QUANTUM MAGNETOTRANSPORT STUDIES OF SEMICONDUCTOR HETEROSTRUCTURE DEVICES

Thesis by

Ronald R. Marquardt

In Partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy

California Institute of Technology
Pasadena, California

1995

(Submitted June 20, 1994)

to my family and friends

Acknowledgments

When attempting to summarize the contributions of many colleagues and friends to a work encompassing such a vast part of your life, it is difficult to know where to begin. Certainly, this thesis is in no small part due to the unique vision and guidance of my advisor, Tom McGill. Long ago he told me there would be something interesting to see if I'd just look at these interband devices in magnetic fields. At times, I admit, I thought him the techno-warrior equivalent of the grizzled old miner who swears by the gold locked deep in some favorite mountain. I was quite gratified to find a nugget or two along the way, and I am indebted to Tom for the experience.

If Tom is the brains behind the operations, then it is Marcia Hudson who is the glue which keeps it from devolving into a chaotic soup of primordial ooze. If you don't believe me, just wait until the next time she's off on vacation... I can honestly say that Marcia's friendship and encouragement have contributed greatly to this thesis. I also must thank her son, David, for keeping me up to date on all things Jurassic. The administrative assistance of Sandy Brooks and Carol McCollum was also appreciated.

It really is cliché to say, after reading so many thesis acknowledgments before this, but the real strength of Tom's research group is the quality and character of its members, and I can honestly claim to have enjoyed working with all that have overlapped with my tenure. David Ting has been a mentor and good friend,

and has always been willing to provide cheerful reminders of my fallibility. Doug Collins grew the samples that I studied and provided great insight into the resulting data. He has also proven to be both an apt guide to the links in Scotland ("Putt, dammit, putt!!") and a good friend. As contemporaries, I owe Mike Wang, Johanes Swenberg, and Rob Miles many thanks for their encouragement, support, and good-natured practical jokes. If I only knew how to get them back... Harold Levy has always been a constant source of witty observations and insightful ideas. His contributions to the group's computer system have been much appreciated. Yixin Liu provided the theoretical support for this study, without which it would not be complete; I have also enjoyed his stories of China and its culture.

I owe much to Mark Phillips, Ed Yu, Mike Jackson, and David Chow, who, as senior students and post-docs provided much-needed assistance to my research. Chris Springfield and Per-Olov Pettersson have, between them, currently taken on responsibility for nearly every piece of equipment the group owns, and should clearly be thanked for their efforts. I have enjoyed working with Erik Daniel and Alicia Alonzo, and I wish them well as they begin their scientific Odyssey. My interactions with Andy Hunter, Gerry Picus, Odgen Marsh, Ron Grant, and Jim McCaldin have been fruitful and rewarding.

Outside the group, many people have made life as a graduate student at Caltech not only possible, but enjoyable. I must thank Lisa Lorden for her continuing faith and support. James and Debbie Larkin have been great friends and roommates, and I will miss their company. I can always count on John and Gemma Iannelli to make me smile, and have truly appreciated their friendship. Carl and Susan Gaines, Pat Ritto, Craig Sangster and David Fields all deserve credit for supporting and encouraging me.

Last but not least, there is my family. Somehow my brother, Jon, managed to put up with me as a roommate for nearly two years during my graduate tenure, and

I am glad that we have become better friends as a result. My parents deserve more credit than all others combined. Their faith and love have been an inspiration.

Abstract

Resonant interband tunneling is perhaps the most generalized example of quantum transport in semiconductor heterosystems, and yet few studies of interband magnetotunneling have been reported. This thesis presents the results of the first in-depth study of quantum magnetotransport through valence band-like well states in broken-gap semiconductor heterostructures. The transport characteristics of a series of InAs/AlSb/GaSb/AlSb/InAs resonant interband tunneling (RIT) diodes were investigated in magnetic fields as great as 8 tesla oriented both parallel and perpendicular to the tunnel current density. The resulting observations advance the understanding of interband transport and reveal interesting and complex phenomena previously unseen in magnetotunneling investigations.

In magnetic fields perpendicular to the tunnel current density, evidence of both low-mass, negative dispersion, and high-mass, positive dispersion states in the GaSb well is shown. The magnetic field in this geometry has no discernible effect on the RIT transport characteristics at low fields, however, and the transition between this low-field regime, and the expected response at higher field strengths is extremely abrupt. Associated with this marked change in behavior is an additional, narrow peak present in the negative differential resistance (NDR) region of the device at the threshold, or critical, magnetic field. This anomalous discontinuity in device transport violates semi-classical theory and is suggestive of a dramatic and fundamental change in resonant quantum transport.

With the field applied normal to the epitaxial layers, Landau levels form and resonant tunneling through them is observed indirectly via Shubnikov-de Haas-like oscillations of the tunneling conductance. The non-conservation of Landau level index in interband magnetotunneling is first proposed theoretically, and subsequently verified experimentally. Evidence for this effect is asserted from the discontinuous changes in the Shubnikov-de Haas oscillatory phase as a function of applied bias. These phase shifts result from bias-dependent changes in the resonant current path through well hole states of differing longitudinal angular momentum. The data are the first observation of Landau level mixing in interband tunnel devices, and only the second report for magnetotunneling in general. The lack of separate Shubnikov-de Haas oscillations for coupling to spin-up and spin-down states is interpreted to be evidence that the well masses are significantly less than the electrode mass, such that the reduced mass $\mu^{\star}$ of the two is approximately equal to the well mass, $m_W^{\star}$.

List of Publications

Work related to this thesis will be published under the following titles:

Resonant magnetotunneling spectroscopy of p-well interband tunneling diodes,
R. R. Marquardt, D. A. Collins, Y. X. Liu, D. Z.-Y. Ting, and T. C. McGill, submitted to Phys. Rev. B.

Evidence of Landau level mixing in resonant interband tunneling diodes,
R. R. Marquardt, D. A. Collins, Y. X. Liu, D. Z.-Y. Ting, and T. C. McGill, to be submitted to Appl. Phys. Lett.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Introduction to Thesis

This thesis details research of quantum transport in semiconductor heterojunction devices in the presence of external magnetic fields. Specifically, the devices studied are epitaxial resonant tunneling diodes (RTD's) fabricated in the InAs/AlSb/GaSb material system. The purpose of this study was two-fold. First, the peculiar nature of the InAs/GaSb band alignments makes the system attractive for basic physics studies of tunneling phenomena. Second, it was hoped that a better understanding of the underlying physics might help in the realization of practical uses and optimizations of this technology for real-world problems. Much of this thesis will focus on the first of these goals, both because of the breadth of the fundamental results and the lack of technological maturity which hinders the adoption of quantum phenomena-based devices.

1.1.1 Motivation

Magnetic fields have proven to be a useful, if not indispensable, tool in the investigation of low-dimensional systems. In particular, the understanding of the physics of resonant tunneling in semiconductor heterostructure devices has been advanced by studies of quantum transport in the presence of magnetic fields. While magnetic fields do no work on individual charge carriers, their effect on the energy and momenta spectra allow subband structure, occupation, and dispersion to be observed experimentally [1]. With the field aligned perpendicular to the growth plane of a semiconductor heterostructure device (parallel to the tunnel current density), the energy spectrum of the quasi-bound states is modified by the addition of discrete ladders of Landau levels. Studies of devices in this configuration [2, 3, 4, 5, 6], and in tilted fields, with magnetic field components both parallel and perpendicular to the epitaxial layers [7, 8], investigate transport properties in systems with increased quantization and reduced dimensionality. The resulting changes in the current-voltage characteristic of a device are indicative of the energies of quantized subbands in the well, the in-plane subband dispersions, and the coupling of the well states to those in the electrodes. In addition, resonant tunneling through Landau levels in the well may be observed through Shubnikov-de Haas-like oscillations of one or more transport variables.

In contrast, with a magnetic field aligned parallel to the growth plane quantization effects due to the field and the crystal potential are coaxial. Except for the highest fields, or widest quantum wells, the magnetic field in this geometry may be treated perturbatively. As a result, the device operates without additional quantization or significant modification to the eigenstates and energies of the system; the only effect of the field, semi-classically, is to bend the carrier trajectories into cyclotron orbits about the axis of the field. This alteration of the carrier distribution in momentum space is exploited in resonant magnetotunneling spectroscopy [9] (RMTS) to probe the energy subband dispersion in semiconductor quantum wells. The shift in k-space induced by the magnetic field is used to experimentally select the parallel k-vector of elastically tunneling electrons. The subsequent effect upon the current-voltage (I-V) behavior of the device, specifically the peak voltage, provides an indirect map of the band structure in the well.

While the aforementioned applications of magnetic fields to the study of quantum transport, and specifically to RTD's, are well-known, relatively few observations have been reported for interband devices employing the InAs/GaSb Type II broken-gap band alignment. The unique properties of InAs/GaSb heterostructures have been known for some time [10], and recently have been exploited to create a new class of quantum transport device [11, 12]. Because the conduction band edge of InAs lies 0.15 eV lower in energy than the valence band edge of GaSb [13], transport through this system involves resonant tunneling between conduction band states in InAs and valence band states GaSb. The few observations of magnetotunneling in devices in this system have focused on the double barrier heterostructure with an InAs well, AlSb barriers, and GaSb electrodes [14, 15, 16], or have investigated the barrierless InAs/GaSb/InAs system [17]. The lack of detailed studies of interband transport in magnetic fields, especially for devices with the rich valence subband structure of GaSb wells, provided the motivation for this thesis. The transport characteristics of a series of InAs/AlSb/GaSb/AlSb/InAs resonant interband tunneling (RIT) diodes were investigated in magnetic fields as great as 8 tesla applied both parallel and perpendicular to the tunnel current density. Magnetotransport in this system combines the complexities of hole subbands, interband tunneling, and quantum magnetic effects. The resulting data represent the first in-depth study of interband quantum magnetotransport through valence band-like states, and the data have revealed interesting and complex phenomena previously unseen in magnetotunneling investigations.

1.1.2 Summary of Results

The resonant magnetotunneling spectroscopy (RMTS) technique, with the magnetic field aligned in the plane of the epitaxial layers, provides a probe of the quasi-bound state in-plane dispersions in double barrier heterostructures. When applied to GaSb-well RIT diodes, anomalous behavior is observed. There exist two distinct regimes of differing character in the RIT data, one for fields less than a critical magnitude, and another for fields at, or exceeding, this same B_{crit}. For fields less than B_{crit} ($\lesssim 5$ T) little change in the current-voltage characteristics of the RIT samples was observed. The exact magnitude of the critical field varies between samples, but does not appear to depend monotonically on well width. Above the critical field, the RIT I-V characteristics are strongly affected by the field, and demonstrate behavior consistent with the RMTS theory. The data indicate the presence of both low-mass, negative dispersion, and high-mass, positive dispersion, states in the well. At the critical field, there is an additional, narrow peak that occurs in the NDR region of the I-V characteristic over a ≈ 300 G range of fields. This second peak is not seen to be dependent upon external circuit parameters, and is therefore not a circuit instability or oscillation effect. Furthermore, the abruptness of its appearance (less than a 50 G change in field at 5 T) is not consistent with a simple, cyclotron orbit resonance. Similar behavior is not seen in InAs/AlSb double barrier structures measured with the same equipment, and thus the phenomenon seems directly related to the interband nature of transport in the InAs/AlSb/GaSb system. The observation of such an abruptly occurring peak has never been reported before in magnetotunneling experiments, and the associated change in magnetotransport behavior at the critical field is suggestive of a dramatic and fundamental change in device transport physics.

As a magnetic field parallel to the tunnel current density is applied to an RTD, Landau levels form and fundamental changes occur in the criteria for resonant elastic transport. As the transverse momentum is replaced by discrete Landau indices, forcing new conservation laws upon the tunneling carriers, resonant Landau level tunneling may be evident by either direct observation of individual I-V peaks corresponding to each magnetic energy state, or Shubnikov-de Haas oscillations of transport-related device characteristics. While there is some direct evidence for Landau level-associated peaks in the RIT I-V's, it is mostly within the NDR region of the device, and suffers from concurrent magnetoresistance effects. Shubnikov-de Haas oscillations of tunneling conductance are also observed, however, at all biases less than the peak voltage, in fields exceeding 2 tesla. In the limit of 0D-0D tunneling between an emitter notch state and the well, these oscillations are attributable to the alternate alignment and misalignment of the Landau levels on either side of the barrier layer. As such, the frequency of these oscillations, and their bias dependence, reflects the character of both the emitter and well states. Specifically, the band edge energy difference between the two subbands and their reduced mass both determine the nature of the Shubnikov-de Haas conductance oscillations. Additionally, however, interband transport introduces the possibility of Landau level index non-conservation through the coupling of states having different angular momentum projections. With a perpendicular field, the total longitudinal angular momentum projection, including that induced by the cyclotron orbits of the carriers, is conserved, and, consequently, tunneling carriers may mix with states of different Landau index. In turn, this change in Landau level index effects the phase of the resulting Shubnikov-de Haas oscillations. Experimental observations of large shifts in Shubnikov-de Haas phase in RIT conductance data are interpreted as direct evidence for Landau level mixing, and conservation of total longitudinal angular momentum. These data are the first observation of Landau level non-conservation in an interband system, and only the second report for magnetotunneling in general. This work also represents both the first theoretical prediction, and subsequent experimental verification, of Shubnikov-de Haas phase shifts due to Landau level mixing in semiconductor heterostructures.

1.2 Thesis Outline

The remainder of the thesis is divided into three chapters and two appendices. Chapter 2 develops background knowledge of semiconductor heterostructures, the InAs/AlSb/GaSb material system, and the effects of magnetic fields upon bulk transport, with the intent to motivate the experimental studies and provide a solid context in which to view them. Chapter 3 provides details of the RMTS experiments on GaSb-well RIT's, and Chapter 4 presents the results of the resonant Landau level tunneling experiments in the same devices. Two appendices are provided for the reference of future students. The history and operation of the superconducting magnet used for this thesis are detailed in Appendix A, and Appendix B develops details of bulk $\vec{k}\cdot\vec{p}$ theory, and the means by which magnetic fields are incorporated in it.

Bibliography

[1] *Molecular Beam Epitaxy and Heterostructures*, proceedings of the NATO Advanced Study Institute on Molecular Beam Epitaxy (MBE) and Heterostructures, Erice, Italy, edited by L. L. Chang and K. Ploog (Martinus Nijhoff, Dordrecht, 1985).

[2] E. E. Mendez, L. Esaki, and W. I. Wang, Phys. Rev. B **33**, 2893 (1986).

[3] V. J. Goldman, D. C. Tsui, and J. E. Cunningham, Phys. Rev. B **35**, 9387 (1987).

[4] L. Eaves *et al.*, Appl. Phys. Lett. **52**, 212 (1988).

[5] C. E. T. Gonçalves de Silva and E. E. Mendez, Phys. Rev. B **38**, 3994 (1988).

[6] E. E. Mendez, H. Ohno, L. Esaki, and W. I. Wang, Phys. Rev. B **43**, 5196 (1992).

[7] J. J. Koning *et al.*, Phys. Rev. B **42**, 2951 (1990).

[8] M. L. Leadbeater, F. W. Sheard, and L. Eaves, Semicond. Sci. Technol. **6**, 1021 (1991).

[9] R. K. Hayden *et al.*, Phys. Rev. Lett. **66**, 1749 (1991).

[10] L. L. Chang and L. Esaki, Surf. Sci. **98**, 70 (1980).

[11] D. H. Chow *et al.*, in *Quantum-Well and Superlattice Physics III*, proceedings of the SPIE **1283** (Bellingham, WA, 1990).

[12] D. A. Collins *et al.*, in *Resonant Tunneling in Semiconductors: Physics and Applications*, edited by L. L. Chang, E. E. Mendez, and C. Tejedor (Plenum, New York, 1991).

[13] G. J. Gualtieri, G. P. Schwartz, R. G. Nuzzo, and W. A. Sunder, Appl. Phys. Lett. **49**, 1037 (1986); G. J. Gualtieri *et al.*, J. Appl. Phys. **61**, 5337 (1987).

[14] E. E. Mendez, in *Resonant Tunneling in Semiconductors: Physics and Applications*, *ibid.*

[15] E. E. Mendez, J. Nocera, and W. I. Wang, Phys. Rev. B **45**, 3910 (1992).

[16] E. E. Mendez, Surf. Sci. **267**, 370 (1992).

[17] T. Takamasu *et al.*, Surf. Sci. **263**, 217 (1992).

Chapter 2

Background and Motivation

2.1 Introduction

This chapter provides a thorough background for the understanding of semiconductor heterostructures (Section 2.2), the InAs/AlSb/GaSb material system (Section 2.3), and the effects of magnet fields upon transport in bulk semiconductors (Section 2.4). The purpose of this treatment is several-fold. First, these discussions set the framework for the research presented in later chapters. The technological and scientific problems surrounding magnetotunneling in broken-gap heterostructures lend motivation for the work, and provide meaning to the results. Second, too often the physics of heterostructures is reduced to simple arguments that map to textbook problems without the added complexities of the real-life problem being truly addressed. Interband tunneling is perhaps the most generalized example of quantum transport in semiconductor heterosystems, and its study requires a firm understanding of the contributing physical processes. Lastly, magnetic fields have a profound impact upon all transport phenomena in semiconductors, and an intuitive discussion of the basic theory highlights the complexities they introduce and justifies their experimental utility. All of these topics are important background for the study of magnetotunneling in interband heterostructure devices, and as such, provide the proper context and motivation for the work contained in this thesis.

2.2 Heterostructures and Quantum Devices

2.2.1 Background

Semiconductor heterostructures are comprised of two or more dissimilar semiconductors physically mated through an epitaxial or pseudomorphic growth process to form a artificially-constructed single crystalline structure. While this definition encompasses a wide variety of possible semiconductor systems, one standout exception happens also to be the most commercially viable at the present: silicon.[1] The reasons for this success are many-fold, including the maturity of silicon processing technology, the quality (near-ideal) of the Si-SiO_2 interface, and the relative immaturity of heterojunction device technology. As optical devices proliferate, the use of non-silicon materials, and heterojunction systems in particular, will flourish, as silicon, an indirect semiconductor, is a notably poor performer in the optical arena. The electrical properties of heterostructures are of great interest as well, and have been extensively studied since their technological inception [3, 4]. Heterostructures offer the designer the flexibility to tailor a given device for optimal performance for a specific set of criteria. Additionally, as device sizes shrink to the scale that quantum mechanics plays a dominant role in their operation, new device

---

[1] Estimates of the world-wide semiconductor market place the total value of shipments in 1992 at roughly $77.3 billion [1]. Of that total, $66.0 billion was spent on integrated circuits exclusively produced in silicon. In contrast, the total semiconductor diode laser market, one of the only commercial markets to include heterostructure devices, was a mere $287 million that same year [2].

(and computational) paradigms must be developed to constructively exploit these quantum effects. Heterostructure-based devices are poised to perform exactly that role.

The feasibility of heterostructure devices must be attributed to the development of growth techniques which allow the precise, ordered deposition of crystalline constituents on an atomic layer scale. The advent of molecular beam epitaxy (MBE) [5], and other growth techniques such as metal-organic chemical vapor deposition (MOCVD), have made the investigation of semiconductor heterostructures possible. MBE growth takes place in a vacuum chamber under ultra-high vacuum (UHV) conditions ($P \lesssim 10^{-10}$ torr). High purity source material (elemental or compound semiconductor constituents) is evaporated in several crucibles and directed toward a single-crystal growth substrate. Growth conditions (substrate temperature, source flux, etc.) are controlled to cause the molecular beams to condense a monolayer at a time on the substrate surface in a lattice ordered to the bulk substrate crystal. Typical growth rates are on the order of a few Å/sec. Because of the UHV environment, the ultra-pure source material, the high-quality semiconductor substrates, and the controlled thermodynamic conditions, the resulting crystals are typically of the highest quality attainable in the laboratory [6]. The absolute quality of MBE deposited thin films depends strongly on the relative difference between the lattice constants of the grown layers and the underlying substrate, the chemistry of the various compounds, and the skill of the operator. The first may be controlled via the selection of appropriately lattice-matched compounds, the second defines areas of active research, and the last involves as much individual artistry as technical skill. In general, the well-developed material systems for MBE growth are all III-V semiconductors, such as AlAs/GaAs, which has been the canonical system for heteroepitaxy research [6].

As discussed earlier, given that they can indeed be grown as requested, heterostructures give device designers a great deal of flexibility in determining the characteristics, both electrical and optical, of a tailored structure. Table 2.1 lists several of the relevant properties of some of the major semiconductor elements and compounds. The use of heterostructures allows the band gap, mobilities, effective masses, and other material and transport parameters to be gradually or abruptly altered. Provided that the interfacial chemistry between two abutting semiconductors is favorable, such heterojunctions may be as sharp as a single monolayer. While the interface plane might exhibit in-plane inhomogeneities (interface roughness), in many systems it is believed that the adjoining layers consist of relatively pure monolayers of the two junction materials [7]. In addition to pure, binary semiconductor compounds, ternary and quaternary alloys of two or more binary systems are commonly grown. This practice is generally used to satisfy the structural and chemical demands of epitaxial growth, but may also be utilized to spatially grade an alloy composition so as to smoothly vary the material properties over many monolayers, rather than create abrupt heterojunctions. In addition to the controlled variation of transport parameters, the differences in band gap between the various layers of a heterostructure can lead to quantum confinement of the charge carriers; the reduced dimensionality of such a system will in turn induce radical changes in its electrical and optical properties.

2.2.2 Unique Issues

There are many issues unique to heterostructures which are not present in traditional homojunction technology. The first conceptual problem encountered when placing two dissimilar semiconductors adjacent is the ambiguous offset between their respective energy scales. While the band gaps of most semiconductors are well known, the experiments which are used to determine them, such as optical absorption or photoluminescence, are relative energy measurements; they are not

		300K Band Gap (eV)	300K Mobility (cm^2/V-sec)	Effective Mass ($m^\star/m_0$)
IV				
	Si	1.12	$\mu_e = 1500$	$m_e^\star = 0.98$
			$\mu_h = 450$	$m_{lh}^\star = 0.16$
III-V				
	AlAs	2.14	$\mu_e = 294$	$m_e^\star = 0.35$
				$m_{lh}^\star = 0.15$
	AlSb	1.58	$\mu_e = 200$	$m_e^\star = 0.39$
			$\mu_h = 420$	$m_{lh}^\star = 0.11$
	GaAs	1.42	$\mu_e = 8,500$	$m_e^\star = 0.065$
			$\mu_h = 400$	$m_{lh}^\star = 0.087$
	GaSb	0.72	$\mu_e = 5,000$	$m_e^\star = 0.049$
			$\mu_h = 850$	$m_{lh}^\star = 0.056$
	InAs	0.36	$\mu_e = 33,000$	$m_e^\star = 0.023$
			$\mu_h = 460$	$m_{lh}^\star = 0.025$
	InSb	0.17	$\mu_e = 80,000$	$m_e^\star = 0.014$
			$\mu_h = 1,250$	$m_{lh}^\star = 0.016$

Table 2.1: Room temperature band gaps, mobilities, and effective masses of selected III-V semiconductors, in comparison to silicon. Although not all of these compounds may be grown epitaxially upon one another, those that can may be used in heterostructures where the spatial variation of material parameters can be used to tailor a device for a given application. The data was compiled from a number of sources [8, 9, 10].

made with respect to a universal energy reference. As a result, although the band gaps of GaAs and AlAs are known to be 1.42 eV and 2.14 eV, respectively, that information alone does not convey the difference in energy (if any) between electrons in the valence band of AlAs, say, and that of GaAs. Band offsets are not easily predicted by theory [11], although a general rule of thumb proposed by McCaldin, McGill and Mead is that two semiconductors sharing the same anion should have, to first order, no energy difference between their respective valence band edges [12]. The known exceptions to this rule are compounds with aluminum cations, such as AlAs. This common anion rule works reasonably well, but experimental observations of band offsets are required to accurately assess the relative band alignments between components of any given heterosystem. This experimental data, such as compiled in Reference [13], is essential for the proper engineering of heterostructure devices.

