Physics of Novel InAs/AlSb/GaSb Resonant Interband Tunneling Structures

A Dissertation
Presented to the Faculty of the Graduate School of Yale University
in Candidacy for the Degree of Doctor of Philosophy

by
John L. Huber

Dissertation Director: Prof. Mark A. Reed

May 1997
Abstract

Physics of Novel InAs/AlSb/GaSb Resonant Interband Tunneling Structures

John L. Huber
Yale University
1997

A resonant tunneling diode consists of a number of abrupt layers of epitaxially grown heterojunctions, usually in III/V semiconductors, with the material composition of the individual layers chosen to create a double barrier structure in the conduction (or valence) band of entire system. As the layer thickness approach the de Broglie wavelength of the electrons in the system, the tunneling probability will be selectively enhanced at certain energies causing a peaked I(V) characteristic. Resonant interband tunneling devices differ from conventional resonant tunneling diodes in that the confined states, accessible to electron (or hole) transport lie in the valence (or conduction) band rather than the conduction band.

This work examines the details of the interband tunneling process in InAs/AlSb/GaSb structures. A simple model of resonant tunneling based on a density of states argument including magnetic field effects is presented. Experimental results from p-type well InAs/AlSb/GaSb/AlSb/InAs structures are presented. The data show that tunneling occurs through multiple subbands, including both light-hole and heavy-hole like subbands. Structures which exhibit both intraband and interband tunneling are investigated. It is shown that the placement of an AlSb tunnel barrier can greatly influence the tunneling characteristics.

Application of resonant tunneling diodes, by means of a simple switching block, to compressed functionality circuits is demonstrated in a binary and a ternary adder circuit. Also, the modification of the turn-on characteristics of a tunneling hot electron transfer amplifier is demonstrated.
Acknowledgments

For my wife, who made it all worth it.

None of this would have been possible without my advisor, Professor Mark Reed, who gave me both the means and guidance to complete my graduate school career at Yale. His insight proved invaluable in deciphering the data and getting it all to make sense. In addition to Prof. Reed, I'd like to thank Profs. Robert Wheeler, T.P. Ma and Lou Guido, the other members of my thesis committee. In particular, Prof. Wheeler, not only for having one of the coolest labs around (everyone should see liquid helium go superfluid in a glass dewar at least once), but for his help in making all the equipment work. I'd like to acknowledge Herb Goronkin and the folks at Motorola for supplying the material for the tunneling structures and for financial support, and Ted Moise (a former Yalie) and the folks at Texas Instruments for the THETA samples.

In addition to my thesis project, I thoroughly enjoyed the time I spent as a teaching assistant. Both Mark and T.P. put up with me shooting my mouth off about how I thought things could be done better, and then allowed me to act on it and rewrite the labs for their classes. I suspect I learned more about electronics from doing that than I learned when I was actually taking classes.

There are many fellow graduate and undergraduate students both past and present I'd like to mention. First, those who directly assisted with this work: Chong-Wu Zhou, Jia Chen and Jim McCormack contributed important pieces of the work with the logic circuits; Theresa Kramer took many of the measurements of the THETA structures and assisted with much of the data analysis; Curt Richter showed me how to make the toys in the dungeon run and how to make a gold o-ring. I'm extremely grateful to my officemate Jeff Sleight, and to Rachel Lombardi for many discussions, both technical and otherwise, and for other highly useful things, like babysitting services. I'd also like to mention Kyle Thomas, who's graduate career paralleled mine and allowed me to see the light at the end.
of the tunnel. Ellen Sleight and Mark Amman also deserve mention for contributing to my time at Yale.

Perhaps the nicest (and also the most difficult and annoying) thing about my experience at Yale was the opportunity to do things not directly related to my project, but without which it never could have been done, ranging from learning more than I ever wanted to know about an awful (but functional) piece of Swedish vacuum equipment to making sample holders in the machine shop (I never thought I'd actually use stuff I learned in metal shop). Most of those things involved assistance from Larry Forster and Dexter Crowley whose assistance was greatly appreciated. I'd also like to thank Sandi Vitale and Arlene Ciociola for all their help with things like paperwork and ordering.

I'd like to thank my parents, whose love, encouragement and support both got me pointed in the right direction to start with and kept me going throughout. Finally, I'd like to thank my wife, Debbie, for putting up with me and letting me drag her across the country into the great unknown, and my children, Jacob and Zachary who make me realize every day that being a backboard beats anything I've ever done.
Contents

List of Figures vii
List of Tables xvi
List of Abbreviations and Symbols xviii

1 Introduction 1

2 Resonant Tunneling Theory 4
   2.1 Introduction ........................................ 4
   2.2 Quantum Size Effects ............................. 6
      2.2.1 Quantum Wells ................................ 6
      2.2.2 Single Band Transmission Coefficient .......... 8
      2.2.3 Resonant Tunneling Diode I(V) Characteristics ..... 10
   2.3 Interband Tunneling .............................. 13
      2.3.1 Resonant Interband Tunneling Structures ......... 13
      2.3.2 Two-band Transmission Coefficient ........... 16
   2.4 Density of states model of resonant tunneling .......... 20
      2.4.1 “Conventional” Resonant Intraband Tunneling .......... 21
      2.4.2 Resonant Interband Tunneling ................ 24
   2.5 Magnetic Field Effects ........................ 26
      2.5.1 Field Parallel to Current .................... 28
2.5.2 Field Perpendicular to Current ............................................ 29
2.6 Summary ..................................................................................... 31

3 Experimental Methods .................................................................. 33
   3.1 Introduction .................................................................................. 33
   3.2 Device Fabrication ......................................................................... 33
   3.3 Electrical Measurements .............................................................. 36
      3.3.1 Cryostat Systems ................................................................. 38
      3.3.2 Computer Controlled Data Acquisition Setup ................. 40

4 Single Well Resonant Interband Tunneling Structure .................. 43
   4.1 Introduction .................................................................................. 43
   4.2 Device Structure ........................................................................... 44
   4.3 GaSb Well Subband Structure ...................................................... 47
   4.4 Magnetotunneling ......................................................................... 47
   4.5 Experimental Results ................................................................. 50
      4.5.1 B parallel to I ...................................................................... 51
      4.5.2 B perpendicular to I ............................................................. 58
   4.6 Discussion ..................................................................................... 61
   4.7 Variable Temperature Measurements ......................................... 73
   4.8 Summary ...................................................................................... 74

5 Multiple Well Resonant Interband Tunneling Structures ........... 80
   5.1 Introduction .................................................................................. 80
   5.2 Device Structure ........................................................................... 81
   5.3 Experimental Results ................................................................. 85
   5.4 Discussion ..................................................................................... 91
   5.5 Summary ...................................................................................... 93
6 Resonant Tunneling Applications

6.1 Introduction ........................................ 94
6.2 RTD/Transistor Switching Block ...................... 96
6.3 Adder Circuits ........................................ 97
   6.3.1 Binary Adder .................................... 97
   6.3.2 Ternary Adder .................................. 100
6.4 Conclusion ........................................... 103

7 Resonant Tunneling in THETA structures ............... 104

7.1 Introduction ......................................... 104
7.2 Device Structure ...................................... 106
7.3 Experimental Results ................................ 106
   7.3.1 Common-Base Characteristics .................. 107
   7.3.2 Common-Emitter Characteristics .............. 109
7.4 Conclusions ......................................... 112

8 Conclusions ........................................... 114

Bibliography ............................................. 117
List of Figures

2.1 Plot of Bandgap vs Lattice Constant for various semiconductors. The solid lines connecting various binary compounds indicate direct bandgaps, while the dashed lines indicate indirect bandgaps. ........................................ 5

2.2 Schematic illustration of quantum size effects in a quasi-two dimensional region. (a) The conduction band has only discrete allowable energies in the confinement direction, but still has a continuous density of states in the transverse direction. (b) The density of states increases in discrete steps corresponding to each discrete energy level in the confinement direction. At energies where the confined energy levels occur, the value of the 2d-density of states is equal to the value of the corresponding 3d-density of states. . . . 7

2.3 (a) Schematic of an RTD. (b) Transmission coefficient for the structure shown in (a). ................................................................. 8

2.4 Band plots of a resonant tunneling structure under different applied bias. By convention, the region to the left of the tunneling structure is referred to as the emitter. The dotted line in the emitter represents the quasi-Fermi level for the emitter. (a) No applied bias. (b) Biased to the peak resonant current (1 V). (c) Biased past resonance (1.5. V). The I(V) characteristics are shown in Figure 2.5 ......................................................... 11
2.5  Expected I(V) characteristics for the structure shown in Figure 2.4. \( V_{\text{peak}} \) corresponds to (b) and \( V_{\text{valley}} \) corresponds to (c). The dotted line connecting the two points corresponds to the NDR region when the device is oscillating. The gray dotted line corresponds to the expected I(V) characteristic of a single tunneling barrier with thickness equal to the sum of the two barriers.

2.6  Band diagram of an Esaki tunnel diode. The dotted line is the quasi-Fermi level in each region. (a) No applied bias (b) Biased beyond the tunnel current cutoff.

2.7  (a) Schematic showing two different types of band alignments. (b) Resonant tunneling structures fabricated from a Type I material system and a Type II material system. The gray lines indicate the confined states.

2.8  Calculated two-band transmission coefficient for two different heterostructures. (a) Single well interband tunneling structure. (b) Double well intra/interband tunneling structure. For both cases, the zero in energy is taken to be the bottom of the InAs conduction band. The gray lines indicate the confined states.

2.9  Graphical representation of a density of states model of resonant tunneling at \( T = 0 \text{K} \). Adapted from reference [37]

2.10  (a) Band diagram of a double barrier tunneling structure. Only the region near the interband tunneling window is shown. \( E_O \) is the energy of the well dispersion at \( k_z = 0 \). (b) Energy vs \( k_z \) for both the emitter and well at four different bias points. The shaded region represents the available carriers (i.e. the emitter Fermi sea). The thick regions on curves #2 and #3 represent those available carriers which can tunnel into the well while conserving transverse momentum. For curves #1 and #4, no tunneling is possible. (c) Current Density vs Energy for the example shown in (b) as described by equation 2.30.
2.11 Effects of a magnetic field applied perpendicular to a 2DEG system. (a) Allowed momentum states for a two dimensional electron gas with no applied magnetic field. The states are evenly spaced in both $k_x$ and $k_y$. (b) With a magnetic field applied perpendicular to a 2DEG electron gas. Allowed states lie on concentric circles of constant radius $k^2 = (2\pi B/\hbar)(n + 1/2)$. An equal number of states lie on each circle. (c) Density of states for no applied field (gray line) and applied field (dark line). The effects of broadening are shown. (d) The gray curve corresponds to Energy vs $k_\parallel$ for a 2DEG with no applied field. The dark segments correspond to energies where the density of states is greater than that of the no field case.

2.12 Simulated I(V) characteristics for the situation illustrated in Figure 2.10c. The units on both current and bias are arbitrary.

2.13 Fan diagram (current peak positions vs magnetic field) for the simulated I(V) characteristics of Figure 2.12.

2.14 Effects of a magnetic field applied parallel to the interfaces of an interband tunneling system. The emitter electrons all receive an extra component of $k_\parallel$ from the applied field.

2.15 Current threshold for various values of $\Delta k_\parallel$ for Figure 2.14. Each trace corresponds to a different magnetic field with $\Delta k_1 < \Delta k_2 < \Delta k_3 < \Delta k_4$. $\Delta k_4$ corresponds to the case where the emitter dispersion boundary, the well dispersion and $E_F$ intersect at the same point.

3.1 Processing sequence for a two-terminal device.

3.2 Experimental setup for low temperature measurements.

3.3 Subcircuits used in the data acquisition setup.

4.1 Schematic bandstructures for a) Sample #1, and b) Sample #2. The emitter corresponds to the top of the sample, and the collector to the substrate. The details of the barrier structure are given in Table 4.3.
4.2 Parallel subband structure for the well structure extrapolated from published calculated data [26]. The dotted lines are parabolic fits. For GaSb, the lattice constant $a = 0.6096$ nm. ............................................. 48

4.3 I(V) characteristics of Sample #1 at 300 K and 1.4 K. The oscillation regions in the 1.4 K curve are not shown. ................................................................. 51

4.4 Reverse bias I(V) characteristics of Sample #1 with $B \parallel I$ before correction for magnetoresistance. ................................................................. 54

4.5 Reverse bias I(V) characteristics of Sample #1 with $B \parallel I$ after correction for magnetoresistance. ................................................................. 54

4.6 Fan diagram generated from peaks in the second derivative of the I(V) characteristics of Sample #1 with $B \parallel I$. Each point is a positive peak in the second derivative, corresponding to a current peak or shoulder. A positive peak corresponds to a point of maximum positive curvature in the I(V) characteristics, indicating a current peak (i.e. a point of more negative current) or shoulder. Each I(V) was measured with 0.5 mV resolution. The field resolution was 0.05 T. ................................................................. 55

4.7 Fan diagram from Figure 4.6. The two sets of dashed lines are fit to the structure. The fan converging to -0.036 V corresponds to structure from HH1, while the fan converging to -0.141 V corresponds to structure from LH1. The slopes are given by $nm$ where $n$ is an integer and $m = 0.01$ for HH1, and $m = 0.065$ for LH1 ................................................................. 56

4.8 Fan diagram from Figure 4.6 with a fan, converging to 0.133 V with slopes given by $nm$ where $n$ is an integer and $m = -0.095$. This fan results from holes in the valence band resonantly tunneling from the collector to the emitter through a light-hole state. ................................................................. 57

4.9 Forward bias I(V) characteristics of Sample #1 with $B \parallel I$ before correction for magnetoresistance. ................................................................. 59
4.10 Forward bias I(V) characteristics of Sample #1 with B||I after correction for magnetoresistance. ................................................................. 59

4.11 Fan diagram generated from peaks in the second derivative of the I(V) characteristics of Sample #1 with B||I. Each I(V) was measured with 0.5 mV resolution. The field resolution was 0.05 T. Each point is a negative peak in the second derivative, corresponding to a current peak (i.e. a local maximum) or shoulder. The fan converges to 0.236 V with slopes given by nm where n is an integer and m = -0.072. This fan results from holes in the valence band resonantly tunneling from the collector to the emitter through a light-hole state. Additional structure in the bias range up to 0.05 V moving to increasing bias with increasing field can be seen. .......................... 60

4.12 Reverse bias I(V) characteristics of Sample #1 with B⊥I before correction for series resistance. ................................................................. 62

4.13 Reverse bias I(V) characteristics of Sample #1 with B⊥I after correction for series resistance. ................................................................. 62

4.14 Fan diagram generated from peaks in the second derivative of the I(V) characteristics of Sample #1 with B⊥I. Both positive and negative peaks in the second derivative are shown. Each I(V) was measured with 0.5 mV resolution. The field resolution was 0.05 T. ................................. 63

4.15 Forward bias I(V) characteristics of Sample #1 with B⊥I before correction for series resistance. ................................................................. 64

4.16 Forward bias I(V) characteristics of Sample #1 with B⊥I after correction for series resistance. ................................................................. 64

4.17 Fan diagram generated from negative peaks in the second derivative of the I(V) characteristics of Sample #1 with B⊥I. The scale is the same as in Figure 4.11. The dense region of points between 0.06 V and 0.1 V results from the subtraction of the magnetoresistive term. Each I(V) was measured with 0.5 mV resolution. The field resolution was 0.05 T. ................................. 65
4.18 Estimated parallel subband structure with experimentally determined critical points. The heavy dashed curve indicates the extent of occupied states in the emitter. The light dashed curve indicates the offset obtained by the application of a magnetic field parallel to the confining interfaces such that the current threshold moves to zero bias. For GaSb, the lattice constant $a = 0.6096 \text{ nm}$. ................................. 67

4.19 Fan diagram generated from the second derivative of the $I(V)$ characteristics obtained using the interband tunneling model of Chapter 2. The $I(V)$ characteristics were modeled using HH1 and LH1 assuming the parabolic fits to the estimated subband structure. The dashed lines are the fans fit to the experimentally measured structure. ................................. 70

4.20 Reverse bias $I(V)$ characteristics of Sample #2 with $B\parallel I$. ................................. 75

4.21 Fan diagram generated from peaks in the second derivative of the $I(V)$ characteristics of Sample #2 with $B\parallel I$. The fan diagram was generated using the same procedure as was done with Figure 4.6. Each $I(V)$ was measured with 1.0 mV resolution. The field resolution was 0.2 T. ................................. 76

4.22 Zero-bias conductance vs. inverse magnetic field for Sample #1. The field is applied parallel to the current direction. ................................. 77

4.23 FFT spectrum for the curve shown in Figure 4.22. The large peak centered at index=14 corresponds to the main frequency component of the curve. The smaller peaks to the right are harmonics of the main peak. ................................. 77

4.24 Zero-bias conductance vs. magnetic field for Sample #1. The field is applied perpendicular to the current direction. ................................. 78

4.25 $I(V)$ characteristic of Sample #1 used for variable temperature measurements. The large valley current in both the forward and reverse bias directions (compare with Figures 4.5 and 4.10 is the result of a parallel conduction path that developed in the device after numerous thermal cyclings. ................................. 78

4.26 $I$ vs 1000/$T$ for and applied bias of -0.2 V, -0.3 V and -0.4 V. ................................. 79
5.1 Band diagrams for the inter/intra band tunneling samples. Transmission coefficients are shown in Figure 5.2.

5.2 Calculated two-band transmission coefficients for the inter/intra band tunneling samples. Transmission coefficients are shown in Figure 5.1.

5.3 I-V characteristics of MWRIT1. The I-V is symmetrical about zero bias, therefore only one bias direction is shown.

5.4 Reverse bias I-V characteristics of MWRIT2

5.5 Forward bias I-V characteristics of MWRIT2

5.6 Reverse bias I-V characteristics of MWRIT3

5.7 Forward bias I-V characteristics of MWRIT3

6.1 Current-Voltage characteristics of an RTD using a single load resistor. The NDR region of the RTD allows two stable operating points to exist for the same circuit. The point lying on the NDR region is not stable.

6.2 (a) Schematic of the switching block. (b) Transfer function of the switching block compared with the RTD transfer function.

6.3 Simulation of the output voltage ($V_{DD} = 5V$) and the transistor base-emitter voltage for different RTD I-V characteristics. The peak current/voltage of the RTD was varied by $\pm 20\%(\pm 5\%)$.

