noise mechanisms which contribute to QWIP performance. A key result found in this work is that when the signal-to-noise ratio (SNR) is limited by either fixed pattern noise or thermal leakage arrival noise, the optimal number of quantum wells for a maximum in the expected QWIP SNR is roughly η_1^{-1}, where η_1 is the quantum efficiency of a QWIP having only one quantum well.

Common QWIP designs used in industry are evaluated. In particular, the commonly used n-QWIP design in which the confinement barriers are comprised of a semiconductor superlattice is considered. This QWIP design is intended to reduce thermionic leakage by pushing the three-dimensional continuum of energy further up in energy by making the miniband transport through the superlattice barrier the means of photocurrent conduction. A Kronig-Penney model presented in this thesis showed that this QWIP design, with a superlattice comprising the QWIP barriers, is expected to have a tunneling leakage which is, at best, commensurate with QWIP barriers which are made of a single semiconductor material but whose band edge is the average value of the band edges of the semiconductors comprising the actual barrier superlattice.

The measured thermionic leakage was found to be in good agreement with a model in which the leakage depends exponentially on an activation energy which varies linearly with the applied bias. A deviation of the measured thermionic leakage from the idealized model is proposed as a quantitative measure of the amount of excessive leakage.

Thesis Supervisor: Clifton G. Fonstad, Jr.
Title: Professor of Electrical Engineering

Acknowledgments

First and foremost, I would like to thank my parents, Nancy N.-T. and Charles L.-S., for all their warmth and support throughout my life. For as long as I can remember, they have emphasized the importance of hard work, perserverance, a good education, and a system of values. They have, and will continue to be, a part of all the milestones in my life. I would like to thank my grandfather, Haw Lin, who unfortunately did not live to see the completion of this thesis, for his love and concern for me throughout my education, for the many wonderful conversations about our cultural heritage and mutual interests. My brother Alan and my sister Diana, having shared similar childhood experiences and similar career paths, have been sources of humor, encouragement, and advice throughout my education and early career.

The completion of this thesis would not have been possible without the extreme patience and understanding of Prof. Clifton G. Fonstad who offered guidance and intellectual freedom as needed. His warm and open-minded personality made him a pleasure to work with. Of course, the use of Clif's labs, research money, and years of experience made this work possible. Prof. Sheila Prasad was a valuable source of intellectual conversations, much warmth, humor and friendship, good coffee, as well as priceless words of wisdom (such as "Hurry up and finish your thesis!"). Dr. Bob Martin at Lockheed Martin and ADIC taught me the value of striving for excellence in several disciplines, as exemplified by his expertise in both device and systems aspects of photodetectors. He gave me invaluable suggestions for making the radiometric measurements in this work, as well as lessons in circuit design and device figures of merit.

I wish to thank Prof. Eugene Fitzgerald and Prof. Leslie Kolodziejski for reading my thesis.

Henry Choy and Dr. Isako Hoshino proved to be incomparable friends as well as colleagues, providing advice, humor, as well as intellectual stimulation. Henry made many astute comments and criticisms about VCSELs, p-i-n devices, as well as QWIPs. Henry also provided priceless insight into dealing with, among other things, people

and the job search. I wish him the best of luck with the Monster, and success in his future career. Isako taught me much about working with UHV chambers and their accessories, and chemistry in the lab. No one helped me more with MacIntosh computers and LaTeX than she has. I thank her for coming to my house and promptly answering email questions on many occasions.

Dr. Giuseppe Lullo and Thomas Knoedl also deserve thanks for keeping me company on those long days in the computer room when I was finishing this thesis. Thomas always came by in the morning or early afternoon. Giuseppe always kept me company (and Henry too) until about 2 in the morning, after which I was all alone in the computer room. I thank them for the fine conversation, and the flavorful Italian and German coffee.

Dr. Paul Martin and Dr. Raj Aggarwal were the primary people who taught me how to grow on and maintain the MBE, and how to process semiconductor devices. Paul also helped me to get started working with QWIPs. Dr. Gale Petrich also gave me advice on growing on and maintaining the MBE, as well as suggestions for use of the equipment in the X-ray and processing labs. Kevin Matney and Matt Wormington at Bede Scientific answered many questions about interpreting glancing incidence X-ray data. Kevin Matney also ran a best fit simulation test software on one of my samples. Dr. Brian Bennett at Naval Research Laboratory shared with me some of his experience in growing strained layer structures on the MBE, as well as measuring and interpreting initial double crystal rocking curves. Dr. Sam Wang at Lockheed Sanders performed some spatially resolved responsivity measurements across a processed die, and mailed me the data. I am also grateful to Dr. Sam Wang, Dr. Jack Ahern, Tom Faska, Mike Taylor, Chris Cooke, and Dr. Barry Lane for answering valuable questions.

I thank Dr. P. Aitor Postigo Resa for sharing with me his experience in MBE growth and surface studies. Dr. Krishna Shenoy, Dr. Yakov Royter, Donald Crankshaw, Joe Ahadian, Praveen Vaidyanathan, and Hao Wang were useful sources of information on the monolithic integration project and on processing of devices. I thank Steve Patterson for good questions which forced me to think harder. I also acknowledge the enthusiastic questions of Joanna London and Karen Waithe.

This work was supported by a fellowship from Rockwell International, by the Vinton Hayes Fund, by a research grant from Lockheed Sanders, as well as by other financial support provided by Prof. Clif Fonstad.

Contents

1 Introduction **22**

1.1 Applications for Infrared Photodetectors 22

1.2 Focal Plane Arrays on III-V Substrates 25

1.3 Thesis Overview 32

2 QWIP Operation **33**

3 QWIP Figures of Merit **37**

3.1 Comparison of Detectivity and Signal-to-Noise Requirements 37

3.2 Noise 39

3.3 Different Operating Regimes 41

3.3.1 Low Temperature Operation 41

3.3.2 Fixed Pattern Noise Limited Devices 43

3.3.3 Thermal Leakage Current Limited Devices 45

3.4 Conclusions 47

4 The Intersubband Absorption Strength 51

4.1 The Polarization of Incident Radiation and QWIP Operating Geometries 52

4.2 $\vec{k} \cdot \vec{p}$ Theory 54

4.2.1 The Kane Hamiltonian and the Bloch Functions 56

4.2.2 Envelope Functions and the Effective Mass Equation 61

4.2.3 Conduction Band Intersubband Transitions 62

4.2.4 Valence Band Intersubband Transitions 70

4.3 Conclusions 78

5 MBE Growth 81

5.1 Flux Calibration 83

5.1.1 Flux calibration by double crystal rocking curve. 83

5.1.2 Flux calibration by measuring beam equivalent pressure (BEP) and RHEED oscillations 83

5.2 Growth Conditions 85

5.3 A sample growth 89

6 Layer Structure Characterization by High Resolution X-ray Diffraction 92

6.1 Introduction 93

6.2 The Diffractometer 94

6.3 Measurement of Double Crystal Rocking Curves 97

6.3.1 Measurement of Epilayer Composition and Layer Thickness . . 97

6.3.2 Relaxation and uniformity of epilayers 102

6.4 Reciprocal Space Mapping 106

6.5 Glancing Incidence X-ray Reflectivity and Scattering Measurements . 113

6.5.1 Glancing Incidence X-ray Reflectivity Measurements 115

6.5.2 Glancing Incidence X-ray Scattering Measurements 125

6.6 Measurement of epitaxial layer width uniformity 127

6.6.1 Introduction 127

6.6.2 Distinguishing layer width variations from interface grading/roughness 130

6.7 Conclusions 137

7 Device Processing 141

8 QWIP Measurement Procedures 147

8.1 Leakage Current Measurements 148

8.2 Photocurrent Measurements 156

8.3 Spectral Measurements 165

8.4 Spatially Resolved Responsivity Measurements 165

8.5 Noise Measurement 172

8.6 Conclusions 183

9 Some Device Designs 186

9.1 Common n-QWIP Designs 187

9.1.1 Kronig-Penney Model of the Superlattice Barriers 189

9.2 Design of p-QWIPs with large quantum efficiency and low leakage current. 192

9.2.1 Three p-QWIP Designs. 194

9.2.2 Some p-QWIP Measurements 199

9.3 Conclusions 209

10 Physical Models **211**

10.1 Introduction 211

10.2 Review of Existing QWIP Models 213

10.2.1 Levine Model 213

10.2.2 Levine Photocurrent Model 213

10.2.3 Levine Leakage Current Model 216

10.2.4 Numerical Models of Rosencher 218

10.2.5 Numerical Models of Ershov 220

10.3 Analytical Expressions from a Physical Model 223

10.3.1 Insufficient Carrier Injection - Intuitive Picture 223

10.3.2 Description of Physical Model 225

10.3.3 Physical Model of Quantum Well Depletion and Accumulation with a QWIP Current of J_{3D} 237

10.3.4 Thermionic Leakage Current Modeling 250

10.3.5 Relation to the Escape Probability 258

10.4 Conclusions 259

11 Conclusions 262

11.1 Thesis Accomplishments 262

11.2 Recent Research Trends 266

List of Figures

1-1 The calculated [1, 2] transmission through 5 km of summer atmosphere at middle latitudes over farmland with a visibility of 5 km and no rain. 24

1-2 The photon flux (Equation (1.2)) radiated from a black body of unit area and temperature near 300K into a unit solid angle and a 1 μm wavelength range. 26

1-3 The photon flux contrast, Equation (1.3), is the fractional change in the photon flux, Equation (1.2), for a unit change in the black body temperature. 27

1-4 The dependence of the array signal-to-noise ratio on the uniformity of the array responsivity and the single pixel quantum efficiency. 30

1-5 An image taken from a 640x480 FPA of QWIPs designed with bariers consisting of a semiconductor superlattice. An optical grating was used to couple radiation from f/2 optics at 30 frames per second onto 50 μm pixels. The measured minimum resolvable temperature was 7 mK. . . 31

2-1 Schematic of one quantum well in a Quantum Well Infrared Phototector (QWIP). 34

3-1 The signal to noise ratio (SNR) at very low operating temperatures. . 44

3-2 The dependence on number of quantum wells of the signal to noise ratio (SNR) for fixed pattern noise limited devices. 46

3-3 The dependence on number of quantum wells of the signal to noise ratio (SNR) for thermal leakage noise limited devices. 48

3-4 The temperature dependence of various contributions to the QWIP minimum resolvable temperature. 50

4-1 Two commonly used geometries for operating QWIPs. 53

4-2 Schematic diagram showing the zinc-blende primitive cell. 55

4-3 Schematic diagram showing a single quantum well and the two confining barriers surrounding it. 57

4-4 Full numerical solution (solid line) and Brillouin-Wigner perturbative solution (dashed line) of the 14×14 Hamiltonian for the electron Bloch function at a finite $\vec{\mathbf{k}}$ in terms of the light electron, heavy electron, and spin-orbit electron Bloch functions at $\vec{\mathbf{k}} = \vec{\mathbf{0}}$. 68

4-5 Full numerical solution (solid line) and Brillouin-Wigner perturbative solution (dashed line) of the 14×14 Hamiltonian for the electron Bloch function at a finite $\vec{\mathbf{k}}$ in terms of the light hole, heavy hole, and spin-orbit hole Bloch functions at $\vec{\mathbf{k}} = \vec{\mathbf{0}}$. 69

5-1 The three chambers of our RIBER 2300 molecular beam epitaxy system. 82

5-2 Double crystal rocking curve for Sample 9443. 84

5-3 The time (horizontal axis) dependence of the intensity (vertical axis) of the specular spot in the RHEED (Reflection High Energy Electron Diffraction) pattern right after the As shutter is closed and then quickly opened under slightly As-rich growth conditions. 87

5-4 The time (horizontal axis) dependence of the intensity (vertical axis) of the specular spot in the RHEED (Reflection High Energy Electron Diffraction) pattern right after the As shutter is closed and then quickly opened under slightly As-deficient growth conditions. 88

5-5 The surface morphology of a sample grown under As-deficient growth conditions. 90

5-6 Scanning electron micrograph (SEM) of Sample 9207, which was comprised of 10 periods of 40 Å GaAs and 500 Å $Al_{0.25}Ga_{0.75}As$. 91

6-1 The triple axis diffractometer, with its three characteristics stages: the beam conditioner, the second (sample) crystal, and the analyzer crystal 95

6-2 Double crystal rocking curve for Sample 9351. 101

6-3 Glancing incidence and glancing exit geometries showing Bragg angle θ and tilt ϕ of planes. 104

6-4 Equivalent asymmetric planes in the substrate and in the epilayer will have a relative tilt of angle $\Delta\phi$. 105

6-5 Reciprocal space map of Sample 9351 near the symmetric (004) Bragg peak. 109

6-6 Reciprocal space map of Sample 9351 near the asymmetric (224) Bragg peak. 111

6-7 Reciprocal space map of Sample 9326 near the symmetric (004) Bragg peak. 112

6-8 Double crystal rocking curve for Sample 9326. 114

6-9 Best fit to a GIXR measurement of Sample 9089 intended to be a superlattice consisting of 50 periods of 80 Å AlAs and 27 Å GaAs, with a cap layer of 130 Å of GaAs. 117

6-10 The positions of the various superlattice diffraction peaks shown in Fig. 6-12. 118

6-11 Simulation of a GIXR measurement of Sample 9089 using a Gaussian distribution for layer width and roughnesses having the same mean and variances as those found from the simulation of the optimized structure shown in Fig. 6-9. 121

6-12 Visual fit, consisting of a superlattice of 50 periods of 75.6 Å AlAs and 28.4 Å GaAs, to a GIXR measurement of Sample 9089. 123

6-13 Comparison of the measured and simulated reflectivities as a function of incident angle for Sample 9089. 124

6-14 Three GIXS measurements of Sample 9089 consisting of a superlattice of 50 periods of 75.6 Å AlAs and 28.4 Å GaAs. 128

6-15 Four GIXS measurements of Sample 9324 consisting of a superlattice of 7 periods of 50 nm $In_{0.521}Al_{0.479}As$ and 6.73 nm $In_{0.532}Ga_{0.468}As$. . . 129

6-16 The dependence of the full-width-half-maxima (FWHM) on the superlattice diffraction order for the reflectivity measurement shown in Fig. 6-12. 133

6-17 Double crystal rocking curve of Sample 9331, which consisted of a 3415 Å of $In_{0.527}Ga_{0.473}As$ top contact, followed by 16 periods of an 79.3 Å $In_{0.527}Ga_{0.473}As$ quantum well, and barriers consisting of a superlattice of 29.6 Å $In_{0.380}Ga_{0.145}Al_{0.475}As$ (6 repetitions) alternated with 63.6 Å $In_{0.587}Al_{0.413}As$ (5 repetitions). 136

6-18 Reciprocal space map of Sample 9331 near the symmetric (004) Bragg peak. 138

6-19 The dependence on the superlattice diffraction order of the full-width-half-maxima (FWHM) along the [001] growth direction for Sample 9331.139

7-1 QWIP processing steps include a mesa etch, oxide deposition and via etch, ohmic metal and bond pad deposition. 142

7-2 QWIP pixels on a processed die after the second mask step. 144

7-3 Single QWIP pixel on a processed die after the second mask step. . . 146

8-1 The curves in the figure are the leakage currents as a function of bias voltage for QWIP 9066 measured between 10K and 80K in steps of 10 Kelvin degrees. 149

8-2 When the upper subband in one quantum well lines up with the lower subband in an adjacent quantum well, as shown in this figure, the sequential resonant tunneling shows a characteristic peak. 150

8-3 Arrhenius plots of the leakage current as a function of inverse QWIP operating temperature for Sample 9066 at negative bias voltages. . . 151

8-4 Arrhenius plots of the leakage current as a function of inverse QWIP operating temperature for Sample 9066 at positive bias voltages. . . . 152

8-5 The voltage dependence of the leakage current activation energy for QWIP 9066. 154

8-6 DC photocurrent measurements of a QWIP are made at cryogenic temperatures with a calibrated black body infrared source. 157

8-7 The current-voltage characteristics for the leakage current (solid lines) measured in 10K increments between 10K and 80K, and the photocurrent (dashed lines) measured at 40K. 159

8-8 Arrhenius plot of the photocurrent as a function of inverse black body temperature for Sample 9066 at 40K. 161

8-9 The points in the figure represent the net photocurrent, equal to the measured photocurrent minus the cold shield current, as a function of the black body target size for a fixed black body temperature of 500K. 163

8-10 The measured responsivity as a function of wavelength obtained for a 100 μm pixel on Sample 9066 at a QWIP operating temperature of 40K, a black body temperature of 1323.2K, and a QWIP bias of -2v. 166

8-11 False contour map of the spatially resolved responsivity at a bias of -2v, $\lambda = 8.4$ μm, and T=80K for a (25 μm)^2 optically active area on Sample 9066. 168

8-12 Spatially resolved responsivity as a function of distance along the x=55 μm cross section of Fig. 8-11. 169

8-13 False contour map of the spatially resolved responsivity at a bias of -2v, $\lambda = 8.4$ μm, and T=80K for a (100 μm)^2 optically active area on Sample 9066. 170

8-14 Spatially resolved responsivity as a function of distance along the x=160 μm cross section of Fig. 8-13. 171

8-15 The source follower circuit mounted inside the DIP package alongside a QWIP to measure the QWIP low frequency response. 175

8-16 Measured frequency response $\frac{V_{OUT}}{V_{QQ}} =$ of the source follower circuit mounted inside the DIP package alongside QWIP 9066. 178

8-17 Measured frequency response $\frac{V_{OUT}}{V_{QQ}} =$ of the source follower circuit mounted inside the DIP package alongside two resistors (R$_{QWIP} = 1\ G\Omega$ and R$_{BB} = 300\ M\Omega$) and no QWIP. 179

8-18 Measured frequency response $\frac{V_{OUT}}{V_{BB}} =$ of the source follower circuit mounted inside the DIP package alongside two resistors (R$_{QWIP} = 1\ G\Omega$ and R$_{BB} = 300\ M\Omega$) and no QWIP. 180

9-1 Some common n-QWIP designs. 188

9-2 The "Design1" design, consisting of 15 periods of 50 nm tensile $In_{0.499}Ga_{0.274}Al_{0.227}As$ barriers and 10.07 nm compressive $In_{0.67}Ga_{0.33}As$ wells, is expected to have large quantum efficiency and large photoconductive gain. 195

9-3 "Design2," consisting of 15 periods of 50 nm compressive $In_{0.532}Al_{0.468}As$ barriers and 6.24 nm tensile $In_{0.45}Ga_{0.55}As$ wells, is expected to have large quantum efficiency and small photoconductive gain. 197

9-4 "Design3," consisting of 15 periods of 8.11 nm lattice matched $In_{0.532}Ga_{0.468}As$ wells and a barrier consisting of five periods of 3.0 nm tensile $In_{0.358}Ga_{0.166}Al_{0.476}As$ barrier1 and 6.6 nm compressive $In_{0.60}Al_{0.40}As$ barrier2, is expected to have large quantum efficiency, small photoconductive gain, and small leakage current. 198

9-5 The measured leakage current as a function of bias voltage for a 50 μm pixel on Sample 9331 at QWIP operating temperatures of 9K, 20K, 30K, 40K, 50K, 60K, 70K, 80K, 90K, 100K, 130K, 160K, 200K (labeled, respectively, from the lowermost the uppermost curve in the figure). . 201

9-6 The measured leakage current as a function of bias voltage for a 50 μm pixel on Sample 9326 at QWIP operating temperatures of 10K, 20K, 30K, 40K, 50K, 60K, 70K, 80K, 90K, 100K, 130K, 160K, 200K, 240K (labeled, respectively, from the lowermost the uppermost curve in the figure). 202

9-7 The voltage dependence of the leakage current activation energy for QWIP 9331 for temperatures between 90K and 200K. 204

9-8 The voltage dependence of the leakage current activation energy for QWIP 9331 for temperatures between 60K and 100K. 205

9-9 The voltage dependence of the leakage current activation energy for QWIP 9326 for temperatures between 90K and 200K. 207

9-10 The voltage dependence of the leakage current activation energy for QWIP 9326 for temperatures between 60K and 100K. 208

10-1 In the commonly used device model for a QWIP, the electric field across each period of the structure is the same. 212

10-2 A more realistic device model for a QWIP shows an inhomogeneous field distribution within the device. 224

10-3 The modeled potential distribution, Fermi levels, and bound state energies within an n-QWIP having ten quantum wells. 240

10-4 The modeled potential distribution, Fermi levels, and bound state energies within an n-QWIP having ten quantum wells. 241

10-5 The modeled potential distribution, Fermi levels, and bound state energies within an n-QWIP having ten quantum wells. The quantum wells in the partially depleted region are uniformly (and abruptly) depleted, with a field-independent (see Equations (10.78) and (10.79)) total charge in the entire partially depleted region. 242

10-6 The modeled potential distribution, Fermi levels, and bound state energies within an n-QWIP having ten quantum wells. The quantum wells in the partially depleted region are uniformly (and abruptly) depleted, with a field-independent (see Equations (10.78) and (10.79)) total charge in the entire partially depleted region. 243

10-7 The modeled potential distribution, Fermi levels, and bound state energies within an n-QWIP having ten quantum wells. At a large enough bias, there is enough carrier injection to make the QWIP electrically neutral in the entire region between the contacts. 244

10-8 The modeled potential distribution, Fermi levels, and bound state energies within an n-QWIP having ten quantum wells. At very large biases, there is enough carrier injection to make the quantum wells slightly accumulated. 245

10-9 The modeled QWIP photocurrent as a function of the applied bias voltage obtained from solving Equation (10.70). 251

10-10The points in the figure denote the measured leakage current for sample a9066b05 as a function of voltage and at operating temperatures of 60K, 70K, and 80K (labeled, respectively, from the lowermost curve). The solid line is a best fit of the measured data points to Equation (10.93) for the thermionic field assisted tunneling through the top (triangular) part of the potential barriers. 255

List of Tables

4.1 Some energy band, electrical, and optical parameters for a selection of III-V semiconductors. 74

6.1 The lattice parameters and Poisson ratios for some III-V semiconductors [3, 4, 5, 6]. 99

6.2 Activation energies, growth temperatures, and fractional flux variations for the Group III effusion cells. 134

8.1 Measured conversion efficiencies for Sample 9066 at 40K as a function of different black body aperture sizes. 162

8.2 Measured DC voltage V_{OUT} at the FET source node as a function of different DC bias voltages V_{QQ} for the source follower circuit of Fig. 8-15. 182

10.1 Measured [7] zero bias escape time ratio, $\left(\frac{\tau_e}{\tau_r}\right)_0$, and effective barrier lowering potential per period, V_p, for four GaAs/Al_xGa_{1-x}As QWIPs. 216

10.2 Measured and theoretical (see Equation (10.89)) values of $|d\Phi_B/dV|$ for a variety of QWIPs. 253

Chapter 1

Introduction

Band gap engineering allows the peak responsivity in the quantum well infrared photodetector (QWIP) absorption spectrum to be designed to be at any place in the infrared longer than about 2 μm. This wavelength regime is useful for the identification of unknown gases as well as for use in the Earth's atmosphere in the transparent spectral regions of 3-5 μm and 8-12 μm. The spectral response of QWIPs is also very narrow, and this narrow response can be used to determine the absolute temperature of a target. Modern epitaxy and processing techniques can achieve high uniformity of semiconductor parameters across entire III-V (GaAs and InP) wafers, which allows for large Focal Plane Arrays (FPAs) of QWIPs with low spatial (fixed) pattern noise.

1.1 Applications for Infrared Photodetectors

Photodetectors which respond to wavelengths in the infrared (2-10 μm) are useful for the identification of unknown chemical species because many chemical species have characteristic vibrational spectra in the infrared. Moreover, such photodetectors are useful for sensing the heat radiated by a target, since the spectrum of the heat radiated by a black body at room temperature has a peak at a wavelength of 10 μm. Photodetectors which respond in this wavelength range are also useful in any

application which requires transmission through the Earth's atmosphere.

Figure 1-1 shows the calculated [1, 2] transmission through 5 km of summer atmosphere at middle latitudes over farmland with a visibility of 5 km and no rain. The principle features in the transmission curve are determined by the absorption of water, carbon dioxide, and particulates in the atmosphere. The highly transmissive "windows" in the infrared which are useful in QWIP applications are the long wavelength IR window between 8 and 12 μm, the "blue spike" in the mid wavelength IR window just shorter than 4 μm, and the "red spike" in the mid wavelength IR window centered at about 4.5 μm. Many surveillance applications require the detection of targets of approximately room temperature through the wide spectral region of high transmission between 8 and 12 μm.

The different IR transmission windows also show differing photon flux contrast ratios. The photon flux contrast is the fractional change in the number of radiated photons for a unit change in temperature. For a black body having a power spectral density W(λ) given by

$$W(\lambda) = \left(\frac{2hc^2}{\lambda^5}\right)\left(\frac{1}{\exp(hc/\lambda k_B T_{BB}) - 1}\right), \tag{1.1}$$

in units of Watt/cm^2-ster-μm, where W(λ) is the power radiated into a unit solid angle and into a unit wavelength interval centered at λ and into both polarizations for a black body of unit area and temperature T_{BB}, the photon flux Φ_P radiated from the black body into the wavelength range $\lambda_L \leq \lambda \leq \lambda_H$ is, in units of 1/cm^2-ster-sec,

$$\Phi_P = \int_{\lambda_L}^{\lambda_H} d\lambda \left(\frac{2c}{\lambda^4}\right)\left(\frac{1}{\exp(hc/\lambda k_B T_{BB}) - 1}\right). \tag{1.2}$$

The photon flux contrast is given by the derivative of Equation (1.2),

$$\frac{1}{\Phi_P}\frac{d\Phi_P}{dT} = \frac{hc}{\lambda k_B T^2}. \tag{1.3}$$

The photon flux (Equation (1.2)) radiated from a black body of temperature near 300K into 1 μm wavelength range is shown in Figure 1-2. Figure 1-3 shows the photon flux contrast, Equation (1.3), the fractional change in the photon flux for a unit change in the black body temperature. The photon flux contrast is seen to

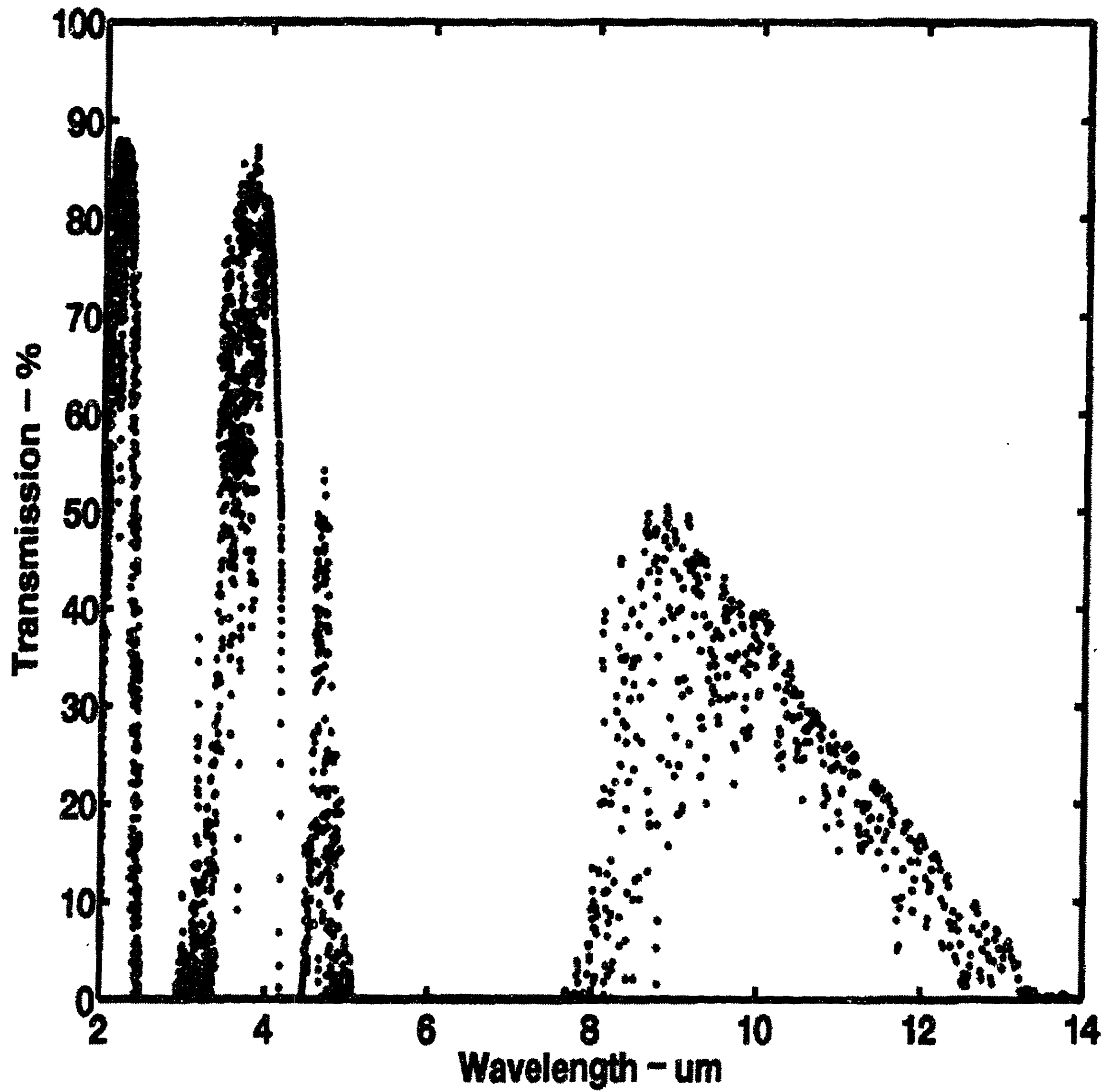


Figure 1-1: The calculated [1, 2] transmission through 5 km of summer atmosphere at middle latitudes over farmland with a visibility of 5 km and no rain.

be larger at the shorter mid IR wavelengths, even though the total photon flux (see Figure 1-2) is smaller at the shorter wavelengths. Of significance is the fact that the flux contrast ratios shown in Figure 1-3 for a unit Kelvin degree temperature change are of the order of 1%. Thus, a design requirement of a minimum resolvable temperature (MRT) of 10 mK corresponds to a flux contrast of about 10^{-4}, where the MRT is related to the signal-to-noise ratio (SNR) through,

$$MRT = \frac{1}{SNR} \frac{\lambda k_B T^2}{hc}. \tag{1.4}$$

The narrow spectral responsivity of QWIPs allows lenses, which are cheaper and smaller than mirrors, to be used, with a minimal amount of chromatic aberration and dispersion, in the optical systems which focus infrared radiation onto QWIPs. This flexibility in the location of the peak in the responsivity spectrum as well as the narrow spectral width of the responsivity is useful in the design of dual band [8] and dual color [9] QWIPs. Detection of the infrared radiation emanating from the target at two different wavelengths makes it possible to ascertain absolute target temperature and to distinguish the target from the "clutter" surrounding it in fast missile seeking detectors.

1.2 Focal Plane Arrays on III-V Substrates

Modern epitaxy techniques can achieve high uniformity of semiconductor parameters across entire III-V (GaAs and InP) wafers, which allows for large Focal Plane Arrays (FPAs) of QWIPs with low spatial (fixed) pattern noise. The figure of merit chosen for a large array of photodetectors is different from that for a single photodetector because the single photodetector is operated under circumstances different from those under which a large array of detectors are operated.

Until very recently, large arrays of photodetectors were not available. Single photodetectors had to be *quickly* scanned over a large solid angle, and photocarriers had to be collected in a short amount of time before they were pointed at a different part of the target. In those days, the figure of merit for infrared photodetectors was chosen

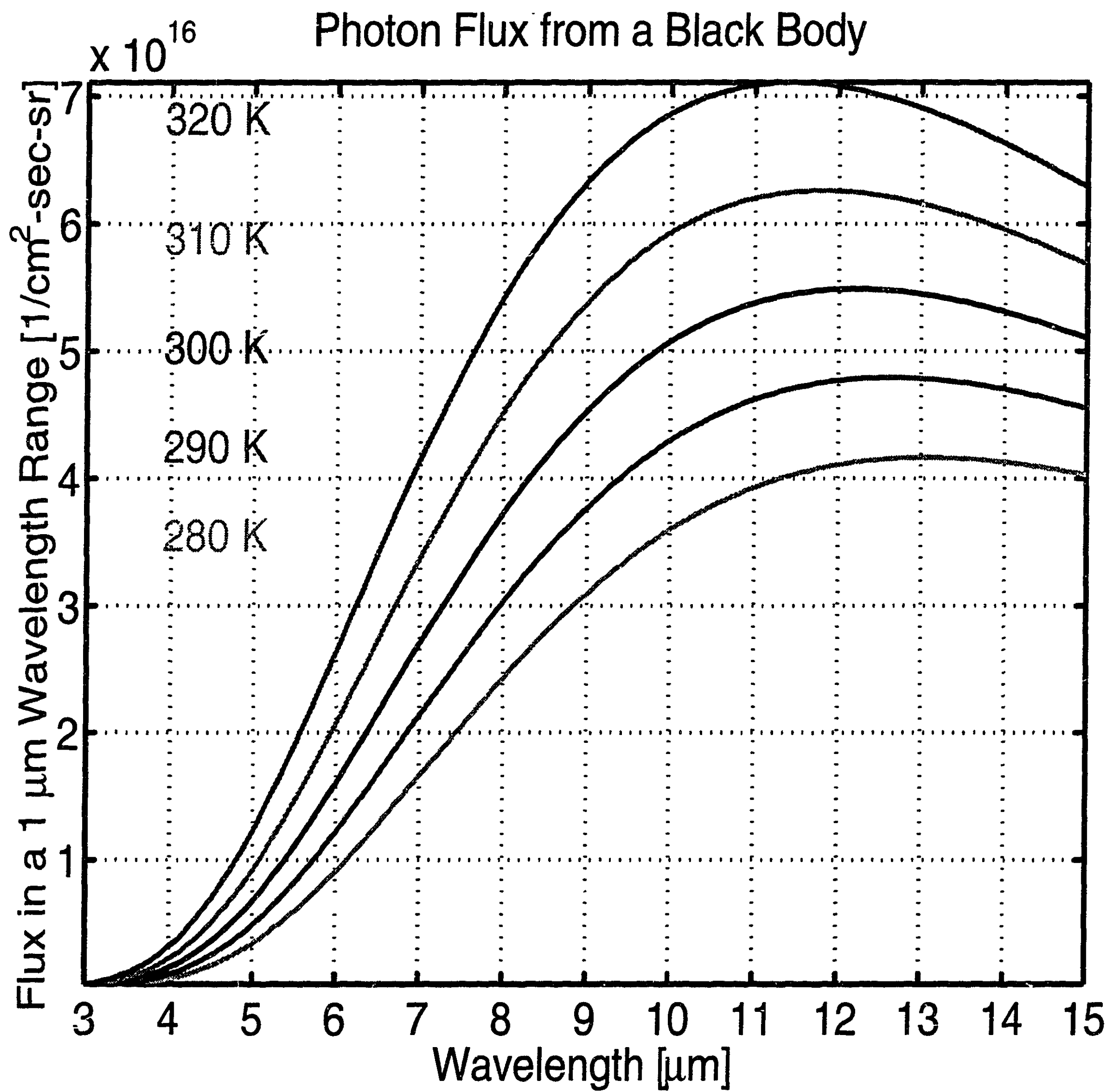


Figure 1-2: The photon flux (Equation (1.2)) radiated from a black body of unit area and temperature near 300K into a unit solid angle and a 1 μm wavelength range. The fractional change in this photon flux for a unit change in the black body temperature is shown in Figure 1-3.

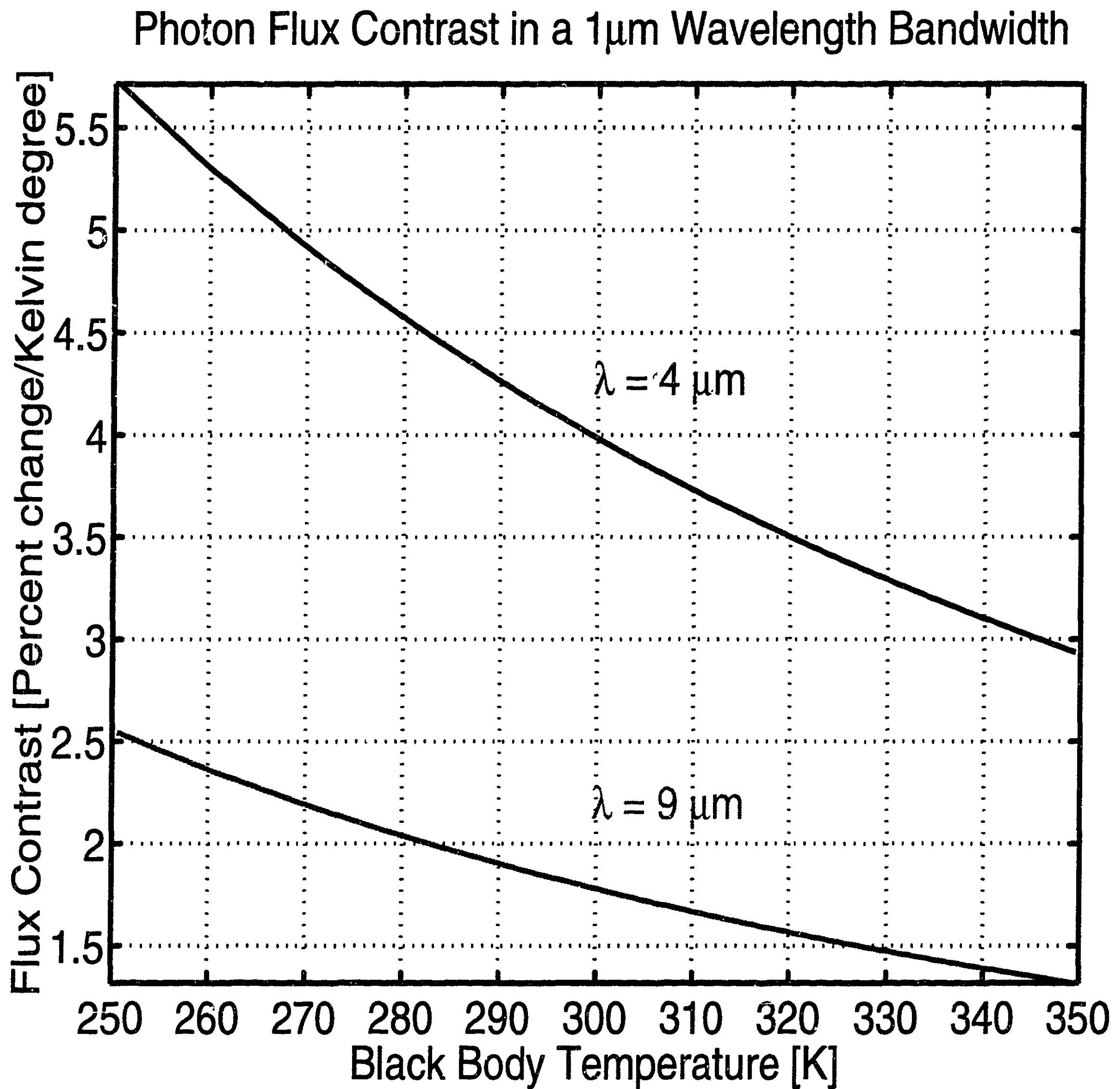


Figure 1-3: The photon flux contrast, Equation (1.3), is the fractional change in the photon flux, Equation (1.2), for a unit change in the black body temperature. The photon flux contrast is larger at the shorter mid IR wavelengths, even though the total photon flux is smaller at the shorter wavelengths.

to be the single device detectivity, where the detectivity D* is defined as [10, 11],

$$D^* = \frac{\sqrt{A\Delta f}}{P_{opt}} SNR, \tag{1.5}$$

where A is the area of a single pixel, $\Delta f = 1/2t_i$ is the measurement bandwidth, t_i is the integration time during which photocurrent is collected, P_{opt} is the incident optical power, and SNR is the measured signal-to-noise ratio. The detectivity differs from the SNR mainly in the presence of the measurement integration time $t_i = 1/2\Delta f$ in Equation (1.5).

With large arrays of QWIPs now becoming available, each photodetector pixel can be pointed at the same region of the target for a much longer duration. SNR, rather than detectivity, then becomes the appropriate figure of merit. Equation (1.5) shows that a significant SNR can be achieved for a single device having a low detectivity if the measurement integration time t_i is made appropriately long (a few milliseconds to hundreds of milliseconds).

The SNR of an entire focal plane array of QWIPs is different from that for a single QWIP detector. The focal plane array signal-to-noise ratio [7] can be written as,

$$SNR = \frac{N_P}{\sqrt{\sigma + N_P + u^2 N_P^2}}, \tag{1.6}$$

where N_P is the number of electrons collected by a sampling capacitor, where σ is the read-out integrated circuit noise floor, and where u is nonuniformity of the QWIP response over the entire focal plane array. The second term in the denominator is the generation recombination noise. The third term in the denominator is the fixed pattern noise associated with nonuniform QWIP responsivity. The nonuniformity of the photodetector responsivity is an important limit to the highest SNR which can be obtained.

Figure 1-4 shows the dependence of the array signal-to-noise ratio on the uniformity of the array responsivity and the single pixel quantum efficiency as calculated from Equation (1.6). In the figure, the higher target temperature indicates a higher photon flux and thus a higher photocurrent signal. It is seen that even though the HgCdTe photodetector has a much higher quantum efficiency (about 70% in this model) than

either the QWIP or the iridium silicide detectors (here modeled as 3.4% and 0.1%, respectively), the HgCdTe photodetector array signal-to-noise ratio is limited by the uniformity of its single pixel responsivity. In fact, Equation (1.6) shows that in the limit of very large quantum efficiency or very large photocurrent signals, the ultimate array signal-to-noise ratio is 1/u, the inverse of the nonuniformity of the photodetector response. The calculations in Fig. 1-4 were done for σ=200 electrons, a 9 μm detection peak, a 50 μm pixel, a QWIP spectral response with a full-width-half-maximum (FWHM) of 1 μm, f/1.5 optics at a 60 Hz frame rate.

This is a very central idea in photodetector technology: photodetectors having a small quantum efficiency (such as less than 1% for PtSi detectors) are still viable in the marketplace for low cost applications if very large, very uniform (less than 0.02% nonuniformity over a 640x480 PtSi array [12]) arrays of photodetectors can be fabricated. The technology for making large focal plane arrays of photodetectors is often limited not by the size of the single detector quantum efficiency but by the uniformity of this response over the entire array of devices, as shown in Fig. 1-4. Moreover, electronic correction for the responsivity nonuniformity across a FPA is never perfect. Even in the best case in which the target and a calibration object have the same temperature and spectral content, the use of a 12 bit A/D converter (ADC) for calibrating the QWIP response over an entire FPA will limit the array nonuniformity to $\frac{1}{\sqrt{12}}\frac{1}{4096} = 0.007\%$.

Figure 1-5 shows an image taken from a 640x480 FPA of QWIPs designed with barriers consisting [13, 14] of a semiconductor superlattice. An optical grating was used to couple radiation from f/2 optics at 30 frames per second onto 50 μm pixels. The measured minimum resolvable temperature was 7 mK. (The image is courtesy of Charles Parton as Lockheed Martin.) The imaged aircraft is a Mexicana Airlines Lockheed L1011. The airplane support structure and engine exhaust are clearly visible.

Growth of QWIPs on GaAs substrates allows for monolithic integration [15] of QWIPs with standard GaAs detector circuits. Monolithic integration removes the necessity of indium bump bonding, an extra processing step which contributes to expense and lowered yield. Monolithic integration is also expected to remove the complications

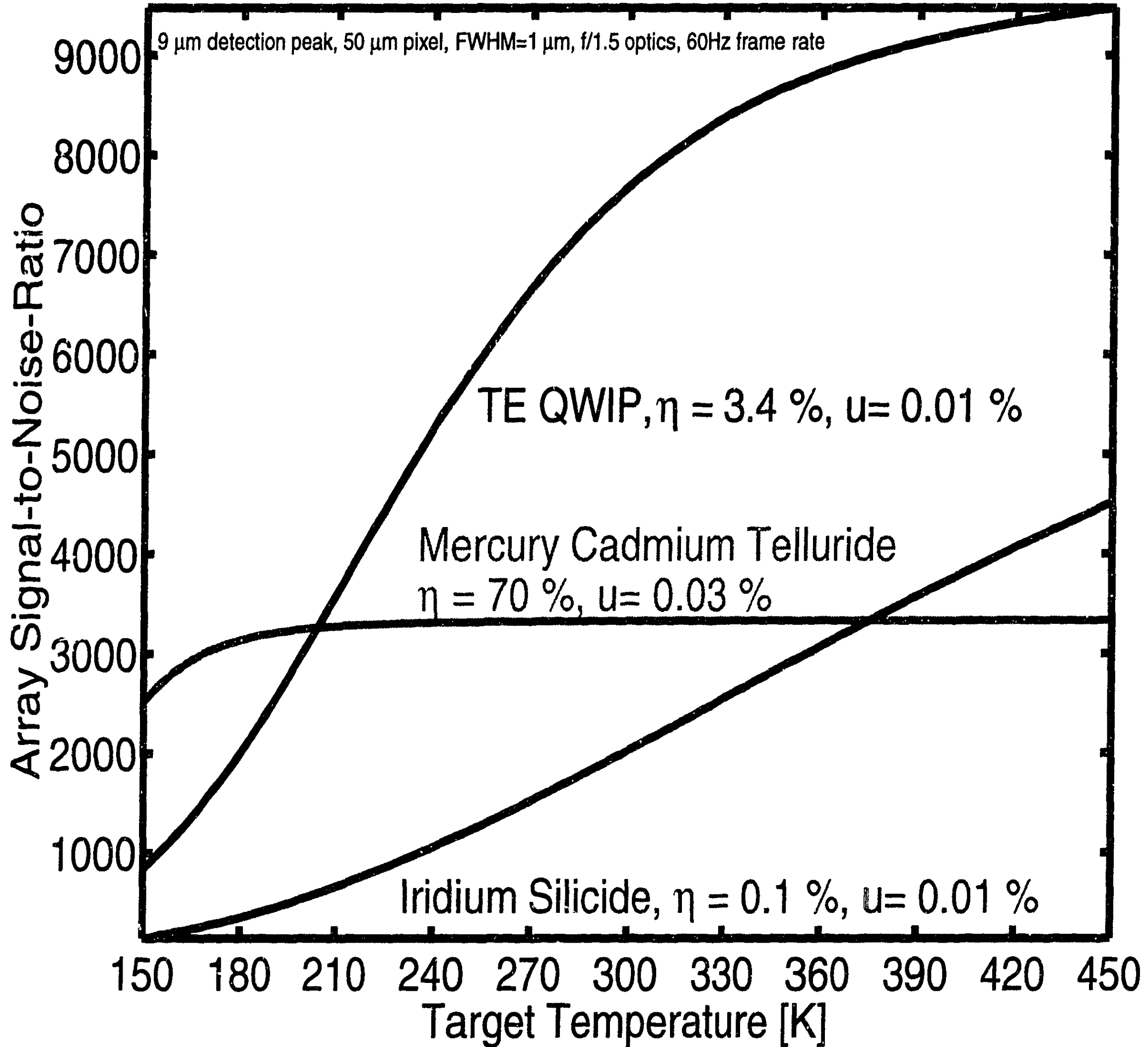


Figure 1-4: The dependence of the array signal-to-noise ratio on the uniformity of the array responsivity and the single pixel quantum efficiency. In the figure, the higher target temperature indicates a higher photon flux and a higher photocurrent signal. It is seen that even though the HgCdTe photodetector has a much higher quantum efficiency (about 70% in this model) than either the QWIP or the iridium silicide detectors (here modeled as 3.4% and 0.1%, respectively), the HgCdTe photodetector array signal-to-noise ratio is limited by the uniformity of its single pixel responsivity. The calculations were done for σ=200 electrons, a 9 μm detection peak, a 50 μm pixel, a QWIP spectral response with a full-width-half-maximum (FWHM) of 1 μm, f/1.5 optics at a 60 Hz frame rate.

Figure 1-5: An image taken from a 640x480 FPA of QWIPs designed with bariers consisting of a semiconductor superlattice. An optical grating was used to couple radiation from f/2 optics at 30 frames per second onto 50 μm pixels. The measured minimum resolvable temperature was 7 mK. (The image is courtesy of Charles Parton as Lockheed Martin.) The imaged aircraft is a Mexicana Airlines Lockheed L1011. The airplane support structure and engine exhaust are clearly visible.

resulting from the thermal mismatch between Si electronics and III-V (GaAs) optical devices.

1.3 Thesis Overview

In Chapter 2, we present general features of QWIP operation. In Chapter 3, we discuss the appropriate figures of merit for various operating regimes of QWIPs, and relate these figures of merit to device design parameters. In particular, we present a discussion of the dependence of different noise mechanisms on QWIP device parameters. In the following chapter, Chapter 4, we present the physics of intersubband transitions. In particular, we present a new analytical expression for the size and selection rules for the aborption of normally incident radiation in p-doped QWIPs (p-QWIPs). Then, in Chapters 5 and 6, we discuss molecular beam epitaxial (MBE) growth and X-ray characterization of QWIPs. In particular, we present a way of characterizing the uniformity of MBE grown epitaxial layers. Chapter 7 discusses the process for making devices. Chapter 8 discusses the different optical measurements which were made in this work on QWIPs. Chapter 9 presents an evaluation of some different device designs. Chapter 10 presents some physical models for QWIPs. These physical models are compared with the existing literature on QWIP models. We conclude with a summary in Chapter 11.

Chapter 2

QWIP Operation

Fig. 2-1 shows a schematic of one quantum well in a Quantum Well Infrared Photodetector (QWIP). The photocurrent density J_P consists of carriers which have been optically excited with a photon flux Φ from a bound lower state to an upper state, usually in the energy continuum above the barrier band edge. The strength of the optical transition is proportional to the optical dipole, and is usually discussed in terms of the quantum efficiency η. Transport of the photoexcited carriers is described in terms of the photoconductive gain g. The leakage current density J_L is the current density flowing through the QWIP in the absence of any incident radiation.

Photodetectors based on intersubband transitions have had a long history. Esaki and Sakaki [16] were the first to suggest GaAs/AlGaAs multiple quantum well photodetectors. Smith et al. [17] did some of the first experimental work. Coon and Karunasiri [18] and Chiu et al. [19, 20] did some of the early theory. West and Eglash [21] made the first unequivocal observation of a strongly absorbing intersubband transition in a stack of 50 quantum wells. Levine [22] was the first to fabricate a QWIP based on the intersubband transition between bound states in a quantum well. Two significant improvements [7] to the original QWIPs were the use of much wider confinement barriers to reduce sequential resonant tunneling and designs placing the upper state in the continuum above the barrier band edge.

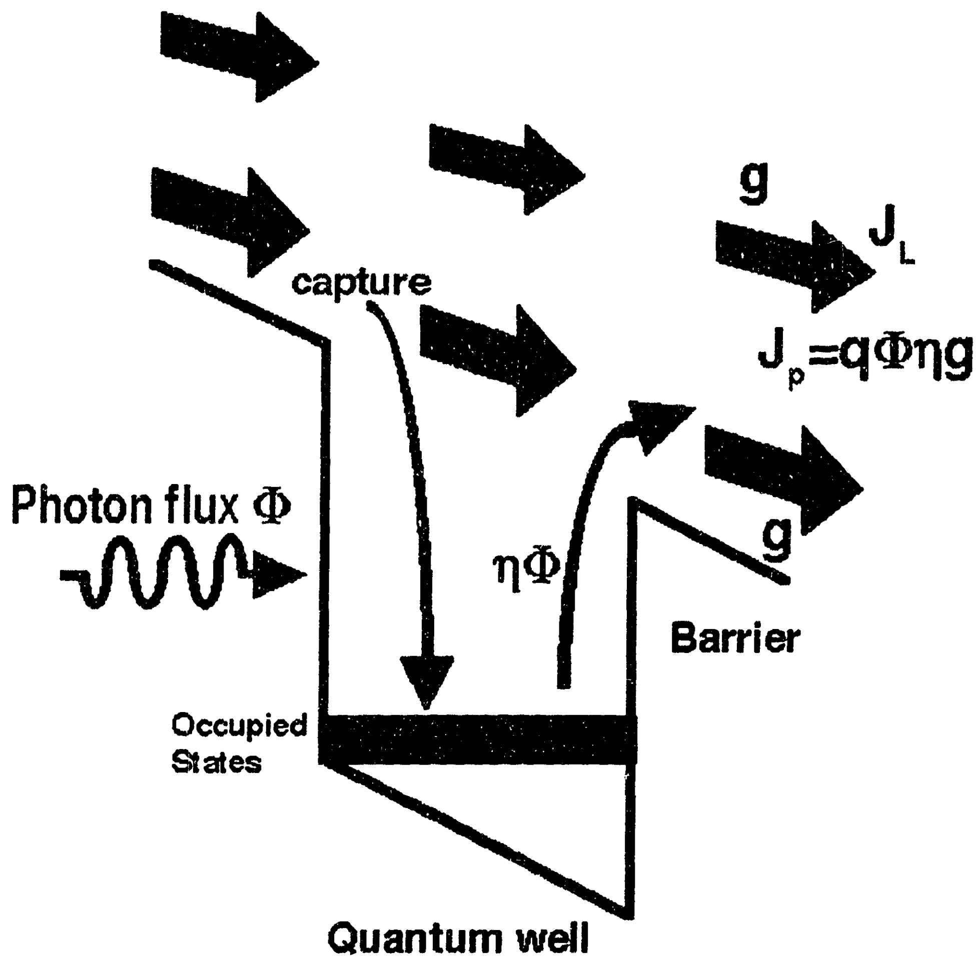


Figure 2-1: Schematic of one quantum well in a Quantum Well Infrared Phototector (QWIP). The photocurrent density J_P consists of carriers which have been optically excited with a photon flux Φ from a bound lower state to an upper state, usually in the energy continuum above the barrier band edge. The strength of the optical transition is proportional to the optical dipole, and is usually discussed in terms of the quantum efficiency η. Transport of the photoexcited carriers is described in terms of the photoconductive gain g. The leakage current density J_L is dominated by the sequential resonant tunneling current at very low temperatures, and by the thermal leakage current at higher (greater than 50K) temperatures.

The QWIP photocurrent is

$$i_p = qF\eta g \tag{2.1}$$

where F is the incident photon current (sec^{-1}), the quantum efficiency

$$\eta = 1 - \exp(-N_w \eta_1) \tag{2.2}$$

describes the strength of the optical absorption for N_w quantum wells, each with a quantum efficiency of η_1, and the photoconductive gain, g, describes the transport properties of the QWIP,

$$g = v_d \tau_L / L_{TOT} \equiv g_1 / N_w, \tag{2.3}$$

where v_d is the drift velocity, τ_L is the lifetime in the upper state (the energy continuum), L_{TOT} is the total length of the QWIP superlattice, and g_1 is the photoconductive gain for one quantum well. The photoconductive gain is related to the the capture probability p_c for a carrier in the upper state through,

$$g_1 = (1 - p_c)/p_c. \tag{2.4}$$

The incident photons arrive with a current F (sec^{-1}) equal to

$$F = \Phi_B \Omega A_{det}, \tag{2.5}$$

where Φ_B (cm^{-1}-$ster^{-1}$-sec^{-1}) is the flux of incident photons emitted by the target, Ω is the solid angle subtended by the target as seen by the QWIP, and A_{det} is the QWIP pixel area.

In photodetectors, the upper subband must also have favorable transport properties, such as a high mobility, and this is usually accomplished [7, 18] by choosing this upper subband to be in the energy continuum above the barriers. For a large absorption strength, the quantum mechanical wavefunction of the upper subband must retain more of the localized nature of a "bound" quantum well state, and less of the delocalized nature of a true continuum state. This is usually accomplished by choosing this upper subband to be barely in the energy continuum, usually right at the band edge of the barrier material.

The leakage current is the current flowing through the QWIP in the absence of any incident radiation. The leakage current is dominated by sequential resonant tunneling at low temperatures (less than 50K), and can be reduced by using smaller applied biases, usually requiring QWIPs to utilize a bound-to-continuum transition [7]. Many applications require the ability to operate the QWIP at as high a temperature as possible, at least at liquid nitrogen temperature (77K). At these temperatures, the leakage current is dominated by the thermal population of the upper subband state by nonradiative processes.

The thermal leakage current i_{th} is

$$i_{th} = n_{th} q v_d A, \tag{2.6}$$

where A is the area of the device, q is the electronic charge, n_{th} is the thermal population of carriers in the energy continuum in the steady state,

$$n_{th} = N_c \exp(-E_{co}/k_B T_{op}) \tag{2.7}$$

where $E_{co} = V_B - E_1$ is the lower cutoff energy of a bound to continuum transition, and the density of states carrier concentration is

$$N_c = 2(2\pi m_{Barr} k_B T_{op}/h^2)^{3/2}. \tag{2.8}$$

Chapter 3

QWIP Figures of Merit

This section describes figures of merit for QWIPs operating in different temperature ranges, where different noise mechanisms are dominant. The figure of merit for QWIP performance is the signal to noise ratio (SNR). The SNR requirement sets a lower bound on the required conversion efficiency, ηg, for any individual device. An upper bound on the conversion efficiency is set by the size of the capacitor used to collect the photocurrent. We shall find in Section 4 that for fixed pattern noise limited or thermal leakage noise limited QWIPs, quantum efficiency is more important than photoconductive gain for increasing the device SNR. Further, using a larger number of quantum wells may result inmore uniform device parameters.

3.1 Comparison of Detectivity and Signal-to-Noise Requirements

The detectivity is defined as

$$D^* = \frac{\sqrt{A\Delta f}}{P_{opt}} SNR = R\frac{\sqrt{A\Delta f}}{i_N}, \tag{3.1}$$

where Δf is the measurement bandwidth, P_{opt} is the incident optical power, the signal-to-noise ratio is $SNR = i_p/i_N$, i_N is the noise current, and the responsivity is

$$R = \frac{q\eta g}{h\nu}, \tag{3.2}$$

the number of photocarriers measured per unit photon of incident radiation. The detectivity differs from the SNR mainly in the presence of the measurement bandwidth Δf in Equation (3.1). Until very recently, detectivity was the figure of merit for the infrared photodetectors fabricated because single photodetectors had to be *quickly* scanned over a large solid angle. Since large arrays of photodetectors were not available then, single photodetectors had to collect photocarriers in a short amount of time before they were pointed at a different part of the target.

With large arrays of QWIPs becoming available, each photodetector pixel can be pointed at the same region of the target for a much longer duration, the inverse of the frame rate. SNR, rather than detectivity, then becomes the appropriate figure of merit, as the integration time for photocarrier collection is relatively long for many applications of focal plane arrays (FPAs). As indicated in Equation (3.1), the same SNR can be obtained from a lower detectivity photodetector if the measurement bandwidth is small enough.

Needed conversion efficiency.

The minimum conversion efficiency, ηg, needed is determined by the required dynamic range, SNR. If a dynamic range of SNR is required, then a total of N_e electrons, where

$$SNR = \sqrt{N_e}, \tag{3.3}$$

must be collected by the integration capacitor in a time t_i. For many applications, a dynamic range of 12 bits, SNR=4096, is required. This requires 16 million electrons to be collected by the integration capacitor. To collect this many electrons during the time t_i requires a total photocurrent of

$$qN_e f_r = i_P = q\Phi_B \Omega A_{det}(\eta g), \tag{3.4}$$

where q is the electronic charge, the frame rate is f_r=60 Hz=$1/2t_i$ typically, where the target radiates with a 300K blackbody flux of Φ_B=6.3×10^{16}/sec-cm^2-ster, the solid angle is Ω=0.1963 for an f/2 lens, the detector area is A_{det}=(22μm)^2. For these numbers, a dynamic range of SNR=4096 would thus require a conversion efficiency of

$$\eta g = 1.58 \text{ percent.} \tag{3.5}$$

3.2 Noise

Since SNR is the appropriate figure of merit for QWIP FPAs, a discussion of the differing noise mechanisms which contribute in various operating regimes is now appropriate.

The generation-recombination noise is the fundamental limiting noise mechanism of photoconductive detectors,

$$i_{GR}^2 = 4qgi_p\Delta f. \tag{3.6}$$

Generation-recombination noise is a result of the random nature of processes that create and destroy photocarriers.

The fixed pattern noise (FPN) is the spatial noise pattern that emerges from a FPA as a result of differences in the leakage current in each QWIP pixel in a FPA. This FPN is proportional to the nonuniformity δ_L, defined as three times the ratio of the standard deviation to the mean, of the leakage current, as well as the change in the leakage current as the operating temperature T_{op} drifts,

$$i_{FPN} = \delta_L \left(\frac{di_L}{dT_{op}}\right) \Delta T_{op}. \tag{3.7}$$

The fixed pattern noise is small when the the QWIP cooling mechanism can hold the operating temperature to be very constant. Typically, for Joule-Thomson (JT) coolers, ΔT_{op}=50 mK. The fixed pattern noise is also small for either small leakage currents or for very uniform leakage currents (δ_L small).

At high enough temperatures, especially above the background limited temperature, the random creation and destruction of the leakage current carriers contribute a thermal leakage noise,

$$i_{th}^2 = 4qgi_L\Delta f. \tag{3.8}$$

The Johnson noise is the noise that results from the uncertainty in the velocities of the excited carriers,

$$i_{n,J}^2 = \left(\frac{4k_BT_{op}}{R}\right)\Delta f. \tag{3.9}$$

The Johnson noise becomes insignificant when the bias voltage is large enough,

$$gV_{bias} > k_BT_{op}/q. \tag{3.10}$$

The read-out (switch) noise,

$$v_{n,rms}^2 = \left(\frac{k_BT_{op}}{C}\right), \tag{3.11}$$

also becomes insignificant for a large enough bias voltage,

$$V_{bias} > v_{n,rms}. \tag{3.12}$$

The photon random arrival noise,

$$i_{photon}^2 = 2q\eta gi_P\Delta f, \tag{3.13}$$

is always less than the generation recombination noise, and is usually not considered, especially for small conversion efficiencies ηg.

1/f noise in silicon has traditionally been associated with traps at the interface between and silicon and its oxide. 1/f noise is less well understood in GaAs, though it is believed [14] to be less of a problem than it is in silicon devices. We will not discuss it further.

3.3 Different Operating Regimes

At very low temperatures, the optimal design is of very large signal-to-noise devices. At all temperatures, fixed pattern noise is always a concern. At intermediate and higher temperatures, the SNR is limited by the random creation and destruction of photocurrent and leakage current carriers.

"Low" and "high" temperatures are defined relative to the Background LImited Performance (BLIP) temperature, which is the operating temperature at which the photocurrent is the same as the leakage current,

$$i_P(T_{BLIP}) = i_L(T_{BLIP}). \tag{3.14}$$

At different operating temperatures, the SNR is dominated by different noise mechanisms. Knowledge of the dominant noise mechanism allows one to choose QWIP designs which maximize the SNR. This section discusses how SNR can be maximized by choosing an appropriate quantum efficiency, photoconductive gain, and quantum well number.

3.3.1 Low Temperature Operation

At low temperatures, the signal to noise ratio is dominated by the generation-recombination noise (GR) noise. In fact, as long as the noise is dominated by device GR noise, smaller photoconductive gains yield higher signal to noise ratios. Indeed, one advantage of photoconductive detectors with respect to photovoltaic detectors is the possibility of photoconductive gains less than unity. The reason is that the shot noise of carriers in a pn junction is $2qi_P\Delta f$, which is greater than the photoconductive GR noise when g>1/4.

Whereas decreasing the photoconductive gain might be a good thing, should the small photoconductive gain come as a result of having many quantum wells or of having a small one quantum well photoconductive gain? To answer this question, one must

consider the Johson noise, as well as the GR noise, as the former becomes important at small photoconductive gains.

When the dominant noise mechanisms are GR noise and Johnson noise, the SNR is

$$SNR = \frac{i_P}{\left\{4\left[qg + \frac{k_B T_{op}}{V_{bias}}\right] i_P \Delta f\right\}^{1/2}}. \tag{3.15}$$

For design purposes, the measurement bandwidth Δf, equal to the frame rate, is not arbitrary, but is commensurate with filling up the integration capacitor with some fixed number of carriers, N_e. In other words, this finite electron storage requirement says that the desired frame rate Δf and the expected photocurrent i_P fixes the size of the integration capacitor to hold some N_e carriers, through the relation,

$$\Delta f = 1/2t_{integration} = i_P/2qN_e. \tag{3.16}$$

The previous SNR can then be written in terms of the integration capacitor size instead of the measurement bandwidth,

$$SNR = \frac{N_e}{\left\{\frac{1}{N_w}\left[g_1 + \frac{g_{1opm}^2}{g_1}\right]\right\}^{1/2}}. \tag{3.17}$$

where we have used the definition of the photoconductive gain to write the optimum photoconductive gain, the value of the photoconductive gain at which the SNR is a maximum, as

$$g_{1opm} = \sqrt{\mu \tau_L V_{th}}/L_P, \tag{3.18}$$

where $v_{th}=k_B T_{op}/q$. This last result comes about because both the photoconductive gain in the GR noise and the Johnson noise depend on the applied bias voltage through the drift velocity.

Clearly, in this model, there is an optimal single quantum well photoconductive gain, but the SNR still increases as the square root of the number of quantum wells used. Thus, QWIPs which operate at low temperature should be designed with the largest number of quantum wells that still allows the device GR and Johnson noise to limit

the SNR. This is shown in Fig. 3-1 for the following typical numbers for an InGaAs pQWIP: μ=100 cm^2/V-s, τ_L=2 ps, k_BT_{op}=6.6 meV, L_P=50 nm, η_1=0.35%, g_{1opm}=0.23, Δf=60 Hz, T_B=300K, N_e=4.5$\times$10^5, F=1.26$\times$10^{11}/s (300K target temperature viewed through f/2 lens with 70% lens transmission with peak responsivity at 8.55μm), A=(40μm)^2.

Fig. 3-1 also shows that when the one quantum well photoconductive gain, g_1, is very small, the SNR is limited by Johnson noise, and when g_1 is very large, the SNR is limited by GR noise.

Of interest is the fact that it is the *total number* of collected photocarriers, N_e, which appears in Equation (3.17), and *not the photocurrent, i_P, by itself.* This is a reflection of our use of the finite size of the integration capacitor, the finite electron storage requirement of Equation (3.16). It is N_e which appears in Equation (3.17), and not i_P, because if the finite integration capacitor is filled up with N_e photocarriers in a time less than the inverse frame rate f_r, then any additional photocurrent cannot be collected by the integration capacitor and thus cannot contribute to the SNR.

3.3.2 Fixed Pattern Noise Limited Devices

At all operating temperatures, fixed pattern noise should be kept small. For example, at temperatures less than the BLIP temperature, the SNR is often determined by both the GR and the fixed pattern noise. To determine which QWIP device parameters to optimize, consider the SNR in this operating temperature regime,

$$SNR = \frac{i_P}{\left\{4qgi_P\Delta f + \left[\delta_L \frac{di_L}{dT_{op}}\right]^2\right\}^{1/2}}. \tag{3.19}$$

It is useful to express this SNR in terms of the noise equivalent change in background temperature (nedt) that would have resulted, had there been no fixed pattern noise,

$$SNR = \frac{1}{\left\{\left[\frac{T_P}{T_B^2}(nedt)\right]^2 + \left[\delta_L \frac{T_P i_L}{T_{op}^2 i_P}\delta T_{op}\right]^2\right\}^{1/2}}. \tag{3.20}$$

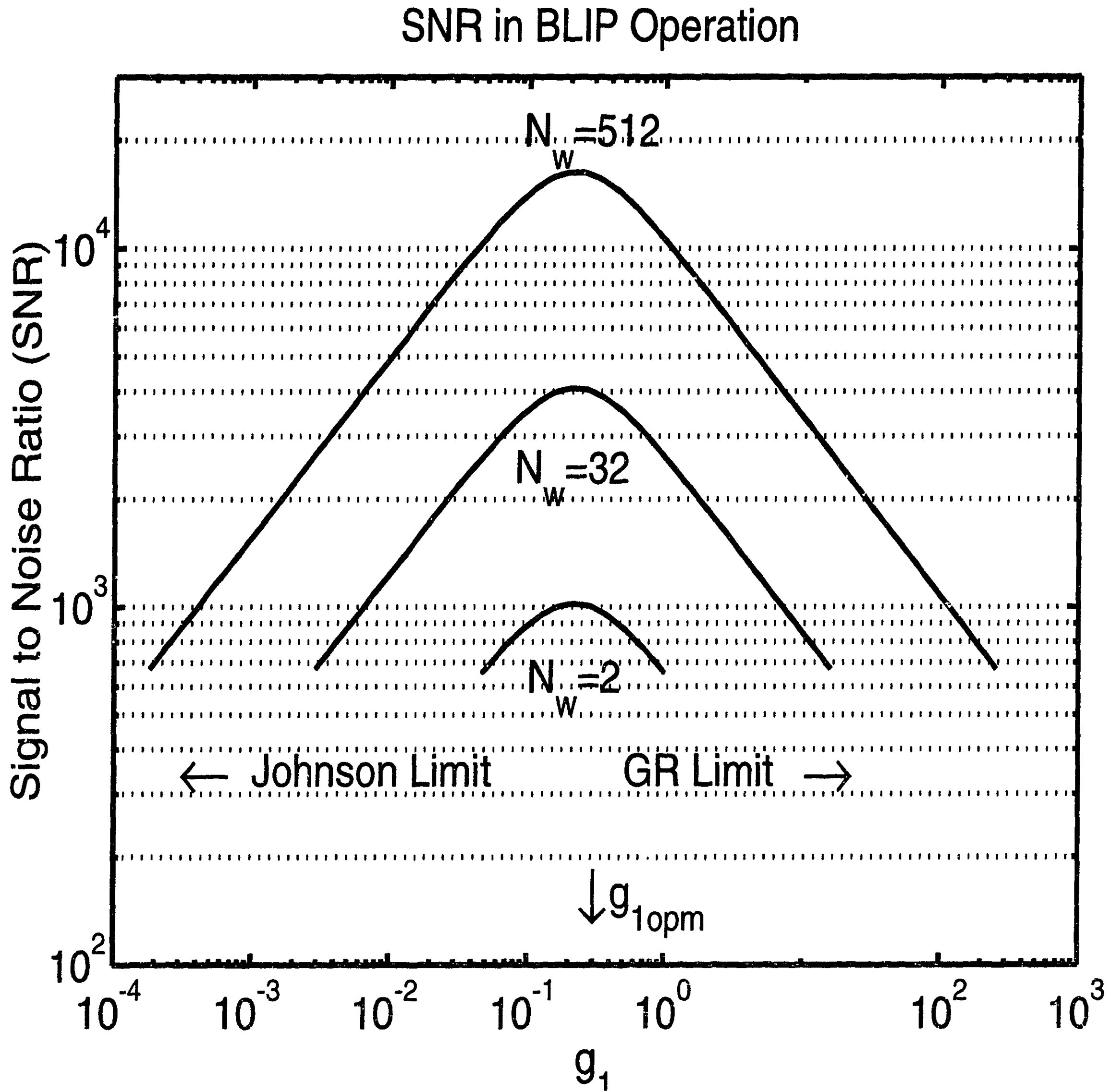


Figure 3-1: The signal to noise ratio (SNR) at very low operating temperatures. At very low operating temperatures (below about 50K), the noise is dominated by the generation-recombination (GR) noise and the Johnson noise. The device parameters and operating conditions used to calculate these curves are given in the text. Of interest is the fact that there is an optimal (single period) photoconductive gain g_{1opm}, below which the Johnson noise is very large, and above which the GR noise is very large.

The fixed pattern noise degrades the expected SNR by the amount given by the second term in the denominator of Equation (3.20).

The SNR is large when the ratio of the conversion efficiency to the nonuniformity of the leakage current, $\eta g/\delta_L i_L$, is large. A good value of $\eta g/\delta_L i_L$ is $250/\mu$A for a SNR of 4000.

The SNR is a maximum at $N_w = 1/\eta_1$, which is the characteristic absorption length in the QWIP. This makes sense because most of the incident radiation is absorbed over the characteristic absorption length, and any additional quantum wells are exposed to very little of the incident radiation. When the total number of quantum wells is greater than $1/\eta_1$, the total photoconductive gain drops as $1/N_w$ but the total quantum efficiency is no longer increasing (see Equation (2.2)). When the total number of quantum wells is less than $1/\eta_1$, the conversion efficiency is independent of the number of quantum wells, and the SNR is limited by nonuniformity, which, by the central limit theorem of statistics, increases as $1/\sqrt{N_w}$.

This is shown in Fig. 3-2 for the following typical numbers for an InGaAs pQWIP: T_B=300K, T_P=1690K, $T_{op}=T_{BLIP}$=68K, ΔT_{op}=50mK, and an initial intended design criterion of nedt=12mK, η_1=0.35%, g_1=0.23, $\delta_L = 3(\exp\{[12.15K/T_{op}]^2\} - 1)^{1/2}/N_w$, L_P=50 nm, Δf=60 Hz, N_e=4.5×10^5, F=1.26×10^{11}/s (300K target temperature viewed through f/2 lens with 70% lens transmission with peak responsivity at 8.55μm), A=(40μm)^2.

3.3.3 Thermal Leakage Current Limited Devices

Near the BLIP temperature, the SNR is limited by the random generation and recombination of both the photocurrent and the leakage current carriers,

$$SNR = \frac{i_P}{\{4qg(i_L + i_P)\Delta f\}^{1/2}}. \qquad (3.21)$$

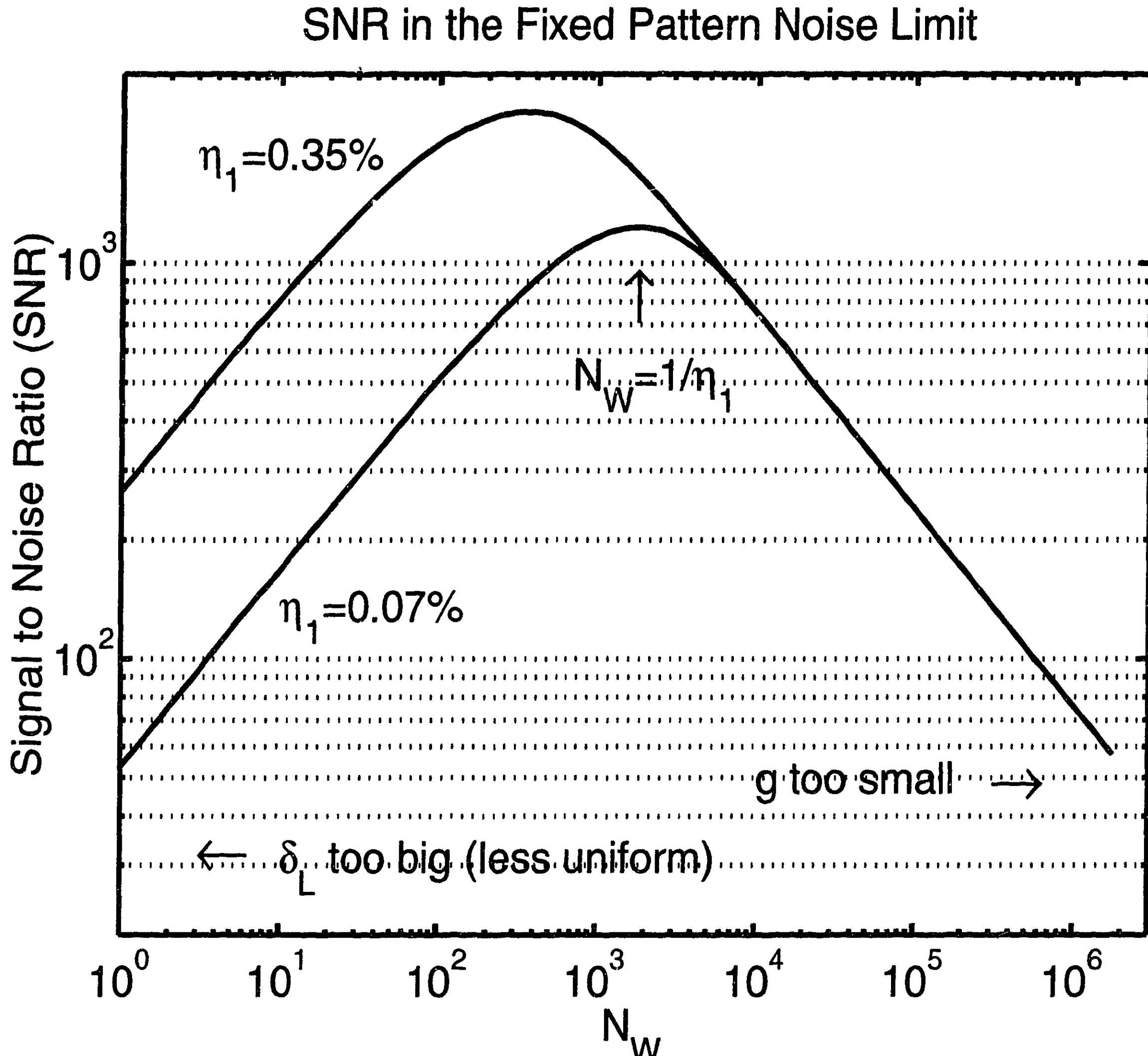


Figure 3-2: The dependence on number of quantum wells of the signal to noise ratio (SNR) for fixed pattern noise limited devices. At intermediate temperatures, QWIP FPA noise could be dominated by fixed pattern noise. In this operating regime, the optimal number of quantum wells to use in a QWIP is determined by minimizing both the GR and the fixed pattern noise. Using fewer quantum wells than this optimal number will make the total quantum efficiency small, according to (2.2). Using too many quantum wells will not increase the total quantum efficiency beyond one hundred percent, but will continue to decrease the total photoconductive gain. The device parameters and operating conditions used to calculate these curves are given in the text.

It is useful to express this SNR in terms of the minimum resolvable target temperature (mrt) that would have resulted, had there been no thermal leakage noise,

$$SNR = \frac{1}{\left\{ \left[\frac{4qgi_L \Delta f}{i_P^2} \right] + \left[\frac{T_P}{T_B^2}(mrt) \right]^2 \right\}^{1/2}}. \tag{3.22}$$

The thermal leakage noise degrades the expected SNR by the amount given by the second term in the denominator of Equation (3.22).

The SNR is large when the second term in the denominator of Equation (3.22) is small. This is equivalent to a large ratio for $\eta^2 g/i_L$.

The SNR is a maximum at $N_w = 2/3\eta_1$, which is roughly the characteristic absorption length in the QWIP. As with the analogous argument for fixed pattern noise limited devices, this makes sense because most of the incident radiation is absorbed over the characteristic absorption length, and any additional quantum wells are exposed to very little of the incident radiation. When the total number of quantum wells is greater than $2/3\eta_1$, the total photoconductive gain drops as $1/N_w$ but the total quantum efficiency is no longer increasing. When the total number of quantum wells is less than $1/\eta_1$, the SNR is limited by the total quantum efficiency which increases as N_w. This is shown in Fig. 3-3 for the following typical numbers for an InGaAs pQWIP: T_B=300K, T_P=1690K, T_{op}=77K, T_{BLIP}=68K, ΔT_{op}=50mK, an initial intended design criterion of mrt=12mK, η_1=0.35%, g_1=0.23, L_P=50 nm, Δf=60 Hz, N_e=4.5×10^5, F=1.26×10^{11}/s (300K target temperature viewed through f/2 lens with 70% lens transmission with peak responsivity at 8.55μm), A=(40μm)^2.

3.4 Conclusions

This section has focussed on the QWIP SNR, and the noise contributions to it. A theoretical study has been made of different noise mechanisms which contribute to QWIP performance: generation-recombination noise, fixed pattern noise, thermal leakage random arrival noise, Johnson noise, read-out (switch) noise, and photon random arrival noise. A key result found in this work is that when the signal-to-

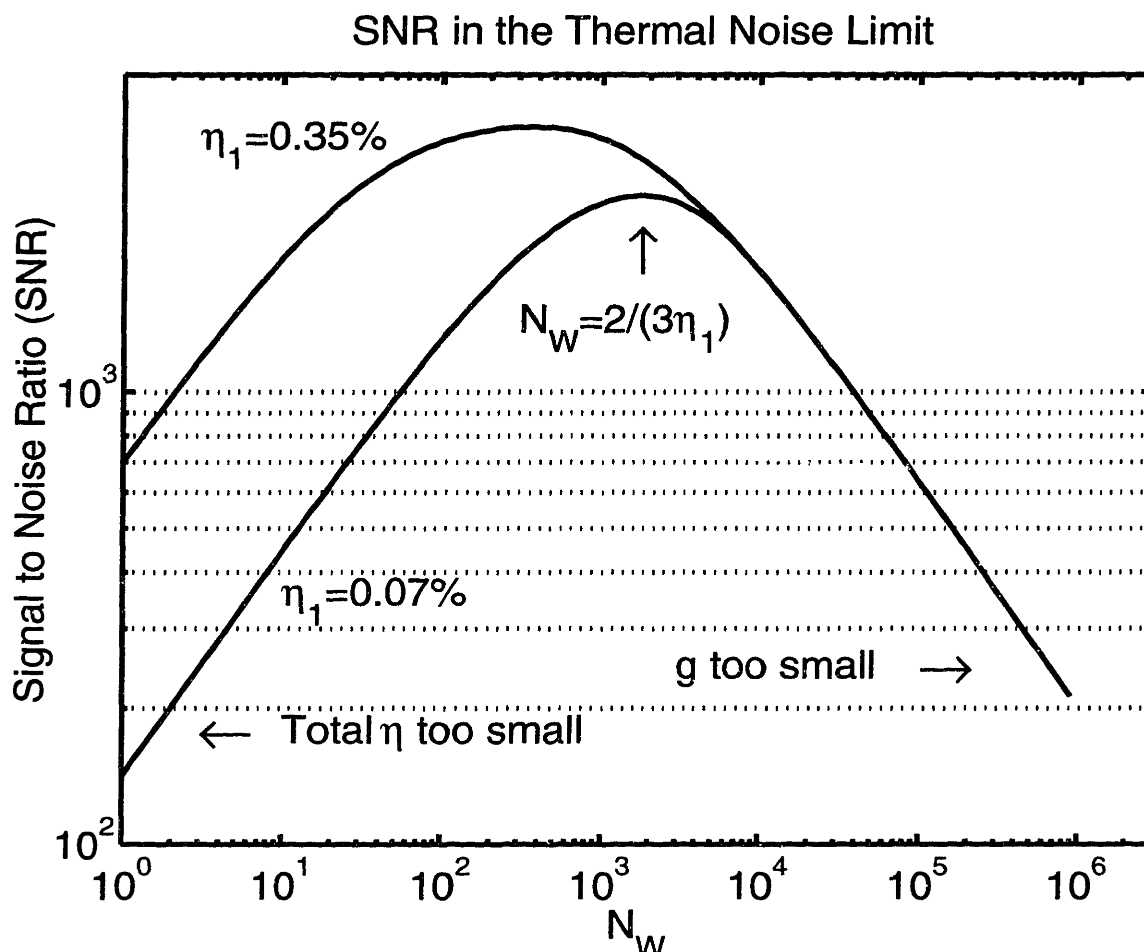


Figure 3-3: The dependence on number of quantum wells of the signal to noise ratio (SNR) for thermal leakage noise limited devices. At intermediate and higher temperatures, QWIP FPA noise could be dominated by the random creation and destruction of both the photocurrent and the thermal leakage carriers. In this operating regime, the optimal number of quantum wells to use in a QWIP is determined by minimizing both the GR and the thermal leakage noise. As with the fixed pattern noise limited devices, the use of too few quantum wells than this optimal number will make the total quantum efficiency small, according to (2.2). Use of too many quantum wells will not increase the total quantum efficiency beyond one hundred percent, but will continue to decrease the total photoconductive gain. The device parameters and operating conditions used to calculate these curves are given in the text.

noise ratio (SNR) is limited by either fixed pattern noise or thermal leakage arrival noise, the optimal number of quantum wells for a maximum in the expected QWIP SNR is roughly η_1^{-1}, where η_1 is the quantum efficiency of a QWIP having only one quantum well. The use of a much larger number of quantum wells is not desirable because the absorption quantum efficiency cannot be increased beyond 100 %, but the photoconductive gain drops with an increase in the number of quantum wells. The use of a much smaller number than the optimal number of quantum wells is not desirable because the absorption quantum efficiency can still be increased considerably with an increase in the number of quantum wells.

SNR is a relevant device figure of merit, as it is a macroscopic device parameter which is directly measured. Many other device figures of merit, such as noise equivalent power, noise equivalent irradiance, and minimum resolvable temperature, are related to the SNR in a very simple way. For example, the minimum resolvable temperature (MRT), a very convenient measure of detector focal plane array performance, describes the smallest change in the target temperature that can be sensed by QWIPs in a FPA in the presence of a total QWIP noise of $\langle i_n \rangle$, and it is related to the SNR through

$$MRT \equiv \frac{\langle i_n \rangle}{d\langle i_P \rangle / dT} = \frac{T_B^2}{T_P} \frac{1}{SNR}, \tag{3.23}$$

where the angle brackets indicate averaging over the different pixels in the FPA.

Fig. 3-4 shows the temperature dependence of various contributions to the QWIP MRT for the following typical numbers for an InGaAs pQWIP: T_B=300K, T_P=1690K, T_{BLIP}=68K, ΔT_{op}=50mK, η_1=0.35%, g_1=0.23, L_P=50 nm, Δf=60 Hz, N_e=4.5×10^5, $\delta_L = 3(\exp\{[12.15K/T_{op}]^2\} - 1)^{1/2}/N_w$, F=1.26×10^{11}/s (300K target temperature viewed through f/2 lens with 70% lens transmission with peak responsivity at 8.55μm), A=(40μm)^2. Fig. 3-4 shows that at very low temperatures, GR noise is the limiting noise mechanism. At higher temperatures, thermal leakage noise becomes important. Depending on the size of the nonuniformity of the leakage current across a QWIP FPA, the fixed pattern noise could also be important at intermediate and higher temperatures.