As differing compounds are grown together, it is especially important that the preferential spatial ordering of the atoms of each semiconductor be compatible, both in lattice constant and crystallographic structure. If this mechanical condition is not met, the resulting crystal will at best be highly dislocated, and may even become polycrystalline or amorphous, and therefore generally of limited use. The crystallographic structures of all the III-V compounds are zincblende, and therefore compatible, but their lattice constants vary considerably. To graphically display the two parameters of immediate significance in the design of heterostructure devices, both the lattice constant data and experimental band offsets, McCaldin introduced a useful plotting paradigm [14]. Figure 2.1 is an example of such for silicon, germanium, and the non-phosphorous III-V semiconductors listed in Table 2.1. Shown are the conduction (triangles) and valence (squares) band edges for these materials plotted against their lattice constant.[2] Thus, in addition

---

[2] In addition, the symbols for the band edges may be open or filled, depending upon the n- and

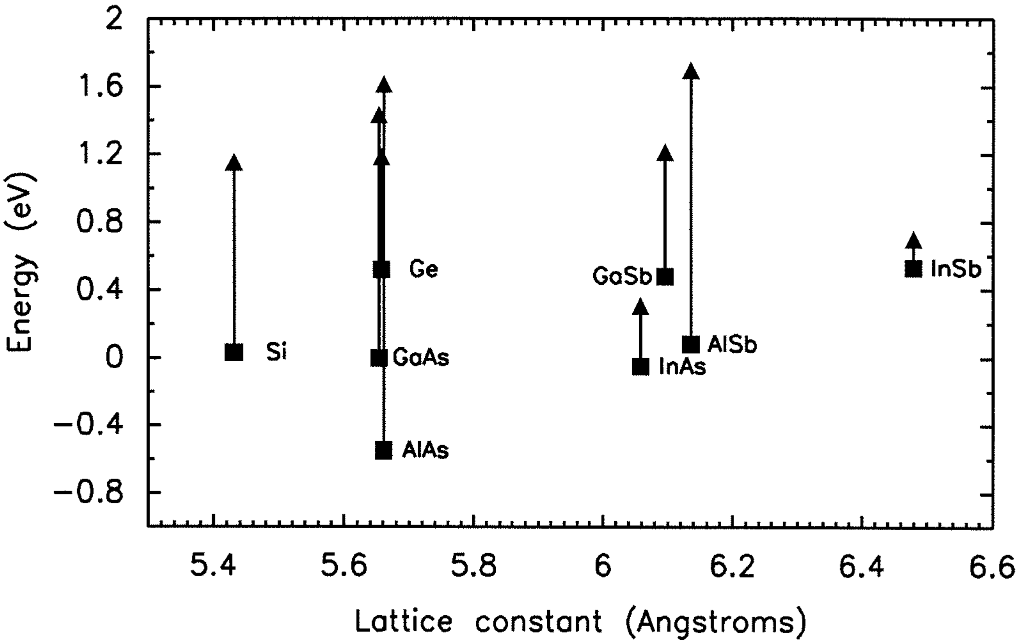


Figure 2.1: McCaldin diagram [14] showing the band alignments, lattice constants, and dopability of silicon, germanium, and the non-phosphorous III-V semiconductors. The triangular and square symbols indicate the conduction and valence band edges, respectively. Were any of these materials undopable either n- or p-type, the appropriate symbol would be left unfilled as an indication.

to identifying which combinations of materials may be grown epitaxially, the band alignments of these systems are simultaneously displayed. From this diagram, it is evident that GaAs and AlAs make a particularly well lattice-matched system. Because of their respective band offsets, AlAs can be used to confine carriers in both the conduction and valence bands of GaAs. Additionally, InAs, GaSb, and AlSb form yet another combination identified as a potential heteroepitaxial system. It is clear from the McCaldin diagram that all three are fairly well lattice-matched, and

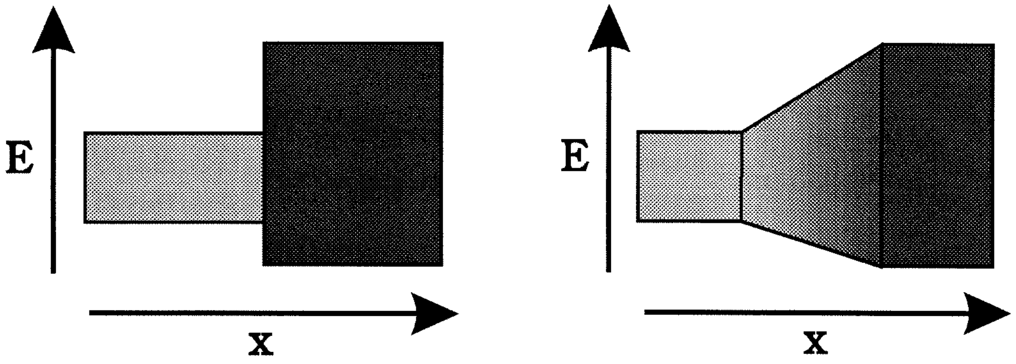
---

p-type dopability of the material, which was important for the II-VI compounds the original article addressed. The II-VI compounds introduce this additional complication for heterostructure device development, which thankfully is not an issue for III-V materials.

AlSb may be used as a barrier for carriers in all but the valence band of InAs. While some strain from mismatched lattices can be tolerated in epitaxial growth, especially for thin layers (thin being less than some experimentally-determined critical thickness), large strains (greater than a few percent difference in lattice constant) are undesirable. Thus, while the low mass and high mobility of InSb look attractive for high-speed electronic devices, its large lattice constant precludes its use in typical heteroepitaxy systems. In addition, strain can have a large effect upon the band offsets, and is often used in device design to engineer appropriate band alignments or quantum confinement.

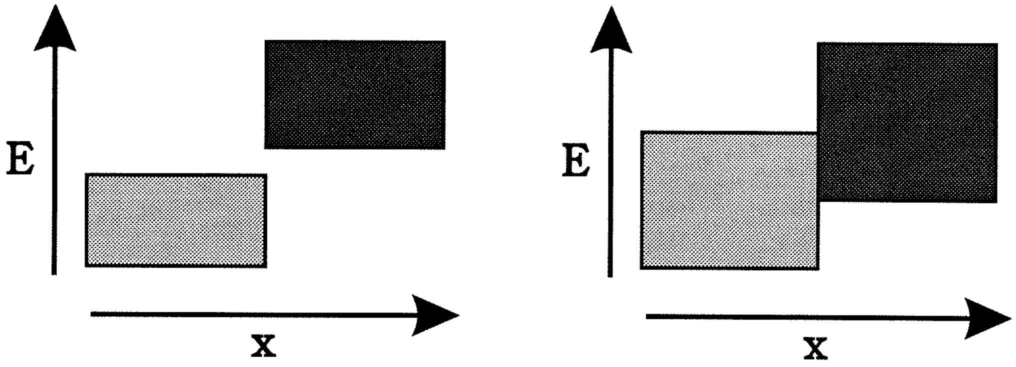
Several types of heterojunctions are possible, as characterized by the band gaps and band offsets of the two adjacent materials. Examples of all three types of offsets are evident in Figure 2.1, but are more explicitly shown in Figure 2.2, where the band gaps (shaded) of two materials are plotted as a function of position perpendicular to the junction (along the growth direction). Figure 2.2(a) shows a Type I heterointerface, in which the band alignments are such that both band edges of the smaller gap material lie between the conduction and valence band edges of the other. Such is the case for the GaAs/AlAs system. Figure 2.2(b) shows, as mentioned earlier, how the junction may be graded so as not to be abrupt. Here a spatially varying alloy concentration is used to gradually grade from one material to another. Figure 2.2(c) shows the broken-gap Type II heterojunction, in which the band gaps of the two materials do not overlap in energy at all. The InAs/GaSb system is a unique example of this interface configuration. Lastly, Figure 2.2(d) shows the remaining possibility, the staggered Type II interface. The InAs/AlSb system is representative of this type of alignment. All of the interfaces are shown under flat band conditions, where the charge transfer between layers due to the energy discontinuities is ignored. Such band bending leads to a built-in electric field at the interface, and subsequent charge accumulation or depletion. Band bending can be accounted for theoretically in most systems through a straightforward Thomas-Fermi model of the electrostatic potential [15]. The Type II broken-gap system of InAs/GaSb is quite pathological for these codes, however. A large amount of charge transfer occurs between these materials, as electrons energetically prefer InAs to GaSb, and there is no barrier to their diffusion, and likewise for holes in GaSb. Consequently, numerical models must be more realistic, and solution algorithms computationally robust. As a general rule, these requirements are necessary for all accurate models of the InAs/GaSb system.

Having said what criteria are important to the design of heterostructure devices, several of the most important types may now be introduced. Figure 2.3 shows the three most prevalent classes of quantum heterostructure devices. In choosing these to discuss, we are deliberately ignoring such devices as heterojunction bipolar transistors (HBT's) and high electron mobility transistors (HEMT's) which may be viewed as heterojunction-based optimizations of normal bipolar and field-effect transistor technology. All three heterostructure device types shown in Figure 2.3 make use of the variation of band gap between the heterolayers to quantize carriers and form two-dimensional (2D) or quasi two-dimensional carrier gases. In the quantum well, shown in Figure 2.3(a), carriers in a central well layer having typical dimensions on the order of the DeBroglie wavelength of the carriers, having an energy between the band edges of the two materials, are spatially confined by the absence of allowed energy states in the barrier layers. Consequently, one or more quantized levels are formed within the well. Quantum wells are important optical systems, and are also used to form channels for in-plane 2D conduction. In contrast, in the double barrier heterostructure (DBH), shown in Figure 2.3(b), well states are only partially confined to the well, and as such are only quasi-2D, so long as the barriers are thin enough for the carrier wavefunctions to commu-



a) Type I Interface

b) Type I Graded Junction



c) Type II Broken-gap Interface

d) Type II Staggered Interface

Figure 2.2: The possible types of heterojunction band alignments under flat band conditions. The Type I and staggered Type II junctions are common interface alignments. The InAs/GaSb system is the unique example among the common semiconductor systems of the broken-gap Type II junction. As a reminder that all heterojunctions need not be abrupt, the graded Type I junction is shown as well.

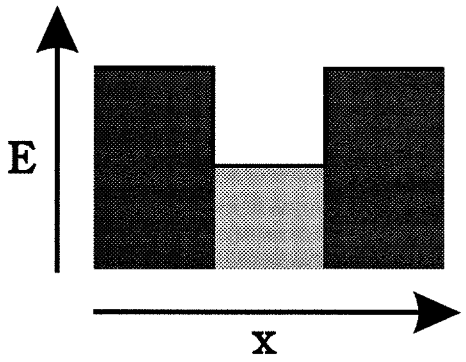
nicate with the electrode material.[3] Perpendicular transport through a double barrier heterostructure involves quantum mechanical tunneling of carriers from the emitter electrode, through quasi-bound well states, to the collector. Resonant tunneling diodes (RTD's) rely upon this heterojunction design to preferentially allow the conduction of carriers at the quasi-bound energies. The superlattice, as demonstrated in Figure 2.3, is a heterosystem in which layers of two materials are symmetrically alternated, defining a long-range, artificial periodicity to the device. This additional periodicity (superperiodicity) results in the creation of minibands, much as the lattice periodicity leads to bulk band structure. Superlattices are important optical systems, as the absorption characteristics may be finely tuned to meet specific operational requirements. While Figure 2.3 shows these devices for Type I heterojunctions, similar devices are possible with Type II offsets. The overlap of conduction and valence band states in Type II broken-gap structures excludes the possibility of a truly confining quantum well in this system.

2.2.3 Electronic Properties

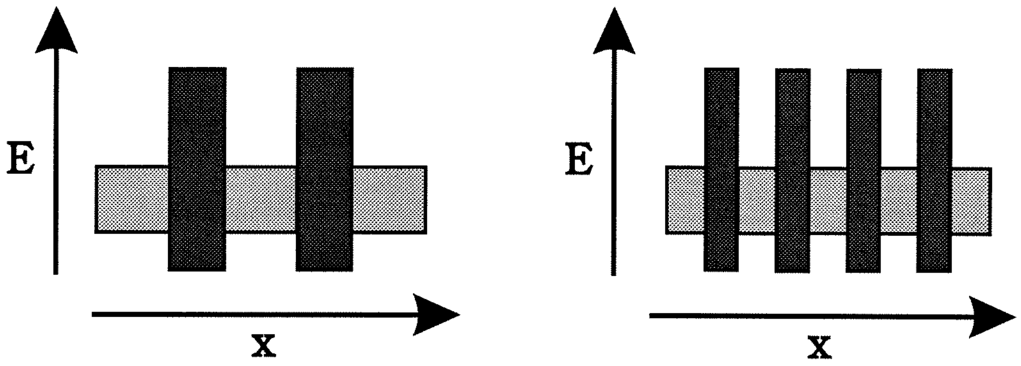
The proper understanding of the quantum mechanical behavior of carriers in heterostructures is essential to the discussion of the electrical properties of these devices. At the simplest level, the variation of band edges from material to material provide classic piecewise linear quantum mechanics problems. Thus, the physics is very accessible. However, additional factors, such as changes in effective mass, multi-dimensional effects, and realistic band structure, provide complications which make the field far richer than any textbook model. As a result, we

---

[3] In the limit that the barriers are thick enough that essentially none of the well wavefunction leaks out to the electrodes, we recover the simple quantum well. Depending upon the barrier height, typically several hundreds of Ångstroms of barrier material must exist, at a minimum, for such a distinction.



a) Quantum Well



b) Double Barrier Heterostructure

c) Superlattice

Figure 2.3: The three major classes of heterostructure devices, showing the band edges as a function of position. The band gap regions of the two constituent materials are shaded. Quantized states in the quantum well, (a), are entirely confined within the well material, while coupling to the electrode regions of the double barrier heterostructure, (b), and between wells of a superlattice, (c), leads to partially-confined quasi-bound states.

first discuss the basics of quantum transport in these systems for Type I unipolar devices, where the physics is more straightforward, and then contrast Type II interband device transport. We will furthermore focus our discussions on resonant tunneling diodes. The physics of quantum wells will be presented as a motivation for the development of RTD's, but superlattices and other heterojunction devices will not be discussed.

Type I Unipolar Devices

To begin, we will treat Type I unipolar devices with a simple, single-band model. For such, the energy dispersion of the bulk materials is assumed to be quadratic, and described by a single effective mass parameter, such that,

$$E(k_x, k_y, k_z) = \frac{\hbar^2}{2m^\star}(k_x^2 + k_y^2 + k_z^2), \tag{2.1}$$

where $m^\star$ is the effective mass. For heterostructures, it is assumed that the bulk effective mass parameters are equally valid for thin layers of material, before the addition of quantization. We treat a quantum well as a central layer of width L lodged between two semi-infinite slabs of cladding barrier material with a positive conduction band edge discontinuity of ΔE from the $E = 0$ band edge of the well. For $0 \leq E_n < \Delta E$, we can calculate the quantized energies in a quantum well, and the wavefunctions, by in addition assuming that the total carrier wavefunction is separable into a Bloch component with the periodicity of the underlying lattice, and a smoothly varying envelope function[16, 17] such that,

$$\psi = \sum_j e^{i\mathbf{k}_{\|}\cdot\mathbf{r}}\, u^j(\mathbf{r})\, \phi_n(x), \tag{2.2}$$

where $\mathbf{k}_{\|}$ is the in-plane wavevector, $u^j(\mathbf{r})$ is the Bloch function of the material indexed by j, and $\phi_n(x)$ is the envelope function for the nth quantized state wavefunction along the growth direction, x. Due to the quantization in x, the wavevector k_x is replaced with the quantum index n. To further simplify the problem, it is usually also assumed that the Bloch periodic functions for the materials of a heterostructure are identical. While this approximation is reasonable, given the similar symmetry properties and material characteristics of semiconductor compounds that are typically grown together, if the problem is solved with more complete generality, deviations from this approximation are observed [18]. Under these assumptions, the Hamiltonian acting on such a wavefunction is itself separable into a Bloch function component dealing with the wavefunction variations over a single lattice site, and therefore containing the physics at an atomic level, and the piecewise linear envelope function Hamiltonian incorporating the large-scale variations due to the heterostructure,

$$-\frac{\hbar^2}{2m_i^\star}\frac{\partial^2\phi_n(x)}{\partial x^2}+V(x)\phi_n(x)=E_n\phi_n(x), \tag{2.3}$$

where $V(x)$ now represents the potential energy variation along the growth axis due to the band edge discontinuities of the heterostructure,

$$\begin{aligned} V(x) &= \Delta E, \quad |x| \geq L/2 \quad \text{(Barrier)} \\ &= 0, \qquad |x| < L/2 \quad \text{(Well)}, \end{aligned} \tag{2.4}$$

for the quantum well with $k_{||}=0$. Because we are interested in the macroscopic quantum properties of the entire heterosystem, the Bloch components and their Hamiltonian are typically ignored. The envelope function Hamiltonian is itself dependent upon position along the growth axis, both through $V(x)$ and the effective masses, and thus changes at a heterointerface. Aside from this complexity, the problem has reduced to a textbook piecewise linear potential problem, and the solution method is identical. Because of the change in effective mass between the different materials, the boundary conditions set by probability current density conservation are,

$$\phi_n^L(x_{int}) \quad = \quad \phi_n^R(x_{int}), \quad \text{and,} \tag{2.5}$$

$$\frac{1}{m_L^\star}\frac{\partial \phi_n^L}{\partial x}(x_{int}) \;=\; \frac{1}{m_R^\star}\frac{\partial \phi_n^R}{\partial x}(x_{int}), \tag{2.6}$$

where L and R refer to the left and right component materials of the interface at position x_{int}. Thus, as in the classic textbook problem, we expect bound states with sinusoidal wavefunction behavior in the well and exponential decay in the barrier materials, as shown in Figure 2.4(a). The carrier described by the wavefunction is completely localized in the well; it cannot classically exist outside the well, and since the cladding layers are semi-infinite, it cannot tunnel out quantum mechanically. Following the textbook solution [19], the stationary state wavefunctions must be of even or odd parity, having a sinusoidal period of $1/k_\perp$, where $\hbar k_\perp = \sqrt{2m_W^\star E_n}$, $m_W^\star$ is the bulk effective carrier mass of the well material, and, similarly, $m_B^\star$ is the bulk effective mass of the barrier. The decay constant in the barriers is given by $\kappa = (1/\hbar)\sqrt{2m_B^\star(\Delta E - E_n)}$. Even solutions exist for,

$$(k_\perp/m_W^\star)\tan(k_\perp L/2) = \kappa/m_B^\star, \tag{2.7}$$

and odd solutions for,

$$(k_\perp/m_W^\star)\cot(k_\perp L/2) = \kappa/m_B^\star. \tag{2.8}$$

Solutions to these boundary-imposed equations yield valid values for $k_\perp$, and therefore the eigenvalues E_n. Unlike the textbook problem, however, quantum wells fabricated from semiconductors are three dimensional, and the potential profile along the x-axis is dependent upon the parallel crystal momentum, $k_\|$,

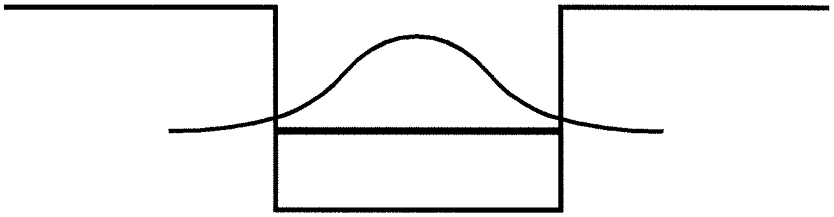
$$\begin{aligned} V(x) &= \Delta E + \frac{\hbar^2 k_\|^2}{2m_B^\star}, \quad |x| \geq L/2 \quad \text{(Barrier)} \\ &= \frac{\hbar^2 k_\|^2}{2m_W^\star}, \qquad\quad |x| < L/2 \quad \text{(Well)}, \end{aligned} \tag{2.9}$$

and as a result, the barrier seen by carriers at arbitrary parallel momentum is given by,

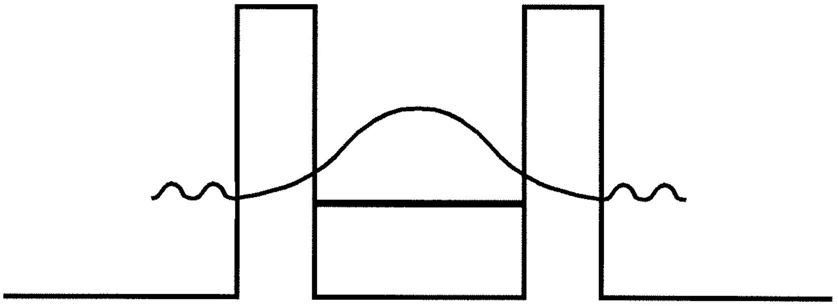
$$\Delta E' = \Delta E - \frac{\hbar^2 k_\|^2}{2m_B^\star}\left(\frac{m_B^\star}{m_W^\star} - 1\right). \tag{2.10}$$

Using this $\Delta E'$ in Equations 2.7 and 2.8 yields the equivalent $k_{\|}$-dependent solutions. Because the effective masses for large band gap materials are typically larger than those with smaller band gaps, the effect of non-zero $k_{\|}$ is to reduce the quantization of the carriers with respect to the $k_{\|} = 0$ problem. However, as the effective mass approximation is only valid for small wavevectors, this model can only treat small perturbations about $k = 0$. A further complication exists for the quantization of holes. Because the valence band exhibits $J = \frac{3}{2}$ angular momentum symmetry, a simple one-band model is usually insufficient to adequately describe the quantum behavior of hole states. As discussed in Appendix B, the dispersion about $k = 0$ must be expanded as a function of the J_x, J_y, and J_z angular momentum operator matrices, following the model of Luttinger [20]. This additional complication leads to multiple hole subbands, with mixing between the states at finite $k_{\|}$. The resulting problem is not analytical, in general, and Equation 2.3 must be solved numerically as a piecewise linear differential equation.

