Quantum Well Intersubband Photodetectors in Focal Plane Arrays

Paul Scott Martin

B.S., California Polytechnic State University
at San Luis Obispo, CA

Submitted to the Department of Physics in partial fulfillment of the requirements for the degree of

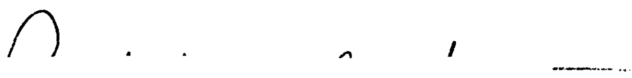
Doctor of Philosophy

at the

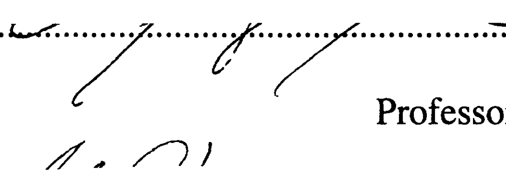
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 1996

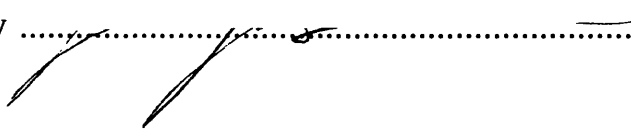
© 1996 Massachusetts Institute of Technology
All Rights Reserved.



Author ....
Department of Physics
3 April 1996



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



Certified by ....
John Joannopoulos
Professor of Physics
Thesis Supervisor

Accepted by ....
George Koster
Chair, Department Committee on Graduate Students
Department of Physics

MASSACHUSETTS INSTITUTE
OF TECHNOLOGY

JUN 05 1996 Science

LIBRARIES

Quantum Well Intersubband Photodetectors in Focal Plane Arrays

by

Paul S. Martin

Submitted to the department of Physics in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Physics

Abstract

Quantum Well Intersubband Photodetectors (QWIPs) are attractive devices for use in very large focal plane arrays (FPAs) because QWIPs can take advantage of both established technology for growing and processing GaAs optical devices and commercially available large area VLSI GaAs IC's. This thesis demonstrates the first normal incidence TE mode QWIP. Detectivity for the normal incidence TE mode devices was as high as 2×10^{10} Jones with responsivities as high as 80mA/W corresponding to quantum efficiencies of 0.069% per quantum well. This value for the quantum efficiency compares well to theoretical calculations which give a value of 0.068% per quantum well. A detectivity of 2×10^{10} Jones is sufficient to produce focal plane arrays with minimum resolvable temperature differences less than 10mK for QWIP FPAs operating with a 300K background in the LWIR band. For current state of the art read-out ICs and calibration techniques these TE mode QWIP FPAs will be uniformity limited, not limited by the detectivity of the individual TE mode QWIP pixels and thus perform as well as FPAs made from devices with much higher detectivity. The recognition that the quantum efficiency of TE mode QWIPs, though low compared to competing devices, is sufficiently high so as not to effect the overall performance of the FPA is an important contribution of the work.

Finally, in discussing issues related to monolithic integration of QWIP FPAs and their companion read out integrated cirtuits (ROICs), this thesis has improved the design of the growth wells used in the epitaxy-on-electronics (E-o-E) integration process and proposed a new integration technique called epitaxy-on-electronics/selective area waferbonding (E-o-E/SAW) which combines the strengths of the E-o-E process with the flexibility of growing on an unprocessed substrate. E-o-E/SAW is quite exciting and very generally applicable to the design of any OEIC which requires integration of VLSI scale electronic circuits with MSI scale optical devices.

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

Thesis Supervisor: John Joannopoulos
Title: Professor of Physics

Acknowledgments

It is a pleasure to acknowledge some of the many people who have made my tenure at MIT both enjoyable and a wonderful learning experience. My thesis advisor Prof. Fonstad has given me both the support and the encouragement to pursue a wide variety of research interests and collaborations. The organization of his research group encourages breadth of experience by grouping students with similar research interests rather than by discipline across many research topics. As a result I've been able to pursue this QWIP project from device design and device modeling, through mask making, MBE growth, material characterization, device fabrication and device testing with detours into a variety of related fields. I have been personally responsible for doing the work at each step and thereby have built up sufficient competence in all of these areas to bring this project to completion. In this system, only after building a foundation in many disparate areas can one hope to push the envelop sufficiently far to complete a Ph.D. One may imagine that this is not the most efficient way of organizing a research group. This is certainly true if the measure of efficiency is number of papers published or number of working devices fabricated. A tremendous amount of time and effort are spent learning the basics of each device development step and correcting the mistakes that are inherent in the learning process. Prof. Fonstad has been willing to sacrifice research efficiency for educational value. Sometimes this is difficult to recognize when the MBE is being vented for the fourth time in a single month due to yet another rookie mistake, but as I learned while interviewing, employers put a high value on breadth of experience and I thank Prof. Fonstad for his foresight and his sacrifice.

As recently recognized by President Clinton who awarded him the National Medal of Science, Prof. Haus is a true giant in field of optics. His ability to look at a problem and clearly identify the boundaries of the solution as well as the subset of important physical effects *before* writing equation 1 is truly inspirational and a model which I am still working to emulate. As this thesis developed process related tasks took up a large fraction of my time, but the time I did spend at Prof. Haus's group meetings and the support he has given me in this work have been both pleasurable and a treasured asset. I would also like to thank Professor Joannopoulos and Professor Grodzins who have taken time out of their own busy schedules to advise me in this work and read this thesis several times over as it has developed. I have benefitted from their experience and support.

Acknowledgments

Several people deserve special mention for contributions to this work. Professor Fitzgerald has been a tremendous help in understanding the issues involved with strained layer heteroepitaxy and a good guy to work with. I look forward to more work together in the future. Dr. Bob Martin at Lockheed/Martin and A/DIC has been a gold mine of information and experience related to QWIP FPAs and the perfect guy to call when in need of a little *nerd* humor. I hope the collaboration started here continues to bear fruit. Robert Shih at Alpha Photonics has been extremely generous with his data and his time and I look forward to him joining us for further work on normal incidence QWIP FPAs. Peter Kannam of ADT has been helpful with some of the measurements in this thesis. Janet Pan has been my collaborator at MIT and heir apparent to this work. She has been a great pleasure to work with bringing both a wealth of knowledge and an enthusiasm for learning new things that is both rare and refreshing. Janet also developed a knack with Wally, our MBE machine, that has lifted a great burden off me and shows great promise for the future of MBE at MIT.

I also wish to thank the many past and present members of the compound semiconductor research group with whom I have had the pleasure of working: Rajni Aggarwal, Joe Ahadian, Eric Braun, Tom Broekaert, Geoff Burns, WooYoung Choi, Yuzo Hirayama, Isako Hoshino, Norio Iizuka, Tanni Kuo, Prof. Tadatsugu Minami, Janet Pan, Steve Patterson, LunHan Peng, Prof. Shiela Prasad, Yakov Royter, Hiroya Sato, Krishna Shenoy, Yosef Singer, Jurgen Smet, Elias Towe, Jim Vlcek and Hao Wang.

Finally, I thank my mother, father and my brother Dave for giving me their love and support through the years it has taken me to get in and out of MIT. I miss them greatly and look forward to moving back home to share our time together. May Lu started out as a teacher for Silicon processing and ended up as a sister-in-law, needless to say her friendship has brought endless joy and I hope to see her and her family join us in California. My mother-in-law Nabu has made it possible for my wife and I to both work and care for a small daughter. Her love, support and Fujianese lessons have enriched my life immeasurably. Finally I would like to acknowledge my wife, Man-Jun, and my daughter Julie who have given me the strength and stability at home without which focussing on research would have been impossible.

Table of Contents

Page #

Abstract 3

Acknowledgments 5

Table of Contents 7

List of Figures 11

List of Tables 19

1 Introduction to Infrared Focal Plane Arrays 21
1.1 Thesis Organization 21
1.2 Historical Introduction 22
1.3 Infrared Physics Background 24
1.4 Figures of Merit for IR detectors and Focal Plane Arrays 29
1.4.1 Responsivity - R_λ, R_{BB}: 30
1.4.2 Noise Equivalent Power - NEP 31
1.4.3 Detectivity - D* 31
1.4.4 Noise Equivalent Temperature Difference - NETD 33
1.4.5 Minimum Resolvable Temperature - MRT 33
1.5 Effect of Calibration on FPA uniformity 43
1.6 Review of Selected IR Detector Technologies 44
1.6.1 Mercury Cadmium Telluride, HgCdTe, FPAs 47
1.6.2 Platinum Silicide Schottky Barrier FPAs 51
1.6.3 Micro-Bolometer FPAs 54
1.6.4 Quantum Well Intersubband Photodetectors 56
1.6.5 Summary of Selected IR Detector Technologies 62
1.7 Conclusion of Introductory Remarks 63

2 Normal Incidence QWIPs 65
2.1 QWIP device parameters and performance trade-offs 66
2.2 TE mode QWIP FPAs 75
2.2.1 Superlattice ***K•p*** theory applied to intersubband optical transitions 77
2.2.2 Potential for using TE absorption in QWIP focal plane arrays 93
2.2.3 Conclusions and avenues to improvement of TE mode absorption 98
2.3 Grating coupled TM mode miniband QWIP FPAs 99
2.4 Alternative normal incidence QWIP technologies 104

2.4.1 P-type normal incidence QWIPs....105
2.4.2 X-band AlAs/AlGaAs QWIPs....107
2.4.3 Concluding remarks on alternative normal incidence QWIPs....109
2.5 Concluding remarks on normal incidence QWIPs....110

3 Growth and Fabrication of Normal Incidence QWIPs....111
3.1 Introduction....111
3.2 Molecular Beam Epitaxy Growth....112
3.3 Material Characterization....113
3.3.1 Double Crystal X-ray Diffraction....113
3.3.2 Photoluminescence....115
3.4 Strained layer heteroepitaxy and critical layer thickness....118
3.4.1 Single layer critical layer thickness....122
3.4.2 Superlattice critical layer thickness....128
3.4.3 Experimental verification of SL critical layer thickness theory....133
3.4.4 Conclusions on strain and SL critical layer thickness....137
3.5 Fabrication of normal incidence QWIPs....137
3.6 Conclusions and future directions in fabrication of QWIPs....142

4 Measurement and Analysis of Normal Incidence QWIPs....145
4.1 Introduction to QWIP measurement....146
4.1.1 Measurement of dark current....149
4.1.2 Measurement of responsivity....153
4.1.3 Measurement of spectral response....160
4.2 Calculation of D^* from measured parameters....162
4.3 Corroborating Measurements done at Alpha Photonics....164
4.4 Conclusions on QWIP measurement....169

5 Modeling and Cryogenic Measurement of VLSI GaAs MESFET Circuits....171
5.1 Introduction to VLSI GaAs MESFETs....171
5.2 Physics and Modeling of GaAs MESFET....173
5.3 Cryogenic Measurement of GaAs MESFETs....185
5.3.1 I-V Characteristics for MESFETs:....186
5.3.2 Threshod Voltage and Transconductance....188
5.3.3 Source and Drain Resistance:....190
5.3.4 Schottky gate barrier height and ideality factor....192
5.4 Cryogenic Measurement of Gate Capacitance....194
5.4.1 Ring Oscillator Design....195
5.4.2 Gate Capacitance extraction....199
5.5 HSPICE parameter extraction:....200
5.6 Conclusions on Cryogenic Characterization of GaAs MESFETs....202

6 Monolithic Integration of Normal Incidence QWIPs....203

6.1 Introduction to Integration of QWIPs 205
6.2 Epi-on-Electronics integration technique for QWIP FPAs 207
6.3 Selective area wafer bonding integration technique 218
6.4 Conclusions and future work on monolithically integrated QWIPs 226

7 Conclusions and Possible Directions for Further Research 229
7.1 Thesis Accomplishments 229
7.2 Possible Directions for Further Research 232

Bibliography 235

List of Figures

Figure #	Figure Caption	Page #

1.1 Black Body Spectral Density 24
1.2 Photon Flux for Black Body near RT with a bandwidth of 1μm....25
1.3 Photon flux for detector with 1μm bandwidth and center frequency in the midwave IR (MWIR) 4μm or long-wave IR (LWIR) 9μm....26
1.4 Contrast Photon Flux for MWIR (4μm) and LWIR (9μm) bands with detector bandwidth of 1μm....27
1.5 MODTRAN3 calculation of atmospheric transmission through a distance of 5km at mid-latitude during summer with visibility of 5km and no rain over farmland. ...28
1.6 Simplified model of an IR FPA Camera....34
1.7 Uniformity limit for FPAs in the LWIR, 8μm to 10μm band. Pixel area 50μm x 50μm, frame rate 30Hz, optics f/# 1.5 AR coated, noise floor 200e^-, $g = 0.5$. 38
1.8 Uniformity limit for FPAs in the MWIR, 3μm to 5μm band. Pixel area 50μm x 50μm, frame rate 30Hz, optics f/# 1.5 AR coated, noise floor 200e^-, $g = 0.5$. 39
1.9 Spatial noise limited regime for uniformities of 0.5%, 0.1%, 0.05% and 0.01% in LWIR 8μm to 10μm band. Pixel area 50μm x 50μm, frame rate 30Hz, optics f/# 1.5 AR coated, noise floor 200e^-, $g = 0.5$. 40
1.10 Spatial noise limited regime for uniformities of 0.5%, 0.1%, 0.05% and 0.01% in MWIR 3μm to 5μm band. Pixel area 50μm x 50μm, frame rate 30Hz, optics f/# 1.5 AR coated, noise floor 200e^-, $g = 0.5$. 41
1.11 Comparison of MWIR, 3μm to 5μm, and LWIR 8μm to 10μm bands. Uniformity 0.02%, pixel area 50μm x 50μm, frame rate 30Hz, optics f/# 1.5 AR coated, noise floor 200e^-, $g = 0.5$....42
1.12 Comparison of D* and MRT as figures of merit for FPAs. LWIR band 8μm to 10μm, f/# 1.5 optics with AR coating, pixel area 50μm x 50μm, noise floor 200e^-, background temperature 300K, $g = 0.5$. 42
1.13 MRT in an InSb photovoltaic focal plane array after two point uniformity correction at 15C and 50C. Array format 160 x 120, *f/#* 2.3.[17]....44
1.14 D_λ^* for MCT detector with area 50μm x 50μm, $\eta = 1$, $t = 5\mu m$ $n = 5\times10^{13}$ cm^{-3} >> p and T as shown....50
1.15 Photoabsorption in a Schottky barrier detector....51
1.16 PtSi 640 x 480 FPA with MRT = 70mK.[32]....53
1.17 BST Pixel Structure. 55
1.18 QWIP Schematic showing 40Å GaAs quantum wells under bias and separated by 300Å $Al_{0.31}Ga_{0.69}As$ barriers. Excited electron mean free path L is represented schematically for an electron generated in the left most well and captured by the right most well. 57
1.19 Measurement geometry for edge coupled QWIP devices....58

1.20 QWIP conduction band diagram. Quantum well is 40Å of GaAs separated by 480Å $Al_{0.25}Ga_{0.75}As$ barriers. 59

2.1 Schematic for operation of one well in QWIP. E_T is the activation energy for thermionic emission, E_P is the photon energy for peak detection, E_F is the Fermi level, E_1 is the energy of the first bound state and ΔE_C is the conduction band offset. ..66
2.2 Multi-well QWIP under bias. 68
2.3 Calculated photoconductive gain for capture probability of p = 0.01, 0.04, 0.07, 0.1 and 0.2. Experimental data taken from Levine[40], Liu[51] and Kane[53]. Figure from H.C. Liu.[50] 70
2.4 Schematic representation of the 14 bulk bands considered. Four energies define the unstrained zone-center bands and six define the strained zone center bands. The matrix element Q couples the six valence band p states to the six conduction band $\bar{p}$ states, the matrix element P_1 couples the conduction band $\bar{s}$ and $\bar{p}$ states and the matrix element P_0 couples the valence band p states to the conduction band $\bar{s}$ states. Overbars refer to antibonding states. 78
2.5 Comparison of the bulk 14-band and 8-band $k \cdot p$ models for GaAs. 85
2.6 A vertical blowup of the region near the unphysical crossing of light and heavy hole bands in the 8-band model.[46] 86
2.7 14 band model calculation of TE and TM absorption coefficients for a superlattice of 40Å GaAs and 300Å $Al_{0.30}Ga_{0.70}As$ at 300K. Absorption length is entire superlattice period, 340Å.[46] 92
2.8 Comparison of TM absorption for a superlattice of 40Å GaAs and 300Å $Al_{0.30}Ga_{0.70}As$ at 300K. Dotted lines are calculations done with 8-band and 14-band models and solid line is experimental data form Levine.[69] Absorption length is entire superlattice period, 340Å.[46] 93
2.9 Measurement geometry for TE mode QWIP devices 94
2.10 Uniformity limit for FPAs in the LWIR, 8μm to 10μm band. Pixel area 50μm x 50μm, frame rate 30Hz, optics f/# 1.5 AR coated, noise floor 200e-, g = 0.5 95
2.11 Uniformity limit for FPAs in the MWIR, 3μm to 5μm band. Pixel area 50μm x 50μm, frame rate 30Hz, optics f/# 1.5 AR coated, noise floor 200e-, g = 0.5 96
2.12 Comparison of MWIR, 3μm to 5μm, and LWIR 8μm to 10μm bands. Uniformity 0.03%, pixel area 50μm x 50μm, frame rate 30Hz, optics f/# 1.5 AR coated, noise floor 200e-, g = 0.5 98
2.13 Schematic for QWIP FPA with 1-D diffraction grating.[73] 100
2.14 Schematic of miniband transport QWIP detector[74] 101
2.15 640 x 480 grating coupled QWIP miniband transport FPA operating in the LWIR. 50μm pixel pitch, f/2 optics, 30 frames/second and MRT = 7mK. Subject is Mexicana Airlines Lockheed L1011. Courtesy Charles Parton (Lockheed/Martin) ...102
2.16 640 x 480 grating coupled QWIP miniband transport FPA operating in the LWIR. 50μm pixel pitch, f/2 optics, 30 frames/second and MRT = 7mK. Image taken in dark room with no illumination. Courtesy Charles Parton (Lockheed/Martin) ..103
2.17 Schematic of band diagram for p-type GaAs/AlGaAs QWIP 106

2.18 Schematic of band diagram for p-type $In_{0.30}Ga_{0.70}As/In_{0.52}Al_{0.48}As$ strained layer QWIP grown on InP.[77] **107**

2.19 Schematic of band diagram for X-band $AlAs/Al_{0.50}Ga_{0.50}As$ QWIP. As shown this is a multiwavelength device with responsivity at 2.5μm, 4.3μm and 14.8μm. Resonances are show with dot-dash lines.[77,79] **108**

3.1 Schematic of DCXD system. Inset shows Bragg angles for GaAs at several common crystal orientations. **114**

3.2 DCXD rocking curve and simulation for QWIP 9065. Quantum well width 45Å $In_{0.13}Ga_{0.87}As$, barrier width 510Å $Al_{0.30}Ga_{0.70}As$. **115**

3.3 Schematic representation of photoluminescence experiment **116**

3.4 Comparison of 25K PL intensity for MQW structure containing 10 50Å quantum wells of GaAs separated by 500Å $Al_{0.30}Ga_{0.70}As$ barriers. Circles are for sample grown on bulk GaAs substrate and dots are for sample grown in E-o-E dielectric growth well. Both samples grown simultaneously at 470C and prepared using hydrogen passivation. **117**

3.5 Bandgaps Available for the AlGaInAs system at Room Temperature[66] Solid lines indicate direct band gaps at the Γ point, dotted lines are indirect **119**

3.6 Schematic diagram for elastically strained epitaxial layer **120**

3.7 Schematic diagram for misfits is relaxed epitaxial layers **121**

3.8 Diagram of dislocation formation in zinc-blende crystals. Hatched region is a {111} plane.[96] **122**

3.9 Critical thickness for single epitaxial layer of InGaAs on GaAs **127**

3.10 Schematic diagram of strained layer superlattice **128**

3.11 Critical layer thickness for superlattice as a function of Indium fraction in the strained layer and number of superlattice periods n. Solid curve $n = 50$,dot-dashed curve $n = 25$ and dashed curve $n = 10$. All curves h_2=500Å. **131**

3.12 Critical number of superlattice periods as a function of Indium fraction in the strained layer and strained layer thickness h_1. Solid curve h_1=50Å,dot-dashed curve h_1=40Å and dashed curve h_1=30Å. All curves h_2=500Å. **132**

3.13 Cathodoluminescence from a 25 period superlattice consisting of 50Å InGaAs and 500Å of AlGaAs. Samples grown by the author using MBE, CL courtesy of M. Bulsara. 2% Indium sample has no features to focus on characteristic of high quality defect free material. 12% Indium sample shows large density of individual misfit dislocations along the two perpendicular <110> crystal directions. 25% Indium sample shows very large density of misfits coalesced into groups which obscure the individual dislocation lines. Note also the large shift in PL peak wavelength for the 25% Indium sample. **134**

3.14 Cathodoluminescence from superlattice consisting of 50Å $In_{0.12}Ga_{0.88}As$ quantum wells and 500Å AlGaAs barriers. Samples grown by the author using MBE, CL courtesy of Mayank Bulsara. 5 well sample has no features to focus on characteristic of high quality defect free material. 15 well sample shows large density of individual misfit dislocations along the two perpendicular <110> crystal directions.

25 well sample also shows large density of individual misfits indicating n is beyond the critical number. Note also the large shift in PL peak wavelength for the 25 well sample. 135
3.15 Five samples grown to test SL critical thickness theory of Equation 3.27 plotted with solid line. "x" indicates beyond critical thickness, "o" indicates less than critical thickness. Samples grown using MBE by author, CL done by Mayank Bulsara. 136
3.16 Step 1: MBE Growth of QWIP. Typical epitaxy consists of ~1μm Si doped $1x10^{18}$ GaAs anode, active superlattice region consisting of 10 pairs of Si doped 8 $x10^{17}$ / cm^3 50Å $In_{0.10}Ga_{0.90}As$ quantum wells separated by 500Å undoped $Al_{0.20}Ga_{0.80}As$ barriers and capped with final $Al_{0.20}Ga_{0.80}As$ barrier and 0.5μm Si doped $1x10^{18}$ GaAs cathode. 138
3.17 Step 2: Pixel Isolation Etch. Cross section through QWIP pixel perpendicular to the <110> direction, bottom and perpendicular to <110> direction, right. 20:1:1, $DIH_2O:H_3PO_4:H_2O_2$, etchant reveals <111> and <221> Ga planes. Figure drawn for pixel array. 139
3.18 Step 3: Contact metallization. Figure 3.18a shows a 100μm x 100μm QWIP pixel after ohmic metallization. Figure 3.18b shows the same device after completion of bond pad metallization. 140
3.19 SEM of 200μmx200μm QWIP pixels with completed metallization. Large spot is crystal defect. 141
3.20 Close-up from same SEM micrograph as used in Figure 3.19. Metallization and both etch profiles are visible at the corner of the device 142

4.1 Comparison of D* and MRT as figures of merit for FPAs. LWIR band 8μm to 10μm, f/# 1.5 optics with AR coating, pixel area 50μm x 50μm, noise floor $200e^-$, background temperature 300K, $g = 0.5$. 145
4.2 QWIP Cryogenic measurement sample mounting system 147
4.3 Dark current path through a QWIP without illumination. Quantum well trapping and thermionic emission combine to maintain a constant electron density in the well 149
4.4 Dark Current for QWIP sample 9066. 10 $In_{0.08}Ga_{0.92}As$ wells of width 54Å doped n-type $5x10^{17}$ cm^{-3} with 450Å barriers of $Al_{0.15}Ga_{0.85}As$. 151
4.5 Dark Current for QWIP sample 9066. 10 $In_{0.08}Ga_{0.92}As$ wells of width 54Å doped n-type $5x10^{17}$ cm^{-3} with 450Å barriers of $Al_{0.15}Ga_{0.85}As$. The lowest current point at 40K shows the noise floor of the measurement. 152
4.6 Photocurrent path through a QWIP with illumination. The Dark current remains the same as in Figure 4.1 however with the addition of photo emission extra injection is required to maintain a constant electron density in the well 154
4.7 Measurement setup for 293K black body. For 100μm square QWIP solid angle, W = 4.4 x 10^{-5} str. Photon flux at 10μm for FWHM = 1.5μm, F_{BB} = 6.7 x 10^{16} photons/cm^2 s str. Incident power 1.5 x 10^{-7} Watts. 155
4.8 Measurement setup for 500K black body. For 100μm square QWIP solid angle, W

= 8.2 x 10^{-6} str. Photon flux at 10μm for FWHM = 1.5μm, F_{BB} = 5.4 x 10^{17} photons/cm^2 s str. Total incident power 2.5 x 10^{-7} Watts. 155
4.9 Photocurrent for sample 9066 illuminated with black body at 500K, solid line and illuminated with 293K black body, dashed line. Bold dashed vertcal lines delineate the knees in the curve which define the tunneling dominated, photocurrent dominated (operating region) and low voltage regions of the photocurrent curves. 156
4.10 Dark current, solid lines from top down at temperatures of 80K, 70K, 60K and 40K. Photo-current, dashed line for black body of 293K. BLIP temperature estimated to be 68K. 158
4.11 Spectral Response Measurement setup at ADT 160
4.12 Sample 9065, dots connected by dashed line and Sample 9056, dots connected by solid line. 9065 contains 5 50Å $In_{0.15}Ga_{0.85}As$ wells and 500Å $Al_{0.30}Ga_{0.70}As$ barriers, target 5.9μm. 9056 contains 4 50Å $In_{0.15}Ga_{0.85}As$ wells and 500Å $Al_{0.20}Ga_{0.80}As$ barriers, target 8.4μm. 161
4.13 Comparison of D* for best and mean TE mode QWIPs. LWIR band 8μm to 10μm, f/# 1.2 optics with AR coating, noise floor 200e^-, background temperature 293K, g = 0.5. Uniformity for 12bit ADC of 0.007% indicated by arrow 163
4.14 Wright Laboratory measured X-ray data for API Sample.[105] 165
4.15 API measured photo current for 200μm diameter device. Top down: solid line 80K, dotted line 70K, dash-dot line 60K dashed line 50K. 166
4.16 WL measured spectral responsivity for API device.[105] 167

5.1 MESFET Cross Section 172
5.2 Comparison of Shockley Model with Measured 20μmx5μm DFET. 176
5.3 Comparison of Shockley Model with Measured 3μmx1.2μm DFET. 177
5.4 Velocity Overshoot and Velocity Saturation in GaAs. 179
5.5 Equivalent Circuit Model for GaAs MESFET. 179
5.6 Comparison of HSPICE Model with Measured 20μmx5μm DFET. 181
5.7 Comparison of HSPICE Model with Measured 3μm x 1.2μm DFET. 181
5.8 Output parameter extraction for 20μm x 5.0μm DFET. V_{gs} from -0.2V to 0.8V in 0.2V steps. See Table 5.2 for output resistance values. 183
5.9 Output parameter extraction for 3μm x 1.2μm DFET. V_{gs} from -0.2V to 0.8V in 0.2V steps. See Table 5.3 for output resistance values. 184
5.10 I-V Characteristics for 20μm x 5.0μm DFET. 187
5.11 I-V Characteristics for 3μm x 1.2μm DFET. 187
5.12 Transconductance and threshold voltage extraction for 3μm x 1.2μm EFET. 188
5.13 Temperature dependence of transconductance. 189
5.14 Temperature dependence of threshold voltage. 189
5.15 Source resistance extraction for 3μm x 1.2μm EFET. 190
5.16 Source resistance temperature dependence. 191
5.17 Drain resistance extraction for 3μm x 1.2μm EFET. 191
5.18 Drain resistance temperature dependence. 192
5.19 Schottky diode ideality extraction for 20μm x 5μm EFET with V_{ds} = 0. Ideality fac-

tors are 4.73, 3.89, 3.26, 2.44, 1.89, 1.40, 1.27, 1.13, 1.07 and 1.05 for temperatures of 30K, 40K, 50K, 70K, 100K, 150K, 200K, 250K, 300K and 350K respectively. Measurement noise floor ~100pA. 193
5.20 Schottky diode ideality factor temperature dependence for 20μm x 5μm EFET with $V_{ds} = 0$. Data from Figure 5.19. 193
5.21 Schematic of Simple Pixel Layout and Circuit Diagram. 194
5.22 Schematic two inverter section of Ring Oscillator. 196
5.23 Graphical load line calculation of inverter characteristics. 197
5.24 Ring Oscillator output stage 198
5.25 Ring Oscillator HSPICE Simulation. 198
5.26 HSPICE Parameter Extraction Proceedure 200

6.1 Schematic of Hybrid Focal Plane Array. 204
6.2 Problems with Hybridization - unzipping due to thermal expansion coefficient mismatch between GaAs, $6.9 \times 10^{-6}\ K^{-1}$ and Silicon, $2.6 \times 10^{-6}\ K^{-1}$. 205
6.3 E-o-E OEIC fabrication flow showing Vitesse Semiconductor fabricated VLSI GaAs MESFET chip with cleaned growth well (a); same chip after MBE growth of LED structure resulting in crystalline material in growth well and polycrystalline material above circuits (b); completed chip after removal of polycrystalline GaAs and upper level metallization (c). Courtesy K.V. Shenoy 209
6.4 Vitesse HGaAs3 metal sheet resistance as a function of 5 hour thermal cycle temperature. Courtesy Eric Braun[122]. 210
6.5 Arrhenius curve set with an activation energy of 3.5eV and experimentally measured Vitesse HGaAs3 metal 1 sheet resistance thermal cycle response. Courtesy Eric Braun.[122] 211
6.6 Schematic comparison of the dielectric growth well (DGW), (a) and metallic growth well (MGW) (b) structures as they appear after returning from Vitesse Semiconductor. For DGW pre-growth processing includes solvent clean and buffered oxide etch (BOE) etch. For MGW pre-growth processing includes solvent clean, metal 1 etch and BOE etch. 212
6.7 Effect of including metal etch stop layer on crystal quality after. 1μm GaAs growth in growth well on VLSI MESFET chip. With Al etch stop layer, crystal shows no defects using Nomarski at 1000x. (a) Without Al etch stop layer, large defects due to RIE surface damage are clearly visible. (b) Note: small features on both photographs are due to microscope contamination. Growth and photographs courtesy J. Ahadian and S. Patterson. 213
6.8 Schematic of QWIP FPA designed using E-o-E. Side view including backside illumination (a). Pixel layout drawn to scale for 12bit dynamic range, 2.7pF capacitor, MESFET gate capacitance of 1.2fF/μm^2 (b). 215
6.9 Schematic of QWIP FPA with lenses designed using E-o-E (a) and SEM of lens array from Smith et al. [137] 216
6.10 Schematic illustration of fabrication flow for E-o-E/SAW integration. QWIP FPAs are etched using either RIE or wet etching to define the pixel size and isolate the

pixels (a). Pillars are grown on pre-fabricated MESFET VLSI ROICs using E-o-E (b). Finally, QWIP FPA is wafer bonded on to prepared VLSI GaAs MESFET ROIC (c). Illumination is from top into QWIP FPA..219

6.11 Forward IV characteristics of n-GaAs/n-InP wafer bonded samples. InP biased positive with respect to GaAs. Wafer bonding done at temperature for 30 minutes in Hydrogen with bonding force of 0.29 N/cm^2. Barrier heights are 0.46, 0.42, 0.40 and 0.37eV for bonding temperatures of 450C, 550C, 650C and 700C respectively.[143] 221

6.12 Cross-sectional transmission electron micrograph of GaAs/InP heterointerface bonded at 650C for 30 minutes. The misfit dislocations are indicated by arrows. From H. Wada.[143] 223

6.13 Schematic illustration of fabrication flow for SAW integration process. QWIP FPAs are etched using either RIE or wet etching twice; first to define the pixel size and isolate the pixels (a) and then to define pillars used for contact to MESFET ROIC (b). Finally, QWIP FPA is wafer bonded on to prepared VLSI GaAs MESFET ROIC (c). Illumination is from top into QWIP FPA..................................224

7.1 Advantages of TE mode normal incidence QWIP FPAs....................................232

List of Tables

Table #	Table Caption	Page #
2.1	Bulk Kane ***k=0*** states for 14 band model.[44,45]	79
2.2	Comparison of parameters for 8 and 14 band models for GaAs[46]	84
2.3	Comparison of experimental and calculated masses for GaAs	84
2.4	***K•p*** SL band structure parameters[46]	90
3.1	Mechanical properties of AlGaInAs material system	126
3.2	Mechanical compliances for selected III-V semiconductors[66]	126
4.1	Dark Current Characteristics for QWIP samples	153
4.2	Electrical and optical properties of QWIPs.	159
5.1	HSPICE parameters for Figures 5.6 and 5.7	182
5.2	Output Characteristics for 20µm x 5µm DFET	183
5.3	Output Characteristics for 5µm x 1.2µm DFET	184
5.4	Gate Capacitance Extraction	199
5.5	HSPICE Parameters for 20µm x 5µm EFET	201
5.6	HSPICE Parameters for 20µm x 5µm DFET	201

Chapter 1

Introduction to Infrared Focal Plane Arrays

1.1 Thesis Organization

The subject of this thesis is infrared (IR) imaging systems. As such the journey starts in Chapter 1 with a broad historical look at the key discoveries which laid the groundwork for modern infrared systems in general and this thesis in particular. Particular attention is paid to the development of the figures of merit which measure progress toward improved IR imaging systems. Figures of merit warrant this emphasis because they are so important in both evaluating progress and determining the direction of future work. To facilitate later analysis of the dependence of imaging system performance on discrete detector performance, a model of an infrared camera that includes discrete detector and system parameters will be developed. Chapter 1 finishes with a discussion of the wide variety infrared detectors using the figures of merit discussed to assess the strengths and weaknesses of the various technologies for a sampling of current and future applications. Chapter 2 focuses on the physics of QWIP focal plane arrays in general and the normal incidence quantum well intersubband photodetectors (QWIPs) which are the central focus of this work in particular, with an emphasis on the trade-offs involved over competing technologies and the specific applications in which the author expects this technology to excel. Chapter 3 looks at the molecular beam epitaxial (MBE) growth of the QWIPs used in this work as well as their fabrication and the specific technological issues brought about by both growth and fabrication methods chosen. Chapter 4 looks at the cryogenic measurements of the world's first normal incidence n-type QWIP without a coupling grating demonstrated as part of this thesis. Specific attention is paid to the adequacy of the devices demonstrated to the applications envisioned in Chapter 2, possible avenues to improved performance and potential pitfalls to final commercialization. Chapter 5 looks at the modeling and cryogenic measurement of the commercially available GaAs VLSI MESFETs proposed for use in monolithic integration of the normal incidence QWIPs with the required read

out integrated circuits (ROICs). Chapter 6 looks at the progress made toward monolithic integration using the epi-on-electronics (E-o-E) technology developed by K.V. Shenoy and C.G Fonstad, Jr.[1] Some mention will be made of the contributions of the author to the E-o-E technology, and the potential of this technology for commercialization of starring FPA cameras. A description is also given of a new selective area wafer bonding (SAW) technology proposed by the author for use in conjuction with E-o-E to address the two weaknesses inherent in the E-o-E technology; requirement for low temperature MBE growth and lateral integration of electrical and optical devices. Finally, Chapter 7 will summarize the accomplishments of this thesis and discuss directions for further work.

1.2 Historical Introduction

Though heat has been sensed and used by humans since the onset of the species, it was not until 1800 that Sir William Herschel[2] (1738 - 1822) identified heat as infrared radiation by dispersing sunlight through a prism and measuring the rise in temperature of a thermometer as a function of spacial position. He found that the largest rise in temperature occurred in the spectral region "below red" and thereby identified "infrared" radiation. The thermometer thus became the first infrared thermal detector. The first infrared photon detector would have to wait more than a century until 1904 when J.C. Bose[3] showed that naturally occurring galena, lead sulfide, exhibited the photovoltaic effect when exposed to infrared radiation. This photon detector was fundamentally different from the thermal detector of Herschel in that its sensitivity varied as a function of incident radiation wavelength, for PbS dropping to zero for infrared radiation with a wavelength longer than about 3μm. Lead salts were the infrared detectors of choice through World War II and continue to be used today, but with the discovery of the transistor at Bell Laboratories in 1947 by John Bardeen and Walter Brattain[4] and the explaination thereof the following year by William Shockley[5], large scale research into the physics and technology of all manner of semiconductors lead to the development of new semiconductor based photon detectors whose spectral response is better suited to remote infrared imaging applications.

Section 1.2 - Historical Introduction

The most important of these semiconductors for infrared imaging has historically been Mercury-Cadmium-Telluride (MCT) because variations in the ternary composition allow the bandgap to be tuned from the visible to beyond 16μm including the three important atmospheric infrared windows. These photon detectors are termed "intrinsic" since the photons absorbed generate electron-hole pairs across the bandgap *both* of which are mobile in the semiconductor and contribute to the electrical photoresponse. "Extrinsic" photodetectors in which a *single* mobile carrier is generated by absorption of a photon at a fixed impurity site are also possible and even preferable in certain applications. A variation of the extrinsic photodetector using intersubband transitions in GaAs based quantum wells, the quantum well intersubband photodetector (QWIP) was first suggested by Esaki and Sakaki at IBM in 1977[6] and demonstrated by Levine[7] at AT&T in 1987.

Simultaneous with the developing understanding of semiconductors the invention of the integrated circuit by Jack Kilby at Texas Instruments and Robert Noyce at Fairchild in 1959 sparked a revolution in the Silicon industry which has driven up the density of Silicon ICs at a rate of one octave every 18 months for more than thirty years. As the understanding of the semiconductors used in IR detectors has matured the emphasis has also shifted from individual devices to the development of imaging systems using the technology first developed for the Silicon industry. The first of these imaging systems used a single IR detector and a mechanical scanning system which used rotating mirrors to scan a two dimensional image onto the single detector. These systems were slow and had poor temperature resolution because the time for collecting data at each object pixel was so short. To get a 30Hz image with a spacial resolution equivalent to a television, 640 x 480 the dwell time at each pixel was only 108nS, much too short to make reasonable images. The next generation of systems used a linear array of devices to increase the dwell time by a factor of 640 to 694μS. Today technology borrowed from the Silicon industry is being applied to the development of the two dimensional infrared focal plane array (FPA), or starring array, used in IR cameras analogous to the common CCD video camera used in visible light cameras. In these arrays each detector element stares at the object for the full frame time, 33mS for a 30Hz system, thus giving the optimum signal to noise ratio.

1.3 Infrared Physics Background

This thesis focuses on infrared imaging using FPAs which implies both that the object to be imaged must emit radiation in the spectral range in which the detector is sensitive and that the infrared radiation emitted can travel with acceptable attenuation from the object to the detector. The start for addressing the issue of emission is the expression for the spectral energy density in a black body postulated by Max Plank in 1900 and for which he received the Nobel prize in physics in 1918.

$$\mu_{\omega} = \frac{h}{\pi^2 c^3} \frac{\omega^3}{\exp\left(\frac{h\omega}{2\pi kT}\right) - 1} \qquad \frac{J}{cm^3 \cdot Hz} \tag{1.1}$$

Figure 1.1 shows the black body spectral density for several temperatures as a function of wavelength. The lowest curve is that for a 300K black body which is a good initial guess for the spectral density of many of the objects; people, tanks, airplanes, etc. that may be of interest for imaging.

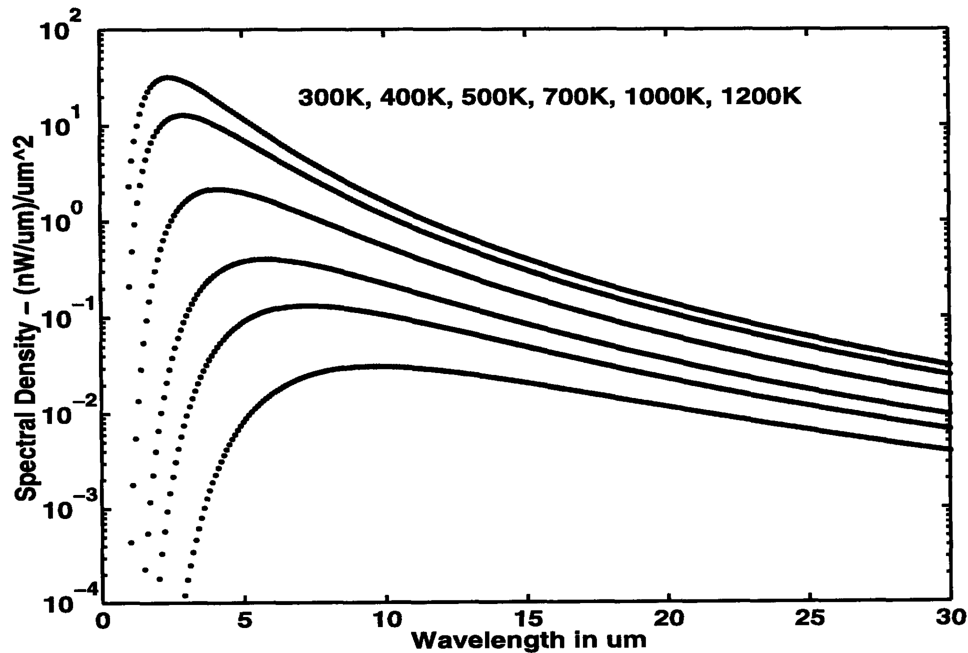


Figure 1.1: Black Body Spectral Density

The peak intensity occurs near 10μm falling off exponentially toward shorter wavelengths and more gradually toward longer wavelengths. An object that emits radiation with a spectral distribution given by Equation 1.1 is said to emit as a grey body. If the object is also a perfect absorber which by Kirchhoff's law also implies that the object is a perfect emitter, emissivity $\varepsilon = 1$, the object is a black body. The power emitted by such an object per unit object area per unit wavelength is given by:

$$W_\lambda = \frac{2\pi c^2 h}{\lambda^5} \frac{1}{\exp\left(\frac{hc}{\lambda kT}\right) - 1} \qquad \frac{Watts}{cm^2 \cdot \mu m} \tag{1.2}$$

Grey bodies emit with the same spectral distribution, but with power given by Equation 1.2 attenuated by the emissivity factor $\varepsilon < 1$. For the special case of a perfect reflector $\varepsilon = 0$; a perfect reflector does not radiate.

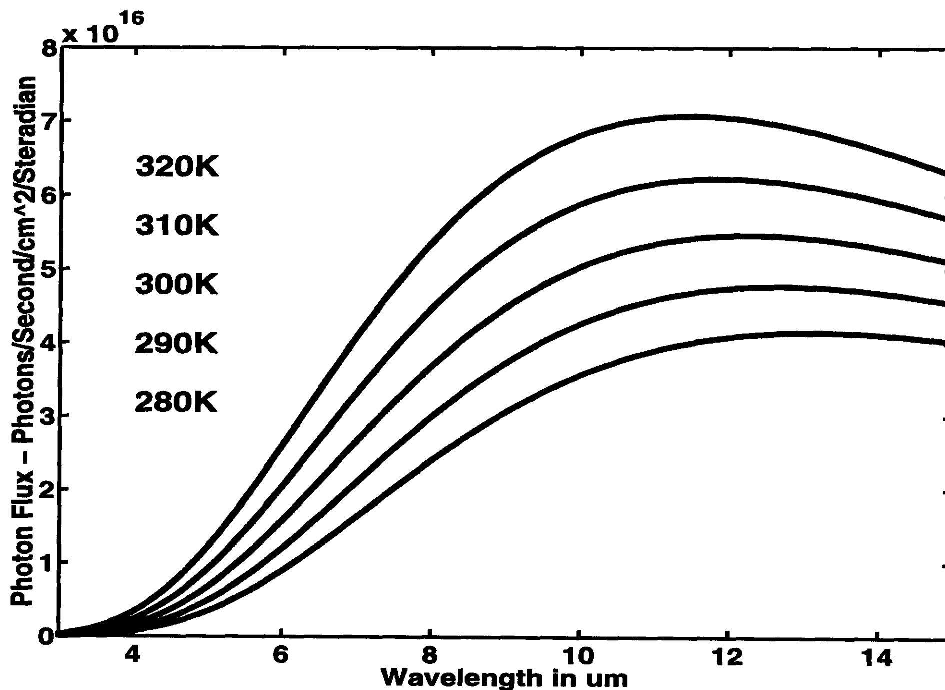


Figure 1.2: Photon Flux for Black Body near RT with a bandwidth of 1μm

For photon counting detectors, it is often more convenient to write Equation 1.2 in terms of the photon number by dividing by the photon energy hc/λ, dividing by 2π to

normalize to the field of view of the detector in Steradians and integrating over the wavelength range over which the detector is sensitive. The result is Equation 1.3.

$$\Phi_P = \int_{\lambda_L}^{\lambda_H} \frac{2c}{\lambda^4\left(\exp\left(\frac{hc}{\lambda kT}\right) - 1\right)} d\lambda \qquad \frac{Photons}{s \cdot cm^2 \cdot Steradian} \tag{1.3}$$

Figure 1.2 shows the result of integrating Equation 1.3 over a one micron bandwidth as a function of center wavelength for a series of black body temperatures near 300K.

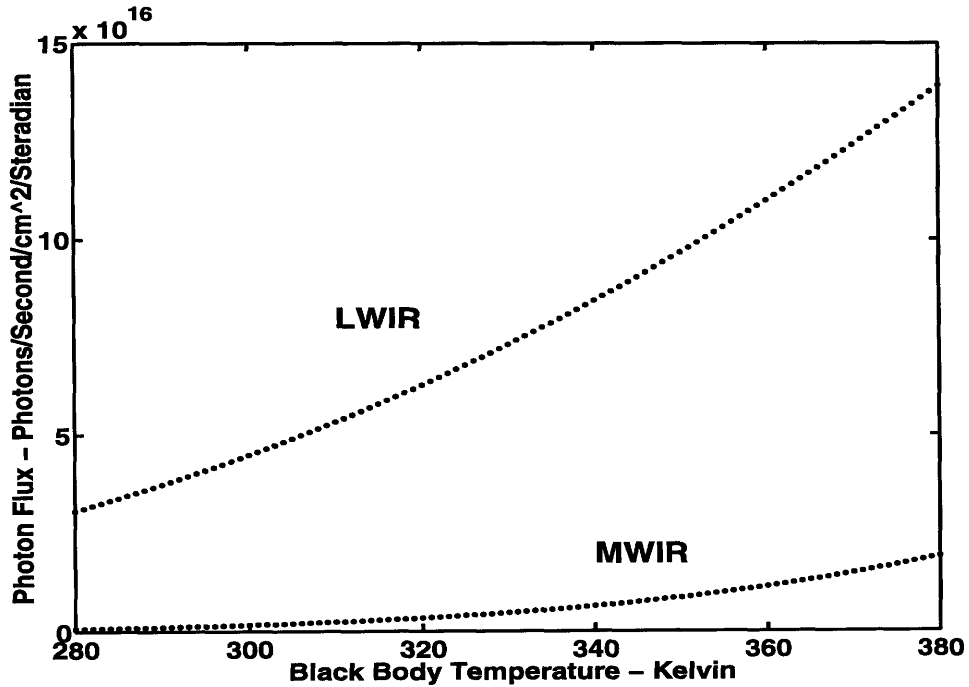


Figure 1.3: Photon flux for detector with 1μm bandwidth and center frequency in the midwave IR (MWIR) 4μm or long-wave IR (LWIR) 9μm.

Figure 1.3 again shows the photon flux for a bandwidth of 1μm, this time as a function of black body temperature with fixed center wavelengths of 4μm, labeled as mid-wave infrared (MWIR), and 9μm, labeled as long-wave infrared (LWIR). For infrared imaging applications in which the camera needs to resolve temperature variations within a scene with an average black body temperature given by the abscissa of Figure 1.3, the ordinate of Figure 1.3 gives the value of the average infrared background flux. Notice that the

LWIR band offers about an order of magnitude more photons than the MWIR band. In addition, the solar background which can be modeled to first order as a black body with a temperature of ~5800K, is lower by about two orders of magnitude.

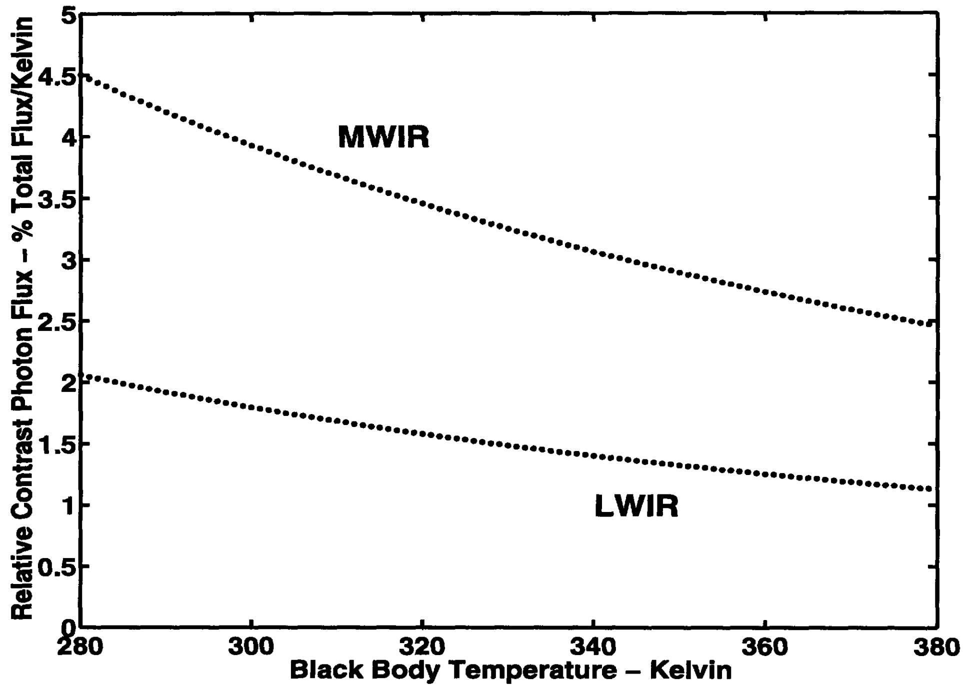


Figure 1.4: Contrast Photon Flux for MWIR (4μm) and LWIR (9μm) bands with detector bandwidth of 1μm

The contrast in photon flux between points within an image that have slightly different temperatures is given by the derivative of Equation 1.3 with respect to temperature.

$$\frac{d\Phi_P}{dT} = \frac{hc}{\lambda k T^2}\Phi_P \tag{1.4}$$

$d\Phi_P/dT$ is shown ratioed to the total background photon flux in Figure 1.4. Notice that the values are on the order of 1 percent per degree Kelvin. Since good quality images typically require resolution on the order of 10mK or better, the desired signal is buried in a background nearly four orders of magnitude larger than the signal itself. Of course this would be a simple matter of subtraction if it were not for the fact that the background is not continuous but rather the result of discrete photons whose arrival time at the detector is

determined by Poisson statistics and therefore introduces shot noise in the background photon flux. Also note that the contrast ratio for the MWIR band is approximately a factor of two larger than for the LWIR band. This can be used to advantage in some MWIR FPAs leading to better performance in spite of the lower photon flux in the MWIR.

The other major factor that determines the wavelength of choice for imaging is the absorption and scattering properties of the atmosphere. The atmosphere is an extremely complicated medium with constituent gases giving a complex spectrum of molecular absorption lines that change in intensity with weather conditions, altitude, latitude, season, ground conditions and many other factors. In addition natural and man-made particulates act as scattering centers for radiation in the infrared. Fortunately, the atmosphere has been studied intensively and very accurate experimental data is available for a wide variety of possible conditions.

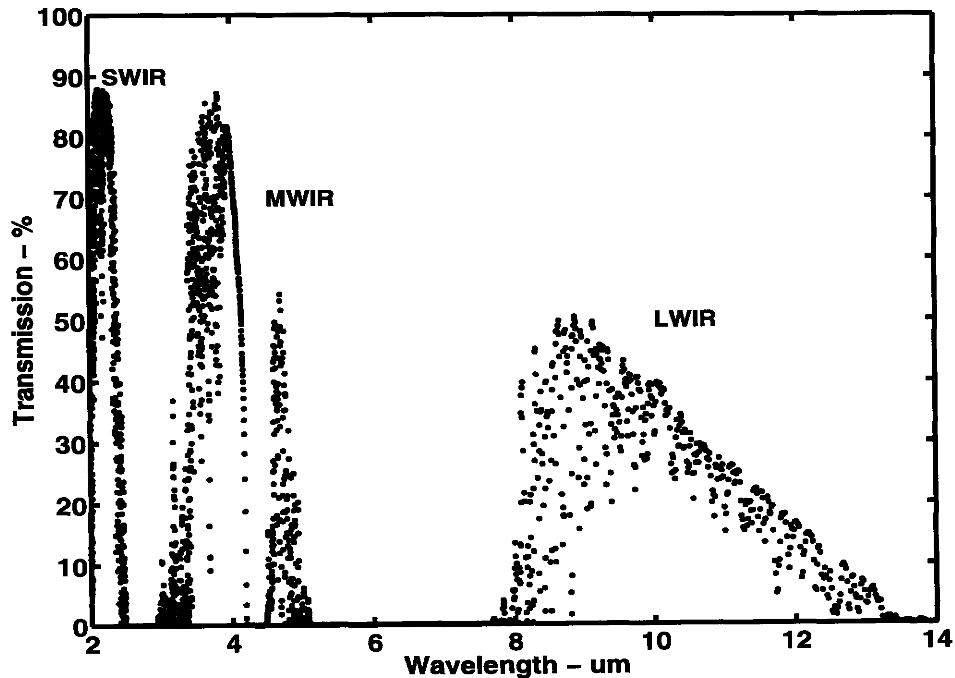


Figure 1.5: MODTRAN3 calculation of atmospheric transmission through a distance of 5km at mid-latitude during summer with visibility of 5km and no rain over farmland.

As an example, Figure 1.5 shows the transmission over a distance of 5km at sea level dur-

ing summer at mid-latitude under clear weather conditions over farmland. The principle features are determined by the absorption of water, CO_2 and particulates. The data was calculated using the US military's MODTRAN3[8,9,10] program which has a resolution of $1cm^{-1}$ over the entire spectral range shown. The three atmospheric transmission "windows" are labeled short-wave infrared (SWIR) from 1.8μm to 2.5μm, mid-wave infrared (MWIR) from 3μm to 5μm and long-wave infrared (LWIR) from 8μm to 12μm. Note the fortuitous coincidence between the peak in the spectral density for a black body at 300K near 10μm and the LWIR transmission band between 8μm and 12μm can be seen. The LWIR spectral band has nearly an order of magnitude larger spectral density than the MWIR spectral band for a 300K black body making LWIR the preferred detector band in many surveillance applications. A more detailed discussion of the criteria used to determine the optimum detector including the detection band will be forthcoming after a very important discussion of the figures of merit appropriate for infrared detection.

1.4 Figures of Merit for IR detectors and Focal Plane Arrays

Fundamental constants like Plank's constant, h, relate physically measurable quantities like photon energy and frequency through the system of units we choose to use. i.e. the physics in the equation $E = h\upsilon$ is the statement that the energy and frequency of a photon are directly proportional to each other. The value of the proportionality constant, h, depends on the units we use; for MKS units this value is $6.626x10^{-34}$ $Joule\ Second$ but once the units are chosen, the value of h is fixed by the physics and in this sense it is a "fundamental" constant. Figures of merit are not fundamental, but rather are attempts to rank a variety of different things on a common scale of "goodness". As an example, contrary to popular belief, Apples and Oranges are comparable if an appropriate figure of merit, for example sugar content, is chosen. One can unequivocally state then that Oranges are "better" than Apples in terms of sugar content. Note the very important point that figures of merit are in some sense subjective and application dependent; one would certainly not want to choose a fruit for a pie in terms of the sugar content figure of merit since it would imply that Oranges were better for making pie than Apples. Clearly to

determine the optimum fruit for making a pie requires a modification of the figure of merit to include other than just sugar content. The discussion of infrared focal plane arrays (FPAs) in the following sections will accentuate the application dependence of figures of merit in the field of infrared detection and again it will be necessary to modify the standard figures of merit to include effects unique to this application.

Typically, figures of merit will attempt to remove the effects of easily adjustable parameters; for infrared devices these might be device area, measurement bandwidth, etc. in favor of what are considered more fundamental device characteristics like signal to noise ratio, where noise is properly defined as having contributions from noise sources like photon shot noise, Johnson noise and array uniformity which are inherent in the operation of the device and/or difficult to control in device fabrication. Note again, that what one determines to be “easily” adjustable is somewhat subjective and is likely to change as technology changes and new applications emerge. For infrared detectors, the initial work on figures of merit was done by R.C. Jones[11,12,13] at Polaroid Corporation in the 1950’s in whose honor the unit $cm \cdot \sqrt{Hz}/Watt$ has been named the “Jones”. This work will be reviewed here as it applies to individual detector elements. It will be shown that for applications involving large arrays of detectors some of Jones’s single detector figures of merit are inappropriate primarily because they ignore spacial variation among detector elements within the array and need to be replaced by other figures of merit more representative of the actual performance observed in large starring arrays.

1.4.1 Responsivity - R_λ, R_{BB}:

The responsivity is a measure of the dependence of the signal output from a detector upon the radiant power incident on the detector. Since detectors are in general not spectrally flat, it is important to specify the type of source. If the source is nearly monochromatic as from the output of a spectrometer one speaks of the spectral responsivity using R_λ for the symbol and specifying the wavelength at which the measurement is made. R_λ^P is a special case of the spectral responsivity where the subscript “P” refers to the peak spectral response. If the source is a black body, one speaks of the black body responsivity

using the symbol R_{BB} and specifying the temperature of the black body source. The black body responsivity is related to the spectral responsivity by:

$$R_{BB} = \frac{\int_{\lambda_1}^{\lambda_2} R_\lambda W_\lambda d\lambda}{\int_{\lambda_1}^{\lambda_2} W_\lambda d\lambda} \qquad Amps/Watt \tag{1.5}$$

where integration is done over the spectral range from λ_1 to λ_2 over which the detector is sensitive and W_λ is given by Equation 1.2. To reduce measurement noise the incident radiation is usually chopped. Since some detectors exhibit 1/f noise, it is important to specify the frequency of measurement. Units for responsivity are *Amps/Watt* for current generating devices like photoconductors or photodiodes and *Volts/Watt* for photovoltaic and most thermal detectors.

1.4.2 Noise Equivalent Power - NEP

NEP is the root mean square, RMS incident radiant power which gives a signal current equal to the RMS noise current. The source, either monochromatic or black body must be specified and then the NEP is inversely related to responsivity by:

$$NEP = \frac{i_N}{R} \qquad \frac{Amps}{Amps/Watt} \tag{1.6}$$

where i_N is the noise current. Since the performance of infrared detectors is limited by noise, NEP which can be thought of as the signal intensity required to give a signal to noise ratio (*SNR*) of 1, is more indicative of image quality than responsivity. However, since NEP does depend on detector area, field of view, noise bandwidth and signal frequency; these parameters must also be specified.

1.4.3 Detectivity - D*

D^* is an attempt to remove from NEP the dependence on device area and measurement bandwidth to facilitate comparison between devices. D^* is defined as:

$$D^* = \frac{\sqrt{A_D \Delta f}}{P} SNR = \frac{R\sqrt{A_D \Delta f}}{\langle i_N \rangle} \qquad Jones \tag{1.7}$$

where A_D is the area of the detector, Δf is the measurement bandwidth, P is the incident

optical power, *SNR* is the signal to noise ratio, i_P/i_N, in either current or voltage, R is the responsivity, i_N is the noise current given below in Equation 1.10 and i_P is the photocurrent given below in Equation 1.9. For a signal to noise ratio of unity this reduces to:

$$D^* = \frac{(A_D \Delta f)^{1/2}}{NEP} \qquad Jones \tag{1.8}$$

The particular functional dependence on A_D and Δf comes from the assumption that the detector is limited by shot noise related to the Poisson distribution in the arrival times of photons at the detector. Consider for example a detector with quantum efficiency η under illumination with a photon arrival rate of Φ_P photons per second per steradian per unit detector area. The current generated in the detector is then:

$$i_P = \eta q A_D \Phi_P \tau_{op} \Omega_{op} g \qquad Amps \tag{1.9}$$

where τ_{op} is the optical transmission efficiency in photons/photon, Ω_{op} is the solid angle subtended by the lens, in terms of the f/#, $\Omega_{op} = \pi/(4(f/\#)^2 + 1)$, A_D is the detector area and Φ_P is the photon flux as given by Equation 1.3. The shot noise rms current associated with this current is simply:

$$i_N = \sqrt{4 q i_P g \Delta f} \qquad Amps \tag{1.10}$$

so that by multiplying the *SNR* by A_D and Δf both the dependence on detector area and measurement bandwidth are eliminated under the assumption that the detector is shot noise limited. As with responsivity and NEP, D^* is dependent on the type of source, black body or monochromatic, used to illuminate the detector. As an example, for a detector whose responsivity as a function of wavelength is Gaussian centered at 9μm with a FWHM of 1.0μm, D^*_λ is a factor of two larger than D^*_{BB}. In terms of detector parameters and the photon flux as given by Equation 1.3, D^*_{BB}. is given by:

$$D^*_{BB} = \frac{\lambda}{2hc}\sqrt{\frac{\eta}{\Phi_P \tau_{op} \Omega_{op}}} \qquad Jones \tag{1.11}$$

1.4.4 Noise Equivalent Temperature Difference - NETD

Noise equivalent temperature difference, NETD, has several definitions in the literature, but will be defined here as the increase in temperature above the background necessary to get a signal to noise ratio of one assuming only *shot* noise. Other noise sources sometimes included in NETD will be dealt with separately. In terms of D*, NETD is:

$$NETD = \frac{(A_D \Delta f)^{1/2}}{D^*_{BB}\left(\frac{dP_{BB}}{dT}\right)} \tag{1.12}$$

where P_{BB} is the integrated blackbody power in the spectral range over which the detector is sensitive as shown in Figure 1.1. As a numerical example, a 50μm x 50μm detector with a D* of 10^{10} Jones at a peak wavelength of 10μm using f/2 optics and a bandwidth of 60Hz has an NETD of 10mK. In terms of detector parameters and the photon flux as given by Equation 1.3, NETD is given by:

$$NETD = \frac{\lambda k T^2}{hc} \frac{1}{\sqrt{A_D \tau_{op} t_f \Omega_{op} g \eta \Phi_P}} \qquad Kelvin \tag{1.13}$$

where A is the area of the detector, τ_{op} is the optical transmission efficiency in photons/photon, t_f is the frame integration time, Ω_{op} is the solid angle subtended by the lens in terms of the f/#, $\Omega_{op} = \pi / (4(f/\#)^2 + 1)$, g is the photoconductive gain in electrons/electron η is the quantum efficiency and Φ_P is the photon flux as given by Equation 1.3.

1.4.5 Minimum Resolvable Temperature - MRT

D^* and NEDT are extremely valuable and widely used figures of merit for discrete detectors however, for FPAs which may contain a million or more discrete detectors, overall image quality is often not determined by the shot noise limit inherent in the definitions of D^* and NEDT. Instead system parameters like detector uniformity, read out integrated circuit (ROIC) noise and pixel capacitor size often limit image quality. The first two of these issues will be addressed here using a model of an infrared camera developed by J.M. Mooney et.al.[14]. The issue of pixel capacitor size is very important but will be left to

the discussion in Chapter 6 on ROICs. From this a new figure of merit, the minimum resolvable temperature, MRT, will be defined which includes the system parameters important for FPAs and which in the limiting case when shot noise is the dominant noise source will reduce to NEDT. Figure 1.6 shows the simplified model of an IR FPA camera. Two sources of electrons contribute to the total charge, N^T, accumulated in one frame time, t_f, at each pixel. N^P is charge accumulated due photons detected at the pixel including background photons and signal photons.

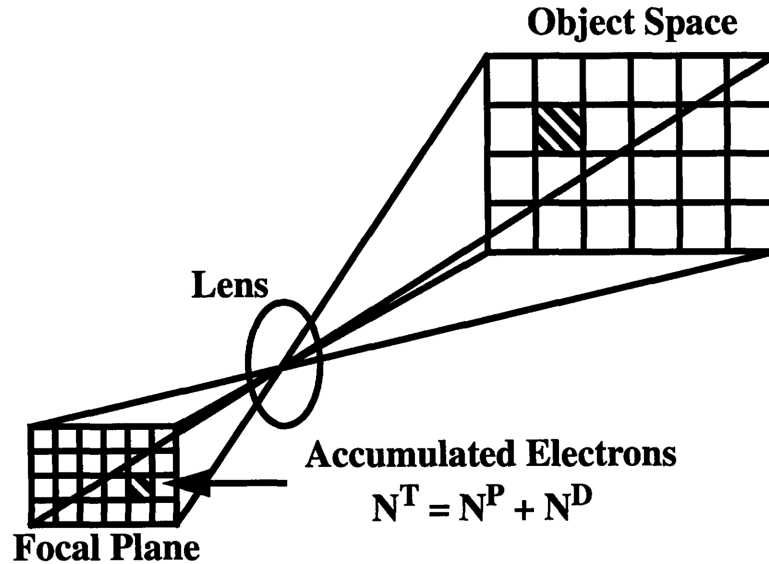


Figure 1.6: Simplified model of an IR FPA Camera

N^D is charge accumulated due to the dark current at each pixel and is independent of photon flux. Electrons due to the ROIC noise floor are also included in this picture in the term N^D. The total electron accumulation at a pixel with coordinates i,j is:

$$N^T{}_{i,j} = R_{i,j} \int_{\lambda_1}^{\lambda_2} \eta_{i,j}(\lambda) \{ \varepsilon_{i,j}(\lambda) W_{i,j}(\lambda, T_{i,j}) + [1 - \varepsilon_{i,j}(\lambda)] W_{i,j}(\lambda, T_B) \} d\lambda + N^D_{i,j} \quad (1.14)$$

where $\eta_{i,j}$ is the net quantum efficiency of the i,j FPA pixel as a function of wavelength, $\varepsilon_{i,j}$ is the emissivity of the i,j object pixel as a function of wavelength, $W_{i,j}$ is the black body spectral response function given in Equation 1.2 as a function of wavelength and object pixel temperature, the term $[1-\varepsilon_{i,j}(\lambda)]$ is the reflectivity of the i,j object pixel included to account for the fact that for objects that are not black bodies the object will have a non-zero reflection as well as emission, $N^{D}_{i,j}$ is the dark current charge accumulated at the i,j FPA pixel during a single frame time and $R_{i,j}^{BB}$ is the black body responsivity of the i,j FPA pixel given by:

$$R_{i,j}^{BB} = A_{i,j}\tau_{op}t_f\Omega_{op}g\cos^4(\theta_{i,j}) \tag{1.15}$$

where $A_{i,j}$ is the area of the i,j FPA pixel, τ_{op} is the optical transmission efficiency in photons/photon, t_f is the frame integration time, Ω_{op} is the solid angle subtended by the lens in terms of the f/#, $\Omega_{op} = \pi/(4(f/\#)^2+1)$. g is the photo-conductive gain in electrons/electron, $\theta_{i,j}$ is the *i,j* FPA pixel angular displacement from the optical axis and the cosine term represents the systematic variation of pixel illumination with position at the FPA.

The integral term in Equation 1.14 is due to the incident photon flux and the $N^{D}_{i,j}$ term is due to detector dark current. In defining the MRT both temporal and spacial variations across the detector FPA need to be included. The temporal noise in each FPA pixel can be represented as the sum of two terms:

$$\sigma^{t}_{i,j} = \sqrt{\left(\sigma^{p}_{i,j}\right)^2 + \left(\sigma^{f}_{i,j}\right)^2} \tag{1.16}$$

The term $\sigma^{p}_{i,j}$ is the photon signal dependent variance, which is typically dominated by shot noise and is therefore from Equation 1.10 equal to the square root of the number of electrons collected $N^{T}_{i,j}$. The term $\sigma^{f}_{i,j}$ is the additive noise floor associated with the ROIC and is therefore independent of $N^{T}_{i,j}$. This term also includes the noise due to the spacial and temporal average dark current $\langle\overline{N^{D}_{i,j}}\rangle$ which is also photon signal independent. The spacial non-uniformity in the dark current is included in the σ^{s}_{FPA} term of Equation 1.16. The spacial noise is measured as a fraction of the mean electron number across the entire array, commonly referred to as the nonuniformity. The nonuniformity, U, is thereby

defined by the expression:

$$\sigma^{s}_{FPA} = U\langle\overline{N^{T}{}_{i,j}}\rangle \tag{1.17}$$

where the subscript FPA on s implies that $\sigma^{s}{}_{FPA}$ is defined for the entire FPA, the angle brackets <> denote spacial averaging and the overbar denotes temporal averaging. This spacial noise includes effects due to variations in quantum efficiency, variations in g factor, variations in quantum efficiency all of which are wavelength dependent as well as variations in device area over the FPA. The total signal dependent portion of the noise is then given by the sum of Equations 1.16 and 1.17 giving a total time and space averaged signal to noise ratio of:

$$SNR = \frac{\langle\overline{N^{T}_{i,j}}\rangle}{\sqrt{\langle\sigma^{t}_{i,j}\rangle^{2} + \langle\sigma^{s}_{FPA}\rangle^{2}}} \tag{1.18}$$

substituting in Equations 1.16 and 1.17 gives:

$$SNR = \frac{\langle\overline{N^{T}_{i,j}}\rangle}{\sqrt{\langle\sigma^{f}_{i,j}\rangle + g\langle\overline{N^{T}_{i,j}}\rangle + U^{2}\langle\overline{N^{T}_{i,j}}\rangle^{2}}} \tag{1.19}$$

Three regions of operation exist for the signal-to-noise ratio. The first is the noise-floor limited region where the $\sigma^{f}{}_{i,j}$ term in the denominator of Equation 1.19 dominates due to large dark current and/or large ROIC noise floor. The SNR for the noise-floor limited region of operation varies linearly with the total number of electrons collected, $\langle\overline{N^{T}_{i,j}}\rangle$. The second region of operation is shot-noise limited where the $\langle\overline{N^{T}_{i,j}}\rangle$ term in the denominator of Equation 1.19 is greater than the other two terms. The SNR for shot-noise limited operation varies as $\sqrt{\langle\overline{N^{T}_{i,j}}\rangle}$. This is typically the region in which discrete single detector elements operate and is the ultimate physical limit for operation of FPAs since shot noise in the arrival of photons is unavoidable. The third region of operation is limited by spacial noise, i.e. uniformity. The SNR for uniformity limited operation is independent of $\langle\overline{N^{T}_{i,j}}\rangle$

and therefore also independent of the frame time t_f. The fact that the SNR is independent of the signal is somewhat surprising and marks a crucial difference between the results for an individual discrete detector in which U=0 by definition and an FPA in which the uniformity and not the performance of the individual devices is often the performance limiting factor.

Equation 1.3 for the photon flux is combined with Equations 1.14 and 1.15 to give the relation between photon flux and $\langle \overline{N^T_{i,j}} \rangle$:

$$\langle \overline{N^T_{i,j}} \rangle = \langle A_{i,j} \rangle \langle \eta_{i,j} \rangle \langle g_{i,j} \rangle t_f \Omega_{op} \Phi_P \qquad electrons \tag{1.20}$$

Since this is an optical system, the *f/#* of the system, the optimum area of the detector, A_D, and the cutoff wavelength, λ_C, of the detector are all related by diffractive optics. Taking the first zero in the Airy function as defining the size of the optical blur spot and setting the area of the detector to match the spot size gives:

$$A_D = (2.44 \times f/\# \times \lambda_C)^2 \tag{1.21}$$

The solid angle subtended by the detector is determined by the *f/#* of the system:

$$\Omega_{op} = \frac{\pi}{4(f/\#)^2 + 1} \approx \frac{\pi}{4(f/\#)^2} \qquad Steradians \tag{1.22}$$

Combining Equations 1.21 and 1.22 then gives the optimum detector area, solid angle product:[128]

$$A_D \Omega_{op} = 4.68\left(\lambda_C^2\right) \qquad cm^2 Steradians \tag{1.23}$$

Finally, a convenient figure of merit for FPA performance can be defined called the minimum resolvable temperature, MRT:

$$MRT = \frac{\sqrt{\langle \sigma^f_{i,j} \rangle + g\langle \overline{N^T_{i,j}} \rangle + U^2 \langle \overline{N^T_{i,j}} \rangle^2}}{\frac{d}{dT}\langle \overline{N^T_{i,j}} \rangle} \qquad Kelvin \tag{1.24}$$

which, since $\langle \overline{N^T_{i,j}} \rangle$ is exponentially dependent on temperature becomes:

$$MRT = \frac{\lambda k T^2}{hc} \frac{1}{SNR} \qquad Kelvin \tag{1.25}$$

Note that for the case where shot noise is dominant, i.e. the discrete detector limit, $MRT = NEDT$ and so can be thought of as an extension of NETD to include the non-uniformity, dark current and ROIC noise inherent in an FPA. In the same shot-noise limit, MRT can be related to D^* using Equation 1.7.

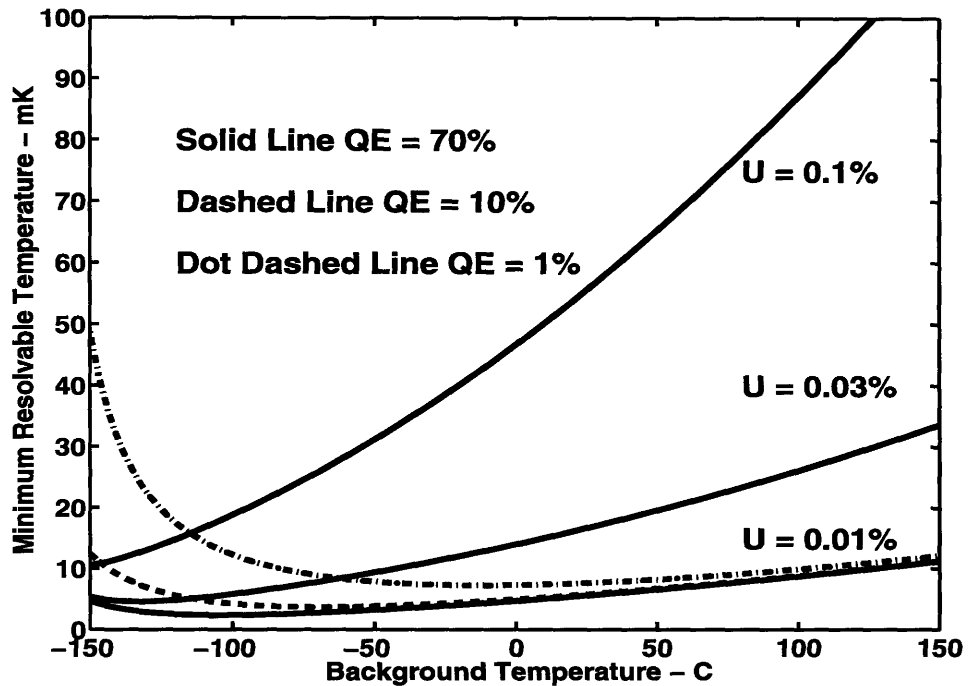


Figure 1.7: Uniformity limit for FPAs in the LWIR, 8μm to 10μm band. Pixel area 50μm x 50μm, frame rate 30Hz, optics f/# 1.5 AR coated, noise floor 200e$^-$, $g = 0.5$.