6.4 Schematic of a full binary adder using the RTD/Transistor switching block.

6.5 Simulation of the full adder. The inputs A,B,C scaled to fit on the same set of axes.

6.6 Schematic of a ternary adder. $V_{DD}$ for the carry output is half that for the sum output.
6.7 Various transfer function schematics for transistors in the ternary adder circuit. The input voltage is taken to be the emitter of Q6. (a) Schematic of transfer function of Q2. (b) Input current to switching block with two RTDs in series (i.e. Q3). (c) Collector voltage for Q3. (d) Sum voltage for the circuit. .......................................................... 102

6.8 Measured transfer function for the ternary adder circuit. The input voltage is taken to be the emitter of Q6. .......................................................... 102

6.9 Schematic of a CMOS ternary adder. High and low output voltages are obtained in a manner similar to that of conventional CMOS in which only one transistor is turned on at any one time. The intermediate voltage is obtained when both transistors are on. .................................................. 103

7.1 Operation of a THETA structure. A voltage placed across the base-emitter junction will inject electrons through the thin tunnel barrier into the base. A voltage across the base-collector junction will vary the height of the energy barrier. Those electrons which traverse the base ballistically (i.e. do not lose energy via inelastic scattering) will be swept into the collector if the top of the energy barrier is sufficiently lowered. In the common-base mode, derivative of the collector current is proportional to the energy distribution of the injected electrons. .................................................. 105

7.2 Self-consistent (Poisson) band diagram of a structure biased to turn on. The inset shows the region near the injection energy. ................................. 107

7.3 Common-base characteristic for a 50 nm base width sample. ............... 108

7.4 Base-emitter junction conductance characteristics \(dI_E/dV_{BE} \text{ vs } V_{BE}\). .... 108

7.5 Common-base characteristic for a 30 nm base width sample. ................. 110

7.6 Self-consistent (Poisson) band diagram for three different values of \(V_{CB}\). with \(V_{CB3} > V_{CB2} > V_{CB1}\). Note the reduction in effective barrier thickness (at the injection energy) in addition to the reduction in barrier height. .... 111
7.7 Common-emitter characteristics for a 50 nm base width sample. . . . . . . 112
7.8 Common-emitter characteristics for a 30 nm base width sample. . . . . . . 113
List of Tables

4.1 Important material parameters for the InAs/GaSb/AlSb system. 44
4.2 Material structure for the single well RIT samples shown in Figure 4.1. The barrier structures are summarized in Table 4.3. 45
4.3 Barrier structure for single well RIT samples summarized in Table 4.2. All layers are undoped. 45

5.1 Material structure for the inter/intra band tunneling samples shown in Figure 5.1. The barrier structures are summarized in Table 5.2. 84
5.2 Barrier structure for the inter/intra band tunneling samples shown in Figure 5.1. All layers are undoped. 84
## List of Abbreviations and Symbols

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>2DEG</td>
<td>Two-dimensional Electron Gas</td>
</tr>
<tr>
<td>A</td>
<td>Area</td>
</tr>
<tr>
<td>a</td>
<td>Lattice Constant</td>
</tr>
<tr>
<td>AlSb</td>
<td>Aluminum Antimonide</td>
</tr>
<tr>
<td>Au</td>
<td>Gold</td>
</tr>
<tr>
<td>B</td>
<td>Magnetic field</td>
</tr>
<tr>
<td>CBE</td>
<td>Chemical Beam Epitaxy</td>
</tr>
<tr>
<td>DOS</td>
<td>Density of States</td>
</tr>
<tr>
<td>DBRTS</td>
<td>Double-barrier resonant-tunneling structure</td>
</tr>
<tr>
<td>$E_F$</td>
<td>Fermi energy</td>
</tr>
<tr>
<td>$E_m$</td>
<td>Energy of maximum current</td>
</tr>
<tr>
<td>$E_O$</td>
<td>Resonance energy</td>
</tr>
<tr>
<td>$E_{th}$</td>
<td>Threshold energy</td>
</tr>
<tr>
<td>FFT</td>
<td>Fast Fourier Transform</td>
</tr>
<tr>
<td>G</td>
<td>Conductance</td>
</tr>
<tr>
<td>GaAs</td>
<td>Gallium Arsenide</td>
</tr>
<tr>
<td>GaSb</td>
<td>Gallium Antimonide</td>
</tr>
<tr>
<td>$h$</td>
<td>Planck's constant</td>
</tr>
<tr>
<td>$\hbar$</td>
<td>Planck's constant divided by $2\pi$</td>
</tr>
<tr>
<td>I</td>
<td>Current</td>
</tr>
<tr>
<td>Symbol</td>
<td>Description</td>
</tr>
<tr>
<td>-------</td>
<td>-------------</td>
</tr>
<tr>
<td>InAs</td>
<td>Indium Arsenide</td>
</tr>
<tr>
<td>$J$</td>
<td>Current density</td>
</tr>
<tr>
<td>$k$</td>
<td>Wavevector</td>
</tr>
<tr>
<td>$k_B$</td>
<td>Boltzmann's constant</td>
</tr>
<tr>
<td>$m^*$</td>
<td>Effective mass</td>
</tr>
<tr>
<td>$m^*_e$</td>
<td>Electron effective mass</td>
</tr>
<tr>
<td>$m^*_lh$</td>
<td>Light-hole effective mass</td>
</tr>
<tr>
<td>$m^*_hh$</td>
<td>Heavy-hole effective mass</td>
</tr>
<tr>
<td>$m^*_E$</td>
<td>Emitter effective mass</td>
</tr>
<tr>
<td>$m^*_W$</td>
<td>Well effective mass</td>
</tr>
<tr>
<td>MBE</td>
<td>Molecular Beam Epitaxy</td>
</tr>
<tr>
<td>nm</td>
<td>nanometer</td>
</tr>
<tr>
<td>NDR</td>
<td>Negative differential resistance</td>
</tr>
<tr>
<td>P/V</td>
<td>Peak-to-valley</td>
</tr>
<tr>
<td>$q$</td>
<td>Electron charge</td>
</tr>
<tr>
<td>RIT</td>
<td>Resonant interband tunneling diode</td>
</tr>
<tr>
<td>RTD</td>
<td>Resonant tunneling diode</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

The concept of a functional device is one where a single device is designed to produce transfer characteristics that otherwise would have taken a number of devices to implement. One example of such a device is the Esaki tunnel diode, which can be fabricated with either a homojunction [11] or a heterojunction [1]. A tunnel diode has a peaked I(V) characteristic and associated region of negative differential resistance (NDR) which, with the appropriate load device, has two stable operating points. The application of the NDR characteristics of Esaki tunnel diodes to digital switching circuits was recognized and demonstrated shortly following the first demonstration of the tunnel diode [4]. The voltage location of the current peak is largely determined by the physical parameters of the materials used. However, the large junction capacitance and difficulty in fabrication of uniform Esaki tunnel diodes proved to be a stumbling block for integrated circuits [53, 56].

A similar I(V) characteristic can be obtained with a resonant tunneling diode (RTD) [5, 48]. An RTD consists of a number of abrupt layers of epitaxially grown heterojunctions, usually in III/V semiconductors, with the material composition of the individual layers chosen to create a double barrier structure in the conduction (or valence) band of system. As the layer thickness approach the de Broglie wavelength of the electrons in the system, quantum size effects become important, and the tunneling probability will be selectively enhanced at certain energies such that a peaked I(V) characteristic is created as carriers are
injected through this energy range. This effect is called resonant tunneling. Because the energy profiles are created by using different epitaxially grown materials rather than diffused dopant, many of the problems associated with Esaki tunnel diodes have been avoided. One new complication, however, is the fact that the energy location of the transmission peak is sensitive to fluctuations of the layer thickness on order of a single atomic layer.

RTDs have possible applications in both high speed switches and oscillators, due to the inherent speed of the tunneling process, and in reduced complexity circuits, resulting from the increased functionality of the device. RTDs have shown oscillations as high as 712 GHz [2], and switching times as short as 1.7 ps [38]. A number of device and circuit designs have been demonstrated that take advantage of this increased functionality [22, 46]. By combining multiple devices, to give a multiple-peaked I-V, functionality can be increased even further [39, 45, 50]. In addition to discrete RTDs, integration of RTDs with conventional transistor structures into a single three-terminal device has been demonstrated [3, 32, 44].

Most RTDs involve carrier transport in either the conduction or valence band, but not both. This is contrary to the case of an Esaki tunnel diode in which both the conduction band and valence band are involved. It is possible to create a tunneling structure that combines features of both RTDs and Esaki tunnel diodes, that is, current peak locations determined by material parameters and high speed associated with resonant tunneling. These structures are called resonant interband tunneling devices (RITs).

RIT devices differ from conventional RTDs in that the confined states, accessible to electron (or hole) transport lie in the valence (or conduction) band rather than the conduction band [24, 51]. A single well resonant interband tunneling structure was first demonstrated by Soderstrom et al. [51]. This study reported on a series of structures, consisting of InAs contacts, a GaSb well with AlSb barriers, exhibiting relatively large peak-to-valley (P/V) current ratios (20 @ 295 K, 88 @ 77 K). At the time this was the largest P/V ratio, which is a significant figure of merit, reported in a tunneling structure. Since then, there have been demonstrations of related structures including polytype structures, where one contact is
n-type InAs and the other p-type GaSb [65, 7, 6, 61, 20]. There have been relatively few experimental investigations into the nature of the interband tunneling process. Because of the opposite sense of the dispersions of the conduction and valence band, and the existence of both light- and heavy-hole subbands, the tunneling process is expected to be somewhat more complicated in an RIT structure than an RTD because of the multiple subbands involved.

This work examines the details of the interband tunneling process in InAs/AlSb/GaSb structures. Chapter 2 introduces resonant tunneling and resonant interband tunneling, and a simple model of such tunneling based on a density of states argument. Starting from the basic principles of quantum size effects, resonant tunneling structures are described. A simple method to model I(V) characteristics related to those effects will be summarized and then applied to interband tunneling structures, including the modification of the tunneling current under influence of an applied magnetic field. Chapter 3 is an overview of the experimental methods used in this work including device fabrication and measurement. Chapter 4 describes measurements on homotype (InAs/AlSb/GaSb/AlSb/InAs) structures which show evidence of interband tunneling through multiple subbands at $k_{\parallel} \neq 0$. The results of the measurements are used to determine critical points of the electronic subband structure of the well, indicating what subbands are involved in the tunneling current. Chapter 5 covers results of the design of various multiple well interband tunneling structures that also exhibit intraband tunneling effects. Chapter 6 describes the operation of two compressed functionality circuits—a binary and a ternary adder—constructed using RTDs. Chapter 7 describes resonant tunneling effects in the turn-on characteristics of a tunneling hot electron transfer amplifier (THETA). A summary if this work is given in Chapter 8.
Chapter 2

Resonant Tunneling Theory

2.1 Introduction

Modern heteroepitaxial growth techniques allow fabrication of atomically precise structures consisting of different materials with the same crystalographic properties (e.g. lattice constant). In general, different materials with the same lattice constant will have different electrical properties; in particular, different bandgaps, as shown in Figure 2.1. Material systems that lie near or at the same point along the horizontal axis of Figure 2.1 are suitable for the growth of such structures. One of the most common material systems used for this purpose is the $\text{Al}_x\text{Ga}_{1-x}\text{As}$. At an interface between two distinct alloys, the difference in bandgap results in abrupt changes in conduction band. By combining thin layers of different materials, arbitrary electronic potential profiles can be created. For thin layers, approaching the electron wavelength in the material, quantum size effects become important.

This chapter will first introduce the basic principles of quantum size effects, in particular, energy quantization in a potential well. Electron tunneling through one of these quantized states, called resonant tunneling, gives a peaked $I(V)$ characteristic. A simple method to model $I(V)$ characteristics related to those effects will be summarized and then applied to interband tunneling structures.
Figure 2.1: Plot of Bandgap vs Lattice Constant for various semiconductors. The solid lines connecting various binary compounds indicate direct bandgaps, while the dashed lines indicate indirect bandgaps.
2.2 Quantum Size Effects

For specific material configurations of thin (≤ 200Å) layers, Quantum Size Effects become significant and can result in dramatic effects in the electrical properties of the structures [64]. The effects result from the spatial variation of the electrical potential approaching the electron wavelength.

2.2.1 Quantum Wells

The simplest structure to show quantum size effects is a quantum well, which consists of a thin layer of a material buried in a second material such that the conduction band edge of the thin layer lies below the conduction band edge of the surrounding material. When the thickness of the layer becomes small (≤ 200Å), the energy separation of allowable electronic wavefunctions in the well (in the confinement direction) is large enough that the electronic structure of the thin layer is no longer continuous and the layer becomes quasi-two dimensional.

For real materials systems (e.g. GaAs/AlGaAs) the barrier heights are large enough that the infinite barrier case accurately describes the position of the lowest lying energy states. The discrete energy levels, shown in Figure 2.2a, are given by

\[ E_n = \frac{(\pi \hbar)^2}{2 m^* L^2} n^2 \]  

(2.1)

where \( m^* \) is the bulk effective mass of the material and \( n \) is an integer.

It is important to note that while the allowed energies in the confinement direction (referred throughout this work as the \( z \)-direction) are discrete, there is no such modification of the transverse direction and therefore the total dispersion relation for the quantum well is given by

\[ E_n(k) = E_n + \frac{\hbar^2 k^2}{2 m^*} \]  

(2.2)

where \( k_t = k_x^2 + k_y^2 \) is the magnitude of the wavevector in the transverse direction.
Figure 2.2: Schematic illustration of quantum size effects in a *quasi-two dimensional* region. (a) The conduction band has only discrete allowable energies in the confinement direction, but still has a continuous density of states in the transverse direction. (b) The density of states increases in discrete steps corresponding to each discrete energy level in the confinement direction. At energies where the confined energy levels occur, the value of the 2d-density of states is equal to the value of the corresponding 3d-density of states.

For a bulk material, the 3d-density of states is given by

$$D_{3d} = \frac{2^{1/2} m^{3/2}}{\pi^2 \hbar^3} E^{1/2}$$  \hspace{1cm} (2.3)

As is implied by the dispersion relation, for each value of $n$, there is a continuous density of states, given by

$$D_{2d} = \frac{m^*}{\pi \hbar^2}$$  \hspace{1cm} (2.4)

Each discrete energy level adds a constant density of states creating a step-like function for the total density of states as shown in Figure 2.2b. The only modification to the energy range of the density of states is that the lowest energy corresponds to the $n = 1$ confined level instead of the bulk conduction band edge. As the width of the well increases, the energy spacing of the confined states decrease and the 2d-density of states approaches the 3d-density of states.
Figure 2.3: (a) Schematic of an RTD. (b) Transmission coefficient for the structure shown in (a).

2.2.2 Single Band Transmission Coefficient

If the confining barrier thickness is small enough, the wavefunctions in the quantum well are able to couple with states outside of the well, shown in Figure 2.3a for a one-dimensional case. Electrons incident on such a structure will see an increased transmission probability when they have an energy equal to one of the discrete states in the well. Such a structure is referred to as a resonant tunneling diode (RTD).

The energies of the resonant states can be determined by calculating a transmission coefficient for an electron through the tunneling structure as described by Tsu and Esaki [60], where the states lie at energies of enhanced transmission through the structure. The results are similar to that shown in Figure 2.3b.

For the structure shown in Figure 2.3a, the electronic potential is divided up into five regions, labeled #1 to #5, with each region having a constant potential, bounded by potential discontinuities. An electron in each region is treated as a plane wave with a
wavefunction of the form

\[ \psi_n(z) = a_n e^{ik_n z} + b_n e^{-ik_n z} \]  

(2.5)

where the subscript \( n \) indicates the \( n^{th} \) layer.

For an electron incident on the structure from the left, the “entering” and “leaving” wavefunctions in regions \#1 and \#5 respectively are given by

\[ \psi_1(z) = e^{ik_n z} + R e^{-ik_n z} \]  

(2.6)

\[ \psi_5(z) = T e^{ik_n z} \]  

(2.7)

where \( R \) is the reflected amplitude of the “entering” wavefunction and \( T \) is the transmitted amplitude “leaving” of the wavefunction.

By successively matching the wavefunction and wavefunction derivatives between adjacent regions, the reflected and transmitted amplitudes for the entire structure can be related. This is given (in matrix form) as

\[
\begin{pmatrix}
T \\
0
\end{pmatrix} = M_1 M_2 M_3 M_4 M_5 \begin{pmatrix} 1 \\ R \end{pmatrix}
\]

(2.8)

where \( M_n \) is a \( 2 \times 2 \) matrix that describes the wavefunction in the \( n^{th} \) region, indicated in Figure 2.3a, given by

\[
M_n = \frac{1}{4} \begin{pmatrix}
e^{ik_{n+2} z_{n+2}} & e^{ik_{n+2} z_{n+2}} \\
-e^{-ik_{n+2} z_{n+2}} & -e^{-ik_{n+2} z_{n+2}}
\end{pmatrix} \times 
\begin{pmatrix}
e^{ik_{n+1} z_{n+1}} & e^{-ik_{n+1} z_{n+1}} \\
-i(k_{n+1}/k_{n+2})e^{ik_{n+1} z_{n+1}} & -i(k_{n+1}/k_{n+2})e^{ik_{n+1} z_{n+1}}
\end{pmatrix} \times 
\begin{pmatrix}
1 + i(k_n/k_{n+1}) & 1 - i(k_n/k_{n+1}) \\
1 - i(k_n/k_{n+1}) & 1 + i(k_n/k_{n+1})
\end{pmatrix}
\]

(2.9)
\[ k_n = \frac{[2m_n^*(V_n - E)]^{1/2}}{\hbar} \]  

(2.10)

where \( z_n \) is the thickness, \( V_n \) is the potential and \( m_n^* \) is the effective mass of the \( n \)th layer.

Taking the "propagation" matrix to be the matrix product

\[ P = M_1 M_2 M_3 M_4 M_5, \]  

(2.11)

the transmission coefficient through the double barrier structure is then given by

\[ T(E) = P_{11} - \frac{P_{12} P_{21}}{P_{22}} \]  

(2.12)

### 2.2.3 Resonant Tunneling Diode I(V) Characteristics

Placing a voltage bias across an RTD allows the electronic structure of the quantum well to be probed. Figure 2.4a shows an RTD with no applied bias.