In contrast to the quantum well, the double barrier heterostructure does not have wholly confined states, as shown in Figure 2.4. A carrier in the well has a wavefunction which penetrates through the narrow barrier regions into the semi-infinite electrodes. As a consequence of these open boundary conditions, the problem is no longer a discrete eigenvalue and eigenvector problem for the quantized energies and wavefunctions, respectively. The behavior of the DBH must be considered in the context of perpendicular quantum transport though the active region containing the well and barriers. Energy- and momentum-dependent transmission coefficients for carriers incident from the emitter electrode may be calculated using standard quantum mechanics. We motivate the standard calculations to distinguish concepts and physical behavior that will be important in further discussions in this thesis, but leave the details to the reference texts [19, 10]. We define our geometry and energies in a manner similar to the discussion of quantum wells, as



a) Quantum Well



b) Double Barrier Heterostructure (Resonant Tunneling Diode)

Figure 2.4: Comparison of bound state wave functions in a quantum well and double barrier heterostructure. The eigenstates in a quantum well are fully confined in the well region, with only partial leakage of the wave function into the semi-infinite barrier layers. In contrast, the quasi-bound well states of a DBH communicate quantum mechanically with the unbound electrode states through the thin barriers.

shown in Figure 2.5, and initially consider the problem under flat band conditions. The electrode and well band edges are assumed to be at $E = 0$, and the barrier band edges ΔE greater. The well is taken to be of width W and the barriers each width a. As for the quantum well, we define the oscillation wavevector and decay constant, respectively, as $\hbar k_\perp = \sqrt{2m_E^\star E_\perp}$ and $\hbar\kappa = \sqrt{2m_B^\star(\Delta E - E_\perp)}$, where $m_E^\star$ is now the effective mass in both the well and the electrodes, and $m_B^\star$ is the effective mass in the barriers. Considering just the first heterointerface, we see that the wavefunction on either side may be written as a sum of positive and negative exponentials. Where the carrier energy is greater than the band edge of the material these are complex exponentials, and therefore traveling waves incident upon, and reflected from, the DBH, respectively. In the barrier, for carrier energies in the band gap, the exponentials are real and consist of growing and decaying states. As the barrier is of finite thickness, the growing exponential states cannot be neglected, as was the case for the quantum well. Consequently, we have at the first interface,

$$\begin{aligned} \phi(x) &= Ae^{ik_\perp x} + Be^{-ik_\perp x}, \quad &&\text{(Electrode)} \\ &= Ce^{-\kappa x} + De^{\kappa x}, &&\text{(Barrier)}. \end{aligned} \tag{2.11}$$

Matching these at the boundary $x = -a/2$, using Equations 2.6, the coefficients of the exponentials may be related as,

$$\begin{pmatrix} A \\ B \end{pmatrix} = \frac{1}{2}\begin{bmatrix} \left(1 + \frac{i\kappa m_E^\star}{k_\perp m_B^\star}\right) e^{\kappa a/2 + ik_\perp a} & \left(1 - \frac{i\kappa m_E^\star}{k_\perp m_B^\star}\right) e^{-\kappa a/2 + ik_\perp a} \\ \left(1 - \frac{i\kappa m_E^\star}{k_\perp m_B^\star}\right) e^{\kappa a/2 - ik_\perp a} & \left(1 + \frac{i\kappa m_E^\star}{k_\perp m_B^\star}\right) e^{-\kappa a/2 - ik_\perp a} \end{bmatrix}\begin{pmatrix} C \\ D \end{pmatrix}. \tag{2.12}$$

If this matrix is inverted, we can express the rightmost wavefunction coefficients as a function of the leftmost,

$$\begin{pmatrix} C \\ D \end{pmatrix} = \mathbf{T}\begin{pmatrix} A \\ B \end{pmatrix}, \tag{2.13}$$

where $\mathbf{T}$ is the transfer matrix linking the wavefunction across the interface. In general, a transfer matrix is necessary to relate the wavefunction across any change in band edge energy, as indicated by the arrows in Figure 2.5. Thus, in a realistic device with band bending, the full transfer matrix of the active region is given by,

$$\mathcal{T} = \prod_i \mathbf{T}_i, \tag{2.14}$$

where each $\mathbf{T}_i$ is a transfer matrix linking adjacent lattice sites at position i. Typically this product must be extend over not only the active quantum heterostructure region, but the regions of the electrodes for which there is significant band bending. Thus, even for an active region a couple hundred Å's in width, typically a half micron or more of material must be included in the calculation. Because the transfer matrix consists of exponentials of the product of very large ($k_\perp$ and κ) and very small (a) numbers, the transfer matrix method, while easy to describe, and intuitive, is numerically unstable and seldom used for device modeling. Using this transfer matrix method analytically, however, it can be shown [10] that the overall transmission coefficient for carriers incident on a symmetric, flat band Type I DBH is given by,

$$\mathcal{T}_{tot} = \frac{\mathcal{T}^2}{|1 + |r|^2 e^{2i(kL+\omega)}|^2}, \tag{2.15}$$

where $\mathcal{T}_{tot}$ is the total transmission coefficient for the DBH, $\mathcal{T}$ is the transmission coefficient for one of the barriers, $|r|^2 \equiv R$ is the complex reflection coefficient for a single barrier, L is the well width, and $\omega = \psi - ka$, where ψ is the transcendental solution to the following:

$$\tan\psi = \frac{1}{2}\left(\frac{km_B^\star}{\kappa m_E^\star} - \frac{\kappa m_E^\star}{km_B^\star}\right)\tanh\kappa a. \tag{2.16}$$

The most important and relevant fact to be gleaned from Equation 2.15 is its behavior as a function of incident energy, shown in Figure 2.6 as calculated by the transfer matrix method for a flat band InAs/AlSb DBH with 36 Å barriers and a 90 Å well. For most values of incident energy, the total transmission coefficient is roughly the square of that for a single barrier; the incident carrier is transmitted as if the barriers and well were absent and a single barrier of twice the width $(2a)$ were instead present. At certain energies, however, the transmission coefficient is unity, and thus the barriers become completely transparent. This condition occurs when the exponential in the denominator of the expression for T_{tot} is -1, or $2(kL + \omega) = (2n + 1)\pi$. At these energies, the carriers resonate with quasi-bound states in the well, and the two electrodes are completely coupled through a resonant tunneling process. In the limit as the barrier widths become infinitely thick, the carriers in the well are completely confined, and the problem reduces to that of the quantum well, with the resonant quasi-bound energies approaching those of the true eigenvalues of an equivalent quantum well.

The addition of $k_{||}$ to the quantum mechanical treatment of a DBH has an effect similar to that described for quantum wells. Because it is a transport problem, however, the additional restriction that $k_{||}$ be conserved must be added. Thus, for resonant tunneling, not only must the incident energy match that of the quasi-bound state, but the parallel wavevectors of the incident electrode state and the final well state into which it tunnels must be identical, and similarly for tunneling from the well to the collector. In addition, the width of the transmission resonances are related to the time required for tunneling to occur through the uncertainty principle, $\delta\tau \ \delta E \approx \hbar$, where $\delta\tau$ is the tunneling time, and δE is the width of the transmission resonance at $T_{tot} = 0.5$. Hole tunneling is again similar, albeit more complex, and involves additional conservation criteria for orbital angular momentum. Coupling between angular momentum states at non-zero $k_{||}$, however, alleviates this restriction through the mixing of hole angular momentum states.

The current-voltage (I-V) characteristic is the single-most important transport property of an RTD. Applying a bias between the emitter and collector of a DBH RTD creates a non-equilibrium carrier distribution in which emitter states are of greater energy than those in the collector. As a result, current flows between the

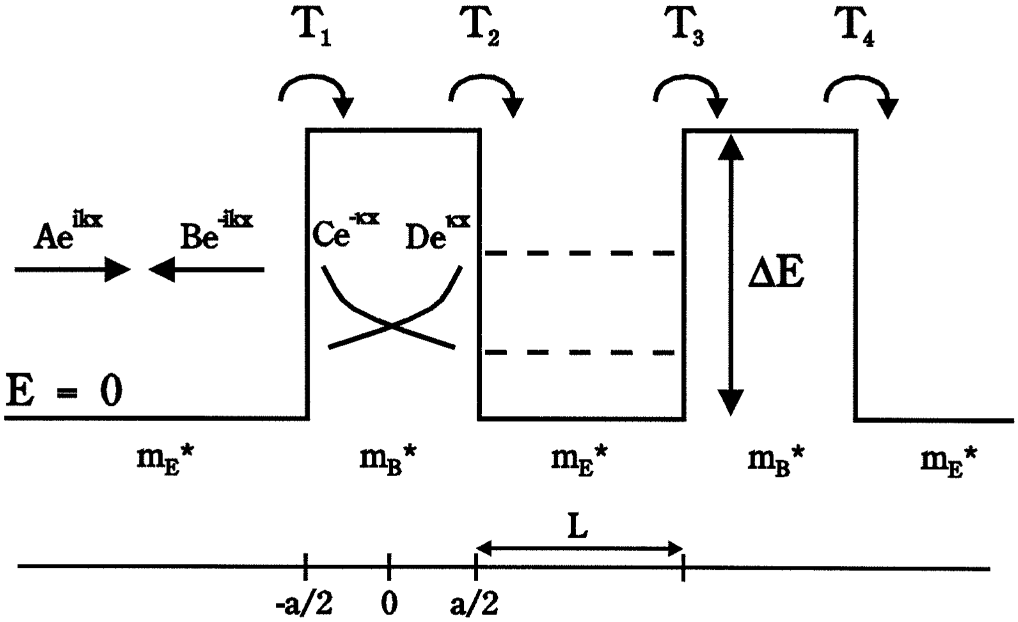


Figure 2.5: Details of the transfer matrix calculation for the transmission coefficient of a double barrier RTD. Incident and reflected states are shown in the emitter region, and exponentially growing and decaying solutions within the first barrier. The quasi-bound states leading to enhanced resonant transport across the device are shown in the well. At each change in potential energy seen by a tunneling electron, a transfer matrix relates the amplitudes of the wavefunction on either side of the interface, as shown. The transfer matrix for transport across the entire device is the product of the individual transfer matrices.

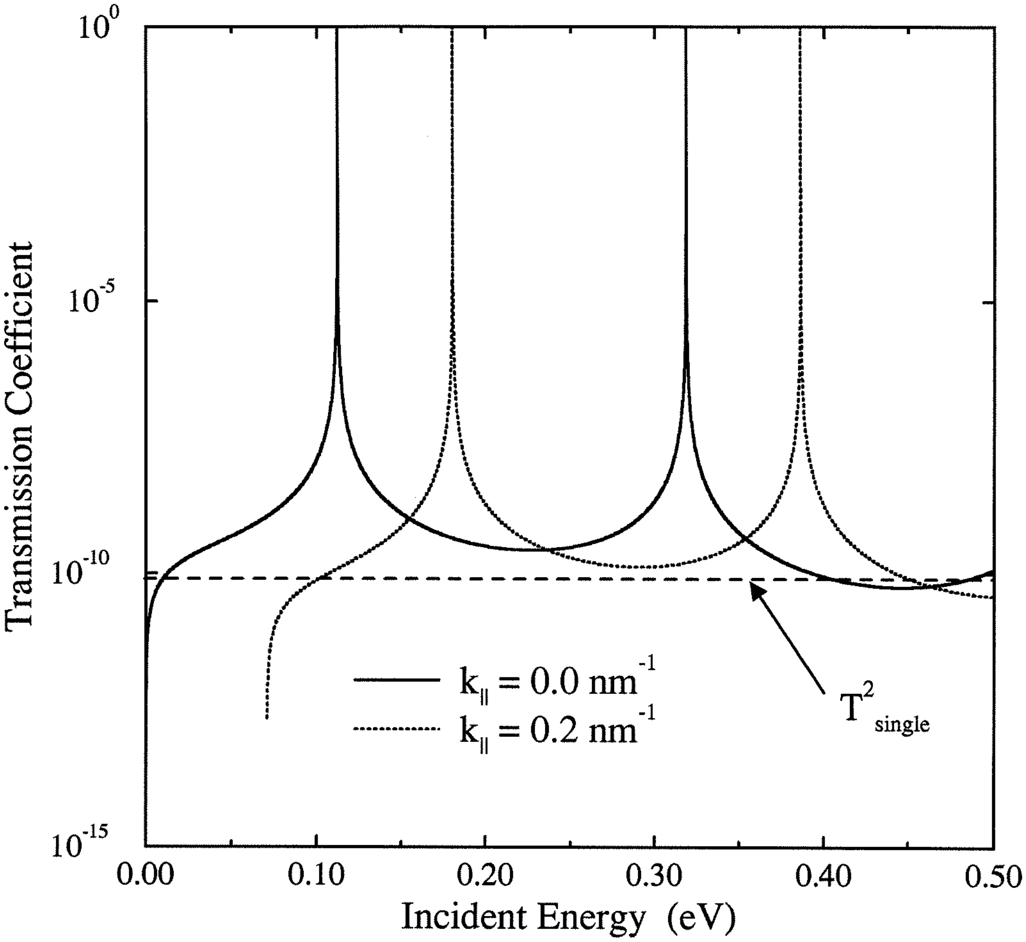


Figure 2.6: Transmission coefficient of an InAs/AlSb double barrier structure having 36 Å barriers and a 90 Å well, calculated as a function of incident energy by the transfer matrix method. The transmission coefficient of the entire device is roughly the square of the transmission coefficient, T_{single}, of a single barrier at all but several isolated energies where the barriers are transparent to the incident carrier. The results are shown for both zero and non-zero $k_{\|}$. The non-zero $k_{\|}(=\Delta k_{\|})$ transmission curve is essentially the $k_{\|}=0$ curve shifted in energy by $\hbar^2\Delta k_{\|}^2/2m_W^\star$.

electrodes as allowed through the limiting active quantum region. As was evident in the flat band, zero-bias analysis of transmission through a Type I unipolar DBH (c.f. Figure 2.6), current transport through such a device has a strong dependence upon the energies of the carriers relative to the quasi-bound states (QBS's) in the well. Consequently, as the bias is increased across the RTD, the magnitude of the current ebbs and flows with the degree to which resonant quantum transmission through the well states is allowed. In Type I unipolar devices, such as GaAs/AlAs or InAs/AlSb double barrier structures, the net current through the structure therefore exhibits a peak at those voltages which cause the preponderance of incident carriers to align in energy with a quasi-bound state in the well. At biases greater than this peak voltage, the transmission coefficient drops precipitously, and the current likewise reaches a minimum. The region over which the current drops results in a negative differential resistance (NDR), which is an inherent property of resonant tunneling devices. A typical I-V curve is shown at the top of Figure 2.7. The ratio of peak and valley currents, or peak-to-valley ratio (PVR), is the single figure of merit quoted to summarize the NDR behavior of a given quantum device. Much of the current transmission in the valley region is experimentally linked to non-resonant, inelastic transport across the structure. Consequently, the PVR is indicative of the degree to which inelastic processes compete with resonant quantum transport to yield the net current. Other parameters of the I-V characteristic which are of importance for the technological application of RTD's are the linearity of the onset regime,[4] the absolute magnitude of the peak current density, and the

---

[4] As indicated in Figure 2.7, the onset of most Type I RTD I-V characteristics is, in fact, exponential, due to the Fermi distribution of emitter carriers. Such behavior should be contrasted with the onset of a resonant interband tunneling (RIT) diode, discussed in Section 2.3 and the subject of much of the experimental work in this thesis. Such a distinction is of great importance in the operation of some classes of systems based upon RTD's.

width of the NDR region.

The theoretical understanding of I-V characteristics in RTD's may be developed on the basis of early work by Tsu and Esaki [21]. The current density can clearly be seen to depend upon two quantities: the transmission probability of a device and the population and distribution of carriers in the electrodes. Each of these, in turn, is dependent upon E_x, the energy due to kinetic motion along the growth direction. The current from emitter to collector, j_{EC}, is given by,

$$j_{EC} = e \sum_{k_x} (\hbar k_x / m_E^\star)\, T_{EC}(E_x)\, N_E(E_x), \tag{2.17}$$

where $\hbar k_x / m_E^\star$ is the semi-classical carrier velocity in the emitter, $T_{EC}(E_x)$ is the emitter-to-collector transmission coefficient, and $N_E(E_x)$ is the carrier population in the emitter at energy E_x. Similarly, the current from emitter to collector, which must in general be included, is,

$$j_{CE} = e \sum_{k_x} (\hbar k_x / m_E^\star)\, T_{CE}(E_x)\, N_C(E_x). \tag{2.18}$$

For a spatially symmetric device, $T(E_x) \equiv T_{EC}(E_x) = T_{CE}(E_x)$, and the total current is given by,

$$j_T \equiv j_{EC} - j_{CE} = \frac{e\hbar k_x}{m_E^\star} \sum_{k_x} T(E_x)\ (N_E(E_x) - N_C(E_x))\,. \tag{2.19}$$

Fermi statistics yields expressions for the electrode populations functions,

$$N_E(E_x) = \frac{kTm_E^\star}{\pi\hbar^2} \ln\left[1 + \exp\left(\frac{E_F - E_x}{kT}\right)\right], \tag{2.20}$$

and,

$$N_C(E_x) = \frac{kTm_E^\star}{\pi\hbar^2} \ln\left[1 + \exp\left(\frac{E_F - E_x - eV}{kT}\right)\right], \tag{2.21}$$

where V is the applied bias. Given accurate transmission coefficients, Equation 2.19 generally predicts peak voltages and currents which, when adjusted for series resistance, agree reasonably well with experiment. Equation 2.19 does not include non-resonant processes, and thus grossly underestimates the valley current. This theory also makes the assumption that the transmission depends only upon the perpendicular motion, and thus all states with a given E_x are treated identically, despite any differences in $k_{||}$. While Figure 2.6 shows that the transmission characteristics of carriers with identical *total* energies, but differing parallel wavevectors, are quite distinct, the assumption in the Tsu and Esaki theory of I-V characteristics is that the transmission is identical if the perpendicular energies are equal. The validity of this approximation is evident in Figure 2.6 as the $k_{||} = 0$ curve can be translated along the x-axis by an amount equal to $\hbar^2 \Delta k_{||}^2 / 2m_E^*$ to lie roughly on top of the $k_{||} = \Delta k_{||}$ transmission curve. This assumption is valid for Type I unipolar conduction band tunneling only. Since additional $k_{||}$ increases the coupling of $m_j = \frac{1}{2}$ and $m_j = \frac{3}{2}$ states, the simple Tsu/Esaki assumption is not valid for hole tunneling and Type II interband tunneling involving conduction and valence band states.

The evolution of a Type I unipolar I-V curve is best developed through an understanding of the process via the simultaneous conservation of energy and momentum. Figure 2.7 shows a sample Type I unipolar (intraband, to contrast it from the later discussions of InAs/AlSb/GaSb interband devices) I-V characteristic. Three points along it are labeled, and the emitter and well in-plane dispersions at these three voltages are detailed in the lower panels. The degree of carrier occupation from Fermi statistics is shown via shading. It is easiest to first consider 3D-2D tunneling involving a continuum of states in the emitter, where the lowest such state, corresponding to the band edge, is shown in the figure. In such a situation, the number of states available for resonant, elastic tunneling in which both energy and parallel wavevector are conserved increases exponentially with applied bias. At voltages less than the peak voltage, the well state band edge (at $k_{||} = 0$) lies higher in energy than the emitter. As the well is brought lower in energy by

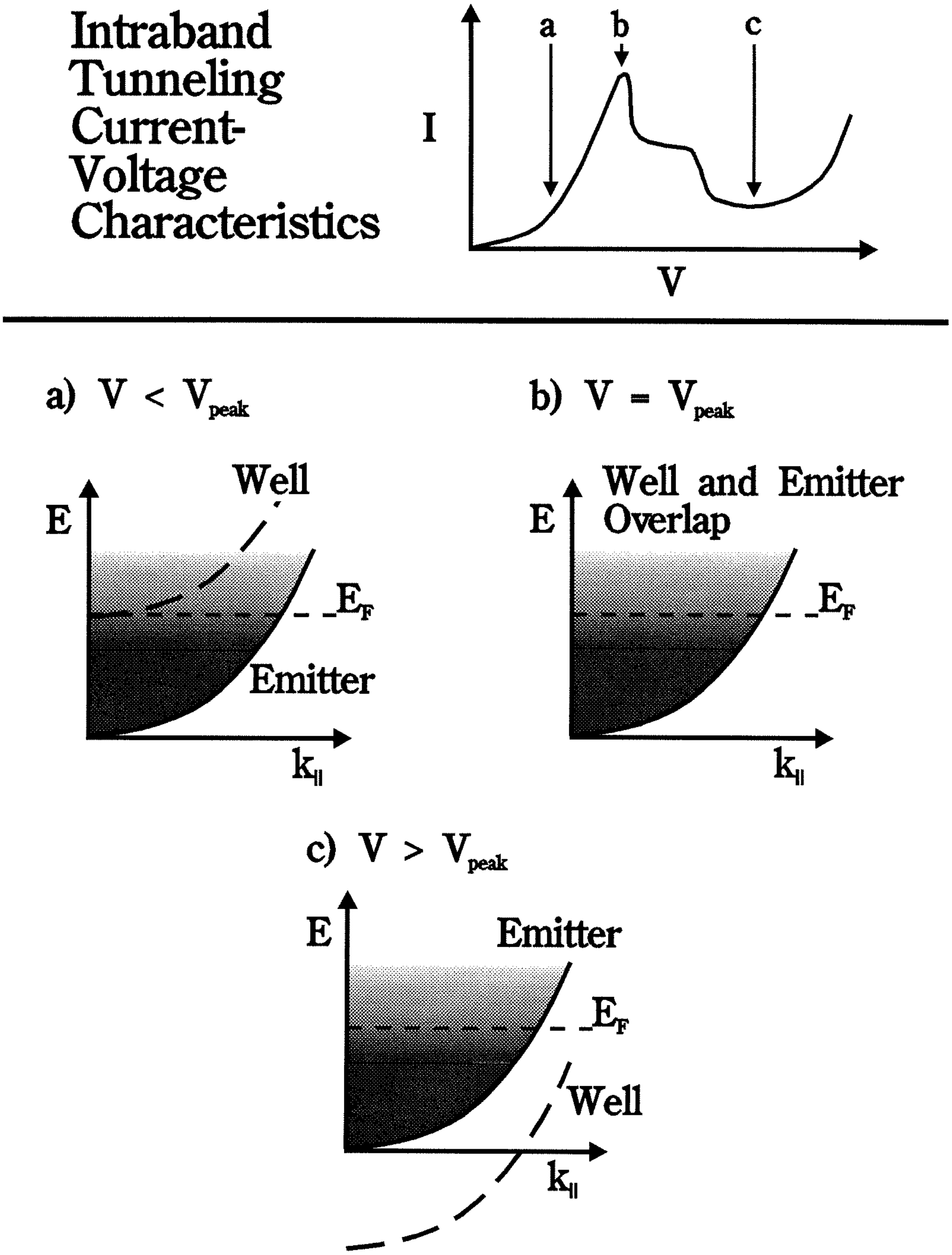


Figure 2.7: The development of a Type I unipolar device I-V characteristic. Shown in the top panel is a representative I-V curve, and beneath, the emitter and well band alignments at each of the three indicated voltages.

the potential difference, it approaches the Fermi level, E_F, and the carrier occupation in the Fermi tail changes as the exponential of the difference between the band edge energy and the Fermi energy. Thus the exponential onset characteristic of intraband tunneling. The peak in the I-V curve must occur when the well and emitter dispersions overlap entirely, and the greatest number of carriers are accessible for elastic transport. At voltages greater than V_{peak}, the well lies below the emitter continuum at all $k_{||}$ and there are no states available to elastically tunnel. As mentioned earlier, however, inelastic process contribute greatly to the current in this regime. At very low temperatures, and also in the case of 2D-2D tunneling where there is a notch state in the emitter accumulation region, inelastic processes also dominate for $V \ll V_{peak}$ (panel (a) in the figure). However, the availability of carriers to inelastically tunnel is *also* limited by the Fermi distribution, and the onset is nonetheless exponential.