To emphasize the importance of including nonuniformity and the noise floor in calculations of FPA performance Figure 1.7 shows a comparison of five different detectors in the LWIR band using MRT as the figure of merit. The three FPAs represented with solid lines have the same high quantum efficiency of 70% with FPA uniformities of 0.1%, 0.03% and 0.01% after correction. The FPA represented with the dashed line has a quantum efficiency of only 10% but a uniformity equal to the best of the other three FPAs, 0.01%. Note that above a background temperature of -110C the FPA with 10% quantum efficiency but a uniformity of 0.01% has a lower MRT than the FPA with 70% quantum

efficiency but a uniformity a factor of 3 worse at 0.03%. The dash dot line represents an FPA with only 1% quantum efficiency but again 0.01% uniformity. Above a background temperature of -60C this FPA has a lower MRT than the FPA with 70% quantum efficiency but a uniformity a factor of 3 worse at 0.03%. All of these FPAs are in fact uniformity limited above a temperature of -50C.

Figure 1.8 shows the same set of FPAs this time operating in the MWIR band. The three FPAs represented with solid lines have the same high quantum efficiency of 70% with FPA uniformities of 0.1%, 0.03% and 0.01% after correction. The FPA represented with the dashed line has a quantum efficiency of only 10% but a uniformity equal to the best of the other three FPA, 0.01%. Note that above a background temperature of -10C the FPA with 10% quantum efficiency but a uniformity of 0.01% has a better MRT than the FPA with 70% quantum efficiency but a uniformity a factor of 3 worse at 0.03%. All of these FPAs are in fact uniformity limited above a temperature of 50C, thus the effect of changing from the MWIR to the LWIR band is to shift the curves to higher temperature.

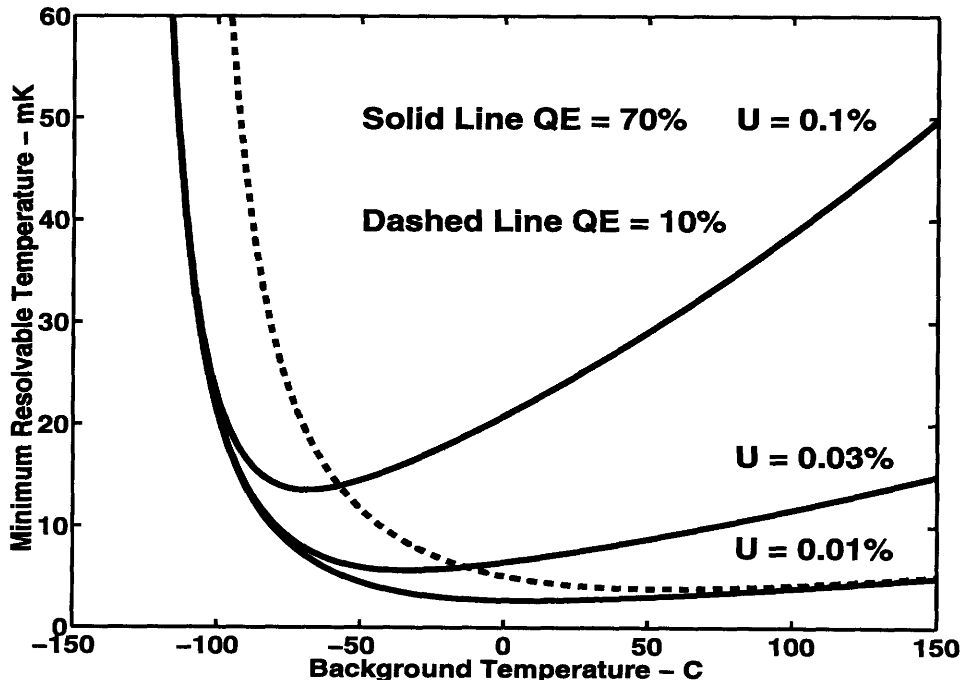


Figure 1.8: Uniformity limit for FPAs in the MWIR, 3μm to 5μm band. Pixel area 50μm x 50μm, frame rate 30Hz, optics f/# 1.5 AR coated, noise floor 200e$^-$, $g = 0.5$.

A more quantitative understanding for the region in which spatial noise is the limiting factor in FPA performance can be achieved by setting the term containing the uniformity, U, in the numerator of Equation 1.24 equal to the other two terms in the numerator of Equation 1.24. Use of Equation 1.20 then gives mean quantum efficiency as a function of uniformity and photon flux:

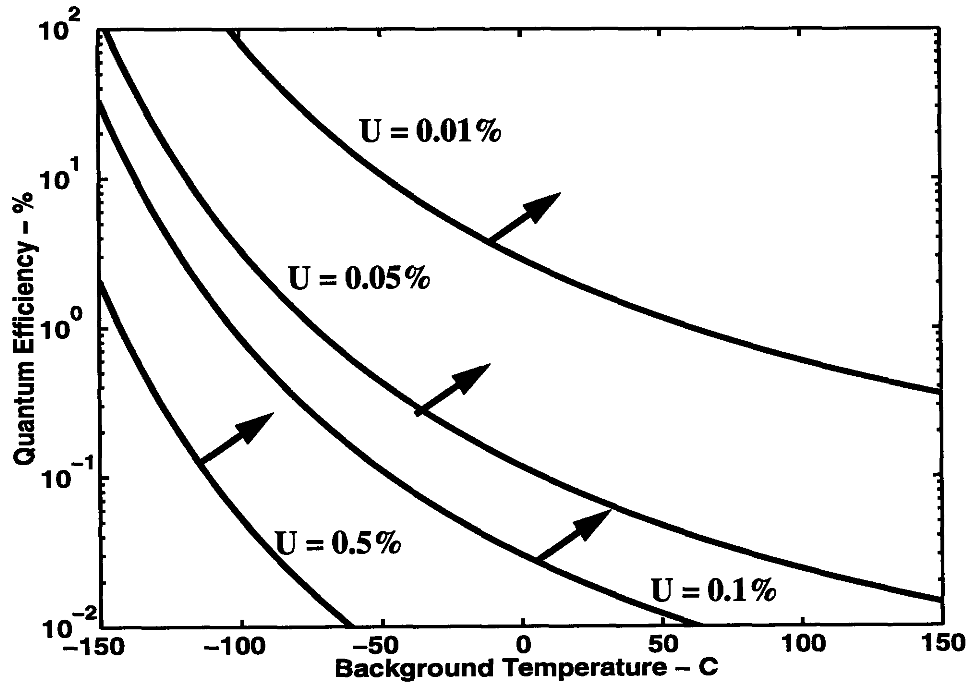


Figure 1.9: Spatial noise limited regime for uniformities of 0.5%, 0.1%, 0.05% and 0.01% in LWIR 8μm to 10μm band. Pixel area 50μm x 50μm, frame rate 30Hz, optics f/# 1.5 AR coated, noise floor 200e$^-$, $g = 0.5$.

$$\langle \eta_{i,j} \rangle = 100 \frac{\left(1 + \sqrt{1 + 4U^2 \langle \sigma^f_{i,j} \rangle^2}\right) \Delta f}{\langle A_{i,j} \rangle U^2 \Omega_{op} \langle g_{ij} \rangle \tau_{op} \Phi_P} \qquad Percent \qquad (1.26)$$

Equation 1.3 can be used to convert the photon flux term to mean background temperature. The result is plotted in Figure 1.9 for the LWIR, 8μm to 10μm, band and Figure 1.10 for the MWIR, 3μm to 5μm, band.

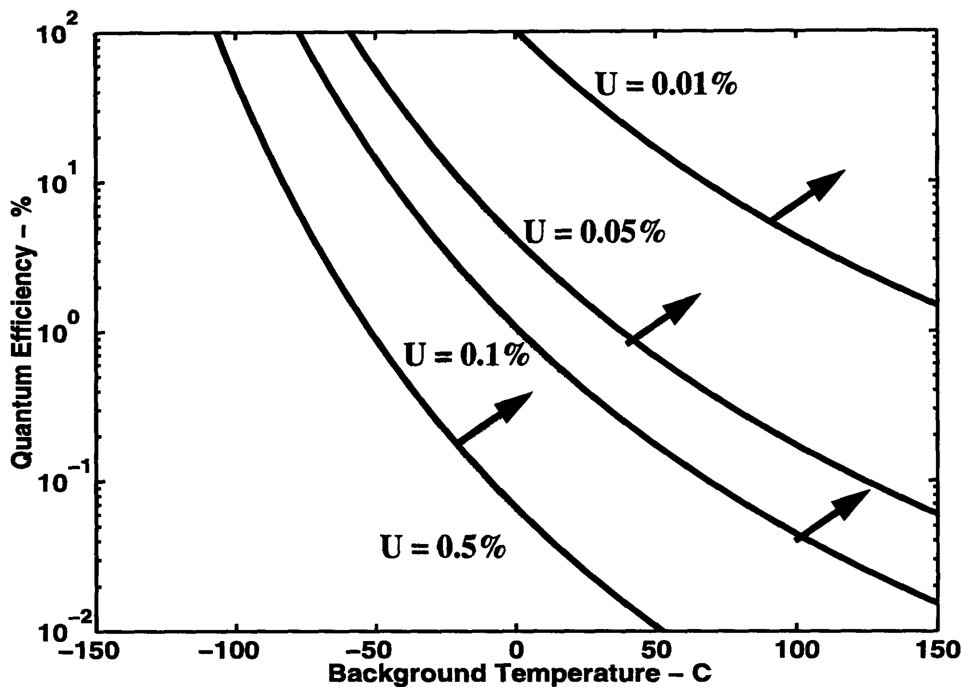


Figure 1.10: Spatial noise limited regime for uniformities of 0.5%, 0.1%, 0.05% and 0.01% in MWIR 3μm to 5μm band. Pixel area 50μm x 50μm, frame rate 30Hz, optics f/# 1.5 AR coated, noise floor $200e^-$, $g = 0.5$.

The MWIR band has greater contrast and fewer photons so that for the same uniformity the FPA does not become uniformity limited until a higher background temperature is reached. To illustrate the effect of the increased photon flux contrast available in the MWIR band verses the LWIR band on FPA MRT, Figure 1.11 shows a comparison between an FPA with a quantum efficiency of 70% operating in the LWIR, solid line, and an FPA with a quantum efficiency of only 1% operating in the MWIR, dashed line. Both FPAs have the same uniformity and are operating in the uniformity limited regime, but due to the improved contrast in the MWIR band the FPA operating in the MWIR has a better MRT for background temperatures above about 30C

Figure 1.12 shows a comparison of MRT and D^* as figures of merit for an FPA. Notice that for the particular FPA parameters chosen little or no improvement in MRT is seen for a D^* greater than about 10^{10} Jones. This shows that the FPAs are limited by uniformity and not shot noise and therefore D^* is not an appropriate figure of merit for the FPA.

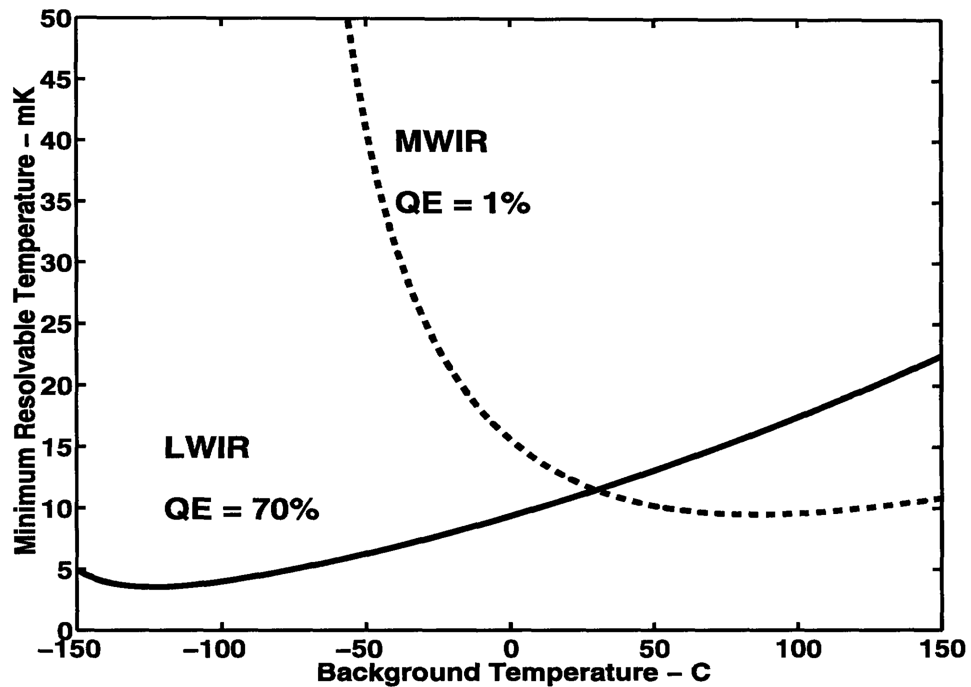


Figure 1.11: Comparison of MWIR, 3μm to 5μm, and LWIR 8μm to 10μm bands. Uniformity 0.02%, pixel area 50μm x 50μm, frame rate 30Hz, optics f/# 1.5 AR coated, noise floor 200e⁻, $g = 0.5$.

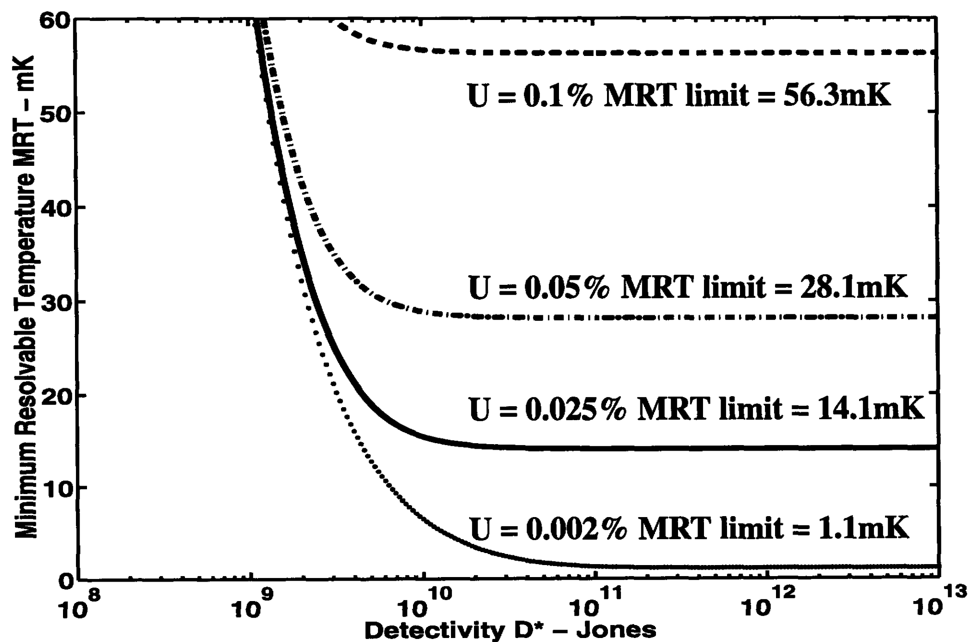


Figure 1.12: Comparison of D* and MRT as figures of merit for FPAs. LWIR band 8μm to 10μm, f/# 1.5 optics with AR coating, pixel area 50μm x 50μm, noise floor 200e⁻, background temperature 300K, $g = 0.5$.

In conclusion, this section has developed a model of a focal plane array camera including individual detector performance, the effects of uniformity on the performance of the entire array and the effect of a noise floor due to background shot noise and/or electronic limits. Using this model it has been shown that in many practical situations the individual detector performance *is not* be the limiting factor in overall FPA performance and therefore individual detector figures of merit like D^* must be replaced by figures of merit like MRT designed for comparison of FPA based cameras.

1.5 Effect of Calibration on FPA uniformity

It is sometimes argued[15,16] that since pixel by pixel uniformity correction is standard in FPAs, uniformity is not the a limiting factor in FPA performance. To investigate this argument Figure 1.13 shows spacial nonuniformity from an InSb array from Cincinnati Electronics Corporation[17] fabricated in 1994. The array was corrected at scene temperatures of 15C and 50C using a two point correction algorithm and a 12bit ADC. If each pixel's output was linear over the scene temperatures used, this two point correction would produce perfectly uniform frames for all scene temperatures. However due to the finite resolution of the digital data and correction coefficients, 12bits, nonlinearity in the pixel responsivities and spacial variations in the pixel responsivities as a function of the spectral content of the scene, which changes with changing scene temperature as shown in Figure 1.2, the spacial uniformity after correction is non-zero.

The minimum MRT from Figure 1.13 is 12mK in spite of individual detector performance in excess of 10^{12} Jones indicating that the overall array performance is limited by spatial noise, ie. non-uniformity. Notice that from Figure 1.12 the MRT of 12mK corresponds very well to the limit of MRT = 14.1mK for an FPA with $U = 0.025$ and any D^* greater than about 10^{10} Jones. Even after correction, the FPA is still not shot noise limited but is limited by residual non-uniformity! It is interesting to calculate what fraction of the measured minimum MRT is due to the use of a 12 bit analog to digital converter, ADC, in the read out integrated circuit, ROIC. For temporal averaged signals, the minimum noise in a 12 bit ADC is $2^{12}/\sqrt{12}$ or 0.006%. Since in the non-uniformity limited regime the

minimum MRT is proportional to the non-uniformity from Figure 1.12 a non-uniformity of 0.006% corresponds to an MRT of 3.3mK or about 25% of the measured MRT of 12mK for this FPA. Also note that this array was calibrated at two points 35C apart and MRT at the extremes of scene temperature are about a factor of two greater than the minimum MRT. To correct to one LSB over the entire useful scene temperature of 150C requires calibration about every 10C which for the 1 megapixel format being designed today means making available 24MB of SRAM memory for storing calibrations. To do better than this quickly gets prohibitively expensive and complex ensuring that for the foreseeable future FPAs, will be limited by residual post-correction uniformity and not by pixel performance.

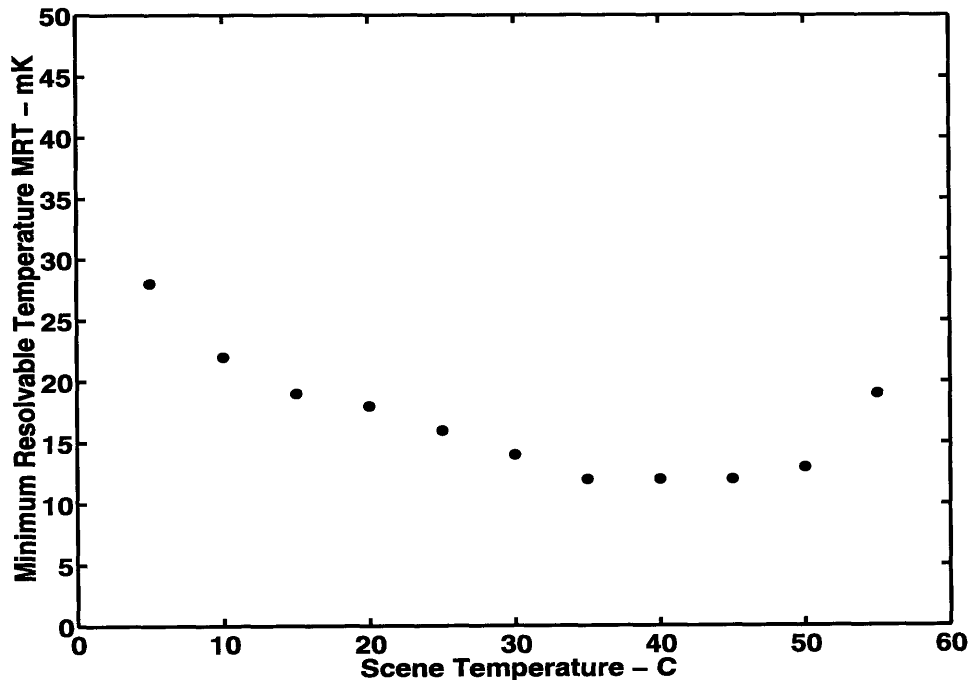


Figure 1.13: MRT in an InSb photovoltaic focal plane array after two point uniformity correction at 15C and 50C. Array format 160 x 120, *f/#* 2.3.[17]

1.6 Review of Selected IR Detector Technologies

The subject of this thesis is primarily quantum well intersubband photodetectors, QWIPs as used in focal plane arrays, FPAs. However, no technology exists in a vacuum unchallenged and so this subsection will be devoted to an introduction to a selected subset of

detector technologies all of which are being applied to FPAs and are to a large extent competing for there own niche in the same relatively small FPA market. Here the author will attempt to put aside his bias for QWIPs and give an objective, but short, evaluation of the strengths and weaknesses of the available technologies as they pertain to the current list of likely applications. Of course as camera cost decreases and availability increases the set of applications is expected to expand and so some attention will also be paid to the technologies which are most likely to drive down camera cost.

Current applications for FPAs can be divided loosely into four categories; space based astronomy, seeker/tracker for missiles and ordinance, cooled surveillance and uncooled surveillance. To a large extent, the category into which a particular technology falls depends on the operating temperature and therefore the cooling mechanism used both because the operating temperature sets an upper limit on the FPA performance and because the cooling system is often a large part of the overall system cost. Options for cooling primarily consist of: closed cycle refrigerators which are big and expensive but are also reliable, very stable and can cool for long periods of time; Joule-Thompson coolers which are small, light-weight with very fast cool down times, but can cool for only a short time, typically minutes; liquid refrigerants like liquid Nitrogen, liquid Argon and liquid Helium sometimes find use where there is easy access to such refrigerants and operation times are on the order of hours; finally, thermo-electric coolers sometimes find use in systems operated near room temperature where moderate cost and availability of electrical power can be taken advantage of.

Space based astronomy systems can take advantage of the very cold temperatures in space to minimize cooling requirements and are therefore often operated near liquid helium temperatures, about 4K. Furthermore, these detectors are aimed away from any nearby bodies and so have as their background only the 2.9K remnant of the big bang. This essentially removes both dark current and background photon current as sources of noise and these FPAs are limited by the ROIC noise floor and the shot noise in the signal. As such, high quantum efficiency is extremely important and these FPAs are almost exclu-

sively made form Mercury Cadmium Telluride, MCT, InSb and lead salts because of their high quantum efficiency.

Seeker/tracker missiles and ordinance use FPAs to acquire an image from the head of a missile after launch and then using digital signal processing, DSP, algorithms compare the acquired image to a set of stored target shapes. Things like the round barrel of a tank gun are favorite targets and once acquired these missiles are extremely accurate. Because of their inherent short life and the need for fast cool down, the cooling system of choice for seeker/tracker missiles is the Joule-Thompson cooler which is lightweight, reliable and relatively cheap. Several technologies including MCT, PtSi and QWIPs are all competing for this market which is currently dominated by MCT.

Cooled surveillance applications typically involve a camera in a building, or on a platform like a tank, ship or airplane. In these applications the camera may need to operate for an extended period of time and so power for a closed cycle refrigerator or liquid refrigerants are often supplied. Here, FPA performance is very important since targets are often at the extreme detection distance and there is a large premium on early detection. Again, QWIPs, MCT and PtSi are all competing for this market with QWIPs and MCT leading somewhat in performance criteria and PtSi leading in the cost area.

Uncooled surveillance applications are just now being developed but seem likely to be the exclusive domain of thermal detectors like the Barium Strontium Titanate, BST, microbolometers being developed at Texas Instruments[18,19]. These FPAs currently have MRTs an order of magnitude larger than cooled detectors, but for applications like rifle sights and hand-held night vision cameras where power and weight limits eliminate the possibility of using any type of cryogenic cooler, these FPAs are showing a lot of promise. In addition, because the technology used in making the microbolometers is largely compatible with standard Silicon CMOS processes, it is expected that these detectors will drive down FPA cost and open up many new applications.

Having discussed the broad outline of applications and requirements, a short description of the operating physics and technology of each for a representative group of FPA devices will be given. To represent the intrinsic photon detectors, MCT has been chosen

because it is the most established of the available technologies. To represent the Schottky barrier type extrinsic detectors, PtSi has been chosen again because it is the most established of the Schottky barrier technologies. To represent the thermal detectors, the BST microbolometer has been chosen because it is currently the most advanced of this very old category of detector. To represent QWIPs, the most common n-type GaAs/AlGaAs QWIP superlattice structure is chosen.

1.6.1 Mercury Cadmium Telluride, HgCdTe, FPAs

Mercury Cadmium Telluride, MCT, is currently the most important material for high performance infrared detectors including FPAs. MCT has earned this title due to its versatility and demonstrated performance over thirty-six years since the first reported photoconductive detector based on MCT was demonstrated by Lawson et. al. in 1959.[20].

The attractive features of this ternary II-VI semiconductor include a direct energy gap that ranges from -0.3eV for the semimetal HgTe goes through zero at x = 0.15 for $Hg_{1-x}Cd_xTe$ and ranges up to 1.648eV for CdTe. The best fit to the energy gap is given by:[21]

$$E_g = -0.302 + 1.93x - 0.81x^2 + 0.832x^3 + 5.32\times10^4(1-2x)\left(\frac{-1822+T^3}{255.2+T^2}\right) \quad eV \tag{1.27}$$

Over this entire range of bandgaps which spans the infrared spectrum down to a wavelength of 0.75μm the crystal lattice constant only varies from 6.4614Å for HgTe to 6.4808Å for CdTe, a change of 0.3% as given by the expression:[22]

$$a = 6.4614 + 0.0084x + 0.01168x^2 - 0.0057x^3 \quad Angstroms \tag{1.28}$$

This very good lattice matching eliminates the problems with growth of mismatched semiconductors and at the same time provides a natural transparent substrate, CdTe, on which MCT detectors covering the entire infrared spectrum can be grown. The intrinsic carrier concentration as a function of composition and temperature can be approximated by:[JRL]

$$n_i = \left(A + Bx + CT + DxT + Fx^2 + GT^2\right) \bullet 10^{14}\left(\frac{hc}{\lambda}\right)^{3/4} T^{3/2} \exp\left(\frac{-hc}{2\lambda kT}\right) \quad (1.29)$$

The electron and light hole effective masses, m^*_e, are very close in the narrow gap Mercury containing compounds and have been modeled with the Kane[23] band model as:[24]

$$\frac{m}{m_e^*} = 1 - 1.6 + \frac{E_p}{3}\left(\frac{2\lambda}{hc} + \frac{1}{\frac{hc}{\lambda} + \Delta}\right) \quad (1.30)$$

where E_p = 19eV, Δ = 1eV and λ is the cutoff wavelength. The high frequency dielectric constant is given by:[25]

$$\varepsilon_\infty = 15.2 - 13.7x + 6.4x^2 \quad (1.31)$$

where x is the Cadmium fraction in the ternary $Hg_{1-x}Cd_xTe$.

Under the assumptions that the noise is dominated by shot noise in the generation and recombination processes, the recombination rate is Auger limited and the detector is in thermal equilibrium implying that the generation and recombinations rates are equal, the spectral current responsivity of the MCT detector is:

$$R_\lambda = \frac{\lambda\eta}{hc} gq \qquad \frac{Amps}{Watt} \quad (1.32)$$

where η is the quantum efficiency, q is the electron charge and g is the gain in electrons collected per electron generated. Assuming the noise is dominated by shot noise in the generation and recombination processes, the noise current is given by:

$$i_n = \sqrt{2\,(G + R)\,A_D t \Delta f q^2 g^2} \qquad Amps \quad (1.33)$$

where G is the electron hole pair generation rate in pairs/cm^{3}s, R is the electron hole pair recombination rate in pairs/cm^{3}s, t is the thickness of the sample and Δf is the noise bandwidth. The detectivity D_λ^* is then, using Equations 1.7, 1.9 and 1.10 given by:

$$D_{\lambda}^{*} = \frac{\lambda}{hc}\frac{\eta}{\sqrt{2t(G+R)}} \qquad Jones \tag{1.34}$$

Assuming that the detector is in equilibrium and therefore the generation and recombination rates are equal and that the recombination rate is Auger limited, the generation rate is given by:

$$G_{Auger} = \frac{1}{2\tau_{A}^{i}}\left(n + \frac{p}{\gamma}\right) \qquad \frac{pairs}{cm^{3} \cdot second} \tag{1.35}$$

where n is the electron concentration, p is the hole concentration, γ is an empirical factor relating hole and electron Auger processes and τ^{i}_{A} is the intrinsic Auger lifetime given approximately by:[26,27]

$$\tau_{A}^{i} = \frac{3.8\times10^{-18}\varepsilon_{\infty}^{2}(1+\mu)^{1/2}(1+2\mu)\exp\left(\frac{hc}{\lambda kT}\right)}{\left(\frac{m_{e}^{*}}{m}\right)|F_{1}F_{2}|^{2}\left(\frac{\lambda kT}{hc}\right)^{3/2}} \qquad seconds \tag{1.36}$$

where μ is the ratio of electron to heavy hole effective mass with the electron effective mass given by Equation 1.30, the heavy hole effective mass will be assumed constant at 0.55, ε_{∞} is given by Equation 1.31 and a value for the overlap integral $|F_1F_2|^2 = 0.25$ will be assumed to fit experimental data.[28] For doped n-type materials this is reduced by the ratio $2\left(\frac{n_i}{n_0}\right)^2$ where n_i is the intrinsic carrier concentration given by Equation 1.29 and n_0 is the doping concentration. This gives the Auger lifetime in doped n-type MCT:

$$\tau_{A} = 2\tau_{A}^{i}\left(\frac{n_{i}}{n_{0}}\right)^{2} \qquad seconds \tag{1.37}$$

Substituting Equation 1.37 into Equation 1.34 gives the Auger limited detectivity:

$$D_{\lambda}^{*} = \frac{\lambda}{2hc}\frac{\eta}{\sqrt{t}}\sqrt{\frac{\tau_{A}}{n+\frac{p}{\gamma}}} \qquad Jones \tag{1.38}$$

Equation 1.38 is plotted in Figure 1.14 as a function of cutoff wavelength for several detector temperatures assuming an n-type doping of $5.0\text{x}10^{13}/\text{cm}^3$ which is close to the currently achievable minimum doping.[29] Note that p-type doping should lead to an improvement in D_λ^* equal to $\gamma^{1/2}$ where γ is a term of fairly high experimental uncertainty, but is usually taken as $3 \leq \gamma \leq 6$.[30] In any case technical difficulties in doping p-type currently prevent taking advantage of this factor.[29]

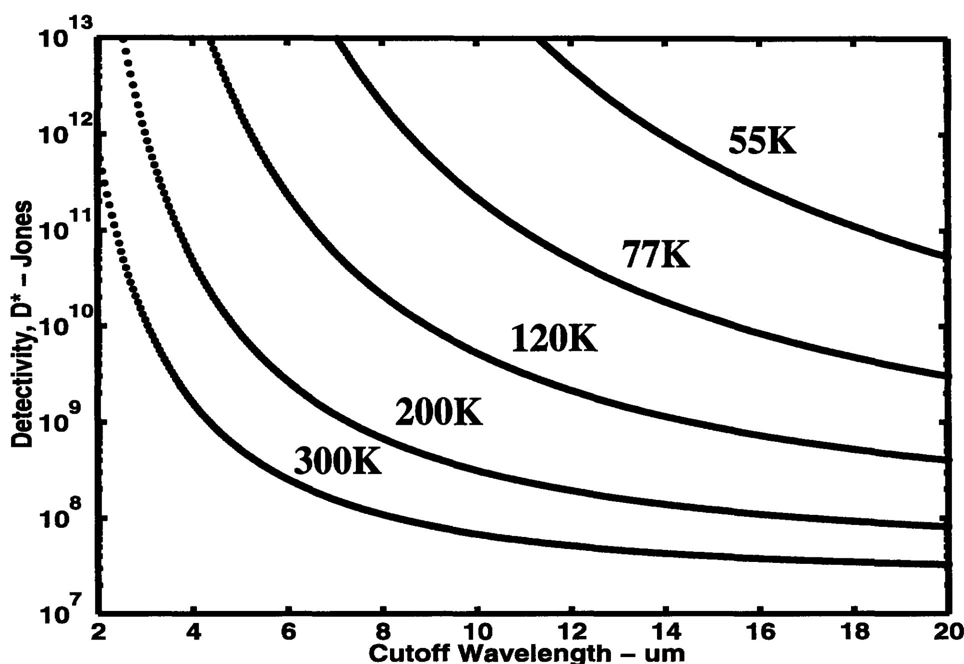


Figure 1.14: D_λ^* for MCT detector with area 50μm x 50μm, $\eta = 1$, $t = 5\mu\text{m}$ $n = 5\text{x}10^{13}\ \text{cm}^{-3} >> p$ and T as shown.

The individual detector performance indeed approaches the limits shown in Figure 1.14 but problems with uniformity in large area material limit FPA performance to orders of magnitude less than that shown. The best FPA performance for an MCT FPA with TV format, 640 x 480, is MRT = 5mK when operated at 50K in the LWIR band and the FPA is limited by uniformity not D_λ^*. Furthermore, due to difficulties growing bulk CdTe substrates larger than 1", current MCT arrays of this format have very poor reliability and extreme expense.

1.6.2 Platinum Silicide Schottky Barrier FPAs

Schottky barrier detectors are extrinsic photon detectors in that only a single carrier contributes to the photo response. When a metal makes intimate contact with an n-type semiconductor, electrons in the semiconductor move into the metal leaving positive immobile dopant charges in the depletion layer of the semiconductor.

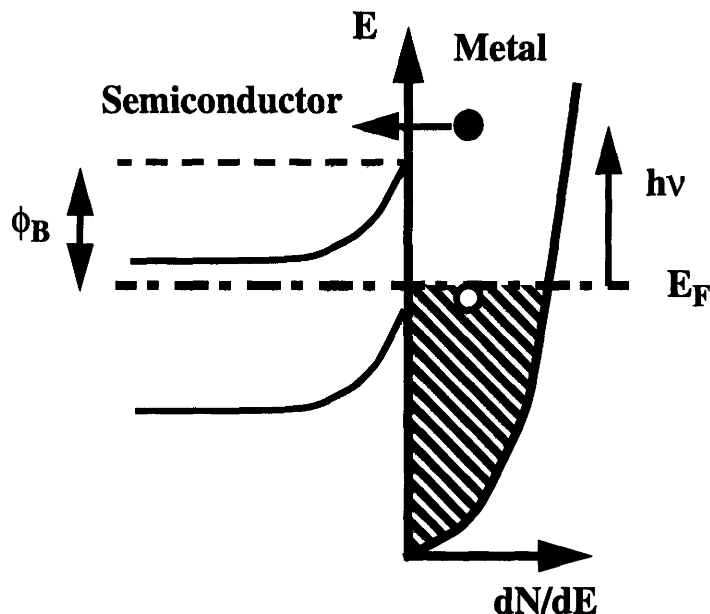


Figure 1.15: Photoabsorption in a Schottky barrier detector

This movement of electrons continues until the Fermi-level in the metal, which defines the chemical potential of the electrons, lines up with that of the semiconductor creating a potential barrier called a Schottky barrier. When a photon whose energy is greater than the Schottky barrier height ϕ_B is absorbed in the metal an energetic electron is created as shown schematically in Figure 1.15. The excited electrons then do a random walk in the metal film until they either reach the semiconductor interface and are transported into the semiconductor or are recombined with a free hole in the metal.

Assuming that the probability of excitation is independent of photon energy and that the sample is near zero Kelvin so that there is an abrupt transition from filled to empty states at the Fermi level, the total number of possible excited states, N_T is:

$$N_T = \int_{E_F}^{E_F + h\nu} \frac{dN}{dE} dE \qquad \frac{states}{cm^3} \tag{1.39}$$

where *dN/dE* is the density of states in the metal. For an excited electron to cross the Schottky barrier into the semiconductor, the state to which it has been excited must have a component of momentum, hv/c, in the direction normal to the interface which corresponds to an energy greater than the Schottky barrier height, ϕ_B. The number of states which meet this criterion, N_E, is given by:

$$N_E = \int_{E_F + \phi_B}^{E_F + h\nu} \frac{dN}{dE} P(E)\, dE \qquad \frac{states}{cm^3} \tag{1.40}$$

where *P(E)* is the photoemission probability for the electron with energy *E* given by:[31]

$$P(E) = \frac{1}{2}\left[1 - \sqrt{\frac{(E_F + \phi_B)}{E}}\right] \tag{1.41}$$

Considering *dN/dE* to be nearly independent of energy over the energy range of interest because the Fermi energy is much greater than the photon energy, the integrals in Equations 1.36 and 1.37 can be replaced by multiplications giving:

$$N_T = \frac{dN}{dE} h\nu \qquad \frac{states}{cm^3} \tag{1.42}$$

$$N_E = \frac{dN}{dE} \frac{(h\nu - \phi_B)^2}{8E_F} \qquad \frac{states}{cm^3} \tag{1.43}$$

Assuming no collisions between excited electrons or energy losses before the excited electrons reach the interface, the internal quantum efficiency, η, is simply the ratio of Equation 1.43 to 1.42, but since these effects are always present and difficult to model, the external quantum efficiency is usually used instead expressed as:

$$\eta = C_1 \frac{(h\nu - \phi_B)^2}{h\nu} \tag{1.44}$$

where C_1 is the experimentally determined quantum efficiency coefficient. For PtSi on p-type Silicon and a PtSi thickness of 20nm $C_1 = 0.07\text{eV}^{-1}$.

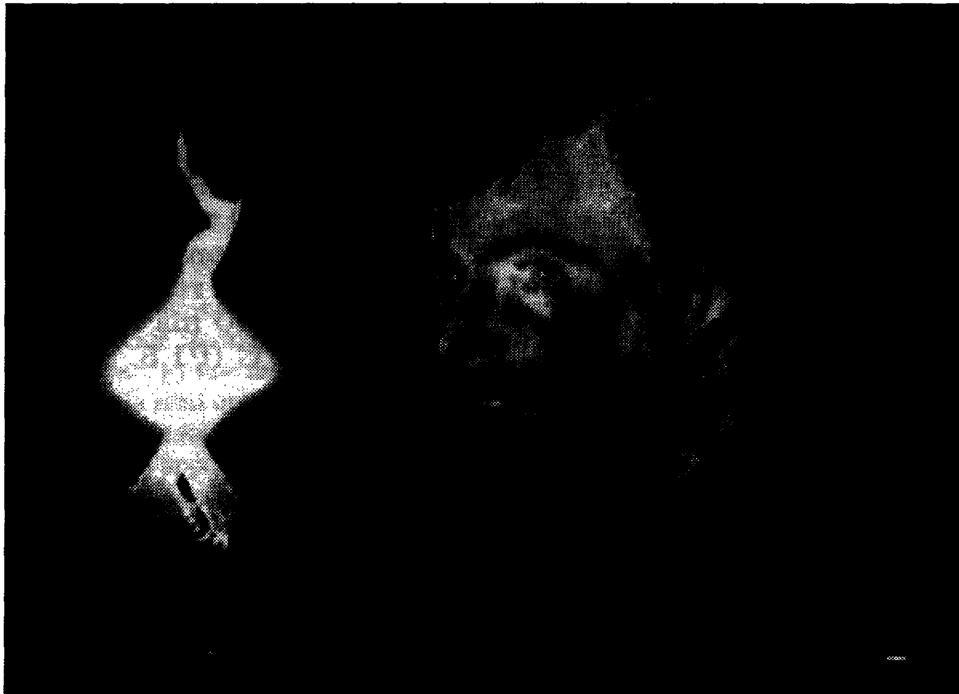


Figure 1.16: PtSi 640 x 480 FPA with MRT = 70mK.[32]

The cutoff wavelength is simply determined by the Schottky barrier height.

$$\lambda_C = \frac{1.24}{\phi_B} \quad \mu m \tag{1.45}$$

For PtSi the cutoff wavelength is about 5μm with an external quantum efficiency of about 1%. This seems like a very low number, but because these detectors are fabricated on large area Silicon using processes developed for the digital Silicon industry, the pixel uniformity can be excellent, < 0.02%. As discussed in section 1.4 FPAs are often operated in the uniformity rather than quantum efficiency regime making PtSi extremely competitive, especially for low cost cooled applications. These detectors do have the disadvantage that they need to be cooled to fairly low temperatures, typically less than 50K to

reduce the dark current to acceptable levels. As an example of the images possible with this technology, Figure 1.16 shows an image taken with a 640 x 480 TV format PtSi array using f/1.5 optics and a frame rate of 30Hz. The MRT is < 70mK.

1.6.3 Micro-Bolometer FPAs

Bolometers are thermal detectors and as such depend on the absorption of electromagnetic energy to raise the temperature of the detector which in turn causes a change in a measurable property such as resistance for bolometers, or gas pressure for the Golay cell, etc. Since the absorption is not measured directly, but rather only the temperature of the detector is measured, thermal detectors are equally sensitive, on an energy scale, to all frequencies of light absorbed. For the absorption part of the process, a thermal detector can be represented by a thermal mass H, coupled via a conductance G to a heat sink at constant temperature T.[33] With no incident radiation, the mean temperature of the detector will also be T. Fluctuations about the mean are determined by the coupling mechanism between the detector. When radiation is incident on the detector, the rise in temperature is determined by solving the equation:

$$\eta I(t) = H\left(\frac{d}{dt}\theta(t)\right) + G\theta(t) \qquad \textit{Watts} \tag{1.46}$$

where I is the power incident on the detector and η is the fraction of the incident power actually absorbed, t is the time and θ is the difference between the detector temperature and the heat sink temperature. Typically the signal incident on the detector is chopped at a frequency ω giving for the power incident on the detector the form:

$$I(t) = I_0 + I_\omega \exp(j\omega t) \qquad \textit{Watts} \tag{1.47}$$

where ω is the chopping frequency.

Plugging in Equation 1.47 into Equation 1.46 and solving for the amplitude, θ_ω, and phase, φ, of the excess temperature $\theta(t)$ gives:

$$\theta_\omega = \eta \frac{I_\omega}{\sqrt{G^2 + \omega^2 H^2}} \qquad \textit{Kelvin} \tag{1.48}$$

$$\varphi = \tan^{-1}\left(\omega\frac{H}{G}\right) \tag{1.49}$$

Equation 1.48 illustrates the important design criteria for thermal detectors with regards to maximizing the temperature change for a given input power. To make θ_ω as large as possible G, which measures the coupling of the detector to the heat sink needs to be made as small as possible with ω being chosen small enough that ωH is on the order of or less than G. In applications where the chopping frequency is fixed by a frame rate or other parameter, H, the thermal mass of the detector must also be made small such that G and ωH are of similar magnitude.

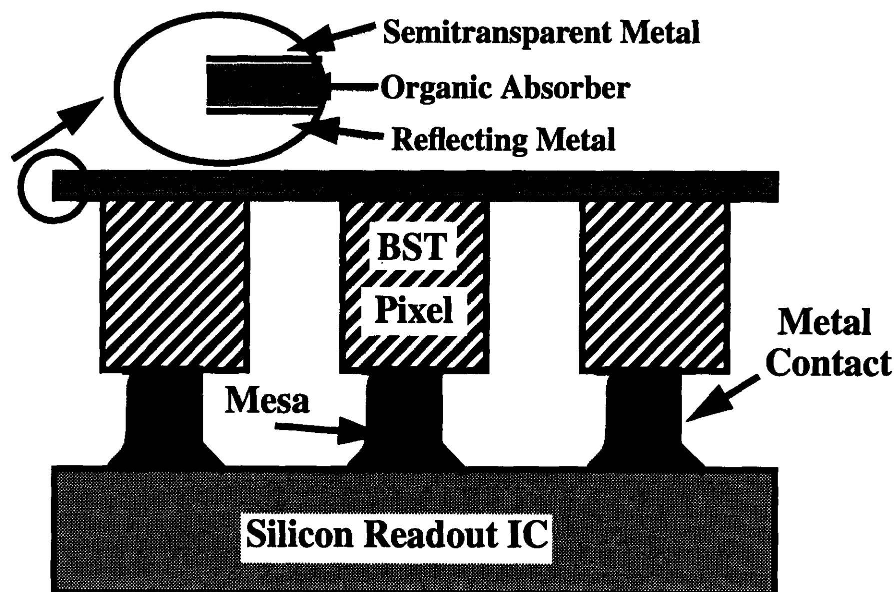


Figure 1.17: BST Pixel Structure.

The micro-bolometer accomplishes large reductions in both H and G by micromachining very small detector elements connected to the heat sink by long thin electrical wires with as small a conductance as possible consistent with the need to carry a small current. Figure 1.17 shows the structure of the Barium Strontium Titanate, BST, fabricated at Texas Instruments[34]. The conversion from temperature change in the pixel to voltage

change in the output is made by passing a small accurately controlled current through the BST pixel and monitoring the change in resistance of the detector. If the radiation produces a change in the pixel temperature θ and then the output voltage signal is:

$$V_S = i\alpha R_{det}\theta \qquad Volts \tag{1.50}$$

where α is the temperature coefficient of resistance $(1/R_{det})\frac{dR_{det}}{dT}$, R_{det} is the detector resistance and i is the current through the detector. Using Equation 1.48 the open circuit output voltage is:

$$V_S = \frac{\eta I_\omega i\alpha R_{det}}{\sqrt{G^2 + \omega^2 H^2}} \qquad Volts \tag{1.51}$$

which gives a responsivity of:

$$R = \frac{\eta i\alpha R_{det}}{\sqrt{G^2 + \omega^2 H^2}} \qquad \frac{Volts}{Watt} \tag{1.52}$$

The BST detector has demonstrated an excellent uncooled MRT of 47mK by taking advantage of semiconductor fabrication techniques to make H and G extremely small and by operating the BST near its paraelectric-ferroelectric phase transition[34] which for the selected composition of BST is near room temperature and results in the maximum pyroelectric coefficient α. Because these detectors are compatible with Silicon manufacturing processes and do not require cooling they are the most likely candidates for hand-held night vision equipment and low cost moderate performance applications.

1.6.4 Quantum Well Intersubband Photodetectors

The possibility of using intersubband transitions in quantum wells consisting of alternating layers of AlGaAs and GaAs to detect infrared radiation was first suggested by Esaki and Sakaki[6] at IBM in 1977 and also Smith at AT&T.[35] This followed closely the technological development of molecular beam epitaxy (MBE) which by the late 1960's Arthur[36] and later Cho[37] had shown was capable of growing single crystal undoped ternary films of AlGaAs with extremely good thickness control. Technical difficulties controlling layer thickness to within the several atomic layers required to get strong

intersubband absorption, controlling interface roughness to the same monolayer tolerance and controlling the dopant levels in the quantum wells proved difficult to surmount and it was not until 1985 that West and Eglash[38] demonstrated the first experimental evidence for strong intersubband absorption in a multi-quantum well structure.

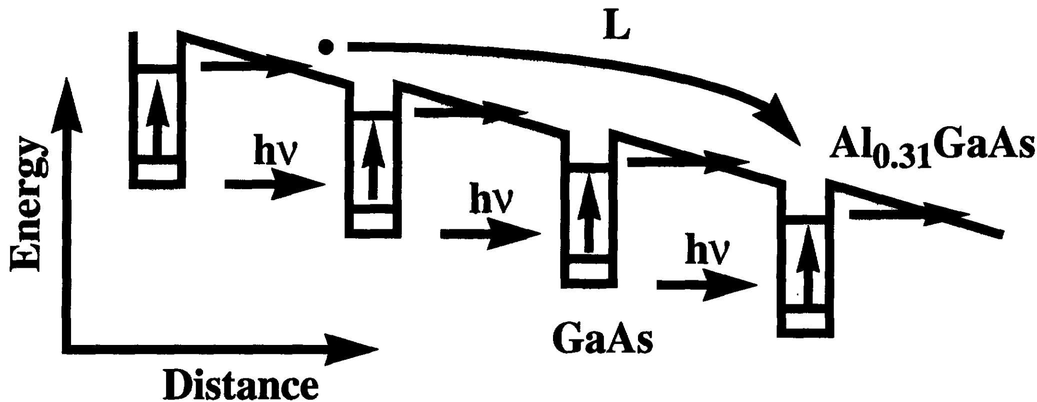


Figure 1.18: QWIP Schematic showing 40Å GaAs quantum wells under bias and separated by 300Å $Al_{0.31}Ga_{0.69}As$ barriers. Excited electron mean free path L is represented schematically for an electron generated in the left most well and captured by the right most well.

This led to a flurry of activity and in 1987, Levine et. al.[7] at AT&T Bell Laboratories demonstrated the first quantum well intersubband photodetector (QWIP). Levine's first QWIP was based on intersubband absorption between two bound states and achieved an impressive peak responsivity of $R_\lambda^P = 0.52A/W$ corresponding to a quantum efficiency of nearly 6% at 10.8μm which is ideally located in the center of the LWIR atmospheric transmission band; see Figure 1.5. This device was grown on a computer controlled MBE machine and consisted of a 50-period superlattice of 65Å GaAs wells, doped $n = 1.4\times10^{18}cm^{-3}$ in the center 50Å, and 95Å $Al_{0.25}Ga_{0.75}As$ barriers. A schematic of the quantum well structure is shown in Figure 1.18 The superlattice was sandwiched between two heavily doped GaAs contact layers and a 45 degree facet was polished on one edge for coupling in the infrared radiation as shown in Figure 1.19.

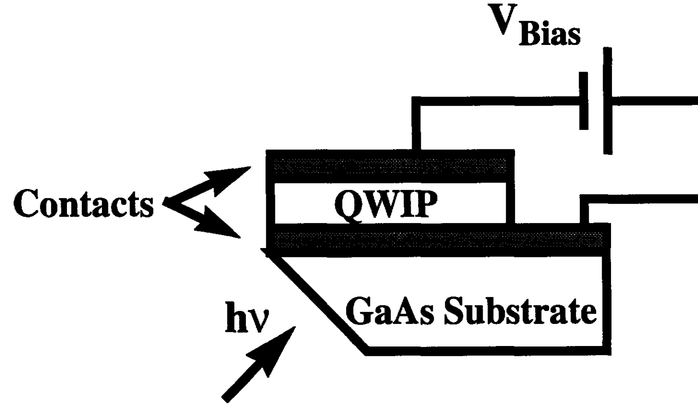


Figure 1.19: Measurement geometry for edge coupled QWIP devices

Following the logic of Kinch and Yariv[15] but using more recent experimental data it is possible to use thermodynamic and radiometric considerations to estimate the limits on performance of individual QWIP detector elements of the type investigated by Levine et.al. [39,40,41,42]. Unlike MCT which, as discussed in Section 1.6.2, is a minority carrier device, QWIPs are majority carrier devices and as such the noise is determined by the variance in the density of majority carriers within the device:

$$\sqrt{\langle \Delta n^2 \rangle} \equiv n = n_t + n_\Phi \tag{1.53}$$

where n_t is the density of thermally generated carriers and n_Φ is the density of photon generated carriers which has a contribution due to the signal (Φ_S) and a contribution due to the background flux (Φ_B) as given by Equation 1.3 and shown in Figure 1.3. The signal is equal to the contrast photon flux, $d\Phi_P/dT$, as given by Equation 1.4 and shown in Figure 1.4, multiplied by the temperature difference between the object of interest and the background. Assuming approximately 300K background and a small, < 1K, temperature difference between the object and the background, $\Phi_S \ll \Phi_B$, the total density of photon generated carriers can be set equal to the density of background generated carriers, $n_{\Phi B}$. The ultimate in single detector performance is achieved when $n_{\Phi B} \geq n_t$, since the total

noise is then dominated by the noise in the background photon flux and background limited performance (BLIP) is achieved.

The thermally generated carrier density is calculated based on the density of states in the quantum well structure shown in Figure 1.20 taken from the experimental work of Levine[7]. The structure is composed of a 50-period multi-quantum well superlattice having a well width L_W of 40Å, a barrier width L_B of 480Å of $Al_{0.25}GaAs$ and an electron doping in the well of $N_D = 1.2\times10^{18}cm^{-3}$. The top 0.5μm and bottom 1μm of the superlattice have n+ contact layers doped $N_{DC} = 1.2\times10^{18}cm^{-3}$.

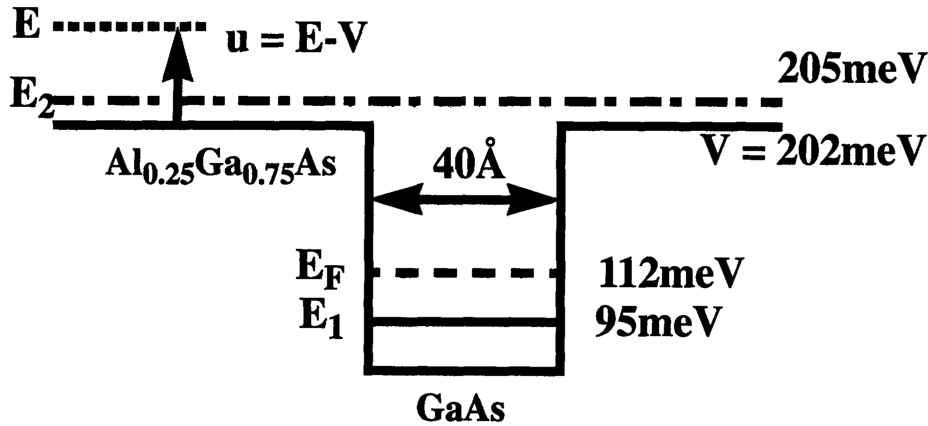


Figure 1.20: QWIP conduction band diagram. Quantum well is 40Å of GaAs separated by 480Å $Al_{0.25}Ga_{0.75}As$ barriers.

An envelope function approximation calculation for this structure yields the energy level diagram also shown in figure 1.20. The first subband lies approximately 95meV above the GaAs conduction band edge and the second subband lies approximately 5meV above the continuum edge of the $Al_{0.25}Ga_{0.75}As$ barrier. Assuming an infinite quantum well barrier, the Fermi level is given approximately by:

$$E_F = \frac{N_D h^2 L_W}{4\pi m^*} \quad eV \tag{1.54}$$

where m^* is the electron effective mass in GaAs and is equal to $0.067m_0$. For the quantum well in question this evaluates to 17meV. The density of states in the continuum

can be approximated by treating the continuum as a wide quantum well of width $L = L_W + L_B$ which for this structure is 520Å. The two dimensional density of states is then:

$$\rho_{2D}(E) = \frac{4\pi m^*}{h^2}\left\{1 + Int\left[L\left(\frac{8m^*(E-V)}{h^2}\right)^{1/2}\right]\right\} \qquad \frac{1}{cm^2 \cdot eV} \tag{1.55}$$

where *Int(x)* is defined as the largest integer less than x. Note that for:

$$L\left(\frac{8m^*(E-V)}{h^2}\right)^{1/2} \gg 1 \tag{1.56}$$

the volume density approaches the bulk value:

$$\rho_{3D}(E) = \frac{\rho_{2D}(E)}{L} = 2\sqrt{\pi}\left(\frac{4m^*}{h^2}\right)^{3/2}\sqrt{E-V} \qquad \frac{1}{cm^3 \cdot eV} \tag{1.57}$$

To first order for the parameters of Figure 1.20, the *Int(x)* term in Equation 1.55 is zero for $u < 3$meV and one for $u < 6$meV and so will be neglected. This slightly underestimates the real two dimensional density of states and the resulting density of thermal generated carriers. The density of thermally generated carriers above the barrier $E = V$ is thus:

$$n_t = \frac{1}{L}\int_V^{\infty} \rho_{2D} \frac{dE}{e^{(E-E_F)/kT} + 1} \qquad \frac{1}{cm^3} \tag{1.58}$$

which after substituting in Equation 1.55 and neglecting the *Int*(x) term becomes:

$$n_t = \frac{4\pi m^* kT}{h^2 L} e^{-\frac{(V-E_F)}{kT}} \qquad \frac{1}{cm^3} \tag{1.59}$$

Combining Equations 1.59 and 1.54 gives the desired expression for the background generated carrier density:

$$n_t = N_W\left(\frac{kT}{E_F}\right)\frac{L_W}{L} e^{-\frac{(V-E_F)}{kT}} \qquad \frac{1}{cm^3} \tag{1.60}$$

For the 45 degree incident experimental configuration of Levine[40], the background generated carrier density is given by:

$$n_{\Phi B} = \frac{\eta \Omega_{op} \Phi_P \tau}{t\sqrt{2}} \qquad \frac{1}{cm^3} \tag{1.61}$$

where t is the thickness of the active region, 2.65μm, Ω_{op} is the solid angle subtended by the lens system, η is the measured quantum efficiency, 20%, and τ is the effective excess carrier lifetime. The effective excess carrier lifetime is a critical parameter with significant uncertainty. Theoretically, the excess carrier lifetime has been estimated to be 70ps assuming energy loss by electron-phonon interactions, primarily the LO phonon with energy $h\nu_1$ = 36.7meV.[43] Experimentally for the QWIP structure of interest here, the excess carrier lifetime can be estimated from the flattening of the responsivity verses bias voltage at which point $\tau \sim \tau_D/2 = t/2v$, where τ_D is the transit time of a hot electron across the device and v is the saturation velocity which for GaAs is v ~ 5×10^6 cm/s at T=77K.[40] The result for the QWIP structure of Figure 1.20 is τ = 27ps. Combining Equations 1.59 and 1.60 the BLIP condition becomes:

$$\frac{\eta \Omega_{op} \Phi_P \tau}{t\sqrt{2}} > n_t \tag{1.62}$$

which for an f/2 system gives a BLIP temperature of 68K. The excess carrier lifetime can also be used to directly estimate D^*. For a photoconductive detector with radiation incident at angle θ to the normal the peak detectivity D^*_P is given by:

$$D^*_P = \frac{\eta}{2h\nu}\sqrt{\left(\frac{\tau \cos\theta}{n_t t}\right)} \qquad Jones \tag{1.63}$$

which evaluates to $D^*_P = 2 \times 10^{10} Jones$ at 77K and a cutoff wavelength of 10μm. For comparison, under the same conditions an Auger limited ideal MCT detector would have a $D^*_P = 3 \times 10^{12} Jones$ due to the much longer Auger lifetime. However, as argued in Section 1.4.5 D^* is not the appropriate figure of merit for FPAs and in fact as shown in Figure 1.12 an FPA of QWIP detectors with $D^*_P = 2 \times 10^{10} Jones$ and nonuniformity limited by

the 12 bit ADCs used for calibration to $1LSB/\sqrt{12}$ or 0.006% would have an MRT = 3.3mK exactly equal to the MRT of an FPA of MCT detectors with $D_P^* = 3\times10^{12} Jones$ and the same post-correction nonuniformity of 0.006% in spite of the more than two orders of magnitude larger D^* for the MCT FPA. The additional D^* for the MCT detectors is of no consequence looking at a background of 300K since the FPA is uniformity limited.

1.6.5 Summary of Selected IR Detector Technologies

In this section a representative selection of current technologies for IR FPAs has been discussed. In fact, there are many other materials that are now or have been used in infrared cameras. MCT was chosen as a representative of the direct bandgap materials because of its wide use and tunable band gap, but other materials like InSb for MWIR cameras, PbS and other lead salts for LWIR cameras and InAs for SWIR cameras are also available in varying degrees of development and sophistication. All being direct bandgap materials, the fundamental physics is the same as that explained for MCT though the specific material parameters vary from material to material. In the area of Schottky barrier detectors, PtSi was emphasized due to its current status as the most advanced of this class of detectors. Other materials such at IrSi for LWIR detection are only recently starting to receive some attention, but again though the material properties vary the physics of all these Schottky barrier type detectors remains the same as that described for PtSi. The field of thermal detectors goes back to the very first experiment by Herschel demonstrating the existence of IR radiation. In the nearly two centuries since that time many thermal detectors have been investigated, but only recently has it been possible to reduce the size and thermal mass of the individual detectors to the point where one can contemplate making video images with thermal detectors. The BST bolometers were chosen to exemplify this class of detectors because they are currently the most advanced of the thermal detector FPAs, however other detectors based on superconductors, ferromagnetic materials and organic materials to name a few are also being actively pursued. For the QWIP FPAs the GaAs/AlGaAs QWIP was described because of its historical position as the first QWIP detector. Since that time many other material systems and device structures have been

investigated. In fact it is the authors belief that the field of QWIP detectors is still in its infancy with many more exciting discoveries yet to be made. Some of the more recent work including the contributions of this thesis will be discussed in the following chapters.

1.7 Conclusion of Introductory Remarks

Chapter 1 began with a broad historical look at the important physical discoveries that laid the groundwork for modern infrared systems. A discussion of infrared radiation from black and grey bodies and atmospheric absorption, was followed by a detailed look at the figures of merit used in comparing performance of discrete detectors and focal plane arrays. It was found that in even the best of modern day FPAs the performance as measured by minimum resolvable temperature, MRT, is limited by array uniformity and not individual detector performance. This is true even after accounting for calibrations done to minimize spatial nonuniformity. For this reason MRT and not D^* will be used as the preferred figure of merit when comparing FPAs in this thesis. Finally, a brief look was taken at applications and at three representative technologies which compete directly or indirectly with the quantum well intersubband photodetectors which are the subject of this thesis. In the remainder of the thesis, the focus narrows to the specific normal incidence QWIPs developed as part of this thesis until the concluding chapter in which once more a broad look at comparing normal incidence QWIPs to the existing technologies discussed in Section 1.6 will be taken and specific applications in which normal incidence QWIPs may be favored are identified.

Chapter 2

Normal Incidence QWIPs

The first quantum well intersubband photodetectors were discrete devices with light coupled into the device through a 45 degree bevel polished into the substrate as shown in Figure 1.19. This geometry allows coupling into the dominant transverse magnetic, (TM) detection mode, but is clearly inappropriate for use in large focal plane starring arrays that may contain upwards of one hundred thousand devices each needing to be coupled into the normally incident photon flux with very high uniformity. It is the subject of this chapter to first discuss the device parameters that effect the performance of quantum well intersubband photodetectors and then to discuss the methods of coupling normally incident transverse electric, (TE) light into the individual pixels that make up a large focal plane starring array. The discussion of device parameters and performance trade-offs is taken up in Section 2.1. Section 2.2 discusses the theory for the direct TE coupling with no gratings which is the primary subject of this thesis. In particular, results from the superlattice ***K•p*** theory of Ehrenreich et al.[44,48,49] will be used to calculate the performance of FPAs based on TE mode QWIPs. It will be found that minimum resolvable temperatures, (MRT) of less than 10mK are possible for TE mode FPAs operating with 300K backgrounds in the LWIR. Section 2.3 discusses the use of diffraction gratings to couple normally incident light into the TM detector mode. This method is currently used in the best commercial FPAs manufactured by Lockheed/Martin and results from these FPAs will be used as an example for comparison with the TE mode QWIPs of this thesis. Because the grating coupled TM mode FPA are uniformity limited, predicted MRT for the TE mode QWIPs of this thesis is comparable to those obtained by Lockheed/Martin. Section 2.4 discusses other options for normal incidence detection. Many of these have not been widely pursued for FPAs because of issues related to integration, like substrate incompatibility. Particular attention will be paid to areas where the epi-on-electronics / selective area wafer bonding, (E-o-E/SAW) integration technique proposed by the author and dis-

cussed in Chapter 6 may facilitate use of these devices in FPAs. Finally, Section 2.5 concludes the discussion of normal incidence QWIPs with a comparison of the various technologies available for use in large format focal plane starring arrays.

2.1 QWIP device parameters and performance trade-offs

The purpose of this section is to develop a theoretical framework that can be used to discuss the fabrication parameters of quantum well intersubband photodetectors and the performance trade-offs associated with their manipulation. Figure 2.1 shows a schematic for the operation of one quantum well in a QWIP device. The four dominant processes in the QWIP: transport above the quantum well, trapping of electrons by the quantum well, thermionic emission out of the quantum well and photoemission out of the quantum well, are all shown in Figure 2.1.

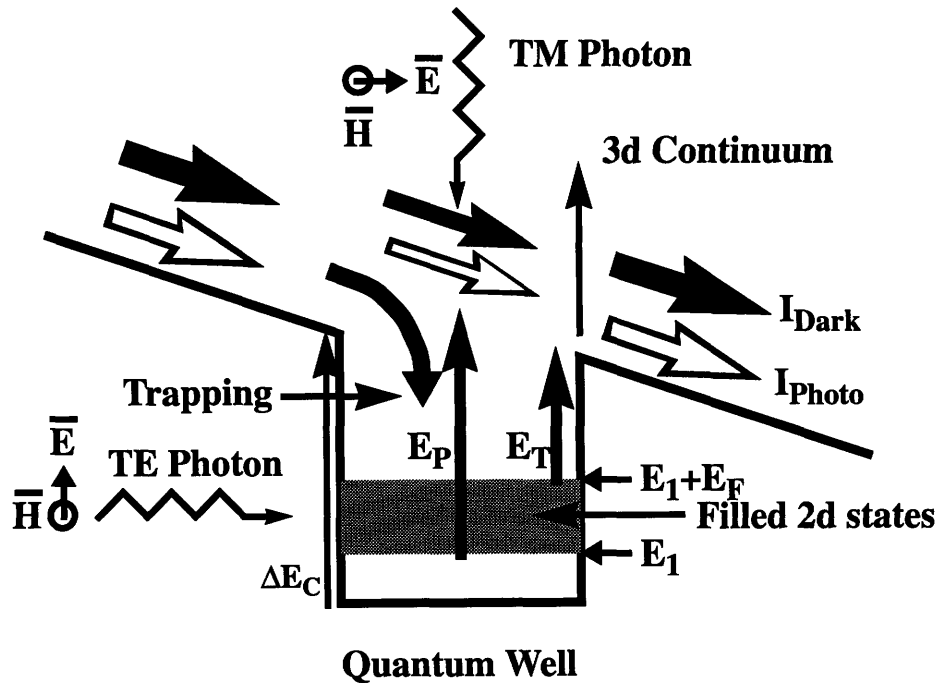


Figure 2.1: Schematic for operation of one well in QWIP. E_T is the activation energy for thermionic emission, E_P is the photon energy for peak detection, E_F is the Fermi level, E_1 is the energy of the first bound state and ΔE_C is the conduction band offset.

If the barrier regions between the quantum wells are sufficiently thick and the applied voltage is sufficiently small so that the contribution from tunneling out of the quantum well is negligible then the major source of dark current is thermionic emission. The thermionic emission current is given by the product of the density of states above the quantum well, the probability of thermal occupation, the area of the device, the electron charge and the velocity of carriers in the barrier as shown in Equation 2.1.

$$I_{Dark} = n_{3d}\exp\left(-\frac{E_T}{kT}\right)Aev(\varepsilon) \qquad Amps \tag{2.1}$$

where A is the area of the device, $v(\varepsilon)$ is the electron velocity as a function of applied electric field, E_T is the thermal activation energy required for an electron to escape the quantum well and is given by $E_T = \Delta E_C - (E_1 + E_F)$, T is the temperature of the device and n_{3d} is the three dimensional density of states in the barrier region into which the thermionic emission occurs given by:

$$n_{3d} = 2\left(\frac{2\pi m^*_{Barrier}kT}{h^2}\right)^{3/2} \qquad \frac{states}{cm^3} \tag{2.2}$$

In terms of the low field mobility and the electron saturation velocity, v_{sat}, the electron velocity as a function of applied field is given by:

$$v(\varepsilon) = \frac{\mu\varepsilon}{\sqrt{1+\left(\frac{\mu\varepsilon}{v_{sat}}\right)^2}} \qquad \frac{cm}{s} \tag{2.3}$$

In equilibrium with no light incident on the detector there are only two processes, thermionic emission and trapping. The thermionic emission of electrons from the well must be balanced by trapping to maintain a constant density of electrons in the well.

When light is incident on the QWIP the dark current remains unchanged, but a photo current is added due to absorption of photons with probability, η^1, where the superscript 1 indicates that this is the quantum efficiency for a single quantum well. In equilibrium, this photo emission of electrons must be balanced by an additional trapping of electrons into

the well to keep the density of electrons in the well constant. Using p as the capture probability for an excited electron traversing a quantum well, the photocurrent directly contributed by photo emission from a single quantum well is given by:[50]

$$i^1_{Photo} = q\Phi_p A\eta^1 (1-p) = q\Phi_p A\eta \frac{(1-p)}{N} \qquad Amps \tag{2.4}$$

where Φ_P is the incident photon flux in photons/cm^2/s, A is the area of the detector, N is the number of quantum wells in the device and $\eta \equiv N\eta^1$ is the total quantum efficiency for all N quantum wells.

Figure 2.2 shows a schematic of a QWIP under bias including N quantum wells and both the emitter and collector contacts. The total directly contributed photocurrent at the collector of the device is given by summing the contributions from each well taking into account that a fraction of the electrons photoexcited by earlier wells will be captured by latter wells. The result is:

$$i_{Photo} = i^1_{Photo} \sum_{n=1}^{N} (1-p)^{n-1} = i^1_{Photo} \frac{1-(1-p)^N}{p} \qquad Amps \tag{2.5}$$

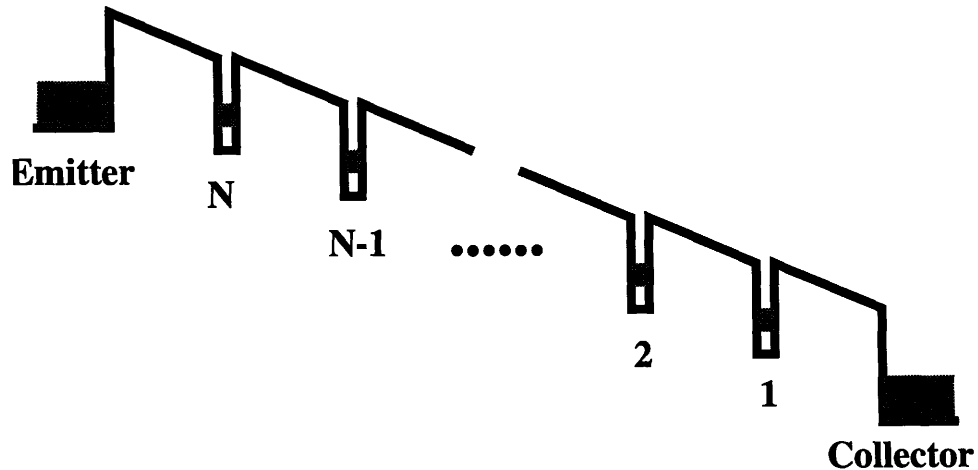


Figure 2.2: Multi-well QWIP under bias.

In addition to the directly contributed photocurrent given by Equation 2.5, there is an extra injection current from the emitter $\delta I_{\mathrm{inject}}$ which must be present in thermal equilibrium in order to maintain a constant density of electrons in each quantum well. This extra injection current must be sufficiently large such that trapping of electrons from this current with probability p exactly cancels the *net* loss electrons due to photoemission and is therefore given by:

$$\delta I_{inject} = \frac{i^1_{Photo}}{p} \qquad Amps \tag{2.6}$$

The fraction of *extra* injection current reaching the collector from all N quantum wells also contributes to the photocurrent seen at the collector since it is indistinquishable from the directly contributed photocurrent. The fraction of *extra* injection current is given by:

$$\delta i = (1-p)^N \delta I_{inject} = i^1_{Photo} \frac{(1-p)^N}{p} \qquad Amps \tag{2.7}$$

The total photocurrent reaching the collector is the sum of Equations 2.5 and 2.7 and is given by:[51]

$$I_{Photo} = i_{Photo} + \delta i = q\Phi_p A\eta\left[\frac{(1-p)}{Np}\right] \qquad Amps \tag{2.8}$$

The term in square brackets is defined as the photoconductive gain, g, for the QWIP.

$$g \equiv \frac{(1-p)}{Np} \tag{2.9}$$

The capture probability, p, for an excited electron traversing a quantum well is clearly a quantum mechanical result but it has a classical analog in bulk photoconductors in which the photoconductive gain is given by the ratio of the carrier lifetime, τ_{life}, to the electrical transit time, τ_{transit}. With this classical definition, the photoconductive gain can be related to electron transport properties in the barrier giving:

$$g_{Classical} = \frac{\tau_{life}}{\tau_{transit}} = \frac{v(\varepsilon)\,\tau_{life}}{NL_P} \tag{2.10}$$

where $v(\varepsilon)$ is the electron velocity as a function of applied field given by Equation 2.3 and L_P is the length of a single superlattice period. Comparing Equations 2.9 and 2.10 under the assumption that the capture probability is small compared to one, $p \ll 1$, gives for the capture probability:

$$p = \frac{L_P}{v(\varepsilon)\tau_{life}} \tag{2.11}$$

Experimentally, the photoconductive gain and therefore the capture probability can be determined by careful measurement of the generation-recombination noise in a photoconductor through use of the noise equation:[52]

$$i_n^2 = 4gqI \qquad Amps^2 \tag{2.12}$$

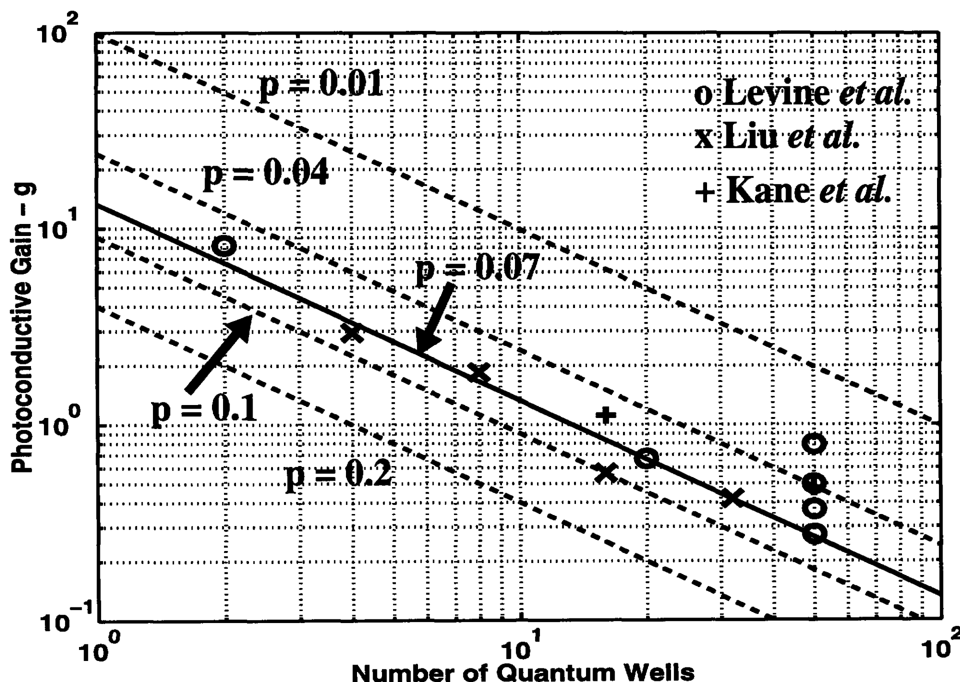


Figure 2.3: Calculated photoconductive gain for capture probability of p = 0.01, 0.04, 0.07, 0.1 and 0.2. Experimental data taken from Levine[40], Liu[51] and Kane[53]. Figure from H.C. Liu.[50]

Values for g calculated from Equation 2.12 for QWIPs and reported by three different laboratories are shown in Figure 2.3 along with lines representing Equation 2.9 with dif-

ferent values for the capture probability. It can be seen that much of the data clusters around the line representing a capture probability of 7%. The value for the capture probability will be used in Chapter 4 for analysis of experimental data collected on the TE mode QWIPs fabricated for this thesis. There is however significant spread in the values which is clear from the five samples with 50 quantum wells which have g values ranging from 0.25 to 0.80 representing capture probabilities ranging from 7% to 2.5%.