As the bias is increased, the confined state is pulled down in energy towards the emitter Fermi level. Because of the increased transmission coefficient, carriers are able to tunnel through the structure by means of the intermediate state. Current will increase, peaking when the confined state approaches the bottom of the emitter conduction band. Once the confined state is pulled below the emitter conduction band edge, the emitter electrons no longer see the increased transmission probability, resulting in a sharp decrease in current. Physically, cutoff of the resonant current occurs because of conservation of transverse momentum.

Superimposed on the resonant current, is a nonresonant background current associated with tunneling through the entire structure taken as a single tunnel barrier such that \( I_{nonres} \sim e^{V_{bias}} \). The I(V) characteristic expected from the structure of Figure 2.4 is shown in Figure 2.5. Further increasing the bias will bring the next confined state into resonance, resulting in another peak in the I(V) relation.
Figure 2.4: Band plots of a resonant tunneling structure under different applied bias. By convention, the region to the left of the tunneling structure is referred to as the emitter. The dotted line in the emitter represents the quasi-Fermi level for the emitter. (a) No applied bias. (b) Biased to the peak resonant current (1 V). (c) Biased past resonance (1.5 V). The I(V) characteristics are shown in Figure 2.5.
Figure 2.5: Expected I(V) characteristics for the structure shown in Figure 2.4. $V_{\text{peak}}$ corresponds to (b) and $V_{\text{valley}}$ corresponds to (c). The dotted line connecting the two points corresponds to the NDR region when the device is oscillating. The gray dotted line corresponds to the expected I(V) characteristic of a single tunneling barrier with thickness equal to the sum of the two barriers.
2.3 Interband Tunneling

The Esaki tunnel diode is a \textit{pm}-junction diode where each side of the junction is doped to degeneracy \cite{11}. Figure 2.6a shows the band diagram for a Esaki tunnel diode with no applied bias. When a small bias is placed across the junction, carriers can tunnel directly from the \textit{n}-type conduction band to the \textit{p}-type valence band. Because of the heavy doping, the tunnel barrier is relatively thin, allowing the tunnel current to be significant compared to the diffusion current in either the conduction or valence band. When the bias is increased, as in Figure 2.6b, the empty states in the \textit{p} side are no longer available to the electrons from the \textit{n} side and the tunnel current is cut off. The resulting \textit{I(V)} characteristic is similar in shape to the RTD \textit{I(V)} characteristic in Figure 2.5, except that \textit{V_{peak}} will occur at a significantly lower voltage.

One of the benefits of the Esaki tunnel diode is the efficiency with which the tunnel current is cut off. Where the current cutoff in an RTD occurs only because of conservation of transverse momentum, the current cutoff in an Esaki tunnel diode occurs because of both conservation of transverse momentum and conservation of energy. The peak/valley current ratios of these, however, are limited by the diffusion current of majority carriers across the \textit{pn} junction. In RTDs at room temperature, inelastic scattering involving optical phonons is a significant source of valley current. For the Esaki tunnel diode, the role of conservation of energy in the tunnel current cutoff reduces the impact of inelastic scattering.

2.3.1 Resonant Interband Tunneling Structures

Most common materials systems are Type I, where the bandgap of one material lies entirely within the bandgap of the other material (e.g. GaAs/Al\textsubscript{\textit{x}}Ga\textsubscript{1-x}As). In some systems, however, the bandgap of one material lies at most partially within the bandgap of the second material. This is called a Type II interface. If the bandgap of the two materials do not overlap, the interface is said to be broken (e.g. InAs/GaSb). Examples of both types of interfaces are shown in Figure 2.7a.
Figure 2.6: Band diagram of an Esaki tunnel diode. The dotted line is the quasi-Fermi level in each region. (a) No applied bias (b) Biased beyond the tunnel current cutoff.
Figure 2.7: (a) Schematic showing two different types of band alignments. (b) Resonant tunneling structures fabricated from a Type I material system and a Type II material system. The gray lines indicate the confined states.
Using a system with a broken Type II band alignment, an electron tunneling structure in which a confined state (accessible to the emitter conduction band) is located in the valence band is possible. One such system is the InAs/AlSb/GaSb. The InAs conduction band edge lies $\sim 150\text{meV}$ below the top of the GaSb valence band, regardless of the inclusion of an intermediate AlSb layer [12]. The lattice mismatch of this system is small enough (see Figure 2.1), that good quality heterostructures can be grown.

An example of such a structure using the InAs/AlSb/GaSb system is shown in Figure 2.7b. In this structure, the resonance occurs as a result of the confined state in the quasi-two dimensional region; it is called an interband tunneling structure because tunneling occurs from a conduction band via a valence band. The structure will produce the peaked $I(V)$ characteristic with just InAs/GaSb interfaces because the mismatch between the wavefunctions at the material interface is sufficient to create a confined state [69]. The addition of the AlSb barriers, however, adds a real tunnel barrier to the structure which greatly reduces the current due to inelastic tunneling, significantly increasing the $P/V$ ratio for the structure.

In the past few years, a number of device structures based on interband tunneling have been proposed and demonstrated [24, 51, 54, 65]. These devices, called Resonant Interband Tunneling (RIT) structures combine the current cutoff mechanism of interband tunneling similar to that of the Esaki tunnel diode, with the growth control of heteroepitaxial structures. Unless otherwise noted, the RITs discussed in this work are similar to that described in [51] and shown in Figure 2.7b.

### 2.3.2 Two-band Transmission Coefficient

In the discussion in Section 2.2.2, only a single band—the conduction band—was involved in the calculation of the RTD transmission coefficient. For an RIT, however, it is necessary to include at least a second band to account for the fact that the confined state lies in the conduction band.

According to $k\cdot p$ theory, the coupling between the conduction band and light-hole
band is much larger than the coupling between the conduction band and heavy-hole band. At \( k = 0 \), there is only coupling between the conduction band and the light-hole band [21]. Therefore, in calculating the one-dimensional transmission coefficient through an interband tunneling structure (at \( k_\parallel = 0 \)), only the conduction band and the light-hole band need be included.

A method, similar to that for the one-band model, for calculating the two-band transmission coefficient has been shown by Yang et. al [66]. A two band Hamiltonian can be written as

\[
\begin{bmatrix}
E_c + \frac{\hbar^2 k^2}{2m_o} & \hbar P k \\
\hbar P k & E_v - \frac{\hbar^2 k^2}{2m_o}
\end{bmatrix}
\begin{bmatrix}
\Psi_c \\
\Psi_v
\end{bmatrix}
= E
\begin{bmatrix}
\Psi_c \\
\Psi_v
\end{bmatrix}
\]

(2.13)

where \( k \) is the wavenumber, \( m_o \) is the free electron mass, \( P \) is the momentum matrix element between the conduction band and light-hole band and \( E_c \) and \( E_v \) are the energies of the conduction and light-hole bands, respectively. \( P \) can be determined from

\[
m^* = \left( m_o^{-1} + \frac{2P^2}{E_c - E_v} \right)^{-1}.
\]

(2.14)

The solutions to (2.13) are of the form

\[
\begin{bmatrix}
\Psi_c \\
\Psi_v
\end{bmatrix}
= \begin{bmatrix}
A e^{ikz} + B e^{-ikz} \\
\beta_k (A e^{ikz} - B e^{-ikz})
\end{bmatrix}
\]

(2.15)

where

\[
\beta_k = \frac{\hbar P k}{(E_c - E_v)}
\]

(2.16)

\[
k = \frac{[(E - E_v)(E - E_v)]^{1/2}}{\hbar P}
\]

(2.17)

The wavefunctions \( \Psi_c \) and \( \Psi_v \) can then be matched across each layer, according to

\[
\begin{bmatrix}
A_1 \\
B_1
\end{bmatrix}
= M
\begin{bmatrix}
A_n \\
B_n
\end{bmatrix}
\]

(2.18)
\[ M = \frac{1}{2} \begin{bmatrix} 1 & \frac{1}{\beta_{k1}} \\ \frac{1}{\beta_{k1}} & -1 \end{bmatrix} \begin{bmatrix} M_{1,1}^P & M_{1,2}^P \\ M_{2,1}^P & M_{2,2}^P \end{bmatrix} \begin{bmatrix} 1 & 1 \\ \beta_{kn} & -\beta_{kn} \end{bmatrix} \]  
(2.19)

\[ M^P = P_2 P_3 \ldots P_{n-1} \]  
(2.20)

\[ P_n = \begin{bmatrix} \cos(k_n d_n) & -i \sin(k_n d_n)/\beta_{kn} \\ -i \beta_{kn} \sin(k_n d_n) & \cos(k_n d_n) \end{bmatrix} \]  
(2.21)

where \( P_n \) is called the propagation matrix for the \( n \)th layer, and \( d_n \) is the thickness of the \( n \)th layer. Note that the contact layers are not included in (2.20). The transmission coefficient can then be written as

\[ T(E) = (\pm) \frac{\beta_{kn}}{\beta_{k1}} |M_{\mu,\nu}|^{-2}. \]  
(2.22)

For tunneling between like bands (e.g. conduction to conduction as in Figure 2.7b), the leading sign is plus and \( \mu = \nu \); for tunneling between unlike bands, the leading sign is negative and \( \mu \neq \nu \).

The calculated two-band transmission coefficient for two different structures are shown in Figure 2.8. The structure in consists of a single GaSb well creating confined states in the valence band accessible to the conduction band of the InAs contact. There are two peaks in the transmission coefficient for this structure, implying that the \( n = 1 \) and \( n = 2 \) states in the well are available for resonant tunneling. Note that both peaks occur within the 150 meV window available for interband tunneling.

The second structure, shown in Figure 2.8b, has two GaSb wells and a single InAs well. The transmission coefficient for this structure has three peaks. Two of the peaks occur within the 150 meV interband tunneling window, while the other occurs at a higher energy.

The structure at \( E = 0.22 \text{eV} \) results from intraband tunneling, that is, tunneling involving only a single band, as discussed in Section 2.2.2. In this case, only the InAs conduction band is involved. Applying the one-band model to this structure would result on only a single peak in the transmission coefficient corresponding to this one.
Figure 2.8: Calculated two-band transmission coefficient for two different heterostructures. (a) Single well interband tunneling structure. (b) Double well intra/interband tunneling structure. For both cases, the zero in energy is taken to be the bottom of the InAs conduction band. The gray lines indicate the confined states.
Though each structure has two peaks in the interband tunneling window, they result from different mechanisms. The two peaks in the single well structure result from two distinct confined states corresponding to $n = 1, 2$. The two peaks in the double well structure result from the doublet state created by the wavefunction overlap from the two wells. Despite the fact that the GaSb layer thicknesses are the same for each structure, the $n = 2$ state is not seen in the double well structure because of the thickness of the AlSb barriers. As the thickness of the barriers decrease, the mixing between the two bands increases, influencing the energy of the confined state [67].

2.4 Density of states model of resonant tunneling

Resonant tunneling can be modeled using a simple density of states argument first proposed by Luryi [25]. This model, which does not require coherence of the electron wavefunction, relies on a transmission coefficient which describes the electrical properties of the quantum well, which can be obtained as outlined in Sections 2.2.2 or 2.3.2. The transmission coefficient is calculated to be a wavefunction across the entire tunneling structure such that an electron coherently tunnels across the entire structure [42]. The treatment by Luryi, however, assumes only that the electron can tunnel into the well and the transmission coefficient need only account for this. Once the electron is in the well, it is assumed that it can easily escape out to the collector. This is commonly called the sequential picture. It has been shown that if the transmission coefficient is much more narrow than the supply of incoming electrons, the coherent and the sequential picture are indistinguishable [63]. This is the case for the structures to be discussed, so the calculated transmission coefficients are valid.

The following discussion follows the expanded treatment by Ohno [37]. This model assumes the current arises from carriers elastically scattered into the well, which implies both conservation of energy and conservation of momentum parallel to the confining interfaces (perpendicular to the direction of current flow). It does not assume coherence of the
electron wavefunction.

2.4.1 "Conventional" Resonant Intraband Tunneling

A general expression for tunneling current density is obtained by summing over all available carriers as follows:

\[ J(E) = q \int N(k) v(k) T(E, k) \, dk \]  \hspace{1cm} (2.23)

where \( N(k) \) is the density of carriers available for tunneling, \( v(k) \) is the carrier velocity, and \( T(E, k) \) is the transmission coefficient through a plane in space perpendicular to the current direction.

For a double barrier tunneling structure, the available carrier densities and velocities in the emitter are described by

\[ N(k) = 2 \left( \frac{1}{2 \pi} \right)^3 \]  \hspace{1cm} (2.24)

\[ v(k) = \frac{1}{\hbar} \frac{\partial E}{\partial k_z} \]  \hspace{1cm} (2.25)

where the current is taken to flow in the \( z \)-direction. A reasonable approximation for the transmission coefficient is to assume a Lorentzian lineshape,

\[ T(E) = |T_O|^2 \frac{\Gamma^2}{(E - E_O)^2 + \Gamma^2}. \]  \hspace{1cm} (2.26)

If the transmission width is much smaller than the incoming supply of electrons (i.e. the emitter Fermi energy), it can further simplified to

\[ T(E) \approx |T_O|^2 \Gamma \delta(E - E_O) \]  \hspace{1cm} (2.27)

where \( \Gamma \) is the resonance width. This allows (2.23) to be written as

\[ J(E) = \frac{q |T_O|^2 \Gamma}{2 \pi^2 \hbar} \int k_i dk_i \int dk_z \frac{\partial E}{\partial k_z} \delta(E_F - E_O) \]

\[ = \frac{q |T_O|^2 \Gamma}{2 \pi^2 \hbar} \int k_i dk_i. \]  \hspace{1cm} (2.28)
The choice of integration limits over $k_q$ is determined by those carriers which can conserve transverse momentum ($k_q$) during the tunneling process. Figure 2.9 is a graphical representation of this condition for the case of $m_{W}^* > m_{E}^*$. This model assumes that the emitter is purely three dimensional, with no confinement effects in the accumulation layer at the emitter/well interface. The emitter Fermi sea is represented by the shaded area enclosed by $E_F$ and $E_E(k_q)$ ($T = 0K$). Current can flow only when there are available states in the well such that both energy and transverse momentum ($k_q$) are conserved. When $E_O \geq E_F$ or $E_O < 0$, there is no overlap between available carriers and available states and therefore, there will be no current flow because energy cannot be conserved. For $E_F > E_O \geq 0$, there is overlap and current will flow. The integration limits in (2.28) are then the extent of this overlap along the horizontal axis. Note that only the range of the horizontal axis is important, the overlap range over the vertical axis is not important. Assuming parabolic bands, (2.28) can then be evaluated to be

$$J(E_O) = \begin{cases} 
0 & E_O \geq E_F \\
\frac{g m_{W}^* |T_o|^2 \Gamma}{2\pi^2 h^3} (E_F - E_O) & E_F > E_O \geq E_m \\
\frac{g \mu |T_o|^2 \Gamma}{2\pi^2 h^3} E_O & E_m > E_O \geq 0 \\
0 & E_O < 0 
\end{cases} \quad (2.29)$$

where $1/\mu = 1/m_{E}^* + 1/m_{W}^*$ and $E_m = E_F (1 - m_{E}^*/m_{W}^*)$. This also describes the case when $m_{W}^* = m_{E}^*$ with the condition that the third item in (2.29) does not apply and $E_m = 0$. The case for $m_{W}^* < m_{E}^*$ is also similar; however, current does not turn off at $E_O = 0$, but continues to flow until $E_O = E_m < 0$. This is because it is still possible to conserve transverse momentum for values of $k_q \neq 0$. In all three cases, current begins to flow as carriers at $k_q = 0$. 
Figure 2.9: (a) Band diagram of a double barrier tunneling structure. The position in energy of the resonant state in the well can be modified by means of an applied bias such that $\Delta E = V_{\text{bias}}/\alpha$. $E_O$ is the energy of the well dispersion at $k_\parallel = 0$. (b) Energy vs $k_\parallel$ for both the emitter and well at four different bias points. The shaded region represents the available carriers (i.e. the emitter Fermi sea). The thick regions on curves #2 and #3 represent those available carriers which can tunnel into the well while conserving transverse momentum. For curves #1 and #4, no tunneling is possible. In this schematic, the effective masses of the emitter and well are not assumed to be the same. (c) Current Density vs Energy for the example shown in (b) as described by equation 2.28. Adapted from reference [37].
2.4.2 Resonant Interband Tunneling

The density of states model of resonant tunneling is easily extensible to interband tunneling. The initial assumptions of Section 2.4.1 are still valid. By comparing Figure 2.7a with Figure 2.8a, it is apparent that the form of the transmission coefficient can still be taken to be that given in (2.26), therefore, (2.28) is still valid. The only modification resulting from interband tunneling is the choice of integration limits in (2.28). The interband tunneling case is shown in Figure 2.10.

For the case of a confined valence band, the assumption of a single parabolic light-hole band is only a crude approximation at best because of the interaction between the light-hole and heavy-hole states [14]. However, as will be discussed in Chapter 4, this approximation gives reasonable qualitative results for certain situations.

In the same manner as for the intraband case, (2.28) can then be evaluated to be

\[
J(E_O) = \begin{cases} 
0 & E_O < 0 \\
\frac{\mu |T_{01}|^2 \Gamma}{2\pi^2 h^3} \left[ E_F \left(1 + \frac{m_W^*}{m_E^*}\right) - E_O \frac{m_W^*}{m_E^*} \right] & E_O \geq E_{th} \\
\frac{\mu |T_{01}|^2 \Gamma}{2\pi^2 h^3} E_O & E_{th} < E_O \geq E_F \\
0 & E_F > E_O \geq 0 \\
0 & E_O < 0 
\end{cases} 
\] (2.30)

where \(1/\mu = 1/m_E^* + 1/m_W^*\) and \(E_{th} = E_F \left(1 + m_E^*/m_W^*\right)\). Again, the assumption of parabolic bands has been made.