2.3 The InAs/AlSb/GaSb Material System

The InAs/AlSb/GaSb material system provides an unique family of semiconductor compounds for the study of quantum heterostructure science and the potential for heterostructure-based technologies. Figure 2.8 shows the band gaps of all three materials, with the conduction and valence bands indicated, and their relative energy alignments, all at room temperature. The InAs/AlSb/GaSb material system exhibits all three band offset possibilities: Type I (InAs/AlSb), staggered Type II (GaSb/AlSb), and broken-gap Type II (InAs/GaSb). It is this latter alignment which makes the system unique, and the subject of a long history of optical and electrical studies. As indicated in the figure, the valence band edge of GaSb lies 0.15 eV above the conduction band edge of InAs [22]. This band alignment results in peculiar heterostructure transport involving conduction band-like states in InAs and valence band-like states in GaSb. Although the unique properties of InAs/GaSb heterostructures have been known for some time [23], only recently were tunnel diodes based upon this material family proposed [24], and subsequently exploited in the creation of a new class of quantum transport device [25, 26]. While the nature of the InAs/GaSb band offset is unique and distinctive, the viability of this system for heterostructure physics and technology is greatly increased by the availability of a nearly lattice-matched barrier material (AlSb) for both holes in GaSb and electrons in InAs. While there is more strain in this system (the lattice constant of GaSb is roughly 0.5% greater than that of InAs, and the AlSb lattice constant is about 1.0% larger still than GaSb) than the near-perfect match of GaAs and AlAs, this trio combine to create one of only a few possible III-V binary compound heterostructure families. One potential drawback is the lack of confinement symmetry provided for by the AlSb barrier. Fortunately, AlSb does provide a barrier for the majority carrier-type in each material, and provides greater confinement for the low mass InAs electrons. However, the valence band of InAs can couple quite strongly to that of GaSb, increasing quantum-mechanical communication between the two across an AlSb barrier, and providing demonstrable effects on the transport of even InAs/AlSb Type I unipolar devices [26]. It has been proposed [27] that strained $\mathrm{In}_x\mathrm{Al}_{1-x}\mathrm{As}$ barrier layers may be used to increase confinement in the valence band of these structures.

Figure 2.9 details the interband tunneling equivalent of the double barrier structure discussed in Section 2.2. Shown is the p-well (so named due to the valence band nature of the GaSb well) resonant interband tunneling (RIT) diode, with InAs electrodes, AlSb barriers, and a GaSb well. Transport through this device involves resonant tunneling of electrons from the InAs emitter, through unoccupied electron states in the subbands of the GaSb well, and subsequently back into the conduction band of the collector. The converse device, with an InAs well

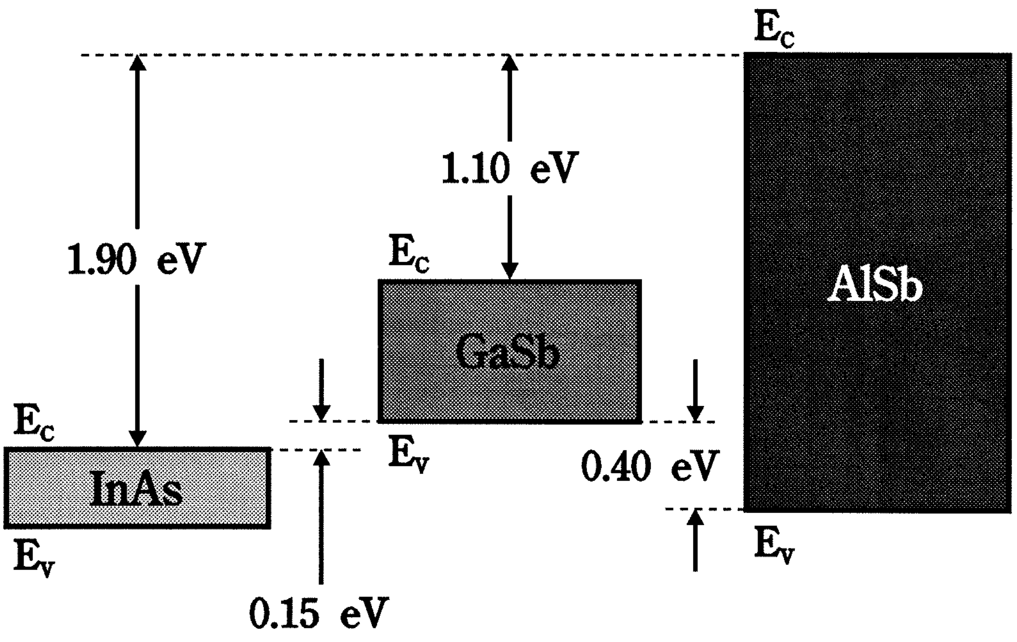


Figure 2.8: Band alignments in the InAs/GaSb/AlSb material system, with the band gap of each material shaded for clarity.

and GaSb electrodes, has been studied as well. The availability of three materials, and the variety of band offsets between them, has produced more than 10 unique heterostructure configurations in this material system, all of which have been fabricated and studied [26]. The unique properties of this material system and interband tunneling devices will be presented in reference to the p-well RIT alone, however, for simplicity. The transport properties of this device, and the applications which result, were the primary motivation behind the research contained in this thesis.

2.3.1 Transport Properties

As indicated previously, transport in interband devices such as the RIT differs significantly from that of Type I unipolar devices. The added complexity of transport

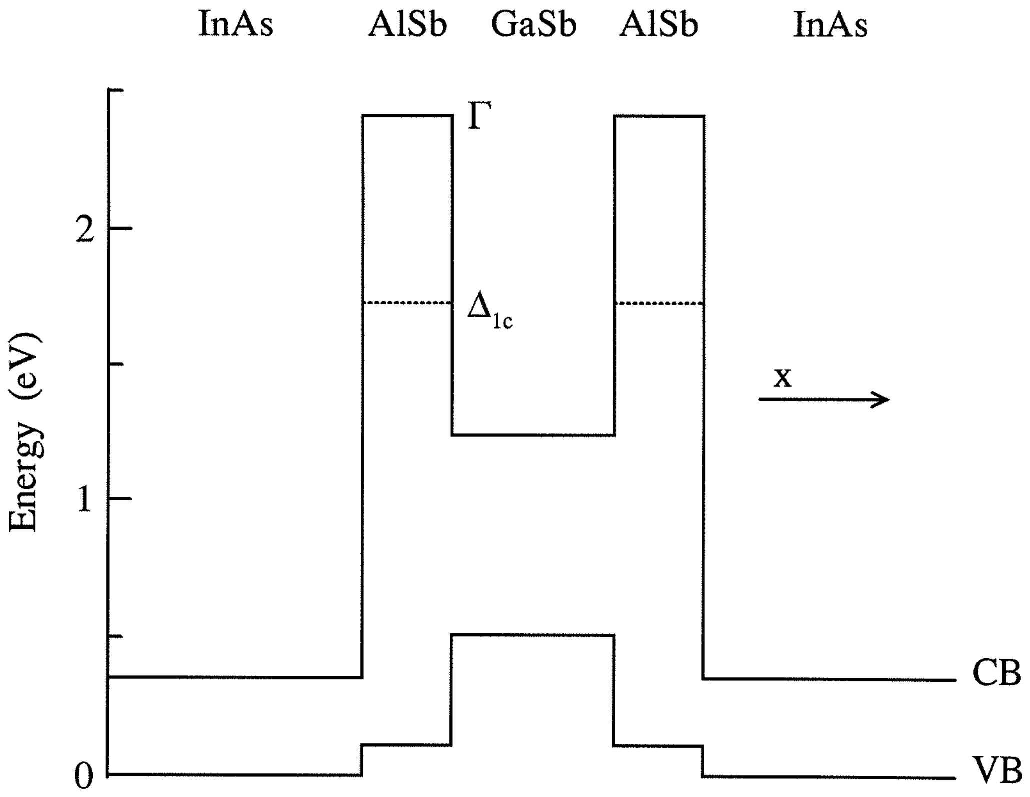


Figure 2.9: Band diagram of a p-well RIT, showing the conduction and valence band edges as a function of position. Quantum transport through this device involves electrons from the emitter conduction band tunneling into empty valence band states in the well, and then out into the conduction band of the collector.

in these structures has both theoretical and experimental implications. Theoretically, quantum transport calculations of interband systems require sophisticated multiband treatments; 1- or 2-band models are not sufficient to explain the physics of these devices. This criterion places further requirements on the stability and efficiency of numerical methods for the solutions of these problems. Transfer matrices are not sufficiently accurate to perform the computations correctly. Thus, better computational techniques are a requirement for the proper simulation of interband tunneling. Experimentally, current-voltage characteristics do not offer a direct measure of the properties of states in the well. Typically, only one I-V peak is observed, even at low temperatures in devices with multiple well subbands. The opposite dispersions of the electrode and well states allow resonant tunneling through multiple well states at a single voltage, convoluting the contributions of the individual subbands, and leading to a single I-V peak.

Figure 2.10 shows the transmission coefficients as a function of energy for a RIT having 36 Å AlSb barriers and a 90 Å GaSb well, at both zero and non-zero $k_{\|}$. These calculations were made with an efficient, stable model developed by Ting *et al.* to accurately model interband devices [28]. The device parameters correspond identically to those of the InAs/AlSb DBH for which calculated transmission coefficients are shown in Figure 2.6, except that the InAs well has been replaced with an identical thickness of GaSb. Note that the RIT transmission characteristics differ greatly from those of a simple unipolar double barrier device, and thus the simple transfer matrix model developed earlier is insufficient to explain the interband problem. At zero parallel wavevector, two peaks exist in the transmission coefficient, one very close to the conduction band edge of InAs (the energy origin), and another near the GaSb valence band edge (at roughly 0.15 eV). Between these peaks, the transmission coefficient for the RIT is more than three orders of magnitude greater than that of the InAs DBH. At incident energies exceeding the GaSb valence band edge ($E > 0.15$ eV), the transmission coefficient drops precipitously (again roughly three orders of magnitude from 0.15 eV to 0.20 eV). This effect has been termed "band gap blocking" and represents a fundamental distinction between interband tunneling and Type I unipolar transport. At these energies, the incident carrier is in the energy gaps of *both* the GaSb and AlSb. Thus, the wavefunction is attenuated greatly, and little transmission through the interband device can occur. This band gap blocking phenomenon is responsible for the large peak-to-valley ratios seen in interband devices (the RIT in particular). PVR's as high as 21:1 at room temperature and 88:1 at 77K have been recorded for p-well RIT's [26].

The additional complexities of heavy hole states is seen in the figure as well, for $k_{\|} \neq 0$. Here, the resonances take on a characteristic Fano-like lineshape, with a minimum closely associated with every maximum. As is true for intraband devices as well, heavy hole states are purely $|\frac{3}{2} \pm \frac{3}{2}\rangle$ only at $k_{\|} = 0$ (where they cannot couple to conduction band states due to the conservation of angular momentum projection, m_j). At non-zero parallel wavevector, the light ($j = \frac{3}{2}, m_j = \pm\frac{1}{2}$) and heavy ($j = \frac{3}{2}, m_j = \pm\frac{3}{2}$) hole states intermix due to their closeness in energy, and similar symmetry. As a result, transmission through heavy hole-like states is permissible at finite $k_{\|}$. While these resonances are typically more narrow and do not contribute as much to the transmitted current, band warping can lead to conduction band-like behavior in heavy hole states, which greatly increases their coupling to the emitter states, as shown in Chapter 3.

The development of the I-V characteristic of a RIT diode is markedly different for an interband device as well. Figure 2.11 shows a sample I-V curve for a p-well RIT as well as the alignment of the electrode and well subband dispersions at the biases indicated in the top panel. Electron occupation at a given energy is indicated by the shading of the region above the emitter curve. First consider

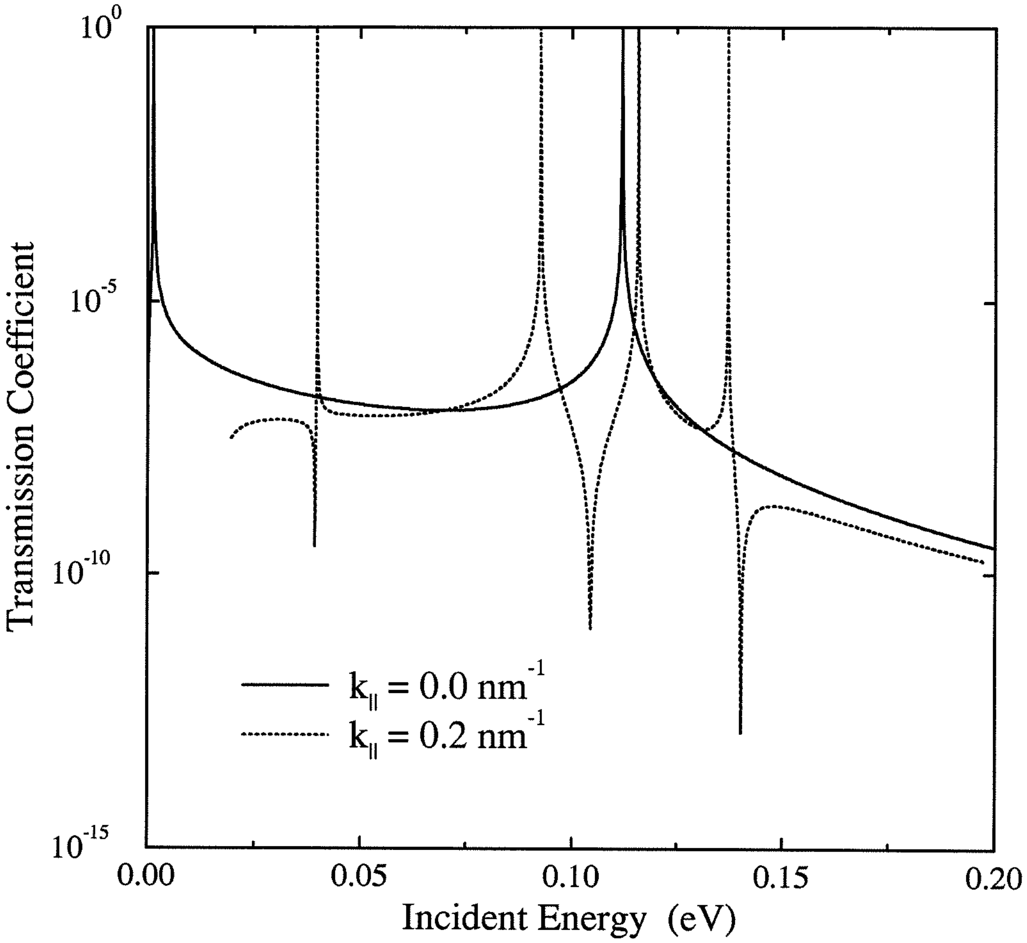


Figure 2.10: Calculated RIT transmission coefficients for a device with 36 Å barriers and a 90 Å well. The energy dependence of the transmission probabilities for an incident carrier are plotted for zero and non-zero $k_{\|}$. The resulting curves are significantly more complex than the results for an identical InAs/AlSb double barrier device. The results were attained through the use of a sophisticated 8-band model attributable to Ting *et al* [28].

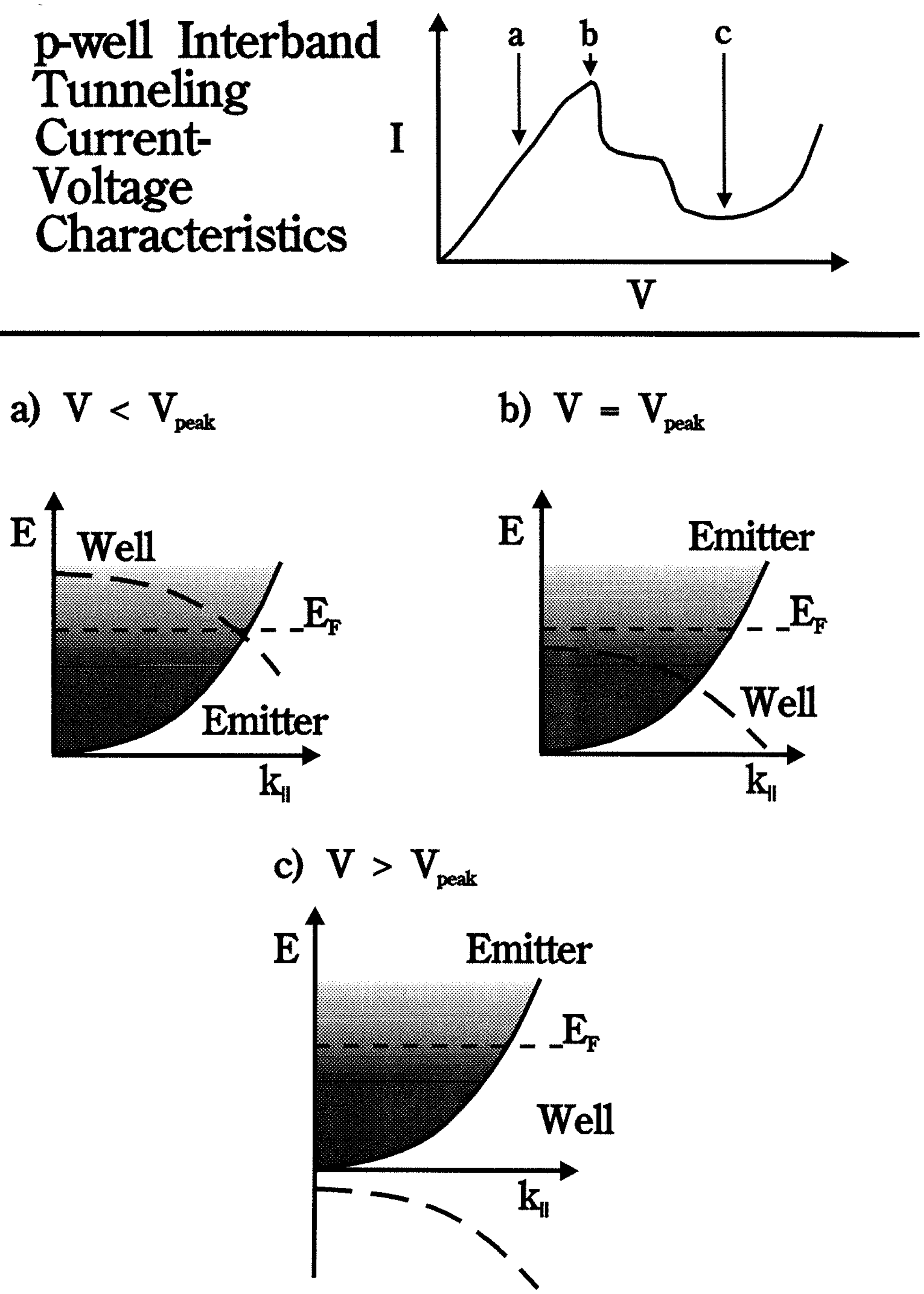


Figure 2.11: The development of an interband p-well RIT I-V characteristic. Shown in the top panel is a representative I-V curve, and beneath, the emitter and well band alignments at each of the three indicated voltages.

the case of 3D-2D tunneling where there is continuum of states, of which only the lowest energy dispersion is indicated in the figure. The first difference that should be noted is the ohmic onset to the interband I-V, as compared to the exponential turn-on of the Type I unipolar I-V curve shown in Figure 2.7. Unlike the Type I unipolar DBH, the RIT does not rely on the thermal excitation of carriers for low voltage transport. As shown in Figure 2.11a, there is overlap between the emitter and well states such that elastic tunneling, for which both energy and parallel wavevector are conserved, is allowed at voltages less than V_{peak}. For some of the well states, this condition may even hold at zero bias. As a result, conduction through RIT diodes at low biases is not limited to the exponential thermal Fermi distribution tail. Another difference, which becomes a significant complication for certain experiments with RIT diodes, is that the current peak does not occur when the $k = 0$ band edges of the two subbands are of equal energy. Clearly, in that limit, the current is nearly zero as there are very few incident carriers to contribute to elastic transport; thus this voltage is somewhere in the valley region. The peak is instead found at some intermediate voltage, as shown in Figure 2.11b. A simple model [29] of interband I-V curves postulates that the resonant current at a voltage $V \leq V_{peak}$ should be proportional to the product of a semiclassical carrier velocity, $v \propto V$, and the number of states available for elastic transport. Such a simple model provides reasonable two-parameter fits to many interband I-V curves [29]. From Figure 2.11, it is clear that these two factors have opposite differential dependencies on the applied voltage; while the carrier velocity is linear in applied voltage, the number of carriers that can elastically tunnel is reduced as V is increased. Analytically, this latter quantity, in the effective mass approximation, is $\propto (V_0 - V)^{-3/2}$ for 3D emitter states, or $\propto (V_0 - V)^{-1/2}$ for 2D emitter states, where V_0 is a constant depending on the bulk effective mass. While the peak voltage may be determined in principle from this model, it is not sophisticated enough to bring full consideration to the quantum aspects of the problem, such as variations in transmission coefficient. Therefore, the exact alignment of the subband dispersions at the peak voltage is not *a priori* known, and can at best only be estimated.