All samples are AlGaAs/GaAs QWIPs with similar barrier compositions and similar barrier widths which from Equation 2.11 eliminates both L_P and $v(\varepsilon)$ as contributing to the spread in capture probabilities leaving only the electron lifetime. In a QWIP the electron lifetime is determined by processes that scatter electrons into the quantum well. Physically these processes include scattering due to impurities and electrons in the quantum well, phonons and interface roughness.[54] Experiments varying doping density in the well suggest that impurity and electron-electron scattering may not be the dominant processes because the photoconductive gain did not vary systematically with doping.[52] Phonon scattering would not be expected to vary substantially from sample to sample leaving interface effects as the likely cause of the variation in g. Interfaces between the GaAs wells and AlGaAs barriers are known to be sensitive to growth conditions and can be very different from sample to sample and from one crystal-growth facility to another.

The responsivity of a QWIP is given by the expression

$$R_\lambda = \frac{q\eta\lambda g}{hc} = \frac{q\lambda\eta^1}{hc}\frac{(1-p)}{p} \qquad \frac{Amps}{Watt} \tag{2.13}$$

Notice that since the total quantum efficiency is proportional to the number of quantum wells, $\eta \equiv N\eta^1$, and the photoconductive gain is inversely proportional to the number of quantum wells, the current responsivity is independent of the number of quantum wells. The responsivity is proportional to $1/p$ though for small p which from Equation 2.11 and 2.3 implies that responsivity is proportional to the carrier mobility in the barriers. Carrier mobility in the barrier material is one important design parameter and can be varied over more than one order of magnitude by changing barrier materials. Changes in device design can also effect the single well quantum efficiency, η_1, in two independent ways.

First, the single well quantum efficiency is proportional to the number of electrons in the quantum well and the optical absorption coefficient per electron.[53] Since the Fermi level in a two dimensional system is proportional to the number of occupied states, this proportionality can be written:

$$\eta^1 = \frac{4\pi m^*}{h^2} A E_F \alpha_0 L_P \qquad (2.14)$$

where m^* is the effective mass in the quantum well, E_F is Fermi level in the quantum well, α_0 is the absorption coefficient per electron and L_P is the superlattice period. Equation 2.14 implies that increasing the doping in the quantum well will linearly increase the responsivity and that changes in the absorption coefficient per electron, as can be expected from changes in polarization of the incident light and/or changes in the band structure of the QWIP material also linearly effect the responsivity. The single well quantum efficiency can also be affected by introduction of a optical cavity which though it does not affect the probability of an electron being emitted per photon passing through the quantum well does affect the number of times a particular photon can pass through a particular quantum well thereby changing the apparent single well quantum efficiency. Optical cavities, including waveguides[55,56,57,58] and random scattering surfaces[59,60] have been used in this manner to increase responsivity by nearly an order of magnitude.

The detectivity for a QWIP has been given in Equation 1.7 and can be written in terms of the responsivity as:

$$D^* = \frac{R\sqrt{A_D \Delta f}}{\langle i_N \rangle} \qquad Jones \qquad (2.15)$$

with R given by Equation 2.13, A_D being the area of the detector, Δf being the measurement bandwidth and with the shot noise current, i_N, given by Equation 1.10. There are two operating regimes for the QWIP photodetector depending on whether the noise current is dominated by dark current, D^*_{Dark}, or the background photocurrent, D^*_{BLIP}. The dividing line between these two regimes is called the background limited performance temperature, T_{BLIP}.

For detectors operating at temperatures below T_{BLIP} the detector performance is limited by background photocurrent which can be written in terms of the background photon flux $i_B = \eta e g \Phi_B A_D$. This gives for the noise current:

$$\langle i_N \rangle = 2eg\sqrt{\eta \Phi_B A_D \Delta f} \qquad Amps \tag{2.16}$$

Combining Equations 2.13, 2.15 and 2.16 then gives for the BLIP detectivity:

$$D^*_{BLIP} = \frac{\lambda}{hc}\sqrt{\frac{N\eta^1}{2\Phi_B}} \qquad Jones \tag{2.17}$$

Notice that all the electron transport factors which affected the responsivity have cancelled since they have a similar effect on the detector noise. The only remaining dependences are on the characteristics of the incident photon flux and a square root dependence on both the number of quantum wells and the single well quantum efficiency. For maximum detectivity under BLIP conditions, one needs to increase the quantum well number as much as practical, increase the single electron absorption coefficient, and include an optical cavity with as large a Q as possible, which effectively replace the quantum efficiency η with the product Q times η. In practice, the single well absorption coefficient is determined by the choice of incident light geometry and material parameters leaving only the quantum well number, quantum well doping and optical cavity design for engineering optimization. Note that though the BLIP detectivity increases with the square root of the Fermi energy, the BLIP temperature decreases beyond a critical Fermi energy discussed below meaning that doping cannot be increased indefinitely without violating the assumptions under which Equation 2.17 was derived.

For detectors operating above T_{BLIP} the noise current is determined by the dark current. Combining Equations 2.1, 2.10 and 2.12 then gives for the detectivity above T_{BLIP}:

$$D^*_{Dark} = \frac{\lambda\eta^1}{hc}\sqrt{\frac{N\tau_{life}}{4L_P n_{3d}\exp\left(-\frac{E_T}{kT}\right)}} \qquad Jones \tag{2.18}$$

where $E_T = \Delta E_C - (E_F + E_1)$ and in the bound to continuum type QWIPs of Figure 2.1

the photon energy, E_P, is approximately, $E_P = \Delta E_C - E_1$. Again, the detectivity is independent of the transport properties of the carriers in the barrier region because transport properties affect both the dark current and the photocurrent cancelling their effect on detectivity. Above T_{BLIP} detectivity increases proportional to the square root of the number of quantum wells and the square root of the excited carrier lifetime. Detectivity decreases inversely proportional to the square root of the density of states in the barrier region which can be affected both by choice of barrier material and by barrier heterostructure design as will be discussed for a particular example in Section 2.3. As discussed in Section 1.6.4 when comparing QWIPs to MCT based detectors, the detectivity is proportional to the square root of the excited carrier lifetime. Finally, for operation above T_{BLIP} there is an optimum doping given by maximizing the equation:

$$D^*_{Dark} \propto E_F \exp\left(-\frac{E_F}{2kT}\right) \qquad Jones \tag{2.19}$$

which gives for the optimum Fermi level $E_F = 2kT$. For GaAs quantum well QWIP operating at 77K this corresponds to a sheet electron density of 3.6 x 10^{11} cm^{-2}.

The operating temperature required for background limited performance, T_{BLIP}, can be found by equating the background photocurrent with the dark current.

$$\frac{\eta^1 \Phi_B \tau_{life}}{L_P} = n_{3d}(T) \exp\left(-\frac{E_T}{kT}\right) \qquad \frac{electrons}{cm^3} \tag{2.20}$$

The dominant temperature dependence comes from the exponential, so that the temperature in the density of states, $n_{3d}(T)$, can be replaced by a geometrical average temperature of 77K to good approximation over the QWIP operating range from 50K to 100K. Equation 2.20 can then be solved for the T_{BLIP} giving:

$$T_{BLIP} = \frac{E_T}{k\left(\ln\left(\frac{\pi h n_{3d}(77K)}{16 m^* \alpha_0}\right) - \ln(E_F)\right)} \qquad Kelvin \tag{2.21}$$

As can be seen from Equation 2.21, the BLIP temperature can be increased by decreasing

the density of states in the barrier material or by increasing the absorption coefficient per electron. The dependence on doping density is weak and the optimum value depends on the particular values of the other device parameters. As a rule of thumb for bound to continuum type QWIPs, the product of T_{BLIP} and the cutoff wavelength, $\lambda_C = \frac{hc}{E_T}$, is given by $\lambda_C T_{BLIP} = 650 \mu m K$.[53]

In conclusion, for most FPA applications operating with 300K backgrounds and coolers the object is to produce the highest BLIP temperature possible to minimize cooling requirements. The detector is then operated near or slightly above T_{BLIP} and the detector is designed to give the highest detectivity possible. This requires growing the maximum number of quantum wells practical, designing the highest single well quantum efficiency possible and building the highest Q optical cavity practical and reducing the barrier density of states through choice of materials and superlattice structure. The remainder of Chapter 2 uses the results presented in this subsection to discuss specific detector designs.

2.2 TE mode QWIP FPAs

The possibility of a significant normal incidence, TE mode, absorption was first pointed out by Shik who noted that there are two contributions when the electric field of the incident photon is in the plane of the quantum well.[61,62,63] The smaller of the two contribution results from the spin orbit interaction in the valence band and gives for the ratio of the dipole matrix element for TE to TM polarized light:

$$\frac{\langle i|p_x|j\rangle}{\langle i|p_z|j\rangle} \cong \frac{1}{3}\left[\frac{\Delta_{so}}{E_g + \frac{2}{3}\Delta_{so}}\right] \tag{2.22}$$

where Δ_{SO} is the spin orbit spliting energy in the valence band and E_{g} is the band gap energy. This contribution to the relative absorption ratio (RAS) which is the square of the ratio of the dipole matrix elements is less than 1%. The second contribution results from the finite in-plane wave vector of the conduction band electrons in a doped quantum well and is given by:

$$\frac{\langle i|p_x|j\rangle}{\langle i|p_z|j\rangle} = \frac{\sqrt{2}m_0 k_x}{2\hbar} \frac{\langle F_i \left| \frac{P_0^2}{E_j - E_{lh}} \right| F_j \rangle}{\langle F_i|p_z|F_j\rangle} \tag{2.23}$$

where k_x is the component of the electron wave vector in the direction of the in-plane dipole, P_0 is the interband Kane matrix element, p_z is the matrix element for TM polarized light, E_j is the energy of the *j*th bound state in the quantum well and F_i and F_j are the envelope portions of the electron wave function which are slowly varying over a unit cell. Because the band parameters, namely the zone center masses and energies, vary abruptly between the well and barrier materials, the integral over the envelope functions is nonzero in the expression above. The contribution from electrons with finite in-plane wave vector is sizeable and for transitions between a the ground state in the quantum well and the first continuum state the result has the form:[62]

$$\frac{\langle i|p_x|j\rangle}{\langle i|p_z|j\rangle} \approx m_w\left(\frac{1}{m_w} - \frac{1}{m_b}\right)\sqrt{\frac{E_F}{E_1}} \tag{2.24}$$

where m_w is the electron effective mass in the quantum well, m_b is the electron effective mass in the quantum well, E_F is the Fermi energy and E_1 is the energy of the ground state in the quantum well. For the case of an GaAs quantum well with $Al_{0.30}Ga_{0.70}As$ barriers and a sheet concentration of 1×10^{12} cm^{-2} this evaluates to a relative absorption coefficient of approximately 10%. In a numerical calculation, Flatte et al.[46] obtained the similar result that the TM polarized absorption is about $1200 cm^{-1}$ and the TE polarized absorption is about 200 cm^{-1} for an ideal superlattice consisting of 40Å GaAs quantum wells and 300Å $Al_{0.30}Ga_{0.70}As$ barriers. This calculation used a 14-band model and included transitions to higher order continuum bands with $n = 3\text{-}5$ which were ignored by Shik and have been pointed out as being important by Yang.[64]

Equation 2.24 also implies a slight change in the optimum doping for a TE mode QWIPs operated above T_{BLIP} compared to that calculated for a TM mode QWIPs in Equa-

tion 2.19. This results from the extra factor of E_F which comes in for TE mode devices giving for the detectivity above T_{BLIP}:

$$D^*_{Dark} \propto E_F^2 \exp\left(-\frac{E_F}{2kT}\right) \qquad Jones \tag{2.25}$$

which gives an optimum doping of $E_F = 4kT$ rather than $2kT$ for TM mode devices. For GaAs quantum well QWIP operating at 77K this corresponds to a sheet electron density of 7.2×10^{11} cm^{-2}.

The calculation of Flatte is discussed in Section 2.2.1 and results from this calculation are used in Section 2.2.2 as well as results of Section 1.4 to show that though the TE absorption coefficient is small it is sufficient to produce excellent focal plane array (FPA) images without the additional complexity introduced by adding a grating to couple the normally incident light into the TM absorption mode. This is the first time the potential for using TE absorption in focal plane arrays has been numerically evaluated using the 14-band $\boldsymbol{K \cdot p}$ theory to calculate TE absorption coefficients and then using these absorption coefficients to calculate the theoretical minimum resolvable temperature (MRT) for an FPA based on TE absorption operating without a grating. Finally Section 2.2.3 summarizes the findings and addresses qualitatively possible enhancements to the TE mode absorption due to the effects of strain, applied field and quantum well asymmetry.

2.2.1 Superlattice $\boldsymbol{K \cdot p}$ theory applied to intersubband optical transitions

In this section the superlattice $\boldsymbol{K \cdot p}$ theory of Ehrenreich et.al. [44,48,49] is applied to the case of a GaAs/AlGaAs QWIP following the 14-band model developed by Flatte, Young, Peng and Ehrenreich[46]. The addition of the antibonding conduction $\bar{p}$ states to the 8-band model which includes the two lowest $\bar{s}$ conduction bands and six uppermost p valence bands produces curvature in the heavy-hole (HH) bands and modifies the momentum matrix elements coupling the bands. The increase in TE absorption relative to that calculated above by Shik is associated with the inclusion of the momentum matrix elements P_1 and Q and the inclusion of transitions to continuum bands with indices n = 2-5. In this theory, the matrix element P_1 couples the $\bar{s}$ and $\bar{p}$ conduction bands and is present in the zinc blende lattice which lacks inversion symmetry, but not in the diamond lattice.

The matrix element Q couples the p valence bands to the $\bar{p}$ conduction bands and is present in both zinc blende and diamond lattices. Q is primarily responsible for the curvature of the HH bands in both zinc blende and diamond lattices. The remainder of this section will describe the theory of Flatte et.al.[46] and the results derived from this theory. The derivation for the bulk 14-band Hamiltonian and the bulk 8-band effective Hamiltonian is done in Section 2.2.1.1, the derivation of the superlattice electronic structure is done in Section 2.2.1.2 and the calculation of the intersubband absorption coefficients for a representative GaAs/AlGaAs QWIP structure is done in Section 2.2.1.3.

2.2.1.1 Bulk 14-band Hamiltonian and Effective 8-band Hamiltonian

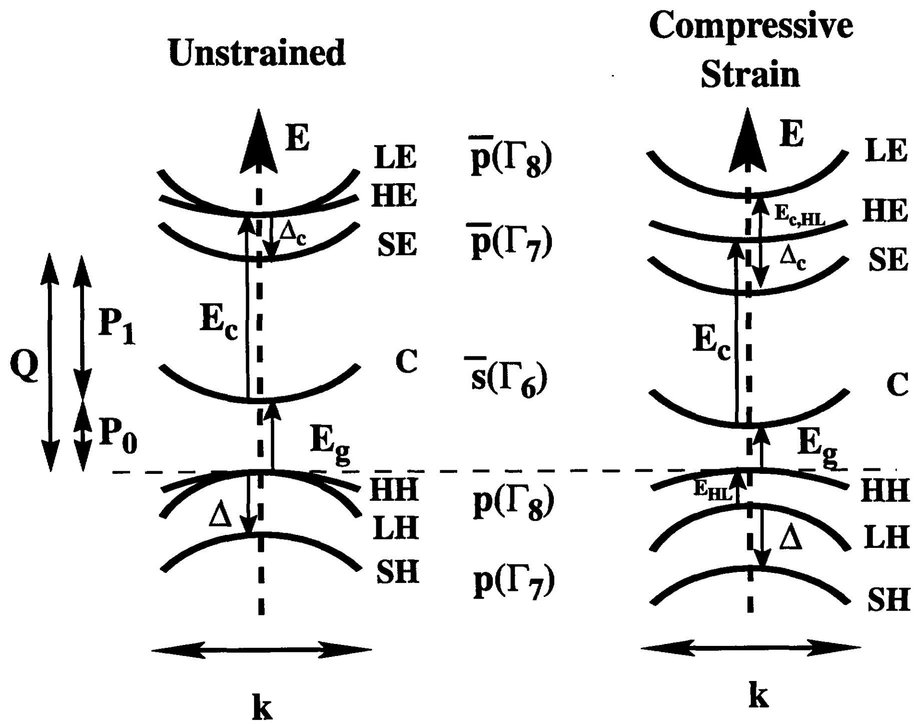


Figure 2.4: Schematic representation of the 14 bulk bands considered. Four energies define the unstrained zone-center bands and six define the strained zone center bands. The matrix element Q couples the six valence band p states to the six conduction band $\bar{p}$ states, the matrix element P_1 couples the conduction band $\bar{s}$ and $\bar{p}$ states and the matrix element P_0 couples the valence band p states to the conduction band $\bar{s}$ states. Overbars refer to antibonding states.

Figure 2.4 shows a schematic representation of the bulk bands including symmetries for the cases with and without strain. The Hamiltonian appropriate for bulk semiconductors is:

$$H(\boldsymbol{r}) = \frac{p^2}{2m} + V(\boldsymbol{r}) + \frac{h}{8\pi(mc)^2}[\sigma \times \nabla V(\boldsymbol{r})] \bullet \boldsymbol{p} \tag{2.26}$$

where m is the free electron mass, c is the velocity of light in a vacuum, h is Plank's constant, σ are the Pauli spin matrices, $\boldsymbol{p}$ is the electron momentum and $V(r)$ is the crystal potential. The 14 $\boldsymbol{k=0}$ basis states $\langle \boldsymbol{r}|n, 0\rangle$, where n is the band index, are defined in Table 2.1 below in terms of the states $|S\rangle$ which transform as atomic s-orbitals, $|X\rangle$, $|Y\rangle$, $|Z\rangle$ which transform as atomic p-orbitals and, $|X^C\rangle$, $|Y^C\rangle$, $|Z^C\rangle$ which also transform as atomic p-orbitals. The subscripts "down" and "up" refer to the spin states.

	$n:|n, 0\rangle$
Γ_6	$1:|S_{up}\rangle$
Γ_8	$2:\sqrt{2/3}|Z_{up}\rangle - \sqrt{1/6}|X_{down} + iY_{down}\rangle$
	$3:\sqrt{1/2}|X_{down} - iY_{down}\rangle$
Γ_7	$4:\sqrt{1/3}|Z_{up}\rangle + \sqrt{1/3}|X_{down} + iY_{down}\rangle$
Γ_6	$5:|S_{down}\rangle$
Γ_8	$6:\sqrt{2/3}|Z_{down}\rangle + \sqrt{1/6}|X_{up} - iY_{up}\rangle$
	$7:\sqrt{1/2}|X_{up} + iY_{up}\rangle$
Γ_7	$8:\sqrt{1/3}|Z_{down}\rangle - \sqrt{1/3}|X_{up} - iY_{up}\rangle$
Γ_8	$9:\sqrt{2/3}|Z_{up}\rangle - \sqrt{1/6}|X_{down} + iY_{down}\rangle$
	$10:\sqrt{1/2}|X^C{}_{down} - iY^C{}_{down}\rangle$
Γ_7	$11:\sqrt{1/3}|Z^C{}_{up}\rangle + \sqrt{1/3}|X^C{}_{down} + iY^C{}_{down}\rangle$
Γ_8	$12:\sqrt{2/3}|Z^C{}_{down}\rangle + \sqrt{1/6}|X^C{}_{up} - iY^C{}_{up}\rangle$
	$13:\sqrt{1/2}|X^C{}_{up} + iY^C{}_{up}\rangle$
Γ_7	$14:\sqrt{1/3}|Z^C{}_{down}\rangle - \sqrt{1/3}|X^C{}_{up} - iY^C{}_{up}\rangle$

Table 2.1: Bulk Kane *k=0* states for 14 band model.[44,45]

The resulting 14 band Hamiltonian has the block diagonal form:

$$\begin{bmatrix} H_{\bar{ps}} & H_{\bar{ps}-\bar{p}} \\ (8x8) & (8x6) \\ H^{\dagger}_{\bar{ps}-\bar{p}} & H_{\bar{p}} \\ (6x8) & (6x6) \end{bmatrix} \tag{2.27}$$

The constituent matrices in Equation 2.27 can be written in diagonal form as follows:

$$H_{\bar{ps}} = \begin{bmatrix} h_{\bar{ps}} & 0 \\ 0 & h_{\bar{ps}} \end{bmatrix} \tag{2.28}$$

where

$$h_{\bar{ps}} = \frac{h^2k^2}{8\pi^2 m} + \begin{bmatrix} E_g & \frac{iP_0hk}{2\pi m} & 0 & \frac{iP_0hk}{2\sqrt{2}\pi m} \\ -\frac{iP_0hk}{2\pi m} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -\frac{iP_0hk}{2\sqrt{2}\pi m} & 0 & 0 & -\Delta \end{bmatrix} \tag{2.29}$$

Equation 2.28 is the standard 8x8 ***k•p*** Hamiltonian with no heavy hole dispersion.

$$H_{\bar{p}} = \begin{bmatrix} h_{\bar{p}} & 0 \\ 0 & h_{\bar{p}} \end{bmatrix} \tag{2.30}$$

where

$$h_{\bar{p}} = \begin{bmatrix} E_c + E_g & 0 & 0 \\ 0 & E_c + E_g & 0 \\ 0 & 0 & E_c + E_g - \Delta_c \end{bmatrix} \tag{2.31}$$

$$H_{p\bar{s}-\bar{p}} = \begin{bmatrix} h_{p\bar{s}-\bar{p}} & 0 \\ 0 & h_{p\bar{s}-\bar{p}} \end{bmatrix} \tag{2.32}$$

where

$$h_{p\bar{s}-\bar{p}} = \begin{bmatrix} \frac{iP_1 hk}{2\pi m} & 0 & \frac{iP_1 hk}{2\sqrt{2}\pi m} \\ 0 & -\frac{Qhk}{2\sqrt{2}\pi m} & 0 \\ \frac{Qhk}{2\sqrt{2}\pi m} & 0 & \frac{Qhk}{2\pi m} \\ 0 & \frac{Qhk}{2\pi m} & 0 \end{bmatrix} \tag{2.33}$$

To simplify the solution of the bulk 14-band model, the block diagonalization procedure of Cohen and Heine[65] is used. In this procedure the matrix:

$$M = \begin{bmatrix} I & -\frac{H_{p\bar{s}-\bar{p}}}{H_{p\bar{s}} - E} \\ -\frac{H^{\dagger}_{p\bar{s}-\bar{p}}}{H_{\bar{p}} - E} & I \end{bmatrix} \tag{2.34}$$

with $H_{p\bar{s}-\bar{p}}$ given by Equation 2.32, and its inverse M^{-1} are inserted into the Schrödinger equation $(H - E)\Phi = 0$ giving:

$$(H - E)MM^{-1}\Phi = 0 \tag{2.35}$$

defining $\Phi^{eff} \equiv M^{-1}\Phi$ this equation becomes:

$$\begin{bmatrix} H^{eff}_{p\bar{s}} - E & 0 \\ 0 & H^{eff}_{\bar{p}} - E \end{bmatrix} \Phi^{eff} = 0 \tag{2.36}$$

where

$$H^{eff}_{p\bar{s}} - E = H_{p\bar{s}} - E - H_{p\bar{s}-\bar{p}}\frac{H^{\dagger}_{p\bar{s}-\bar{p}}}{H_{\bar{p}} - E} \equiv H_{p\bar{s}} - E + \delta H_{p\bar{s}} \tag{2.37}$$

The structure of Equation 2.36 breaks the 14-component Φ^{eff} into an 8-component $\Phi^{eff}_{p\bar{s}}$ and a 6-component $\Phi^{eff}_{\bar{p}}$ with $\Phi_{p\bar{s}} = \Phi^{eff}_{p\bar{s}}$ and $\Phi_{\bar{p}} = -(H_{\bar{p}} - E)^{-1}H^{\dagger}_{p\bar{s}-\bar{p}}\Phi^{eff}_{p\bar{s}}$ where the solutions of interest which reduce the 14-band problem to an effective 8-band problem are $\Phi^{eff}_{\bar{p}} = 0$. This reduces the numerical complexity sufficiently to justify the extra effort in folding down. The matrix $H^{eff}_{p\bar{s}}$ is block diagonal and of the form:

$$H^{eff}_{p\bar{s}} = \begin{bmatrix} h^{eff}_{p\bar{s}} & 0 \\ 0 & h^{eff}_{p\bar{s}} \end{bmatrix} \tag{2.38}$$

where $h^{eff}_{p\bar{s}} = h_{p\bar{s}} + \delta h_{p\bar{s}}$ and $\delta h_{p\bar{s}}$ is given by:

$$\delta h_{p\bar{s}} = \frac{h^2k^2}{8\pi^2 m}\begin{bmatrix} \frac{P_1^2}{m}\left[\frac{2}{E-E_c-E_g} + \frac{1}{E-E_c-E_g+\Delta_c}\right] & 0 & \frac{iP_1Q}{m}\left[\frac{\sqrt{2}}{E-E_c-E_g} - \frac{\sqrt{2}}{E-E_c-E_g+\Delta_c}\right] & 0 \\ 0 & \frac{Q^2}{m\left[E-E_c-E_g\right]} & 0 & -\frac{\sqrt{2}Q^2}{m\left[E-E_c-E_g\right]} \\ -\frac{iP_1Q}{m}\left[\frac{\sqrt{2}}{E-E_c-E_g} - \frac{\sqrt{2}}{E-E_c-E_g+\Delta_c}\right] & 0 & \frac{Q^2}{m}\left[\frac{1}{E-E_c-E_g} + \frac{2}{E-E_c-E_g+\Delta_c}\right] & 0 \\ 0 & -\frac{\sqrt{2}Q^2}{m\left[E-E_c-E_g\right]} & 0 & \frac{Q^2}{m\left[E-E_c-E_g\right]} \end{bmatrix} \tag{2.39}$$

where rows and columns of Equation 2.39 are labeled sequentially with the states C (conduction band), LH (light hole), HH (heavy hole) and SH (Split-off hole). Equation 2.39 gives the additional terms in the standard 8x8 $\boldsymbol{k \cdot p}$ Hamiltonian[44] represented in Equation 2.28 which result from the inclusion of the antibonding $\bar{p}$ conduction band states to

the basis containing the antibonding $\bar{s}$ conduction band and the bonding $\bar{p}$ valence band states.

At this point Equation 2.36 can be solved using second-order perturbation theory by replacing E with the zone-center energy of the relevant band. Notice that the major effect of the antibonding $\bar{p}$ conduction band states is to make the heavy-hole mass finite; i.e. the (3,3) matrix element in the 8x8 of Equation 2.28 is zero corresponding to infinite heavy-hole mass but becomes finite in the 14x14 band model due to the non-zero (3,3) element in Equation 2.39. Masses for all four particles including the effects of the antibonding $\bar{p}$ conduction band states are:

$$\frac{m}{m_C} = 1 + \frac{P_0^2 (3E_g + 2\Delta)}{mE_g (E_g + \Delta)} - \frac{P_1^2 (3E_c - 2\Delta_c)}{mE_c (E_c - \Delta_c)} \tag{2.40}$$

$$\frac{m}{m_{HH}} = 1 - \frac{Q^2 (3E_c + 3E_g - \Delta_c)}{m (E_c + E_g) (E_c + E_g - \Delta_c)} \tag{2.41}$$

$$\frac{m}{m_{LH}} = 1 - \frac{2P_0^2}{mE_g} - \frac{Q^2}{m (E_c + E_g)} \tag{2.42}$$

$$\frac{m}{m_{SH}} = 1 - \frac{P_0^2}{m (E_g + \Delta)} - \frac{Q^2}{m (E_c + E_g + \Delta)} \tag{2.43}$$

where the three momentum matrix elements which couple the states away from the zone center in the $\boldsymbol{k \cdot p}$ formalism are defined as:

$$P_0 = -i\sqrt{2/3} \langle S|p_z|Z^v \rangle \tag{2.44}$$

$$P_1 = -i\sqrt{2/3} \langle S|p_z|Z^c \rangle \tag{2.45}$$

$$Q = -i\sqrt{2/3} \langle X^v|p_y|Z^c \rangle = i\sqrt{2/3} \langle X^c|p_y|Z^v \rangle \tag{2.46}$$

Since the electron mass in Equation 2.40 only weakly depends on the value of the matrix element P_1, P_1 is determined from the psuedopotential calculation of Cardona et.al.[45] which yields $P_1 = 0.48P_0$. The experimental values of m_C and m_{HH}[47] then determine the remaining parameters P_0 and Q. Table 2.2 gives a comparison of the band parameters for the 8 band model[44], the 14 band model[46] and the psuedopotential calculations[45]. E_g is the fundamental bandgap, E_c is the zone center gap between $\bar{s}$ and $\bar{p}$ states, Δ is the spin-orbit splitting for the valence band $\bar{p}$ states and Δ_c is the spin-orbit splitting for the antibonding conduction band $\bar{p}$ states.

	8 bands[44]	14 bands[46]	Psuedopotential[45]
E_g (eV)	1.519	1.519	1.519
Δ (eV)	0.341	0.341	0.340
E_c (eV)	N/A	3.140	3.140
Δ_c (eV)	N/A	0.171	0.171
$E_{P_0} = 2\|\langle S\|p_z\|Z^v\rangle\|^2/m$ (eV)	24.1	25.7	26.0
$E_{P_1} = 2\|\langle S\|p_z\|Z^c\rangle\|^2/m$ (eV)	N/A	5.9	6.0
$E_Q = 2\|\langle X^v\|p_z\|Z^c\rangle\|^2/m$ (eV)	N/A	13.5	12.0
m_{HH}/m	0.51	N/A	N/A

Table 2.2: Comparison of parameters for 8 and 14 band models for GaAs[46]

Table 2.3 shows a comparison of the four experimental masses[66] m_c, m_{LH}, m_{HH} and m_{SH} with the values calculated from the 14-band model used here and the psuedopotential of Cardona. Considering the accuracy of the experiments, agreement is reasonable.

	Experiment[47]	14 bands[46]	Psuedopotential[45]
m_C/m	0.067	0.067	0.066
m_{LH}/m	0.082	0.089	0.089
m_{HH}/m	0.51	0.51	0.61
m_{SH}/m	0.154	0.22	0.22

Table 2.3: Comparison of experimental and calculated masses for GaAs

Figure 2.5 shows a comparison of the 8-band model[44] which includes an empirical term for the heavy hole mass with the 14-band model[46] for the bulk band structure of GaAs. The most striking feature is the unphysical crossing of the light hole and heavy hole bands in the 8-band model at moderate values of k. This is eliminated in the 14-band model with the inclusion of the antibonding conduction band $\bar{p}$ states which couple to the bonding valence band p states through the matrix element Q and eliminate the unphysical crossing of heavy hole and light hole bands. Figure 2.6 shows a blowup of the region around the unphysical crossing of the heavy hole and light hole bands in the 8-band model and the correct anti-crossing behavior of the 14-band model. Because this crossing occurs at fairly large k, $k > 0.16\text{Å}^{-1}$, the 8-band model is considered satisfactory for many calculations.

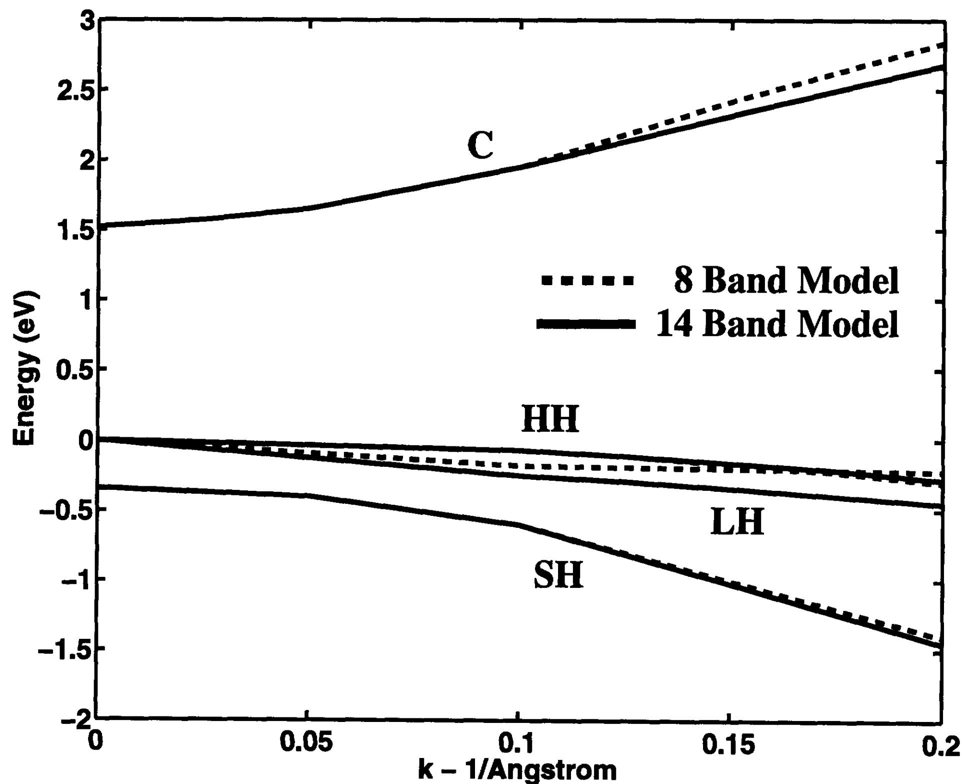


Figure 2.5: Comparison of the bulk 14-band and 8-band $\boldsymbol{k \cdot p}$ models for GaAs.

The band structure shown in Figures 2.5 and 2.6 for GaAs and a similar band structure for AlGaAs are the inputs into the superlattice (SL) $\boldsymbol{K \cdot p}$ calculation for calculation of the

superlattice wave functions and subsequently the SL optical absorption coefficients. The following subsection describes the method of Flatte, Young, Peng and Ehrenreich in calculation of these superlattice wave functions.[46] It will be shown that the 14-band model leads to a TM absorption of $1200cm^{-1}$ in excellent agreement with published experiments and a TE absorption of $200cm^{-1}$. The TE absorption is about a factor of two larger than that calculated in Subsection 2.2.2 using the method of Shik and as will be shown in Section 2.2.4 is sufficient to give excellent FPA images without the requirement for fabrication of a coupling grating.

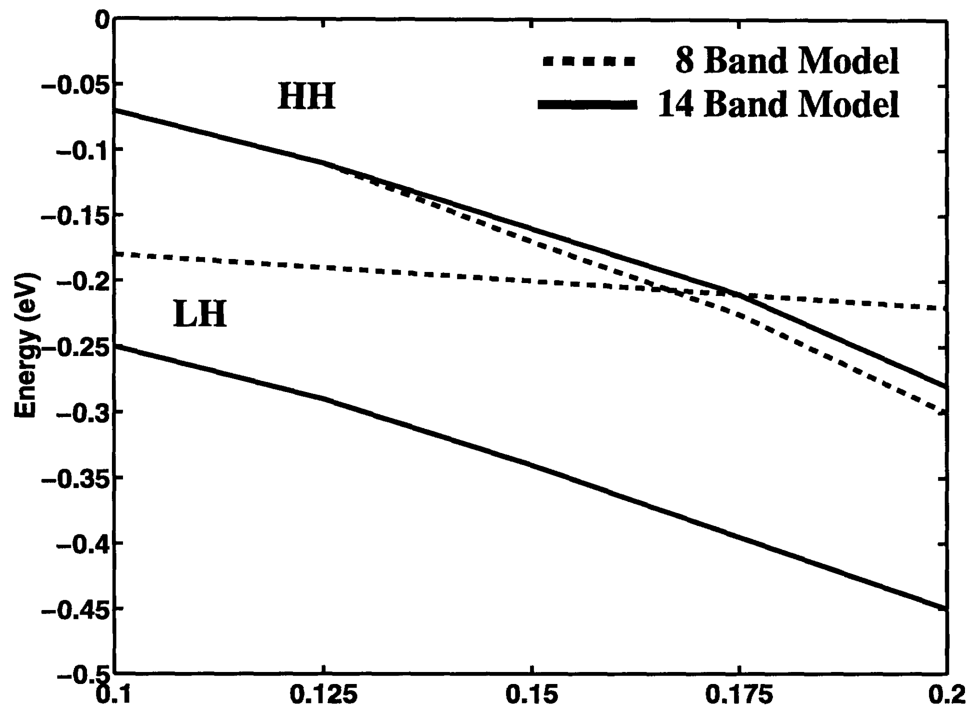


Figure 2.6: A vertical blowup of the region near the unphysical crossing of light and heavy hole bands in the 8-band model.[46]

2.2.1.2 Superlattice Electronic Structure

Having calculated the14-band electronic structure for bulk semiconductors, solutions to the band structure of a superlattice (SL) consisting of alternating layers, A and B, of two bulk semiconductors are sought. The layers A and B are alternately stacked in the z direction with period P. In the limit of infinite extent, the SL is a perfectly periodic, though

highly anisotropic, crystal with wave functions $\Psi_{SL}(L, \boldsymbol{K})$ that can be characterized by SL wave vector $\boldsymbol{K}$ and band index L and which obey Bloch's theorem. Assuming that the superlattice period is large compared to the underlying bulk crystal lattice constant, the $\boldsymbol{K=0}$ states of the superlattice can be written as products of slowly varying envelope functions $F_n(z)$ and the $\boldsymbol{k=0}$ Bloch functions $\langle \boldsymbol{r}|n, 0\rangle$ corresponding to band n of the 14 bands calculated for bulk semiconductors in Section 2.2.1.1.

$$\langle \boldsymbol{r}|\Psi_{SL}(L, \boldsymbol{K}=0)\rangle \equiv \langle \boldsymbol{r}|L, \boldsymbol{K}=0\rangle = \sum_{n} F_n(L, \boldsymbol{K}=0;z)\,\langle \boldsymbol{r}|n, 0\rangle \qquad (2.47)$$

Substituting the envelope-function expression of Equation 2.47 along with the Hamiltonian of Equation 2.26 into the Schrödinger equation $(H - E)\Phi = 0$ and assuming that the wave functions of the bulk states $\langle \boldsymbol{r}|n, 0\rangle$ are nearly identical in the SL layers A and B due to the similarity of the two materials, the multiband Schrödinger equation for the $\boldsymbol{K=0}$ envelope functions can now be written as:

$$\left[H_{A(B)}\left(k_x, k_y, k_z \Rightarrow i\frac{\partial}{\partial z}\right) - E\right]\boldsymbol{F}(L, \boldsymbol{K}=0;z) = 0 \qquad (2.48)$$

Solutions are sought to Equation 2.48 for the energies and envelope functions which satisfy continuity of cell averaged current across the interface between layers A and B. These boundary conditions for the envelope functions $\boldsymbol{F}(L, \boldsymbol{K}=0;z)$ are obtained by integrating Equation 2.48 across the interface between layers A and B:[46]

$$\sum_{n'} \langle n, 0|p_z|n', 0\rangle F_{n'}(z_+) + \frac{1}{m}\frac{\partial}{\partial z}F_n(z_+) = \sum_{n'} \langle n, 0|p_z|n', 0\rangle F_{n'}(z_-) + \frac{1}{m}\frac{\partial}{\partial z}F_n(z_-) \qquad (2.49)$$

where $\langle n, 0|p_z|n', 0\rangle$ is the momentum matrix element connecting bulk bands.

Having determined the $\boldsymbol{K=0}$ envelope functions, a superlattice $\boldsymbol{K \bullet p}$ theory[44] is used to express the finite $\boldsymbol{K}$ SL states $|L, \boldsymbol{K}\rangle$ in terms of the $\boldsymbol{K=0}$ envelope function states.

$$\begin{aligned} \langle \boldsymbol{r}|L, \boldsymbol{K}\rangle &= e^{i\boldsymbol{K}\bullet\boldsymbol{r}} \sum_{n} F_n(L, \boldsymbol{K};z)\,\langle \boldsymbol{r}|N, 0\rangle \\ &\equiv e^{i\boldsymbol{K}\bullet\boldsymbol{r}} \sum_{N} c_{LN}(\boldsymbol{K})\,\langle \boldsymbol{r}|N, 0\rangle \end{aligned} \qquad (2.50)$$

Applying the Hamiltonian of Equation 2.26 to the states in Equation 2.50 gives the SL $\boldsymbol{K}\bullet\boldsymbol{p}$ equation for the c_{LN}'s as a function of $\boldsymbol{K}$.

$$\sum_N \left[\left(E_N(0) + \frac{h^2 k^2}{8\pi^2 m} - E_L(\boldsymbol{K}) \right) \delta_{NN'} + \langle N', 0|H'|N, 0\rangle \right] c_{LN'}(\boldsymbol{K}) = 0 \tag{2.51}$$

where H' is the $\boldsymbol{K}\bullet\boldsymbol{p}$ part of the Hamiltonian and describes dispersion in the SL bands

$$\langle N', 0|H'|N, 0\rangle = \frac{h\boldsymbol{K}}{2\pi m} \langle N', 0|\boldsymbol{p}|N, 0\rangle \tag{2.52}$$

and the momentum matrix elements connecting the superlattice $\boldsymbol{K=0}$ states are

$$\begin{aligned} \langle L', 0|\boldsymbol{p}|L, 0\rangle &= \frac{1}{V}\sum_{nn'} \Big(\Big[\int d\boldsymbol{r} F^*_{n'}(L', 0;z) F_n(L, 0;z)\Big] \langle n', 0|\boldsymbol{p}|n, 0\rangle \\ &\qquad + \Big[\int d\boldsymbol{r} F^*_{n'}(L', 0;z) \boldsymbol{p} F_n(L, 0;z)\Big] \delta_{nn'} \Big) \\ &\equiv \sum_{nn'} (\alpha_{nn'}(L, L') \langle n', 0|\boldsymbol{p}|n, 0\rangle + \Pi_n(L, L') \delta_{nn'}) \end{aligned} \tag{2.53}$$

with V being the crystal volume. The Π terms have been shown to be smaller by a factor m_C/m relative to the α terms and will subsequently be ignored.[44]. At this point the folding down procedure of Flatte et.al.[46] will be applied to the SL bands in a manner analogous to that done for the bulk bands in Section 2.2.1.1 in order to concentrate on the effects of the $\bar{p}$ on the conduction band and valence bands of interest. The SL K=0 states are written here in terms of both the 14-component F_n and the 8-component F_n^{eff}.

$$\begin{aligned} \langle \boldsymbol{r}|\Psi_{SL}\rangle &= \sum_{n=1}^{14} F_n(L, K = 0;z) \langle \boldsymbol{r}|n, 0\rangle \\ &= \sum_{n=1}^{8} F_n^{eff}(L, K = 0;z) \langle \boldsymbol{r}|n, 0\rangle^{eff} \end{aligned} \tag{2.54}$$

where analogous to the bulk case, the 14-component

$$F = \begin{bmatrix} F_{\bar{ps}} \\ F_{\bar{p}} \end{bmatrix} \quad (2.55)$$

is expressed in terms of the 8-component F^{eff} for K away from the zone center by:

$$F_{\bar{ps}}(L, K) = F^{eff}_{\bar{ps}}(L, K) \quad (2.56)$$

$$F_{\bar{p}}(L, K) = -\left(H_{\bar{p}} - E_L\right)^{-1} H^{\dagger}_{\bar{ps}-\bar{p}} F_{\bar{ps}}(L, K) \quad (2.57)$$

The envelope function equation for the effective 8-component model is then:

$$H^{k \bullet p}_{A\,(B)}\left(k_x, k_y, k_z \Rightarrow i\frac{\partial}{\partial z}\right) F_{\bar{ps}}(L, K = 0; z) = E F_{\bar{ps}}(L, K = 0; z) \quad (2.58)$$

where the Hamiltonian for the bulk materials A and B comes from Equation 2.38 and including the z dependence due to the alternating A and B SL layers is given by:

$$\begin{bmatrix} E_{\Gamma_6}(z) & \frac{ihP_0k_z}{2\pi m} & 0 & \frac{ihP_0k_z}{\sqrt{2}2\pi m} \\ -\frac{ihP_0k_z}{2\pi m} & E_{\Gamma_8}(z) & 0 & 0 \\ 0 & 0 & E_{\Gamma_8}(z) - k_z\frac{h^2}{8\pi^2 m_{HH}(z)}k_z & 0 \\ -\frac{ihP_0k_z}{\sqrt{2}2\pi m} & 0 & 0 & E_{\Gamma_7}(z) \end{bmatrix} \begin{bmatrix} F_C(z) \\ F_{LH}(z) \\ F_{HH}(z) \\ F_{SH}(z) \end{bmatrix} = E \begin{bmatrix} F_C(z) \\ F_{LH}(z) \\ F_{HH}(z) \\ F_{SH}(z) \end{bmatrix} \quad (2.59)$$

In addition to the bulk parameters of Table 2.1, the solution of Equation 2.51 for the superlattice ***K•p*** band structure requires values for the layer thicknesses and the valence band offset $\Lambda = E_{\Gamma_8}(A) - E_{\Gamma_8}(B)$. For the GaAs/AlGaAs quantum well structure discussed in Sections 2.2.1.3 and 2.2.2 these values are given below in Table 2.4.

Using the procedure described in this section and the parameters of Table 2.4, Flatte et.al.[46] have solved the superlattice ***K•p*** band structure for the case of a 40Å GaAs SL A layer and a 300Å $Al_{0.30}Ga_{0.70}As$ SL B layer. The results of this calculating are described

in the following section in terms of the resulting optical matrix elements and the absorption coefficients for intersubband optical absorption.

	GaAs	$Al_{0.3}GaAs$	GaAs/ $Al_{0.3}GaAs$
E_g RT fundamental gap (eV)	1.420	1.786	
Δ spin-orbit splitting (eV)		0.328	
m_{HH}/m in (100) direction		0.48	
Λ valence band offset (eV)			0.138

Table 2.4: *K•p* SL band structure parameters[46]

2.2.1.3 Intersubband Optical Absorption from 14-band SL *K•p* theory

The superlattice momentum matrix elements discussed in Section 2.2.1.2 involving p_z and p_x determine respectively, the intersubband optical absorption coefficients for light polarized with electric field perpendicular to the superlattice growth direction, referred to as TM, and for light polarized with electric field in the plane of the superlattice layers, referred to as TE. The expressions for the TE and TM intersubband absorption coefficients in the notation of Section 2.2.1.2 are

$$\begin{aligned} \alpha_{x(z)} = {} & \frac{2\pi^2 e^2}{m^2 c\omega V}\sum_{\boldsymbol{K}}\sum_{L,L'}[f_n(E_{L'}(\boldsymbol{K})) \\ & - f_n(E_L(\boldsymbol{K}))]\left|\langle L,\boldsymbol{K}|p_{x(z)}|L',\boldsymbol{K}\rangle\right|^2 \\ & \times\delta[E_L(\boldsymbol{K}) - E_{L'}(\boldsymbol{K}) - h\nu] \end{aligned} \tag{2.60}$$

and

$$\begin{aligned} \langle L,\boldsymbol{K}|p_{x(z)}|L',\boldsymbol{K}\rangle & = \boldsymbol{F}^{\dagger}(L,\boldsymbol{K})\boldsymbol{p}_{x(z)}\boldsymbol{F}(L',\boldsymbol{K}) \\ & = \boldsymbol{F}^{\dagger}_{p\bar{s}}(L,\boldsymbol{K})\boldsymbol{p}_{x(z),p\bar{s}\,p\bar{s}}\boldsymbol{F}_{p\bar{s}}(L',\boldsymbol{K}) \\ & + \boldsymbol{F}^{\dagger}_{p\bar{s}}(L,\boldsymbol{K})\boldsymbol{p}_{x(z),p\bar{s}\,\bar{p}}\boldsymbol{F}_{\bar{p}}(L',\boldsymbol{K}) \\ & + \boldsymbol{F}^{\dagger}_{\bar{p}}(L,\boldsymbol{K})\boldsymbol{p}_{x(z),p\bar{s}-\bar{p}}\boldsymbol{F}_{p\bar{s}}(L',\boldsymbol{K}) \end{aligned} \tag{2.61}$$

where for TE absorption the subscript x is used and for TM absorption the subscript z is used and the matrix elements between superlattice states $\langle L, K|p_{x(z)}|L', K\rangle$ are given by Equation 2.61. The 14-band bulk basis has been decomposed into an 8x8 matrix $p_{x(z), p\bar{s}}$, a 6x6 matrix $p_{x(z), \bar{p}}$ and an 8x6 matrix $p_{x(z), p\bar{s}-\bar{p}}$.

Before evaluating these Equations 2.60 and 2.61, insight into the expected results can be obtained from the f-sum rule for intersubband transitions derived by Johnson et.al.[67]. This rule relates the matrix elements involving p_x and p_z to the effective masses along and perpendicular to the superlattice stacking direction z as follows:

$$(m/m_L)_\alpha = 1 + {\sum_{L'}}' f^{\alpha}_{L',L} \tag{2.62}$$

where the prime on the summation indicates that the $L' = L$ term is excluded and the oscillator strengths $f^{\alpha}_{L',L}$ are given by:

$$f^{\alpha}_{L',L} = \frac{2}{m} \frac{\langle L,0|p_\alpha|L',0\rangle}{E_L(0) - E_{L'}(0)} \tag{2.63}$$

Cyclotron resonance measurements of the electron mass parallel and perpendicular to the superlattice stacking direction[68] show a 50% difference between the C1 effective masses m^z_{C1} and m^x_{C1} which Johnson et al. used in combination with the f-sum rule of Equation 2.62 to argue that $f^z_{C2,C1} \gg f^x_{C2,C1}$.

This leads to the conclusion that the $C1 \Rightarrow C2$ intersubband TM absorption determined by p_z is considerably larger than the intersubband TE absorption determined by p_x. This fact remains unchanged in the 14-band model as shown in Figure 2.7 which shows both the TE and TM absorption as a function of photon energy for a superlattice consisting of a 40Å GaAs quantum well and a 300Å $Al_{0.30}Ga_{0.70}As$ barrier. The absorption length for this calculation is the entire superlattice period, i.e. 340Å and the relative absorption strength (RAS) defined as the ratio of TM/TE absorption coefficients is 0.19. The RAS value of 0.19 for the 14-band model including continuum states n = 2-5 compares to a RAS value of 0.10 for the calculation of Shik[62]. Figure 2.8 shows a comparison of the

8-band and 14-band calculations with the experimental data of Levine et al.[69] for the same superlattice consisting of a 40Å GaAs quantum well and a 300Å $Al_{0.30}Ga_{0.70}As$ barrier.

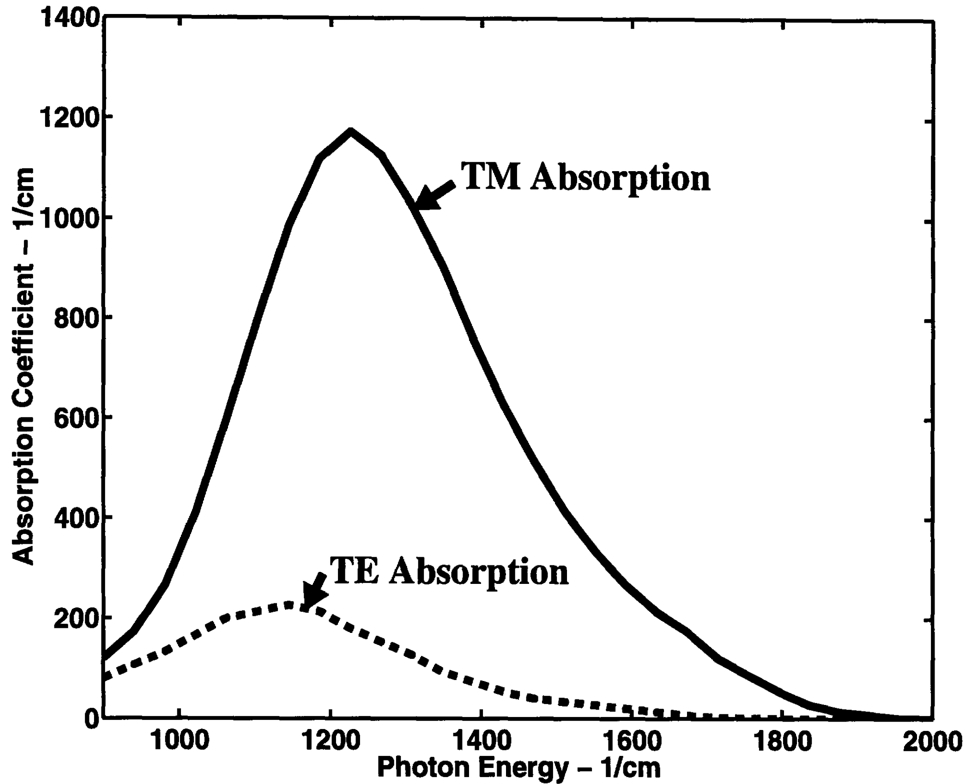


Figure 2.7: 14 band model calculation of TE and TM absorption coefficients for a superlattice of 40Å GaAs and 300Å $Al_{0.30}Ga_{0.70}As$ at 300K. Absorption length is entire superlattice period, 340Å.[46]

Note that the already good agreement between the experiment and the 8-band model is improved significantly by addition of the antibonding $\bar{p}$ conduction band states in the 14-band model. Since the *f*-sum rule for intersubband transitions discussed above links the TE and TM absorption coefficients, the excellent agreement between the TM absorption coefficient calculated with the 14-band model also improves confidence in the TE absorption coefficient calculated with the same 14-band model. As an example of the implications of the small, but significant TE absorption coefficient calculated in the 14-band model, the total percent absorption for a superlattice consisting of 50 periods of 40Å GaAs quantum wells and 300Å $Al_{0.30}Ga_{0.70}As$ barriers is 3.4% in the 14-band model. This is

small compared to the more than 20% absorption for the TM mode, but as shown for the first time in the following section even this relatively small absorption can be used in infrared focal plane array cameras to make surprisingly good quality images.

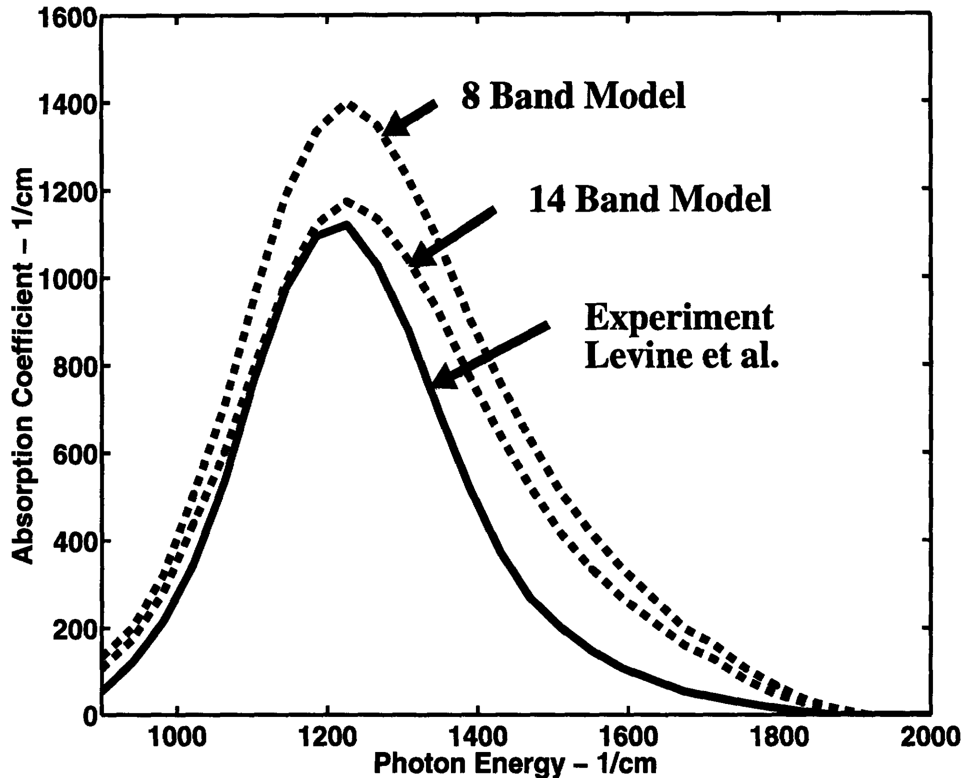


Figure 2.8: Comparison of TM absorption for a superlattice of 40Å GaAs and 300Å $Al_{0.30}Ga_{0.70}As$ at 300K. Dotted lines are calculations done with 8-band and 14-band models and solid line is experimental data form Levine.[69] Absorption length is entire superlattice period, 340Å.[46]

2.2.2 Potential for using TE absorption in QWIP focal plane arrays

In Section 1.4.5 a model of an IR focal plane array (FPA) camera was described that included the effects of non-uniformity in the FPA. Here this model will be applied to the case of an FPA using quantum well intersubband photodetector (QWIP) pixels operating using the TE absorption mode. This FPA therefore has no coupling grating or any other structure designed to excite the TM absorption mode. The TE mode QWIP pixel has a very simple measurement geometry in which the light is incident perpendicular to either the front or back of the detector as shown in Figure 2.9.

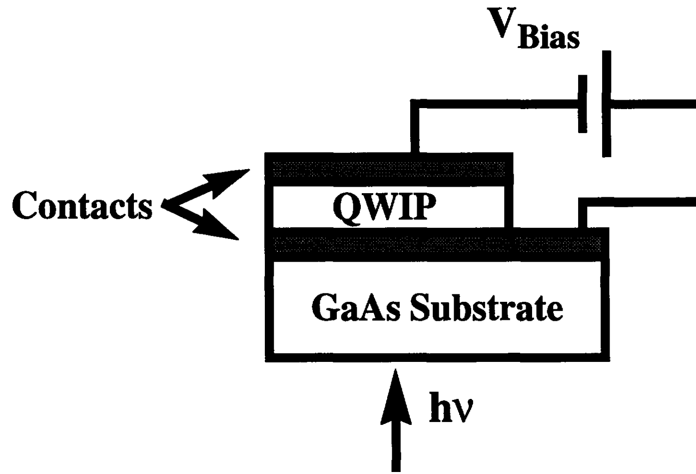


Figure 2.9: Measurement geometry for TE mode QWIP devices

Using the TE absorption coefficient of $200 cm^{-1}$ for a superlattice consisting of a 40Å GaAs quantum well and a 300Å $Al_{0.30}Ga_{0.70}As$ barrier calculated by Flatte et al.[46] and described above in Sections 2.2.1 - 2.2.3 the total absorption in a 50 quantum well structure is:

$$(50x340\text{Å})\,x\left(1\times10^{-8}\frac{cm}{\text{Å}}\right)x\left(200\frac{1}{cm}\right) = 3.4\% \qquad (2.64)$$

Using Equation 1.21 for the minimum resolvable temperature (MRT) in a focal plane array including the effects of non-uniformity, Figure 2.10 shows a comparison of four different detectors in the LWIR band using MRT as the figure of merit. The FPA represented by the solid line has a high quantum efficiency of 70% with FPA non-uniformity after correction of 0.03% typical of research grade MCT arrays. The FPA represented with the dotted line has a quantum efficiency of 20% but a non-uniformity of 0.01%, values typical of research grade TM mode QWIP FPAs.

Note that above a background temperature of -110C the TM QWIP FPA with 20% quantum efficiency but a uniformity of 0.01% has a lower MRT than the MCT FPA with 70% quantum efficiency but a uniformity a factor of 3 worse at 0.03%. The dashed line represents an FPA with only 3.4% quantum efficiency but again 0.01% non-uniformity. This line is meant to represent the TE mode QWIP FPA with quantum efficiency calcu-

lated from Equation 2.64 and non-uniformity equal to that demonstrated in TM QWIP FPAs.

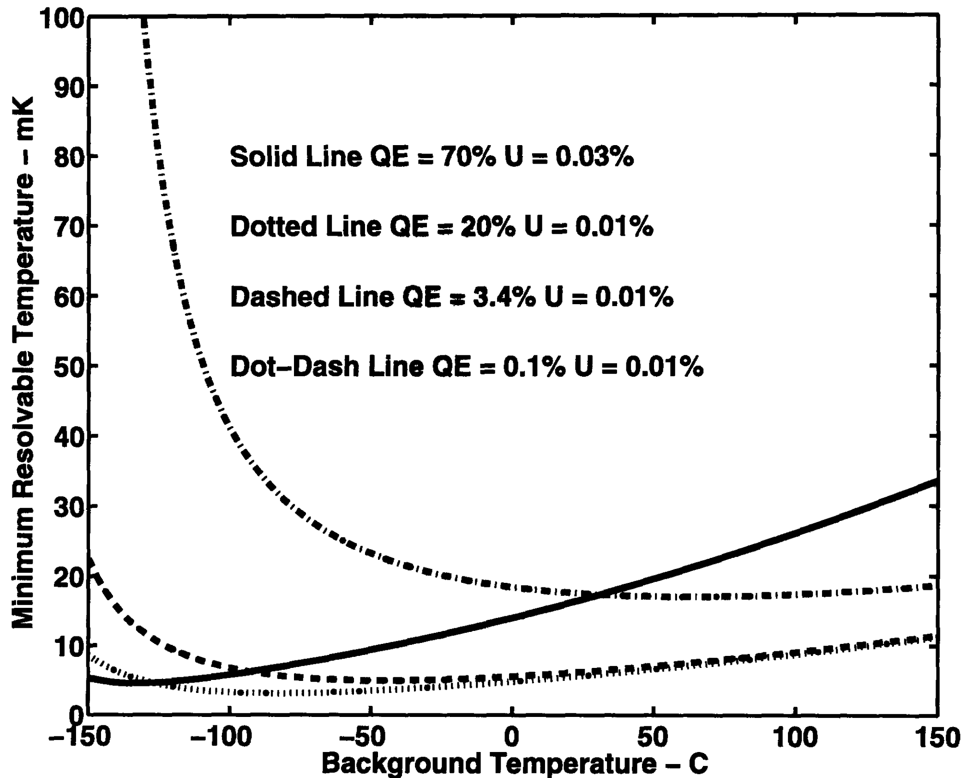


Figure 2.10: Uniformity limit for FPAs in the LWIR, 8μm to 10μm band. Pixel area 50μm x 50μm, frame rate 30Hz, optics f/# 1.5 AR coated, noise floor $200e^-$, g = 0.5.

Above a background temperature of -80C the TE QWIP FPA has a lower MRT than the MCT FPA with 70% quantum efficiency but a non-uniformity a factor of 3 worse at 0.03%. Finally, the FPA represented by the dot-dash line has a quantum efficiency of 0.1% and a non-uniformity of 0.01%, values predicted for Iridium Silicide, a Schottky barrier type detector sensitive to the LWIR band. This FPA has about a factor of three worse MRT at room temperature than the two QWIP based FPAs, but above room temperature it still out performs the MCT FPA in spite of a quantum efficiency nearly three orders of magnitude lower.

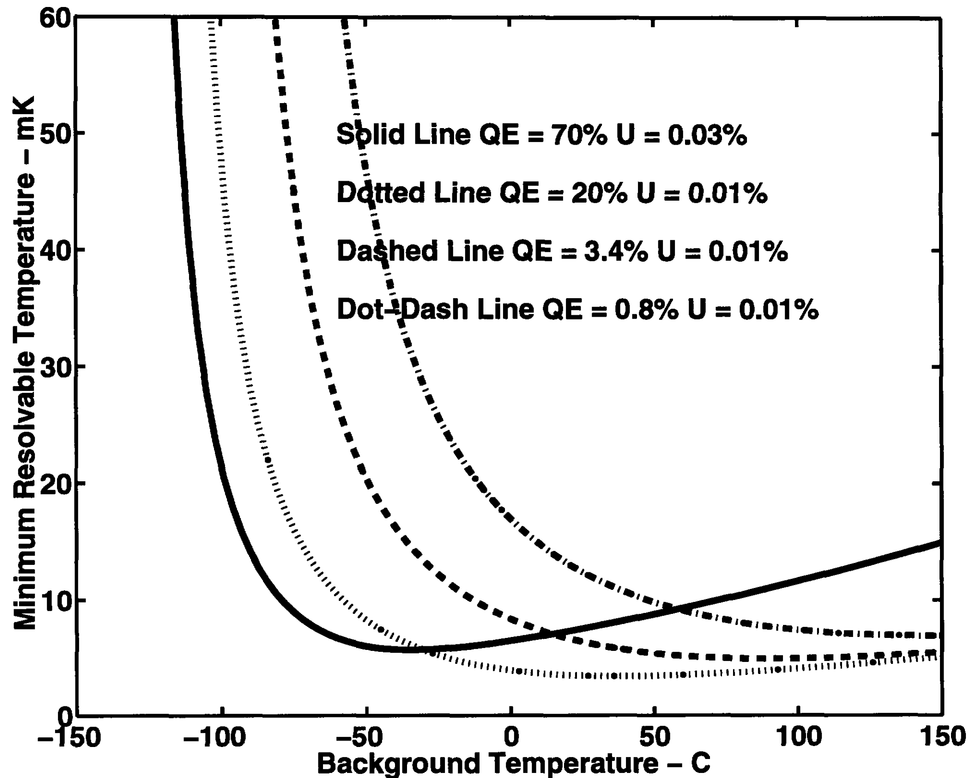


Figure 2.11: Uniformity limit for FPAs in the MWIR, 3μm to 5μm band. Pixel area 50μm x 50μm, frame rate 30Hz, optics f/# 1.5 AR coated, noise floor 200e$^-$, g = 0.5.

The conclusion here is that even though TE mode QWIPs have quantum efficiencies a factor of six lower than TM mode QWIPs; for FPAs operating in the LWIR with background temperatures above about -100C the FPAs are uniformity limited and there is no penalty for using the lower quantum efficiency detectors. In fact the grating used to couple normally incident light into the TM mode QWIP detectors is the largest source of non-uniformity in TM QWIP FPAs[70] and therefore in the region of operation where MRT is proportional to non-uniformity, TE QWIP FPAs with no grating should have better MRT than the TM QWIP FPAs in spite the factor of six lower quantum efficiency of TE QWIPs verses TM QWIPs as calculated in Section 2.2.1. Figure 2.11 goes through the same arguments in the MWIR band where background photon fluxes are smaller, Figure 1.3, and photon flux contrast ratios are higher, Figure 1.4. Figure 2.11 shows a comparison of four different detectors in the MWIR band using MRT as the figure of merit. The FPA repre-

sented by the solid line has a high quantum efficiency of 70% with FPA non-uniformity after correction of 0.03% typical of research grade MCT and InSb arrays. The FPA represented with the dotted line has a quantum efficiency of 20% but a non-uniformity of 0.01%, values typical of research grade TM mode QWIP FPAs. Note that above a background temperature of -30C the TM QWIP FPA with 20% quantum efficiency but a uniformity of 0.01% has a lower MRT than the MCT FPA with 70% quantum efficiency but a uniformity a factor of 3 worse at 0.03%. The dashed line represents an FPA with only 3.4% quantum efficiency but again 0.01% non-uniformity. This line is meant to represent the TE mode QWIP FPA with quantum efficiency calculated from Equation 2.64 and non-uniformity equal to that demonstrated in TM QWIP FPAs. Above a background temperature of 15C the TE QWIP FPA has a lower MRT than the MCT FPA with 70% quantum efficiency but a non-uniformity a factor of 3 worse at 0.03%. Finally, the FPA represented by the dot-dash line has a quantum efficiency of 0.8% and a non-uniformity of 0.01%, values typical for Platinum Silicide, a Schottky barrier type detector sensitive to the MWIR band and described in Section 1.6.2. This FPA has about a factor of three worse MRT at room temperature than the two QWIP based FPAs, but above a background of about 50C it still out performs the MCT FPA in spite of a quantum efficiency two orders of magnitude lower. Again, the point is made that even in the MWIR band where the background flux is lower at a given background temperature and therefore uniformity is relatively less important than it is in the LWIR band, TE QWIP FPAs can produce images of a quality similar to that of MCT and TM QWIP FPAs with all the advantages of inherent in using a GaAs substrate and without the need for using a TM coupling grating.

To complete this section, Figure 2.12 shows a comparison between an FPA with a quantum efficiency of 70% operating in the LWIR band typical of MCT, solid line, and an FPA with a quantum efficiency of only 3.4% operating in the MWIR band, dashed line. Both FPAs have the same non-uniformity of 0.03% for comparison purposes and are operating in the uniformity limited regime, but due to the improved photon flux contrast in the MWIR band, Figure 1.4, the FPA operating in the MWIR has a better MRT for background temperatures above about -20C.

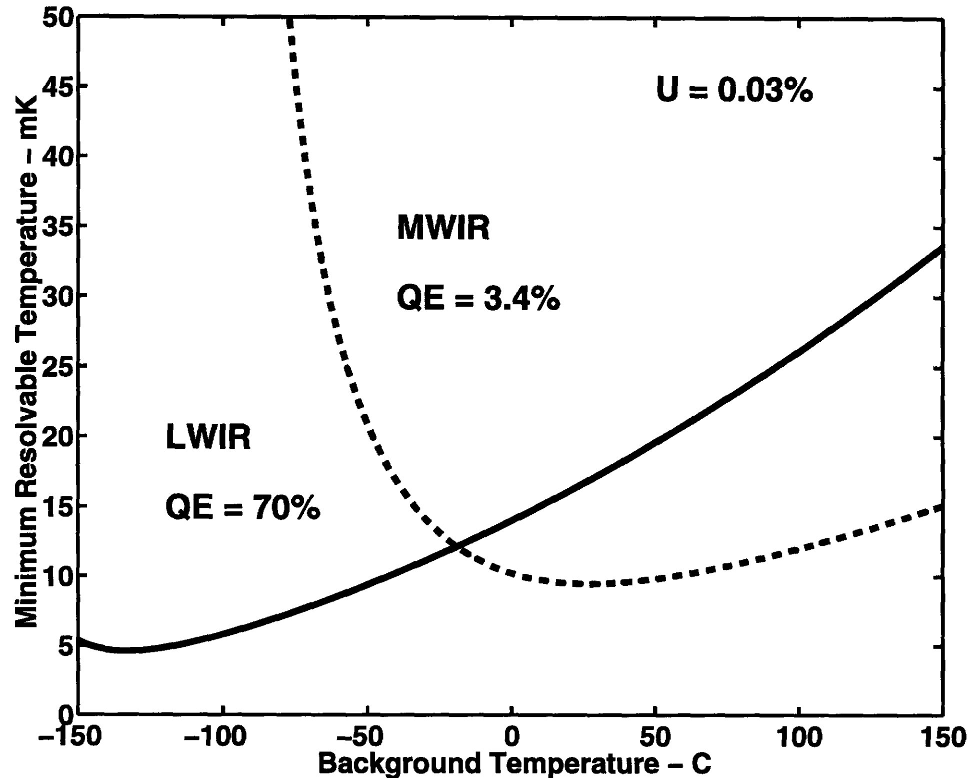


Figure 2.12: Comparison of MWIR, 3μm to 5μm, and LWIR 8μm to 10μm bands. Uniformity 0.03%, pixel area 50μm x 50μm, frame rate 30Hz, optics f/# 1.5 AR coated, noise floor $200e^-$, g = 0.5.

In conclusion, even for the rather unfavorable case of TE absorption in square GaAs/AlGaAs quantum wells, the small absorption coefficient can be combined with improvements in uniformity relative to both direct band systems like MCT FPAs and grating coupled TM QWIP FPAs to produce TE QWIP FPAs with excellent image quality. This is done in a structure that is extremely simple with no need for a coupling grating and therefore lends itself well to manufacturing.

2.2.3 Conclusions and avenues to improvement of TE mode absorption

So far in this section, the TE mode absorption coefficient and its application to TE mode QWIP FPAs have been discussed in the context of ideal square quantum wells of GaAs with lattice matched barriers of AlGaAs. In these structures however, non-ideal quantum wells may offer another avenue to increasing the TE absorption. As an example, Tidrow et al. have seen large TE/TM absorption ratios in systems having large applied

fields along the stacking direction for the superlattice.[71] This is particularly interesting because QWIP detectors operate with rather large bias fields of $few \times 10^4 \ V/cm$. In fact, Young et al.[48] have shown that fields of this order in GaAs/AlGaAs quantum wells can account for the appearance of normally forbidden transitions in the exciton absorption spectrum and similar effects should occur in the intersubband absorption spectra as well. The addition of strain in the quantum well structure by adding Indium to the GaAs quantum well could change the TE absorption coefficients in two ways. First, the reduction in symmetry of the bulk band structure will alter the momentum matrix elements coupling the various bands and second, the well known segregation of Indium along the growth direction results in an asymmetric quantum well similar in effect to the addition of a large built in field.[49] These effects have not been systematically studied here, but qualitatively they all can lead to a larger TE absorption coefficient and so the calculations done above in Section 2.2.2 can be considered lower limits for TE mode QWIP FPAs.