Contrary to the case of intraband tunneling, current begins to flow before \(E_O = E_F\). This is true regardless of the relationship of the two effective masses, resulting from the difference in sign of \(m_E^*\) and \(m_W^*\). For the situation shown in Figure 2.10, current will begin to flow at some energy threshold \(E_{th}\) which is reached when both energy and \(k_z\) are conserved. For curve #1, no current will flow because though energy can be conserved, \(k_z\) cannot be conserved [28]. This is opposite from the intraband case, in which the limiting factor is energy conservation. The maximum current is reached for \(E_O = E_F\). Also, for all cases, current is cutoff for \(E_O < 0\).
Figure 2.10: (a) Band diagram of a double barrier tunneling structure. Only the region near the interband tunneling window is shown. $E_O$ is the energy of the well dispersion at $k_\parallel = 0$. (b) Energy vs $k_\parallel$ for both the emitter and well at four different bias points. The shaded region represents the available carriers (i.e. the emitter Fermi sea). The thick regions on curves #2 and #3 represent those available carriers which can tunnel into the well while conserving transverse momentum. For curves #1 and #4, no tunneling is possible. (c) Current Density vs Energy for the example shown in (b) as described by equation 2.30.
2.5 Magnetic Field Effects

When $\omega_c \tau \gg 1$, where $\omega_c = qB/m^*$ is the cyclotron frequency and $\tau$ is the average collision time, a magnetic field will have a significant effect on a two-dimensional electron gas (2DEG). For a 2DEG with no applied field, momentum states are uniformly distributed over $k_x$ and $k_y$, as shown in Figure 2.11a, leading to a constant density of states, given by (2.4). With the application of a magnetic field perpendicular to the 2DEG, the allowed momentum states are constrained to concentric circles of constant radius (Figure 2.11b)

$$k_{nn}^2 = k_x^2 + k_y^2 = \frac{2qB}{\hbar} \left( n + \frac{1}{2} \right)$$ (2.31)

which gives a discrete density of states

$$E_{nk} = \hbar \omega_c \left( n + \frac{1}{2} \right)$$ (2.32)

called Landau levels. Each Landau level is degenerate with $qB/\hbar$ states per level. Despite the change in the density of states, the total number of states is not changed. Ideally, the 2DEG will now only take on discrete values of $k_n$; however, in a real system there will be both broadening associated with each level, and localized states so the density of states will be as shown in Figure 2.11c. The corresponding Energy vs $k_n$ curve is no longer continuous. In Figure 2.11d, the gray curve is what is expected for the no-field case. The thick segments represent the spread of the energies where the density of states is greater than that for the no field case in Figure 2.11c.

In a three dimensional system, quantization will still occur in the plane, but because there still is a continuous density of states in the $z$-direction, the delta-function like density of states like that in Figure 2.11c will not occur. The three-dimensional density of states will be continuous, but will have the periodic modulation of the $x$-$y$ density of states superimposed on the original density of states. For these magnetic effects to occur, the carrier scattering time must be sufficiently long. If the scattering time is short enough such
Figure 2.11: Effects of a magnetic field applied perpendicular to a 2DEG system. (a) Allowed momentum states for a two dimensional electron gas with no applied magnetic field. The states are evenly spaced in both $k_x$ and $k_y$. (b) With a magnetic field applied perpendicular to a 2DEG electron gas. Allowed states lie on concentric circles of constant radius $k^2 = (2qB/\hbar)(n + 1/2)$. An equal number of states lie on each circle. (c) Density of states for no applied field (gray line) and applied field (dark line). The effects of broadening are shown. (d) The gray curve corresponds to Energy vs $k_\parallel$ for a 2DEG with no applied field. The dark segments correspond to energies where the density of states is greater than that of the no field case.
that $\omega_c \tau \sim 1$, the energy broadening will be enough such that the Landau effects will be washed out.

### 2.5.1 Field Parallel to Current

In (2.23), the term $T(E, k)$ contains all the information about the quantum well. For the no-field case, the 2d-DOS is constant, resulting in no specific dependence of $T$ on $k$. The magnetic field effects can be taken into account mathematically by adding a $k_\parallel$ dependent term to the transmission coefficient. For simplicity, this was chosen to be

$$T(k_\parallel) = T_b + T_l(k_\parallel)$$  \hspace{1cm} (2.33)

$$T_l(k_\parallel) = \begin{cases} T_{l_0} & (k_{in}^2 - \Gamma_l)^{1/2} < k_\parallel < (k_{in}^2 + \Gamma_l)^{1/2} \\ 0 & \text{otherwise} \end{cases} \hspace{1cm} (2.34)$$

where $T_b$ is the background transmission coefficient related to either localized states or Landau level broadening, $T_{l_0}$ is the transmission coefficient for a Landau level and $\Gamma_l$ is the energy broadening associated with each Landau level.

Simulated I(V) characteristics for the case of Figure 2.10b with $B \parallel I$ are shown in Figure 2.12 for different values of field. The similarity to the predicted current of Figure 2.10c is apparent, with fluctuations of the magnetic field effects superimposed on the no-field I(V) characteristics. The peak (and shoulder) positions of the I(V) characteristics as a function of field are plotted in Figure 2.13, resulting in a fan diagram.

In the fan diagram, two distinct fans are apparent. The low bias fan comes from region #2 of Figure 2.10c, while the high bias fan comes from region #3. The current peaks from the low bias fan are a consequence of the fluctuating density of states in the well being pulled down through the Fermi level, with the fan converging to the point where the top of the well dispersion (at $k_\parallel = 0$) crosses $E_F$. The high bias fan results from the fluctuating density of states moving through the bottom of the emitter conduction band, with the fan converging to the point where the top of the well dispersion crosses the bottom
Figure 2.12: Simulated I(V) characteristics for the situation illustrated in Figure 2.10c. The units on both current and bias are arbitrary.

of the emitter conduction band.

2.5.2 Field Perpendicular to Current

A 2DEG is only collapsed into Landau levels when the magnetic field is applied perpendicular to the confining interfaces. When the field is applied parallel to the interfaces, no such effect occurs. The only first order perturbation to the emitter electron population is an extra component of \( k_\parallel \) while the electron energy distribution remains the same [16]. This is illustrated in Figure 2.14 which shows the emitter dispersion shifted by

\[
\Delta k_\parallel = \frac{qB \Delta l}{\hbar}
\]  

(2.35)

where \( \Delta l \) is the distance in the \( z \)-direction from the average emitter electron position to the center of the well. One result of this is a change in the threshold voltage for the turn-on of tunneling current for some particular band. Because of the opposite polarity of the emitter and well, the initial current flow involves carriers located away from \( k_\parallel = 0 \). As the applied bias is increased, the point where the well dispersion crosses the emitter quasi-Fermi energy
Figure 2.13: Fan diagram (current peak positions vs magnetic field) for the simulated I(V) characteristics of Figure 2.12.
moves horizontally along the $k_z$ axis towards $k_z = 0$. As indicated by (2.35), the magnetic field will move the emitter electrons horizontally towards the intersection point. By varying the strength of the field, $E_{th}$ will map out that portion of the dispersion curve.

### 2.6 Summary

Quantum size effects can lead to tunneling structures with novel I(V) characteristics. Transmission resonances through double barrier tunneling structures leads to a peaked I(V) characteristic. The position in energy can be calculated by considering the incident electrons to be plane waves. Using this transmission peak, a simple model can be developed to model the I(V) characteristics through the tunneling structure. While the tunneling structures usually involve only a single band (e.g. the conduction band in the GaAs/AlAs system) Some material systems have band alignments such that both the conduction and valence bands are involved (e.g. InAs/AlSb/GaSb). The model developed in this chapter is capable of describing this situation.

A magnetic field applied to the tunneling structure will modify the subband structure
Figure 2.15: Current threshold for various values of $\Delta k_4$ for Figure 2.14. Each trace corresponds to a different magnetic field with $\Delta k_1 < \Delta k_2 < \Delta k_3 < \Delta k_4$. $\Delta k_4$ corresponds to the case where the emitter dispersion boundary, the well dispersion and $E_F$ intersect at the same point.

in the well, and therefore, the transmission characteristics, with different effects occurring depending on the field orientation. These effects make it possible to experimentally determine critical points in the subband structure.
Chapter 3

Experimental Methods

3.1 Introduction

The experimental work done included all phases of fabrication and data acquisition except for the epitaxial growth. The Sb-based structures were grown by Motorola. The THETA devices described in Chapter 7 were grown and fabricated at Texas Instruments. A custom computer controlled data acquisition system, running over a GPIB network, capable of acquiring multiple I-V sweeps with resolution limited by the resolution of the individual instruments (e.g. DMM) was created. The details of the fabrication and data acquisition system, and the cryostats used are given below.

3.2 Device Fabrication

The two-terminal InAs/AlSb/GaSb devices were fabricated using standard lithographic techniques following the procedure illustrated in Figure 3.1. All fabrication was done in the Yale Microfabrication Facility. The lithography was done using an HTG mask aligner, all metal evaporation was done using an electron-beam evaporator and the plasma processing (RIE/PECVD) was done using a Vacutec system. The measured intensity of the lamp in the HTG mask aligner was 5.2 mW/cm² @ 365 nm, 8.6 mW/cm² @ 405 nm.
The mesa structures were etched using $\text{H}_2\text{SO}_4:\text{H}_2\text{O}_2:\text{H}_2\text{O}$ (1:8:160), a standard GaAs wet etch. The etch rate of InAs was similar to the predicted etch rate of GaAs ($\sim$2500 Å/min). When the Sb-based layers were reached, etching slowed significantly, though progress could be seen by observing a change in interference color on the sample. Etching through the tunneling structure (100 Å~150 Å) took $\sim$45 sec.

The bonding pads were deposited by electroplating. This was done to give a thick (> 1 μm) gold film to improve the reliability of wire bonding to the sample. Initial samples, which were fabricated with evaporated bonding pads, had problems with both adhesion of wire bonds and failure after thermal cycling. The mask set used for electroplated bonding pads was designed to produce air bridges between the mesa contact and the bonding pad improving device reliability after thermal cycling. The plating solution used was Orotemp24 Gold Plating Solution, a proprietary mixture based on $\text{KAu(CN)}_2$ manufactured by Technic, Inc.

Electroplating requires a continuous film over the entire sample to allow electrical continuity. A photoresist layer is patterned on the sample, then thin film of metal is evaporated over the surface. The film is evaporated at a slight angle, then the sample is rotated 180° and the evaporation is repeated. Also, no chlorobenzene soak is done. This insures that the film is continuous over the step edges of the patterned photoresist. After the evaporation is completed, another layer is then patterned with photoresist exposing only those areas where plating is desired. Except for the small gap which creates the air bridge, this pattern is identical to the previous pattern. The top layer of resist and the thin metal film are removed (this also removes the equivalent thickness from the electroplated metal). The bottom layer of resist is left as a protective layer.

Each sample to be plated was about 1 cm². The solution was heated to 60 °C. The sample (cathode) and a platinum anode were placed in the solution with a 3 V bias and a 1 kΩ current limiting resistor. For a 1 cm² sample, this gave a current of 1.2 mA. This deposited Au at a rate of about 1 μ/20 min.

Using the following definitions:
• Sample clean
  1. 5 min TCA @ 50 °C
  2. 5 min methanol @ 50 °C
  3. 2 min D.I. Water

• Standard lithography
  1. Spin on 1400-27 resist 40 sec @ 4000 rpm
  2. Bake 240 sec @ 90 °C
  3. 8 sec expose
  4. 240 chlorobenzene soak
  5. 50 sec develop H₂O:351 concentrate (5:1)
  6. 30 sec O₂ plasma (60 sccm, 100 mTorr, 100 W)

The exact processing sequence is as follows:

• Sample clean

• Define Emitter pattern
  1. Standard lithography

• Emitter evaporation/lift-off
  1. 60 sec H₂O:NH₄OH (20:1)
  2. Deposit metal
    - 200 Å Ni @ 5 Å/sec
    - 200 Å Ge @ 5 Å/sec
    - 500 Å Au @ 5 Å/sec
  3. Lift-off in hot acetone

• Sample clean

• Mesa etch
  1. 150 sec etch in H₂SO₄:H₂O₂:H₂O (1:8:160)

• Define Collector pattern
  1. Standard lithography

• Collector evaporation/lift-off
  1. 60 sec H₂O:NH₄OH (20:1)
  2. Deposit metal
    - 200 Å Ni @ 5 Å/sec
- 200 Å Ge @ 5 Å/sec
- 500 Å Au @ 5 Å/sec

3. Lift-off in hot acetone

- Sample clean
- SiNₓ insulating layer
  1. Deposit 1500 Å SiNₓ
     - He:500 sccm
     - NH₃:250 sccm
     - SiH₄:10 sccm
     - 30 min @ 500 mTorr, 40 W
- Define and etch Via pattern
  1. Standard lithography
- Sample clean
- Define bottom Pad and Bond layer
  1. Standard lithography without chlorobenzene soak
  2. Deposit metal
     - 100 Å Ti @ 5 Å/sec (sample at ~10 deg angle)
     - 500 Å Au @ 5 Å/sec (sample at ~10 deg angle)
     - Rotate sample 180 degrees
     - 500 Å Au @ 5 Å/sec (sample at ~10 deg angle)
  3. Standard lithography without chlorobenzene soak
- Electroplate sample
  1. Heat solution to 60 °C.
  2. 30 min @ 3 V with a 1 kΩ current limiting resistor
  3. Remove top layer of resist with flood expose and develop
  4. 25 sec Gold etch
  5. 25 sec Buffered Oxide Etch (to remove Ti)

3.3 Electrical Measurements

Electrical measurements on all two-terminal devices were carried out at from room temperature down to liquid helium temperatures (4.2 K) in different cryostat systems using a common computer controlled data acquisition setup, illustrated in Figure 3.2. For the three
Figure 3.1: Processing sequence for a two-terminal device.
terminal device measurements in Chapter 7, only a HP 4145B Semiconductor Parameter Analyzer was used.

3.3.1 Cryostat Systems

Two different cryostat systems, one for magnetic field and one for variable temperature, were used for all the measurements described in this work. In both systems, the sample was mounted on a nonmagnetic 24-pin header, with lead resistance from the top of the cryostat to the header less than 1Ω.

The 9 T Magnet System

Low temperature magnetic field measurements were done in a liquid helium cryostat, a double glass dewar system where the liquid helium dewar sits inside a liquid nitrogen dewar. A 9 T superconducting magnet with a 3/4" bore sits inside the liquid helium dewar. The sample sits inside a third stainless steel dewar, which fits inside the magnet bore, designed to condense liquid ³He. The sample is positioned to sit in the center of the magnet. The ³He feature was not used, so the stainless steel dewar was not sealed, allowing ⁴He to condense inside the ³He dewar, immersing the sample. In this situation, the sample temperature can be varied from 1.1 K to 4.2 K by varying the pressure over the sample.

The magnetic field was set by means of a voltage controlled current supply. The field (in the magnet center) was calculated by measuring the voltage drop across a 0.01 Ω sense resistor. The field was changed by changing the control voltage at a known rate for a fixed time. Maximum resolution for the field steps with the measurement setup used was 500 Gauss.

Janis Varitemp System

A Janis Research Supervaritemp cryostat was used for the variable temperature measurements. This system consists of two of stainless steel dewars, which are an outer liquid
Figure 3.2: Experimental setup for low temperature measurements. The sample sits in a cryostat connected to the electronics at room temperature. The dotted lines represent the actual current loop for the device.
nitrogen dewer, separated from an inner liquid helium dewer by a vacuum wall. The sample to be measured sits in a separate chamber surrounded by the helium dewar. The sample chamber is connected to the helium reservoir by a small capillary with a needle valve which is used to control the flow of liquid helium.

A continuous range from 1.2 K to 300 K is possible with this system using three different methods in three distinct temperature ranges. The temperature is measured using a LakeShore 330 Temperature Controller with a calibrated silicon diode. From 1.4 K to 4.2 K, the sample is immersed in liquid helium. The temperature is then varied by controlling the pressure in the sample chamber. This is done by means of a collapsible latex membrane in the pumping line surrounded by a constant pressure. Temperatures down to ~1.2 K (estimated from the sample chamber pressure) are obtainable, but the thermometer is only calibrated down to 1.4 K.

Temperatures from 4.2 K to ~80 K are obtained by means of a heater on the capillary which heated the helium which was then used to heat the sample. From there up to room temperature, the sample was simply allowed to warm as the liquid helium and nitrogen evaporated, or conversely allowed to cool as the liquid nitrogen dewar was filled. The minimum sample temperature obtainable in a reasonable amount of time during cool down using only liquid nitrogen was ~140 K. The rate was slow enough that for a 1 minute sweep, the temperature would vary by ~1 K or less. Typical warm up times from 80 K to 300 K (once the liquid nitrogen completely evaporated) were ~12 hours. Cool down times from 300 K to 140 K were ~18 hours.

3.3.2 Computer Controlled Data Acquisition Setup

Data acquisition was done using a computer controlled setup, shown in Figure 3.2, using a Macintosh computer with a National Instruments GPIB card as a controller. I-V sweeps were obtained from a 4-point measurement. The current was sensed using the circuit shown in Figure 3.3a. In this circuit, the input current is converted by the feedback resistor into a voltage which is then sensed by a DMM. The voltage across the device was sensed by an
Figure 3.3: Subcircuits used in the data acquisition setup. (a) Current ammeter: The output voltage is related to the input current by $V_I = -I/R$ with $V_I$ referenced to the voltage source of the setup. Different values of $R$ were placed on a rotary switch so the output voltage range could be adjusted. (b) Voltage amplifier: The voltage drop across the device was measured. A switch connecting pin 3 with either pin 11, 12 or 13 controlled the gain of the amplifier.
instrumentation amplifier (Figure 3.3b) with leads connected to the sample header. The output voltage is then sensed by a DMM. For both circuits, a supply voltage of ±18 V was generated using 9 V batteries.

The voltage source used was a Yokogawa 7651 Programmable DC Source capable of 10 μV steps and the DMMs used were Keithley 2001 Multimeters using 7 digit resolution for the 20 V range. I(V)s were taken by stepping the DC source to the desired voltage. When the output reach the programmed voltage, a signal would trigger the DMMs which would then read the two DMMs. Using this method, I(V)s could be acquired at about 3 points/second. The minimum voltage step used was 0.3 mV.

The current loop, indicated by the dotted line, had a series resistance of about 2 Ω for each cryostat. This, combined with the zero-load of the current sense circuit gave an essentially vertical load line for the setup.

The software could also sense and change the magnetic field as described in Section 3.3.1. The field was measured then compared to the desired field. Then, a constant voltage was then sent to an integrating amplifier for a set amount of time. The resistor voltage was measured using a HP 34401A Multimeter.