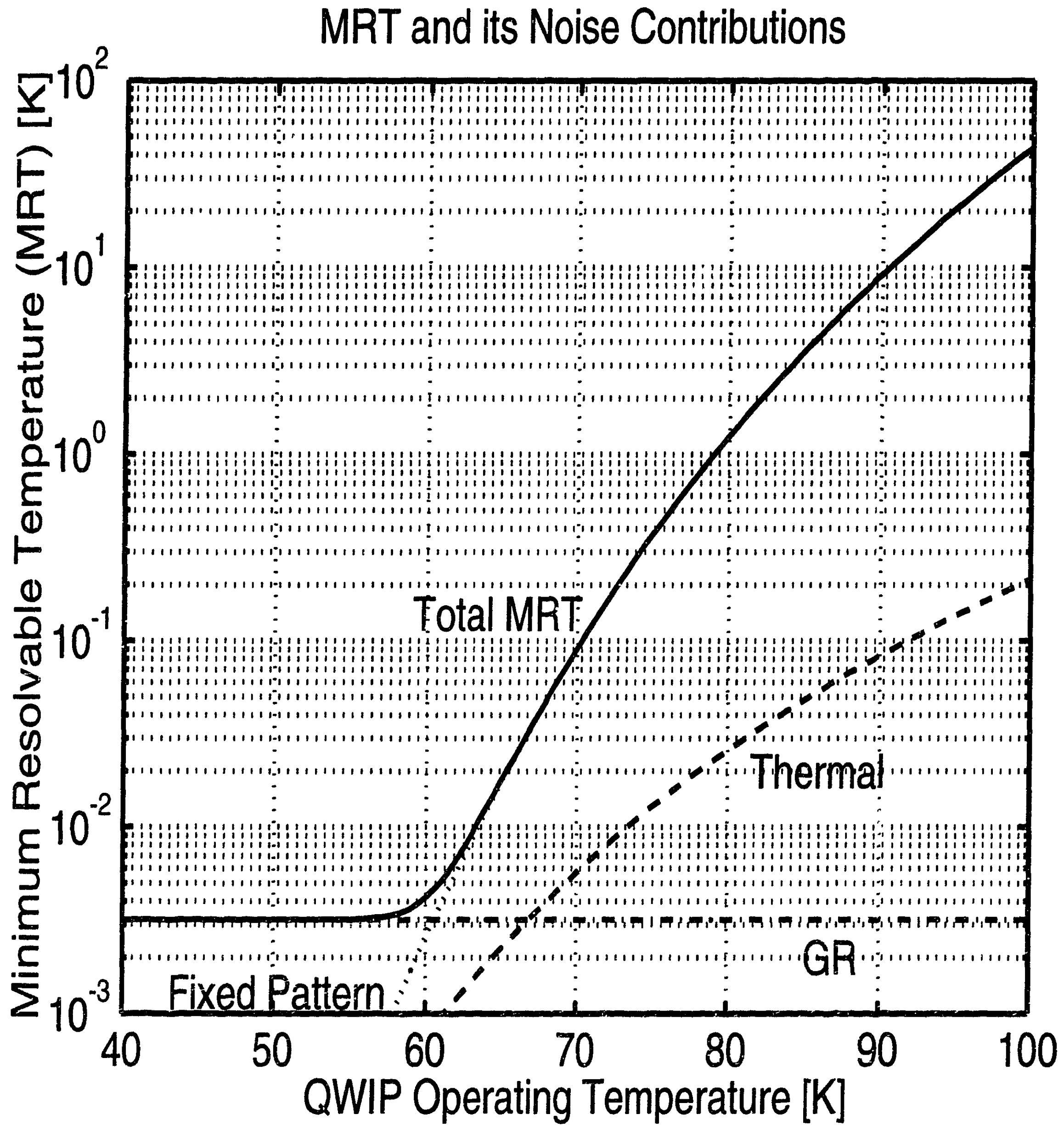


Figure 3-4: The temperature dependence of various contributions to the QWIP minimum resolvable temperature. The device parameters and operating conditions used to calculate these curves are given in the text. Under these operating conditions, the generation-recombination noise is the limiting noise mechanism at low temperatures. At intermediate and higher temperatures, the limiting noise mechanism is the fixed pattern noise resulting from nonuniformity of the leakage current across a QWIP focal plane array.

Chapter 4

The Intersubband Absorption Strength

It is desirable to study QWIPs which respond strongly to normally incident radiation without the use of an optical grating, since the elimination of the processing steps for the optical grating results in lower cost, higher yield, and more uniform QWIP FPAs. This requires the study of the microscopic physics of quantum wells in order to elucidate the physical origin of the polarization selection rules of intersubband absorption.

The well known selection rules and optical dipoles for a conduction intersubband transition are reviewed in Section 4.2.3. The possibility of the absorption of normally incident radiation by electrons in an n-QWIP is also investigated in that section.

A key contribution of this thesis was the derivation of selection rules within the framework of $\vec{k}\cdot\vec{p}$ theory for the intersubband absorption of normally incident radiation by holes in a p-QWIP (p-doped QWIP) in the absence of an optical grating. This is done in Section 4.2.4. It was found that the absorption of normally incident radiation by holes in a p-QWIP in the absence of an optical grating is largest for transitions from a heavy hole state to a light hole state. This result is compared with the literature.

It was also found that $\vec{k}\cdot\vec{p}$ theory predicts that uniaxial strain does not have a

large effect on the strength or the selection rules of intersubband absorption because the Hamiltonian describing uniaxial strain has the same (tetragonal) symmetry as that describing the confinement of carriers in the quantum wells along the growth direction. This is discussed in Section 4.2.4, where we also relate the transport and leakage current properties of p-QWIPs to the hole subband dispersion relations.

4.1 The Polarization of Incident Radiation and QWIP Operating Geometries

Two commonly used geometries for operating QWIPs are shown in Figure 4-1. In both QWIP operating geometries, the electron transport is along the growth direction.

The waveguide geometry is shown on the left side of the figure. In the waveguide geometry, radiation is incident upon the edge of the sample wafer. This radiation propagates in a direction parallel to the epitaxial layers and perpendicular to the growth direction. Moreover, if the electric field is polarized in the growth direction, the radiation is TM polarized. The intersubband absorption of TM polarized radiation is very large. Unfortunately, the necessity of TM polarized radiation to propagate in a direction perpendicular to the growth direction is inconvenient for operating large arrays of devices.

The operation of a large array of QWIPs requires that the incident radiation propagate in a direction normal to the array and parallel to the growth direction. This QWIP operating geometry is shown on the right side of Fig. 4-1. Consequently, the electric field is in the plane of the quantum wells and is perpendicular to the growth direction. Thus, normally incident radiation is always TE polarized. The absorption of TE polarized radiation is small in n-QWIPs, as discussed in Section 4.2.3, but could be large in p-QWIPs, as discussed in Section 4.2.4. The fabrication of an optical grating on the semiconductor surface is required in order to couple normally incident radiation into an optical mode which is TM polarized within the semiconductor.

It is desirable to study QWIPs which respond strongly to normally incident radiation

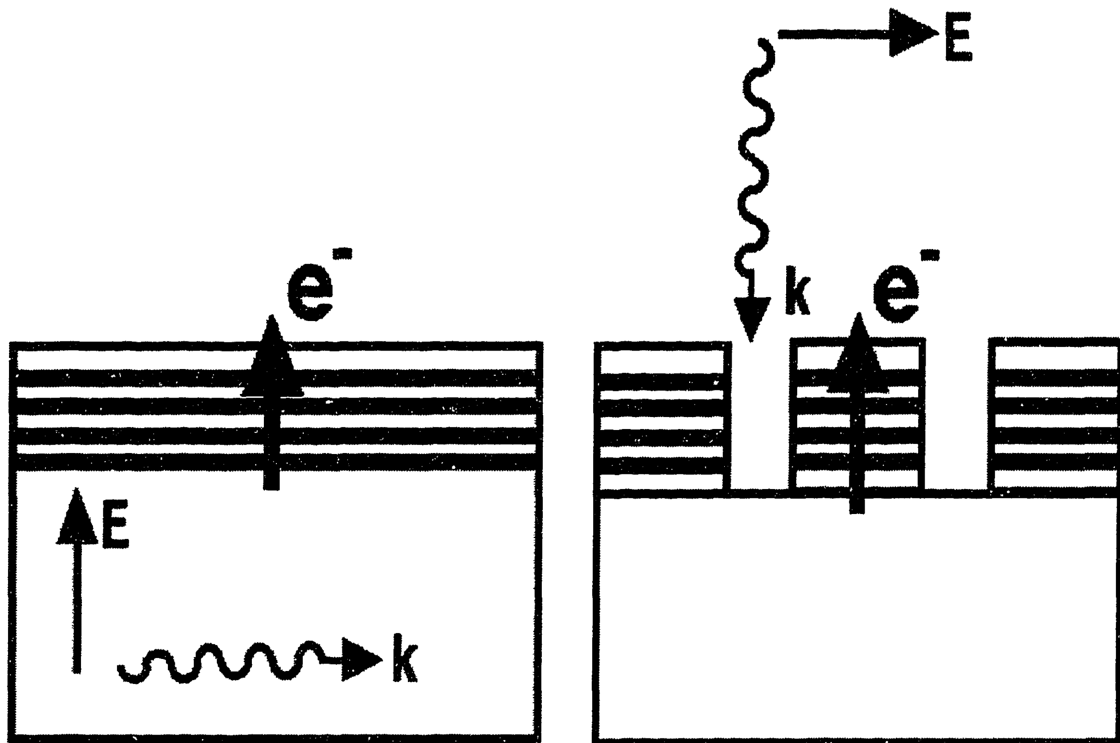


Figure 4-1: Two commonly used geometries for operating QWIPs. The waveguide geometry is shown in the left schematic in the figure. In the waveguide geometry, radiation is incident upon the edge of the sample wafer. This radiation propagates in a direction parallel to the epitaxial layers and perpendicular to the growth direction. Moreover, if the electric field is polarized in the growth direction, the radiation is TM polarized. The intersubband absorption of TM polarized radiation is very large. The second geometry for operating QWIPs is shown on the right side of the figure. In this geometry, radiation is normally incident on the sample wafer with the radiation wave vector parallel to the growth direction. Consequently, the electric field is in the plane of the quantum wells and is perpendicular to the growth direction. This is TE polarized radiation. The absorption of TE polarized radiation is small in n-QWIPs but could be large in p-QWIPs. This second QWIP operating geometry, in which radiation is normally incident on the sample surface, is convenient for operating large focal plane arrays of devices. In both QWIP operating geometries, the electron transport is along the growth direction.

without the use of an optical grating, since the elimination of the processing steps for the optical grating results in lower cost, higher yield, and more uniform QWIP focal plane arrays (FPAs). This requires the study of the physical origin of the polarization selection rules of intersubband absorption, as is done in the rest of this chapter.

4.2 $\vec{k} \cdot \vec{p}$ Theory

In order to understand the size of and polarization selection rules of intersubband absorption, one must understand the electronic wave functions in semiconductor quantum wells. The electron and hole wave functions in quantum wells contain information about both the bulk semiconductor which comprises the quantum well, as well as the confining nature of the potential barriers surrounding the quantum well.

Many III-V compounds, such as GaAs, have the zinc-blende crystal structure, shown schematically in Figure 4-2 below. There are 8 electrons per unit cell (3 from the Ga and 5 from the As) which contribute to chemical bonding, and these arise from sp^3 orbitals. In a zinc-blende crystal, where the sp^3 orbitals from neighboring atoms mix, these hybridized sp^3 orbitals form broad bands of energy whose associated wave functions have a known symmetry. As discussed in Section 4.2.1 below, $\vec{k} \cdot \vec{p}$ theory gives explicitly the symmetry of a wave function associated with a particular energy in the broad energy bands. It is well known that the bonding s-orbitals in GaAs are tightly bond, and completely filled (2 electrons per unit cell). At zero Kelvin (O K), the bonding p orbitals are completely filled with the remaining 6 electrons per unit cell. These bonding p orbitals comprise the *valence* band. At zero Kelvin, the higher energy anti-bonding levels are unoccupied. The lowest energy anti-bonding states are denoted the *conduction* band, and the lowest conduction band is known to have s symmetry.

A quantum well differs from a bulk semiconductor in that along the growth direction, the quantum well has only a finite spatial extent, beyond which a barrier material has been grown by molecular beam epitaxy (MBE). This is shown schematically in Figure 4-3, where an InGaAs quantum well has been grown between two confining

electron: s (antibonding) state

hole: p (bonding) state

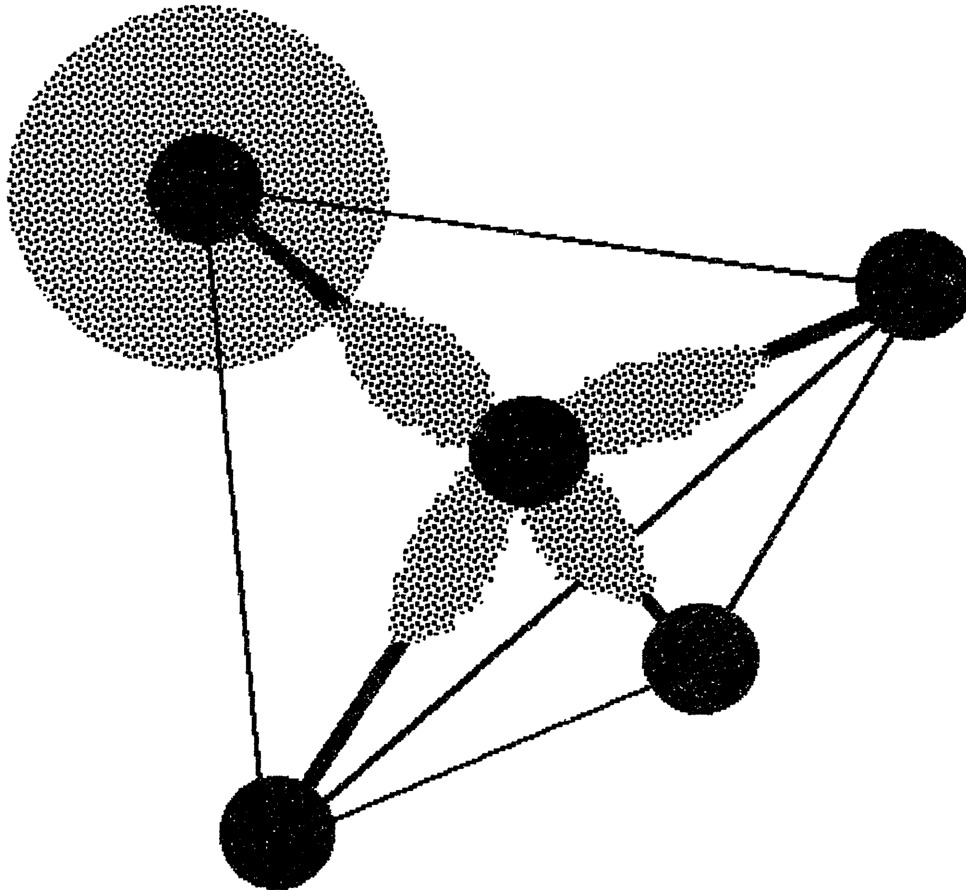


Figure 4-2: Schematic diagram showing the zinc-blende primitive cell. Also shown schematically are the $\vec{\mathbf{k}} = \vec{0}$ lowest energy unoccupied conduction band anti-bonding $|S\rangle$ state and the $\vec{\mathbf{k}} = \vec{0}$ highest energy occupied valence band bonding $|P\rangle$ state ($|X\rangle$, $|Y\rangle$ or $|Z\rangle$).

AlGaAs barrier layers. The confinement of electrons or holes to the quantum well by the difference in the quantum well and the barrier energy band edges can be described by the "envelope" portion of the full wave function. This envelope function is slowly varying over a single unit cell, and is a solution of the effective mass equation, as described in Section 4.2.2 below. A quantum well which is symmetric about its center will be described by envelope functions which have a definite (either even or odd) parity. Both the parity of the envelope function and the symmetry of the Bloch function will determine the intersubband selection rules, as described in Sections 4.2.3 and 4.2.4 below.

4.2.1 The Kane Hamiltonian and the Bloch Functions

Bloch's theorem can be used to understand the symmetry of the wave functions in a bulk semiconductor. Bloch's theorem states that the translational symmetry of a bulk semiconductor allows the wave functions $\Psi_{n,\vec{\mathbf{k}}}(\vec{\mathbf{r}})$ in the semiconductor to be written in the Bloch form,

$$\Psi_{n,\vec{\mathbf{k}}}(\vec{\mathbf{r}}) = \exp(i\vec{\mathbf{k}} \cdot \vec{\mathbf{r}}) u_{n,\vec{\mathbf{k}}}(\vec{\mathbf{r}}), \tag{4.1}$$

where $\vec{\mathbf{k}}$ is an eigenvalue of the translation operator, where n is a band index, and where the Bloch functions, $u_{n,\vec{\mathbf{k}}}(\vec{\mathbf{r}})$, have the periodicity of the lattice.

In a classic paper, E. O. Kane [23] has shown that the Bloch function at any particular value of the translation eigenvalue $\vec{\mathbf{k}}$ can be expanded in the complete basis formed by the Bloch functions at the zone center $\vec{\mathbf{k}} = \vec{\mathbf{0}}$,

$$u_{n,\vec{\mathbf{k}}}(\vec{\mathbf{r}}) = \sum_{m} c_{mn}(\vec{\mathbf{k}}) u_{m,\vec{\mathbf{0}}}(\vec{\mathbf{r}}). \tag{4.2}$$

In many applications, only a small range of $\vec{\mathbf{k}}$ is accessible to free carriers. In such situations, the summation in Equation (4.2) can be approximated by including only a few bands, such as the conduction and valence bands only. The coefficients c_{mn} in Equation (4.2), for m,n=conduction and valence bands only, can then be found from degenerate perturbation theory.

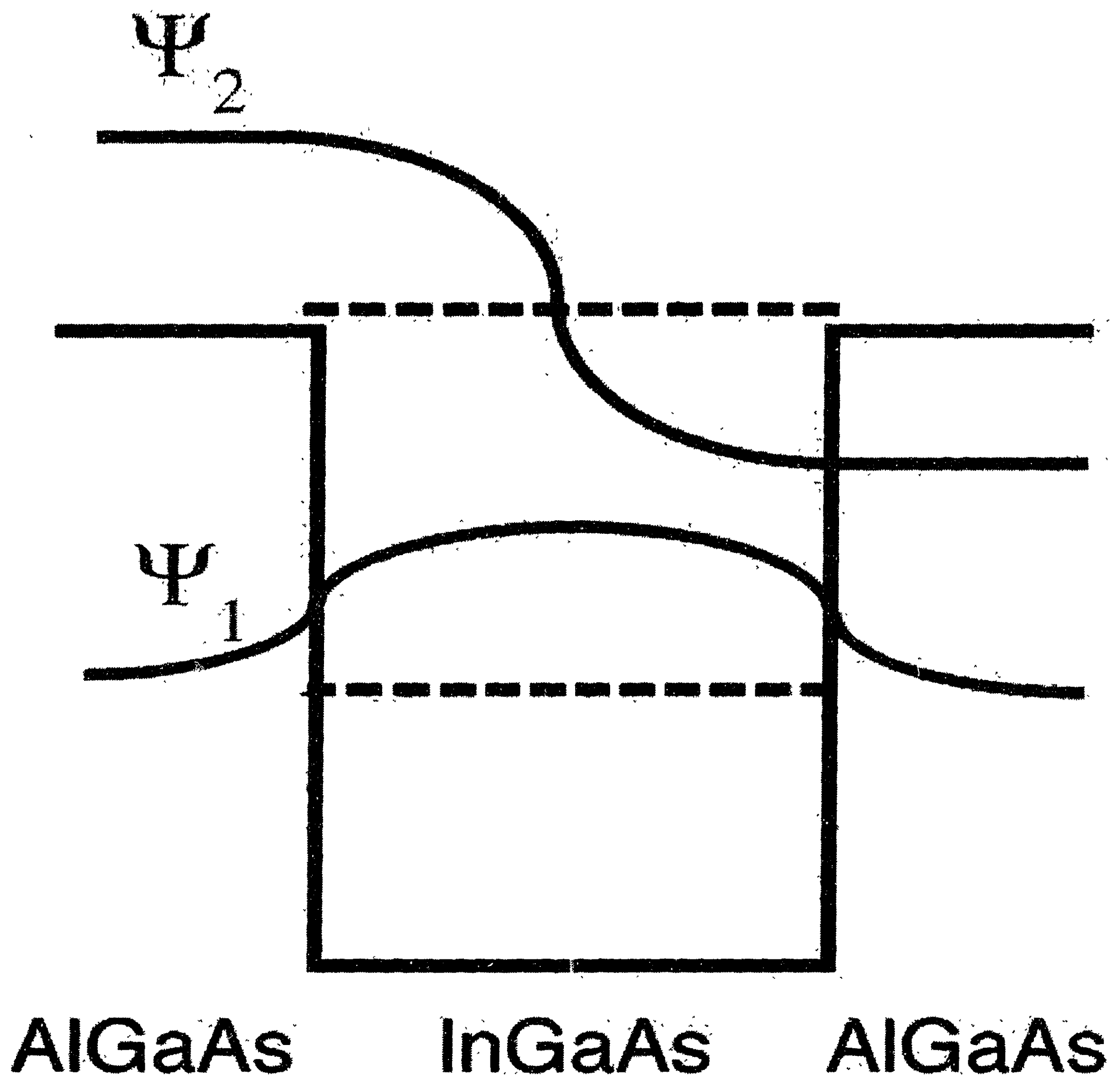


Figure 4-3: Schematic diagram showing a single quantum well and the two confining barriers surrounding it. Also shown is the envelope function which is slowly varying over a single zinc-blende unit cell and which is a solution of the effective mass equation describing the molecular beam epitaxially grown quantum well and barriers.

The Kane $\vec{k}\cdot\vec{p}$ perturbation Hamiltonian [23] is, for $k_x = k_y = 0$ and finite k_z,

$$\mathcal{H} = \begin{bmatrix} H & 0 \\ 0 & H \end{bmatrix}, \tag{4.3}$$

in the basis $|iS\downarrow\rangle$, $\left|\frac{1}{\sqrt{2}}(X - iY)\uparrow\right\rangle$, $|Z\downarrow\rangle$, $\left|\frac{1}{\sqrt{2}}(X + iY)\uparrow\right\rangle$, $|iS\uparrow\rangle$, $\left|\frac{-1}{\sqrt{2}}(X + iY)\downarrow\right\rangle$, $|Z\uparrow\rangle$, $\left|\frac{1}{\sqrt{2}}(X - iY)\downarrow\right\rangle$, respectively, where

$$H = \begin{bmatrix} E_s & 0 & Pk_z & 0 \\ 0 & E_p - \Delta/3 & \sqrt{2}\Delta/3 & 0 \\ Pk_z & \sqrt{2}\Delta/3 & E_p & 0 \\ 0 & 0 & 0 & E_p + \Delta/3 \end{bmatrix}, \tag{4.4}$$

where E_s, E_p are the eigenvalues of H in the absence of the $\vec{k}\cdot\vec{p}$ perturbation, where $E_s = E_G$ and E_G is the band gap energy, where $E_p = -\Delta/3$ and Δ is the spin-orbit splitting of the valence band,

$$\Delta = \frac{3i\hbar}{4m_0^2c^2}\left\langle Z\left|\frac{\partial V}{\partial x}p_y - \frac{\partial V}{\partial y}p_x\right|Y\right\rangle, \tag{4.5}$$

where V is the periodic crystal potential, where the Kane momentum matrix element P is

$$P = -i\frac{\hbar}{m_0}\langle S|p_z|Z\rangle. \tag{4.6}$$

The spin-orbit interaction results from the interaction of the magnetic field seen in the frame of reference of an accelerating electron with the intrinsic magnetic moment of that electron. Kane shows that the eigenvalues of Equation (4.4) can be found from diagonalizing Equation (4.4),

$$\begin{gathered} E'_{hh} = 0, \\ E'(E' - E_G)(E' + \Delta) - P^2k_z^2(E' + \frac{2}{3}\Delta) = 0 \end{gathered} \tag{4.7}$$

where we have defined

$$E' = E - \frac{\hbar^2k_z^2}{2m_0}. \tag{4.8}$$

In the limit of small k_z, the eigenvalues of the Kane Hamiltonian, Equation (4.7), yield parabolic bands,

$$\begin{aligned}
E_c &= E_G + \frac{\hbar^2 k_z^2}{2m_0} + \frac{P^2 k_z^2}{3}\left(\frac{2}{E_G} + \frac{1}{E_G + \Delta}\right) \\
E_{hh} &= \frac{\hbar^2 k_z^2}{2m_0} \\
Elh &= \frac{\hbar^2 k_z^2}{2m_0} - \frac{2P^2 k_z^2}{3E_G} \\
E_{soh} &= -\Delta + \frac{\hbar^2 k_z^2}{2m_0} - \frac{P^2 k_z^2}{3(E_G + \Delta)}.
\end{aligned} \tag{4.9}$$

Diagonalization [23] of the Hamiltonian in Equation (4.3) yields the following expressions for the Bloch functions,

$$\begin{aligned}
u_{i,\vec{\mathbf{k}}} &= a_i|iS\uparrow\rangle + b_i\left|\frac{-1}{\sqrt{2}}(X+iY)\downarrow\right\rangle + c_i|Z\uparrow\rangle \\
u_{hh,\uparrow,\vec{\mathbf{k}}} &= \left|\frac{1}{\sqrt{2}}(X+iY)\uparrow\right\rangle \\
u_{i,\vec{\mathbf{k}}} &= a_i|iS\downarrow\rangle + b_i\left|\frac{1}{\sqrt{2}}(X-iY)\uparrow\right\rangle + c_i|Z\downarrow\rangle \\
u_{hh,\downarrow,\vec{\mathbf{k}}} &= \left|\frac{1}{\sqrt{2}}(X-iY)\downarrow\right\rangle,
\end{aligned} \tag{4.10}$$

where the index i=c,lh,soh denotes the band, and where the coefficients a_i, b_i, c_i are,

$$\begin{aligned}
a_i &= \frac{1}{N}\left[Pk_z(E_i' + \frac{2}{3}\Delta)\right] \\
b_i &= \frac{1}{N}\left[\frac{\sqrt{2}\Delta}{3}(E_i' - E_G)\right] \\
c_i &= \frac{1}{N}\left[(E_i' - E_G)(E_i' + \frac{2}{3}\Delta)\right],
\end{aligned} \tag{4.11}$$

where N is a normalization constant equal to the square root of the sum of the squares of the numerators in Equation (4.11). In the limit of zero wave vector, the

Bloch functions of Equation (4.10) have the well known form,

$$\begin{aligned}
\text{conduction band } &= \left|S, \frac{1}{2}, \frac{1}{2}\right\rangle = u_{c,\vec{0}} = |iS\uparrow\rangle \\
\text{light hole band } &= \left|P, \frac{3}{2}, \frac{1}{2}\right\rangle = u_{lh,\vec{0}} = +\frac{1}{\sqrt{3}}\left|\frac{-1}{\sqrt{2}}(X+iY)\downarrow\right\rangle + \sqrt{\frac{2}{3}}|Z\uparrow\rangle \\
\text{split off hole band } &= \left|P, \frac{1}{2}, \frac{1}{2}\right\rangle = u_{soh,\vec{0}} = \sqrt{\frac{2}{3}}\left|\frac{-1}{\sqrt{2}}(X+iY)\downarrow\right\rangle - \frac{1}{\sqrt{3}}|Z\uparrow\rangle \\
\text{heavy hole band } &= \left|P, \frac{3}{2}, \frac{3}{2}\right\rangle = u_{hh,\vec{0}} = \left|\frac{1}{\sqrt{2}}(X+iY)\uparrow\right\rangle, \qquad (4.12)
\end{aligned}$$

(and similarly for the other 4 Bloch wave functions), where we have explicitly written the Bloch functions at $\vec{\mathbf{k}} = \vec{\mathbf{0}}$ as eigenfunctions, $|J, J_z\rangle$, of total (spin plus orbital) angular momentum J and its z-component J_z.

For a finite value of k_z, Equation (4.10) can be used to write the conduction band Bloch function in terms of the Bloch functions at zero wave vector,

$$\begin{aligned}
u_{c,\vec{\mathbf{k}}} &= \frac{1}{N}\left(|iS\uparrow\rangle + \frac{Pk_z(E_c' + \frac{2}{3}\Delta)}{E_c'(E_c'+\Delta)}|Z\uparrow\rangle + \frac{Pk_z}{E_c'}\frac{\sqrt{2}\Delta}{3(E_c'+\Delta)}\left|\frac{-1}{\sqrt{2}}(X+iY)\downarrow\right\rangle\right) \\
&\approx \frac{1}{N}\left(|iS\uparrow\rangle + \frac{Pk_z}{E_G}|Z\uparrow\rangle + \frac{Pk_z}{E_G}\frac{\sqrt{2}\Delta}{3(E_G+\Delta)}\left|\frac{-1}{\sqrt{2}}(X+iY)\downarrow\right\rangle\right) \qquad (4.13)
\end{aligned}$$

where $E_c' > E_G$ is measured from the heavy hole band edge $E_{hh}(\vec{\mathbf{k}} = \vec{\mathbf{0}})$ at the zone center. Equation (4.10) can also be used to write the light hole Bloch function, at a finite value of k_z, in terms of the Bloch functions at zero wave vector,

$$\begin{aligned}
u_{lh,\vec{\mathbf{k}}} &= \frac{1}{N}\left(|Z\uparrow\rangle + \frac{Pk_z}{E_{lh}' - E_G}|iS\uparrow\rangle + \frac{\sqrt{2}\Delta}{3(E_{lh}' + \frac{2}{3}\Delta)}\left|\frac{-1}{\sqrt{2}}(X+iY)\downarrow\right\rangle\right) \\
&\approx \frac{1}{N}\left(\sqrt{\frac{2}{3}}|Z\uparrow\rangle + \sqrt{\frac{1}{3}}\left|\frac{-1}{\sqrt{2}}(X+iY)\downarrow\right\rangle - \sqrt{\frac{2}{3}}\frac{Pk_z}{E_G}|iS\uparrow\rangle\right) \qquad (4.14)
\end{aligned}$$

where E_{lh}' is measured from the heavy hole band edge $E_{hh}(\vec{\mathbf{k}} = \vec{\mathbf{0}})$ at the zone center. Equations (4.13) and (4.14) are key to understanding Bloch function symmetries and the intersubband absorption selection rules.

4.2.2 Envelope Functions and the Effective Mass Equation

Luttinger and Kohn [24, 25] have shown that in the presence of a weak and slowly varying potential $V(\vec{\mathbf{r}})$, the envelope function (portion of the total wave function which is slowly varying over a unit cell) satisfies the effective mass equation,

$$\sum_m [H_{mn}(-i\nabla) + V(\vec{\mathbf{r}})\delta_{mn}] = f_m(\vec{\mathbf{r}}) = Ef_n(\vec{\mathbf{r}}) \tag{4.15}$$

where $H_{mn}(-i\nabla)$ is the $\vec{k}\cdot\vec{p}$ Hamiltonian, as in Equation (4.3) above. The total wave functions can be written as linear combinations of the envelope functions and the zone center Bloch functions,

$$\Psi_n(\vec{\mathbf{r}}) = \sum_m c_{mn}(\vec{\mathbf{k}})f_m(\vec{\mathbf{r}})u_{m,\vec{0}}(\vec{\mathbf{r}}). \tag{4.16}$$

If the total wave function is dominated by only one term in Equation (4.16), as is the case for conduction band electrons and light and heavy holes for small k, then the effective mass equation reduces to a scalar form,

$$\left[-\frac{\hbar^2}{2}\nabla\cdot\left(\frac{1}{m^*}\nabla\right) + V(\vec{\mathbf{r}})\right] f_n(\vec{\mathbf{r}}). = Ef_n(\vec{\mathbf{r}}). \tag{4.17}$$

This is an important result, as it says that the details of a weak and slowly varying crystal potential can be summarized by the parameter m^*.

Boundary Conditions At the boundary z=0 between layer A and layer B, the envelope function is continuous as is $1/\mathrm{m}^*$ times its first derivative,

$$f_n^A(-\epsilon) = f_n^B(+\epsilon) \tag{4.18}$$

$$\frac{1}{m_A^*}\frac{\partial f_n^A}{\partial z}_{-\epsilon} = \frac{1}{m_B^*}\frac{\partial f_n^B}{\partial z}_{+\epsilon} \tag{4.19}$$

The boundary condition of continuity of $1/\mathrm{m}^*$ times the first derivative of the envelope function is equivalent to requiring conservation of probability current.

For conduction band electrons, White and Sham [26] show that Equation (4.19) is equivalent to requiring that the $|S\rangle$ symmetric and $|P\rangle$ symmetric parts of the electron

wave functions (the first two terms in Equation (4.20)) are *each* continuous across the boundary of the heterojunction.

Equation (4.19) is not the only possible boundary condition. Since Schrodinger's equation requires that the *full* wave function (not just the envelope function) and its first derivative be continuous across a boundary, Equation (4.19) needs to be revised to account for the Bloch function. There have been many proposed methods to do this, all with some degree of success. The question of more sophisticated boundary conditions is beyond the scope of this work, but can be found in the literature [27, 28, 29, 30, 31, 32, 33].

4.2.3 Conduction Band Intersubband Transitions

The total wave function for a conduction band electron of quantum number n is in the Kane model,

$$\Psi_{c,n} = \frac{1}{N}\left(|iS\uparrow\rangle f_n(\vec{\mathbf{r}}) - \frac{P}{E_G}|Z\uparrow\rangle i\nabla f_n(\vec{\mathbf{r}}) + \frac{P}{E_G}\frac{\sqrt{2}\Delta}{3(E_G+\Delta)}\left|\frac{X+iY}{\sqrt{2}}\downarrow\right\rangle i\nabla f_n(\vec{\mathbf{r}})\right), \tag{4.20}$$

where N is a normalization constant. For quantum wells of uniform composition in the absence of any external applied potential, the envelope function $f_n(\vec{\mathbf{r}})$ is a linear combination of plane waves.

The Optical Dipole for Absorption of TM Polarized Radiation

In the limit of infinitely large potential barriers surrounding the quantum well located between z=0 and z=L_W, the envelope function for conduction band electrons is,

$$f_n(z) = \sqrt{\frac{2}{L_W}}\sin(n\pi z/L_W), \tag{4.21}$$

where n is the integral quantum number. The literature shows that the optical dipole for the absorption of TM polarized radiation (electric field parallel to the growth direction) is dominated by the first term in Equation (4.20), and for a transition from

a bound state of quantum number n to one of quantum number m, the optical dipole $\langle c, m|z|c, n\rangle$ has the value [21],

$$\langle c, m|z|c, n\rangle = \begin{cases} Lw\frac{8}{\pi^2}\frac{mn}{(m^2-n^2)^2} & \text{for m-n odd,} \\ 0 & \text{otherwise,} \end{cases} \tag{4.22}$$

which also shows the well known selection rule requiring that m-n be odd for conduction intersubband transitions in a symmetric square well potential.

The optical dipole $\langle z\rangle$ is related to the imaginary part χ'' of the susceptibility through [34],

$$\chi'' = \frac{e^2|\langle z\rangle|^2}{2\epsilon_0\hbar}(N_1 - N_2)g(\nu), \tag{4.23}$$

where e is the electronic charge, ϵ_0 is the dielectric constant of free space, $g(\nu)$ is the absorption lineshape as a function of the frequency ν, and $N_1 - N_2$ is the population difference between the lower state (N_1) and the upper state (N_2). The absorption coefficient $\gamma(\nu)$ is defined as the $\exp(-\gamma(\nu)z)$ exponential decay of the power, and is related to χ'' through [34],

$$\gamma(\nu) = \frac{k_{fs}\chi''(\nu)}{n_r}, \tag{4.24}$$

where k_{fs} is the wave vector in free space of the incident radiation, and n_r is the refractive index of the semiconductor. The optical dipole is related to the Einstein A coefficient for the spontaneous emission rate through [34],

$$A = \frac{e^2\omega^3 n_r^3}{3\pi c^3\hbar\epsilon}|\langle z\rangle|^2, \tag{4.25}$$

where ω is the frequency of the radiation, and ϵ is the dielectric constant of the semiconductor.

The Effect of Higher (Anti-Bonding $|P\rangle$ Symmetric) Conduction Bands

In the recent QWIP literature, there has been some discussion [35, 36, 37, 38, 39, 40] as to whether the conduction band Bloch function contains any symmetries other

than the $|S\rangle$ and the light hole Bloch function symmetries. For example, Yang et al. [39, 40] and Flatte et al. [35] have considered the amount of the conduction band Bloch function which is an admixture of the upper conduction (anti-bonding $|P\rangle$) states. This is accomplished by considering a $\vec{k}\cdot\vec{p}$ Hamiltonian which explicitly includes the effect of the upper conduction bands,

$$\mathcal{H} =$$

$$\begin{vmatrix}
G_1' & 0 & 0 & \frac{1}{\sqrt{3}}P_1k_- & \frac{1}{3}\bar{\Delta} & \frac{1}{\sqrt{3}}Qk_z & 0 & 0 & 0 & 0 & -\frac{\sqrt{2}}{\sqrt{3}}P_1k_z & 0 & -\frac{\sqrt{2}}{\sqrt{3}}Qk_- & Qk_+ \\
 & G_1' & 0 & -P_1k_+ & -\frac{1}{\sqrt{3}}Qk_z & \frac{1}{3}\bar{\Delta} & -\frac{\sqrt{2}}{\sqrt{3}}Qk_z & 0 & 0 & 0 & 0 & -\frac{\sqrt{2}}{\sqrt{3}}Qk_- & 0 & \frac{1}{\sqrt{3}}Qk_- \\
 & & E_1' & \frac{\sqrt{2}}{\sqrt{3}}P_1k_- & 0 & \frac{\sqrt{2}}{\sqrt{3}}Qk_z & -\frac{2}{3}\bar{\Delta} & 0 & 0 & 0 & \frac{1}{\sqrt{3}}P_1k_z & -Qk_+ & \frac{1}{\sqrt{3}}Qk_- & 0 \\
 & & & 0 & \frac{1}{\sqrt{3}}P_0k_+ & -P_0k_- & \frac{\sqrt{2}}{\sqrt{3}}P_0k_+ & \frac{\sqrt{2}}{\sqrt{3}}P_1k_z & 0 & -\frac{1}{\sqrt{3}}P_1k_z & 0 & \frac{\sqrt{2}}{\sqrt{3}}P_0k_z & 0 & -\frac{1}{\sqrt{3}}P_0k_z \\
 & & & & E_0' & 0 & 0 & 0 & \frac{\sqrt{2}}{\sqrt{3}}Qk_- & -Qk_+ & -\frac{\sqrt{2}}{\sqrt{3}}P_0k_z & 0 & 0 & 0 \\
 & & & & & E_0' & 0 & \frac{\sqrt{2}}{\sqrt{3}}Qk_- & 0 & -\frac{1}{\sqrt{3}}Qk_- & 0 & 0 & 0 & 0 \\
 & & & & & & G_0' & Qk_+ & -\frac{1}{\sqrt{3}}Qk_- & 0 & \frac{1}{\sqrt{3}}P_0k_z & 0 & 0 & 0 \\
 & & & & & & & G_1' & 0 & 0 & \frac{1}{\sqrt{3}}P_1k_+ & \frac{1}{3}\bar{\Delta} & \frac{1}{\sqrt{3}}Qk_z & 0 \\
 & & & & & & & & G_1' & 0 & P_1k_- & -\frac{1}{\sqrt{3}}Qk_z & \frac{1}{3}\bar{\Delta} & -\frac{\sqrt{2}}{\sqrt{3}}Qk_z \\
 & & & & & & & & & E_1' & \frac{\sqrt{2}}{\sqrt{3}}P_1k_+ & 0 & \frac{\sqrt{2}}{\sqrt{3}}Qk_z & -\frac{2}{3}\bar{\Delta} \\
 & & & & & & & & & & 0 & \frac{1}{\sqrt{3}}P_0k_- & P_0k_+ & \frac{\sqrt{2}}{\sqrt{3}}P_0k_- \\
 & & & & & & & & & & & E_0' & 0 & 0 \\
 & & & & & & & & & & & & E_0' & 0 \\
 & & & & & & & & & & & & & G_0'
\end{vmatrix}, \tag{4.26}$$

where G_1', E_1', E_0', G_0', are, respectively, the heavy (and light) electron energy, the spin-orbit electron energy, the heavy (and light) hole energy, and the spin-orbit hole band energy, where $k_\pm = (k_x \pm ik_y)/\sqrt{2}$, where [41] the hole spin-orbit splitting is Δ_0=-0.341 eV, where the upper conduction band spin-orbit splitting is Δ_1=0.171 eV, where the so-called interband spin-orbit splitting is $\bar{\Delta}$=-0.061 eV,
$E_{P_0}=2m_0P_0^2/\hbar^2$=27.86 eV, $E_{P_1}=2m_0P_1^2/\hbar^2$=2.36 eV, $E_Q=2m_0Q^2/\hbar^2$=15.56 eV, where for Equation (4.26) we have written the basis in the order, u_1, u_9, u_3, u_{11}, u_5, u_{13}, u_7,

u_8, u_2, u_{10}, u_4, u_{12}, u_6, u_{14}, where

$$
\begin{aligned}
\text{light electron band} &= \left|P', \frac{3}{2}, \frac{1}{2}\right\rangle = u_1 = +\frac{1}{\sqrt{3}}\left|\frac{-1}{\sqrt{2}}(X'+iY')\downarrow\right\rangle + \sqrt{\frac{2}{3}}|Z'\uparrow\rangle \\
\text{heavy electron band} &= \left|P', \frac{3}{2}, \frac{3}{2}\right\rangle = u_2 = \left|\frac{1}{\sqrt{2}}(X'+iY')\uparrow\right\rangle, \\
\text{split off electron band} &= \left|P', \frac{1}{2}, \frac{1}{2}\right\rangle = u_3 = \sqrt{\frac{2}{3}}\left|\frac{-1}{\sqrt{2}}(X'+iY')\downarrow\right\rangle - \frac{1}{\sqrt{3}}|Z'\uparrow\rangle \\
\text{conduction band} &= \left|S, \frac{1}{2}, \frac{1}{2}\right\rangle = u_4 = |iS\uparrow\rangle \\
\text{light hole band} &= \left|P, \frac{3}{2}, \frac{1}{2}\right\rangle = u_5 = +\frac{1}{\sqrt{3}}\left|\frac{-1}{\sqrt{2}}(X+iY)\downarrow\right\rangle + \sqrt{\frac{2}{3}}|Z\uparrow\rangle \\
\text{heavy hole band} &= \left|P, \frac{3}{2}, \frac{3}{2}\right\rangle = u_6 = \left|\frac{1}{\sqrt{2}}(X+iY)\uparrow\right\rangle, \\
\text{split off hole band} &= \left|P, \frac{1}{2}, \frac{1}{2}\right\rangle = u_7 = \sqrt{\frac{2}{3}}\left|\frac{-1}{\sqrt{2}}(X+iY)\downarrow\right\rangle - \frac{1}{\sqrt{3}}|Z\uparrow\rangle \\
\text{light electron band} &= \left|P', \frac{3}{2}, -\frac{1}{2}\right\rangle = u_8 = +\frac{1}{\sqrt{3}}\left|\frac{-1}{\sqrt{2}}(X'-iY')\uparrow\right\rangle + \sqrt{\frac{2}{3}}|Z'\downarrow\rangle \\
\text{heavy electron band} &= \left|P', \frac{3}{2}, -\frac{3}{2}\right\rangle = u_9 = \left|\frac{1}{\sqrt{2}}(X'-iY')\downarrow\right\rangle, \\
\text{split off electron band} &= \left|P', \frac{1}{2}, -\frac{1}{2}\right\rangle = u_{10} = \sqrt{\frac{2}{3}}\left|\frac{-1}{\sqrt{2}}(X'-iY')\uparrow\right\rangle - \frac{1}{\sqrt{3}}|Z'\downarrow\rangle \\
\text{conduction band} &= \left|S, \frac{1}{2}, -\frac{1}{2}\right\rangle = u_{11} = |iS\downarrow\rangle \\
\text{light hole band} &= \left|P, \frac{3}{2}, -\frac{1}{2}\right\rangle = u_{12} = +\frac{1}{\sqrt{3}}\left|\frac{-1}{\sqrt{2}}(X-iY)\uparrow\right\rangle + \sqrt{\frac{2}{3}}|Z\downarrow\rangle \\
\text{heavy hole band} &= \left|P, \frac{3}{2}, -\frac{3}{2}\right\rangle = u_{13} = \left|\frac{1}{\sqrt{2}}(X-iY)\downarrow\right\rangle, \\
\text{split off hole band} &= \left|P, \frac{1}{2}, -\frac{1}{2}\right\rangle = u_{14} = \sqrt{\frac{2}{3}}\left|\frac{-1}{\sqrt{2}}(X-iY)\uparrow\right\rangle - \frac{1}{\sqrt{3}}|Z\downarrow\rangle
\end{aligned}
\tag{4.27}
$$

where $|X\rangle, |Y\rangle, |Z\rangle$ denote the $|P\rangle$-symmetric hole states and $|X'\rangle, |Y'\rangle, |Z'\rangle$ denote the $|P\rangle$-symmetric anti-bonding higher conduction band states, where $\Delta_0, \Delta_1, \bar{\Delta}$ are,

respectively,

$$\Delta_0 = \frac{3i\hbar}{4m_0^2c^2}\left\langle Z\left|\frac{\partial V}{\partial x}p_y - \frac{\partial V}{\partial y}p_x\right|Y\right\rangle,$$
$$\Delta_1 = \frac{3i\hbar}{4m_0^2c^2}\left\langle Z'\left|\frac{\partial V}{\partial x}p_y - \frac{\partial V}{\partial y}p_x\right|Y'\right\rangle,$$
$$\bar{\Delta} = \frac{3i\hbar}{4m_0^2c^2}\left\langle Z\left|\frac{\partial V}{\partial x}p_y - \frac{\partial V}{\partial y}p_x\right|Y'\right\rangle, \tag{4.28}$$

where the different momentum matrix elements are,

$$P_0 = -i\frac{\hbar}{m_0}\langle S|p_z|Z\rangle,$$
$$P_1 = -i\frac{\hbar}{m_0}\langle S|p_z|Z'\rangle,$$
$$Q = -i\frac{\hbar}{m_0}\langle X|p_y|Z'\rangle = i\frac{\hbar}{m_0}\langle X'|p_y|Z\rangle. \tag{4.29}$$

In the literature [41, 42], the Hamiltonian in Equation (4.26) is often solved numerically for either the band structure or the energy dependent (nonparabolic) effect mass. However, the approach of this thesis is to avoid an entirely numerical computation of either the wave functions or the optical dipoles. Rather, the approach of this thesis is to use Brillouin-Wigner perturbation theory to write down a *form* for the electron wave functions, which is numerically approximate but from which the symmetries of the wave function is explicit. The solution to the $\vec{k}\cdot\vec{p}$ Hamiltonian, Equation (4.26), which corresponds to the conduction band wave function can be written down by inspection in Brillouin-Wigner [43] perturbation theory,

$$\begin{aligned}\Psi_{c,m} = {} & |iS\uparrow\rangle f_{c,m} + \frac{P_0}{(E_c - E_{lh})}|Z\uparrow\rangle(-i\nabla f_{c,m}) + \frac{\sqrt{2}P_0}{3(E_c - E_{lh})^2}\Delta_0|R_+\downarrow\rangle(-i\nabla f_{c,m}) + \\ & + \frac{P_0k_-}{(E_c - E_{lh})}|R_-\uparrow\rangle f_{c,m} + \frac{P_0k_+}{(E_c - E_{lh})}|R_+\downarrow\rangle f_{c,m} + \\ & + \frac{P_1}{(E_c - E_{le})}|Z'\uparrow\rangle(-i\nabla f_{c,m}) + \frac{\sqrt{2}P_1}{3(E_c - E_{le})^2}\Delta_1|R'_+\downarrow\rangle(-i\nabla f_{c,m}) + \\ & + \frac{P_1k_-}{(E_c - E_{le})}|R'_-\uparrow\rangle f_{c,m} + \frac{P_1k_+}{(E_c - E_{le})}|R'_+\downarrow\rangle f_{c,m},\end{aligned} \tag{4.30}$$

where we have used the notation $R_\pm = \mp(X \pm iY)/\sqrt{2}$, and where the subscript "le" refers to the light electrons. To find the bound states in a quantum well, the boundary

conditions of Section 4.2.2 are invoked on $f_{c,m}$, the envelope function associated with the dominant term in Equation (4.30).

Figures 4-4 and 4-5 show that the analytical, perturbative expansion in Equation (4.30) is a very accurate approximation to the numerically determined eigenfunctions of the Hamiltonian in Equation (4.26). Both figures show the full numerical solution (solid line) and the Brillouin-Wigner perturbative solution (the components of Equation (4.30) as the dashed line) of the 14×14 Hamiltonian in Equation (4.26) for the amount at a finite $\vec{\mathbf{k}}$ of the electron Bloch function which has the light electron, heavy electron, and spin-orbit electron, light hole, heavy hole, and spin-orbit hole Bloch symmetries at $\vec{\mathbf{k}} = \vec{\mathbf{0}}$. The four plots in each figure show the overlap integral ($\langle u_i | u_{c,\vec{\mathbf{k}}} \rangle$ for i=8,2,10,4 in Fig. 4-4, and for i=12,6,14,4 in Fig. 4-5, where the u_i are defined in Equation (4.27)) of the electron Bloch function at a finite $\vec{\mathbf{k}}$ with the Bloch functions at $\vec{\mathbf{k}} = \vec{\mathbf{0}}$. (The dashed lines showing $\langle u_2 | u_{c,\vec{\mathbf{k}}} \rangle$ and $\langle u_6 | u_{c,\vec{\mathbf{k}}} \rangle$ were calculated in second order perturbation theory, and all other quantities were calculated in first order perturbation theory.) The Brillouin-Wigner perturbative solution, which can be written down by inspection, is seen to be a numerically accurate approximation to the full numerical solution.

The Optical Dipole

The conduction band intersubband transitions in quantum wells is known to be strongest [7] for TM polarized radiation, with its large electric field component along the growth direction. The reason is that the conduction band wave function is dominated by the first two terms in (4.30) which have cell periodic Bloch parts that are s-symmetric or p_z-symmetric.

The optical dipole for the intersubband absorption of normally incident radiation can be calculated from Equation (4.30) to be,

$$\langle c, n | x | c, m \rangle = \begin{cases} \langle f_{c,n} | -i\nabla | f_{c,m} \rangle \frac{P}{3(E_c - E_{lh})^2} \Delta_0 \langle iS | x | X \rangle & \text{for m-n=odd} \\ \langle f_{c,n} | f_{c,m} \rangle \frac{P k_x}{(E_c - E_{lh})} \langle iS | x | X \rangle & \text{for m-n=even} \end{cases}, \quad (4.31)$$

where we have made explicit use of the fact that $\langle S | z | Z \rangle = \langle Z | z | S \rangle$, and that $P_0 \gg P_1$.

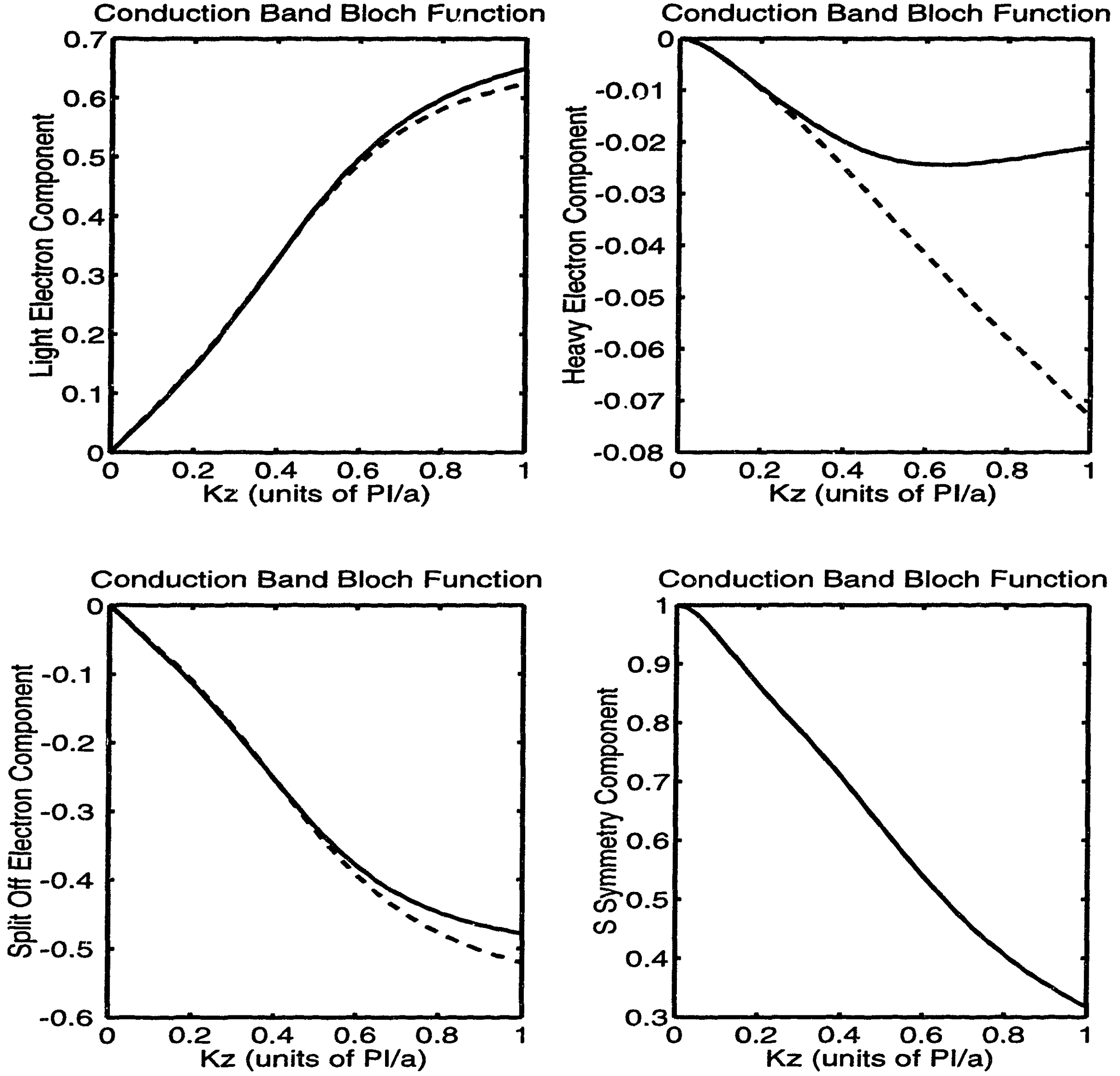


Figure 4-4: Full numerical solution (solid line) and Brillouin-Wigner perturbative solution (the components of Equation (4.30) as the dashed line) of the 14×14 Hamiltonian for the amount at a finite $\vec{k}$ of the electron Bloch function which has the light electron, heavy electron, and spin-orbit electron Bloch symmetries at $\vec{k} = \vec{0}$. The four plots show the overlap integral ($\langle u_i | u_{c,\vec{k}} \rangle$ for i=8,2,10,4, where the u_i are defined in Equation (4.27)) of the electron Bloch function at a finite $\vec{k}$ with the upper conduction Bloch functions at $\vec{k} = \vec{0}$. The units of k_z are π/a, where "a" is the lattice constant. The Brillouin-Wigner perturbative solution, which can be written down by inspection, is seen to be a numerically accurate approximation to the full numerical solution.

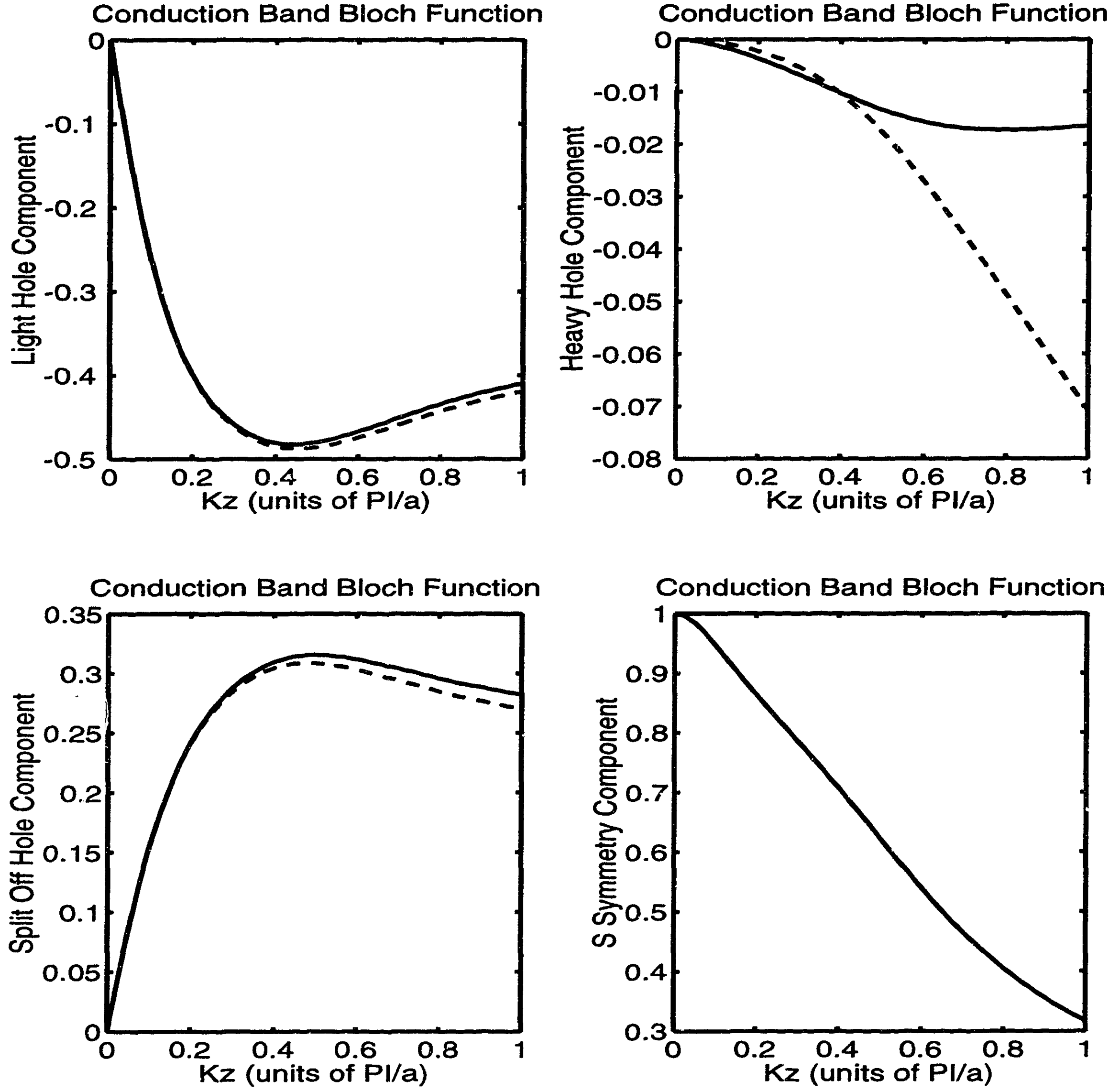


Figure 4-5: Full numerical solution (solid line) and Brillouin-Wigner perturbative solution (the components of Equation (4.30) as the dashed line) of the 14×14 Hamiltonian for the amount at a finite $\vec{k}$ of the electron Bloch function which has the light hole, heavy hole, and spin-orbit hole Bloch symmetries at $\vec{k} = \vec{0}$. The four plots show the overlap integral ($\langle u_i | u_{c,\vec{k}} \rangle$ for i=12,6,14,4 where the u_i are defined in Equation (4.27)) of the electron Bloch function at a finite $\vec{k}$ with the hole Bloch functions at $\vec{k} = \vec{0}$. The units of k_z are π/a, where "a" is the lattice constant. The Brillouin-Wigner perturbative solution, which can be written down by inspection, is seen to be a numerically accurate approximation to the full numerical solution.

The dipoles $\langle f_{c,n}| - i\nabla|f_{c,m}\rangle$ and $\langle iS|x|X\rangle$ are calculated in Equations (4.37) and (4.35) below.

The strength of the intersubband absorption of normally incident radiation depends on the size of the $|X\rangle$ or $|Y\rangle$ component of the electron wave function. For conduction band electrons, the $|X\rangle$ or $|Y\rangle$ component of the wave function is proportional to either $\frac{\Delta_0}{E_c - E_{lh}}$ or $\frac{Pk_x}{E_c - E_{lh}}$, both of which are small, since the spin orbit energy is (for GaAs) Δ_0=340 meV and the quantized conduction subband to valence band energy difference is $E_c - E_{lh}$ ~1.7 eV large, and the in-plane energies, set by the doping levels are also small (~10 meV for doping levels on the order of $10^{12}\mathrm{cm}^{-2}$). This small size of the $|X\rangle$ and $|Y\rangle$ component of the electron wave function is the reason for the weak strength of the intersubband absorption of normally incident radiation.

Equation (4.31) is in qualitative agreement with the numerical work of Flatte et al. [35] and with the direct diagonalization of the Hamiltonian in Equation (4.26) done by Yang [39, 40]. Both Flatte et al. [35] and Yang [39, 40] note the necessity of a finite electron in-plane wave vector for the absorption of normally incident radiation by an n-QWIP without an optical grating. Yang [39, 40] also noted that semiconductors with a large spin-orbit splitting (relative to the band gap energy) would exhibit larger absorption of normally incident radiation without the use of an optical grating, in agreement with our Equation (4.31). Of course, our derivation of Equation (4.31) has the virtue of simplicity, as it was obtained, by inspection, from Brillouin-Wigner perturbation theory. Shik [36, 37, 38] has also arrived at a result very similar to Equation (4.31) by considering the conduction and valence bands only (and not the upper conduction bands).

4.2.4 Valence Band Intersubband Transitions

The heavy hole wave function is

$$\begin{aligned}\Psi_{hh,m} &= (u_{hh,\vec{0}})f_{hh,m} \\ &= \left|\frac{1}{\sqrt{2}}(X + iY)\uparrow\right\rangle f_{hh,m},\end{aligned} \tag{4.32}$$

where we have used Equation (4.12) for $u_{hh,\vec{0}}$, and where the envelope function associated with the quantum number m is $f_{hh,m}$, and the latter has a definite parity (either odd or even) for a symmetric quantum well. The light hole wave function is

$$\Psi_{lh,n} = (u_{lh,\vec{0}})f_{lh,n} + \sqrt{\frac{2}{3}}\frac{P}{E_{lh}-E_c}(u_{c,\vec{0}})\left(-i\frac{\partial f_{lh,n}}{\partial z}\right)$$
$$= \frac{1}{N}\left[\left(\sqrt{\frac{2}{3}}|Z\uparrow\rangle + \sqrt{\frac{1}{3}}\left|\frac{-1}{\sqrt{2}}(X+iY)\downarrow\right\rangle\right)f_{lh,n} - \sqrt{\frac{2}{3}}\frac{P}{E_c-E_{lh}}|iS\uparrow\rangle\left(-i\frac{\partial f_{lh,n}}{\partial z}\right)\right] \tag{4.33}$$

where we have used Equation (4.12) for $u_{lh,\vec{0}}$ and $u_{c,\vec{0}}$, and where the envelope function associated with the quantum number n is $f_{lh,n}$, and the latter has a definite parity (either odd or even) for a symmetric quantum well.

The Optical Dipole within the Kane Hamiltonian

The optical dipole describing the absorption of normally incident, TE circularly polarized radiation from a heavy hole state of quantum number m to a light hole state of quantum number n is,

$$\left\langle lh,n\left|\frac{x-iy}{\sqrt{2}}\right|hh,m\right\rangle = -\frac{1}{N}\left\langle f_{hh,m}\left|-i\frac{\partial}{\partial z}\right|f_{lh,n}\right\rangle\sqrt{\frac{2}{3}}\frac{P}{E_c-E_{lh}}\langle X|x|iS\rangle, \tag{4.34}$$

where the first factor is a normalization constant equal to $N^2 = 1+\frac{2}{3}\left(\frac{Pk_z}{E-E_c}\right)^2$. Equation (4.34) can be evaluated quantitatively by noting that the optical dipole for an interband (conduction band to valence band) transition is known to be [23, 41],

$$\langle X|x|S\rangle = \frac{P}{E_G} \quad (=7.2\text{ Å in GaAs}), \tag{4.35}$$

where we have used the following numbers for GaAs [41]: E_G=1.424 eV and $E_P=2m_0P^2/\hbar^2$=27.86 eV. In the limit of infinitely large potential barriers surrounding the quantum well (located between z=0 and z=L_W) and in the limit of small mixing of the light and heavy hole states, the envelope function for both the light holes and the heavy holes is,

$$f_n(z) = \sqrt{\frac{2}{L_W}}\sin(n\pi z/L_W), \tag{4.36}$$

where n is the integral quantum number. The integral $\left\langle f_{hh,m} \left| -i\frac{\partial}{\partial z} \right| f_{lh,n} \right\rangle$ has already been evaluated in the literature [21] on conduction intersubband transitions in a symmetric square well potential,

$$\left\langle f_{hh,m} \left| -i\frac{\partial}{\partial z} \right| f_{lh,n} \right\rangle = \begin{cases} \frac{-4i}{L_W} \left(\frac{mn}{m^2-n^2}\right) & \text{for m-n odd,} \\ 0 & \text{otherwise.} \end{cases} \tag{4.37}$$

Thus, the optical dipole describing the absorption of normally incident, TE circularly polarized radiation from a heavy hole state of quantum number m to a light hole state of quantum number n is,

$$\left\langle lh, n \left| \frac{x - iy}{\sqrt{2}} \right| hh, m \right\rangle = \begin{cases} \frac{1}{N} \frac{4i}{L_W} \left(\frac{mn}{m^2-n^2}\right) \sqrt{\frac{2}{3} \frac{P^2}{E_G^2}} & \text{for m-n odd,} \\ 0 & \text{otherwise,} \end{cases} \tag{4.38}$$

in the limit of very large potential barriers surrounding a symmetric, square quantum well potential, and in the limit of very little mixing of the light and heavy hole states. When the quantum well cannot be modelled as a symmetric, square quantum well with very large potential barriers surrounding it, the full expression in Equation (4.34) must be used (together with Equation (4.35)) with the integral $\left\langle f_{hh,m} \left| -i\frac{\partial}{\partial z} \right| f_{lh,n} \right\rangle$ explicitly evaluated.

Equation (4.38) is a key result of this work. This result has never been explicitly stated in the QWIP literature, but is implicit in the original Kane [23] $\vec{k} \cdot \vec{p}$ theory. (The only analytical expression we have found in the literature for the optical dipole in an intervalence band transition is that of Sugimura [44], who estimated the optical dipole for a heavy hole to *spin-orbit hole* transition in a *bulk* semiconductor.) Equation (4.38) shows that the absorption of normally incident radiation by holes in a p-QWIP in the absence of an optical grating is largest for transitions from a heavy hole state to a light hole state. The absorption of normally incident radiation by holes is also largest for transitions in which n-m=1 because of the term in parentheses (which arises from the symmetry of the envelope portion f_n of the bound states in a symmetric, square quantum well) in Equation (4.38). In fact, for a transition from the m=1 heavy hole state to the n=2 light hole state, Equation (4.38) evaluates to

$$\langle lh, 2|x|hh, 1\rangle = \text{ 2.3 Å for a } L_W\text{=50Å GaAs quantum well,} \tag{4.39}$$

which is comparable to the optical dipole for a standard interband transition, Equation (4.35). For comparison, it is well known [21] that the optical dipole for the intersubband absorption of TM polarized radiation from the n=1 to the n=2 conduction subbands is $\langle c,2|x|c,1\rangle = 0.18L_W$=8.8Å for a L_W=50Å GaAs quantum well in the limit of infinitely large confining potential barriers surrounding the quantum well.

The value of knowing explicitly the different symmetries contained in the light and heavy hole wave functions, Equations (4.33) and (4.32) respectively, lies in being able to understand how to increase the hole intersubband absorption of normally incident radiation. One common suggestion in the literature is to design devices that increase the mixing of the light and heavy hole states. From the expressions derived above for the light and heavy hole wave functions (Equations (4.32) and (4.33)), it should be clear that the absorption of normally incident radiation is largest for a transition from a *pure* heavy hole to a *pure* light hole state. A large mixing of the light and heavy hole states does NOT increase the absorption of normally incident radiation because the absorption of normally incident radiation must involve a state having $|S\rangle$ symmetry. (The latter is a result of the following argument. The optical dipole for the absorption of normally incident radiation polarized in the $\hat{x}$ direction is nonzero only for the term $\langle X|x|S\rangle$ because all other terms are zero by symmetry: $\langle X|x|X\rangle = \langle X|x|Y\rangle = 0$.) Since light hole states have much more $|S\rangle$ character than heavy hole states and since the heavy hole state has no $|S\rangle$ character within the Kane model (see (4.32)), the absorption of normally incident radiation is strongest for a transition involving a pure light hole state. If this light hole state is now strongly mixed with a heavy hole state, as in energy degenerate perturbation theory, the strongly mixed states, denoted as Ψ_1, Ψ_2, must have the form,

$$\Psi_{1,2} = \frac{1}{\sqrt{2}}(\Psi_{lh} \pm \Psi_{hh}), \tag{4.40}$$

where Ψ_{lh}, Ψ_{hh} are the pure light hole and pure heavy hole states existing in the absence of the perturbation. Clearly, Equation (4.40) shows that the mixed states Ψ_1, Ψ_2 each have half of the original light hole character Ψ_{lh}.

Equation (4.38) is consistent with recent numerical work [45, 46] in the literature. Chang [46] et al. even state in the beginning of their paper that heavy hole to light

	InAs	GaAs	AlAs	InP
energy gap E_g (eV)	0.36	1.424	2.15	1.35
electron effective mass m_e/m_0	0.023	0.063	0.14	0.08
heavy hole effective mass m_{hh}/m_0	0.6	0.62	0.76	0.85
light hole effective mass m_{lh}/m_0	0.027	0.082	0.15	0.089
electron mobility μ_e (cm^2/V-s)	28000	8500		4500
hole mobility μ_h (cm^2/V-s)	450	400		130
static dielectric constant ϵ_s/ϵ_0	14.6	12.85	8.2	12.4
high frequency dielectric constant $\epsilon_{\text{inf}}/\epsilon_0$	12.3	10.95	10.1	9.55

Table 4.1: Some energy band, electrical, and optical parameters for a selection of III-V semiconductors.

hole transitions are allowed in the bulk for both polarizations of incident radiation.

Much of the theoretical work [46] on hole intersubband transitions has been numerical, and has focussed on holes with a large value of in-plane wave vector. Following this theoretical work, much of the original experimental work done by Levine [7, 47] also emphasized p-QWIP designs where the hole in-plane wave vector is designed to be large, either through heavy doping of the quantum wells or through the use of asymmetric quantum wells. However, the selection rules and optical dipoles derived above show that the heavy hole to light hole absorption is strong even in the limit of very small hole in-plane wave vector. The reason is that the absorption is strong when there exists a large confinement wave vector, k_z in Equation (4.38) (describing the amount of the light hole Bloch function which has the $|S\rangle$ symmetry). Moreover, even for a large doping of $1\times10^{12}\text{cm}^{-2}$ in the quantum wells, the hole in-plane wave vector k_x is still much smaller than the confinement wave vector k_z ($\frac{k_x}{k_z} = \sqrt{\frac{E_F}{E_{\text{upper}}}} \approx \sqrt{\frac{3.8\text{ meV}}{200\text{ meV}}}$, where E_F, E_{upper} are respectively the Fermi energy and the upper bound state energy). (We have used the band structure parameters given in Table 4.1.) Thus, it is the large size of the confinement wave vector k_z of the upper bound state in a quantum well which allows the light hole state to have some $|S\rangle$ character and which allows the optical dipole in Equation (4.38) to be sizeable.

The central idea is that strong hole intersubband absorption of normally incident radiation must involve a light hole state. This light hole state should be chosen to have a large admixture of $|S\rangle$ symmetry through either a large k_z or a large k_x,k_y. In quantum wells illuminated by a moderately powerful radiation source, such as a room temperature black body, the light holes involved in the optical transition have a k_z much larger than k_x,k_y, so that just being a bound state in a quantum well will introduce a large admixture of $|S\rangle$ symmetry to a light hole state.

From the form of Equations (4.32) and (4.33), it is clear that heavy hole to heavy hole transitions and light hole to light hole transitions (in the limit of small mixing of the light and heavy holes) follow the usual intersubband selection rule of being nominally forbidden for normally incident radiation in the absence of an optical grating. Thus, the optical dipole describing the absorption of TM polarized radiation has the well known [21] form,

$$\begin{aligned}\langle hh,m|z|hh,n\rangle &= \langle lh,m|z|lh,n\rangle \\ &= \begin{cases} L_W \frac{8}{\pi^2}\frac{mn}{(m^2-n^2)^2} & \text{for m-n odd,} \\ 0 & \text{otherwise,}\end{cases}\end{aligned} \tag{4.41}$$

for a symmetric, square well potential in the limit of infinitely large potential barriers (and in the limit of small mixing of the light and heavy holes). Equation (4.41) shows that the optical dipole drops as $1/m^3$ for a transition from a heavy hole state of quantum number n=1 to a heavy hole state of (large) quantum number m. Thus, the absorption of TM polarized radiation from a heavy hole ground state of quantum number n=1 to a heavy hole upper state in the continuum above the barrier band edge could be weaker than the absorption of TE polarized radiation from the same heavy hole ground state to a *light* hole state in the same continuum because the upper light hole state would have a smaller quantum number m and the latter transition also obeys a different selection rule, Equation (4.38) (and not Equation (4.41)). This may explain why Levine [7] has observed the absorption of TE polarized radiation to be twice as large as that of TM polarized radiation for one of his p-QWIPs which was measured in the waveguide geometry.

The value of these selection rules is in designing QWIPs which have a reproducible

absorption spectrum. The latter requires the choice of an intersubband transition which is strongly absorbing and *not* one which is nominally forbidden. For example, a good choice for a hole intersubband transition which exhibits a strong absorption of normally incident radiation would be an intersubband transition involving a (pure) light hole state. Equation (4.38), which is a new result of this work, is an easy selection rule to remember, and it is consistent with existing numerical and experimental work on p-QWIPs.

The Hole Subband Dispersion Relations

The Luttinger-Kohn $\vec{k} \cdot \vec{p}$ Hamiltonian [48] for the hole subbands with the Pikus-Bir description of uniaxial strain is, in the absence of interaction with the conduction band states,

$$H = \begin{pmatrix} a_+ & b & ib/\sqrt{2} & -i\sqrt{2}c & c & 0 \\ b^* & a_- & if & i\sqrt{\frac{3}{2}}b & 0 & c \\ -ib^*/\sqrt{2} & -if^* & d & 0 & i\sqrt{\frac{3}{2}}b & i\sqrt{2}c \\ i\sqrt{2}c^* & -i\sqrt{\frac{3}{2}}b^* & 0 & d & -if^* & ib/\sqrt{2} \\ c^* & 0 & -i\sqrt{\frac{3}{2}}b^* & if & a_- & -b \\ 0 & c^* & -i\sqrt{2}c^* & -ib^*/\sqrt{2} & -b^* & a_+ \end{pmatrix}, \tag{4.42}$$

where

$$a_+ = -\frac{\hbar^2}{2m_0}\left[(\gamma_1+\gamma_2)(k_x^2+k_y^2)+(\gamma_1-2\gamma_2)k_z^2\right]+V_{hh} \tag{4.43}$$

$$a_- = -\frac{\hbar^2}{2m_0}\left[(\gamma_1-\gamma_2)(k_x^2+k_y^2)+(\gamma_1+2\gamma_2)k_z^2\right]+V_{lh} \tag{4.44}$$

$$b = i\sqrt{3}\frac{\hbar^2}{m_0}\gamma_3 k_z(k_x-ik_y) \tag{4.45}$$

$$c = -\frac{\sqrt{3}}{2}\frac{\hbar^2}{m_0}\left[\gamma_2(k_x^2-k_y^2)-2i\gamma_3 k_x k_y\right] \tag{4.46}$$

$$d = -\frac{\hbar^2}{2m_0}\gamma_1(k_x^2+k_y^2+k_z^2)+Aa_v\epsilon-\Delta_{so}+V_{so} \tag{4.47}$$

$$f = \sqrt{2}\frac{\hbar^2}{2m_0}\gamma_2(k_x^2+k_y^2-2k_z^2)-\frac{\sqrt{2}}{3}Bb_v\epsilon \tag{4.48}$$

$$A = 2\left(\frac{C_{11}-C_{12}}{C_{11}}\right) \tag{4.49}$$

$$B = 3\left(\frac{C_{11}+2C_{12}}{C_{11}}\right) \tag{4.50}$$

$$\epsilon = \frac{a_{\|}-a}{a}, \tag{4.51}$$

where m_0 is the free electron mass, V_{hh}, V_{lh} are respectively the heavy hole and light hole band edges, $\gamma_1, \gamma_2, \gamma_3$ are Luttinger parameters, a_v and b_v are, respectively, the hydrostatic and uniaxial valence band deformation potentials, ϵ is the strain, C_{11} and C_{12} are elastic constants, $a_{\|}$ and a are the strained and unstrained in-plane lattice constants, and where the order of the total angular momentum $|J, J_z\rangle$ basis used in (4.42) is: $|3/2,+3/2\rangle$ (heavy hole), $|3/2,+1/2\rangle$ (light hole), $|1/2,+1/2\rangle$ (spin orbit hole), $|1/2,-1/2\rangle$ (spin orbit hole), $|3/2,-1/2\rangle$ (light hole), and $|3/2,-3/2\rangle$ (heavy hole).

A detailed discussion of the changes in the hole wave function symmetry which result from strain and which are relevant to hole intersubband transitions has been lacking. The hole wave functions in a strained quantum well can be found from diagonalizing the Pikus-Bir Hamiltonian [49] $H_{PB}^{(001)}$: for growths on a (001) substrate, $H_{PB}^{(001)} = -a\sum_i \epsilon_{ii} - b\sum_i \epsilon_{ii}[J_i^2 - \frac{1}{3}J^2] - \frac{2d}{\sqrt{3}}[\epsilon_{xy}\{J_xJ_y\} + c.p.]$, where ϵ_{ij}, i=x,y,z is the strain tensor, a is the hydrostatic deformation potential, b, d are uniaxial deformation potentials, where J, J_i are the total angular momentum and the component of the angular momentum along the direction i(=x,y,z) respectively, and "c.p." stands for

cyclic permutation. Terms in the Pikus-Bir Hamiltonian [49] containing ϵ_{ij} appear in exactly the same way as terms in the Luttinger-Kohn $\vec{k} \cdot \vec{p}$ Hamiltonian containing $k_i k_j$. This is a reflection of the fact that both quantum confinement and axial strain in the growth direction choose the growth direction as the direction of physical significance, and both effects impart the multiple quantum well with the same (tetragonal) symmetry.

This symmetry allows the form of the Pikus-Bir Hamiltonian to be written down directly, with the mixing of the $|S\rangle$ and $|Z\rangle$ states expressed as $\langle S|H_{PB}^{(001)}|Z\rangle = Pk_z + P\epsilon_{zz}k_z$ for strained layers grown on an (001) substrate. Thus, the optical dipole describing absorption from a heavy hole to a light hole state in strained epilayers grown on an (001) substrate can be obtained from Equation (4.38) by replacing the Kane $\vec{k} \cdot \vec{p}$ perturbative energy Pk_z in the absence of strain with the sum $Pk_z + P\epsilon_{zz}k_z$. Therefore, the heavy hole to light hole transition dipole in strained quantum wells is larger than the value without strain given in Equation (4.38) by the factor,

$$\frac{|\langle lh|x|hh\rangle| \text{ with strain}}{|\langle lh|x|hh\rangle| \text{ without strain}} = 1 + \epsilon_{zz}. \tag{4.52}$$

For typical strained InGaAlAs layers grown on either InP or GaAs substrates, ϵ_{zz} is less than 0.01.

Solution of the Pikus-Bir Hamiltonian shows that the hole states which are light hole-like along the growth direction have heavy hole-like dispersion relations in-plane. In other words, $E = \frac{\hbar^2(k_x^2+k_y^2)}{2m_1} + \frac{\hbar^2 k_z^2}{2m_2}$, where $m_1 = m_0/(\gamma_1 + \gamma_2)$ and $m_2 = m_{hh} = m_0/(\gamma_1 - 2\gamma_2)$ for heavy holes, and $m_1 = m_0/(\gamma_1 - \gamma_2)$ and $m_2 = m_{lh}/(\gamma_1 + 2\gamma_2)$ for light holes.

4.3 Conclusions

Studies of the microscopic physics of quantum wells have been presented to elucidate the physical origin of the intersubband absorption of normally incident radiation. It is desirable to study QWIPs which respond strongly to normally incident radiation

without the use of an optical grating, since the elimination of the processing steps for the optical grating results in lower cost, higher yield, and more uniform QWIP FPAs.

A key contribution of this thesis was the derivation, in Equation (4.38), of selection rules within the framework of $\vec{k}\cdot\vec{p}$ theory for the intersubband absorption of normally incident radiation by holes (in the limit of small hole in-plane wave vector) in a p-QWIP (p-doped QWIP) in the absence of an optical grating. It was found that the absorption of normally incident radiation by holes in a p-QWIP in the absence of an optical grating is largest for transitions from a heavy hole state to a light hole state. This result is implicit in the original Kane $\vec{k}\cdot\vec{p}$ theory, but it has never been explicitly stated in the QWIP literature. Heavy hole to heavy hole transitions and light hole to light hole transitions follow the usual intersubband selection rule of being nominally forbidden for normally incident radiation in the absence of an optical grating. The value of these selection rules is in designing QWIPs which have a reproducible absorption spectrum. The latter requires the choice of an intersubband transition which is strongly absorbing and *not* one which is nominally forbidden.

Much of the theoretical work on hole intersubband transitions has been numerical, and has focussed on holes with a large value of in-plane wave vector. Following this theoretical work, much of the original experimental work done by Levine also emphasized p-QWIP designs where the hole in-plane wave vector is designed to be large, either through heavy doping of the quantum wells or through the use of asymmetric quantum wells. However, the selection rules and optical dipoles derived here show that the heavy hole to light hole absorption is strong, even in the limit of very small hole in-plane wave vector, if the confinement wave vector k_z is large.

The well known selection rules and optical dipoles for a conduction intersubband transition have been reviewed. The possibility of the absorption of normally incident radiation by electrons in an n-QWIP has been investigated. In so doing, it was found that the cell periodic Bloch function for electrons having a finite wave vector can be accurately described numerically within $\vec{k}\cdot\vec{p}$ theory by a Brillouin-Wigner perturbative expansion in terms of the Bloch functions at $\vec{\mathbf{k}} = \vec{\mathbf{0}}$. This expansion is analytical, and easy to use. The intersubband absorption of normally incident radiation by electrons in an n-QWIP in the absence of an optical grating was found within $\vec{k}\cdot\vec{p}$ theory

to be much smaller than that in a p-QWIP. The size of the electron intersubband absorption of normally incident radiation in the absence of an optical grating is found to be proportional to the size of the electron in-plane wave vector.

It was also found from $\vec{k}\cdot\vec{p}$ theory that strained layer growth on an (001) substrate does not significantly affect the intersubband absorption strength. Since the Hamiltonian describing uniaxial strain has the same (tetragonal) symmetry as that describing the confinement of carriers in the quantum wells along the growth direction, then the heavy hole to light hole absorption in strained epilayers grown on an (001) substrate has an optical dipole which is larger than that in unstrained layers by the amount given in Equation (4.52), which is only slightly different from unity.

Chapter 5

MBE Growth

Figure 5-1 shows the three chamber RIBER 2300 system used for the molecular beam epitaxial (MBE) growth of the III-V semiconductor devices described in this work. The growth chamber is the chamber at the center and right side of the figure. The cell flange, shown on the right side of the figure, could accommodate five Group III Knudsen effusion cells (for the indium, two gallium, and two aluminum sources), two dopant cells (for beryllium and silicon), and a solid As source with a valved As cracker. The long black pipes, positioned vertically in the photograph, carry liquid nitrogen to and from the Meissner trap which is cooled during growth to provide a base pressure of 8×10^{-10} Torr when the cells were at their growth temperatures and the As valve was closed. The electron gun used for Reflection High Energy Electron Diffraction (RHEED) is shown at the very center of the photograph, attached to a part of the upper growth chamber near the cell flange. The growth chamber is pumped by an ion pump, a titanium sublimation pump, (both not visible in the photograph) and a CTI Cryogenics Cryo-Torr8 cryopump (in the lower left portion of the photograph). The preparation chamber is situated between two gate valves, and its lower exterior is covered with aluminum foil in this photograph. The load chamber is on the left side of the figure.

The thermocouple readings of the effusion cells were kept constant to within about 0.1 °C with Eurotherm Model 818 temperature controllers.

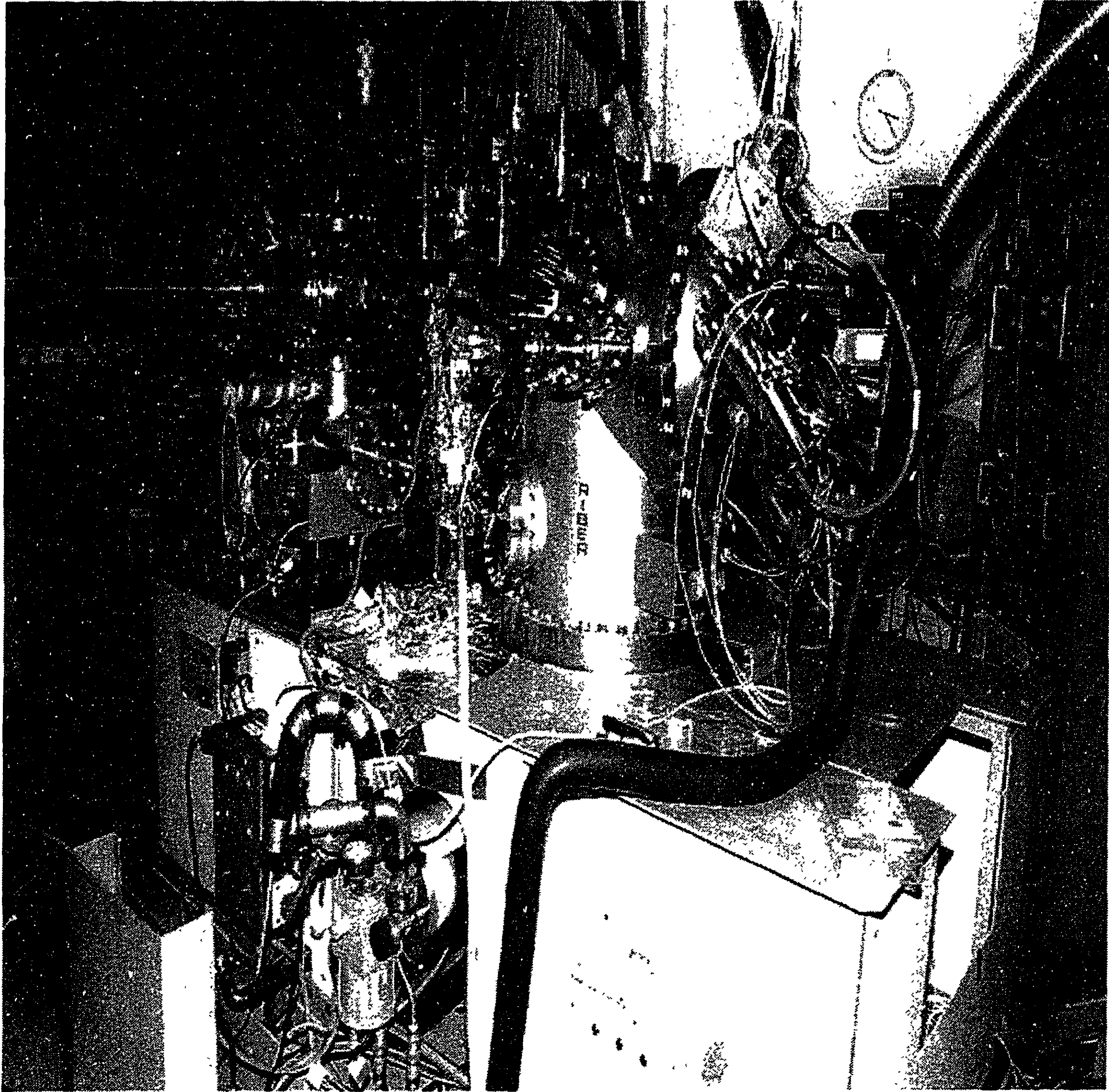


Figure 5-1: The three chambers of our RIBER 2300 molecular beam epitaxy system. (The photograph was taken by Dr. Isako Hoshino.)