In conclusion, this section has for the first time calculated the potential for using TE mode QWIPs in focal plane array cameras in terms of the ultimate FPA performance using minimum resolvable temperature (MRT) as the figure of merit. The TE mode QWIP absorbances were calculated by Flatte et al.[46] for an ideal superlattice consisting of 50 periods of a 40Å GaAs quantum well and a 300Å $Al_{0.30}GaAs$ barrier in a 14-band superlattice $\boldsymbol{K \cdot p}$ theory including the effects of the antibonding $\bar{p}$ conduction band states. Including the effects of non-uniformity, it has been shown that in the LWIR band, TE QWIP FPAs should produce excellent quality images with MRTs lower than those for MCT FPAs for background temperatures above -100C. In the MWIR band TE QWIP FPAs should produce excellent quality images with MRTs lower than those for MCT FPAs for background temperatures above 15C.

2.3 Grating coupled TM mode miniband QWIP FPAs

All of the high performance QWIP FPAs demonstrated to date have used diffraction gratings to couple normally incident light, transverse electric (TE), into the high absorption coefficient transverse magnetic (TM) mode of n-type QWIPs.[72] This section will dis-

cuss one particular grating coupled QWIP design developed by Lockheed/Martin which in addition to a coupling grating uses a superlattice barrier design to reduce the dark current. This miniband transport type QWIP is an excellent example of the potential for "bandgap engineering" in III-V semiconductor superlattices which does not exist in bulk photodetectors such as those based on MCT.

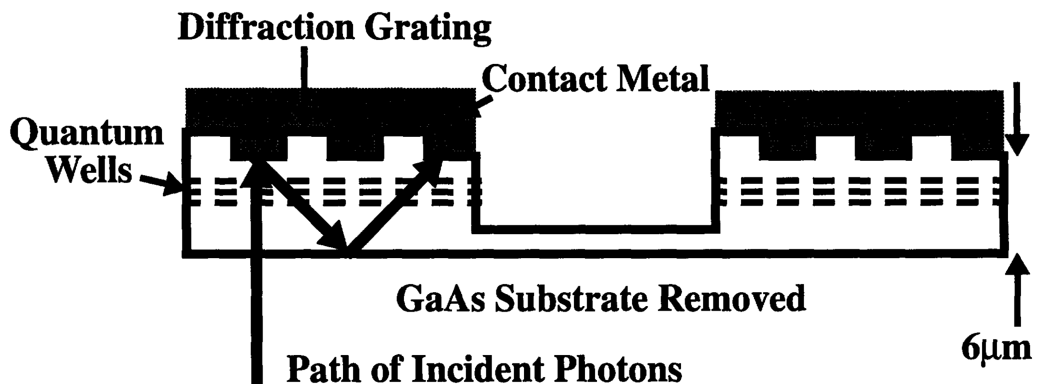


Figure 2.13: Schematic for QWIP FPA with 1-D diffraction grating.[73]

The gratings in this design are etched into the top side of the QWIP and metalized both to provide an ohmic contact for biasing the device and to improve the efficiency of the diffraction grating. The IR light is then incident from the substrate side of the device and is absorbed after diffraction from the grating. One dimensional gratings of the type shown in Figure 2.13 couple to only one of the two orthogonal TE polarizations and so have a peak theoretical coupling approaching 50%. Two dimensional gratings have also been used with similar geometries to allow coupling to both TE polarizations, but these require more complex lithography.[58]

The best FPA performance for any large format FPA has recently been demonstrated by Lockheed/Martin using the pixel structure shown in Figure 2.13 with a one dimensional grating.[73,74] The substrate has been removed after hybridization to the Silicon read-out integrated circuit using an epitaxial lift-off technique.[75] This eliminates cross-talk between pixels due to reflections at the substrate/air interface and also improves the

reliability of the hybridization process as will be discussed in Chapter 6. The quantum well active region uses a bound to miniband transition rather than the bound to continuum transition depicted in Figure 2.1 and used in the majority of QWIP FPAs.[39] In this structure, the barrier region between the quantum wells is not simply a bulk wide bandgap semiconductor like AlGaAs, but rather is a short period superlattice consisting of thin alternating layers of wide bandgap and narrow bandgap semiconductors as show in Figure 2.14. Transport of photoexcited carriers and dark current occurs in the miniband formed by coupling between overlapping quantum mechanical wavefunctions in the narrow quantum wells which make up the barrier.

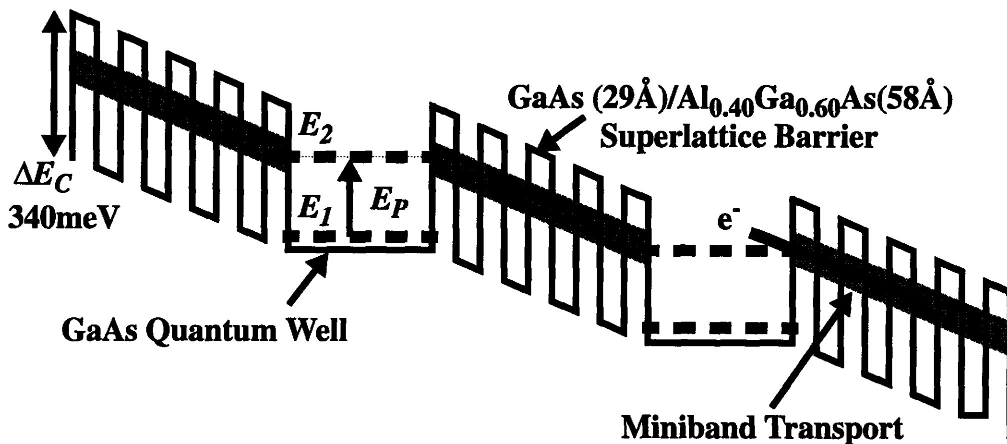


Figure 2.14: Schematic of miniband transport QWIP detector[74]

The advantage of the miniband transport comes from decoupling the conduction band offset, ΔE_C, and the photon energy, E_P. In the bound to continuum type QWIPs of Figure 2.1, the photon energy is determined from the equation:

$$E_P = \Delta E_C - E_1 \tag{2.65}$$

which along with the doping density determine the thermionic emission as shown in Equation 2.1. For the miniband transport QWIP the coupling between the photon energy and the conduction band offset is broken; instead, the photon energy is determined by the difference in energy between the ground state in the quantum well and the miniband energy labeled E_2 in Figure 2.14. This allows a large reduction in thermionic emission

which now occurs for transitions between the heavily populated ground state in the quantum well and the three dimensional continuum which has been lifted up in energy due to the increase in the conduction band offset.

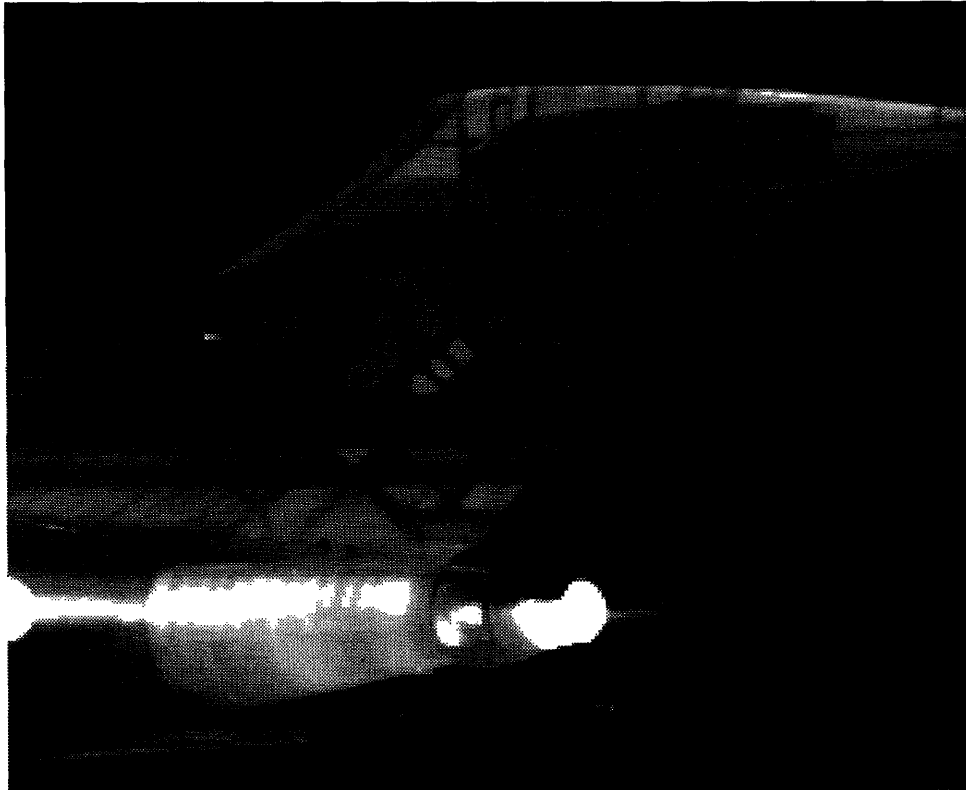


Figure 2.15: 640 x 480 grating coupled QWIP miniband transport FPA operating in the LWIR. 50µm pixel pitch, f/2 optics, 30 frames/second and MRT = 7mK. Subject is Mexicana Airlines Lockheed L1011. Courtesy Charles Parton (Lockheed/Martin)

The decrease in dark current can be easily calculated from Equation 2.1. Comparing a bound to continuum QWIP with a 10µm cutoff wavelength, E_T = 120meV, and the miniband transport QWIP of Figure 2.14 with E_T = 300meV, Equation 2.1 gives a twelve fold decrease in dark current. The decrease in dark current also allows an increase in T_{BLIP} with the result that the Lockheed/Martin miniband transport QWIP FPA has a T_{BLIP} of 85K for a cutoff wavelength of 10µm.

Figure 2.16: 640 x 480 grating coupled QWIP miniband transport FPA operating in the LWIR. 50μm pixel pitch, f/2 optics, 30 frames/second and MRT = 7mK. Image taken in dark room with no illumination. Courtesy Charles Parton (Lockheed/Martin)

Figure 2.15 shows a single frame from a 640x480 Lockheed/Martin miniband transport QWIP FPA. The minimum resolvable temperature for this FPA is 7mK limited by the uniformity of the array and not by the quantum efficiency which is 26%. The temperature resolution is clearly spectacular showing details of the support frame in the tail of the Lockheed L1011 airliner as well as variations in emissivity caused by the painted markings which identify the airplane as belonging to Mexicana Airlines. Equally important for image quality is the lack of blooming in the very hot jet exhaust region which saturates the detector but does not adversely effect the resolution in neighboring regions. Figure 2.16 shows a single frame from the same QWIP FPA taken with a 300K black background in a dark room. Again both the temperature and spatial resolutions are excellent allowing for clear identification of an individual with no illumination.

In conclusion this section has presented results from a state-of-the-art grating coupled QWIP miniband transport FPA. The 7mK minimum resolvable temperature is seen to produce excellent image quality and it is this standard against which the TE mode grating free QWIPs which are the subject of this thesis will be measured. In addition to image quality, this FPA exemplifies two other properties relevant to this thesis. First, this miniband transport type QWIP is an excellent example of the potential for "bandgap engineering" in III-V semiconductor superlattices which does not exist in bulk photodetectors such as those based on MCT. This and other applications of bandgap engineering will likely lead to improvements in TE mode QWIP devices as well. Finally, the images shown here with a 300K background are limited in MRT by the spatial nonuniformity and not by the individual pixel performance. As discussed in Chapter 1 this implies that a reduction in quantum efficiency which results in a reduction in pixel performance can be traded off for another desirable quantity with no deterioration in overall FPA performance as measured by MRT. The TE mode QWIP proposed to take advantage of this excess quantum efficiency to eliminate the coupling grating, which is one of the most difficult steps in FPA manufacture, thus lowering the quantum efficiency but also dramatically simplifying the manufacture of the FPA with no performance penalty for uniformity limited operation.

2.4 Alternative normal incidence QWIP technologies

In addition to the widely used grating coupled TM mode QWIPs and newly developed TE mode QWIPs there are several other III-V based QWIP structures that lead have potential utility in QWIP FPAs. The alternative structures all attempt to eliminate the need for a coupling grating while retaining a high quantum efficiency by using an alternative to the n-type quantum well grown on <100> oriented substrates. Two of these alternative normal incidence QWIP technologies will be discussed in this section with a short description of the physical principles and experimental performance data and a discussion of issues related to the integration. These particular structures are chosen because they best exemplify the potential for bandgap engineering and growth manipulation inherent in QWIP structures, but lacking in the other detector technologies discussed in Section 1.6;

Mercury Cadmium Telluride, Schottky barrier detectors and bolometers. Section 2.4.1 discusses p-type QWIPs which use the complexities of the valence band in III-V semiconductors to achieve normal incidence absorption without a grating and Section 2.4.2 discusses AlAs/AlGaAs X-point QWIPs grown on <110> oriented substrates which use the complexities of the conduction band at the X-point to achieve normal incidence absorption without a grating. Finally, Section 2.4.3 presents concluding remarks on the promise of these alternative normal incidence QWIP Technologies.

2.4.1 P-type normal incidence QWIPs

P-type QWIPs were first demonstrated at AT&T where much of the early experimental work on QWIPs was done.[76] The first devices consisted of 50 periods of a superlattice with 40Å GaAs quantum wells and 300Å $Al_{0.30}Ga_{0.70}As$ barriers grown on <100> GaAs substrate and are shown schematically in Figure 2.17. The quantum wells were doped p-type at $4x10^{18}$ cm^{-3} with Beryllium giving a cutoff wavelength of 7.9μm. The strong mixing in the valence band between the light and heavy holes away from $k=0$ creates a large absorption coefficient for TE polarized incident light between the occupied heavy hole bound state and the unoccupied light hole quasibound state with no need for a diffraction grating.

The quantum efficiency per well was η_1=0.65% for a total quantum efficiency for all 50 wells of η=28%, comparable to the values for n-type grating coupled TM QWIPs and an order of magnitude larger than TE mode QWIPs. The photoconductive gain was g=0.034, an order of magnitude smaller than the photoconductive gain for 50 well n-type QWIPs due to the smaller hole mobility in the AlGaAs barriers and larger hole mass which leads to a much larger excited carrier capture probability, p=37%, compared to p=7% found in the n-type QWIPs. Taken together, these parameters explain the low value of the responsivity measured at 39mA/W. Detectivity for these devices was $3.1x10^{10}$ Jones in spite of the low responsivity because the small gain effects both responsivity and noise.[76]

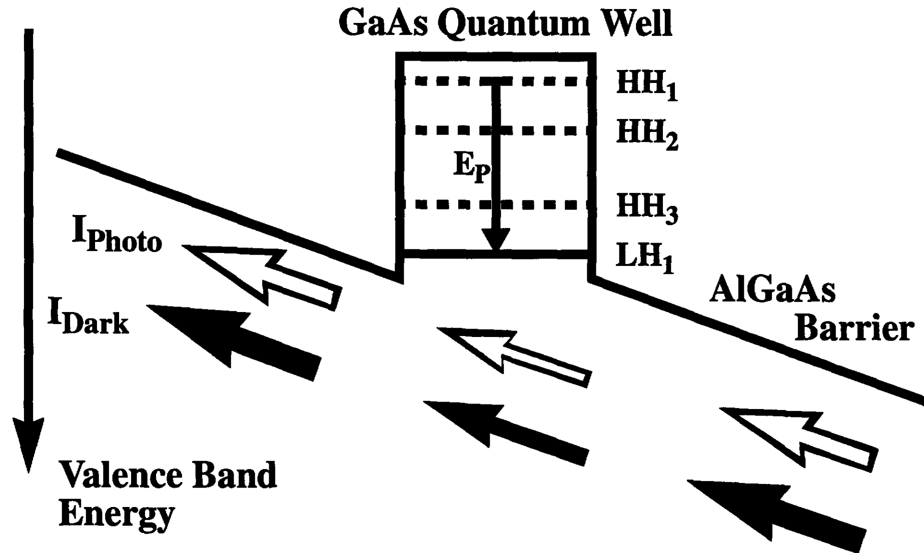


Figure 2.17: Schematic of band diagram for p-type GaAs/AlGaAs QWIP

One of the important advantages of QWIPs comes from the ability to engineer the band structure through the growth of superlattices. This means that unlike the case for bulk semiconductors, the band structure is not a fixed material parameter, but rather can be modified within limits through the use of strain, variation in material composition and variation in layer thickness. Bandgap engineering is particularly powerful for p-type devices because of the complexity of the valence band in III-V semiconductors. An example is the p-type strained layer $In_{0.30}Ga_{0.70}/In_{0.52}Al_{0.48}As$ QWIP grown on <100> InP shown in Figure 2.18.[77]

The tensile strain in the quantum well causes the normally degenerate Γ-point light hole (LH) and heavy hole (HH) bands to split forcing the light hole band down in energy relative to the heavy hole band. For appropriate choice of strain and quantum well width the light hole states become the occupied states in a p-doped quantum well and the heavy hole states are left unoccupied.[78] The optical transition occurs between the occupied $\langle J|m_J\rangle = \langle\frac{3}{2}|\frac{1}{2}\rangle$ light hole states and the unoccupied $\langle\frac{3}{2}|\frac{3}{2}\rangle$ heavy hole continuum states above the barriers. Since the light hole has a small effective mass perpendicular to the

superlattice growth direction and a large in-plane density of states, the optical absorption and responsivity can be significantly increased.

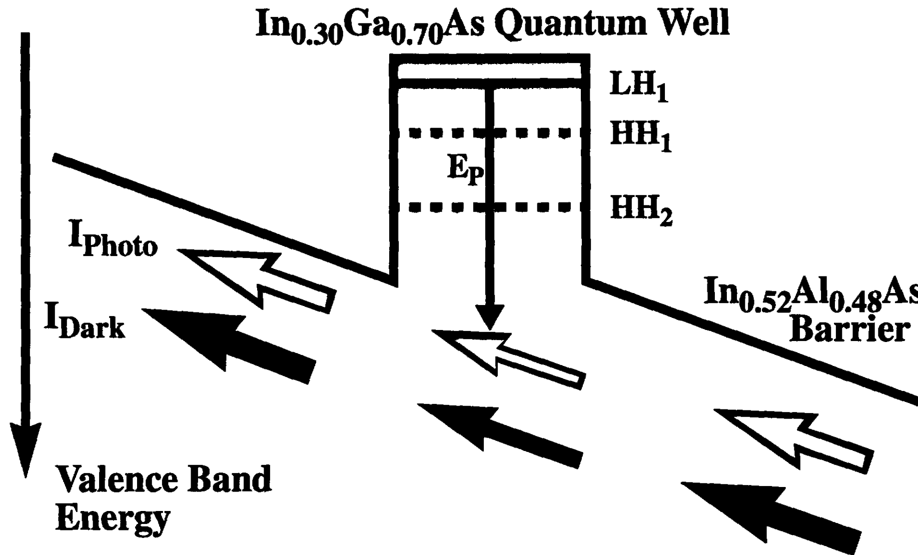


Figure 2.18: Schematic of band diagram for p-type $In_{0.30}Ga_{0.70}As$/ $In_{0.52}Al_{0.48}As$ strained layer QWIP grown on InP.[77]

The device of Figure 2.18 consists of twenty 40Å quantum wells of $In_{0.30}Ga_{0.70}As$ doped with Beryllium at $1x10^{18}$ cm^{-3} and separated by 450Å of $In_{0.52}Al_{0.48}As$ lattice matched to the InP substrate. The quantum efficiency per well was 0.5% for a total quantum efficiency of 10% assuming a capture probability of 37% and a resulting photoconductive gain from Equation 2.9 of g=0.085. The responsivity was 51mA/W for a cutoff wavelength of 8.8μm at 77K resulting in a BLIP detectivity at 77K of $5.9x10^{10}$ Jones. T_{BLIP} was not measured, but the dark current at 77K was more than an order of magnitude smaller than the photocurrent which at 3K per octave for thermionic emission gives T_{BLIP} of 87K. This is excellent performance comparable to the best TM mode grating coupled miniband transport QWIPs discussed in Section 2.3.

2.4.2 X-band AlAs/AlGaAs QWIPs

All of the QWIP devices discussed to this point have been grown on <100> oriented substrates and have used the Γ-point band structure for confining and transporting carriers. This is quite common in all III-V semiconductor devices because the energy minimum for

both holes and electrons in GaAs occurs at the Γ-point. However, for $Al_xGa_{1-x}As$ with the Aluminum fraction x greater than 40% the conduction band minimum moves to the X-point and the semiconductor becomes indirect. Figure 2.19 shows an n-type QWIP which takes advantage of the lower symmetry of the X-point to achieve normal incidence detection with no grating in an n-type device.[77]

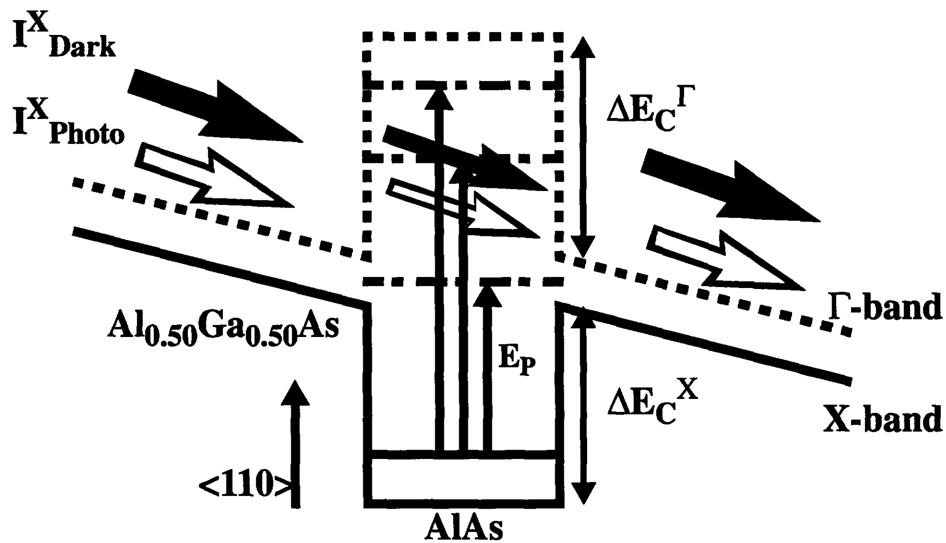


Figure 2.19: Schematic of band diagram for X-band $AlAs/Al_{0.50}Ga_{0.50}As$ QWIP. As shown this is a multiwavelength device with responsivity at 2.5μm, 4.3μm and 14.8μm. Resonances are show with dot-dash lines.[77,79]

In order to form the X-point quantum wells, a twenty period superlattice of 30Å of AlAs and 500Å of $Al_{0.50}Ga_{0.50}As$ was grown on a <110> oriented GaAs wafer. The AlAs layers which form the quantum wells at the X-point were doped $2x10^{18}$ with Silicon. Because of multiple resonances above the edge of the quantum well these devices actually have multiwavelength detection capability with responsivities of 110A/W, 18.3A/W and 24mA/W at cutoff wavelengths of 2.5μm, 4.3μm and 14.8μm. The huge responsivities at 2.5μm and 4.3μm are due to photoconductive gains calculated to be 630 and 3200 respectively caused by resonant transfer of photoexcited electrons from the X-band into the Γ-band where they can not be captured by the X-band quantum well. Detectivities at the

three bands were 1.1×10^{12} Jones, 3.0×10^{11} Jones and 1.1×10^{9} Jones at 77K. BLIP temperature was not measured, but was less than 77K due to a large dark current.

The very large responsivities and low BLIP temperatures of the X-band QWIPs discussed here are not amenable to use in FPAs which require the charge generated to be stored in a small capacitor, however the multiwavelength capability is potentially useful and with proper choice of materials and superlattice parameters the resonances which led to the very high photoconductive gains could possibly be suppressed.

2.4.3 Concluding remarks on alternative normal incidence QWIPs

Fabrication of the coupling gratings used in high performance large format QWIP FPAs constitutes a large part of the FPA chip complexity and cost. In addition, the coupling grating is often times the single largest contributor to nonuniformity in the array. [80] This section has discussed several alternatives for normal incidence QWIPs that do not require coupling gratings. The GaAs based p-type QWIPs, InP based p-type strained layer QWIPs and the n-type X-band QWIPs grown on <110> GaAs have all been shown to have detectivities well in excess of the 1×10^{10} Jones discussed in Chapter 1 as being the threshold for uniformity limited performance. Hybrid integration of these devices using Indium bump bonds as will be discussed in Chapter 6 for TM grating coupled QWIPs should work well for any of these devices however, if monolithic integration is the goal these devices all have difficulties. In particular for the InP based p-type QWIPs one would need to develop high density, high performance InP based electronic devices to integrate with the InP based QWIPs. Though low density integration has been demonstrated on InP for microwave applications the VLSI density necessary for QWIP FPAs has yet to be demonstrated. For the n-type QWIPs based on <110> oriented substrates the low mobility in the X-band leaves no incentive to develop electronic devices on <110> substrates leaving hybridization as the only alternative. The p-type <100> GaAs QWIP could possibly be integrated with p-HEMT devices now being developed for cellular phones, but this technology offers no obvious choice for the integration capacitor and also no VLSI density circuits. Only the TE mode n-type <100> GaAs based QWIPs which are the subject of this thesis have the potential for monolithic integration with already existing VLSI density

MESFET circuits. Issues related to integration of these devices will be discussed in Chapter 6.

2.5 Concluding remarks on normal incidence QWIPs

This chapter has developed the theoretical basis for design of normal incidence QWIPs. The results from the ***K•p*** calculations of Flatte et al. have been used to calculate the quantum efficiency for a prototypical TE mode QWIP. Using this absorption coefficient and the FPA performance model developed in Chapter 1 it has been found that minimum resolvable temperatures, (MRT) of less than 10mK are possible for TE mode FPAs operating under uniformity limited conditions with 300K backgrounds in the LWIR. This is the first time that calculated TE quantum efficiencies have been used to demonstrate that FPAs based on TE mode QWIPs can compete with the more standard TM type devices under conditions in which the 300K background causes FPAs based on both types of devices to be uniformity limited. In order to understand the current state of QWIP FPAs, Section 2.3 discussed the use of diffraction gratings to couple normally incident light into miniband transport type TM mode FPAs developed by Lockheed/Martin. These FPAs represent the state-of-the-art in FPA performance with 640x480 format arrays demonstrating MRTs of less than 10mK. The largest format MCT based FPA have similar MRTs, but with a format of only 256x256, nearly five times fewer devices. PtSi arrays as large as 1024x1024 have been fabricated, but with MRTs more than an order of magnitude larger. The potential for TE mode FPAs to compete with these TM mode FPAs demonstrated here bodes well for their acceptance in marketplace. Subsequent chapters will discuss the experimental work on TE mode QWIPs, including fabrication of devices based on the analysis presented in this chapter and their characterization.

Chapter 3

Growth and Fabrication of Normal Incidence QWIPs

3.1 Introduction

This chapter discusses the techniques used for fabrication of normal incidence QWIPs in the GaInAlAs material system. Where the technique used is novel and/or the results are unusually enlightening some extra time will be spent to expand the explanation. This is particularly true for Section 3.4 on strained layer heteroepitaxy and critical layer thickness. Section 3.4 has particular relevance to the fabrication of QWIP devices because as derived in Section 2.1 a good QWIP has "as many quantum wells as practical". Section 3.4 discusses the issues related to determining how many quantum wells are possible when working with strained layer superlattices. The other two material sections, Section 3.2 on molecular beam epitaxy (MBE) and Section 3.3 on material characterization very briefly describe the techniques used and present some of the relevant results. Their brevity is a reflection of one of the great advantages present in QWIPs verses other FPA technologies, namely that high quality III-V semiconductors have been grown by MBE for more than twenty years. This has resulted in a very large established knowledge base that can be leveraged to produce and characterize high quality material in a relatively short period of time. In addition, high quality GaAs substrates are readily available in sizes up to 6" diameter from a variety of commercial vendors and a wealth of experimental data has been published with which to compare characterization results. Combine this with a material, GaAs, that is fundamentally easy to work with and has none of the subtleties inherent in II-VI materials and it is no surprise that in the ten years since the first QWIP was demonstrated QWIP FPAs have surpassed the performance of MCT based FPAs in spite of forty years of work and a much larger investment in MCT. Finally, Section 3.4 describes the actual fabrication sequence used for the devices presented in this thesis and Section 3.5 concludes with a discussion of future improvements in QWIP fabrication.

3.2 Molecular Beam Epitaxy Growth

Samples for this thesis were grown on a Riber 2300 three chamber MBE system. Details of the techniques used for substrate preparation and growth are quite standard and are given in the MIT MBE manual[81] so only a couple of key points will be pursued here. In particular, determination of the temperature of the substrate during MBE growth is critical and must be consistent across many different growths. An optical pyrometer with a bandpass filter centered at 2.3μm was used to measure relative changes in temperature above 300C. The critical step is the calibration for the emissivity of the sample. This was done using a phase transition in the surface of GaAs which is visible using insitu reflection high energy electron diffraction (RHEED). With the Arsenic shutter closed and the beam equivalent pressure of Arsenic flux incident on the sample less than 10^{-7} torr, two orders of magnitude lower than the typical Arsenic flux during growth, the GaAs surface undergoes a phase transition from an Arsenic stable 2x RHEED pattern below 640C to a Gallium stable 4x RHEED pattern above 640C. Watching for this phase transition while slowly ramping up the substrate temperature can consistently determine the temperature 640C which is then used to calibrate the pyrometer. Since this temperature is easily repeatable, a constant of the material and a direct measurement of the sample being grown, most of the experimental factors are eliminated and consistent results can be obtained over long periods of time. Note that for the particular pyrometer being used, the GaAs substrate is transparent to the wavelength of the pyrometer 2.3μm. This means that for samples mounted with Indium the pyrometer is actually looking at the metallic Indium on the backside of the sample. In fact this is nearly ideal since the Indium is in intimate contact, in fact alloyed into, the GaAs substrate and the emissivity of the sample in this temperature range is nearly independent of the materials being grown all of which have bandgaps larger than 2.3μm for the InGaAlAs material system used in this thesis. The efficacy of this measurement can be seen insitu under the unfortunate circumstances when the Indium on the back side of the sample de-wets the Molybdenum sample mounting block. In this case the pyrometer, which is focused to a spot about 5mm in diameter, will read large variations in temperature as it move from areas where the Indium has de-wetted

from the Moly block thus reducing heat conduction to the GaAs from the hot Moly block and areas where the Indium still makes good contact between the Moly block and the GaAs sample. Needless to say, samples where this effect is severe are trash.

In addition to good control of temperature, the QWIPs grown in this thesis consisted of multiple quantum wells which require good control of the layer thicknesses to get maximum effect. Good control in this context means average variations in quantum well width of one monolayer or less which at a typical growth rate of one micron per hour or one monolayer per second translates into time variations for shutter operations of one second or less. This was achieved using a computer system and software designed by a previous graduate student, J. Vlcek.[82,83] The software, code named "Molly", allows simple control over both shutter times and effusion cell temperatures. Significant improvements in the reliability of the hardware system made by the author and the use of the Molly software have enabled students to grow as many as fifty quantum wells with the high degree of repeatability required to achieve optimum results.

3.3 Material Characterization

Material characterization done on unprocessed material is an extremely valuable tool for determining the potential for a particular chip prior to investing a large effort into the fabrication and testing of individual devices. Two of the most valuable non-destructive characterization tools, double crystal x-ray and photoluminescence will be discussed in this section with examples relevant to QWIP device fabrication.

3.3.1 Double Crystal X-ray Diffraction

All X-ray sources produce radiation with a finite dispersion in both angle and energy. Energy dispersion is determined by the atomic transitions in the target, typically $K\alpha_1$ in Copper. The angular dispersion is determined by the geometry of the target and collimating slits and for practical systems is limited to a minimum of about one-tenth of a degree. Double crystal X-ray diffraction improves this resolution more than two orders of magnitude and removes the energy dispersion by reflection of the X-ray beam emitted from the collimator from a high purity single crystal. For the work in this thesis in which GaAs

based devices were analyzed, the first crystal was chosen to be a GaAs substrate which lead to a minimum resolution of ~10 arc seconds.[84] A schematic of the DCXD system is shown in Figure 3.1 along with a selection of relevant angles for GaAs.

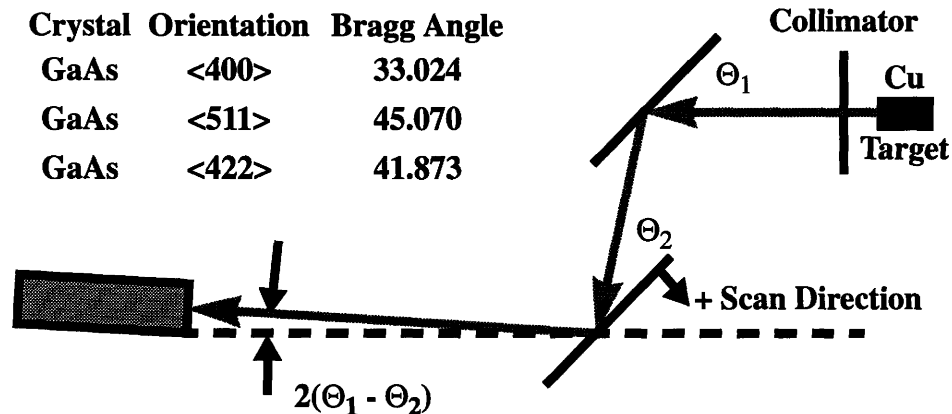


Figure 3.1: Schematic of DCXD system. Inset shows Bragg angles for GaAs at several common crystal orientations.

The additional resolution afforded by the DCXD allows accurate determination of both layer thickness and material composition for QWIP samples in a completely non-destructive manner. The results for QWIP 9065 are shown below in Figure 3.2 in which the measured DCXD rocking curve is shown on top and the simulated rocking curve calculated using Bede Corporations RADS program[85] is shown below. Superlattice parameters are determined by varying the inputs to the simulator and finding the best fit to the measured rocking curve. The periodicity of the large peaks determines accurately the periodicity of the crystal superlattice which in this case is determined to be 555Å +/- 5Å. The position of the fringes relative to the large substrate peak determines the Indium composition in the quantum wells, 13% +/- 2% and the splitting in the large peak determines the Aluminum fraction in the barriers, 30 +/- 5%. The Aluminum fraction in the barriers is the least accurate number because the lattice constant of AlAs and GaAs are so close that relatively large changes in Aluminum fraction are required to move the Bragg angle for the AlGaAs material out from under the GaAs substrate peak. DCXD scans similar to Figure 3.2 are quite valuable, but they do require significant resources since the signal to

noise ratio improves with the square root of the number of counts. To bring out the small fringes requires count times of several minutes per angle point totaling 8-12 hours for a single run not including time for setup and allignment.

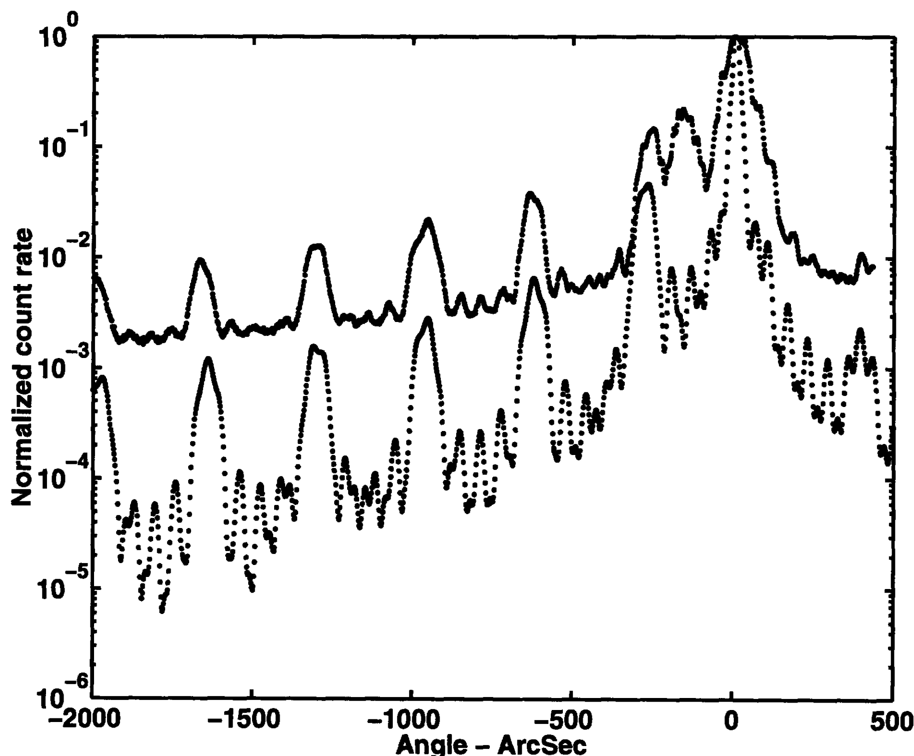


Figure 3.2: DCXD rocking curve and simulation for QWIP 9065. Quantum well width 45Å $In_{0.13}Ga_{0.87}As$, barrier width 510Å $Al_{0.30}Ga_{0.70}As$.

3.3.2 Photoluminescence

Photoluminescence, PL, is a non-destructive technique that can be very powerful in determining the optical properties of quantum well structures. In particular, because the interband relaxation time is nearly three orders of magnitude slower than the intraband relaxation, carrier excited high into the conduction band by absorption of a visible photon, as from an Argon laser, quickly relax to the lowest levels in the conduction band by phonon emission and then emit photons to relax back to the valence band. The spectral purity of the emitted photons as measured by a grating spectrometer system similar to that shown in Figure 3.3 gives information about the starting and ending energy levels of the

relaxing electron. Particularly when done at low temperature to reduce thermal broadening of the peaks and dissociation of weakly bound particles, this technique gives sensitive information on the existence of cooperative quasi-particles like excitons and defect levels in the quantum wells. Some of these defect levels such as those associated with donors in n-type quantum wells or acceptors in p-type material are inevitable, but others such as those associated with quantum well interfaces and excitonic binding are indicative of the optical properties of the quantum well.

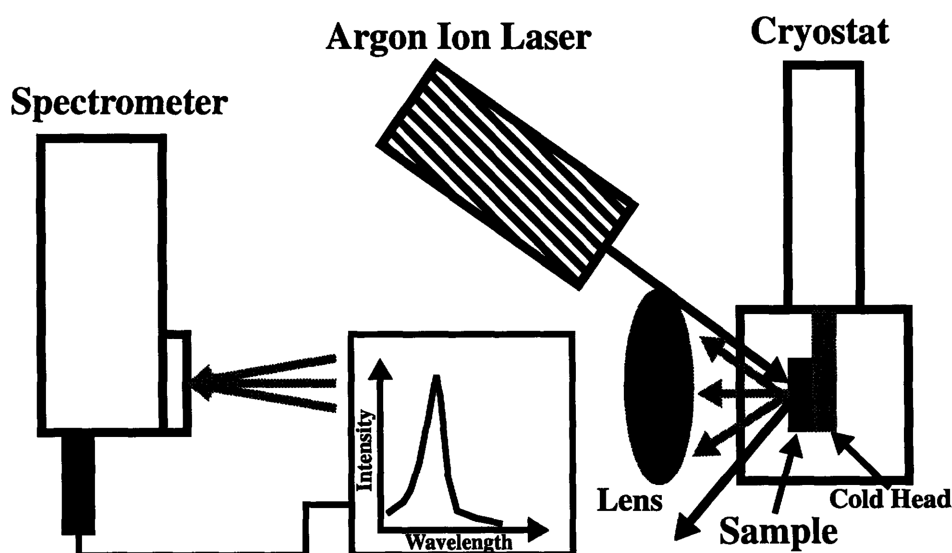


Figure 3.3: Schematic representation of photoluminescence experiment

Qualitative information can also be gleaned from PL as shown in Figure 3.4 where a comparison of two simultaneously grown multi-quantum well structures is shown for PL taken at 25K. The circles are data from a multi-quantum well sample consisting of ten 50Å GaAs quantum wells separated by 500Å $Al_{0.30}Ga_{0.70}As$ barriers grown on a bulk n-type substrate with a growth temperature of 470C. The dots are data taken from the same multi-quantum well structure grown simultaneously in the dielectric growth well of a VLSI MESFET chip as described in Chapter 6 for the epi-on-electronics, E-o-E, optoelectronic integration technique. The data have been normalized to a peak intensity of one to correct for variations in alignment and incident laser power. The growth temperature was

chosen to be compatible with the MESFET circuitry.[1,86] The native oxide on the GaAs was removed using hydrofluoric acid which passivates the surface with Hydrogen. The Hydrogen is then desorbed in the MBE system by heating the substrate up to the 470C growth temperature. The similarity of the two peaks demonstrates both the viability of the Hydrogen passivation technique and the quality of the substrate surface underneath the electronic circuits. However, the absolute width of the PL peak is nearly 45meV, an order of magnitude larger than that obtained from similar samples grown at the more typical growth temperature of 630C. Possible causes include interface roughness[87] and oxygen incorporation.[88] The specific causes for this increased PL width have not been determined, but this result clearly indicates both the potential of the E-o-E technique and the difficulties associated with growing at low temperatures.

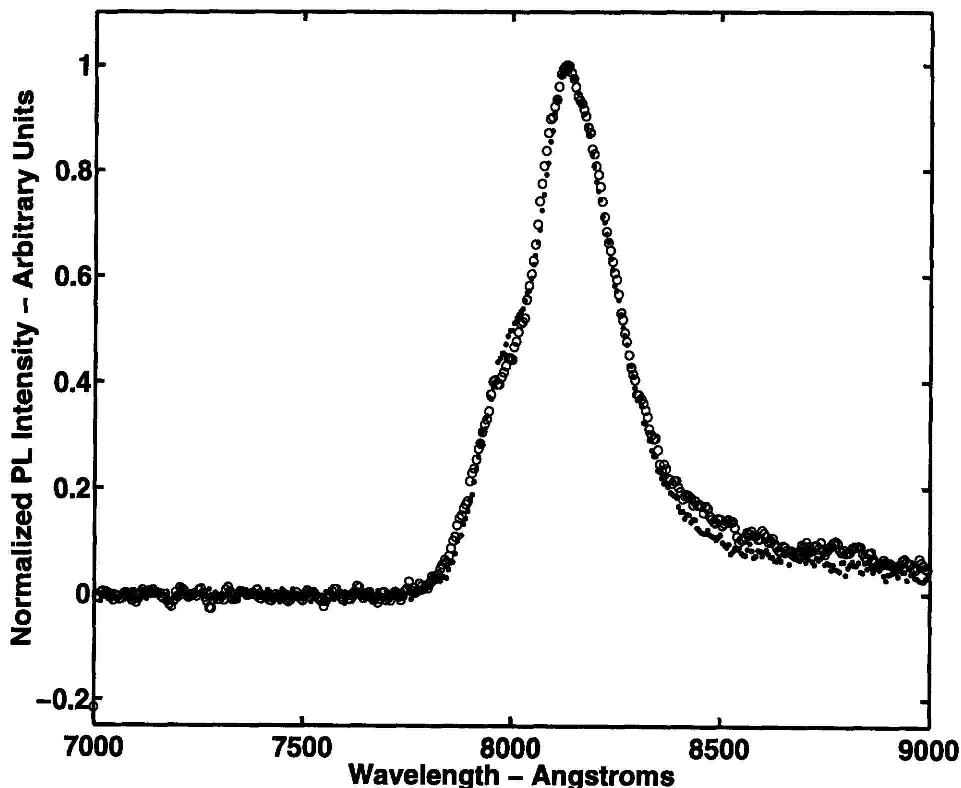


Figure 3.4: Comparison of 25K PL intensity for MQW structure containing 10 50Å quantum wells of GaAs separated by 500Å $Al_{0.30}Ga_{0.70}As$ barriers. Circles are for sample grown on bulk GaAs substrate and dots are for sample grown in E-o-E dielectric growth well. Both samples grown simultaneously at 470C and prepared using hydrogen passivation.

3.4 Strained layer heteroepitaxy and critical layer thickness

This section has particular relevance to the fabrication of QWIP devices because as derived in Section 2.1 a good QWIP has "as many quantum wells as practical". For the case of unstrained superlattices, what is practical is limited only by the reliability of the growth equipment and the perseverance of the grower. However, for the case of strained superlattices, there is a *critical layer thickness* beyond which threading dislocations present at a density of about 1000 per square centimeter in high quality vertical gradient freeze (VGF) grown semi-insulating substrates[89] begin to glide forming misfits in the material. The effect of these misfits on QWIP performance has not been explicitly determined, but Section 3.4 indisputably establishes their existence and presents a theory by Prof. Fitzgerald for the critical thickness which is consistent with the measured results.[90] Thus Section 3.4 lays the groundwork for a future study of the effect of dislocations on QWIP performance. In comparison it is not uncommon to find reports of strained superlattices that are claimed to be defect free for layer structures well beyond the critical thickness.[91] This conclusion is often based on the use of relatively insensitive measurement techniques, like X-ray, which can not determine the existence of defects at the densities often present. The result though is that since the layers are claimed to be defect free, no effort is put in to determining the effect of defects on device performance thus squandering an opportunity to better understand critical issues in device performance, reliability and uniformity. Clearly the applicability of this work is not limited to only QWIPs.

Defects in the crystalline structure of semiconductors can dramatically change the electrical and optical properties of the constituent material. Optical devices in particular tend to be extremely sensitive to defects in the crystal structure which can act as non-radiative recombination centers in optical emitters and sources of increased dark current in optical detectors such as QWIPs. At the same time, the availability of high quality bulk single crystal semiconductor substrates with low defect density is largely limited to Silicon, GaAs, GaP and InP. These substrates serve as templates for epitaxial growth in techniques such as MBE described above in Section 3.2. If the epitaxial layers grown on top

of the substrate consist of the either the same semiconductor material as the substrate, "homoepitaxy", or a different semiconductor material that is chosen to have the same lattice constant, "lattice matched heteroepitaxy" there are no physical limits to the thickness of the epitaxial layers. Common examples of homoepitaxy are GaAs p-n junction light emitting diodes and Silicon p-i-n photodetectors. In both cases, the only difference between layers of the device is the doping type and/or concentration; the semiconductor material is the same throughout. Common examples of lattice matched heteroepitaxy are GaAs/AlGaAs infrared lasers grown on GaAs substrates for CD players and $In_{0.53}Ga_{0.47}As/In_{0.52}Al_{0.48}As$ 1.55μm infrared lasers grown on InP for optical fiber communications.

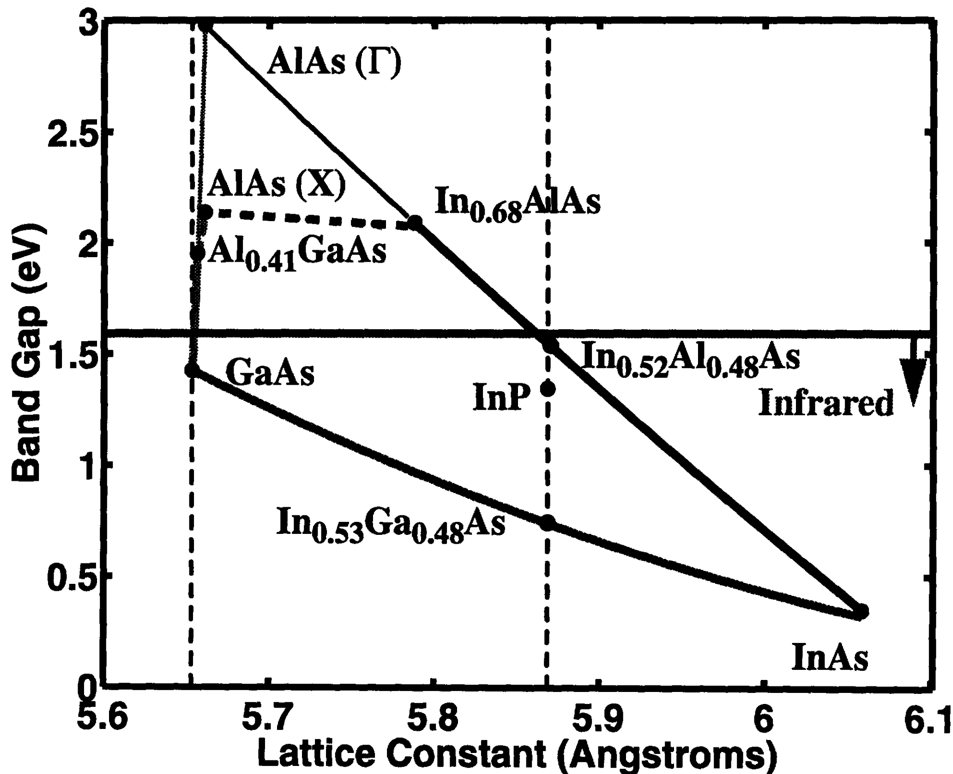


Figure 3.5: Bandgaps Available for the AlGaInAs system at Room Temperature[66] Solid lines indicate direct band gaps at the Γ point, dotted lines are indirect.

In both cases the epitaxial layers are lattice matched to the substrate material but in the GaAs/AlGaAs case the AlGaAs ternary material is lattice matched to the GaAs substrate

for any arbitrary fraction of Aluminum whereas in the $In_{0.53}Ga_{0.47}As/In_{0.52}Al_{0.48}As$ case, only the particular ternary compositions given are lattice matched to the InP substrate. Figure 3.5 shows a plot of band gap verses lattice constant for the AlGaInAs quaternary material system with dotted vertical lines indicating the lattice constants for the two substrates, GaAs and InP, with lattice constants in the range accessible by this material system.Notice that for a GaAs substrate direct band gaps from 1.42eV to 1.9eV are available at room temperature in the AlGaAs material system whereas for the InP substrate direct band gaps from 0.75eV to 1.55eV are available using the quaternary InGaAlAs.

To access band gaps outside these ranges, epitaxial layers must be grown that have lattices constants different from that of the substrate material. Growth of epitaxial layers on non-lattice matched substrates is termed "strained layer heteroepitaxy" and is the subject of this section. Epitaxial layers grown strained on a substrate with a different lattice constant can be either elastically strained or partially relaxed as shown schematically in Figures 3.6 and 3.7. Because energy is required to form misfit dislocations, in sufficiently thin or sufficiently lattice matched epitaxial layers the material is elastically strained as shown in Figure 3.6.

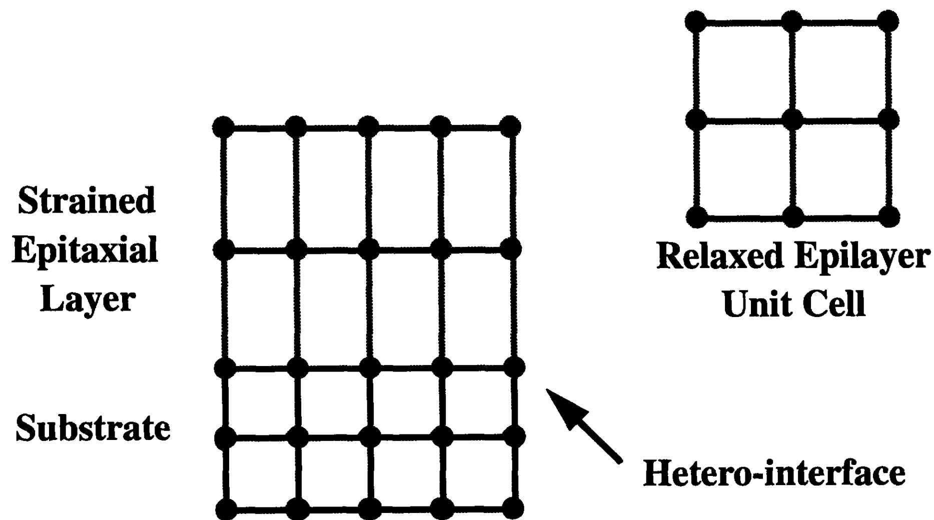


Figure 3.6: Schematic diagram for elastically strained epitaxial layer

However at some thickness, termed the critical thickness, h_c, the strain energy released by the formation of a misfit dislocation equals the energy required to form a misfit dislocation. Beyond this critical thickness threading dislocations already in the substrate and threading through the epitaxial layer will glide laterally producing a misfit dislocation in the epitaxial material. This mechanism for misfit formation is shown in Figure 3.7 as first suggested by Jesser and Mathews[92,93,94] who also calculated the theoretical critical thickness for this mechanism. The first experimental evidence in semiconductor systems is from GaAs/GaAsP superlattices grown on GaAs by Mathews and Blakeslee[95] for whom the critical thickness, h_c, is named the "Mathews-Blakeslee critical thickness".

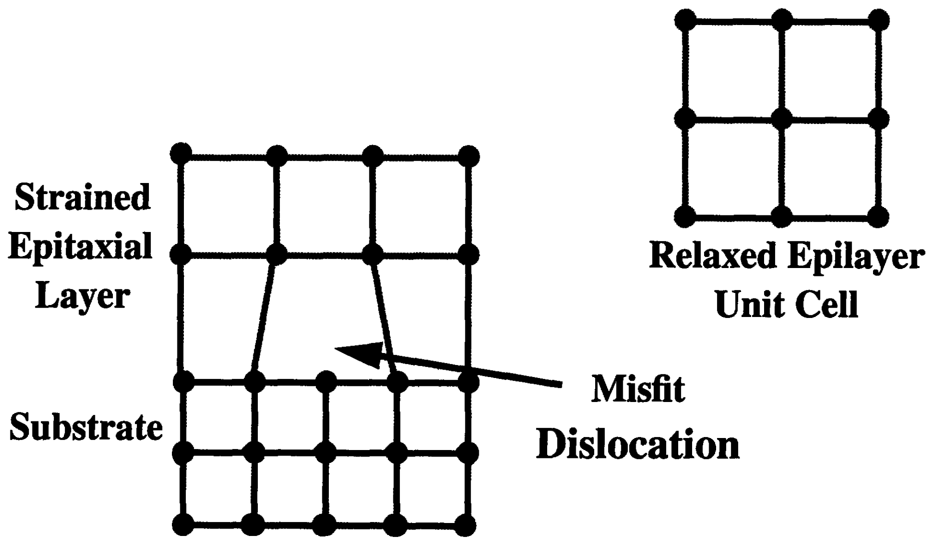


Figure 3.7: Schematic diagram for misfits is relaxed epitaxial layers

In this section the critical thickness will be calculated first for a single strained epitaxial layer in sub-section 3.4.1 following the theoretical work of Mathews[95] as reviewed by Fitzgerald.[96] Then in sub-section 3.4.2 the theory will be extended for use in strained layer superlattices following the work of Fitzgerald.[96,90] In subsection 3.4.3 results of an experimental study done by the author in collaboration with Prof. Eugene Fitzgerald and Mayank Bulsara will be shown in which cathodoluminescence is used to measure the onset of misfit glide and thereby determine the critical layer thickness in a

superlattice. This study very clearly corroborates the superlattice theory of Fitzgerald and sets new limits on the number of superlattice periods that can be grown without misfit formation. Finally, subsection 3.4.4 concludes the remarks on strain and superlattice critical thickness.

3.4.1 Single layer critical layer thickness

The material systems of interest in this work are all III-V semiconductors crystallizing in the Zinc-blende crystal lattice. For this crystal structure the slip system is {111}<110>. For epitaxial layers grown on (100) substrates this results in an orthogonal set of misfit dislocations that form along <110> directions. The dominant type of misfit dislocation observed in strained layer heteroepitaxy with small lattice constant mismatches are the 60° dislocations which have Burgers vectors 60° from the interface <110> dislocation line direction and can glide on a {111} plane. A diagram of the dislocation flow process is shown in Figure 3.8.

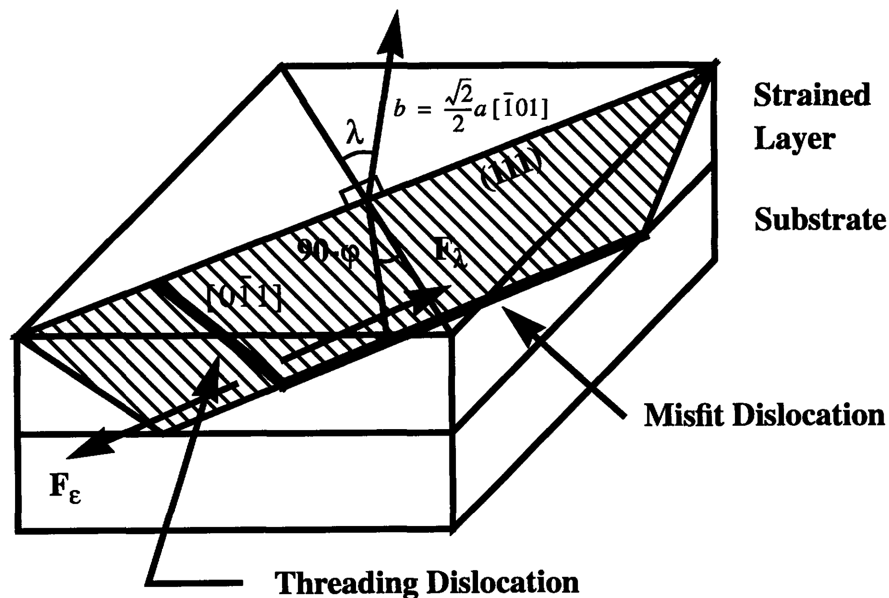


Figure 3.8: Diagram of dislocation formation in zinc-blende crystals. Hatched region is a {111} plane.[96]

First, the critical thickness for a single epitaxial overlayer will be calculated based on the energy considerations of Mathews. The energy per unit area of an elastically strained layer of thickness h is: [97]

$$E_{\varepsilon} = \varepsilon^2 Yh \qquad \frac{dyne}{cm} \tag{3.1}$$

where ε is the elastic strain in the overlayer defined by:

$$\varepsilon = \frac{a_s - a_o}{a_o} \tag{3.2}$$

with a_o the lattice parameter for the unstrained overlayer and a_s the lattice parameter for the substrate. The epitaxial overlayer is attached to the substrate at an interfacial plane, so Y is Young's modulus under biaxial stress and is defined by:

$$Y = \frac{1}{2}\left(C_{11} + 2C_{12}\right)\left[3 - \frac{C_{11} + 2C_{12}}{C_{11} + 2\left(2C_{44} - C_{11} + C_{12}\right)\left(l^2m^2 + m^2n^2 + n^2l^2\right)}\right] \qquad \frac{dyne}{cm^2} \tag{3.3}$$

where l, m and n are the directional cosines that relate the axis normal to the interface to the cube axes of the zinc-blende crystal and the C_{ij} are the elastic constants of the epitaxial overlayer. For the specific case considered here of growth on (100) oriented substrates Young's modulus becomes:

$$Y = C_{11} + C_{12} - \frac{2C_{12}^2}{C_{11}} \qquad \frac{dyne}{cm^2} \tag{3.4}$$

If it is assumed, as did Mathew's, that the semiconductors are isotropic, Young's modulus for biaxial stress reduces to:

$$Y = 2G\frac{(1 + \nu)}{1 - \nu} \qquad \frac{dyne}{cm^2} \tag{3.5}$$

where G is the Shear modulus and ν is Poisson's ratio. However, the zinc-blende crystals of interest here are mildly anisotropic so subsequently Equation 3.4 will be used in favor of Equation 3.5.

Misfit dislocations can be introduced to reduce the strain as shown in Figure 3.7 with a cost in energy per unit area of:[96]

$$E_d = 2\left(\frac{1}{S}\right)\left[\frac{1}{2}Db\left(1 - \nu cos^2\alpha\right)\right]\left[\ln\left(\frac{R}{b}\right) + 1\right] \qquad \frac{dyne}{cm} \tag{3.6}$$

where the fist factor of 2 is for the 2 sets of orthogonal dislocations in each of the <110> directions, 1/S is the number of dislocations per unit length along the interface, b is the magnitude of the Burgers vector of the dislocations, $b = a_s/\sqrt{2}$ for the case of (100) oriented substrates, ν is Poisson's ratio, R is the outer cut-off radius of the dislocation energy shown in Equation 3.10, α is the angle between the Burgers vector and the dislocation line and D is the average shear modulus of the interface given by:

$$D = \frac{G_o G_s b}{\pi (G_o + G_s)(1 - \nu)} \qquad \frac{dyne}{cm} \tag{3.7}$$

where G_o and G_s are the shear moduli of the epitaxial overlayer and the substrate respectively and are given in terms of the elastic compliances as:

$$G = C_{44} - \frac{1}{3}(2C_{44} + C_{12} - C_{11}) \qquad \frac{dyne}{cm^2} \tag{3.8}$$

and Poisson's ratio n is given in terms of the elastic compliances as:

$$\nu = \frac{C_{12}}{C_{12} + C_{11}} \tag{3.9}$$

If misfit dislocations exist at the interface then the total misfit including both elastic and plastic deformation is:

$$f \equiv \frac{a_s - a_o}{a_o} = \varepsilon + \delta \tag{3.10}$$

where ε is the elastic strain and δ is the plastic strain and the interface dislocation spacing is given by:

$$S = \frac{b_{eff}}{\delta} \qquad cm \tag{3.11}$$

where b_{eff} is the effective Burgers vector defined as the interface-plane component of the Burgers vector in the direction of the spacing S. For the case of 60° dislocations which have Burgers vectors 60° from the interface <110> dislocation line direction and can glide on a {111} plane $b_{eff} = b/2$ and the angle α is $60°$.

The total energy of the system is the sum of the strain energy given by Equations 3.1 and the dislocation energy given by Equation 3.6, $E_{tot} = E_\varepsilon + E_\delta$. The equilibrium value of strain, ε, in the system is found by differentiating the total energy with respect to strain and setting the result equal to zero giving for the equilibrium strain ε^*:

$$\varepsilon^* = \frac{D\left(\frac{b}{b_{eff}}\right)\left(1 - \nu cos^2\alpha\right)\left[\ln\left(\frac{R}{b}\right) + 1\right]}{2Yh} \tag{3.12}$$

The value of R depends on the dislocation density. If the interface dislocation spacing S divided by 2 is greater than the epitaxial overlayer thickness h, then the outer cut-off radius R is the height of the epitaxial overlayer giving for the equilibrium strain:

$$\varepsilon^* = \frac{D\left(\frac{b}{b_{eff}}\right)\left(1 - \nu cos^2\alpha\right)\left[\ln\left(\frac{h}{b}\right) + 1\right]}{2Yh} \tag{3.13}$$

which is simply Equation 3.10 with R replace by h. This corresponds to the low misfit density case and is the proper equation to use when calculating the critical thickness, h_c, which is the epitaxial overlayer thickness at which the addition of the first misfit dislocation becomes energetically favorable. Below and up to this critical thickness, the epitaxial overlayer is totally strained with no plastic deformation, i.e. $f = \varepsilon^*$ and $\delta = 0$. The value of the critical thickness can be calculated from a transcendental equation derived from Equation 3.11 by setting $f = \varepsilon^*$, $h = h_c$, $b = 2b_{eff}$ and solving for h_c:

$$h_c = \frac{D\left(1 - \nu cos^2\alpha\right)\left[\ln\left(\frac{h_c}{b}\right) + 1\right]}{Yf} \tag{3.14}$$

Equation 3.12 is the critical layer thickness for a single epitaxial overlayer grown strained on a thick substrate and is the desired result of this subsection. Epitaxial layers

with misfit *f* given by Equation 3.8 can be grown on a thick substrate to thickness h_c given by Equation 3.12 without introducing misfit dislocations. Table 3.1 gives values of the required parameters for the binary constituents of the AlGaInAs material system.

Material	Y ($x10^{11}$ dyne/cm^2)	G ($x10^{11}$ dyne/cm)	ν	a (Å)
GaAs	12.34	4.14	0.311	5.6532
AlAs	16.50	2.59	0.274	5.6623
InAs	7.93	5.10	0.352	6.0584
InP	9.49	3.02	0.357	5.8687
GaP	14.78	4.96	0.306	5.4505
AlP	20.76	5.27	0.263	5.4635

Table 3.1: Mechanical properties of AlGaInAs material system

These have been derived using Equation 3.4 for Young's modulus, Equation 3.8 for the Shear modulus and Equation 3.9 for Poisson's ratio from data for the elastic compliances from Landolt and Bornstein[66] and summarized in Table 3.2.

Material	Compliances all $x10^{11}$ dyne/cm^2			a (Å)
	C_{11}	C_{12}	C_{44}	
GaAs	11.81	5.32	5.94	5.6532
AlAs	12.02	5.70	5.89	5.6623
InAs	8.33	4.53	3.96	6.0584
InP	10.11	5.61	4.56	5.8687
GaP	14.05	6.20	7.03	5.4505
AlP	18.83	6.71	3.69	5.4635

Table 3.2: Mechanical compliances for selected III-V semiconductors[66]

Parameters for the ternary and quaternary components of this material system can be calculated from the binary parameters using Vegard's law which for the parameter *P* in a ternary of the form $Al_xGa_{1-x}As$ is:

$$P_{Al_xGa_{1-x}As} = (x)P_{AlAs} + (1-x)P_{GaAs} \tag{3.15}$$

and in a quaternary of the form $Al_xGa_yIn_{1-x-y}As$ is:

$$P_{Al_xGa_yIn_{1-x-y}As} = (x)P_{AlAs} + (y)P_{GaAs} + (1-x-y)P_{InAs} \tag{3.16}$$

Figure 3.9 shows a plot of Equation 3.12 for the case of InGaAs on GaAs over the range of Indium fractions from 10% to 30% Indium in InGaAs. The critical thickness for a single epitaxial layer is seen to be less than 200Å over the entire range of Indium fractions plotted.

Having derived the critical thickness for a single epitaxial overlayer using the energy balance approach of Mathews and Blakeslee, the next subsection uses a slightly different approach based on balancing of the force on a misfit dislocation to once again derive the critical thickness for a single epitaxial overlayer and then extend this derivation to the case of a superlattice which is of particular interest to this thesis. The discussion follows that of Fitzgerald.[96,90]

Figure 3.9: Critical thickness for single epitaxial layer of InGaAs on GaAs

3.4.2 Superlattice critical layer thickness

This subsection follows the work of Fitzgerald in deriving the critical layer thickness for a superlattice with n pairs of alternating layers, each pair consisting of a strained material with layer thickness h_1 and an unstrained material with layer thickness h_2.

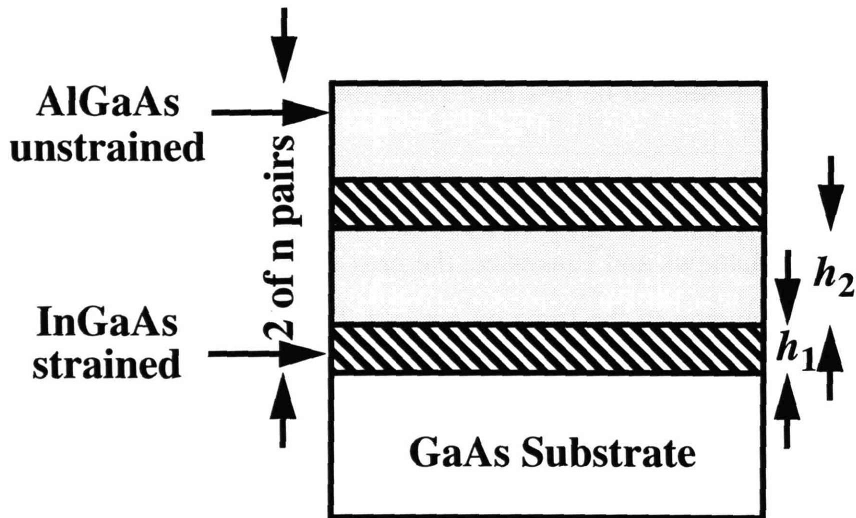


Figure 3.10: Schematic diagram of strained layer superlattice

The strained layer is the first to be grown on the substrate and subsequent layers are grown as shown schematically in Figure 3.10. The derivation here is equivalent to the approach used in Section 3.4.1 when the superlattice consists of only a single strain layer, but this time the *force* on a pre-existing threading dislocation due to stress i n the overlayer will be balanced with the line-tension force of the dislocation segment residing in the interface to determine the overlayer thickness which causes a pre-existing threading dislocation to glide forming a misfit dislocation. The issue of where the pre-existing threading dislocations come from will not be addressed other than to say that the typical two inch semi-insulating and p-type GaAs substrates from American Xtal Technology (AXT) grown by vertical gradient freeze (VGF) and used in this thesis had areal densities of threading dislocations specified to be less than 1000 per square centimeter. n-type GaAs substrates also

grown by AXT using VGF had threading dislocation densities of less 100 per square centimeter.

The glissile dislocations in zinc-blende crystals are the 60° dislocations which have a {111} slip plane. Glide for such a dislocation threading through a strained overlayer is governed by the component of stress which acts in the direction of the dislocation slip. The angles involved are shown in Figure 3.8 with resolved shear stress given by:

$$\sigma_{res} = \sigma \cos\lambda \cos\varphi \qquad \frac{dyne}{cm^2} \tag{3.17}$$

where the $\cos\lambda\cos\varphi$ term is the Schmid factor which resolves the stress, σ in the strained overlayer into the glide system of the dislocation. For the zinc-blende crystals of interest here, $\cos\lambda = 1/2$ and $\cos\varphi = \sqrt{2/3}$.[92] The force per unit length on the dislocation in the overlayer is given by:

$$F_{res} = \sigma_{res} b = \sigma b \cos\lambda \cos\varphi \qquad \frac{dyne}{cm} \tag{3.18}$$

where b is the magnitude of the Burgers vector and the stress in the overlayer, σ, is given for a biaxially strained layer as:

$$\sigma = \sigma_x = \sigma_y = Y\varepsilon \qquad \frac{dyne}{cm^2} \tag{3.19}$$

For overlayers with thickness less than or equal to the critical layer thickness dislocation glide has not begun and so all the misfit is incorporated as strain in the overlayer. Inserting Equation 3.19 into Equation 3.18 then gives:

$$F_{res} = Y\varepsilon b \cos\lambda \cos\varphi \qquad \frac{dyne}{cm} \tag{3.20}$$

The total lateral force on the dislocation in the overlayer is the lateral component of the force per unit length given by Equation 3.20 multiplied by the lateral length of the dislocation threading through a single strained overlayer of thickness h which is $h/\cos\varphi$. The total lateral force on the dislocation in the overlayer, F_o, is then given by:

$$F_o = \frac{F_{res}h}{\cos\varphi} = Y\varepsilon bh\cos\lambda \qquad dyne \tag{3.21}$$

The force F_o which acts on the point of intersection between the threading dislocation and the interface plane is balance by the line tension force F_l which is given by:[94]

$$F_l = \frac{Db}{2}\left(1 - \upsilon cos^2\alpha\right)\left[\ln\left(\frac{h}{b}\right) + 1\right] \qquad dyne \tag{3.22}$$

noting that $\cos\lambda = 1/2$ and prior to the initiation of plastic deformation $f = \varepsilon$ combining Equations 3.21 and 3.22 gives the critical thickness for a single strained overlayer equivalent to Equation 3.14 derived based on energy considerations in section 3.4.1.

$$h_c = \frac{D\left(1 - \nu cos^2\alpha\right)\left[\ln\left(\frac{h_c}{b}\right) + 1\right]}{Yf} \qquad cm \tag{3.23}$$

To extend this theory to superlattices of the form shown in Figure 3.10, Fitzgerald[96] noted that only the strained layers contribute to the force in the overlayer as given by Equation 3.21. This requires that in a superlattice with n pairs of alternating layers, each pair consisting of a strained material with layer thickness h_1 and an unstrained material with layer thickness h_2 the total thickness of strained material is nxh_1, having strain ε_1 from Equation 3.10 and thus Equation 3.12 must be replace by:

$$F_o = \frac{F_{res}nh_1}{\cos\varphi} = Y\varepsilon bnh_1\cos\lambda \qquad dyne \tag{3.24}$$

Furthermore, the line tension force balancing this strain-resolved force, F_l, must include both the strained and unstrained layers. Noting that the first misfit will occur at the interface with a strained layer on the bottom and an unstrained layer on top, this then requires that the h is Equation 3.22 for the line tension force be replaced by $(n-1)\ (h_1 + h_2) + h_1$ giving for the line tension force the new equation:

$$F_l = \frac{Db}{2}\left(1 - \upsilon cos^2\alpha\right)\left[\ln\left(\frac{(n-1)\left(h_1 + h_2\right) + h_1}{b}\right) + 1\right] \qquad dyne \tag{3.25}$$

Setting Equations 3.24 and 3.25 equal and solving for the critical thickness for the strained layer with the number of periods in the superlattice fixed gives:

$$h_{1c} = \frac{D\left(1 - \nu cos^2\alpha\right)\left[\ln\left(\frac{(n-1)\left(h_{1c}+h_2\right)+h_{1c}}{b}\right)+1\right]}{Yf_1 n} \quad cm \quad (3.26)$$

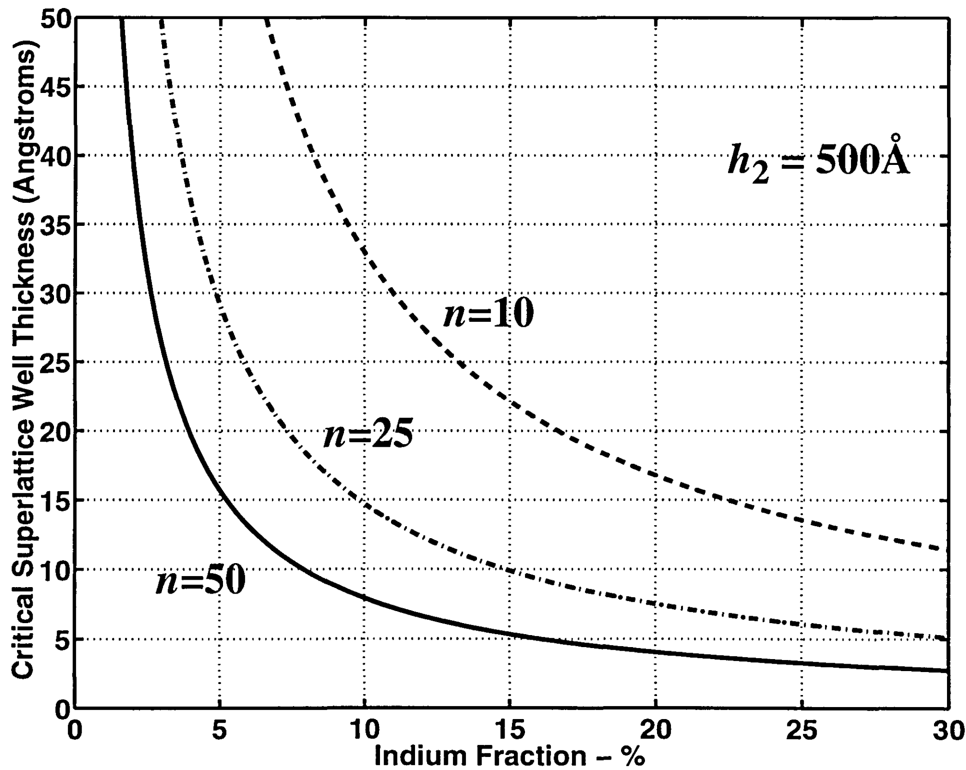


Figure 3.11: Critical layer thickness for superlattice as a function of Indium fraction in the strained layer and number of superlattice periods n. Solid curve $n = 50$,dot-dashed curve $n = 25$ and dashed curve $n = 10$. All curves h_2=500Å.