The technical advantages that the InAs/AlSb/GaSb system offer stem from these unique transport properties. Due to band blocking in the valley region, both elastic and inelastic transport through the device at these voltages is severely curtailed, leading to PVR's at 300 K which rival Type I unipolar device parameters at cryogenic temperatures. Additionally, the onset of these devices is not exponentially dependent upon temperature. Interband devices are thus far more likely to provide convenient room temperature operation for solutions to real-world problems than are GaAs/AlAs or other Type I devices. This fact is of great consequence; as interband resonant tunneling devices have a greater likelihood of making positive economic contributions, the understanding of their transport properties, and ensuing optimizations, have significance beyond academic study. The existence of many potential applications for these devices emphasizes their importance.

2.3.2 Technological Applications

Because of their unique band alignments, and the resulting performance improvement at room temperature, interband quantum effect devices fabricated in the InAs/AlSb/GaSb material system have been proposed for use in a wide variety of technologies. One advantage of devices in this system is their intrinsic switching speed. An InAs/AlSb double barrier device grown by the McGill group at Caltech has been observed to oscillate at frequencies as high as 712 GHz at room temperature, representing the current record for a solid-state electronic oscillator [30]. Extrapolated device performance yielded an estimated maximum oscillation frequency of 1.24 THz. The measured power density of this device at 360 GHz, 90 W cm^{-2}, is fifty times greater than that generated by GaAs/AlAs diodes at similar frequencies. While this diode was a purely Type I unipolar DBH, it nonetheless demonstrates the technological advantages enjoyed in this material system. A Stark effect transistor, consisting of an InAs collector, a GaSb emitter, and a GaSb base, has been demonstrated with room temperature current gains in excess of 50 [31]. This unique device was first proposed in the AlAs/GaAs system [32], but later only demonstrated transistor action at cryogenic temperatures (7 K) [33]; the performance advantages of the InAs/AlSb/GaSb material system therefore take the Stark effect device from the position of scientific curiosity to that of real technical feasibility unattainable in Type I systems.

The RIT diode has also provided large performance benefits in computational systems based upon neural paradigms. A single, vertically-integrated RIT diode replaces 16 silicon CMOS transistors for the generation of a resistive fuse-like I-V characteristic for the extraction of spatial discontinuities in artificial retinas [34]. It does so with lower power, far less lithographic real estate, and without the common mode difficulties of active devices. RIT diodes have also been used by Levy *et al.* to store analog weights in feedforward artificial neural networks [35]. A three-layer prototype using RIT's provided room temperature computation of the fundamental logic functions of XOR, AND, OR, and NOT. Clearly, interband tunneling diodes have demonstrated competitive solutions to current and future computational needs with room temperature, quantum effect-based operation.

Given the success of interband devices in other arenas, it was believed very early in the magnetotunneling investigations that RIT diodes might make highly sensitive, tunable detectors of magnetic fields for such technological applications as magnetic media read sensors for digital data storage. Many empirical issues contributed to this presumption. First, semiconductor heterostructure devices may be engineered for precise control of the quantized energy states, and the energies of these states, which fundamentally depend upon the magnitude of external magnetic fields, in turn determine the nature of the diode current-voltage characteristic. Furthermore, the interband nature of the RIT diode intuitively appears more sensitive to changes in magnetic fields, as conduction band states are raised, and valence band states lowered, by external fields. Since these devices are also capable of duplicating the cryogenic results of other systems at technologically-realistic temperatures, the RIT appeared to be ideal as a room temperature B-field detector. Several realities, however, became evident that have lessened the applicability of RIT's as a technological solution. First, the anomalous response of RIT's to perpendicular fields (c.f. Chapter 3), all but eliminates them for magnetic sensing at technologically-realistic fields. For room temperature operation, the perpendicular field orientation is necessary to induce any appreciable change in the I-V curve, but the critical field phenomenon discussed in Chapter 3 eliminates this orientation as a viable detection geometry. The peculiarities of transport in this system therefore may hinder, as well as improve, device performance in certain technological applications. In addition to the intrinsic limitations of the RIT diodes, the development of reproducible devices utilizing giant magnetoresistance (GMR) phenomena have greatly raised the metric for device performance, and further weakened the case for magnetic field detectors based upon current InAs/AlSb/GaSb technology. However, as nanostructures with inherent 0D-like carrier behavior are developed and perfected, the size and potential detectivity advantages of antimonide-based resonant tunneling magnetic field sensors may ultimately lead to their adoption.

2.4 Magnetic Fields in Bulk Semiconductors

2.4.1 Landau Levels

The derivation of the effect of an external magnetic field upon the quantum mechanics of charge carriers in semiconductor heterostructures begins with the modification of the single-particle Hamiltonian to include an arbitrary constant field. The field enters the Hamiltonian through both potential and kinetic energy effects. The former arise from the interaction between the field and the carrier's intrinsic angular momentum (spin). An additional term, $g\mu_B\,\vec{\sigma}\cdot\vec{B}$, is added to the Hamiltonian in the effective mass approximation, where μ_B is the Bohr magneton, and g is the gyromagnetic ratio relating the carrier spin to the intrinsic magnetic moment. As with electric fields, a magnetic field does not enter the kinetic portion of the Hamiltonian directly, but rather is incorporated through a related vector potential, $\vec{A}$, defined by $\vec{B} = \vec{\nabla} \times \vec{A}$. This definition of $\vec{A}$ requires that it be a unique vector field only to within the divergence of a scalar function, $\vec{\nabla} f$. The specification of this function f yields a particular choice of gauge; the physics of the problem are the same in all gauges, but the ease of solution depends upon a wise selection of $\vec{A}$. The use of the Landau gauge, $\vec{A} \equiv -By\,\hat{x}$, or, equivalently, $\vec{A} \equiv Bx\,\hat{y}$, for a constant field $\vec{B} = B\,\hat{z}$, simplifies the kinetics, reducing the field-dependence to a one dimensional eigenproblem. In general, for all $\vec{A}$, the kinetic portion of the Hamiltonian transforms as $\mathcal{H}(\vec{p}) \longrightarrow \mathcal{H}(\vec{p} + e\vec{A})$, where $\vec{p}$ is the momentum operator, and the negative charge of the electron is implicitly included in the expression. Thus, for a charge carrier in the effective mass approximation, the Hamiltonian in the presence of an external field becomes,

$$\mathcal{H} = \frac{1}{2m^\star}\left(\vec{p} + e\vec{A}\right)^2 + g^\star\frac{\hbar e B}{m^\star}, \tag{2.22}$$

where $m^\star$ is the carrier effective mass, and $g^\star \equiv m^\star g/2m_e$ is the effective Landé factor. To determine the change in kinetic behavior, we will ignore the spin-related term for the remainder of the calculation. This term is a constant for carriers of a given spin state, and has little effect on the quantum mechanics of the problem. The magnitude of this term, proportional to $m^\star/m_e$, is negligible in many semiconductors. Considering simply the kinetic portion of the Hamiltonian, and choosing the gauge $\vec{A} = -By\,\hat{x}$, the Schrödinger equation for the wavefunction ψ becomes,

$$-\frac{\hbar^2}{2m^\star}\left(\frac{\partial}{\partial x}\hat{x} + \frac{\partial}{\partial y}\hat{y} + \frac{\partial}{\partial z}\hat{z} - \frac{eBy}{i\hbar}\hat{x}\right)^2 \psi = \epsilon\psi, \tag{2.23}$$

or,

$$-\frac{1}{2m^\star}\left[\hbar^2\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right) - 2i\hbar eBy\frac{\partial}{\partial y} - (eBy)^2\right]\psi = \epsilon\psi. \tag{2.24}$$

Taking the anzatz that ψ is of the form, $\psi(\vec{r}) = e^{-i(k_x x + k_z z)}\phi(y)$, substitution into Equation 2.24 yields a differential equation for the unknown function $\phi(y)$,

$$\frac{d^2\phi}{dy^2} + \frac{2m^\star}{\hbar^2}\left[\epsilon - \frac{\hbar^2\left(k_x^2 + k_z^2\right)}{2m} + \frac{eB}{m}yk_x - \frac{1}{2m}(eBy)^2\right]\phi = 0. \tag{2.25}$$

Defining $\omega_c = eB/m^\star$ (the cyclotron frequency), $y' = y - k_x/s^2$, where $s^2 = eB/\hbar$, and $\epsilon' = \epsilon - \hbar^2 k_z^2/2m^\star$, and substituting in the differential equation for $\phi(y)$, we obtain the Schrödinger equation for a quantum harmonic oscillator (QHO) in the primed variables:

$$-\frac{\hbar^2}{2m^\star}\frac{d^2\phi}{dy'^2} + \frac{1}{2}m^\star\omega_c^2 y'^2\phi = \epsilon'\phi. \tag{2.26}$$

The solutions of this equation are well-known,

$$\epsilon' = \left(l + \frac{1}{2}\right)\hbar\omega_c,\ l = 0, 1, 2\ldots, \tag{2.27}$$

and,

$$\phi_l(y) = \left(\frac{s^2}{\pi}\right)^{1/4}\frac{1}{2^{l/2}\sqrt{l!}}e^{-s^2y^2/2}H_l(sy), \tag{2.28}$$

where $H_n(y)$ is the 1^{th} Hermite polynomial, and s, as defined earlier, is now clearly an inverse scaling length. The discrete levels indexed by l are called Landau levels, and are characteristic of the quantum-mechanical transverse kinetic energy in the presence of a constant magnetic field.

Since the behavior of a semiconductor carrier in a magnetic field has reduced to a simple, well-known 1D problem, the raising operator, which acts on a state ϕ_l to yield a state proportional to ϕ_{l+1}, may defined in the same manner as a simple QHO,

$$\begin{aligned} a^\dagger &\equiv \frac{1}{\sqrt{2}}\left[sy' - i\frac{p_y}{\hbar s}\right], && (2.29)\\ &= \frac{1}{\sqrt{2}}\left[s\left(y - \frac{k_x}{s^2}\right) - i\frac{p_y}{\hbar s}\right], && (2.30)\\ &= -\frac{1}{\sqrt{2}s}\left[\overline{k}_x + i\overline{k}_y\right], && (2.31) \end{aligned}$$

where $\hbar\overline{k}_i \equiv p_i + eA_i$ is the generalized k-vector in a magnetic field. The lowering operator, a, is simply the complex conjugate of $a^\dagger$. These operators have the expected result when operating on the Landau level wavefunctions, $\phi_l(y)$,

$$\begin{aligned} a^\dagger\phi_l &= \sqrt{l+1}\phi_{l+1}, && (2.32)\\ a\phi_l &= \sqrt{l}\phi_{l-1}. && (2.33) \end{aligned}$$

The functional form of a and $a^\dagger$ emphasize the the fundamental effect of the field on the quantum mechanics of the carriers; since $\vec{A}$ modifies the momentum-dependent term of the Hamiltonian, and therefore $q\vec{A}/\hbar$ is, in a very real sense, a field- and position-dependent crystal momentum vector, the essential effect of the field is to directly modify the allowed states in k-space. Consequently, as developed in Appendix B, magnetic field effects are best incorporated in rigorous band structure models with k-space ($\vec{k} \cdot \vec{p}$) methods, rather than real-space (tight binding) techniques. Because the vector potential may be viewed as an entity in k-space, yet is itself proportional to a position in real-space along a perpendicular axis, there is an interesting duality to magnetic field problems. Despite the reduction of the Landau level problem to a single 1D problem in one of the two coordinate variables perpendicular to the field, the problem is equally well-formulated along the other transverse axis. This fact is easiest seen in the equivalent gauges, $\vec{A} = Bx\,\hat{y}$ and $\vec{A} = -By\,\hat{x}$. The raising and lowering operators in the latter gauge may be expressed as rotations of those derived earlier,

$$a^\dagger \longrightarrow ia^\dagger, \tag{2.34}$$

$$a \longrightarrow -ia. \tag{2.35}$$

Note that under this rotation, the number operator, defined as $\hat{N} \equiv a^\dagger a$, in terms of which the transverse Hamiltonian may be expressed, is invariant, and thus the energy spectra are unaltered by the difference in gauge. Thus, while although expressed as a 1D problem, the magnetic field scrambles k-space in both directions normal to the field; choosing the symmetric gauge $\vec{A} \equiv -\frac{1}{2}By\hat{x} + \frac{1}{2}Bx\hat{y}$ is perhaps the most intuitive way of expressing the general problem, but requires a more complicated derivation to achieve the same results. Thus, while the use of the 1D gauges yield results that appear to apply to only a single transverse direction, they may be extended, with caution, to the entire k-space plane normal to the field.

Therefore, using the derived raising and lowering operators, we can find expressions for the position, y', and momentum, p'_y, operators in a magnetic field. Rewriting Equations 2.31, y' and p'_y may be reexpressed in terms of $a^\dagger$ and a as,

$$y' = \frac{1}{\sqrt{2}s}\left(a^\dagger + a\right), \tag{2.36}$$

and,

$$p'_y = \frac{i}{\sqrt{2}}\hbar s\left(a^\dagger - a\right), \tag{2.37}$$

respectively. It is immediately obvious that $\langle y' \rangle$ and $\langle p_y \rangle$ are zero, as $\langle \phi_l | \phi_{l'} \rangle = \delta_{ll'}$. In general, every $a^\dagger$ must be matched with an a to yield non-zero matrix elements.

The square average deviations of both variables, however, are non-zero, and given by,

$$(\Delta y')^2 = (l+\frac{1}{2})\frac{1}{s^2}, \tag{2.38}$$

$$(\Delta p'_y)^2 = (l+\frac{1}{2})\hbar^2 s^2, \tag{2.39}$$

yielding RMS expectations for the mean deviation in y' and $k_y = p_y/\hbar$ of $\sqrt{l+\frac{1}{2}}\,\frac{1}{s}$ and $\sqrt{l+\frac{1}{2}}\,s$, respectively. Semi-classically, these values correspond to the orbital radii in real- and k-space. As the field increases, the orbits get tighter in real space and simultaneously move to larger values of transverse crystal momentum.

2.4.2 Shubnikov-de Haas Oscillations

The variation of orbital radius in k-space as a function of the applied field results in periodic variations in the Fermi energy at low temperatures, and subsequent oscillations in the bulk transport properties. With zero magnetic field, the transverse k states form a periodic 2D reciprocal lattice with regular spacing between each of the allowed k-values. When a magnetic field is applied in the z-direction, these states begin to coalesce onto circular orbits in the k_x–k_y plane, forming coaxial cylinders of radii $\sqrt{l+\frac{1}{2}}\,s$. The radius of each cylinder thus increases linearly with the magnitude of the applied field. The degeneracy of each cylindrical cross-section in the k_x–k_y plane can be determined from the number of cyclotron orbits possible in a rectilinear crystal of dimensions L_x and L_y transverse to the field. In the gauge $\vec{A} = -By\,\hat{x}$, k_x is a valid quantum number, and the eigenstates of the Hamiltonian keep their plane wave-like nature along this axis. The density of states along k_x is thus $L_x/2\pi$ and is unaffected by the magnetic field. In the y-direction, the orbits are centered about $y_0 \equiv y - y' = k_x/s^2$. Because these must lie within the confines of the crystal, $0 < y_0 < L_y$, the total number of orbits allowed along the k_y-axis is $L_y s^2$; therefore the total density of states per unit area per Landau level is given by

$s^2/2\pi = eB/h$, and likewise increases linearly with field. It should be noted that the average number of states in zero field contained in the area in k-space between two Landau levels at field B is given by $\Delta E\,\mathcal{D}(E) = (\hbar\omega_c)(m^\star/(2\pi\hbar^2)) = eB/h$, and so the average areal density of allowed states is unchanged, although their distribution is greatly altered.

With the carriers lying upon these cylinders in k-space, the degeneracy and radius of which are linear functions of the field, the Fermi wavevector (and consequently the Fermi energy) oscillates as a function of applied field. At an arbitrary field chosen such that the upper most Landau level is only partially populated, the Fermi wavevector must lie at the radius of the outermost k-space orbit at zero temperature. The radius of this orbit increases linearly with applied field, and the Fermi level must follow as well. As the field is increased, however, the degeneracies of the lower-lying levels also increase linearly, and carriers from the upper Landau level must drop into these lower states to satisfy the Fermi population criterion. At a certain threshold field, the last carrier drops into a lower level, and the Fermi level shifts from the l to the $l-1$ Landau level. The field at which this shift occurs is such that the total degeneracy of the lowest l_{max} levels equals the total volume charge density in the crystal. Since the degeneracy per unit area of each Landau level is given by eB/h, and the density of states along the direction of the field is $L_z/2\pi$, the Fermi level is expected to jump whenever $l_{max}eBL_z/2\pi h = N_{3D}$, where N_{3D} is the three-dimensional carrier density, or,

$$\frac{1}{B_l} = \frac{eL_z}{2\pi h N_{3D}}\,l, \tag{2.40}$$

where $1/B_l$ is the field at which the Fermi level shifts to the Landau level indexed by l. Thus, the shift in Fermi level is periodic in $1/B$, and any transport characteristic of the bulk material pertaining to the behavior of carriers at E_F should be periodic as well. Shubnikov-de Haas oscillations in the measured bulk resistivity of a semiconductor in a magnetic field are due to the increased carrier scattering rates at fields where the Fermi energy drops down to that of a fully occupied Landau level. In two-dimensional carrier gas systems, Shubnikov-de Haas oscillations also are observed for in-plane transport, although only the component normal to the 2D gas, $B\cos\theta$, contributes to the periodicity. Other transport variables are observed to oscillate as well, and Shubnikov-de Haas-like phenomena occur in quantum transport devices such as those studied in this thesis, as discussed in Chapter 4.

Bibliography

[1] J. Farrel, "Economic Indicator," Semiconductor International **17** (4), 49 (1994).

[2] M. Levitt, Proc. SPIE - Int. Soc. Opt. Eng. **1719**, 1 (1992).

[3] *Molecular Beam Epitaxy and Heterostructures*, proceedings of the NATO Advanced Study Institute on Molecular Beam Epitaxy (MBE) and Heterostructures, Erice, Italy, edited by L. L. Chang and K. Ploog (Martinus Nijhoff, Dordrecht, 1985).

[4] G. Bastard, *Wave Mechanics Applied to Semiconductor Heterostructures* (Les Éditions de Physique, Les Ulis, France, 1988).

[5] A. Y. Cho and J. R. Arthur, Prog. Solid State Chem. **10**, 157 (1975).

[6] *The Technology and Physics of Molecular Beam Epitaxy*, edited by E. H. C. Parker (Plenum Press, New York, 1985).

[7] R. M. Feenstra *et al.*, Phys. Rev. Lett. **72**, 2749 (1994).

[8] *Semiconductors: Group IV Elements and III-V Compounds*, edited by O. Madelung (Springer-Verlag, Berlin, 1991).

[9] S. M. Sze, *Physics of Semiconductor Devices*, (John Wiley & Sons, New York, 1981).

[10] C. Weisbuch and B. Vinter, *Quantum Semiconductor Structures: Fundamentals and Applications*, (Academic Press, New York, 1991).

[11] R. L. Anderson, Solid-State Electronics **5**, 341 (1962).

[12] J. O. McCaldin, T. C. McGill, and C. A. Mead, Phys. Rev. Lett. **36**, 56 (1976).

[13] E. T. Yu, J. O. McCaldin, and T. C. McGill, Solid State Physics **46**, 1 (1992).

[14] J. O. McCaldin, J. Vac. Sci. Technol. A **8**, 1188, (1990).

[15] M. W. Wang *et al.*, J. Appl. Phys. **73**, 4660 (1993).

[16] G. Bastard and J. A. Brum, IEEE J. Quantum Electron. **QE-22**, 1625 (1986).

[17] G. Bastard, Phys. Rev. B **25**, 7584 (1982).

[18] C. Aversa and J. E. Sipe, Appl. Phys. Lett. **63**, 1975 (1993).

[19] E. Merzbacher, *Quantum Mechanics*, (Wiley and Sons, New York, 1970).

[20] J. M. Luttinger, Phys. Rev **102**, 1030 (1956).

[21] R. Tsu and L. Esaki, Appl. Phys. Lett. **22**, 562 (1973).

[22] G. J. Gualtieri, G. P. Schwartz, R. G. Nuzzo, and W. A. Sunder, Appl. Phys. Lett. **49**, 1037 (1986); G. J. Gualtieri et al., J. Appl. Phys. **61**, 5337 (1987).

[23] L. L. Chang and L. Esaki, Surf. Sci. **98**, 70 (1980).

[24] M. Sweeny and J. Xu, Appl. Phys. Lett. **54**, 546 (1989).

[25] D. H. Chow *et al.*, in *Quantum-Well and Superlattice Physics III*, proceedings of the SPIE **1283** (Bellingham, WA, 1990).

[26] D. A. Collins *et al.*, in *Resonant Tunneling in Semiconductors: Physics and Applications*, edited by L. L. Chang, E. E. Mendez, and C. Tejedor (Plenum, New York, 1991).

[27] R. R. Marquardt, unpublished.

[28] D. Z.-Y. Ting, E. Yu, and T. C. McGill, Phys. Rev. B **45**, 3583 (1992).

[29] D. A. Collins, Ph.D. Thesis, California Institute of Technology, 1993.

[30] E. R. Brown *et al.*, Appl. Phys. Lett. **58**, 2291 (1991).

[31] D. A. Collins, D. H. Chow, and T. C. McGill, Appl. Phys. Lett. **58**, 1673 (1991).

[32] A. R. Bonnefoi, D. H. Chow, and T. C. McGill, Appl. Phys. Lett. **47**, 888 (1985).

[33] F. Beltram *et al.*, Appl. Phys. Lett. **53**, 219 (1988).

[34] H. J. Levy, D. A. Collins, and T. C. McGill, in Proceedings of the 1992 IEEE International Symposium on Circuits and Systems, San Diego, CA (1992).

[35] H. J. Levy and T. C. McGill, IEEE Trans. Neural Nets **4**, 427 (1993).

Chapter 3

Resonant Magnetotunneling Spectroscopy

3.1 Introduction

The resonant magnetotunneling spectroscopy technique was developed for the study of hole tunneling in AlAs/GaAs double barrier structures [1]. While simple in concept, it is an experimentally powerful method for probing the in-plane energy dispersion of quantized carrier states in semiconductor double barrier structures. As it implicitly relies upon the tunneling of carriers into the quantizing region, it is not a technique which can be used in the study of quantum wells. It is also typically used to study subbands in which the effective mass approximation is not expected to hold for all magnitudes or directions in the $k_{\|}$-plane. With a magnetic field aligned parallel to the growth plane quantization effects due to the field and the crystal potential are coaxial. Except for the highest fields, or widest quantum wells, the magnetic field in this geometry may be treated perturbatively. As a result, the device operates without additional quantization or significant modification to the eigenstates and energies of the system; the only effect of the field, semi-classically, is to bend the carrier trajectories into cyclotron orbits about the axis of the field. This alteration of the carrier distribution in momentum space, and the resulting effect it has upon the elastic tunneling of carriers through the device, is exploited in resonant magnetotunneling spectroscopy (RMTS) [1] to probe the energy subband dispersion of quantized carriers in semiconductor heterostructures.