5.1 Flux Calibration

5.1.1 Flux calibration by double crystal rocking curve.

The group III fluxes were calibrated by growing thin layers of InGaAs and InAlAs on an InP substrate in a fixed duration, typically 12 minutes, at a fixed estimated growth rate, typically 1 μm /hour. Double crystal rocking curves were measured to obtain empirical compositions and layer thicknesses, from which growth rates were obtained. Figure 5-2 shows the measured rocking curve of such a calibration growth. The desired layer was 200 nm of $In_{0.521}Al_{0.479}As$, and the measured layer was 168 nm of $In_{0.508}Al_{0.492}As$. The choice of thickness for a calibration growth and the estimation of its composition are discussed in Chapter 6.

5.1.2 Flux calibration by measuring beam equivalent pressure (BEP) and RHEED oscillations

Measurement of the Group III beam equivalent pressure (BEP) or RHEED oscillation frequency as a function of the inverse temperature gives an Arrhenius dependence whose slope is an activation energy for the line-of-sight evaporation of the Group III sources into a perfect vacuum [54]. (See Equation (6.40) below.) This activation energy yields the relative growth rates for the Group III sources.

Absolute growth rates are obtained by taking RHEED oscillations. RHEED oscillations are taken of AlGaAs growth on GaAs, to get Ga and Al growth rates. The In growth rate is obtained by growing fully relaxed InAs on InP. To obtain growth rates on the substrate InP based on these RHEED oscillation measurements, one must account for the difference in lattice constants of the underlying substrates. Since the number of atoms in one monolayer (ML) of an epitaxial growth is inversely proportional to the square of the lattice constant, the growth rates in ML/s differ according

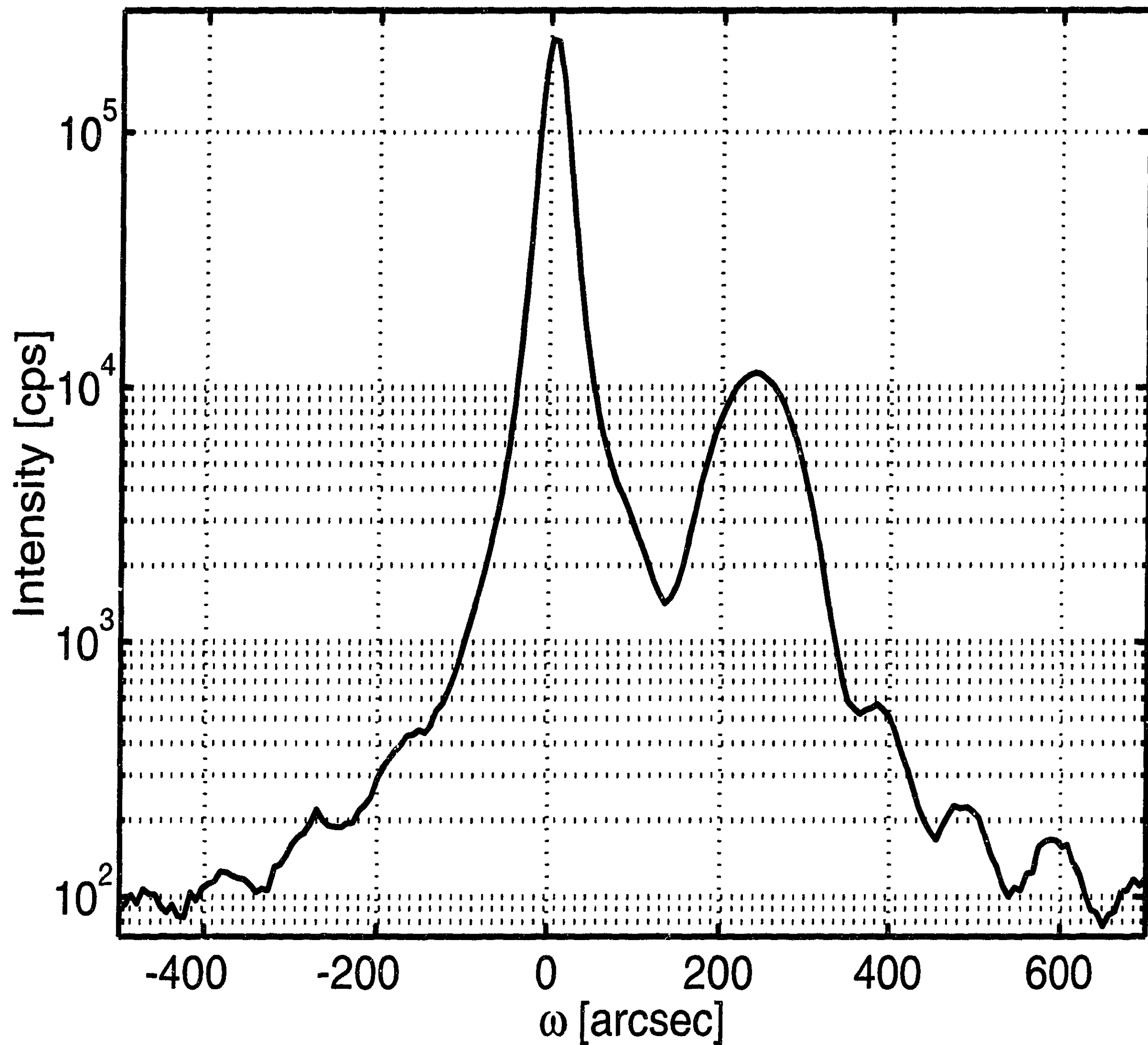


Figure 5-2: Double crystal rocking curve for Sample 9443. The desired layer was 200 nm of $In_{0.521}Al_{0.479}As$, and the measured layer thickness was 189 nm of $In_{0.514}Al_{0.486}As$. The analysis of this type of measurement is discussed in Chapter 6.

to,

$$(\text{growth of matched InAs in ML/s on InP} =)$$
$$(\text{growth of fully relaxed InAs in ML/s})(a_{\text{InP}}/a_{\text{InAs}})^2; \quad (5.1)$$
$$(\text{growth of matched GaAs in ML/s on InP}) =$$
$$(\text{growth of matched GaAs in ML/s on GaAs})(a_{\text{InP}}/a_{\text{GaAs}})^2. \quad (5.2)$$

Since the perpendicular lattice constants also differ, the growth rates in μm/hr differ according to,

$$(\text{growth of matched Ga in InGaAs in } \mu\text{m/hr on InP}) =$$
$$(\text{growth of matched GaAs ML in } \mu\text{m/hr on GaAs})(a_{\text{InP}}/a_{\text{GaAs}}) \quad (5.3)$$
$$(\text{growth of matched In in InGaAs in } \mu\text{m/hr on InP}) =$$
$$(\text{growth of fully relaxed InAs ML in } \mu\text{m/hr})(a_{\text{InP}}/a_{\text{InAs}}) \quad (5.4)$$

So the ratio of lattice matched-growth rates in μm/hr on InP to lattice-matched growth rates on GaAs is $(a_{InP}/a_{GaAs})^3$, and the ratio of lattice matched-growth rates in μm/hr on InP to fully relaxed growth rates of InAs is $(a_{InP}/a_{InAs})^3$.

5.2 Growth Conditions

The substrate temperature maintained during most of the growths on InP substrates was about 500°C, which is a temperature high enough for the aluminum adatom mobility to be adequate during growth but low enough for the indium not to desorb from the sample surface during growth. Growth of AlGaAs/GaAs structures having low aluminum mole fraction on GaAs substrates was done at substrate temperatures between 580 and 600°C. Substrate temperature was maintained constant by using a constant output from the power supply. Reproducibility of the substrate temperature was obtained by calibrating an optical pyrometer with a band pass at 2.3μm. Calibration of the optical pyrometer was usually done with a GaAs substrate: the pyrometer setting corresponding to 640°C was determined by observing when the As stable RHEED 2x pattern changed to a Ga stable 4x pattern in less than 10 sec with

the As shutter closed and the As beam equivalent pressure (BEP) less than 10^{-7} Torr, about two orders of magnitude less than the usual [55, 56] As BEP during growth. The pyrometer could also be calibrated by looking for the 2×4 to 4×2 transition on InP at 518°C.

Growth interruptions of 5-90 seconds were used in some of these growths to change the cell temperatures, as there was only one Ga and one Al cell when most of these samples were grown. As such, the As overpressure is also critical, since a larger As overpressure appears to keep the In from desorbing at 500°C. This was noticed as a better lattice-matched structure when a higher As overpressure was used during the growth interruptions. Too high an As overpressure was avoided, as that would impede the aluminum adatom mobility during growth. The typical ratio of the Group V to Group III BEPs during growth was between 15:1 and 25:1.

The growth rate can be monitored in-situ by observing the oscillations in the intensity of the RHEED specular spot during growth. One period of such RHEED oscillations corresponds to the growth of one monolayer. Examples of RHEED oscillations are shown in Figures 5-3 and 5-4. In these two figures, the RHEED oscillations begin right after the As shutter is closed and then quickly opened.

Figs. 5-3 and 5-4 also show that the optimal As overpressure can be monitored in-situ by observing the RHEED specular intensity right after the As shutter is closed and then quickly opened. If the RHEED specular intensity brightens right after the As shutter is closed, then the growth conditions are slightly As-rich and optimal. This is shown in Fig. 5-3. However, if the RHEED specular intensity does not increase right after the As shutter is closed, then the growth conditions are As-deficient. This is shown in Fig. 5-4.

The growth of a sample under slightly As-deficient conditions leaves the sample surface with the characteristic morphology shown in Figure 5-5. The Group III element is believed to segregate during growth under As-deficient conditions, forming ripples in the $(1\bar{1}0)$ direction and the Group III-rich regions shown in the figure. The ripples shown in Figure 5-5 result from the orientation [57] dependence of the interfacial energy of Group III-rich regions on the compound semiconductor surface. Such As-

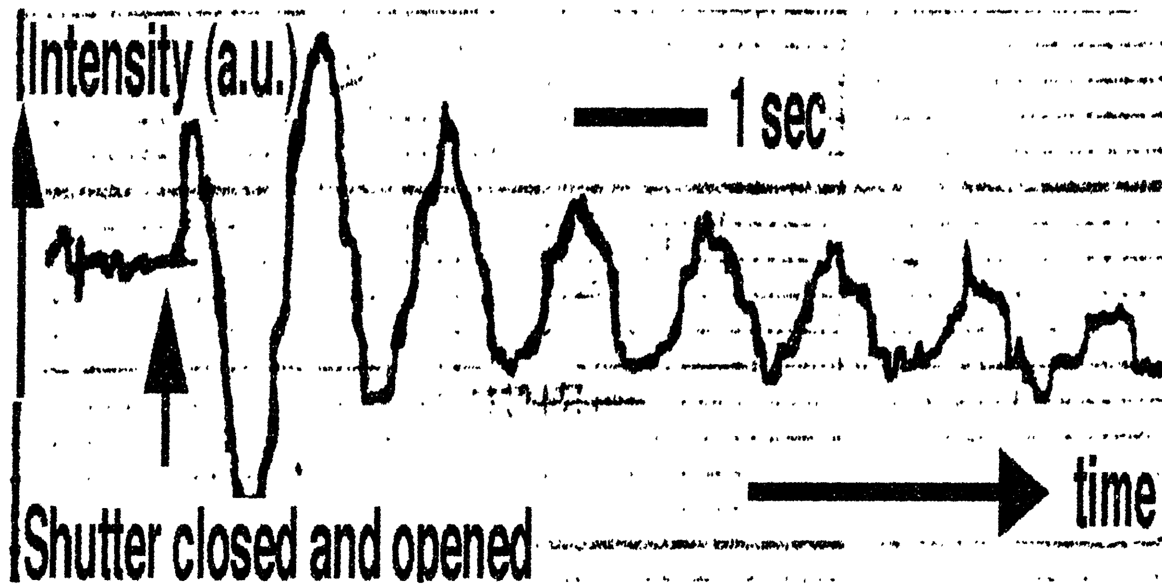


Figure 5-3: The time (horizontal axis) dependence of the intensity (vertical axis) of the specular spot in the RHEED (Reflection High Energy Electron Diffraction) pattern right after the As shutter is closed and then quickly opened under slightly As-rich growth conditions. The RHEED specular intensity is observed to rise right after the As shutter is closed, thus indicating slightly As-rich growth conditions.

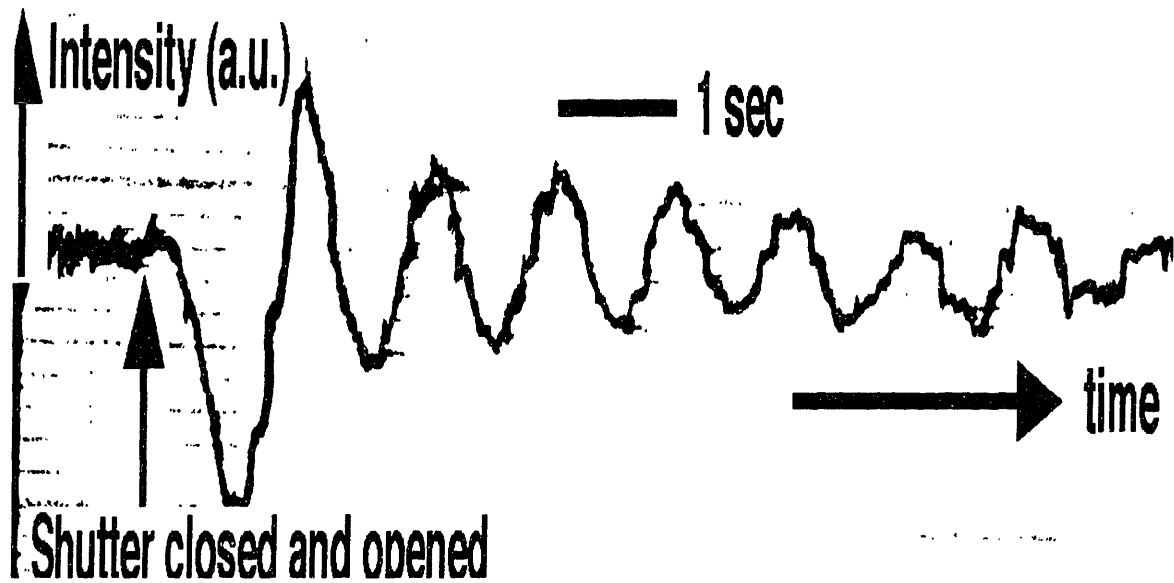


Figure 5-4: The time (horizontal axis) dependence of the intensity (vertical axis) of the specular spot in the RHEED (Reflection High Energy Electron Diffraction) pattern right after the As shutter is closed and then quickly opened under slightly As-deficient growth conditions. The RHEED specular intensity is observed to drop right after the As shutter is closed, thus indicating slightly As-deficient growth conditions.

deficient growth conditions were observed, for example, in June 1997 when the solid As source was depleted after more than two and a half years of use.

5.3 A sample growth

Figure 5-6 shows a scanning electron micrograph (SEM) of a typical QWIP, Sample 9207, which was comprised of 10 periods of 40 Å GaAs and 500 Å $Al_{0.25}Ga_{0.75}As$. The horizontal width of the micrograph corresponds to an in-plane dimension of 1.13 μm. The magnification used for this SEM was 100,000×. In this figure, the surface is located near the top of the figure, and the substrate is located at the bottom of the figure. The 40 Å GaAs quantum wells are the narrow, dark, horizontal lines in Fig. 5-6, and the 500 Å $Al_{0.25}Ga_{0.75}As$ barriers are the wider, lighter, horizontal lines in the figure.

The central portion of Fig. 5-6 appears brighter than the rest of the micrograph, and the quantum wells appear to be "wiggly" lines instead of straight horizontal lines in this portion of the micrograph. This is an artifact of the focusing of the SEM.

A stain etch was used to enhance the contrast between the GaAs and the AlGaAs layers in Fig. 5-6. The stain etch selectively etches the GaAs but not the AlGaAs. Prior to the stain etch, the sample was dipped for 5 seconds in buffered oxide etch (roughly 50% hydrofluoric acid and 50% ammonium fluoride) followed by a quick rinse in deionized (DI) water. The stain etch itself was a 20 second immersion in a solution of 4 drops of ammonium hydroxide dissolved in 20 milliliters of hydrogen peroxide.

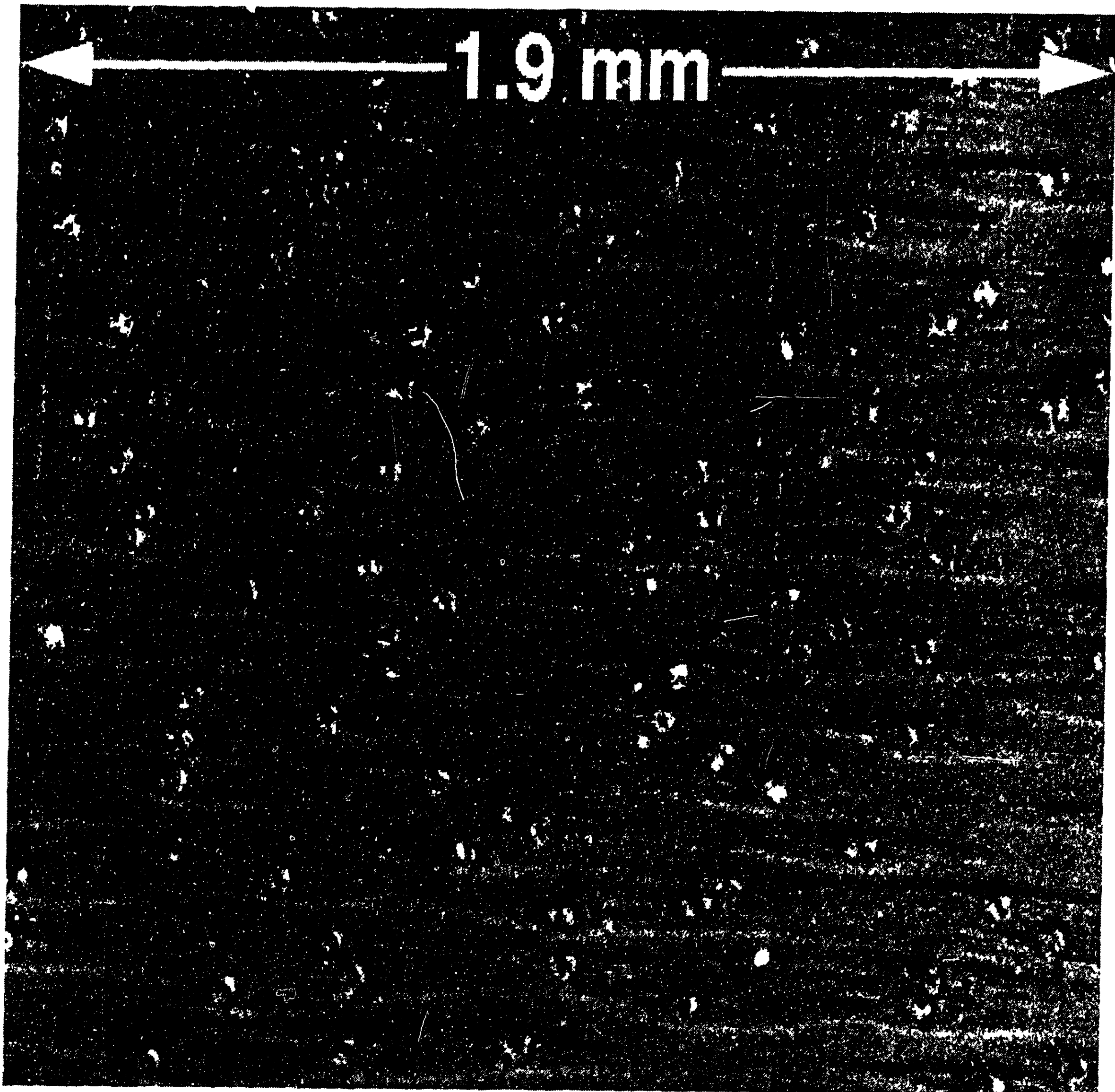


Figure 5-5: The surface morphology of a sample grown under As deficient growth conditions. This Normarski image was taken under 50× magnification, and the horizontal distance across the image corresponds to 1.9 mm.

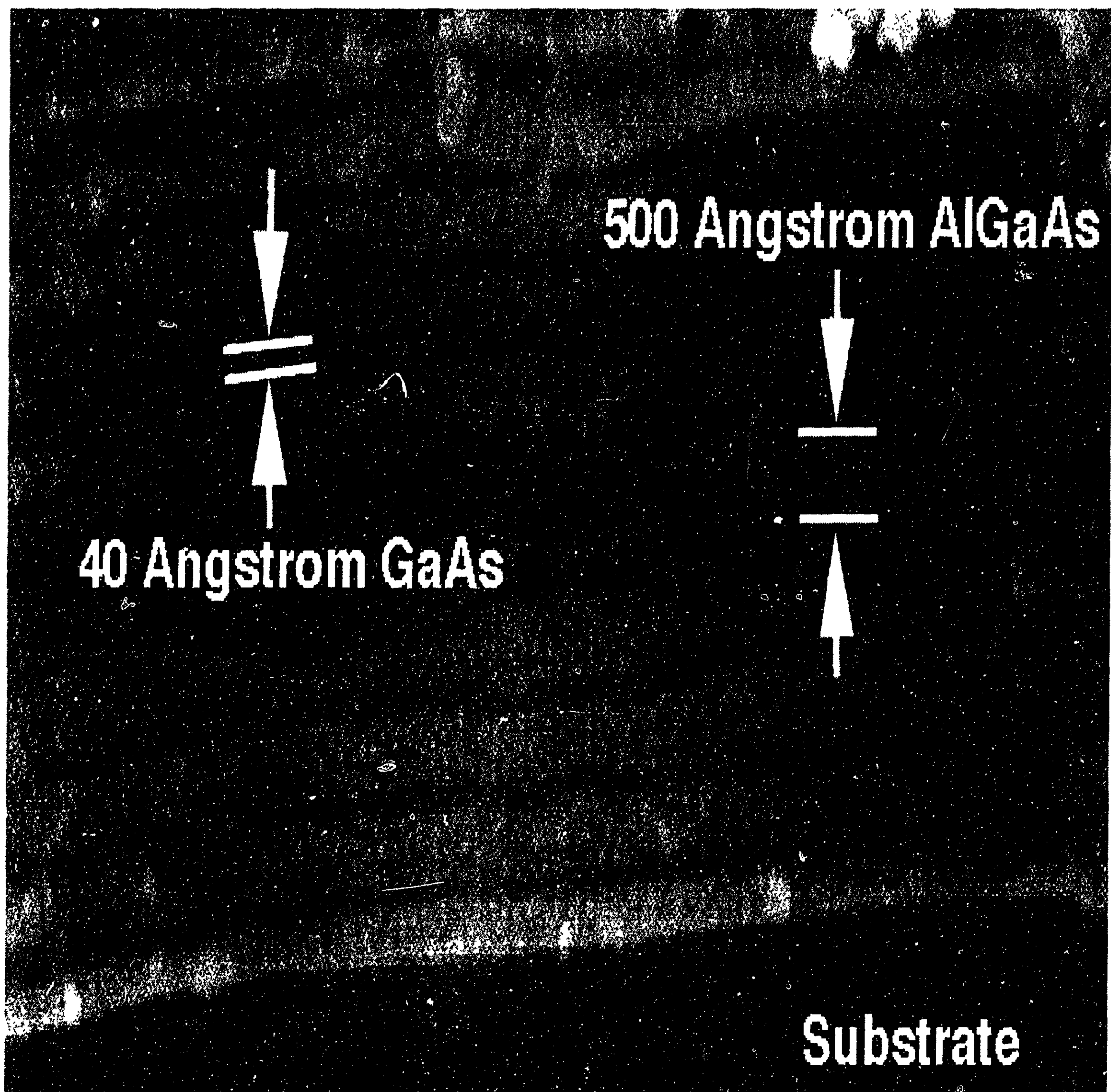


Figure 5-6: Scanning electron micrograph (SEM) of Sample 9207, which was comprised of 10 periods of 40 Å GaAs and 500 Å $Al_{0.25}Ga_{0.75}As$. The horizontal width of the micrograph corresponds to an in-plane dimension of 1.13 μm. The magnification used for this SEM was 100,000×.

Chapter 6

Layer Structure Characterization by High Resolution X-ray Diffraction

Characterization of the epitaxial layers by High Resolution X-ray Diffraction (HRXRD) allows an empirical measurement of the compositions and layer widths, the amount of relaxation of strained layers, some assessment of the uniformity of the materials parameters throughout device, the estimation of the sharpness of the semiconductor heterointerfaces resulting from either smoothly graded interfaces or three dimensional roughness at the interfaces. When the epitaxial layers are fully pseudomorphic, a single rocking curve measurement yields information about both the layer compositions and thicknesses. This is discussed in Section 6.3.1. When the epitaxial layers have started to relax, a single rocking curve measurement is not enough to determine both the layer compositions and the amount of relaxation. A combination of several rocking curve measurements can be used to determine the amount of relaxation, as discussed in Section 6.3.2. Measurement of the reciprocal space map of the epitaxial layers allows one to distinguish between variations in the perpendicular lattice constant resulting from variations in the composition during the course of the growth. This is discussed in Section 6.4. Section 6.5 discusses the sharpness of the semiconductor

heterointerfaces resulting from either smoothly graded composition at the interfaces or three dimensional structures at the interfaces. Epitaxial layer width uniformity can be ascertained from either a reciprocal space map or a glancing incidence X-ray reflectivity measurement, as discussed in Section 6.6. The diffractometer used for making these X-ray measurements is described in Section 6.2.

6.1 Introduction

Diffraction of an incident, monochromatic X-ray beam of wavelength λ and incident wave vector $\vec{\mathbf{k}}_{\mathbf{i}}$ produces a diffracted beam with the diffracted wave vector $\vec{\mathbf{k}}_{\mathbf{d}}$ and an electric field amplitude E_s of [58, 59]:

$$E_s = E_{s,1} \sum_{\vec{\mathbf{R}}_{\mathbf{i}}} \exp[i(\vec{\mathbf{k}}_{\mathbf{d}} - \vec{\mathbf{k}}_{\mathbf{i}}) \cdot \vec{\mathbf{R}}_{\mathbf{i}}] \sum_{\vec{\mathbf{r}}_{\mathbf{n}}} \exp[i(\vec{\mathbf{k}}_{\mathbf{d}} - \vec{\mathbf{k}}_{\mathbf{i}}) \cdot \vec{\mathbf{r}}_{\mathbf{n}}] f_n(\vec{\mathbf{R}}_{\mathbf{i}}) \tag{6.1}$$

$$\equiv E_{s,1} \sum_{\vec{\mathbf{R}}_{\mathbf{i}}} \exp[i(\vec{\mathbf{k}}_{\mathbf{d}} - \vec{\mathbf{k}}_{\mathbf{i}}) \cdot \vec{\mathbf{R}}_{\mathbf{i}}] F_{\vec{\mathbf{R}}_{\mathbf{i}}} \tag{6.2}$$

$$\equiv E_{s,1} F_{\text{epilayer+substrate}}, \tag{6.3}$$

where [58] $E_{s,1}$ is the amplitude of the scattered electric field resulting from one electron, where $\vec{\mathbf{R}}_{\mathbf{i}}$ denote the position vectors of all unit cells which comprise the entire crystal (the epilayer plus the substrate), where

$$F_{\text{epilayer+substrate}} = \sum_{\vec{\mathbf{R}}_{\mathbf{i}}} \exp[i(\vec{\mathbf{k}}_{\mathbf{d}} - \vec{\mathbf{k}}_{\mathbf{i}}) \cdot \vec{\mathbf{R}}_{\mathbf{i}}] F_{\vec{\mathbf{R}}_{\mathbf{i}}} \tag{6.4}$$

is the total crystal structure factor, where $\vec{\mathbf{r}}_{\mathbf{n}}$ denote the position vectors of all the atoms which comprise a single unit cell, where

$$F_{\vec{\mathbf{R}}_{\mathbf{i}}} = \sum_{\vec{\mathbf{r}}_{\mathbf{n}}} \exp[i(\vec{\mathbf{k}}_{\mathbf{d}} - \vec{\mathbf{k}}_{\mathbf{i}}) \cdot \vec{\mathbf{r}}_{\mathbf{n}}] f_n(\vec{\mathbf{R}}_{\mathbf{i}}) \tag{6.5}$$

is the structure factor for the unit cell located at the position $\vec{\mathbf{R}}_{\mathbf{i}}$, and where $f_n(\vec{\mathbf{R}}_{\mathbf{i}}) = \int \exp[i(\vec{\mathbf{k}}_{\mathbf{d}} - \vec{\mathbf{k}}_{\mathbf{i}}) \cdot \vec{\mathbf{r}}_{\mathbf{a}}] \rho(\vec{\mathbf{r}}_{\mathbf{a}}) d^3\vec{\mathbf{r}}_{\mathbf{a}}$, is the atomic scattering factor for the atom located at $\vec{\mathbf{r}}_{\mathbf{n}}$ within the unit cell located at $\vec{\mathbf{R}}_{\mathbf{i}}$, and where ρ is the density of the electron cloud associated with this atom.

The structure factor approach is equivalent to a kinematical [58] model of X-ray diffraction, and is valid [60, 61] for most superlattice measurements because the intensity of the satellite peaks is much weaker than that of the central zero-order peak.

High Resolution X-Ray Diffraction (HRXRD) measurements of epitaxial layers grown by molecular beam epitaxy yield information about the lattice constants in both the growth direction and the in-plane directions, the widths of the constituent epitaxial layers, the amount of relaxation, the average variations in the layers widths and compositions, and the interface sharpness. All of this information is contained in Equation (6.4), the total crystal structure factor, as Equation (6.4) is observed to be the Fourier transform of the electron *spatial distribution* within the total crystal comprised of both the epitaxial layers and the substrate.

6.2 The Diffractometer

Figure 6-1 shows the triple axis diffractometer, with its three characteristics stages: the beam conditioner, the second (sample) crystal, and the analyzer crystal.

The X-ray source is a water-cooled rotating copper anode, typically operated at 50 kV and 200 mA. The emitted Cu-$K_{\alpha 1}$ line is at 1.54 Å. The cross sectional area of the X-ray beam as measured by the detector is $(250\ \mu m)^2$.

The beam conditioner provides the incident X-ray beam with more resolution in both angle and wavelength. In high intensity mode, angular resolution is achieved with four reflections and subsequent diffraction of the incident beam by a silicon channel cut crystal. These four reflections reduce the streak in reciprocal space along the growth direction which results from the semi-infinite extent of the beam conditioner crystal. The slits following the channel cut collimator in the beam conditioner block the Cu-$K_{\alpha 2}$ line. The monochromator provides more wavelength resolution.

The Bede Enhanced Dynamic Range X-ray Detector is composed of a scintillator, a photomultiplier tube, and amplification circuitry. X-rays diffracted by the sample are passed through a beryllium window and strike a scintillator crystal sensitive to pho-

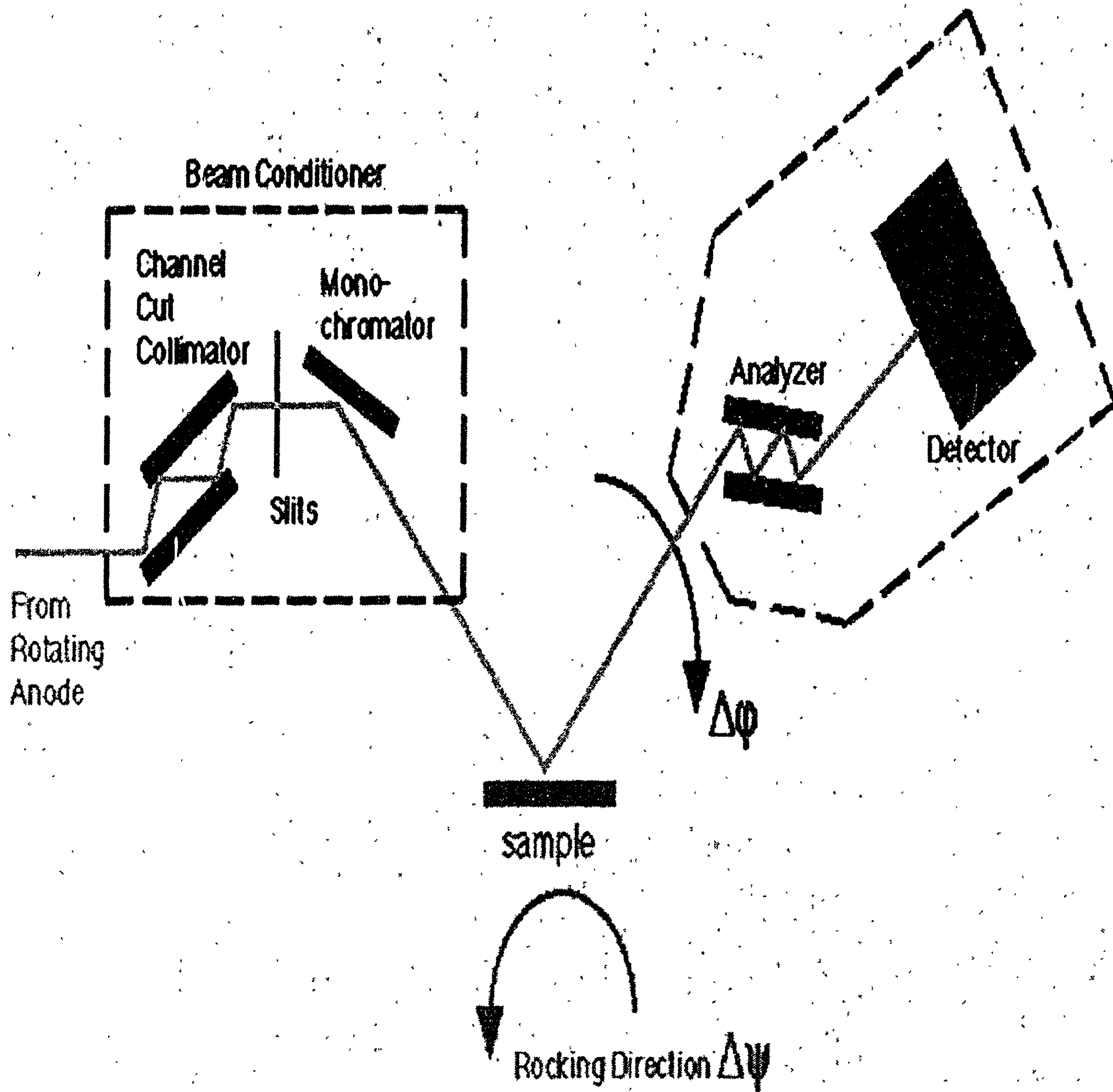


Figure 6-1: The triple axis diffractometer, with its three characteristics stages: the beam conditioner, the second (sample) crystal, and the analyzer crystal

tons with energies of 4-20 keV. The scintillator crystal, here made of sodium iodide, converts with a high efficiency each X-ray photon to a visible photon of wavelength 0.1 μm. This induced luminescence has a decay time of 200 ns, and thus allows high speed detection. The photomultiplier tube, usually operated at 1000 V, is comprised of an active ten-dynode chain, and is linear over the dynamic range of the detector. The Enhanced Dynamic Range Detector exhibits a background count rate of 0.15 cps and peak count rates greater than 450,000 cps. Count rates as low as 1 cps are easily discerned. The integration time of the detector is chosen to achieve a desired signal-to-noise ratio. Typical single scans took one half hour to finish.

HRXRD refers to any system employing two or more crystals. The two crystals referred to in a standard double crystal rocking curve measurement are the monochromator and the sample crystals shown in Figure 6-1. A third crystal, a four reflection silicon (111) analyzer crystal, placed after the sample and before the detector gives very high resolution to the measured angle of the diffracted beam. When the analyzer crystal is used in the detector arm of the diffractometer, information about the lattice constant in the growth direction (i.e., epilayer composition and its variations) can be separated from information about the lattice constant in the in-plane direction (i.e., local tilt of the lattice planes resulting from some amount of relaxation).

In a standard double crystal rocking curve measurement, the detector angle is fixed while the sample crystal is "rocked" (scanned) through the Bragg angle. In the literature, this scan is often called an ω-scan. When the analyzer crystal is not used in an ω-scan, the detector observes a large section of the Ewald sphere, and information about the lattice constant along the growth direction cannot be distinguished from the in-plane lattice constant. When the analyzer crystal is used in an ω-scan, the variation of the in-plane lattice constant is observed, thus yielding information about the substrate curvature or the local tilt, possibly due to epilayer relaxation, of the lattice planes from the normal direction.

In a θ-2θ scan, the detector is scanned in unison with the sample but at twice the scan rate of the sample. In such a scan, as the sample angle is "rocked" by an amount $\delta\theta$ from the Bragg angle, the detector angle is simultaneously moved by an amount -$2\delta\theta$, so that the peaks in the diffracted intensity arise from those epitaxial layers whose

lattice constant in the growth direction deviates from the substrate lattice constant by an amount that is determined by the Bragg condition (see Equation (6.8) below). The θ-2θ scan thus yields information about the lattice constant (compositional variations) along the growth direction. The FWHM of the diffraction peak measured in a θ-2θ scan from a high quality silicon sample is about 5 arcseconds or 12 arcseconds, depending on whether the beam conditioner is operated in high resolution or high intensity mode. This sample FWHM measured in a θ-2θ scan is limited by the divergence of the input beam.

When the analyzer crystal is used in an ω-scan of a high quality silicon sample (second crystal), the FWHM of the sample peak is measured to be about 5 arcseconds or 12 arcseconds, depending on whether the beam conditioner is operated in high resolution or high intensity mode. The sample FWHM measured with the analyzer crystal in place in an ω-scan is limited by the crystalline perfection of the sample.

Alignment is critical. If the incident X-ray beam does not intersect the sample crystal at the axis of rotation of the sample crystal, the diffracted peak widths will appear lower and broader when the X-ray detector has a large aperture size. The reason is that as the sample is rocked, the diffraction peaks arise from both different lateral positions in the incident beam cross section and different lateral positions along the sample surface.

6.3 Measurement of Double Crystal Rocking Curves

6.3.1 Measurement of Epilayer Composition and Layer Thickness

Peaks in the intensity (proportional to the square of Equation (6.4)) of a beam diffracted from a crystal occur when the radiation scattered by the atoms in consecutive planes constructively interfere according to the Bragg condition,

$$n\lambda = 2d \sin \theta_B, \tag{6.6}$$

where n is the order of the diffraction, λ is the wavelength of the incident radiation, d is the spacing between the consecutive planes involved in the constructive interference, and θ_B is the Bragg angle. For diffracting planes of indices (hkl), d is related to the lattice parameter a through,

$$d = \frac{a}{\sqrt{h^2 + k^2 + l^2}}. \tag{6.7}$$

When the central (zero order) epitaxial layer peak in the rocking curve is shifted from the substrate peak by an amount $\Delta\theta$, the epitaxial layer has an average lattice constant along the growth direction that deviates from the substrate lattice constant by the amount $\Delta a_\perp$

$$\left(\frac{\Delta a}{a}\right)_\perp = \frac{a_\perp - a_{\text{substrate}}}{a_{\text{substrate}}} = \frac{\sin\theta_B}{\sin(\theta_B + \Delta\theta)} - 1 \approx -\left(\frac{\Delta\theta}{\tan\theta_B}\right) \tag{6.8}$$

for a symmetric reflection, in which the incidence angle is equal to the exit angle of the X-rays.

The average composition of the epitaxial layers is related to the average epilayer lattice constant in the growth direction and the average amount of epilayer relaxation through [62],

$$\left(\frac{\Delta a}{a}\right)_r = \frac{a_r - a_{\text{substrate}}}{a_{\text{substrate}}} = \frac{1-\nu}{1+\nu}\left(\frac{\Delta a}{a}\right)_\perp + \frac{2\nu}{1+\nu}\left(\frac{\Delta a}{a}\right)_{||}, \tag{6.9}$$

where a_r is the lattice constant of a fully relaxed (bulk) crystal having the same composition as that of the epilayer being measured, where $a_{||}$ is the average in-plane lattice constant of the epilayer, and where ν is the Poisson ratio. The lattice parameters and Poisson ratios for some [3, 4, 5, 6] III-V semiconductors are shown in Table 6.1. The amount of relaxation is proportional to $\Delta a_{||}$: a fully coherent layer has $a_{||}$=$a_{textsubstrate}$, and a fully relaxed layer has $a_{||}$=$a_\perp$=a_r.

Layer Composition. Calculation of the lattice constant a_r of a fully relaxed crystal having the same composition as that of the sample epilayer allows the average composition of the epilayer to be extracted from Vegard's Law (here written for InGaAs):

$$a_{r,In_xGa_{1-x}As} = xa_{r,InAs} + (1-x)a_{r,GaAs}, \tag{6.10}$$

Semicon-ductor	Lattice Parameter (Angstroms)	Poisson ratio ν
InAs	6.0584	0.352
GaAs	5.6532	0.311
AlAs	5.6622	0.274
InP	5.8688	0.360

Table 6.1: The lattice parameters and Poisson ratios for some III-V semiconductors [3, 4, 5, 6].

which says that the bulk ternary lattice constant can be found by linearly interpolating the constituent bulk binary lattice constants. Vegard's law [63, 64, 65] has been found to give good agreement with experiment.

Layer Thicknesses. The contribution to the structure factor, equation (6.4), of a thin, pseudomorphic epilayer is, in the kinematic theory of X-ray diffraction,

$$F_{\text{epilayer}} = \frac{\sin(Nu/2)}{\sin(u/2)} \tag{6.11}$$

where $u = (k_{z,d} - k_{z,i})R_{z,\text{epilayer}}$ and N is the number of unit cells in the epilayer along the growth direction z. Equation (6.11) shows that around the central zero-order epilayer peak will appear a number of interference or Pendellosung fringes, whose angular spacing $\Delta\theta$ is related to the thickness t of the epilayer by [62]:

$$\Delta\theta = \frac{\lambda|\gamma_h|}{t\sin(2\theta_B)} \tag{6.12}$$

$$= \frac{\lambda}{2t\cos\theta_B} \qquad \text{for symmetric reflections,} \tag{6.13}$$

where γ_h is the cosine of the angle between the diffracted wave vector and the surface normal, and $\gamma_h = \cos(90 - \theta_B)$ for symmetric reflections. Equation (6.13) is the spacing between consecutive Pendellosung peaks of fringe order ± 2 or higher; the fringe spacing between the central epilayer peak and either of the ± 1 peaks is 1.5 times Equation (6.13).

An example of a rocking curve obtained from a thin InAlAs epilayer grown on an InP substrate was shown in Fig. 5-2 of Section 5. The Pendellosung fringe spacing was measured to be 111 arcsec, corresponding to an epilayer thickness of 168 nm. The central epilayer peak was measured to be 237 arcsec away (in the direction of In-deficient compositions) from the InP substrate peak, corresponding to a measured composition of $In_{0.508}Al_{0.492}As$. Very good lattice matching is observed, with the substrate peak differing from the central epilayer peak by 237 arcsec. For comparison, this separation between the epilayer and substrate diffracted intensity peaks can be compared with the separation of 390 arcsec between the diffracted intensity peaks contributed by a pseudomorphic AlAs epilayer and its underlying GaAs substrate.

For the purpose of calibration of the growth rate, previous growers [66] within the group have found that 200 nm was an optimal strained epilayer thickness for the unambiguous interpretation of a single rocking curve measurement of a calibration growth sample. If a thinner epilayer were grown as the calibration sample, the Pendellosung fringes became smaller in intensity and more widely spaced. Larger errors in the estimation of the growth rate were likely when interpreting low, broad Pendellosung fringes. If a thicker epilayer were grown as the calibration growth sample, then relaxation of the epilayer was possible, making Pendellosung fringes invisible and the epilayer composition unknown, since the amount of relaxation could not be ascertained from a single rocking curve measurement.

A superlattice example. Epitaxial layer widths and compositions for a superlattice are easy to obtain from a double crystal rocking curve. Shown in Figure 6-2 is the double crystal rocking curve for Sample 9351 which was designed to be a top contact of 350 nm of $In_{0.532}Ga_{0.468}As$, followed by fifteen periods of the structure, which consisted of 50 nm of a $In_{0.499}Ga_{0.274}Al_{0.227}As$ barrier and 10.08 nm of a $In_{0.67}Ga_{0.33}As$ well.

To extract the actual layer widths and compositions, we observe the following. The substrate peak is at -7074.5 arcsec. The peak at -6993.5 arcsec yields the composition of the InGaAs contacts and buffer layer. The zero order epilayer peak is at -7135 arcsec. The barrier composition is given by the peak at -6590.75 arcsec. The

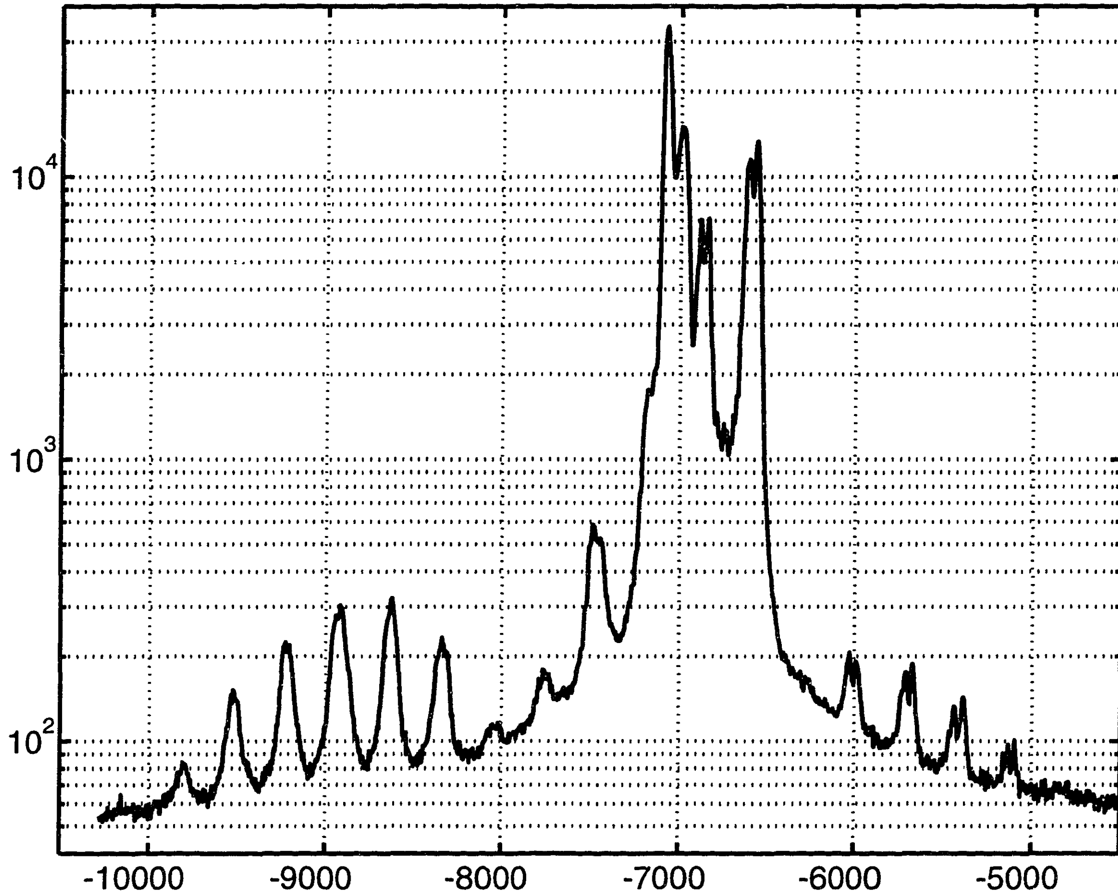


Figure 6-2: Double crystal rocking curve for Sample 9351. The desired epitaxial layer was a top contact of 350 nm of $In_{0.532}Ga_{0.468}As$, followed by fifteen periods of the structure: 50 nm of a $In_{0.499}Ga_{0.274}Al_{0.227}As$ barrier and 10.08 nm of a $In_{0.67}Ga_{0.33}As$ well. The actual structure was a top contact of 377 nm of $In_{0.530}Ga_{0.470}As$, followed by fifteen periods of the structure: 53.5 nm of a $In_{0.500}Ga_{0.276}Al_{0.224}As$ barrier and 10.84 nm of a $In_{0.668}Ga_{0.332}As$ well.

well composition is given by the low broad peak at -8915 arcsec. The superlattice periodicity is given the spacing (290 arcsec) between the superlattice fringes. Thus, the measured superlattice period is 64.34 nm, compared to a desired period of 60.08 nm, and thus the actual total growth rate was about 7 percent larger than desired.

Since the low broad peak denoting the well composition does not have a clearly defined peak, we use the positions of the other peaks (the substrate, the InGaAs top contact, the barrier, and the average superlattice peaks) to calculate the ratio of the actual to the desired growth rates on the day this sample was grown. Any two of these peak positions gives the gallium and the aluminum incorporation rates relative to the indium incorporation rate. These peak positions in combination with the measured superlattice periodicity yields the ratio of the actual to the desired growth rates for the indium, gallium, and aluminum cells on that day to be 1.072, 1.081, and 1.055, respectively. So, the structure actually grown on that day was a top contact of 377 nm of $In_{0.530}Ga_{0.470}As$, followed by fifteen periods of the structure: 53.5 nm of a $In_{0.500}Ga_{0.276}Al_{0.224}As$ barrier and 10.84 nm of a $In_{0.668}Ga_{0.332}As$ well.

6.3.2 Relaxation and uniformity of epilayers

Epitaxial layers grown in the InGaAlAs materials system may be relaxed, sometimes intentionally (under the assumption that a small amount of relaxation does not deteriorate device behavior by much) and sometimes unintentionally (when the source fluxes are slightly different from the intended ones). A knowledge of the amount of relaxation of the epitaxial layers may allow an estimate of the variation of the layer compositions and thicknesses throughout growth. For example, a large value of the full width at half maximum (FWHM) of the epilayer diffraction peaks could come from either a relaxing of the epilayers or composition/thickness variations of the epilayers during growth. To obtain the amount of relaxation of the epilayers, so that the layer compositions and the amount of layer variation throughout a growth can be obtained, a combination of several rocking curves must be measured at both asymmetric and symmetric Bragg peaks.

To measure the amount of relaxation of an epitaxial layer, the in-plane lattice constant must be measured. For a partially relaxed epitaxial layer, the in-plane lattice constants along the [110] and [1$\bar{1}$0] directions could differ because misfit dislocations form preferentially along the [1$\bar{1}$0] direction. For a partially relaxed epitaxial layer, the in-plane lattice constant differs from the lattice constant of the underlying substrate.

An asymmetric Bragg reflection is used to measure the in-plane lattice constant. Here, the normal to the diffracting planes differs from that to the sample surface by a finite angle ϕ. If the angle between the incident wave vector and the surface of the diffracting plane is denoted by θ_B, then the angle between the incident wave vector and the sample surface is $\theta_B - \phi$ for glancing incidence measurements and $\theta_B + \phi$ for glancing exit measurements. This is shown in Figure 6-3.

For asymmetric reflections, the Bragg angles associated with the substrate and the epilayer will differ by $\Delta\theta_B$,

$$\Delta\theta_B = -\left[\left(\frac{\Delta a}{a}\right)_\perp \cos^2\phi + \left(\frac{\Delta a}{a}\right)_\parallel \sin^2\phi\right] \tan\theta_B \tag{6.14}$$

and the equivalent *asymmetric* planes in the substrate and the epilayer will have a relative tilt of angle $\Delta\phi$,

$$\Delta\phi = \left[\left(\frac{\Delta a}{a}\right)_\perp - \left(\frac{\Delta a}{a}\right)_\parallel\right] \sin\phi\cos\phi, \tag{6.15}$$

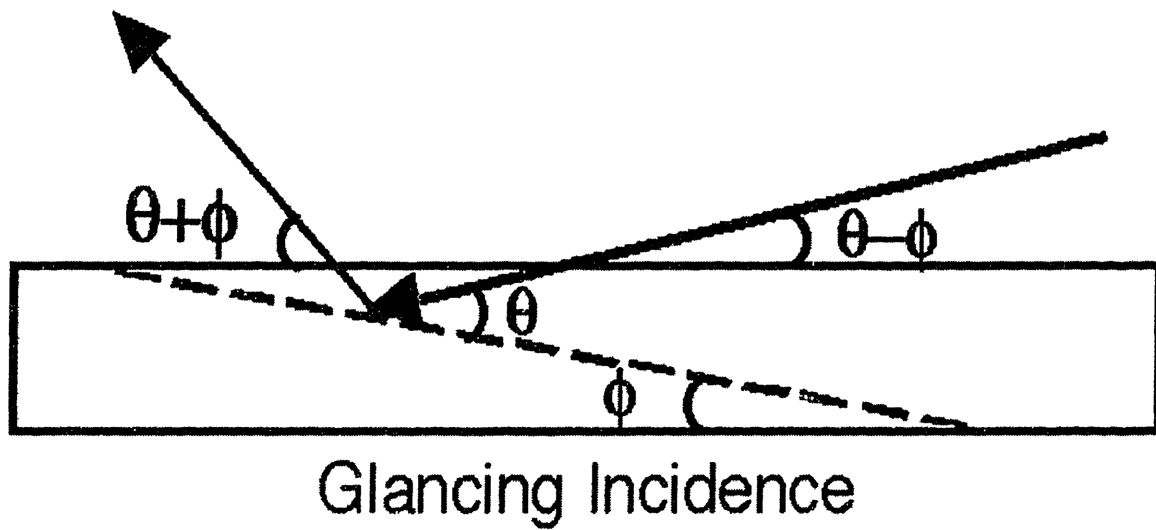
as shown in Figure 6-4. Thus, the substrate and the epilayer peaks measured in an asymmetric Bragg reflection will differ by

$$\Delta\theta = \Delta\theta_B \mp \Delta\phi \tag{6.16}$$

$$= \left(\frac{\Delta a}{a}\right)_\perp \left(-\cos^2\phi\tan\theta_B \pm \sin\phi\cos\phi\right) + \left(\frac{\Delta a}{a}\right)_\parallel \left(-\sin^2\phi\tan\theta_B \mp \sin\phi\cos\phi\right), \tag{6.17}$$

where the upper signs refer to glancing exit measurements and the lower ones to glancing incidence measurements.

To assess the amount of partial relaxation of strained layers, the author measured rocking curves around the (224) asymmetric reflection. For an (001) substrate,



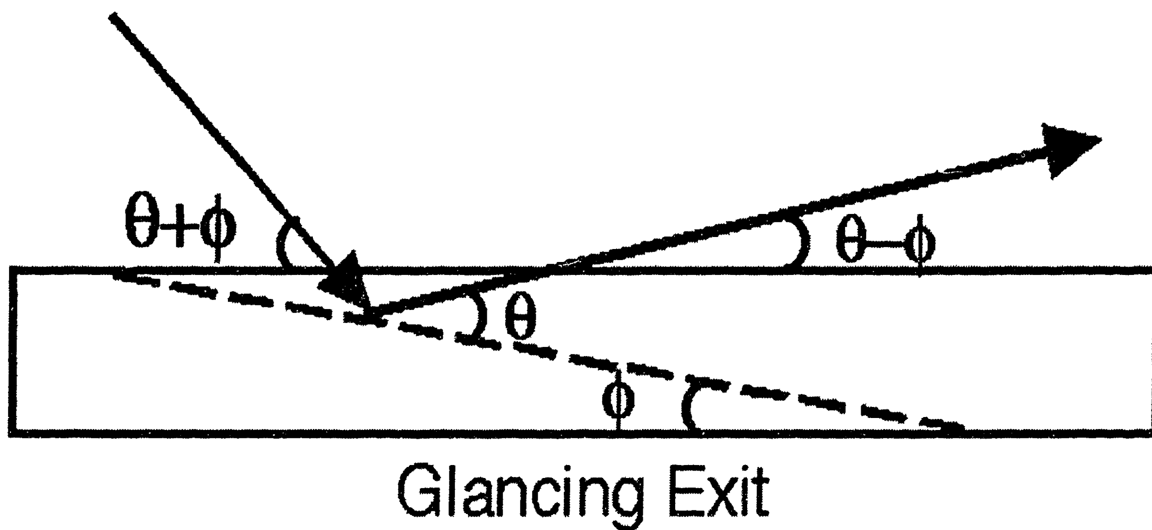


Figure 6-3: Glancing incidence and glancing exit geometries showing Bragg angle θ and tilt ϕ of planes.

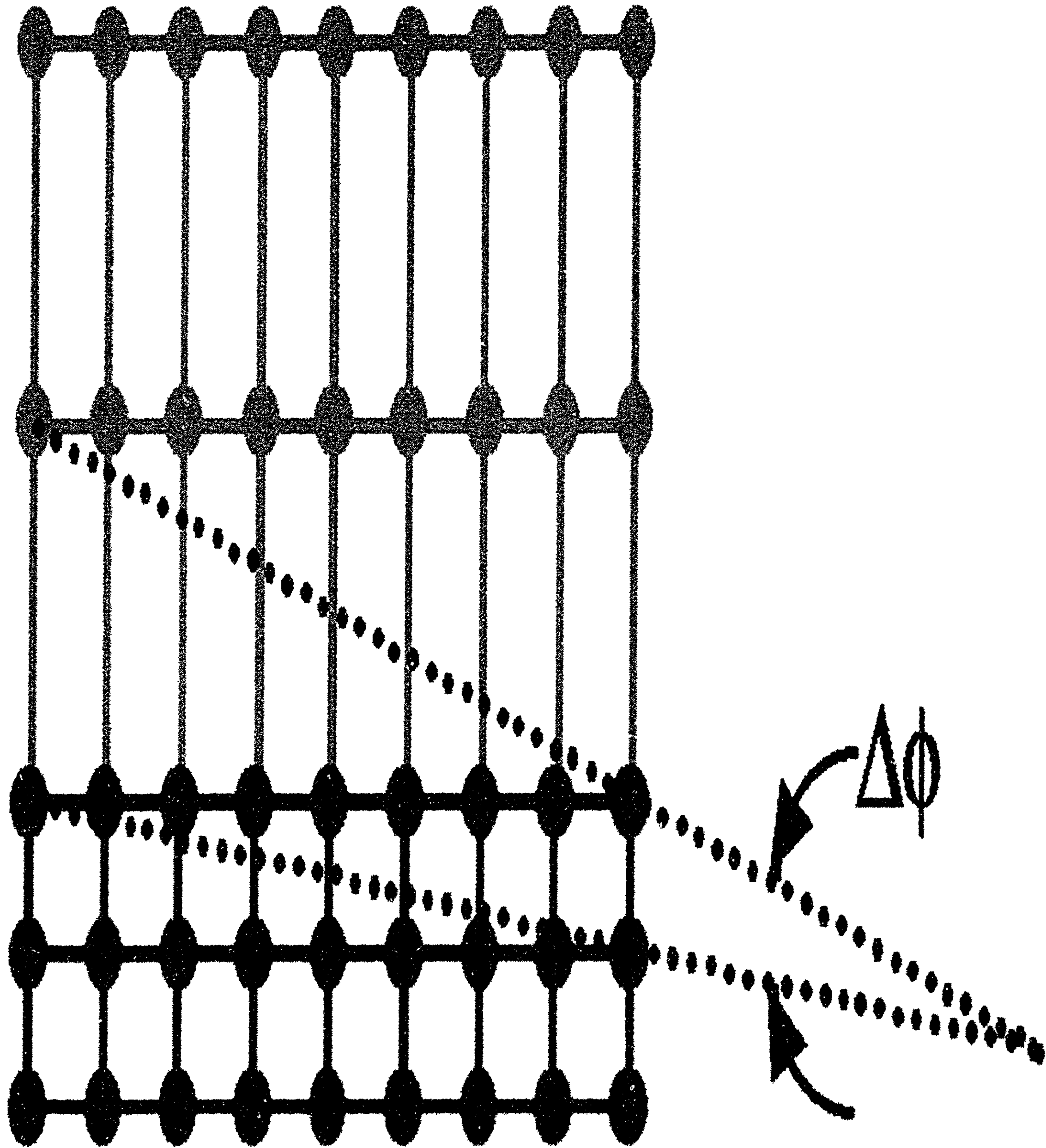


Figure 6-4: Equivalent asymmetric planes in the substrate and in the epilayer will have a relative tilt of angle $\Delta\phi$.

$\phi(224)$=35.26°. For an InP substrate, the separation between the epilayer and substrate peaks is [66]

$$\Delta\theta = -1.031\left(\frac{\Delta a}{a}\right)_{\perp} + 0.191\left(\frac{\Delta a}{a}\right)_{\|} \tag{6.18}$$

for a (224) glancing incidence measurement, and is

$$\Delta\theta = -0.089\left(\frac{\Delta a}{a}\right)_{\perp} - 0.751\left(\frac{\Delta a}{a}\right)_{\|} \tag{6.19}$$

for a (224) glancing exit measurement. For a simple tetragonal distortion, there are two unknowns: $\left(\frac{\Delta a}{a}\right)_{\perp}$ and $\left(\frac{\Delta a}{a}\right)_{\|}$. These can be found from either two asymmetric measurements or a combination of one asymmetric and one symmetric measurement. If the substrate has a miscut angle, then four measurements are made: the same measurements described in the previous sentence are made at two different in-plane azimuths of the sample, and then averaged.

For orthorhombic distortion, the in-plane lattice constants differ, and are characterized by $\left(\frac{\Delta a}{a}\right)_{\|[110]}$ and $\left(\frac{\Delta a}{a}\right)_{\|[1\bar{1}0]}$. The epitaxial layer composition must be found from

$$\left(\frac{\Delta a}{a}\right)_{r} = \frac{1-\nu}{1+\nu}\left(\frac{\Delta a}{a}\right)_{\perp} + \frac{\nu}{1+\nu}\left[\left(\frac{\Delta a}{a}\right)_{\|[110]} + \left(\frac{\Delta a}{a}\right)_{\|[1\bar{1}0]}\right], \tag{6.20}$$

and three measurements are needed to find $\left(\frac{\Delta a}{a}\right)_{\perp}$, $\left(\frac{\Delta a}{a}\right)_{\|[110]}$, and $\left(\frac{\Delta a}{a}\right)_{\|[1\bar{1}0]}$. If the sample has a miscut angle, then six measurements are needed to determine the lattice parameters.

6.4 Reciprocal Space Mapping

Reciprocal space maps are made by compiling a set of θ-2θ scans (or an equivalent set of ω scans) near a Bragg peak. In an ω scan, the sample is rocked with its axis of rotation perpendicular to the plane of incidence, as indicated in Fig. 6-1, with the detector fixed. In a θ-2θ scan, the detector and sample are rotated in unison, with the axis of rotation perpendicular to the plane of incidence, and with the detector scanned at twice the angular frequency as the sample angular scan frequency.

Reciprocal space maps allow [67] strain or perpendicular lattice parameter distributions to be distinguished from mosaic tilt or interface roughness. Mosaic tilt manifests itself as a broadening in the θ-2θ direction as a result of the shape of the mosaic block, and also a further broadening of the superlattice peaks in the ω direction, with the broadening larger for the larger order satellite peaks. In the presence of mosaicity, the direction of the broadening of the satellite peaks is concentric with circles centered around the origin (000) of reciprocal space.

Interface roughness broadens the superlattice peaks in directions perpendicular to the growth direction, with the broadening larger for larger order satellite peaks. Interface roughness does not tilt the peaks with respect to the growth direction. The width of the zeroth order superlattice peak is unaffected by the presence of interface roughness.

Variations in the layer thickness broadens the satellite peaks along the growth direction, with the zeroth order satellite peak width unaffected by variations in the layer thicknesses.

Near the (004) Bragg peak, the deviation of the specimen angle ($\Delta\psi$) and the deviation of the detector angle ($\Delta\varphi$) from their nominal Bragg positions yields the diffracted intensity at the point in reciprocal space given by [67]

$$q_{[001]} = \Delta\varphi \cos\theta_B/\lambda \tag{6.21}$$

$$q_{[110]} = (2\Delta\psi - \Delta\varphi)\sin\theta_B/\lambda, \tag{6.22}$$

where [001] is the growth direction. Thus, the θ-2θ scan, with $2\Delta\psi = \Delta\varphi$, is a scan in reciprocal space parallel to the growth direction, [001], and yields information about layer thicknesses or lattice constants along the growth direction. The ω scan, with $\Delta\varphi$=0, is a scan in reciprocal space perpendicular to the growth direction, and yields information about the tilt of the growing planes with respect to the nominal growth direction. The standard double crystal rocking curve is a scan in reciprocal space perpendicular to the growth direction, with intensities which are the summations of the diffracted intensities over the wide detector angular aperture ($\Delta\varphi$) along the Ewald sphere.

The analyzer crystal provides more resolution in the ω scan by reducing the aperture

size $\Delta\varphi$. The analyzer crystal also reduces the instrumental streak in reciprocal space.

Near an asymmetric Bragg peak, the deviation of the specimen angle ($\Delta\psi$) and the deviation of the detector angle ($\Delta\varphi$) from their nominal Bragg positions yields the diffracted intensity at the point in reciprocal space given by

$$q_{[001]} = \frac{\sin[\theta_B + \varphi + \Delta\psi] + \sin[\theta_B - \varphi + \Delta\varphi - \Delta\psi]}{\lambda} - \frac{\sin[\theta_B + \varphi] + \sin[\theta_B - \varphi]}{\lambda} \tag{6.23}$$

$$q_{[110]} = \frac{-\cos[\theta_B + \varphi + \Delta\psi] + \cos[\theta_B - \varphi + \Delta\varphi - \Delta\psi]}{\lambda} - \frac{-\cos[\theta_B + \varphi] + \cos[\theta_B - \varphi]}{\lambda} \tag{6.24}$$

where [001] is the growth direction. When the epitaxial layers are coherent, the in-plane lattice constant throughout the epitaxial layer is the same as that in the substrate, and the peaks associated with the epitaxial layer must lie parallel to the $q_{[001]}$-axis. To measure a reciprocal space map near an asymmetric Bragg peak in a direction parallel to the $q_{[001]}$-axis, the detector angle $\Delta\varphi$ must be varied in unison with the sample angle $\Delta\psi$ according to

$$\Delta\varphi = \left[1 + \frac{\sin(\theta_B - \varphi)}{\sin(\theta_B + \varphi)}\right] \Delta\psi \quad = 1.085\Delta\psi \qquad \text{near (224) on an InP substrate.} \tag{6.25}$$

Figure 6-5 shows the reciprocal space map near the symmetric Bragg peak (004) obtained by triple axis diffractometry for the Sample 9351, which was discussed earlier to consist of a top contact of 377 nm of $In_{0.530}Ga_{0.470}As$, followed by fifteen periods of the structure: 53.5 nm of a $In_{0.500}Ga_{0.276}Al_{0.224}As$ barrier and 10.84 nm of a $In_{0.668}Ga_{0.332}As$ well. The peak widths along the [001] growth direction indicate layer width variations.

The peak widths along the [110] ω direction indicates either mosaic tilt or interface roughness of the epitaxial layers. The very narrow FWHM along the ω direction of 12 arcsec remains constant even for the increasing superlattice diffraction orders shown in Fig. 6-5, and this FWHM corresponds to the resolution limit of the diffractometer. This very narrow FWHM in the [110] direction and the absence of an

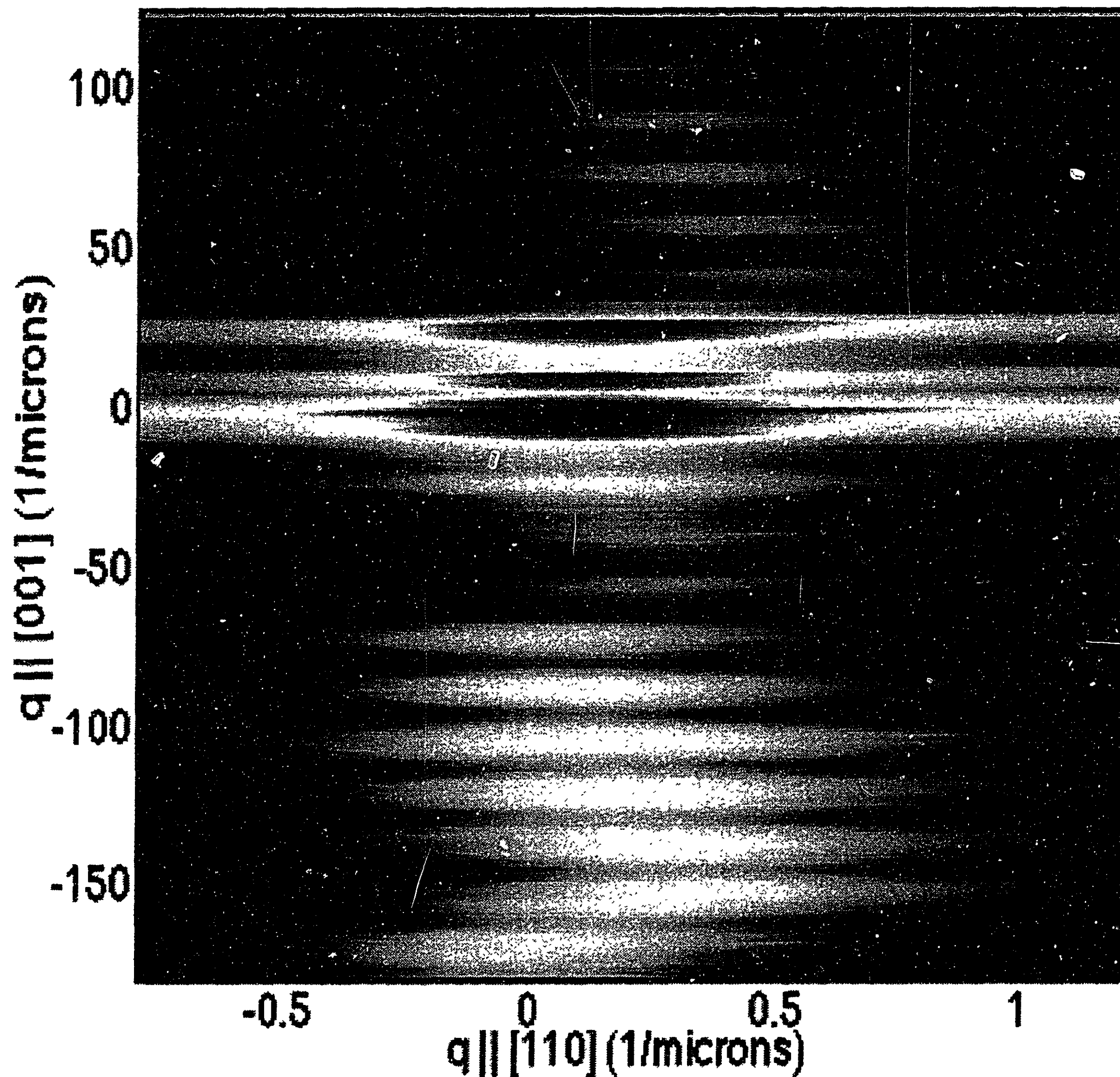


Figure 6-5: Reciprocal space map of Sample 9351 near the symmetric (004) Bragg peak. The peaks along the [001] growth direction indicate the different perpendicular lattice constants and thus the superlattice periodicity. The peak widths along the [001] growth direction indicate layer width variations. The peak widths along the [110] ω direction indicate mosaic tilt. The very narrow FWHM along the ω direction of 12 arcsec is the same for all superlattice diffraction orders shown, and it corresponds to the resolution limit of the diffractometer. The sample and detector angular positions are related to a point in reciprocal space through Equation (6.24).

increase of this FWHM with increasing diffraction order indicate an amount of mosaic tilt or epitaxial interface roughness which is too small to be resolved in this measurement.

The reciprocal space maps near asymmetric Bragg peaks, such as near (224) in Figure 6-6 for Sample 9351, yields information about the in-plane lattice constant and thus the lattice relaxation. If the superlattice had fully relaxed, the superlattice peaks in Fig. 6-6 would lie along the [224] direction, signified by the black line in the reciprocal space map. Since the superlattice peaks in Fig. 6-6 lie parallel to [001], the superlattice is pseudomorphic. Further, the very narrow FWHM along the [110] direction near the asymmetric (224) Bragg peak and the absence of an observed increase in this FWHM with increasing diffraction order are indications of an unresolvably small amount of mosaic tilt or interface roughness.

This is to be contrasted with the reciprocal space map of a superlattice which shows partial relaxation. Shown in Fig. 6-7 is such a structure, Sample 9326, whose nominal layer structure is a top contact of 350 nm of $In_{0.532}Ga_{0.468}As$, followed by fifteen periods of the structure: 50 nm of a $In_{0.532}Al_{0.468}As$ barrier and 6.24 nm of a $In_{0.45}Ga_{0.55}As$ well. The measured FWHM of the peaks along the [110] (ω) direction for Sample 9326 is about 100 arcsec. This FWHM is much wider than that measured for Sample 9351, and indicates partial relaxation of the epitaxial layers.

The wealth of information in a reciprocal space map is thus illustrated. The large FWHM in the [110] direction in the reciprocal space map of Sample 9326 was a clear indication of partial relaxation. That the epilayers in this sample had partially relaxed was not otherwise clear. This sample was perfectly shiny, even though more than 2.5 microns of epitaxial layers were grown. Its double crystal rocking curve, shown in Fig. 6-8, showed many narrow superlattice peaks.

Only by carefully considering the actual compositions obtained during the growth of Sample 9326 was it clear that the epitaxial layers had partially relaxed because they were thicker than the Matthews-Blakeslee [68] critical limit. The top contact alone, designed to be 350 nm of $In_{0.532}Ga_{0.468}As$, was actually 343 nm of $In_{0.5183}Ga_{0.4817}As$, which is quite a bit larger than the Matthew-Blakeslee critical limit of 231 nm for

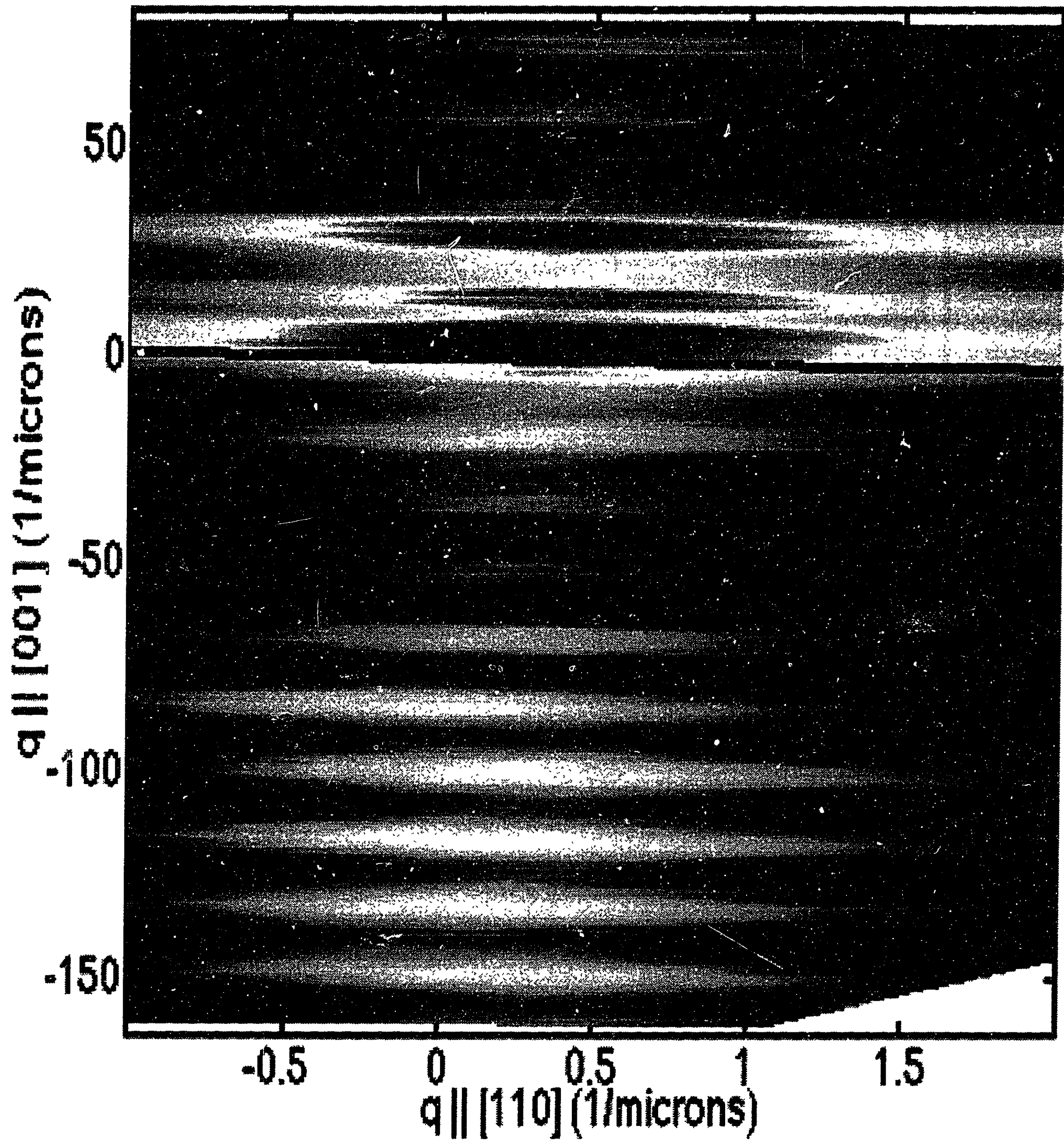


Figure 6-6: Reciprocal space map of Sample 9351 near the asymmetric (224) Bragg peak. If the QWIP superlattice had fully relaxed, the superlattice peaks would lie along the [224] direction, signified in the reciprocal space map by the black line. The sample and detector angular positions are related to a point in reciprocal space through Equation (6.24).

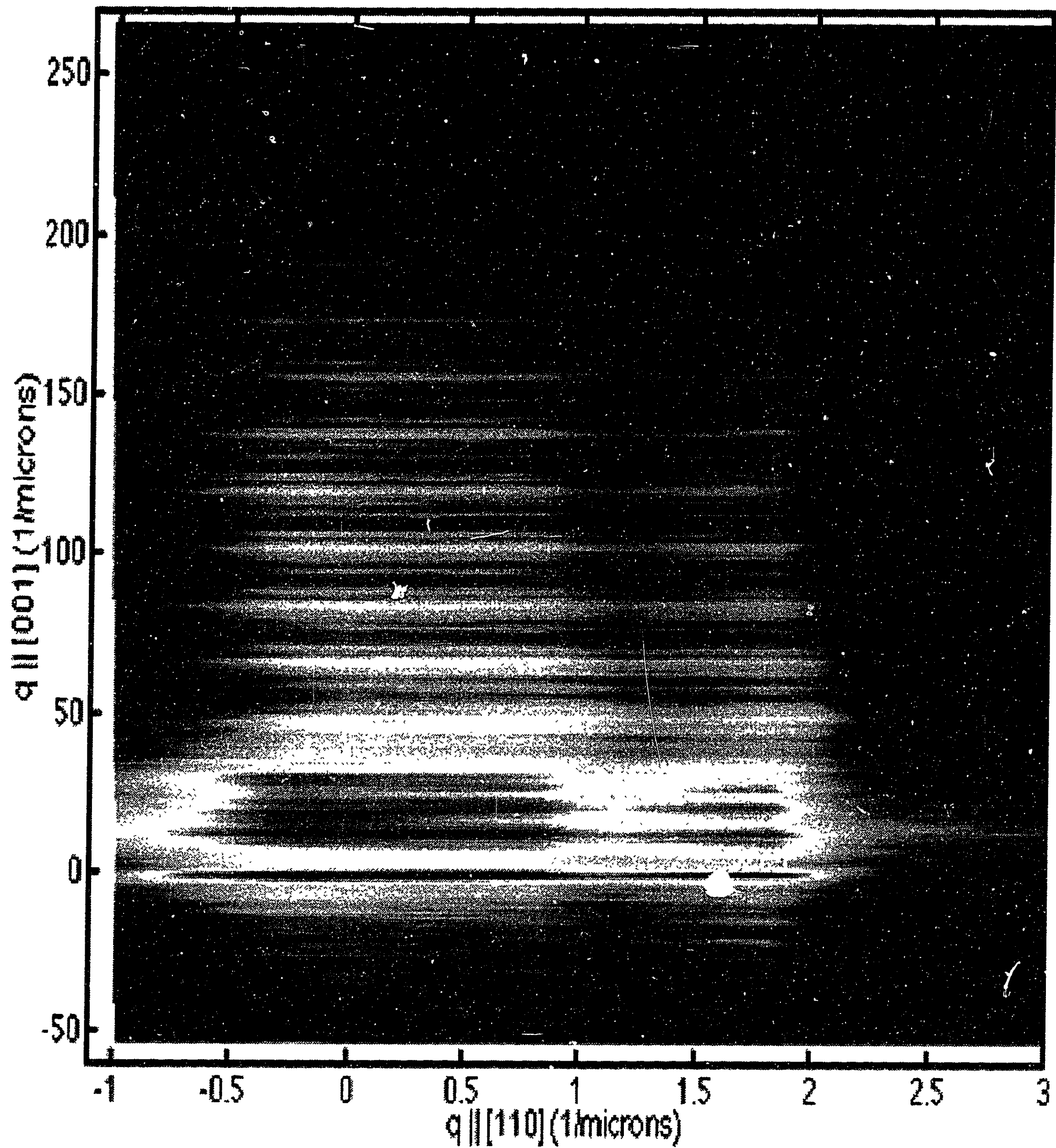


Figure 6-7: Reciprocal space map of Sample 9326 near the symmetric (004) Bragg peak. The measured FWHM of the peaks along the [110] (ω) direction is about 100 arcsec. This FWHM is much wider than that measured for Sample 9351, and indicates partial relaxation of the epitaxial layers. The sample and detector angular positions are related to a point in reciprocal space through Equation (6.24).

growing $In_{0.5183}Ga_{0.4817}As$. (The composition of the top contact can be obtained from noting its position at -6499 arcsec in Fig. 6-8 to be 243 arcsec in the indium deficient direction above the InP substrate peak at -6742 arcsec. The composition of the barriers can be obtained from the peak at -6364 arcsec to be $In_{0.5088}Al_{0.4912}As$, instead of the desired $In_{0.532}Al_{0.468}As$. The superlattice fringe spacing of 332 arcsec corresponds to a measured superlattice period of 56.20 nm, which is very close to the design goal of 56.24 nm. The ratio of the actual to the desired incorporation of the indium, gallium, and aluminum in Sample 9326 was 0.95425, 1.00797, and 1.04738, respectively.)

6.5 Glancing Incidence X-ray Reflectivity and Scattering Measurements

Glancing Incidence X-ray Reflectivity (GIXR) measurements are very sensitive to the epilayer interface grading/roughness, and allow the interface grading/roughness to be distinguished from the layer thickness variations. GIXR measurements ascertain interface grading and roughness without separating the two. Glancing incidence X-ray scattering (GIXS) measurements allow the interface roughness to be unambiguously ascertained. Comparison of the simulation and measurement of GIXR allows the determination of device layer thicknesses and variations thereof, compositions, and interface grading/roughness.

GIXR measurements are made in the same manner as θ-2θ scans, but are made near $\theta=0°$ instead of at the Bragg angle. Alignment is critical. After roughly aligning the detector and sample angular positions, an initial reflectivity measurement is made. The detector is then moved to twice the incident angle for an epilayer diffraction peak. Fine adjustments of the sample alignment are made by scanning the sample angular position through the incident angle for an epilayer diffraction peak. Alignment is achieved for that sample angular position at which the diffracted intensity (near glancing incidence) is a maximum.

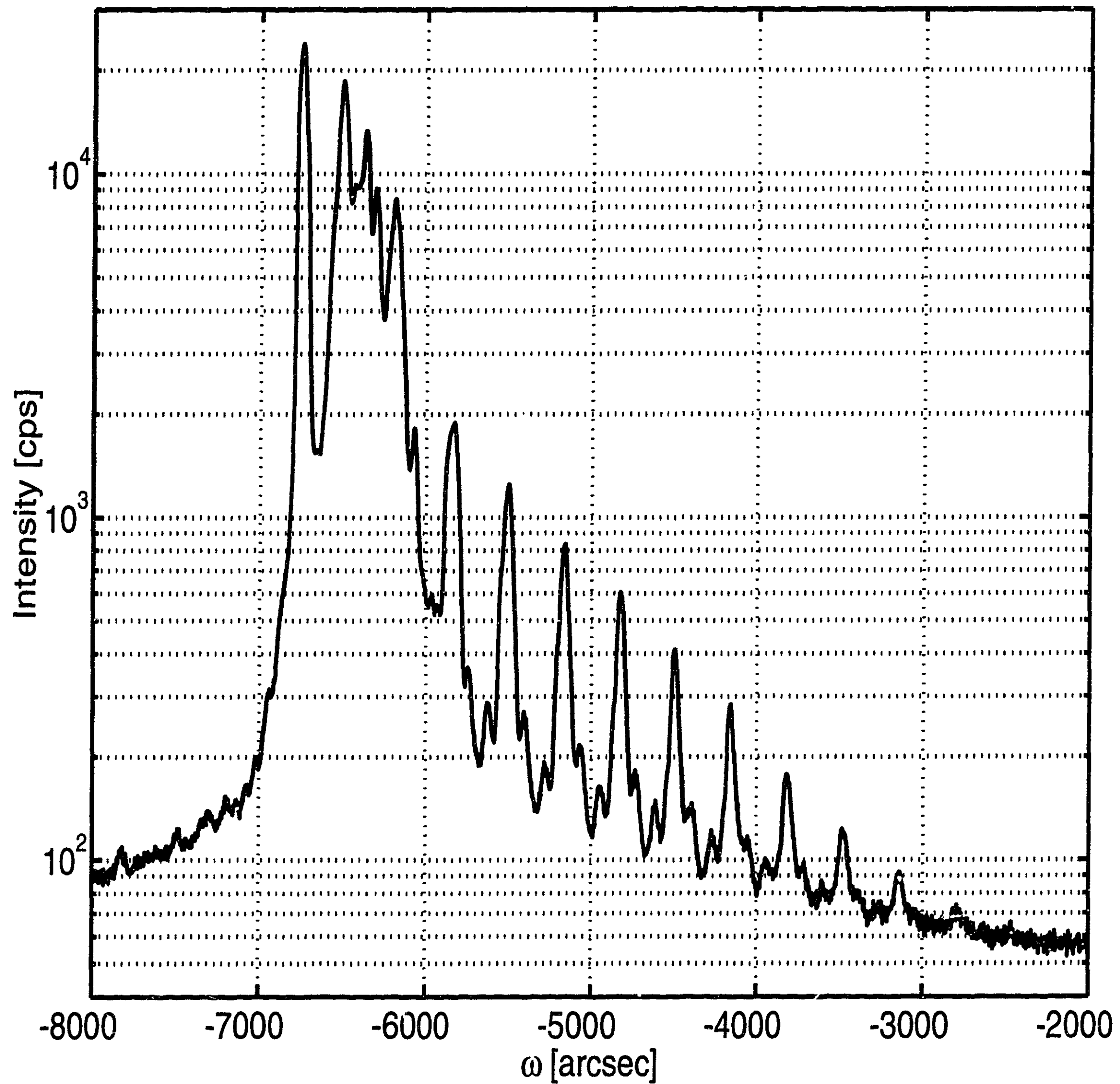


Figure 6-8: Double crystal rocking curve for Sample 9326. The desired epitaxial layer was a top contact of 350 nm of $In_{0.532}Ga_{0.468}As$, followed by fifteen periods of the structure: 50 nm of a $In_{0.532}Al_{0.468}As$ barrier and 6.24 nm of a $In_{0.45}Ga_{0.55}As$ well. The actual structure was a top contact of 343 nm of $In_{0.5183}Ga_{0.4816}As$, followed by fifteen periods of the structure: 50 nm of a $In_{0.5088}Al_{0.4912}As$ barrier and 6.14 nm of a $In_{0.436}Ga_{0.563}As$ well.