Figure 3.11 shows a plot of Equation 3.26 for the case of a superlattice grown on a GaAs substrate with InGaAs strained layers separated by unstrained AlGaAs layers of thickness 500Å with the strained layer thickness plotted on the abscissa. Equivalently, with fixed strain and layer thicknesses Equations 3.24 and 3.25 can be solved for the critical number of superlattice periods which are stable with respect to dislocation glide.

$$n_c = \frac{D\left(1 - \nu\cos^2\alpha\right)\left[\ln\left(\frac{(n_c - 1)\left(h_1 + h_2\right) + h_{1c}}{b}\right) + 1\right]}{Yf_1 h_1} \tag{3.27}$$

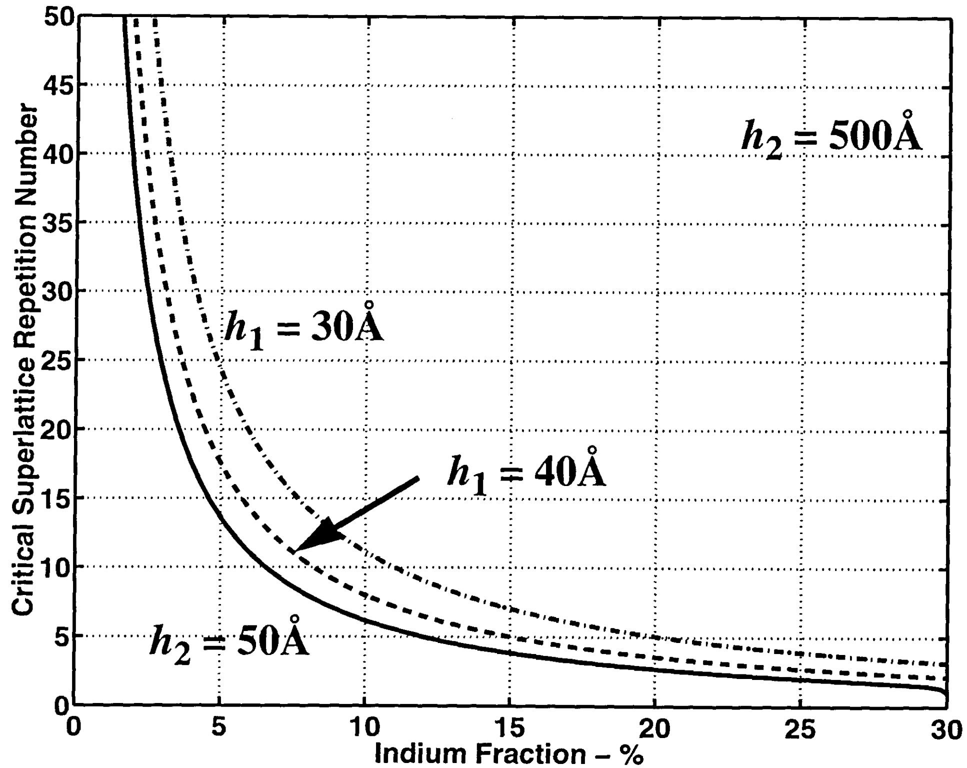


Figure 3.12: Critical number of superlattice periods as a function of Indium fraction in the strained layer and strained layer thickness h_1. Solid curve h_1=50Å,dot-dashed curve h_1=40Å and dashed curve h_1=30Å. All curves h_2=500Å.

Figure 3.12 shows a plot of Equation 3.27 for the case of a superlattice grown on a GaAs substrate with InGaAs strained layers separated by unstrained AlGaAs layers of thickness 500Å with the number of superlattice periods plotted on the abscissa.

In this subsection the theory for calculating the critical layer thickness in semiconductor superlattices has been developed. Since this theory calculates the thickness as which the first threading dislocation begins to glide, care must be taken in experimentally verifying the theory to use an experimental technique sensitive to very low densities of misfit

dislocations. Cathodoluminescence is one such technique and a small study done by the author in collaboration with Prof. Fitzgerald and Mayank Bulsara using Cathodoluminescence to characterize InGaAs/GaAs superlattices will be described in the following subsection.

3.4.3 Experimental verification of SL critical layer thickness theory

Attempts to experimentally verify the critical layer thickness in superlattices have been fraught with theoretically unjustifiable claims and use of experimental techniques that are insufficiently sensitive to properly evaluate the onset of misfit glide that identifies the critical layer thickness. In particular, theoretical claims that the superlattice critical thickness can be calculated from the single layer critical thickness, Equation 3.23, by taking the average strain in superlattice as ε_{avg} defined by taking the strain in each layer and weighting with the layer thickness and then solving Equation 3.23 for the superlattice critical thickness are inconsistent with the correct expression Equation 3.26 and overestimate the critical thickness.[98] Also, often heard claims of the form "If a single strained layer is below the Mathew's Blackeslee limit and the strained layers are separated by a sufficiently large unstrained layer the strained layers are uncoupled and so there is no limit to the number of superlattice periods that can be grown." simply beg the question of how large is "sufficiently large". Equation 3.27 tells exactly how large "sufficiently large" is and not only is there always a limit to the number of superlattice periods that can be grown without misfit dislocations, but the inclusion of the unstrained layer h_2 though helpful comes in within a natural log term and given the slow growth rates of most semiconductor growth systems, ~1μm/hr for MBE or up to ~10μm/hr for MOCVD, is quite limited in its stabilizing ability for reasonable layer thicknesses.

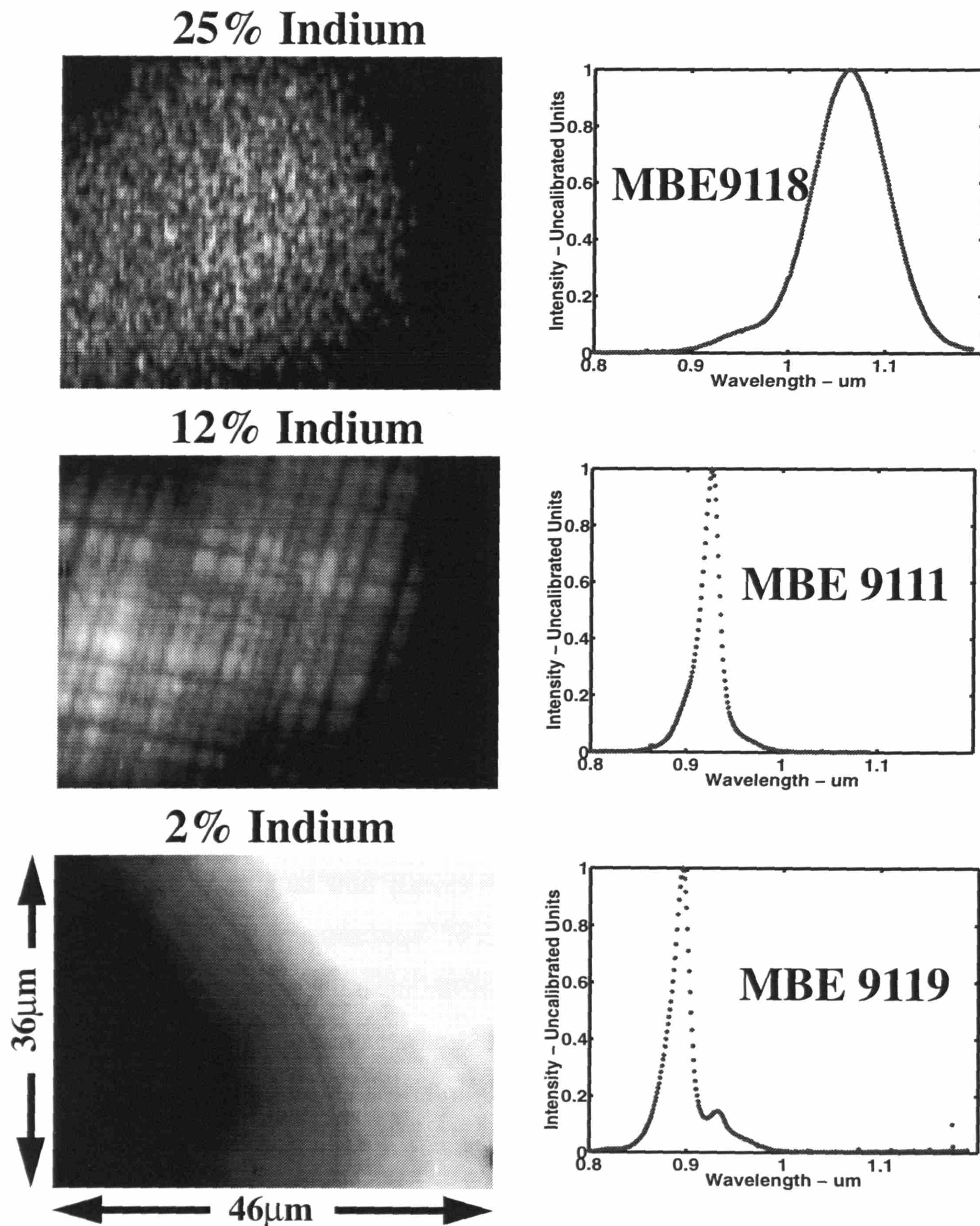


Figure 3.13: Cathodoluminescence from a 25 period superlattice consisting of 50Å InGaAs and 500Å of AlGaAs. Samples grown by the author using MBE, CL courtesy of M. Bulsara. 2% Indium sample has no features to focus on characteristic of high quality defect free material. 12% Indium sample shows large density of individual misfit dislocations along the two perpendicular <110> crystal directions. 25% Indium sample shows very large density of misfits coalesced into groups which obscure the individual dislocation lines. Note also the large shift in PL peak wavelength for the 25% Indium sample.

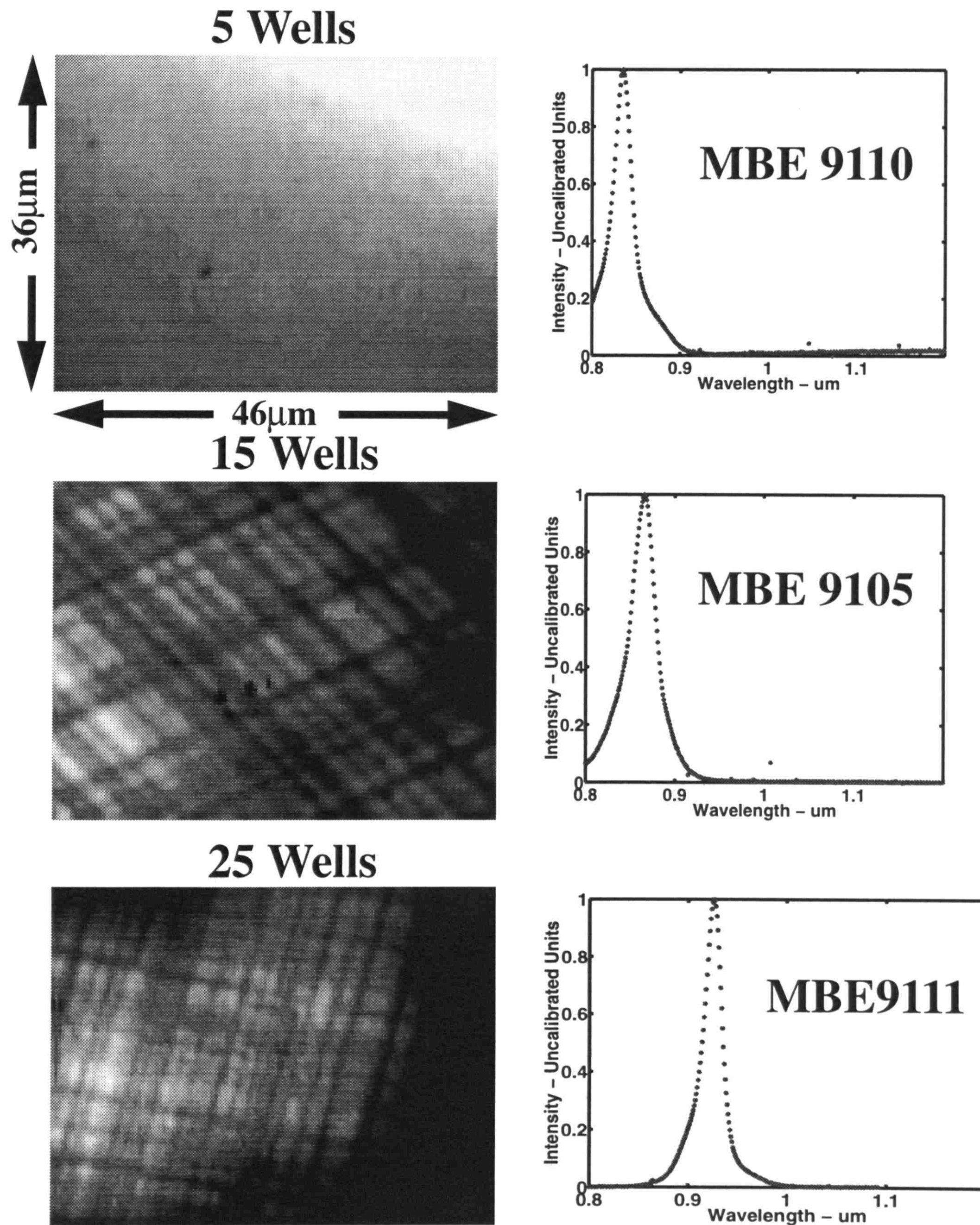


Figure 3.14: Cathodoluminescence from superlattice consisting of 50Å $In_{0.12}Ga_{0.88}As$ quantum wells and 500Å AlGaAs barriers. Samples grown by the author using MBE, CL courtesy of Mayank Bulsara. 5 well sample has no features to focus on characteristic of high quality defect free material. 15 well sample shows large density of individual misfit dislocations along the two perpendicular <110> crystal directions. 25 well sample also shows large density of individual misfits indicating n is beyond the critical number. Note also the large shift in PL peak wavelength for the 25 well sample.

Experimental claims of critical layer thicknesses significantly larger than those given by Equations 3.26 and 3.27 are typically based on X-ray data which because of the penetrating nature of the X-rays and the large spot sizes of the X-ray beams are sensitive primarily to the average plastic deformation which may be only 1 part in 10 million for the first dislocation in a 4mm X-ray spot, i.e. much below the resolution of the measurement technique. For this reason cathodoluminescence which can image with submicron resolution individual misfit dislocations has been used to compare with the critical layer thickness theory described above.

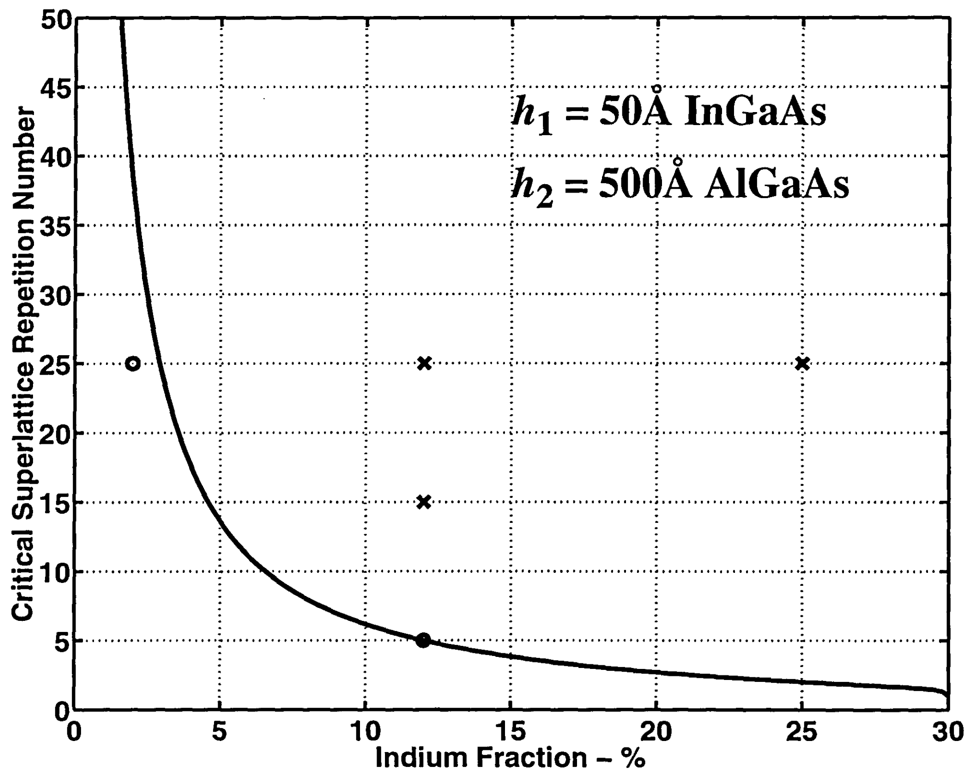


Figure 3.15: Five samples grown to test SL critical thickness theory of Equation 3.27 plotted with solid line. "x" indicates beyond critical thickness, "o" indicates less than critical thickness. Samples grown using MBE by author, CL done by Mayank Bulsara.

Finally, Figure 3.15 shows the calculated critical number of superlattice periods for 50Å InGaAs quantum wells grown on GaAs and separated by 500Å AlGaAs barriers with a solid line and superimposes the experimental data with "o" indicating samples grown

with no misfit dislocations and "x" indicating samples grown with misfit dislocations as taken from the data of Figures 3.13 and 3.14.

3.4.4 Conclusions on strain and SL critical layer thickness

In conclusion, this section has developed a theory for calculating the critical layer thickness both in terms of superlattice period number and strained layer thickness. The theory is based on balancing the force on a threading dislocation due to the stress in the strained layer and the opposing force on the threading dislocation at the interface between the substrate and the superlattice. The result is a theoretical extension of the Mathew's Blakeslee critical thickness for a single strained overlayer. Five InGaAs/AlGaAs superlattices with varying number of superlattice periods and Indium concentrations were grown by the author using MBE to test the theory and Cathodoluminescence measurements done by Mayank Bulsara are consistent with the theory developed. More work is currently in progress by Prof. Fitzgerald to add SiGe/Si strained layer superlattice data to the current data set presented here.

3.5 Fabrication of normal incidence QWIPs

Fabrication of the QWIP devices for this thesis was intentionally kept simple with several options at each step to allow for optimization of the process for specific applications and to allow for working around the inevitable broken piece of equipment. The basic flow for a typical etch isolated LWIR InGaAs/AlGaAs QWIP is shown below in Figures 3.16 through 3.18. Figure 3.16 describes the structure used for MBE growth. The GaAs substrate can be either semi-insulating or n-type, though n-type substrates are sometimes preferred for their lower density of threading dislocations.[89] For substrates mounted with Indium the Indium on the back side of the substrate was removed with $HCl:H_20$ as described in the MIT MBE manual[81].

Figure 3.17 shows a schematic for the result of the pixel isolation etch which was typically done using a weak phosphoric acid based crystallographic etch consisting of $DIH_2O:H_3PO_4:H_2O_2$ in a ratio of 20:1:1 which consistently gave an etch rate for GaAs of 0.24μm/min at room temperature.[99,100] Side walls for <$\bar{1}$10> oriented mesa segments

reveal $\langle\bar{1}\bar{1}1\rangle$ Ga planes which sloped outward from the mesa wall. For $\langle 110\rangle$ oriented mesa segments the walls are undercut revealing $\langle 2\bar{2}\bar{1}\rangle$ Ga planes. Sidewall orientation is particularly important for designs that require metallization over the edge of the mesa. Metal lines traversing the $\langle\bar{1}10\rangle$ oriented mesas lay continuously over the $\langle\bar{1}\bar{1}1\rangle$ Ga planes, but metal lines traversing the $\langle 110\rangle$ oriented mesas are broken at the edge of the mesa.

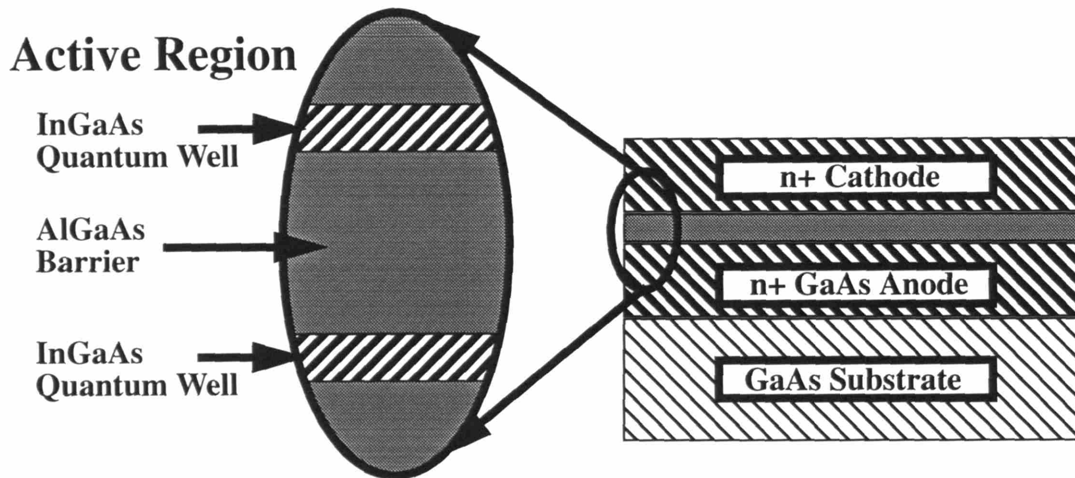


Figure 3.16: Step 1: MBE Growth of QWIP. Typical epitaxy consists of ~1μm Si doped $1\mathrm{x}10^{18}$ GaAs anode, active superlattice region consisting of 10 pairs of Si doped $8\ \mathrm{x}10^{17}\ /\mathrm{cm}^3$ 50Å $In_{0.10}Ga_{0.90}As$ quantum wells separated by 500Å undoped $Al_{0.20}Ga_{0.80}As$ barriers and capped with final $Al_{0.20}Ga_{0.80}As$ barrier and 0.5μm Si doped $1\mathrm{x}10^{18}$ GaAs cathode.

Contact metallization was done using Ni/AuGe/Ni/Au contact metal layers, thermal evaporation and a lift-off technique. The metallization was performed in two steps as shown in Figure 3.18. The first step defined the ohmic metal for contact to the anode and cathode of the QWIP device. This was done by depositing ~2000Å of PECVD Silicon Dioxide over the entire wafer, patterning the oxide using a positive photoresist and then thermally evaporating the ohmic contact metallization. Slightly overetching the dielectric layer leaves an overhang of resist which facilitates the lift-off of the unneeded metal. For the cathode the ohmic metal consists of a ring around the center of the active region.

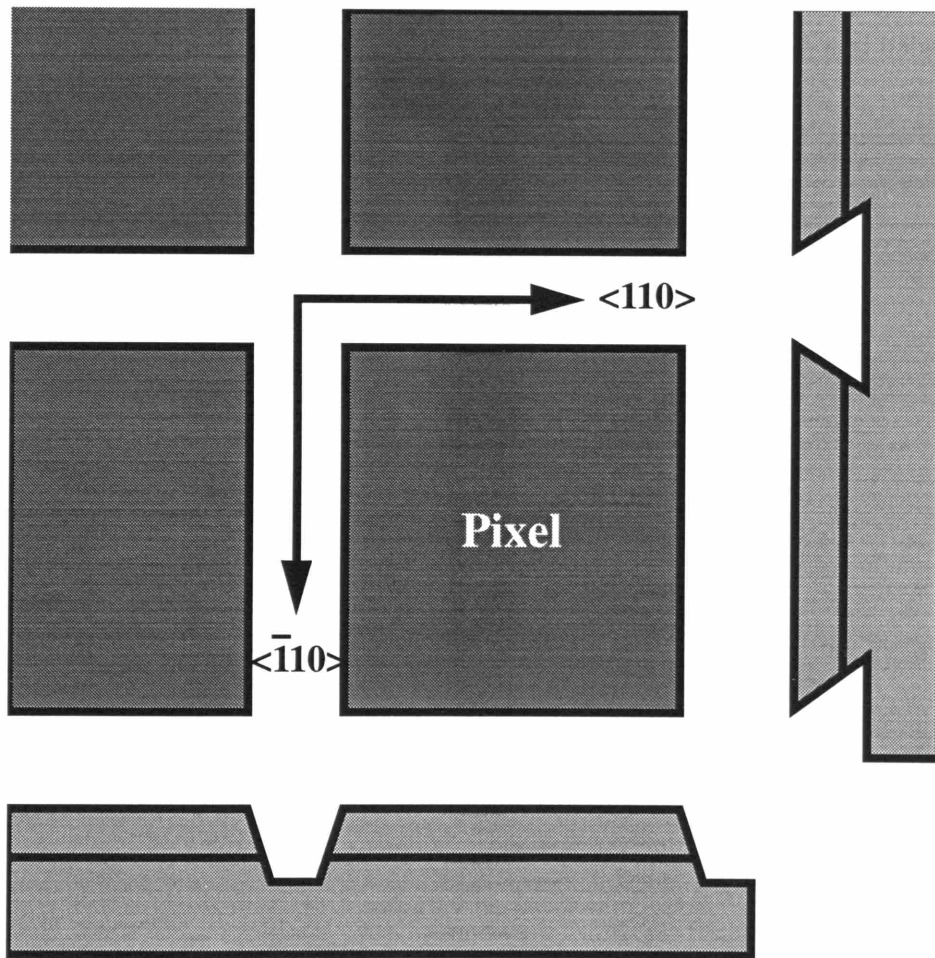


Figure 3.17: Step 2: Pixel Isolation Etch. Cross section through QWIP pixel perpendicular to the <$\bar{1}$10> direction, bottom and perpendicular to <110> direction, right. 20:1:1, $DIH_2O:H_3PO_4:H_2O_2$, etchant reveals <$\bar{1}\bar{1}$1> and <2$\bar{2}\bar{1}$> Ga planes. Figure drawn for pixel array.

For the anode the ohmic metal is a large pad located near the device. As designed, the anode is common to all the devices on each chip allowing a single anode contact to contact all the devices for efficient bonding and pin usage. The ohmic metal layers consisted of 100Å of Ni which formed a barrier layer to the diffusion of Au into the semiconductor, 2500Å of Au/Ge(12%) eutectic co-evaporated from a commercial source, 250Å of Ni to form a barrier layer which prevents intermixing of upper level Au with Au/Ge eutectic and

2000Å of Au to prevent oxidation of the Ni. After evaporation and liftoff the ohmic metal was annealed in a rapid thermal annealer (RTA) for 30 seconds at 380C.

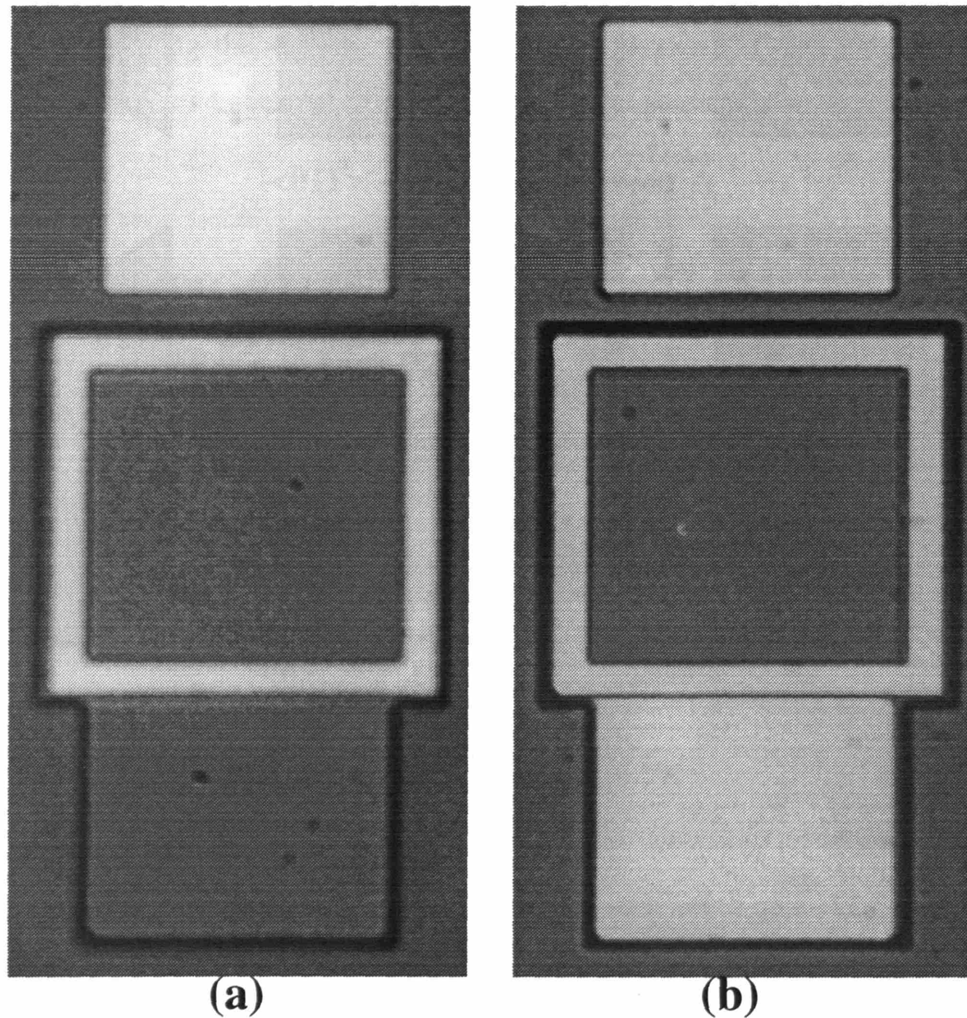


Figure 3.18: Step 3: Contact metallization. Figure 3.18a shows a 100μm x 100μm QWIP pixel after ohmic metallization. Figure 3.18b shows the same device after completion of bond pad metallization.

Annealing time and temperature are critical, as is the thickness of the first Ni layer, to get low resistance ohmic contacts that do not short through from the cathode to the anode. Since metal thickness and annealer calibration vary substantially from run to run, the transmission line measurement technique[101] was used to characterize both anode and

cathode contacts for each chip of devices tested. Figure 3.18a shows a single pixel after ohmic metal deposition but before bond metal deposition.

Bond pad metallization was also done using a lift-off technique, this time using Futurrex NR-8 negative photoresist to form the reentrant profile.[102] Futurrex NR-8 was spun at a thickness of ~3µm, patterned and slightly overdeveloped leaving an overhang profile ideal for lift-off. Total metal thicknesses as large as 1µm have been lifted off with no edge problems using this resist. Bond pad metallization consisted of ~0.5µm to 1µm of pure gold evaporated using thermal evaporation. This thickness was chosen to allow easy wire bonding using 2mil gold wire and a ball bonder with a heated stage. Figure 3.18b shows a single 100µm x 100µm pixel after completed metallization just prior to bonding into a DIP package. Figures 3.19 and 3.20 show an SEM of completed 200µm x 200µm devices including both metallization and both orientations of edge walls.

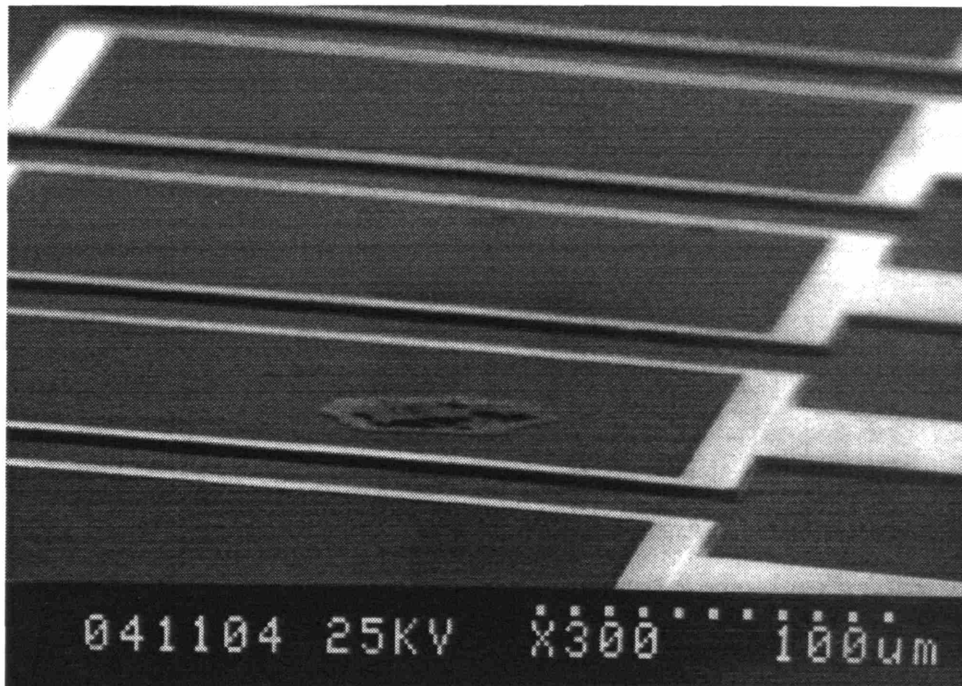


Figure 3.19: SEM of 200µmx200µm QWIP pixels with completed metallization. Large spot is crystal defect.

After completing metallization, processed chips which were typically 1cm x 1cm in size were cleaved into smaller rectangular die ~2mm x 3mm. Individual die were then

mounted into 14pin DIP packages using superglue and bonded using 2mil gold wire and a ball bonder with the stage heated to ~130C. One bond was made to the anode of the die and the remaining thirteen pins were used to bond to up to thirteen individual devices. Where pins were available, bonding was also done to TLM patterns for confirming the performance of ohmic contacts in the packaged devices.

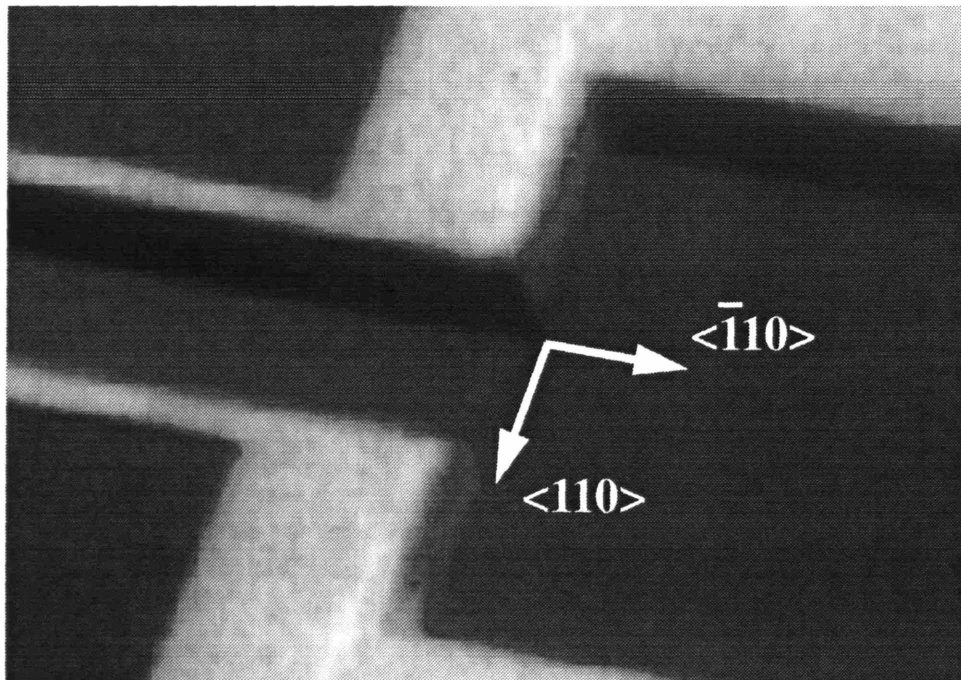


Figure 3.20: Close-up from same SEM micrograph as used in Figure 3.19. Metallization and both etch profiles are visible at the corner of the device.

3.6 Conclusions and future directions in fabrication of QWIPs

The fabrication used for the QWIP devices in this thesis was intentionally kept simple and flexible to maximize the potential for success. In concluding, mention will be made of some of the next generation complexities that can be added to realize higher performance devices more suitable for use in large FPAs.

As described above, the MBE grown QWIP structure is very basic consisting of a set of quantum wells and two ohmic contacts. Significant improvements in this structure can be obtained by putting the quantum well active region within an optical cavity formed of

mirrors above and below the quantum wells. This results in light being trapped in the cavity between the mirrors effectively increasing the quantum efficiency of the QWIP. Possibilities for mirrors include two air/GaAs surfaces formed by thinning the QWIP to a thickness of several microns as is done with the Lockheed/Martin FPAs discussed in Chapter 6, the inclusion of semiconductor superlattices[55] or the inclusion of oxidized AlAs mirrors. All of these techniques improve effective quantum efficiency at the cost of increased MBE growth complexity, often a reasonable trade-off considering that the MBE growth is done over an entire wafer which can produce many QWIP devices.

As described above, the pixel isolation was done with wet etching. This has the advantage of simplicity and good surface morphology, but it is somewhat messy and does not lend itself to large scale manufacture. An alternative to wet etching is reactive ion etching, RIE, which is a plasma etch process using Boron Trichloride and Silicon Tetrachloride. This process has been implemented in Micro-fabrication Shared Experimental Facility at MIT by the author for etching a wide variety of III-V semiconductors and is the logical next step for pixel isolation. Another alternative is oxygen implantation which has been demonstrated by Levine at AT&T[39] and has the advantage of producing a completely planar structure. Implantation is widely used both the Silicon CMOS and Gallium Arsenide MESFET industries, but is not available on campus at MIT.

Finally, AuGe was used as the metallization because of its low contact resistance and relative ease of use. It does however suffer from two problems; morphology tends to be rough after annealing causing difficulties with bonding and extreme care must be taken to prevent the Au from spiking through from the top cathode contact to the anode thus shorting out the device. Other possibilities with better morphology and better repeatability may include Pd/Ge[103] and Ti/Pt[104], both of which are smooth after annealing and are very shallow.

In conclusion, the simplicity of the fabrication structure chosen for this work has been extremely beneficial in demonstrating TE mode QWIP devices. Future work needs to concentrate on maximizing performance of the devices through choice of growth and fab-

rication parameters and on improving reliability, repeatability and uniformity in the fabrication of large numbers of QWIP devices.

Chapter 4

Measurement and Analysis of Normal Incidence QWIPs

In this chapter measurements done on discrete normal incidence TE mode quantum well intersubband photodetectors (QWIPs) will be presented. The goal is to demonstrate that the discrete device performance already achieved meets the criteria laid down in Chapter 2 for fabrication of high resolution large format focal plane arrays (FPAs) operating in the long wave infrared (LWIR) atmospheric transmission band from 8μm to 12μm with minimum resolvable temperatures (MRTs) in the 10mK range. Since discrete devices are being fabricated the detectivity, D^*, is the appropriate figure of merit. However, as discussed in Section 1.4 and shown in Figure 4.1 below, because the application is FPAs, uniformity of the array sets a limit on the range of D^* which lead to substantial increases in MRT. For the applications considered here the limiting value of D^* is ~10^{10} Jones and thus TE mode QWIP with detectivities >10^{10} Jones give excellent results.

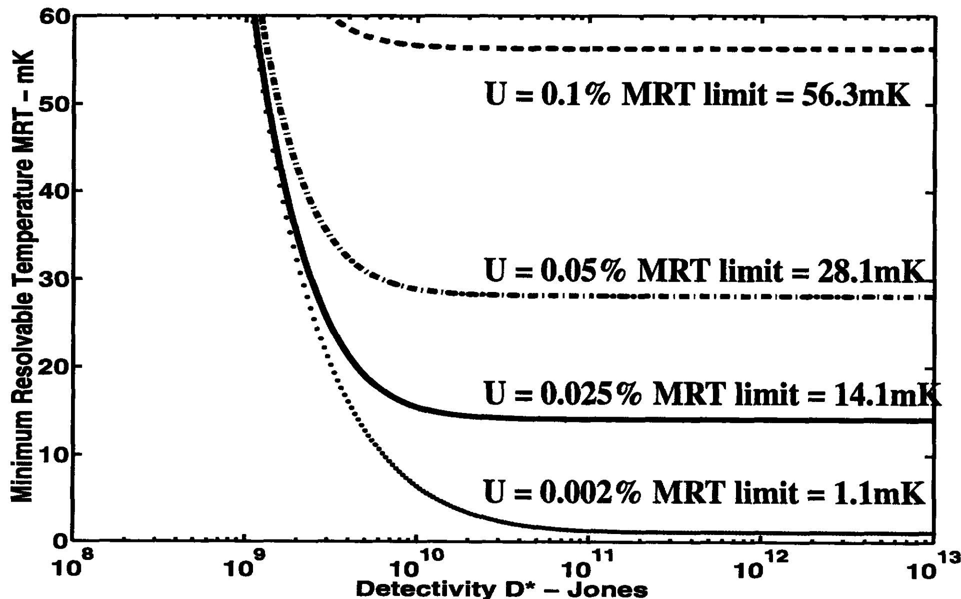


Figure 4.1: Comparison of D* and MRT as figures of merit for FPAs. LWIR band 8μm to 10μm, f/# 1.5 optics with AR coating, pixel area 50μm x 50μm, noise floor 200e⁻, background temperature 300K, g = 0.5.

The devices were grown by the author at MIT using molecular beam epitaxy (MBE) as described in Section 3.2. Reflection high energy electron diffraction (RHEED) and double crystal x-ray (DCXD) were used to determine the material composition and layer thicknesses as described in Sections 3.2 and 3.3.1. The devices were fabricated by the author using the fabrication sequence described in Section 3.5. This chapter begins with a general description of the cryogenic measurement systems common to all QWIP measurements. Section 4.1.1 then describes the dark current measurements done by the author to determine the device cutoff wavelength and background limited performance (BLIP) temperature. Section 4.1.2 describes black body measurements done by the author which determine the responsivity of the discrete QWIPs. Section 4.1.3 describes spectral response measurements done in collaboration with Advanced Device Technology (ADT) that demonstrate the wavelength selectivity and tunability of two of the devices grown by the author and measured at ADT. Section 4.2 puts together all the measured data to determine the D^* for the discrete devices fabricated, compares these results to the theoretical predictions of Chapter 2 and discusses the implications for FPA fabrication. Section 4.3 describes independent corroborating measurements done on similar TE mode QWIP devices fabricated by Alpha Photonics and measured at Wright-Patterson AFB. These results were first published at the same Device Research Conference at which the MIT results were published and resulted in a sharing of the claim for first TE mode QWIP.[106,107] The similarity in the results adds important support to both the explanations presented in this thesis and the claims of feasibility. Finally, Section 4.4 concludes the discussion of measurements done on normal incidence QWIPs and points to possible new avenues for improving individual devices and expanding the set of applications.

4.1 Introduction to QWIP measurement

Some elements are common to all QWIP measurements and it is the focus of this introduction to discuss these common elements, QWIP operating temperature and signal current, prior to getting into the specifics of dark current, responsivity and spectral response. As discussed in Chapter 2, QWIP operating temperatures can be characterized

by the background limited performance (BLIP) temperature defined as the temperature at which the photocurrent at a given background photon flux Φ_B equals the thermally activated dark current through the device. The physics behind the BLIP temperature was discussed in Chapter 2, so the experimental rule of thumb relating the cutoff wavelength, λ_C and the BLIP temperature, T_{BLIP} for QWIPs operating from a bound quantum well state into a three dimensional continuum

$$\lambda_C T_{BLIP} \approx 650\ \mu m\, K \tag{4.1}$$

will be used to note that for QWIPs operating in the center of the LWIR band with cutoff wavelengths about 10μm, T_{BLIP} is about 65K, a temperature most conveniently reached with a closed cycle Helium refrigerator. Furthermore, because the background photon flux Φ_B increases at a rate of one octave per 3K, temperature stability during QWIP measurements is crucial.

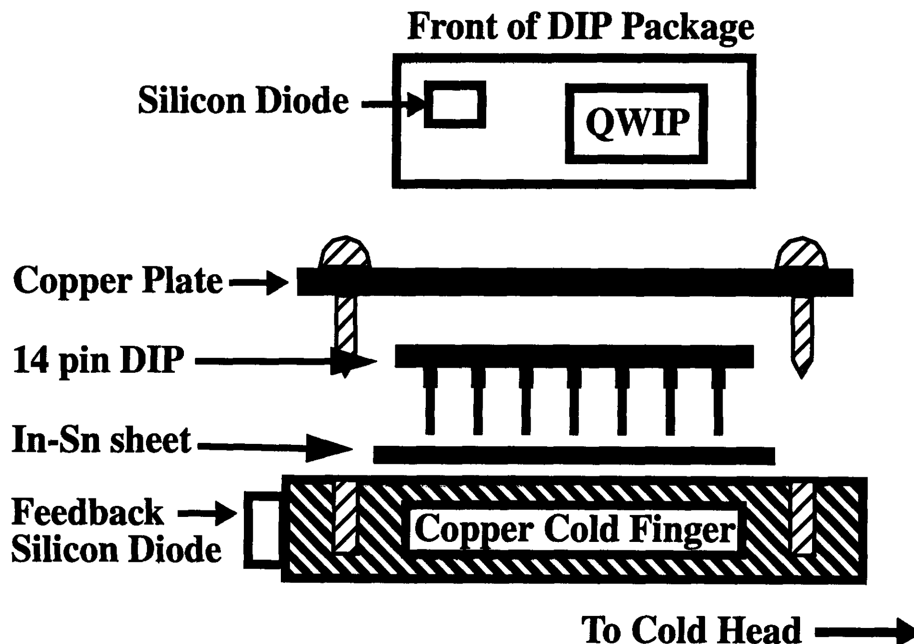


Figure 4.2: QWIP Cryogenic measurement sample mounting system

To meet the requirements of low temperature and good temperature stability, the QWIPs devices used in this thesis were mounted into a high thermal conductivity 14pin

DIP package which in turn was plugged into a mating connector on the copper cold finger of a closed cycle Helium refrigerator. A Silicon diode temperature sensor was attached to the end of the copper cold finger and fed into an RMC cryostat control system.

To improve thermal contact between the DIP package and the copper cold finger a thin Indium-Tin metal sheet was placed between the DIP package and the copper cold finger and the entire sandwich of DIP package, Indium-Tin metal sheet and copper cold finger was held in intimate contact with a 1/8" copper plate that covered the front of the DIP packaged and was screwed into the copper cold finger with two brass bolts. To let in light for responsivity measurements the copper plate had a hole through the center. A schematic of the mounting hardware is shown in Figure 4.2. To insure the accuracy of the temperature readout a second calibrated Silicon diode temperature sensor, a CY7-SD7 from Omega Engineering, was bonded into a 14 pin DIP package in a fashion identical to that used for the QWIP samples. Comparison of the temperature read from the Silicon diode mounted in the DIP package with the Silicon diode mounted on the cold finger show differences of less than 1K demonstrating excellent thermal conductivity between QWIP samples and the copper cold finger.

The total photocurrent through the QWIP photoconductor is very small, typically on the order of 10's of nano-Amps when looking at a black body at 300K. Dark current for temperatures below BLIP drops from this value at a rate of about one octave per 3K. In order to measure such small currents, the author used an HP4145B semiconductor parameter analyzer with all measurements taken using the longest integration time. This resulted in an experimental noise floor in the 10's of pico-Amps range allowing dark current measurements to be taken down to about 40K before falling below the noise floor.

In the remainder of Section 4.1 the specifics of measurements to characterize the QWIPs devices will be discussed. Section 4.1.1 discusses the measurement of dark current, Section 4.1.2 discusses the measurement of black body responsivity and Section 4.1.3 discusses the measurement of spectral response.

4.1.1 Measurement of dark current

Dark current is the current through the QWIP photoconductor when not exposed to incident photons. As discussed in Section 1.6.4, if the interwell barriers in the QWIP superlattice are sufficiently thick to eliminate tunneling through the barrier at the operating bias, the major source of dark current is thermal excitation of electrons out of the quantum wells. This current flows above the barriers and is affected by trapping in the quantum wells and thermionic emission out of the quantum wells which balance to keep the electron density in the quantum wells constant in time as illustrated in Figure 4.3. Note that since there is no incident photon flux, the dark current only depends on the physical characteristics of the QWIP, the bias voltage and the device temperature.

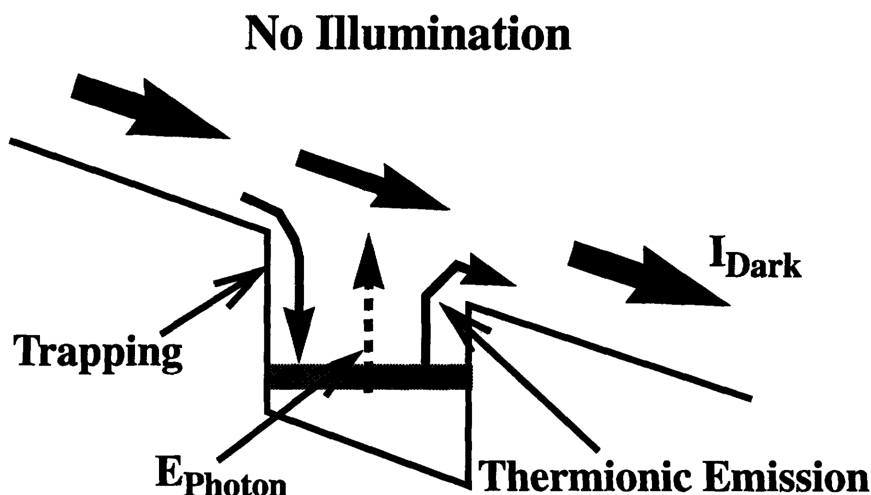


Figure 4.3: Dark current path through a QWIP without illumination. Quantum well trapping and thermionic emission combine to maintain a constant electron density in the well.

The dark current density is given by the expression:

$$J = n_{3d} e v(\varepsilon) \qquad \frac{Amps}{cm^2} \tag{4.2}$$

where $v(\varepsilon)$ is the electron velocity at the electric field ε and n_{3d} is the electron density in the continuum states above the top of the quantum well. Because the barriers are thick,

typically ~500Å, relative to the size of the quantum wells n_{3d} can be taken as the density of electrons in the bulk barrier material and is given by the expression:

$$n_{3d} = \frac{1}{4}\left(\frac{8\pi m^* kT}{h^2}\right)^{3/2} \exp(-E_a) \qquad \frac{electrons}{cm^3} \tag{4.3}$$

where E_a is the activation energy required for an electron to escape from the quantum well into the continuum and is given by the difference between the energy at the top of the well and the Fermi energy in the well, $E_a = E_{photon} - E_F$, m^* is the effective mass of the electron in the bulk barrier material, T is the temperature of the QWIP and the other constants have their conventional definitions. The electron velocity including the effects of velocity saturation is approximated using the expression:

$$v(\varepsilon) = \frac{\mu\varepsilon}{\sqrt{1 + \left(\frac{\mu\varepsilon}{v_{sat}}\right)^2}} \qquad \frac{cm}{second} \tag{4.4}$$

where μ is the low field mobility, ε is the electric field and v_{sat} is the saturation velocity of the electron with a value of about 10^7cm/s in undoped GaAs.

The dark current as a function of bias voltage was measured for the QWIPs grown by the author using the measurement setup of Figure 4.2 with the addition of an Aluminum foil cover between the top of the DIP package and the copper plate. The Aluminum foil cover prevented any radiation from illuminating the QWIP from outside the cryostat and in addition was kept at same temperature as the QWIP thus reducing the black body emission from the Aluminum foil cover to negligible levels. The dark current curve for sample 9066 is shown in Figure 4.4. This sample contained 10 quantum wells of width 54Å with composition $In_{0.08}Ga_{0.92}As$ and doped n-type at 5×10^{17} cm^{-3} separated by barriers of width 450Å and composition $Al_{0.15}Ga_{0.85}As$. Measurement floor is ~100pA.

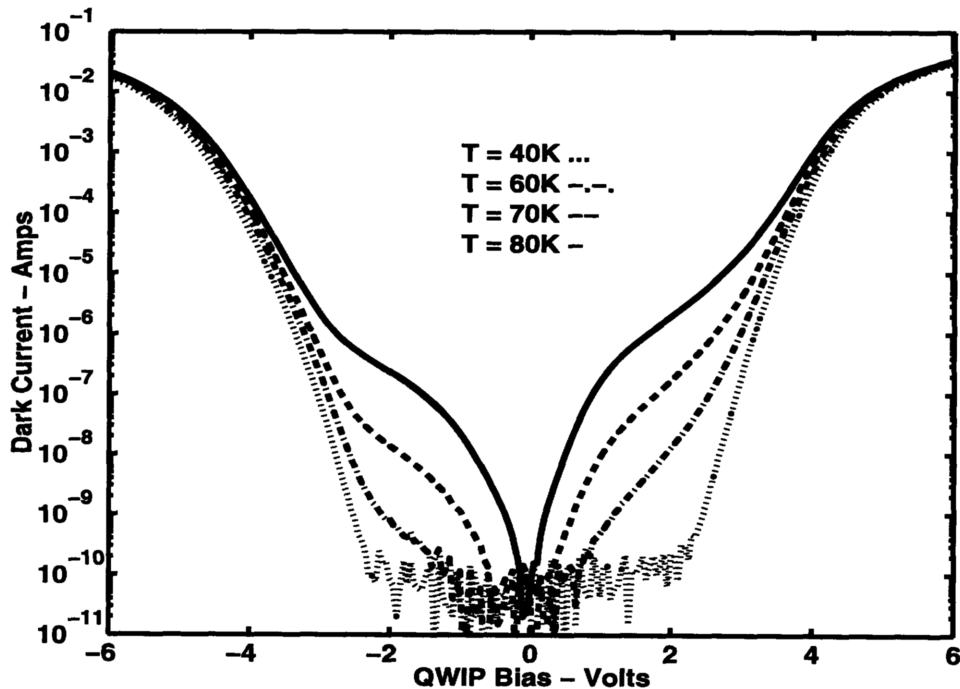


Figure 4.4: Dark Current for QWIP sample 9066. 10 $In_{0.08}Ga_{0.92}As$ wells of width 54Å doped n-type $5x10^{17}$ cm^{-3} with 450Å barriers of $Al_{0.15}Ga_{0.85}As$.

The region up to the first knee is the sweep-out region in which the current of thermally generated carriers collected at the contacts is exponentially dependent on the field and rises very rapidly up to the point where electrons thermally excited from a quantum well have negligible chance of being reabsorbed by the same quantum well. The dark current then enters the thermal excitation rate limited regime which as shown in Equation 4.2 is proportional to the electron velocity. If thermionic emission were the only mechanism contributing to the dark current in the sweep-out region, the curves of Figure 4.4 would be flat between the two knees, however, thermionic assisted tunneling gives electrons with not quite enough energy to escape the quantum well classically a probability of tunneling out the quantum well thus contributing to a rise in the dark current.[52] The sharp rise in the dark current beyond about 3V, corresponding to $6x10^5$ V/cm, is due to the onset of field induced tunneling though the barriers and puts a limit on the usable range of bias voltages. QWIPs are best operated in the center region between the two knees since

photo-generated carriers are swept out of the quantum well active region and tunneling current which adds noise but no photo-signal is negligible.

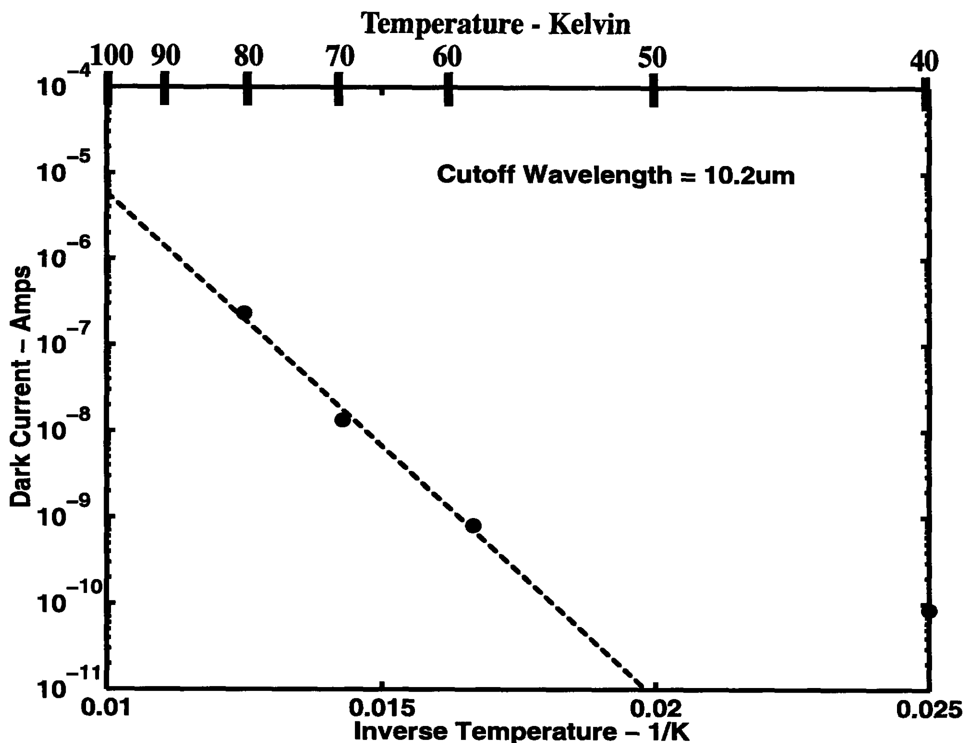


Figure 4.5: Dark Current for QWIP sample 9066. 10 $In_{0.08}Ga_{0.92}As$ wells of width 54Å doped n-type $5x10^{17}$ cm^{-3} with 450Å barriers of $Al_{0.15}Ga_{0.85}As$. The lowest current point at 40K shows the noise floor of the measurement.

The asymmetry in the dark current curve of Figure 4.4 is likely due to the MBE growth kinetics which cause dopants to diffuse out of the well into the barrier in the growth direction[108] and with the addition of strain in the quantum well system leads to diffusion of Indium out of the well in the growth direction. Both effects tend to lower the quantum well barrier height in the direction of the growth leading to increased currents in the positive voltage direction which corresponds to positive bias on the top, last to be grown, contact of the QWIP device.

The cutoff wavelength of the QWIP photodetector can also be obtained from the exponential dependence of the dark current on the activation energy, E_a, at constant bias as given by Equation 4.3. Figure 4.5 shows a semilog plot of dark current verses inverse temperature for sample 9066 at a bias of -2V. The slope of the curve gives an activation

energy of 112meV which when combined with a Fermi energy of 10meV gives a cutoff wavelength of 10.2μm in reasonable agreement with the designed wavelength of 9.5μm for this QWIP photoconductor. Table 4.1 shows a summary of dark current data for the samples used in this thesis and measured by the author at MIT.

Sample	N_{2d} (cm^{-2})	N_{well}	In_{well} (%)	$Al_{barrier}$ (%)	W_{well} (Å)	$W_{barrier}$ (Å)	E_{Fermi} (meV)	λ_{cutoff} (μm)
9034	$3x10^{11}$	15	7	15	60	500	10	11.4
9062	$5x10^{11}$	2	10	12	50	500	18	10.5
9066	$2x10^{11}$	10	8	15	45	450	7	10.2
9104	$1.5x10^{11}$	10	9	15	50	450	5	11.3

Table 4.1: Dark Current Characteristics for QWIP samples

4.1.2 Measurement of responsivity

The responsivity in a QWIP photoconductor is the ratio of the photo-current generated by the device to the optical power incident on the device. Responsivity is given by:

$$R_{\lambda} = \frac{e\eta g}{h\nu} \qquad \frac{Amps}{Watt} \qquad (4.5)$$

where η is the quantum efficiency of the QWIP photoconductor, e is the electron charge and g is the photoconductive gain. The quantum efficiency is proportional to the number of quantum wells, N, and is given by:

$$\eta = N\eta_1 = 100(1 - \exp(-\alpha L_{optical})) \qquad \% \qquad (4.6)$$

where η_1 is the quantum efficiency for a single quantum well, α is the absorption coefficient of the QWIP active region and $L_{optical}$ is the optical path length through the active region. The photoconductive gain was derived in Chapter 2 and is given by:

$$g = \frac{(1-p)}{Np} \qquad \frac{electrons\ collected}{electrons\ generated} \qquad (4.7)$$

where p is the capture probability for an excited electron traversing a quantum well.

Figure 4.6 shows a schematic of the dominant processes in a QWIP photoconductor under bias. Note that in order to keep a constant electron density in the quantum well photo-emitted electrons must be replaced by extra trapping from the stream of excited electrons resulting in a photoconductive gain given by Equation 4.7.

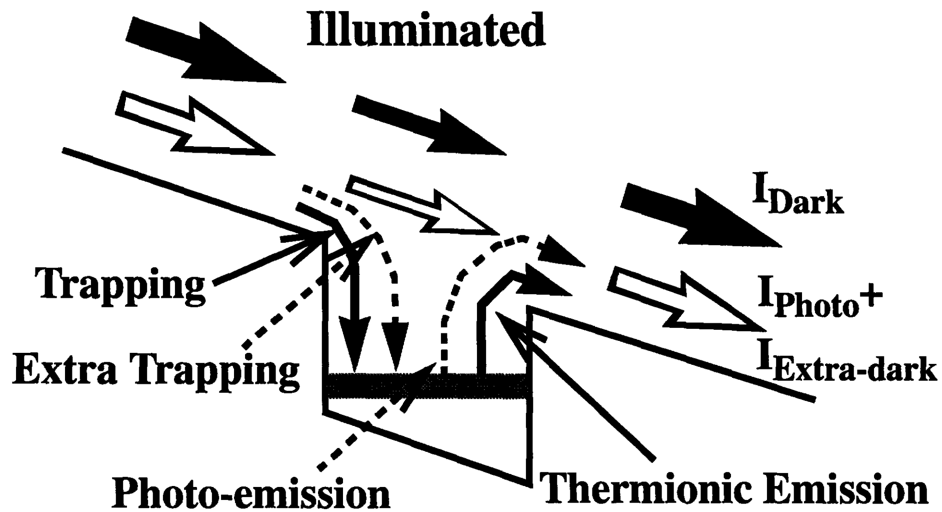


Figure 4.6: Photocurrent path through a QWIP with illumination. The Dark current remains the same as in Figure 4.1 however with the addition of photo emission extra injection is required to maintain a constant electron density in the well.

The responsivity of the TE mode QWIP photoconductors used in this thesis was measured with the setup of Figure 4.2 with the addition of a ZnSe window through which the QWIP was exposed to IR light from a black body at a controlled temperature. For measurements done at room temperature, a black cloth sock covered the front of the ZnSe window to approximate a black body at 293K as shown in Figure 4.7. For measurements done at 500K, a commercial black body was used which was kept at a controlled 500K and illuminated the QWIP photoconductor through a circular stop as shown in Figure 4.8.

The photocurrent as a function of bias voltage for sample 9066 at two black body temperatures, 293K and 500K is shown in Figure 4.9. This sample contained 10 quantum wells of width 54Å with composition $In_{0.08}Ga_{0.92}As$ and doped n-type at $5x10^{17}$ cm^{-3} separated by barriers of width 450Å and composition $Al_{0.15}Ga_{0.85}As$.

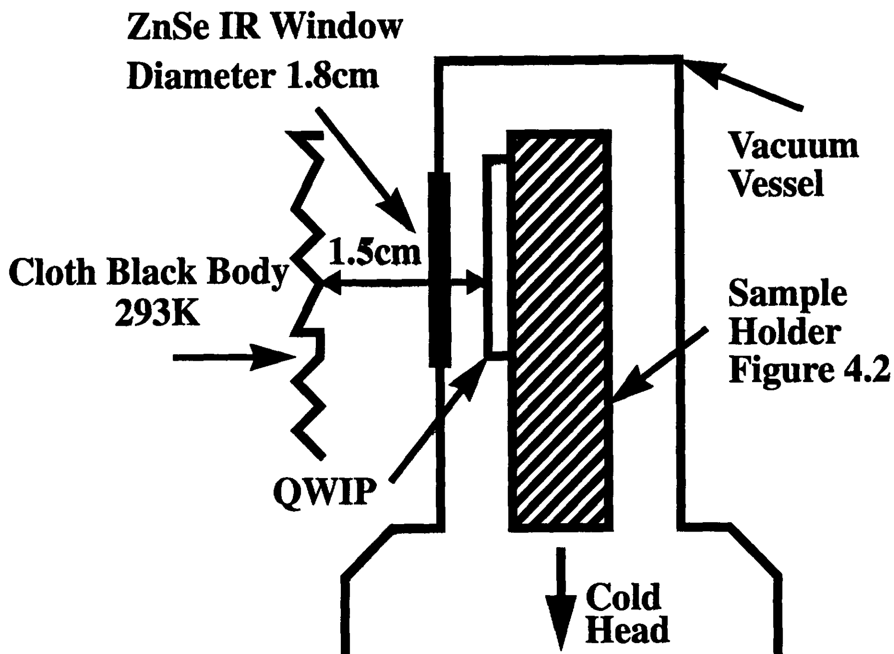


Figure 4.7: Measurement setup for 293K black body. For 100μm square QWIP solid angle, $\Omega = 4.4 \times 10^{-5}$ str. Photon flux at 10μm for FWHM = 1.5μm, $\Phi_{BB} = 6.7 \times 10^{16}$ photons/cm^2 s str. Incident power 1.5×10^{-7} Watts.

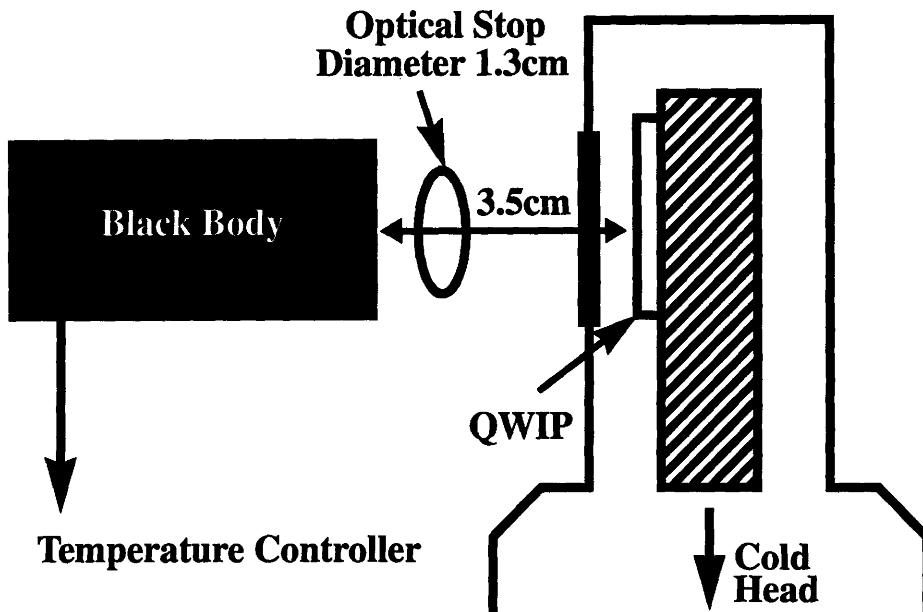


Figure 4.8: Measurement setup for 500K black body. For 100μm square QWIP solid angle, $\Omega = 8.2 \times 10^{-6}$ str. Photon flux at 10μm for FWHM = 1.5μm, $\Phi_{BB} = 5.4 \times 10^{17}$ photons/cm^2 s str. Total incident power 2.5×10^{-7} Watts.

Device size was approximately 100μm x 100μm. Measurements of photocurrent were done with the device held at 30K at which temperature the discussion of Section 4.1.1 showed the dark current to be below the noise floor of the experiment. This guaranteed that all the current was photon related.

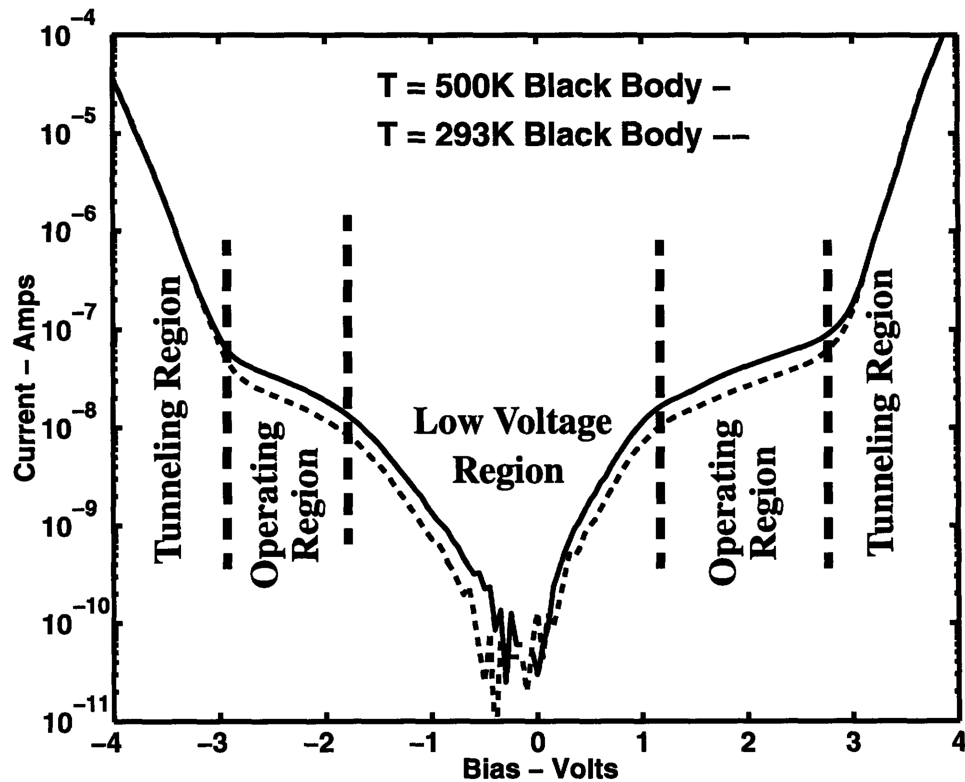


Figure 4.9: Photocurrent for sample 9066 illuminated with black body at 500K, solid line and illuminated with 293K black body, dashed line. Bold dashed vertcal lines delineate the knees in the curve which define the tunneling dominated, photocurrent dominated (operating region) and low voltage regions of the photocurrent curves.

To convert photocurrent to responsivity the incident power must be calculated from the geometry of the measurement system and the black body emission curves of Chapter1, Equation 1.3. The geometry used with the 293K black body is shown in Figure 4.7. For a 100μm x 100μm detector with a peak responsivity at 10μm and a FWHM of 1.5μm typical of bound to continuum type QWIPs and confirmed by spectral response measurements described in Section 4.1.3, the result for the power incident on the QWIP photoconductor from the 293K black body is 1.5 x 10^{-7} Watts. The geometry used with the 500K black

body is shown in Figure 4.8. For the same FWHM of 1.5μm, peak responsivity at 10μm and 100μm x 100μm area the total power incident is 2.5 x 10^{-7} Watts including both the power from the 500K black body and the power from the uncooled region around the optical stop which is at the ambient temperature, 293K.

If photoemission were the only source of electrons the curves of Figure 4.9 would be flat in the photocurrent dominated operating region between the two knees where all the photo-excited electrons get swept away from their original quantum well. However, since we are using a black body to illuminate the QWIP photoconductor photo-assisted tunneling due to lower energy photons also occurs leading to a slight slope in the photocurrent dominated region. To eliminate this effect, the value of the photocurrent used for calculating responsivity is taken just beyond the knee separating the low voltage region from the photocurrent dominated operating region. For sample 9066 shown in Figure 4.9 this gives values at a -2 Volt bias of 80 mA/W and 76 mA/W for the 300K and 500K black bodies respectively. At a 1.25 Volt bias responsivities are 80mA /W and 72 mA/W for the 293K and 500K black bodies respectively. Because the responsivity of the devices is not expected to change with illumination intensity, the agreement to within 10% between the responsivities for the 293K and 500K black bodies is a check on the calculations for the relative power incident on the QWIP photodetectors for the two setups.

Having calculated the responsivity, Equation 4.5 can then be rearranged to give for the quantum efficiency per quantum well, η_1:

$$\eta_1 = 100 \times \frac{h\nu R}{Neg} \qquad \frac{\%}{well} \tag{4.8}$$

where N is the number of wells, e is the electron charge and g is the photoconductive gain given by Equation 4.7. Note that since g is inversely proportional to the number of quantum wells, both η_1 and R are independent of the number of quantum wells. For sample 9066 taking the responsivity to be 80 mA/W at 10μm and assuming a probability of capture for excited electrons traversing a quantum well of 7% as discussed Chapter 2, the single well quantum efficiency, η_1, is 0.069% in excellent agreement with the theoretical calculations of Flatte[46] described in Section 2.3 which result in η_1 = 0.068%. The

agreement between theory and measurement for this sample is so good as to beg the question of exactly how accurate the two numbers are. Table 4.2 shows results from several devices of various sizes taken from different samples. The standard deviation in single well quantum efficiencies is about 50% which is a rough estimate of the accuracy of the experimental number. The error in the theoretical quantum efficiency is more difficult to estimate because it depends on overlaps between states high in the conduction band. However, assuming the error here is also about 50%[109] gives agreement between theory and experiment to within a factor of two in about 60% of the samples, quite reasonable considering the small size of the numbers involved.