The controlling software was capable of either a single sweep or an automated series of sweeps with varying temperature or magnetic field. When using the varitemp cryostat, the software would sense the temperature from the LakeShore 330 controller then trigger a sweep when the designated temperature was reached. For magnetic field runs, the software could move and sense the field, then trigger a sweep after each field move.
Chapter 4

Single Well Resonant Interband Tunneling Structure

4.1 Introduction

A single well resonant interband tunneling structure was first demonstrated by Soderstrom et al. [51]. That study reported on a series of structures, consisting of InAs contacts, a GaSb well with AlSb barriers, exhibiting relatively large peak-to-valley (P/V) current ratios (20 @ 295 K, 88 @ 77 K). At the time this was the largest P/V ratio, which is a significant figure of merit, reported in a tunneling structure. Despite modifications of this basic structure, including polytype structures, where one contact is n-type InAs and the other p-type GaSb [65, 7, 6, 61, 20], the only significant improvement in P/V was achieved with the addition of AlAs to the AlSb [57]. The AlAs bandgap extends below the InAs valence band edge creating a tunneling barrier to holes leaking out of the GaSb well, which was shown to be the main cause of valley current [47]. Unfortunately, the lattice mismatch between AlAs and AlSb is so large (see Figure 2.1) that only a few monolayers can be grown before the lattice in no longer able to accommodate the strain. This fact limits the practical applications of InAs/AlSb/GaSb interband tunneling structures.

Initially, attempts to explain the operation of the structures took into account only the
<table>
<thead>
<tr>
<th>Material</th>
<th>$E_g$ (eV)</th>
<th>$m_e^*$</th>
<th>$m_{lh}^*$</th>
<th>$m_{hh}^*$</th>
<th>$a$ (nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>InAs</td>
<td>0.35</td>
<td>0.024</td>
<td>0.026</td>
<td>0.3</td>
<td>0.6058</td>
</tr>
<tr>
<td>GaSb</td>
<td>0.75</td>
<td>0.041</td>
<td>0.05</td>
<td>0.3</td>
<td>0.6096</td>
</tr>
<tr>
<td>AlSb</td>
<td>2.3 (d)</td>
<td>0.1</td>
<td>0.3</td>
<td>0.6136</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.6 (i)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.1: Important material parameters for the InAs/GaSb/AlSb system.

InAs conduction band, and the GaSb light-hole band [52, 66] on the basis that $k\cdot\mathbf{P}$ theory at $k = 0$, there is only coupling between the conduction band and the light-hole band [21]. More detailed treatment by Ting et al. demonstrated that not only can there be significant contributions to the transmission coefficient from the heavy-hole band for $k_i \neq 0$ [58], but that there can be significant contribution to the I(V) characteristics including additional transmission resonances [59]. Studies involving polytype interband tunneling structures have shown evidence supporting this idea [34, 23].

This chapter describes experiments on a homotype (InAs/AlSb/GaSb/AlSb/InAs) structures which show evidence of interband through multiple subbands at $k_i \neq 0$. The experimental data is superimposed on the approximate subband structure of the GaSb well showing good agreement with calculated results. In addition, the results show unexplained structure which implies an even more complex situation.

### 4.2 Device Structure

Measurements from two different samples will be discussed in this chapter. The samples are both single p-type well InAs/AlSb/GaSb interband tunneling structures. Bandstructures for both samples are shown in Figure 4.1. Material parameters for the InAs/AlSb/GaSb system are shown in Table 4.1.

The structures were grown in a Varian (Intevac) Gen II solid-source MBE system. The growth rates of GaSb, AlSb, and InAs were 1.0, 1.0, and 0.8 $\mu$m/h, respectively. Before growth the wafers were heated to 610 °C for 15 minutes to desorb the oxide. The first epitaxial layer, 0.2 $\mu$m of GaAs, was deposited at $T_{sub} = 580$ °C to smooth the
### Table 4.2: Material structure for the single well RIT samples shown in Figure 4.1. The barrier structures are summarized in Table 4.3.

<table>
<thead>
<tr>
<th>Material</th>
<th>Thickness</th>
<th>Doping</th>
</tr>
</thead>
<tbody>
<tr>
<td>n⁺-InAs</td>
<td>2500 Å</td>
<td>1×10^{18} cm⁻³</td>
</tr>
<tr>
<td>n-InAs</td>
<td>500 Å</td>
<td>2×10^{16} cm⁻³</td>
</tr>
<tr>
<td>InAs</td>
<td>100 Å</td>
<td>undoped</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Barrier Structure</td>
</tr>
<tr>
<td>InAs</td>
<td>100 Å</td>
<td>undoped</td>
</tr>
<tr>
<td>n-InAs</td>
<td>500 Å</td>
<td>2×10^{16} cm⁻³</td>
</tr>
<tr>
<td>n⁺-InAs</td>
<td>1 μm</td>
<td>1×10^{18} cm⁻³</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Substrate</td>
</tr>
</tbody>
</table>

### Table 4.3: Barrier structure for single well RIT samples summarized in Table 4.2. All layers are undoped.

<table>
<thead>
<tr>
<th>Sample #1</th>
<th>Sample #2</th>
</tr>
</thead>
<tbody>
<tr>
<td>25 Å AlSb</td>
<td>6 Å AlAs</td>
</tr>
<tr>
<td>65 Å GaSb</td>
<td>12 Å AlSb</td>
</tr>
<tr>
<td>15 Å AlSb</td>
<td>6 Å AlAs</td>
</tr>
<tr>
<td></td>
<td>65 Å GaSb</td>
</tr>
<tr>
<td></td>
<td>9 Å AlSb</td>
</tr>
<tr>
<td></td>
<td>6 Å AlAs</td>
</tr>
</tbody>
</table>

wafer surface. A 0.5 μm superlattice buffer region, consisting of a GaSb/AlSb superlattice followed by 0.4 μm of AlGaSb, was then grown at T_{sub} = 530 °C. The T_{sub} was then lowered to 500 °C for the growth of the active region of the structure, which consists of 1 μm n⁺-InAs (2 × 10^{18} cm⁻³), 50 nm of lightly doped InAs (2 × 10^{16} cm⁻³), the double barrier layer sequence, 50 nm of lightly doped InAs (2 × 10^{16} cm⁻³) and then a cap layer of 0.2 μm of n⁺-InAs (2 × 10^{18} cm⁻³).

The material structure for the two samples is described in Table 4.2. However, most of the experimental results come from Sample #1 and will only be from Sample #2 where explicitly noted.

For the following discussion, forward bias is defined as electrons initially tunneling through 1.5 nm barrier (larger peak current) and reverse bias as electrons going through 2.5 nm barrier (smaller peak current)
Figure 4.1: Schematic bandstructures for a) Sample #1, and b) Sample #2. The emitter corresponds to the top of the sample, and the collector to the substrate. The details of the barrier structure are given in Table 4.3.
4.3 GaSb Well Subband Structure

The GaSb well subband structure for both Sample #1 and Sample #2 are expected to be similar because both wells are nominally identical; only the barrier composition is different. The well subband structure can be estimated by using extrapolated data from structures with similar well thickness, published by Marquardt et. al. [26]. The estimated subband structure, shown in Figure 4.2 for a 6.5 nm well is obtained by linearly extrapolating the calculated structure from a 7 nm and an 8 nm well. The obtained curves are the solid lines.

The dashed curve represents the boundary $E$ vs $k_z$ dispersion for electrons in the emitter. The zero energy point is taken to be the bottom of the InAs conduction band. Electrons occupy those states from $k_z = 0$ out to the dashed line for energies up to the Fermi energy.

HH1 is the lowest index heavy-hole like subband, LH1 is the lowest indexed light-hole like subband. The dotted lines are fits, calculated using the bulk effective heavy-hole mass ($m^* = 0.3$) for HH1, and $m^* = 0.08$ for LH1 (for the bulk light-hole, $m^* = 0.05$). This value for LH1 gives the best approximation over the range of $k_z$ which overlap occupied $k_z$ states in the emitter.

4.4 Magnetotunneling

The current through a resonant tunneling structure (both intraband and interband) is determined to a large extent by the electronic subband structure of the two-dimensional well. Any modification of this structure will consequently modify the I(V) characteristics through the structure. One method of modifying the structure is by means of a magnetic field.

Magnetotunneling was first used by Mendez et al. [27] to probe an intraband resonant tunneling structure. A magnetic field applied perpendicularly to the confining interfaces of a two-dimensional structure ($B \parallel I$) quantized the allowed values of $k_z$ resulting in a
Figure 4.2: Parallel subband structure for the well structure extrapolated from published calculated data [26]. The dotted lines are parabolic fits. For GaSb, the lattice constant $a = 0.6096$ nm.
quantized density of electronic states. This quantization results in a series of peaks which vary with magnetic field appearing in the I(V) characteristics through a structure. When the peak positions are plotted against the applied magnetic field, a “fan” like structure, converging to the position of the state at zero field. Each branch of the fan corresponds to a single Landau level. Calculated I(V) characteristics by Goncalves da Silva and Mendez subsequently confirmed the interpretation of the experimental results [15]. Application of the field parallel to the confining interfaces ($B \perp J$) has no such quantizing effect. The incident carrier distribution on the structure, however, will be modified, allowing the $k_\parallel$ dispersion of the structure to be probed [16].

Magnetic field quantization of a GaSb/AlSb/InAs/AlSb/GaSb tunneling structure has been shown to result in a fan diagram similar to that of intraband tunneling structures [29]. Other studies have also used magnetic fields to probe similar structures with interband tunneling into a quantized InAs well [33, 35].

To date, very few magnetotransport studies involving interband tunneling structures with GaSb wells have been done. Oscillatory structure in the I(V) characteristics of a GaSb well structure with [36] and without [55] AlSb barriers has been observed, but no further analysis has been done. For a structure with interband tunneling occurring via a confined GaSb well, more than one band will be involved. Evidence for this has been reported for a polytype GaSb/AlSb/GaSb/AlSb/InAs structure [34] and for a homotype InAs/AlSb/GaSb/AlSb/InAs structure [26]. However, in both cases, the field was only applied parallel to the confining interfaces. In this chapter, magnetotunneling results with the field applied both parallel and perpendicular to the confining interfaces will be used to map out critical points in the $k_\parallel$ dispersion of the GaSb well.

Magnetotunneling data from Sample #1 is taken with field both parallel and perpendicular to current direction. Structure in the I(V) characteristics resulting from the magnetic field will appear as peaks, or shoulders in the data. Fan diagrams from the second derivative of the data are plotted versus the applied magnetic field. There is no physical significance to the second derivative in the sense that the first derivative represents the conductance.
but this method allows for more accurate determination of shoulders in the data.

4.5 Experimental Results

Current-voltage characteristics under different applied magnetic fields were measured using the experimental setup described in Chapter 3. Initially the characteristics with $B \parallel I$ were done, followed by the case with $B \perp I$. Only a limited number of devices were available, and those devices had problems with thermal cycling. Only one device of Sample #1 survived enough cycles to have measurements with both field orientations, which results will be described here. Other devices, however, did have similar I(V) characteristics. This device had a bond wire come off after the $B \parallel I$ measurements, requiring a different collector contact be used. This introduced an extra 70Ω of series resistance. This was subtracted out from all $B \perp I$ measurements to allow for comparison with the $B \parallel I$ data.

The I(V) characteristics for Sample #1 are shown in Figure 4.3. Except for the variable temperature measurements, all data was taken at 1.4 K. The device was a mesa structure with a diameter of 15 μm. In the forward bias direction ($T = 1.4$ K), the peak current is 2.9 mA occurring at 0.15 V, while for the reverse bias direction, the peak current is 0.65 mA occurring at 0.06 mA. The asymmetry is expected because of the asymmetry of the barriers, with the larger peak voltage occurring for the thinner barrier being the initial (and therefore controlling) barrier [40].

Measurements were taken using a four-point method, to eliminate any effects from series resistance in the leads. This, however, did not eliminate any resistive effects from either the bond wires or contacts themselves. Thus any measurement will actually include the combination of the tunneling characteristics of the structure and the series resistance. If the series resistance is larger than the structure resistance, it will dominate and the current will be controlled by the value of the resistance.

For no applied magnetic field, this does not appear to be the case, with the device having a minimum dynamic resistance of about 36 Ω in the reverse direction and 26 Ω in
Figure 4.3: I(V) characteristics of Sample #1 at 300 K and 1.4 K. The oscillation regions in the 1.4 K curve are not shown.

The forward direction. With increasing magnetic field (for \( B \parallel I \)), the minimum dynamic resistance increases with field (see following discussion). A magnetic field will cause an increase in resistivity according to

\[
R = R_o + \kappa B^2
\]  

(4.1)

where \( R_o \) is the zero-field resistance, \( B \) is the applied field, and \( \kappa \) is a constant dependent on the specific material. For both forward and reverse bias with \( B \parallel I \) a magnetoresistive term with \( \kappa = 1.14 \Omega/T^2 \) was subtracted. This value was empirically fit to the peak current, which at high fields was limited by the series resistance component.

### 4.5.1 B parallel to I

**Reverse bias**

The reverse bias I(V) characteristics at various values of \( B \parallel I \) for Sample #1 are shown in Figure 4.4. As the applied magnetic field is increases, structure appears, increasing in strength with field. The peak current also increases over the zero-field case, reaching similar
values at both 6 T and 9 T, but moving to a larger voltage. This behavior is consistent with a series resistance component dominating the dynamic resistance and is the justification for subtracting out a magnetoresistive component as described above. The I(V) characteristics, corrected for magnetoresistance are shown in Figure 4.5.

A fan diagram generated from the I(V) data after correction for magnetoresistance is shown in Figure 4.6. Each point in the diagram is a positive peak in the second derivative of each I(V) trace, which for negative values of current correspond to either a peak or a shoulder in the current magnitude. The voltage region between -0.07 V and -0.15 V corresponds to the oscillation region for all traces, so no points are shown.

The fan diagram has structure that is immediately apparent. Based on the model discussed in Chapter 2, the fan is expected to move to lower bias with increasing magnetic field, converge to a single point at zero field, and to spacing between points to increase linearly with increasing field. Using these assumptions, two distinct fans can be drawn, shown in Figure 4.7. The first converges to -0.141 V with slopes given by \( m \times 0.065 \text{ V/T} \) and the second converges to -0.036 V with slopes given by \( m \times 0.01 \text{ V/T} \) where \( m \) is an integer.

For the fan converging to -0.141 V, the straight line fits are in excellent agreement with the data, particularly for \( m > 4 \). For lower values of \( m \), oscillations occur in the data, but the trend follows the straight line fits. Similar oscillations were seen previously and attributed to field induced charge transfer from the well to the contacts [29]. For \( m < 7 \), there are extra lines between each of the straight line fits. In general, strength of these features are smaller than the lines corresponding to the fits. This, in addition to the excellent agreement with the straight line fits, however, suggests the “extra” structure results from some related, but secondary of effect. Also, the agreement of the fits implies the validity of the subtraction of the magnetoresistive component.

There is a third fan that can be drawn that moves to higher bias with increasing field. In this case, the structure occurs after the main NDR, converging to 0.133 V with slopes given by \( m \times -0.095 \text{ V/T} \). This fan is shown in Figure 4.8. The structure responsible for this fan is
approximately an order of magnitude smaller than that responsible for the other fans once the current magnitudes are taken into account. The fact that the structure occurs after the interband current is shut off (i.e. post resonance) indicates it results from something not connected with interband tunneling. In addition, the fact that the structure moves to a larger bias, which for reverse bias means a more negative voltage, with increasing field requires that the carrier source and confined 2DEG are associated with the same band.

This fan is consistent with holes back tunneling from the collector to the emitter via Landau levels from a light-hole band. The accumulation at the emitter and collector contacts create triangular shaped barriers in the valence band which create a double barrier structure (see Figure 4.1. This structure also occurs in Sample #2, which has AlAs barriers extending into the InAs valence band. Bandprof places the $n = 2$ GaSb light-hole state 0.15 eV above the top of the InAs valence band. Because the holes are injected from the top of the InAs valence band, the state is at a positive energy relative to the injection energy. For negative applied voltages (reverse bias), the fan structure should converge to a positive voltage, which is the case.

**Forward bias**

The forward bias $I(V)$ characteristics at various values of $B \parallel I$ for Sample #1 are shown in Figure 4.9. As with the reverse bias case, with increasing field, structure appears, increasing in strength with field. The peak current does not increase over the zero-field case, but does move to a larger voltage with increasing field. This behavior is still consistent with a series resistance component increasing with field and thus, justification for subtracting out a magnetoresistive component. The $I(V)$ characteristics, using the same correction factor as the reverse bias case are shown in Figure 4.10.

A fan diagram generated from the $I(V)$ data after correction for magnetoresistance is shown in Figure 4.11. Each point in the diagram is a negative peak in the second derivative of each $I(V)$ trace, which for positive values of current correspond to either a peak or a shoulder in the current magnitude. The voltage region corresponding to the oscillation
Figure 4.4: Reverse bias I(V) characteristics of Sample #1 with $B \parallel I$ before correction for magnetoresistance.

Figure 4.5: Reverse bias I(V) characteristics of Sample #1 with $B \parallel I$ after correction for magnetoresistance.
Figure 4.6: Fan diagram generated from peaks in the second derivative of the I(V) characteristics of Sample #1 with B||I. Each point is a positive peak in the second derivative, corresponding to a current peak or shoulder. A positive peak corresponds to a point of maximum positive curvature in the I(V) characteristics, indicating a current peak (i.e. a point of more negative current) or shoulder. Each I(V) was measured with 0.5 mV resolution. The field resolution was 0.05 T.
Figure 4.7: Fan diagram from Figure 4.6. The two sets of dashed lines are fit to the structure. The fan converging to -0.036 V corresponds to structure from HH1, while the fan converging to -0.141 V corresponds to structure from LH1. The slopes are given by \( nm \) where \( n \) is an integer and \( m = 0.01 \) for HH1, and \( m = 0.065 \) for LH1.
Figure 4.8: Fan diagram from Figure 4.6 with a fan, converging to 0.133 V with slopes
given by $nm$ where $n$ is an integer and $m = -0.095$. This fan results from holes in the
valence band resonantly tunneling from the collector to the emitter through a light-hole
state.
region for all traces are not shown.

A single fan, converging to 0.236 V with slopes given by $m \times -0.072$ V/T is shown. As with the reverse bias case, for $m < 4$ "extra" branches in the fan occur. The agreement with the straight line fits, however, along with their absence for larger values of $m$ suggest a secondary effect. Also, the agreement with the fits confirms the validity of the subtraction of the magnetoresistive component. Structure moving to higher bias with increasing field is present at biases from 0 V to about 0.08 V, however, this will be shown to apparently not be related to Landau quantization in the well. Structure occurring after the NDR either is not present, or cannot be resolved.