Since the penetration depth of the X-rays in GIXR measurements was estimated to be just several thousand angstroms, the best samples for GIXR measurements were those with only a thin or no cap layers. Cap layers should not be removed by etching, since it was noted that the roughness added to the sample surface by etching tends to dominate the roughness inherent in the layer interfaces.

6.5.1 Glancing Incidence X-ray Reflectivity Measurements

It is well known that the index of refraction must be accounted for in the determination of the epitaxial layer thicknesses from glancing incidence X-ray measurements [67, 69]. The index of refraction in most semiconductors is less than unity at the X-ray wavelengths. If the refractive index in air is assumed to be unity, and the refractive index in the interior of the semiconductor is assumed to be n_I, then the critical angle θ_C (in air measured from the sample surface) for total external reflection of X-rays is found from,

$$\cos\theta_C = n_I. \tag{6.26}$$

If the incident angle (exterior to the sample and measured from the sample surface) is denoted as θ_E, then the refracted X-ray beam in the interior of the semiconductor makes an angle of θ_I with respect to the sample surface, where θ_I is related to θ_E and θ_C through,

$$\cos\theta_I \cos\theta_C = \cos\theta_E. \tag{6.27}$$

For small (glancing) incident angles, this last equation is approximately

$$\theta_I^2 = \theta_E^2 - \theta_C^2. \tag{6.28}$$

When the epitaxial layers consist of a superlattice having a periodicity of D, the GIXR measurements show superlattice diffraction peaks at the superlattice Bragg condition:

$$\sin\theta_{I,n} = n\lambda/2D, \tag{6.29}$$

where n is the diffraction order and $\theta_{I,n}$ is the angle with respect to the sample surface of the refracted X-ray beam in the *interior* of the semiconductor. Using Equation (6.28), the superlattice diffraction peaks occur at incident angles (in the exterior of the semiconductor) satisfying,

$$\theta_{E,n}^2 = n^2 \left(\frac{\lambda}{2D}\right)^2 + \theta_C^2. \tag{6.30}$$

An example of a GIXR measurement is shown in Figure 6-9 for Sample 9089 which consisted of a superlattice intended to be 50 periods of 80 Å AlAs and 27 Å GaAs, with a cap layer of 130 Å of GaAs. The points in Fig.6-10 denote the measured superlattice diffraction peak positions as a function of the diffraction order as obtained from Fig. 6-9. The solid line in Fig.6-10 is a least squares best fit of the measured points to Equation (6.30). The slope of the solid line yields the average measured superlattice periodicity, which was found to be 104 Å for Sample 9089. Extrapolation of this curve to the zeroth diffraction order (the intercept of this curve with the ordinate axis) yields [69] the critical angle for total external reflection, which was measured to be about 1100 arcsecs=0.31° for Sample 9089. It is well known [67] that the critical angle for total external reflection of Cu Kα radiation from a GaAs substrate is 0.31°.

The solid line in Fig. 6-9 represents the reflectivity calculated [70] for a superlattice in which the layer widths and roughnesses are optimized to minimize the magnitude of the difference in the logarithms of the measured and calculated reflectivities. The modeled superlattice is comprised of a GaAs cap layer, an oxide layer, followed by 7 repetitions of a 7 period superlattice, for which the layer widths and roughness/interfacial grading are optimized for a best fit. The best fit for the 7-period superlattice was found to have layers with a mean width of 75.9 Å for the AlAs layers and 28.2 Å for the GaAs layers, with a standard deviation of 1.50 Å and 0.565 Å for the AlAs and GaAs layers, respectively. The interfaces between the GaAs and AlAs layers are modeled as not sharp, and were found to have an interfacial grading or roughness that spans a mean distance 2.69 Å and 2.45 Å for the AlAs and GaAs, respectively, with a standard deviation of 0.90 Å and 0.53 Å for the AlAs and GaAs, respectively. The best fit to the measured data points required a GaAs cap layer of

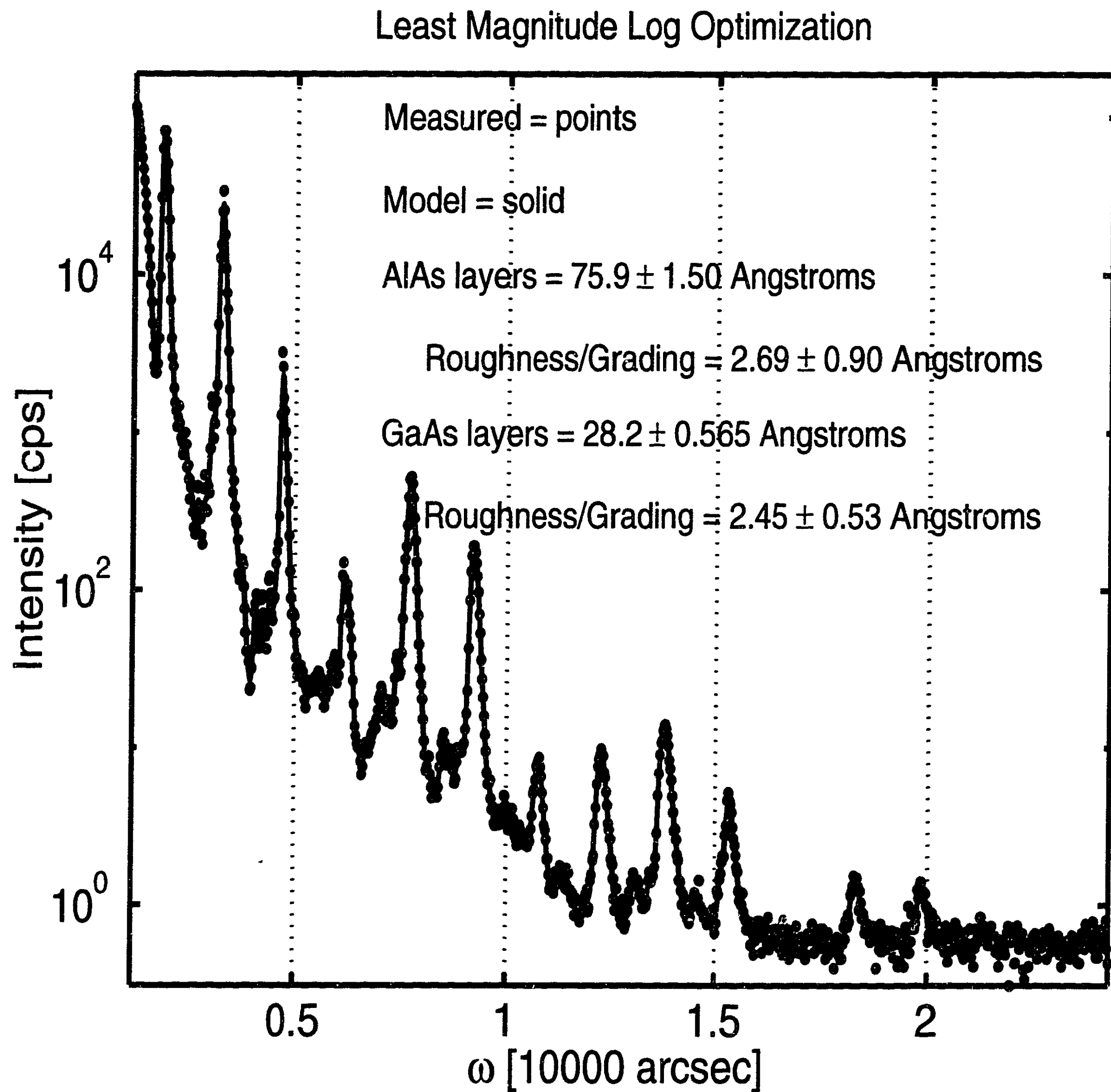


Figure 6-9: The points in this figure represent a GIXR measurement of Sample 9089 intended to be a superlattice consisting of 50 periods of 80 Å AlAs and 27 Å GaAs, with a cap layer of 130 Å of GaAs. The solid line in the figure represents the reflectivity calculated for a superlattice in which the layer widths and roughnesses are optimized to minimize the magnitude of the difference in the logarithms of the measured and calculated reflectivities.

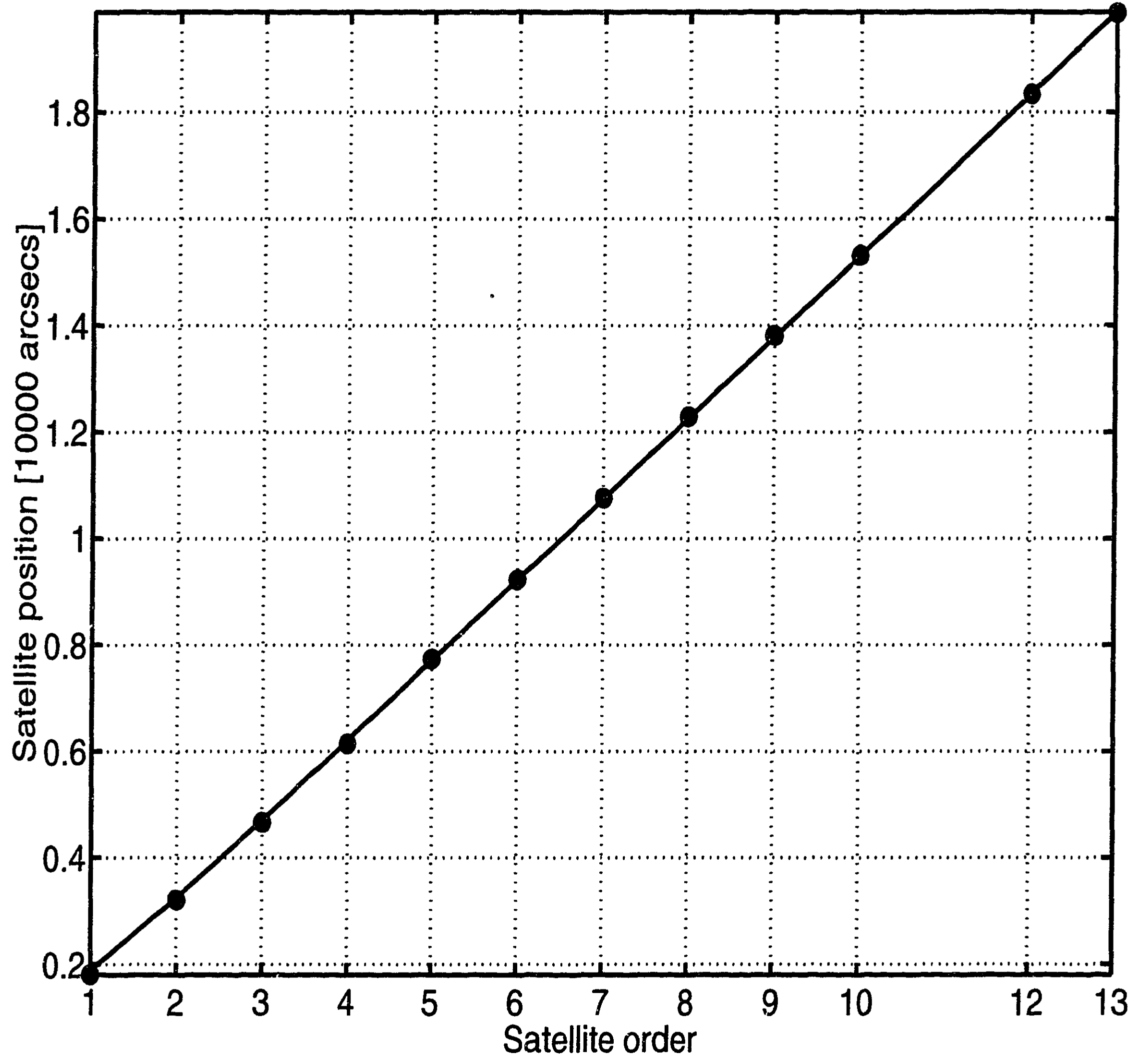


Figure 6-10: The points in this figure denote the measured superlattice diffraction peak positions obtained from Fig. 6-12. The solid line in this figure is a least squares best fit of the measured points to Equation (6.30). The spacing between the diffraction peaks (the slope of the solid line) yields the average measured superlattice periodicity, which was found to be 104 Å for Sample 9089. Extrapolation of this curve to the zeroth diffraction order (the intercept of this curve with the ordinate axis) yields the critical angle for total external reflection, which was measured to be about 1100 arcsecs=0.31° for Sample 9089.

150 Å on which there is an 22.91 Å oxide layer of roughness 6.80 Å. The need to include a thin oxide layer to get a better match of the simulated to the measured GIXR data has been reported previously in the literature [71]. (The optimized 7 period superlattice was found by Kevin Matney at Bede Scientific using their test software.)

The question arises as to the uniqueness of the layer structure found from this optimization procedure. The answer is that the period of the superlattice and the individual layer widths comprising each period are found with small uncertainty. The period is inferred from the spacing of the diffraction orders, as shown in Fig. 6-10, and an uncertainty in the inferred period manifests itself as a deviation of the measured data points from the best fit straight line in Fig. 6-10. The individual layer widths comprising each period are found from a comparison of measured and the simulated strengths of the different diffraction peaks, as shown in Fig. 6-9. Though the mean layer widths are found from the simulation with small uncertainty, the simulation may not yield definitive unique values for the layer width and roughness variations as well as the mean roughnesses. The statistics of the layer width and roughnesses found in the simulation of Fig. 6-9 should not be taken too literally. For one thing, there is an extra periodicity coming from the repetition of the 7-period superlattice seven times in the simulation of Fig. 6-9. This extra periodicity in the simulation appears in the figure as the seven small peaks (having a separation of about 214 arcsec) which appear between the much stronger diffraction peaks (having a separation of about 1500 arcsec).

The question also arises as to how sensitive the simulated reflectivity is to this artificial periodicity introduced by the simulation and to the exact order of the layers in the model superlattice. To test how sensitive the simulated reflectivities are to the exact layer structure, different 50-period superlattices were generated with layer widths and roughnesses having a Gaussian distribution of the same mean and variances as those in the simulation of Fig. 6-9. These different model superlattices were found to yield simulated reflectivities that reproduce the peak positions shown in Fig. 6-9, but the peak intensities are not quite reproduced. One such simulation is shown in Figure 6-11. Of particular note is that the two highest diffraction peaks (of order 12 and 13)

simulated in Fig. 6-11 are lower in intensity than the measured intensities.

In Section 6.6 below, we show that the spread of the full-width-half-maxima (FWHM) with diffraction order for the higher order diffraction peaks is a measure of the layer width variation. This deterioration of the fit for the higher order diffraction peaks of Fig. 6-9 affects the assessment of the layer width variation. In fact, in Section 6.6 below, we find that the superlattice period has a distribution with a FWHM of 2 Å. The best fit model of Fig. 6-9 corresponds to a superlattice period having a distribution with a FWHM of $2\sqrt{2\log(2)(1.50^2+0.565^2)}$=3.77 Å. The agreement is not perfect because the fit shown in Fig. 6-9 is not as good for the higher order diffraction peaks.

The reason for the apparent sensitivity of the simulated reflectivity to the precise layer structure is believed to be insufficient statistics in solving for the properties of only 7 periods of the superlattice. If the layer widths and roughnesses of all 50 periods were found directly in the optimization procedure, the simulated reflectivity would probably better match the measured reflectivity, but the computation time would be prohibitive.

In fact, when the mean value of the AlAs and GaAs layer widths in the simulations were changed slightly to 75.6 Å and 28.4 Å, respectively (keeping all other simulation parameters the same as in Fig. 6-9), different simulated superlattices having the same Gaussian distributions for the layer widths and roughnesses all yielded very similar reflectivities. One such structure is studied in Figure 6-12. The solid line in the figure represents a calculated reflectivity [70] with a good visual fit to the measured data in which the superlattice is modeled as having a Gaussian layer width distribution of mean 75.6 Å and 28.4 Å for the AlAs and GaAs, respectively) and some standard deviation (0.68 Å and 1.01 Å for the AlAs and GaAs, respectively). The interfaces between the GaAs and AlAs layers are modeled as not sharp, and have an interfacial grading or roughness that spans a distance which is Poisson distributed with some mean (3.52 Å and 2.98 Å for the AlAs and GaAs, respectively) and some standard deviation (1.73 Å and 1.60 Å for the AlAs and GaAs, respectively). The model also assumed a GaAs cap layer of 120.4 Å on which there is an 22.9 Å oxide layer of roughness 11.68 Å. Better agreement than in Fig. 6-11 is observed in Fig. 6-12 between

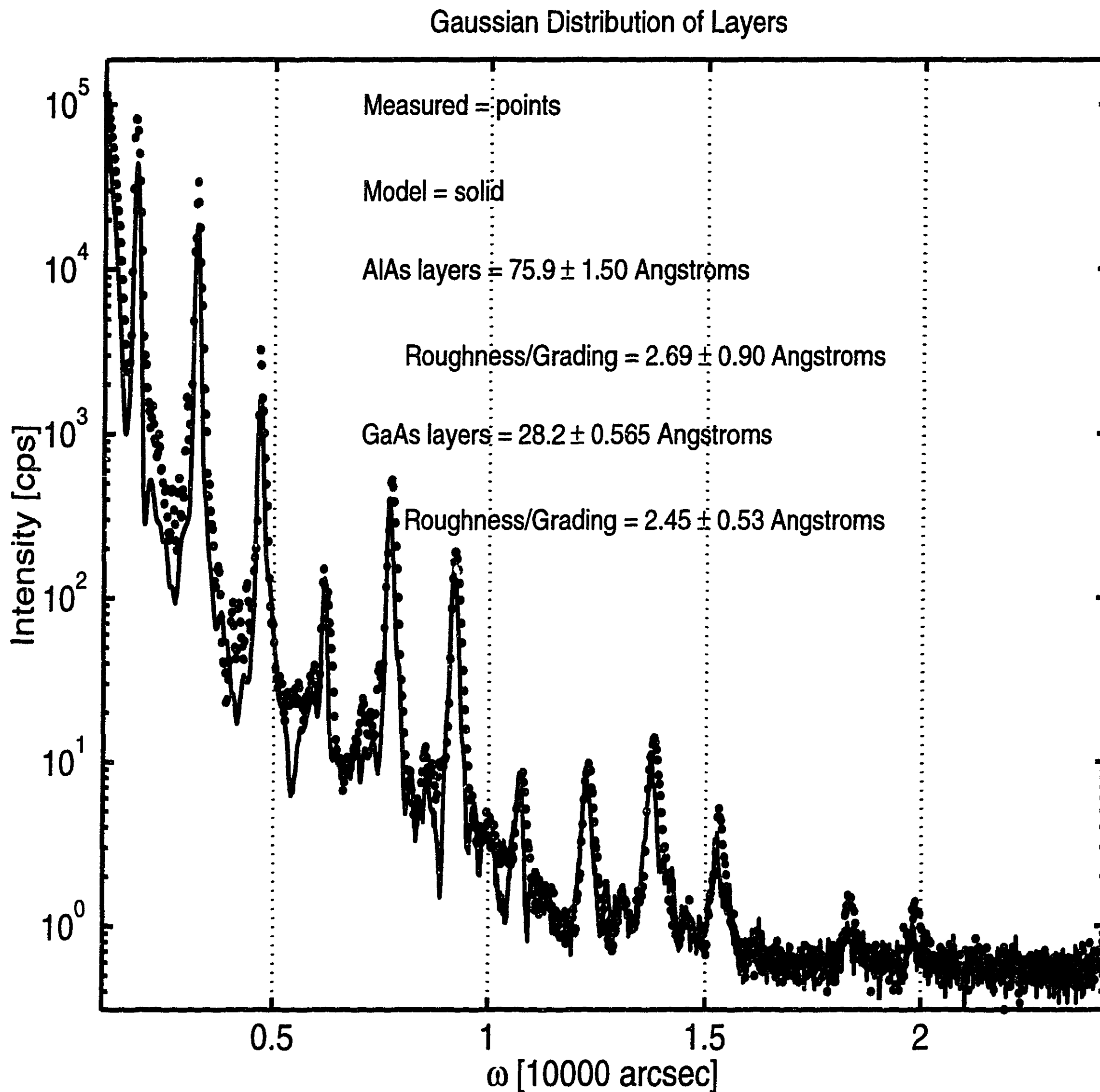


Figure 6-11: The points in this figure represent a GIXR measurement of Sample 9089 intended to be a superlattice consisting of 50 periods of 80 Å AlAs and 27 Å GaAs, with a cap layer of 130 Å of GaAs. The solid line in the figure represents the reflectivity calculated for a superlattice in which the layer widths and roughnesses are Gaussian distributed with the same mean and variances as those found from the simulation of the optimized structure shown in Fig. 6-9.

the measurement and the modeled reflectivity for the highest order (the twelfth and thirteenth) diffraction peak intensities.

That the simulation of Fig. 6-12 is a good fit to the measured data regardless of the exact order of the layers used in the simulation (as long as the simulated structures all have the same Gaussian distribution of layer parameters) shows that we have a good idea of what the layer structure is. As a comparison, two other layer structures are simulated in Figure 6-13 as examples of what Sample 9089 is NOT comprised. The two upper curves in Fig. 6-13 are the measured and simulated reflectivities shown previously in Fig. 6-12. The third curve in Fig. 6-13 shows the simulated reflectivity of a structure in which all 50 superlattice periods have the same layer widths and roughnesses. Such a simulated [72] reflectivity shows diffraction peaks which will always be too narrow to match the measurement, especially at higher diffraction order. The fourth curve in Fig. 6-13 is for a structure having the same Gaussian distribution for the layer parameters as used for the second curve in the figure, but the layer widths have been ordered to increase monotonically from the substrate to the sample surface. Such a structure has a simulated reflectivity whose diffraction peaks are much broader and smaller (in peak intensity) than the measured ones. Thus, the layers comprising Sample 9089 have a Gaussian distribution of layer parameters, and the layer widths are randomly ordered.

The distribution (mean and standard deviation) of the layer roughnesses used in the simulation of Fig. 6-12, which gave a good visual fit to the measured data, was taken to be slightly different from that used in the best fit simulation of Fig. 6-9. The layer widths chosen in the simulation shown in Fig. 6-12 were changed slightly to give a better fit to the peak positions than in Figs. 6-9 and 6-11, which show simulated peak positions which are always at a slightly lower angle than the measured peak position. The roughness variations chosen in the simulation shown in Fig. 6-12 are larger than those used in Figs. 6-9 and 6-11, in an effort to get the simulated intensities of the twelfth and thirteenth order diffraction peaks to match experiment. (The simulation of a roughness that was the same for all layers required a value that was unrealistically small (0.5 monolayer for all layers) to fit the measured data.) The roughness of (1.0 ± 0.5) monolayer seen in these simulations is in good agreement

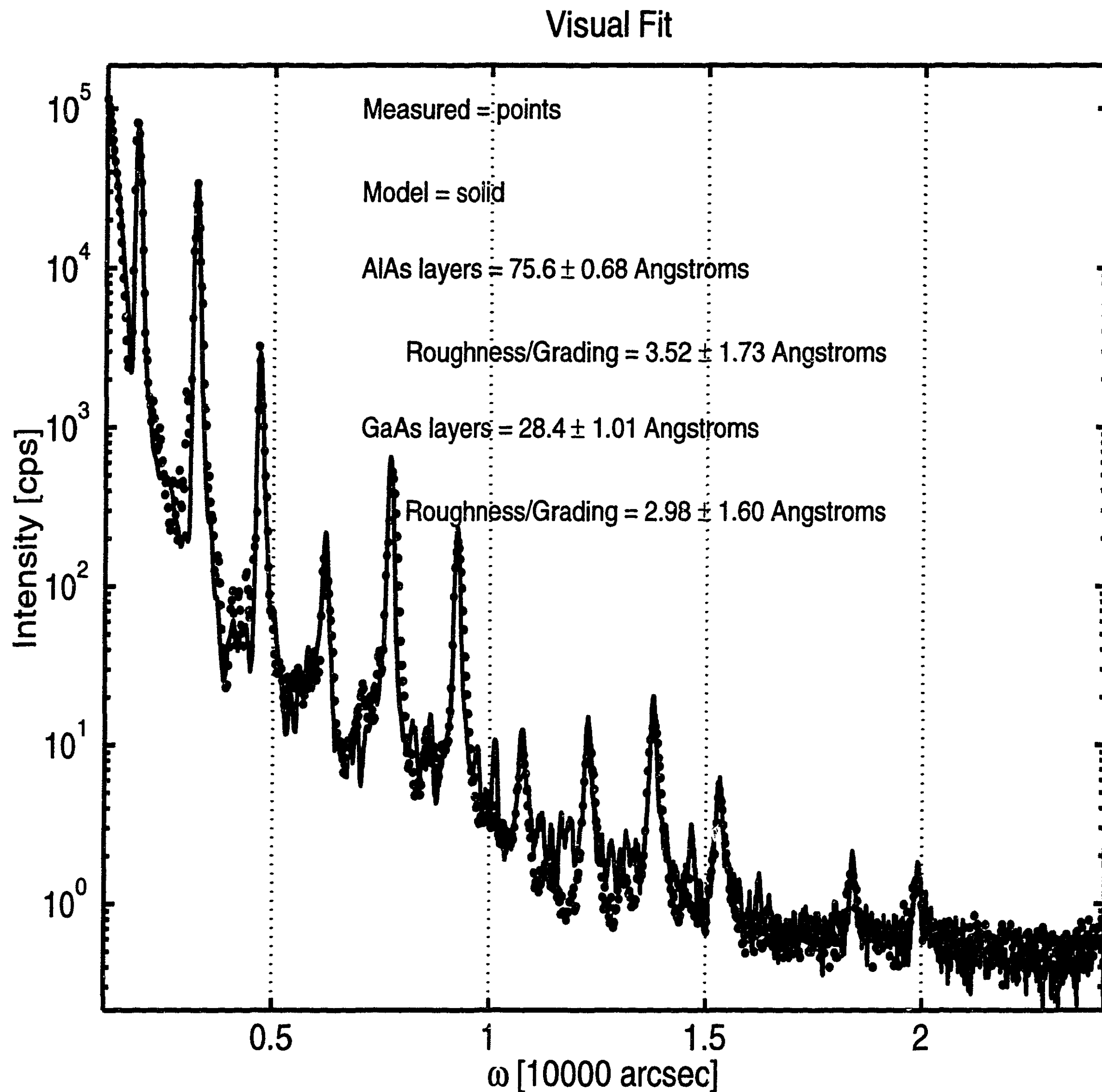


Figure 6-12: The points in this figure represent a GIXR measurement of Sample 9089 intended to be a superlattice consisting of 50 periods of 80 Å AlAs and 27 (0.68 Å and 1.01 Å for the AlAs and GaAs, respectively). The solid line in the figure represents a calculated reflectivity with a good visual fit to the measured data in which the superlattice is modeled as having a Gaussian layer width distribution of some mean (75.6 Å and 28.4 Å for the AlAs and GaAs, respectively) and some standard deviation (0.68 Å and 1.01 Å for the AlAs and GaAs, respectively).

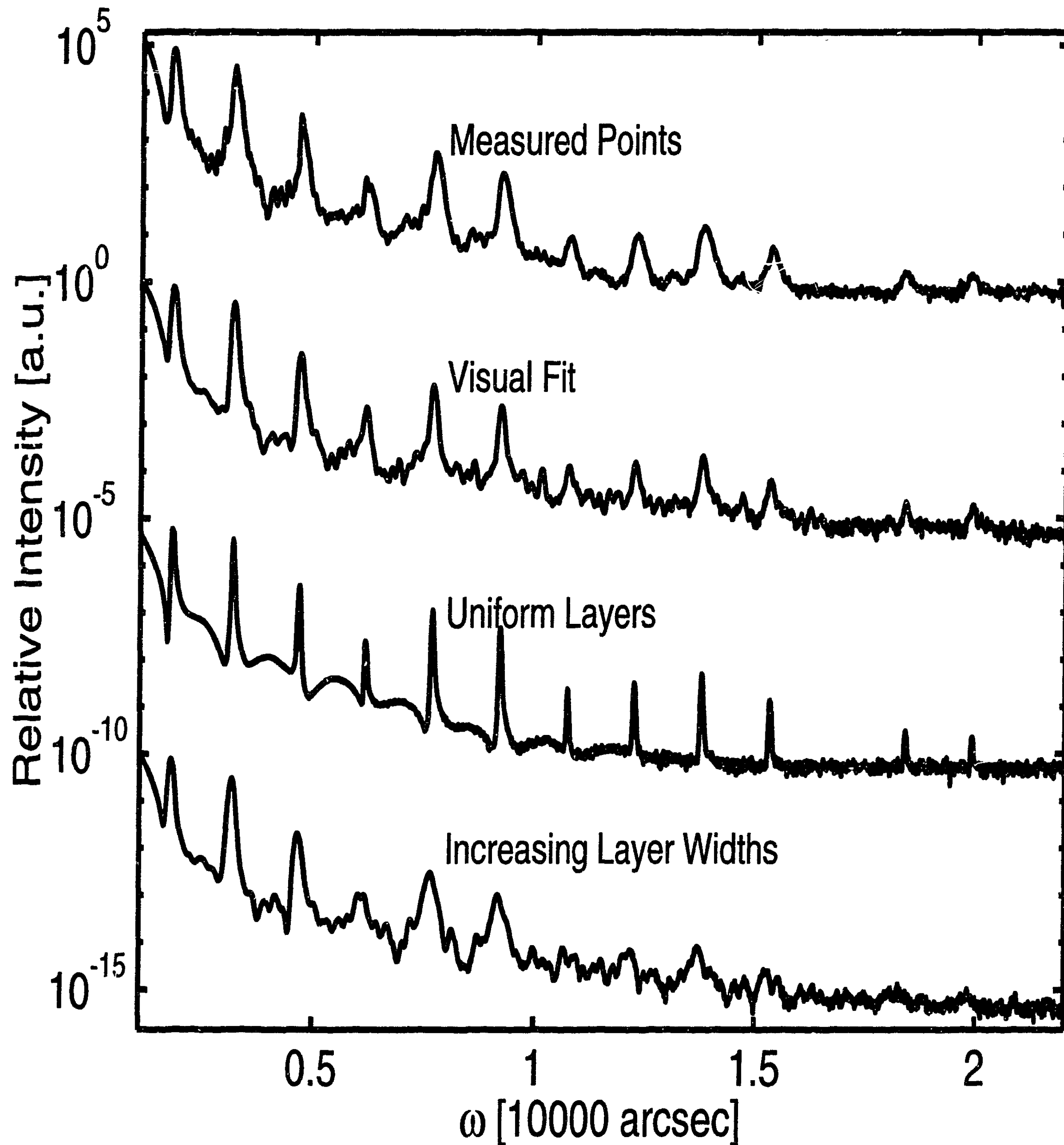


Figure 6-13: Comparison of the measured and simulated reflectivities as a function of incident angle for Sample 9089. The upper curve is the measured reflectivity. The second curve is the simulation from Fig. 6-12, showing a good visual fit to the measured data. The third curve was calculated by a simulation which assumed that all the layer widths were identical and equal to the average measured layer widths. The fourth curve was calculated by a simulation which assumed the same set of layer widths as for the second curve, but which also assumed that the layer widths were increasing monotonically during growth.

with the literature. The reports of *roughness* of growths by MBE [67] seem to be in agreement for the *absolute size* of the roughness: about 1-3 monolayers. This is not surprising, as the standard MBE growth front involves about 3 monolayers. Thus, the roughness of the sample interfaces may be improved with different growth techniques, such as growth interruptions or migration enhanced epitaxy.

Modeling of the GIXR is done by solution of the Parratt (transmission line) equations for the electromagnetic fields in a layered structure. The grading and roughness at the interfaces is accounted for in a distorted wave Born approximation [73, 70] by replacing the standard Fresnel reflection coefficient at an interface $r_{\text{Fresnel}} = \frac{k_{z,m} - k_{z,m+1}}{k_{z,m} + k_{z,m+1}}$, where $k_{z,m}$ is the component of the wave vector along the growth direction z in the layer designated "m," by $\langle r \rangle$,

$$\langle r \rangle = r_{\text{Fresnel}} \exp(-2\sqrt{k_{z,m}k_{z,m+1}}\sigma^2) \int_{-\infty}^{\infty} g(z) \exp(2i\sqrt{k_{z,m}k_{z,m+1}}z) dz, \qquad (6.31)$$

where σ is the standard deviation of a Gaussian distribution for the lateral (in-plane) distribution of the roughness, and g(z) is the normalized first derivative of the electric susceptibility. In equation (6.31), the first exponential denotes the roughness of the interface, and the Fourier integral denotes the grading of the electron distribution across the interface. Roughness and grading cannot be distinguished in a GIXR measurement because both effects appear in the product in equation (6.31).

For an ideal crystal, it is easy to show from the Fresnel equations for reflection that the glancing incidence reflectivity drops as the fourth power of the incident angle [67], as is necessary for a crystal which has a semi-infinite extent. Interface grading manifests itself in a sharper drop of the glancing incidence X-ray reflectivity with respect to the incident angle.

6.5.2 Glancing Incidence X-ray Scattering Measurements

Glancing Incidence X-ray Scattering (GIXS) measurements distinguish between interface grading and roughness. GIXS measurements are made in the same manner as ω scans, but are made near θ=0° instead of at the Bragg angle. Interface grading and

roughness can be distinguished because rough surfaces enhance the diffuse scattering of the glancing incidence X-rays. In contrast, smooth but compositionally graded interfaces enhance the transmission into (and not the scattering from) the sample.

Calculation of the diffuse scattering of glancing incidence X-rays is very involved. A key contribution to the diffuse scatter is the surface structure factor S($\vec{\mathbf{q}}$) for a given scattered wave vector $\vec{\mathbf{q}}$ [74]:

$$S(\vec{\mathbf{q}}) = \frac{1}{|q_z|^2} \exp\left(-\left[q_z^2 + q_z^{*\,2}\right]\sigma^2/2\right) \int_{\text{illuminated surface}} \left[\exp\left\{|q_z|^2 C(x)\right\} - 1\right] \exp(iq_x x)dx, \tag{6.32}$$

where C(x) is a function describing whether correlations in the height (in the z or growth direction) of an interface exist with an in-plane separation of x,

$$C(x) = \langle z(0)z(x)\rangle \equiv \sigma^2 \exp\left[-(|x|/\xi)^{2h}\right], \tag{6.33}$$

where σ is the root-mean-square roughness, ξ is the correlation length of the roughness, and h describes how fast C(x) drops for large $|x|$. Equation (6.32) shows that the diffuse scatter structure factor is the Fourier transform of the height correlation function.

Peaks in the diffuse scatter, known as Yoneda [75] wings, are observed when either the incident or exit angle of the X-rays is equal to the critical angle for total external reflection, for which there is a large enhancement of the surface electric field.

Three GIXS measurements of Sample 9089 are shown in Fig. 6-14. The three curves shown correspond to measurements in which the difference between the scattered angle and the incident angle are, respectively, 4430, 6230, and 9496 arcsec. In such a measurement, the detector angle is fixed at twice the angle at which the peaks occur in Fig. 6-14. (The peaks in Fig. 6-14 occur when the detector angle is exactly twice the incident angle, and is the situation in which a glancing incident reflection measurement was made.) The sample angle is then varied from glancing (zero angle of) incidence to glancing (zero angle of) exit conditions. The Yoneda wings in Fig. 6-14 are very small because the measurement was made at a superlattice diffraction peak, for which many superlattice periods are making significant contributions to the

reflected/scattered radiation, while the relative contribution of the surface roughness of any one particular layer is small.

Four GIXS measurements are shown in Fig. 6-15 for Sample 9324 consisting of a superlattice of 7 periods of 50 nm $In_{0.521}Al_{0.479}As$ and 6.73 nm $In_{0.532}Ga_{0.468}As$. The four curves shown correspond to measurements in which the difference between the scattered angle and the incident angle are, respectively, 3134, 3868, 4860, and 5398 arcsec. Large Yoneda wings (the cusps) are visible because Sample 9324 is believed to have rough interfaces. The first InGaAs layer is believed to contain rough three dimensional structures because the Ga shutter did not open for about 10 seconds. Thus, the first 5 monolayers of the first InGaAs layer is believed to be InAs.

6.6 Measurement of epitaxial layer width uniformity

6.6.1 Introduction

This section describes the measurement of epitaxial layer width variations. The amount of layer width variation can be separated from the amount of interface grading/roughness through the measurement of either a reciprocal space map or a combination of a glancing incidence reflectivity and scattering measurements.

An example of this is given for a superlattice of binary alloys: an AlAs/GaAs superlattice. Measuring the layer width variation for a superlattice of binary alloys gives a measure of the size of the composition variations expected for a ternary. A comparison is given for a superlattice of ternary alloys.

For QWIPs, composition variations affect the height of the confining potential barrier, and thus affect the thermal leakage current uniformity. Both composition variations and roughness affect the layer width uncertainty, which affects the variations in the responsivity, and sequential resonant and thermionic field assisted tunneling. The

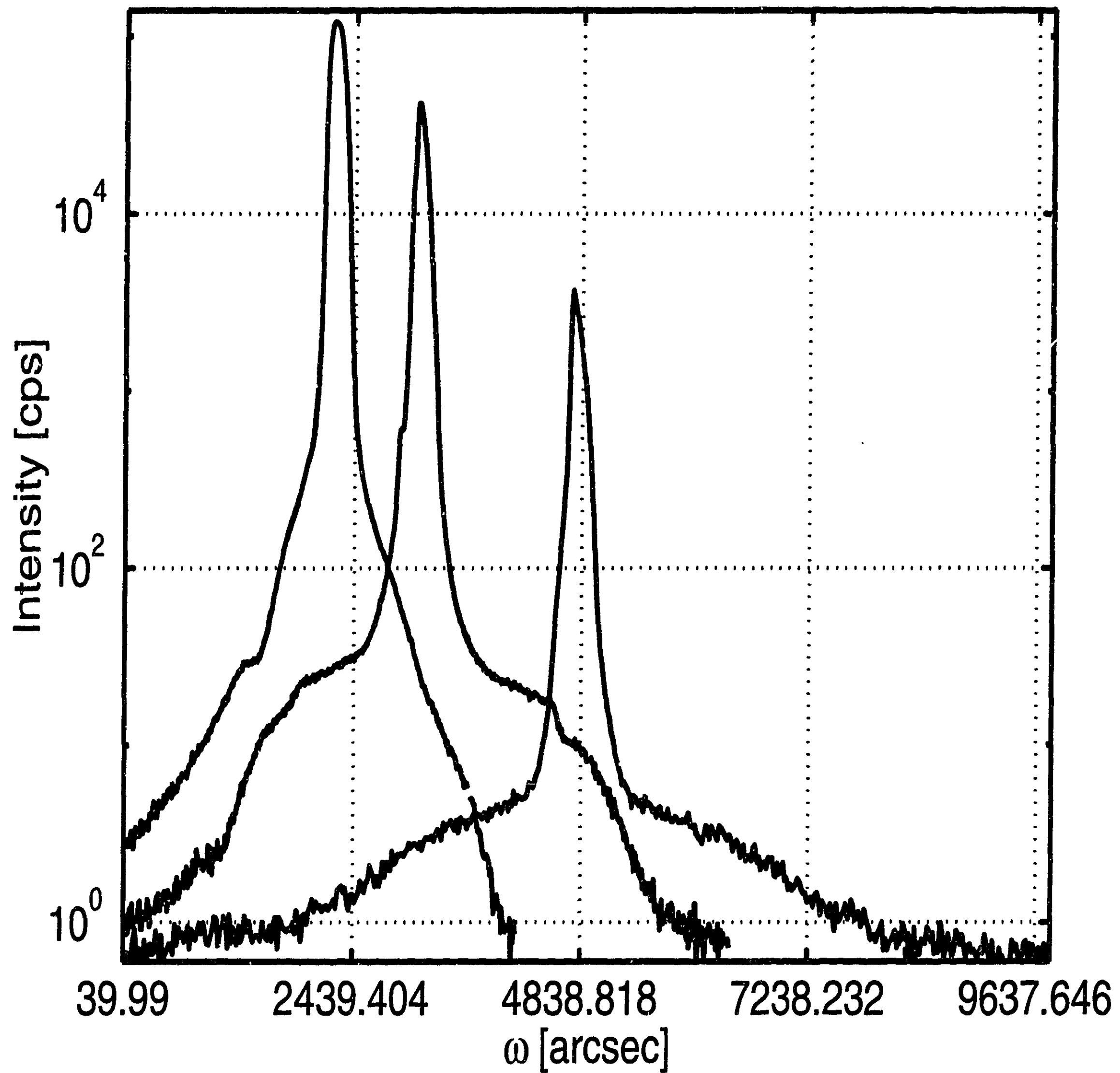


Figure 6-14: Three GIXS measurements of Sample 9089 consisting of a superlattice of 50 periods of 75.6 Å AlAs and 28.4 Å GaAs. The scattered intensity is plotted as a function of the incident angle. The three curves shown correspond to measurements in which the difference between the scattered angle and the incident angle are, respectively, 4430, 6230, and 9496 arcsec. This superlattice was designed to be 50 periods of 80 Å AlAs and 27 Å GaAs, with a cap layer of 130 Å of GaAs.

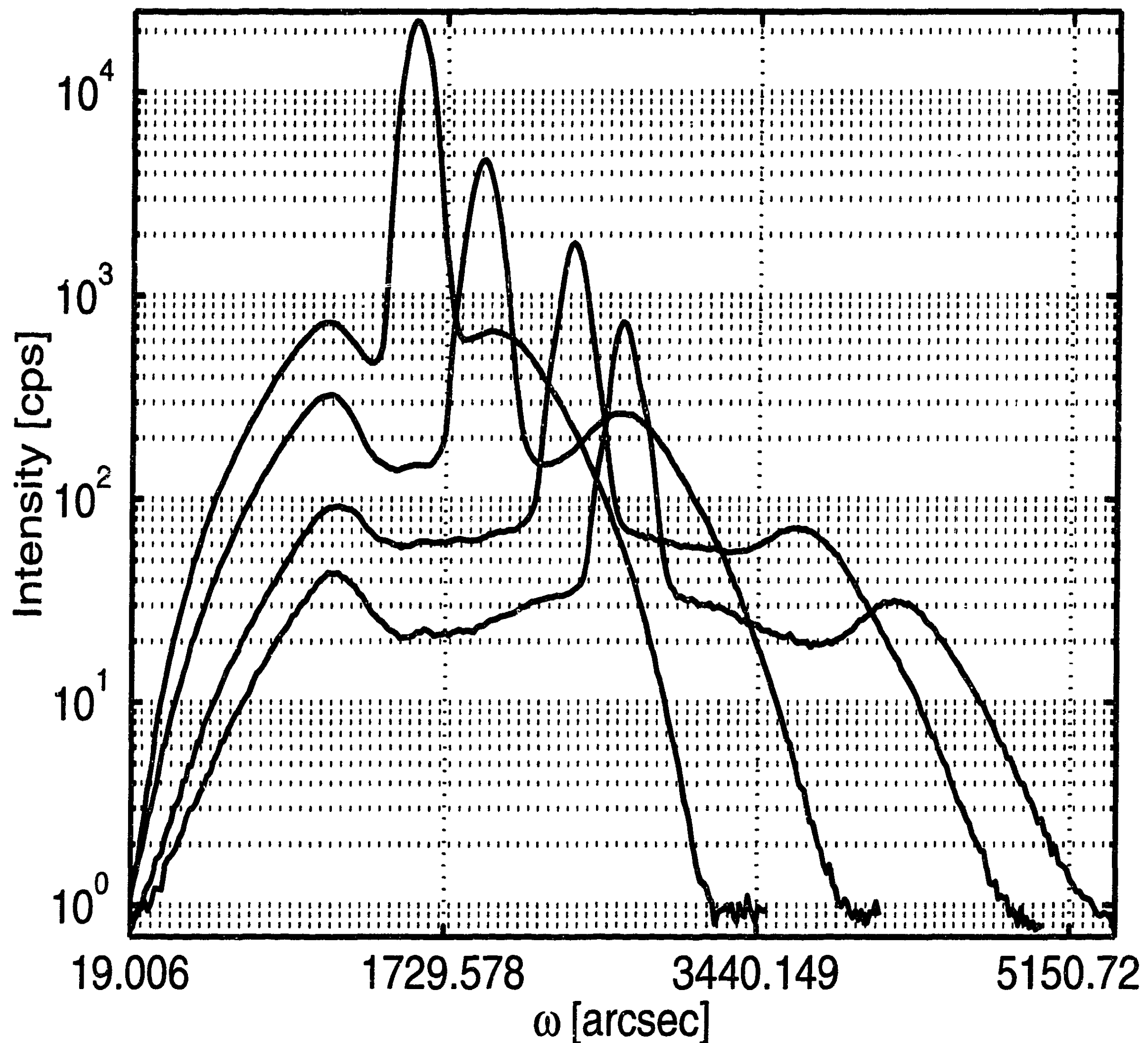


Figure 6-15: Four GIXS measurements of Sample 9324 consisting of a superlattice of 7 periods of 50 nm $In_{0.521}Al_{0.479}As$ and 6.73 nm $In_{0.532}Ga_{0.468}As$. The scattered intensity is plotted as a function of the incident angle. The four curves shown correspond to measurements in which the difference between the scattered angle and the incident angle are, respectively, 3134, 3868, 4860, and 5398 arcsec. Large Yoneda wings (the cusps) are visible when either the incident or exit angle equals the critical angle of total external reflection. Sample 9324 is believed to have rough interfaces because the first InGaAs layer contained rough three dimensional structures: since the Ga shutter did not open for about 10 seconds, the first 5 monolayers of the first InGaAs layer is believed to be InAs.

lateral size of the roughness affects the uniformity across large arrays of QWIPs.

6.6.2 Distinguishing layer width variations from interface grading/roughness

The separation between superlattice diffraction peaks in a GIXR or a double crystal rocking curve or a θ-2θ measurement near a symmetric Bragg peak yields the periodicity D of the superlattice, according to the Bragg law for the superlattice,

$$(\theta_n - \theta_m) = \frac{(n-m)\lambda}{2D\cos([\theta_n + \theta_m]/2)}, \tag{6.34}$$

where θ_m is the angular position of the m-th order diffraction peak, and the argument of the cosine in Equation (6.34) can be approximated as zero degrees for a GIXR measurement and the Bragg angle for a rocking curve or θ-2θ measurement.

Fewster [76, 77] pointed out that if the superlattice period varies throughout the growth, having some statistical distribution of mean D and variation ΔD, then the full-width-half-maximum (FWHM) of the diffraction peaks will increase with the diffraction order. The reason is that each superlattice period contributes to the position of the n-th order diffraction peak as indicated in Equation (6.34), so that a variation in the superlattice period D is magnified at the higher diffraction orders, (n-m), according to Equation (6.34). If variable period thicknesses ΔD are the dominant contribution to the FWHM $\Delta\theta_{n,\text{variable thickness}}$ of the n-order diffraction peak, then the increase in the FWHM of the diffraction peaks with the diffraction order (n-m) can be found by differentiating Equation (6.34) with respect to D. The result,

$$(\Delta\theta_{n,\text{variable thickness}} - \Delta\theta_{m,\text{variable thickness}}) = \frac{2\Delta D(\theta_n - \theta_m)^2\cos([\theta_n + \theta_m]/2)}{(n-m)\lambda}, \tag{6.35}$$

was obtained by Fewster [76, 77]. A simpler form for this equation can be obtained by substituting Equation (6.34) into Equation (6.35), resulting in,

$$(\Delta\theta_{n,\text{variable thickness}} - \Delta\theta_{m,\text{variable thickness}}) = (\theta_n - \theta_m)\left(\frac{\Delta D}{D}\right). \tag{6.36}$$

Thus, the FWHM of the n-th diffraction peak, when expressed as a fraction of the position of the n-th diffraction peak, is directly proportional to the variation of the superlattice period throughout the growth.

Since most semiconductors have a refractive index less than unity at the X-ray wavelengths, the incident X-ray beam is refracted upon entry into the semiconductor according to Equation (6.28). To analyze a GIXR measurement, the semiconductor refractive index must be accounted for, and the equation analogous to Equation (6.36),

$$\Delta\theta_{I,n,\text{variable thickness}} = \theta_{I,n}\left(\frac{\Delta D}{D}\right), \tag{6.37}$$

is written for the angle $\theta_{I,n}$ of the X-ray beam in the interior of the semiconductor. Since the diffraction peaks are measured as a function of the incident angle $\theta_{E,n}$ in the exterior of the semiconductor, then the FWHM of the measured diffraction peaks must be obtained from both equations (6.28) and (6.37), and is related to variations ΔD in the superlattice periodicity through,

$$\Delta\theta_{E,n,\text{variable thickness}} = \frac{\theta_{I,n}^2}{\theta_{E,n}}\left(\frac{\Delta D}{D}\right). \tag{6.38}$$

In order to get a good fit to Equation (6.38), the contributions to the FWHM which are unrelated to layer width variations must be removed. If we denote by $\Delta\theta_0$ the contributions to the FWHM which are unrelated to layer width variations, then the measured FWHM is the sum (in quadrature) of the contributions related and unrelated to the layer width variations,

$$\Delta\theta_{n,\text{measured}}^2 = \Delta\theta_{n,\text{variable thickness}}^2 + \Delta\theta_0^2. \tag{6.39}$$

The subscript "0" in $\Delta\theta_0$ denotes the zeroth order diffraction peak, as the FWHM of the zeroth order diffraction peak is unaffected by layer width variations. Contributions to $\Delta\theta_0$ come from instrumental broadening and wafer curvature. Wafer curvature is an important factor at small incident angles, as the portion of the sample surface illuminated by the X-ray beam goes as (X-ray beam size)/$\sin\theta_E$. For an X-ray beam of cross sectional size (250 μm)^2 at an incident angle of $\theta_E = 0.31°$, the portion of the sample surface illuminated is 250 μm$\times$4.6 cm. This large area of the sample surface illuminated by the X-ray beam indicates that GIXR measurements yield information

about uniformity of the epitaxial layers in the transverse (along the wafer surface) as well as the growth directions.

An example of a superlattice of binary alloys

Figure 6-16 shows the dependence of the FWHM on the superlattice diffraction order for the reflectivity measurement of Sample 9089 shown in Fig. 6-12. The points in this figure are measured data, and the solid line is a least squares best fit of the measured data to Equation (6.39). The best fit of the measured data yields the parameters $\Delta\theta_0$=88 arcsecs and ΔD=2.0 Å.

While reports of the roughness of MBE growths are in agreement over the *absolute size* (in numbers of monolayers) of the roughness, the reports of the layer width variations of MBE growths are in agreement over the *fractional* variation of the layer widths. The measured *fractional* layer width variation of ΔD/D=2.0 Å/104 Å=1.92% for our Sample 9089 is very similar to the *fractional* layer width variation measured for our other superlattice growths, one of which is described below to have a layer width variation of 2.05% (a 11.6 Å variation for a 565.7 Å period). This fractional layer width variation of superlattice growths is also in agreement with the literature. Matney and Goorsky [71] report that for a distributed Bragg reflector (DBR) consisting of 15 periods of 857 Å AlAs/718 Å GaAs, the variation (=FWHM=$2\sqrt{2\log 2}\times$ standard deviation) of the DBR period was 31.9 Å which is a 2.03 % layer width variation. The *absolute size* of the layer width variation is different for these three structures (2.0 Å, 11.6 Å, and 31.9 Å, respectively), but the *fractional* layer width variation is about 2% for all three growths.

Since both our Sample 9089 and the DBR structure of Matney and Goorsky involve superlattices of binary (AlAs/GaAs) alloys, then both our layer width variations are a direct measure of the variation in the incorporation of one single Group III species during MBE growth. This measured fractional layer width variation of 2 percent is about the size of the fractional variation in the Group III flux that arises from small changes in the effusion cell temperatures during growth. More quantitatively, the Group III flux F during growth can be modeled as line-of-sight evaporation [54] into

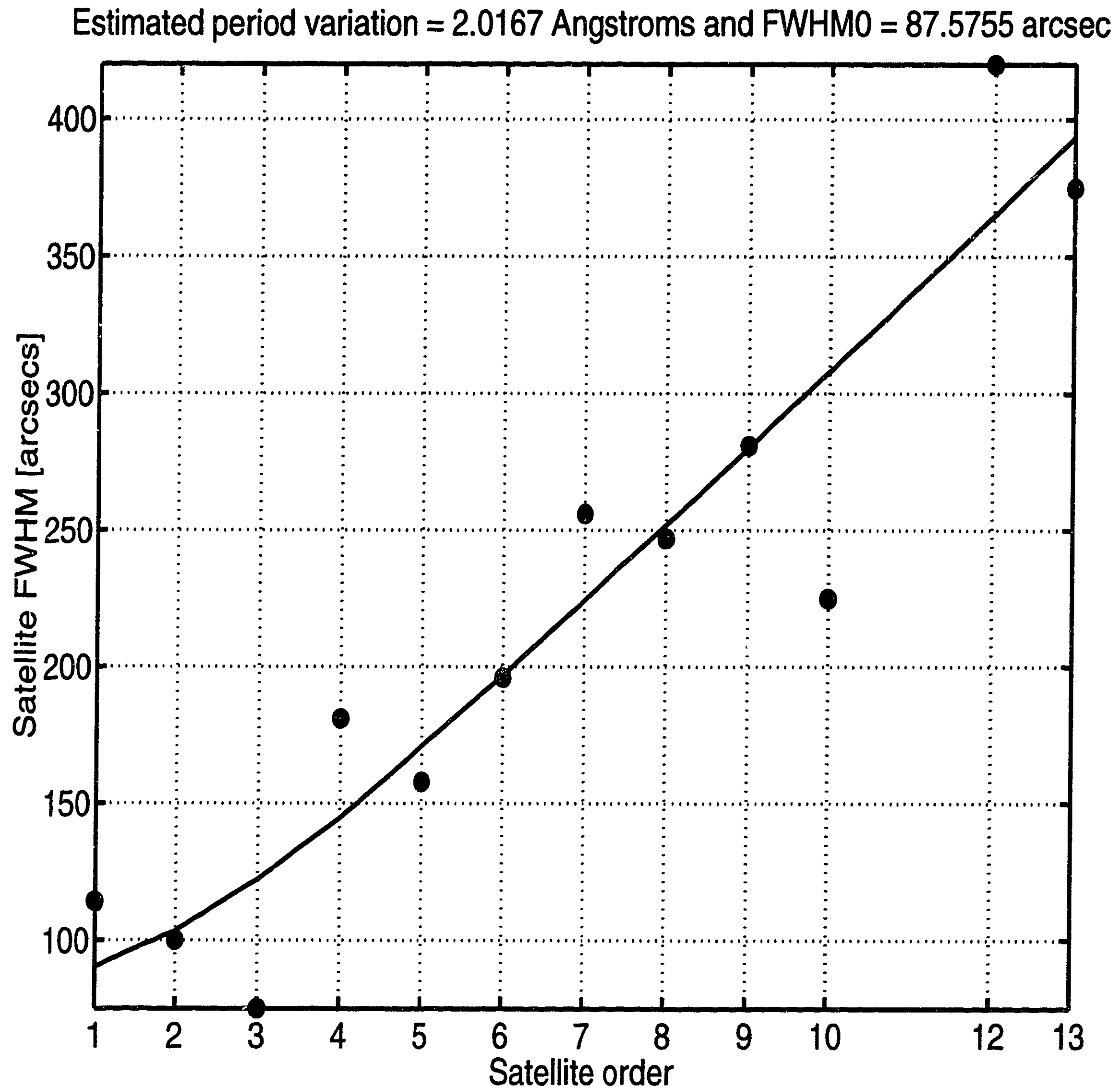


Figure 6-16: The dependence of the full-width-half-maxima (FWHM) on the superlattice diffraction order for the reflectivity measurement shown in Fig. 6-12. The points in this figure are measured data, and the solid line is a least squares best fit of the measured data to Equation (6.39). This best fit shows a layer width variation of 2.0 Å for a measured superlattice period of 104 Å.

Group III Source	Activation Energy E_A (eV)	Growth Temperature T (°C)	Flux Variation $\Delta F/F$ (percent)
Indium	2.26	780	2.36
Gallium	2.50	890	2.14
Aluminum	2.85	1080	1.80

Table 6.2: Activation energies, growth temperatures, and fractional flux variations for the Group III effusion cells. The activation energy was found from Arrhenius plots of either the measured flux or the measured RHEED oscillations as a function of the effusion cell temperature. The typical growth temperatures correspond to 0.5μm/hour for growth of the binary semiconductor. The fractional flux variation was for an effusion cell temperature variation of 1°C.

a perfect vacuum, and varies with the effusion cell temperature T according to the Arrhenius equation,

$$F = F_0 \exp\left(-\frac{E_A}{k_B T}\right), \tag{6.40}$$

where $F_0 = P_0/\sqrt{2\pi M k_B T}$ (P_0 is the equilibrium vapor pressure and M is the mass of the species) is approximately independent of the cell temperature, and E_A is the activation energy. The measured activation energies and typical growth temperatures for our Group III effusion cells are given in Table 6.2. The fractional variation in the Group III flux,

$$\frac{\Delta F}{F} = \left(\frac{E_a}{k_B T}\right)\left(\frac{\Delta T}{T}\right), \tag{6.41}$$

for a cell temperature variation of 1°C is about 2%, as calculated in Table 6.2. This size flux variation has been measured directly with an ionization gauge in our growth chamber right after a cell shutter is opened or when a cell is ramped in temperature [78].

An example of a superlattice of ternary alloys

Another sample for which layer width variations were measured was Sample 9331 whose double crystal rocking curve is shown in Figure 6-17. The position of the InGaAs peak and the different superlattice peaks showed this sample to consist of a 3415 Å of $In_{0.527}Ga_{0.473}As$ top contact, followed by 16 periods of an 79.3 Å $In_{0.527}Ga_{0.473}As$ quantum well, and barriers consisting of a superlattice of 29.6 Å $In_{0.380}Ga_{0.145}Al_{0.475}As$ (6 repetitions) alternated with 63.6 Å $In_{0.587}Al_{0.413}As$ (5 repetitions). The intended structure was a 3500 Å of $In_{0.532}Ga_{0.468}As$ top contact doped with Be at 3×10^{18} cm^{-3}, followed by 16 periods of an 81.3 Å $In_{0.532}Ga_{0.468}As$ quantum well doped with Be at 7×10^{17} cm^{-3}, and unintentionally doped barriers consisting of a superlattice of 30 Å $In_{0.388}Ga_{0.145}Al_{0.467}As$ (6 repetitions) alternated with 64.8 Å $In_{0.596}Al_{0.404}As$ (5 repetitions).

This sample was interesting for studying the layer width variations which result from the fast temperature ramps of the Group III effusion cells during growth interruptions which were used after each layer. Before the growth of the quantum well, 91 seconds were allowed to elapse while the Ga cell was ramped up 43°C at 30°C/min. The last 5 seconds of this 91-second growth interruption was used to allow the Ga cell to reach a steady state. Between the growth of the different layers of the superlattice barrier, 65 seconds were allowed to elapse while the Al cell was ramped up or down 28°C at 30°C/min. The last 7 seconds of this 65-second growth interruption was used to allow the Al cell to reach a steady state. Direct measurements [79] with a flux gauge during ramping of the effusion cells showed the Ga cell transients to be in excess of 20 seconds, and the Al cell transients to be about 4 seconds. Therefore, the largest errors in the layer widths in Sample 9331 are expected for the Ga-rich layers (the quantum well).

Figure 6-18 shows the measured reciprocal space map of Sample 9331 near the symmetric (004) Bragg peak. The peak separation along the [001] growth direction indicates a superlattice periodicity of 565.7 Å, which is slightly less than the intended period of 585.3 Å. The very narrow FWHM along the [110] ω direction of 12 arcsec remains constant for all superlattice diffraction orders shown, thus indicating an

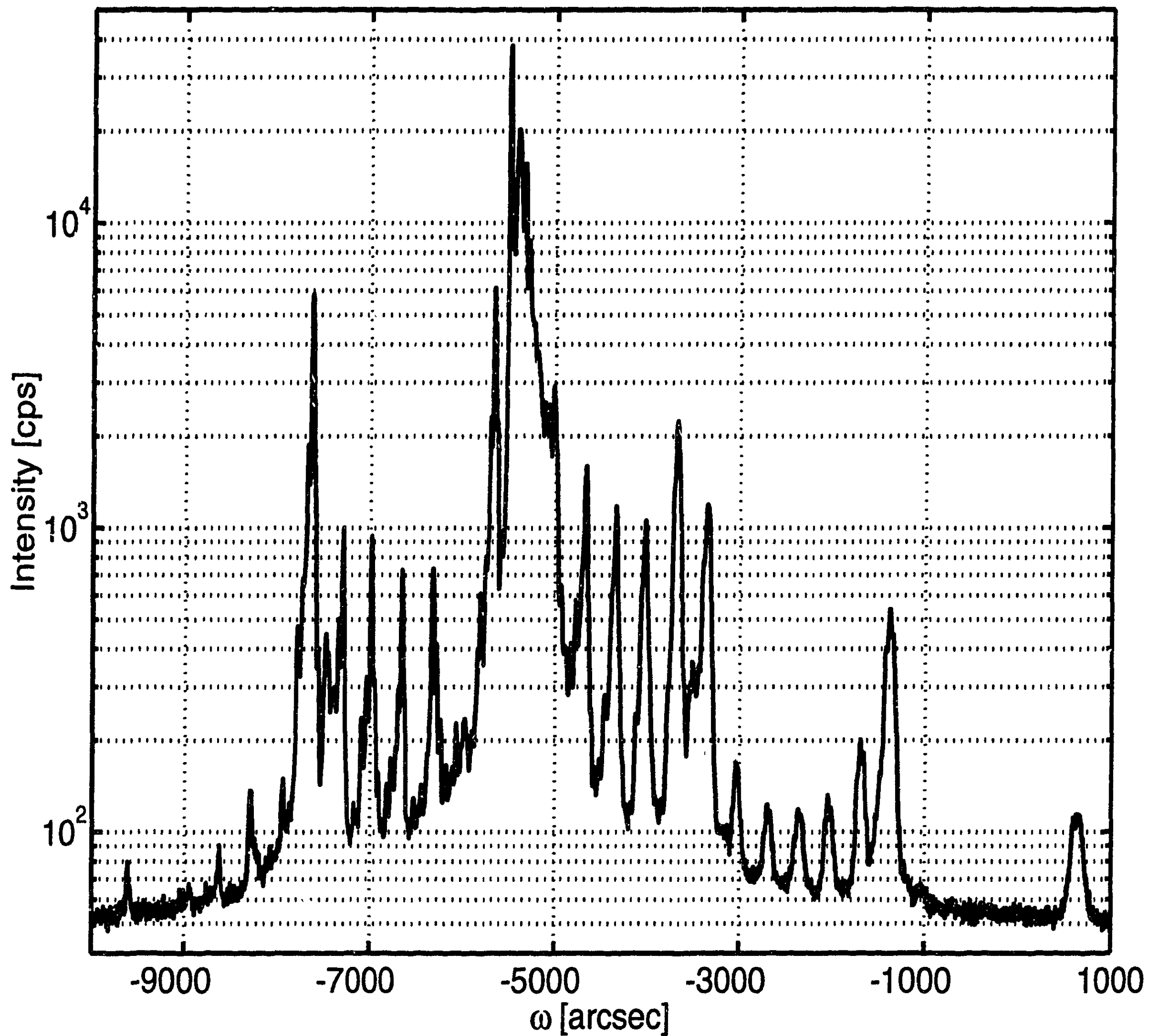


Figure 6-17: Double crystal rocking curve of Sample 9331, which consisted of a 3415 Å of $In_{0.527}Ga_{0.473}As$ top contact, followed by 16 periods of an 79.3 Å $In_{0.527}Ga_{0.473}As$ quantum well, and barriers consisting of a superlattice of 29.6 Å $In_{0.380}Ga_{0.145}Al_{0.475}As$ (6 repetitions) alternated with 63.6 Å $In_{0.587}Al_{0.413}As$ (5 repetitions). The intended structure was a 3500 Å of $In_{0.532}Ga_{0.468}As$ top contact, followed by 16 periods of an 81.3 Å $In_{0.532}Ga_{0.468}As$ quantum well, and unintentionally doped barriers consisting of a superlattice of 30 Å $In_{0.388}Ga_{0.145}Al_{0.467}As$ (6 repetitions) alternated with 64.8 Å $In_{0.596}Al_{0.404}As$ (5 repetitions). As a result of variations in the epitaxial layer widths, the higher order diffraction peaks are observed to have a larger FWHM than the lower order diffraction peaks, as would be expected from Equation (6.36).

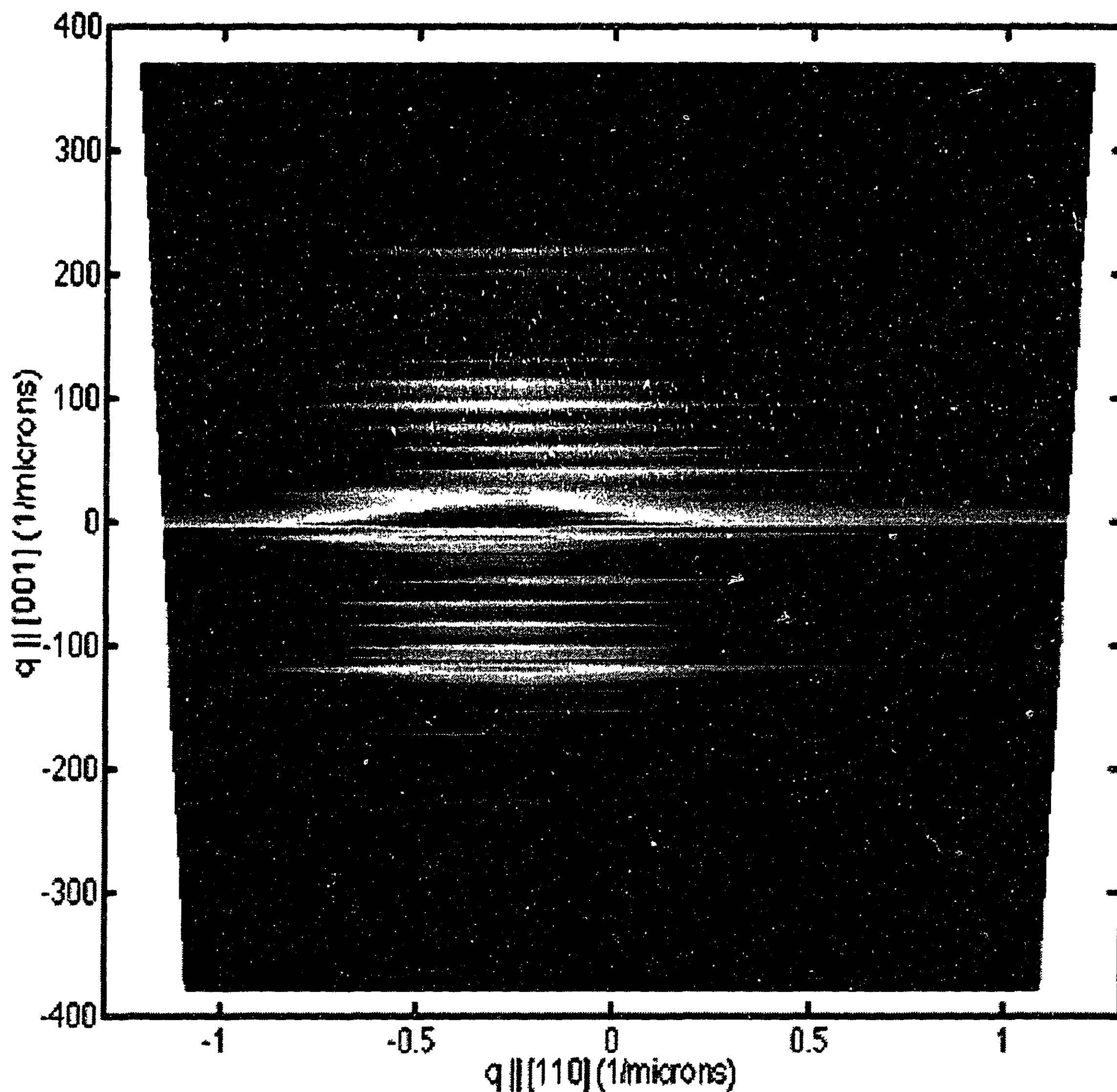


Figure 6-18: Reciprocal space map of Sample 9331 near the symmetric (004) Bragg peak. The peak separation along the [001] growth direction indicates a superlattice periodicity of 565.7 Å. The increase of the peak widths along the [001] growth direction with increasing satellite order indicates layer width variations. The very narrow FWHM along the [110] ω direction of 12 arcsec remains constant for all superlattice diffraction orders shown, thus indicating an unresolvably small mosaic tilt and interface roughness. The sample and detector angular positions are related to a point in reciprocal space through Equation (6.24).

unresolvably small mosaic tilt and interface roughness. However, the increase of the peak widths along the [001] growth direction with increasing satellite order indicates layer width variations.

Figure 6-19 shows the dependence on the superlattice diffraction order of the full-width-half-maxima (FWHM) along the [001] growth direction for Sample 9331. The points in this figure are measured data, and the solid line is a least squares best fit of the measured data to Equation (6.39). This best fit shows a layer width variation of 11.6 Å for a measured superlattice period of 565.7 Å. Again, the *fractional* layer width variation of $\Delta D/D$=11.6 Å/565.7 Å= 2.05% is similar to that observed in Sample 9089 and to that reported by Matney and Goorsky.

The effect of the long Ga cell transients is to make the InGaAs layer widths slightly different from the intended widths. However, the Ga cell *transients* are probably very similar from one superlattice period to the next, as the same effusion cell ramp rates are used during the growth of each superlattice period. Thus, the fractional layer width variation of 2.05% for Sample 9331 is very similar to that of 1.92% for Sample 9089, for which none of the effusion cells were ramped during growth.

6.7 Conclusions

Characterization of the epitaxial layers by High Resolution X-ray Diffraction (HRXRD) allows an empirical measurement of the compositions and layer widths and the variations thereof through a double crystal rocking curve, a θ-2θ measurement, or a glancing incidence reflectivity measurement. The amount of relaxation of strained layers can be measured through measurement of either a reciprocal space map or a combination of symmetric or asymmetric double crystal rocking curves.

The amount of layer width variation can be separated from the amount of interface grading/roughness through the measurement of either a reciprocal space map or a combination of a glancing incidence reflectivity and scattering measurements. The X-ray spot size in a glancing incidence measurement is very large (about 4.6 cm$\times$250 μm),

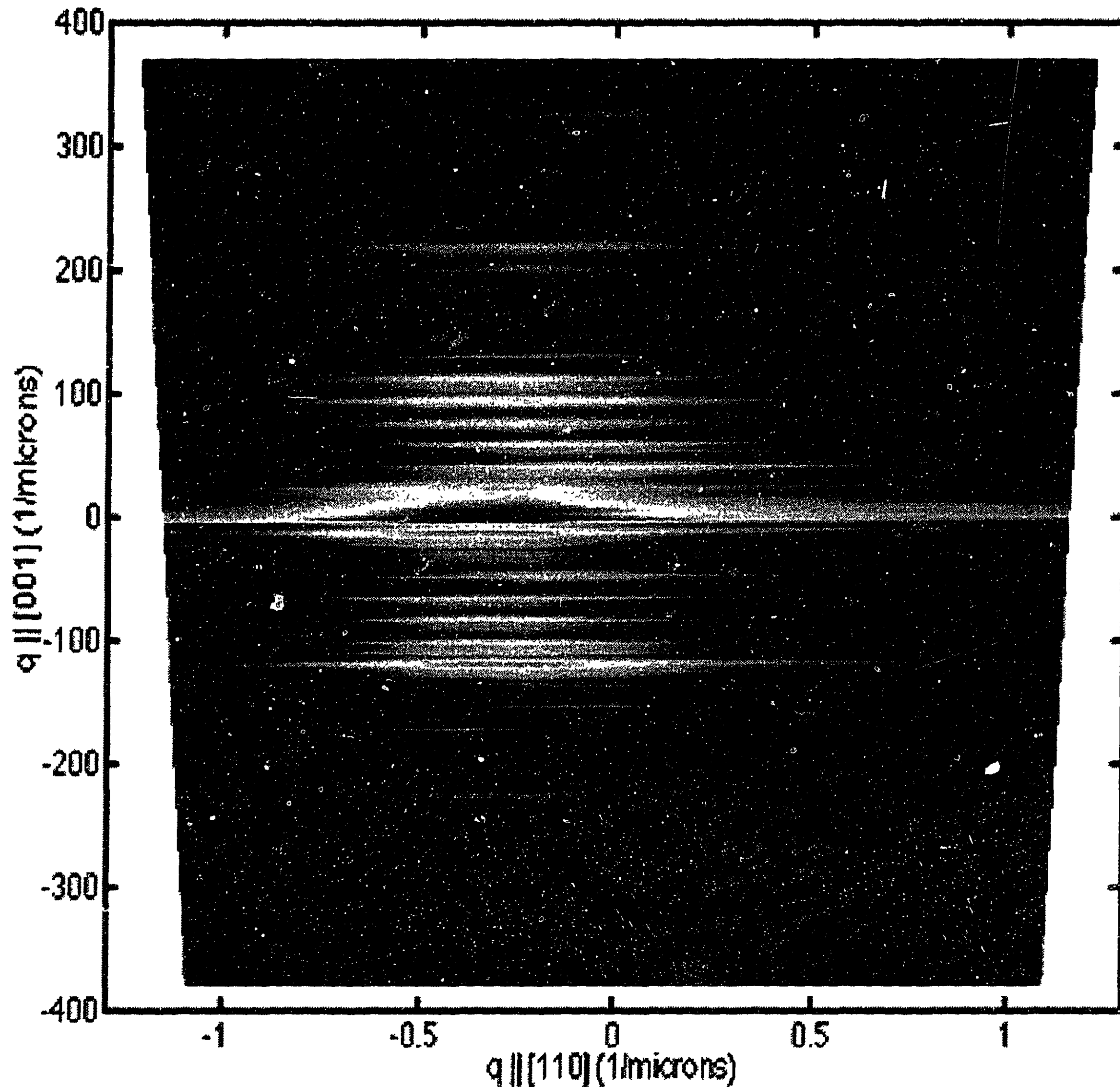


Figure 6-18: Reciprocal space map of Sample 9331 near the symmetric (004) Bragg peak. The peak separation along the [001] growth direction indicates a superlattice periodicity of 565.7 Å. The increase of the peak widths along the [001] growth direction with increasing satellite order indicates layer width variations. The very narrow FWHM along the [110] ω direction of 12 arcsec remains constant for all superlattice diffraction orders shown, thus indicating an unresolvably small mosaic tilt and interface roughness. The sample and detector angular positions are related to a point in reciprocal space through Equation (6.24).

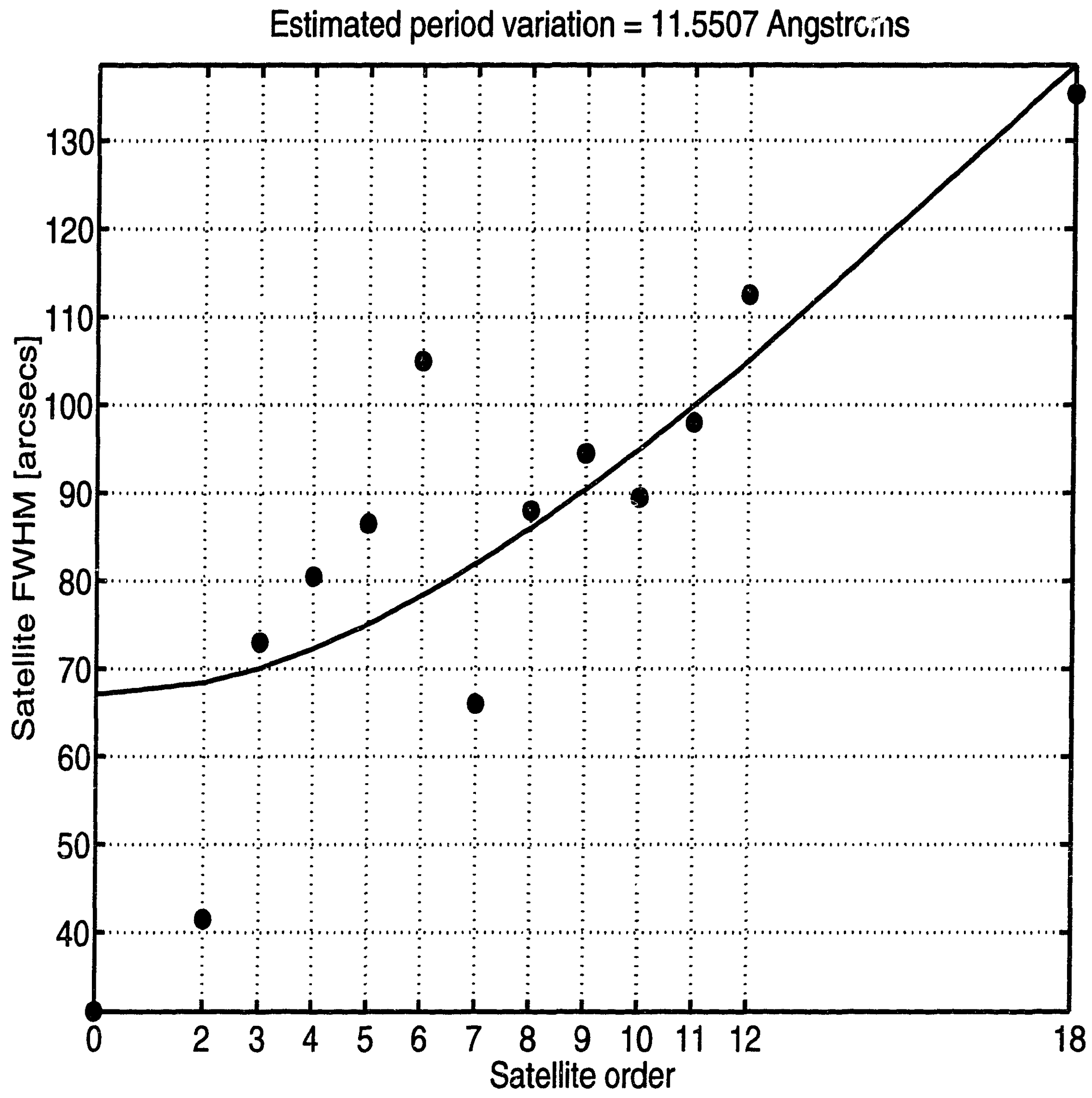


Figure 6-19: The dependence on the superlattice diffraction order of the full-width-half-maxima (FWHM) along the [001] growth direction for Sample 9331. The points in this figure are measured data, and the solid line is a least squares best fit of the measured data to Equation (6.39). This best fit shows a layer width variation of 11.6 Å for a measured superlattice period of 565.7 Å.

and thus yields information about epitaxial layer uniformity across a wafer. An example presented here of a glancing incidence X-ray reflectivity measurement of an AlAs/GaAs superlattice with a period of 104 Å shows a layer width variation of 1.92%. Another example presented here of a reciprocal space map measured near the (004) Bragg peak of an InGaAs/InAlAs/InGaAlAs superlattice having a period of 565.7 Å shows a layer width variation of 2.05%. This is in agreement with a report [71] in the literature of a measured layer width variation of 2.03% for an AlAs/GaAs distributed Bragg reflector having a period of 1575 Å. The *absolute size* of the layer width variation is different for these three structures (2.0 Å, 11.6 Å, and 31.9 Å, respectively), but the *fractional* layer width variation is about 2% for all three growths. This layer width variation is consistent with an effusion cell temperature variation of 1°C during growth.

For QWIPs, composition variations affect the height of the confining potential barrier, and thus affect the thermal leakage current uniformity. Both composition variations and roughness affect the layer width uncertainty, which affects the variations in the responsivity, and sequential resonant and thermionic field assisted tunneling. The lateral size of the roughness affects the uniformity across large arrays of QWIPs.

Chapter 7

Device Processing

Figure 7-1 shows the process flow for fabricating QWIPs.

After MBE growth, a positive photoresist, such as Shipley 1400-27 or S1813, is spun onto the sample at 4500 rpm for 40 seconds, exposed at 300mJ for 10 sec on a Karl Suss MJB contact aligner, and developed for 60 sec with Shipley MF319. The etch used is $H_20{:}H_3PO_4{:}H_2O_2$ in the ratio of 20:1:1. The etch rate is nominally 2400Å/sec in both InGaAlAs and AlGaAs layers. The duration of the etch is chosen so that the bottom contact is exposed. The mesas on the mask set used in this work have optically active areas of $(25\ \mu m)^2$, $(50\ \mu m)^2$, $(100\ \mu m)^2$, and $(200\ \mu m)^2$. The different mesa sizes are useful for determining any area and mesa edge dependences of the leakage and photocurrents.

After the mesa etch, silicon dioxide is deposited on the entire processed wafer. The oxide prevents the bond pads on the top contact from shorting to the bottom contact. The oxide also serves as both an anti-reflection coating and a passivating layer for dangling InGaAlAs or AlGaAs bonds. Dangling bonds are believed to contribute traps and 1/f noise [80]. The oxide can be deposited by sputtering, plasma enhanced chemical vapor deposition (PECVD), or electron beam evaporation.

A second mask is used to create openings, also known as vias, in the oxide for an ohmic contact metal. Again, a positive photoresist is used. Overexposing the sample

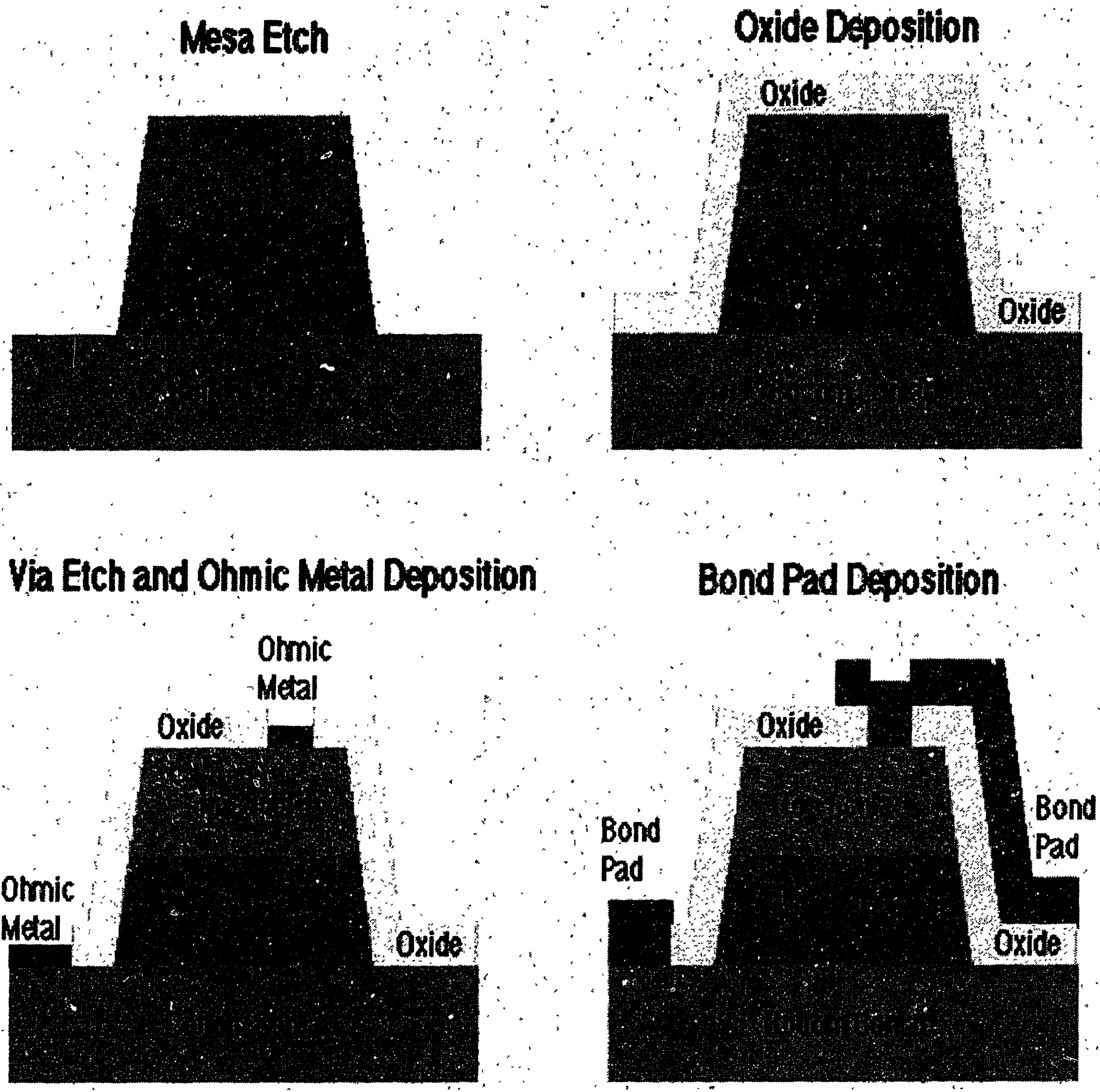


Figure 7-1: QWIP processing steps include a mesa etch, oxide deposition and via etch, ohmic metal and bond pad deposition.

was found to be necessary to develop photoresist out of regions in the bottom contact, which is located 1-3 μm further below the mask than the top contact. The trade-off in doing this is a roughness along the edges of regions that are developed out.

An 100mT oxygen plasma is used for 30sec at 100 watts to remove excess photoresist before deposition of the ohmic metal.

Buffered oxide etch (BOE) (roughly 50 percent hydrofluoric acid and 50 percent ammonium fluoride) is used to etch the oxide at roughly 2000 Å/sec. The duration of the etch is determined by observing when a control wafer appears hydrophobic after the BOE dip. A slight undercut of the oxide under the positive photoresist helps to ease lift-off of the ohmic metal. Lift-off is achieved with ultrasound in acetone.

The ohmic metal for n-type contacts is Ni:Ge:Au:Ni:Au (50 Å:250 Å:500 Å:100 Å:3000 Å). The rapid thermal anneal consisted of a quick ramp at 100°C/min of the sample temperature from room temperature to 250°C, a steady hold of the sample temperature at 250°C for 30 seconds, and a rapid rise in the sample temperature at 100°C/min to the annealing temperature. The Ni:Ge:Au:Ni:Au contact is annealed at the AuGe eutectic melting temperature of 363°C [81] for 30 seconds. (During the anneal, it is believed [82, 83, 84, 85] that the lower Ni layer reacts with the GaAs to form NiGaAs. Gallium is believed to react with the first Au layer to form the β-AuGa phase, and Ge is believed to move into the Ga vacancies, thus forming NiGeAs. The initial Ni layer on the GaAs sample surface is believed to provide a smoother interface between the alloyed contact and the GaAs [82].) The ohmic metal for p-type contacts is Ti:Au (250 Å:3000 Å) or Cr:Au (250 Å:3000 Å) or Ni:AuZn:Ni:Au (250 Å:500Å:300 Å:3000 Å). The Ni:AuZn:Ni:Au contact is annealed at the AuZn eutectic melting temperature of 420°C [81] for 30 seconds. The Ti:Au and Cr:Au contacts are not annealed.