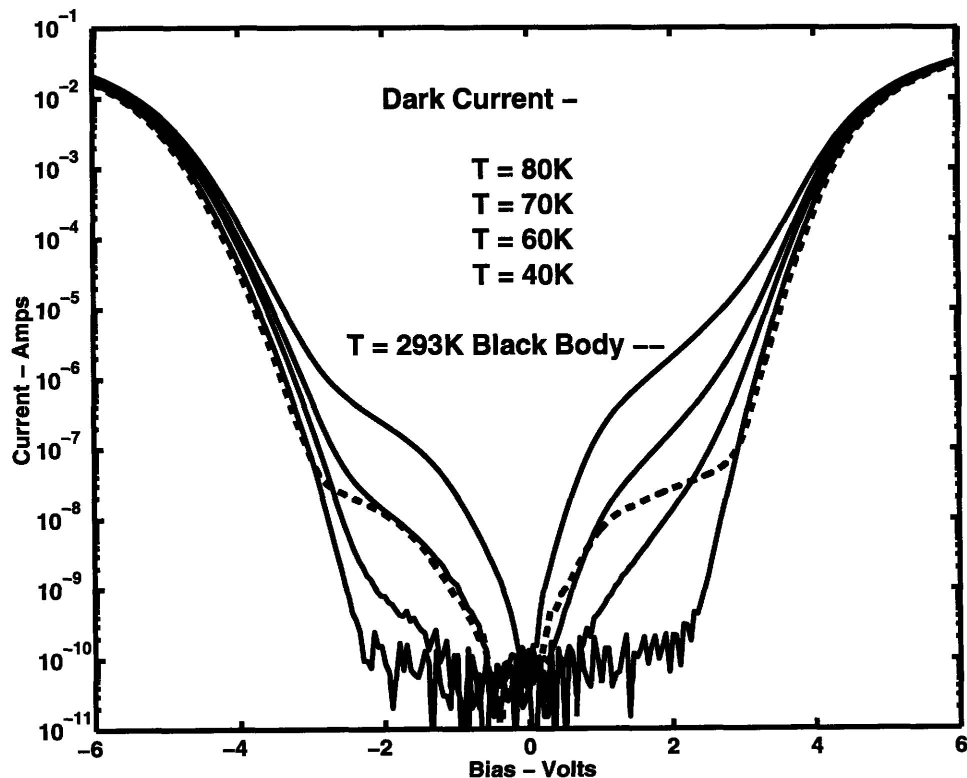


Figure 4.10: Dark current, solid lines from top down at temperatures of 80K, 70K, 60K and 40K. Photo-current, dashed line for black body of 293K. BLIP temperature estimated to be 68K.

Combining the dark current curves of Figure 4.4 with the photocurrent curve of Figure 4.9 for the black body at 293K gives an experimental measure of the background limited performance (BLIP) temperature for the QWIP photoconductor.

nple	N_{well}	N_{2d} (cm^{-2})	In_{well} (%)	$Al_{barrier}$ (%)	W_{well} (Å)	$W_{barrier}$ (Å)	Area ($x10^{-4}$ cm^2)	λ_{cutoff} (µm)	R (mA/W)	T_{BLIP} (K)	η_1 (%)	D^*_{BLIP} (Jones)
34	15	$3x10^{11}$	10	15	60	500	1	11.4	13	61	0.011	$9.7x10^9$
34	15	$3x10^{11}$	10	15	60	500	4	11.4	18	61	0.016	$1.1x10^{10}$
62	2	$5x10^{11}$	10	12	50	500	1	10.5	33	58	0.022	$5.7x10^9$
62	2	$5x10^{11}$	10	12	50	500	4	10.5	25	58	0.029	$5.0x10^9$
66	10	$2.5x10^{11}$	10	15	45	500	1	10.2	80	68	0.069	$2.0x10^{10}$
66	10	$2.5x10^{11}$	10	15	45	500	4	10.2	47	68	0.041	$1.2x10^{10}$
66	10	$2.5x10^{11}$	10	15	45	500	1	10.2	67	68	0.058	$1.8x10^{10}$
04	10	$1.5x10^{11}$	10	15	50	500	4	11.3	30	65	0.026	$1.2x10^{10}$
04	10	$1.5x10^{11}$	10	15	50	500	1	11.3	53	65	0.046	$1.6x10^{10}$
04	10	$1.5x10^{11}$	10	15	50	500	4	11.3	42	65	0.036	$1.4x10^{10}$
56[1]	4	$2.5x10^{11}$	15	20	50	500	1	9.6			0.023	$2.0x10^{9*}$
65[1]	5	$2.5x10^{11}$	15	30	40	500	1	7.5	25		0.052	$6.0x10^{9*}$
ha[2]	50	$8x10^{11}$	30	0	40	300	1	12	1500	45	0.160	$2.0x10^{10}$
age[3]									40.8		0.035	$1.2x10^{10}$
ma[3]									21.4		0.019	$3.7x10^9$

Table 4.2: Electrical and optical properties of QWIPs.

e 1: Samples 9056 and 9065 measured by Peter Kannam at ADT. D* for these samples measured at 77K.
e 2: Sample “Alpha” taken from WL measured data courtesy Robert Shih of Alpha Photonics Inc [105]

Experimentally, the BLIP temperature is simply the temperature at which the dark current equals the photocurrent for a black body at the background temperature, taken by convention as 293K. Figure 4.10 shows the dark current and photo-current curves for sample 9066 plotted on top of each other. Looking at the negative voltage side of the curve because the positive voltage side has larger dark current, the BLIP temperature for this device is estimated to be 68K. Since from Figure 4.5 a cutoff wavelength of 10.2μm was estimated from the dark current, this is consistent with rule of thumb given in Equation 4.1 that the cutoff wavelength multiplied by the BLIP temperature is approximately 650 μmK for the bound to continuum type QWIP photoconductors studied here.

4.1.3 Measurement of spectral response

Spectral measurements were done in collaboration with Peter Kannam at ADT to determine the wavelength dependence of the responsivity. The experimental setup is shown in Figure 4.11. The sample is held at 77K using a liquid nitrogen dewar and illuminated with the output from a spectrometer which when scanned gives the QWIP responsivity as a function of wavelength.

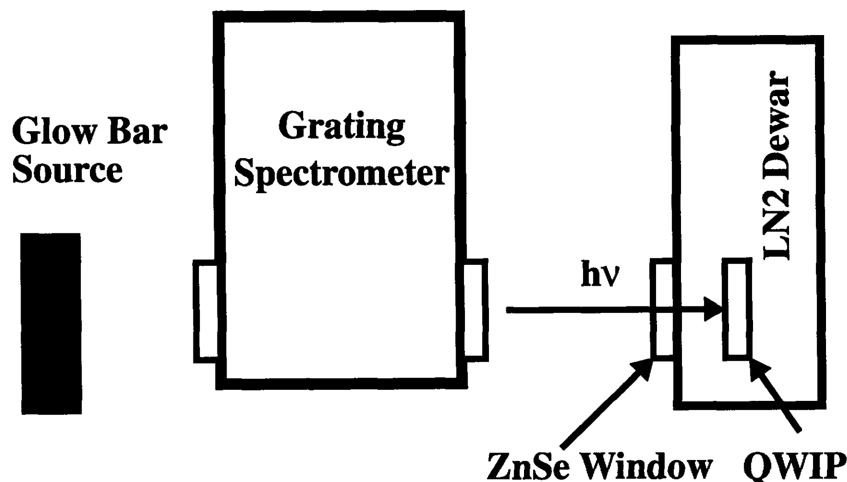


Figure 4.11: Spectral Response Measurement setup at ADT

Results of the relative spectral response for sample 9056 is shown in Figure 4.12 as dots connected by a solid line. This device consisted of 4 50Å quantum wells of composition $In_{0.15}GaAs$ and doping density $5x10^{17}$ cm^{-3} separated by 500Å barriers of composition $Al_{0.20}GaAs$. The cutoff wavelength is shown to be 9.6μm with a peak responsivity at 8.7μm and a FWHM of 1μm. The relative spectral response for sample 9065 is also shown in Figure 4.12 but as dots connected by a dashed line. This device consisted of 5 40Å quantum wells of composition $In_{0.15}Ga_{0.85}As$ and doping density $5x10^{17}$ cm^{-3} separated by 500Å barriers of composition $Al_{0.30}Ga_{0.70}As$. The cutoff wavelength is shown to be 7.5μm with a peak responsivity at 6.5μm and a FWHM of 1.5μm. Both devices have nice narrow peaks characteristic of intersubband transitions and ideal for applications that require wavelength selectivity.

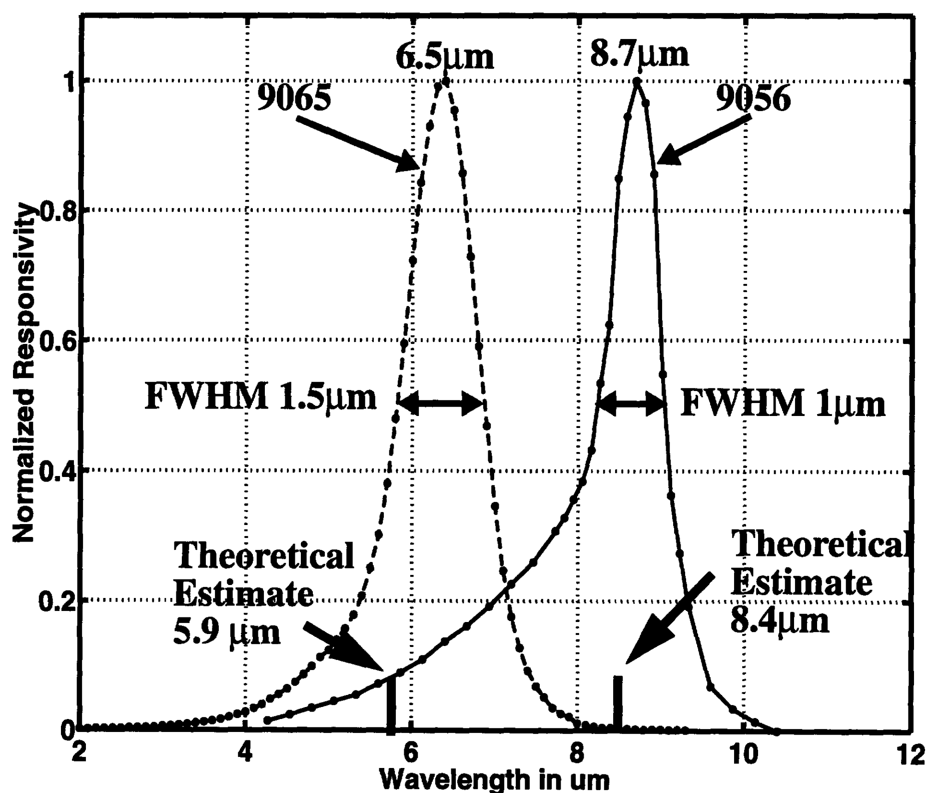


Figure 4.12: Sample 9065, dots connected by dashed line and Sample 9056, dots connected by solid line. 9065 contains 5 50Å $In_{0.15}Ga_{0.85}As$ wells and 500Å $Al_{0.30}Ga_{0.70}As$ barriers, target 5.9μm. 9056 contains 4 50Å $In_{0.15}Ga_{0.85}As$ wells and 500Å $Al_{0.20}Ga_{0.80}As$ barriers, target 8.4μm.

The difference in the peak absorption wavelengths for the two devices is due to the different well widths and barrier compositions and demonstrates one of the strengths of III-V based QWIP devices, namely the ability to tune the peak of the response based on changes in the well widths and compositions.

4.2 Calculation of D* from measured parameters

Having measured dark current and photocurrent it is now possible to calculate the detectivity of the TE mode QWIP photoconductor under BLIP conditions and use this number in combination with Figure 4.1 to predict the performance of an FPA of TE mode QWIPs. The black body detectivity was discussed in Section 1.4.3 and is given in terms of the black body responsivity as:

$$D^*_{BB} = \frac{R_{BB}\sqrt{A\Delta f}}{\langle i_n \rangle} \qquad Jones \tag{4.9}$$

where R_{BB} is the black body responsivity in Amps/Watt, A is the area of the detector, Δf is the frequency range of the measurement and i_n is the noise current. Under BLIP conditions the noise current is due to generation-recombination noise in the photocurrent generated by the background photon flux and is given by:

$$\langle i_n \rangle = ge\sqrt{2\eta\Phi_B A\Delta f} \tag{4.10}$$

where g is the photoconductive gain, η is the quantum efficiency and Φ_B is the background photon flux. Combining Equation 4.10 with Equation 4.5 for R_{BB} gives:

$$D^*_{BB} = \frac{1}{h\nu}\sqrt{\frac{\eta}{2\Phi_B}} \qquad Jones \tag{4.11}$$

Notice that the detectivity is independent of photoconductive gain, g, and therefore independent of the electron transport properties of the device. Detectivity is however proportional to the square root of the number of quantum wells since the total quantum efficiency of the device is proportional to the number of quantum wells. This dependency

comes not from the responsivity which is independent of the number of quantum wells but from the noise in the background photocurrent.

Using Equation 4.11 with Φ_B taken from the geometry of Figure 4.7 the BLIP detectivity for each of the QWIP photodetectors measured at MIT has been calculated in the last column of Table 4.2. The best QWIP has a BLIP detectivity of 2×10^{10} Jones and the mean for the 10 devices measured is 1.2×10^{10} Jones with a standard deviation of 3.7×10^{9} Jones. These values nicely meet the criteria set forth for use in high quality FPA and as shown in Figure 4.13 for uniformity set by the 12bit ADC used in calibration will result in focal plane arrays with excellent minimum resolvable temperatures less than 10mK for detection in the LWIR band.

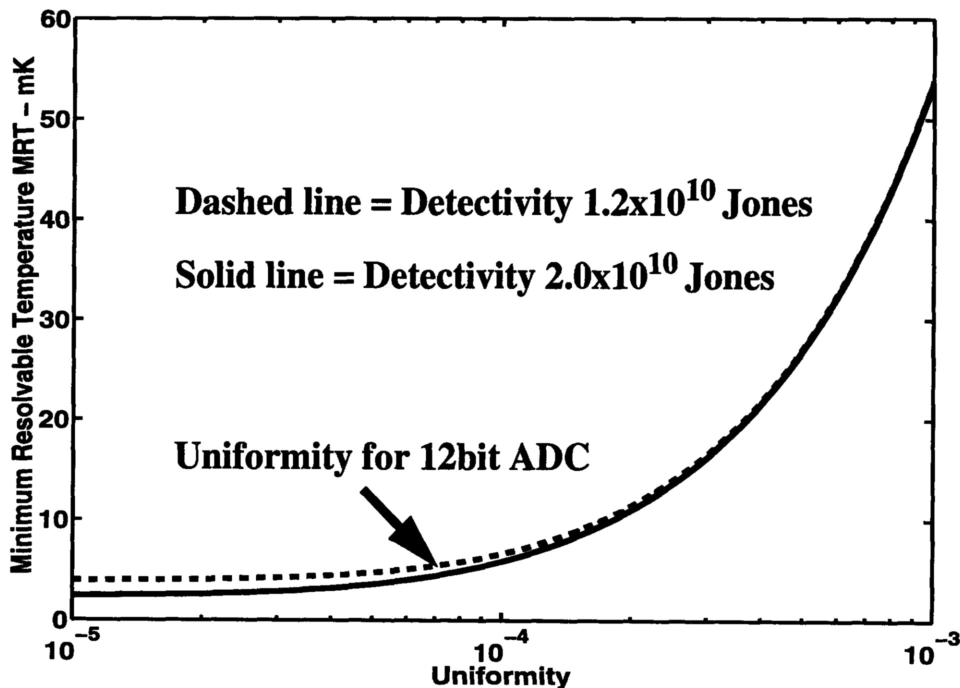


Figure 4.13: Comparison of D* for best and mean TE mode QWIPs. LWIR band 8μm to 10μm, f/# 1.2 optics with AR coating, noise floor 200e$^-$, background temperature 293K, $g = 0.5$. Uniformity for 12bit ADC of 0.007% indicated by arrow.

This concludes the discussion of measurements done on TE mode QWIPs fabricated and tested at MIT. The results are quite encouraging and seem to indicate both a good theoretical understanding of the dominant processes involved in the detection of normal inci-

dence radiation and a device performance sufficient to make excellent quality FPAs without the fabrication complexity involved in using coupling gratings. The following section discusses corroborating measurements done on TE mode QWIPs with a similar structure designed and fabricated by Alpha Photonics. This work was done simultaneous to the work at MIT but the two groups were unknown to each other until the first results from both groups were published at the Device Research Conference in June, 1995.

4.3 Corroborating Measurements done at Alpha Photonics

Simultaneous to the work done at MIT, Alpha Photonics Incorporated (API) has done fabrication of TE mode normal incidence QWIP photoconductors using InGaAs quantum wells and GaAs barriers.[106,110,105] In this section the work of API will be compared and contrasted with the work done at MIT. A piece of the API sample was also fabricated into TE mode QWIPs at Wright Laboratory (WL), Wright-Patterson AFB, and some of these results will be used as well in discussion the API sample.[105] Though there are some important differences in the results, these differences can be explained based on differences in material choices and device structures. The resulting similarity in the detectivities which is insensitive to these differences, $D^*_{BLIP} = 2.0\times10^{10}$ for Alpha Photonics and $D^*_{BLIP} = 1.2\times10^{10}$ for the mean of the MIT devices, lends strength to the claim in this thesis that the TE mode QWIP photoconductors have sufficient individual device performance to be useful in focal plane arrays.

The API QWIP devices consist of fifty 40Å $In_{0.30}Ga_{0.70}As$ quantum wells doped n-type at 2×10^{18} cm^{-3} and separated by 300Å barriers of undoped GaAs. Compared to the MIT devices which were based on InGaAs quantum wells with AlGaAs barriers, the API devices have no Aluminum in the barriers and to keep a similar conduction band offset, which along with the quantum well width determines the wavelength of sensitivity, the Indium fraction in the quantum wells is increased to 30%. InGaAs has a different lattice constant than GaAs and as discussed in Section 3.4 there are limits on how thick a strained layer can be grown before the onset of dislocation glide. At 30% Indium the critical thick-

ness is about 50Å as shown in Figure 3.5 therefore a single well of 40Å is thermodynamically stable with respect to dislocation glide.

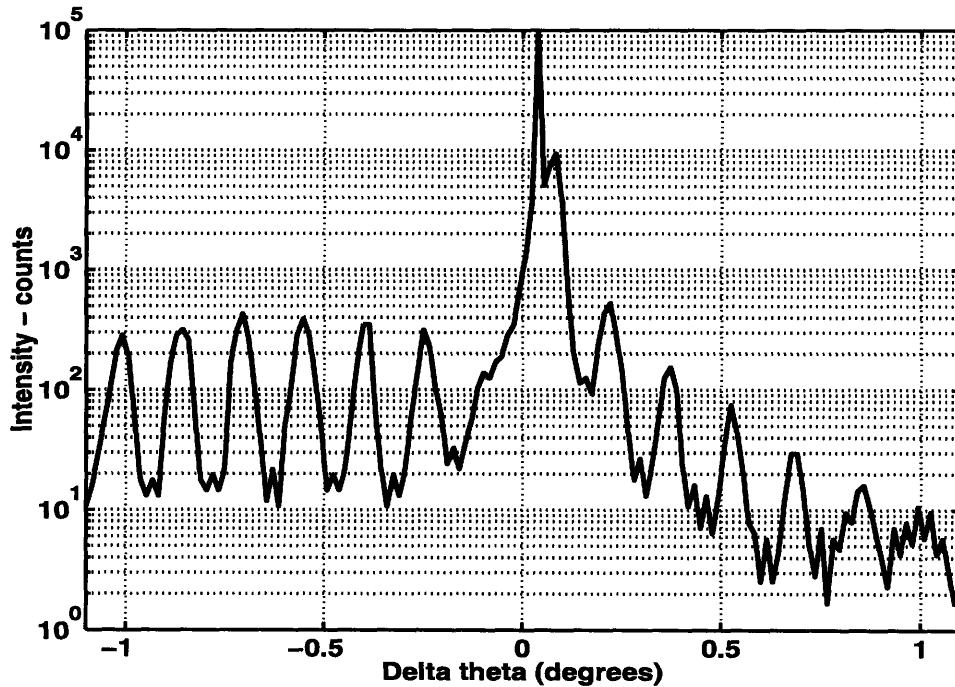


Figure 4.14: Wright Laboratory measured X-ray data for API Sample.[105]

However, if the superlattice consisting of 40Å $In_{0.30}Ga_{0.70}As$ quantum wells separated by 300Å of GaAs is repeated fifty times as is done in the API sample, it is clear from Figure 3.6 that the entire superlattice structure is not stable. This implies that in spite of the excellent double crystal x-ray results shown below in Figure 4.14 the API sample clearly has a large number of misfit dislocations. The interesting point that has not been made before is that these dislocations seem to have no deleterious effect on the QWIP device performance, in particular the dark current. One can speculate that this is due to the in-plane nature of the misfit dislocations which therefore may not contribute to excess dark current and/or the unipolar nature of the QWIP devices that inhibits electron-hole recombination at the misfits, but clearly this is an area that needs more research in the future.

The dark photocurrent for the API devices is shown in Figure 4.15. This data was taken with the device looking at a 300K black body similar to the geometry of Figure 4.7

used at MIT. The curves appear slightly less asymmetric than those for the MIT grown devices indicating possibly a lower growth temperature which would limit the diffusion of impurities and Indium out of the quantum well region and also possibly better control over the quantum well interface region. BLIP temperature can be estimated by noting that all the curves below 50K lie on top of each other indicating background limited operation and a BLIP temperature estimated at 50K. This is reasonably consistent with the rule of thumb given in Equation 4.1 for a cutoff wavelength of 12μm.

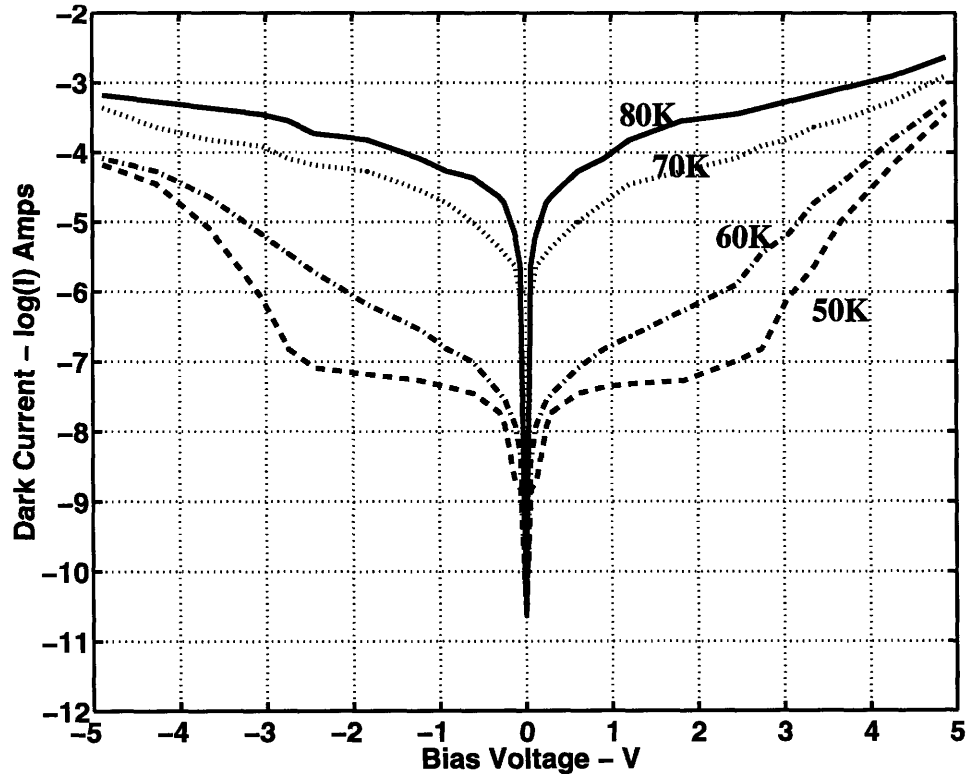


Figure 4.15: API measured photo current for 200μm diameter device. Top down: solid line 80K, dotted line 70K, dash-dot line 60K dashed line 50K.

The spectral responsivity of the API sample is shown in Figure 4.16 as measured by WL. The cutoff wavelength is 11μm with a peak near 9μm and a FWHM of about 3μm. The data agrees reasonably well with the FTIR measured peak in the absorption at a wavelength of 9.7μm. The FWHM is about a factor of two larger than that measured for the MIT samples indicative of the position of the n=2 resonance state which is further above

the edge of the quantum well in the API sample. The peak responsivity measured by API is 600mA/W using a 300K black body illuminated at 45 degrees to the sample edge, but due to the measurement geometry this number is expected to be less than the actual value. For a piece of the same growth processed and measured at WL in normal incidence as done at MIT, the responsivity is 1500mA/W using a 300K black body.

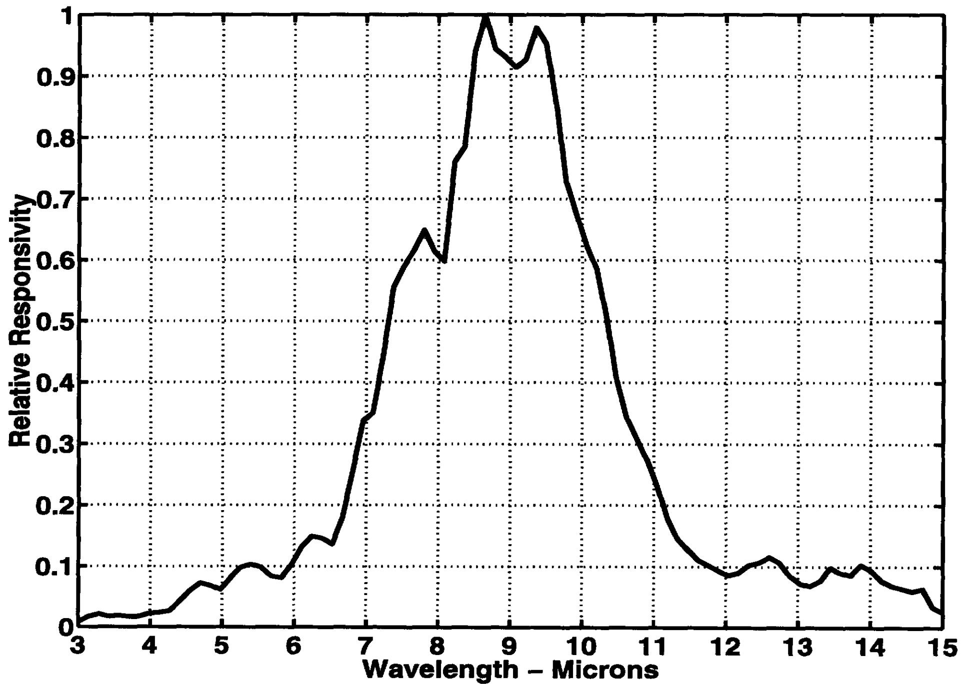


Figure 4.16: WL measured spectral responsivity for API device.[105]

The responsivity for this device is more than an order of magnitude larger than the responsivity for the devices measured at MIT, however there are two crucial differences between the API and MIT samples that explain the difference. First, note from the discussion of Section 2.1 that the photoconductive gain, g, is proportional to the ratio of the excited electron lifetime, τ_{life}, to the transit time through the device, $\tau_{transit}$, and the transit time of the device is inversely proportional to the electron velocity giving for g:

$$g = \frac{\tau_{life}}{\tau_{transit}} = \frac{v_e(\varepsilon)}{NL_P}\tau_{life} \qquad \frac{electrons\ collected}{electrons\ generated} \tag{4.12}$$

where $v_e(\varepsilon)$ is the electron velocity as a function of electric field given by Equation 4.4 and L_P is the length of one period in the superlattice active region of the device and N is the number of quantum wells in the device.

The important point is that the barriers in the API device are GaAs which has a mobility about five times larger than the mobility in the $Al_{0.15}Ga_{0.85}As$ barriers used in the MIT samples.[111] This leads to a 5x increase in the photoconductive gain and a 5x increase in the responsivity. Actual measurements comparing QWIPs with GaAs barriers to similar devices with AlGaAs barriers show responsivities for the QWIPs with GaAs barriers about 10x larger compared to QWIPs with $Al_{0.25}Ga_{0.75}As$ barriers so the 5x increase in responsivity predicted from the bulk mobility measurements is a lower limit.[53] The second important difference is that the API device used 300Å barriers compared with 500Å barriers. Since photoconductive gain is inversely proportional to the superlattice periodicity this leads to a further increase in the responsivity of the API devices of 1.6x compared to the MIT devices. Taking all these factors into account, the single well quantum efficiency for the API devices is 0.16% if the 5x increase in g is taken from bulk mobility data or 0.081% if the 10x increase in g is taken from QWIP measurements of Kane et al. Both numbers are larger, but reasonably close to the value of 0.069% measured for sample 9066 made at MIT. The remainder of the difference can be explained by noting that the API devices have $2d$ sheet carrier densities a factor of 3x larger than that of sample 9066 grown at MIT. Since responsivity is proportional to $2d$ electron density through its dependence on the quantum efficiency, this explains the remainder of the difference between the responsivity of the API QWIPs and those grown and fabricated at MIT.

Though the responsivity of the API devices is twenty times larger than that of the MIT QWIPs for the reasons explained above, the BLIP detectivity is nearly identical with a value of 2×10^{10} Jones for the API devices equal to the BLIP detectivity for the best MIT QWIP. This is explained simply by noting that all of the factors that lead to a larger responsivity for the API device also lead to a larger background photocurrent noise. Thus detectivity is independent of the photoconductive gain, independent of the superlattice

period and only weakly dependent on the position of the 2*d* sheet carrier density through its dependence on the fermi level, E_F:

$$D^*_{BLIP} \propto E_F \exp\left(-\frac{E_F}{2kT}\right) \qquad Jones \tag{4.13}$$

Essentially all the devices are near enough to the peak in Equation 4.13 of $E_F = 2kT$, which for GaAs gives a 2*d* sheet carrier density of 3×10^{11} cm^{-2} at 65K, that the relative effect on the BLIP detectivity is minor.

In concluding this comparison of API and MIT QWIP devices it is important to point out that in a practical FPA system with a limited capacitor size available to collect electrons at each pixel, all equal detectivities are *not* created equal. In particular, as shown in the comparison of the API and MIT QWIPs, devices with dramatically different responsivities can end up with the same BLIP detectivity since many factors that affect the responsivity of a device also affect the noise in the device and the two affects tend to cancel when calculating the detectivity. In a system with a fixed capacitor size, BLIP detectivity is inversely proportional to the square root of the photoconductive gain because the number of electrons that can be stored is fixed. Thus the lower g, lower responsivity QWIPs of the style grown and fabricated at MIT are favored over the QWIPs grown and fabricated by API in spite of the detectivity being the same.

4.4 Conclusions on QWIP measurement

In this chapter measurements done on discrete normal incidence TE mode QWIPs have been presented. The results summarized in Table 4.2 show an average responsivity of 41 mA/W and a single well quantum efficiency of 0.035% consistent with the single well quantum efficiency of 0.068% of Section 2.3. The BLIP detectivity resulting from these measurements has a mean value of 1.2×10^{10} Jones for the devices grown, fabricated and measured at MIT. As discussed in Chapter 1 this is sufficient to produce uniformity limited FPAs with excellent MRTs less than 10mK without the additional fabrication complexity and expense involved with fabricating diffraction gratings. The devices fabricated at MIT were also compared with devices from API with slightly different structures but

very similar BLIP detectivities. The excellent agreement between the measurements done at MIT and those done at API and WL demonstrates the good understanding of the theory for TE mode QWIPs as well as the competitiveness of the growth, fabrication and measurement techniques developed as part of this thesis.

Chapter 5

Modeling and Cryogenic Measurement of VLSI GaAs MESFET Circuits

5.1 Introduction to VLSI GaAs MESFETs

The most important difference between gallium arsenide and silicon for fabrication of electronic circuits is existence of a high quality low surface state oxide for Silicon, namely Silicon Dioxide, and the lack of any comparable oxide for Gallium Arsenide. This simple observation explains why for GaAs electronics we must talk about metal semiconductor field effect transistors, MESFETs, rather than the more common Silicon based metal oxide field effect transistors, MOSFETs. The additional complexities in dealing with a rectifying MES based gate verses an insulating MOS based gate also go a long way in explaining the overwhelming dominance of Silicon based electronics in the market place.

Despite it's lack of a high quality oxide GaAs has several properties that have provided niche markets from which GaAs based technology has been able to develop and expand. First, unlike silicon, GaAs is a direct bandgap semiconductor which means that electroluminescence is highly efficient which has led to the dominance of GaAs and related III-V semiconductors in markets for light emitting diodes, LEDs, and semiconductor lasers. Second, there is a fortuitous alignment of the lattice constant of GaAs and AlAs which allows growth of very high quality pseudomorphic heterostructures of $Ga_xAl_{1-x}As$ on a GaAs substrate with none of the strain related problems associated with growing SiGe on Si. Advantage of this property will be taken in discussions related to integrating VLSI GaAs MESFETs with QWIPs in this thesis. Finally, and most important to the topic of this chapter, the electron mobility for GaAs MESFETs is typically a factor of about ten higher than that of Silicon MOSFETs at room temperature and the electron saturation velocity is typically a factor of two to three higher for GaAs MESFETs. Both of these fac-

tors lead to higher switching speeds and lower propagation delays for a GaAs MESFET verses a Silicon MOSFET of similar dimensions.

This chapter will first be concerned with developing a physical device model for the GaAs MESFET which is then used to understand the operation of the device. In addition, since these devices are intended for use in QWIP staring arrays which operate between 40 and 80 Kelvin, the results of measurements made on individual GaAs MESFETs at cryogenic temperatures will be analyzed to extract the relevant device parameters as a function of operating temperature based on the above mentioned model. Because one of the critical parameters, the gate capacitance, is not easily extracted from individual devices due to its extremely small value, about 1.2fF/um^2 for an EFET, and the consequent dominance of the measurement setup, a circuit model of a ring oscillator will be developed from which it will be shown that the oscillation frequency of the ring oscillator is directly proportional to the gate capacitance and other parameters which are easily extracted from the DC MESFET characterization. Measurement of ring oscillator frequency as a function of temperature and the subsequent extraction of the gate capacitance completes the characterization of individual MESFETs at cryogenic temperatures and lays the groundwork for future QWIP starring array multiplexer designs.

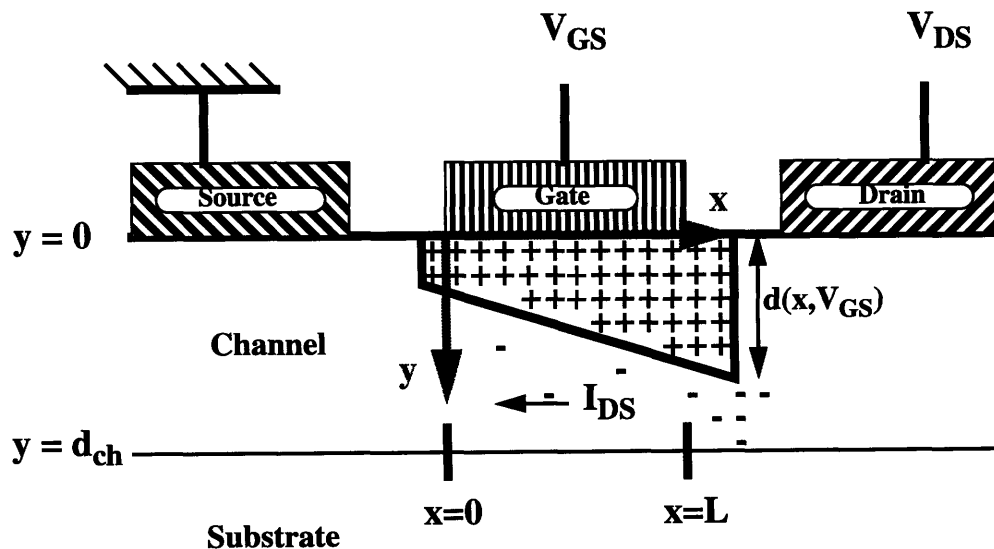


Figure 5.1: MESFET Cross Section

5.2 Physics and Modeling of GaAs MESFET

The GaAs MEtal Semiconductor Field Effect Transistor (MESFET) consists of a doped channel contacted on each end by ohmic contacts whose thickness is controlled by the depletion layer below the Schottky metal gate as shown in Figure 5.1. The key element of this device is the Schottky diode formed between the metal gate and the doped channel. For GaAs, the density of electron surface states at the interface between the gate metal and the semiconductor is such that the Fermi level is pinned near mid gap at an energy largely independent of the metal used for the gate[112] and creating a built-in reverse bias V_{BI} between the gate and the channel. This bias depletes the channel below the gate to an extent found by integrating Poison's Equation,

$$V(y) = \iint \frac{\rho(y)}{\varepsilon_s} dy^2 \tag{5.1}$$

with y the distance into the channel down from the gate and boundary conditions given by the depletion approximation. Assuming a uniformly doped channel for simplicity, and zero drain-to-source voltage, V_{DS}, the depletion depth as a function of applied gate voltage V_{GS} is

$$d(V_{GS}) = \sqrt{\frac{2\varepsilon_s(V_{bi} - V_{GS})}{qN_D}} \tag{5.2}$$

where ε_S is the permittivity for GaAs,$\varepsilon_S = 13.1$, $\varepsilon_0 = 1.14\,(fF/cm)$, q is the electron charge, N_D is the volume density of ionized dopant atoms and V_{BI} is the built in junction potential in the Schottky gate which for GaAs is $V_{BI} \approx 0.75\,Volts$. The applied V_{GS} subtracts from V_{BI} to modulate the extent of the depletion layer and thus the drain-to-source current, I_{DS}. With non-zero drain-to-source voltage, the largest extent of the depletion region, the largest longitudinal electric field and the constriction in the channel all occur on the drain side of the gate due to the increasing voltage drop along the channel which leads to a maximum in the reverse gate-to-channel bias on the drain side of the gate. With zero applied gate voltage, depending on the dopant density N_D and the channel depth, d_{CH}, the channel can be either completely depleted, $I_{DS} = 0$, or only partially depleted,

$I_{DS} \neq 0$. If the channel is completely depleted with zero applied gate voltage, a positive gate voltage $V_{GS} > 0$ must be applied to bring the device into conduction and this device is called an Enhancement mode FET (EFET). If the channel is conducting even with zero applied gate voltage, a negative gate voltage $V_{GS} < 0$ must be applied to shut off the device and this device is called a Depletion mode FET (DFET). The applied gate voltage above which the channel becomes conducting is called the threshold voltage V_T and is calculated by setting d in Equation 1.2 equal to the channel depth, d_{CH}, and then solving for $V_{GS} = V_T$, we find:

$$V_T = V_{BI} - \frac{qN_D d_{ch}^2}{2\varepsilon_s} \tag{5.3}$$

For EFETs the channel doping and depth are chosen such that $V_T > 0$ and for DFETs the channel doping and depth are chosen such that $V_T < 0$. It is important to note that in GaAs both EFETs and DFETs use electrons and not holes as their current carriers and in both cases the channel is doped n-type.

To get a basic understanding of the channel current, I_{DS}, a simple model attributed to Schockley[113] called the gradual channel approximation will be used. This model assumes the electron drift mobility, μ_{drift}, is constant with applied electric field giving a drift velocity

$$v_{drift} = \mu_{drift} E \tag{5.4}$$

The model also assumes that the gate junction is a slowly varying function of position, the conducting channel is neutral, the region under the gate is fully depleted, the electric field in the channel is oriented in the x-direction, the electric field in the depletion region is oriented in the y-direction, the boundary between the channel and the depletion region is sharp, the channel doping is uniform and the potential across the channel varies so slowly that the thickness of the depletion region at any point along the length of the gate can be found by integrating Poison's equation, Equation 5.1, valid for a one-dimensional junction. All of these assumptions are incorporated in Figure 5.1 and result in an incremental change in the channel potential, $V(x)$, given by

$$dV(x) = \frac{I_{DS}dx}{q\mu N_D W[d_{ch} - d(x, V_{GS})]} \tag{5.5}$$

where W is the gate width and $d(x, V_{GS})$ is the depletion depth as a function of distance along the gate length and is by assumption the depletion width below the Schottky gate due to the gate bias, Equation 5.2, with the additional bias due to the drain-to-source bias, $V(x)$ and is therefore given by

$$d(x, V_{GS}) = \sqrt{\frac{2\varepsilon_s[V(x) + V_{bi} - V_{GS}]}{qN_D}} \tag{5.6}$$

Substituting $d(x, V_{GS})$ into $dV(x)$ and integrating with respect to x from the source side of the gate, $x = 0$, to the drain side of the gate, $x = L$, gives the desired result for the channel current, I_{DS}, as a function of applied gate-to-source voltage, V_{GS}, and applied drain-to-source voltage, V_{DS}

$$I_{DS} = g_0\left(V_{DS} - \frac{2}{3}\frac{(V_{DS} + V_{BI} - V_{GS})^{3/2} - (V_{BI} - V_{GS})^{3/2}}{V_{PO}^{1/2}}\right) \tag{5.7}$$

where g_0 is the channel conductance when no gate metal is present given by

$$g_0 = \frac{q\mu N_D W d_{ch}}{L} \tag{5.8}$$

V_{PO} is the pinch off voltage given by

$$V_{PO} = \frac{qN_D d_{ch}^2}{2\varepsilon_s} \tag{5.9}$$

and the series resistance of the drain-to-gate, R_D, and gate-to-source, R_S regions have been neglected for now. For a given V_{GS} this result is valid from $V_{DS} = 0$ up to $V_{DS} = V_{sat} \equiv V_{PO} - V_{BI} + V_{GS}$ where the channel is completely pinched off at the drain side of the gate and the saturation channel current is from Equation 5.7

$$I_{DS}^{sat} = g_0\left(\frac{1}{3}V_{PO} + \frac{2}{3}\frac{(V_{BI} - V_{GS})^{3/2}}{V_{PO}^{1/2}} - V_{BI} + V_{GS}\right) \tag{5.10}$$

For increases in V_{DS} beyond V_{sat} the Schockley model assumes that I_{DS} remains constant. Taking the derivative of Equation 5.10 with respect to applied gate voltage, V_{GS}, gives the Shockley transconductance of the MESFET in the saturation region.

$$g_m^S = \frac{dI_{DS}^{sat}}{dV_{GS}} = g_0\left(1 - \left(\frac{V_{BI} - V_{GS}}{V_{PO}}\right)\right) \tag{5.11}$$

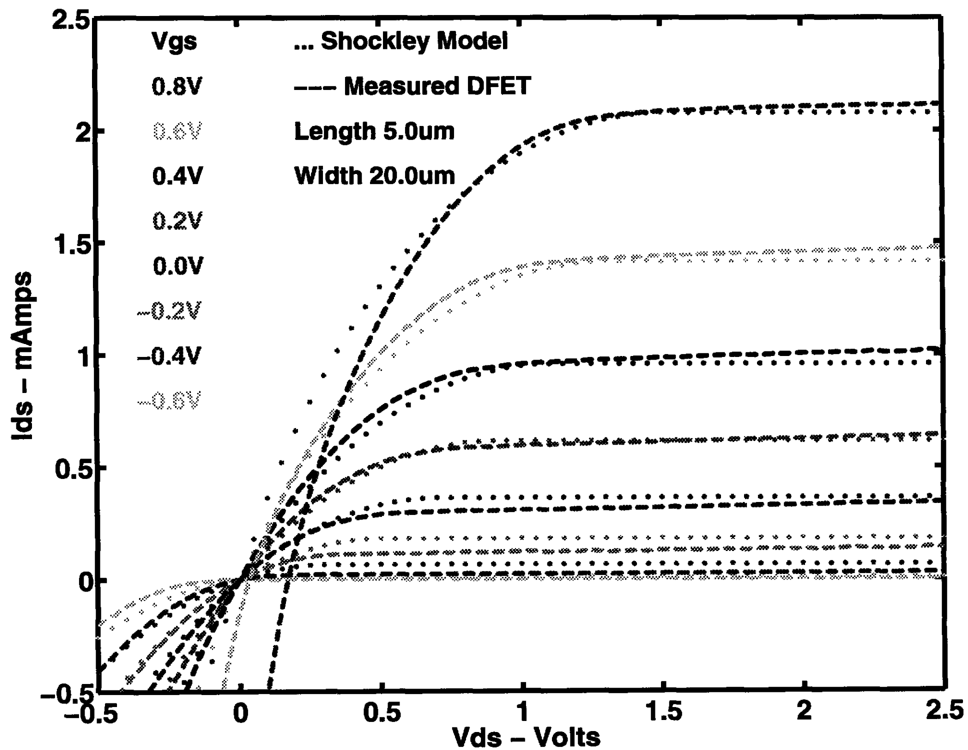


Figure 5.2: Comparison of Shockley Model with Measured 20µmx5µm DFET.

To asses the accuracy of the Shockley model, Figure 5.2 plots the measured room temperature V_{DS} vs. I_{DS} curves for a 20µm x 5µm DFET. The Shockley model used a doping density $N_D = 1.0e17/cm^3$, a channel depth of $d_{ch} = 0.142\mu m$, an electron drift mobil-

ity of $\mu_{drift} = 4000\ cm^2/V-S$, and a built in bias of $V_{BI} = 0.9\ V$. It is clear from Figure 5.2 that despite its simplicity the Shockley model gives an excellent fit to the drain current for this relatively long gate DFET throughout the linear and saturation regions as well as in the critical transition region near V_{sat}. The most striking deviation occurs at high gate bias where the Schottky gate begins to conduct creating an offset in V_{DS} equal to the gate-to-source current I_{GS} multiplied by the source resistance R_S. Since the Shockley model does not include the characteristics of the conducting gate and in our derivation we have set $R_S = 0$ the deviation is not surprising. The positive slope in I_{DS} for V_{DS} beyond the saturation voltage, V_{sat}, is the result of increased current injection into the pinch off region as a function of increased drain voltage. This leads to finite output resistance and finite gain.

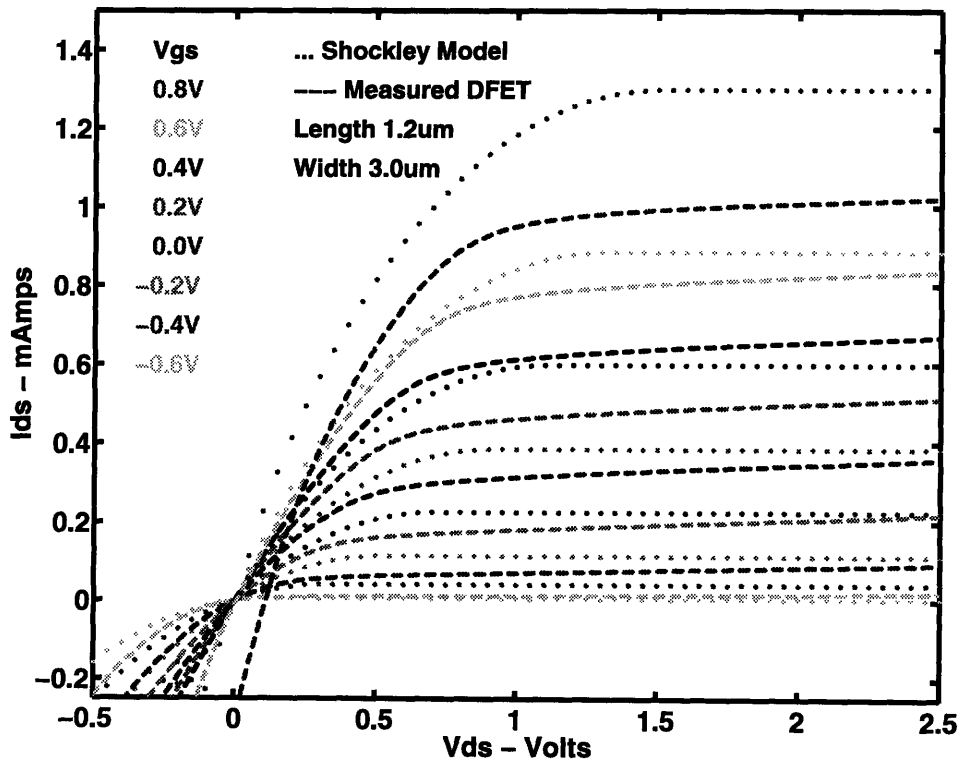


Figure 5.3: Comparison of Shockley Model with Measured 3μmx1.2μm DFET.

The fit for the 3μm x 1.2μm DFET shown in Figure 5.3 is much less satisfying. In addition to the shifts due to gate conduction, the saturation voltage, V_{sat}, the saturation

current, I_{sat}, and the transconductance, g_m, are all poorly modeled due to “short channel effects”. Short channel effects are the result of the breakdown of both the gradual channel approximation and the constant electron drift mobility assumption on which the Shockley model is based. As can be seen from a plot of experimentally measured electron velocity as a function of applied electric field[114], Figure 5.4, electron drift velocity is proportional to electric field only for electric fields $E < 3kV/cm = 0.3V/\mu m$. For higher electric fields typical of short gate MESFETs, the electron velocity initially “overshoots” and then drops down to a “saturation” level in which velocity is only weakly dependent on electric field. Physically, the velocity overshoot and saturation are both due to electrons being scattered out of the high mobility Γ valley and into the higher energy and lower mobility X and L valleys[115]. Several MESFET models incorporating the physics of short channel effects have been developed[116,RP_THC], but since the interest here is in modeling rather than detailed physical understanding a circuit model that takes advantage of the physical similarities between MESFETs and JFETs to use the JFET modeling capabilities of HSPICE[117] will be described. This model has the important advantage that it accurately models the output characteristics of the MESFETs. In particular, the small signal output resistance, defined in the pinch off region as:

$$\frac{1}{r_{out}} \equiv \frac{\partial I_D}{\partial V_{DS}} = \frac{V_A}{I_{DS}} \tag{5.12}$$

and the maximum small signal gain given by:

$$a = r_{out} g_m = \frac{2V_A}{V_{GS} - V_T} \tag{5.13}$$

are properly modeled in the HSPICE JFET model. Since the output resistance and maximum gain are critical parameters in the design of the analog readout circuits for the focal plane arrays which are the objective of this thesis, the HSPICE JFET model is a very important tool. The HSPICE equivalent circuit used to describe the Vitesse ion-implanted, self-aligned, refractory gate MESFET is shown in Figure 5.5.

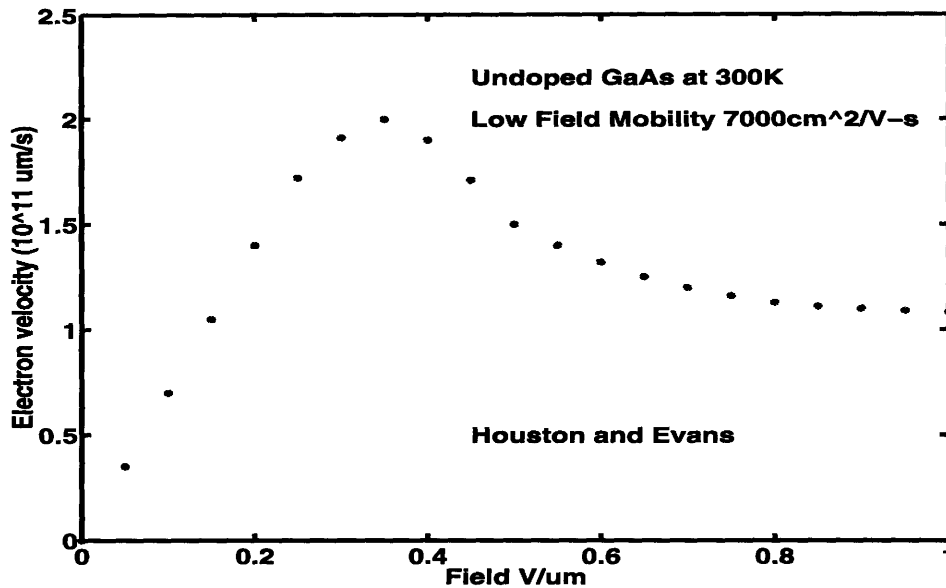


Figure 5.4: Velocity Overshoot and Velocity Saturation in GaAs.

The current source, I^0_{DS}, describes the channel current as a function of threshold voltage, V_T, device geometry, backgate voltage and the applied voltages, V_G, V_S and V_D. The total gate current, I_{DS}, includes I^0_{DS} plus the current due to the Schottky gate which is divided into gate-source and gate-drain Schottky diodes.

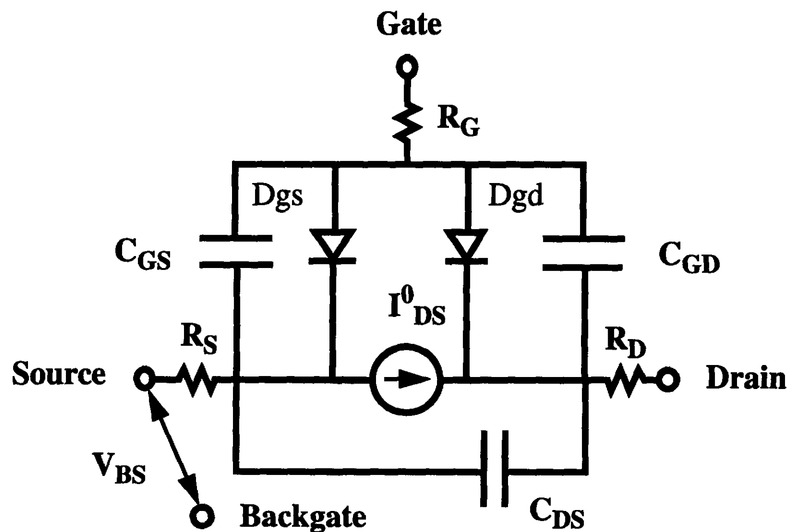


Figure 5.5: Equivalent Circuit Model for GaAs MESFET.

The total capacitance includes the capacitance associated with the two Schottky gate diodes, C_{GS} and C_{GD}, which are voltage dependent as well as the geometrical capacitance,

C_{DS}, associated with the source and drain. The threshold voltage and channel current models are both taken from Statz[118] as modified by Vitesse[119] and are given by:

$$V_T = V_T^0 + (GAMDS \times V_{DS}) + K_1 (V_{BS}) \qquad (5.14)$$

where V^0_T, is the threshold voltage given by Equation 5.3, $GAMDS$ accounts for the bias dependence and $K_1 (V_{BS})$, accounts for the backgating effect. The channel current is:

$$I_{DS} = \beta_{eff}(V_{GS} - V_T)^{VG}(1 + \lambda V_{DS})\left[1 - \left(1 - \frac{\alpha V_{DS}}{3}\right)^{SAT}\right] + I_{SUB} \qquad (5.15)$$

where β_{eff}, is the transconductance parameter which includes the effects of velocity saturation, α is the drain voltage multiplier which determines the slope of the linear region of the I_{DS} vs. V_{DS} curve, λ is the channel length modulation parameter and both VG and SAT are contents in Statz's model which have been parameterized to better fit the data. The transconductance parameter β_{eff} is related to the intrinsic transconductance parameter β and the critical field for onset of velocity saturation, $VCRIT$ by

$$\beta_{eff} = \frac{\beta}{1 + VCRIT(V_{GS} - V_T)} \qquad (5.16)$$

Finally, I_{SUB}, is the subthreshold current as modeled by Lee et.al[120]

$$I_{SUB} = I_0 e^{ND \times V_{DS}} e^{-NG \times V_{DS}} \qquad (5.17)$$

where I_0, ND and NG are fitting parameters for the empirical model.

Figure 5.6 shows the fit of the HSPICE model described above to the same measured data as shown in Figure 5.2 for a 20μm x 5μm DFET. The parameters were allowed to vary within one standard deviation of the Vitesse fabrication process and then optimized using HSPICE. The resulting HSPICE parameters are shown in Table 5.1, Column 2. The model clearly shows a very good agreement in both the linear and saturation regimes including a proper modeling of the V_{DS} offset for large gate-source voltages. The discrepancy for large V_{GS} is likely due to parasitic gate resistance in the measurement apparatus which reduces the true applied V_{GS} from the measured value.

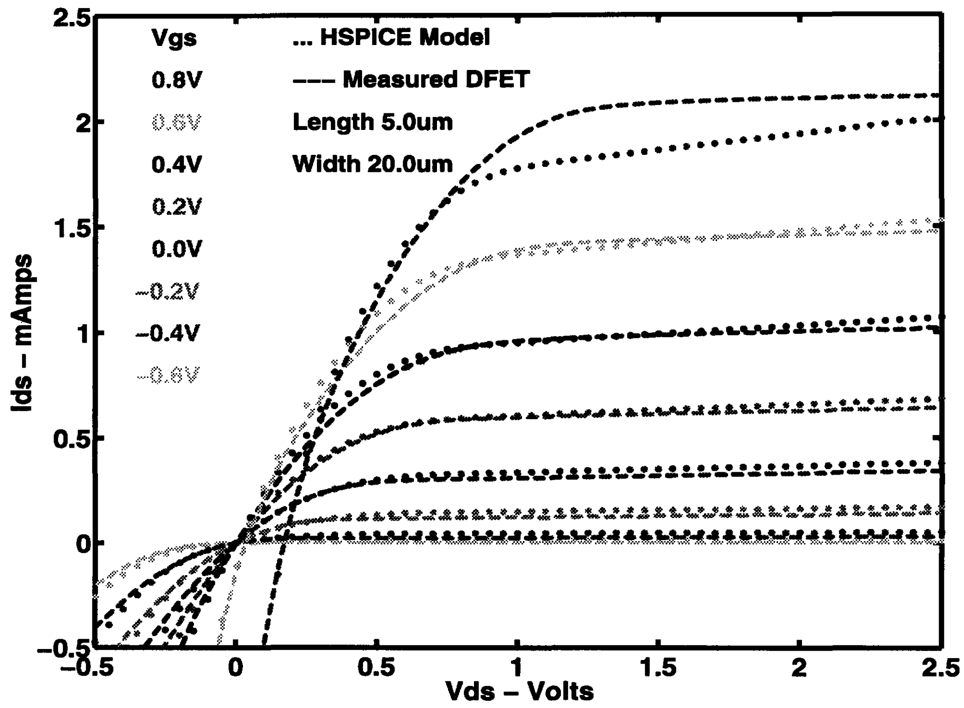


Figure 5.6: Comparison of HSPICE Model with Measured 20μmx5μm DFET.

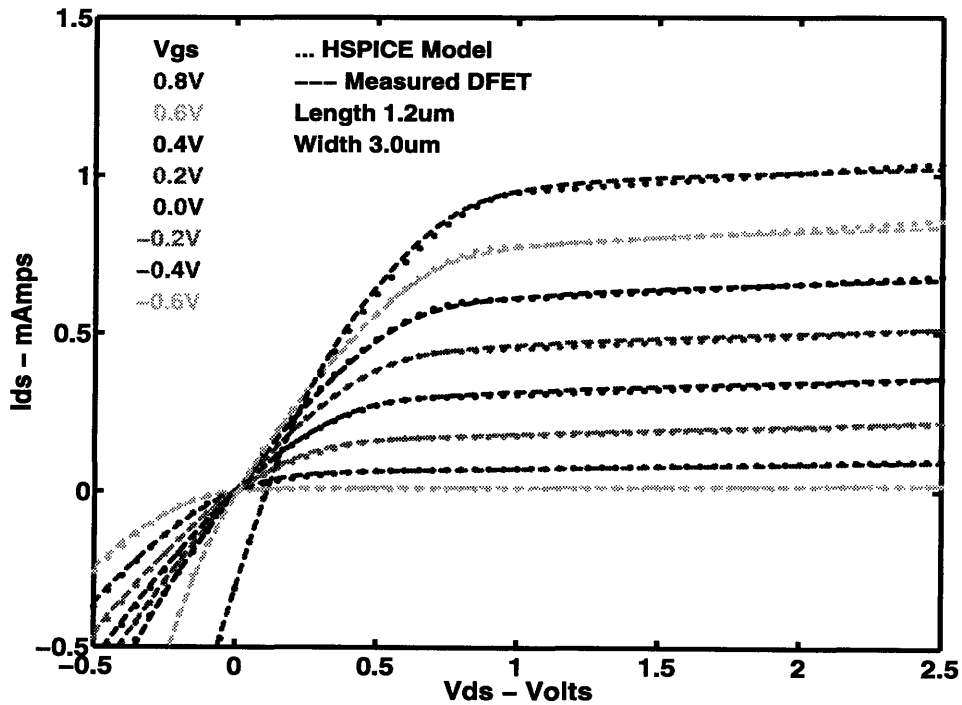


Figure 5.7: Comparison of HSPICE Model with Measured 3μm x 1.2μm DFET.

Figure 5.7 shows the fit of the HSPICE model to the same measured data as shown in Figure 5.3 for a 3μm x 1.2μm DFET. The resulting HSPICE parameters are shown in Table 5.1 column 3. Again the fit is excellent over the entire I_{DS} vs. V_{DS} curve showing that the HSPICE model appropriately accounts for the short channel effects that reduced the usefulness of the Shockley model at short gate lengths. In particular, the small signal output resistance calculated by HSPICE is within a small factor of that obtained by direct measurement of the MESFET DC characteristics.

HSPICE Parameter	DFET (L x W)		Units
	5um x20um	3um x 1.2um	
V^0_T	645	795	mV
β	190	190	uA/V^{VG}
ΔL	-0.40	-0.4	um
GAMDS	0.0	-0.065	#
λ	0.065	0.065	V^{-1}
VCRIT	0.0	0.5	V^{-1}
α	3.5	3.5	V^{-1}
ΔW	-0.15	-0.15	um
NG	1.1	1.1	V^{-1}
ND	0.2	0.2	V^{-1}
K1	0.35	0.35	V
VG	2.0	2	#
SAT	3.5	3.5	#

Table 5.1: HSPICE parameters for Figures 5.6 and 5.7

V_{gs} (Volts)	V_a (Volts)	R_{out} ($x10^4$ Ohms)	a_{max} (dB)
0.8	-96.6	4.7	41.9
0.6	-39.7	2.9	35.4
0.4	-27.3	2.9	33.5
0.2	-19.6	3.5	32.3
0.0	-12.4	4.4	30.3
-0.2	-8.3	7.5	29.5

Table 5.2: Output Characteristics for 20μm x 5μm DFET

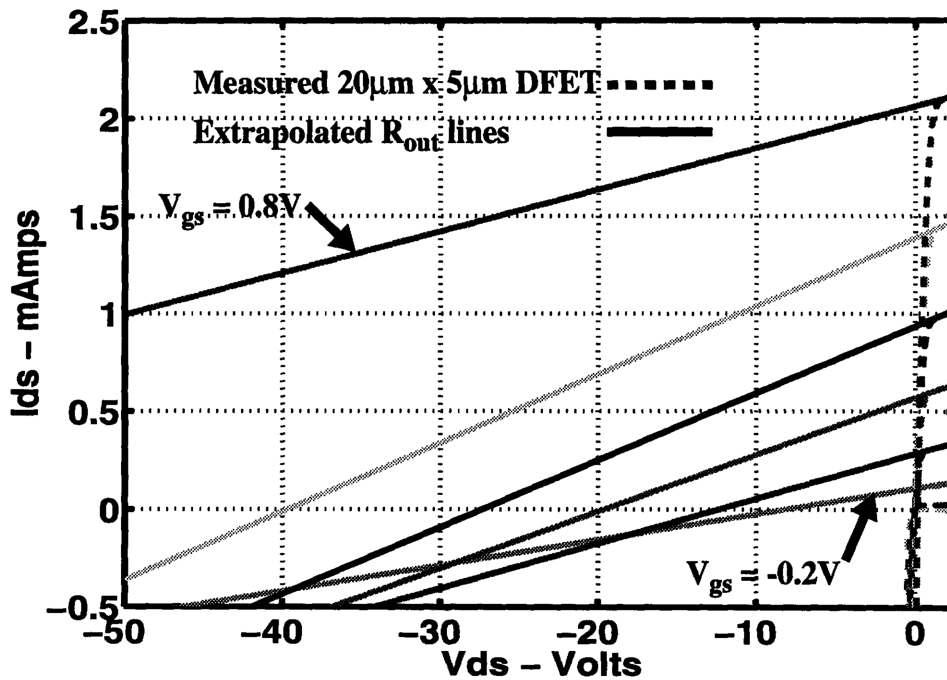


Figure 5.8: Output parameter extraction for 20μm x 5.0μm DFET. V_{gs} from -0.2V to 0.8V in 0.2V steps. See Table 5.2 for output resistance values.

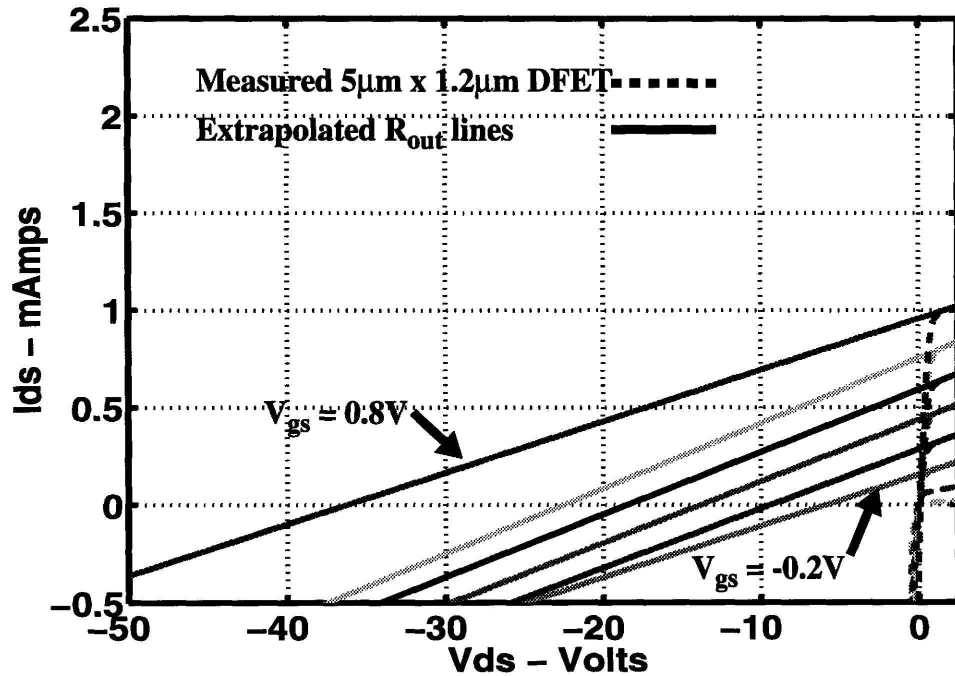


Figure 5.9: Output parameter extraction for 3μm x 1.2μm DFET. V_{gs} from -0.2V to 0.8V in 0.2V steps. See Table 5.3 for output resistance values.

V_{gs} (Volts)	V_a (Volts)	R_{out} ($x10^4$ Ohms)	a_{max} (dB)
0.8	-36.3	3.8	33.4
0.6	-22.5	3.0	30.4
0.4	-18.5	3.1	30.1
0.2	-13.8	3.2	29.2
0.0	-9.4	3.3	27.9
-0.2	-5.9	3.8	26.5

Table 5.3: Output Characteristics for 5μm x 1.2μm DFET

The small signal output resistance given by Equation 5.12 and the maximum gain given by Equation 5.13 of the MESFETs were also measured from the DC characteristics. Figure 5.8 shows the results for the 20μm x 5.0μm DFET and Figure 5.9 shows the results for

the 3μm x 1.2μm DFET. It is interesting to note that in spite of the nearly four fold difference in the channel length, the output resistance changes only by a factor of about three. This contrasts dramatically with Silicon CMOS where for a comparable 1.2μm n-well process an increase in the channel length from 1.2μm to 5.0μm leads to an increase of more than two orders of magnitude in both the small signal output resistance and the maximum small signal gain.

5.3 Cryogenic Measurement of GaAs MESFETs

The HSPICE model described above has been used by Vitesse to build a data base for simulation of MESFET circuits over the military temperature range (-55C to +125C). Here, measurements were performed which extend this range down to the operating temperature for QWIP FPAs typically 50K to 70K, (-223C to -203C). Individual device measurements were done on three MESFETs two DFETs and one EFETs with geometries of 20μm x 5μm and 3μm x 1.2μm and 3μm x 1.2μm respectively and which are part of the Process Control Monitor, PCM, for the OEVLSI-1 chip designed by K.V. Shenoy[121]. The measurements performed were used to extract threshold voltage, V_T, transconductance, g_m, source resistance, R_s, drain resistance, R_D, Schottky barrier height, ϕ_b and Schottky barrier ideality factor, n all as a function of device temperature. Measurements on ring oscillators used to extract the gate capacitances, C_{GS} and C_{GD} are described in Section 5.4. All measurements were performed on devices bonded into 14-pin DIP packages and cooled using a CTI closed cycle Helium refrigerator with a Cryogenics cryostat and cryogenic temperature controller. Special care was taken to accurately measure the temperature at the position of the device under test. This included a copper plate which covered the 14-pin DIP package and was screwed into the Helium refrigerator cold finger forcing intimate contact between both the front and the back of the 14-pin DIP package and the cold finger. To confirm our temperature readings an Omega CY7-SD7 calibrated silicon diode temperature sensor was bonded into the 14-pin DIP package along side the device under test. Without the copper plate, temperature readings from the controller taken with a temperature diode located at the end of the cold finger were as much as 100K lower than the tem-

perature readings inside the 14-pin DIP package, but use of the copper plate reduced this difference to less than 1K which is within the quoted accuracy of the Omega CY7-SD7 silicon diode. A Hewlett-Packard 4145 semiconductor parameter analyzer controlled by a PC was used to collect the electrical data from the cooled devices. Measurement techniques are based on those used by Vitesse[119] and described in detail by Eric Braun[122].

5.3.1 I-V Characteristics for MESFETs:

Before extracting the temperature dependence of specific parameters in the MESFET model described above, a quick look at the I-V characteristics as a function of temperature will give an important qualitative understanding of what to expect. Figure 5.10 shows the output I-V characteristics for a 20μm x 5.0μm DFET at 50K, 70K and 300K. Figure 5.11 shows the output I-V characteristics for a 3μm x 1.2μm DFET at the same temperatures. Both curves show an decrease in the source to drain voltage required to get zero source to drain current at high gate bias indicating that the Schottky barrier height is increased at low temperature reducing the gate leakage current. Both curves also show an increase in the change in the source to drain current for a fixed increase in the gate bias indicating an increase in the device transconductance. There is a small decrease in the slope of the I-V characteristic in the saturation region for both curves as the temperature is increased indicating an increase in the small signal output resistance and small signal gain.

Comparing Figures 5.10 and 5.11 for the case of zero gate bias gives an average drain current of:

$$I_{DS} = 90\mu A \times \frac{W}{L} \tag{5.18}$$

at 70K with only a very small change as a function of temperature.

Finally, it is important to point out that this is the first temperature dependent characterization of the Vitesse HGaAsIII devices deonstrating that they work well at the temperatures required for operation of QWIP FPAs and though rigorous reliability testing has not been done, repeated temperature cycling during the measurement had no noticeable effect on the I-V characteristics. Threshold Voltage and Transconductance:

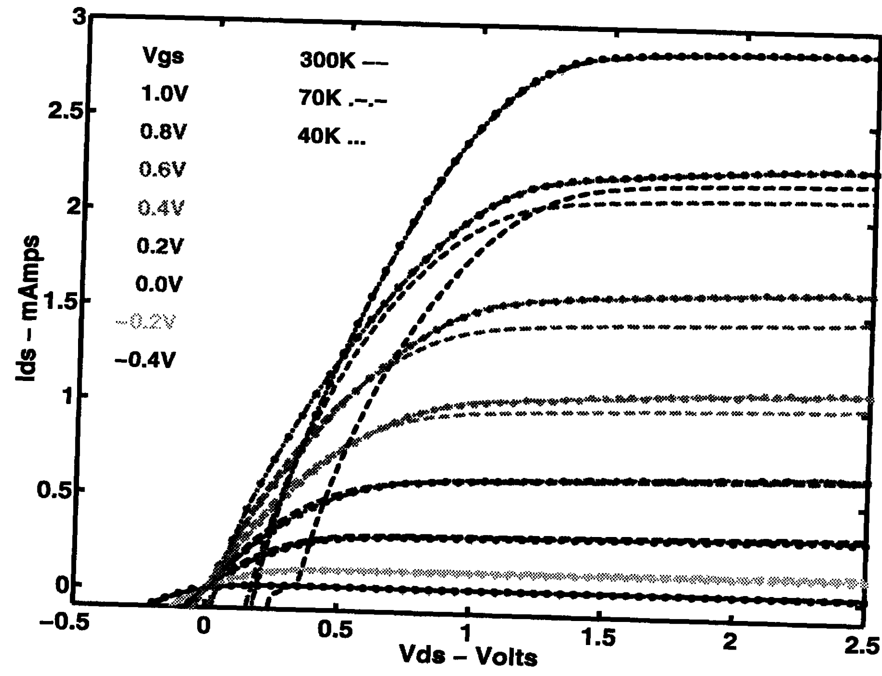


Figure 5.10: I-V Characteristics for 20μm x 5.0μm DFET.