4.5.2 B perpendicular to I

Before measurements with $B \perp I$ were made, the original collector bonding wire failed. Another “good” contact existed, but it had an extra series resistance. Using a value of 70 $\Omega$ for the reverse bias direction and 58 $\Omega$ for the forward bias direction, the zero-field $I(V)$ characteristics could be made to match previous measurements. The fan diagrams for both bias directions were made after correction for this resistance. Because of this series resistance, the magnetoresistance present in the $B \parallel I$ was not as significant, and was not taken into account. This is reasonable since in this case the quantitative data extracted comes from a low current region where the structure has a larger dynamic resistance.

Reverse bias

The reverse bias $I(V)$ characteristics with $B \perp I$ are shown in Figure 4.12. The data, after correction for the extra series resistance are shown in Figure 4.13. A fan diagram generated from this data is shown in Figure 4.14, points from both negative and positive curvatures are shown. For negative currents, the positive peaks correspond to current magnitude peaks and shoulders, while the negative peaks correspond to current magnitude minima.

The fan diagram shows two distinct features which will be discussed. The first is a current minima at -0.013 V monotonically decreases to lower bias with increasing magnetic
Figure 4.9: Forward bias I(V) characteristics of Sample #1 with B∥I before correction for magnetoresistance.

Figure 4.10: Forward bias I(V) characteristics of Sample #1 with B∥I after correction for magnetoresistance.
Figure 4.11: Fan diagram generated from peaks in the second derivative of the I(V) characteristics of Sample #1 with B||I. Each I(V) was measured with 0.5 mV resolution. The field resolution was 0.05 T. Each point is a negative peak in the second derivative, corresponding to a current peak (i.e., a local maximum) or shoulder. The fan converges to 0.236 V with slopes given by \(nm\) where \(n\) is an integer and \(m = -0.072\). This fan results from holes in the valence band resonantly tunneling from the collector to the emitter through a light-hole state. Additional structure in the bias range up to 0.05 V moving to increasing bias with increasing field can be seen.
field. This initial minima indicates a current threshold occurring at the initial current increase. The positive peak at -0.002 V is related to the experimental setup. The second is related to the positive peak which occurs at 0.058 V at zero field. This is the main current peak occurring just before the main NDR. Starting about 5 T, there is a distinct change in the behavior of this peak. The single peak splits into two, with one, which is still the peak just before the main NDR, sharply increasing in bias with increasing field while a second shifts to lower bias, reaching 0.02 V at 9 T. This peak can be seen in the I(V) characteristics of Figure 4.13. This behavior is similar to that seen by Marquardt et al. [26].

Forward bias

The forward bias I(V) characteristics with $B \perp I$ are shown in Figure 4.15. The data, after correction for the extra series resistance are shown in Figure 4.16. A fan diagram generated from this data is shown in Figure 4.17, only the negative peaks are shown. For positive currents, the negative peaks correspond to current magnitude peaks and shoulders.

From 0.06 V, the fan diagram is noisy, and no structure can be seen. Below this voltage, however, fan-like structure, similar to the forward bias direction with $B \parallel I$ can be seen. A monotonically decreasing with field minima (not shown), similar to that seen in the reverse bias direction can be seen. This minima approaches zero bias at the same field as the reverse bias case, indicating they are the same feature scaled to different voltages resulting from the asymmetric tunneling structure.

4.6 Discussion

Experimentally Determined Critical Points

Using the simple interband tunneling model developed in Chapter 2, a number of critical points on the GaSb subband structure can be determined. The points are shown, superimposed on the subband structure in Figure 4.18. The size of the markers indicating these points are consistent with the experimental error of ± 2 mV (data points were 0.5 mV and
Figure 4.12: Reverse bias I(V) characteristics of Sample #1 with B \perp I before correction for series resistance.

Figure 4.13: Reverse bias I(V) characteristics of Sample #1 with B \perp I after correction for series resistance.
Figure 4.14: Fan diagram generated from peaks in the second derivative of the $I(V)$ characteristics of Sample #1 with B.I.I. Both positive and negative peaks in the second derivative are shown. Each $I(V)$ was measured with 0.5 mV resolution. The field resolution was 0.05 T.
Figure 4.15: Forward bias I(V) characteristics of Sample #1 with \( B \perp I \) before correction for series resistance.

Figure 4.16: Forward bias I(V) characteristics of Sample #1 with \( B \perp I \) after correction for series resistance.
Figure 4.17: Fan diagram generated from negative peaks in the second derivative of the I(V) characteristics of Sample #1 with B.I. The scale is the same as in Figure 4.11. The dense region of points between 0.06 V and 0.1 V results from the subtraction of the magnetoresistive term. Each I(V) was measured with 0.5 mV resolution. The field resolution was 0.05 T.
a smoothing window of 9 points was used prior to peak finding).

As the allowed values of \( k_\parallel \) (and therefore, the density of states), for the different subbands are quantized by the application of \( B \parallel I \), a series of fluctuations in the I(V) characteristics occur, resulting in a fan like structure, with the fan converging to a single point at zero-field. Figure 4.7 exhibits two distinct fans, one converging to -0.141 V and the other to -0.036 V. This discussion will make the initial assumption that the Fermi level lies below the intersection of HH1 and the boundary of the occupied \( k_\parallel \) states in the emitter. This assumption will later be shown to be consistent with the experimental data.

Given the position of the Fermi level, the two fans arise from two different mechanisms. Each the zero-field point of each fan results from a critical transition. The zero-field point at -0.036 V occurs when the \( k_\parallel = 0 \) point of HH1 crosses the Fermi energy. The zero-field point at -0.141 V occurs when the \( k_\parallel = 0 \) point for LH1 crosses the bottom of the InAs conduction band. While theoretically there should also be a fan resulting from HH1 crossing the bottom of the InAs conduction band, it will be much weaker due to the relative transmission coefficient magnitudes between HH1 and LH1 [59] and will therefore not be resolvable.

The voltage of the LH1 zero-field point can be related to the subband energy by a conversion factor \( \alpha \) with units eV/V. This energy corresponds to Point #1 on Figure 4.18. The zero-field voltage of this point is -0.141 V, and the estimated value of the zone center of LH1 is 0.053 V, giving \( \alpha = 0.38 \).

Application of a magnetic field parallel to the confined well (\( B \perp I \)) results in a shift in \( k_\parallel \) for the emitter electrons according to (2.35). This shift is illustrated by the thin dotted curve in Figure 4.18. Given the position of the Fermi energy, at low biases, emitter electrons cannot tunnel into the unoccupied well states because of conservation of momentum [28]. This condition can occur by either shifting the emitter electron distribution in energy with an applied bias, or shifting the distribution with an applied magnetic field.

Once this occurs, by either method, or a combination of both, there will be a turn-on threshold in the I(V) characteristics as the new conduction channel is opened. This
Figure 4.18: Estimated parallel subband structure with experimentally determined critical points. The heavy dashed curve indicates the extent of occupied states in the emitter. The light dashed curve indicates the offset obtained by the application of a magnetic field parallel to the confining interfaces such that the current threshold moves to zero bias. For GaSb, the lattice constant $a = 0.6096$ nm.
behavior can be seen in Figure 4.14 with the negative curvature peak that starts at -0.013 V and moves to zero bias at 3.5 T. Given the emitter electron distribution and the curvature of HH1, this can be used to estimate the Fermi level by noting the value of magnetic field where the threshold goes to zero bias.

The quantity $\Delta l$ in (2.35) is the average distance traveled by an electron while it is tunneling from the emitter to the well. The final position is assumed to be the center of the well, and the initial position is taken to be the position of the maximum of the electron accumulation layer in the emitter [16]. This is estimated to be 5 nm using Bandprof. giving $\Delta l = 5 + 2.5 + 3.3 = 10.8$ nm. This gives the shift of $\Delta k_{||} = 0.46 \% (\pi/a)$ indicated in Figure 4.18 and allows the placement of the Fermi level at $E_F = 0.112$ eV, indicated by Point #2.

Given the determined Fermi energy the zero-field threshold should occur at an energy shift of 0.005 eV, indicated by Point #3. Using $\alpha = 0.38$ the zero-field threshold should occur at 0.013 eV, which is where it occurs in Figure 4.14. The zero-field point of the HH1 fan (Figure 4.7 corresponds to the top of HH1 crossing the Fermi level. This occurs at -0.036 V. Again, using $\alpha = 0.38$, this gives the top of HH1 to be at 0.126 eV, indicated by Point #4.

The preceding analysis can be repeated with the forward bias case for all four points. Using a value of $\alpha = 0.22$, all four points fall in the same location. The fan used in determining Point #4 is not present, presumably washed out by the unexplained structure. However, a peak does appear at zero-field at the expected voltage 0.061 mV.

**Modeled Fan Diagram**

The expected 2nd derivative fans for the reverse bias case of $B || I$ were modeled using the derived band parameters ($E_F$ and $m^*$). The resulting fan is shown in Figure 4.19. There is good agreement with the calculated fan for HH1 with fan fit to the data, especially at field values greater than 4 T. The expected contribution from LH1, however has fewer traces (by about a factor of 2.5) than the calculated fans. This also suggests a connection with
the “extra” structure in the fan compared to the straight line fits.

The most likely reason for this is the fact that the three dimensional emitter density of states will also be modified, resulting from the quantization of $k_z$ as described in Chapter 2. This extra modulation in the I(V) characteristics would result in “extra” branches in the fan. It would not occur in the HH1 fan, since that structure results from well states crossing the Fermi level. Since only the carriers at the Fermi surface are involved, no extra modulation, and therefore structure will occur.

Unexplained Structure

In a resonant tunneling structure when the emitter and confined potential are in band of opposite polarity (i.e. conduction and valence) a magnetic field, applied perpendicular to a two-dimensionally confined region (i.e. the $z$-direction), splits the quantized states, resulting in a series of converging peaks that move to lower applies bias with increasing magnetic field. The same type of behavior will also result from quantizing the motion in the $x$-$y$ plane of the three-dimensional electrons. This behavior does not occur in the two-dimensional layer when the field is parallel to the confining layers.

In the forward bias fan diagrams with the field both parallel and perpendicular to the direction of current flow (Figures 4.11 and 4.17) there is converging structure which moves to higher bias. The structure in both field orientations is similar, both in position and slope. This structure is different than the post-resonant structure visible in the $B \parallel I$ reverse bias fan, which was attributed to back tunneling of holes in the valence band. This structure is not the same as the fan in the reverse bias $B \parallel I$ case because it occurs before resonance.

While the origin of this structure is not known, it is not sample specific. Magneto-transport measurements were made on Sample #2 with $B \parallel I$. The I(V) characteristics for three different fields are shown in Figure 4.20. These are qualitatively similar the I(V) characteristics of Sample #1 shown in Figure 4.4. This is expected because the total barrier thickness are essentially the same, and the well is the same width. A fan diagram of
Figure 4.19: Fan diagram generated from the second derivative of the I(V) characteristics obtained using the interband tunneling model of Chapter 2. The I(V) characteristics were modeled using HH1 and LH1 assuming the parabolic fits to the estimated subband structure. The dashed lines are the fans fit to the experimentally measured structure.
from peaks in the 2nd derivative were generated using the same procedure as previously discussed as is shown in Figure 4.21. No correction for magnetoresistance was done.

The structure in the fan diagram is nearly identical to that of Sample #1, shown in Figure 4.6. There appears to be less structure for B<3 T. This, however, is a result of the reduced resolution of the data (1 mV steps for the I(V) traces every 0.2 T). Much of the structure in Sample #1 only appears because of the increased resolution. For Sample #2, the three sets of explained structure, resulting from the lowest index heavy- and light-hole like bands and the back tunneling of valence band holes are easily apparent, along with the unexplained structure.

**Zero-bias Conductance**

An effect analogous to the Shubnikov-de Hass oscillations can be seen in the zero-bias conductance as a function of magnetic field. Shubnikov-de Hass oscillations occur in the in plane conductance of a two dimensional electron gas (2DEG) where the conductance goes through peaks and valleys corresponding to the Fermi level lying either in, or in between Landau levels. The effect only occurs for the case where the field is perpendicular to the plane of the 2DEG. When the field is parallel to the plane, Landau quantization does not occur, and the oscillatory conductance does not occur. This is essentially the same reason the fan diagrams occurred in the previous discussion for $B \parallel I$ but not for $B \perp I$.

Each Landau level in a 2DEG has the same degeneracy given by $qB/h$ where $q$ is the electron charge, and $h$ is Planck’s constant. One consequence of this is the fact that zero-bias conductance is periodic with $1/B$. Since the degeneracy is known, the carrier density can be determined by

$$N_s = (q/h) / \Delta(1/B)$$

(4.2)

where $\Delta(1/B)$ is the period of the conductance plotted versus $1/B$.

The zero-bias conductance characteristics for $B \parallel I$ plotted versus $1/B$ for Sample #1 are shown in Figure 4.22. The FFT spectrum of the data is shown in Figure 4.23. The
FFT spectrum has a single strong peak and significantly smaller peaks with increasing index. When the procedure for the conductance is repeated for different applied biases, the smaller peaks track identically with the strong peak, and they are equally spaced. This implies the smaller peaks are simply harmonics of the larger peak, resulting from the finite nature of the FFT.

For the periodic nature of the zero-bias conductance to be indicative of a two dimensional system, the oscillations should be dependent on the direction of the magnetic field. The data for $B \perp I$ plotted versus $B$ are shown in Figure 4.24. This clearly does not show oscillatory behavior, though it is not without some structure.

The previous discussion showed that more than one subband was involved in the current flow for the structure. However, only one strong peak is seen in Figure 4.23, implying only one subband is involved in the zero-bias conductance. At zero bias, only those electrons lying on the Fermi surface are available for transport and thus, only one subband is involved (see Figure 4.2).

Using the data plotted in Figure 4.22, 4.2, gives a carrier density of $6.3 \times 10^{11} cm^{-2}$ or, assuming a uniform carrier density in the well, about $1 \times 10^{18} cm^{-3}$. A self-consistent electrostatic calculation from Bandprof give a density of about $1.5 \times 10^{18} cm^{-3}$ for the well.

The case of $B \perp I$, shown in Figure 4.24 is as expected given the assumed electronic subband structure. As the field is increased, the electron distribution in the emitter receives an additional component of $k_{\parallel}$. As this approaches the occupied value of $k_{\parallel}$, the conductance is expected to increase. Once this value is exceeded, the conductance will peak, then decrease as an increasing fraction of the carriers in the emitter have excess $k_{\parallel}$ and are not able to participate in tunneling. This type of behavior has been observed previously in a GaSb/AlSb/InAs/AlSb/GaSb structure [28].
4.7 Variable Temperature Measurements

The main contribution to the valley current of InAs/AlSb/GaSb/AlSb/InAs interband tunneling structures has recently been shown to be related to holes escaping from the GaSb well [47]. Though the cause at all temperatures is hole current, the mechanisms at high and low temperature are different. Holes in the well are confined by a barrier on order of the bandgap of InAs. As a bias is placed across the structure, the barrier takes on a triangular shape as a result of band bending in the emitter.

For high temperatures (T > 250 K), thermal energy causes an appreciable population of higher energy holes (i.e. the energy distribution of holes extends down toward the top of the InAs conduction band. These holes can then be swept out of the well, causing appreciable valley current. Since this is a thermal process, an exponential dependence on temperature is expected.

For lower temperatures (T < 100 K), the valley current becomes essentially independent of temperature, but is expected to increase with increasing bias voltage. This is because as the bias increases, band bending in the emitter increases, effectively narrowing the barrier seen by the confined holes. Once the barrier is thin enough, holes escape the well by direct Fowler-Nordheim tunneling through the barrier.

This allows estimates of the Fermi level and the amount of band bending in the emitter by analysis of the activation energy at high temperature, and the magnitude of the valley current versus bias at low temperatures, respectively. Unfortunately, the samples analyzed in this chapter did not thermally cycle an indefinite number of times. Sample #2 failed completely, and Sample #1 developed a parallel conduction path, as shown in Figure 4.25. Because the exact magnitude of the current contribution from the path could not be determined, quantitative data could not be extracted. However, the general qualitative trend could be seen, shown in Figure 4.26.
4.8 Summary

A single GaSb well interband tunneling structure has been studied using magnetotransport to probe the subband structure of the well. Critical points in the results were extracted then compared to the approximate calculated electronic subband dispersion. By selecting the appropriate alpha (meV/mV) good agreement with the calculated results is achieved. This is a good validation of the fact that multiple subbands at $k_\parallel \neq 0$ are involved in the tunneling process. Comparison to the calculated fan diagrams imply some of the structure in the fan diagram results from modulation of the three dimensional density of states in the emitter. There is similar unexplained structure in the forward bias fan diagrams of both field orientations. The fact that the structure occurs in both orientations indicates that it is not a result of Landau quantization of the two dimensional well.
Figure 4.20: Reverse bias I(V) characteristics of Sample #2 with B||I.
Figure 4.21: Fan diagram generated from peaks in the second derivative of the I(V) characteristics of Sample #2 with B||I. The fan diagram was generated using the same procedure as was done with Figure 4.6. Each I(V) was measured with 1.0 mV resolution. The field resolution was 0.2 T.
Figure 4.22: Zero-bias conductance vs. inverse magnetic field for Sample #1. The field is applied parallel to the current direction.

Figure 4.23: FFT spectrum for the curve shown in Figure 4.22. The large peak centered at index=14 corresponds to the main frequency component of the curve. The smaller peaks to the right are harmonics of the main peak.
Figure 4.24: Zero-bias conductance vs. magnetic field for Sample #1. The field is applied perpendicular to the current direction.

Figure 4.25: I(V) characteristic of Sample #1 used for variable temperature measurements. The large valley current in both the forward and reverse bias directions (compare with Figures 4.5 and 4.10) is the result of a parallel conduction path that developed in the device after numerous thermal cyclings.
Figure 4.26: $I$ vs $1000/T$ for applied bias of $-0.2\,\text{V}$, $-0.3\,\text{V}$ and $-0.4\,\text{V}$. 
Chapter 5

Multiple Well Resonant Interband Tunneling Structures

5.1 Introduction

By combining multiple devices, to give a multiple-peaked I-V, functionality can be increased even further [39, 45, 50]. “Conventional” RTDs usually consist of a double barrier, single quantum well semiconductor heterostructure where the center region has quasi-bound quantum well states accessible to electron (or hole) transport [5]. RIT devices differ from conventional RTDs in that the confined states, accessible to electron (or hole) transport lie in the valence (or conduction) band rather than the conduction band [24, 51]. The peaked nature of the I-V characteristics of both RTD and RIT structures is potentially useful for increased functionality of electronic devices [3, 22, 45, 46, 50]. Because of the difference in origin of the NDR, it is possible to achieve a multiple-peaked I-V characteristic using both effects in the same device.