Figure 7-2 shows QWIP pixels on a processed die after the second mask step. The left side of the photograph shows a 4×4 array of (50 μm)^2 detectors. Shown are 16 small square ring contacts which are electically connected to the top contact (on the sample surface) for each of the (50 μm)^2 detectors. The much larger square ring surrounding the entire 4×4 array is the bottom (substrate) contact. On the right

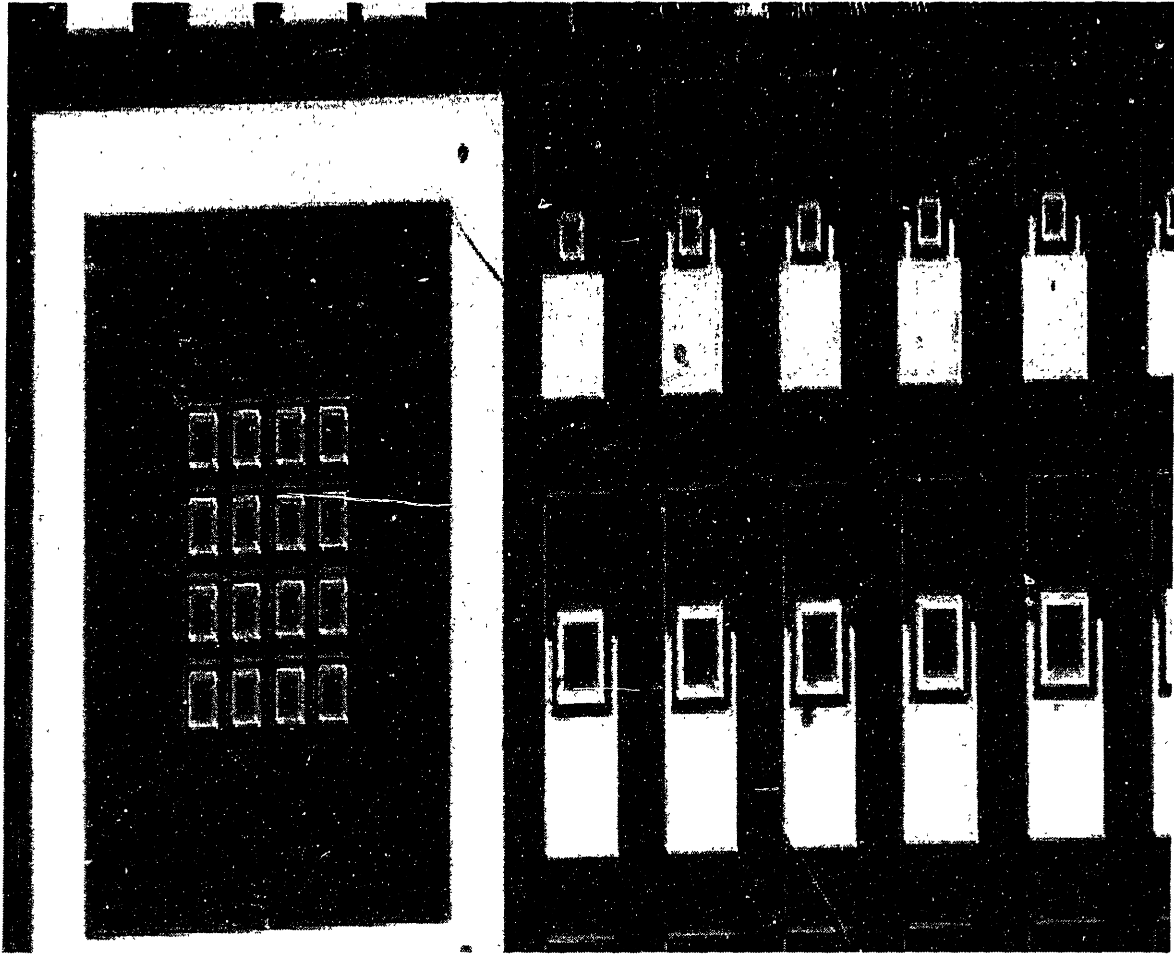


Figure 7-2: QWIP pixels on a processed die after the second mask step.

side of the photograph are shown two rows of individual pixels. The area within the small square ring (top) contacts is $(25\ \mu\text{m})^2$ and $(50\ \mu\text{m})^2$, respectively, for devices along the top and second rows. The bottom contacts are the large rectangular metallic regions right below each ring contact. The mesas defined by the etch in the first mask step are actually much larger than the area within the small ring (top) contacts. The mesas also include a $(100\ \mu\text{m})^2$ area adjacent to the ring contacts. This $(100\ \mu\text{m})^2$ part of the mesa area will seat the bond pads deposited in the third mask step. Unfortunately, this $(100\ \mu\text{m})^2$ area of the mesa under the bond pads contributes no photocurrent, as the bond pads will block the infrared radiation.

Figure 7-3 is a detailed view of one single QWIP pixel on a processed die after the second mask step. The mesa defined by the etch in the first mask step is on the right side of the photograph. The top contact is the square ring contact. The area within the square ring contact is (50 μm)^2. The bottom (substrate) contact is the large metallic contact on the left side of the photograph.

After this processing step, the contacts are probed at room temperature for electrical continuity across the device.

The third mask is used to deposit bond pads on top of the ohmic metal. Three μm thick NR8-3000 Futurex negative photoresist is spun on at 4500 rpm for 40 seconds, exposed for 40 seconds on at 300mJ on a Karl Suss MJB contact aligner, and developed for about 5 minutes in Futurex RD2. The bond pad metal is either Ti:Au or Cr:Au, and is deposited inside the electron beam or the thermal evaporator.

During the cryogenic measurement of the QWIPs following the last processing step (previous paragraph), it was found that many of the devices were short circuits. We believe that Ge- and Zn-rich columns diffuse deep into the devices, thus shorting them. Since the literature reports [82, 86] that the alloyed region is between 1000 and 3000 Å deep, fewer short circuited devices were expected with the use of thicker top contacts (much greater than 3000 Å) and the lowest possible anneal temperatures.

Large numbers of oval defects (greater than 500/cm^2=5/mm^2) or large numbers of group III-rich defects on the sample surface are also associated with devices that are permanently short circuits or that show large leakage currents, even at low temperature.

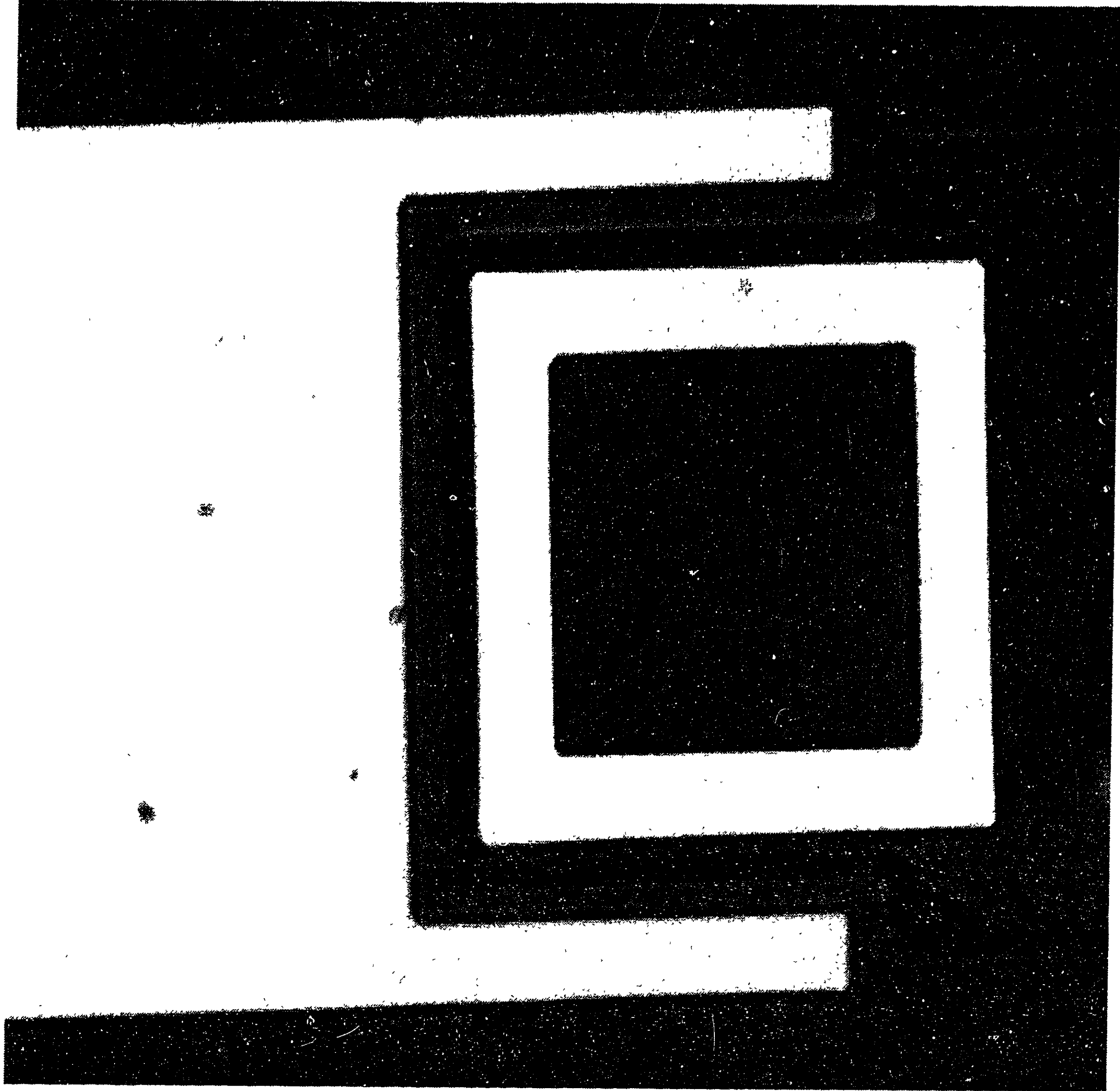


Figure 7-3: Single QWIP pixel on a processed die after the second mask step.

Chapter 8

QWIP Measurement Procedures

This chapter describes general procedures for measuring QWIP device parameters. The example of an n-type QWIP (n-QWIP) will be used throughout the chapter.

Temperature measurements of the leakage give the QWIP thermal leakage activation energy, as well as the temperature dependences for the sequential resonant tunneling and thermal leakage currents. Total photocurrent measurements as a function of the black body target size yield the empirical conversion efficiency, and the product of the quantum efficiency and the photoconductive gain. The wavelength of the QWIP responsivity peak is measured by holding the QWIP operating temperature constant and changing the temperature of the black body target. Spectral measurements demonstrate that a QWIP, with its characteristic narrow spectrum of responsivity, has been achieved. Noise measurements give the value of the photoconductive gain.

The current through the QWIPs is measured as a function of the applied voltage with a HP4145 semiconductor parameter analyzer at cryogenic temperatures. Cryogenic temperatures can be reached with an RMC (at 3450 South Broadmont Drive, Suite 100, Tucson, AZ 85713, 520-903-9366 or 520-889-7900) cold head cooled within a closed cycle refrigerator.

8.1 Leakage Current Measurements

Leakage current measurements are made by covering the QWIP dual-in-line package (DIP) with aluminum foil, and lowering both the DIP package and the aluminum foil to cryogenic temperatures. The QWIP is then exposed only to the heat radiated by the aluminum foil, which is at the same temperature as the QWIP. Figure 8-1 shows the leakage currents as a function of bias voltage for n-QWIP 9066 measured between 10K and 80K in steps of 10 Kelvin degrees. Sample 9066 was comprised of 10 quantum wells of 54 Å $In_{0.08}Ga_{0.92}As$ doped with Si at 5×10^{17} cm^{-3} alternated with 11 barriers of 450 Å $Al_{0.15}Ga_{0.85}As$.

At very low temperatures (below about 50K), the leakage current is dominated by sequential resonant tunneling [7, 87, 88, 89, 90, 91]. The sequential resonant tunneling shows characteristic peaks as a function of the applied bias as the upper subband in one quantum well lines up with the lower subband in the adjacent quantum well in sequence, as shown in Figure 8-2. The peaks and valleys in the (sequential resonant tunneling) leakage current at temperatures below 50K are quite reproducible upon repeated measurements of the same device even over many months. The peaks and valleys in the current-voltage characteristics of the sequential resonant tunneling are stronger and more distinct at lower temperatures and in devices with thinner barriers.

At intermediate temperatures, above 50K, the leakage current is dominated by thermionic leakage. The slope of an Arrhenius plot of the leakage current as a function of the inverse absolute QWIP operating temperature yields an activation energy or equivalently the activation wavelength. Two examples of this are shown in Figure 8-3 and Figure 8-4 for Sample 9066 at negative and positive applied bias voltages, respectively. The points indicate measured data and the straight lines are best fit Arrhenius dependences. The leakage currents in Figs. 8-3 and 8-4 both show activation energies which decrease with the magnitude of the applied bias. Figure 8-5 shows this voltage dependence of the leakage current activation energy for QWIP 9066.

The leakage activation energy is a measure of the position of the Fermi level in the quantum wells. The thermionic leakage current in a multiple quantum well structure

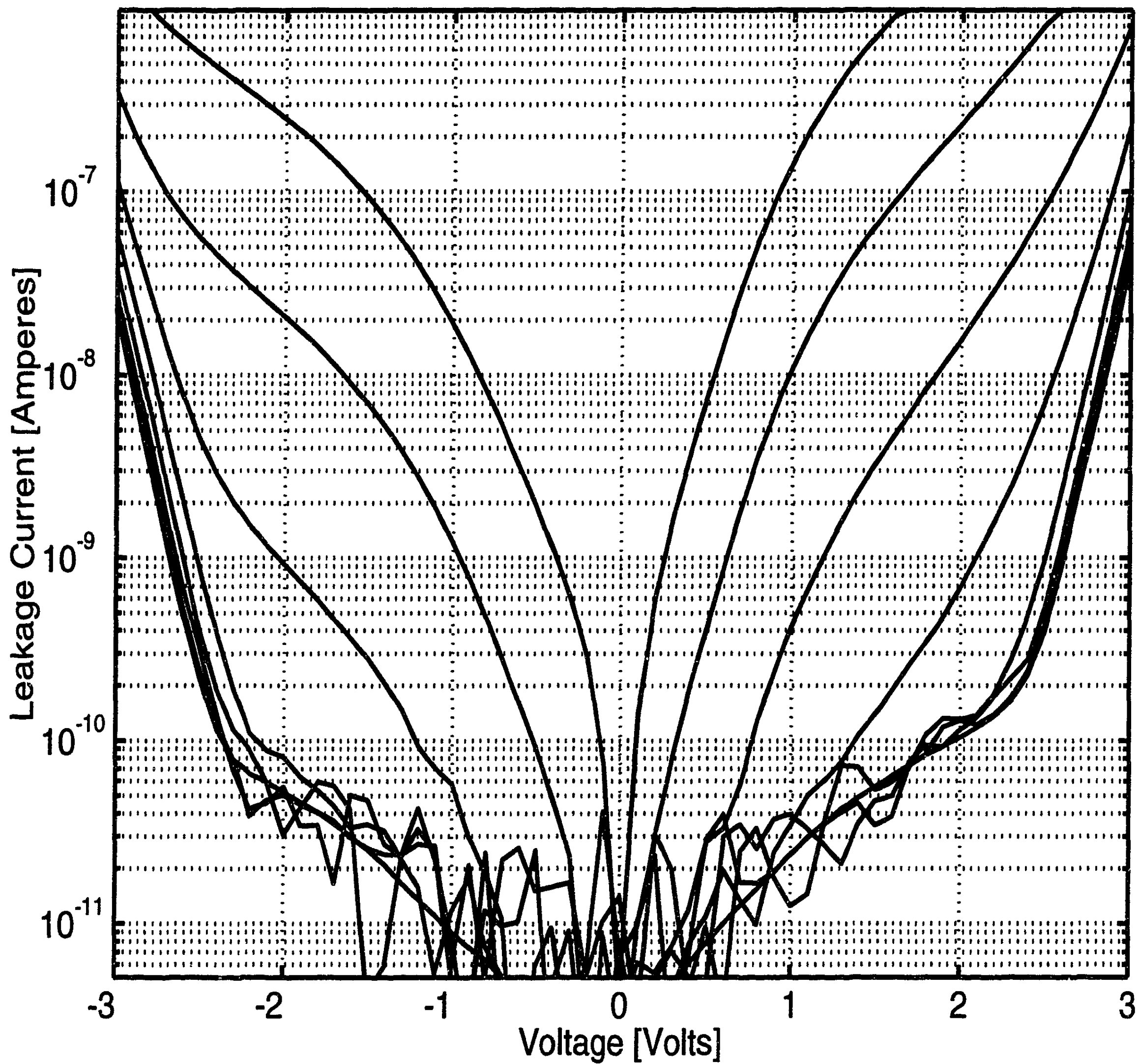


Figure 8-1: The curves in the figure are the leakage currents as a function of bias voltage for QWIP 9066 measured between 10K and 80K in steps of 10 Kelvin degrees. At low temperatures (less than 50K), the leakage is dominated by sequential resonant tunneling, and at intermediate temperatures, the leakage is dominated by thermionic or thermionic field assisted tunneling.

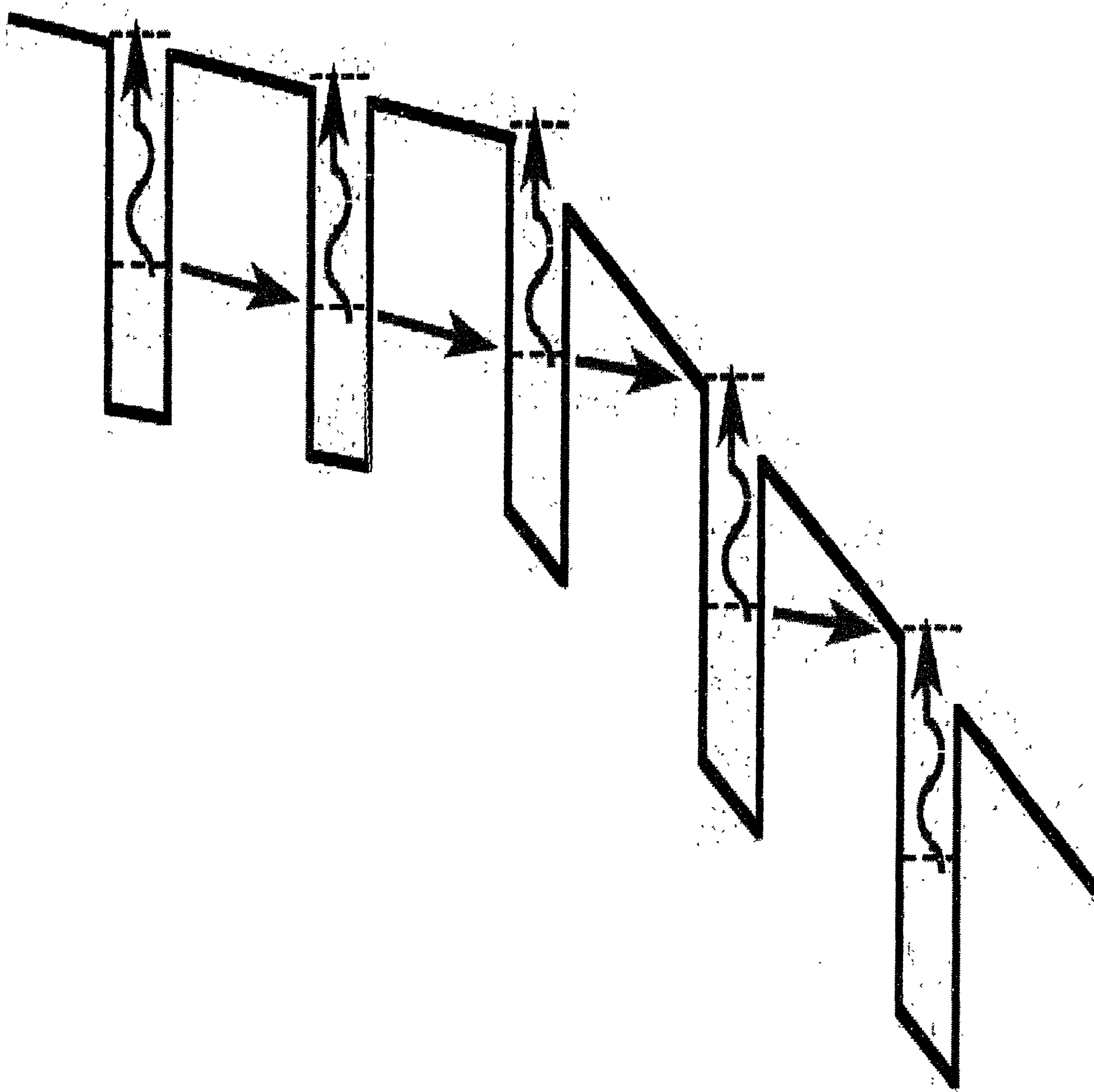


Figure 8-2: When the upper subband in one quantum well lines up with the lower subband in an adjacent quantum well, as shown in this figure, the sequential resonant tunneling shows a characteristic peak.

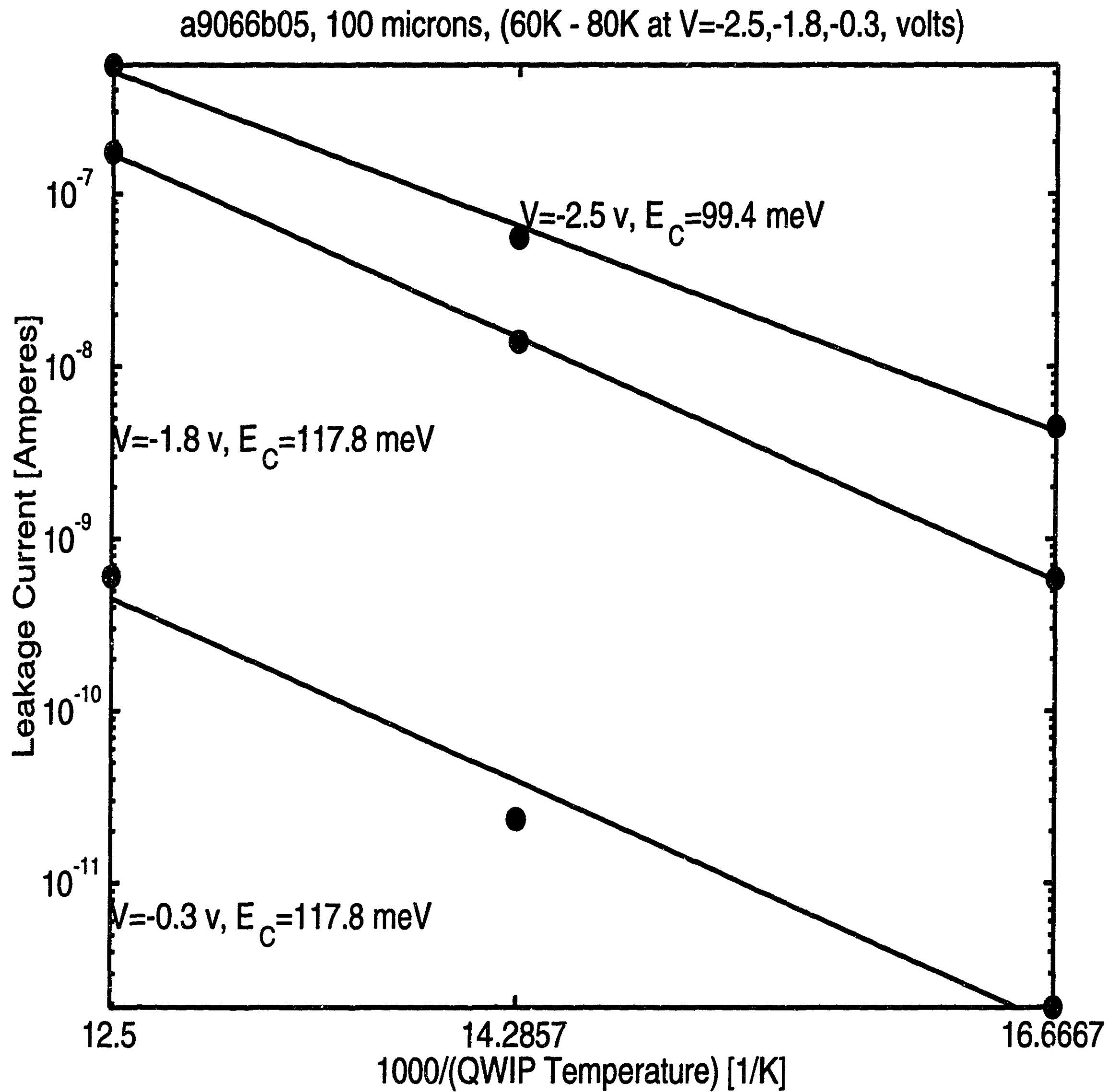


Figure 8-3: Arrhenius plots of the leakage current as a function of inverse QWIP operating temperature for Sample 9066 at negative bias voltages. The points indicate measured data and the straight lines are best fit Arrhenius dependences. The slope of the best fit curves yields the thermal leakage activation energies.

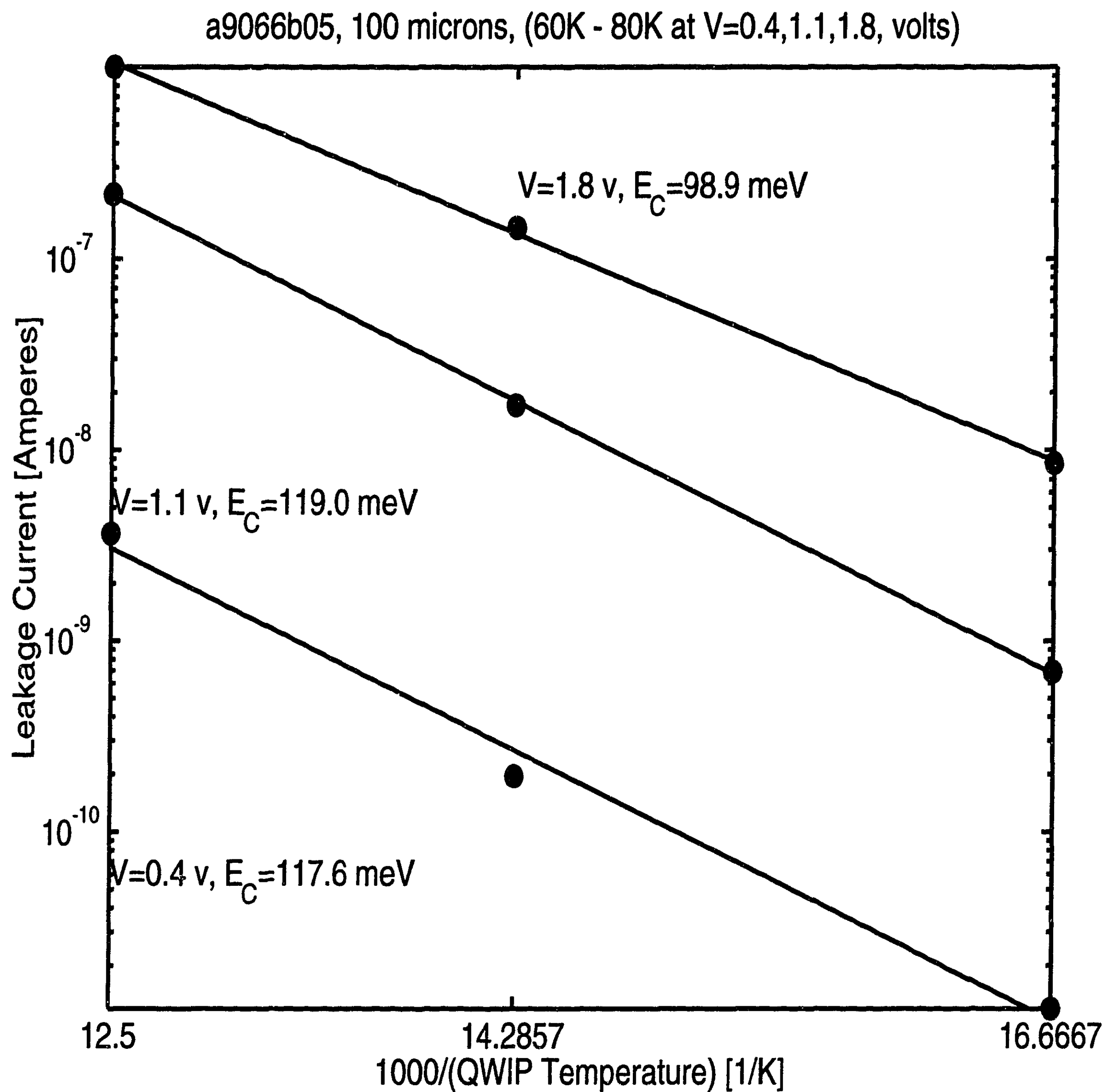


Figure 8-4: Arrhenius plots of the leakage current as a function of inverse QWIP operating temperature for Sample 9066 at positive bias voltages. The points indicate measured data and the straight lines are best fit Arrhenius dependences. The slope of the best fit curves yields the thermal leakage activation energies.

is given by [92, 93, 94]

$$J = qN_C v_d \exp\left(-\frac{\Delta E_C - E_F}{k_B T}\right)\left[\exp\left(\frac{qFL_W}{k_B T}\right) - 1\right], \tag{8.1}$$

$$= qN_C v_d \exp\left(-\frac{q\Phi_B}{k_B T}\right)\left[1 - \exp\left(\frac{-qV}{k_B T}\right)\right], \tag{8.2}$$

where q is the electron charge, N_C is the density of states in the barrier, v_d is the drift velocity, ΔE_C is the band offset between the quantum well and barrier conduction band edges, T is the QWIP operating temperature, F is the average field in the structure, L_W is the well width, and Φ_B is the activation energy measured at large voltages ($V > k_B T$),

$$\Phi_B = (\Delta E_C - E_F) - qVL_W/L_{TOT}, \tag{8.3}$$

where L_{TOT} is the length of the entire QWIP structure.

Fig. 8-5 shows that at small bias voltages, the measured activation energy varies from very small (<20 meV) to very large (>140 meV) values. In this range, it is probably unreliable primarily because the leakage current at small bias voltages may be below the noise floor of the HP 4145 Semiconductor Parameter Analyzer.

Fig. 8-5 shows that at intermediate values of the negative bias voltage (between -0.6 v and -2.2 v), the measured activation energy is best fit by the relation

$$\Phi_B = 131.2 \text{ meV} + (7.60 \text{ meV/V}) \times V. \tag{8.4}$$

A comparison of Equation (8.4) with Equation (8.3), we can infer a measured value of $\Delta E_C - E_F$=131.2 meV. This measured value is indeed close to the calculated value. For Sample 9066, the doping level yields an E_F=10.2 meV above the bound state subband edge E_1, and we calculate $\Delta E_C - E_F$=135.6 meV.

The slope of Equation (8.4) indicates a measured value of qL_W/L_{TOT}=7.60 meV/V, upon comparison with Equation (8.3). This also is close to the calculated: For Sample 9066, L_W=54 Å and L_B=480 Å, and we calculate qL_W/L_{TOT}=9.28 meV/V. This 20 % discrepancy between the measured and calculated voltage dependence of the activation energy may result in part from uncertainties in the growth parameters.

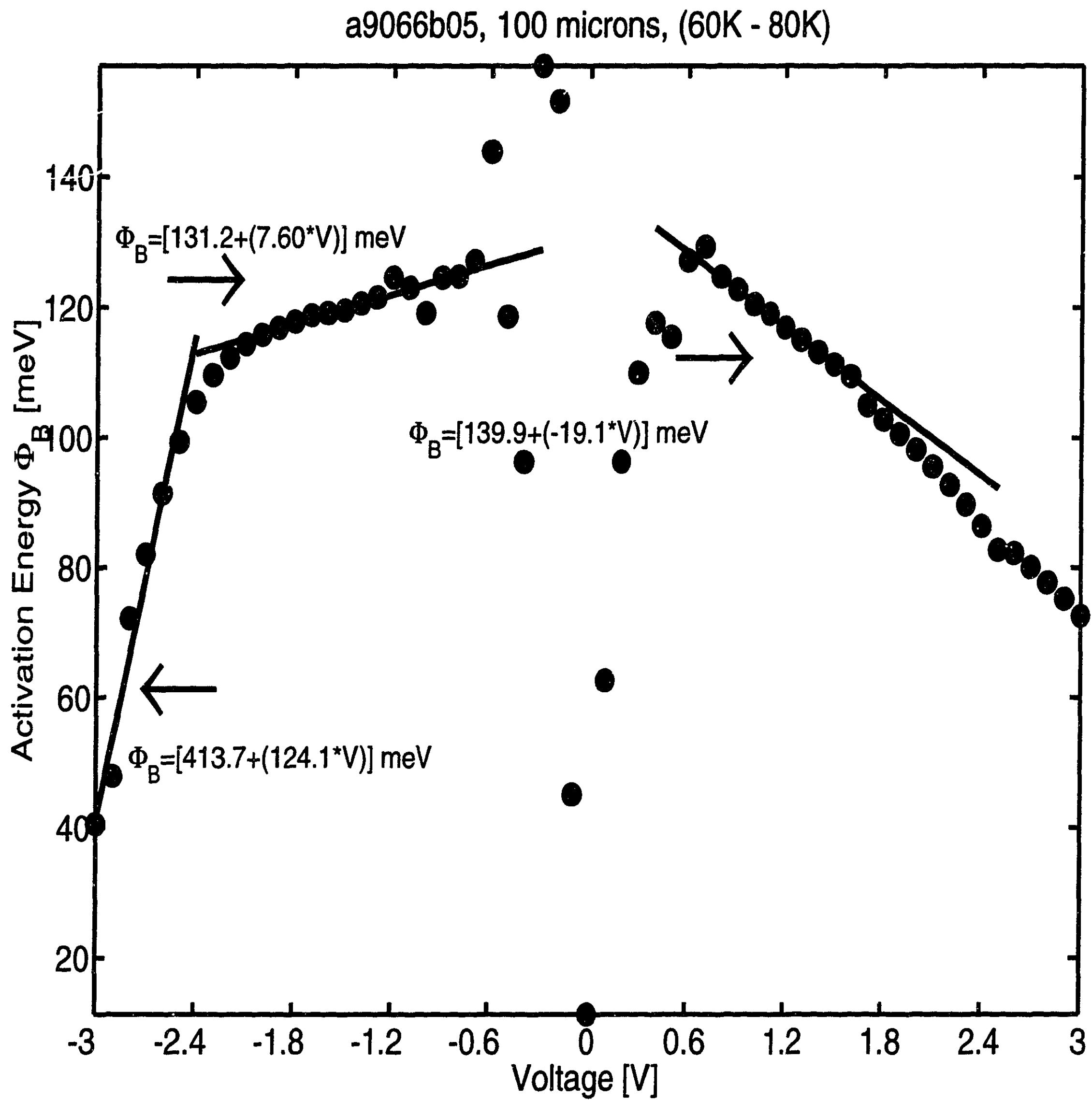


Figure 8-5: The voltage dependence of the leakage current activation energy for QWIP 9066. The activation energy was obtained by fitting the measured leakage current to the thermionic form in Equation (8.2) for temperatures between 60K and 80K.

However, models to be presented in Chapter 10 reveal how the voltage dependence of this activation energy Φ_B may be related to the injection mechanisms in the QWIP.

This voltage dependence of the activation energy Φ_B in Equation (8.3) has not been studied in the QWIP literature. Pelve et al. [95] has made a plot very similar to Fig. 8-5. However, he interprets the decrease of the activation energy Φ_B with finite bias in terms of Fowler-Nordheim tunneling, and not thermionic leakage, Equation (8.3) above. The slope of the measured activation energy with respect to small negative biases can be measured from Fig. 3 of Pelve's work to be 5 meV/V. This is to be compared with a calculated value of qL_W/L_{TOT}=2.31 meV/V which would be predicted from the growth structure given in the paper (50 periods of wells of width L_W=40 Å and barriers of width L_B=300 Å). Thus, the measured leakage current shown in Pelve's paper is larger than would be expected from thermionic leakage, Equation (8.3), alone. This is probably a result of their use of relatively narrow barriers of width L_B=300 Å).

The quality of fit of the measured leakage to a thermionic leakage model indicates the presence or absence of excessive leakage. Williams [96] has measured leakage currents that are much larger than expected for barrier widths greater than 300 Å even at very low cryogenic temperatures, which he attributes to defect assisted tunneling. Levine [7] reports a very good match to Equation 8.2 even for barriers of width 500 Å at (small) fields of 5.81×10^3 V/cm. The good fit of the measured leakage currents in Figs. 8-1 and 8-5 at small and intermediate negative bias to the thermionic leakage model of Equations (8.2) and (8.3) is an indication of the absence of excessive leakage, such as that which might result from defect assisted tunneling right through our (large) 500 Å barriers.

Excessive leakage currents are observed in Fig. 8-1 at positive voltage biases. It is clear from Fig. 8-1 that the leakage current rises more quickly with voltage at positive biases than at negative biases. This is indicated in Fig. 8-5 by a sharper drop in the activation energy Φ_B with increasing positive bias than with increasing negative bias. This asymmetry in the leakage current-voltage characteristics has been observed many times in the literature [7] and is believed to result from asymmetries in the MBE growth conditions: the dopant diffuses towards the surface during growth,

and the quality of the AlGaAs on GaAs interface differs from the GaAs on AlGaAs interface.

At very large voltage magnitudes (greater than 2.7 volts in Fig. 8-1), the leakage current is dominated by Fowler-Nordheim [92] tunneling. This is indicated in Fig. 8-5 by a very sharp drop in the activation energy Φ_B at very large bias magnitudes.

8.2 Photocurrent Measurements

DC photocurrent measurements are made with the HP4145 by exposing the QWIP through a ZnSe window on the room temperature shield of the cryostat to a calibrated black body, as shown in Figure 8-6. At the typical black body operating temperatures of between 300K and 600K, the black body spectra has a peak at wavelengths between 6 and 10 μm. (The peak black body wavelength in microns occurs at 2897.8/T_{BB}, where T_{BB} is the black body temperature in Kelvin degrees.) Use of higher black body temperatures tends to warm up the cryostat cold head, so that the measured QWIP DC current is a combination of both the photocurrent and the larger thermal leakage current resulting from the warmer QWIP operating (cold head) temperature.

QWIP DC photocurrent measurements are not easy to make because of the presence of a large background DC current which is comparable to the photocurrent resulting from the blackbody radiation source. This large background DC current has two main sources: the room temperature heat generated by a ZnSe viewport right in front of the QWIP, and the relatively warm (nominally 77K but actually 220K) radiation shield inside the cryostat.

The small optical f/# of the cryostat when compared with the small solid angle subtended by the blackbody orifice means that much of the room temperature background heat generated by the ZnSe viewport in front of the QWIP is also contributing to the QWIP DC photocurrent. The solution to this problem is to increase the cryostat f/# to correspond to the f/# of the solid angle subtended by the blackbody target. This is accomplished by decreasing the size of the aperture in the cryostat

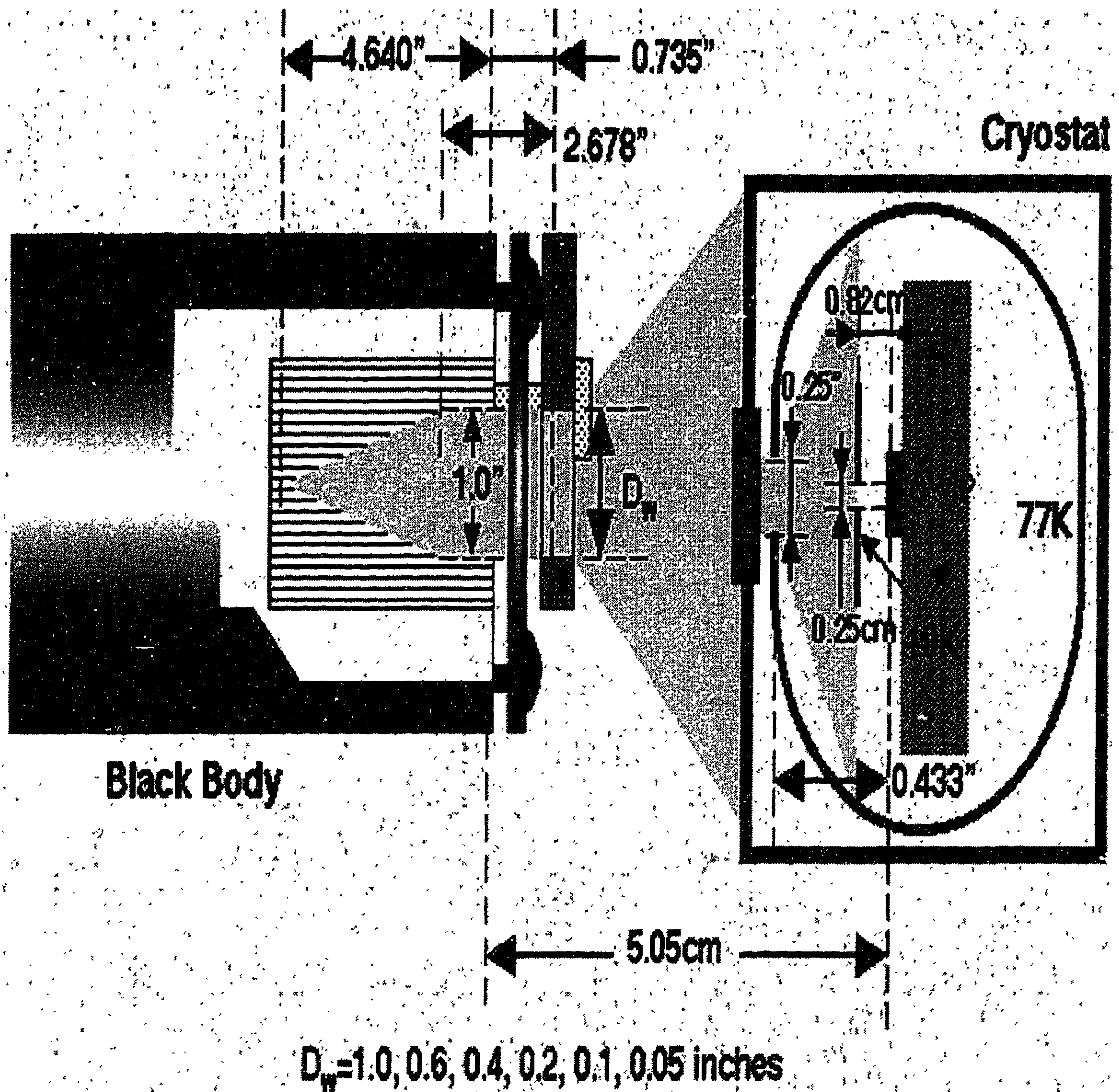


Figure 8-6: DC photocurrent measurements of a QWIP are made at cryogenic temperatures with a calibrated black body infrared source. Precise distances are given for aperture sizes and spacings between objects so that a precise incident photon flux can be calculated.

77K radiation shield which exposes the QWIP to the ZnSe window.

The radiation shield inside the cryostat, which has a nominal temperature of 77K, was found to have an actual temperature of 220K. Since a 77K target radiates very little heat, about 8 orders of magnitude less than a 300K target, it was expected that when the QWIP was exposed to only the cryostat radiation shield, the QWIP DC current should be the same as the sequential resonant tunneling current seen at 10K. Instead, it was found that the cryostat radiation shield generated enough heat to increase the QWIP DC current to about two orders of magnitude higher than the sequential resonant tunneling current. To reduce this QWIP DC current resulting from the warm 220K cryostat radiation shield, a copper block was placed in thermal contact with the cryostat cold head right in front of the QWIP. This copper block was chosen to have a large enough thickness and a small enough aperture size to be commensurate with the f/# associated with the solid angle subtended by the black body target. In this way, the QWIP is exposed to as little of the cryostat radiation shield as possible, while still seeing the entire black body target.

Figure 8-7 shows the current-voltage characteristics for the leakage current (solid lines) measured in 10K increments between 10K and 80K, and the photocurrent (dashed lines) measured at 40K. The QWIP has an 180° field of view of a 220K background during the measurement of the photocurrent. Both the magnitude and shape of the photocurrent i-v characteristic is observed to remain virtually unchanged between 10K and 50K: at a bias of -1.0 volt, the photocurrent increases by 50% between 10K and 50K, and at a bias of -2.0 volts, the photocurrent changes by less than 10% between 10K and 50K. Under these measurement conditions, the QWIP is observed to have a Background LImited Performance (BLIP) temperature of 62K, at which the leakage current is equal to the photocurrent. Measurement of the QWIP responsivity is usually done at bias voltages (between -2.0 and -2.6 volts in Fig. 8-7) where the photocurrents are no longer sharply rising. At these bias voltages, the photocurrent and responsivities are large, but the Fowler-Nordheim tunneling leakage is not yet dominant.

In order to measure the wavelength of the QWIP responsivity peak, DC photocurrent measurements are made at many different black body target temperatures for a fixed

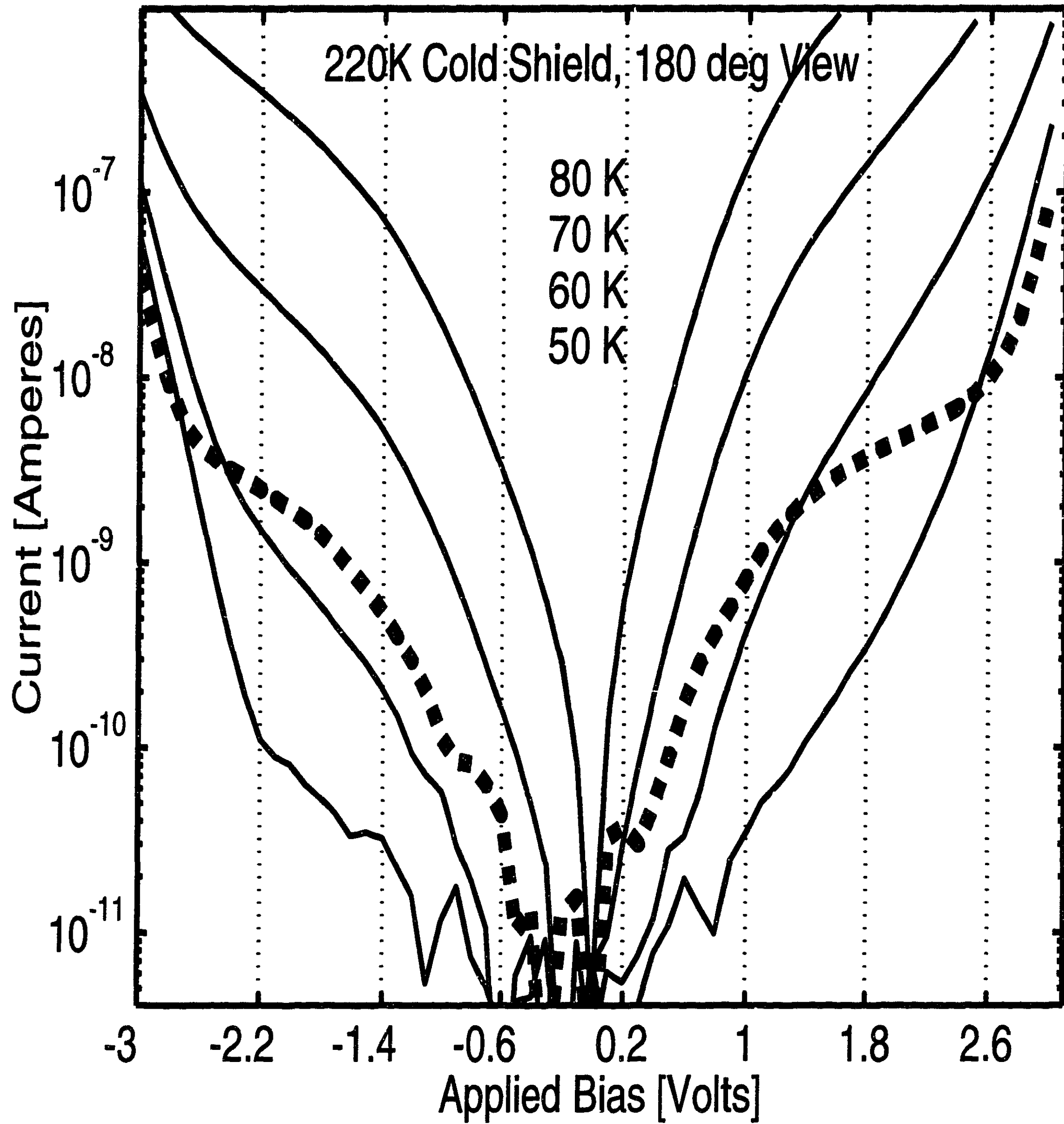


Figure 8-7: The current-voltage characteristics for the leakage current (solid lines) measured in 10K increments between 10K and 80K, and the photocurrent (dashed lines) measured at 40K. The QWIP has an 180° field of view of a 220K background during the measurement of the photocurrent.

QWIP operating temperature and a fixed black body target size. An Arrhenius plot of the measured photocurrent as a function of the inverse absolute black body temperature yields the wavelength of the responsivity peak. An example of this is shown in Fig. 8-8. The slope of the curve in Fig. 8-8 predicts Sample 9066 to have a responsivity peak at a wavelength of 8.7 μm. The points in Fig. 8-8 which were measured at lower black body temperatures do not lie on the best fit curve because the infrared radiation generated by the room temperature ZnSe viewport and the 220K radiation shield (which occupies a very large solid angle) is much larger than the radiation generated by the small black body target at these temperatures.

The conversion efficiency can be obtained from a measurement of the net photocurrent, equal to the measured photocurrent minus the cold shield current, as a function of the black body target size for a fixed black body temperature of 500K. The geometry shown in Fig. 8-6 was used to obtain measured conversion efficiencies for Sample 9066 at 40K as summarized in Table 8.1. The straight line in Figure 8-9 is a best fit to the measured data points of Table 8.1, which for this sample yields a best fit conversion efficiency of 4%. This size conversion efficiency corresponds to a responsivity of 270 mA/W or, equivalently, a detectivity of 4×10^{10}cm-$\sqrt{Hz}$/Watt, if we assume a capture probability of 7% as is commonly measured in the literature [97].

This detectivity of 4×10^{10}cm-$\sqrt{Hz}$/Watt is significant as it was obtained on an n-QWIP without the use of an optical grating. Of particular importance is the fact that this detectivity is also large [7] enough for focal plane array performance to be limited by the uniformity of processing rather than the size of the single pixel detectivity. (See Chapter 1.) Moreover, n-QWIPs fabricated without an optical grating involve fewer processing steps, and could lead to higher yield, lower cost focal plane arrays. The observation of a significant detectivity for the absorption of normally incident radiation by an n-QWIP fabricated without an optical grating was reported concurrently by our group [98] and two other groups [99, 100, 101].

The measurement of the QWIP conversion efficiency or detectivity requires the use of a black body because the latter has a precisely known photon distribution. The

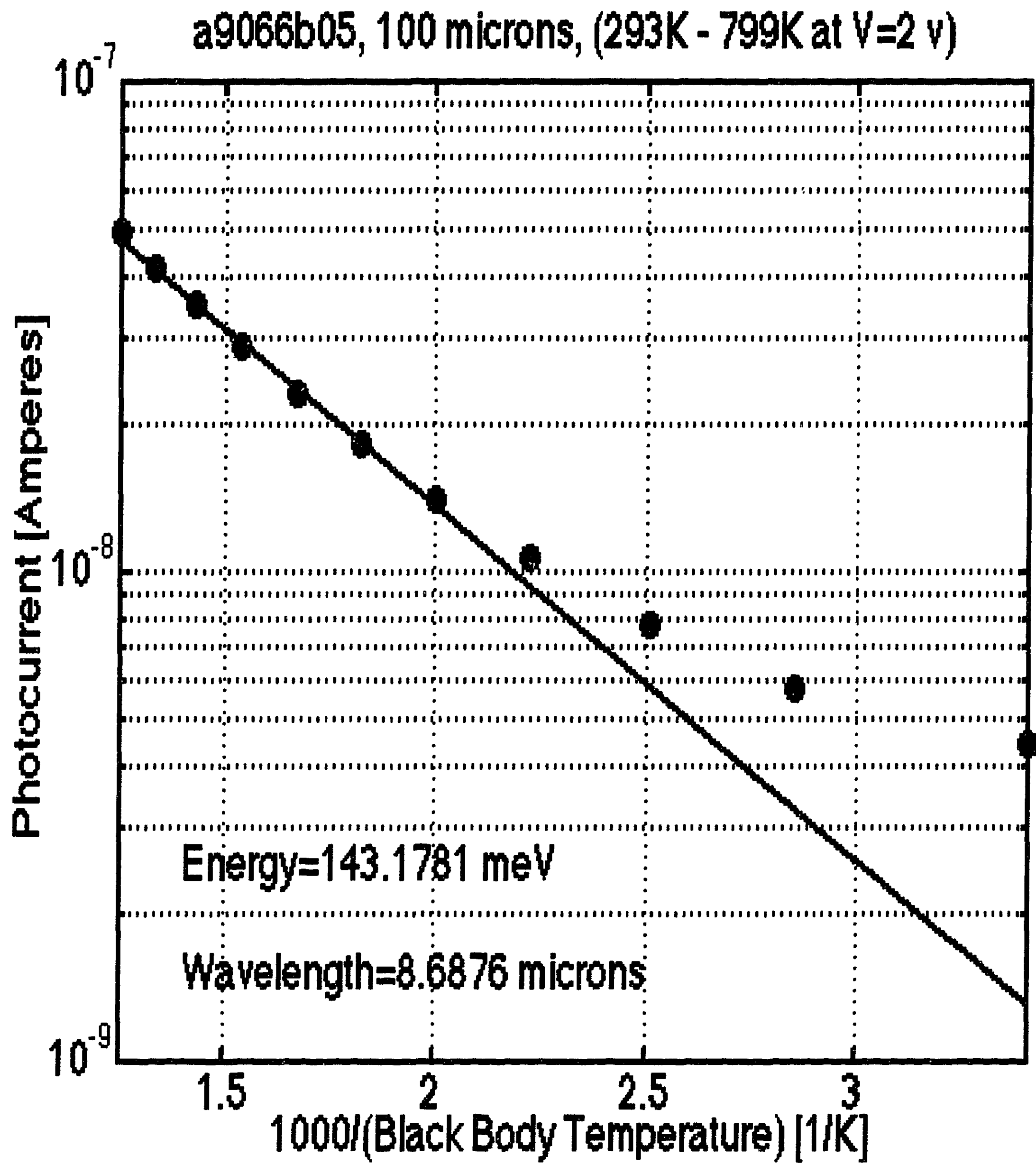


Figure 8-8: Arrhenius plot of the photocurrent as a function of inverse black body temperature for Sample 9066 at 40K. The slope of this curve predicts Sample 9066 to have a responsivity peak at a wavelength of 8.7 μm.

Wheel Diameter (cm)	Effective BB Diameter (cm)	Net Photo-current (Amps)	Measured Conversion Efficiency (percent)
0.1270	0.2015	1.3090e-10	3.3008
0.2540	0.4030	7.3850e-10	4.6597
0.5080	0.8059	2.8629e-09	4.5321
1.0160	1.5396	8.1721e-09	3.5894
1.5240	1.5396	9.1921e-09	4.0374
2.5400	1.5396	9.6371e-09	4.2329

Table 8.1: Measured conversion efficiencies for Sample 9066 at 40K as a function of different black body aperture sizes. The net photocurrent, equal to the measured photocurrent minus the cold shield current, measured as a function of the black body target size can by used together with the expressions for the measured photocurrent, Equation (8.8), and for the QWIP responsivity, Equation (3.2), to allow accurate measurements of the QWIP conversion efficiencies.

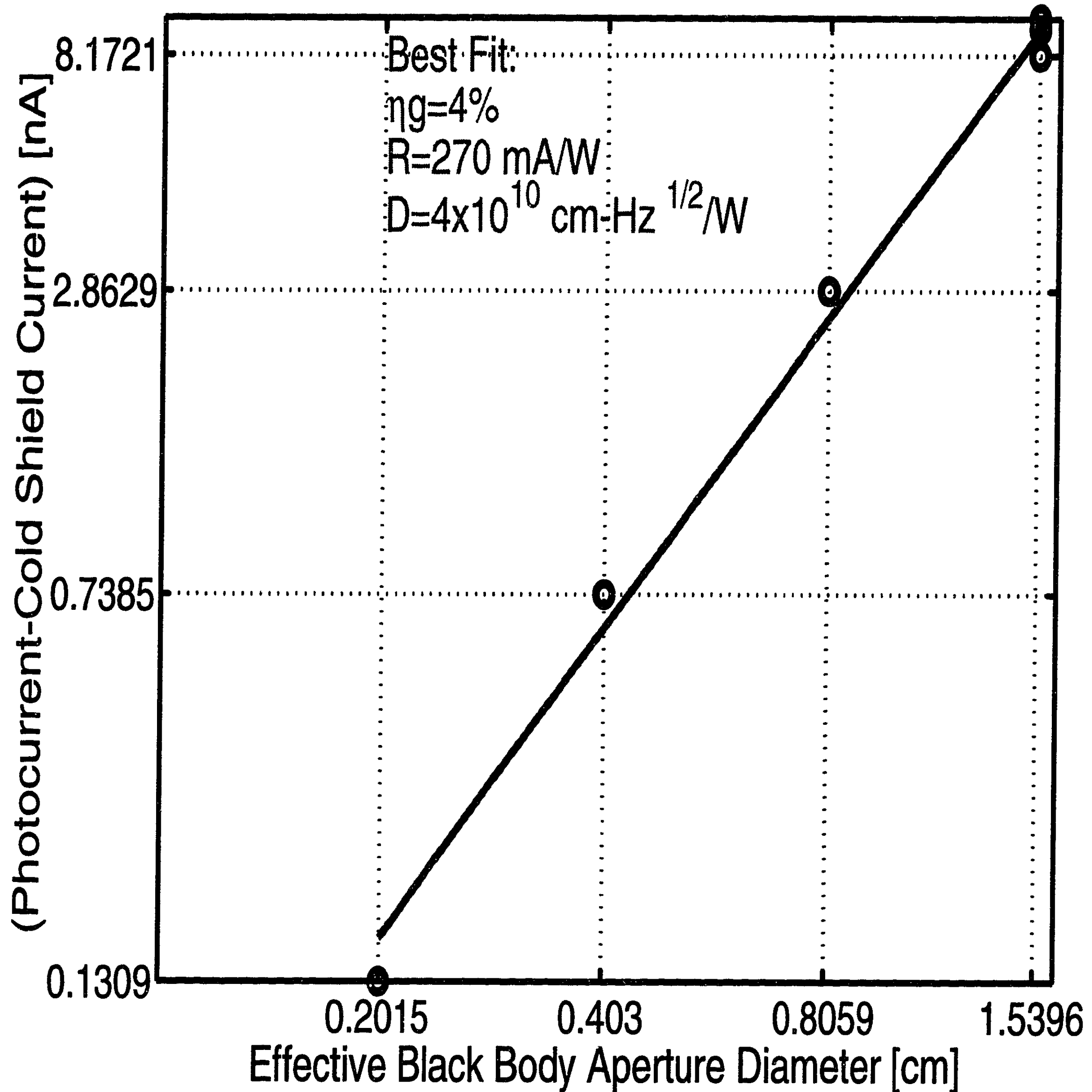


Figure 8-9: The points in the figure represent the net photocurrent, equal to the measured photocurrent minus the cold shield current, as a function of the black body target size for a fixed black body temperature of 500K. The solid line represents a best fit of the measured points to a straight line, and yields a best fit conversion efficiency of 4%. This corresponds to a measured detectivity of 4×10^{10}cm-$\sqrt{Hz}$/Watt for Sample 9066 at 40K.

black body power spectral density W(λ) is given by

$$W(\lambda) = \left(\frac{2hc^2}{\lambda^5}\right)\left(\frac{1}{\exp(hc/\lambda k_B T_{BB}) - 1}\right) \tag{8.5}$$

in units of Watt/cm^2-ster-μm. W(λ) is the power radiated into a unit solid angle and into a unit wavelength interval centered at λ and into both polarizations for a black body of unit area and temperature T_{BB}. The black body power P(λ) per unit wavelength incident [7] upon the detector is

$$P(\lambda) = W(\lambda)\Omega A_{DET} C_F \cos\Theta, \tag{8.6}$$

where A_{DET} is the area of the detector, where the coupling factors are all lumped into C_F=T_f(1-r), where T_f=0.7 is the transmission of the ZnSe viewport in front of the QWIP, where r=0.28 is the reflectivity of the GaAs surface, where $\Theta = 0°$ is the angle of incidence for normally incident radiation, and where

$$\Omega = \pi \sin^2(\omega/2) = \pi/[4(f/\sharp)^2 + 1] \tag{8.7}$$

is the solid angle subtended by a field of view of ω as seen by the QWIP. (If the black body has an aperture of radius R_{BB} and is located a distance D from the QWIP, then $\tan(\omega/2) = R_{BB}/D$ and 2f/#=D/R_{BB}.) The measured photocurrent is then

$$i_P = \int d\lambda P(\lambda) R(\lambda) \tag{8.8}$$

where the integration is over wavelengths where the QWIP responsivity $R(\lambda)$ is sizeable. The expressions for the measured photocurrent, Equation (8.8), and for the QWIP responsivity, Equation (3.2), along with the measurement geometry shown in Fig. 8-6 allow accurate measurements of the QWIP conversion efficiencies.

To prevent heating of the cryostat cold head by the 500K black body used in the measurement of Fig. 8-9, the black body aperture is closed as soon as the measurement of each point in Fig. 8-9 is accomplished. The duration of the measurement of each point in Fig. 8-9 is a few seconds. The possibility of heating of the cold head is further minimized by making the measurement of the points in Fig. 8-9 in order of *increasing* black body aperture size. The largest black body aperture sizes, which flood the cold head with the largest amount of heat, are not used until the end of the measurement

session so that any possible heating of the cold head does not complicate the earlier measurements which involve smaller photocurrents, smaller black body aperture sizes, and smaller possibility of warming up of the cold head.

8.3 Spectral Measurements

Photocurrent spectroscopy measurements for QWIPs are obtained by using a calibrated black body as the radiation source. Both the QWIP operating temperature and the black body temperature are held fixed as a filtered portion of the black body spectrum is incident upon the QWIP. The function of a spectrometer is obtained with a circular variable filter, which passes a calibrated wavelength. The f/# of the cryostat is used to determine the spatial extent of that part of the black body radiation which is incident on the circular variable filter and which also arrives at the QWIP. This spatial extent of the radiation on the circular variable filter together with the measurement geometry determines the actual wavelength distribution arriving at the QWIP.

Figure 8-10 shows the measured responsivity as a function of wavelength thus obtained for a 100 μm pixel on Sample 9066 at a QWIP operating temperature of 40K, a black body temperature of 1323.2K, and a QWIP bias of -2v. The responsivity of the device shows a very narrow FWHM of 1 μm, as is expected [7] for a QWIP with an upper state right at the barrier band edge. The peak in the responsivity spectrum of Fig. 8-10 is at 8.7 μm, as would be expected from the Arrhenius dependence measured in Fig. 8-8.

8.4 Spatially Resolved Responsivity Measurements

Dr. Sam Wang at Lockheed Sanders has measured the response of our n-QWIP 9066 to a spatially resolved radiation source. The spatially resolved responsivity measurement sheds light on the nature of the absorption of normally incident radiation by an

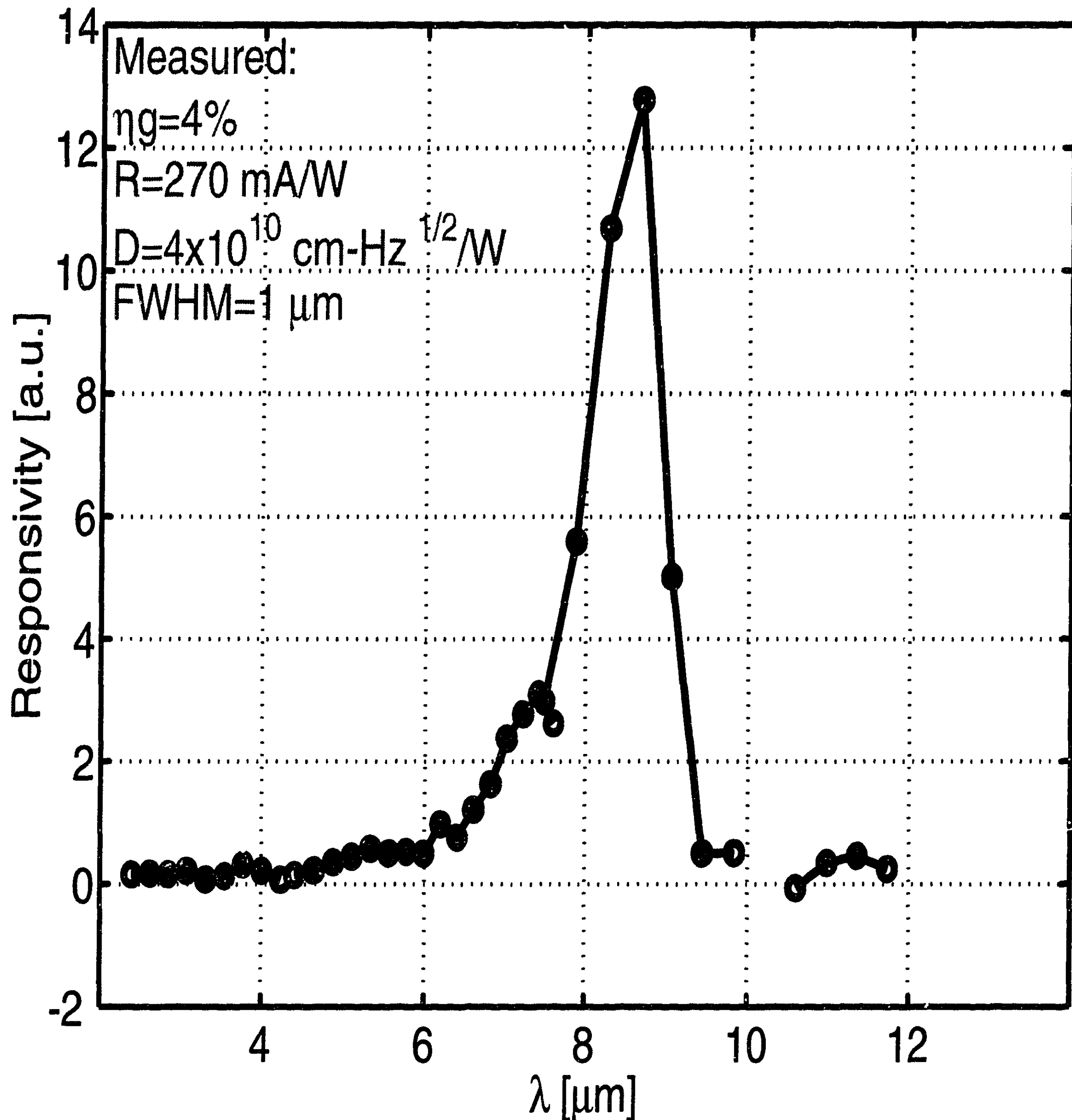


Figure 8-10: The measured responsivity as a function of wavelength obtained for a 100 μm pixel on Sample 9066 at a QWIP operating temperature of 40K, a black body temperature of 1323.2K, and a QWIP bias of -2v. The responsivity of the device shows a very narrow FWHM of 1 μm, as is expected [7] for a QWIP with an upper state right at the barrier band edge. The peak in the responsivity spectrum of Fig. 8-10 is at 8.7 μm, as would be expected from the Arrhenius dependence measured in Fig. 8-8.

n-QWIP which has been processed without an optical grating. The results are shown in Figures 8-11, 8-12, 8-13, and 8-14. Several observations are in order. First, TE polarized radiation having a small spot size normally incident upon a sample in the middle of a large optically active area yields a finite response. This response is three times as large as the measurement background, which we take to be the response resulting from radiation incident on the sample between QWIP pixels. Second, radiation incident upon the pixels at the pixel boundaries, which are (111) planes, yields a large response, which is presumed to be dominated by the response due to TM polarized radiation. The measured photocurrent is thus a response due to both TE and TM polarizations. Finally, there is a small amount of responsivity crosstalk between adjacent QWIP pixels.

Figure 8-11 shows the response at a QWIP bias of -2v, $\lambda = 8.4$ μm, T=80K for a (25 μm)^2 optically active area on Sample 9066 to a spatially resolved radiation source of spot size of 20 μm which is scanned in increments of 1.3 μm. This measurement was accomplished with f/1 reflective optics through a fiber. The relative size of the photoresponse as a function of distance along the x=55 μm cross section of Fig. 8-11 is shown in Fig. 8-12.

Figs. 8-11 and 8-12 show that the responsivity is largest when the radiation is focused along the boundaries of the (25 μm)^2 optically active area. (The mesa and the optically active area for this pixel size is shown in the upper right side of Fig. 7-2.) At the boundaries of the optically active area, the radiation is no longer normally incident, but rather is refracted into the pixel by the (111) planes which are created by the etch at the mesa boundaries. This is significant because refraction of the radiation at the mesa boundaries produces both TE and TM polarized radiation propagating through the QWIP pixel. The photoresponse thus could result from either polarization of radiation, and it is not possible in this measurement to quantitatively isolate the photoresponse resulting from one particular polarization. However, the photoresponse resulting from radiation incident on the pixel boundaries is probably dominated by the absorption of TM polarized radiation as the latter is much stronger than the absorption of TE polarized radiation.

Figure 8-13 shows the response at a QWIP bias of -2v, $\lambda = 8.4$ μm, T=80K for

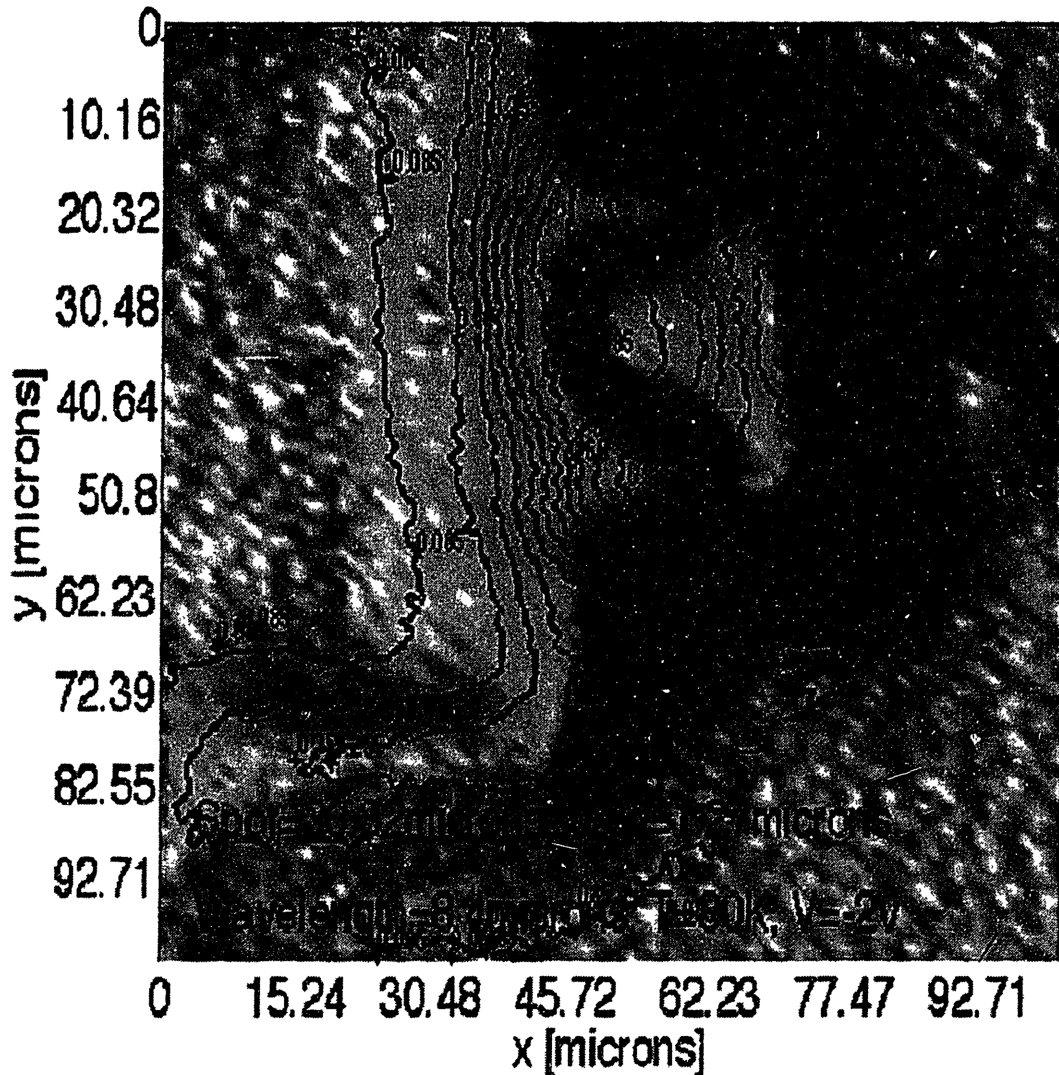


Figure 8-11: False contour map of the spatially resolved responsivity at a bias of -2v, $\lambda = 8.4\ \mu$m and T=80K, for a $(25\ \mu\text{m})^2$ optically active area on Sample 9066. This measurement was accomplished with f/1 reflective optics through a fiber to obtain a spot size of 20 μm which is scanned in increments of 1.3 μm. (The mesa and the optically active area for this pixel size are shown on the upper right side of Fig. 7-2. This measurement was performed by Dr. Sam Wang at Lockheed Sanders.)

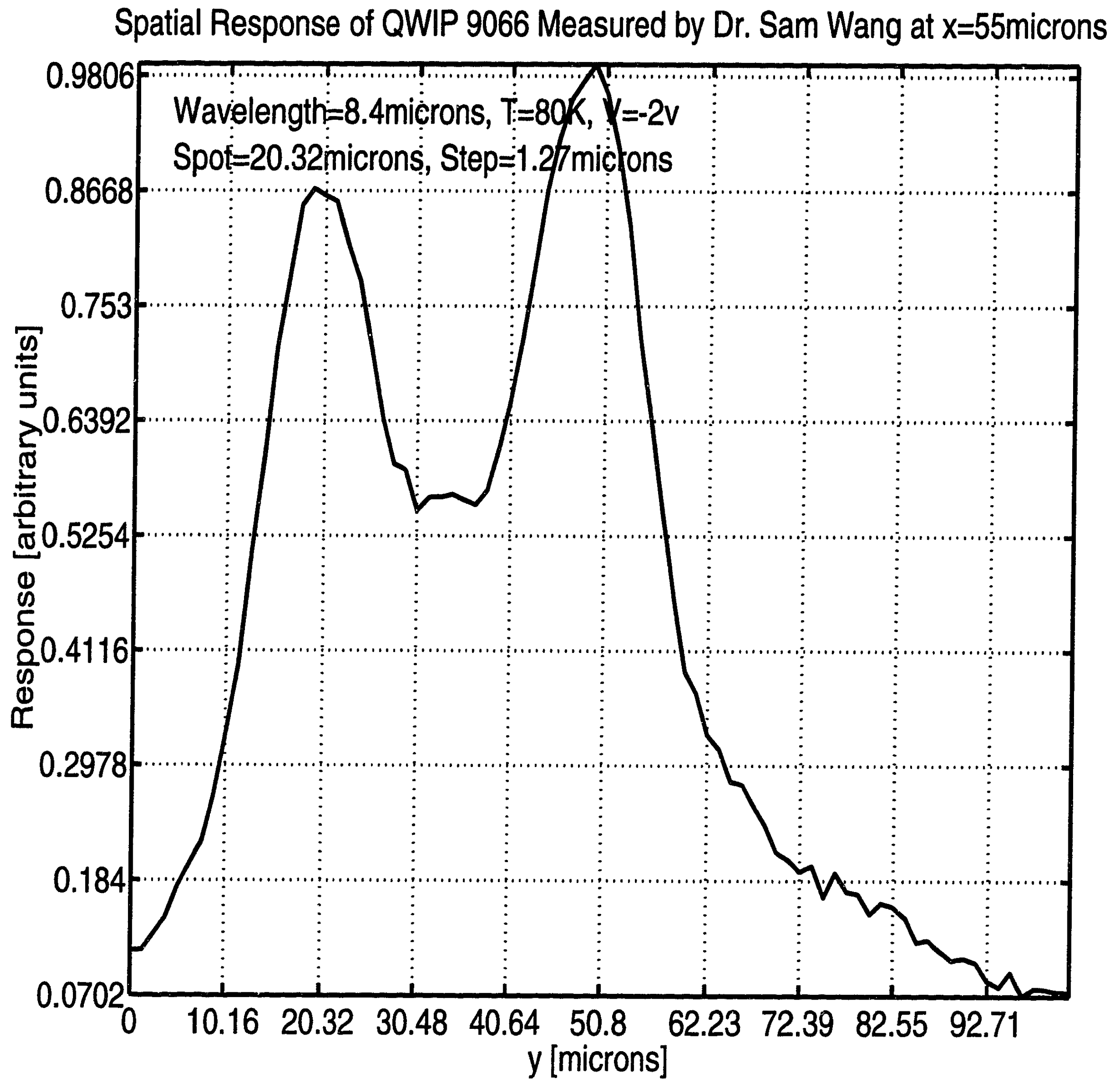


Figure 8-12: Spatially resolved responsivity as a function of distance along the x=55 μm cross section of Fig. 8-11.

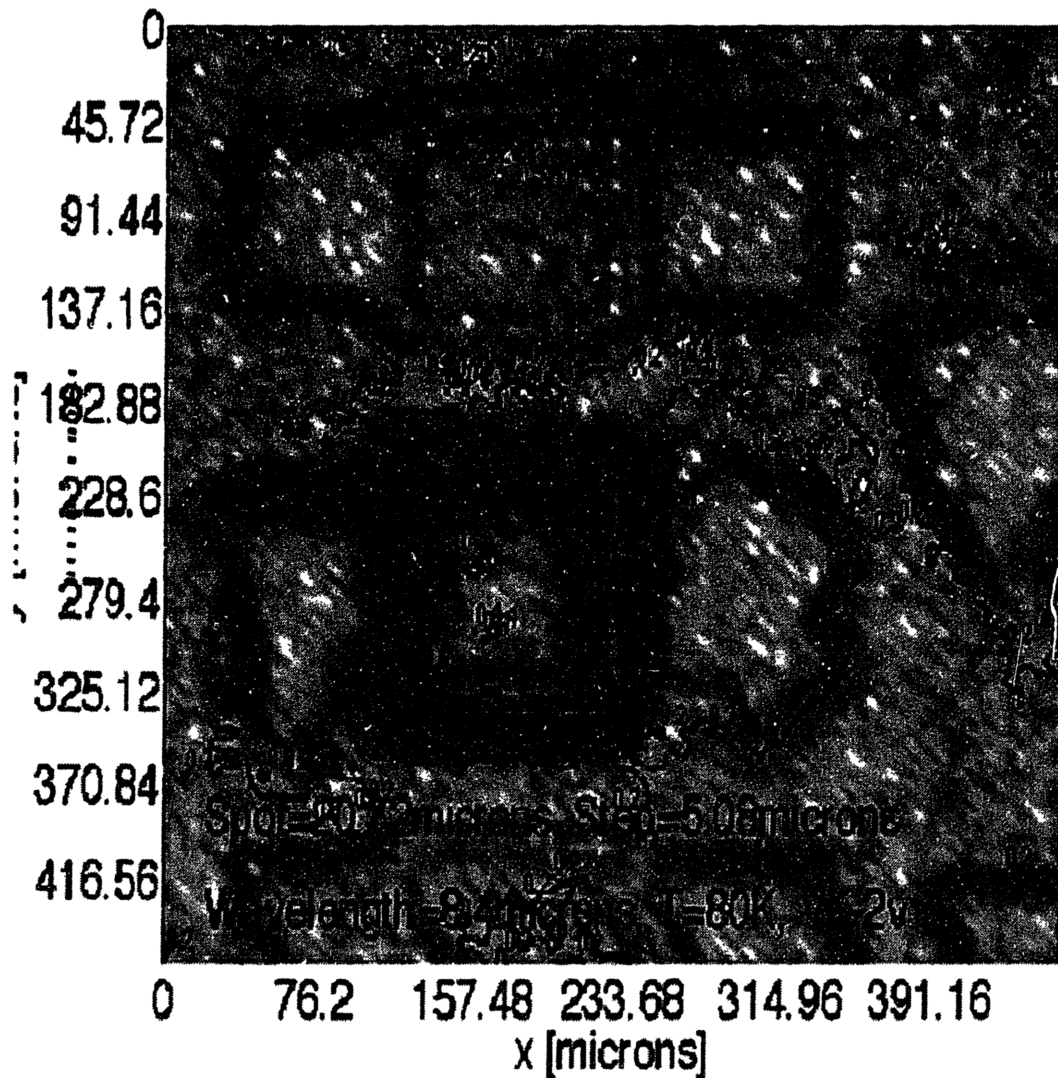


Figure 8-13: False contour map of the spatially resolved responsivity at a bias of -2v, $\lambda = 8.4\ \mu$m, and T=80K for a $(100\ \mu\text{m})^2$ optically active area on Sample 9066. This measurement was accomplished with f/1 reflective optics through a fiber to obtain a spot size of 20 μm which is scanned in increments of 5.1 μm. (This measurement was performed by Dr. Sam Wang at Lockheed Sanders.)

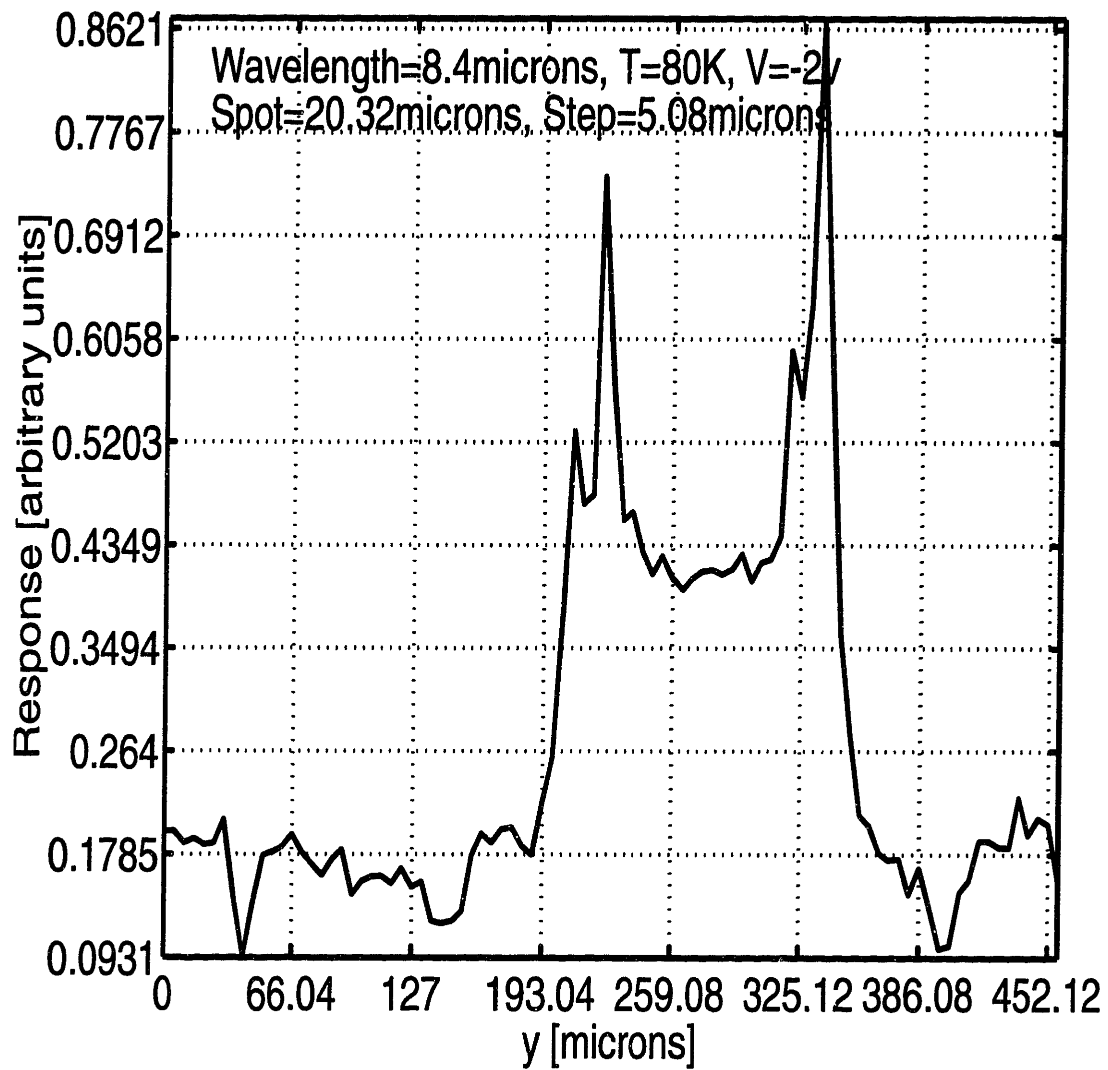


Figure 8-14: Spatially resolved responsivity as a function of distance along the x=160 μm cross section of Fig. 8-13.

a (100 μm)^2 optically active area on Sample 9066 to a spatially resolved radiation source of spot size of 20 μm which is scanned in increments of 5.1 μm. The relative size of the photoresponse as a function of distance along the x=160 μm cross section of Fig. 8-13 is shown in Fig. 8-14.

Analogous to Figs.8-11 and 8-12, Figs. 8-13 and 8-14 show that the responsivity is largest when the radiation is focused along the boundaries of the optically active area. However, unlike Fig. 8-11, the spot size of 20 μm used in the measurement of Fig. 8-13 is much smaller than the optically active area, which is (100 μm)^2 in Fig. 8-13 but is only (25 μm)^2 in Fig. 8-11. Since the focused radiation spot size is much smaller than the optically active area in the measurement corresponding to Fig. 8-13, the QWIP responsivity in this measurement is finite for TE polarized (normally incident) radiation. This response to normally incident radiation is three times as large as the measurement background, which we take to be the response resulting from radiation incident on the sample between QWIP pixels, as is clear from Fig. 8-14.

Finally, there exists a small amount of crosstalk in the response of adjacent QWIP pixels in Fig. 8-13. In the measurement corresponding to Fig. 8-13, the photocurrent running though the second (lower) QWIP pixel is being measured, and remains finite even when the 20 μm focused spot of radiation is scanned through the adjacent QWIP pixel lying above it in the figure. Dr. Sam Wang [102] at Lockheed Sanders reports that this type of crosstalk between adjacent pixels on a processed wafer is removed after thinning of the substrate.

8.5 Noise Measurement

Noise measurements are needed to obtain a measured value for the photoconductive gain, assuming that the noise is dominated by either generation-recombination noise or thermal leakage random arrival noise. Measured values for the photoconductive gain, along with measured values of the conversion efficiency, yield measured values of the quantum efficiency. The purpose of this section is to show that though a noise

measurement can always be made, the *interpretation* of that noise measurement may be difficult as a result of the presence of large stray capacitances in our measurement setup.

Noise current is usually measured [102, 103] by amplifying it with a low noise transimpedance amplifier, like a Keithley model 428, and then measuring the noise spectrum with a spectrum analyzer. The generation-recombination noise is usually taken from the flat portion of the noise spectrum between 300 Hz and 3 kHz. At low frequencies (lower than 300 Hz), the noise is observed to be dominated by 1/f noise. At high frequencies (higher than 3 kHz), the noise measurements are limited by the bandwidth of the transimpedance amplifier.