A vast majority of work with the RMTS technique has been with traditional Type I unipolar intraband tunneling diodes [1, 2, 3, 4, 5], where either valence band states or in-plane band warping was studied. Few quantum magnetotransport observations of any type have been reported for interband heterostructures employing the InAs/GaSb Type II broken-gap band alignment. Previous magnetotunneling studies in this material system have focused primarily on the n-well RIT, with an InAs well and GaSb electrodes [9, 10, 11], or have investigated the barrierless InAs/GaSb/InAs system with the magnetic field perpendicular to the layers [12]. This chapter details the experimental results of quantum transport through InAs/AlSb/GaSb/AlSb/InAs heterostructures having well widths of 7.0 nm, 8.0 nm, and 11.9 nm, respectively, in magnetic fields of up to 8.0 tesla aligned in the RMTS geometry parallel to the epitaxial growth planes. Unlike the previous magnetotunneling experiments in InAs/AlSb/GaSb, this study investigates the rich valence subband structure of GaSb wells, and reveals correspondingly interesting and complex quantum transport phenomena. The results of these studies are detailed in the remainder of this chapter. In Section 3.2 we examine the RMTS technique, and provide a simple theory for the interpretation of experimental data. The measurements are presented in Section 3.3, and analyzed in Section 3.4. Section 3.5 summarizes and concludes the chapter.

3.2 Technique

3.2.1 Basic Model

The theory of resonant magnetotunneling spectroscopy is similar for tunneling mechanisms involving initial and final states with either the same, or opposite, energy dispersions. Thus, the theory for both may be developed in tandem. Referring to Figure 3.1, we take the x-axis to lie along the growth direction, while the y- and z-axes both lie in the plane parallel to the heterointerfaces. With the B-field along the z-axis, $\vec{B} = B\,\hat{z}$, and using the Landau gauge vector potential $\vec{A} = Bx\,\hat{y}$, the sole effect of the field on the Hamiltonian is to replace the canonical crystal momentum operator $\vec{\mathbf{p}}$ with $\vec{p} - q\vec{A}$. Assuming the vector potential term can be treated as a perturbation, a valid assumption provided the cyclotron orbital radius is significantly larger than the width of the well ($1/s \ll L_W$), and the cyclotron energy is small relative to the quantizing material potential ($\hbar\omega_c \ll \Delta E$), the eigenstates of the Hamiltonian remain unchanged to first order, and the mechanical momentum along the y-axis is increased by $eB\langle x\rangle/\hbar$, where $\langle x\rangle$ is the expected value of the position operator averaged over a given localized state under the perturbation. The origin along the x-axis is arbitrary; a change in origin amounts to a change of gauge, for which the physics is invariant. Changes in $\langle x\rangle$, however, such as tunneling from a localized emitter notch state to a quasi-bound state in the well, result in an increased mechanical momentum along the y direction. The conjugate momentum, $p_y \equiv \hbar k_y$, however, remains conserved in elastic transport. The parallel energy dispersions of the well states, dependent upon the mechanical momentum, $M_y = p_y + eB\langle\Delta x\rangle$, are shifted in k-space when plotted against the conserved quantity, k_y, as shown in Figure 3.2. As the magnetic field is increased, and the mechanical momentum is likewise boosted, the well states are further shifted along k_y, changing the quasi-bound states available for elastic tunneling processes conserving both E and k_y. An ensemble of electrons, thermalized in an emitter subband such that $|k_y| \leq k_F$ and $M_y = p_y = \hbar k_y$, will have mechanical momenta distributed from $-k_F + eB\langle\Delta x\rangle/\hbar$ to $k_F + eB\langle\Delta x\rangle/\hbar$ after tunneling an expected distance $\langle\Delta x\rangle$ to a localized quasi-bound state in the well. As a result, the energy of the ensemble has been altered, and the change in energy must be compensated by a change in the applied voltage required to meet the resonance condition for tunneling. The magnetic field, doing no work on the charge carriers, cannot alter their energies. As the well states are shifted, the energy dispersion of the destination states in the well changes the energy required for resonance at a given k_y. Thus, as a function of applied magnetic field, the current-voltage (I-V) characteristics of the device should be affected in a manner which is fundamentally related to the energy dispersion of the subbands in the well. It is important to note that this analysis rests upon two implicit assumptions: first, that tunneling may be pragmatically viewed as a sequential process, in which a carrier tunnels first into the well, and, subsequently, from it, and, secondly, that there exists a quasi-bound state in the emitter accumulation region (a notch state) such that the quantity $\langle\Delta x\rangle$ is well-defined. In the absence of either of these conditions, the semi-classical tunneling distance $\langle\Delta x\rangle$ becomes ill-defined, and the theory loses validity.

For devices in which the initial and final states are of similar dispersion (Type I unipolar intraband tunneling), peaks in the I-V curves can be associated with the energy of a given subband dispersion at a $k_{||}$ value given by the increase in mechanical momentum induced by the applied magnetic field, $\Delta M_y = eB\langle\Delta x\rangle$ [1]. Graphically, this criterion is shown for intraband tunneling devices in Figure 3.2. This figure shows the energy dispersion for both the emitter and well quantized states along the k_y-direction (perpendicular to both the tunnel current and the applied field) in the effective mass approximation for a Type I unipolar DBH

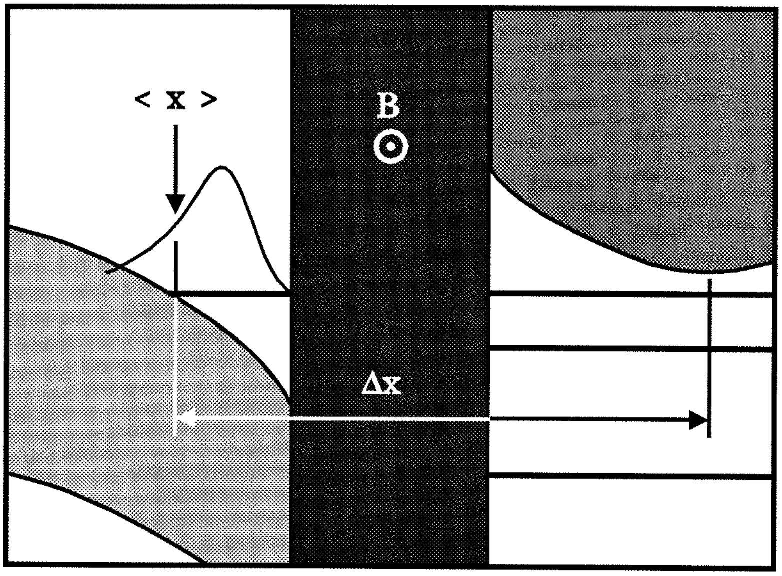


Figure 3.1: Energy-band diagram of an InAs/AlSb/GaSb/AlSb/InAs RIT diode at 300K, in the emitter-barrier region of the device. Electrons from the quasi-bound state in the emitter tunnel an expected distance $\langle \Delta x \rangle$ through the barrier into empty valence band states in the well. In the presence of a magnetic field perpendicular to this tunnel current, carriers experience a semi-classical mechanical momentum shift perpendicular to both the applied field and the growth direction.

at the peak voltage. In the absence of any magnetic field, the peak in the I-V characteristic corresponds to the condition that the $k_{||}$ band edges of the well and electrode dispersions are coincident in energy, as shown. In a magnetic field, the well state is shifted along the k_y-axis, and, as the figure shows, the I-V peak again occurs at $k_y = 0$ (with $|k_x| \leq k_F$), where the electrode has the greatest number of carriers available for tunneling. The degree to which the peak voltage is shifted is therefore given by the well dispersion in the k_y-direction at a relative wavevector $eB\langle \Delta x \rangle / \hbar$ from the well state band edge. Thus, plotting the change in peak voltage (or, equivalently, simply the peak voltage itself) as a function of the applied magnetic field traces out the parallel band structure of the quasi-bound state in the well. From this argument, it is clear that the well dispersion need not obey the effective mass approximation, and the technique is capable of mapping complicated well band structures.

The mechanics of interband tunneling are, unfortunately, far more complex, and they obscure the direct interpretation of the experimental results. The opposite dispersions of the electrode and well result in elastic tunneling through a broad distribution of states in k-space at a single bias. As shown in Figure 3.3, in the absence of a magnetic field the two dispersions have coincident energies and momenta at two values along the k_y-axis (for $k_x = 0$), symmetrically located about $k_y = 0$. With the consideration of non-zero k_x as well, the overlap between the two bands involves all k_y values between these extrema. It can be shown in the effective mass approximation that, for an arbitrary magnetic field, the set of elastically-allowed states is a circle in the $k_{\|}$ plane, centered along the k_x-axis at $k_x = 0$, and along the k_y-axis at,

$$K_0 = \frac{m_E^\star \, \delta}{m_W^\star + m_E^\star}, \tag{3.1}$$

having a radius,

$$K_R = \sqrt{\frac{\Delta E_0 - \left(eV + (\hbar^2\delta^2)/(2(m_W^\star + m_E^\star))\right)}{\hbar^2/2\mu^\star}}, \tag{3.2}$$

where ΔE_0 is the zero-field difference in the $k_{\|} = 0$ band edges, $m_E^\star$ and $m_W^\star$ are the electrode and well effective masses, respectively, δ is the field-induced increase in the mechanical momentum along the k_y direction ($= eB\langle\Delta x\rangle/\hbar$), V is the applied bias, and $\mu^\star$ is the reduced mass of the electrode and well dispersions. As shown in the figure, this intersection occurs over a range of energies as well. As the magnetic field is increased, k-space is sampled in a region centered about a k_y value that is dependent upon the masses of both dispersions and linearly proportional to the applied field. Thus, while this shift in k_y results in changes to many of the factors determining the tunnel current, including the number of states

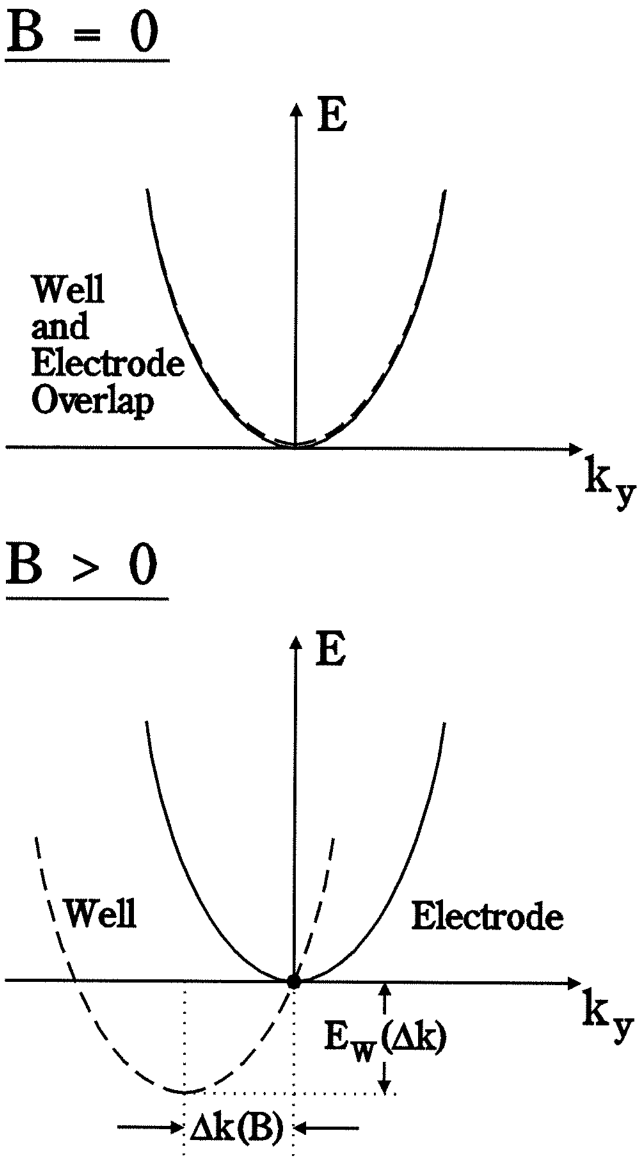


Figure 3.2: Electrode and well dispersions in the k_y-direction (for $k_x = 0$) in the effective mass approximation for an intraband double barrier structure, with and without a magnetic field applied perpendicular to the tunnel current. Both cases are shown at the peak current bias, with states meeting the elastic tunneling condition at the highlighted overlap. The well subband dispersion is shifted by a $\Delta k(B)$ linearly proportional to the magnetic field, increasing the peak voltage by $E_W(\Delta k)$.

involved, and their quantum mechanical transmission probabilities, the change in peak voltage is related, albeit in a complex manner, to the well (and electrode) dispersions.

3.2.2 Experimental Setup

Three p-well RIT samples, grown by molecular beam epitaxy (MBE) [6, 7, 8] on (100) n^+-GaAs substrates, were investigated. The three devices, having well widths of 7.0 nm (device A), 8.0 nm (device B), and 11.9 nm (device C), were otherwise grown with identical layer thicknesses and doping. These device parameters are summarized in Table 3.1. In addition, an InAs/AlSb Type I unipolar RTD with a 16.0 nm well (designed to have roughly identical confinement as that for light holes in the 8.0 nm RIT) was investigated for comparison. For each device, mesas were defined by standard photolithography and Au/Ge liftoff, and etched approximately 500 nm with a 1:8:80 H_2SO_4:H_2O_2:H_2O wet etch, leaving a thick InAs electrode beneath to minimize series resistance. Au/Ge lateral contacts were then deposited in close proximity to the mesas. Au/Ge was used solely for its beneficial adhesion properties; InAs has a negative Schottky barrier for electrons, and the formation of ohmic contacts requires no special processing. Devices A and B each had octagonal mesa cross-sections, 150 μm across, with a total area of 17,300 μm^2. Device C was fabricated with a 100 μm square mesa geometry. Devices A, B, and C had zero-field peak current densities of 126 A/cm^2, 44 A/cm^2, and 138 A/cm^2, respectively, at 4.5 K. The higher current densities in samples A and C may be indicative of greater non-resonant leakage current in these devices. The peak-to-valley ratios (PVRs) of all three were low, typically 2–8:1, at 4.5 K. These PVRs are low even for devices of this size, and are indicative of enhanced inelastic transport across the quantum-well region. Once fabricated, the samples were mounted with silver paint on custom, non-magnetic 8-pin TO-5 headers manufactured by Coors Ceramic Corporation.

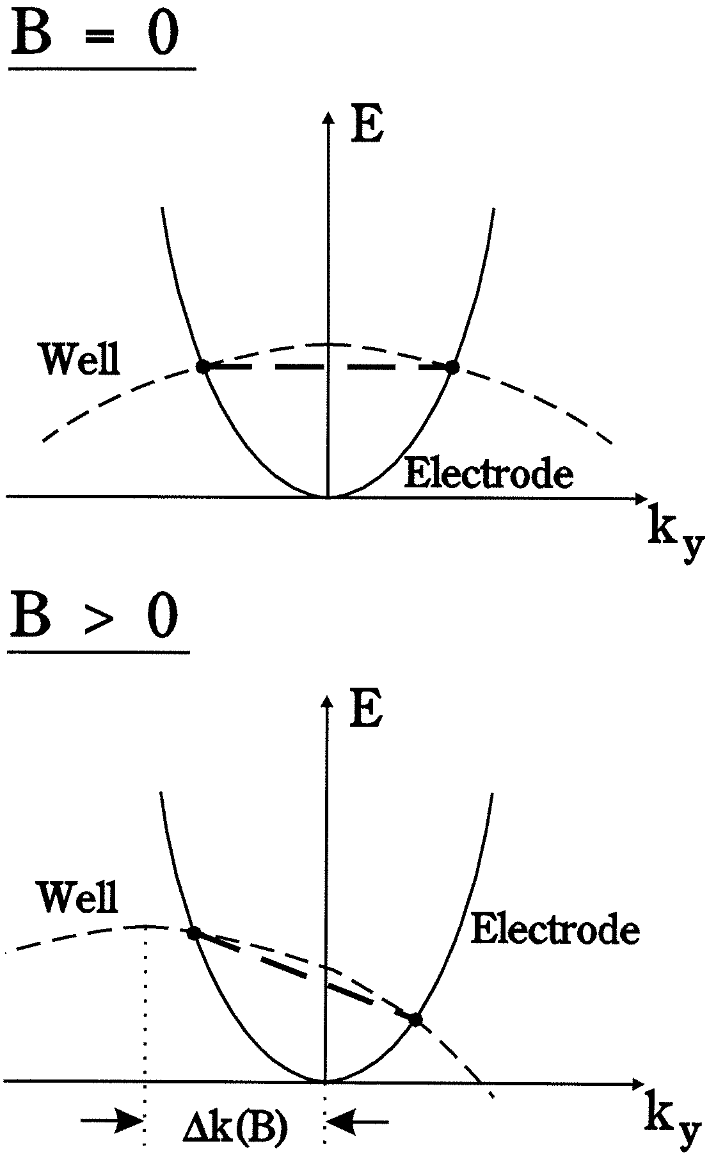


Figure 3.3: Electrode and well dispersions in the k_y-direction (for $k_x = 0$) in the effective mass approximation for the investigated p-well interband RIT structure, with and without a magnetic field applied perpendicular to the tunnel current. The overlap at any given bias involves a distribution of values in k_y, as indicated by the thick dashed lines. Experimental RMTS data for this device therefore cannot be attributable to a single, field-dependent parallel wavevector.

Table 3.1: RIT growth parameters

Material	Thickness	Doping
n^+-InAs	>0.25 μm	3×10^{18} cm^{-3}
n-InAs	50.0 nm	1×10^{17} cm^{-3}
InAs	5.0 nm	undoped
AlSb	4.0 nm	undoped
GaSb	7.0, 8.0, & 11.9 nm	undoped
AlSb	4.0 nm	undoped
InAs	5.0 nm	undoped
n-InAs	50.0 nm	1×10^{17} cm^{-3}
n^+-InAs	>1.2 μm	3×10^{18} cm^{-3}

These headers consisted of an Al_2O_3 ceramic compound base with copper pins coated in a Cu/Ag euthectic braze alloy.

All of the measurements were performed in a cryogenic dewar housing a Nb/Ti superconducting magnet and a variable temperature cryostat. The sample temperature was monitored with a carbon glass thermometer, and was maintained at 4.5 K during all observations. The angle of the sample with respect to the applied field was continuously adjustable from 0° to 360°. A Hall probe mounted adjacent to the sample was used to align the sample surface parallel to the field to within less than 0.5°. The magnet coil was rated for a maximum field of 8.0 tesla, with a homogeneity of ±0.1% over a 1 cm diameter central region. A voltmeter, monitoring a calibrated shunt voltage, was used to determine the magnitude of the applied field.

Current-voltage (I-V) curves were obtained with a Hewlett-Packard 4145A semiconductor parameter analyzer. All I-V curves were scanned from low to high voltage, although typically no more than 2-3 mV of hysteresis was observed when the scan direction was reversed. Most of the mesas displayed nearly-symmetric I-V curves, but in all cases the mesa and polarity with the highest PVR at 4.5 K was selected for study. The header-to-4145A series resistance was measured to be 1.4 Ω at 4.5 K, and the total resistance in series with the devices was estimated not to exceed 2.5 Ω from point-to-point measurements on the lower InAs electrode layer. For each device, I-V curves from 0.0 to 0.5 V were taken in 1 mV steps at regularly spaced ($\leq$ 0.1 T) magnetic fields from 0.0 to 8.0 tesla.

3.3 Experimental Results

3.3.1 I-V Peak and Shoulder

A representative I-V curve from the InAs DB sample at zero applied magnetic field and a temperature of 4.5 K is shown in Figure 3.4. As marked by arrows in the figure, there are four well states identifiable from peaks or inflections in the I-V characteristic. As the field is increased, not only do these states move as per the RMTS theory mentioned earlier, they broaden as the well dispersion is shifted away from $k_y = 0$. As a result, the location of only two of these states could be ascertained accurately for fields greater than several tenths of a tesla, and only the resonance at the highest voltage could be tracked for $B > 2.1$ T. The peak voltages of the two prevalent states are plotted as a function of magnetic field in Figure 3.5. Both obey the effective mass approximation, as expected for conduction band states, with effective masses of $0.044 \pm 0.008\, m_e$ and $0.006 \pm 0.002\, m_e$, assuming a $\langle \Delta x \rangle$ of 25.0 ± 5.0 nm as estimated from band bending calculations and a one-band

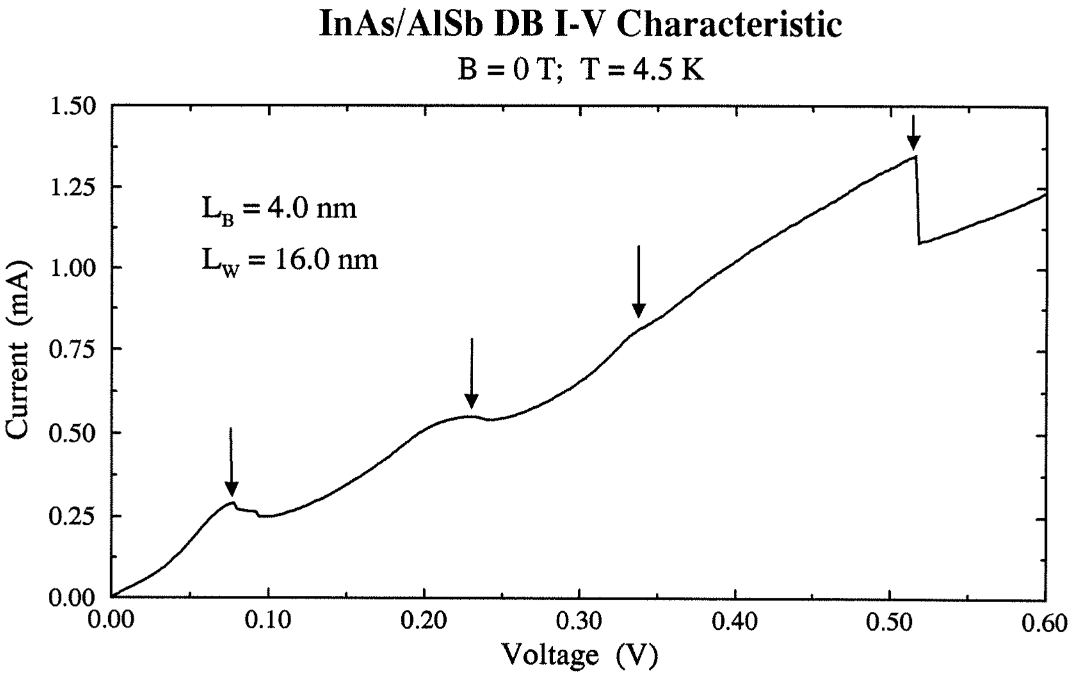


Figure 3.4: Zero-field I-V characteristic of an InAs/AlSb double barrier structure at 4.5 K. Four resonant states are obvious, two of which persisted at non-zero magnetic fields.