Here, a series of structures in which a coupled double well exists in the valence band and a single well exists in the conduction band is studied. Tunneling through both of these regions was observed. These types of structures may have applications in three terminal operation of interband tunneling structures because the InAs well confined state
is separated from the interband states both spatially and in energy. Therefore, it will be possible to vary the potential of the interband tunneling structure by adding charge to the well state.

5.2 Device Structure

The structures were grown in a Varian (Intevac) Gen II solid-source MBE system. The growth rates of GaSb, AlSb, and InAs were 1.0, 1.0, and 0.8 μm/h, respectively. Before growth the wafers were heated to 610 °C for 15 minutes to desorb the oxide. The first epitaxial layer, 0.2 μm of GaAs, was deposited at T_{sub} = 580 °C to smooth the wafer surface. A 0.5 μm superlattice buffer region, consisting of a GaSb/AlSb superlattice followed by 0.4 μm of AlGaSb, was then grown at T_{sub} = 530 °C. The T_{sub} was then lowered to 500 °C for the growth of the active region of the structure, which consists of 1 μm n⁺-InAs (2 × 10^{18} cm\(^{-3}\)), 50 nm of lightly doped InAs (2 × 10^{16} cm\(^{-3}\)), the double barrier layer sequence, 50 nm of lightly doped InAs (2 × 10^{16} cm\(^{-3}\)) and then a cap layer of 0.2 μm of n⁺-InAs (2 × 10^{18} cm\(^{-3}\)). Growth conditions were such that the InAs/GaSb interfaces were InSb like.

Three similar samples, summarized in Table 5.1 designed to show both intraband and interband tunneling were investigated. The samples were nominally identical except for the barrier structure, which varied by the placement of an AlSb barrier as listed in Table 5.2. Schematic band diagrams for the structures are shown in Figure 5.1.

At low biases, because of the broken bandgap at the InAs/GaSb interface, the structures are RTIs with tunneling occurring through a coupled double well structure in the GaSb valence band. The GaSb layer thickness was chosen such that only one light hole quantum well state exists in the GaSb valence band above the InAs conduction band edge [67]. The InAs well thickness was chosen such that the lowest energy quantum well state would be above the GaSb valence band edge. Thus, at higher biases, the structures are conventional RTDs, with tunneling occurring through a single InAs quantum well state.
Figure 5.1: Band diagrams for the inter/intra band tunneling samples. Transmission coefficients are shown in Figure 5.2.
Figure 5.2: Calculated two-band transmission coefficients for the inter/intra band tunneling samples. Transmission coefficients are shown in Figure 5.1.
<table>
<thead>
<tr>
<th>Material</th>
<th>Thickness</th>
<th>Doping</th>
</tr>
</thead>
<tbody>
<tr>
<td>n⁺-InAs</td>
<td>2500 Å</td>
<td>$1 \times 10^{18}$ cm⁻³</td>
</tr>
<tr>
<td>n-InAs</td>
<td>500 Å</td>
<td>$2 \times 10^{16}$ cm⁻³</td>
</tr>
<tr>
<td>InAs</td>
<td>100 Å</td>
<td>undoped</td>
</tr>
<tr>
<td><strong>Barrier Structure</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>InAs</td>
<td>100 Å</td>
<td>undoped</td>
</tr>
<tr>
<td>n-InAs</td>
<td>500 Å</td>
<td>$2 \times 10^{16}$ cm⁻³</td>
</tr>
<tr>
<td>n⁺-InAs</td>
<td>1 μm</td>
<td>$1 \times 10^{18}$ cm⁻³</td>
</tr>
<tr>
<td><strong>Substrate</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5.1: Material structure for the inter/intra band tunneling samples shown in Figure 5.1. The barrier structures are summarized in Table 5.2.

<table>
<thead>
<tr>
<th>$MWRIT1$</th>
<th>$MWRIT2$</th>
<th>$MWRIT3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>95 Å GaSb</td>
<td>25 Å AlSb</td>
<td>95 Å GaSb</td>
</tr>
<tr>
<td>55 Å InAs</td>
<td>95 Å GaSb</td>
<td>25 Å AlSb</td>
</tr>
<tr>
<td>95 Å GaSb</td>
<td>55 Å InAs</td>
<td>55 Å InAs</td>
</tr>
<tr>
<td></td>
<td>95 Å GaSb</td>
<td>95 Å GaSb</td>
</tr>
</tbody>
</table>

Table 5.2: Barrier structure for the inter/intra band tunneling samples shown in Figure 5.1. All layers are undoped.

The flatband transmission coefficient for each structure was calculated using the two-band model described in Chapter 2. The coefficient is similar for each structure with a broad, double peak at low energies corresponding to transmission through the coupled GaSb valence band wells, and a more narrow single peak corresponding to transmission through the InAs well conduction band state. The effect of the AlSb layer in two of the samples can be seen by comparing the magnitudes of the corresponding transmission peaks for each device. The condition of the transmission coefficient being unity for the symmetric sample (#1) can be viewed as a mathematical artifact. Any distribution of electrons incident on the tunneling structure will have some finite energy width which will tend to smear out the transmission coefficient. For the case where inelastic scattering is allowed, the resonance occurs at the peak transmission coefficient, but the peak current at resonance is related only to the transmission coefficient of the first barrier seen by the tunneling electrons [63], thus both the interband and intraband peak currents for Samples #2 and #3 will be different for different bias directions because of the difference in thickness of the tunnel barriers.
Figure 5.3: I-V characteristics of MWRIT1. The I-V is symmetrical about zero bias, therefore only one bias direction is shown.

5.3 Experimental Results

MWRIT1

The I-V characteristics for MWRIT1, measured at 300 K and 4 K are shown in Figure 5.3. The I-V is symmetrical about zero bias, therefore only one bias direction is shown. At 300 K, a single resonance at 0.41 V is visible. At 4 K, however, the resonance shifts to 0.35 V, and three other strong resonances at 0.02 V, 0.06 V and 0.70 V become visible. Three of the resonances show no hysteresis when the bias is swept up from, then back down to 0 V. The resonance at 0.70 V, however, shows significant hysteresis between the two
sweep directions. Other than a small dip at 0.16 V, no additional structure can be seen in the conductance.

The two resonances at low bias are interband resonances where the lowest bias resonance (0.02 V) results from the GaSb well farthest from the emitter, and the next highest resonance (0.06 V) results from the near well. The resonance at 0.35 V is the intraband resonance resulting from the InAs well.

**MWRIT2**

The reverse and forward bias I-V characteristics for MWRIT2, measured at 300 K and 4 K are shown in Figure 5.4 and Figure 5.5, respectively. Reverse (negative) bias voltages correspond to the case where electrons initially tunnel through the AlSb barrier.

At 300 K, only a single resonance can be seen in the reverse bias direction at -0.06 V, and in the forward bias direction at 0.90 V. At 4 K, additional resonances appear. In the reverse bias direction, there are resonances at -0.07 V, -0.66 V and -0.90 V. In the forward bias direction, resonances occur at 0.11 V, 0.18 V, 0.68 V and 0.89 V.

The intraband resonances for the structure appear at -0.90 V and 0.68 V (there are also small features, at -0.96 V and 0.89 V slightly above each of the intraband resonances). The interband resonances into the near GaSb valence band well appear at -0.07 V and 0.18 V. Additional “extra” features, visible only in the conductance, occur below the large interband tunneling resonance. For the negative bias interband resonance, the peak-to-valley ratio is higher and the peak current smaller than for the forward bias resonance. This is expected from, and consistent with, previous observations for other InAs/AlSb/GaSb tunneling structures [8] since electrons initially tunnel through the AlSb barrier, which serves as a real tunnel barrier, reducing inelastically scattered valley current. The strengths of the extra features are also smaller than for the forward bias direction. The strengths of the “extra” features are also smaller than for the forward bias direction. The position and peak current of the intraband resonances are consistent with the asymmetric barriers created by the addition of the AlSb layer. The peak voltage is increased and the peak
Figure 5.4: Reverse bias I-V characteristics of MWRIT2
Figure 5.5: Forward bias I-V characteristics of MWRT2
current reduced for the bias direction where electrons initially tunnel into the thicker barrier (29 mA at 0.68 V, 13 mA at -0.90 V).

**MWRT3**

The reverse and forward bias I-V characteristics for MWRT3, measured at 300 K and 4 K are shown in Figure 5.6 and Figure 5.7, respectively. Reverse (negative) bias voltages correspond to the case where electrons initially tunnel through the AlSb containing barrier (electrons move left to right in Figure 5.1).

In the reverse bias direction, the intraband resonance occurs at -0.70 V. Weaker structures occur at -0.21 V, -0.46 V, and -0.96 V, and very weak structure, visible only in the
Figure 5.7: Forward bias I-V characteristics of MWRIT3
conductance, occur at -0.17 V, -0.28 V, and -0.59 V. In the forward bias direction, strong NDR regions occur at 0.04 V and 0.18 V. These two features are qualitatively similar, each with a small peak at a slightly higher bias (0.13 V and 0.23 V). Another small feature of similar strength is located at 0.31 V. A strong NDR region occurs at 0.36 V, also with a weaker feature at the slightly larger bias (0.47 V). A number of features appear at reverse biases greater than 0.50 V. The forward bias intraband resonance is not shown in Figure 5.7. It will occur at a voltage larger than the reverse bias direction because of the lever arm effect. Devices usually burned out at about the voltage this was expected because of the sharp (exponential-like) increase in current.

5.4 Discussion

To assist in the labeling of the various NDR regions, we have used the band profiles at various biases generated by a self-consistent Poisson solver. The energy levels of the InAs quantum well states were then calculated using a single band model. This method is precise for modeling the intraband resonant state energies because the InAs quantum well states lie above the window between the InAs conduction band and the GaSb valence band; little interaction between the various conduction and valence bands is expected. For zero bias, the calculated energy for the lowest InAs well confined state matches reasonably well with the energy of the transmission peak calculated from the two-band model. The interband resonant state energies calculated with a single band model are not precise because of the coupling between the electron-like states in the InAs contacts and the hole-like states in GaSb valence band wells. For these structures, however, the single band model predicts that only one confined state in each GaSb valence band lies above the InAs conduction band edge, and at energies similar to those predicted by the two-band model, and therefore can be used to approximate the resonant voltage.

This method predicts a total of three resonant states for each structure: two interband and one intraband. The two interband resonances result from the localizing of the light-hole
states as a bias is applied across the structure. For each structure, tunneling through the far well (relative to the emitter) occurs at a lower bias than tunneling through the near well, while intraband tunneling occurs at a much higher bias. For Samples #2 and #3, there are more features in the interband tunneling region than are predicted by the simple two-band model, especially those in MWRIT3. At $k_i = 0$, selection rules only allow coupling between the conduction band and the light-hole band, thus the justification for using only a two-band model to account for the interband coupling. Away from $k_i = 0$, coupling to the heavy hole band is allowed [58]. For $k_i \neq 0$, significant contributions to current can occur (see Section 2.4.2) it is reasonable to conclude that these “extra” resonances result from contributions from heavy-hole bands. Similar structures have also shown similar extra structure in the I-V characteristics. Transmission coefficients calculated including $k_i \neq 0$ have indicated these are the result of heavy-hole band contributions [23].

MWRIT1 has a unique feature that is not seen in the other two samples. The resonance at 0.70 V has a large hysteresis giving a shift of 0.09 V depending on sweep direction, with an absolute peak current for the upward sweep of about 21 mA. The resonance at 0.35 V is similar in peak current to the one at 0.70 V but shows no sign of hysteresis. The load line for the I-V measurement, estimated from the 0.35 V resonance has a resistance of 2 Ω which will not account for the hysteresis. This implies that the hysteresis must be a result of a bistability inherent to the device. Current bistability has been observed in InAs/Al$_x$Ga$_{1-x}$Sb tunneling structures resulting from charge collection in valence band wells [9]. In those structures, the composition of the Al$_x$Ga$_{1-x}$Sb was such that the interfaces were Type II, but not broken. Trapped charge, which is present depending on the electrical history of the structure, in the resulting collector well modifies the potential profile resulting in the bistability [43]. This exact situation will not occur in MWRIT1 because the bandgap of the barriers is broken, not allowing charge to be trapped in the collector well. However, it should be possible to trap charge in the emitter well. The charge can be removed when the emitter well comes in resonance with the InAs well. This is the most likely explanation for the hysteresis.
Given this argument, it should be expected to see similar effects in MWRT2 and MWRT3. MWRT2 has a small, unexplained resonance after the intraband resonance at 0.89 V and one at -0.95 V that can only be seen in the conductance. MWRT1 has a similar, small resonance that is only visible in the downward sweep of the resonance that shows the hysteresis, which also occurs after the intraband resonance at 0.62 V. The similarity of the two small features in both size and position suggest the may be related and that the presence of the AlSb barrier somehow prevents the bistability indicated by the hysteresis.

5.5 Summary

InAs/GaSb/AlSb heteroepitaxial tunneling structures in which both interband and intraband tunneling occur, dependent on injection energy have been realized. The structures, which may be adaptable for three terminal operation, consist of a single InAs well with GaSb barriers which serve as quantum wells for interband tunneling and barriers for intraband tunneling. The addition of a single, thin AlSb barrier variably throughout the double barrier structure can greatly influence the nature of the features in the I-V/conductance characteristics.

The number of observed interband features is greater than can be accounted for by simple conduction band-heavy hole band tunneling. When contributions for $k_4 \neq 0$ are taken into account, extra features from heavy-hole bands can be present. The extra features are consistent with heavy-hole related tunneling seen in other structures.
Chapter 6

Resonant Tunneling Applications

6.1 Introduction

Two of the most promising applications for RTDs are high speed switches and oscillators and in reduced complexity of circuits. RTDs have shown oscillations as high as 712 GHz [P], and switching times as short as 1.7 ps [38]. The complex current-voltage characteristic of RTDs can be used to reduce the number of circuit elements required to implement logic functions.

An RTD has an I-V characteristic similar to that of an Esaki tunnel diode which allows implementation of both binary and multistate logic with fewer devices than conventional transistor circuits. The simplest demonstration of this is when a resistor acts as a load for a tunnel diode (Figure 6.1). If the resistor has a larger value than the average negative differential resistance (NDR) the load line intersects the RTD I-V characteristics more than once, allowing two stable operating points for the circuit exist.

The application of the NDR characteristics of Esaki tunnel diodes to digital switching circuits was recognized and demonstrated shortly following the first demonstration of the tunnel diode [4]. However, the large junction capacitance and difficulty in fabrication of uniform Esaki tunnel diodes proved to be a stumbling block for integrated circuits [53, 56].

Precisely controlled semiconductor heterostructure epitaxial growth techniques have
Figure 6.1: Current-Voltage characteristics of an RTD using a single load resistor. The NDR region of the RTD allows two stable operating points to exist for the same circuit. The point lying on the NDR region is not stable.

lead to the development of the resonant tunneling diode (RTD) which has a similar peaked I-V characteristic [5, 48]. Multiple RTDs can be integrated in series to produce a multiple peaked structure [39]. RTDs do not suffer from the same problems as Esaki tunnel diodes and have been used to implement a number of different circuit designs [22, 45, 46, 50].

In addition to discrete RTDs, integration of RTDs with conventional transistor structures into a single three-terminal device has been demonstrated [3, 32, 44]. The advantage of this approach is the decoupling of the output from the input that is associated with three-terminal operation while retaining the peaked I-V characteristics. The disadvantage is that of increased fabrication complexity for certain configurations.

A simple yet flexible switching configuration can be made using a single discrete RTD and a single transistor. In this configuration, the RTD acts as a switching device but does not set the operating logic levels. This approach eliminates uncertainty in logic levels due to fluctuations in individual RTD characteristics, increasing the noise margin for each state.
Figure 6.2: (a) Schematic of the switching block. (b) Transfer function of the switching block compared with the RTD transfer function.

### 6.2 RTD/Transistor Switching Block

The basic switching block is shown in Figure 6.2a. A schematic of the transfer function is shown in Figure 6.2b. The RTD is the actual switching element, while the transistor acts as both an output buffer to the RTD, and a source of gain. An input buffer (not shown) is required for a constant input impedance.

When $V_{in}$ is zero, the transistor is off, and $V_{out}$ will be $V_{DD}$. As $V_{in}$ is increased, the voltage across the transistor base-emitter junction will linearly increase as determined by the shunt resistance, assuming the base impedance of the transistor is much greater than shunt resistance. Once the input current is large enough that the voltage drop across the shunt resistor is equal to the turn-on voltage for the base-emitter junction, the transistor will begin to draw the excess input current, the transistor will be forced into saturation.
and $V_{out}$ will be forced to $V_{CE\text{sat}}$.

At some point, the voltage drop across the RTD will increase such that it is operating in the NDR region, and the input current will drop sharply, turning off the transistor, thus switching $V_{out}$ back to $V_{DD}$. When the input current again exceeds the threshold value, the transistor will turn on, and $V_{out}$ will switch back to $V_{CE\text{sat}}$.

Small fluctuations on the order of a single monolayer in the epitaxial structure used to form an RTD can cause relatively large fluctuations in both its peak current and peak voltage characteristics [40, 41]. The uncertainty of these parameters reduces the robustness of RTD circuits using resistive and RTD loads. The threshold nature of the switching element described here helps to reduce the effects of those fluctuations. The voltage at which the threshold is reached can be set by choosing the appropriate value for the shunt resistor. Figure 6.3 shows the output of the switching block for different RTD I-V characteristics. The peak current (voltage) can vary by as much as $\pm 20\% (\pm 5\%)$ and still produce the same output voltage for input voltages of, for example, 1.25, 2.5, 3.75 and 5 V which are easily producible by using a resistive ladder network to generate $V_{in}$. Because the output voltages are generated by sources external to the RTD, there is no uncertainty in operating voltage levels resulting from fluctuations in the RTD I-V characteristics, and hence, errors are not propagated through a circuit consisting of a number of switching blocks in series.

A disadvantage to the switching block as described is that it relies on forcing the output of a transistor into hard saturation. This can be eliminated while still obtaining identical functionality, at the cost of more transistors, by replacing the transistor with an emitter coupled pair.