A question arises as to whether directly connecting the QWIP (inside the closed cycle refrigerator) to a transimpedance amplifier (outside the closed cycle refrigerator) really amplifies the generation-recombination noise spectrum of the QWIP. The question arises because the QWIP impedance is very large, and the noise spectrum will roll off at a low frequency determined by the product of the QWIP resistance and any stray capacitances across the QWIP. At typical operating temperatures and biases, the QWIP differential resistance is between 500 MΩ and 1 GΩ. Directly connecting the QWIP to a transimpedance amplifier with a one-foot coaxial cable, which has a capacitance of 30 pF/foot, results in an RC roll-off at a frequency of $[(500\ \mathrm{M}\Omega)(30\ \mathrm{pF})]^{-1}=2\pi(10.6\ \mathrm{Hz})$.

To fix this problem of a low frequency RC roll-off resulting from a large QWIP differential resistance, a source follower circuit, shown in Figure 8-15, is used to measure the noise. The source follower circuit is designed to achieve nearly unity gain, so that the source voltage V_{OUT} follows the gate voltage V_G, according to

$$\frac{V_{OUT}}{V_G} = \frac{g_m Z_s}{g_m Z_s + 1} \tag{8.9}$$

$$= g_m \left(g_m^{-1} \parallel R_{SS} \parallel r_0 \parallel \frac{1}{C_L s} \right). \tag{8.10}$$

The second line in Equation (8.10) is obtained by substituting for the impedance Z_s seen at the FET source node: $Z_s = R_{SS} \parallel r_0 \parallel \frac{1}{C_L s}$, where C_L is the stray capacitance measured at the source node, g_m is the JFET transconductance, R_{SS} is the resistor

connected to the source of the JFET, and r_0 is the JFET output resistance. Any stray capacitance (C_L in the figure), perhaps associated with a coaxial cable or an oscilloscope probe, does not see the large QWIP resistance R_{QWIP}. Rather, the capacitance C_L sees the output resistance of the source follower, which is the parallel combination $g_m^{-1} \parallel R_{SS} \parallel r_0$. For most of the measurements made, this parallel resistance is about equal to g_m^{-1}=1-3 kΩ, which is much smaller than the QWIP resistance. In using this source follower circuit, care must be taken not to put an oscilloscope probe or a function generator connection at the gate of the FET because this would add a stray capacitance (C_G in the figure, of about 13 pF for a good probe plus 30 pF/foot for the coaxial cable) in parallel with the large QWIP resistance, thus limiting the measured bandwidth to very low frequencies.

The circuit elements in Fig. 8-15 were chosen in the following manner. The general purpose, surface mount, depletion mode JFET was a Motorola MMBF5459LT1 SOT-23. These JFETs have low leakage current (nanoamperes at room temperature) and capacitances (C_{iss} <7 pF and C_{rss} <5 pF at room temperature). R_{BB} was chosen to be 300 MΩ, which is about equal to the QWIP resistance (a few hundred megaohms to a few gigaohms) so that the voltages across R_{QWIP} and R_{BB} both make appreciable contributions to the large and small signal gate voltage. The high value resistors were bought from the company OHMCRAFT Precision Resistors (3800 Monroe Avenue, Pittsford, NY 14534, 716-586-0824). R_{SS} should be chosen large enough to achieve about unity gain from the source follower, but small enough not to contribute much to the RC roll-off of the cable (at 30pF/foot). R_{SS} was chosen to be 100 kΩ for most of these measurements. V_{DD} can have a wide variety of values in the range of V_{DS} where I_{DS} does not change much. V_{BB} and V_{QQ} are chosen to satisfy desired values for V_{GS} and V_{DET}.

The measured transfer function $|V_{OUT}/V_{QQ}|$ is expected to verify the low-frequency-pass behavior of the source follower in Fig. 8-15. In this way, it is expected that noise measurements can be made with the source follower circuit at frequencies of about one hundred Hz, where the noise is dominated by the QWIP generation recombination noise and not 1/f noise. The QWIP geometric capacitance is about 2 pF for a (100 μm)$^{(2)}$ QWIP pixel having 10 periods of 500 Å barriers. The low frequency roll-

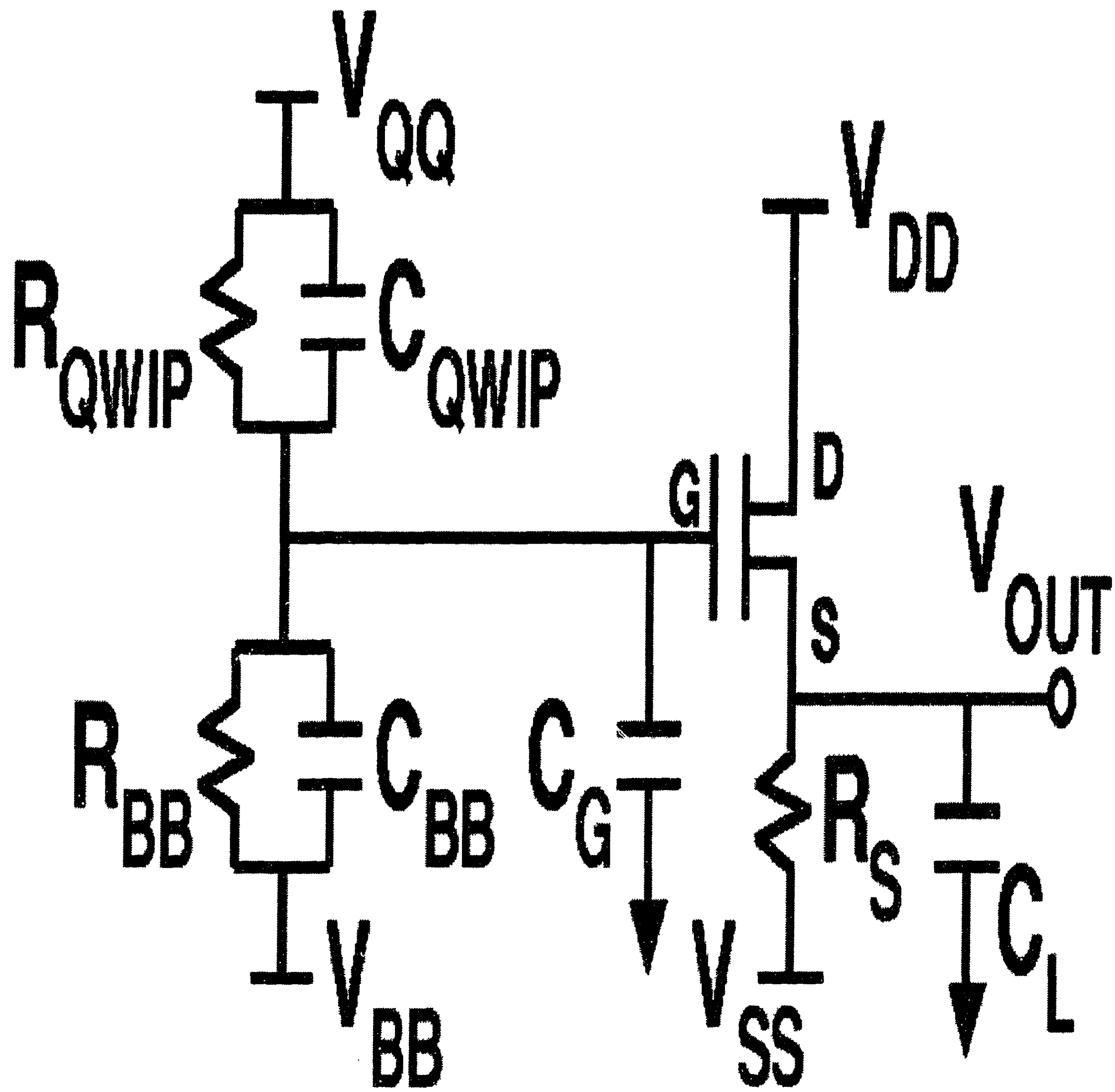


Figure 8-15: The source follower circuit mounted inside the DIP package alongside a QWIP to measure the QWIP low frequency response.

off is expected to occur at $(R_{QWIP}C_{QWIP})^{-1} \approx [(500\ M\Omega)(2\ pF)]^{-1} = 2\pi(160\ Hz)$.

In fact, the measured transfer function $|V_{OUT}/V_{QQ}|$ is expected to yield values for stray capacitances as well as the QWIP capacitance C_{QWIP}. This is done by fitting the measured transfer function to that expected for the circuit elements of Fig. 8-15:

$$\frac{V_G}{V_{QQ}} = \frac{\left[\frac{R_{BB}}{R_{BB}+R_{QWIP}}\right][R_{QWIP}C_{QWIP}s+1]}{\left[\left(\frac{R_{BB}R_{QWIP}}{R_{BB}+R_{QWIP}}\right)(C_B+C_{QWIP})s+1\right]}. \tag{8.11}$$

The ratio of the QWIP to bias resistances (R_{QWIP} to R_{BB}) can be obtained from the low frequency behavior of the transfer function,

$$\left|\frac{V_G}{V_{QQ}}\right|_{f\to 0} = \frac{R_{BB}}{R_{BB}+R_{QWIP}}. \tag{8.12}$$

The ratio of the QWIP capacitance C_{QWIP} to the stray capacitances $C_B = C_{BB} + C_G$ can be obtained from the high frequency behavior of the transfer function,

$$\left|\frac{V_G}{V_{QQ}}\right|_{f\to \infty} = \frac{C_{QWIP}}{C_{BB}+C_G+C_{QWIP}}, \tag{8.13}$$

where C_{BB} is the stray capacitance associated with the resistor R_{BB}, and C_G is the stray capacitance connecting the FET gate to ground. (See Fig. 8-15.) Individual values for the capacitances can be obtained from the break point frequencies; at frequencies above

$$2\pi f_{\text{upper}} = \frac{1}{R_{QWIP}C_{QWIP}}, \tag{8.14}$$

Equation (8.13) is valid, and at frequencies below

$$2\pi f_{\text{lower}} = \frac{1}{[R_{BB} \parallel R_{QWIP}][C_{QWIP}+C_{BB}+C_G]}, \tag{8.15}$$

Equation (8.12) is valid.

Figure 8-16 shows the measured low frequency response $\frac{V_{OUT}}{V_{QQ}} =$ of the source follower circuit in Fig. 8-15 mounted inside the DIP package alongside QWIP 9066, the bias resistor $R_{BB} = 300\ M\Omega$, and the JFET. The resistor $R_{SS} = 100\ k\Omega$ is located outside the DIP package and the cryostat because R_{SS} does not limit the bandwidth of the source follower. The sinusoidal input is denoted v_{qq}, and is applied

at the V_{QQ} node, where the total voltage is $v_{QQ} = V_{QQ} + v_{qq}$. The input voltage is 800 mV peak-to-peak, and the output voltage V_{OUT} is measured at the source node of the JFET. The measured frequency response shown in Fig. 8-16 was achieved with biases of V_{QQ}=0.67 v, V_{BB}=0.00 v, V_{DD}=4.97 v, and V_{SS}=-0.30 v. The FET is then operated at V_{DS}=2.0 v, V_{GS}=-2.9 v, i_D=32.68 μA. Direct measurement of the FET i-v characteristics at this operating point yields a measured r_0=367.6 kΩ and g_m^{-1}=3.13 kΩ.

The best fit of the measured data points in Fig. 8-16 to Equation (8.11) is indicated by the solid line in Fig. 8-16. The best fit values for the QWIP and bias resistors and the QWIP and stray capacitances are R_{QWIP}=365.3 MΩ, R_{BB}=321.1 MΩ, C_{QWIP}=60.5 pF, and $C_{BB} + C_G$=483.1 pF, respectively. These values of the resistances are the same as those found by direct measurement of the i-v characteristics at the measurement temperature. The stray capacitances however seem unusually large. The QWIP capacitance is expected to be about 2 pF, as discussed above, and the FET capacitances are expected to be at most a few picofarads at low temperature. The stray capacitances due to the DIP pins are perhaps 1 pF. There is no known stray capacitance of the order of 60 pF or 500 pF.

The frequency responses shown in Fig. 8-17 and Fig. 8-18 were measured to get a better understanding of the stray capacitances inferred from the measurement of Fig. 8-16. The measurements were accomplished with biases of V_{DD}=8.65 v, and V_{SS}=-8.65 v. The FET is then operated at V_{DS}=6.0 v, V_{GS}=-2.7 v, and i_D=113.0 μA. Direct measurement of the FET i-v characteristics at this operating point yields a measured r_0=235.3 kΩ and g_m^{-1}=1.406 kΩ. Also, V_{QQ}=0.95 v, V_{BB}=-0.358 v for the measurement of Fig. 8-17, and V_{QQ}=-3.3 v, V_{BB}=0.95 v for the measurement of Fig. 8-18. Again, $R_{BB} = 300\ M\Omega$ (inside the DIP package) and $R_{SS} = 100\ k\Omega$ (outside the DIP package). The measured frequency responses shown in Fig. 8-17 and Fig. 8-18 were both accomplished by replacing the QWIP with a 1 GΩ resistor. The sinusoidal input is 1 v peak-to-peak and is applied at the V_{QQ} node for the measurement of Fig. 8-17 and at the V_{BB} node for that of Fig. 8-18.

The purpose of replacing the QWIP with a 1 GΩ resistor is to obtain separate measurements of the different stray capacitances contributing to the measurement of

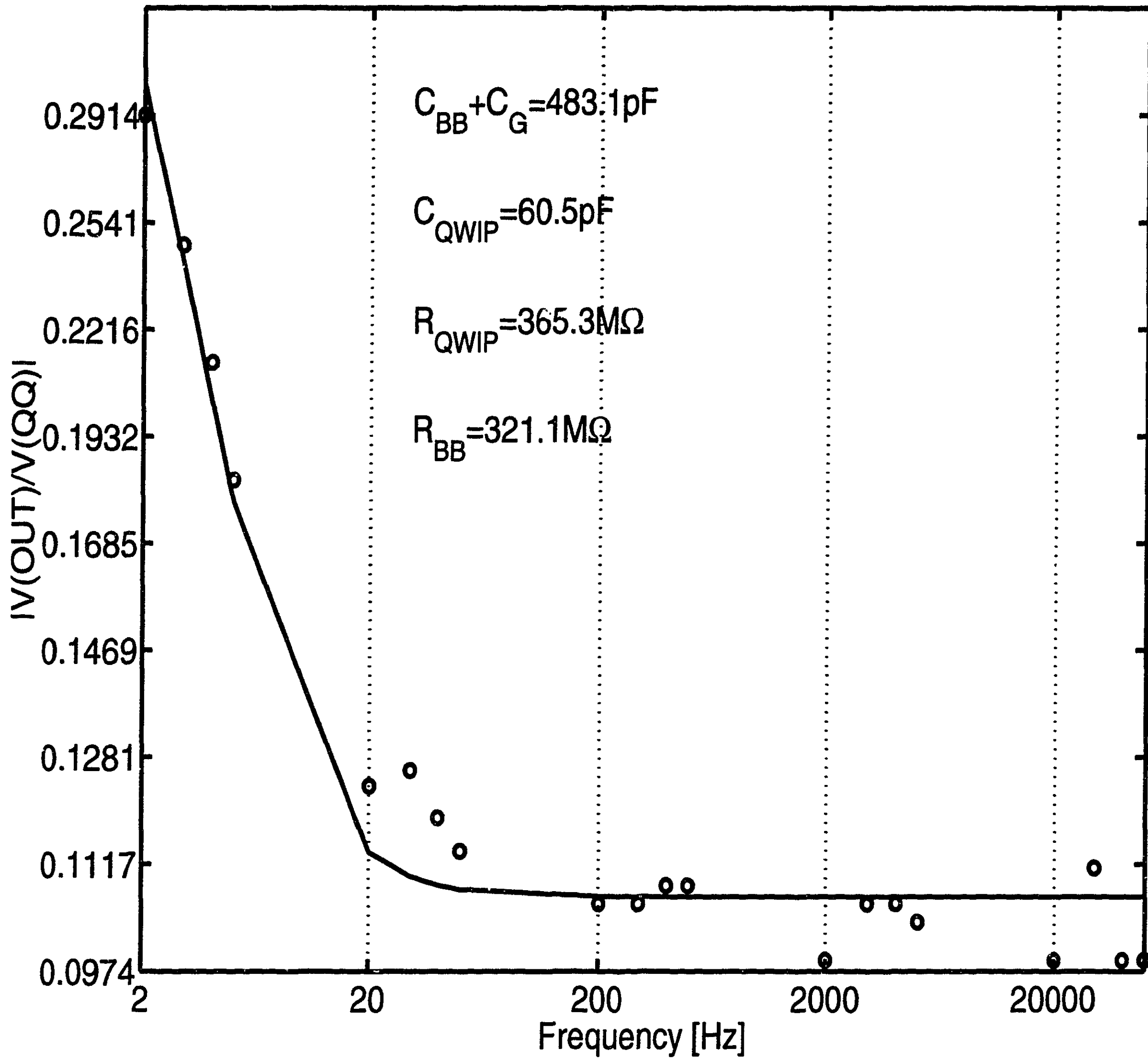


Figure 8-16: The points in the figure are the measured frequency response $\frac{V_{OUT}}{V_{QQ}} =$ of the source follower circuit mounted inside the DIP package alongside QWIP 9066. The solid line is the best fit of the measured data to Equation (8.11).

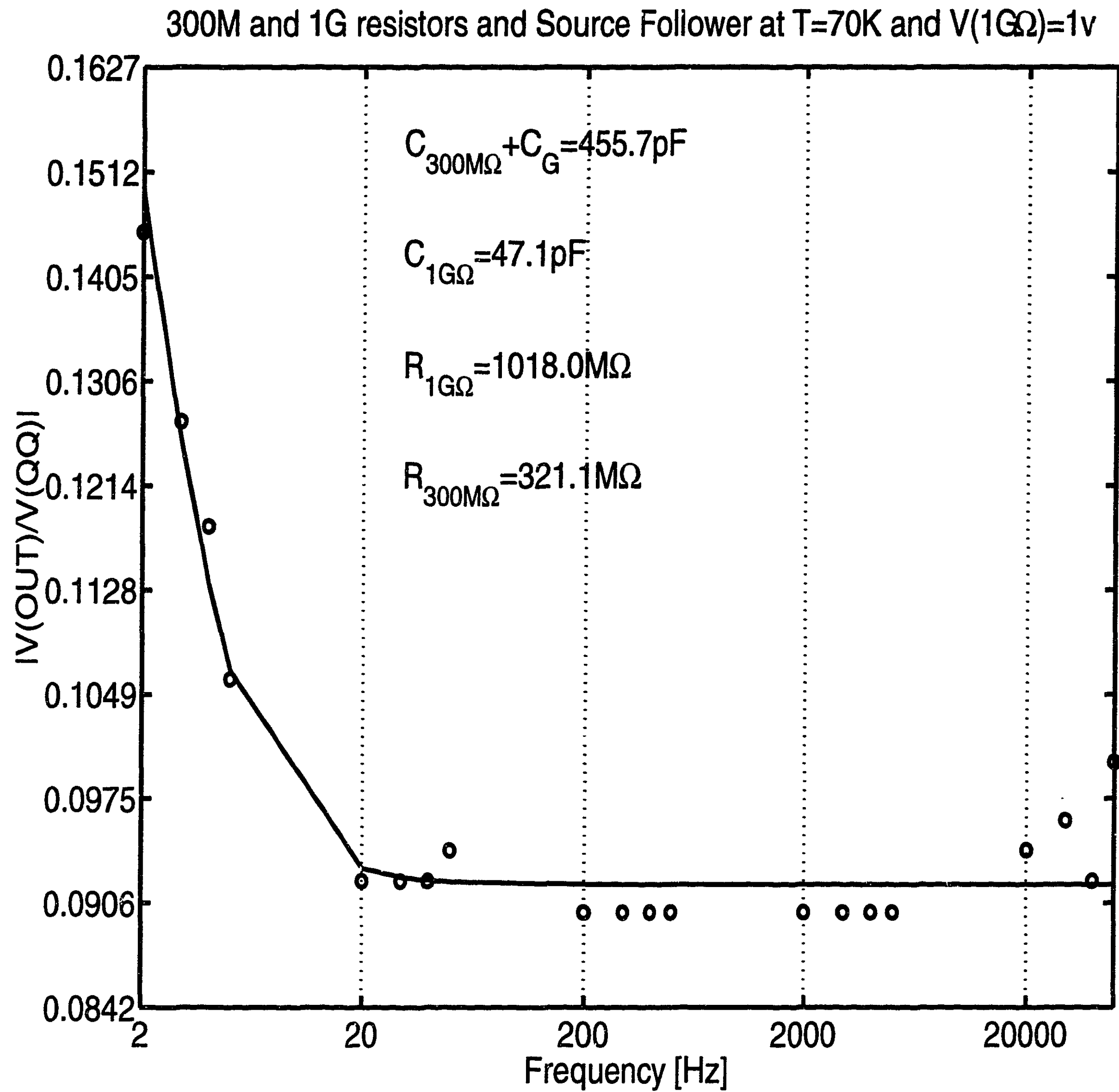


Figure 8-17: The points in the figure are the measured frequency response $\frac{V_{OUT}}{V_{QQ}} =$ of the source follower circuit mounted inside the DIP package alongside two resistors ($R_{QWIP} = 1\ G\Omega$ and $R_{BB} = 300\ M\Omega$) and no QWIP. The sinusoidal input is denoted v_{qq}, and is applied at the V_{QQ} node, where the total voltage is $v_{QQ} = V_{QQ} + v_{qq}$. The solid line is the best fit of the measured data to Equation (8.11).

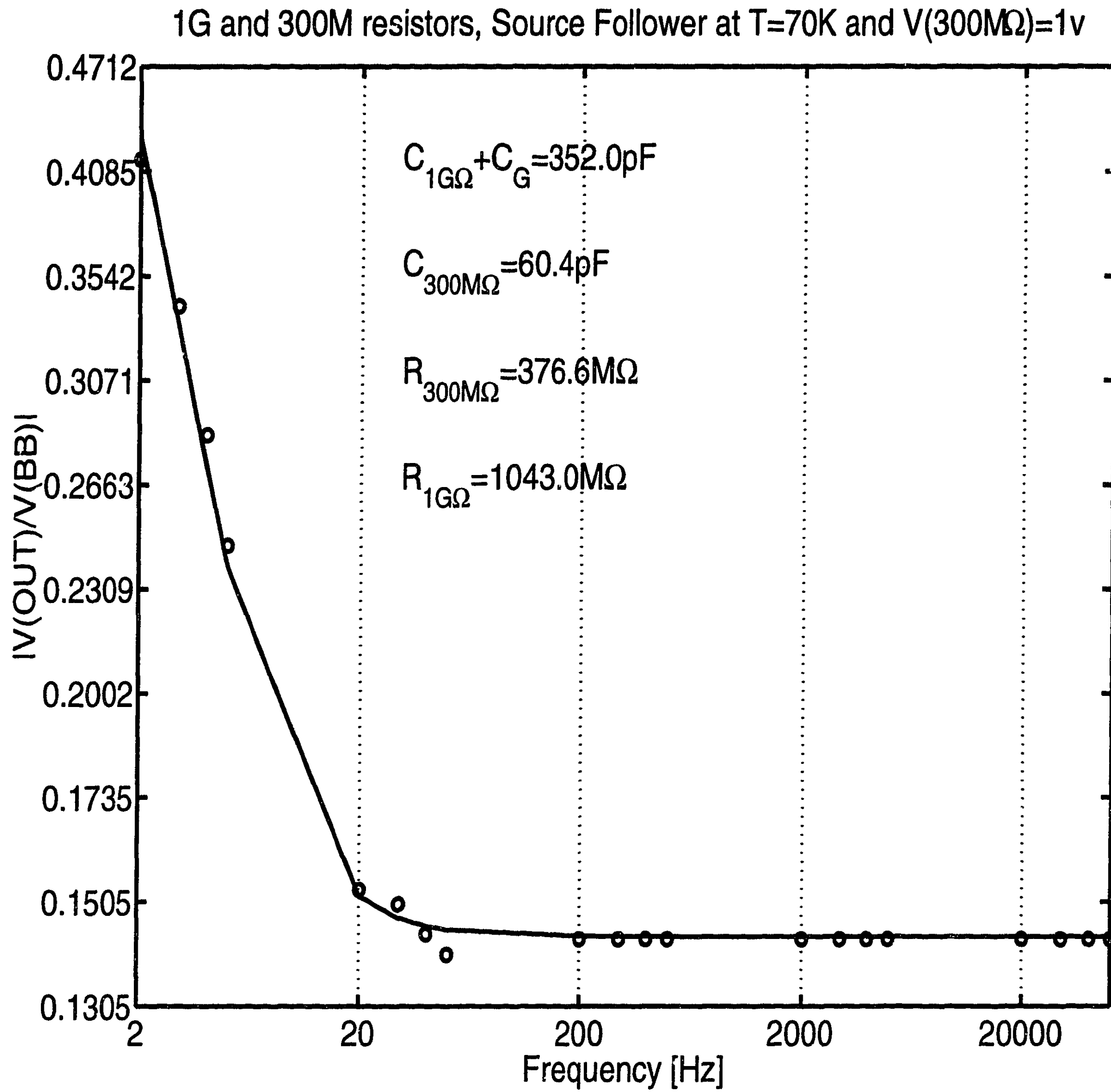


Figure 8-18: The points in the figure are the measured frequency response $\frac{V_{OUT}}{V_{BB}} =$ of the source follower circuit mounted inside the DIP package alongside two resistors ($R_{QWIP} = 1\ G\Omega$ and $R_{BB} = 300\ M\Omega$) and no QWIP. The sinusoidal input is denoted v_{bb}, and is applied at the V_{BB} node, where the total voltage is $v_{BB} = V_{BB} + v_{bb}$. The solid line is the best fit of the measured data to Equation (8.11).

Fig. 8-16. Replacing the QWIP with an OHMCRAFT resistor in Fig. 8-15 is important for assessing the low frequency response of the biasing circuit in the *absence* of the QWIP.

A frequency measurement complementary to that shown in Fig. 8-17 is summarized in Figure 8-18. The measurement shown in Fig. 8-18 is obtained with the same circuit as that used for the measurement of Fig. 8-17, where the QWIP was replaced by a 1 GΩ resistor. The difference between the two measurements is that the 1 V peak-to-peak input sinusoid is injected at the V_{QQ} node in Fig. 8-17 and at the V_{BB} node in Fig. 8-18. The importance of having both measurements is that Fig. 8-17 yields a measured value for $C_{300M\Omega} + C_G$, and Fig. 8-18 yields a measured value for $C_{1\ G\Omega} + C_G$. Both measurements together yield measured values for all *three* capacitances $C_{1\ G\Omega}$, $C_{300\ M\Omega}$, and C_G. The best fit values for all three stray capacitances, taken as a best fit to the values obtained from *both* Fig. 8-17 and Fig. 8-18, are $C_{1\ G\Omega}$=24.5 pF, $C_{300\ M\Omega}$=83 pF, and C_G=350.1 pF.

These measured values of the stray capacitances are unexpectedly large. At this time, the origin of these stray capacitances is unclear. The FET capacitances are expected to be at most a few picofarads at low temperature. The stray capacitances due to the DIP pins are perhaps 1 pF. There is no known stray capacitance of the order of 50 pF or 400 pF. In using the source follower circuit, care was taken not to put an oscilloscope probe or a function generator connection at the gate of the FET because this would add to the stray capacitance C_G at the FET gate. These stray capacitances may arise from the way the QWIP and the source follower circuit are mounted inside the DIP package. The superglue used to attach the QWIP, the FET, and the resistors to the DIP package was measured to have a resistance of about 1 MΩ at low temperature. Perhaps the way these electrical components have been glued into the DIP package adds extra capacitance to the frequency response.

The unexplained stray capacitances inferred from the measurements of Figs. 8-16, 8-17, and 8-18 seem to be purely capacitive in nature, with no significant resistive component. This can be deduced from the low frequency behavior of the source follower circuit in Fig. 8-15. The expected low frequency behavior (Equation (8.12)) of the source follower circuit can be used to check the resistances obtained from the

DC Bias V_{QQ} (volts)	DC Output V_{OUT} (volts)
0.85	2.628
0.95	2.650
1.05	2.675
1.15	2.695

Table 8.2: Measured DC voltage V_{OUT} at the FET source node as a function of different DC bias voltages V_{QQ} for the source follower circuit of Fig. 8-15. All other bias voltages and resistors used in this circuit are the same as those used in the measurement of Fig. 8-17.

best fit models (the solid lines) of Figs. 8-16, 8-17, and 8-18. Unfortunately, at the measured frequencies of above 2 Hz in Figs. 8-16, 8-17, and 8-18, the low frequency behavior (Equation (8.12)) is not observed. To check the resistive components of the source follower circuit requires another means, that of measuring DC transfer functions. One such measurement of the measured DC output voltage V_{OUT} as a function of various DC bias voltages V_{QQ} is shown in Table 8.2 for the circuit used in the measurement of Fig. 8-17. The numbers in Table 8.2 show a small signal DC transfer function $\Delta V_{OUT}/\Delta V_{QQ}$ of about 0.235, which is consistent with Equation (8.12) and the resistances inferred from the best fit procedure shown in Fig. 8-17. Similar measurements of the DC transfer functions of the circuits used in the measurements of Figs. 8-16 and 8-18 show that the low frequency behavior of both source follower circuits are completely explained by the bias resistors which were deliberately mounted into the DIP package.

In summary, measurement of the QWIP generation recombination noise is needed to get a measured value for the photoconductive gain. Directly connecting the QWIP (inside the closed cycle refrigerator) to a transimpedance amplifier (outside the closed cycle refrigerator) is not desirable because the QWIP impedance is very large (500 MΩ-1 GΩ), and the capacitance of the coaxial cable will cause the noise spectrum

to roll off at too low a frequency for a reliable measurement of the generation recombination noise. The noise spectrum can still be measured if a source follower circuit is mounted alongside the QWIP in the closed cycle refrigerator. The noise spectrum can then be measured at higher frequencies because the capacitance of the coaxial cable sees the output resistance of the source follower (about $g_m^{-1} \approx 1\ k\Omega$) instead of the much larger QWIP differential resistance. Unfortunately, in our measurement setup, the use of the source follower circuit did not increase the measured noise bandwidth because of unusually large stray capacitances associated with the QWIP, the bias resistors and the JFET. The origin of the stray capacitances is unclear at this time.

8.6 Conclusions

This chapter described general procedures for measuring QWIP device parameters. An example of an n-type QWIP (n-QWIP) was used throughout the chapter.

Leakage current measurements are made as a function of QWIP operating temperature and bias. At low operating temperatures, the leakage is dominated by sequential resonant tunneling. At large QWIP bias voltages, the leakage is dominated by Fowler-Nordheim tunneling. At intermediate operating temperatures (above 50K) and bias voltages, the leakage is dominated by thermionic leakage. An Arrhenius plot of the measured leakage as a function of inverse QWIP operating temperature reveals an activation energy for the leakage. This activation energy can be used to determine the position of the Fermi level within the occupied quantum well subband. The bias dependence of this activation energy Φ_B has not been studied in the QWIP literature, and the models to be presented in Chapter 10 reveal how the bias dependence of this activation energy may be related to the injection mechanisms in the QWIP. The quality of fit of the measured leakage to a thermionic leakage model indicates the absence of excessive leakage. Excessive leakage was not observed in this n-QWIP example at intermediate negative bias voltages.

The conversion efficiency, the product of the quantum efficiency and the photoconductive gain, can be measured by measuring the net photocurrent as a function of different

black body target sizes for a fixed QWIP operating temperature and a fixed black body temperature. The typical n-QWIP we measured had a conversion efficiency of 4% or, equivalently, a responsivity of 270 mA/W or, equivalently, a detectivity of 4×10^{10}cm-$\sqrt{Hz}$/Watt without the use of an optical grating. This size conversion efficiency or detectivity is large enough for focal plane array performance to be limited by the uniformity of processing rather than the size of the single pixel detectivity.

Measurement of the QWIP photocurrent (for a fixed QWIP operating temperature and a fixed black body target size) as a function of the inverse black body temperature reveals an Arrhenius dependence whose slope is the energy of the responsivity peak. The energy of the responsivity peak is separately verified by a direct measurement of the responsivity spectrum using a circular variable filter to isolate particular wavelengths from a black body infrared source.

The spatially resolved QWIP responsivity measurements shed light on the nature of the absorption of normally incident radiation by an n-QWIP which has been processed without an optical grating. First, TE polarized radiation having a small spot size of 20 μm normally incident upon a sample in the middle of a large 100 μm optically active area yields a finite response. This response is three times as large as the measurement background, which we take to be the response resulting from radiation incident on the sample between QWIP pixels. Second, radiation incident upon the pixels at the pixel boundaries, which are (111) planes, yields a large response, which is presumed to be dominated by the response due to TM polarized radiation. The measured photocurrent is thus a response due to both TE and TM polarizations. Finally, there is a small amount of responsivity crosstalk between adjacent QWIP pixels.

Measurement of the QWIP generation recombination noise is needed to get a measured value for the photoconductive gain. Attempts to make noise measurements on QWIPs were inconclusive. Though a noise measurement can always be made, the *interpretation* of that noise measurement may be difficult as a result of the presence of large stray capacitances. Directly connecting the QWIP (inside the closed cycle refrigerator) to a transimpedance amplifier (outside the closed cycle refrigerator) is not desirable because the QWIP impedance is very large (500 MΩ-1 GΩ), and the

capacitance of the coaxial cable will cause the noise spectrum to roll off at too low a frequency for a reliable measurement of the generation recombination noise. The noise spectrum can still be measured if a source follower circuit is mounted alongside the QWIP in the closed cycle refrigerator. The noise spectrum can then be measured at higher frequencies because the capacitance of the coaxial cable sees the output resistance of the source follower (about $g_m^{-1} \approx 1\ k\Omega$) instead of the much larger QWIP differential resistance. Unfortunately, in our measurement setup, the use of the source follower circuit did not increase the measured noise bandwidth because of unusually large stray capacitances associated with the QWIP, the bias resistors and the JFET. The origin of the stray capacitances is unclear at this time.

Chapter 9

Some Device Designs

QWIPs exhibiting a large SNR are one possible design objective. As discussed in Chapter 3, for very low operating temperatures (below T_{BLIP}), a large SNR can be obtained with a large single quantum well quantum efficiency, a large number of quantum wells, and an optimal single quantum well photoconductive gain.

In many cases, the ability to operate a QWIP at the highest possible temperature while maintaining a particular SNR is a more practical goal, as the cost of buying a cooler or the time it takes to reach the QWIP operating temperature is the biggest operating expense. In such cases, maximizing the SNR of a QWIP at some operating temperature above T_{BLIP} is the design goal. This can be accomplished by increasing the conversion efficiency or decreasing the leakage current or increasing the device parameter uniformity across a FPA.

Section 9.1 presents some commonly used n-QWIP designs. Section 9.2 discusses three p-QWIP designs, and relates their device parameters to the device figures of merit.

9.1 Common n-QWIP Designs

Some common n-QWIP designs [14, 13] are shown in Figure 9-1. In all three designs shown in the figure, the intersubband transition energy ($E_2 - E_1$) is fixed by the application.

The uppermost schematic in the figure shows an intersubband transition from a bound state in a QWIP to an upper state in the continuum of energies above the barrier band edge. This design is often referred to as the bound-to-continuum QWIP design, and was first demonstrated at Bell Laboratories [104]. Of the three designs shown in the figure, the bound-to-continuum QWIP design displays the largest dark current and the largest absorption spectral full-width-half-maximum (FWHM). This design shows the largest dark current because the confining potential, V_B, is smallest for a fixed $E_2 - E_1$ of the three designs shown. This design shows a large absorption FWHM because transitions are possible from the quantum well bound state to many states in the continuum.

The middle schematic shows an intersubband transition from a bound state in a QWIP to an upper state at the barrier band edge. This design is often referred to as the bound-to-quasicontinuum QWIP design, and has been studied by Levine [7], Liu [105], and Gunapala [106]. This design has a higher quantum efficiency than the uppermost design in Fig. 9-1 because the the upper state in the middle schematic still retains much of its bound state character. (Transitions between bound states in a quantum well show stronger[105] absorption than transitions from a bound state to a state in the continuum.) Carriers excited to the quasicontinuum also show good escape probability since the uppermost state is at the barrier band edge. This design shows smaller leakage than the uppermost design in Fig. 9-1 because the confining potential, V_B, is larger for the middle schematic at a fixed $E_2 - E_1$.

In the middle schematic shown in Fig. 9-1, the upper state E_2 is designed to be at the barrier band edge. The optimal barrier width in this design was found experimentally by Levine [7] to be 500 Å, because QWIPs made with thinner barriers exhibited large sequential resonant tunneling. However, QWIPs made with wider barriers (1000 Å)

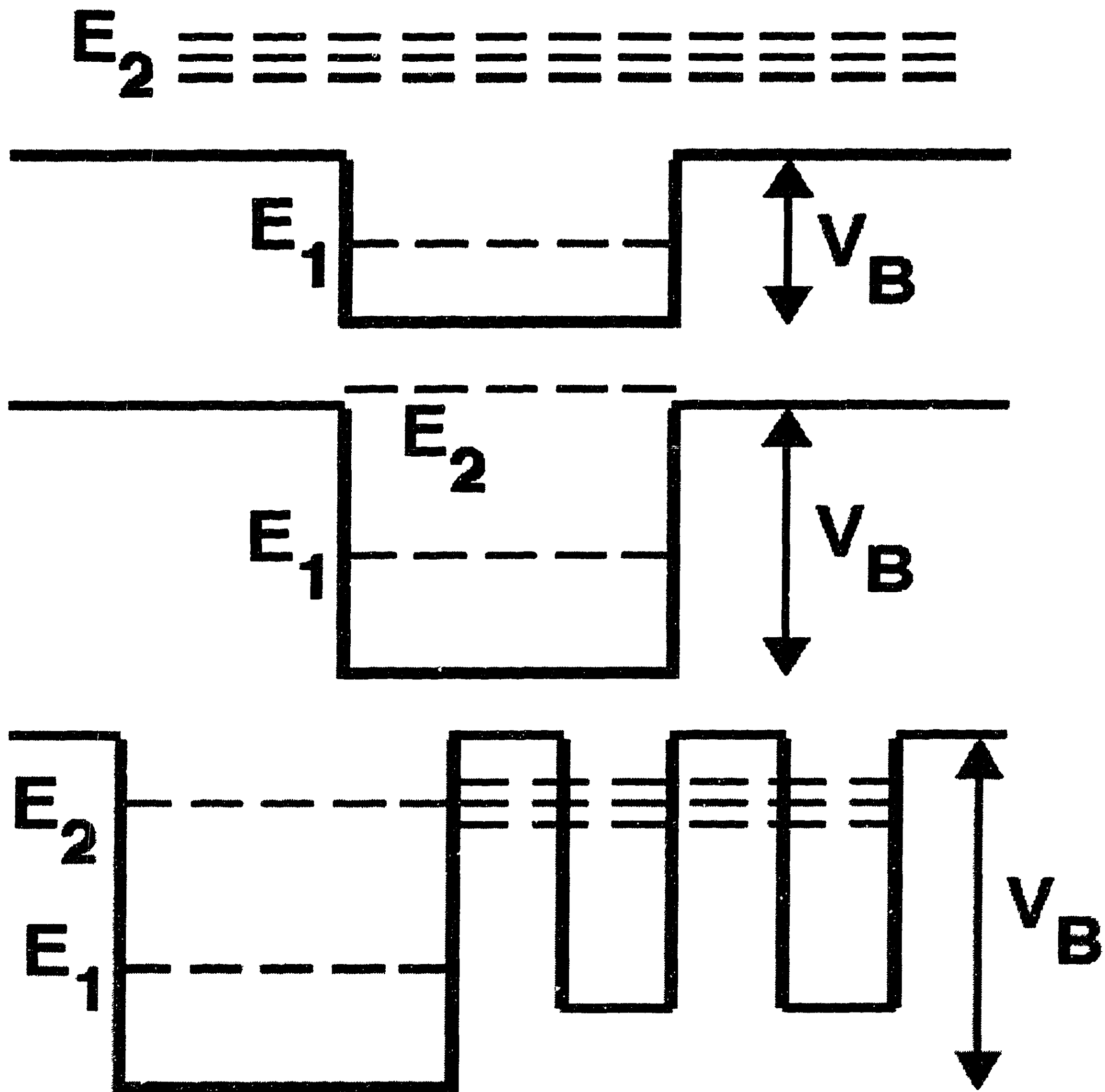


Figure 9-1: Some common n-QWIP designs. The uppermost schematic shows an intersubband transition from a bound state in a QWIP to an upper state in the continuum of energies above the barrier band edge. The middle schematic shows an intersubband transition from a bound state in a QWIP to an upper state at the barrier band edge. The lower schematic shows an intersubband transition between two bound states in a quantum well, where photocurrent transport is via miniband transport through the superlattice barrier.

did not show less [7] sequential resonant tunneling because larger bias voltages were required to obtain the same photocurrents which were measured for the 500 Å barrier QWIPs.

The lower schematic shows an intersubband transition between two bound states in a quantum well, where photocurrent transport is via miniband transport through the superlattice barrier. This design is often referred to as the bound-to-miniband QWIP design, and it was developed by researchers at Lockheed Martin [107, 13]. This design has a high quantum efficiency, since it involves a bound-to-bound intersubband transition. Photoexcited carriers show a high escape probability because the upper quantum well state is energetically degenerate with the miniband in the superlattice barrier. This design is expected to show the smallest thermionic leakage of the three designs shown in Fig. 9-1 because the confining potential V_B is largest for this design at a fixed $E_2 - E_1$). More quantitatively, if the thermionic leakage for the middle design shown in Fig. 9-1 is $J_{\text{th,B-to-QC}}$ and the thermionic leakage for the last design shown in the figure is $J_{\text{th,B-to-MiniB}}$, it is expected that,

$$\frac{J_{\text{th,B-to-MiniB}}}{J_{\text{th,B-to-QC}}} = \exp(-[V_B - E_2]/k_B T). \tag{9.1}$$

The infrared image shown in Chapter 1 was measured with a 640x480 focal plane array made from a bound-to-miniband design.

9.1.1 Kronig-Penney Model of the Superlattice Barriers

In this section, we calculate the tunneling through the superlattice barrier shown in the last design in Fig. 9-1 by comparing the superlattice barrier with a Kronig-Penney model. For ease of calculation, the periodic potential for these superlattice barriers is modeled as,

$$V(z) = V_{av} L \delta(z - nL) \tag{9.2}$$

where L is the periodicity of the superlattice barrier, and $V_{av} = \frac{1}{L}\int_{-L/2}^{L/2} V(z)dz$ is the average potential barrier (band edge) of the semiconductors comprising superlattice

barrier. Here, we will assume that $V_{av} > 0$ and the carrier energy $E > 0$, so that we are considering carriers incident upon a lattice of narrow, tall potential barriers. Narrow, tall potential barriers, as in Equation (9.2), are chosen to model the QWIP superlattice barriers because the QWIP superlattice barrier widths must be chosen to be narrow enough to allow the formation of a miniband for electron transport.

For a periodic potential, the solutions $\Psi(z)$ of Schrodinger's equation satisfy Bloch's theorem,

$$\Psi(z) = \exp(ikz)u(z) \tag{9.3}$$

where $u(z) = u(z + L)$ has the periodicity of the superlattice, and where k is the eigenvalue of the translation operator. For $E > 0$, we define

$$q^2 = 2mE/\hbar^2. \tag{9.4}$$

Kronig and Penney [43, 108] show that the energy eigenvalues form a miniband of energies whose k value satisfies,

$$\cos kL = \cos qL + \frac{mV_{av}L^2}{\hbar^2}\frac{\sin qL}{qL}. \tag{9.5}$$

It is easy to show [43] that minibands of energies are formed when,

$$\begin{aligned}(2n+1)\pi - \epsilon_n \leq qL &\leq (2n+1)\pi \\ 2n\pi \leq qL &\leq 2n\pi + \epsilon_n \\ &\text{for integral n}\end{aligned} \tag{9.6}$$

where ϵ_n is the energy width of the miniband. The width of the miniband ϵ_n is very narrow in the tight binding limit where V_{av} is large. For bound-to-miniband QWIPs, it desirable that the miniband be wide enough in energy to yield a reasonable photoconductive gain at a particular bias voltage.

Much of the work [109, 110] on the bound-to-miniband QWIPs has concentrated on calculating the miniband energies. What has not appeared in the literature are detailed, analytical expressions for the tunneling through the minigap regions, the regions between the miniband energies in Equation (9.6). Equation (9.5) can be used

to show that the minigap regions are centered about,

$$\text{qL=0 and qL=} \left(n + \frac{1}{2}\right)\pi \quad \text{for integral n.} \tag{9.7}$$

For the superlattice barriers in Fig. 9-1 having a finite width L_B (i.e., the superlattice does not extend to z=±∞), Bloch's theorem [108] also allows solutions to Schrodinger's equation of the form,

$$\Psi(z) = \exp(\pm\mu z)u(z), \tag{9.8}$$

so that the tunneling through the superlattice barrier through the minigap regions can be approximated as,

$$T \approx \exp(-2\mu L_B) \equiv \exp\left(-2L_B\sqrt{\frac{2m}{\hbar^2}(V_{tunn} - E)}\right), \tag{9.9}$$

where V_{tunn} defines an effective tunneling barrier which is derived below, and where the imaginary part of the wave vector μ satisfies the Kronig-Penney equation, Equation (9.5),

$$\cosh\mu L = \cos qL + \frac{mV_{av}L^2}{\hbar^2}\frac{\sin qL}{qL}. \tag{9.10}$$

Equation (9.10) can be solved numerically, but more insight can be obtained by getting an analytical solution. Right in the middle of the minigap, Equation (9.7), the inverse, μ, of the decay length satisfies,

$$\begin{aligned}\cosh\mu L &= 1 + \frac{mV_{av}L^2}{\hbar^2} \quad \text{near qL=0}\\ \cosh\mu L &= \frac{mV_{av}L^2}{\hbar^2}\frac{1}{(n+\frac{1}{2})\pi} \quad \text{near qL=}\left(n+\frac{1}{2}\right)\pi.\end{aligned} \tag{9.11}$$

Equation (9.11) can be cast in a form which is easy to remember. It is easy to show from Equation (9.10) that,

$$\begin{aligned}\mu^2 &= \frac{2m}{\hbar^2}\left(\left[V_{av}\left(1 - \frac{(qL)^2}{6}\right)\right] - E\right)\\ &\quad \text{for } \frac{mV_{av}L^2}{\hbar^2} < 1 \text{ and qL}\sim 0\\ \mu^2 &\ll \frac{2m}{\hbar^2}\left(\left[V_{av}\left(1 - \frac{(qL)^2}{6}\right)\right] - E\right)\\ &\quad \text{for } \frac{mV_{av}L^2}{\hbar^2} \gg 1 \text{ and qL}\sim 0\end{aligned} \tag{9.12}$$

where we have used Equation (9.4) for E. Equation (9.12) is physically significant, and is a key result of this chapter. It says that an electron which has an energy deep within the first minigap (with qL≈0) and which is incident on the superlattice barrier, will have a probability for tunneling as given in Equation (9.9) with a tunneling barrier of,

$$V_{tunn} \leq V_{av}\left(1 - \frac{(qL)^2}{6}\right) \quad \text{near qL=0.} \tag{9.13}$$

This tunneling barrier V_{tunn} can be much smaller than the average superlattice potential, $V_{av} = \frac{1}{L}\int_{-L/2}^{L/2} V(z)dz$, especially for barriers designed to be highly confining (for $\frac{mV_{av}L^2}{\hbar^2} \gg 1$). In such cases, the exact expression, Equation (9.11), should be used to calculate μ.

Equation (9.11) can be used to show that

$$\mu^2 = \frac{2m}{\hbar^2}\left[\frac{V_{av}}{(n+\frac{1}{2})\pi} - \frac{\hbar^2}{mL^2}\right]$$
$$\text{at the center of the minigap at qL=}\left(n+\frac{1}{2}\right)\pi. \tag{9.14}$$

Equation (9.14) shows significant tunneling.

9.2 Design of p-QWIPs with large quantum efficiency and low leakage current.

Both the quantum confinement and the strain contributions to the Hamiltonian have the same (tetragonal) symmetry. This means that qualitatively, both change the light hole wave function in exactly the same way: by adding s-symmetry to the nominally p-symmetric light hole wave functions. In the literature, both quantum confinement [47] and uniaxial strain [50, 53] have been used to increase the quantum efficiency by increasing the strength of the absorption of normally incident radiation in the heavy hole to light hole transition.

This work differs from other works in the literature in our choice of the use of strain. Whereas both quantum confinement and uniaxial strain have the same qualitative

effect, that of adding s-symmetry to the nominally p-symmetric light hole, these two effects differ in their size and strength. Since quantum confinement energies (typically several hundred meV) are much larger than strain splittings (a few tens of meV), quantum confinement is much more effective than uniaxial strain in changing the symmetry of the light holes and increasing the quantum efficiency. However, quantum confinement energies are always positive because zero point energies are always positive. In contrast, strain splittings can have either sign: for tensile strain, the light hole band is closer to the conduction band than is the heavy hole band, and for compressive strain, the heavy hole band is closer to the conduction band. Moving the light hole band relative to the heavy hole has important consequences for the transport of current carriers and also for the size of the leakage currents.

In our work, uniaxial strain is not used for increasing quantum efficiency. The quantum efficiency for absorption of normally incident radiation by holes is already very large because of the large size of quantum confinement energies. Rather, strain is used to change the transport properties of the photocurrents and also to reduce leakage currents.

By pulling the light hole continuum below the heavy hole continuum, the much smaller light hole mass becomes the transport mass. Kuroda and Garmire [111] have already noted this in a theoretical discussion of p-QWIPs designed to have larger photoconductive gains.

Moreover, continuum heavy hole states contribute little photocurrent because the bound state heavy hole to continuum heavy hole absorption is very weak for normally incident radiation. In an unstrained semiconductor, the heavy holes dominate the density of states in the continuum of energies above the barrier, thus contributing little to the photocurrent and contributing mostly to the leakage current for normally incident radiation upon a QWIP without an optical grating. By pulling the light hole continuum closer to the conduction band and away from the heavy hole continuum, the heavy hole continuum, which contributes mostly leakage current, is moved away from the states (the light hole continuum) contributing mostly photocurrent. This reduces the leakage current resulting from thermal population of the heavy hole

continuum. The reduction in the heavy hole thermal leakage current is,

$$\frac{J_{\text{th,hh,with strain}}}{J_{\text{th,hh,without strain}}} = \exp(-|\Xi|/k_B T_{op}), \tag{9.15}$$

where Ξ is the splitting of the light hole and heavy hole continua.

The p-QWIP designs presented in this thesis have the new feature that strain is used to lift the light and heavy hole valence degeneracy in the energy continuum, so that the lowest energy hole band in the continuum of energy states can be chosen as light-hole-like for both large photoconductive gain and large quantum efficiency. The use of InP as the substrate allows for both tensile and compressive strain to be possible with InGaAlAs layers. The photocurrents and leakage currents of devices grown with either tensilely or compressively strained layers are to be compared. The ability to have both compressively and tensilely strained layers also allows the entire growth to have zero net strain.

9.2.1 Three p-QWIP Designs.

Three QWIP structures were grown to increase size of the quantum efficiency and photoconductive gain, or to decrease the leakage current. All structures are designed to have a peak responsivity at 156 meV (8 μm) and zero net strain.

The first one, known as "Design1," consists of 15 periods of 50 nm tensile $In_{0.499}Ga_{0.274}Al_{0.227}As$ barriers and 10.07 nm compressive $In_{0.67}Ga_{0.33}As$ wells, as shown in Figure 9-2. It has a large quantum efficiency because it involves a heavy hole (n=1) to light hole (m=2) transition. The lowest continuum energy states will have a large photoconductive gain, as the transport effective mass is the light hole mass. Both the photocurrent and the leakage currents are expected to be larger than that for an unstrained p-QWIP because of this large photoconductive gain. The splitting of the light and heavy hole continua in the barrier region is calculated to be 16.8 meV. For comparison, $k_B T$=6.6 meV at T=77 K.

The second structure, known as "Design2," consists of 15 periods of 50 nm compressive $In_{0.532}Al_{0.468}As$ barriers and 6.24 nm tensile $In_{0.45}Ga_{0.55}As$ wells, as shown in

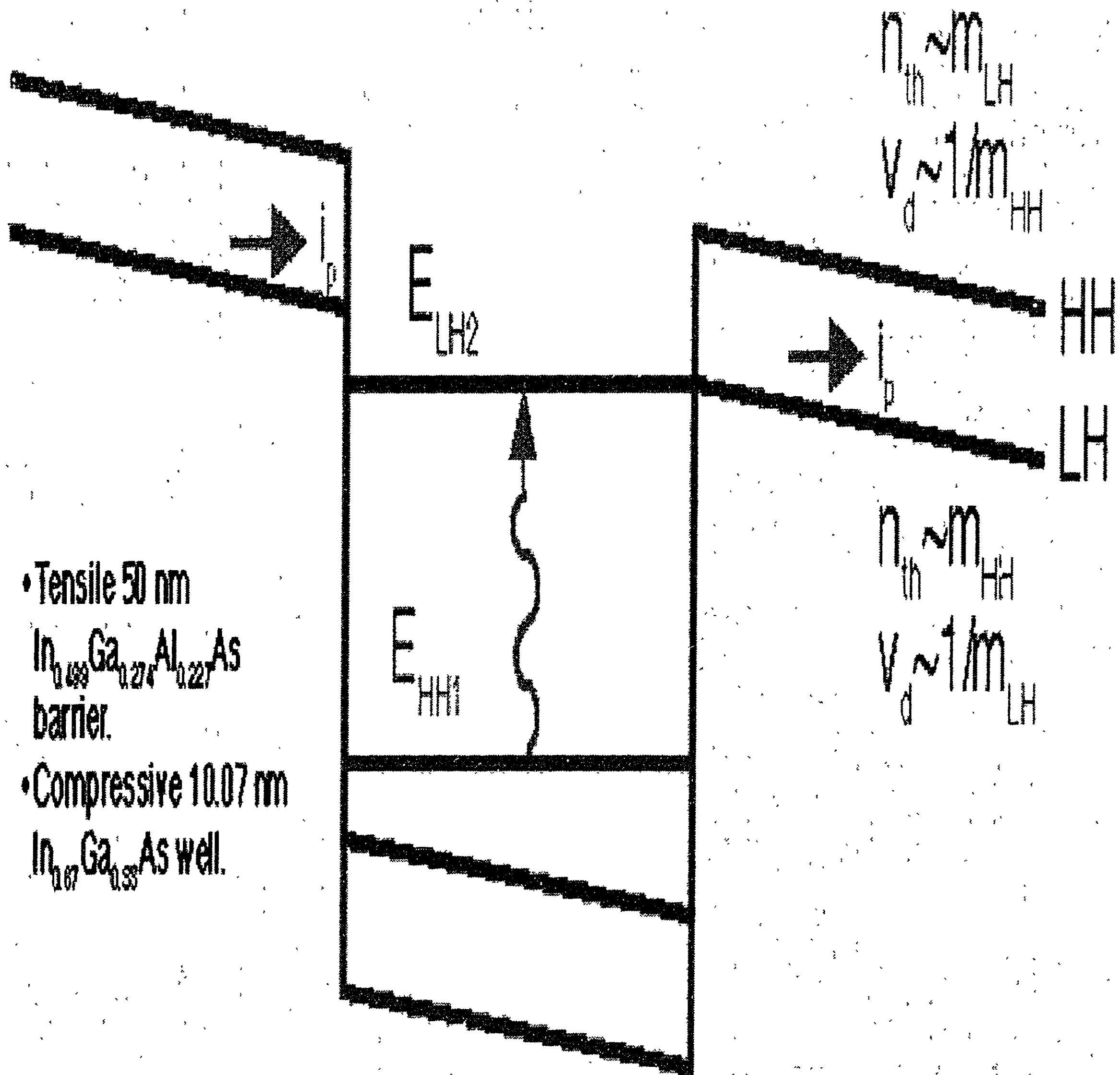


Figure 9-2: The "Design1" design, consisting of 15 periods of 50 nm tensile $In_{0.499}Ga_{0.274}Al_{0.227}As$ barriers and 10.07 nm compressive $In_{0.67}Ga_{0.33}As$ wells, is expected to have large quantum efficiency and large photoconductive gain.

Figure 9-3. It has a large quantum efficiency because it involves a light hole (n=1) to heavy hole (m=2) transition. It will have a low photoconductive gain because the lowest continuum band, the strained heavy hole band, has a large transport effective mass, that associated with the bulk heavy hole band. This structure was designed to have a small leakage current since the strain induced splitting between the light hole and heavy hole bands in the energy continuum above the barriers forces the strained light hole band, which has a large density of states (in-plane) effective mass, away from the strained heavy hole band, through which the photocurrent is flowing in the barrier. This strain induced splitting of the light and heavy hole bands in the energy continuum forces the leakage current to photocurrent ratio down a factor of $\exp(-[E_{LH} - E_{HH}]/k_B T_{op})$ compared with a similar structure for which the barrier light and heavy hole bands are energetically degenerate. In order to maintain a zero net strain for the entire growth, the largest strain induced splitting of the light and heavy hole bands in the energy continuum that is possible with designs consisting of barriers having just one single composition ($In_{0.532}Al_{0.468}As$ in this case) is 5.6 meV.

The third structure, known as "Design3," consists of 15 periods of 8.11 nm lattice matched $In_{0.532}Ga_{0.468}As$ wells and a barrier consisting of five periods of 3.0 nm tensile $In_{0.358}Ga_{0.166}Al_{0.476}As$ barrier1 and 6.6 nm compressive $In_{0.60}Al_{0.40}As$ barrier2, as shown in Figure 9-4. It has a large quantum efficiency because it involves a heavy hole (n=1) to light hole (m=2) transition. It will have a low photoconductive gain because the lowest continuum band, the strained heavy hole band, has a large transport effective mass, that associated with the bulk heavy hole band. This structure was designed to have a small leakage current since the strain induced splitting between the light hole and heavy hole bands in the energy continuum above the barriers forces the strained light hole band, which has a large density of states (in-plane) effective mass, away from the strained heavy hole band, through which the photocurrent is flowing in the barrier. This strain induced splitting of the light and heavy hole bands in the energy continuum forces the leakage current to photocurrent ratio down a factor of $\exp(-[E_{LH} - E_{HH}]/k_B T_{op})$ compared with a similar structure for which the barrier light and heavy hole bands are energetically degenerate . "Design3" differs from "Design2" in that a larger strain induced splitting of 30 meV between the light and heavy hole bands in the energy continuum is obtained.

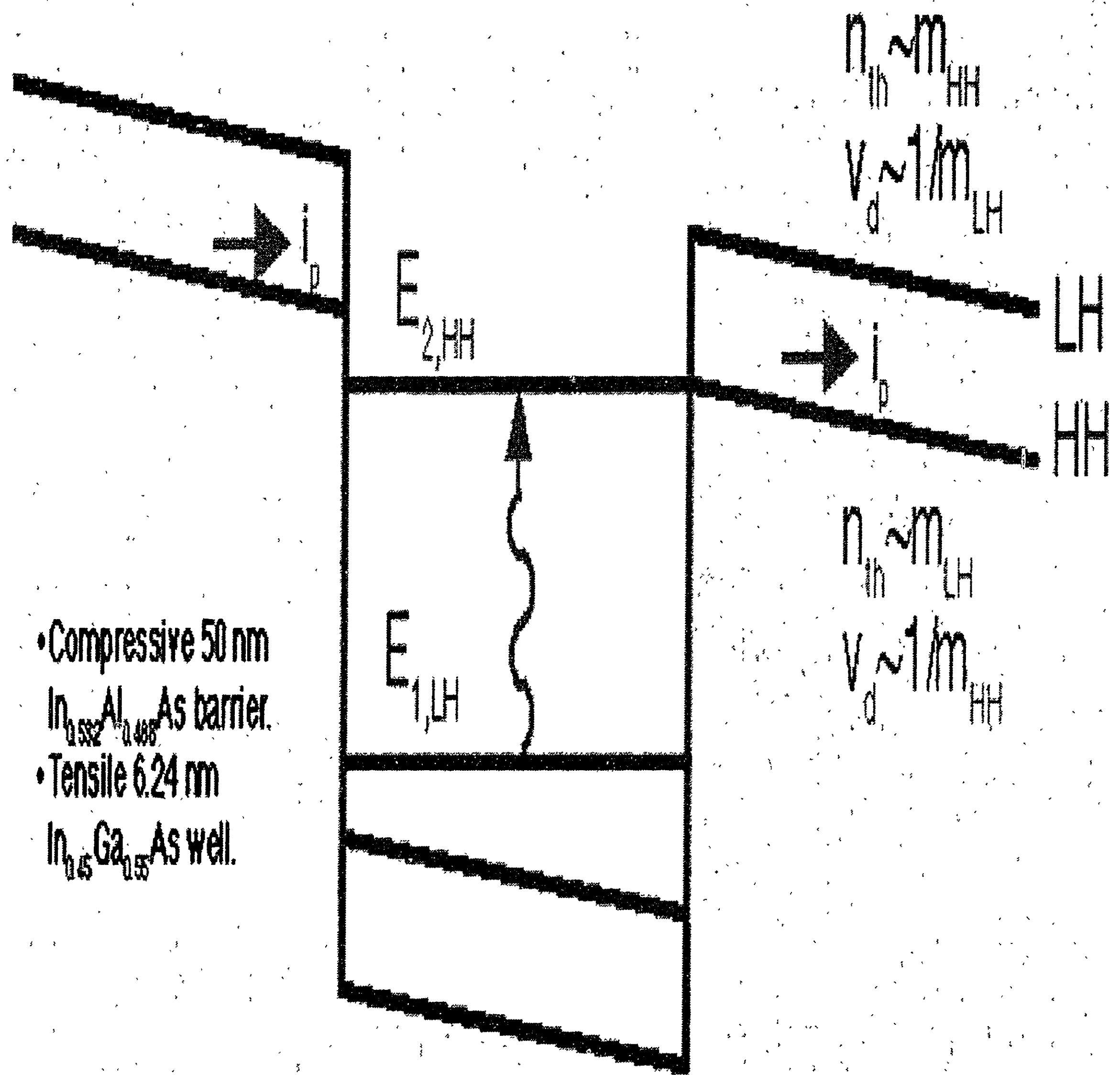


Figure 9-3: "Design2," consisting of 15 periods of 50 nm compressive $In_{0.532}Al_{0.468}As$ barriers and 6.24 nm tensile $In_{0.45}Ga_{0.55}As$ wells, is expected to have large quantum efficiency and small photoconductive gain.

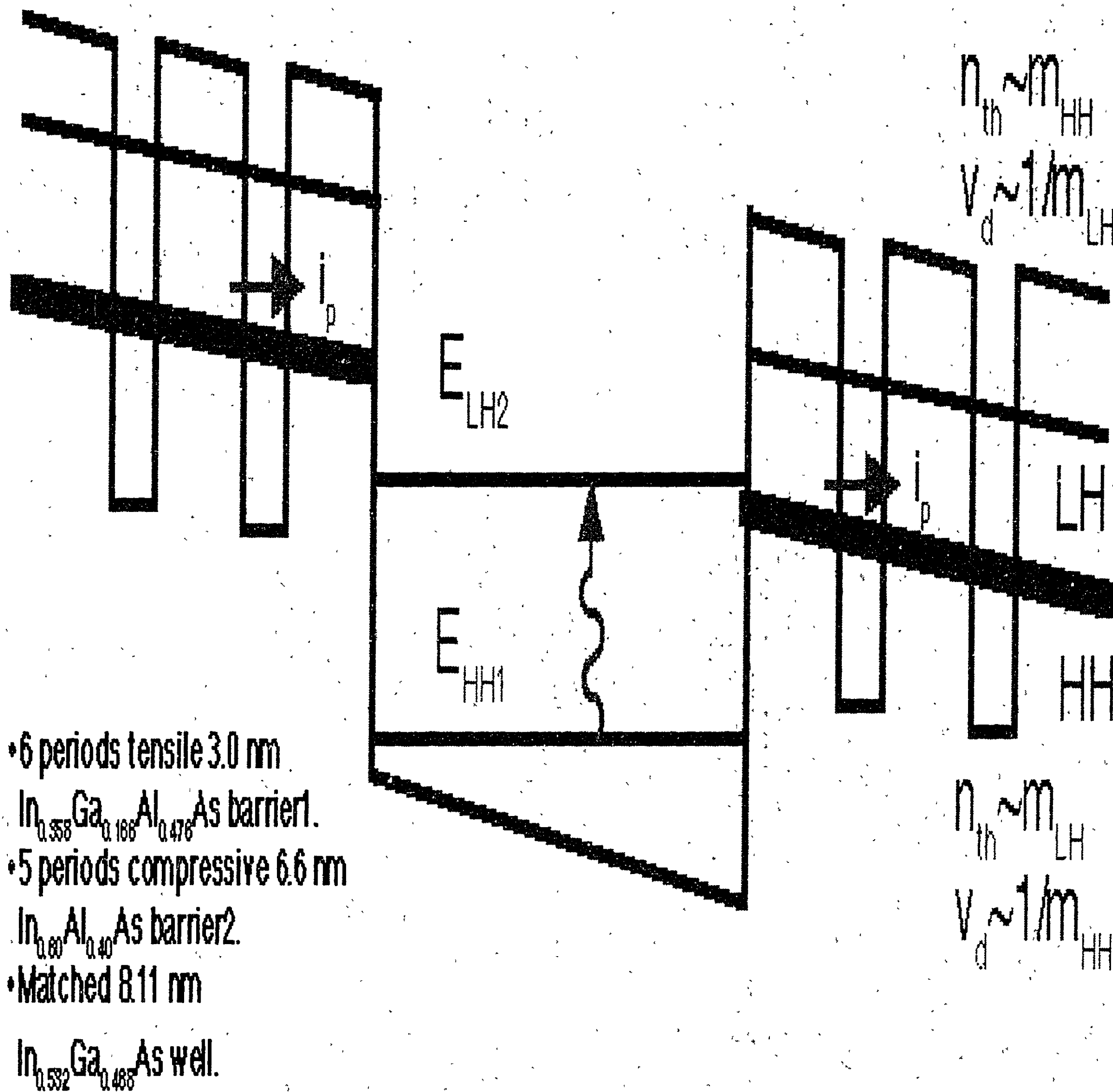


Figure 9-4: "Design3," consisting of 15 periods of 8.11 nm lattice matched $In_{0.532}Ga_{0.468}As$ wells and a barrier consisting of five periods of 3.0 nm tensile $In_{0.358}Ga_{0.166}Al_{0.476}As$ barrier1 and 6.6 nm compressive $In_{0.60}Al_{0.40}As$ barrier2, is expected to have large quantum efficiency, small photoconductive gain, and small leakage current.

9.2.2 Some p-QWIP Measurements

Each of the three designs discussed in the previous section were grown at least twice, each with X-ray rocking curves which showed very good lattice matching. (The rocking curves showed that the average epilayer peak was within 100 arcsec of the substrate peak.) After processing the wafers, the devices were measured at cryogenic temperatures.

The samples which were grown after August 1996 were all found to be short circuits (for 100% of the pixels measured). The samples all appeared very, very "dusty" even to the naked eye. Under a Nomarski microscope, all of these samples showed a large number of Group III droplets on the sample surfaces. Samples which were comprised of much more InAlAs than InGaAs showed far fewer Group III droplets on the sample surfaces. It is believed that the Group III droplets are Ga droplets which originated from the lip of the Ga crucible.

Many Ga droplets accumulated on the lip of the Ga crucible after an accident in August 1996 in which the Eurotherm temperature controller for the Ga cell failed in the middle of the growth and had to be quickly replaced by an available Eurotherm. This temperature controller replacement was accomplished as quickly as possible, but not before the Ga cell temperature had dropped and then quickly returned to the growth temperature of 880 °C. Since the Eurotherm had failed, it is not known how many degrees the Ga cell temperature had plummeted before the Eurotherm replacement was completed. The quick rise in temperature of the Ga cell following the replacement of the Eurotherm probably resulted in a large increase in the number of Ga droplets on the lip of the Ga crucible.

Subsequent replacement of the RIBER Ga cell with a hot-lipped Ga cell and a hot-lipped Ga "Sumo" cell, both of which were purchased from EPI in St. Paul, MN, did not solve the problem of the large number of Ga droplets on the wafer surfaces right after growth. (Samples which were grown with very little Ga still showed far fewer Group III droplets on the sample surfaces.) It is believed that the hot lips on both Ga cells purchased from EPI were not hot enough. EPI has since changed the

designs of both hot-lipped Ga cells.

The samples grown before August 1996 were processed into p-QWIP devices. Many of the p-QWIP pixels measured on these wafers were also short circuits. About 20% of the pixels measured did show sufficiently low leakage current to warrant continued measurement. This yield of about 20% for the p-QWIPs was much smaller than the yield of about 80% on most of the n-QWIPs measured. In our experience, when the yield was very low for a particular wafer, the "working" pixels still showed a leakage current which was too large. This is now discussed.

Figures 9-5 and 9-6 show the measured leakage currents as a function of bias voltage for 50 μm pixels on Samples 9331 ("Design 3") and 9326 ("Design 1") at QWIP operating temperatures between 10K and 240K. Samples 9331 and 9326 showed BLIP (Background LImited Performance) temperatures of 45K and 52K, respectively, for a 90° field of view of a 293K background through a ZnSe viewport.

To investigate the size of the leakage currents shown in Figs. 9-5 and 9-6, the leakage currents in these two figures were fit to the thermionic form in Equation (8.2) of Chapter 8. It was found that for both samples, the leakage current rises slowly for temperatures below 90K and then rises sharply above 90K. For both samples, the leakage current was found to be larger for negative biases of the top contact, in which case the holes travel toward the sample surface. This larger leakage for negative bias is believed [7] to result from dopant assisted hole tunneling, and the fact that dopants diffuse towards the sample surface during growth. To remove the complication of analyzing the dopant assisted tunneling, we will consider the p-QWIP leakage currents for positive biases only.

Fig. 9-7 shows the voltage dependence of the leakage current activation energy Φ_B, defined in Equation (8.3), for QWIP 9331 for temperatures between 90K and 200K. The best fit of the leakage current to Equation (8.2) yields an activation energy for positive biases of,

$$\Phi_B = 122.0 \text{ meV} - (10.7 \text{ meV/V}) \times V. \tag{9.16}$$

This measured value of 122 meV at V=0 volts is a bit smaller than the expected value

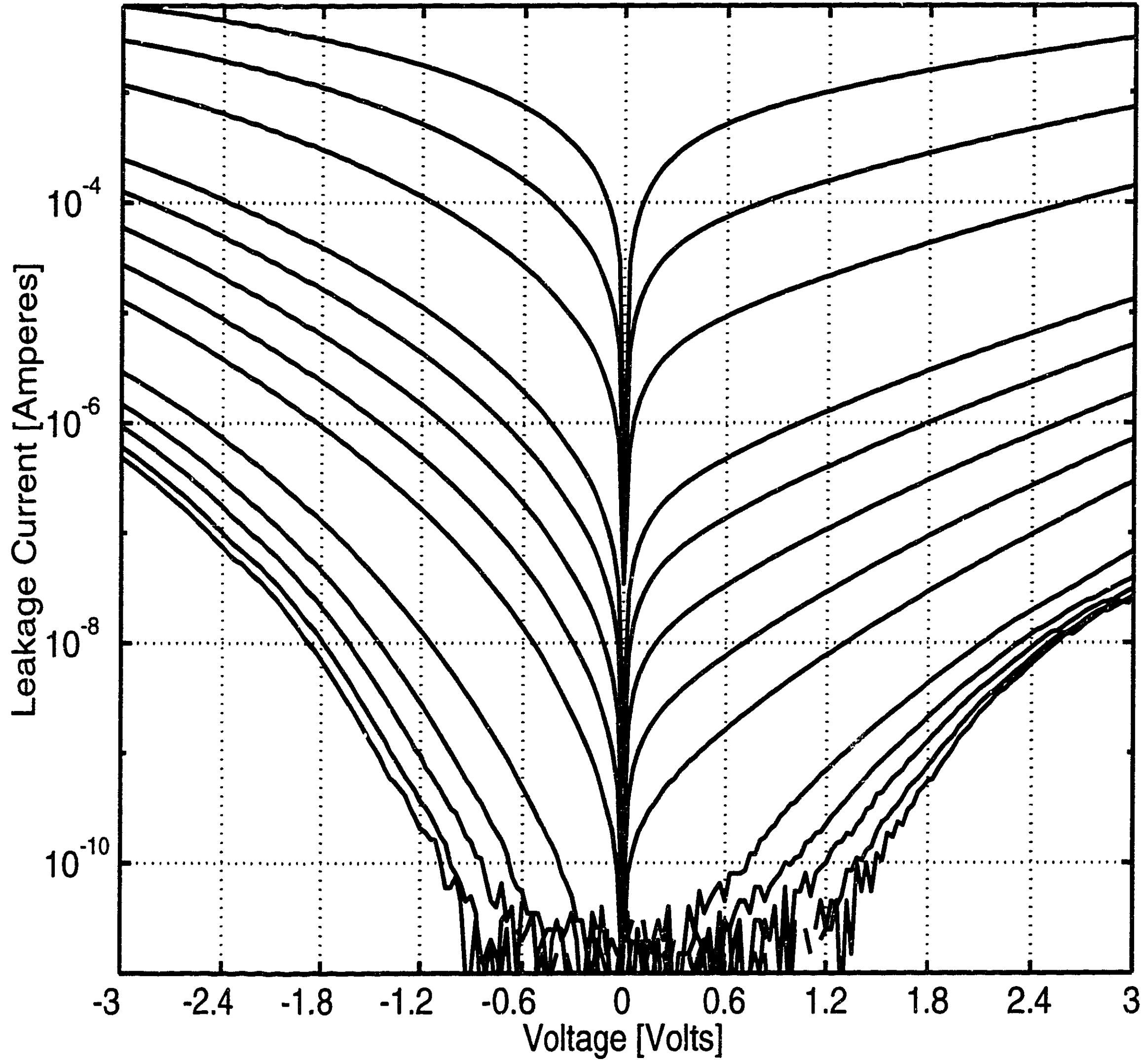


Figure 9-5: The measured leakage current as a function of bias voltage for a 50 μm pixel on Sample 9331 at QWIP operating temperatures of 9K, 20K, 30K, 40K, 50K, 60K, 70K, 80K, 90K, 100K, 130K, 160K, 200K (labeled, respectively, from the lowermost the uppermost curve in the figure).

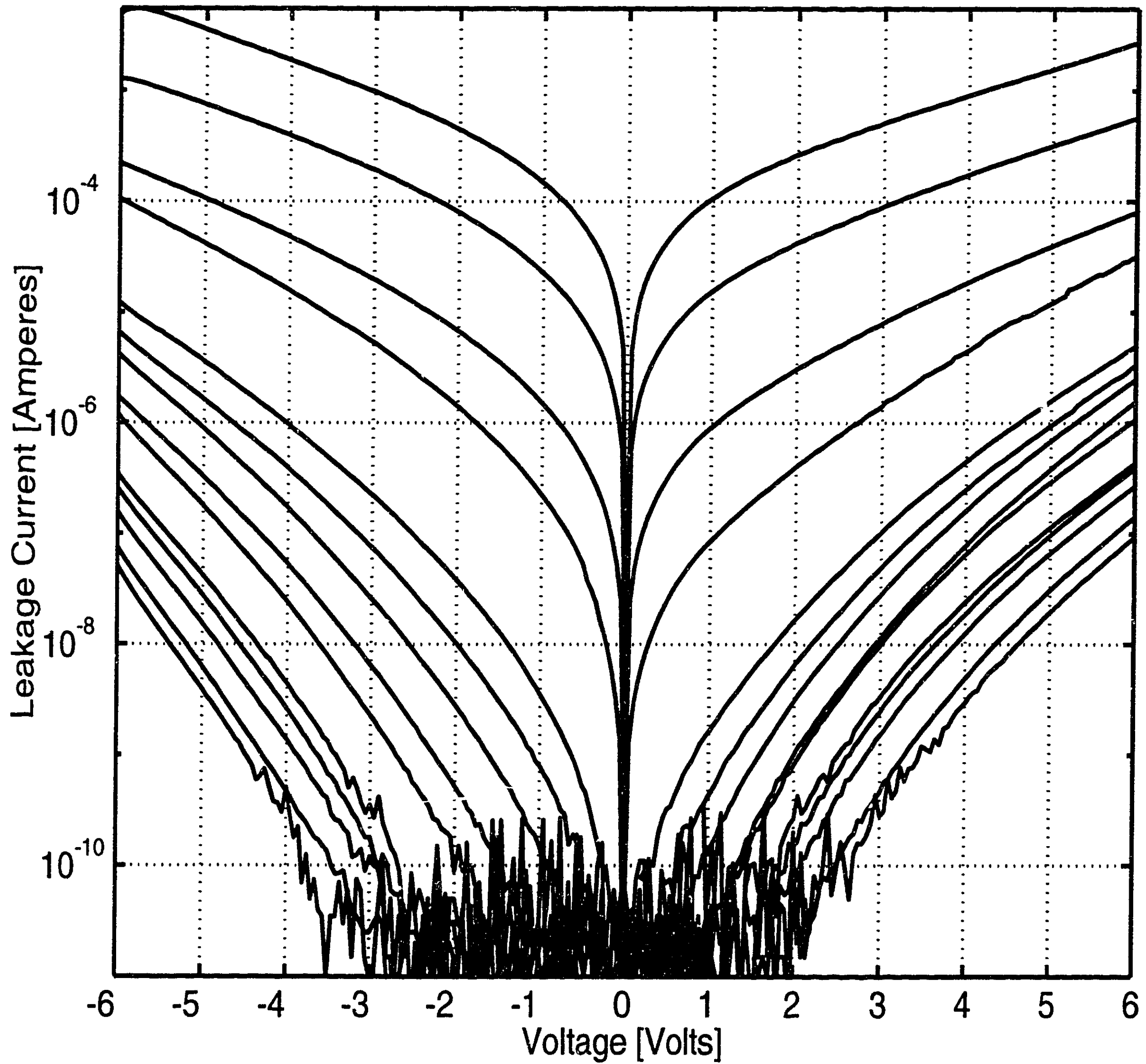


Figure 9-6: The measured leakage current as a function of bias voltage for a 50 μm pixel on Sample 9326 at QWIP operating temperatures of 10K, 20K, 30K, 40K, 50K, 60K, 70K, 80K, 90K, 100K, 130K, 160K, 200K, 240K (labeled, respectively, from the lowermost the uppermost curve in the figure).

of $V_B - E_F$, which was designed to be 155 meV. The measured value of 10.7 meV/V for the slope of the activation energy is a bit larger than the expected design value of qL_W/L_{TOT}=8.7 meV/V. The 20% discrepancy between both of the measured values and the designed values could result from uncertainties in the expected values of the band parameters, as well as large tunneling through the superlattice barrier , as discussed in Section 9.1.1 above.

Fig. 9-8 shows the voltage dependence of the leakage current activation energy Φ_B for QWIP 9331 for temperatures between 60K and 100K. The best fit of the leakage current to Equation (8.2) yields an activation energy for positive biases of,

$$\Phi_B = 80.4 \text{ meV} - (12.2 \text{ meV/V}) \times V. \tag{9.17}$$

This measured value of 80.4 meV at V=0 volts is puzzling, as it is much smaller than the expected value of $V_B - E_F$, which was designed to be 155 meV. This very small value (80.4 meV) of the activation energy could be a result of the small tunneling barrier calculated in Equation (9.12) above. Indeed, 80.4 meV is about the size of the average value of the band edges (relative to the quantum well valence band edge) of the semiconductors comprising the superlattice barrier. The measured value of 12.2 meV/V for the slope of the activation energy is a bit larger than the expected design value of $qL_W/L_{TOT} = 8.7\ meV/V$. This discrepancy could result from uncertainties in the expected values of the band parameters, but is probably a result of the large tunneling through the superlattice barrier even for energies below that of the superlattice miniband, as discussed in Section 9.1.1 above.

Fig. 9-9 shows the voltage dependence of the leakage current activation energy Φ_B for QWIP 9326 for temperatures between 90K and 240K. The best fit of the leakage current to Equation (8.2) yields an activation energy for positive biases of,

$$\Phi_B = 190.5 \text{ meV} - (15.3 \text{ meV/V}) \times V. \tag{9.18}$$

This measured value of 190.5 meV at V=0 volts is much larger than the expected value of $V_B - E_F$, which was designed to be 155 meV. This discrepancy may be a result of large uncertainties in the valence band lineup or effective masses in InAlAs/InGaAs, or of the strain induced valence band splitting. The measured value of 15.3 meV/V

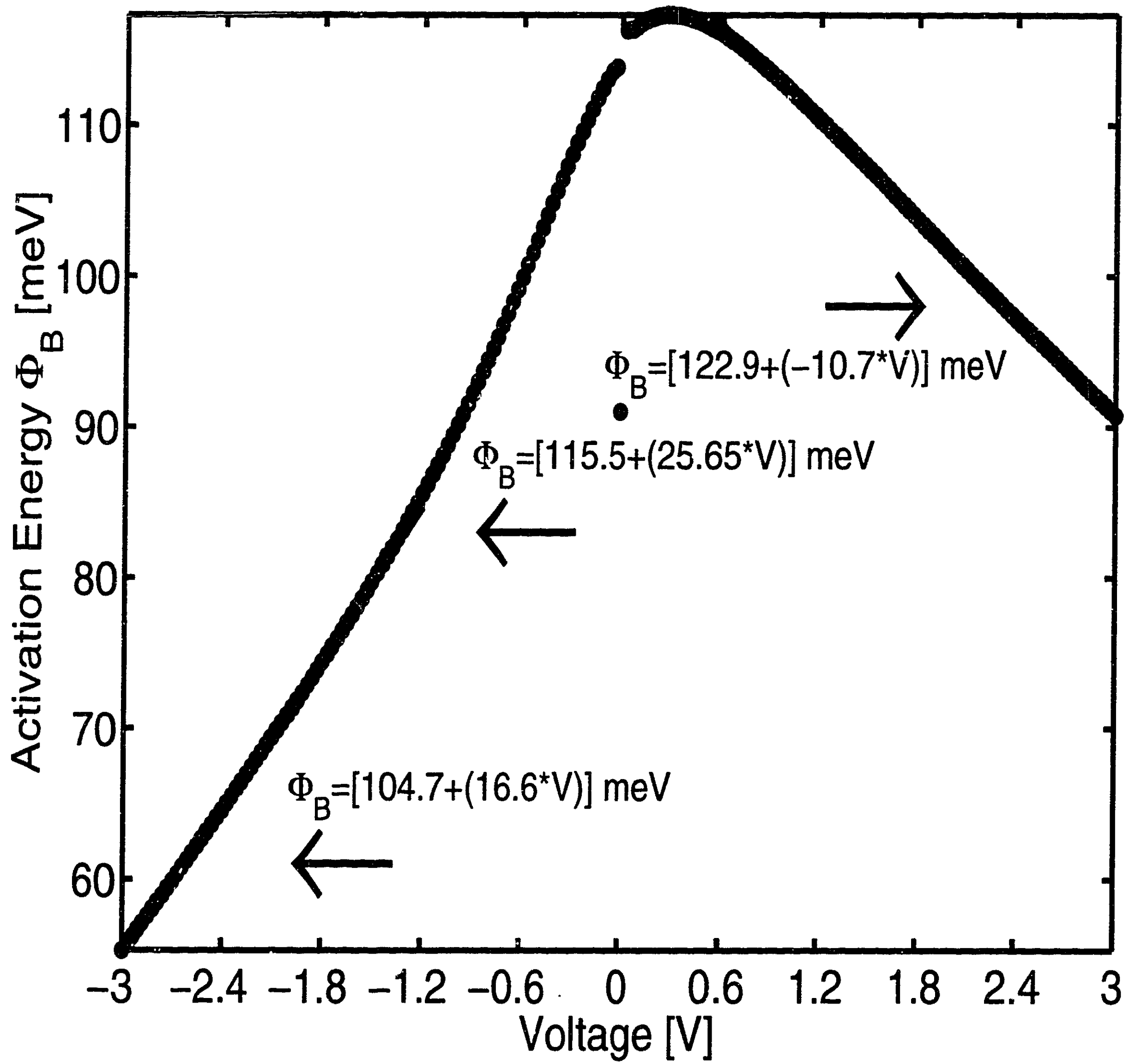


Figure 9-7: The voltage dependence of the leakage current activation energy for QWIP 9331 for temperatures between 90K and 200K. The activation energy was obtained by fitting the measured leakage current to the thermionic form in Equation (8.2) in Chapter 8 for temperatures between 90K and 200K.

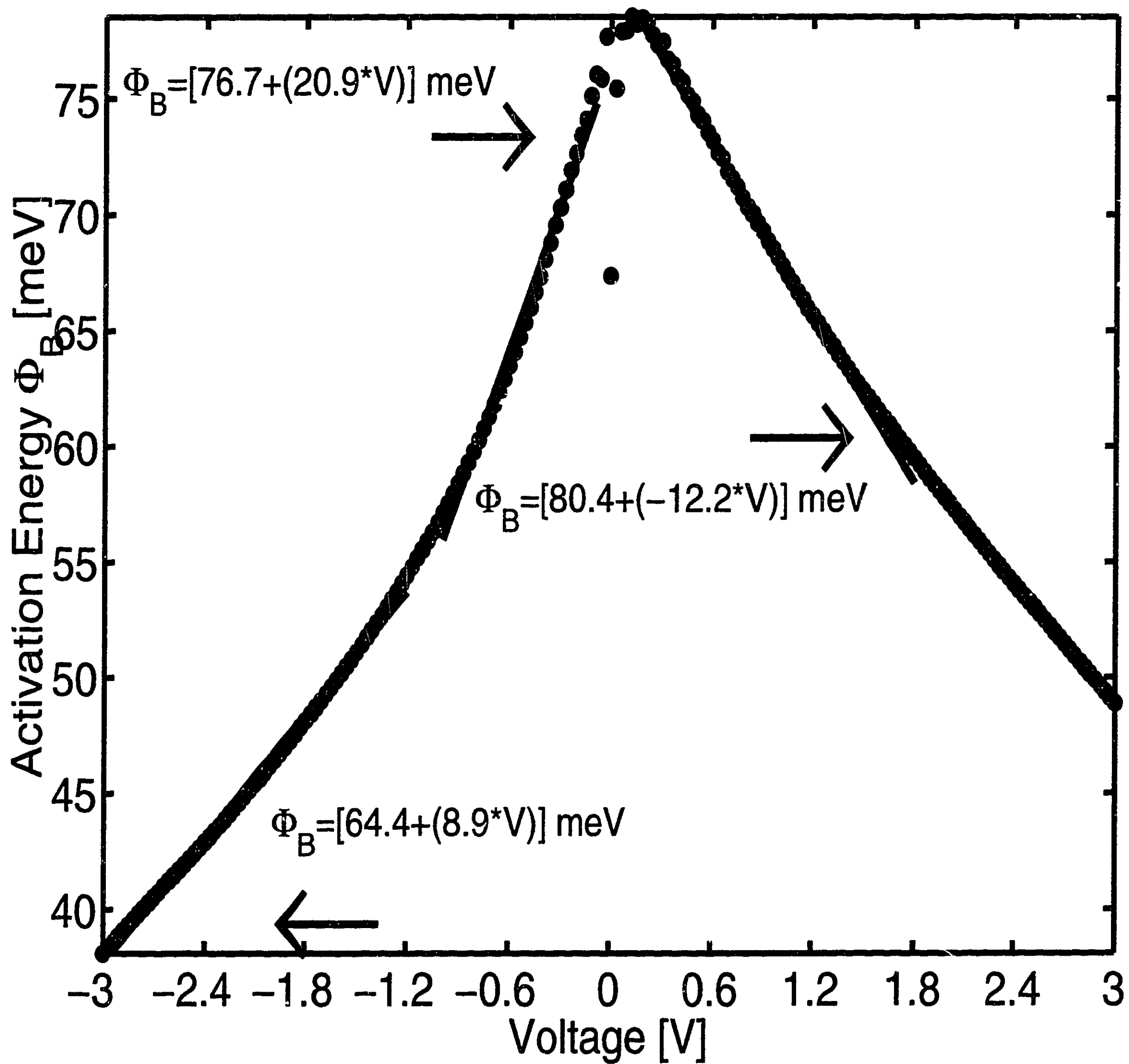


Figure 9-8: The voltage dependence of the leakage current activation energy for QWIP 9331 for temperatures between 60K and 100K. The activation energy was obtained by fitting the measured leakage current to the thermionic form in Equation (8.2) in Chapter 8 for temperatures between 60K and 100K.

is a bit larger than the expected design value of qL_W/L_{TOT}=10.6 meV/V. The discrepancy could result from a tunneling through the barriers resulting from impurities which are unaccounted for. Williams [96] has also observed much larger tunneling through a 500 Å barrier for his n-QWIP than was expected.