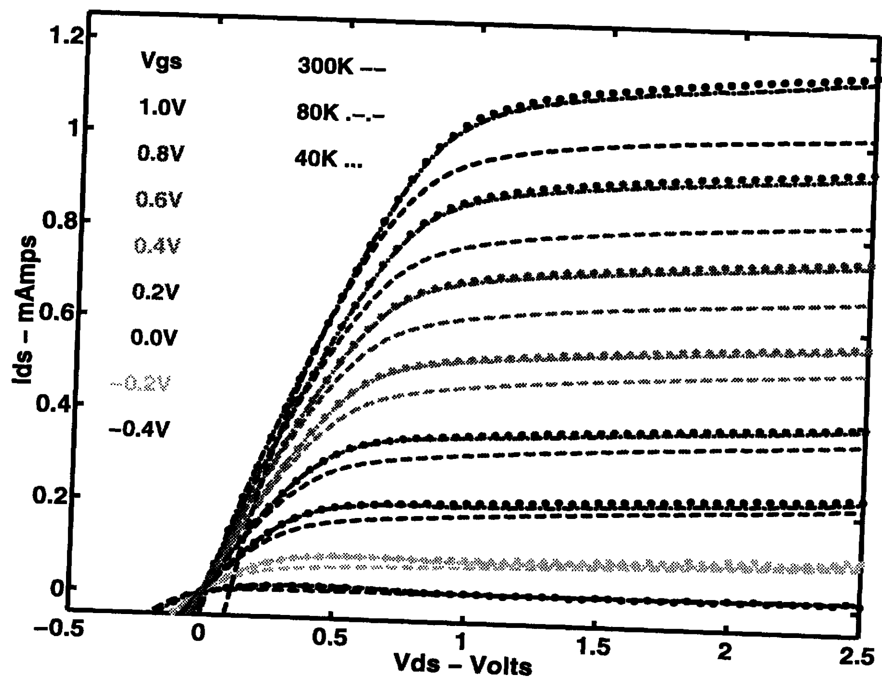


Figure 5.11: I-V Characteristics for 3μm x 1.2μm DFET.

5.3.2 Threshod Voltage and Transconductance

For each measured temperature, the threshold voltage and transconductance were extracted from a plot of I_{DS} vs. V_{GS} with $V_{DS} = 1\,Volt$. The maximum slope of this curve was used as the device transconductance. To find the threshold voltage, the linear region of the I_{DS} vs. V_{GS} curve was extrapolated and its intersection with the V_{GS} axis, V_{GS}^0, was found.

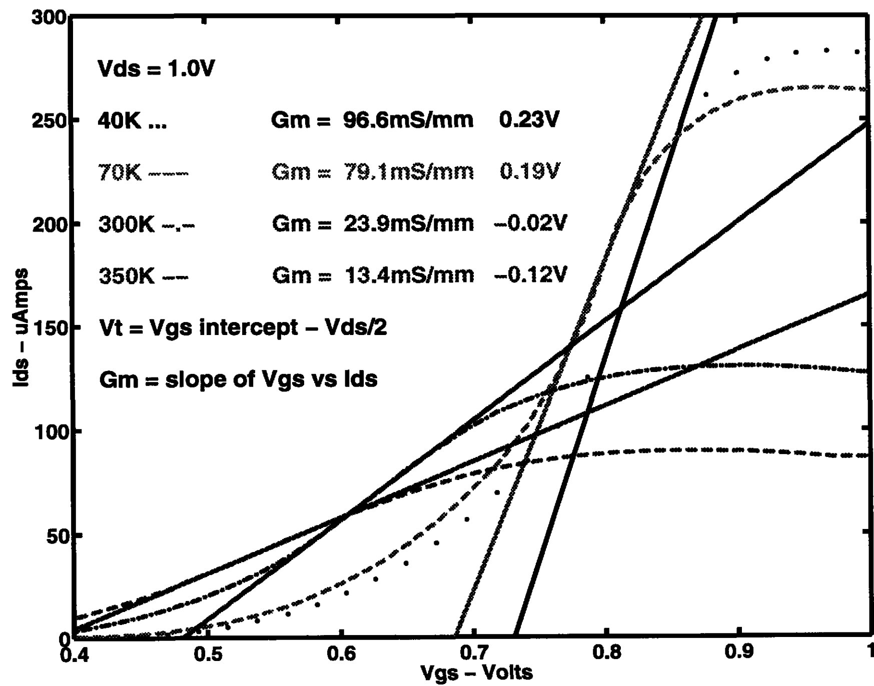


Figure 5.12: Transconductance and threshold voltage extraction for 3µm x 1.2µm EFET.

The threshold voltage was defined as $V_T = V_{GS}^0 - V_{DS}/2$. Figure 5.12 shows the extraction for the 3µm x 1.2µm EFET, Figure 5.13 shows the temperature dependence of the transconductance for the all three devices and Figure 5.14 shows the temperature dependence of the threshold voltage for all three devices. The transconductance is seen to increase slowly at a rate of about -0.15mS/mm/K as the temperature is lowered and the threshold voltage is seen to increase slowly at a rate of about -1mV/K, making all the devices more EFET like.

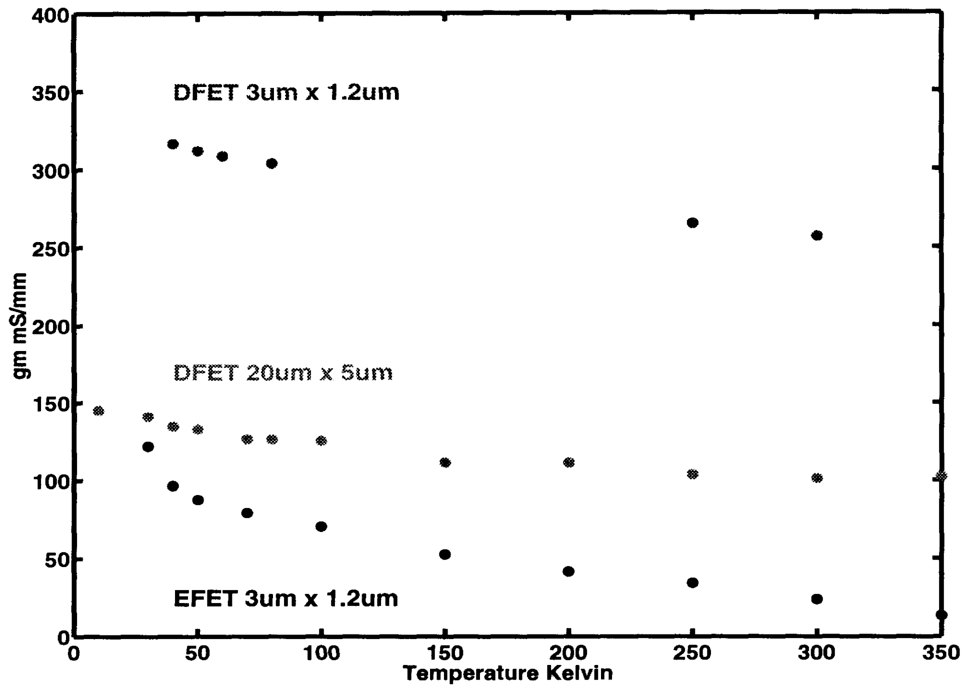


Figure 5.13: Temperature dependence of transconductance.

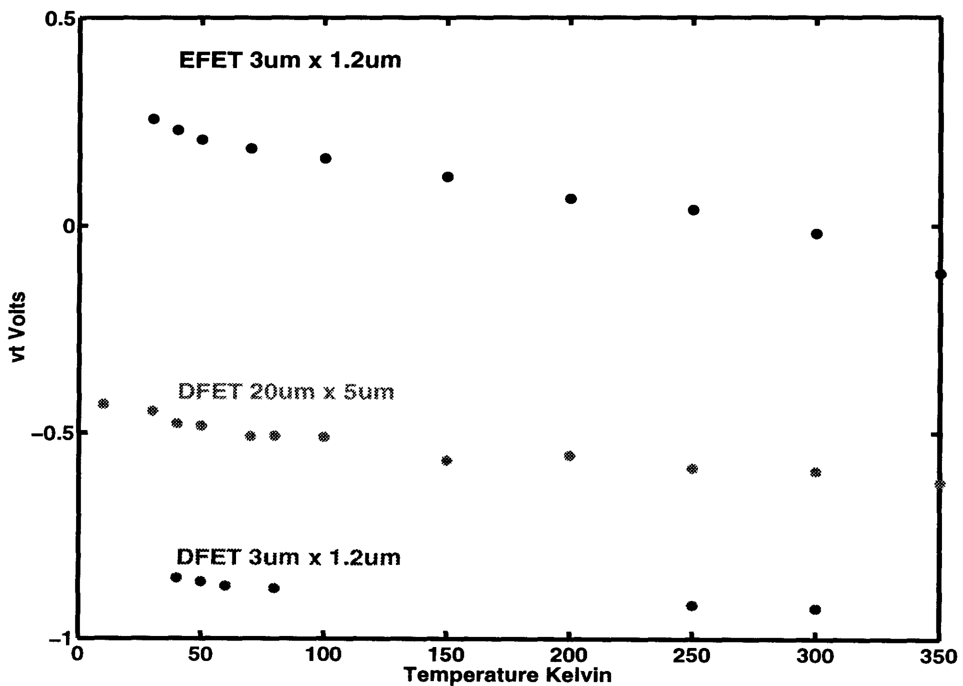


Figure 5.14: Temperature dependence of threshold voltage.

5.3.3 Source and Drain Resistance:

The source and drain resistance were extracted using the method of Chaundhuri and Das[123]. In this method, the Schottky gate diode is forward biased forcing a current, I_{GS}, through the source resistor. A small drain current ($I_{DS} = 1\mu A$) was applied to stabilize the measurement. The source resistance was identified as the slope of the I_{GS} vs. I_{DS} curve in the central linear region. Figure 5.15 shows the source resistance for the 3μm x 1.2μm EFET, Figure 5.16 shows the temperature dependence of the source resistance for all four devices. The drain resistance was extracted in an analogous manner and is shown for the 3μm x 1.2μm EFET in Figure 5.17. Figure 5.18 shows the temperature dependence of the drain resistance for all three devices. Note that all these devices are symmetric, $R_S = R_D$, within the measurement error.

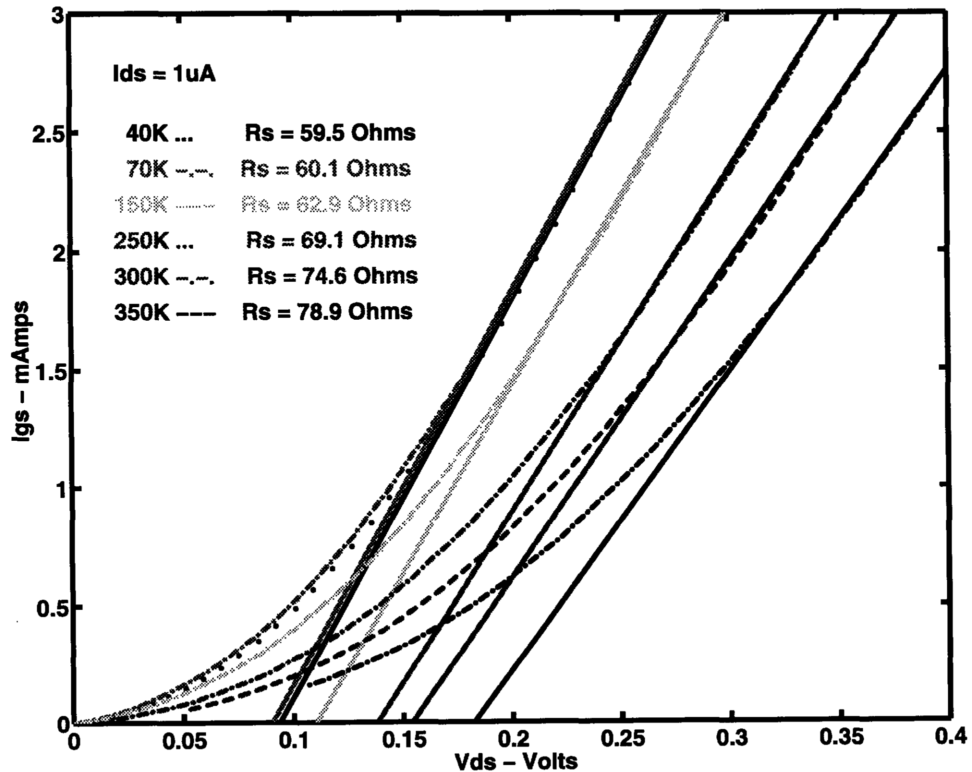


Figure 5.15: Source resistance extraction for 3μm x 1.2μm EFET.

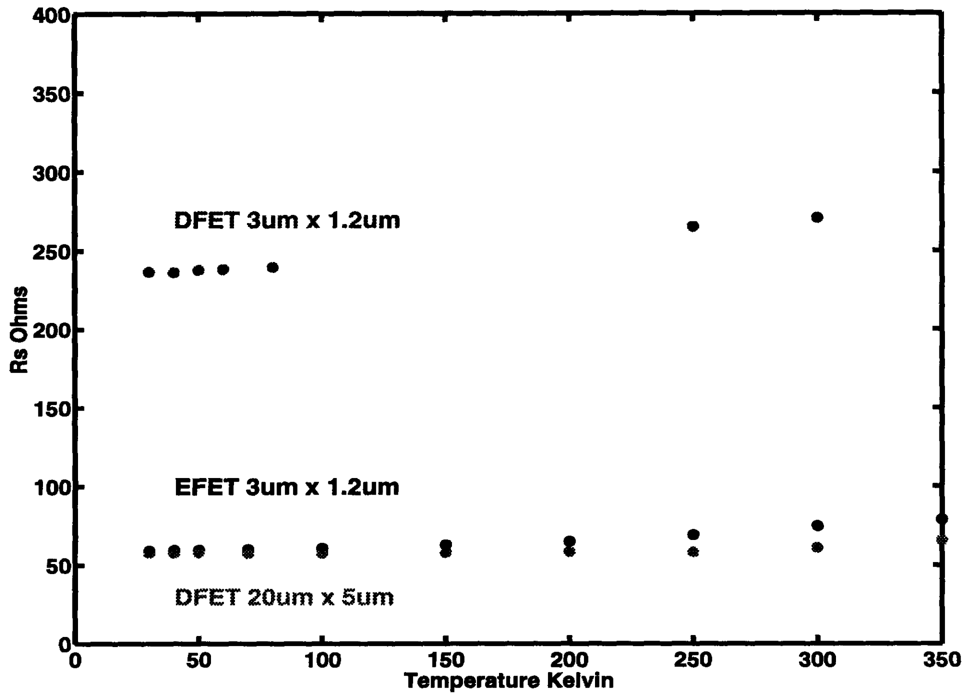


Figure 5.16: Source resistance temperature dependence.

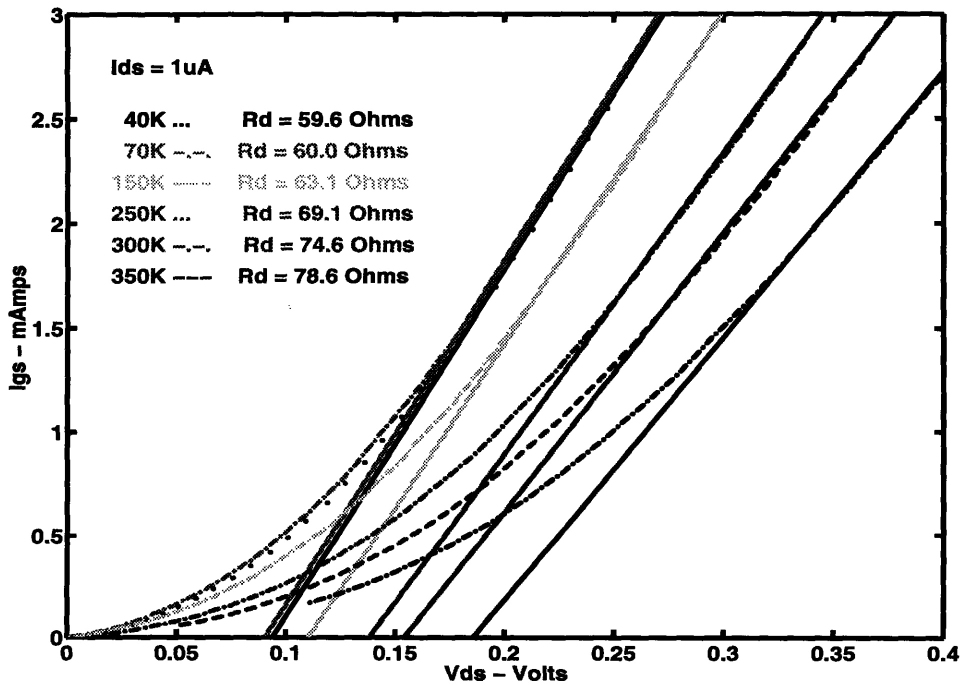


Figure 5.17: Drain resistance extraction for 3μm x 1.2μm EFET.

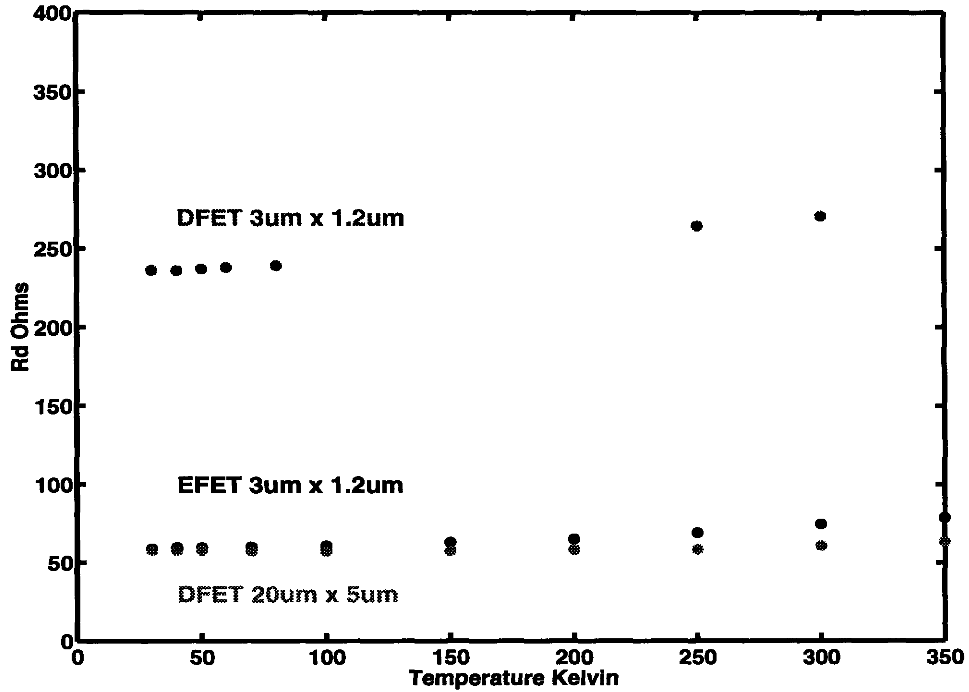


Figure 5.18: Drain resistance temperature dependence.

5.3.4 Schottky gate barrier height and ideality factor

The Schottky gate barrier height and ideality were measured by tying the source and drain to ground and applying a forward bias to the gate. Using the approximate expression for the total current density through a Schottky diode

$$J_{DS} = \frac{I_{DS}}{A} = J_{sat}\left[\exp\left(\frac{qV_{GS}}{nkT}\right) - 1\right] \tag{5.19}$$

where J_{sat} is the saturation current density and plotting the natural log of the gate current, log (I_{DS}) vs gate voltage, V_{GB} gives a line whos slope is the ideality factor:[124]

$$n \equiv \frac{q}{kT} \frac{1}{\left(\frac{dJ_{DS}}{dV_{GS}}\right)} \tag{5.20}$$

The extraction of the ideality factor for the 20μm x 5μm EFET is shown in Figure 5.19. The current measurement noise floor is about 100pA but for currents higher than this, there are the expected linear regions from which the ideality factor may be extracted.

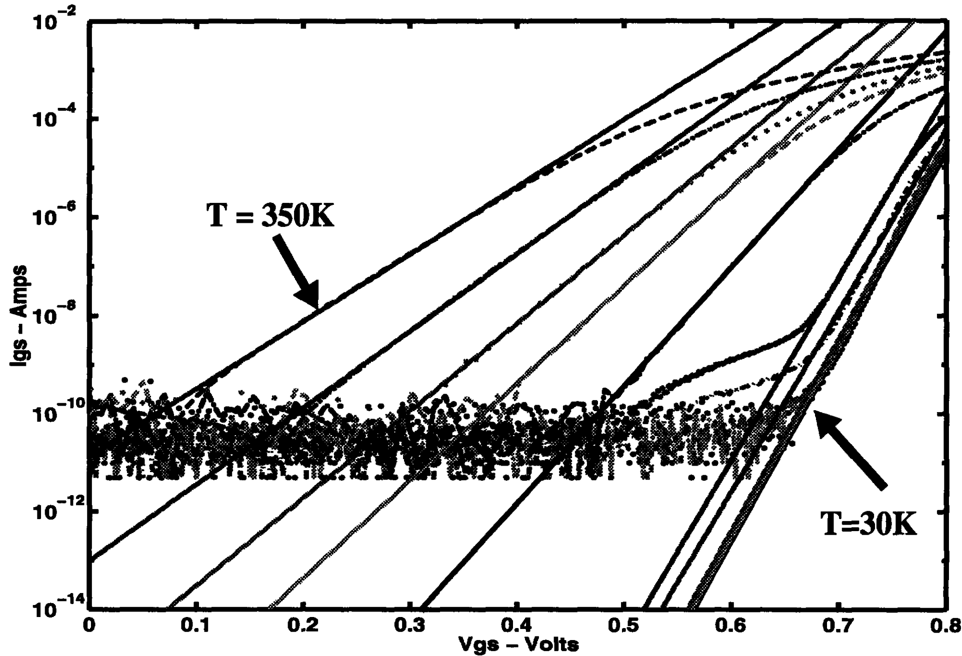


Figure 5.19: Schottky diode ideality extraction for 20µm x 5µm EFET with $V_{ds} = 0$. Ideality factors are 4.73, 3.89, 3.26, 2.44, 1.89, 1.40, 1.27, 1.13, 1.07 and 1.05 for temperatures of 30K, 40K, 50K, 70K, 100K, 150K, 200K, 250K, 300K and 350K respectively. Measurement noise floor ~100pA.

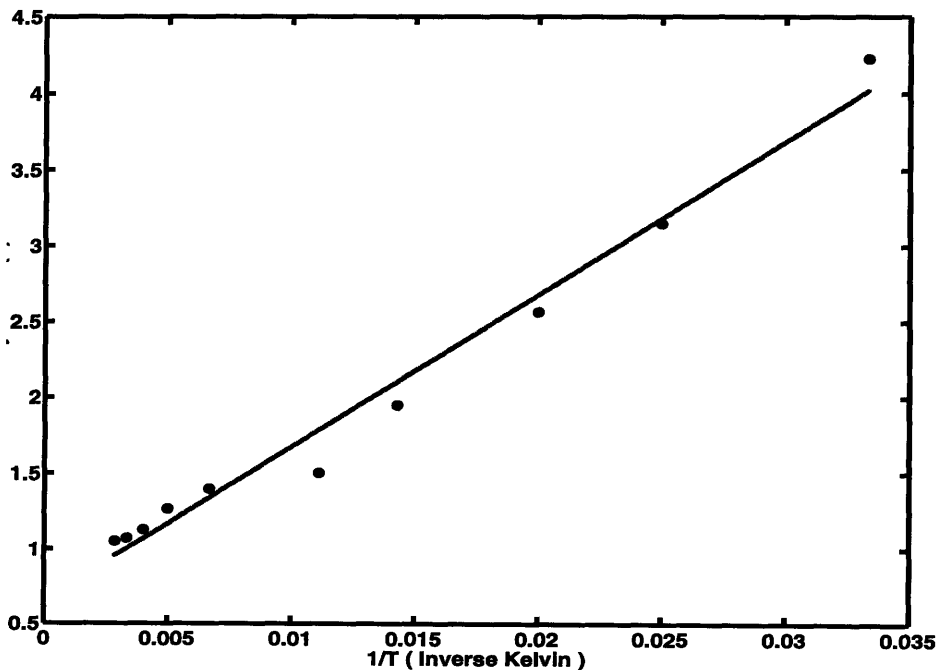


Figure 5.20: Schottky diode ideality factor temperature dependence for 20µm x 5µm EFET with $V_{ds} = 0$. Data from Figure 5.19.

Figure 5.20 shows the ideality factor as a function of temperature for the 20μm x 5μm EFET with $V_{ds} = 0$. The fairly strong increase in the ideality factor at low temperature was at first surprising, but can be understood by a more detailed look at not only the thermionic emission current which dominates near room temperature and results in the simplified Equation 5.19, but also including the tunneling current which is important for low temperatures and high channel dopings.[125,126,127]

5.4 Cryogenic Measurement of Gate Capacitance

The gate capacitance is a critical parameter in the design of FPA readouts because the gate to channel capacitor forms the highest capacitance and most uniform capacitor available in the HGaAsIII fabrication process. These capacitors will be used in the readouts to store charge collected by each single pixel over a single frame cycle. A schematic diagram of a simple pixel is shown in Figure 5.21.

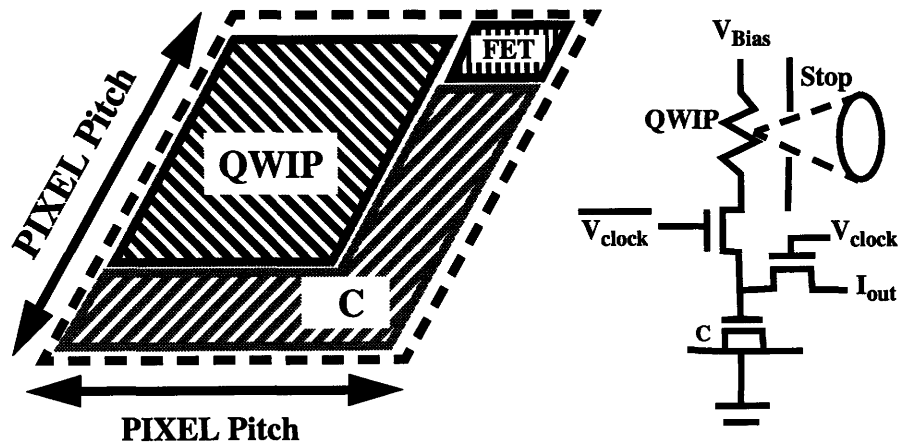


Figure 5.21: Schematic of Simple Pixel Layout and Circuit Diagram.

The approximate size of the capacitor required can be estimated from the relation:[128]

$$Dynamic\ Range \approx \sqrt{Number\ Stored\ electrons} \qquad (5.21)$$

For a typical dynamic range of 12bits (72dB) this requires that about 17million electrons be stored for each pixel. Assuming a capacitor voltage of 1Volt gives a capacitance per pixel on the order of 2.7pF. At room temperature the gate capacitance for the Vitesse

HGaAsIII process is 2fF/μm^2 [119] giving a required capacitor area of 1360 μm^2. Since this large, more than half of the desired pixel size, 50μmx50μm, more complex pixel designs that do not require as large capacitors are quite valuable.[129] At this point having demonstrated the importance of the gate capacitance value an attempt will be made to measure the gate capacitance near the QWIP operating temperature. At 2fF/μm^2 the largest MESFETs available on the PCM, 20μm x 5.0μm have a capacitance of only 0.2pF. Since the bond pads, package and dewar together have stray capacitance several orders of magnitude higher than this, direct measurement of the gate capacitance is not possible. Instead, after a short discussion of ring oscillators, the oscillation frequency of two ring oscillators designed by Eric Braun[122] will be used to indirectly measure the gate capacitance as a function of temperature.

5.4.1 Ring Oscillator Design

A ring oscillator is fabricated by stringing together an odd number of inverters in a ring such that for an n inverter ring, the output of the nth inverter is tied to the input of the 1st inverter. If the inverters have sufficient voltage gain, any small transient noise source at any inverter input will be quickly amplified to the full output swing of the inverters. When there are an odd number of inverters, no stable operating point exists and the voltage level at each inverter input will oscillate with a period determined by the propagation delay through the entire ring. Assuming a propagation delay through a single inverter of τ_{inv} then a ring of n inverters will oscillate with a period $\tau_{ring} = n\tau_{inv}$ where n is an odd integer. The propagation delay can be estimated [130], using:

$$\tau_{inv} = 2\frac{\Delta V \times C_L}{I_C} \tag{5.22}$$

where ΔV is the output voltage swing, C_L is the capacitance at the output node, I_C is the current charging the output capacitance and τ_{inv} is the propagation delay through a single inverter defined as average of the time required to charge C_L from 10% to 90% of ΔV and the time required to discharge C_L from 90% to 10% of ΔV.

The specific inverter design used in this work is a simple common source amplifier with an active load consisting of a diode connected DFET shown in Figure 5.22. The small signal gain for this inverter is:

$$a = -gm_{EFET} x (r_{o,EFET} \| r_{o,DFET}) \quad (5.23)$$

Using gm_{EFET} = 350μS from Figure 5.13, $r_{0,EFET}$=50kΩ and $r_{0,DFET}$=40kΩ from Figure 5.8 gives a gain of about 10 at 300K.

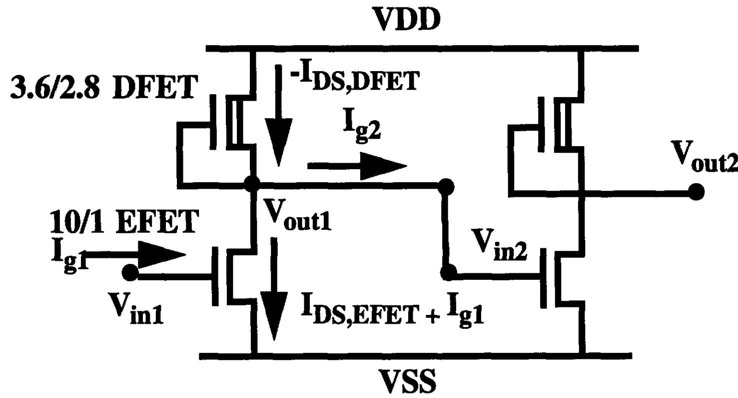


Figure 5.22: Schematic two inverter section of Ring Oscillator.

The output characteristic of each inverter is easily found by applying Kirchoff's Current Law to (KCL) to the output node of inverter.

$$-I_{DS,DFET} = I_{DS,EFET} + I_{g2} \quad (5.24)$$

and solving graphically for the intersections. Figure 5.23 shows the steps involved in this procedure as calculated using HSPICE for the specific inverter used in this project at 300K. From Figure 5.23 it is clear that the output voltage swing is determined by the Schottky barrier height which determines the voltage at which the gate diode turns on. At room temperature this gives $\Delta V \cong 0.7V$. The average current charging the output capacitor is approximately $I_C \cong 81\mu A$ as read off from Figure 5.23b. The load capacitor value includes the parasitic capacitance extracted from the layout $C_P = 8fF$, the gate capacitance of the following common source stage given by:

$$C_G = 1.2 \frac{fF}{\mu m^2} \times (W_{drawn} - \Delta W) \times (L_{drawn} - \Delta L) \qquad (5.25)$$

where $\Delta W = 0.2\mu m$, $\Delta L = 0.4\mu m$ and the drawn dimensions are given in Figure 5.22 resulting in $C_G = 7fF$ and the drain capacitance of the inverter EFET and DFET estimated to be one half of the gate capacitance given in Equation 5.25 or $C_{GS} \approx C_{GD} = 3.5fF$ for a total load capacitance of $C_L = 22fF$. Plugging these numbers into Equation 5.22 gives a propagation delay for each inverter of $\tau_{inv} = 190ps$ or multiplying by the number of inverters, 23, the ring oscillator period is $\tau_{ring} = 4.4ns$ which is about 12% smaller than the HSPICE simulated value of 5ns as shown in Figure 5.24.

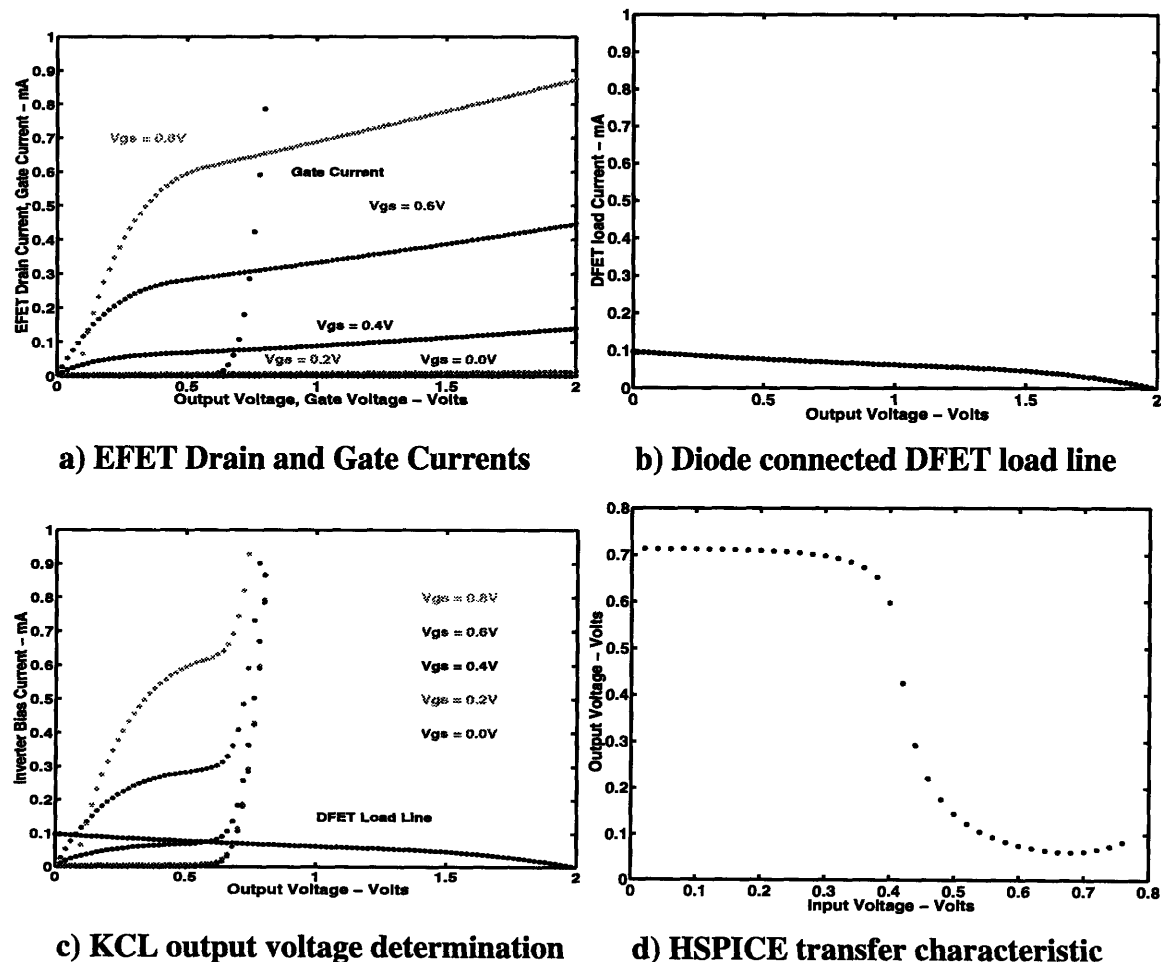


Figure 5.23: Graphical load line calculation of inverter characteristics.

This justifies the use of Equation 5.22 to calculate the change in gate capacitance as a function of temperature from the measured period of the ring oscillator

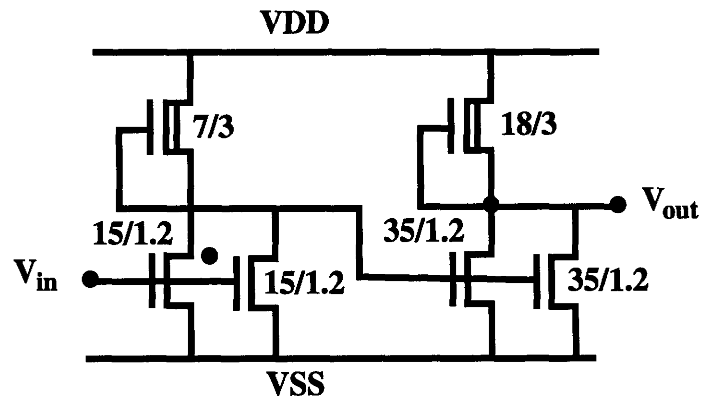


Figure 5.24: Ring Oscillator output stage.

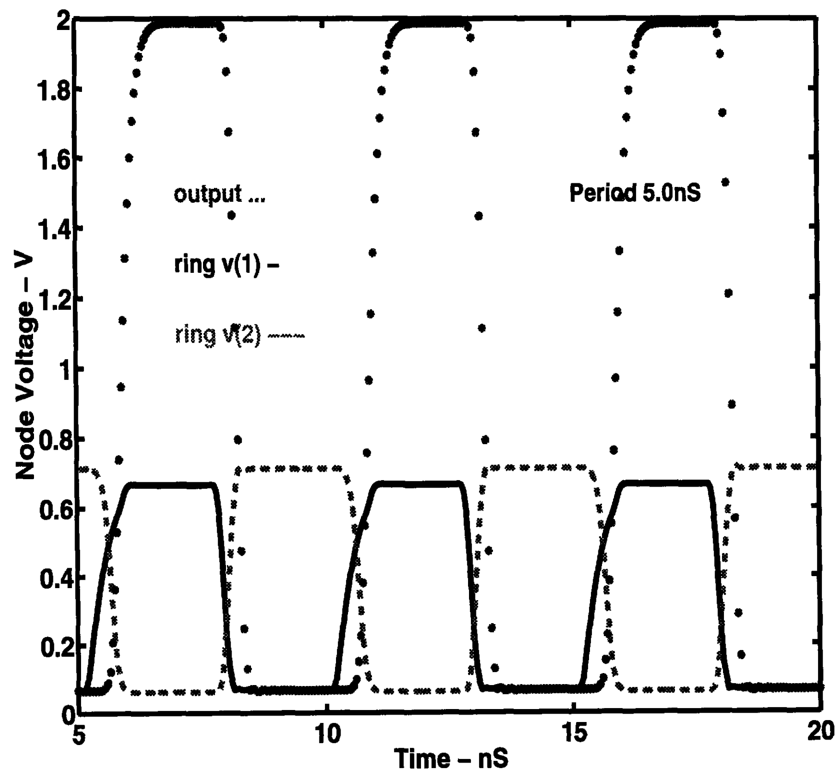


Figure 5.25: Ring Oscillator HSPICE Simulation.

. In addition to showing the input to the first two stages of the ring oscillator, labeled "ring(1)" and "ring(2)", Figure 5.25 shows the output of the an additional output stage which is attached in parallel to the output of the first stage in the ring oscillator and whos purpose is to both drive the large capacitances associated with going off chip, bond pads, bond wires, etc. and also to bring the amplitude of the output of the ring oscillator up to the full power supply range, in this case 2V. The schematic for the output stage is shown in Figure 5.24. It is this output that will be measured when actually measuring the ring oscillator frequency as a function of temperature.

5.4.2 Gate Capacitance extraction

Gate capacitance was extracted as a function of temperature by bonding the ring oscillator into a 14-pin DIP package and mounting into the temperature controlled Helium cryostat described above. The output from the ring oscillator was fed into a fast oscilloscope and the period of the oscillations was measured by hand. The DFET load current was taken from Figure 5.11 and scaled to the appropriate W/L ratio. The output voltage swing was taken as $0.7V$ at 300K and increase by $1.5mV/K$ [119]. Equation 5.22 was then used to extract the load capacitance. The results are given in Table 5.4:

Temperature (K)	Period (ns)	ΔV (V)	Current (μA)	Capacitance (fF)
30	4.42	1.11	85	14.7
50	4.40	1.08	82	14.5
80	4.24	1.03	80	14.3
100	4.20	1.0	78	14.2
300	3.10	0.7	73	14.0

Table 5.4: Gate Capacitance Extraction

From the above analysis it is clear that even at the lowest temperatures, the load capacitance changes by only 5% from it's value at room temperature which, taking into account that 40% of this load capacitance is due to layout parasitics means that the gate capacitance changes by less than 10% over the range from room temperature to the operating temperature of the QWIP FPAs.

5.5 HSPICE parameter extraction:

To allow accurate modeling of MESFET circuits to be used in FPA readouts, the above cryogenic temperature measurements were combined and used to fit the I-V characteristics of individual devices in the temperature range from 30K to 100K. The proceedure used was as shown in Figure 5.26. The results from this analysis for the 20µmx5µm EFET and 20µmx5µm DFET are sumarized in Tables 5.5 amd 5.6. It can be seen from the data that within the 30K to 100K operating range for QWIP FPA's, the parameter variation is within the measurement error and thus a single set of parameters for each device size and type should be sufficient to cover the entire range.

1. Adjust source and drain resistance parameters, **rsh** and **rshl**, based on measurements of Section 5.3.3
2. Adjust gate capacitance parameter, **gcap**, based on Section 5.4.2.
3. Adjust Schottky diode ideality factor, **n**, and Schottky diode saturation current, $\mathbf{i_s}$, based on Section 5.3.4. Note: for temperatures below 150K, the saturation current is to small to enter into HSPICE, so it was replace with zero. This has no effect for gate biases less than about 1V.
4. Set temperature related parameters, **bex, tcv, trs** and **trd** to zero.
5. Set temperature, using **.temp**, to appropriate temperature
6. Fit remaining parameters, **vto, beta, gamds, lamda** and **alpha** to measured I-V characteristic at temperature of interest. Note: **vto** and **beta** should be consistent with values measured in Section 5.3.2.

Figure 5.26: HSPICE Parameter Extraction Proceedure

Section 5.5 - HSPICE parameter extraction:

Parameter	30K	50K	70K	80K	100K	300K
fit	1.3	2.1	5.7	2.3	12.5	1.6
vto (mV)	496	490	484	500	480	341
beta ($\mu A/V^2$)	398	487	489	430	425	250
gamds(mV^{-1})	-9.2	-3.5	-0.1	-7.7	-1.7	7.6
lamda (mV^{-1})	100	100	116	120	128	101
alpha (V^{-1})	6.76	4.75	4.0	5.1	6.0	6.1
rsh Ω	168	168	169	169	170	210
rshl $k\Omega$	4.0	4.0	4.1	4.1	4.1	5.0
n (#)	4.73	3.26	2.44	2.30	1.89	1.07
is ($fA/\mu m^2$)	0	0.0	0.0	0.0	0.0	0.5
gcap ($fF/\mu m^2$)	1.26	1.24	1.23	1.23	1.22	1.20

Table 5.5: HSPICE Parameters for 20μm x 5μm EFET

Parameter	30K	50K	70K	80K	100K	300K
fit	5.9	3.7	8.9	3.7	3.8	58.3
vto (mV)	-503	-500	-508	-500	-501	-645
beta ($\mu A/V^2$)	285	286	281	286	286	194
gamds(mV^{-1})	0	0	0	0	0	0.0
lamda (mV^{-1})	69	67	71	64	61	30
alpha (V^{-1})	2.1	2.4	2.0	2.4	2.4	3.5
rsh Ω	159	159	160	160	160	210
rshl $k\Omega$	2.3	2.3	2.4	2.4	2.4	3.0
n (#)	4.73	3.26	2.44	2.30	1.89	1.16
is ($fA/\mu m^2$)	0	0.0	0.0	0.0	0.0	9.2
gcap ($fF/\mu m^2$)	1.79	1.77	1.75	1.74	1.72	1.70

Table 5.6: HSPICE Parameters for 20μm x 5μm DFET

5.6 Conclusions on Cryogenic Characterization of GaAs MESFETs

Measurements have been performed on a set of three enhancement mode and depletion mode GaAs MESFET's to determine their operating characteristics in the 30K to 100K temperature range appropriate for FPA operation. All the devices performed well even after repeated temperature cycling indicating that the dielectric cracking and void formation problems evident in high temperature processing [122] are not a problem for low temperature operation even though the temperature differences from room temperature are similar, about 200K. DC measurements have been used extract HSPICE model parameters for both EFETs and DFETs fabricated using the Vitesse HGaAsIII process. The results show that Vitesse Semiconductor's refractory gate VLSI GaAs MESFETs are ideal for fabrication of monolithic FPA readout electronics.

Chapter 6

Monolithic Integration of Normal Incidence QWIPs

Fabrication of high quality, high uniformity quantum well intersubband photodetectors is only one step in the process of fabricating large format focal plane arrays. The individual detector pixels must then be biased and connected to a read-out integrated circuit (ROIC) which converts the current from each QWIP pixel into a digital number proportional in the light intensity at the pixel and formats the data from all the pixels into a bit stream for transfer off-chip to a signal processing and display unit. In linear arrays or small two dimensional arrays with a few 10's of elements or less, individual pixels can be wire bonded and the signals taken immediately off chip. However, in large format FPAs the total number of pixels varies from 10's of thousands to well over a million per FPA chip. Furthermore, to maximize the sensitivity of the FPA the photocurrent from each pixel must be integrated separately and simultaneously with the photocurrents of all the other pixels in a so-called *staring* array format. The large number and density of pixels as well as the requirement for simultaneous current collection requires that staring FPAs have a capacitor to collect the photocurrent and a simple readout circuit to sample the charge on the capacitor for each pixel. At the start of a frame collection time the capacitor for each individual pixel is *dumped* by grounding both plates. The current is then allowed to collect on the capacitor for one frame time causing the voltage across the capacitor to *ramp* at which point the voltage on the capacitor is *sampled* and the process repeats at a rate typically between 30 and 300Hz. In the current state-of-the-art infrared cameras, the "dump-ramp-sample" ROIC is fabricated in Silicon CMOS because of the availability of low leakage, high capacitance per unit area capacitors and high speed digital and analog circuits. The Silicon ROIC is then bonded to the GaAs based QWIP FPA chip using Indium bumps at each pixel as shown schematically in Figure 6.1. This hybridization process can be made to work as evidenced by the 256x256 Lockheed-Martin QWIP FPAs described in

Section 2.3 however, there are potentially significant problems with manufacturability and reliability in hybrid FPA systems.

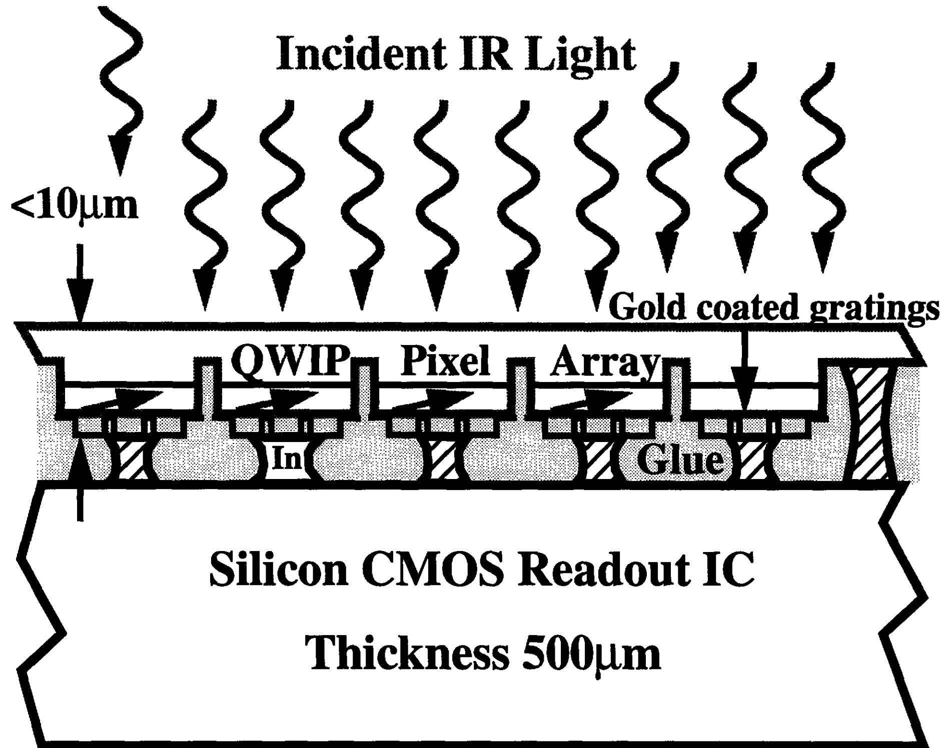


Figure 6.1: Schematic of Hybrid Focal Plane Array.

Having built a theoretical understanding for TE mode QWIPs in Chapter 2, demonstrated high quality TE mode QWIP pixels in Chapters 3-4 and characterized the low temperature performance of refractory metal gate VLSI GaAs MESFETs in Chapter 5 it is the focus of Chapter 6 to look ahead and discuss the issues involved with fabrication of complete IR camera subsystems including both the QWIP FPA and the ROIC. In particular, problems associated with hybridization and then two possible solutions to these problems both using monolithic integration of QWIP FPAs on VLSI GaAs MESFETs will be discussed. Section 6.1 discusses problems associated with hybrid FPA systems. Section 6.2 gives a short description of the epi-on-electronics (E-o-E) integration technique developed by Shenoy and Fonstad[1,131] as it applies to QWIP FPAs and with particular attention paid to the critical well cleaning procedure developed by the author to improve reliability

in the E-o-E process. Section 6.3 describes the selective area wafer bonding (SAW) technique proposed by the author to work in combination with the E-o-E technique to address specific issues of integration peculiar to QWIP FPAs. Finally, Section 6.4 summarizes the discussion on monolithic integration of QWIP FPAs and looks to possible avenues for further research.

6.1 Introduction to Integration of QWIPs

The primary difficulty associated with hybridization using Silicon CMOS ROICs is due to the thermal expansion coefficient difference between the GaAs substrate used for the QWIP FPA and the Silicon substrate used for the ROIC in combination with the low operating temperature of the QWIP FPA. The thermal expansion coefficient of GaAs is $6.9 \times 10^{-6}\ K^{-1}$ which is more than twice the thermal expansion coefficient for Silicon which is $2.6 \times 10^{-6}\ K^{-1}$.

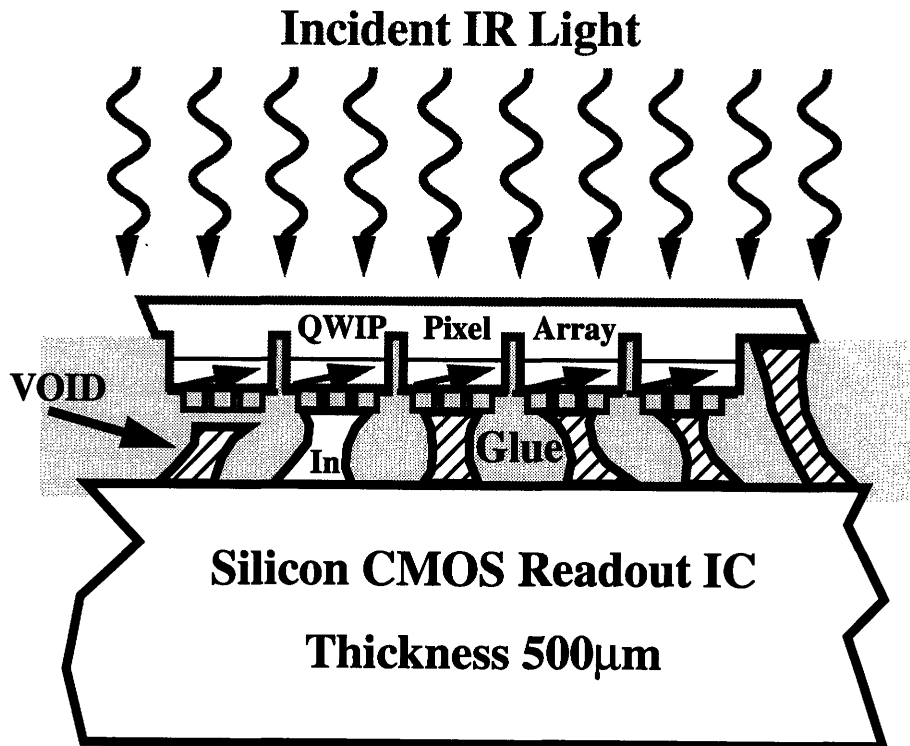


Figure 6.2: Problems with Hybridization - unzipping due to thermal expansion coefficient mismatch between GaAs, $6.9 \times 10^{-6}\ K^{-1}$ and Silicon, $2.6 \times 10^{-6}\ K^{-1}$.

This 19μm length difference implies that pixels near the periphery of the QWIP FPA which were originally aligned on top of their respective ROIC pixels are displaced by about 9.5μm or 20% of the pixel size when at the QWIP FPA operating temperature. The stress induced by this strain results in formation of voids or *unzipping* of the QWIP FPA from the Silicon ROIC as shown schematically in Figure 6.2. To combat the unzipping problem, QWIP FPAs are often thinned either by lapping or by selectively etching away the entire substrate to thicknesses on the order of less than 10μm. This allows deformation of the QWIP FPA in the vertical as well as horizontal directions lessening, but not eliminating the unzipping problem. In addition, the thinning process adds significant complexity to the processing of the QWIP FPAs and is one of the more important yield limiting factors.

One alternative to using hybridization is monolithic integration which uses the same substrate for both QWIP FPA and ROIC. Monolithic integration is not plagued by any thermal expansion issues because both the QWIP FPA and the ROIC expand and contract at the same rate with changes in temperature. Monolithic integration thereby eliminates both the bump bonding and wafer thinning steps used in hybridization. As discussed in Section 1.6.2, for monolithic integration with Silicon CMOS ROICs Platinum Silicide Schottky barrier MWIR detectors and Iridium Silicide Schottky barrier LWIR detectors are a natural choice, however their relatively low quantum efficiency compared even to TE mode QWIPs, 0.1% for IrSi in the LWIR verses 3.4% for TE mode QWIPs in the LWIR, means that FPAs based on monolithic integration on Silicon substrates are limited in performance. Schottky barrier detectors also tend to operate at very low temperatures, <40K, requiring more expensive and cumbersome cooling arrangements. To take advantage of higher performance GaAs based QWIP FPAs the ROICs must also be fabricated in GaAs. This can be achieved by using GaAs based MESFETs such as those fabricated by Vitesse Semiconductor and characterized for cryogenic temperature performance in Chapter 5. At the operating temperatures for QWIP FPAs the Schottky barrier gate also forms a low leakage, high capacitance per unit area capacitor suitable for use as the integration capacitor in dump-ramp-sample GaAs based MESFET ROICs.

Assuming a QWIP FPA operating temperature of 80K with a pixel pitch of 50μm and an array size of 256x256 pixels, the difference in the thermal expansion coefficients for Silicon and GaAs of $4.3x10^{-6}$ K^{-1} combined with a temperature difference of 350K between the melting point of Indium at 157C and the operating point of the QWIP FPA at 80K results in a length difference of 19μm between the GaAs and Silicon chips.

There are several possible avenues to integration of QWIP FPAs and MESFET ROICs on GaAs. For the highest performance one might consider first growing the QWIP FPAs on a GaAs substrate and then growing a semi-insulating cap region into which the GaAs MESFETs would be fabricated on top of the QWIP FPAs. This "stacked device" technique has the advantage that the optical and electronic circuits are vertically integrated, but requires dedicated, high yield and very expensive fabrication facilities for both the QWIP FPAs and the MESFET ROICs. One alternative called Epi-on-Electronics (E-o-E) has been developed by Fonstad and Shenoy[132] which takes already existing commercially fabricated VLSI scale GaAs MESFET circuits and grows optical devices on top of these circuits using MBE. This technique allows the use of relatively low cost, high quality GaAs MESFETs and thereby leverages the knowledge and capital built up in fabricating GaAs MESFETs for high speed electronic applications. A short description of this technique as it applies to QWIP FPAs will be described in Section 6.2. The second alternative to the stacked device technique is based on selective area wafer bonding, (SAW) and is proposed by the author to be used in combination with the E-o-E technique to specifically address the issue of fill factor, the ratio of optically active pixel area to total pixel area, which is somewhat compromised by the planar E-o-E integration technique. The SAW integration technique is discussed in Section 6.3. Section 6.4 does some comparison between the monolithic integration options and attempts to layout a path for future work.

6.2 Epi-on-Electronics integration technique for QWIP FPAs

The epi-on-electronics (E-o-E) technique is an attempt to leverage the large capital and intellectual investments of the VLSI MESFET industry to quickly and economically fabricate optoelectronic circuits. This technology allows integration of optical devices on

the medium scale integration (MSI) level, ~thousands of optical devices per square centimeter, with VLSI level, ~millions of electrical devices per square centimeter, electrical devices. Production of an E-o-E optoelectronic integrated circuit (OEIC) is done in four steps: design of VLSI GaAs MESFET electronic circuits using standard layout software (e.g. Cadence, Mentor Graphics, ...) with selected areas reserved for optical devices and appropriate simulations using foundry supplied HSPICE parameters, commercial GaAs foundry fabrication of electronic circuits, epitaxial growth of optical devices in the reserved areas on the same GaAs substrate used for the electronic devices and finally optical device fabrication and interconnection with electronic circuits. The first two steps, up to the point where completed electronic circuits have been fabricated are equivalent to the fabrication of an electronic GaAs IC and in fact all of the circuits used at MIT have been fabricated alongside, on the same wafer, purely electronic circuits designed by other universities through the ARPA/NSF sponsored MOS Implementation System (MOSIS) circuit broker. Only after the circuits have been fabricated and diced into chip form does MIT begin the remaining epitaxy and optical device formation steps which result in a completed OEIC. Figure 6.3 shows a schematic of the E-o-E process flow used by K.V. Shenoy in the fabrication of OEICs with LEDs being the optical devices. Figure 6.3a shows the electronic chip as received by MIT after circuit fabrication and dielectric growth well cleaning done at Vitesse Semiconductor. Figure 6.3b shows the same chip after MBE growth of the LED. In the region of the growth well where the crystalline substrate has been exposed high quality epitaxial material is grown which is suitable for fabrication of optical devices, in this case LEDs. In the region above the electronic circuits where the substrate is not exposed the MBE material growth is polycrystalline. Finally, Figure 6.3c shows the completed OEIC after removal of the unwanted polycrystalline material on top of the electronic circuits and interconnection of the electronic and optical devices. Growth wells as small as 50μm have been demonstrated implying a density of optical devices on the MSI level.

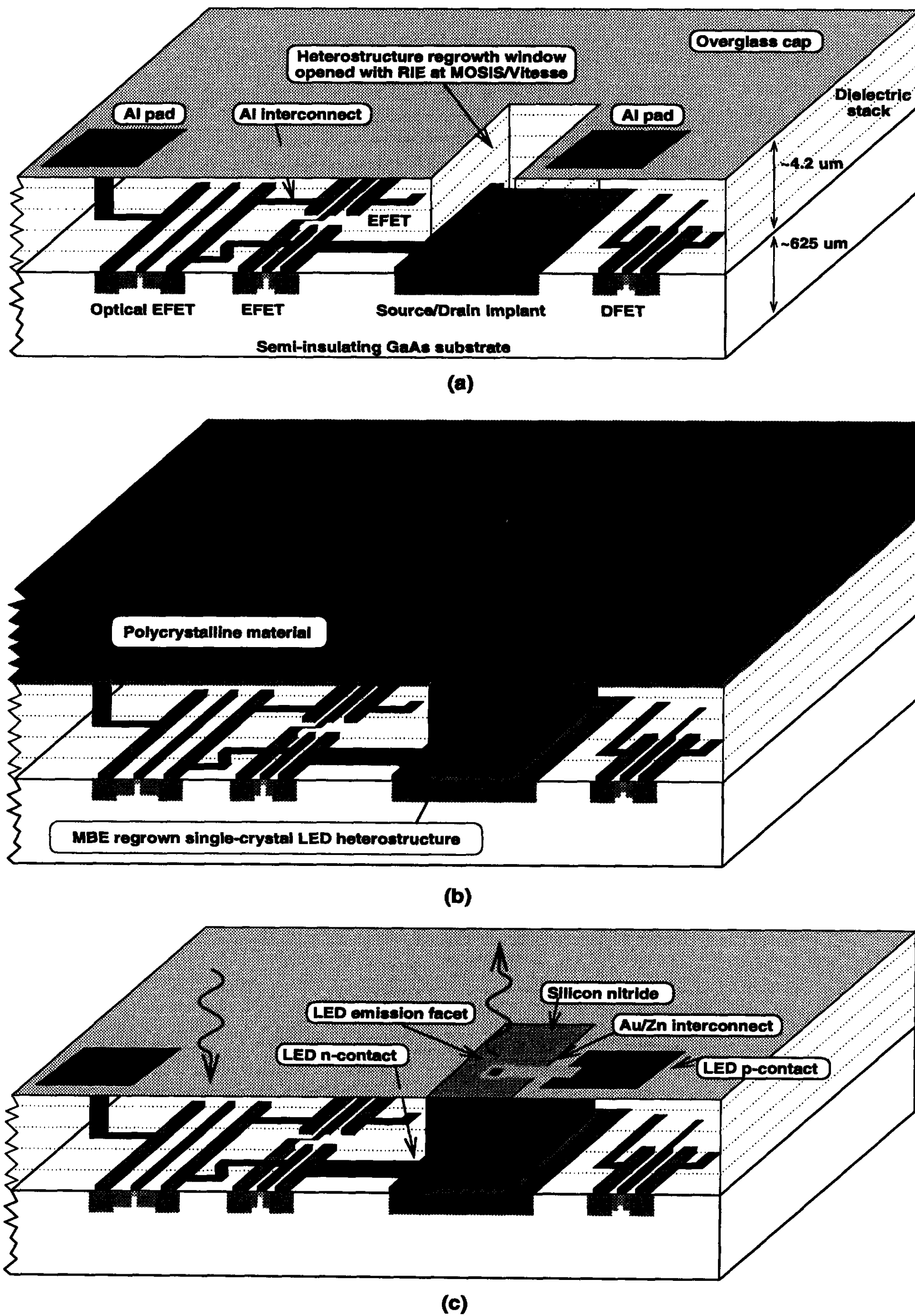


Figure 6.3: E-o-E OEIC fabrication flow showing Vitesse Semiconductor fabricated VLSI GaAs MESFET chip with cleaned growth well (a); same chip after MBE growth of LED structure resulting in crystalline material in growth well and polycrystalline material above circuits (b); completed chip after removal of polycrystalline GaAs and upper level metallization (c). Courtesy K.V. Shenoy

Success of the E-o-E integration technique depends on two characteristics of the electronic circuits; first, the circuits must be able to survive the temperatures of MBE growth for a period long enough to grow usable optical devices and second, the circuits must be designed to allow exposure of the undamaged GaAs substrate in regions where the optical devices are to be grown. The time/temperature characteristics of Vitesse fabricated refractory metal gate GaAs MESFETs have been investigated by Shenoy[131] and Braun[86] who found that a metallurgical reaction between the AlCu core and the WN_x clad of the electrical interconnects limits the time which a circuit can be exposed to elevated temperatures without permanent damage to approximately 5 hours at 470C as shown in Figure 6.4. The metallurgical reaction between the AlCu core and the WN_x clad is characterized by an Arrhenius characteristic with an activation energy of 3.5eV and linear thermal cycle time dependence with a saturation at an interconnect resistance 15 times the original value as shown in Figure 6.5.

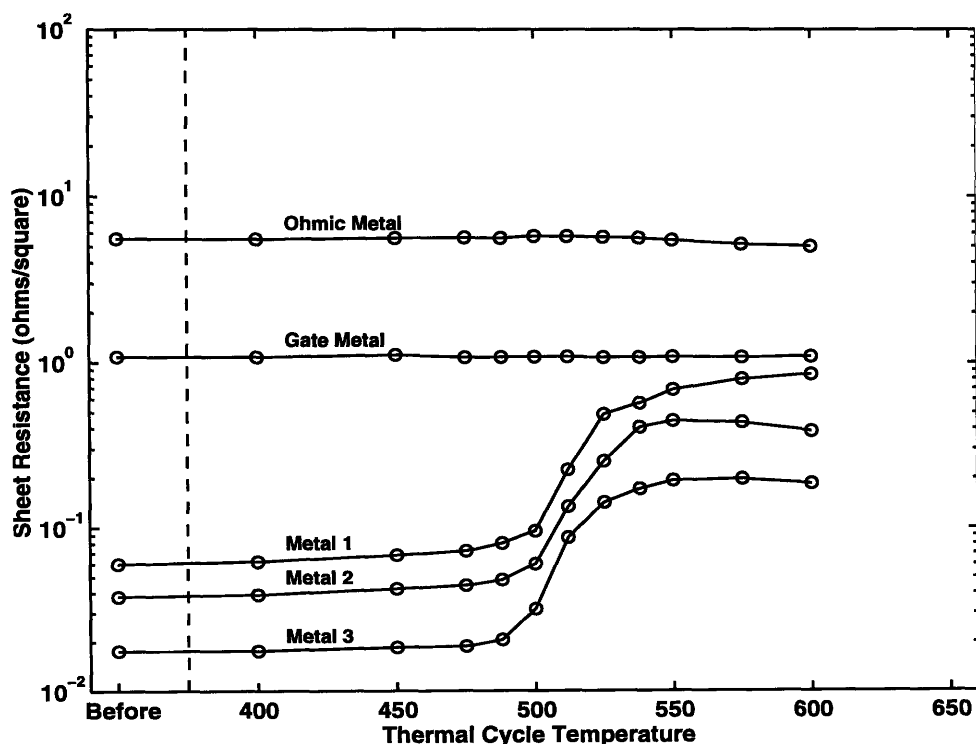


Figure 6.4: Vitesse HGaAs3 metal sheet resistance as a function of 5 hour thermal cycle temperature. Courtesy Eric Braun[122].

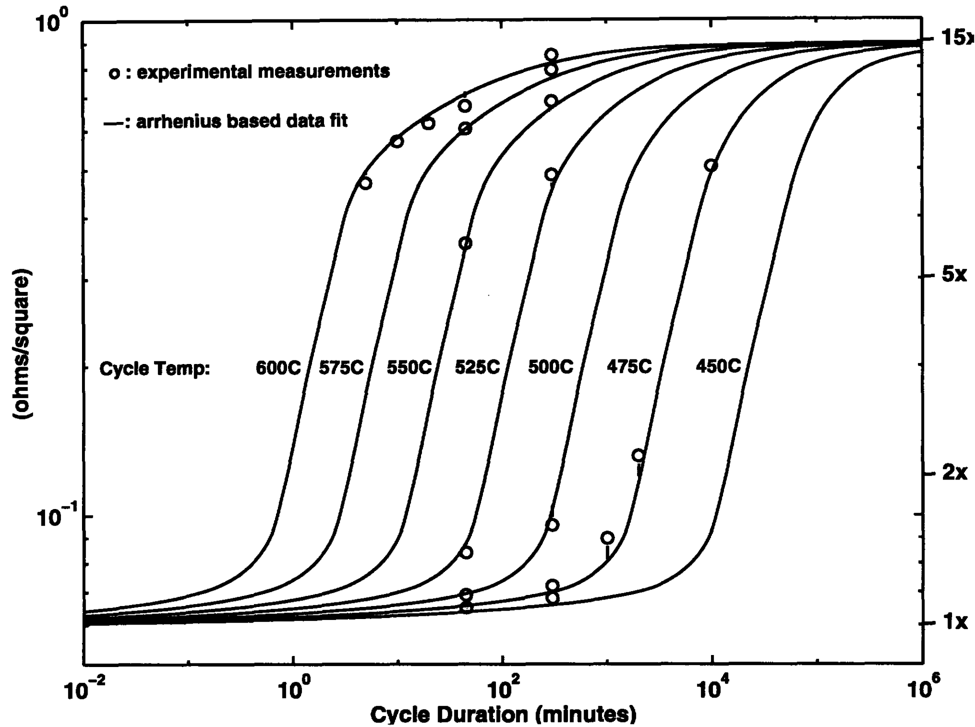


Figure 6.5: Arrhenius curve set with an activation energy of 3.5eV and experimentally measured Vitesse HGaAs3 metal 1 sheet resistance thermal cycle response. Courtesy Eric Braun.[122]

Five hours at 470C is sufficient to grow a wide variety of optical devices including QWIPs with total grown layer thicknesses of 5μm or less.

The performance of optical devices, both emitters and detectors, is critically dependent on the crystal quality of the device material which in-turn is dependent on the surface quality of the substrate. Design and preparation of the dielectric growth well (DGW) regions in the electronic circuits into which the optical devices will be grown is therefore critical to the success of the E-o-E integration technique. As originally developed by Shenoy[132] the growth wells consisted of about six microns of intermetal dielectric on top of a source/drain implant region which acts as the bottom interconnect between the electronic and optical circuits, Figure 6.6a. As the final step in the processing of these circuits at Vitesse Semiconductor a *scribe line* etch is done using RIE which removes this dielectric material and exposes the substrate. Prior to growth a solvent clean and ashing in an oxygen plasma remove any organic material left on the substrate by the RIE dielectric

etch process. A very short, 10 second, dip in buffered hydrofluoric acid removes the oxide of GaAs which forms on the surface of the substrate after exposure to air leaving a GaAs surface appropriate for MBE growth. This well cleaning process was used by K. Shenoy and others[133,134,135] to successfully integrate a variety of optical devices but was eventually found to result in surfaces with sufficient damage from RIE overetch of the GaAs substrate to limit uniformity and performance.

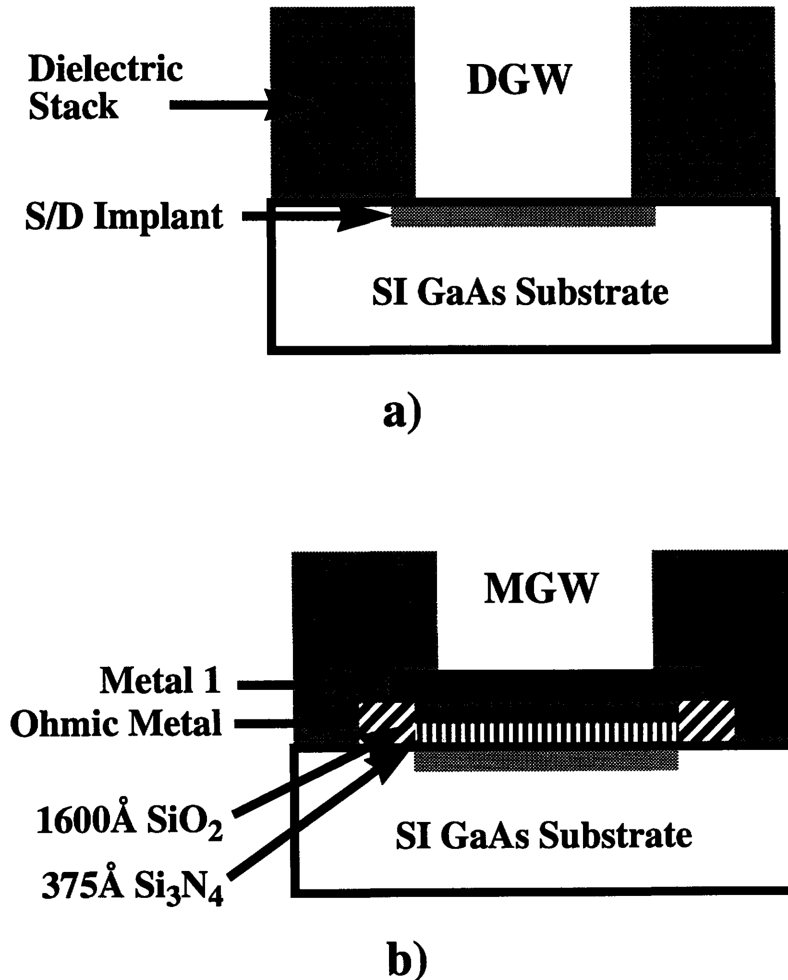
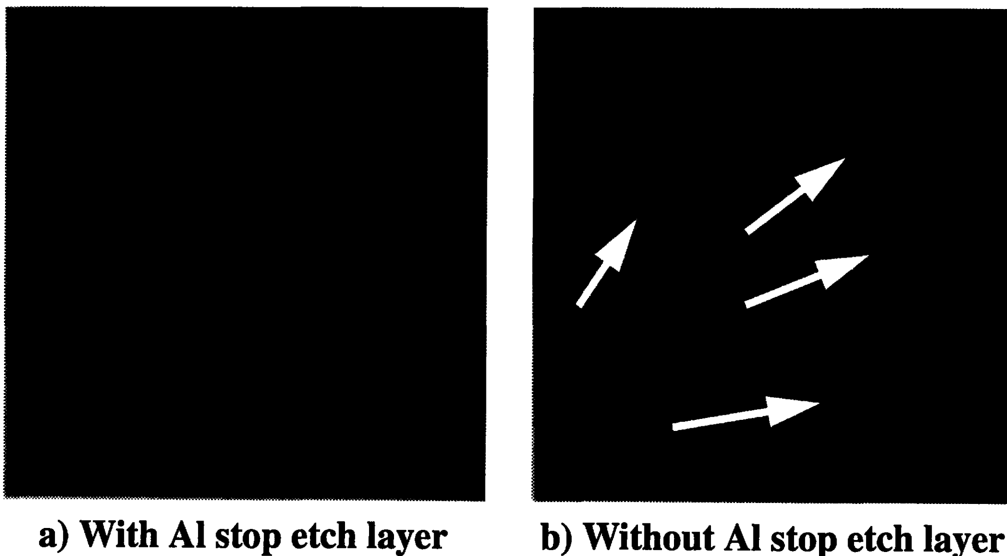


Figure 6.6: Schematic comparison of the dielectric growth well (DGW), (a) and metallic growth well (MGW) (b) structures as they appear after returning from Vitesse Semiconductor. For DGW pre-growth processing includes solvent clean and buffered oxide etch (BOE) etch. For MGW pre-growth processing includes solvent clean, metal 1 etch and BOE etch.