### 6.3 Adder Circuits

#### 6.3.1 Binary Adder

The usefulness of this switching block is demonstrated in Figure 6.4, which is a full adder, including carry bit. The simulated output is shown in Figure 6.5. A full adder, built using
Figure 6.3: Simulation of the output voltage ($V_{DD} = 5V$) and the transistor base-emitter voltage for different RTD I-V characteristics. The peak current/voltage of the RTD was varied by ±20% (±5%).

Conventional architectures requires 25~30 transistors. The circuit shown here uses 5, plus one RTD. Of those 5, three function as buffers and/or inverters. The sum bit is simply the switching element with an inverter to give the correct polarity. The carry bit is a saturable amplifier. Advantages of the reduced device count are more functionality for the same on chip area, or reduced fabrication tolerances for a given amount of functionality area density.

With proper choice of output stages, the lower output voltage could be set very near to 0 V, implying a large on/off voltage ratio. Assuming the output is a bipolar transistor which is turned off or on, the output peak/valley current ratio could be several orders of magnitude. While this compares to the on/off current ratio of a CMOS inverter, it is important to note that this does not achieve the same low power dissipation as a CMOS inverter. This is because in the circuit, even when the output voltage is low, there is still current flowing through the RTD (and the shunt resistor). The best room temperature peak/valley ratios for tunneling structures that have been achieved are on order of 100 [10]. Thus the valley current of the RTD is the limiting factor to low power dissipation, not the output stage.
Figure 6.4: Schematic of a full binary adder using the RTD/Transistor switching block.

Figure 6.5: Simulation of the full adder. The inputs A, B, C scaled to fit on the same set of axes.
6.3.2 Ternary Adder

Increasing the number of logic levels used in a circuit can further reduce the number of devices required for a given functionality and/or complexity of the device interconnect scheme [49]. An RTD based ternary adder has previously been demonstrated [62]. That circuit was a clock based circuit where the output voltages were determined directly by a resistive load line on an RTD.

A ternary adder can be built using two of the previously described switching blocks, one with a slight modification with respect to the other. The circuit, shown in Figure 6.6, has the same advantages as those described for the binary adder. The first block consists of Q1, and an inverter, Q2. The transfer function for this stage only is shown in Figure 6.7a. When two RTDs are connected in series, a multiple peaked I-V characteristic is obtained, as shown in Figure 6.7b. This adds an extra switching cycle to the transistor Q3. The thresholds are set such that Q3 initially switches before Q1 (and hence Q2). The full circuit transfer function is obtained as follows: When Q1 is off, Q2 is on and the sum is low, regardless of Q3. When Q1 is on, and Q2 is off, the output is high if Q3 is off. However, if Q2 is off and Q3 is on, the output is midway between high and low, assuming the collector resistors for Q2 and Q3 are equal. The measured transfer function for the circuit (sweeping the emitter of Q6) is shown in Figure 6.8.

Compressed functionality circuits as described above are only advantageous if some overall quantity can be improved on, such as power dissipation or wafer area. For a binary application, CMOS inverters dissipate essentially no power compared to the switching block of Figure 6.2 and thus will always be advantageous over the RTD based circuit. The power dissipation of the switching block can be improved by replacing the shunt resistor with a second RTD. This is equivalent to replacing the resistive load in the circuit shown in Figure 6.1 with an RTD load. This allows the low voltage operating point to dissipate less power. While this still does not approach CMOS power dissipation, the circuit may still be useful if for example, speed or inclusion in a III/V based optoelectronic circuit is required.
Figure 6.6: Schematic of a ternary adder. $V_{DD}$ for the carry output is half that for the sum output.

rather than low power dissipation.

For ternary applications, the simplest possible CMOS ternary circuit is shown in Figure 6.9. The high and low states are obtained in a manner equivalent to a binary CMOS inverter where only one transistor is on at a time. Since no current flows in either of these states, no power is dissipated. The intermediate state, however, requires both transistors to be on, allowing current to flow through the resistor to give the intermediate output voltage. In this light, the RTD based circuit begins to approach the power dissipation characteristics of the CMOS circuit because the intermediate voltage is generated in the same fashion (the corresponding resistors in Figure 6.6 are the collector resistors of Q2 and Q3) because the circuit in Figure 6.6 is the complete adder circuit, while any CMOS based ternary circuit will require numerous ternary inverter circuits to implement useful functions.
Figure 6.7: Various transfer function schematics for transistors in the ternary adder circuit. The input voltage is taken to be the emitter of Q6. (a) Schematic of transfer function of Q2. (b) Input current to switching block with two RTDs in series (i.e. Q3). (c) Collector voltage for Q3. (d) Sum voltage for the circuit.

Figure 6.8: Measured transfer function for the ternary adder circuit. The input voltage is taken to be the emitter of Q6.
Figure 6.9: Schematic of a CMOS ternary adder. High and low output voltages are obtained in a manner similar to that of conventional CMOS in which only one transistor is turned on at any one time. The intermediate voltage is obtained when both transistors are on.

6.4 Conclusion

The switching characteristics of a simple RTD/transistor switching block have been demonstrated. The configuration switches when a threshold current through the RTD is obtained, decreasing the variability from RTD peak current and voltage fluctuations. The switching block was used to build both a binary and a ternary adder with a minimum number of active circuit elements.
Chapter 7

Resonant Tunneling in THETA structures

7.1 Introduction

One of the limiting factors in high-speed operation of semiconductor devices is the transit time of carriers across some particular region (e.g. the base of a bipolar transistor). In general, the transit time is limited by the mobility of the material. If the region to be traversed is shorter than the average scattering length, ballistic transport will dominate and the transit time will be much shorter than expected based on the bulk mobility of the material. In semiconductors, this distance is short enough that quantum size effects can become important.

The tunneling hot electron transfer amplifier (THETA) is a unipolar device based on ballistic transport across a thin base region [19, 17, 68]. Transistor action for a THETA is achieved in a manner analogous to a bipolar transistor, shown in Figure 7.1.

Carriers are injected through a thin tunnel barrier into the base at an energy significantly above the base Fermi level. Turn-on occurs when the top of the base-collector barrier is lowered such that the injected electrons can cross over the barrier and an electric field is established which will draw injected electrons to the collector contact [30].
Figure 7.1: Operation of a THETA structure. A voltage placed across the base-emitter junction will inject electrons through the thin tunnel barrier into the base. A voltage across the base-collector junction will vary the height of the energy barrier. Those electrons which traverse the base ballistically (i.e. do not lose energy via inelastic scattering) will be swept into the collector if the top of the energy barrier is sufficiently lowered. In the common-base mode, derivative of the collector current is proportional to the energy distribution of the injected electrons.
THETA structures fabricated from the InGaAlAs system have been shown to exhibit room temperature operation at both dc and high-frequency [31]. In these structures, there is no graded region at the collector-base barrier. This increases the lever arm effect (a larger collector-base voltage is required to lower the collector barrier) and enhances the effect of quantum mechanical confinement in the base.

The effect of this confinement near the turn-on region of operation is investigated at low temperature (1.4 K) in a series of InGaAlAs THETA structures. The turn-on characteristics are shown to be modified by carriers resonantly tunneling through a confined state in the base at energies lower than ballistic turn-on.

### 7.2 Device Structure

The emitter of all three structures consists of n-type In$_{0.52}$Al$_x$Ga$_{0.48-x}$As, where $x = 0.4$ for the 40 nm base width sample and $x = 0.48$ for the 30 nm and 50 nm samples. A 1 nm AlAs tunnel barrier separates this from the base, which is a 2 nm layer of undoped In$_{0.53}$Ga$_{0.47}$As followed by a layer of n-type In$_{0.53}$Ga$_{0.47}$As ($1 \times 10^{18}$ cm$^{-3}$) of 30 nm, 40 nm or 50 nm for the different samples. The collector-base isolation layer is a 250 nm layer of undoped In$_{0.52}$Al$_x$Ga$_{0.48-x}$As where $x = 0.2$ for the 40 nm base width sample and $x = 0.3$ for the 30 nm and 50 nm samples. The collector is a 50 nm layer of undoped In$_{0.53}$Ga$_{0.47}$As and a thick n-type layer of In$_{0.53}$Ga$_{0.47}$As ($5 \times 10^{18}$ cm$^{-3}$). The self-consistent (Poisson) conduction band of the structure [13], biased to turn-on, is shown in Figure 7.2.

### 7.3 Experimental Results

The turn-on characteristics of both the common-base and common-emitter configurations of the structures were investigated. All measurements were done at 1.4 K.
Figure 7.2: Self-consistent (Poisson) band diagram of a structure biased to turn on. The inset shows the region near the injection energy.

### 7.3.1 Common-Base Characteristics

For an arbitrary THETA structure, the common-base conductance \(dI_C/dV_{CB}\) is expected to be a single peak corresponding to injection above base-collector barrier. The conductance peak will shift to lower energy with increasing emitter current reflecting the increase in energy of the injected carriers, while the FWHM of the peak remains essentially constant [19]. The common-base characteristics for the 50 nm base width sample are shown in Figure 7.3. With increasing emitter current, the FWHM of the conductance peak increases linearly as a result of series resistance in the base.

The base-emitter conductance characteristics \(dI_E/dV_{BE}\) vs \(V_{BE} (V_{CB} = 0 \text{V})\), are shown in Figure 7.4. The peaks in the conductance result from the confined states in the base [18] with the relative strengths of the peaks increasing with decreasing well thickness. Up to \(V_{BE} \approx 0.25 \text{V}\), the conductance of these samples shows a linear trend (plotted on a log scale) as expected with a single tunnel barrier. At higher voltages, however, where the current approaches \(\sim 1 \text{mA}\), the conductance saturates as series resistance dominates.

\(V_{BE}\) includes the voltage drop across the series resistance in addition to the potential drop across the junction, and therefore, it must be added to \(V_{CB}\) for the true potential drop across the collector-base junction. Because the emitter current is fixed, this extra voltage
Figure 7.3: Common-base characteristic for a 50 nm base width sample.

Figure 7.4: Base-emitter junction conductance characteristics ($dI_E/dV_{BE}$ vs $V_{BE}$).
drop varies linearly with the change in collector current (and corresponding change in base current). This causes the width of the common-base conductance peak to increase with emitter current. The base series resistance can be determined by measuring the change in base voltage as a function of base current before and after turn-on. For the 50 nm device at 1.4 K, the total base-emitter series resistance was measured to be $\sim 110 \, \Omega$ and the base series resistance was measured to be $\sim 65 \, \Omega$.

A nominally identical structure, with a 30 nm rather than a 50 nm base width, does not show the single, expected peak in the common base conductance (Figure 7.5). While there is a strong peak, associated with the ballistic turn-on, near $V_{CB} = 0.5 \, V$, there is a prominent shoulder at a lower $V_{CB}$. As the collector voltage is increased, the effective barrier thickness is reduced as with Fowler-Nordheim tunneling (Figure 7.6). Just before the expected turn-on, the base-collector barrier thickness (for the injected electrons) is comparable to that of the emitter-base injection barrier, creating a double-barrier tunneling structure. If a confined state exists at an energy within the occupied energy region of the emitter conduction band, those electrons will see an enhanced transmission probability through the entire double barrier structure at the energy of the confined state (i.e. the electrons can resonantly tunnel), which results in an increase in the collector current.

As the emitter current is increased, the strength of the shoulder is decreased relative to the main peak. This is because for larger emitter currents, the confined base state lies farther down in energy (relative to the emitter Fermi energy) and the transmission efficiency is reduced [37]. The 40 nm base width also shows this effect.

### 7.3.2 Common-Emitter Characteristics

The existence of the confined state near the injection energy also influences the common-emitter turn-on. In this configuration, there is an offset voltage for ballistic turn-on resulting from the voltages required to both bias the emitter-base junction and lower the base-collector barrier. With increasing base current, the bias voltage across the emitter-base junction is larger, so the collector-emitter voltage must be larger to achieve the required
Figure 7.5: Common-base characteristic for a 30 nm base width sample.

voltage drop across the base-collector junction, causing the offset voltage to increase with increasing base current. This is readily apparent for the 50 nm base width sample, shown in Figure 7.7.

In the 30 nm base width sample, shown in Figure 7.8, the offset voltage does not increase, but rather, decreases with increasing base current. As the collector-emitter voltage is increased, the effective double-barrier condition is created. Because the confined base state lies below the top of the base collector barrier, it will reach the emitter Fermi level before turn-on, allowing resonant tunneling to occur. With increasing base current, and hence emitter base voltage, the confined state will reach the emitter Fermi level sooner relative to ballistic turn-on, causing the offset voltage to decrease with increasing base current.

When the separation between resonant tunneling and ballistic turn-on (in collector-emitter bias) is large enough, the resonant tunneling results in a region of negative differential conductance (NDC) in the collector current which is visible for the 30 nm base width sample. Presumably, the strength of the NDC will decrease if the base current is too large
Figure 7.6: Self-consistent (Poisson) band diagram for three different values of $V_{CB}$, with $V_{CB3} > V_{CB2} > V_{CB1}$. Note the reduction in effective barrier thickness (at the injection energy) in addition to the reduction in barrier height.
since confined state will be in resonance before appreciable reduction of the base collector barrier occurs. This, however, was not observed because the required base current is large enough to destroy the device. The 40 nm base width structure shows the decrease in offset voltage with increasing base current, but not the NDC.

7.4 Conclusions

It has been shown that in a THETA structure, resonant tunneling through a confined state in the base lying just below the top of the base collector barrier can occur. This resonant tunneling leads to a soft turn-on in both the common-base and common-emitter configurations, and to a region of NDC in the common-emitter configuration.

In previously studied THETA structures the base collector barrier is initially graded over a few nanometers to prevent an abrupt junction, reducing the amount of quantum mechanical reflections in the base. In the current structures, the presence of such a graded
Figure 7.8: Common-emitter characteristics for a 30 nm base width sample.

region at the base collector interface presumably would reduce or possibly eliminate the
effects discussed here.
Chapter 8

Conclusions

The main focus of this work has been the resonant interband tunneling characteristics of single well InAs/GaSb/AlSb interband tunneling structures, which showed early promise as candidates for high peak-to-valley (P/V) current ratio devices. In “conventional” resonant tunneling diodes (RTDs), resonant tunneling current is cutoff when momentum cannot be conserved by the tunneling electrons. In resonant interband tunneling (RIT) structures, cutoff of the resonant tunneling current results from conservation of energy, occuring when carriers in the emitter contact are blocked by the bandgap of the well. This cutoff mechanism greatly reduces the inelastically scattered tunnel current, which is the largest source of valley current in conventional RTDs, increasing the P/V ratio.

For an InAs/AlSb/GaSb/AlSb/InAs structure, in which the confined states in the valence band of the GaSb well are accessible for resonant tunneling, the largest component of resonant tunneling current arises from tunneling via the light-hole states. However there is also the possibility of tunneling via the heavy-hole states. Because of the opposite sense of the dispersion curves for the tunneling electrons and the confined hole states, tunneling via light- and heavy-hole states will occur simultaneously (at $k_\parallel \neq 0$) despite the fact that the confined states lie at different energies (at $k_\parallel = 0$). The current cutoff for each state will occur at a different applied bias, effectively increasing the valley current associated with a particular state. Thus the P/V ratio for the main current peak, associated with a
light-hole state, can be reduced by tunneling current via heavy-hole states. While heavy-hole state tunneling current has been predicted theoretically, and indirectly demonstrated experimentally, this work shows direct experimental evidence of tunneling through a heavy-hole state in a single p-type (i.e. GaSb) well RIT. The different behavior of behavior of light- and heavy-hole states (because of different effective masses) under the influence of a magnetic field allowed the contributions of different states to be uniquely identified in the I(V) characteristics of the RIT. These measurements are described in Chapter 4.

Simple attempts to describe the I(V) characteristics of RITs have been limited to one-dimensional models, using only the conduction band and the light-hole band in the well. Attempts to include three dimensions have resulted in very complex models. There is no coupling between conduction band and heavy-hole states at $k_{\perp} = 0$ so most of the current through the heavy-hole state occurs away from $k_{\parallel} = 0$ and it can be a significant fraction of the total current. In Chapter 2, a simple model which takes this into account is presented. The model qualitatively describes the I(V) characteristics through the tunneling structure under the influence of a magnetic field and agrees well with the experimental data.

While "leakage" current via heavy-hole states appears to be significant at low temperatures, it has been demonstrated elsewhere that the main component of valley current at room temperature is hole diffusion out of the well due to a relatively small hole confining potential (i.e. the InAs/GaSb valence band discontinuity). So far, the only solution to this is to incorporate AlAs into the AlSb tunnel barriers. The AlAs bandgap extends down into the InAs bandgap increasing the size of the potential barrier seen by the holes in the GaSb. However, because of the large lattice mismatch between AlAs and the InAs/AlSb/GaSb system, the thickness of the AlAs layer is limited to a few monolayers, which is not sufficient to provide a significant barrier. This appears to limit the room temperature P/V ratio to what has already been achieved (30).

To take advantage of the efficient current cutoff mechanism inherent in interband tunneling, different device structures will be necessary. This, however, will require further investigation of more complex structures, such as those described in Chapter 5, to better
understand the coupling mechanism between light- and heavy-hole states. In these structures, it is shown that the placement of a single thin AlSb tunneling barrier results in a large variation in the I(V) characteristics. More complex device structures will also require careful design to avoid other unexpected effects such as the turn-on characteristics of a THETA structure described in Chapter 7.

Because of the room temperature characteristics of the simple structures, and the large variations in I(V) characteristics of the more complex structures, one of the more likely applications of RTDs in general will be in the form of threshold elements in larger compressed functionality circuits, such as the simple RTD/transistor switching block demonstrated in Chapter 6. One advantage of compressed functionality circuits is the smaller required area achieved by reducing the required number of devices. In low power applications, by reducing the number of devices, the effect of a large valley current on the off-state current can be diminished. Also, because the RTD (or RIT) does not have to supply drive current, the total current (by means of device area) can also be reduced, further reducing the consequences of a large valley current.

While there is a wide body of literature investigating the nature of resonant tunneling, there has been relatively little involving resonant interband tunneling, and most of what does exist has been theoretical in nature. Much of the experimental studies of RIT structures has been carried out at 77K and above, where the valley current is dominated by hole leakage out of the GaSb well. Clearly more low temperature studies of these structures is required for a full understanding of the nature of the resonant interband tunneling process.
Bibliography