Fig. 9-10 shows the voltage dependence of the leakage current activation energy Φ_B for QWIP 9326 for temperatures between 70K and 100K. The best fit of the leakage current to Equation (8.2) yields an activation energy for positive biases of,

$$\Phi_B = 70.5 \text{ meV} - (9.0 \text{ meV/V}) \times V. \tag{9.19}$$

This measured value of 70.5 meV for the extrapolated activation energy at zero bias is very small, and cannot be accounted for by any obvious energy in the system. However, this measured activation energy should not be taken too seriously, as the value of Φ_B for small positive biases (below 1.2 volts in Fig. 9-10) appears very different, in both slope and intercept, from Equation (9.19). The dominant leakage current mechanism seems to change at about 1.2 volts in Fig. 9-10. There appears to be an unknown leakage mechanism in Sample 9326 which depends on both QWIP bias voltage and operating temperature, and which is not accounted for in the simple thermionic leakage model of Equation (8.2) in Chapter 8.

In summary, at temperatures above 90K, Sample 9331 seems to have a leakage current which can be accounted for within the framework of the thermionic leakage model of Equation (8.2). Both the measured activation energy of $V_B - E_F$ and the measured derivative of the activation energy with the applied bias of $(-qL_W/L_{TOT})$ were in reasonable agreement with the expected design values for Sample 9331 at temperatures above 90K. At temperatures between 60K and 90K, the measured leakage currents for Sample 9331 showed an activation energy which was too small (about 80 meV). This activation energy could be explained by the lowered tunneling barrier, as seen by bound carriers in the quantum well, commensurate with a superlattice described in the Kronig-Penney model (see Equations (9.11) and (9.12) in Section 9.1.1). At temperatures above 90K, Sample 9326 could also be fit to the thermionic leakage model of Equation (8.2), but with a measured value of 190 meV for $V_B - E_F$, which is a little larger than expected. This could be a result of uncertainties in the band lineup or strain induced energy splitting. At temperatures between 60K and 100K, Sample

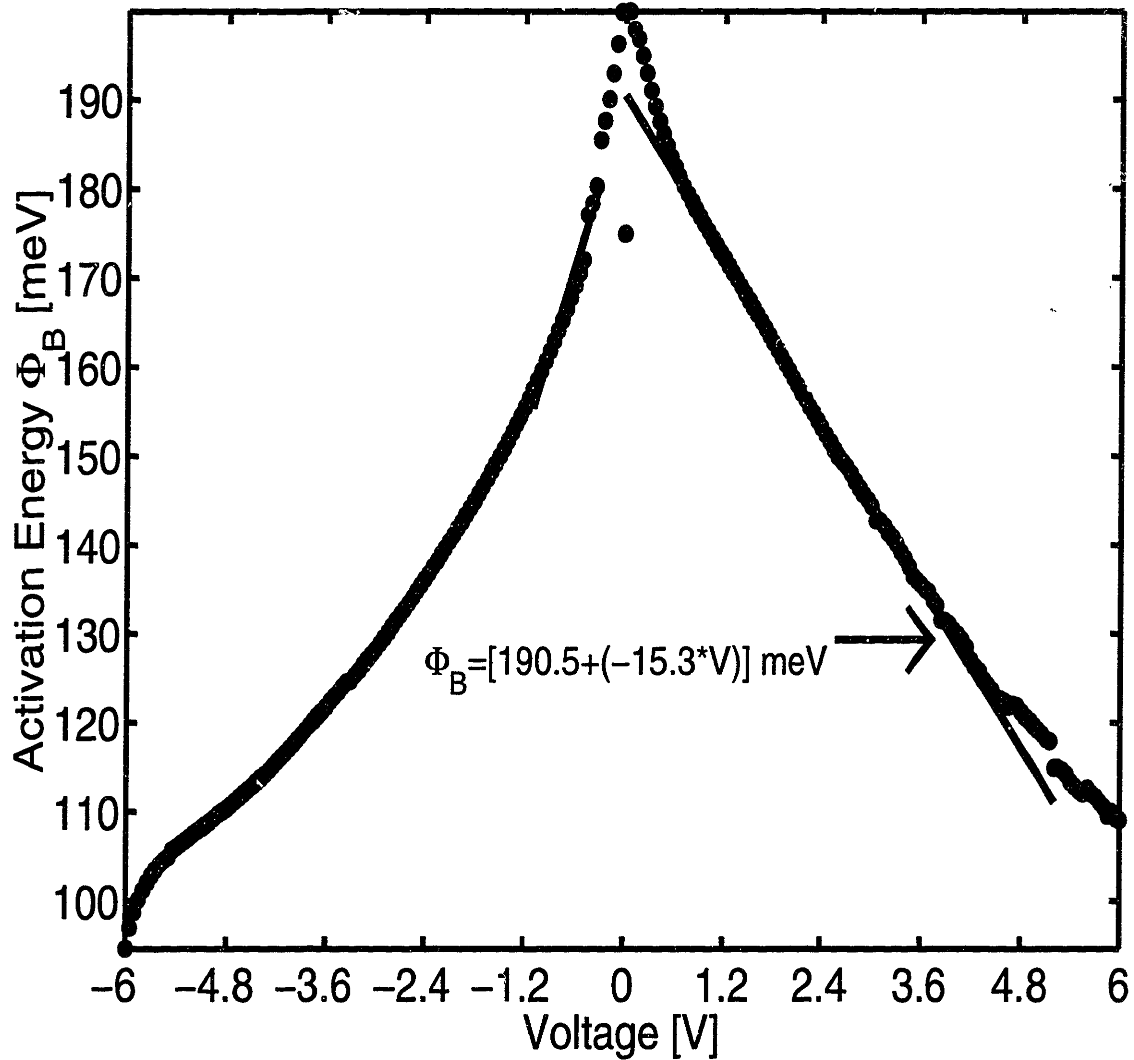


Figure 9-9: The voltage dependence of the leakage current activation energy for QWIP 9326 for temperatures between 90K and 240K. The activation energy was obtained by fitting the measured leakage current to the thermionic form in Equation (8.2) in Chapter 8 for temperatures between 90K and 200K.

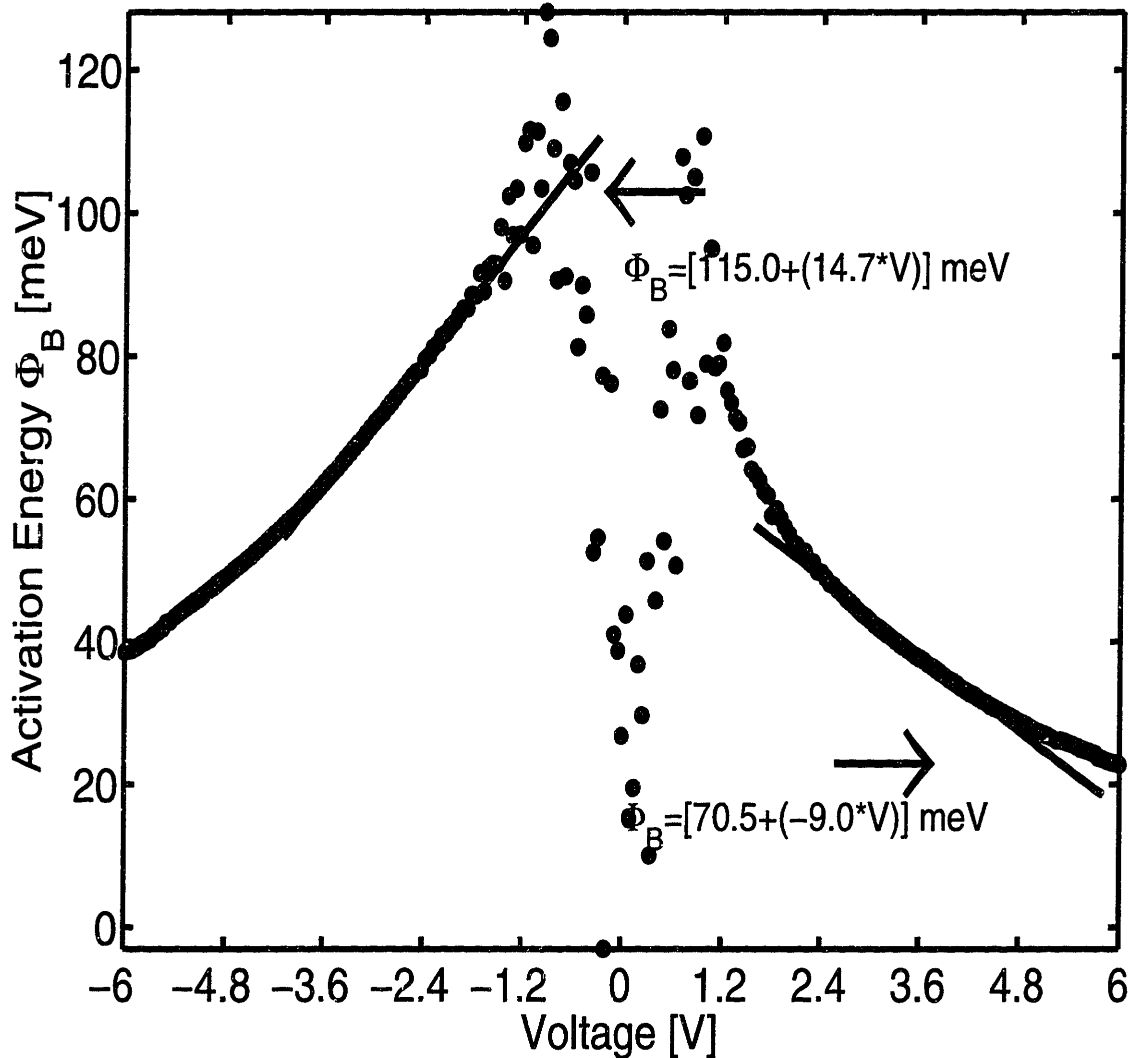


Figure 9-10: The voltage dependence of the leakage current activation energy for QWIP 9326 for temperatures between 60K and 100K. The activation energy was obtained by fitting the measured leakage current to the thermionic form in Equation (8.2) in Chapter 8 for temperatures between 60K and 100K.

9326 was fit to the thermionic leakage model with an inferred activation energy of 70 meV. This 70 meV activation energy is not understood at this time, as no states are believed to exist within the barrier which are 70 meV above the bound quantum well state.

It is difficult to say whether the growth or the processing can account for the large leakage current in Sample 9326. Both samples 9331 and 9326 were grown in the same week, and both samples were processed together. Both samples showed a relatively low processing yield of about 20%. Many devices showed extremely low resistances at the end of the processing. This yield of 20% is troubling. The n-QWIPs which showed significant detectivity all showed a high yield of at least 80%. The n-QWIPs grown after August 1996 (after the accident with Eurotherm temperature controller using either the RIBER cell or the hot lip EPI Ga or Sumo EPI Ga cell) all showed a much smaller yield of about 20%. Of more than 150 n-QWIP pixels processed on wafers grown on 15 separate occasions after August 1996, *only one single* device showed significant detectivity, though about 20% of these devices showed leakage currents which were small enough to warrant a closer study of the leakage current. Some of these n-QWIPs had the same design as those measured in Chapter 8, but they showed insignificant detectivity. It appeared as though the small processing yield of about 20% for both the n-QWIPs and the p-QWIPs came at about the same time that n-QWIP detectivities were unexpectedly low.

9.3 Conclusions

Some common n-QWIP designs used in industry have been evaluated. Some new p-QWIP designs are presented and measured.

First, we discussed the most common n-QWIP design, in which the barrier is made of one material, usually AlGaAs and usually of 500 Å width. Second, the commonly used n-QWIP design in which the confinement barriers are comprised of a semiconductor superlattice is considered. This QWIP design is intended to reduce thermionic leakage by pushing the three-dimensional continuum of energy further up in energy by making

the miniband transport through the superlattice barrier the means of photocurrent conduction.

A Kronig-Penney model presented in this chapter showed that this QWIP design, with a superlattice comprising the QWIP barriers, is expected to have a larger tunneling leakage than QWIPs whose confinement barriers are comprised of a single semiconductor material. The reason is that the tunneling through the superlattice barrier is, at best, commensurate with tunneling through a QWIP barrier comprised of a single semiconductor material but whose band edge is the average value of the band edges of the semiconductors comprising the actual barrier superlattice.

An experimental investigation was carried out in this work for p-QWIPs which were designed for reduced thermionic leakage and also for large absorption of normally incident radiation without the use of an optical grating. These p-QWIP designs have an additional new feature in that strain is used to lift the light and heavy hole valence degeneracy in the energy continuum, so that the lowest energy hole band in the continuum of energy states can be chosen as light-hole-like for both large photoconductive gain and large quantum efficiency. Electrical measurement of these p-QWIPs showed excessive leakage current, some of which may be accounted for within the framework of the Kronig-Penney model.

Chapter 10

Physical Models

10.1 Introduction

The device model for the quantum well infrared photodetector (QWIP) is usually taken to be that for a standard photoconductor, and thus is described by two parameters: the photoconductive gain and the quantum efficiency. In these models, both the photoconductive gain and the internal electric field are taken to be a constant throughout the entire structure of the device, as shown in Figure 10-1. This photoconductive gain is equal to the ratio of the drift velocity to the quantum well capture velocity, or equivalently the ratio of the capture time into a quantum well to the transport time through the entire QWIP structure.

Recent numerical models [112, 113, 114, 115], which account for the photoexcited, tunneling, and drift currents, show that there is an electric field inhomogeneity in the QWIP structure. These numerical models are an extension of the numerical models [116, 117, 118] which were used to explain experimental observations of the responsivity and gain in QWIPs having only a single quantum well. Key to understanding these QWIPs having only a single quantum well was an understanding of the charge distribution and nonuniform potential distribution within these devices under different bias and illumination conditions. This electric field inhomogeneity is distinct

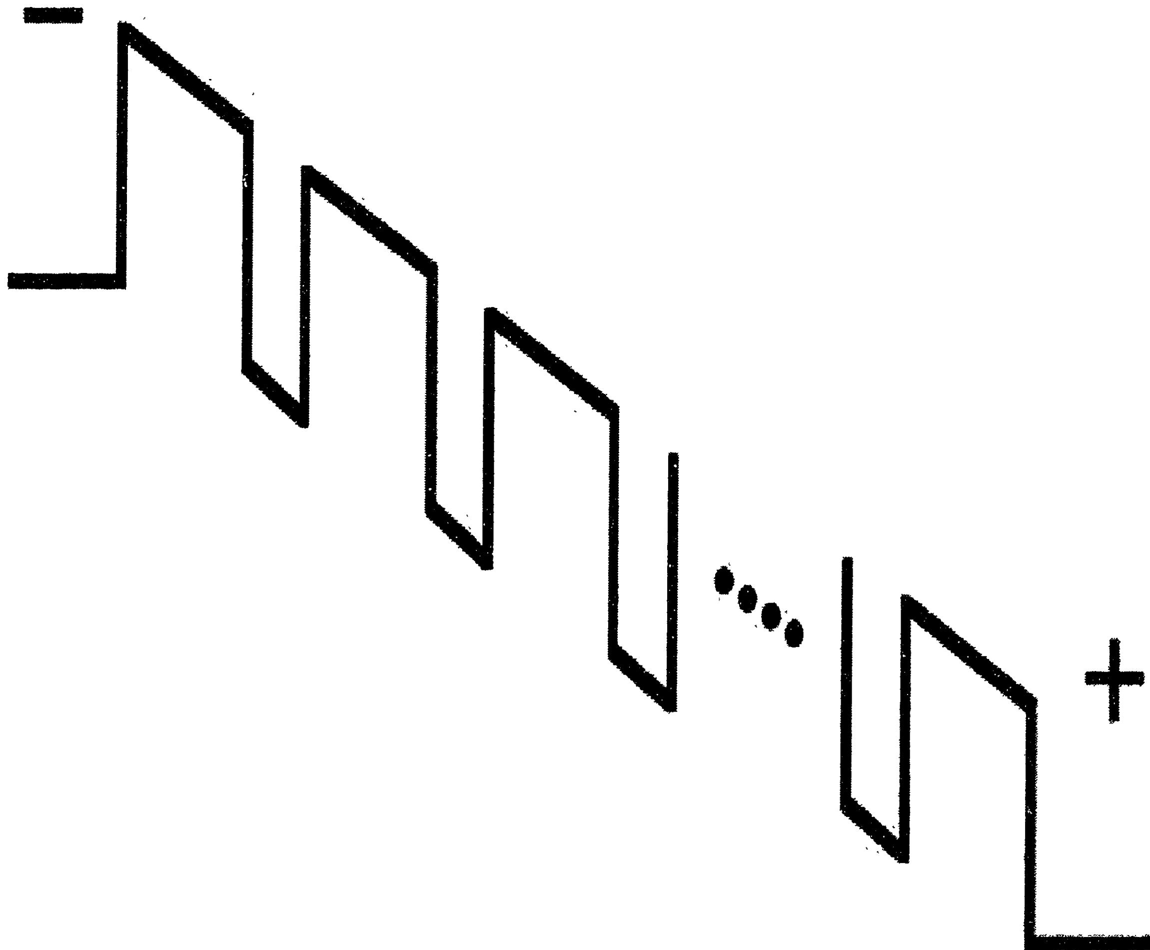


Figure 10-1: In the commonly used device model for a QWIP, the electric field across each period of the structure is the same.

from the kind of high field domains that form at low temperature (below 10K) as a result of resonant tunneling from the ground state in one well to an upper state in an adjacent well. (See Fig. 8-2 in Chapter 8.) Rather, this electric field inhomogeneity occurs at the more usual QWIP operating temperatures (40K or higher), and is the result of inadequate injection of carriers from the emitter contact.

The purpose of this chapter is to develop a numerically accurate physical model which explains the electric field inhomogeneity that exists in a QWIP through which current is flowing. Solution of the equations describing this physical model is not computationally intensive, and can be done on a simple calculator. This is described in section 10.3.

We also review three existing models which are commonly used to describe current flow through a QWIP: Levine's original model (section 10.2.1), Rosencher's model (section 10.2.4), and Ershov's rate equation model (section 10.2.5). We compare the different models. We formulate a new rate equation model which is a combination of Ershov's and Rosencher's models, and reduces to Levine's model under appropriate circumstances.

10.2 Review of Existing QWIP Models

10.2.1 Levine Model

10.2.2 Levine Photocurrent Model

Key points, for later discussion, in the Levine model of QWIP photocurrent are the photoconductive gain g and the excited carrier escape probability p_e. This is now reviewed.

Following Levine and Beck [7], the photoconductive gain is now derived. (A similar result was obtained by Liu [97].) The contribution to the photocurrent from

photoexcitation out of *one* quantum well is

$$i_p(1) = q\Phi\eta_1, \tag{10.1}$$

where Φ is the incident photon flux (in $\sec^{-1}$), and η_1 is the one quantum well quantum efficiency. In the steady state, a fraction p_c of the total photocurrent I_P is captured by any one particular quantum well, but each single quantum well will also contribute $i_p(1)$ to the total photocurrent,

$$I_P = I_P - p_c I_P + i_p(1) \tag{10.2}$$

so that the total photocurrent I_P is related to the single quantum well contribution $i_p(1)$ through

$$I_P = \frac{i_p(1)}{p_c} \equiv g_1\, i_p(1) \tag{10.3}$$

where the capture probability p_c is related to the single quantum well photoconductive gain g_1 through,

$$p_c = \frac{1}{g_1} = \frac{L_p}{\tau_L v_d} = \frac{\tau_{T1}}{\tau_L}, \tag{10.4}$$

where $L_p, \tau_L, v_d, \tau_{T1}$ are, respectively, the multiple quantum well (MQW) superlattice period, the lifetime and the drift velocity of the carriers in the upper state, and the transport time through one period of the QWIP. The total photoconductive gain [7] g is $1/N_W$ times the single quantum well photoconductive gain g_1, where N_W is the total number of quantum wells in the entire QWIP structure,

$$g = \frac{g_1}{N_W}. \tag{10.5}$$

Another feature of the Levine QWIP photocurrent model is that the quantum efficiency η is not exactly the same as the optical absorption efficiency η_a. Rather, Levine [7] finds that in order to obtain quantitative agreement between theory and experiment, the quantum efficiency η must be equal to the product of the optical absorption efficiency η_a and a phenomenological escape probability p_e,

$$\eta = \eta_a p_e, \tag{10.6}$$

where the optical absorption efficiency is

$$\eta_a = (1 - \exp(-\eta_1 N_W)), \tag{10.7}$$

and where the phenomenological escape probability p_e is found to have the form

$$p_e = \frac{1}{1 + \frac{\tau_e}{\tau_r}}. \tag{10.8}$$

By measuring the noise to obtain g, and by measuring the absorption to get η_a, and by measuring the photocurrent to get η, Levine finds the following empirical form for $\frac{\tau_e}{\tau_r}$,

$$\begin{aligned} \frac{\tau_e}{\tau_r} &= \left(\frac{\tau_e}{\tau_r}\right)_0 \exp(-V_{TOT}/V_0) \\ &\equiv \left(\frac{\tau_e}{\tau_r}\right)_0 \exp(-V_{1\text{period}}/V_p), \end{aligned} \tag{10.9}$$

where $V_{TOT}, V_0, V_{1\text{period}}, V_p$ and $\left(\frac{\tau_e}{\tau_r}\right)_0$ are, respectively, the total bias on the entire QWIP, N_W times the effective barrier lowering potential, the bias on one period of the QWIP, the effective barrier lowering potential, and the zero bias escape time ratio. Some measured values for these quantities for a variety of QWIPs measured by Levine [7] are shown in Table 10.1.

Rosencher et al. [119] obtained a similar dependence of the quantum efficiency η on the applied bias, but they interpreted this dependence in terms of a statistical distribution for the upper state in the intersubband transition.

A central point here is that the photoconductive gain g_1 is large [7], between 10 and 20, so that the total photocurrent I_P flowing through the device is much larger than the contribution $i_p(1)$ from any single quantum well. Another key idea is embodied in Equation (10.3), which says that in the steady state, the photoconductive gain can be obtained from requiring the continuity of the flux of carriers captured into a quantum well. Yet, another key point is that the quantum efficiency is the product of the optical absorption efficiency and a phenomenological escape probability p_e, where the phenomenological escape probability is found by fitting theory to experiment.

Levine Sample	L_W Å	x	Doping (10^{18}cm^{-3})	Doping Type	Periods	λ_p (μm)	Inter-sub-band Transi-tion	$\left(\frac{\tau_e}{\tau_r}\right)_0$	V_p (meV/ per-iod)
A	40	0.26	1	n	50	8.95	B-C	1.9	13
B	40	0.25	1.6	n	50	9.8	B-C	0.94	17
C	60	0.15	0.5	n	50	13.2	B-C	1.9	16
D	70	0.10	0.3	n	50	16.6	B-C	1.6	16

Table 10.1: Measured [7] zero bias escape time ratio, $\left(\frac{\tau_e}{\tau_r}\right)_0$, and effective barrier lowering potential per period, V_p, for four GaAs/Al$_x$Ga$_{1-x}$As QWIPs. L_W is the well width. The intersubband transitions are: bound-to-bound (B-B), bound-to-quasicontinuum (B-QC), and bound-to-continuum (B-C).

10.2.3 Levine Leakage Current Model

Levine [7] writes the leakage current as

$$J_L = qn_{3d}(V)v_d(V), \tag{10.10}$$

where v_d is the drift velocity,

$$v_d(V) = \frac{\mu F}{\sqrt{1 + \left(\frac{\mu F}{v_{sat}}\right)^2}}, \tag{10.11}$$

where μ, F, v_{sat} are, respectively, the mobility, the average field, and the saturated drift velocity, where n_{3d} is the effective number of carriers which are excited out of the well as a function of the bias voltage V,

$$n_{3d}(V) = \left(\frac{m_W}{\pi\hbar^2 L_p}\right)\int_{E_1}^{\infty} f(E)T(E,F)dE \tag{10.12}$$

where m_W, L_p, E_1, E_F are, respectively, the effective mass in the quantum well, the period of the QWIP multiple quantum well structure, the lowest bound state measured

from the quantum well band edge, and the two-dimensional Fermi level measured from E_1, and $f(E)$ is the Fermi-Dirac distribution,

$$f(E) = \frac{1}{1 + \exp(E - E_1 - E_F)/k_B T} \tag{10.13}$$

where the tunneling coefficient is, in the WKB approximation,

$$T(E,F) = \begin{cases} 1 \\ \text{for E} > V_B \\ \\ \exp\left(-\frac{4}{3qF}\left[\frac{2m_B}{\hbar^2}\right]^{1/2}[V_B - E]^{3/2}\right) \\ \text{for } V_B - qFL_B \le E < V_B \\ \\ \exp\left(-\frac{4}{3qF}\left[\frac{2m_B}{\hbar^2}\right]^{1/2}\left([V_B - E]^{3/2} - [V_B - E - qFL_B]^{3/2}\right)\right) \\ \text{for } E < V_B - qFL_B \end{cases} \tag{10.14}$$

where V_B is the potential barrier.

The energy appearing in the transmission coefficient $T(E,F)$ in Equation (10.12) is the total energy E of the bound carrier. The reasoning commonly given in the literature [7] is that in a real system, electron scattering causes the electron wave function to decay in the barriers according to the total energy E instead of the bound state E_1.

The values of mobility and the saturated drift velocity used in Equation (10.11) are obtained [7] by fitting theoretically modeled to experimentally measured leakage currents.

At a very low bias of 0.1 volts on a 49 periods of GaAs wells of width $L_W = 76$ Å and 50 periods of $Al_{0.27}Ga_{0.73}As$ barriers of width $L_W = 88$ Å, Levine finds a good fit of the sequential resonant tunneling (for $qFL_p < \hbar/\tau_1$, where τ_1 is the ground state scattering time) to

$$J_{srt} = \frac{qk_BT}{\hbar L_W^2} T(E_1, F) \ln\left(\frac{1 + \exp(E_F/k_BT)}{1 + \exp((E_F - qFL_B/k_BT)}\right), \tag{10.15}$$

which dominates the leakage current at low temperatures, and a good fit of the thermionic leakage current to,

$$J_{th} = \frac{q^2 m_W}{\pi \hbar^2} \frac{v_d}{L_W} (qFL_B) \exp\left[-(V_B - qFL_W - E_F - E_1)/k_B T\right] \tag{10.16}$$

A key point here is that the mobility and the saturated drift velocity appearing in Equation (10.12) are numerically adjusted to match modeled leakage currents to measured leakage currents. Another key point is that it is the total energy E which appears in the transmission coefficient $T(E, F)$ inside the integral in Equation (10.12) because it is believed [7] that in a real system, electron scattering is sufficient to cause the electron wave function to decay in the barriers according to the total energy E instead of the bound state E_1.

10.2.4 Numerical Models of Rosencher

A key point in Thibaudeau's [115] work (in Rosencher's group) is that the contribution $J_{1\text{well}}$ of *each* quantum well to the total leakage current or the total photocurrent is used to solve for the self-consistent field distribution throughout the QWIP in terms of the injected current. More specifically, the electric field over each period of the QWIP is found from,

$$p_c J_{inj} = J_{1\text{well}} \quad \text{at } \textit{each} \text{ well,} \tag{10.17}$$

where the contribution $J_{1\text{well}}$ from each quantum well is calculated from,

$$J_{1\text{well}} = \frac{q m_W k_B T}{2\pi^2 \hbar^3} \int_{E_1}^{\infty} T(E, F) \ln\left[\frac{1 + \exp\left(\frac{E_F - E}{k_B T}\right)}{1 + \exp\left(\frac{E_F - qFL_p - E}{k_B T}\right)}\right] dE$$

$$\text{(scattering as in 3D reservoir).} \tag{10.18}$$

The two natural logarithm terms in Equation (10.18) correspond to the transmitted current in both the forward direction (from well i to $i+1$) and the reverse direction (from well $i+1$ to i)). Equation (10.18) makes the significant assumption that the current emitted by a two dimensional electron reservoir can be modeled as that emitted by a three dimensional electron contact with an appropriate choice of Fermi level.

This is believed to be a valid assumption [119, 120] when there is sufficient scattering [121] of the electrons within the quantum well to make the quantum well equivalent to a contact as far as calculation of the tunneling is concerned.

An alternative assumption is to model the quantum well electrons as a two dimensional electron reservoir, as in

$$J_{1\text{well}} = \frac{qm_W}{\pi\hbar^2} \int_{E_1}^{\infty} \nu T(E,F) dE \left[\frac{1}{1+\exp\left(\frac{E-E_F}{k_B T}\right)} - \frac{1}{1+\exp\left(\frac{E+qFL_p-E_F}{k_B T}\right)} \right]$$
$$\text{(scattering as in 2D reservoir).} \tag{10.19}$$

Of significance is the impinging frequency ν [7, 115, 122] appearing in Equation (10.19),

$$\nu = \begin{cases} \frac{1}{L_W}\left(\frac{E_1}{2m_W}\right)^{1/2} \approx \frac{\hbar\pi}{2m_W L_W^2} & \text{(Oppenheimer-Bohr frequency of Rosencher)} \\ \frac{v_d}{L_p} & \text{(Levine).} \end{cases} \tag{10.20}$$

Rosencher's group [116, 122] use the Oppenheimer-Bohr oscillation frequency of an electron on a quantized level, and Levine uses the drift velocity divided by the period of the multiple quantum well structure. The use of the Oppenheimer-Bohr oscillation frequency of an electron on a quantized level yields a tunneling transmission which is very similar to theoretical values obtained from a more involved [122] phase-shift formalism. Levine [7] obtains a good fit of the measured and modeled leakage currents by varying the mobility and saturation velocities as free parameters.

Thibaudeau's [115] model has one adjustable physical parameter, the capture probability p_c, which is then fit to the functional form,

$$p_c(F) = u \exp[v/(F+w)], \tag{10.21}$$

where F is the electric field across a particular quantum well within the QWIP, and where the three free parameters, u, v, w, are adjusted until there is good agreement between the measured leakage current and the modeled leakage current.

Both Pelve et al. [95] and Martinet et al. [122] have reported a good fit of the data, respectively, from leakage current measurements and from transient-capacitance spectroscopy (on carefully designed single quantum wells) to the *voltage* dependence of

Equation (10.14) *evaluated at* $E = E_1$=bound state energy. This has important consequences on the use of Equations (10.18) and (10.19), since T(E,F) in both equations is evaluated at the total energy E and not $E = E_1$=bound state energy. Moreover, Martinet et al. [122] have reported that the Oppenheimer-Bohr frequency in Equation (10.20) for the impinging of quantum well carriers on the barrier may be two orders of magnitude larger than the impinging frequency inferred from transient-capacitance spectroscopy on a carefully designed single quantum well device. Both of these observations (the use of the total energy in T(E,F) in both Equations (10.18) and(10.19), and the use of the Oppenheimer-Bohr frequency in Equation (10.20)) may explain why the measured tunneling currents are sometimes [123] smaller than predicted.

A key point in Rosencher's model is that the contribution of *each* quantum well to the total leakage current or the total photocurrent is used to solve for the self-consistent field distribution throughout the QWIP in terms of the injected current. An important difference between Levine's model and Rosencher's model for tunneling leakage is Rosencher's use of the Oppenheimer-Bohr impinging frequency for bound state carriers onto a barrier. An important feature of both Levine's and Rosencher's models is that the total tunneling leakage is calculated by assuming that the wave function decay within the barrier drops as the total energy rather than just the bound state energy of the carrier in the quantum well.

10.2.5 Numerical Models of Ershov

A complete numerical model of QWIPs would include the various radiative and non-radiative capture and absorption mechanisms into a quantum well, plus the differing amounts of drift and tunneling transport. One version of this has recently been done by Ershov [112, 113, 114]. In this model, we assume that the quantum wells have only one bound state, and that the upper state in the intersubband transition is in the continuum of energies above the barrier band edge.

The Poisson equation is written down to describe the concentration N_{2D} of bound

carriers remaining in each quantum well, in terms of the concentration N_D^+ of donor impurities, and the free carriers n_{3D} having an energy in the continuum above the barrier band edge,

$$\frac{d^2\phi}{dz^2} = -\frac{q}{\epsilon_B}\left[N_D^+ - N_{2D} - n_{3D}L_p\right], \tag{10.22}$$

where q is the magnitude of the electronic charge, and where ϕ is the electric potential throughout the QWIP. The continuity equation for the flow of carriers into the bound state in the i-th quantum well is is written as,

$$\frac{\partial N_{2D}(i)}{\partial t} = L_W\left[r^i - g_0^i - g_T^i\right], \tag{10.23}$$

and the continuity equation for the flow of carriers into the upper state from the i-th quantum well is written as,

$$\frac{\partial n_{3D}}{\partial t} = [G_0 + G_T - R] - \frac{1}{q}\nabla\cdot\vec{J}_{3D}, \tag{10.24}$$

where the total current density J_{3D} of carriers in the continuum above the barrier band edge,

$$J_{3D} = qn_{3D}\mu F + qD\frac{\partial n_{3D}}{\partial z}, \tag{10.25}$$

is the sum of the drift and diffusion components, where μ, F, D are, respectively, the mobility, the electric field and the diffusion constants within the barrier, where g_0, G_0, g_T, G_T, r, R are, respectively, appropriately normalized fluxes for the photoexcited carriers to the continuum, thermally excited carriers to the continuum, and thermally relaxing carriers from the continuum,

$$g_0 = \sigma I\, N_{2D}/L_W \qquad g_T = N_{2D,eq}/(\tau_e L_W) \qquad r = n_{3D}v_{QW}/L_W \tag{10.26}$$

$$G_0 = \sigma I\, N_{2D}/L_B \qquad G_T = N_{2D,eq}/(\tau_e L_B) \qquad R = n_{3D}v_{QW}/L_B, \tag{10.27}$$

where $\sigma, I, v_{QW}, \tau_e$ are, respectively, the photoexcitation cross section, the incident photon flux, the recombination velocity [119] from the continuum into each quantum well, and the thermal excitation time constant.

Ershov et al. [112, 113, 114] write detailed balance for the g and r coefficients as,

$$v_{QW}\tau_e = \frac{N_{2D,eq}}{n_{3D,eq}(E_{F,W,eq} - V_B)\exp(\Delta E_{F,W}/k_BT)} \tag{10.28}$$

where $\Delta E_{F,W}$ is the difference between the quasi-Fermi level $E_{F,W}$ and the Fermi level $E_{F,W,eq}$, in the presence of and in the absence of the applied bias, respectively, associated with the two dimensional electron gas in the quantum wells,

$$\Delta E_{F,W} = E_{F,W} - E_{F,W,eq} \tag{10.29}$$

where

$$N_{2D,eq} = \frac{m_W}{\pi \hbar^2} \ln\left[1 + \exp(E_{F,W,eq}/k_B T)\right], \tag{10.30}$$

$$n_{3D,eq}(E_{F,W,eq} - V_B) = 2\left(\frac{m^* k_B T}{2\pi\hbar^2}\right)^{3/2} \exp([E_{F,W,eq} - V_B]/k_B T). \tag{10.31}$$

The key points in the Ershov model are the use of the rate equations in Equation (10.23) and (10.24), the use of the Poisson equation in Equation (10.22), and the explicit appearance in the divergence term in Equation (10.24) of the current density J_{3D} of the carriers in the continuum. In principle, the use of the rate equations, together with detailed balance, should yield the same result as directly writing down the leakage currents and photocurrents, as was done by Levine [7]. The solution of the rate equations in the vicinity of *each* well should be equivalent to Thibaudeau's heuristic approach of calculating the flux into and out of *each* well, as in Equation (10.17). These two observations allow one to check the results of the Ershov model by comparing these results with those of the Levine and Rosencher models. The explicit appearance in the divergence term in Equation (10.24) of the current density J_{3D} of the carriers in the continuum, together with the explicit *absence* of a similar divergence term for the tunneling current from the bound state of one quantum well to an adjacent well in Equation (10.23), indicates that Ershov's model implicitly assumes that the total current through the QWIP comes mainly from J_{3D} and not from the tunneling current through the bound states in the quantum wells.

10.3 Analytical Expressions from a Physical Model

10.3.1 Insufficient Carrier Injection - Intuitive Picture

Before presenting the physical model developed in this work, we present our intuitive interpretation of the numerical work of Ershov [112, 113] and Thibaudeau [115].

Figure 10-1 shows the commonly used device model for a QWIP, where the electric field across each period of the structure is the same. QWIPs which are grown with the same barrier width and barrier composition in each period of the multiple quantum well structure cannot be described by the uniform field distribution shown in Fig. 10-1. The reason is that if J_1 is the current density supplied by one quantum well to the total current density, then the total current density is gJ_1, where g is either the dark current gain or the photoconductive gain, and this total current density gJ_1 is too large to be supplied by the emitter contact in the model shown in Fig. 10-1. (The photoconductive gain [7] has been measured to be on the order of 10 or 20, depending on the applied bias.) To see that this total current density gJ_1 is too large to be supplied by the emitter contact, one need only observe that within the model shown in Fig. 10-1, the barrier on the positive voltage side of each quantum well has the same height, the same width, and the same field across it as the barrier adjacent to the emitter. Thus, if each quantum well is contributing a current density of J_1 (through or above the barrier on the positive voltage side of it), then the emitter contact cannot be injecting the much larger current density of gJ_1 through or above an identical barrier.

Figure 10-2 shows a more realistic device model for a QWIP shows an inhomogeneous field distribution within the device. When a sufficiently large photocurrent is excited in the body of the QWIP, the quantum wells near the emitter contact must deplete in order to supply the necessary carriers. The depletion of carriers from those quantum wells closest to the emitter contact will supply the necessary photoexcited carriers by accomplishing two feats: this depletion near the emitter will lower the fraction of the applied voltage that drops over the main part (those quantum wells that are not too near the emitter contact) of the QWIP structure, thus lowering the fraction of

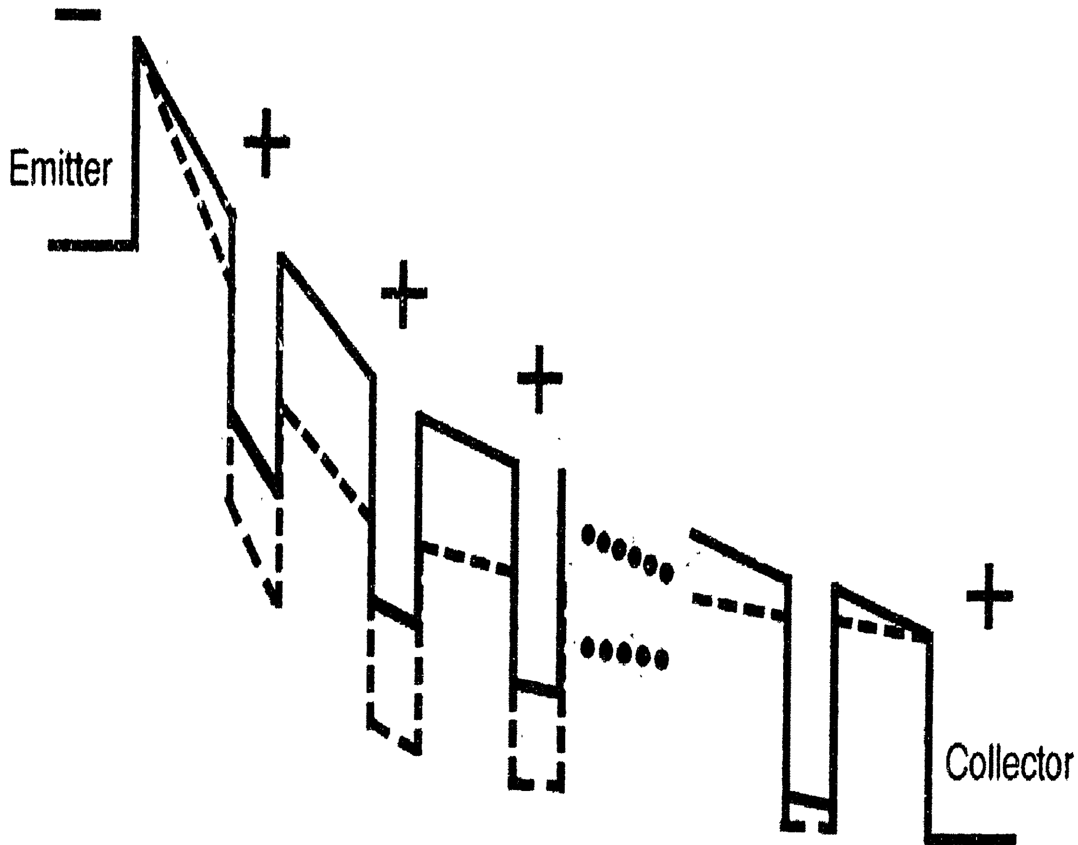


Figure 10-2: A more realistic device model for a QWIP shows an inhomogeneous field distribution within the device. This field distribution results from inadequate carrier injection from the emitter contact. In the steady state, current continuity and Gauss's law require that enough carriers deplete from the quantum wells near the emitter contact in order for the emitter to inject enough carriers to supply the photocurrent within the bulk of the QWIP. The energy band diagram for the QWIP under dark conditions and under illuminated conditions are indicated by the solid line and the dashed line, respectively, in the figure. For a fixed applied bias voltage, more carriers are depleted from the quantum wells near the emitter (to make the transmission larger through the barrier closest to the emitter) under illuminated conditions (where the emitter supplies both photocurrent and leakage current) than under dark conditions (where the emitter supplies only the leakage current).

the photoexcited current that reaches the collector contact; this depletion also lowers the barrier near the emitter which is seen by the tunneling injected carriers, thus supplying more carriers into the QWIP.

The energy band diagram for the QWIP under dark conditions and under illuminated conditions are indicated by the solid line and the dashed line, respectively, in the Fig. 10-2. For a fixed applied bias voltage, more carriers are depleted from the quantum wells near the emitter (to make the transmission larger through the barrier closest to the emitter) under illuminated conditions (where the emitter supplies both the photocurrent and leakage current) than under dark conditions (where the emitter supplies only the leakage current).

This depletion of the quantum wells near the emitter contact is undesirable, particularly if the depletion is already large under dark conditions. Under these conditions, the photoconductive gain is small, especially at small voltages. A larger voltage will increase the photoconductive gain, but it will also increase the leakage current. We believe that this is reason why the optimal QWIP design for the 8-12 micron wavelength of operation is usually found experimentally [7] to be one with 500 Å barriers. Use of a larger QWIP barrier increases the depletion of the quantum wells near the emitter barrier, and thus requires higher operating voltages to get an adequate photoconductive gain. At these higher operating voltages, the leakage currents are higher than in the optimal QWIP design.

10.3.2 Description of Physical Model

The purpose of this chapter is to develop a numerically accurate physical model which explains the electric field inhomogeneity that exists in a QWIP through which current is flowing. Solution of the equations describing this physical model is not computationally intensive, and can be done on a simple calculator.

In the device model developed in this work, it is found that the inhomogeneous field distribution in a QWIP is determined by Poisson's equation, current continuity, and the rate equations for the concentration of bound and free carriers in a QWIP. The

physical model developed in this work is very similar to Ershov's rate equation model. However, in the physical model developed in this work, a distinction is made between two types of tunneling currents: the tunneling between ground states in adjacent quantum wells, as well as tunneling from a ground state into an upper (continuum) state.

We start with Poisson's equation in Equation (10.22). We write the continuity equation for the flow of carriers into the bound state as,

$$\frac{\partial N_{2D}(i)}{\partial t} = [n_{3D}(i-1)v_{QW} - N_{2D}(i)\left(G_{opt}(i) + G_{tfa}(i) + G_{th}(i)\right)] - \frac{1}{q}L_p(i)\nabla \cdot \vec{J}_{1,\text{tunn}}, \tag{10.32}$$

and the continuity equation for the flow of carriers into the upper state (in the continuum) as,

$$\frac{\partial n_{3D,i}L_{p,i}}{\partial t} = [-n_{3D}(i-1)v_{QW}(i) + N_{2D,i}\left(G_{opt}(i) + G_{tfa}(i) + G_{th}(i)\right)] - \frac{1}{q}L_{p,i}\nabla \cdot \vec{J}_{3D}, \tag{10.33}$$

where the index i denotes the quantum well number, as measured by its distance from the emitter contact, where $J_{1\text{tunn}}, G_{opt}(i), G_{tfa}(i), G_{th}(i)$ are, respectively, the tunneling current from the ground state in one quantum well to the ground state in an adjacent well, the optical excitation rate, the net thermionic field assisted tunneling rate in the forward direction, and the thermal excitation rate (resulting from scattering from phonons and other electrons). (These G coefficients in our model are different from Ershov's G coefficients in Equation (10.27), as the latter are carrier excitation *fluxes* and not excitation rates.)

In Equations (10.32) and (10.33), the current flowing above the barrier band edge is,

$$J_{3D} = q\, n_{3D} v_d. \tag{10.34}$$

In our notation, detailed balance requires that the coefficients in Equations (10.32) and (10.33) satisfy,

$$\begin{aligned} \frac{G_{th}}{v_{QW}} &= \frac{n_{3D,eq}(E_{F,W,eq} - V_B - qFL_W)}{N_{2D,eq}} \\ &= \frac{1}{N_{2D,eq}} 2\left(\frac{m^* k_B T}{2\pi\hbar^2}\right)^{3/2} \exp([E_{F,W,eq} - V_B + qFL_W]/k_BT), \end{aligned} \tag{10.35}$$

where N_{2D} is given in Equation (10.30) and $n_{3D,eq}(E_{F,W,eq} - V_B - qFL_W)$ is Equation (10.31) evaluated with a potential barrier which is not V_B, but is $V_B - qFL_W$.

The tunneling current $J_{1\text{tunn}}$ in Equations (10.32) from the ground state in one quantum well to the ground state in an adjacent well can be taken to be either of,

$$J_{1\text{tunn}}(F) = \frac{qm_W k_B T}{2\pi^2\hbar^3} \int_{E_1}^{V_B - qFL_p} T(E,F) \ln\left[\frac{1+\exp\left(\frac{E_F - E}{k_B T}\right)}{1+\exp\left(\frac{E_F - qFL_p - E}{k_B T}\right)}\right] dE$$

(scattering as in 3D reservoir) (10.36)

$$J_{1\text{tunn}}(F) = \frac{qm_W}{\pi\hbar^2} \int_{E_1}^{V_B - qFL_p} \nu T(E,F) dE \left[\frac{1}{1+\exp\left(\frac{E - E_F}{k_B T}\right)} - \frac{1}{1+\exp\left(\frac{E + qFL_p - E_F}{k_B T}\right)}\right]$$

(scattering as in 2D reservoir), (10.37)

where we have given two commonly used forms for the tunneling current: one assumes that scattering is sufficient to model the 2D (quantum well) electron reservoir appear like a 3D reservoir, and the other assumes that quantum well electrons do indeed look like a 2D electron reservoir. (See the discussion following Equations (10.18) and (10.19).) Similarly, the thermionic field assisted tunneling from the ground state into the continuum can be taken to be either of

$$N_{2D}G_{tfa} = \frac{m_W k_B T}{2\pi^2\hbar^3} \int_{V_B - qFL_p}^{V_B} T(E,F) \ln\left[\frac{1+\exp\left(\frac{E_F - E}{k_B T}\right)}{1+\exp\left(\frac{E_F - qFL_p - E}{k_B T}\right)}\right] dE$$

(scattering as in 3D reservoir) (10.38)

$$N_{2D}G_{tfa} = \frac{m_W}{\pi\hbar^2} \int_{V_B - qFL_p}^{V_B} \nu T(E,F) dE \left[\frac{1}{1+\exp\left(\frac{E - E_F}{k_B T}\right)} - \frac{1}{1+\exp\left(\frac{E + qFL_p - E_F}{k_B T}\right)}\right]$$

(scattering as in 2D reservoir). (10.39)

Equations (10.37) and (10.39) differ from the contribution $J_{1\text{well}}$ made by each quantum well to the total current, given in Equations (10.18) and (10.19), in the limits of the energy integral: whereas the energy integrals in Equations (10.18) and (10.19) extend from E_1 to ∞, the energy integral in Equation (10.37) extends from E_1 to $V_B - qFL_p$, and the energy integral in Equation (10.39) extends from $V_B - qFL_p$ to

V_B. This distinction between the two types of tunneling currents, the net forward tunneling current $J_{1\text{tunn}}$ between ground states in adjacent quantum wells, as well as the tunneling current $N_{2D}G_{tfa}$ from a ground state into an upper (continuum) state, is of particular importance in Equations (10.32) and (10.33), in the correct expression of continuity of the flux of carriers into each of the two subband states (the bound quantum well state, and the upper continuum state).

In the steady state, all the time derivatives in Equations (10.32) and (10.33) are zero. If we assume that the electric fields are uniform within any single barrier and are denoted by F_i for barrier i, then in the steady state, Poisson's equation (Equation (10.22)) becomes,

$$F_i - F_{i-1} = \frac{q}{\epsilon_B}\left[N_D^+(i) - N_{2D}(i) - n_{3D}(i-1)L_p(i-1)\right]. \qquad (10.40)$$

In the steady state, the current continuity equation for the carrier flow into the upper state (the continuum) becomes,

$$\begin{aligned}&[-n_{3D}(i-1)v_{QW}(i) + N_{2D}(i)(G_{opt}(i) + G_{tfa}(i) + G_{th}(i))] \\ &\qquad + \frac{1}{q}[J_{3D}(i-1) - J_{3D}(i)] = 0. \end{aligned} \qquad (10.41)$$

In the steady state, the current continuity equation for the carrier flow into the lower (quantum well bound) state becomes,

$$\begin{aligned}&[n_{3D}(i-1)v_{QW}(i) - N_{2D}(i)(G_{opt}(i) + G_{tfa}(i) + G_{th}(i))] \\ &\qquad + \frac{1}{q}[J_{1\text{tunn}}(i-1) - J_{1\text{tunn}}(i)] = 0. \end{aligned} \qquad (10.42)$$

The sum of Equation (10.41) and Equation (10.42) shows that in the steady state, the total current density J_{TOT},

$$J_{TOT} = J_{3D}(i) + J_{1,\text{tunn}}(i), \qquad (10.43)$$

is uniform throughout the device. This is as it should be because in the steady state, continuity of the total current requires that $\nabla \cdot J_{TOT} = -\frac{\partial \rho}{\partial t} = 0$.

Total Current Dominated by either J_{3D} or $J_{1,\text{tunn}}$

A key assumption which is very realistic is now made. If we have,

$$
\begin{aligned}
&|\nabla \cdot \vec{J}_{3D}| \ll |\nabla \cdot \vec{J}_{1\text{tunn}}| \\
&\text{or} \\
&|\nabla \cdot \vec{J}_{3D}| \gg |\nabla \cdot \vec{J}_{1\text{tunn}}|,
\end{aligned}
\tag{10.44}
$$

then we have from either Equation (10.41) or (10.42),

$$
\begin{aligned}
n_{3D}(i-1)v_{QW} &= N_{2D}(i)\left(G_{opt}(i) + G_{tfa}(i) + G_{th}(i)\right) \\
&\equiv N_{2D}(i)G_{up}(i),
\end{aligned}
\tag{10.45}
$$

where we have defined the term in parentheses on the right hand side of Equation (10.45) as $G_{up}(i)$, the total upward transition rate. Equations (10.44) and (10.45) show that if either $J_{1,\text{tunn}}$ or J_{3D} is divergence-free, then in the steady state, the flux of carriers which are captured from the continuum into the quantum wells is balanced by the flux of carriers which are excited out of the quantum wells into the continuum.

One situation in which the assumption in Equation (10.44) is valid occurs when

$$
\begin{aligned}
&|\vec{J}_{3D}| \ll |\vec{J}_{1\text{tunn}}|, \\
&\text{or} \\
&|\vec{J}_{3D}| \gg |\vec{J}_{1\text{tunn}}|,
\end{aligned}
\tag{10.46}
$$

which is the case when the current through the entire QWIP is either dominated by the tunneling current through the barriers or dominated by the current flowing above the barrier band edge (which is either the photocurrent or the thermionic leakage). Thus, the assumption in Equation (10.44) is often valid.

As noted earlier, the explicit appearance in the divergence term in Equation (10.24) of the current density J_{3D} of the carriers in the continuum, together with the explicit *absence* of a similar divergence term for the tunneling current $J_{1\text{tunn}}$ from the bound state of one quantum well to an adjacent well in Equation (10.23), indicates that Ershov's model implicitly assumes that the total current through the QWIP comes mainly from J_{3D} and not from the tunneling current through the bound states in the quantum wells.

Total Current Dominated by J_{3D}

If the total current is dominated by J_{3D}, which is the current flowing above the barrier band edge (either photocurrent or thermionic leakage), then

$$\begin{aligned} J_{TOT} &\approx J_{3D} \\ &= q\, n_{3D}(i) v_d(i) \\ &\approx q\, N_{2D}(i) G_{up}(i) \frac{v_d(i)}{v_{QW}} \end{aligned} \tag{10.47}$$

where we have used Equation (10.45) in arriving at the last line above. Equation (10.47) as well as Poisson's Equation (Equation (10.40)), and the steady state flux in Equation (10.45) easily yield,

$$N_{2D}(i) = \frac{1}{1 + \frac{G_{up}(i)}{v_{QW}(i)} L_p(i-1)} \left[N_D^+ - \frac{\epsilon_B}{q}(F_i - F_{i-1}) \right], \tag{10.48}$$

$$\begin{aligned} n_{3D}(i) &= N_{2D}(i) \frac{G_{up}(i)}{v_{QW}(i)} \\ &= \frac{G_{up}(i)}{v_{QW}(i)} \frac{1}{1 + \frac{G_{up}(i)}{v_{QW}(i)} L_p(i-1)} \left[N_D^+ - \frac{\epsilon_B}{q}(F_i - F_{i-1}) \right] \end{aligned} \tag{10.49}$$

$$J_{TOT} = \frac{q\, v_d(i) G_{up}(i)}{v_{QW}(i)} \frac{1}{1 + \frac{G_{up}(i)}{v_{QW}(i)} L_p(i-1)} \left[N_D^+ - \frac{\epsilon_B}{q}(F_i - F_{i-1}) \right]. \tag{10.50}$$

Equation (10.49) shows that the quasi-Fermi level for the n_{3D} electrons flowing in the continuum above the barrier band edge is different from the quasi-Fermi level for the N_{2D} bound carriers within the quantum wells. The reason is that out of all the physical (radiative, tunneling, collisional and phonon) processes which contribute to the total excitation rate, $G_{up}(i) = G_{opt}(i) + G_{tfa}(i) + G_{th}(i)$, out of a quantum well, the radiative and tunneling processes which contribute to $G_{opt}(i) + G_{tfa}(i)$ are NOT inverse, in the detailed balance sense, to the processes which contribute to $v_{QW}(i)$.

Relation to Levine and Rosencher Photocurrents To compare the photocurrent calculated in this rate equation model with the photocurrent calculated in either the Levine or the Rosencher models, we set $G_{up} \approx G_{opt}$ in Equation (10.47), which is

appropriate when the photocurrent is a lot larger than the leakage current. In this case, Equation (10.47) becomes the photocurrent density J_P,

$$\begin{aligned} J_{TOT} &\approx J_P \\ &= q\, N_{2D}(i) G_{opt}(i) \frac{v_d(i)}{v_{QW}}. \end{aligned} \tag{10.51}$$

Making the observation that carriers are photoexcited at the rate of $\Phi\eta_1$ in Equations (10.1) and (10.3), and at the rate of $N_{2D}(i)G_{opt}(i)$ in Equation (10.51), then we conclude that Equations (10.3) and Equation (10.51) are consistent when we identify the photoconductive gain with,

$$g_1 = \frac{1}{p_c} = \frac{v_d}{v_{QW}}, \tag{10.52}$$

as was noted by Rosencher et al. [119].

The point is that g_1 in Levine's model is related to the capture velocity v_{QW} in Rosencher's model through the Equation (10.52).

Relation to Levine and Rosencher Leakage Currents If the leakage current is dominated by thermionic field assisted tunneling through the top (the triangular) part of the barrier, then $G_{up} \approx G_{tfa}$ in the rate equation model of Equation (10.47), and thus,

$$\begin{aligned} J_{3D} &= q\, N_{2D}(i) \frac{G_{tfa}(i)}{v_{QW}} v_d(i) \\ &= q \frac{v_d(i)}{v_{QW}} \frac{m_W}{\pi\hbar^2} \int_{V_B - qFL_p}^{V_B} \nu T(E,F) dE \left[\frac{1}{1+\exp\left(\frac{E_F - E}{k_B T}\right)} - \frac{1}{1+\exp\left(\frac{E_F + qFL_B - E}{k_B T}\right)} \right] \end{aligned} \tag{10.53}$$

where we have used Equation (10.39) for $N_{2D}G_{tfa}$, and we use Rosencher's value for the impinging frequency ν (the Oppenheimer-Bohr frequency in Equation (10.20)).

A comparison of Equation (10.51) for the photocurrent and Equation (10.53) for the thermionic field assisted tunneling current shows that both are proportional to v_d/v_{QW}, and this merely says that all the n_{3D} carriers flowing in the continuum see

the same functional form for the gain, v_d/v_{QW}, regardless of the previous history (thermionic field assisted tunneling or photoexcitation past) of the carrier in the continuum.

The thermionic field assisted tunneling leakage of Equation (10.53) in the rate equation model can be obtained from Equations (10.12) and (10.10) in Levine's model (for $V_B - qFL_p < E < V_B$) by replacing $\frac{v_d}{L_p}$ in Levine's model with $\frac{v_d}{v_{QW}}\nu$. The origin of this difference is in Rosencher's use of Equation (10.20) for the frequency at which carriers impinge on the barrier. Levine's group [7] does obtain a good fit of modeled to measured leakage currents with the use of reasonable fitting parameters, even though Rosencher's [115] method of using the Oppenheimer-Bohr impinging frequency may seem correct from first principles.

Carrier Injection and Quantum Well Carrier Depletion when the Total Current is Dominated by J_{3D}

The field distribution within the QWIP is found by iteratively solving Equation (10.50) for F_i in terms of F_{i-1}, the fields across the barriers closer to the emitter,

$$\frac{J_{TOT} v_{QW}(i)}{q\, G_{up}(i) \mu_i F_i} = \frac{1}{1 + \frac{G_{up}(i)}{v_{QW}(i)} L_p(i-1)} \left[N_D^+ - \frac{\epsilon_B}{q}(F_i - F_{i-1}) \right]$$

$$\text{for } \mu_i F_i < v_{sat}$$

$$\frac{J_{TOT} v_{QW}(i)}{q\, G_{up}(i) v_{sat}} = \frac{1}{1 + \frac{G_{up}(i)}{v_{QW}(i)} L_p(i-1)} \left[N_D^+ - \frac{\epsilon_B}{q}(F_i - F_{i-1}) \right]$$

$$\text{for } \mu_i F_i > v_{sat}. \qquad (10.54)$$

(In Equation (10.54), it is almost always true that $G_{up}(i)L_p(i-1) < v_{QW}(i)$. For room temperature black body radiation incident on a (40 μm)^2 QWIP pixel, G_{opt}=314/sec. Typical [7] designs have L_p=550 Å and typical models [7, 115, 112] use v_{QW}=10^5cm/s.)

The physical model developed here yields analytical expressions, as in Equation (10.54), for the number of, and the distance over which, carriers are depleted from quantum wells whenever the photocurrent is larger than the leakage current. For designs having the same periodic structure (the same quantum well and barrier compositions and

layer widths) throughout the QWIP, the carriers in the quantum wells can be shown to deplete abruptly whenever the photocarrier drift velocity is large (the lower condition in Equation (10.54)), and to deplete linearly whenever the photocarrier drift velocity is linear in the electric field (the upper condition in Equation (10.54)).

All that is needed now is to find F_1, the field across the barrier closest to the emitter contact. This is found from the boundary condition,

$$J_{inj}(F_1) = J_{TOT}, \tag{10.55}$$

that enough carriers will deplete from the quantum wells closest to the emitter contact and enough carriers will accumulate in the emitter contact so that the field F_1, which is needed to supply the current (J_{TOT}) from the emitter, drops across the barrier next to the emitter.

To find the field F_1, one needs an expression for the injected current density $J_{inj}(F_1)$. Both Levine's [117] group and Rosencher's [116] group write an equation very similar to Equation (10.18) (with E_1 replaced by the conduction band edge E_{em} in the emitter, and with E_F replaced by the Fermi level $E_{F,em}$ in the emitter contact) for $J_{inj}(F_1)$,

$$J_{inj}(F_1) = \frac{q m_{em} k_B T}{2\pi^2 \hbar^3} \int_{E_{em}}^{\infty} T(E, F_1) \ln \left[\frac{1 + \exp\left(\frac{E_{F,em} - E}{k_B T}\right)}{1 + \exp\left(\frac{E_{F,em} + qF_1 L_B - E}{k_B T}\right)} \right] dE. \tag{10.56}$$

When current is injected from the emitter contact predominantly by thermionic emission over the top of the barrier next to the emitter contact, then [124] Equation (10.56) can be approximated as,

$$\begin{aligned} J_{inj}(F_1) &\equiv J_{\text{thermionic}} \\ &= \frac{q m_{em} (k_B T)^2}{2\pi^2 \hbar^3} \exp\left[-\left(\frac{V_{B,em} - E_{F,em}}{k_B T} \right) \right] \left[\exp\left(\frac{W_{em}}{k_B T} \right) - 1 \right], \end{aligned} \tag{10.57}$$

where m_{em} is the effective mass in the emitter material, where $V_{B,em}$ is the band edge (relative to the emitter band edge) of the barrier next to the emitter, where W_{em},

$$W_{em} = E_{F,em} - E_{F,em,0}, \tag{10.58}$$

is the rise in the Fermi level $E_{F,em}$ in the emitter contact with respect to its value $E_{F,em,0}$ in equilibrium (in the absence of any net current flowing through the device). W_{em} is approximately

$$\begin{aligned} W_{em} &= F_1\sqrt{\frac{\epsilon_B}{2\,g(E_{F,em})}}, \\ &\equiv \frac{1}{2}q\,F_1 L_n \\ &\propto E_{F,em}^{-1/4} \end{aligned} \tag{10.59}$$

where we have assumed abrupt accumulation of carriers in the emitter contact over the distance L_n, where F_1 is the field across the barrier next to the emitter, and where $g(E_{F,em})$ is the density of states at the Fermi level in the emitter contact. It is easy to show that W_{em} varies very slowly with respect to the emitter contact Fermi level.

It is easy to show from Equation (10.56) that the injection current resulting from thermionic field assisted tunneling through the triangular part of the barrier is approximately,

$$\begin{aligned} J_{inj}(F_1) &\equiv J_{\text{tfa}} \\ &= J_{\text{tfa},0} T(E_{F,em}, F_1) \\ &= \frac{4qE_{F,em}}{3\pi\hbar}\left(\frac{m_{em}}{\pi\hbar^2}\right)[\Delta E_{F,em} + k_BT]\,[1 - \exp(-qF_1L_p(1))]\,T(E_{F,em}, F_1) \\ &\quad \text{for } V_{B,em} - qF_1L_B < E_{F,em}, \end{aligned} \tag{10.60}$$

where $T(E_{F,em}, F_1)$ is the transmission coefficient in Equation (10.14) evaluated at the Fermi level in the emitter contact, where the terms $\Delta E_{F,em}$ and k_BT in the first square brackets in Equation (10.60) come from energies in the integral in Equation (10.56) which are respectively less than and greater than $E_{F,em}$, and where $\Delta E_{F,em}$ refers to an energy width centered right below $E_{F,em}$ within which the transmission coefficient in Equation (10.14) is sizeable and $\Delta E_{F,em}$ satisfies,

$$\begin{aligned} T(E_{F,em} - \Delta E_{F,em}, F_1) &= \frac{1}{2}T(E_{F,em}, F_1) \\ \Delta E_{F,em} &= \ln 2\left[\frac{3}{4}qF_1\right]\left[\frac{\hbar^2}{2m_{em}}\right]^{1/2}\frac{1}{[V_B - E_{F,em}]^{1/2}}. \end{aligned} \tag{10.61}$$

It is easy to show that thermionic emission (rather than thermionic field assisted tunneling) dominates the injection current from the emitter when

$$J_{\text{thermionic}} > J_{\text{tfa}} \quad \text{when} \quad k_B T > \Delta E_{F,em}. \tag{10.62}$$

Numerically, when $F_1 = 4 \times 10^4$ V/cm, V_B=224 meV, and $E_{F,em}$=75 meV, the injection current is dominated by thermionic emission over the barrier near the emitter ($J_{\text{thermionic}} > J_{\text{tfa}}$) when T>70K.

It should be easy to see from Fig. 10-2 that the depletion of carriers from the quantum wells near the emitter contact is large when the photocurrent is much larger than the leakage current. The reason is that when the photocurrent is much larger than the leakage current, much more current must be pulled from the emitter contact through the first barrier under illuminated conditions than under dark conditions at a fixed bias. This latter situation requires that the electric field across the barriers near the emitter is larger under illuminated than under dark conditions, and thus more carriers must deplete from the quantum wells near the emitter under illuminated than under dark conditions. This carrier depletion from the quantum wells near the emitter barrier may cause the applied bias to drop only across the barriers near the emitter (for small values of the applied bias).

At a large enough applied bias voltage, there is enough field across the barriers near the emitter contact in order for the emitter to inject enough current through these barriers to supply a significant photocurrent through the QWIP. At these values of the applied bias, a larger fraction of the applied bias drops over the main body of the QWIP, than at smaller values of the applied bias.

To get an idea of how large the electric field must be across the barrier next to the emitter contact in order for the emitter contact to supply a significant photocurrent, consider a QWIP at T=40K. At these temperatures, the photocurrent to leakage current ratio can easily be 30 (see Fig. 8-7 in Chapter 8). At these temperatures, current injection is by thermionic field assisted tunneling, Equation (10.60), through the barrier next to the emitter contact. To find the value of the field F_1 across the barrier next to the emitter which is needed to supply a small photocurrent of $J_{\text{small photocurrent}} = q\, N_{2D} G_{opt} \frac{0.01 v_{sat}}{v_{QW}}$ (and a drift velocity within the body of the QWIP

which is $0.01v_{sat}$), we require that the current injected from the emitter contact match this required photocurrent,

$$J_{inj}(F_1) = J_{\text{small photocurrent}}$$
$$J_{\text{tfa},0}T(E_{F,em}, F_1) = q\, N_D^+ G_{opt}\frac{0.01v_{sat}}{v_{QW}}, \tag{10.63}$$

where we have used Equations (10.47) and (10.60) above for, respectively, the photocurrent flowing the QWIP and the injected current from the emitter contact.

To evaluate F_1 from Equation (10.63), we observe that most of F_1 the dependence in Equation (10.63) comes from the transmission coefficient $T(E_{F,em}, F_1)$, so that if we write Equation (10.63) as,

$$\frac{[V_B - E_{F,em}]^{3/2}}{W_{em}} = \frac{3}{2}\left[\frac{\hbar^2}{2m_B L_n^2}\right]^{1/2} \ln\left|\frac{J_{tfa,0}}{J_{\text{small photocurrent}}}\right|, \tag{10.64}$$

then most of the F_1 dependence comes from the left hand side of the above equation. In Equation (10.64), W_{em} is the rise in the emitter Fermi level (relative to its value in equilibrium, as given in Equation (10.58)) as a result of a finite current flow through the QWIP, where $L_n = \sqrt{\frac{2\epsilon_{em}}{q^2 g(E_{F,em})}} \propto E_{F,em}^{-1/4}$ is the distance over carriers are accumulated in the emitter contact, and where $J_{tfa,0}$ was defined in Equation (10.60).

For a typical [7] n-QWIP having a responsivity peak at 8 μm, a spectral FWHM of 1 μm, V_B=224 meV, $E_{F,em,0}$=90 meV for a doping of 2x10^{18}cm^{-3} in the emitter contact, L_n=70 Å, L_B=500 Å, L_W=50 Å, a detector size of $A_{DET}=(40\mu\text{m})^2$, a transmission coefficient of 0.75 for the GaAs to air interface, a single quantum well quantum efficiency of η_1=0.35%, so that an incident, room temperature, black body flux of $\Phi_B=4\times10^{16}$/s-cm^2-ster filling a solid angle of Ω=0.1965 ster yields an optical excitation rate of

$$N_{2D}G_{opt}A_{DET} \equiv \eta_1 \Phi_B \Omega A_{DET} = 4.4 \times 10^8/s, \tag{10.65}$$

and G_{opt}=27.5/sec. For these numbers, Equation (10.64) can be solved on a calculator to yield $W_{em} = \frac{1}{2}q\, F_1 L_n$=12.6 meV, and

$$F_1 = 3.6 \times 10^4 \text{ V/cm} \qquad \frac{\epsilon_B}{q}F_1 = 2.5 \times 10^{11} \text{ cm}^{-2}, \tag{10.66}$$

for, respectively, the internal field across the barrier next to the emitter, and for the excess charge drawn into the emitter contact to supply the current through the QWIP.

This value of the internal electric field (about $F_1 = 3.6 \times 10^4$ V/cm) within a QWIP, which results from carrier depletion from the quantum wells, is in numerical agreement with that calculated in the numerical models of Ershov et al. [112, 113, 114] and Thibaudeau et al. [115].

The space charge resulting from depletion of the quantum wells near the emitter contact has been observed [116, 125] photocapacitance measurements and in leakage current measurements.

10.3.3 Physical Model of Quantum Well Depletion and Accumulation with a QWIP Current of J_{3D}

To calculate the potential distribution within a QWIP resulting from the depletion of the quantum wells near the emitter contact for any given QWIP current J_{3D}, we consider the case of a QWIP with a large number N_W of quantum wells. (The generalization to a smaller number N_W of quantum wells will be obvious.)

The electric field near the collector contact (i.e., F_i for large i) is determined by the excitation rate within the main body of the QWIP (for a given value of J_{3D}),

$$J_{3D} = q\, N_D^+ G_{up} \frac{v_d(F_{\text{large } i})}{v_{QW}}. \tag{10.67}$$

The electric field near the emitter contact (i.e., F_1) is determined by injection past the barrier next to the emitter (for a given value of J_{3D}),

$$J_{inj}(F_1) = J_{3D}, \tag{10.68}$$

where $J_{inj}(F_1)$ is given by Equation (10.56) or one of the approximations, Equation (10.57) or Equation (10.60). The charge within the body of the QWIP (between

the emitter and the collector) contacts is clearly,

$$\text{charge within body of QWIP} = \frac{\epsilon_B}{q}(F_1 - F_{\text{large } i}). \tag{10.69}$$

The distribution of this charge through the QWIP can be found from Equation (10.54),

$$\begin{aligned} \frac{\epsilon_B}{q}(F_i - F_{i-1}) &= N_D^+ - \frac{J_{TOT} v_{QW}(i)}{q\, G_{up}(i) \mu_i F_i} \\ &\text{for } \mu_i F_i < v_{sat} \\ \frac{\epsilon_B}{q}(F_i - F_{i-1}) &= N_D^+ - \frac{J_{TOT} v_{QW}(i)}{q\, G_{up}(i) v_{sat}} \\ &\text{for } \mu_i F_i > v_{sat}. \end{aligned} \tag{10.70}$$

Equations (10.67) and (10.70) show that for large i (near the collector and N_W large), the quantum wells are not depleted at all. (The right hand side of Equation (10.70) is much smaller than N_D^+.) Equation (10.70) also shows that for an excitation rate G_{up} which is large compared to the injected current J_{TOT}, the quantum wells near the emitter (i small) are almost completely depleted. (The right hand side of Equation (10.70)) is almost exactly N_D^+.)

For designs having the same periodic structure (the same quantum well and barrier compositions and layer widths) throughout the QWIP, the carriers in the quantum wells can be shown to deplete abruptly whenever the photoexcited carrier drift velocity is large (the lower equation in Equation (10.70)), and to deplete linearly whenever the photocarrier drift velocity is linear in the electric field (the upper equation in Equation (10.70)). This can be shown by solving Equation (10.70) for F_i iteratively in terms of the field F_{i-1} in the adjacent barrier.

We now consider a n-QWIP having the same quantum well and barrier width throughout the layer structure and having a responsivity peak at 8 μm, a spectral FWHM of 1 μm, V_B=224 meV, $E_{F,em,0}$=71 meV for a doping of 2x10^{18}cm^{-3} in the emitter contact, L_n=70 Å, L_B=500 Å, L_W=50 Å, a detector size of $A_{DET}=(40\mu\text{m})^2$, a single quantum well quantum efficiency of η_1=0.35%, so that an incident, room temperature, black body flux of $\Phi_B=4\times10^{16}$/s-cm^2-ster produces an optical excitation rate of G_{opt}=27.5/sec for each of the N_{2D} quantum well carriers. Each quantum well is

assumed to be doped at 5x10^{11}cm^{-2}, and the barrier is assumed to have a mobility and saturated drift velocity consistent with the literature [7]: μ=1000 cm^2/V-s and v_{sat}=5x10^6 cm/s. We also assume a quantum well capture velocity of [119] v_{QW}=1x10^5 cm/s. This v_{QW} is assumed to be independent of field because the measured photoconductive gain, $g_1 = v_d/v_{QW}$, is found to be independent of field when $v_d = v_{sat}$, as would be expected from a field-independent v_{QW}.

The numerical solution of Equation (10.70) is presented in Figures 10-3, 10-4, 10-5, 10-6, 10-7, 10-8.

Very small applied bias regime. The modeled photocurrent is negligibly small when the bias voltage is less than qF_1L_p. F_1 is given approximately by Equation (10.63), and is the minimum electric field which must drop over the emitter barrier before any appreciable photocurrent can be injected from the emitter contact through this emitter barrier.

At these very small bias voltages, the bias voltage drops only over the barrier next to the emitter. This is shown in Figure 10-3. The electric field lines all originate from the quantum well closest to the emitter, and terminate on the emitter contact. The quantum well closest to the emitter is depleted, and has a net charge of

$$N_D^+ - N_{2D}(1) = \frac{\epsilon_B}{q}F_1. \tag{10.71}$$

At these very small bias voltages, the QWIP behaves like a capacitor having the capacitance associated with just one potential barrier, $\epsilon_B A_{DET}/L_B(1)$.

If a very large field F_1 is required to inject carriers from the emitter, which might occur for a QWIP with a layer structure different from the one we are now considering, then several of the quantum wells near the emitter contact may be depleted. The amount of depleted charge can be found from Equation (10.70). Several quantum wells near the emitter will deplete abruptly if F_1 is large (the lower equation in Equation (10.70)), and they will deplete linearly for smaller values of F_1 (the upper equation in Equation (10.70)).

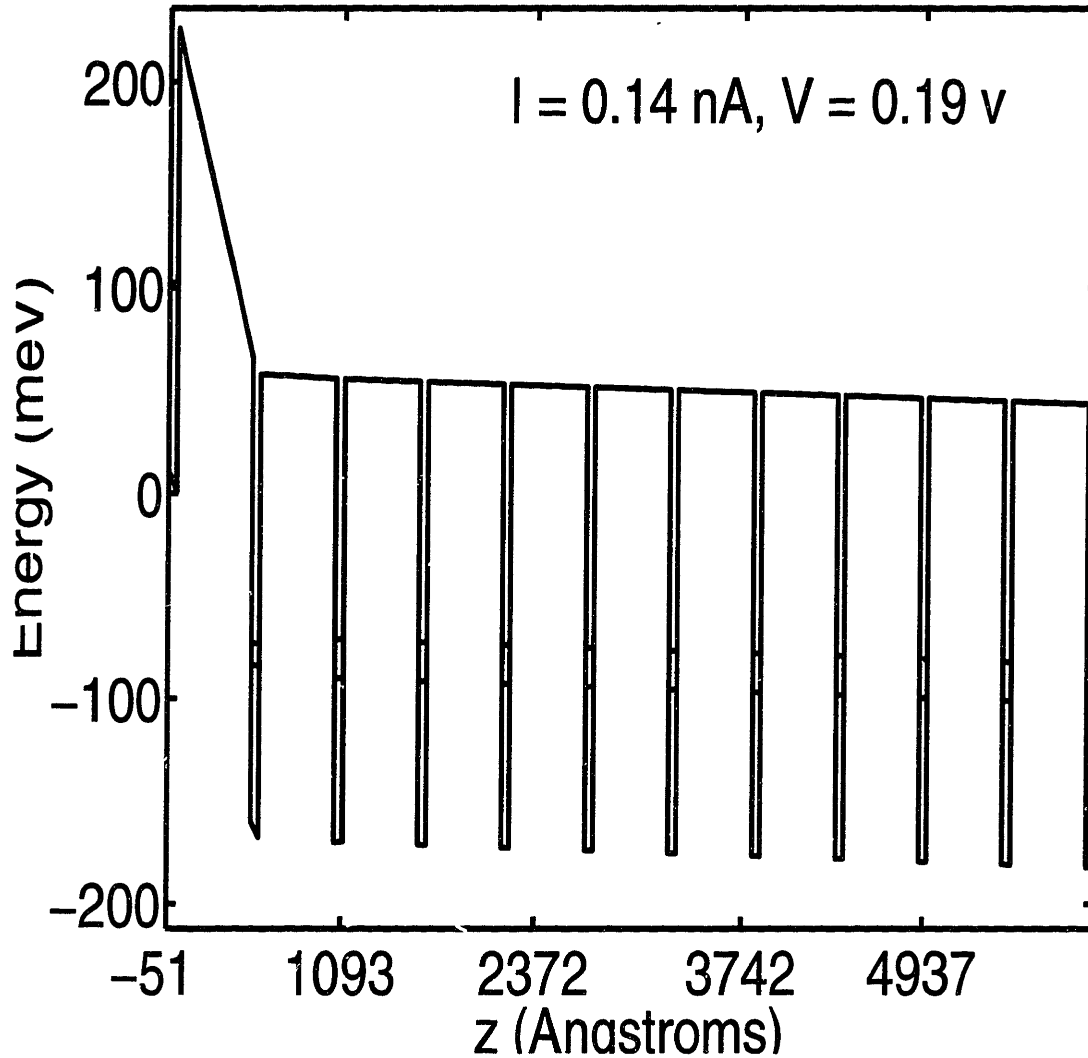


Figure 10-3: The modeled potential distribution, Fermi levels, and bound state energies within an n-QWIP having ten quantum wells. The modeled photocurrent is negligibly small when the bias voltage is less than qF_1L_p, where F_1 is given approximately by Equation (10.63). In the bias regime, the applied bias drops only over the barrier next to the emitter. At these very small bias voltages, the QWIP behaves like a capacitor having the capacitance associated with just one potential barrier, $\epsilon_B A_{DET}/L_B(1)$.

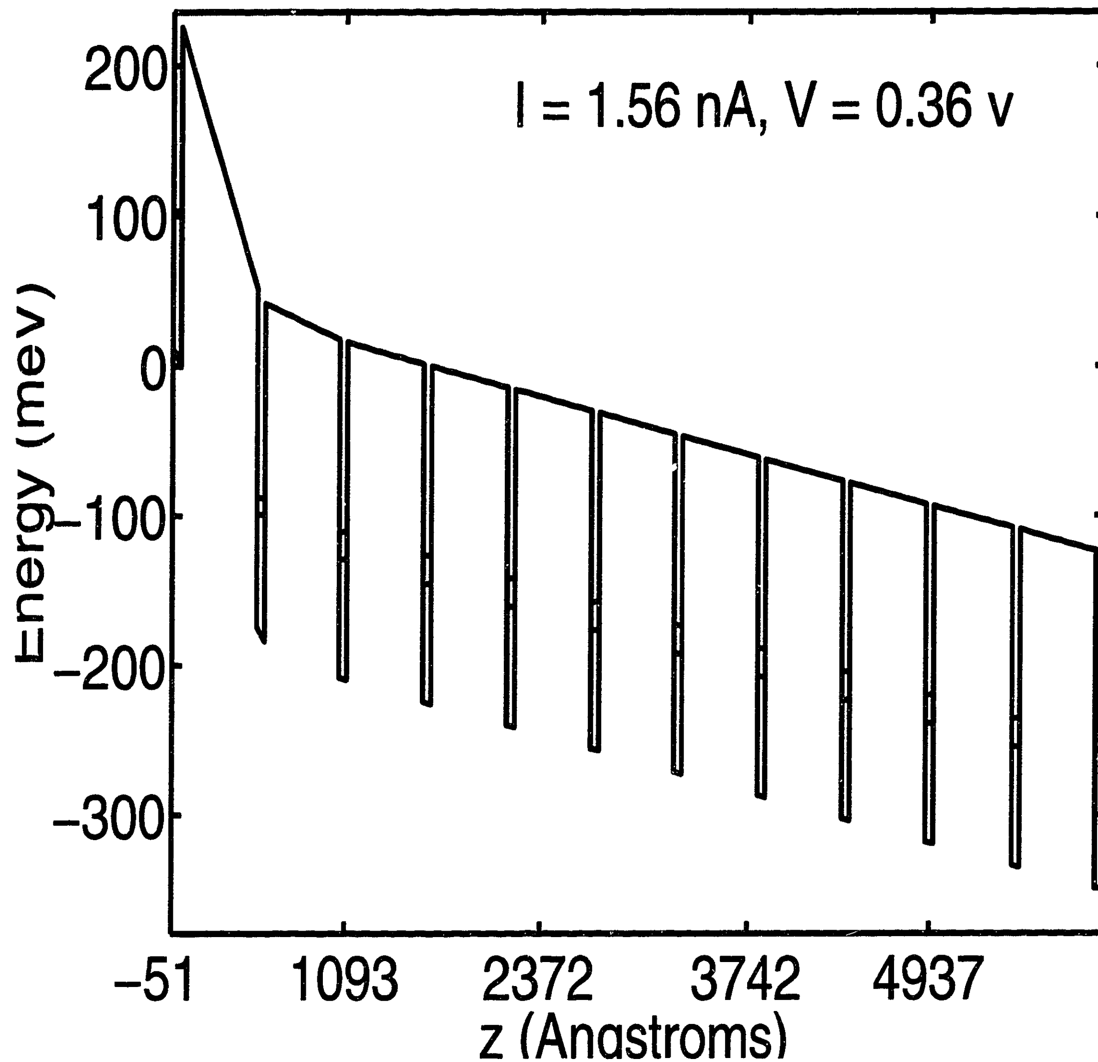


Figure 10-4: The modeled potential distribution, Fermi levels, and bound state energies within an n-QWIP having ten quantum wells. For voltages $qF_1L_p < V < qL_p(NF_{sat} + F_1)$ where $F_{sat} = v_{sat}/\mu$, some of the applied bias drops over the main body of the QWIP, and the photocurrent rises linearly with the applied bias. The photocarrier drift velocity is linear in the electric field in this applied bias regime.

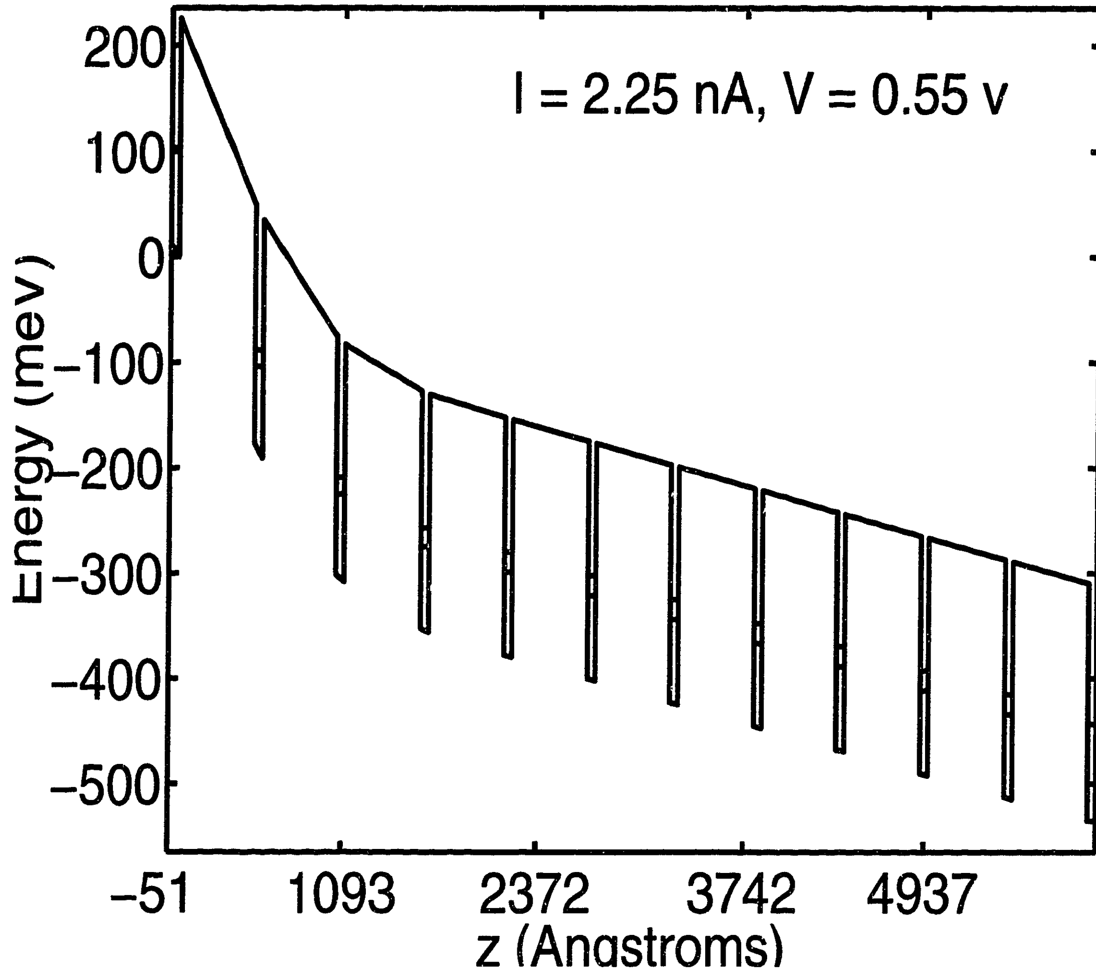


Figure 10-5: The modeled potential distribution, Fermi levels, and bound state energies within an n-QWIP having ten quantum wells. For voltages $qL_p(N_W F_{sat} + F_1) < V < qL_p(N_W + 1)F_1$, the quantum wells have a net charge of $N_D^+ - \frac{J_{TOT} v_{QW}}{q G_{up} v_{sat}}$ in each of the quantum wells in the partially depleted region. The quantum wells in the partially depleted region are uniformly (and abruptly) depleted, with a field-independent (see Equations (10.78) and (10.79)) total charge in the entire partially depleted region. In this figure, the partially depleted region extends over the three quantum wells closest to the emitter. In this bias regime, the photocarriers travel at the saturated drift velocity, and the slow rise of the photocurrent with voltage comes from extra injected carriers through a more transmissive emitter barrier.