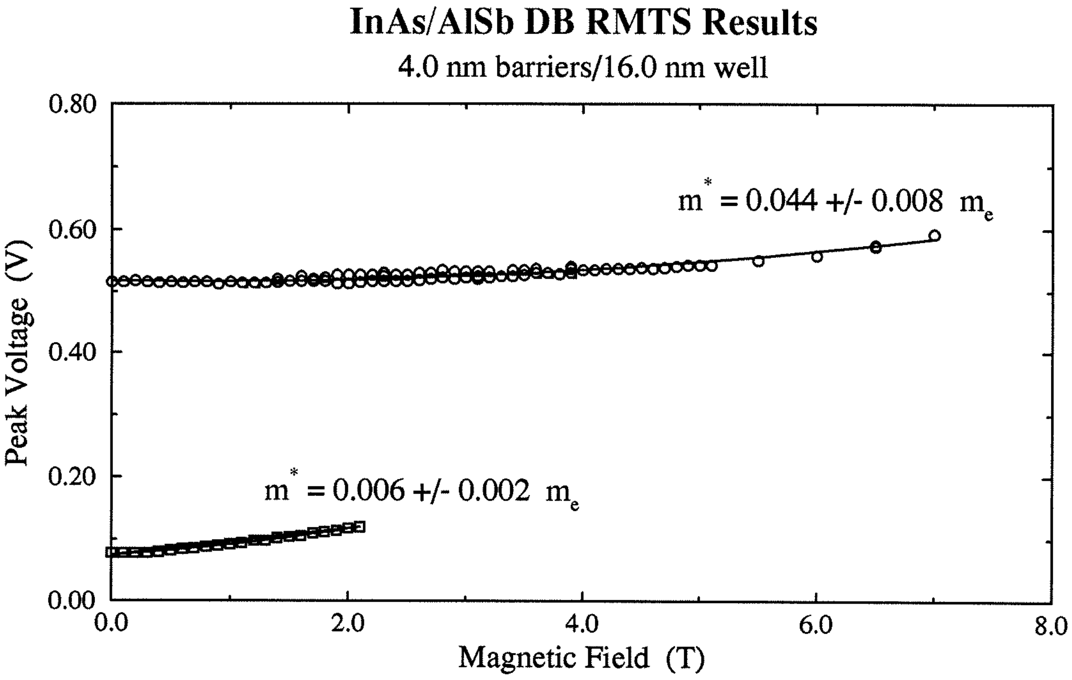


Figure 3.5: Peak positions of resonant states of an InAs/AlSb double barrier device as a function of magnetic field applied perpendicular to the tunnel current. The resulting curves are expected to follow the transverse energy dispersion of the states in the InAs well, both of which appear to obey the effective mass approximation.

model of the emitter quasi-bound state. These values should be compared to the bulk InAs effective mass of $0.023\,m_e$.

In contrast, the I-V characteristics of the interband RIT devices exhibit only a single peak, despite several quantized states in each well. Because of the opposite dispersions, and small energy range over which resonant transport occurs, the presence of a single peak is typical. These interband devices differ from the InAs DB as well in their response to an applied perpendicular magnetic field. A typical zero-field current-voltage characteristic from device B at 4.5 K is shown in Figure 3.6a. The I-V curve of the same device in an 8 T field is shown in Figure 3.6b. Many of the general trends seen in all three samples are evident in this data. First, at high fields, a shoulder formed and moved rapidly to lower voltages as the applied magnetic field was increased. In addition, the peak of the I-V curve shifts, less dramatically, to higher bias. The peak current was observed to decrease with increasing applied field in samples A and B, while in sample C the peak current was constant until it was observed to increase in magnetic fields above 6 T. The peak current was never observed to change by more than $\approx 10\%$ of its zero-field value.

Figure 3.7 shows the peak and shoulder positions for each sample plotted as a function of applied field. The shoulder position was determined from a local minimum of the conductance curves for each sample. In cases where the broad peak appeared to contain two or more maxima, each position is shown. For all samples, the behavior is similar; little, if any, appreciable change at low fields until some critical field (B_{crit}) is exceeded, at which point both the main peak and subsidiary shoulder are observed to shift in voltage with changing applied magnetic field. The shift of the main peak saturates at high fields and is not seen to change in any of the samples at fields greater than 7 T. The shoulder position, however, appears to be a linear function of magnetic field in all three samples. For

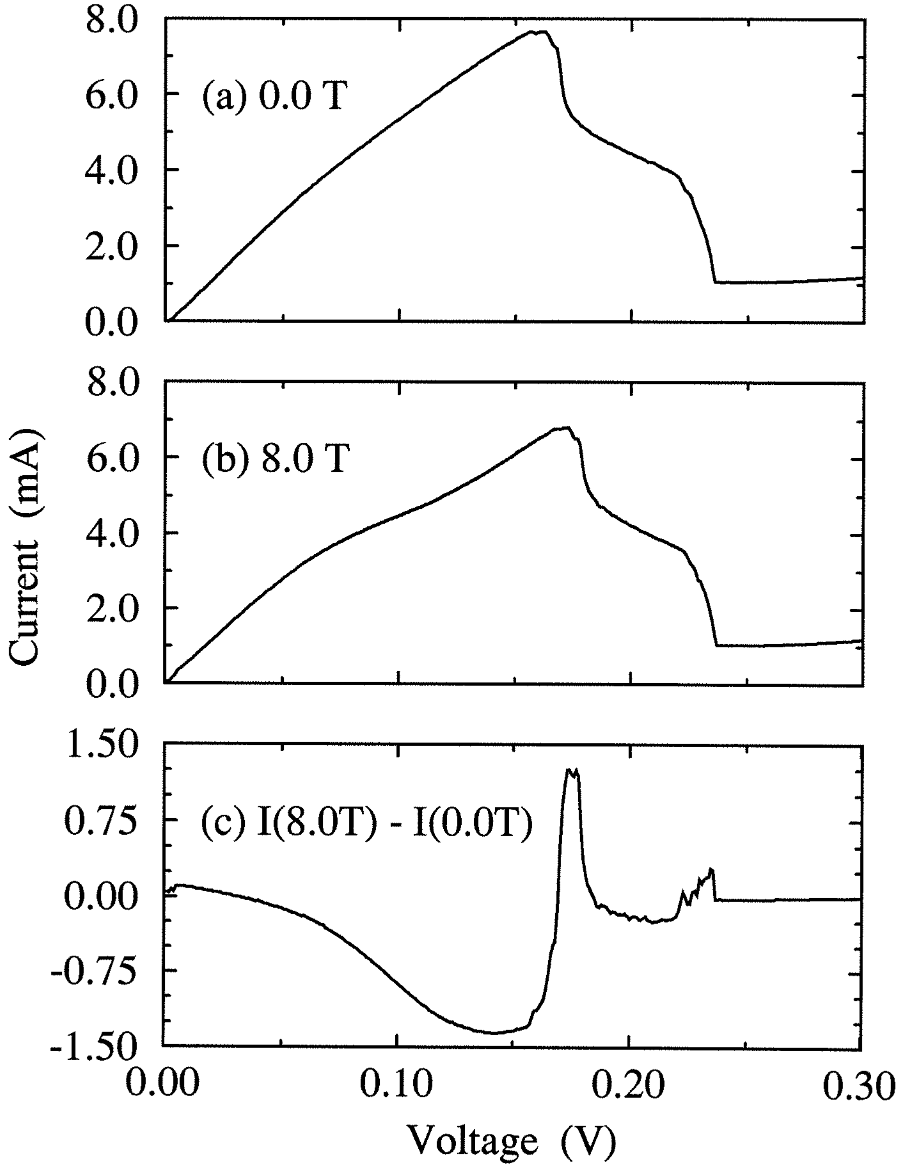


Figure 3.6: Example current-voltage curves for device B at (a) 0.0 T, and (b) 8.0 T. The magnetic field shifts the peak to higher voltage, while pushing out a shoulder to lower bias. In addition, the peak current has decreased. The position and magnitude of the valley current is unchanged by the magnetic field. The difference spectrum, shown in (c), is formed by the subtraction of the 0.0 T data from the 8.0 T data.

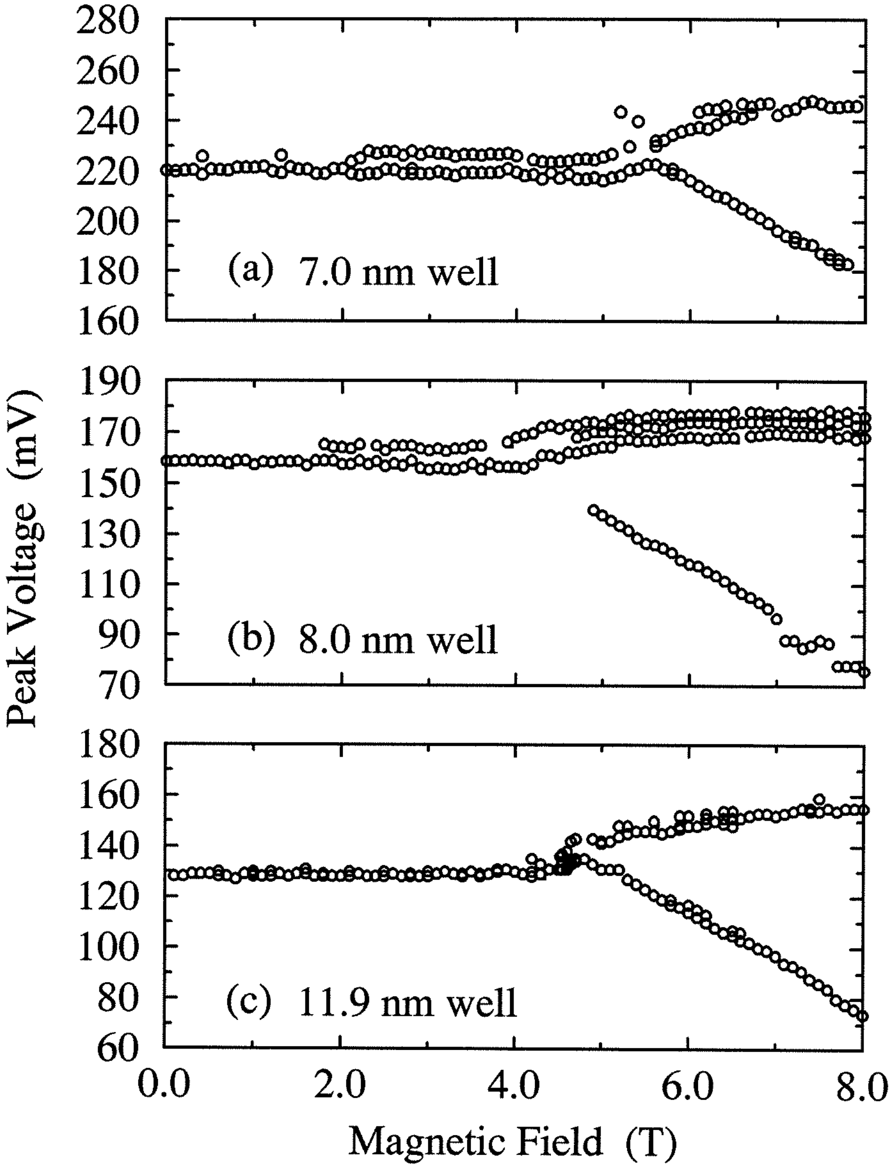


Figure 3.7: Peak and shoulder voltages as a function of applied magnetic field for devices A, B, and C, respectively. Little change is observed for fields below some sample-dependent critical value. Above this critical field, the peak voltage appears to saturate at high fields, whereas the location of the shoulder is a linear function of applied field for all three devices.

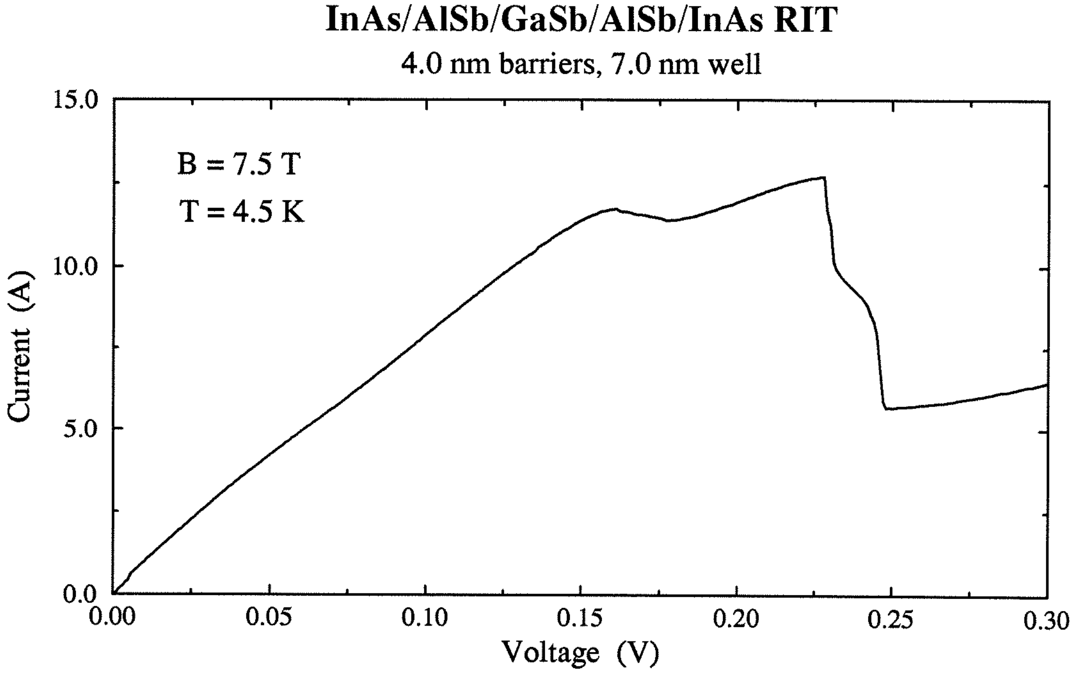


Figure 3.8: Observation of the shoulder in a RIT I-V form new second peak at low voltages. Presumably the contacts had lower series resistance and there was less inelastic leakage current in the device when this data was taken. Under less ideal conditions, as were normally experienced, the second peak was never observed as more than a broad, flat shoulder in the I-V curves at low voltages.

device A, a linear regression of the shoulder position yielded a slope of -18.9 ± 0.3 mV/T, with a squared correlation coefficient, r^2, of 0.996. Regressions of devices B and C found slopes of -21.3 ± 0.5 mV/T ($r^2 = 0.986$) and -18.8 ± 0.3 mV/T ($r^2 = 0.992$), respectively. The slopes for devices A and C are therefore statistically identical, whereas that of device B is more than 10% greater. Recent investigations of Sample A actually have shown a separate peak where a shoulder was previously observed, apparently as a result of less inelastic leakage current in the device. This data is shown in Figure 3.8. The movement of this second peak was found to be in good agreement with that previously noted for the shoulder.

3.3.2 Difference Spectra

Unlike the data from RMTS investigations of traditional intraband tunnel devices, much of the structure of the subbands in the GaSb well is missing from the data in Figures 3.6a, 3.6b and 3.7. Whereas, in intraband devices, multiple peaks, each associated with a separate subband, can be tracked with applied magnetic field to trace the $k_{\|}$ well dispersion, the opposite dispersions of the electrode and the well in interband devices, as well as the narrow voltage window in which NDR can occur, result in resonant transmission through several well states at the same bias. The resulting I-V curve is a complicated convolution of the contributions of the various well states. Consequently, it was useful to examine the difference spectra, such as shown in Figure 3.6c, of the three devices over the range of investigated fields, so as to separate the changes in conduction induced by the applied field from the complexities of the zero-field transport. Additionally, each of these spectra details the effect of the magnetic field more completely than the simple plot of peak positions shown in Figure 3.7. Each difference spectrum was formed by subtracting the zero-field current-voltage data for each device from the I-V curve at a given magnetic field. This spectrum may be positive or negative since the shift in carrier distribution caused by the magnetic field both adds and removes conduction channels at a given bias relative to the zero-field equilibrium distribution. The difference spectrum in Figure 3.6c is typical of all three samples. The NDR region is demarked by two peaks in the spectrum, each indicating additional current at those biases. When the peak or valley voltage shifts, as in this case with an 8.0 T field, the subtraction necessary to form the difference spectrum mimics the calculation of the finite element approximation to the local derivative, and thus in regions of large $|dI/dV|$ the difference spectrum has narrow peaks (or valleys, depending upon the sign of the peak shift). Comparison of the behavior of the difference spectra for the three samples highlights their differing subband structure and the effect of the magnetic field on resonant transport.

Figure 3.9 shows the low-field (1.0 and 3.0 T) difference spectra for all three samples. These magnetic fields are less than the critical fields of the three devices. The difference spectra indicate that there are field-induced changes in the current-voltage characteristics not fully evident in Figure 3.7. The peak of device B shifted to lower voltage at both 1.0 and 3.0 T, as evidenced by the negative spikes in the difference spectrum. The peak of device C was pushed to higher voltage, however. The device with the narrowest well, sample A, showed little shift in the peak voltage at either field, but a negative shift of the valley position at 3.0 T. All of these shifts are small, no more than 10 mV. In addition, the non-resonant current beyond the NDR region was affected in devices A and C. Device A showed a decrease in current at voltages greater than 0.3 V in the 3.0 T field, while device C showed a broad increase in conduction at voltages greater than 0.2 V in the 1.0 T field. In contrast, sample B showed no change in current at voltages beyond its NDR region (> 0.25 V). Samples A and C had peak-to-valley ratios significantly lower than sample B, and the additional transverse momentum provided by the magnetic field appears to have affected the inelastic transport responsible for these

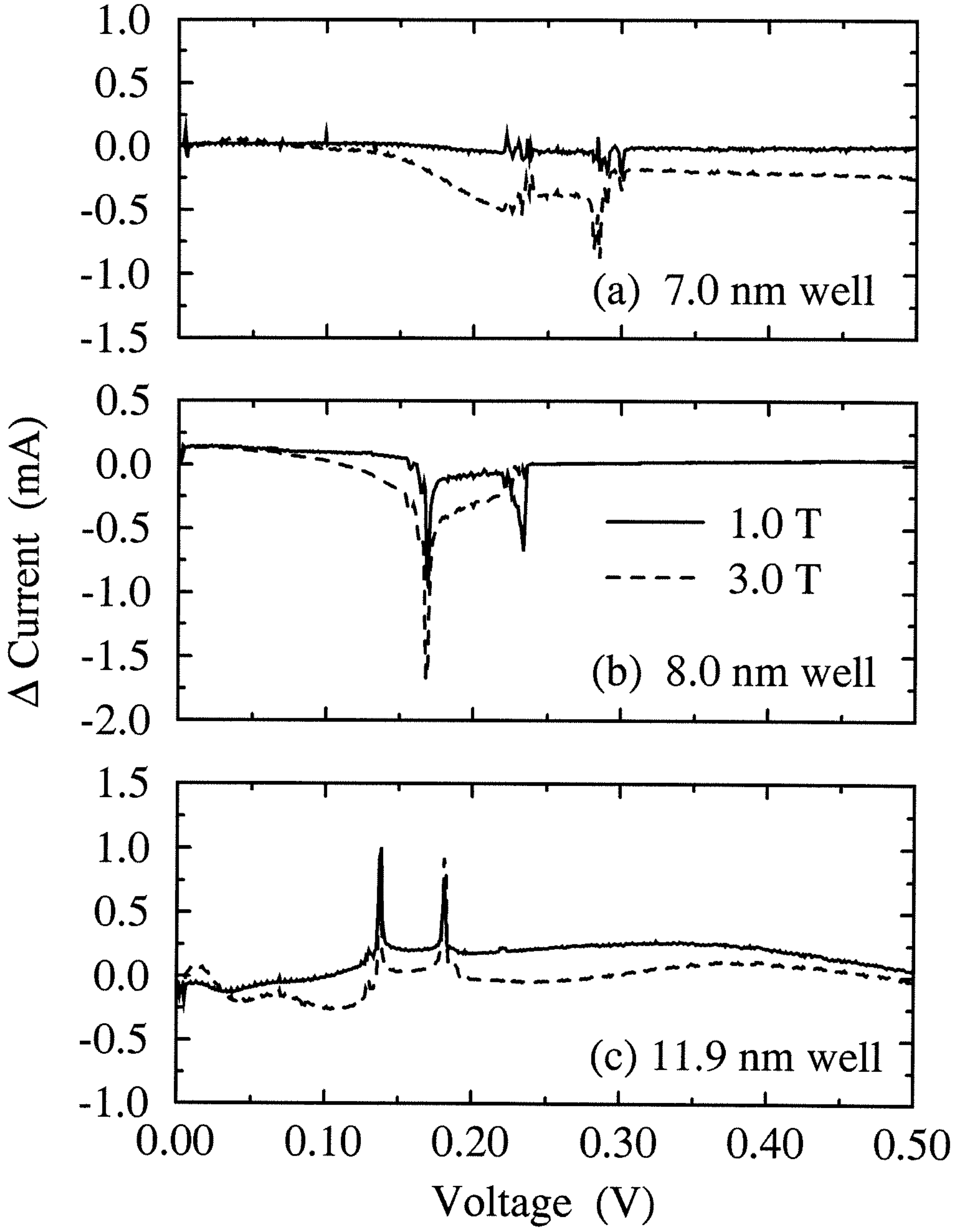


Figure 3.9: Difference spectra for samples A, B, and C, respectively, at both 1.0 and 3.0 tesla, below the critical field of each sample. Each spectrum is obtained by subtracting the zero-field current-voltage characteristic from the same data taken at the given magnetic field. Below the critical field, there are slight changes in the observed behavior of each sample not evident in the I-V peak position.

low PVRs.

The difference spectra for the three samples in high magnetic fields (at and greater than B_{crit}) are shown in Figure 3.10. It is in this regime that the behavior of each sample is distinct. Referring to Figure 3.10a, we first examine the data for device A. From the 4.0 T data, and the previous results shown in Figure 3.9, it is clear that both the peak and valley of device A are shifted to lower voltage at intermediate magnetic fields. At 5 T, this trend is continued, with additional loss of current across the entire NDR region. In fields greater than 5 T, however, there is a sudden shift of the peak to higher voltages, while the valley is similarly shifted lower, shrinking the NDR region by $> 50\%$ at 8 T. The difference lineshape at the peak and valley are both nearly symmetric, and at 5.0 T and greater, the spectra all meet and cross at 270 mV. For voltages less than the peak, there is a broad decrease in current, widening, and extending down to lower voltages with increasing magnetic field. In contrast, the data in Figure 3.10b for sample B shows significantly less change in transport across this sample due to the field. While the I-V peak is shifted to higher voltages by the field, the width of the difference spectra peak is roughly one third that of sample A, and less dependent upon the magnitude of the magnetic field. While not obvious from the figure, the difference peak height and width of the 8.0 T data are both slightly less than data taken at 7 T. In addition, the valley, while shifted to slightly higher voltages than the zero-field data, is independent of the field in this regime. The NDR region also shows little change in conduction, and the decrease in current at voltages below the peak is broader than that of sample A. The data for the sample with the widest well, device C, is shown in Figure 3.10c. Here, a peak in the difference spectrum occurs at 5 T, with no shift in the valley and no reduction in conduction for voltages below the peak. At greater than 5 T, the peak in the difference spectrum encompasses the entire NDR region, growing in magnitude and width with increasing field.