Because uniformity is particularly important in QWIP FPAs, the author proposed a new metallic growth well (MGW) design in which a metal 1 etch stop layer is introduced underneath the majority of the dielectric stack, Figure 6.6b. The scribe line etch at Vitesse then removes the dielectric down to metal 1 but stops prior attacking the GaAs substrate leaving behind metal 1, and a highly uniform dielectric layer consisting of 375Å of Si_3N_4 and 1600Å SiO_2. At MIT the MGW pre-growth preparation then consists of three steps; a solvent clean and ashing in an oxygen plasma to remove any organic material left on the substrate by the RIE dielectric etch process, removal of metal 1 with an Aluminum etch, either an HCl wet etch or a $SiCl_4/BCl_3$ RIE plasma etch with a mask covering everything except the MGW region to be cleaned in order to prevent damage to the exposed Aluminum bonding pads, and removal of the remaining thin dielectrics in buffered hydrofluoric acid to expose the pristine GaAs substrate.



a) With Al stop etch layer **b) Without Al stop etch layer**

Figure 6.7: Effect of including metal etch stop layer on crystal quality after. 1μm GaAs growth in growth well on VLSI MESFET chip. With Al etch stop layer, crystal shows no defects using Nomarski at 1000x. (a) Without Al etch stop layer, large defects due to RIE surface damage are clearly visible. (b) Note: small features on both photographs are due to microscope contamination. Growth and photographs courtesy J. Ahadian and S. Patterson.

The addition of an ohmic metal/metal 1 ring around the MGW as shown in Figure 6.6b acts as a lateral etch stop and prevents damage to the surrounding circuits during the buffered hydrofluoric acid etch. This new MGW process has recently been implemented in the design of MIT-OEIC4 by Joe Ahadian.[136] Figure 6.7 shows a comparison of the surface of two regions on the same OEIC chip with different growth well preparations after growth of ~1μm GaAs. Figure 6.7b shows the result for the original DGW process with defects caused by RIE damage to the GaAs substrate surface highlighted with arrows. Figure 6.7a shows the result for the new MGW process which has no such defects indicating that the metal 1 etch stop has successfully eliminated RIE damage and the resulting crystal defects in the epitaxial material.

A schematic for a QWIP FPA designed using the E-o-E integration technique is shown in Figure 6.8. Figure 6.8a shows a side view of the chip with the electronics indicated by left slanting hash marks and the quantum well active region of the QWIP indicated by a thin horizontal line in the MBE epi region. Note that the bias is applied between a bottom contact consisting of a source/drain implant fabricated as part of the Vitesse MESFET process and a top contact fabricated after MBE growth of the optical devices. In this design light is incident from the back side of the substrate which can be polished and coated with an anti-reflection layer for improved optical coupling. Figure 6.8b shows a top view of a single pixel in the QWIP FPA. The absolute size of the pixel is determined by the desired spatial resolution and operating wavelength of the QWIP FPA. The relative size of the components is determined by the desired dynamic range of the signal response, and the complexity of the ROIC. For the example of Figure 6.8b, the parameters for the Javelin seeker/tracker missile program have been used. This missile has a pixel pitch of 60μm, an operating wavelength between 8μm and 9.5μm, a dynamic range of 12bits and a frame rate adjustable from 30 to 300Hz. The 12bit dynamic range requires the storage of 17 million electrons per frame which, assuming 1V operating voltage requires a capacitance of 2.7pF. Using for the capacitance of a MESFET gate the value measured in Chapter 5, ~1.3fF/μm^2, gives a capacitor area of 2200μm^2 or approximately 61% of the pixel area.

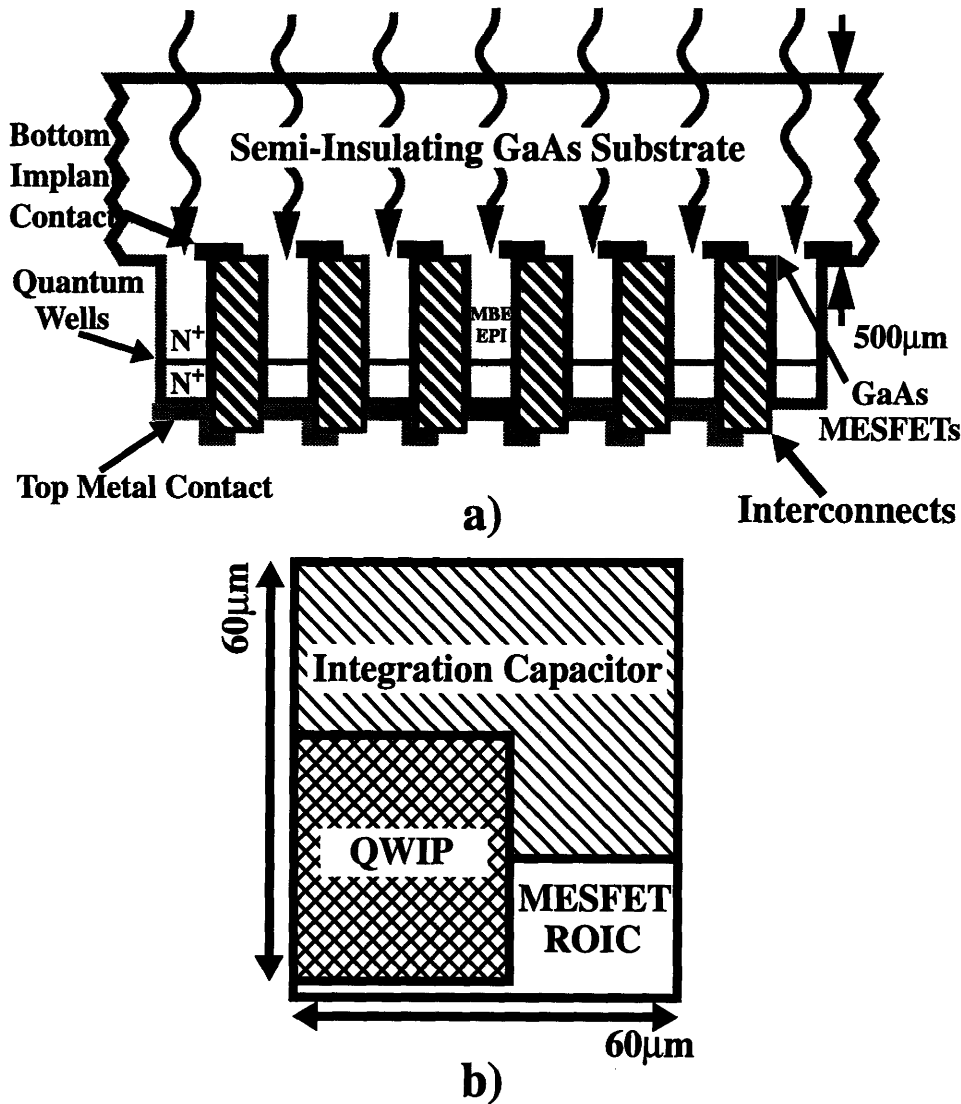


Figure 6.8: Schematic of QWIP FPA designed using E-o-E. Side view including backside illumination (a). Pixel layout drawn to scale for 12bit dynamic range, 2.7pF capacitor, MESFET gate capacitance of $1.2fF/\mu m^2$ (b).

This leaves only $1400\mu m^2$ for the QWIP photoconductor, the MESFET ROIC and appropriate interconnects. Assuming a minimum of $200\mu m^2$ for the MESFET ROIC leavs only about 33% or $1200\mu m^2$ for the QWIP photodetector. The fill factor, ratio of QWIP area to total pixel area, can now be calculated and for this example comes to only 33%. The small fill factor is an important drawback in the design of QWIP FPAs using the standard E-o-E technology which places optical and electrical devices next to each other on the same sub-

strate. This is in contrast to the hybrid technologies which stack optical and electrical devices on top of each other resulting in fill factors exceeding 90%.

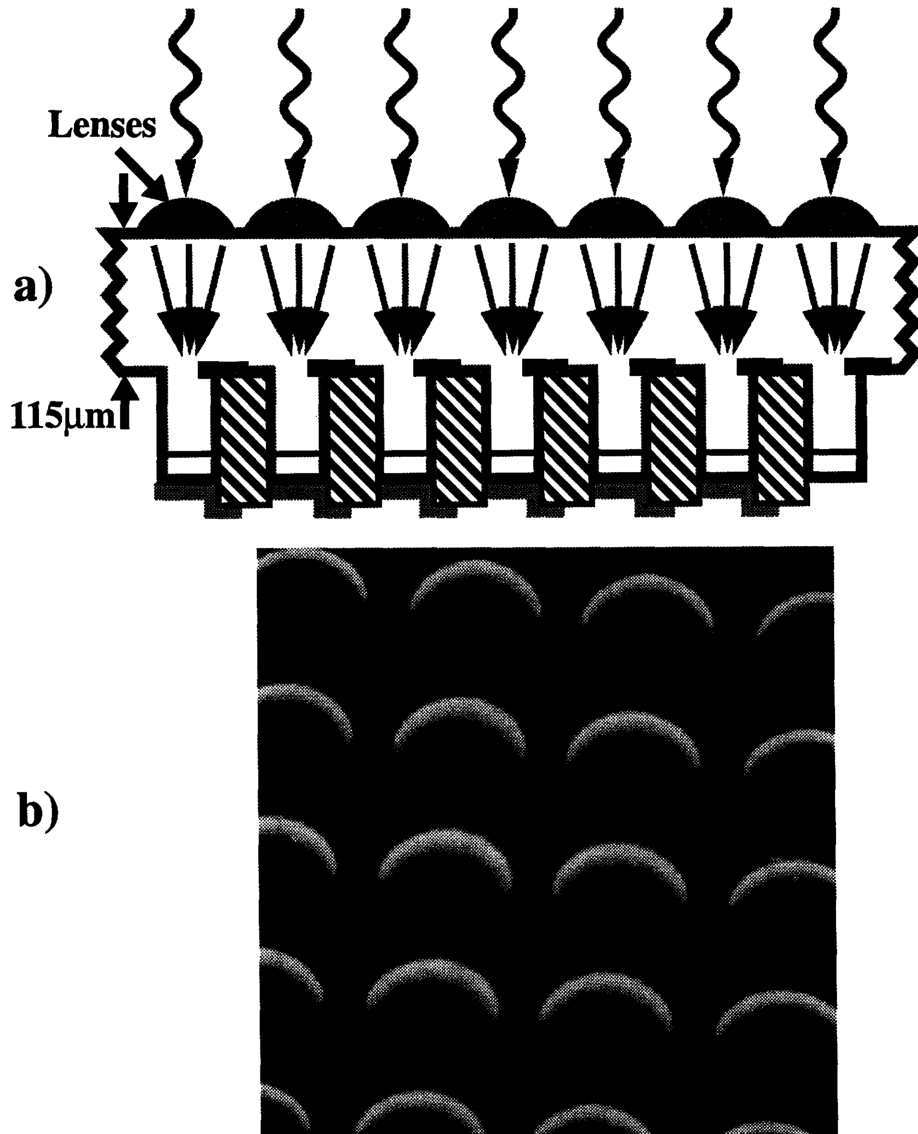


Figure 6.9: Schematic of QWIP FPA with lenses designed using E-o-E (a) and SEM of lens array from Smith et al. [137]

To combat the fill factor limitation the author has proposed using lenses that are etched into the back side of the QWIP FPA. This technique has previosly been proposed to improve the performance of MCT FPAs[138] and the formation of good quality lens

arrays has been demonstrated by Jones[137] using a resist flow and etch back technique. Figure 6.9a shows a schematic of the proposed QWIP FPA using E-o-E in combination with lenses etched on the back side of the OEIC substrate and Figure 6.9b shows an SEM of a lens array formed using this technique. The monolithically integrated lenses could potentially improve the fill factors to about 78%, $\pi/4$. In addition, because the lenses concentrate the light while reducing the active area of the QWIP, the background current in the QWIP which is proportional to the device area could be reduce by a factor of 3, for the QWIP FPA of Figure 6.8, resulting in an increase in the BLIP operating temperature of about 5C, a small but important advantage. One might consider that the premium placed on uniformity would put very tight fabrication tolerances on the lens array, however within the narrow bandwidth of sensitivity for the QWIPs, lenses with small fabrication non-uniformities should have different absolute coupling efficiencies, but negligible variation in coupling efficiency with temperature and spectral content of the scene. As discussed in Chapter 1 it is the variation with spectral content which puts such strict uniformity restraints on the QWIP devices themselves and the constant coupling coefficient of the lenses can be easily compensated for during array calibration.

In concluding the discussion of QWIP FPAs fabricated using the E-o-E integration technique, it should be emphasized again that monolithic integration of QWIP FPAs with their ROICs is extremely attractive. Improved manufacturability and reliability due to the elimination of the Silicon substrate and the entire hybridization process can potentially lead to smaller, cheaper and more sophisticated infrared cameras. E-o-E is particularly attractive because it leverages the investments of the VLSI MESFET industry to economically fabricate QWIP FPA OEICs. Issues such as the low fill factor inherent in the E-o-E process are important but soluble using techniques such as the integrated lenses discussed in this section. There is however a tremendous amount of work yet to be done to jointly optimize the performance of the electronic circuits and the QWIP FPAs and to demonstrate that the epitaxial material grown in the E-o-E growth wells is of sufficient quality and uniformity to produce large area QWIP FPAs. Recognizing this the author has also proposed a new monolithic integration technique called selective area wafer bonding

(SAW) which in combination with the E-o-E integration technique can dramatically reduce the requirements for high quality epitaxy in the E-o-E growth wells while at the same time providing fill factors greater than 90% and separately optimized QWIP FPA devices. This technique is the subject of Section 6.3.

6.3 Selective area wafer bonding integration technique

The selective area wafer bonding (SAW) integration technique is designed to build on the strengths of the E-o-E integration technique while at the same time providing for vertical integration and relaxing some of more severe requirements for high crystal quality in the low temperature MBE material. In particular, the greatest advantage of E-o-E over all previous integration methods is that it leverages the large capital and intellectual investments of the VLSI GaAs MESFET industry to quickly and economically fabricate OEICs using commercially fabricated VLSI GaAs MESFET circuits. There are two basic weaknesses in the E-o-E approach when applied to QWIP FPAs; the first being that the optical devices and electrical devices are integrated in the same plane, as shown in Figure 6.8b this leads to low fill factors for QWIP FPAs, the second weakness is that to prevent damage to the MESFET circuits during subsequent MBE growth the MBE must be done at low temperature, 470C verses the more typical GaAs growth temperature of 580C, requiring extreme care and special growth techniques to get good quality optical material. Recently there has been an impressive amount of progress in developing E-o-E, particularly with the use of Phosphorus based III-V semiconductors like $In_{0.50}Ga_{0.50}P$ which is lattice matched to GaAs but which can be grown with high quality on GaAs in the 450C to 470C temperature range.[139] A larger set of optical devices could be quickly added to the E-o-E repertoire however, if there were a way to leverage the large capital and intellectual investment in already existing III-V optical devices based on InGaAlAs, InGaAsP and other material systems that are preferably grown above 470C. Some of these devices like QWIPs which depend critically on carrier lifetime, recall from Equation 1.60 that D^* is proportional to the square root of the carrier lifetime, are particularly susceptible to possible degradations in material quality introduced because of the low growth temperature.

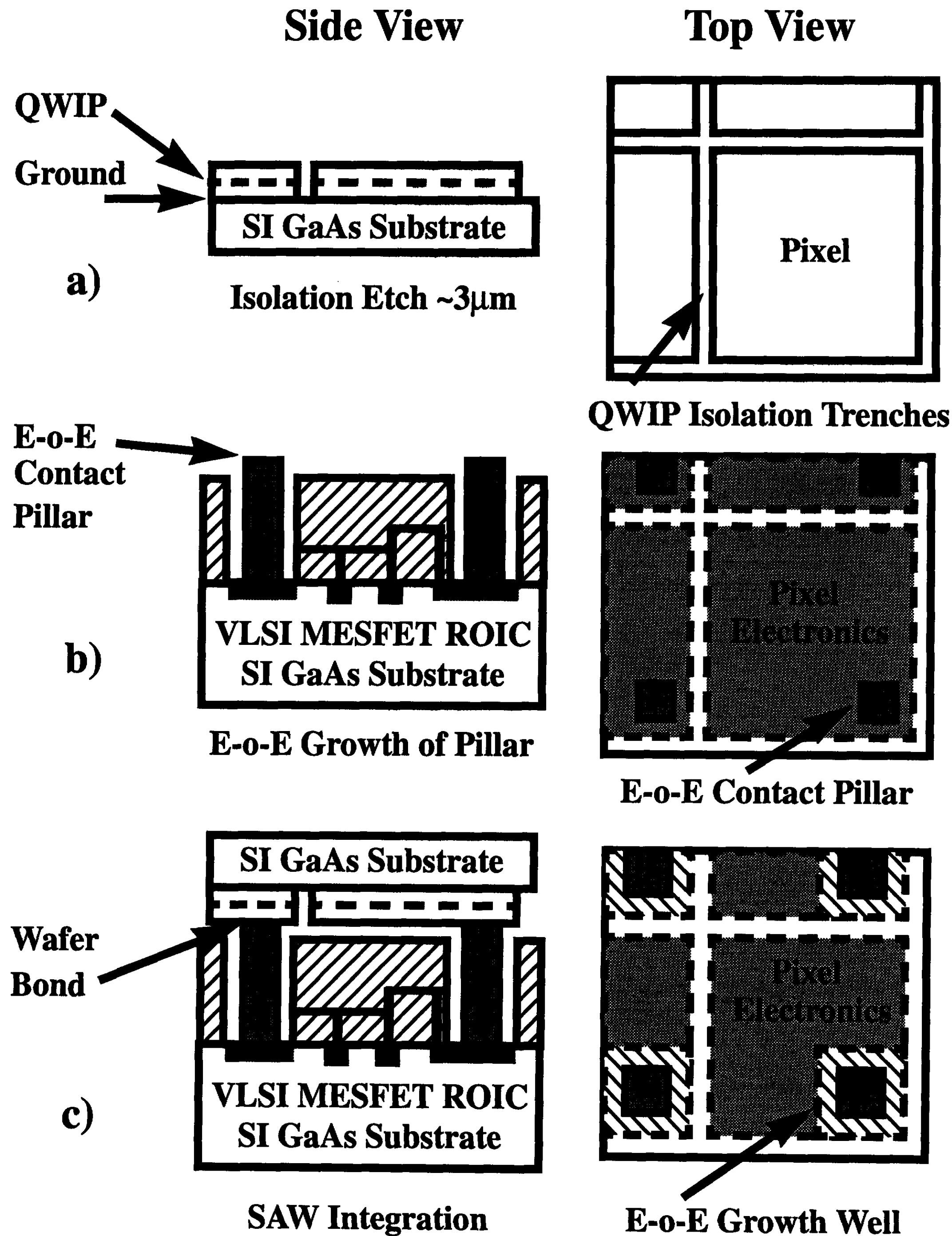


Figure 6.10: Schematic illustration of fabrication flow for E-o-E/SAW integration. QWIP FPAs are etched using either RIE or wet etching to define the pixel size and isolate the pixels (a). Pillars are grown on pre-fabricated MESFET VLSI ROICs using E-o-E (b). Finally, QWIP FPA is wafer bonded on to prepared VLSI GaAs MESFET ROIC (c). Illumination is from top into QWIP FPA.

Separate optimization and growth of the optical and electronic devices with no restrictions on optical device growth temperature is an important advantage of E-o-E/SAW.

The key to the SAW technology is wafer bonding. The idea behind wafer bonding is that by placing two pristine semiconductors in intimate contact and with the application of heat and pressure the two semiconductors can be made to bond.[140,141] This is a technique which has received a lot of attention lately in the Silicon analog device industry because it can be used to fabricate Silicon on Insulator (SOI) devices with highly linear characteristics due to the decoupling of the doped substrate and thus the elimination of voltage dependent substrate coupling. Wafer bonding has also been applied to III-V semiconductors. Wafer bonding is being used in production of high luminosity transparent substrate orange LEDs fabricated by Hewlett-Packard[142] and has been used to demonstrate 1.3mm GaInAsP lasers grown on InP substrates and then wafer bonded to a GaAs substrate.[143] Of particular importance here is that wafer bonds with excellent electrical characteristics determined by the band discontinuities in the constituent semiconductors have been obtained for wafer bonding temperatures as low as 450C for 30 minutes. This is well within the time/temperature tolerance of the Vitesse refractory gate VLSI GaAs MESFET circuits as shown in Figures 6.4 and 6.5.

The combination of E-o-E with the SAW integration technique (E-o-E/SAW) proposes to keep the advantages of E-o-E while at the same time enabling vertical integration and freeing the active portion of the optical devices to be grown with no restrictions on growth temperature. A schematic diagram of the fabrication flow is shown in figure 6.10. Figure 6.10a shows a QWIP FPA grown on a semi-insulating GaAs substrate. This substrate is a blank epi-ready substrate as was used in this thesis and the QWIP FPA can be grown at whatever temperature and using whatever growth technique yields the best device performance. After growth the wafer is then processed with a single mask step which etches narrow isolation trenches between each pixel thus defining the spacial resolution and pixel format of the QWIP FPA. Figure 6.10b shows a companion ROIC chip which has been fabricated by Vitesse Semiconductor and contains the VLSI GaAs MESFET ROIC circuits along with a small, i.e. 20μm x 20μm, E-o-E growth well. The MGWs have been

appropriately cleaned and inside each growth well has been grown using E-o-E a contact pillar which has been etched to remove the surrounding polycrystalline material and to electrically isolate each pillar. Figure 6.10c then shows the processed QWIP FPA after wafer bonding to the processed E-o-E chip with contact pillars.

The only function of the contact pillar is to electrically connect the ROIC to the QWIP FPA. As such the crystal quality of the pillar material is not critical. With sufficiently high doping even highly dislocated material totally inappropriate for direct fabrication of optical devices will work as an excellent electrical conductor and after wafer bonding provide excellent contact to the QWIP FPA. The material used for the pillar can even be selected specifically to reduce the contact resistance at the wafer bonded interface through interdiffusion which can be enhanced by using Phosphide based materials with increased mass transport.

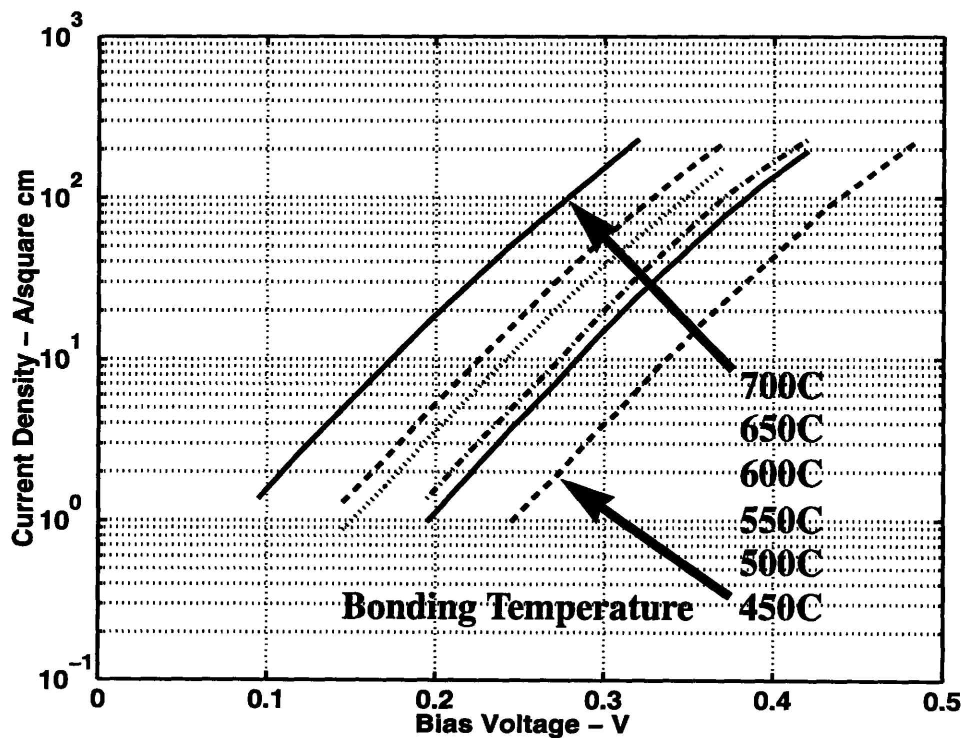


Figure 6.11: Forward IV characteristics of n-GaAs/n-InP wafer bonded samples. InP biased positive with respect to GaAs. Wafer bonding done at temperature for 30 minutes in Hydrogen with bonding force of 0.29 N/cm^2. Barrier heights are 0.46, 0.42, 0.40 and 0.37eV for bonding temperatures of 450C, 550C, 650C and 700C respectively.[143]

As a numerical estimate of the effect on QWIP performance of a wafer bonded interface, the data of Wada et al.[143] for the bonding of n-type InP to n-type GaAs can be used. Figure 6.11 shows the I-V characteristics of an n-type InP sample wafer bonded to an n-type GaAs sample at temperatures from 450C to 700C for 30 minutes in Hydrogen.

The data are consistent with a theoretical model based on thermionic emission over the barrier due to the conduction band discontinuity at the GaAs/InP interface. The barrier heights determined from the curves of Figure 6.11 are 0.46eV, 0.42eV, 0.40eV and 0.37eV for samples bonded at 450C, 550C, 650C and 700C with an ideality factor of 1.8. and assuming an effective Richardson constant of 8.4 Amps/(cm^2*K^2). Both wafers were bulk material doped 2x10^{18} cm^{-3}. All values are higher than the predicted conduction-band discontinuity of 0.3eV[144] in pristine GaAs/InP junctions presumably due to defects at the interface.

To calculate the voltage drop across such a GaAs/InP waferbonded interface in a QWIP the current density can be expressed from thermionic emission theory as:[112]

$$J = AT^2 \exp\left(-\frac{qV_B}{kT}\right)\left[\exp\left(\frac{qV_F}{nkT}\right) - \exp\left(\frac{qV_R}{nkT}\right)\right] \qquad \frac{Amps}{cm^2} \tag{6.1}$$

where A is the effective Richardson, 8.4 Amps/(cm^2*K^2), T is the QWIP operating temperature, V_B is the barrier height calculated from Figure 6.11, n is the ideality factor also calculated from Figure 6.11 to be 1.8, V_F is the portion of the applied voltage appearing in the GaAs which determines the forward bias characteristic and V_R is the portion of the applied voltage appearing in the InP substrate which determines the reverse bias characteristic. From Chapter 4 the BLIP current density for the TE mode QWIPs of this thesis requires approximately 0.1mA/cm^2 in the pillar which when inserted into Equation 6.1 along with the 450C waferbonding barrier height of 0.46eV and a BLIP temperature of 68K gives a voltage drop across the wafer bonding interface of 0.6V, significantly less than the QWIP operating voltage of 2V and therefore not a limit to the QWIP device performance. Homojunction, GaAs/GaAs, voltage drops, voltage drops across narrow band-gap materials like InAs/InAs and voltage drops across heavily doped materials in which

tunneling currents can be as large or larger than the currents due to thermionic emission are expected to be even lower. Near ideal waferbonding interfaces could be achieve by doing all the waferbonding within the post-growth UHV environment. In this case after processing, wafers to be bonded would be loaded back into the vacuum chamber for a Hydrogen plasma clean identical to that now used prior to MBE growth and then manipulated together and heat treated all within the UHV environment.

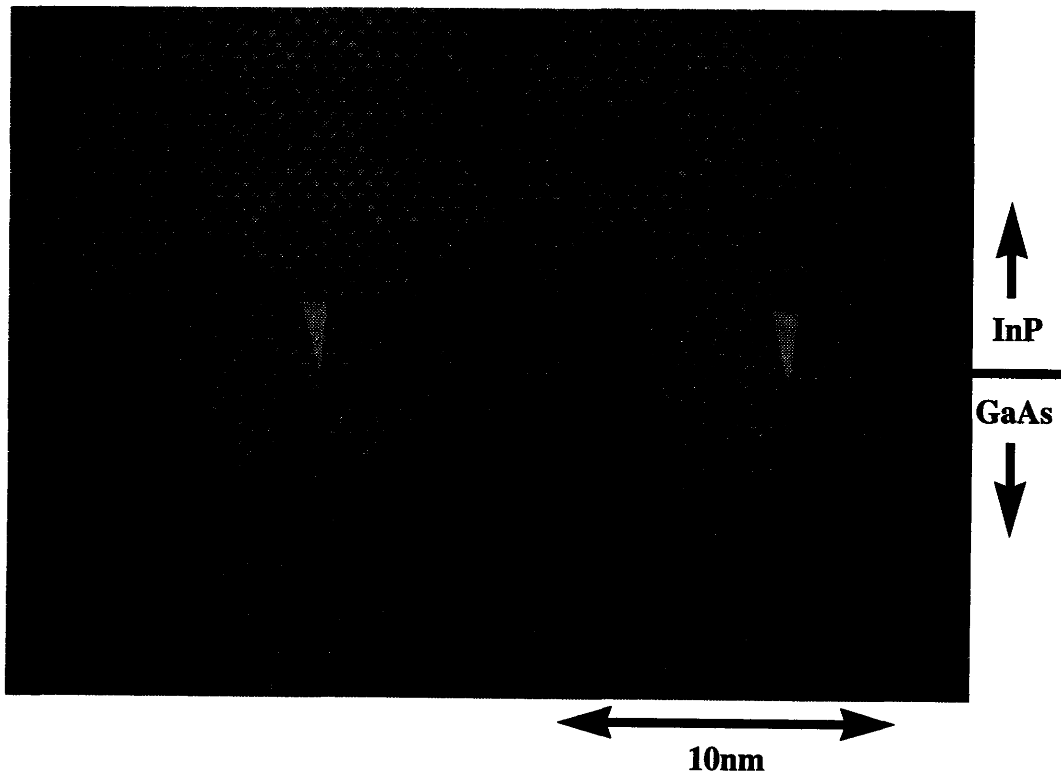


Figure 6.12: Cross-sectional transmission electron micrograph of GaAs/InP heterointerface bonded at 650C for 30 minutes. The misfit dislocations are indicated by arrows. From H. Wada.[143]

The quality of the wafer bonded interface on the atomic scale is indicated by the cross-sectional TEM from Wada[143] shown in Figure 6.12. The high resolution of the TEM image allows one to count the number of atomic planes between the two misfit dislocations, there are 27. This corresponds to a 3.7% lattice mismatch, exactly what one gets by calculating the lattice mismatch between bulk GaAs with a lattice constant of 5.6532Å and InP with a lattice constant of 5.8687Å.

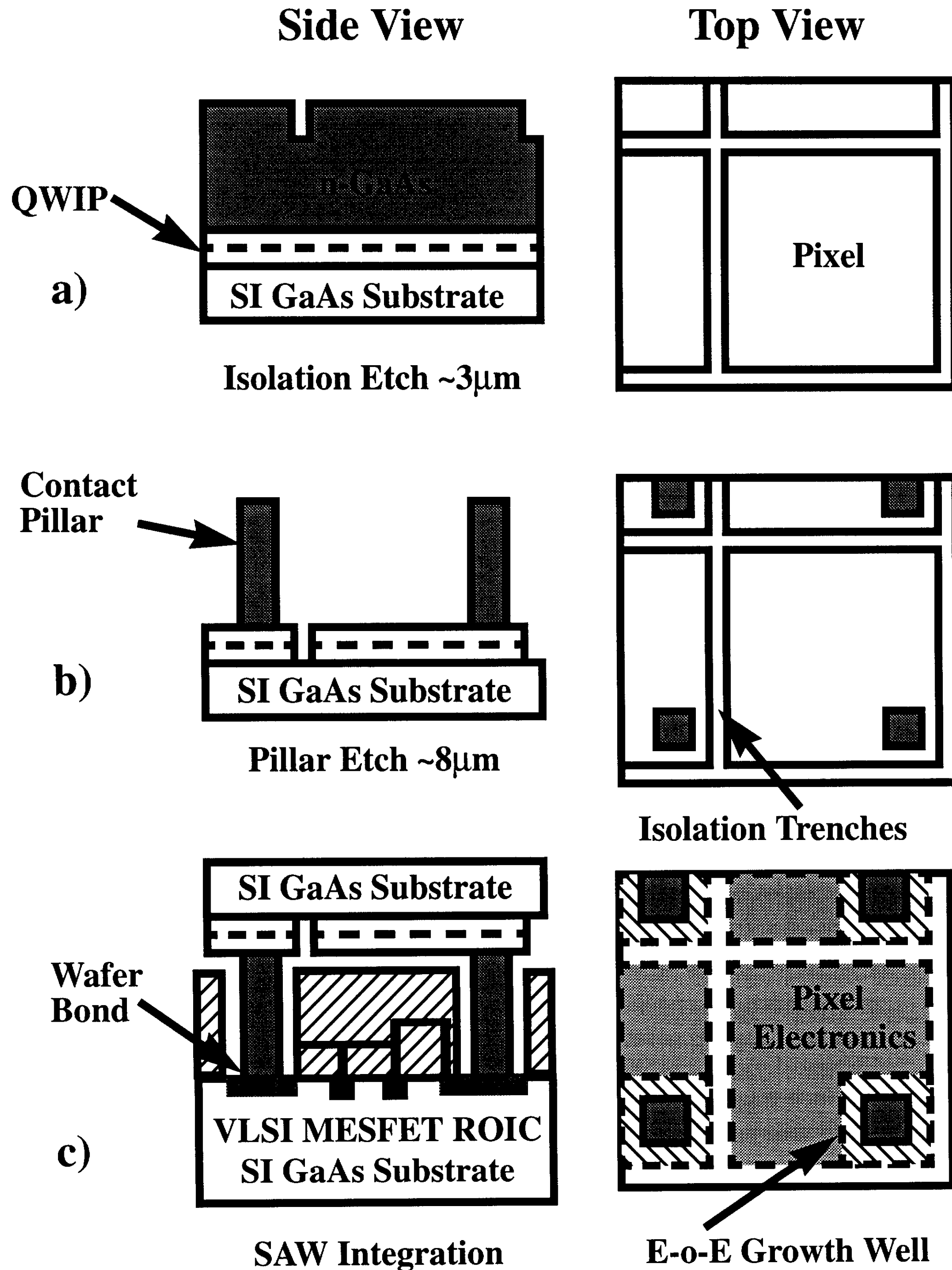


Figure 6.13: Schematic illustration of fabrication flow for SAW integration process. QWIP FPAs are etched using either RIE or wet etching twice; first to define the pixel size and isolate the pixels (a) and then to define pillars used for contact to MESFET ROIC (b). Finally, QWIP FPA is wafer bonded on to prepared VLSI GaAs MESFET ROIC (c). Illumination is from top into QWIP FPA.

Note also that there are no threading dislocations propagating perpendicular to the wafer bonded interface. These two characteristics, complete relaxation and zero threading dislocations are in fact the ideal characteristics of a grown heterointerface with thickness of the epilayer grown thicker than the critical thickness discussed in Chapter 3. The continuity of the atomic layers through the interface also implies that the wafer bonding process results in a single semiconductor crystal indistinguishable from an ideally relaxed grown heterointerface and therefore worthy of the name monolithic (undifferentiated) integration.

For some applications it may be preferable to limit or even eliminate the need for growing on top of the chip containing the circuits. In this case the wafer bonding can be done directly onto the cleaned growth well through the use of pillars etched into the mating optical device as shown schematically in Figure 6.13. Figure 6.13a shows a QWIP FPA with approximately 8μm of doped pillar material grown on top of the active device after pixel isolation trenches have been etched approximately 3μm down into the pillar material. Figure 6.13b shows the same QWIP FPA after masking of the pillars and a further 8μm of pillar material has been etched away leaving isolated pixels with one 8μm high contact pillar per pixel. Figure 6.13c shows the companion VLSI MESFET ROIC which has had the MGW cleaned to reveal the pristine substrate, but has not been grown on wafer bonded to the QWIP FPA. This time the wafer bond occurs on the substrate of the VLSI MESFET ROIC with the pillars acting as an electrical connection and physical spacer between the ROIC and the QWIP FPA.

In both the E-o-E/SAW integration technique and the pure SAW integration technique the optical and electrical circuits are vertically integrated allowing for pixel fill factors approaching 100%. In addition because any E-o-E material required is only used for electrical interconnection the requirements for good crystal quality are dramatically reduced and therefore the growth and fabrication conditions for the optical and electrical circuits can be separately optimized. In effect this allows E-o-E/SAW to leverage both the large investment of the VLSI GaAs MESFET industry and the large investment of the III-V epi-

taxy industry in a monolithic OEIC that has only semiconductor bonds throughout and is indistinquishable on an atomic scale from a single heterostructure.

6.4 Conclusions and future work on monolithically integrated QWIPs

In this chapter issues related to the integration of QWIP FPAs with the ROICs that are required to fabricate IR cameras have been discussed. Monolithic integration based on the epitaxy-on-electronics technique developed by Shenoy and Fonstad[132,121] has been compared to current hybrid integration techniques and found to have significant advantages. Finally, a new proposal for monolithic integration based on a combination of E-o-E and selective area wafer bonding has been put forward to address the issues of crystal quality and pixel fill factor which are weaknesses in E-o-E integration. This combination E-o-E/SAW integration keeps the best aspects of the E-o-E process while at the same time allowing independent optimization of the QWIP FPAs and VLSI MESFET circuits and improving pixel fill factors to above 90%.

The combination of E-o-E and SAW is very attractive for QWIP FPAs and other OEICs. Future work toward demonstrating the utility of this technique must start with a better operational understanding of the parameters which affect the wafer bonding interface. In particular, effects of doping, material composition and bonding temperature need to be characterized with an eye toward minimizing the time and temperature required to get high quality, low resistance wafer bonds. As shown in Figures 6.3 and 6.13 10μm square posts are to be aligned within 20μm square growth wells. This requires an alignment tolerance of +/- 5μm, tight, but still smaller than the +/- 1μm tolerance required for alignment of optical fiber lasers and single mode fibers. As integration density increases and the MGW size shrinks, the alignment technique used to align the two wafers to be bonded may become an important limiting factor. Fixtures for alignment underneath a microscope similar to those used in the alignment of laser diodes and single mode optical fibers need to be developed possibly with the addition of lithographically defined pins and sockets to aide in accurate placement. Finally, though the E-o-E/SAW technique has been discussed in terms of QWIP FPAs and GaAs VLSI MESFETs investigations aimed at

extensions to other optical material systems like MCT and possibly other electronic circuits like Silicon CMOS should be interesting.

Chapter 7

Conclusions and Possible Directions for Further Research

7.1 Thesis Accomplishments

Quantum Well Intersubband Photodetectors (QWIPs) are attractive devices to compete with Mercury Cadmium Telluride (MCT) narrow bandgap semiconductors for use in very large focal plane arrays (FPAs) because QWIPs can take advantage of both established technology for growing and processing GaAs optical devices and commercially available large area VLSI GaAs IC's. The key disadvantage of high detectivity n-type QWIPs has been their requirement for a secondary coupling mechanism, typically an etched grating or roughened surface, in order to couple normally incident TE polarized light into the TM active quantum wells. Though the etched couplers have been demonstrated to be very efficient, they add complexity to the fabrication process, non-uniformity to the FPAs and problems with cross-talk between adjacent array elements. In addition, all current FPAs use Silicon based multiplexer (MUX) circuits which require bump bonding of the MUX to the active array elements. This bump bonding process is inherently very low yield and suffers from "un-zipping" at cryogenic temperatures due to the factor of three mismatch between the thermal expansion coefficients of GaAs and Silicon. All of these factors limit system performance as measured by minimum resolvable temperature, MRT, and reduce yields greatly increasing finished array prices. This thesis demonstrates the first TE active normal incidence pseudomorphic n-type QWIP. By using TE active devices the need for couplers is eliminated simplifying the fabrication process, improving uniformity and eliminating coupler induced cross talk. The monolithic integration of the QWIPs with VLSI GaAs MESFETs in which the MUX is fabricated is discussed and a new monolithic integration technique is proposed called E-o-E/SAW which combines the strengths of the epi-on-electronics technique with those of an epi only integration process. The E-o-E/SAW technique uses wafer bonding to monolithically integrate optical devices onto an E-o-E

electronic circuit chip with epitaxially grown connecting pillars. This allows separate optimization of the growth for optical devices and connecting pillars and removes the requirement for high quality MBE to be grown at low temperature. The E-o-E/SAW technique also allows stacking of the electronic and optical circuits providing for QWIP FPA fill factors in excess of 90%. E-o-E/SAW should be generally applicable to any OEIC, but for QWIP FPAs it has special importance because it eliminates the need for bump bonding as well as all incompatibilities between the MUX and the active FPA. IR cameras fabricated from this technology should have state-of-the-art performance with higher reliability, lower cost and better scaling.

Growth of TE active QWIPs has been done using solid source Molecular Beam Epitaxy (MBE) on <100> GaAs substrates using the AlGaInAs material system. A large number of growths has been done using a variety of device structures and growth conditions to investigate the effects of device design and MBE growth on device performance. Characterization of the grown material using Double Crystal X-ray Diffraction (DCXD), Photoluminescence (PL) and Cathodoluminescence (CL) has been important to understanding the properties of the grown material. Fabrication of the QWIPs uses all the common semiconductor processing technologies as well as several less common techniques implemented at MIT by the author. Among the latter are silicon tetra-chloride / boron trichloride Reactive Ion Etching (RIE) of GaAlAs, Plasma Enhanced Chemical Vapor Deposition (PECVD) of low stress silicon oxy-nitride dielectrics and negative resist based metal lift-off. These techniques have been central to the development of the TE active QWIPs.

A $\boldsymbol{k \cdot p}$ analysis of the TE sensitivity for GaAs/AlGaAs quantum wells shows that discontinuities at the quantum well interface including coupling between the p-type valence band and the s-type conduction band lead to about a 20% ratio of TE to TM absorption coefficient between the occupied n=1 bound state and the unoccupied n=2 quasi-bound state for the quantum wells used in this work. The resulting TE absorption coefficient of about $200 cm^{-1}$ corresponds to a quantum efficiency of 0.068% per quantum well in a superlattice with 50Å quantum wells and 300Å barriers.

Extensive measurement of individual TE active QWIP devices has been done at cryogenic temperatures to characterize dark current, spectral response and black body responsivity of the devices. Background Limited Performance (BLIP) temperatures as high as 68K have been demonstrated with leakage currents at BLIP of less than $0.2mA/cm^2$. Responsivities as high as 80mA/W have been measured using a black body at 300K for a QWIP with 10 $In_{0.10}GaAs$ 50Å quantum wells doped n type $5x10^{17}$ cm^{-3} and separated by 500Å $Al_{0.20}GaAs$ barriers. The detectivity of this device is $2x10^{10}$ Jones at BLIP and the quantum efficiency per quantum well is 0.068%. This value for the quantum efficiency per quantum well is in excellent agreement with the ***k•p*** theoretical predictions of the TE sensitivity.

An important accomplishment of this thesis is the recognition that the quantum efficiency of TE mode QWIPs, though low compared to competing devices, is sufficiently high so as not to effect the overall performance of the FPA. A critical analysis of the requirements for FPAs shows that improvements in overall FPA cost, reliability and performance can be gained by trading off individual device detectivity for improvements in uniformity and simplifications in the manufacturing process. The TE active QWIPs demonstrated here do exactly that and calculations based on measured devices show that state-of-the-art minimum resolvable temperature differences, MRTs, of less than 10mK are feasible using these devices while at the same time eliminating the need to manufacture high quality diffraction gratings. Use of the E-o-E/SAW technology directly addresses the problems of yield and reliability in existing bump-bonding hybrid integration techniques to eliminate the need for bump bonding and wafer thinning.

Since the VLSI MESFETs used in the FPA MUXs are commercial products designed to work within the range -55C to 125C systematic characterization of the component EFETs and DFETs at cryogenic temperatures has been carried out. All the components continue to work at temperatures as low as 15K and show no performance or reliability degradation even after repeated temperature cycling. The primary effects of lowering temperature are an increase in the transconductance by a factor of about three for a tem-

perature change from 300K to 60K, an increase in the threshold voltage at a rate of about 1.5mV/K for both EFETs and DFETs and an increase in the Schottky diode gate ideality factor from near one at 300K to near four at 15K due to an increase in the tunneling of carriers across the barrier.

In concluding, Figure 7.1 shows the motto for this work and emphasizes the focus that has been placed on applying interesting new technology to the problems of manufacturing infrared cameras.

A New Technology for Improving
Quantum Well Intersubband Photodetector (QWIP)
Focal Plane Array (FPA)
Performance and Manufacturability

Figure 7.1: Advantages of TE mode normal incidence QWIP FPAs

7.2 Possible Directions for Further Research

This thesis has presented several important results which lay the groundwork for further progress in QWIP FPA research. These results fall into two broad categories, related to the performance of individual devices and related to the integration of QWIP FPAs with

ROICs. In terms of the individual devices, increasing the BLIP temperature always pays important dividends because of the large cost associated with cryogenic coolers. As discussed in Chapter 2, the use of short period superlattices as barriers instead of bulk material has the advantage that the density of states is significantly reduced, from 3D to quasi 2D, thereby reducing thermionic emission and increasing BLIP temperature. This bound to miniband type QWIP transition has proved very successful in the TM grating coupled devices and can be expected to have a similar effect on thermionic emission in TE mode devices thereby increasing the BLIP temperature though the effect on TE mode absorption coefficient has not been calculated. The devices investigated in this thesis were all chosen to have Indium compositions and well widths below the critical thickness as discussed in Chapter 3. This was done on the assumption that any defects introduced by misfit dislocations would adversely affect carrier lifetime and dark current which in turn deteriorate the detectivity. However, the one device fabricated by Alpha Photonics had 50 quantum wells consisting of 40Å of $In_{0.30}Ga_{0.70}As$ placing it well beyond the critical thickness. It is possible that the in-plane nature of these defects has little or no effect on carrier lifetime and dark current, a fortuitous result that if confirmed by further work would allow the number of quantum wells to be chosen based on other performance criteria like the photoconductive gain which decreases with increase quantum well number. The devices investigated in this thesis were also all in the LWIR band partly because the large background photon fluxes in this band puts a premium on uniformity and thereby discounts quantum efficiency which is known to be low in TE mode devices and partly because equipment for measuring in the LWIR band was readily available. Further work now being pursued by MIT and ADT to expand the range of TE mode devices to shorter as well as longer wavelengths and to vertically integrate devices with different spectral responses should be quite interesting.

In the area of integration, a new integration technique called E-o-E/SAW has been proposed, but there is still a tremendous amount of further research to be done to demonstrate that high quality OEICs can be produced using this technique. First, though wafer bonding is becoming more wide spread in industry it appears to be somewhat of a "black art"

with some groups reporting much better results than others. MIT needs to demonstrate that the know-how exists here to produce high quality wafer bonds compatible with the time/temperature tolerances of the VLSI GaAs MESFETs used for electronic circuits. Research into the growth of specially designed epitaxial bonding layers containing narrow bandgap materials and/or high vapor pressure materials such as InP and InGaP should also prove fruitful in designing the optimum interface for the E-o-E/SAW process. Finally, though E-o-E/SAW has been conceived in the context of QWIP FPAs integrated on VLSI GaAs MESFETs the beauty of the technique is that it if widely applicable. Further research into its use with Silicon CMOS circuits, InP based optical devices and even MCT based optical detectors for low background temperature applications may prove fruitful.

Bibliography

[1] K. Shenoy and C.G. Fonstad, Jr. Monolithic Optoelectronic Circuit Design and Fabrication by Epitaxial Growth on Commercial VLSI GaAs MESFETs. *IEEE Photonics Techechnology Letters*, 7(5):508–510, May 1995.

[2] Sir William Herschel. Experiments on the Refrangibility of the Invisible Rays of the Sun. *Royal Society of London*, 90:255, 1800.

[3] J. Bose. U.S. Patent 755840, 1904.

[4] J. Bardeen and W. Brattain. The Transistor, A Semiconductor Triode. *Physical Review*, 74:230, 1948.

[5] W. Shockley. The Theory of p-n Junctions in Semiconductors and p-n Junction Transistors. *Bell System Technical Journal*, 28:435, 1949.

[6] L. Esaki and H. Sakaki. New Photoconductor. *IBM Technical Disclosure Bulleten*, 20:2456–2457, 1977.

[7] B. Levine, K. Choi, C. Gethea, J. Walker, and R. Malik. New 10mm infrared detector using intersubband absorption in resonant tunneling GaAlAs superlattices. *Applied Physics Letters*, 50:1092–1094, 1987.

[8] A. Berk, L. Bernstein, and D. Robertson. MODTRAN: A Moderate Resolution Model for LOWTRAN 7. Technical Report GL-TR-89-0122, US Army, 1989.

[9] G. Anderson, J. Chetwynd, J. Theriault, P. Acharya, A. Berk, D. Robertson, F. Kneizys, M. Hoke, L. Aberu, and E. Shettle. MODTRAN: Suitability for Remote Sensing. *SPIE*, 1984, 1993.

[10] 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.

[11] R. Jones. Proposal of the Detectivity D** for Detectors Limited by Radiation Noise. *Journal of the Optical Society of America*, 50(11):1058–1059, 1960.

[12] R. Jones. Performance of Detectors for Visible and Intrared Radiation. In L. Martin, editor, *Advances in Electronics*, volume 5, pages 1–96. Plenum, 1953.

[13] R. Jones. Noise in Radiation Detectors. *Proceedings of IRE*, 47:1481–1486, 1959.

[14] J. Mooney, F. Shepherd, W. Ewing, J. Murguia, and J. Silverman. Responsivity nonuniformity limited performance of infrared staring cameras. *Optical Engineering*, 28(11):1151–1161, November 1989.

[15] M. Kinch and A. Yariv. Performance limitations of GaAs/AlGaAs infrared superlattices. *Applied Physics Letters*, 55(20):2093–2095, Novermber 1989.

[16] B. Levine. Comment on Performance limitations of GaAs/AlGaAs infrared superlattices. *Applied Physics Letters*, 56(23):2354–2356, June 1990.

[17] M. Greiner, R. Smith, and H. Timlin. Uniformity and stability in 2-dimensional infrared focal plane arrays. *SPIE*, 2225:176–184, 1994.

[18] H. Beratan, C. Hanson, and E. G.Meissner. Low-Cost Uncooled Ferroelectric Detector. *SPIE*, 2274:147–156, 1994.

[19] R. Owen, S. Frank, and C. Daz. Producibility of uncooled IRFPA detectors. *SPIE*, 1683:74–80, 1992.

[20] W. Lawson, S. Nielson, E. Putley, and A. Young. Preparation and Properties of HgTe and Mixed Crystals of HgTe-CdTe. *Journal of Physical Chemistry of Solids*, 9:325–329, 1959.

[21] J. Lowney, D. Seiler, C. Littler, and I. Yoon. Intrinsic carrier concentration of narrow-gap mercury cadmium telluride based on the nonlinear temperature dependence of the band gap. *Journal of Applied Physics*, 71(3):1253–1258, February 1992.

[22] W. Higgins, G. Pultz, R. Roy, and R. Lancaster. Standard relationships in the properties of $Hg_{1-x}Cd_xTe$. *Journal of Vacuum Science and Technology*, A7:271–275, 1989.

[23] E. O. Kane. Band Structure of Indium Antimonide. *Journal of Physics and Chrmistry of Solids*, 1:249–261, 1957.

[24] M. Weiler. Magnetooptical properties of $Hg_{1-x}Cd_xTe$ alloys. In R. Willardson and A. Beer, editors, *Semiconductors and Semimetals*, volume 16, pages 119–191. Academic Press, New York, NY, 1981.

[25] A. Rogalski and J. Piotrowski. Intrinsic infrared detectors. *Progress in Quantum Electronics*, 12:87–289, 1988.

[26] P. Petersen. Auger Recombination in $Hg_{1-x}Cd_xTe$. *Journal of Applied Physics*, 41(8):3465–3467, July 1970.

[27] P. Petersen. Auger recombination in Mercury Cadmium Telluride. *Semiconductors and Semimetals*, 18:121–155, 1981.

[28] M. Kinch, M. Brau, and A. Simmons. Recombination Mechanisms in 8-14μm HgCdTe. *Journal of Applied Physics*, 44(4):1649–1663, April 1973.

[29] J. Piotrowski. $Hg_{1-x}Cd_xTe$ Detectors. In A. Rogalski, editor, *Infrared Photon Detectors*, chapter 11, pages 391–494. SPIE Optical Engineering Press, Bellingham, WA, 1995.

[30] T. Casselman and P. Peterson. A comparison of the dominant Auger transitions in p-type (Hg,Cd)Te. *Solid State Communications*, 33:615–619, 1980.

[31] M. Kimata and N. Tubouchi. Schottky Barrier Photoemisive Detectors. In A. Rogalski, editor, *Infrared Photon Detectors*, chapter 9, pages 299–349. SPIE Optical Engineering Press, Bellingham, WA, 1995.

[32] 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.

[33] R. Keyes. *Optical and Infrared Detectors*. Springer-Verlag, New York, NY, 1980.

[34] C. Hanson, H. Beratin, R. Owen, M. Corbin, and S. McKenney. Uncooled thermal imaging at Texas Instruments. *SPIE*, 1735:17–26, 1992.

[35] J. Smith, L. Chiu, S. Margalit, A. Yariv, and A. Cho. A new infrared detector using electron emission from multiple quantum wells. *Journal of Vacuum Science and Technology B*, 1:376–378, 1983.

[36] J. Arthur. Vapor pressures and phase equilibria in the Ga-As system. *Journal of Physics and Chemistry of Solids*, 28:2257–2267, 1967.

[37] A. Cho and J. Arthur. Molecular Beam Epitaxy. *Progress in Solid State Chemistry*, 10:157–191, 1975.

[38] L. West and S. Eglash. First observation of an extremely large-dipole infrared transition within the conduction band of a GaAs quantum well. *Applied Physics Letters*, 46(12):1156–1158, 1985.

[39] B. Levine. Quantum-well infrared photodectors. *Journal of Applied Physics*, 74(15):R1–R81, October 1993.

[40] B. Levine, C. Bethea, G. Hasnian, V. Shen, E. Pelve, R. Abbott, and S. Hsieh. High sensitivity low dark current 10μm GaAs quantum well infrared photodetectors. *Applied Physics Letters*, 56(9):851–853, 1990.

[41] B. L. G. Hasnain, C. Bethea, and N. Chard. Broadband 8-12μm high-sensitivity GaAs quantum well infrared photodetector. *Applied Physics Letters*, 54(26):2704–2706, June 1989.

[42] B. Levine, S. Gunapala, and R. Kopf. Photovoltaic GaAs quantum well infrared detectors at 4.2μm using indirect $Al_xGa_{1-x}As$ barriers. *Applied Physics Letters*, 58(14):1551–1553, April 1991.

[43] B. Ridley. The electron-phonon interaction in quasi-two-dimensional semiconductor quantum-well structures. *Journal of Physics C*, 15:5899–5917, 1982.

[44] N. Johnson, H. Ehrenerich, P. Hui, and P. Young. Electronic and optical properties of III-V and II-VI semiconductor superlattices. *Physical Review B*, 41(6):3655–3669, 1990.

[45] M. Cardona, N. Christensen, and G. Fasol. Relativistic band structure and spin-orbit splitting of zinc-blende type semiconductors. *Physical Review B*, 38(3):1806–1827, July 1988.

[46] M. Flatte, P. Young, L. Peng, and H. Ehrenreich. Generalized superlattice ***K•p*** theory and intersubband optical transitions. *Physical Review B*, 53(4):1–16, 1996.

[47] O. Madelung, editor. *Semiconductors: Group IV Elements and III-V Compounds*. Springer, 1991.

[48] P. Young, E. Runge, M. Ziegler, and H. Ehrenreich. Optical absorption and exciton linewidths of Zn_{1-x} Cd_x Se quantum wels. *Physical Review B*, 49(11):7424–7431, March 1994.

[49] P. Young and H. Ehrenreich. Evidence for quantum well asymmetry in optical absorption. *Applied Physics Letters*, 61(9):1069–1071, August 1992.

[50] H. Liu. Photoconductive gain mechanism of quantum-well intersubband infrared detectors. *Applied Physics Letters*, 60(12):1507–1509, March 1992.

[51] H. Liu, B. Levine, and J. Andersson, editors. *Quantum well Intersubband Transition Physics and Devices*. NATO ASI Series. Kluwer Academic, Netherlands, 1993.

[52] S. Gunapala, B. Levine, L. Pfeiffer, and W. West. Dependence of the performance of GaAs/AlGaAs quantum well infrared photodetectors on doping and bias. *Journal of Applied Physics*, 69(9):6517–6520, 1991.

[53] M. Kane, S. Millidge, M. Emeny, D. Lee, D. Guy, and C. Whitehouse. Performance trade offs in the Quantum Well Infrared Detector. In E. Rosencher, editor, *Intersubband Transitions in Quantum Wells*, pages 31–42. Plenum, New York, NY, 1992.

[54] H. Liu, A. Steele, M. Buchanan, and Z. Wasilewski. Effects of the Upper State Position and the Number of Wells on the Performance of Intersubband Quantum Well Detectors. In E. Rosencher, editor, *Intersubband Transtions in Quantum Wells*, pages 57–63. Plenum, New York, NY, 1992.

[55] J. Andersson and L. Lundqvist. Near-unity quantum efficiency of AlGaAs/GaAs quantum well infrared detectors using a waveguide with a doubly periodic grating coupler. *Applied Physics Letters*, 59(7):857–859, August 1991.

[56] J. Andersson, L. Lundqvist, J. Borglind, and D. Haga. Performance of Grating Coupled AlGaAs/GaAs Quantum Well Infrared Detectors and Detector Arrays. In H. Liu, editor, *Quantum Well Intersubband Transition Physics and Devices*, pages 13–27. Kluwer Academic, Netherlands, 1994.

[57] J. Andersson, L. Lundqvist, and Z. Paska. Quantum efficiency enhancement of AlGaAs/GaAs quantum well infrared detectors using a waveguide with a grating coupler. *Applied Physics Letter*, 58(20):2264–2266, May 1991.

[58] K. Andersson and L. Lundqvist. Near-unity quantum efficiency of AlGaAs/GaAS quantum well infrared detectors using a eaveguide with a boubly periodic grating coupler. *Applied Physics Letters*, 59(7):857–859, August 1991.

[59] E. Yablonovitch and G. D. Cody. Intensity Enhancement in Textured Optical Sheets for Solar Cells. *IEEE Transactions on Electron Devices*, ED-29(2):300–305, February 1982.

[60] B. Levine, G. Sarusi, S. Pearton, K. Bandara, and R. Leibenguth. Random Scattering Optical Couplers for Quantum Well Infrared Photodetectors. In H. Liu, editor, *Quantum Well Intersubband Transition Physics and Devices*, pages 1–12. Kluwer Academic, Netherlands, 1994.

[61] A. Shik. Optical absorption at a heterojunction. *Soviet Physics Semiconductor*, 22(10):1165–1167, November 1988.

[62] A. Shik. Intraband photoconductivity of quantum-well heterostructures. *Soviet Physics Semiconductor*, 20(9):1002–1006, September 1986.

[63] A. Shik. Theory of Optical Intersubband Transitions. In E. Rosencher, editor, *Intersubband Transtions in Quantum Wells*, pages 319–328. Plenum, New York, NY, 1992.

[64] R. Yang, J. Xu, and M. Sweeny. Selection rules of intersubband transitions in conduction-band quantum wells. *Physical Review B*, 50(11):7474–7482, September 1994.

[65] M. Cohen and V. Heine. The Fitting of Pseudopotentials to Experimental Data and Their Subsequent Application. In H. Ehrenreich, editor, *Solid State Physics*, volume 24, pages 38–246. Academic Press, 1970.

[66] O. Madelung, editor. *Landolt-Bornstein Numerical Data and Functional Relationships in Science and Technology - Semiconductors: Intrinsic Properties of Group IV Elements and III-V, II-VI and I-VII Compounds*, volume 22a. Springer Verlag, 1986.

[67] N. Johnson, H. Ehrenerich, and a. T. M. K.C. Hass. *f*-Sum Rule and Effective Masses in Superlattices. *Physical Review Letters*, 59(20):2352–2355, November 1987.

[68] T. Duffield, R. Bhat, M. Koza, F. DeRose, D. Hwang, P. Grable, and S. Allen. Electron mass tunneling along the growth direction of (Al,Ga)As/GaAs semiconductor superlattices. *Phyical Review Letters*, 56(25):2724–2727, 1986.

[69] B. Levine, C. Bethea, G. Hasnain, J. Walker, and R. Malik. High-detectivity $D^*=1.0\times10^{10}$ $cmHz^{1/2}/W$ GaAs/AlGaAs multiquantum well $\lambda=8.3\mu m$ infrared detector. *Applied Physics Letters*, 53(4):296–298, 1988.

[70] B.F.Levine, C.G.Bethea, K.G.Glogovsky, J.W.Stay, and R.E.Leibenguth. Long-wavelength 128x128 GaAs quantum well infrared phootodetector arrays. *Semiconductor Science and Technology*, 6:C114–C119, 1991.

[71] M. Tidrow, K. Choi, C. Fariey, and F. Chang. Multicolor infrared detection using a voltage tunable bandpass filter. *Applied Physics Letters*, 65(23):2996–2998, December 1994.

[72] G. Hasnain, B. Levine, C. Bethea, R. Logan, J. Walker, and R. Malik. GaAs/AlGaAs multiquantum well infrared detector arrays using etched gratings. *Applied Physics Letters*, 54:2515–2517, 1989.

[73] W. Beck, J. Little, A. Goldberg, and T. Faska. Imaging performance of LWIR miniband transport multiple quantum well infrared focal plane array. In H. Liu, editor, *Quantum Well Intersubband Transition Physics and Devices*, pages 55–68. Kluwer Academic, Netherlands, 1994.

[74] L. Yu and S. Li. A metal grating coupled bound to miniband transition GaAs mulitquantum well superlattice infrared detector. *Applied Physics Letters*, 59(11):1332–1334, September 1991.

[75] E. Yablonovitch, T. Sands, D. Hwang, I. Shnitzer, T. Gmitter, S. Shastry, D. Hill, and J. Fan. Vander Waals bonding of GaAs on pd leads to a permanent, solid-phasetopotaxial, metallurgical bond. *Applied Physice Letters*, 59(9):3159–3161, December 1991.

[76] B. Levine, S. Gunapala, J. Kuo, S. Pei, and S. Hui. Normal incidence hole intersubband absorption long wavelength GaAs/$Al_xGa_{1-x}As$ quantum well infrared photodetectors. *Applied Physics Letters*, 59:1864–1866, October 1991.

[77] S. Li and Y. Wang. Novel grating coupled and normal incidence III-V quantum well infrared photodetectors with background limited performance at 77K. In H. Liu, editor, *Quantum Well Intersubband Transition Physics and Devices*, pages 29–42. Kluwer Academic, Netherlands, 1994.

[78] H. Xie, J. Katz, and W. Wang. Infrared absorption enhancement in light-and heavy-hole inverted $Ga_{1-x}In_xAs/Al_{1-y}In_yAs$ quantum wells. *Applied Physics Letters*, 59(27):3601–3603, December 1991.

[79] E. Brown and S. Eglash. Calculation of the intersubband absorption strength in ellipsoidal-valley quantum wells. *Physical Review*, 41(11):7559–7568, April 1990.

[80] B. Levine, C. Bethea, V. Swaminathan, J. Stayt, R. Leibenguth, K. Glogovsky, and W. Gault. Producibility of GaAs Quantum Well Infrared Photodetector Arrays. *SPIE*, 1683:41–45, 1992.

[81] The MIT MBE Manual, 1995.

[82] J. C. Vlcek. *Molecular beam epitaxy growth and applications of graded bandgap InGaAlAs semiconductor alloys*. PhD thesis, Massachusetts Institute of Technology, 1991.

[83] J. C. Vlcek and C. G. Fonstad, Jr. Precise computer control of the MBE process - Application to graded InGaAlAs/InP alloys. *Journal of Crystal Growth*, 111:56–63, 1991.

[84] V. Swaminathan and A. T. Macrander. *Materials Aspects of GaAs and InP*, volume Prentice-Hall. Prentice-Hall, Englewood Cliffs, NJ, 1991.

[85] Bede Scientific Instruments LTD. *RADS - Rocking Curve Analysis by Dynamical Simulations*, 1990.

[86] E. Braun, K. Shenoy, and C.G. Fonstad, Jr. Elevated Temperature Stability of GaAs Digital Integrated Circuits. *IEEE Electron Device Letters*, EDL-16:1–3, January 1995.

[87] K. Xie, C. R. Wie, J. A. Varriano, and G. W. Wicks. Interface traps and interface recombination in AlGaAs/GaAs quantum well laser diodes. *Applied Physics Letters*, 60:428–430, 1992.

[88] K. A. Prior, G. J. Davies, and R. Heckingbottom. The thermodynamics of oxygen incorporation into III-V semiconductor compounds and alloys in MBE. *Journal of Crystal Growth*, 66:55–67, 1984.

[89] American Xtal Technology. Dublin, CA.

[90] E. Fitzgerald. Work in progress.

[91] H. Chui and J. Harris. Growth studies on $In_{0.5}Ga_{0.5}As$/AlGaAs quantum wells grown on GaAs with a linearly graded InGaAs buffer. *Journal of Vacuum Science and Technology*, B12(2):1019–1022, April 1994.

[92] W. Jesser and J. Matthews. Evidence for Psuedomorphic Growth of Iron on Copper. *Philosphical Magazine*, 15:1097–1106, 1967.

[93] W. Jesser and J. Matthews. Psuedomorphic Deposits of Cobalt on Copper. *Philosphical Magazine*, 17:461–473, 1968.

[94] W. Jesser and J. Matthews. Psuedomorphic Growth of Iron on Hot Copper. *Philosphical Magazine*, 17:595–602, 1968.

[95] J. Matthews and A. Blakeslee. Defects in Epitaxial Mulitlayers. *Journal of Crystal Growth*, 27:118–125, 1974.

[96] E. Fitzgerald. Dislocations in strained layer epitaxy: theory, experiment and applications. *Materials Science Reports*, 17(3):87–142, 1991.

[97] J. Matthews. Defects associated with the accommodation of misfit between crystals. *Journal of Vacuum Science and Technology*, 12:126–133, 1975.

[98] R. Hull, J. Bean, F. Cerdeira, A. Fiory, and J. Gibson. Stability of semiconductor strained-layer superlattices. *Applied Physics Letters*, 48(1):56–58, January 1986.

[99] S. Adachi and K. Oe. Chemical Etching Characteristics of GaAs. *Journal of the Electrochemical Society: Solid-State Science and Technology*, 130:2427–2435, December 1983.

[100] D. W. Shaw. Localized GaAs Etching with Acidic Hydrogen Peroxide Solutions. *Journal of the Electrochemical Society: Solid-State Science and Technology*, 128:874–880, April 1981.

[101] H. H. Berger. Models for contacts to planar devices. *Solid State Electronics*, 15:145–158, 1972.

[102] Z. P. Sobczak. Technical Bulletin, Negative Resist NR8-3000. Technical report, Futurrex Corporation, Newton, NJ, 1992.

[103] W. X. Chen, S. C. Hsueh, P. K. L. Yu, and S. S. Lau. Solid-Phase Epitaxial Pd/Ge Ohmic Contact to $In_{1-x}Ga_xAs_yP_{1-y}$/InP. *IEEE Electon Device Letters*, EDL-7(8):471–473, 1986.

[104] L. G. Shantharama, H. Schumacher, H. P. LeBlanc, R. Esagui, R. Bhat, and M. Koza. Evaluation of Single Ohmic Metalizations for Contacting Both p- and n-type GaInAs. *Electronics Letters*, 26:1127–1130, 1990.

[105] 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.

[106] G. Karunasiri, R. Shih, and J. Chen. Normal Incident Intersubband Infrared Detector using n-type InGaAs/GaAs Quantum Wells. In *53rd Device Research Conference*, Charlotte, VA, June 19-21 1995.

[107] P. Martin, J. Pan, C.G. Fonstad, Jr. and P. Kannam. Normal Incidence Pseudomorphic InGaAs QWIP. In *Late News Paper, 53rd Device Research Conference*, Charlotte, VA, June 19-21 1995.

[108] Y. H. Chai. *Impurity Segregation and Diffusion Study in GaAs/AlGaAs Quantum Wells Grown by Molecular Beam Epitaxy.* PhD thesis, University of Massachusetts, Lowel, 1995.

[109] H. Ehrenreich. Personal Communication.

[110] G. Karunasiri, J. Park, J. Chen, R. Shih, J. Scheiling, and M. Dodd. Normal incident InGaAs/GaAs multiple quantum well infrared detector using electron intersubband transitions. *Applied Physics Letters*, 67(18):2600–2602, 1995.

[111] G. Stringfellow and H. Kunzel. Electron mobility in compensated GaAs and $Al_xGa_{1-x}As$. *Journal of Applied Physics*, 51:3254–3258, 1980.

[112] S. Sze. *Physics of Semiconductor Devices*. Wiley, New York, NY, 2nd edition edition, 1981.

[113] W. Shockley and G. Pearson. Modulation of Conductance of Thin Films of Semiconductors with Surface Charges. *Physical Review*, 74:232, 1948.

[114] P. Houston and A. Evans. Electron Drift Velocity in n-GaAs at High Electric Fields. *Solid State Electronics*, 20:197–204, 1977.

[115] Y. Pozhela. Electron transport properties in GaAs at high electric fields. *Solid State Electronics*, 23:927–933, 1980.

[116] G. Dacey and I. Ross. The Field Effect Transistor. *Bell System Technical Journal*, 34:1149–1189, 1955.

[117] Meta-Software. HSPICE User's Manual, 1992. Version H92.

[118] H. Statz, P. Newman, I. Smith, R. Pucel, and H. Haus. GaAs FET Device and Circuit Simulation in SPICE. *IEEE Transactions on Electron Devices*, ED-34:160–169, February 1987.

[119] V. S. Inc. Foundry Design Manual, 1993. Version 6.0.

[120] S. Lee, R. Vahenkamp, G. Kaelin, L. Hou, R. Zucca, C. Lee, and C. Kirkpatrick. Modeling of GaAs MESFETs. In *Process GaAs IC Symposium*, page 74, 1983.

[121] K. Shenoy and C.G. Fonstad, Jr. MIT-OEIC-3/NCIPT-OEVLSI-1 Optoelectronic VLSI GaAs Chip Design and Testing Manual, December 1994.

[122] E. K. Braun. Elevated Temperature Stability of Gallium Arsenide Integrated Circuits. Master's thesis, Massachusetts Institute of Technology, 1995.

[123] S. Chaundhuri and M. Das. On the Determination of Source and Drain Series Resistance of MESFETs. *IEEE Electron Device Letters*, EDL-5(7):224–246, 1984.

[124] C. Crowell and S. Sze. Current Transport in Metal-Seminconductor Barriers. *Solid-State Electronics*, 9:1035–1048, 1966.

[125] C. Y. Chang and S. M. Sze. Carrier Transport Across Metal-Semiconductor Barriers. *Solid-State Electronics*, 13:727–740, 1970.

[126] C. Crowell and S. Sze. Electron-Optical Phonon Scattering in the Emitter and Collector Barriers of Semiconductor-Metal-Semiconductor Structures. *Solid-State Electronics*, 8:979–990, 1965.

[127] C. Crowell and S. Sze. Quantum Mechanical Reflection of Electrons at Metal-Semiconductor Barriers: Electron Transport in Semiconductor-Metal-Semiconductor Structures. *Journal of Applied Physics*, 37(7):2683–2689, June 1966.

[128] R. Martin. Infrared Focal Plane Array Read Out Integrated Circuits. Personal Communication.

[129] R. Martin, C. Partin, and G. Milne. Dump, Ramp, Sample IR FPA ROIC with Switched Capacitor Filtering. In *IRIS*, July 1995.

[130] L. J. Herbst. *High Speed Digital Electronics*. Prentice Hall, New York, NY, 1992.

[131] K. Shenoy, C.G. Fonstad, Jr. and J. Mikkelson. High Temperature Stability of Refractory metal VLSI GaAs MESFETs. *IEEE Electron Device Letters*, EDL-15(3):106–108, March 1994.

[132] K. V. Shenoy. *Monolithic Optoelectronic VLSI Circuit Design and Fabrication for Optical Interconnects*. PhD thesis, Massachusetts Institute of Technology, 1995.

[133] K. Shenoy, C.G. Fonstad, Jr., 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.

[134] R. Aggarwal, K. Shenoy, and C.G. Fonstad, Jr. A Tecnology for monolithic integration of high-Indium fraction resonant tunneling diodes with commercial MESFET VLSI electronics. In *Proceedings of the 1995 InP and Related Materials Conference*, pages 85–88, Japan, May 1995.

[135] A. Grot, D. Psaltis, K. Shenoy, and C.G. Fonstad, Jr. Integration of LEDs and GaAs Circuits by MBE regrowth. *IEEE Photonics Technology Letters*, 6(7):819–821, July 1994.

[136] J. Ahadian and C.G. Fonstad, Jr. MIT OEIC-4/NCIPT OEVLSI 2 Optoelectronic VLSI GaAs Chip Design and Testing Manual, 1996.

[137] C. Jones, B. Matthews, D. Purdy, and N. Metcalfe. Fabrication and assessment of optically immersed CdHgTe detector arrays. *Semiconductor Science and Technology*, 6:C110–C113, 1991.

[138] N. Gordon. Design of $Hg_{1-x}Cd_xTe$ infrared detector arrays using optical immersion with microlenses to achieve a higher operating temperature. *Semiconductor Science and Technology*, 6:C106–C109, 1991.

[139] J. Ahadian. Phosphide-based optical emitters for monolithic integration with GaAs MESFETs. Master's thesis, Massachusetts Institute of Technology, 1995.

[140] M. Shimbo, K. Furukawa, K. Fukuda, and K. Tanzawa. Silicon-to-Silicon direct bonding method. *Journal of Applied Physics*, 60(8):2987–2989, 1986.

[141] R. Stengl, T. Tan, and C. Gosele. A Model for the Silicon Wafer Bonding Process. *Japanese Journal of Applied Physics*, 28:1735–1741, 1989.

[142] F. Kish, F. Steranka, D. Defevere, D. Vanderwater, K. Park, C. Kuo, T. Osentowski, M. Peanasky, J. Yu, R. Fletcher, D. Steigerwald, and V. Robbins. Very high-efficiency semiconductor wafer-bonded transparent-substrate $(Al_xGa_{1-x})_{0.5}In_{0.5}$ P/GaP light-emitting diodes. *Applied Physics Letters*, 64(21):2839–2841, May 1994.

[143] H. Wada, Y. Ogawa, and T. Kamijoh. Electrical characteristics of directly-bonded GaAs and InP. *Applied Physics Letters*, 62(7):738–740, February 1993.

[144] D. Nolte. Band offsets for pseudomorphic InP/GaAs. *Applied Physics Letters*, 54(3):259–261, 1989.

6213-13