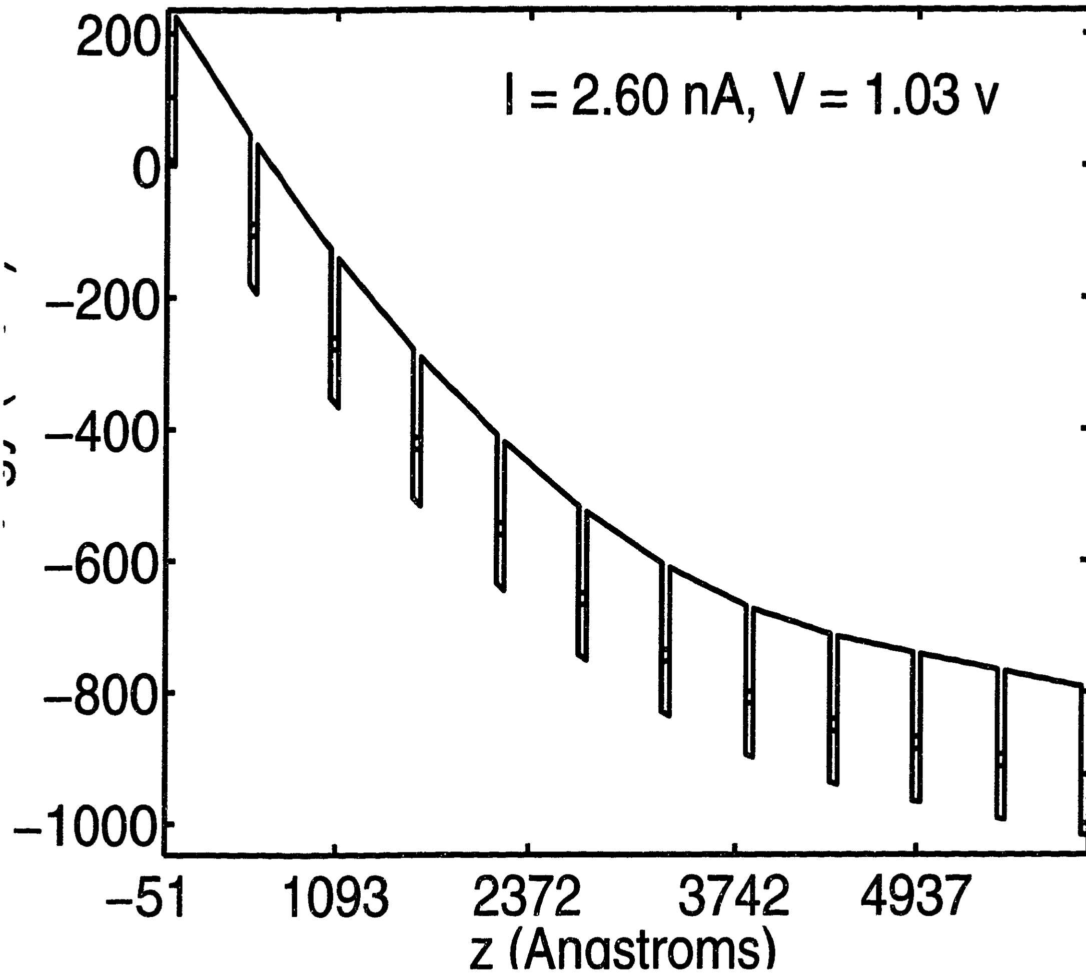


Figure 10-6: The modeled potential distribution, Fermi levels, and bound state energies within an n-QWIP having ten quantum wells. For voltages $qL_p(N_W F_{sat} + F_1) < V < qL_p(N_W + 1)F_1$, the quantum wells have a net charge of $N_D^+ - \frac{J_{TOT} v_{QW}}{q G_{up} v_{sat}}$ in each of the quantum wells in the partially depleted region. The quantum wells in the partially depleted region are uniformly (and abruptly) depleted, with a field-independent (see Equations (10.78) and (10.79)) total charge in the entire partially depleted region. In this figure, the partially depleted region extends over the seven quantum wells closest to the emitter. In this bias regime, the photocarriers travel at the saturated drift velocity, and the slow rise of the photocurrent with voltage comes from extra injected carriers through a more transmissive emitter barrier.

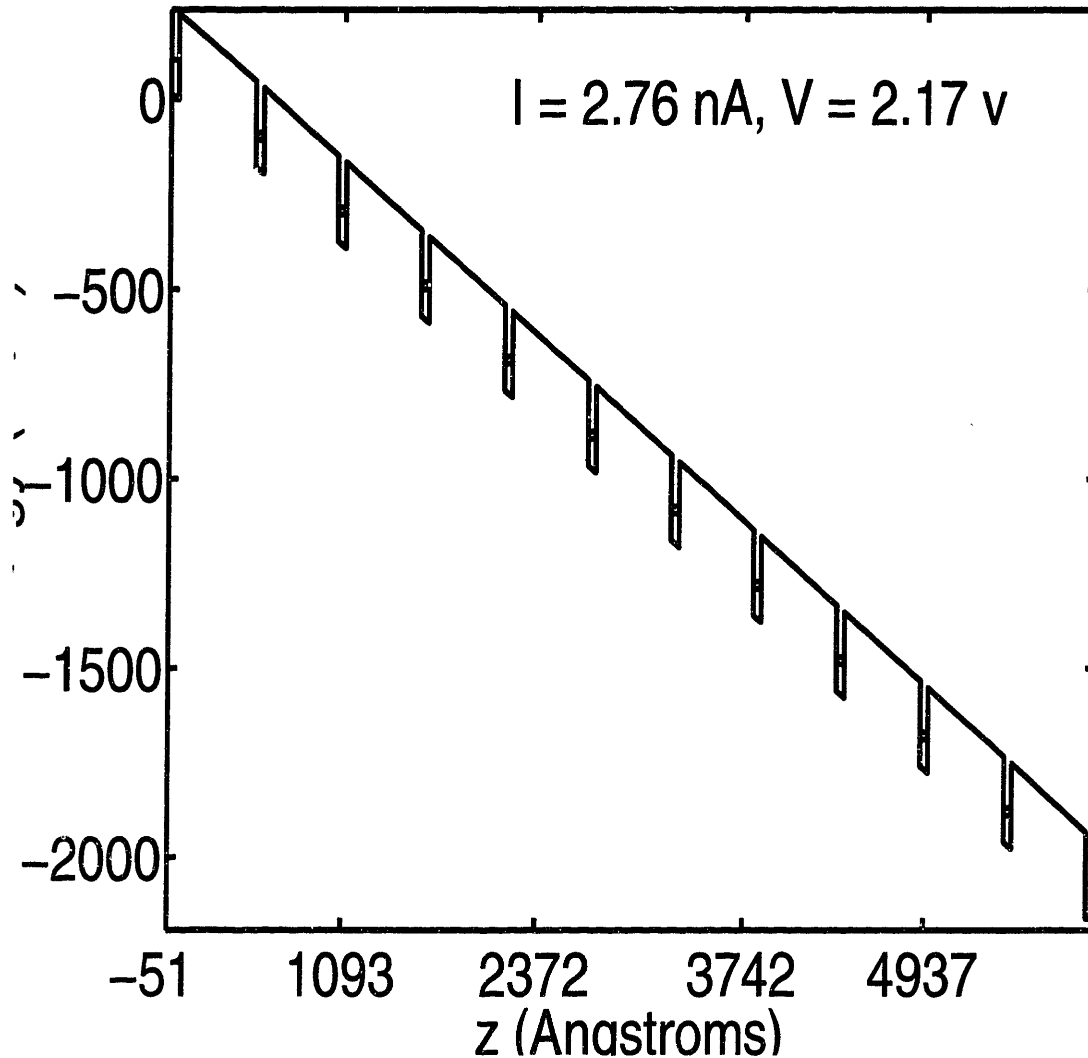


Figure 10-7: The modeled potential distribution, Fermi levels, and bound state energies within an n-QWIP having ten quantum wells. At a large enough bias, there is enough carrier injection to make the QWIP electrically neutral in the entire region between the contacts. This occurs at an applied bias of NF_1L_p, where F_1 is given approximately by Equation (10.63). The current flowing through the device is exactly $J_{TOT} = q\,N_D^+ G_{up} \frac{v_{sat}}{v_{QW}}$.

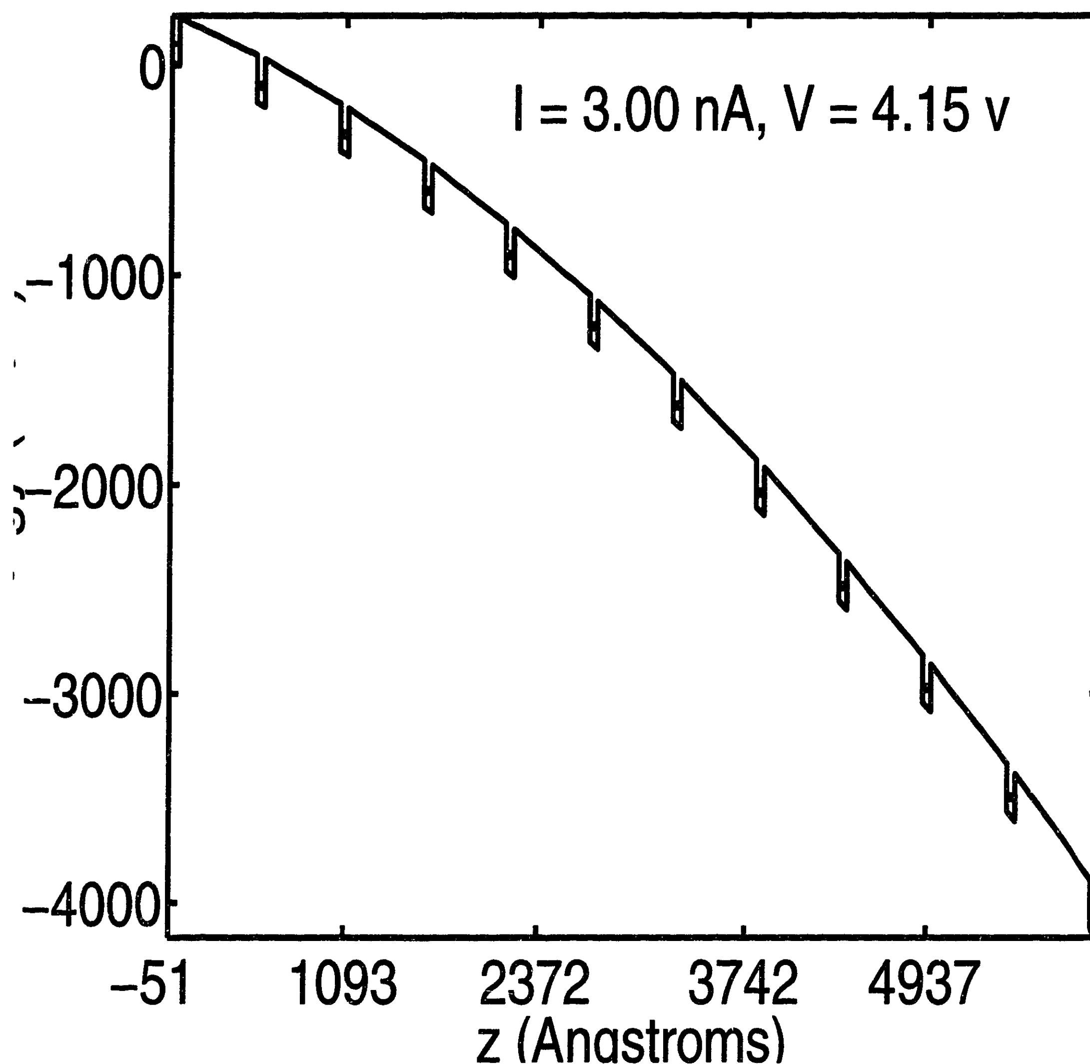


Figure 10-8: The modeled potential distribution, Fermi levels, and bound state energies within an n-QWIP having ten quantum wells. At very large biases, there is enough carrier injection to make the quantum wells slightly accumulated. At these biases, the carrier drift velocity is v_{sat} and is a constant throughout the QWIP structure. The free carrier concentration, n_{3D}, (and thus the quantum well population, N_{2D}, and the accumulated charge density, $N_D^+ - N_{2D}$) is thus uniform throughout the QWIP layer structure.

Small drift velocity regime. For slightly larger bias voltages, the field F_1 across the barrier next to the emitter does not need to change much in order to inject the necessary photocurrent, and thus some of the applied bias drops over the main body of the QWIP. This is shown in Figure 10-4. In this bias regime, the photocurrent rises linearly with the applied bias according to,

$$J_{TOT} = q\,\frac{N_D^+ G_{up}}{v_{QW}} \mu \left(\frac{V}{L_{TOT}} - F_1 \right) \tag{10.72}$$

where $L_{TOT} = N_W L_p + L_B$ is length of the entire multiple QWIP quantum well structure. For these relatively small applied biases, the photocarrier drift velocity is linear in the electric field $\frac{V}{L_{TOT}} - F_1$ across the main body of the QWIP,

$$v_d = \mu \left(\frac{V}{L_{TOT}} - F_1 \right). \tag{10.73}$$

This bias regime is valid when all the photocarriers have drift velocities smaller than the saturated drift velocity v_{sat}, and this corresponds to applied bias voltages satisfying,

$$qF_1 L_p < V < qL_p(N_W F_{sat} + F_1) \tag{10.74}$$

where

$$F_{sat} = \frac{v_{sat}}{\mu}. \tag{10.75}$$

For the numbers given above, F_{sat} evaluates to 5x10^3, which is much smaller than the typical F_1 (see Equation (10.66)) which is needed to inject a sizeable photocurrent.

Saturated drift velocity regime - depleted quantum wells. When the applied bias voltage is increased into the regime,

$$qL_p(N_W F_{sat} + F_1) < V < qL_p(N_W + 1)F_1, \tag{10.76}$$

all the photoexcited carriers drift at the velocity v_{sat}. When the total current through the QWIP is a little lower than the expected value of $q\,N_D^+ G_{up} \frac{v_{sat}}{v_{QW}}$,

$$J_{TOT} \leq q\,N_D^+ G_{up} \frac{v_{sat}}{v_{QW}}, \tag{10.77}$$

the quantum wells near the emitter contact are still depleted. However, it is easy to show from Equation (10.70)) that many of the quantum wells near the emitter are depleted and that the total amount of depleted charge remains fixed. This is shown in Figures 10-5 and 10-6.

Mathematically, if the first N_W^{depl} quantum wells near the emitter are uniformly (and partially) depleted, then the net charge in each quantum well is,

$$q[N_D^+ - N_{2D}^i] = \begin{cases} q\left[N_D^+ - \frac{J_{TOT} v_{QW}}{q\, G_{up} v_{sat}}\right], & \text{for } 1 \leq i \leq N_W^{depl} \\ 0 & \text{for } N_W^{depl} \leq i \leq N_W. \end{cases} \tag{10.78}$$

In Equation (10.81) below, we will use Equation (10.70)) to related the number, N_W^{depl}, of partially depleted quantum wells to the applied bias voltage.

Equation (10.70)) can be used to show that the total charge, $q[N_D^+ - N_{2D}]N_W^{depl}$, in the partially depleted region is,

$$\begin{aligned} q[N_D^+ - N_{2D}]N_W^{depl} &= -\epsilon_B \left[F_1 - F_{sat} \frac{J_{TOT} v_{QW}}{q\, N_D^+ G_{up} v_{sat}}\right] \\ &\approx -\epsilon_B (F_1 - F_{sat}). \end{aligned} \tag{10.79}$$

Since F_{sat} and F_1 change very little with the applied bias, the total charge in the depleted region remains fixed in this operating regime. (Physically, F_1 changes very little with the applied bias in this operating regime where the photocurrent is roughly linear in the applied bias, because a linear change in F_1 would result in an exponential change in the injected current. Physically, $F_{\text{large } i} \approx F_{sat}$ and changes very little with the applied bias in the operating regime defined by Equation (10.76), because current continuity requires that $N_{2D}(\text{large } i) v_d(\text{large } i) \approx N_D^+ \mu F_{sat}$. For thick L_B, it also turns out that $F_1 \gg F_{sat}$.) Thus, the total charge in the partially depleted region changes very little with applied bias, even though the number of partially depleted quantum wells increases with increasing applied bias (in this operating regime).

Poisson's equation can be used to find the electric field (negative for electrons traveling towards the cathode in the figures above) within the QWIP,

$$F_i = \begin{cases} F_1 + \frac{q}{\epsilon_B}\left[N_D^+ - \frac{J_{TOT} v_{QW}}{q\, G_{up} v_{sat}}\right](i-1), & \text{for } 1 \leq i \leq N_W^{depl} \\ F_{sat}\left(\frac{J_{TOT} v_{QW}}{q\, N_D^+ G_{up} v_{sat}}\right) & \text{for } N_W^{depl} \leq i \leq N_W. \end{cases} \tag{10.80}$$

The electric potential throughout the QWIP multiple quantum well structure is then found to increase with distance from the emitter according to,

$$V = \begin{cases} \left(F_1 + \frac{q}{\epsilon_B}\left[N_D^+ - \frac{J_{TOT}v_{QW}}{q\,G_{up}v_{sat}}\right]\frac{(i-1)}{2}\right) iL_p, & \text{for } 1 \leq i \leq N_W^{depl} \\ \left(F_1 + \frac{q}{\epsilon_B}\left[N_D^+ - \frac{J_{TOT}v_{QW}}{q\,G_{up}v_{sat}}\right]\frac{(N_W^{depl}-1)}{2}\right) N_W^{depl} L_p & \\ +F_{sat}\left(\frac{J_{TOT}v_{QW}}{q\,N_D^+ G_{up}v_{sat}}\right)(i - N_W^{depl}) & \text{for } N_W^{depl} \leq i \leq N_W, \end{cases} \tag{10.81}$$

where

$$N_{2D} = \frac{J_{TOT}v_{QW}}{q\,G_{up}v_{sat}} \leq N_D^+. \tag{10.82}$$

Equation (10.81) shows that the electric potential rises parabolically with distance from the emitter in the partially depleted region, and rises linearly with distance in the electrically neutral regions closer to the collector contact. Equation (10.81) also shows that the current rises linearly and slowly with voltage, in this operating regime where the QWIP is slightly depleted and the carriers drift at the saturated drift velocity.

Saturated drift velocity regime - accumulated quantum wells. At a large enough bias, there is enough carrier injection to make the QWIP electrically neutral in the entire region between the contacts, as shown in Figure 10-7. This occurs at an applied bias of NF_1L_p, where F_1 is given approximately by Equation (10.63). The current flowing through the device is then exactly $J_{TOT} = q\,N_D^+ G_{up}\frac{v_{sat}}{v_{QW}}$.

At very large biases, the carrier drift velocity is always v_{sat}, and not larger than v_{sat}. Thus, the increase in the carrier injection with increased bias manifests itself as an accumulation of charge within the QWIP. This is shown in Figure 10-8. At these biases, the carrier drift velocity is a constant (v_{sat}) throughout the QWIP structure. The free carrier concentration, n_{3D}, (and thus the quantum well population, N_{2D}, and the accumulated charge density, $N_D^+ - N_{2D}$) is thus uniform throughout the QWIP layer structure.

Negative charges accumulate in the quantum well regions when the injected current

is large enough to satisfy,

$$J_{TOT} > q\,N_D^+ G_{up} \frac{v_{sat}}{v_{QW}}, \tag{10.83}$$

for which the accumulated charge in each quantum well is,

$$q[N_D^+ - N_{2D}] = q\left[N_D^+ - \frac{J_{TOT} v_{QW}}{q\,G_{up} v_{sat}}\right], \tag{10.84}$$

and is uniform (the same for all quantum well periods) throughout the QWIP layer structure. The field across the i-th period of the QWIP multiple quantum well structure ($1 \leq i \leq N_W+1$) thus increases linearly with distance from the emitter according to,

$$F_i = F_1 + \frac{q}{\epsilon_B}\left[N_D^+ - \frac{J_{TOT} v_{QW}}{q\,G_{up} v_{sat}}\right](i-1), \tag{10.85}$$

where F_1 is negative. The electric potential at the i-th period of the QWIP multiple quantum well structure thus increases parabolically with distance from the emitter according to,

$$V = F_1 L_{TOT} + \frac{q}{\epsilon_B}\left[N_D^+ - \frac{J_{TOT} v_{QW}}{q\,G_{up} v_{sat}}\right] L_p \left(\frac{N_W(N_W+1)}{2}\right). \tag{10.86}$$

Equation (10.86) also shows that the current rises linearly and slowly with voltage, in this operating regime where the QWIP is slightly accumulated and the carriers drift at the saturated drift velocity.

In making the measurements shown in Fig. 8-7 of Chapter 8, the signal-to-noise ratio was observed to be largest when the photocurrent measurement was made at an applied bias which corresponds to a slight accumulation of carriers within the QWIP, as shown in Figure 10-8.

Figure 10-9 shows the modeled QWIP photocurrent as a function of the applied bias voltage obtained from solving Equation (10.70). The modeled photocurrent is seen to be negligibly small when the voltage drop over the emitter barrier is too small (voltages less than qF_1L_p, which is 0.2 volts in Fig. 10-9). For voltages $qF_1L_p < V < Nq(v_{sat}/\mu)L_p$ (of which the latter corresponds to about 0.5 volts in Fig. 10-9), some of the applied bias drops over the main body of the QWIP, and the

photocurrent rises linearly with the applied bias. The photocarrier drift velocity is linear in the electric field in this applied bias regime. For voltages $V > Nq(v_{sat}/\mu)L_p$, the quantum wells have a net charge of $N_D^+ - \frac{J_{TOT}v_{QW}}{qG_{up}v_{sat}}$ in each of the quantum wells in the abruptly depleted (or accumulated) region. In this bias regime, the photocarriers travel at the saturated drift velocity, and the slow rise of the photocurrent with voltage comes from extra injected carriers through a more transmissive emitter barrier.

The curve in Fig. 10-9 has a shape very similar to the measured photocurrent current-voltage characteristic shown in Fig. 8-7 in Chapter 8. However, the details of the photocurrent curves in the two figures are not exactly the same. At low voltages, the measured curve in Fig. 8-7 rises more slowly (exponentially) than the modeled curve in Fig. 10-9 (which rises linearly). This is a result of the finite escape probability p_e which can be small, and is not included in our physical model. We chose not to model the escape probability because the escape probability appears to depend on the specific QWIP layer structure: if the optical transition is chosen with its upper state deep into the continuum of energies above the barrier band edge, then the observed p_e is [7] measured to be larger than that for optical transitions involving bound-to-bound or bound-to-quasi-continuum transitions. At larger applied biases, the measured curve rises a little more quickly than the modeled curve (though still linearly). We believe this to come from thermionic field assisted tunneling of the bound carriers from the quantum wells, which is not included in this physical model.

10.3.4 Thermionic Leakage Current Modeling

The leakage current is most often modeled by Equations (10.10) and (10.12), as was given by Levine [7]. A key point is that it is the total energy E which appears in the transmission coefficient $T(E, F)$ inside the integral in Equation (10.12) because it is believed [7] that in a real system, electron scattering is sufficient to cause the electron wave function to decay in the barriers according to the total energy E instead of the bound state energy E_1. This dependence of the transmission coefficient $T(E, F)$ on the total energy E rather than on the bound state energy E_1 makes the computation

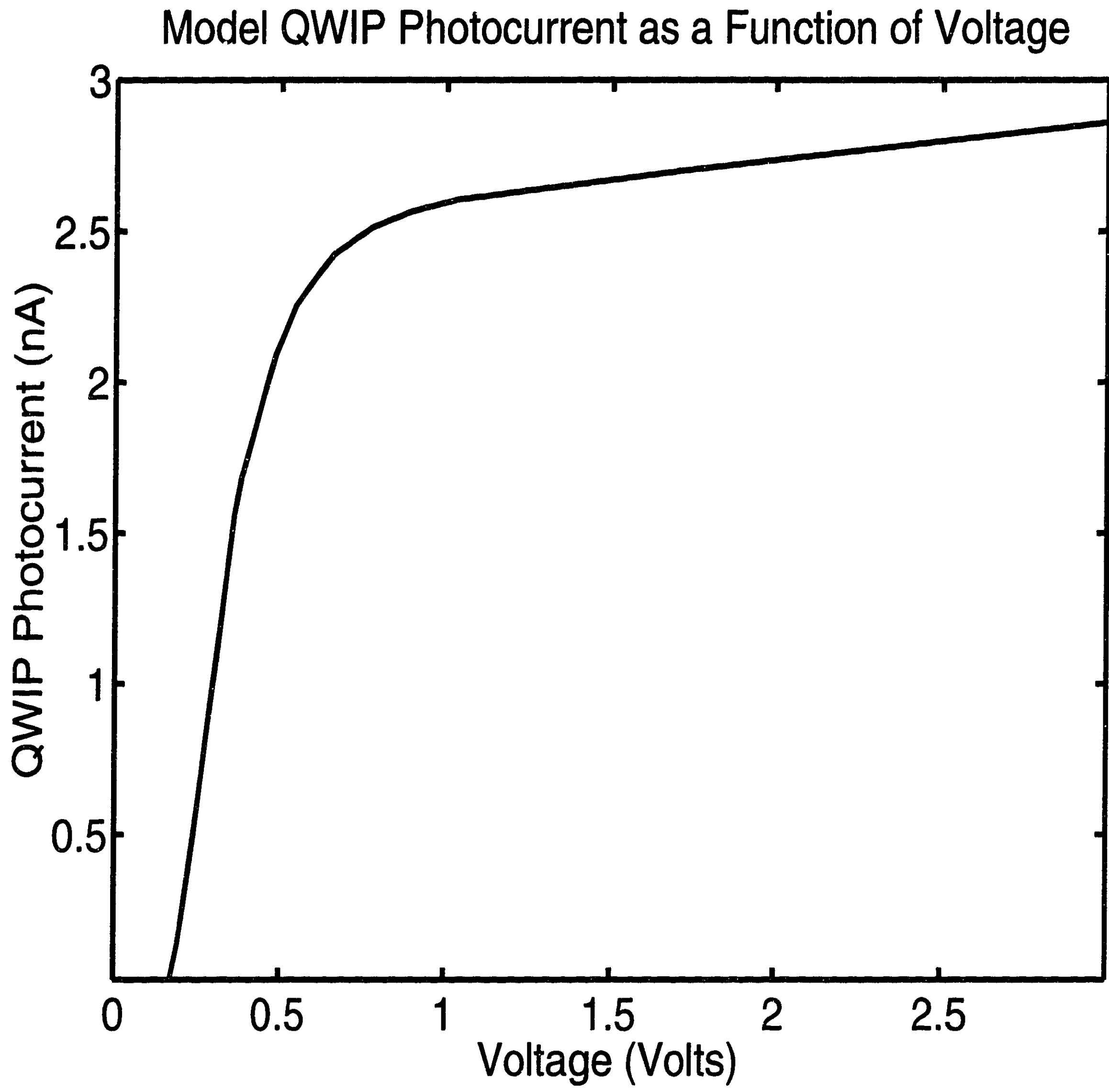


Figure 10-9: The modeled QWIP photocurrent as a function of the applied bias voltage obtained from solving Equation (10.70).

of the integral in Equation (10.12) laborious. Thus, it is not expected that at any arbitrary applied bias, the thermionic leakage current will have the simple form given in Equation (10.16) by Levine for a small applied bias. In fact, Levine's [7] calculations show that at intermediate values of the applied bias, there is a very large contribution to the thermionic leakage coming from the tunneling of carriers through the top part (the triangular part) of the barriers.

In Fig. 8-3 of Chapter 8, we have plotted an Arrhenius plot of the logarithm of the measured leakage current as a function of the inverse temperature. In Chapters 8 and 9, we used these Arrhenius plots to fit the measured thermionic leakage currents to the form [92, 93, 94],

$$J = qN_C v_d \exp\left(-\frac{\Delta E_C - E_F}{k_B T}\right)\left[\exp\left(\frac{qFL_W}{k_B T}\right) - 1\right], \tag{10.87}$$

$$= qN_C v_d \exp\left(-\frac{q\Phi_B}{k_B T}\right)\left[1 - \exp\left(\frac{-qV}{k_B T}\right)\right], \tag{10.88}$$

where q is the electron charge, N_C is the density of states in the barrier, v_d is the drift velocity, ΔE_C is the band offset between the quantum well and barrier conduction band edges, T is the QWIP operating temperature, F is the average field in the structure, L_W is the well width, and Φ_B is the activation energy measured at large voltages ($V > k_B T$),

$$\Phi_B = (\Delta E_C - E_F) - qVL_W/L_{TOT}, \tag{10.89}$$

where L_{TOT} is the length of the entire QWIP structure. Equation (10.88) actually has the same simple form as Equation (10.16) (apart from the $\exp\left(\frac{-qV}{k_B T}\right)$ term, which is the reverse current flowing towards the QWIP cathode, within the square brackets of Equation (10.88)). Both Equations (10.88) and (10.16) show that the leakage activation energy is a measure of the position of the Fermi level in the quantum wells: for a voltage drop of VL_W/L_{TOT} across each well, the Fermi level is closer to the barrier band edge by the amount qVL_W/L_{TOT}.

Table 10.2 shows the measured and theoretical values of $|d\Phi_B/dV|$ for a variety of QWIPs, taken from both this work and the literature. The measured thermionic leakage is found to be close to the model given above, in which the activation energy is

Sample	Well/Barrier	N_W	Inter-Sub-band Transition	L_W/L_B Å/Å	Designed qL_W/L_{TOT} (meV/V)	Measured $\lvert d\Phi_B/dV\rvert$ (meV/V)
Pelve [95] nQWIP	GaAs/ $Al_{0.3}Ga_{0.7}As$	50	B-C	40/300	2.31	5
Jelen [103] nQWIP	$In_{0.53}Ga_{0.47}As$/ InP	20	B-C	50/500	4.3	2.8
nQWIP 9066	$In_{0.08}Ga_{0.92}As$/ $Al_{0.15}Ga_{0.85}As$	10	B-QC	54/450	9.3	7.6
pQWIP 9331	$In_{0.527}Ga_{0.473}As$/ superlattice	15	B-MB	79/510	8.7	10.7
pQWIP 9326	$In_{0.45}Ga_{0.55}As$/ $In_{0.532}Al_{0.468}As$	15	B-QC	62/500	10.6	15.3

Table 10.2: Measured and theoretical (see Equation (10.89)) values of $|d\Phi_B/dV|$ for a variety of QWIPs. The barriers in Sample 9331 consisted of a superlattice of 29.6 Å $In_{0.380}Ga_{0.145}Al_{0.475}As$ (6 repetitions) alternated with 63.6 Å $In_{0.587}Al_{0.413}As$ (5 repetitions)

given $\Phi_B = V_B - qFL_W - E_F$. A large deviation of the measured thermionic leakage from this idealized model leakage current, as in perhaps the first and last entries in Table 10.2, is proposed as a quantitative measure of the amount of excessive leakage.

Our nQWIP 9066 and Jelen's [103] published data showed a thermionic leakage current dependence which has a larger (better) activation energy than the value of $\Phi_B = V_B - qFL_W - E_F$ in Equation (10.88) over a very large range of bias voltages. This raises the question of how accurately the measured leakage currents are described by Equation (10.88). Our measurements show that Equation (10.88) accurately describes the *temperature* dependence (as exemplified by the voltage dependence of the activation energy Φ_B) of the thermionic leakage current. However, Equation (10.88) does not give a good description of the *voltage* dependence of the thermionic leakage

current at a *fixed temperature.*

The points in Figure 10-10 denote the measured leakage current for sample a9066b05 as a function of voltage and at operating temperatures of 60K, 70K, and 80K (labeled, respectively, from the lowermost curve). The measured data points are taken from Fig. 8-1 in Chapter 8. The solid line is a best fit of the measured data points to the form shown in Equation (10.93), which is derived below.

To get a good fit of the measured leakage current in Figure 10-10 to Equation (10.88), we found that it was necessary for the drift velocity v_d in Equation (10.88) to be directly proportional to the electric field. However, for the applied biases $|V|$ >0.25 volts in Fig. 10-10, the drift velocity is no longer proportional to the electric field, but is *equal to the saturated drift velocity,* v_{sat}. (For typically used values of the saturated drift velocity, $v_{sat} = 5x10^6$ cm/s, and mobility, $\mu = 10^3$ cm^2/V-s, the drift velocity, v_d, is about equal to the saturated drift velocity, v_{sat}, for fields F$\geq 5x10^3$ V/cm, which in Fig. 10-10 corresponds to $|V|$ >0.25 volts for a multiple quantum well structure of total width $L_{TOT} = 0.5$ μm.) Thus, in order for the coefficients multiplying the exponential in Equation (10.88) to be linearly proportional to the electric field, the physics behind Equation (10.88) must be carefully examined.

We believe that in order for the measured thermionic leakage currents in Figure 10-10 to be accurately modeled, Equation (10.88) must be modified to include the thermionic field assisted tunneling through the top (triangular) parts of the barriers. This thermionic field assisted tunneling through the top parts of the barriers was calculated by Levine [7] in his numerical evaluation of the integral in Equation (10.12) for the effective number of carriers which are excited from each quantum well. In this work, we take a different approach which does not involve directly calculating Equation (10.12). We shall see below that the number of carriers in the quantum well which escape by thermionic field assisted tunneling through the top (triangular) part of the barrier is directly proportional to the applied electric field.

The contribution of thermionic field assisted tunneling to the total leakage current

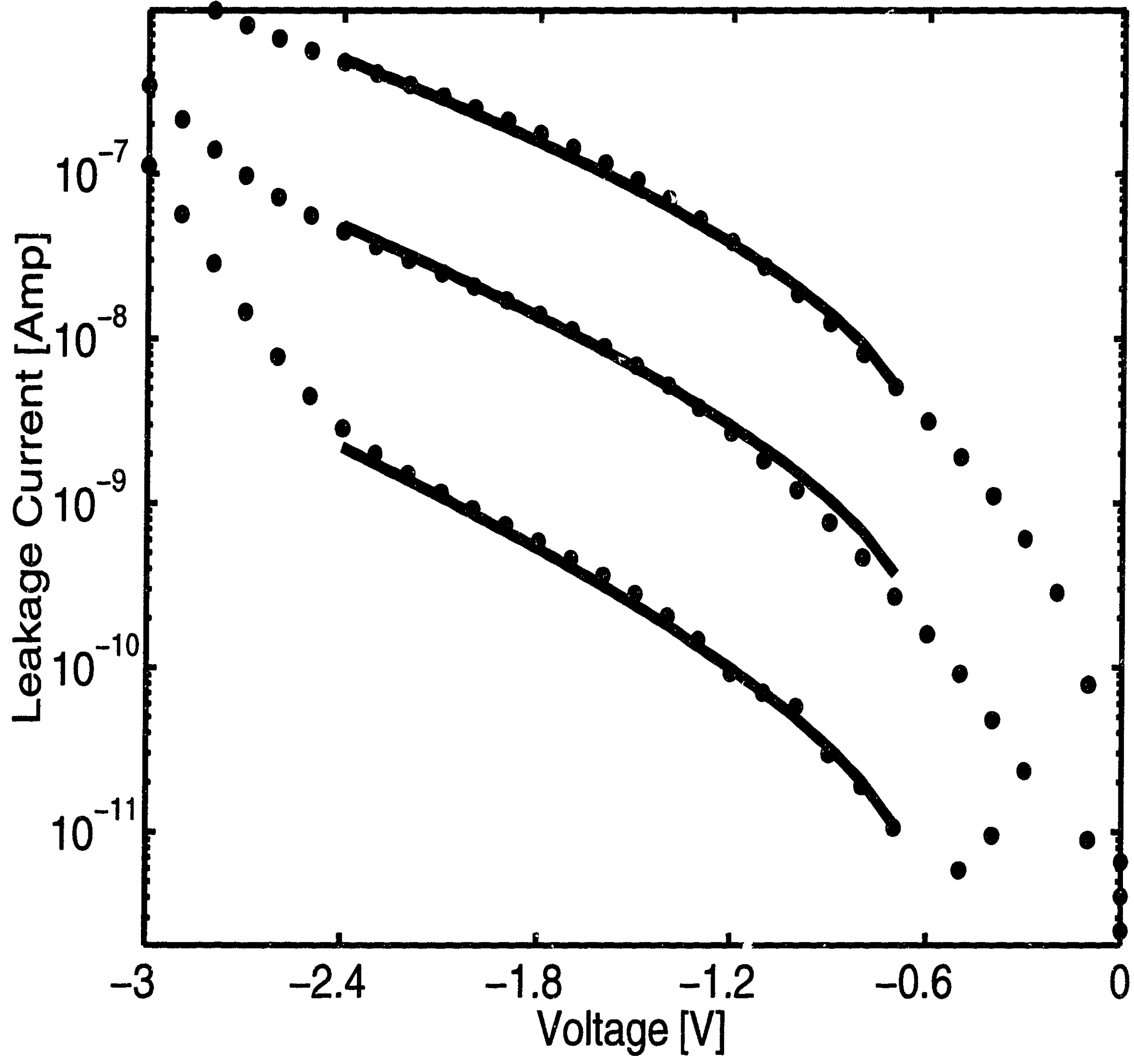


Figure 10-10: The points in the figure denote the measured leakage current for sample a9066b05 as a function of voltage and at operating temperatures of 60K, 70K, and 80K (labeled, respectively, from the lowermost curve). The solid line is a best fit of the measured data points to Equation (10.93) for the thermionic field assisted tunneling through the top (triangular) part of the potential barriers. The measured data points are taken from Fig. 8-1 in Chapter 8.

was given in Equation (10.53),

$$J_L = qN_{2D}G_{tfa}\frac{v_d}{v_{QW}}, \tag{10.90}$$

where we can approximate the thermionic field assisted tunneling flux in Equation (10.53) as,

$$\begin{aligned} N_{2D}G_{tfa} &= \frac{m_W}{\pi\hbar^2}\int_{V_B-qFL_p}^{V_B} \nu T(E,F)\exp\left(-\frac{E-E_F-\frac{qFL_W}{2}}{k_BT}\right)\left[1-\exp\left(-\frac{qFL_p}{k_BT}\right)\right] \\ &\approx \frac{m_W}{\pi\hbar^2}\left(\frac{qFL_{\text{tunn}}}{2}\right)\frac{\nu}{2}\exp\left(-\frac{V_B-\frac{qFL_{\text{tunn}}}{2}-E_F-\frac{qFL_W}{2}}{k_BT}\right)\left[1-\exp\left(-\frac{qFL_p}{k_BT}\right)\right] \end{aligned} \tag{10.91}$$

where the energy separation, $V_B - E_F - \frac{qFL_W}{2}$, between the Fermi level and the barrier band edge is smaller by the amount [92, 93, 94] $\frac{qFL_W}{2}$ in the presence of the electric field. (The Stark shift of the bound state energies in a symmetric quantum well is quadratic in the electric field, so that we can assume that to first order, the quantum well bound state energies do not change with the electric field.) We have also assumed that the largest contribution to the integral in Equation (10.91) comes from those energies which satisfy $1 > T(E,F) > \frac{1}{2}$, and these latter energies correspond to $E > V_B - \frac{qFL_{\text{tunn}}}{2}$, where L_{tunn} is found from $T(E = V_B - \frac{qFL_{\text{tunn}}}{2}, F) = \frac{1}{2}$ to be,

$$\frac{L_{\text{tunn}}}{2} = \left(\frac{3}{4}\ln 2\right)^{2/3}\left[\frac{\hbar^2}{2m_BqF}\right]^{1/3} = \begin{cases} 70 \ \text{Å for F=3.5x}10^3\text{V/cm} \\ 33 \ \text{Å for F=3.5x}10^4\text{V/cm} \end{cases} \tag{10.92}$$

Using Equation (10.91), the contribution of thermionic field assisted tunneling to the total leakage current becomes,

$$J_L \approx q\frac{v_{sat}}{v_{QW}}\frac{m_W}{\pi\hbar^2}\left(\frac{qFL_{\text{tunn}}}{2}\right)\nu\exp\left(-\frac{V_B-\frac{qFL_{\text{tunn}}}{2}-E_F-\frac{qFL_W}{2}}{k_BT}\right)\left[1-\exp\left(-\frac{qFL_p}{k_BT}\right)\right]. \tag{10.93}$$

The modeled leakage current shown in Figure 10-10 is obtained by fitting the *form* (the voltage and temperature dependences) of Equation (10.93) to the measured leakage current. An arbitrary constant multiplying Equation (10.93) is taken as a free

parameter. Without this arbitrary constant multiplying Equation (10.93), it is not possible to match Equation (10.93) to the measured data in Fig. 10-10 for the typically used values of $v_{sat} = 5x10^6$ cm/s and $v_{QW} = 10^5$ cm/s. Of all the constants appearing in Equation (10.93), there has been some discussion in the literature as to why the impinging frequency ν (given in Equation (10.20)), though pedagogically correct, may be the parameter whose magnitude is in doubt. Martinet et al. [122] have noted that the use of the Oppenheimer-Bohr impinging frequency in Equation (10.93) yields values of the modeled tunneling leakage which are much larger than the measured leakage for a single quantum well tunneling diode.

As we have noted earlier, Levine's group [7] does obtain a good fit of the modeled to the measured leakage currents with the use of reasonable fitting parameters by replacing $\frac{v_d}{v_{QW}}\nu$ in Equation (10.93) with $\frac{v_d}{L_p}$, even though the Oppenheimer-Bohr impinging frequency in $\frac{v_d}{v_{QW}}\nu$ may have seemed correct from first principles.

Our Equation (10.93) actually has the same form as Levine's original equation for thermionic leakage, Equation (10.16) above, which he [7] found to be accurate for very small applied biases. Equation (10.93) differs from Equation (10.16) in the appearance of the tunneling distance L_{tunn} in the former equation and the appearance of the barrier width L_B in the latter equation.

The point is that the fit of the measured leakage in Fig. 10-10 to the modeled thermionic leakage currents shows a significant contribution from thermionic field assisted tunneling through the top (triangular) parts of the barriers. The thermionic field assisted tunneling is indicated by the *form* of Equation (10.93): the factor of F multiplying the exponential (even for a field-independent carrier drift velocity of v_{sat}), and the activation energy appearing in the exponential in Equation (10.93) of $\Phi_B = V_B - \frac{qFL_{\text{tunn}}}{2} - E_F - \frac{qFL_W}{2}$. Finally, even though Equation (10.93) does have a voltage and temperature *dependence* which closely matches the measured thermionic leakage, the use of the Oppenheimer-Bohr impinging frequency ν in Equation (10.93) makes the numerical value of Equation (10.93) too large to match the measured values of the leakage current. By replacing $\frac{v_d}{v_{QW}}\nu$ in Equation (10.93) with $\frac{v_d}{L_p}$, the fit of Equation (10.93) to the measured leakage current is better, even though the Oppenheimer-Bohr impinging frequency ν may have seemed correct from first prin-

ciples.

10.3.5 Relation to the Escape Probability

We saw earlier in this chapter that Levine's measured QWIP photocurrents indicated a quantum efficiency, Equation (10.6), which was best fit by a product of the optical absorption efficiency, Equation (10.7), and a phenomenological escape probability, Equation (10.8) with the escape time ratio given by Equation (10.9). We also saw that Rosencher et al. [119] measured a similar dependence of the quantum efficiency η on the applied bias, but they interpreted this dependence in terms of a statistical distribution for the upper state in the intersubband transition.

We now present an explanation of the functional form (the voltage dependence) of the escape probability, p_e, in the expression, Equation (10.6), for the total quantum efficiency. We believe that the form of the expressions, Equation (10.8) and Equation (10.9), for the escape probability comes from the finite tunneling transmission through the top part of a barrier of a photoexcited carrier.

A carrier which is photoexcited to an upper bound state or quasi-bound state will tunnel out of the quantum well with the transmissivity [126],

$$T(E,F) = \frac{1}{1+\exp(2K)} \tag{10.94}$$

where

$$\exp(-2K) = \exp\left(-\frac{4}{3qF}\left[\frac{2m_B}{\hbar^2}\right]^{1/2}[V_B - E]^{3/2}\right) \text{for } V_B - qFL_B \leq E < V_B. \tag{10.95}$$

Equation (10.94) is known to be exact for a barrier which varies parabolically with distance [126].

The functional dependence of Equations (10.94) and (10.95) can be matched to that in Equations (10.8) and (10.9) by choosing the barrier lowering potential, V_p, appearing

in Equation (10.9) to be,

$$V_p = \frac{4}{3q}\left[\frac{2m_B L_p^2}{\hbar^2}\right]^{1/2}[V_B - E]^{3/2}. \tag{10.96}$$

Table 10.1 shows that many of the bound-to-continuum n-QWIPs have values of V_p satisfying,

$$V_p = 12\text{-}17 \text{ meV} \quad \text{thus corresponding to } [V_B - E] = 2.3\text{-}2.9 \text{ meV}, \tag{10.97}$$

for a typical L_p of 550 Å. Equation (10.97) is meaningful in the sense that the measured values for V_p must correspond to a realistic value for $V_B - E$. Equation (10.97) thus shows that the measured values of V_p correspond to upper states in the optical transition which are within about 2 meV of the top of the barrier.

Levine [7] has also measured the barrier lowering potential, V_p, for many p-QWIPs. The typical barrier lowering potential V_p for bound-to-continuum p-QWIPs was measured to be about twice or three times the values given in Table 10.1. This observation is consistent with Equation (10.95) in that p-QWIPs have a heavy hole mass which is typically about 10x the typical conduction band electron mass, and Equation (10.96) shows that V_p should vary as the square root of the mass in the barrier.

10.4 Conclusions

An important part of this work was the development of numerically accurate physical models yielding simple analytical expressions for the QWIP leakage current and photocurrent. These physical models do not require the work involved in the fully numerical solutions already existing in the literature. The physical model developed in this work is derived in Section 10.3 from Gauss's law, Equation (10.22), and current continuity, Equations (10.32) and (10.33), for the flux of carriers into the quantum well bound states and into the continuum of energies above the barrier band edge.

The physical model developed here yielded analytical expressions, as in Equation (10.70) and Section 10.3.3, for the number of, and the distance over which, carriers are depleted from quantum wells whenever the photocurrent is larger than leakage current.

For designs having the same periodic structure (the same quantum well and barrier compositions and layer widths) throughout the QWIP, the carriers in the quantum wells are found to deplete abruptly whenever the photocarrier drift velocity is large, and they are found to deplete linearly whenever the photocarrier drift velocity is linear in the electric field. At very large electric fields, the quantum wells in the QWIP are accumulated.

The depletion of the quantum wells near the emitter barrier is expected to be important in situations where the photocurrent is much larger than the leakage current. This depletion is also important for QWIPs designed with a small number of quantum wells, as would be the case for the optimal design (see Chapter 3) when the single quantum well quantum efficiency is large or when an optical cavity is used. Moreover, this depletion capacitance is expected to be important at high frequencies as in optical heterodyne applications [127, 128, 129, 130, 131, 132].

To complete this physical picture, we present a discussion of current injection via thermionic emission over (see Equation (10.57)) and thermionic field assisted tunneling through (see Equation (10.60)) the barrier next to the emitter.

The measured thermionic leakage is found to be in good agreement with a model in which the leakage depends exponentially on an activation energy which varies linearly with the applied bias. This measured activation energy appears to be $\Phi_B = V_B - \frac{qFL_{\text{tunn}}}{2} - \frac{qFL_W}{2} - E_F$. A fit of the measured to the modeled thermionic leakage currents shows a significant contribution from thermionic field assisted tunneling through the top (triangular) parts of the barriers. A large deviation of the measured thermionic leakage from this idealized model leakage current is proposed as a quantitative measure of the amount of excessive leakage.

The finite tunneling transmission of a photoexcited carrier through the top part of a barrier is used to explain the functional form of the phenomenological escape probability, defined by Levine to account for the discrepancy between the measured quantum efficiency and the measured optical absorption efficiency.

We have reviewed several commonly used models for describing the current flow

through a QWIP. An important difference between Levine's model and Rosencher's model for tunneling leakage is Rosencher's use of the Oppenheimer-Bohr impinging frequency for bound state carriers onto a barrier. This difference may need more study. An important feature of both Levine's and Rosencher's models is that the total tunneling leakage is calculated by assuming that the wave function decay within the barrier drops as the total energy rather than just the bound state energy of the carrier in the quantum well. The rate equation model developed in this work is very similar to Ershov's rate equation model. However, in the physical model developed in this work, a distinction is made between two types of tunneling currents: the tunneling between ground states in adjacent quantum wells, as well as tunneling from a ground state into an upper (continuum) state. All models reduce to Levine's model under appropriate circumstances.

Chapter 11

Conclusions

Quantum well infrared photodetectors (QWIPs) have many potential applications because their narrow responsivity peak can be designed to be at any place in the infrared longer than about 2 μm. This is a wavelength regime useful for the identification of many chemical species, for sensing radiated heat, for determination of the absolute temperature of an object, and for the transmission of radiation through the transparent spectral regions of the Earth's atmosphere. The use of GaAs substrates allows QWIPs to take advantage of modern epitaxy and mature processing technologies as well as possible monolithic integration with GaAs circuits.

11.1 Thesis Accomplishments

This thesis discusses the figures of merit for, the design of, the growth, the (materials, electrical, and optical) characterization of, and the physical modelling of QWIPs.

QWIPs which respond to normally incident radiation without the need for an optical grating are of particular interest because they can be fabricated with fewer process steps and increased expected yield. An important contribution of this work is the demonstration of the first n-type QWIP (n-QWIP) which showed a significant detectivity of $4\text{x}10^{10}\text{cm-}\sqrt{Hz}/\text{Watt}$ without the use of an optical grating. This de-

tectivity corresponds to a conversion efficiency of 4 % or, equivalently, a responsivity of 270 mA/W. This detectivity is significant because it is large [7] enough for focal plane array performance to be limited by the uniformity of processing rather than the size of the single pixel detectivity.

An important part of this work was the development of numerically accurate physical models yielding simple analytical expressions for the QWIP leakage current and photocurrent. These physical models do not require the work involved in the fully numerical solutions already existing in the literature. The physical model developed in this work is derived from Gauss's law, current continuity, and rate equations for the number of bound and free carriers in a QWIP. This physical model yielded analytical expressions for the number of, and the distance over which, carriers are depleted from quantum wells whenever the photocurrent is larger than leakage current. For designs having the same periodic structure (the same quantum well and barrier compositions and layer widths) throughout the QWIP, the carriers in the quantum wells are found to deplete abruptly whenever the photocarrier drift velocity is large, and they are found to deplete linearly whenever the photocarrier drift velocity is linear in the electric field. This depletion capacitance is expected to be important at high frequencies, as in optical heterodyne applications, or when the photocurrent is much larger than the leakage current, or when QWIPs are designed with a small number of quantum wells.

The measured thermionic leakage is found to be in good agreement with a model in which the leakage depends exponentially on an activation energy which varies linearly with the applied bias. A fit of the measured to the modeled thermionic leakage currents shows a significant contribution from thermionic field assisted tunneling through the top (triangular) parts of the barriers. A deviation of the measured thermionic leakage from the idealized model leakage current is proposed as a quantitative measure of the amount of excessive leakage.

Studies of the microscopic physics of quantum wells are presented to elucidate the physical origin of the intersubband absorption of normally incident radiation. A key result of this work is the derivation within the framework of $\vec{k} \cdot \vec{p}$ theory of selection rules for the intersubband absorption of normally incident radiation by holes in a

p-QWIP (p-doped QWIP) in the absence of an optical grating. It is found that the absorption of normally incident radiation by holes in a p-QWIP in the absence of an optical grating is largest for heavy hole to light hole transitions. The intersubband absorption of normally incident radiation by electrons in an n-QWIP in the absence of an optical grating is found within $\vec{k} \cdot \vec{p}$ theory to be much smaller than that in a p-QWIP. The size of the electron intersubband absorption of normally incident radiation in the absence of an optical grating is found to be proportional to the size of the electron in-plane wave vector. It is also found that $\vec{k} \cdot \vec{p}$ theory predicts that uniaxial strain does not have a large effect on the strength or the selection rules of intersubband absorption because the Hamiltonian describing uniaxial strain has the same (tetragonal) symmetry as that describing the confinement of carriers in the quantum wells along the growth direction.

Nonuniformity of device parameters across an array of QWIPs is an important issue. High Resolution X-ray Diffraction (HRXRD) was used to measure the layer width variations of QWIPs grown by molecular beam epitaxy. The spread of the measured full-width-half-maxima of superlattice diffraction peaks with the diffraction order was used with Bragg's Law to obtain the measured layer width variation in the growth direction. A glancing incidence reflectivity measurement is particularly useful for measurement of layer width variations because the large X-ray spot size (about 4.6 cm$\times$250 μm) also yields information about epitaxial layer width uniformity *across* a wafer. It was found that the *absolute size* of the layer width variation was different for three superlattice examples, but the *fractional* layer width variation was about 2% for all three growths. This layer width variation is consistent with an effusion cell temperature variation of 1°C during growth.

Appropriate figures of merit are discussed for different QWIP operating regimes, and related to photodetector device parameters. A theoretical study has been made of different noise mechanisms which contribute to QWIP performance: generation-recombination noise, fixed pattern noise, thermal leakage random arrival noise, Johnson noise, read-out (switch) noise, and photon random arrival noise. A key result

found in this work is that when the signal-to-noise ratio (SNR) is limited by either fixed pattern noise or thermal leakage arrival noise, the optimal number of quantum wells for a maximum in the expected QWIP SNR is roughly η_1^{-1}, where η_1 is the quantum efficiency of a QWIP having only one quantum well. The use of a much larger number of quantum wells is not desirable because the absorption quantum efficiency cannot be increased beyond 100 %, but the photoconductive gain drops with an increase in the number of quantum wells. The use of a much smaller number than the optimal number of quantum wells is not desirable because the absorption quantum efficiency can still be increased with an increase in the number of quantum wells.

Common QWIP designs used in industry are evaluated. In particular, the commonly used n-QWIP design in which the confinement barriers are comprised of a semiconductor superlattice is considered. This QWIP design is intended to reduce thermionic leakage by pushing the three-dimensional continuum of energy further up in energy by making the miniband transport through the superlattice barrier the means of photocurrent conduction. A Kronig-Penney model presented in this thesis showed that this QWIP design, with a superlattice comprising the QWIP barriers, is expected to have a tunneling leakage which is, at best, commensurate with QWIP barriers which are made of a single semiconductor material but whose band edge is the average value of the band edges of the semiconductors comprising the actual barrier superlattice. An experimental investigation of QWIPs with reduced thermionic leakage was proposed in this work for p-QWIPs, which are also designed for large absorption of normally incident radiation without the use of an optical grating. These p-QWIP designs have an additional new feature in that strain is used to lift the light and heavy hole valence degeneracy in the energy continuum, so that the lowest energy hole band in the continuum of energy states can be chosen as light-hole-like for both large photoconductive gain and large quantum efficiency. Electrical measurement of these p-QWIPs showed excessive leakage current, some of which may be accounted for within the framework of the Kronig-Penney model.

11.2 Recent Research Trends

Recent trends in the development of QWIPs having better performance have focussed on two main applications: tactical applications, which require the detection of bright objects at relatively short distances and relatively high QWIP operating temperatures; and space based strategic sensors, which require the detection of faint objects at large distances and low QWIP operating temperatures.

Tactical applications require the detection of a bright target, such as a room temperature black body with a photon flux of $10^{16}\mathrm{cm}^{-2}\mathrm{s}^{-1}$, from a short distance (less than a kilometer) away. The detection of bright targets from a short distance away involve relatively large signals, and thus, relatively large noise may be tolerated. For tactical applications, an important goal is to be able to use the highest possible QWIP operating temperature, so that the time and cost of cooling the QWIP is minimized. Thus, for tactical applications, the QWIP is operated at temperatures above that for radiation Background LImited Performance (BLIP). At such temperatures, it is desirable to minimize the QWIP thermionic leakage current.

An important advance in the development of QWIP FPAs for relatively high temperature operation (about 80K) is the novel grating design of Schimert et al. [13, 133] at Lockheed Martin Vought Systems. The optical grating in Schimert's design is not made of metallic lines. Rather, the optical grating is obtained by etching the grating lines right into the QWIP pixel through the top contact and through all the quantum wells. This has the obvious advantage of reducing the total area of the QWIP pixel without reducing the available optical area. The total leakage current is easily reduced by a factor of four. By using a metallic reflector with the optical grating, the quantum wells are placed inside a resonant optical cavity, and Schimert estimates a doubling of his quantum efficiency with respect to the value in the absence of the optical cavity. By using fewer quantum wells, he also estimates an increase in the photoconductive gain in his design by about a factor of two. Schimert then estimated an increase of the detectivity in his design by at least a factor of five with respect to the detectivity obtained in other designs.

The key point in Schimert's design is the reduction in the total leakage current through each QWIP pixel. (The leakage current for a (35 μm)^2 pixel was measured to be 150μA/cm^2 at 77K for a bias voltage of 1.5 volts over 18 wells.) The improvement in the QWIP detectivity is therefore large only when the noise is dominated by thermal leakage noise. Thus, the projected improvement in the QWIP detectivity at temperatures *below the BLIP* temperature is only about 40%. The reason is that the limiting noise mechanism at these lower temperatures is generation recombination noise and not thermal leakage noise.

Space based strategic detection applications [134] require the detection of faint targets (radiating a photon flux between 10^9cm^{-2}s^{-1} and 10^{13}cm^{-2}s^{-1}) from large distances (thousands of kilometers) away. The small photon flux radiated by faint targets requires a correspondingly low leakage current [134] (less than 1 nA/cm^2) for a significant SNR. This requires low temperature (BLIP) QWIP operation, for which QWIPs are limited by tunneling leakage rather than thermionic leakage. Read-out circuit noise also becomes important in the detection of small [134] photon fluxes (less than 10^{12}cm^{-2}s^{-1}).

Spaced based strategic detection applications have stringent requirements [134] on QWIP response linearity and FPA uniformity. The reason is that the detection of faint targets is accomplished by precise radiometry (rather than imaging), and responsivity linearity is important for precise radiometry. Strategic detection applications utilize larger QWIP pixel sizes, of the order of the spot size of the focussing optics, because absolute radiometry does not require the small pixels needed for imaging over short distances. The use of larger pixel sizes requires uniformity over a larger wafer area.

Rogalski [135] has accomplished a study of QWIP and HgCdTe photodiode performance at temperatures below 77K. He concluded that QWIPs have a higher detectivity at long wavelengths (cutoff wavelengths greater than 8 μm) and low operating temperatures (below 50K). The reason is that at these wavelengths and operating temperatures, n$^+$-p HgCdTe photodiodes are limited by trap assisted tunneling, whereas QWIPs are in the BLIP regime. These observations, in combination with the mature III-V processing technology, makes QWIPs a viable option for highly uniform and highly reproducible, low cost, large area FPAs on a radiation hard substrate.

Bibliography

[1] G. Anderson, J. Chetwynd, J. Theriault, P. Acharya, A. Berk, D. Robertson, F. Kneizys, M. Hoke, L. Abreu, and E. Shettle. Modtran: Suitability for remote sensing. *SPIE*, page 1984, 1993.

[2] G. Anderson, J. Chetwynd, F. Kneizys, L. Hall, L. Kinball, L. Bernstein, P. Acharya, A. Berk, D. Robertson, E. Shettle, L. Abreu, K. Minschwaner, and J. Conant. Modtran3: Suitability as a flux-divergence code. In *Proc. of the 4th ARM Science Team Meeting*, 1994.

[3] In O. Madelung and M. Schulz, editors, *Landolt-Bornstein Numerical Data and Functional Relationship in Science and Technology*, New Series III/22a. Springer-Verlag, Berlin, 1987.

[4] In *Properties of Indium Phosphide.* INSPEC, the Institution of Electrical Engineers, London, 1991.

[5] In M. R. Brozel and G. E. Stillman, editors, *Properties of gallium arsenide.* INSPEC, the Institution of Electrical Engineers, London, 1996.

[6] In P. Bhattacharya, editor, *Properties of lattice-matched and strained indium gallium arsenide.* INSPEC, the Institution of Electrical Engineers, London, 1993.

[7] B. F. Levine. Quantum well infrared photodetectors. *J. Appl. Phys.*, 74:R1–R81, 1993.

[8] J. C. Chiang, S. S. Li, and A. Singh. A two stack indirect barrier/triple coupled quantum well infrared detector for mid-wavelength and long wavelength infrared dual band detection. *Appl. Phys. Lett.*, 71(24):3546–48, Dec. 1997.

[9] K. L. Tsai, K. H. Chang, C. P. Lee, K. F. Huang, J. S. Tsang, and H. R. Chen. Two color infrared photodetector using gaas/algaas and strained ingaas/algaas multiquantum wells. *Appl. Phys. Lett.*, 62(26):3504–3506, Jun. 1993.

[10] R. Jones. Proposal of the detectivity d** for detectors limited by radiation noise. *J. Opt. Soc. America*, 50(11):1058–9, 1960.

[11] R. Jones. Noise in radiation detectors. *Proc. IRE*, 47:1481–6, 1959.

[12] T. Villani, B. Esposito, T. Pletcher, D. Sauer, P. Levine, F. Shallcross, G. Meray, and J. Tower. Performance of generation iii 640x480 ptsi mos array. *SPIE*, 2225:2–18, 1994.

[13] L. T. Claiborne. Improved performance iii-v quantum well ir photodetectors: review of current and potential focal plane technology. In G. J. Brown and M. Razeghi, editors, *Photodetectors: Materials and Devices II*, number Vol. 2999 in Proc. of SPIE, pages 94–102, San Jose, CA, Feb. 1997. SPIE, SPIE - Intern. Soc. for Opt. Eng.

[14] T. N. Casselman. State of infrared photodetectors and materials. In G. J. Brown and M. Razeghi, editors, *Photodetectors: Materials and Devices II*, number Vol. 2999 in Proc. of SPIE, pages 2–10, San Jose, CA, Feb. 1997. SPIE, SPIE - Intern. Soc. for Opt. Eng.

[15] K. V. Shenoy, Jr. C. G. Fonstad, B. Elman, F. Crawford, and J. Mikkelson. Laser diodes and refractory metal gate vlsi gaas mesfets for smart pixels. In *IEEE/LEOS Annual Meeting Digest*, Boston, MA, 1992. IEEE/LEOS.

[16] L. Esaki and H. Sakaki. *IBM Tech. Disc. Bull.*, 20:2456, 1977.

[17] J. S. Smith, L. C. Chiu, S. Margalit, A. Yariv, and A. Y. Cho. A new infrared detector using electron emission from multiple quantum wells. *J. Vac. Sci. Tech. B*, 1:376–78, 1983.

[18] D. D. Coon and R. P. G. Karunasiri. New mode of ir detection using quantum wells. *Appl. Phys. Lett.*, 45(6):649–651, 1984.

[19] L. C. Chiu, J. S. Smith, S. Margalit, A. Yariv, and A. Y. Cho. Application of internal photoemission from quantum well and heterojunction superlattices to infrared photodetectors. *Infrared Physics*, 23(2):93–97, 1983.

[20] L. C. Chiu, J. S. Smith, S. Margalit, and A. Yariv. Internal photoemisision form quantum well heterojunction superlattices by phononless free carrier absorption. *Appl. Phys. Lett.*, 43(4):331–2, 1983.

[21] L. C. West and S. J. Eglash. First observation of an extremely large dipole infrared transition within the conduction band of a gaas quantum well. *Appl. Phys. Lett.*, 46:1156–58, 1985.

[22] B. F. Levine, K. K. Choi, C. G. Bethea, J. Walker, and R. J. Malik. *Appl. Phys. Lett.*, 50:1092–94, 1987.

[23] E. O. Kane. Band structure of indium antimonide. *J. Phys. Chem. Sol.*, 1:249–61, 1957.

[24] J. M. Luttinger and W. Kohn. Motion of electrons and holes in perturbed periodic fields. *Phys. Rev.*, 97:869–883, 1955.

[25] W. Kohn and J. M. Luttinger. Quantum theory of electrical transport phenomena. *Phys. Rev.*, 108:590, 1957.

[26] S. R. White and L. J. Sham. Electronic properties of flat-band semiconductor heterostructures. *Phys. Rev. Lett.*, 47:879–882, 1981.

[27] G. Bastard. Theoretical investigations of superlattice band structure in the envelope-function approximation. *Phys. Rev. B*, 25:7584–7597, 1982.

[28] M. F. H. Schuurmans and G. W. 't Hooft. Simple calculations of confinement states in a quantum well. *Phys. Rev. B*, 31(12):8041–8048, 1985.

[29] G. T. Einevoll and L. J. Sham. Boundary conditions for envelope functions at interfaces between dissimilar materials. *Phys. Rev. B*, 49(15):10533–10542, 1994.

[30] J. P. Cuypers and W. van Haeringen. Connection rules for envelope functions at semiconductor-heterostructure interfaces. *Phys. Rev. B*, 47(16):10310–10318, 1993.

[31] T. Ando, S. Wakahara, and H. Akera. Connection of envelope functions at semiconductor heterointerfaces. i. interface matrix calculated in simplest models. *Phys. Rev. B*, 40(17):11609–11618, 1989.

[32] T. Ando and H. Akera. Connection of envelope functions at semiconductor heterointerfaces. ii. mixings of Γ and x valleys in gaas/$\text{al}_x\text{ga}_{1-x}$as. *Phys. Rev. B*, 40(17):11619–11633, 1989.

[33] B. Laikhtman. Boundary conditions for envelope functions in heterostructures. *Phys. Rev. B*, 46(8):4769–4774, 1992.

[34] A. Yariv. *Quantum electronics.* Wiley, New York, 1989.

[35] M. Flatte, P. Young, L. Peng, and H. Ehrenreich. Generalized superlattice $\vec{k}\cdot\vec{p}$ theory and intersubband optical transitions. *Phys. Rev. B*, 53(4):1–16, 1996.

[36] A. Shik. Optical absorption at a heterojunction. *Sov. Phys. Semicond.*, 22:1165–67, 1988.

[37] A. Shik. Intraband photoconductivity of quantum well heterostructures. *Sov. Phys. Semicond.*, 20:1002–6, 1986.

[38] A. Shik. *Theory of optical intersubband transitions in Intersubband transitions in quantum wells*, volume ed. by E. Rosencher, pages 319–28. Plenum Press, N.Y., N.Y., 1992.

[39] R. Q. Yang, J. M. Xu, and M. Sweeny. Selection rules of intersubband transitions in conduction band quantum wells. *Phys. Rev. B*, 50:7474, 1994.

[40] R. Q. Yang. Can coupling with remote conduction bands cause a significant normal incidence absorption in n-type direct gap semiconductor quantum wells? In G. J. Brown and M. Razeghi, editors, *Photodetectors: Materials and Devices II*, number Vol. 2999 in Proc. of SPIE, pages 161–169, San Jose, CA, Feb. 1997. SPIE, SPIE - Intern. Soc. for Opt. Eng.

[41] P. Pfeffer and W. Zawadzki. Conduction electrons in gaas: Five-level k.p theory and polaron effects. *Phys. Rev. B*, 41:1562, 1990.

[42] M. Cardona, N. E. Christensen, and G. Fasol. Relativistic band structure and spin-orbit splitting of zinc-blende type semiconductors. *Phys. Rev. B*, 38:1806, 1988.

[43] G. A. Baym. *Lectures on Quantum Mechanics.* W. A. Benjamin, Reading, MA, 1974.

[44] A. Sugimura. Band-to-band auger recombination effect on ingaasp laser threshold. *IEEE J. Quant. Electron.*, QE-17(5):627–635, 1981.

[45] B. W. Kim and A. Majerfeld. *J. Appl. Phys.*, 77:4552, 1995.

[46] Y. C. Chang and R. B. James. Saturation of intersubband transitions in p-type semiconductor quantum wells. *Phys. Rev. B*, 39(17):12672–12681, 1989.

[47] B. F. Levine, S. D. Gunapala, J. M. Kuo, S. S. Pei, and S. Hui. Normal incidence hole intersubband absorption long wavelength gaas/$al_x ga_{1-x}$as quantum well infrared photodetectors. *Appl. Phys. Lett.*, 59:1864, 1991.

[48] E. P. O'Reilly. *Semicond. Sci. Tech.*, 4:121, 1989.

[49] G. E. Pikus and G. L. Bir. *Fiz. Tverd. Tela*, 1:1642, 1959.

[50] H. Xie, J. Katz, and W. I. Wang. Infrared absorption enhancement in light and heavy hole inverted $ga_{1-x}in_x$as/$al_{1-y}in_y$as quantum wells. *Appl. Phys. Lett.*, 59:3601–3, 1991.

[51] H. Xie, J. Katz, W. I. Wang, and Y. C. Chang. *J. Appl. Phys.*, 71:2844, 1992.

[52] S. S. Li and J. T. Chu. Effect of compressive strain on the performance of p-type quantum well infrared photodetectors. In G. J. Brown and M. Razeghi, editors, *Photodetectors: Materials and Devices II*, number Vol. 2999 in Proc. of SPIE, pages 132–143, San Jose, CA, Feb. 1997. SPIE, SPIE - Intern. Soc. for Opt. Eng.

[53] S. S. Li and Y. H. Wang. *Quantum Well Intersubband Transition Physics and Devices*, volume ed. by H. C. Liu, pages 29–42. Klüwer Academic Publishers, Netherlands, 1994.

[54] M. A. Hermann and H. Sitter. *Molecular Beam Epitaxy.* Springer-Verlag, N.Y., N.Y., 1989. pages 303–305.

[55] J. Arthur. Vapor pressures and phase equilibria in ga-as system. *J. Phys. Chem. Sol.*, 28:2257–67, 1967.

[56] A. Cho and J. Arthur. Molecular beam epitaxy. *Prog. Sol. St. Chem.*, 10:157–91, 1975.

[57] R. A. Stall, J. Zilko, V. Swaminathan, and N. Schumaker. Morphology of gaas and $al_x ga_{1-x} as$ grown by molecular beam epitaxy. *J. Vac. Sci. Tech. B*, 3(2):524–27, Mar./Apr. 1985.

[58] B. E. Warren. *X-ray Diffraction.* Dover Publications, Mineola, N.Y., 1990.

[59] S. Bates. Advanced interpretation of x-ray interference phenomena. In *Materials Research Society Short Course C.23*, MRS 1990 Fall Meeting, Boston, MA, 1990. Materials Research Society.

[60] I. C. Bassignana. Other applications - x-ray topography. In *Materials Research Society Short Course C.23*, MRS 1990 Fall Meeting, Boston, MA, 1990. Materials Research Society.

[61] I. C. Bassignana. Industrial applications x-ray diffraction. In *Materials Research Society Short Course C.23*, MRS 1990 Fall Meeting, Boston, MA, 1990. Materials Research Society.

[62] V. Swaminathan and A. T. Macrander. *Materials aspects of GaAs and InP.* Prentice-Hall, Englewood Cliffs, NJ, 1991.

[63] E. F. Hockings, T. E. Seidel, C. M. Schmelz, and E. F. Steigmeier. Thermal and electrical transport in inas-gaas alloys. *J. Appl. Phys.*, 37:2879, 1966.

[64] R. E. Nahory, M. A. Pollack, W. D. Johnston, and R. L. Barns. Band gap versus composition and demonstration of vegard's law for ingaasp lattice matched to inp. *Appl. Phys. Lett.*, 33:659, 1978.

[65] K. Nakajima, A. Yamaguchi, K. Akita, and T. Kotani. Composition dependence of the band gaps of ingaasp quaternary solids lattice matched on inp substrates. *J. Appl. Phys*, 49:5944, 1978.

[66] B. R. Bennett. *Molecular beam epitaxial growth and characterization of mismatched InGaAs and InAlAs layers on InP.* PhD dissertation, MIT, Department of Material Science and Engineering, 1993.

[67] A. Krost. *Optical Characeristics of Epitaxial Semiconductor Layers*, volume ed. by G. Bauer, page 287. Springer-Verlag, N.Y., 1996.

[68] J. Matthews and A. Blakeslee. Defects in epitaxial multilayers. *J. Crys. Growth*, 27:118–25, 1974.

[69] P. F. Miceli, D. A. Neumann, and H. Zabel. X-ray refractive index: A tool to determine the average composition in multilayer structures. *Appl. Phys. Lett.*, 48:24–26, 1986.

[70] Matthew Wormington. *GIXR v2.0 - Glancing incidence X-ray reflectivity dynamical simulations.* Bede Scientific Incorporated, Boulder, CO, 1995.

[71] K. Matney and M. S. Goorsky. Determining period variations in a distributed bragg reflector through high resolution x-ray analysis. *J. Crys. Growth*, 148:327–35, 1995.

[72] A. Krol, H. Resat, C. J. Sher, S.C. Woronick, W. Ng, Y. H. Kao, T. L. Cole, A. K. Green, C. K. Lowe-Ma, T.-W Nee, and V. Rehn. Investigation of interfacial roughness of $in_x ga_{1-x} as$ epitaxial layers on gaas and inp substrates by soft x-ray reflectivity. *J. Appl. Phys.*, 69:949–53, 1991.

[73] M. Wormington and D. K. Bowen. Simulation of grazing incidence x-ray techniques. *Bede Scientific Unpublished Internal Report*, 1996.

[74] D. K. Bowen and M. Wormington. Glancing incidence x-ray characterization of materials. *Adv. X-ray Anal.*, 36:171, 1993.

[75] Y. Yoneda. *Phys. Rev.*, 131:2010, 1963.

[76] P. F. Fewster. Interface roughness and period variations in mqw strucures determined by x-ray diffraction. *J. Appl. Cryst.*, 21:524–9, 1988.

[77] P. F. Fewster. X-ray diffraction from multiple qunatum well structures. *Philips J. Res.*, 41:268–289, 1986.

[78] J. C. Vlcek and C. G. Fonstad. Precise computer control of the mbe process - application to graded ingaalas/inp alloys. *J. Cryst. Growth*, 111:56–60, 1991.

[79] K. H. Choy. Aluminum and gallium effusion cell transients in molecular beam epitaxy. Private communication.

[80] E. L. Dereniak and G. D. Boreman. *Infrared detectors and systems.* Wiley, New York, 1996.

[81] H. Okamoto and T. B. Massalski. In O. Madelung and M. Schulz, editors, *Landolt-Bornstein Numerical Data and Functional Relationship in Science and Technology*, New Series IV/5a, pages 368–371, 441–445, 457. Springer-Verlag, Berlin, 1991.

[82] T. C. Shen, G. B. Gao, and H. Morkoc. Recent developments in ohmic contacts for iii-v compound semiconductors. *J. Vac. Sci. Tech. B.*, 10(5):2113–2132, Sept./Oct. 1992.

[83] A. Piotrowska and E. Kaminska. Ohmic contacts to iii-v compound semiconductors. *Thin Solid Films*, 193/194:511–527, 1990.

[84] A. Kulkarni and J. T. Lukowski. Effect of annealing process parameters on the properties of auge ohmic contacts to gaas. *J. Appl. Phys.*, 59(8):2901–2904, Apr. 1986.

[85] J. M. Woodall, N. Braslau, and J. L. Freeouf. Contacts to gaas devices. In *Physics of Thin Films*, 13, pages 199–225. Academic Press, New York, 1987.

[86] N. Braslau. Alloed ohmic contacts to gaas. *J. Vac. Sci. Tech.*, 19(3):803–807, Sept./Oct. 1981.

[87] K. K. Choi, B. F. Levine, C. G. Bethea, J. Walker, and R. J. Malik. Photoexcited coherent tunneling in a double-barrier superlattice. *Phys. Rev. Lett.*, 59(21):2459–62, Nov. 1987.

[88] K. K. Choi, B. F. Levine, C. G. Bethea, J. Walker, and R. J. Malik. Multiple quantum well 10 μm gaas/al$_x$ga$_{1-x}$as infrared detector with improved responsivity. *Appl. Phys. Lett.*, 50(25):1814–1816, Jun. 1987.

[89] A. Shakouri, I. Grave, Y. Xu, A. Ghaffari, and A. Yariv. Control of electric field domain formation in multiquantum well structures. *Appl. Phys. Lett.*, 63:1101–03, 1993.

[90] A. Shakouri, Y. Xu, I. Grave, and A. Yariv. Electon transport, negative differential resistance and domain formation in very weakly coupled quantum wells. In H. C. Liu, editor, *Quantum Well Intersubband Transition Physics and Devices*, page 123. Klüwer Academic Publishers, Netherlands, 1994.

[91] Y. Xu, A. Shakouri, and A. Yariv. Quantum interference effects and electric field domain formation in quantum well infrared photodetectors. *Appl. Phys. Lett.*, 66:3307–09, 1995.

[92] S. M. Sze. *Physics of Semiconductor Devices.* Wiley, New York, second edition, 1981. pp. 259-263.

[93] X. Letartre, D. Stievenard, and M. Lannoo. Admittance spectroscopy measurement of band offset in gaas-gaalas multiquantum well. *J. Appl. Phys.*, 68(1):116–119, Jul. 1990.

[94] R. E. Cavicchi, D. V. Lang, D. Gershoni, A. M. Sergent, J. M. Vandenberg, S. N. G. Chu, and M. B. Panish. Admittance spectroscopy measurement of band offsets in strained layers of in$_x$ga$_{1-x}$as grown on inp. *Appl. Phys. Lett.*, 54(8):739, Feb. 1989.

[95] E. Pelve, F. Beltram, C. G. Bethea, B. F. Levine, V. O. Shen, S. J. Hsieh, and R. R. Abbott. Analysis of the dark current in doped well multiple quantum well algaas infrared photodetectors. *J. Appl. Phys.*, 66:5656–58, 1989.

[96] G. M. Williams, R. E. DeWames, C. W. Farley, and R. J. Anderson. Excess tunnel currents in algaas/gaas multiple quantum well infrared detectors. *Appl. Phys. Lett.*, 60(11):1324–26, Mar. 1992.

[97] H. C. Liu. Photoconductive gain mechanism of quantum well intersubband infrared detectors. *Appl. Phys. Lett.*, 60:1507–9, 1992.

[98] P. S. Martin, J. L. Pan, Jr. C. G. Fonstad, and P. Kannam. Normal incidence pseudomorphic ingaas qwip. In *Proc. 53rd Device Research Conference, Late news paper*, Charlottesville, VA, 19-21 Jun. 1995.

[99] R. Shih, J. Scheihing, M. Dodd, J. Chen, and G. Karunasiri. Normal incidence intersubband infrared detector using n-type ingaas/gaas quantum wells. In *IRIS*, 1995.

[100] G. Karunasiri, R. Shih, and J. Chen. Normal incident intersubband infrared detectors using n-type ingaas/gaas quantum wells. In *Proc. 53rd Device Research Conference*, Charlottesville, VA, 19-21 Jun. 1995.

[101] G. Karunasiri, J. Park, J. Chen, R. Shih, J. Scheihing, and M. Dodd. Normal incident ingaas/gaas multiple quantum well infrared detector using electron intersubband transitions. *Appl. Phys. Lett.*, 67:2600–02, 1995.

[102] S. C. H. Wang. Private communication.

[103] C. Jelen, S. Slivken, T. David, M. Razeghi, and G. J. Brown. Noise performance of ingaas-inp quantum well infrared photodetectors. *IEEE. J. Quant. Electron.*, 34(7):1124–28, Jul. 1998.

[104] B. F. Levine, C. G. Bethea, G. Hasnain, V. O. Shen, E. Pelve, R. R. Abbott, and S. J. Hsieh. High sensitivity low dark current 10 μm gaas quantum well infrared photodetectors. *Appl. Phys. Lett.*, 56:851–3, 1990.

[105] H. C. Liu, A. Steele, M. Buchanan, and Z. Wasilewshi. *Effects of the upper state position and the number of wells on the performance of intersubband quantum well detectors in Intersubband transitions in quantum wells*, volume ed. by E. Rosencher, pages 57–63. Plenum Press, N.Y., N.Y., 1992.

[106] S. D. Gunapala, J. S. Park, G. Sarvsik, T-L Lin, J. K. Lieu, P. D. Maker, R. E. Muller, C. A. Shott, and T. Hoelter. 15 μm 128x128 gaas/algaas quantum well infared photodetector focal plane array camera. *IEEE Trans. Electron. Devices*, ED-44:45–50, 1997.

[107] W. Beck, J. Little, A. Goldberg, and T. Faska. *Imaging performance of LWIR miniband transport multiple quantum well infrared focal plane arrays in Quantum Well Intersubband Transition Physics and Devices*, volume ed. by H. C. Liu, pages 55–68. Klüwer Academic Publishers, Netherlands, 1994.

[108] R. A. Smith. *Wave Mechanics of Crystalline Solids.* Chapman and Hall, Ltd., London, UK, 1961.

[109] L. S. Yu and S. S. Li. *Appl. Phys. Lett.*, 59:1332–4, 1991.

[110] L. S. Yu, Y. H. Wang, S. S. Li, and P. Ho. *Appl. Phys. Lett.*, 60:992–4, 1992.

[111] R. T. Kuroda and E. Garmire. *Infrared Physics*, 34:153, 1993.

[112] M. Ershov, V. Ryzhii, and C. Hamaguchi. Contact and distributed effects in quantum well infrared photodetectors. *Appl. Phys. Lett.*, 67:3147–49, 1995.

[113] M. Ershov and V. Ryzhii. Modelling of multiple ingaas/gaas quantum well infrared photodetectors. In *Proc. of the Seventh International Conference on Narrow Gap Semiconductors*, Institute of Physics Conference Series Number 144, pages 353–58, Santa Fe, New Mexico, 8–12 Jan. 1995.

[114] M. Ershov, H. C. Liu, and M. Buchanan Z. R. Wasilewski. Nonlinear photoresponse of quantum well infrared photodetectors at high excitation power. In G. J. Brown and M. Razeghi, editors, *Photodetectors: Materials and Devices II*, number Vol. 2999 in Proc. of SPIE, pages 153–160, San Jose, CA, Feb. 1997. SPIE, SPIE - Intern. Soc. for Opt. Eng.

[115] L. Thibaudeau, Ph. Bois, and J. Y. Duboz. A self consistent model for quantum well infrared photodetectors. *J. Appl. Phys.*, 79:446–54, 1996.

[116] E. Rosencher, F. Luc, Ph. Bois, and S. Delaitre. Injection mechanism at contacts in a quantum well intersubband infrared detector. *Appl. Phys. Lett.*, 61–63:468–71, 1992.

[117] K. M. S. V. Bandara, B. F. Levine, R. E. Leibenguth, and M. T. Asom. *J. Appl. Phys.*, 74:1826, 1993.

[118] H. C. Liu, G. C. Aers, M. Buchanan, Z. R. Wasilewski, and D. Landheer. *J. Appl. Phys.*, 70:935, 1991.

[119] E. Rosencher, B. Vinter, F. Luc, L. Thibaudeau, Ph. Bois, and J. Nagle. Emission and capture of electrons in multiquantum well structures. *IEEE Trans. Quant. Elect.*, 30(12):2875–2888, 1994.

[120] K. M. S. V. Bandara, B. F. Levine, and M. T. Asom. *J. Appl. Phys.*, 74:346, 1993.

[121] S. V. Meshkov. *Sov. Phys. JETP*, 64:1337, 1986.

[122] E. Martinet, E. Rosencher, F. Chevoir, J. Nagle, and P. Bois. Direct determination of the electron tunneling escape time from a gaas/$Al_xGa_{1-x}As$ quantum well by transient-capacitance spectroscopy. *Phys. Rev. B*, 44(7):3157–61, 1991.

[123] M. Kane, S. Millidge, M. Emeny, D. Lee, D. Guy, and C. Whitehouse. *Performance trade offs in the Quantum Well Infrared Detector in Intersubband transitions in quantum wells*, volume ed. by E. Rosencher, pages 31–42. Plenum Press, N.Y., N.Y., 1992.

[124] A. A. Grinberg and S. Luryi. Space-charge limited current and capacitance in double junction diodes. *J. Appl. Phys.*, 61:1181–89, 1987.

[125] E. Rosencher, N. Vodjdani, J. Nagle, P. Bois, E. Costard, and S. Delaitre. Photocapacitance spectroscopy of gaas/algaas multiquantum wells. *Appl. Phys. Lett.*, 55:1853–55, 1989.

[126] M. Brack and R. K. Bhaduri. *Semiclassical Physics.* Addison-Wesley, N.Y., N.Y., 1997.

[127] E. R. Brown, K. A. McIntosh, F. W. Smith, and M. J. Manfra. Coherent detection with a gaas/algaas multiple quantum well structure. *Appl. Phys. Lett.*, 62:1513–15, 1993.

[128] H. C. Liu, G. E. Jenkins, E. R. Brown, K. A. McIntosh, K. B. Nichols, and M. J. Manfra. Optical heterodyne detection and microwave rectification up to 26 ghz using quantum well infrared photodetectors. *IEEE Elect. Dev. Lett.*, 16:253–55, 1995.

[129] H. C. Liu, Jianmeng Li, E. R. Brown, K. A. McIntosh, K. B. Nichols, and M. J. Manfra. Quantum well intersubband heterodyne infrared detection up to 82 ghz. *Appl. Phys. Lett.*, 67:1594–96, 1995.

[130] H. C. Liu, Jianmeng Li, M. Buchanan, and Z. R. Wasilewski. High frequency quantum well infrared photodetectors measured by microwave rectification technique. *IEEE J. Quant. Elect.*, 32:1024–28, 1996.

[131] S. Ehret, H. Schneider, J. Fleissner, P. Koidl, and G. Bohm. Ultrafast intersubband photocurrent response in quantum well infrared photodetectors. *Appl. Phys. Lett.*, 71:641–43, 1997.

[132] I. Melngailis, W. E. Keicher, C. Freed, S. Marcus, B. E. Edwards, A. Sanchez, T. Y. Fan, and D. L. Spears. Laser radar component technolocy. *Proc. IEEE*, 84:227–67, 1996.

[133] T. R. Schimert, S. L. Barnes, A. J. Brouns, F. C. Case, P. Mitra, and L. T. Claiborne. Enhanced quantum well infrared photodetector with novel multiple quantum well grating structure. *Appl. Phys. Lett.*, 68:2846–48, 1996.

[134] A. Singh and D. A. Cardimona. Design issues relating to low temperature dark current in quantum well infrared photodetectors. In G. J. Brown and M. Razeghi, editors, *Photodetectors: Materials and Devices II*, number Vol. 2999 in Proc. of SPIE, pages 46–54, San Jose, CA, Feb. 1997. SPIE, SPIE - Intern. Soc. for Opt. Eng.

[135] A. Rogalski and K. Jozwikowski. Gaas/algaas quantum well infrared photoconductors versus hgcdte photodiodes for long wavelength infrared applications. *Opt. Eng.*, 33(5):1477–84, May 1994.

THESIS PROCESSING SLIP

FIXED FIELD: ill. ______ name ______

index ______ biblio ______

► COPIES: (Archives) Aero Dewey (Eng) Hum

Lindgren Music Rotch Science

TITLE VARIES: ►☐ ______

NAME VARIES: ►☐ ______

IMPRINT: (COPYRIGHT) ______

► COLLATION: 281p

► ADD: DEGREE: ______ ► DEPT.: ______

SUPERVISORS: ______

NOTES:

cat'r: date:

page: F52

► DEPT: E.E.

► YEAR: 1999 ► DEGREE: Ph.D.

► NAME: PAN, Janet Lin