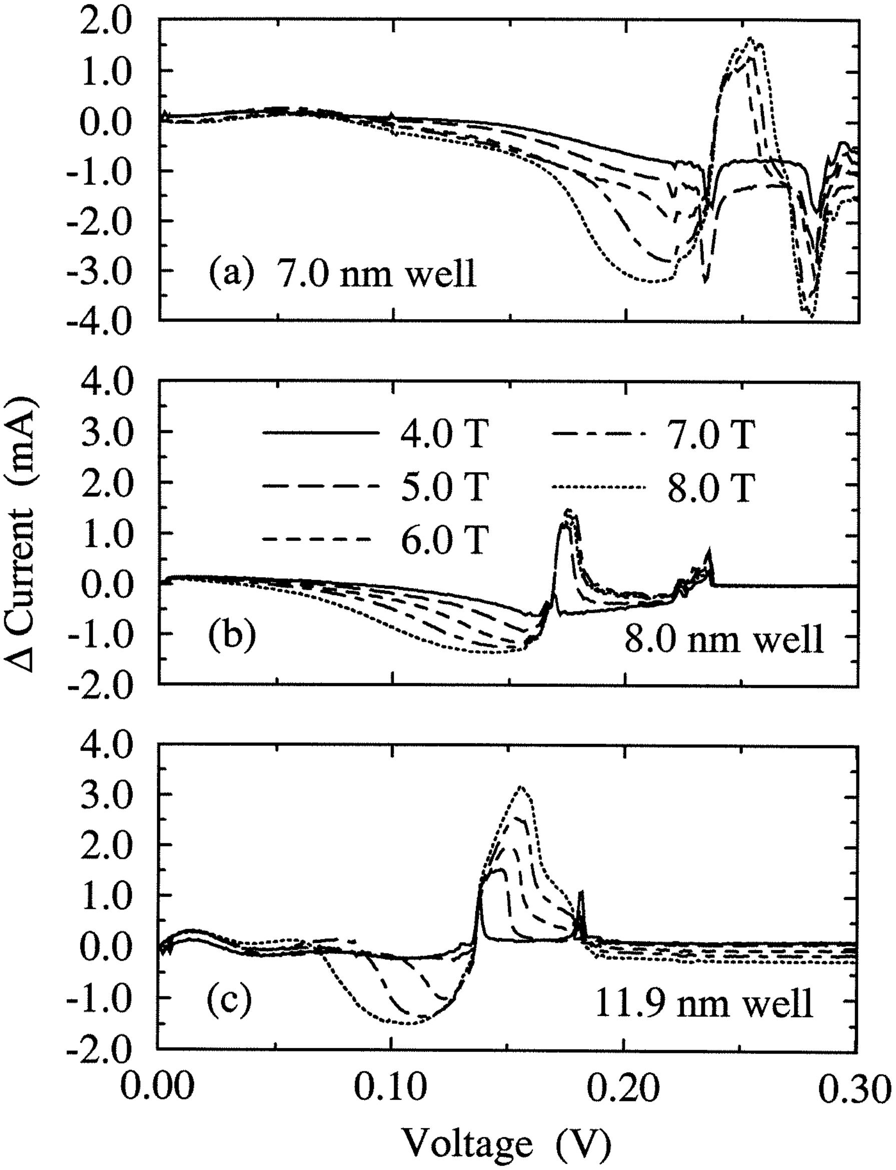


Figure 3.10: Difference spectra for samples A, B, and C, respectively, at 4.0, 5.0, 6.0, 7.0, and 8.0 tesla, at or above the critical field of each sample. Each spectrum is obtained by subtracting the zero-field current-voltage characteristic from the same data taken at the given magnetic field. In this regime, at or above the critical magnetic field, each sample displays distinct behavior. The transport characteristics of all three samples, however, show a far greater dependence on magnetic fields greater than B_{crit}.

The valley voltage remains constant throughout this range of magnetic fields, and there is a wide region of lower conduction just below the peak voltage. This latter behavior results in wide parabolic minima in the spectra. In contrast, the minima in Figure 3.10a are only semi-parabolic, and the broad features in Figure 3.10b are not at all quadratic. The lineshape of the difference spectra peak is also distinctly different than that of devices A or B. The peak is asymmetric, with a hunchback feature at lower voltages and a plateau-like shoulder above. There is also a slight maximum, roughly independent of field, at the lowest voltages, not present in either of the other devices. Attempts were made to model these difference spectra using a simple, intuitive picture of interband tunneling [13], but they were unsuccessful. As is often true for simulations in this system, even sophisticated, computationally-intensive 8-band models yield results with only qualitative resemblance to the actual data. Without a more complete understanding of all the physical processes involved, elastic and inelastic, further quantitative details of the parallel dispersions are unattainable from the difference spectra.

3.3.3 Critical Field and NDR I-V Peak

The unique, and unexpected, aspect of this data was the existence of a threshold, or critical, magnetic field, below which little change was affected by the field upon quantum transport in these devices, and above which peak shifts attributable to the simple RMTS theory are observed. While the critical field, B_{crit}, for each sample can be estimated from Figure 3.7, it is more easily obtained by looking at the integrated area of the difference spectra in the NDR region of the sample, as a function of applied field. This data is shown in Figure 3.11; the three critical fields are marked by arrows. Devices A and C show very sharp discontinuities in the region of their critical fields, between 5.1-5.2 T and 4.5-4.6 T, respectively. These values are in good agreement with the data in Figure 3.7. The data for device B shows no discontinuity, but rather a smooth transition, at a critical field of approximately 4.1-4.2 T. The critical field does not, therefore, seem to depend monotonically on well width. Samples A and C also displayed additional behavior at their respective critical fields not evident in sample B. Figure 3.12 shows the NDR region of the I-V curves for sample A at 5.1, 5.2, and 5.3 T. In this regime near the critical field, a significant second peak, located in the device's NDR region, was observed to appear abruptly as the field was increased above B_{crit}. A similar peak was observed at the critical field of sample C. The presence of this peak was wholly repeatable in each sample, across many mesas on each die. In both samples, this secondary peak was only evident over a narrow range of fields ($\approx 300G$), above which the main peak would broaden and envelope the other, as shown in Figure 3.12. In addition, in Sample A, for which this phenomena was most greatly studied, the critical field was observed to be dependent upon the polarity of sample bias, being almost 1.0 T greater for substrate positive bias ($B_{crit} \approx 6.2T$). This polarity-dependence was also repeatable across the die, and in each polarity the critical fields for different mesas of the same sample were identical to within the resolution of the magnet. The addition of series resistance to the measurement circuit merely shifted both peaks to higher voltage and did not decrease the magnitude of either. Additionally, as the temperature was increased, the width of the secondary peak increased, until around 55 K it was no longer distinguishable from the main resonance, as shown in Figure 3.13. For these reasons, we do not believe that the secondary peak is related to the circuit oscillations and hysteresis known to exist in the NDR region of resonant tunneling diodes [13, 14]. Given the nature of this peak, it seems apparent that it appears at exactly the critical field for a given sample. As its appearance was so abrupt (the exact field at which it appears can be determined to the resolution limit of the magnet, roughly 50 gauss), it is the cause of the discontinuities seen in Figure 3.11. The extra current from the new peak adds to the total integral over the NDR region discontinuously as the peak occurs. Both samples in which this second peak was observed (A and C) have discontinuities in the NDR integral, whereas sample B displays neither of these behaviors.

3.4 Analysis

3.4.1 $B > B_{crit}$

Because of the dichotomous nature of the data, these experiments are best analyzed separately for fields below and above the critical field. At fields larger than B_{crit}, the expected RMTS behavior was observed, and although the nature of interband tunneling obscures direct quantitative analysis of the valence subband structure, significant qualitative information may be inferred. Figures 3.14 and 3.15 show the theoretical parallel band structure of the three samples investigated experimentally, at 0.0 and 8.0 tesla, respectively. These dispersions were calculated in our group by Yixin Liu using an efficient, numerically stable 8-band $\vec{k} \cdot \vec{p}$ method generalized to include magnetic fields.[15] Figure 3.14 demonstrates the complexity of the well subbands in this device structure, even in the absence of magnetic fields. Despite the fact that several subbands are present, only one peak is evident in the current-voltage characteristic due to resonant transmission through multiple subbands at a single bias. As can be seen in Figure 3.15, the addition of the magnetic field complicates the situation even further. Clearly, the field has more effect than to simply shift the bands in k-space. There is a great deal of interaction between the subbands, in addition to lifting of the spin degeneracy. From the basis of the well width dependent shift of the band edges, theory predicts that $\langle \Delta x \rangle$ is $\approx 11\,\mathrm{nm} + b + w/2$, where b and w are the barrier and well widths, respectively.

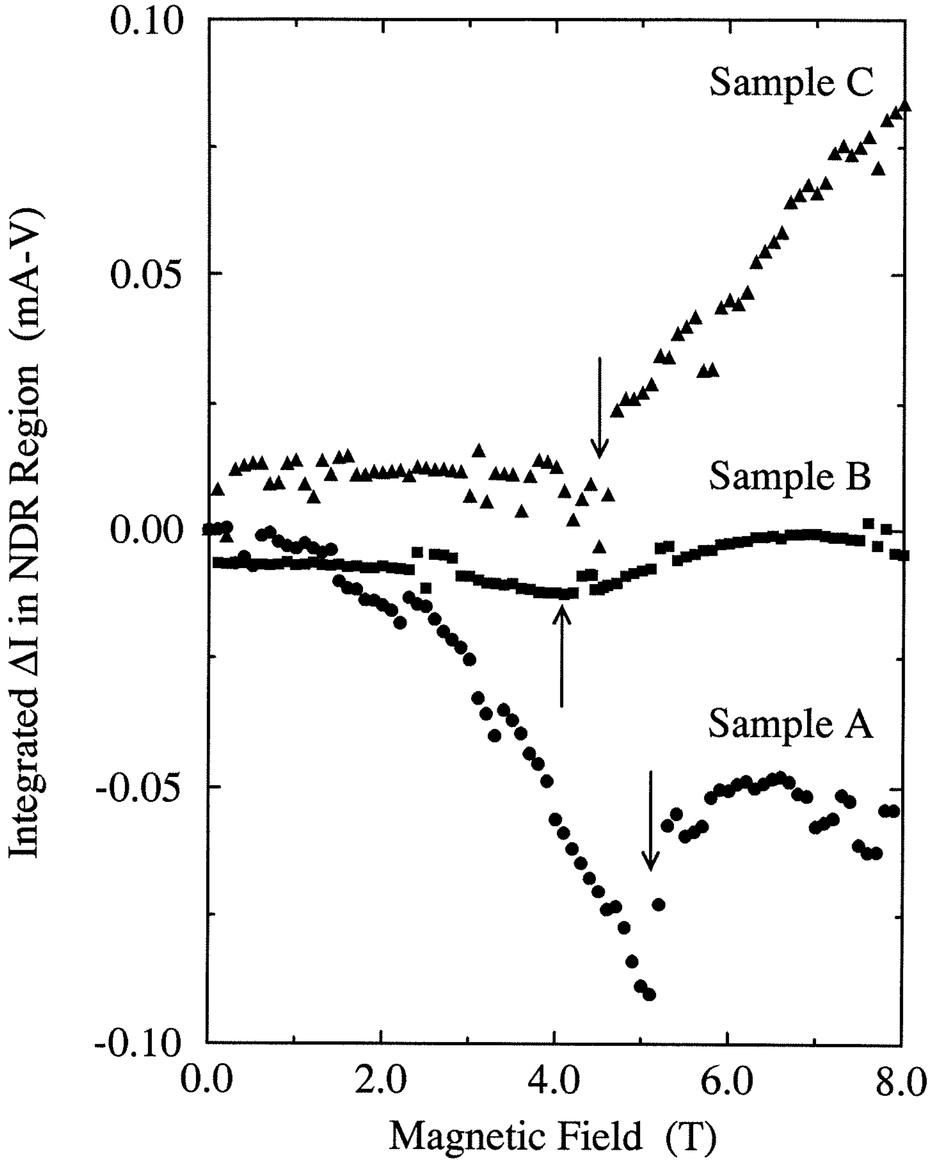


Figure 3.11: The change in device current integrated over the NDR region (as defined for each sample) plotted as a function of magnetic field. The apparent critical magnetic fields are indicated by arrows. The critical fields so obtained agree well with values derived from Figure 3.7. The abrupt nature of this transition is evident in samples A and C.

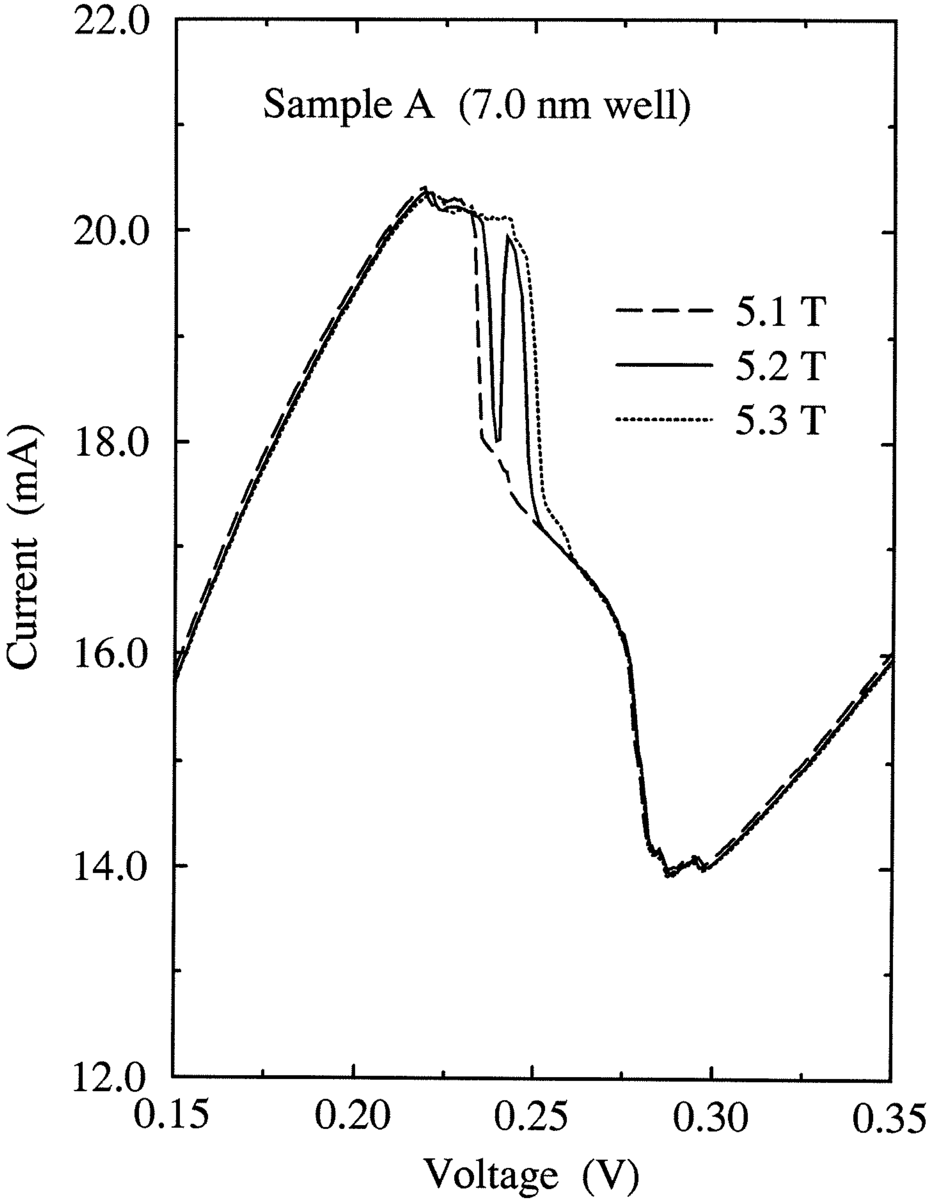


Figure 3.12: Current-voltage characteristics of device A near the critical magnetic field. In both devices A and C, a second, narrow peak was observed in the NDR region at B_{crit}. In fields only $\approx$ 300 gauss greater, the main peak was seen to widen and encompass it entirely.

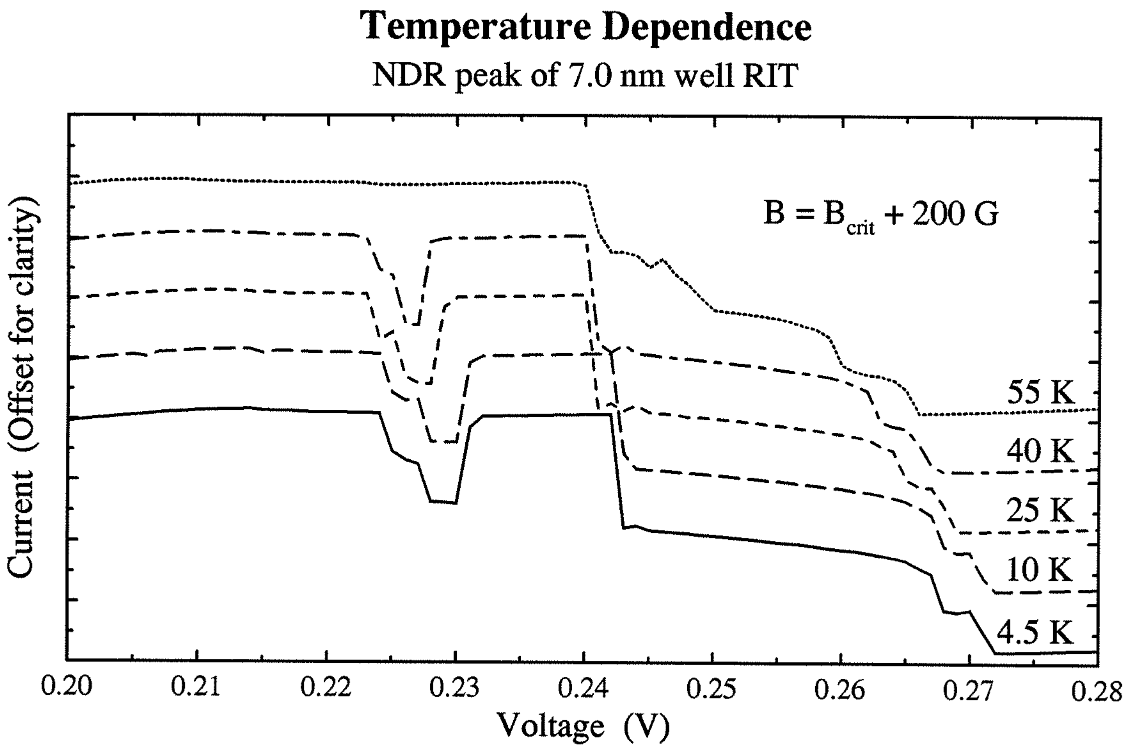


Figure 3.13: Observed temperature dependence of the I-V peak seen in the NDR region of device A at its critical field. The peak here is broader and wider than that seen in Figure 3.12 because the magnetic field was roughly 100G greater for these I-V traces. At a temperature of roughly 55 K, the secondary peak in the NDR region was indistinguishable from the main I-V peak.

In addition to the complexities of the subband structure, the quantum mechanical couplings of the subbands to the electrode states vary quite strongly between bands and, to a lesser degree, as a function of $k_{||}$. The calculated transmission coefficients for the narrow-well device, and their dependence upon applied field, are shown in Figure 3.16. Due to band mixing, at non-zero magnetic field, or equivalently, non-zero $k_{||}$, the states are strongly mixed, and the LH (light hole) and HH (heavy hole) labels are strictly valid only for the band symmetries at $k = 0$. From this data, however, it is clear that the greatest coupling from emitter occurs to the HH2 and LH1 bands, and not to the HH1, HH3, or HH4 bands shown in Figures 3.14 and 3.15. Consequently, these bands provide the greatest transmission channels for current. It is expected that the LH1 band would couple well to the conduction band states of the emitter, due to its symmetry. The HH2 couples more strongly than the other HH bands since it has a strongly positive dispersion for small $k_{||}$, and is therefore more conduction band-like. As a result, it is suggestive that the movements of the peak and shoulder in the experimental I-V data for fields greater than B_{crit} correspond in some manner to the HH2 (for the peak) and LH1 (for the shoulder) subbands, subject to the previously mentioned caveats. The movement of the shoulder was in fact consistent with a lighter mass negative dispersion, and that of the peak similarly consistent with a heavier, positive dispersion, as found in Figures 3.14 and 3.15 for the LH1 and HH2 subbands, respectively. While the thick barriers used in the experiments make the calculation of a theoretical I-V curve intractable, Figure 3.17 shows the results of an I-V curve calculation for a RIT with a 7.0 nm GaSb well and significantly thinner 1.5 nm AlSb barriers. Even at zero field, a shoulder is evident, and as the magnetic field is increased, the theory predicts behavior similar to that experimentally observed (compare to Figure 3.8). Additionally, the nature of the low voltage shoulder can be theoretically traced to transmission through the LH1 state. The theory also indicates that for the widest well, the LH1 and HH3 bands interact strongly, but it is doubtful that the experimental results should be capable of distinguishing this behavior from the general, low mass trend of the coupled bands. In general, for fields larger than B_{crit}, the experimental data appear well-understood, and indicate the presence of a low, positive mass dispersion (for $k_{\|}$ values accessible at these fields) and a separate high, negative mass dispersion at lower well energies. These observations agree well with the theoretically predicted behavior of the HH2 and LH1 states, respectively.

3.4.2 $B \leq B_{crit}$

The behavior of the experimental results at fields equal to and below the critical field is incongruent with the semi-classical analysis presented. The magnetic field has little outward effect on the experimental results at values less than B_{crit}. Additionally, the transition at the critical field is hyperabrupt, experimentally distinguishable down to the resolution limit of the magnet, roughly 50 gauss. The sharp distinction between behavior above and below the critical field, and the extra, narrow peak occurring within the NDR region of the device at the critical field raise a number of questions. This phenomenon has not been reported in earlier RMTS investigations of intraband devices, and was not seen with our apparatus in the investigation of transport in an InAs/AlSb double barrier heterostructure. It would therefore seem that this behavior is attributable to the interband nature of transport in the InAs/GaSb/AlSb system. Either the quasi-bound states in the GaSb well have infinite effective masses for some set of wavevectors, or the basic model itself is invalid for fields less than B_{crit}. The latter possibility introduces a number of potential explanations. As the RMTS theory assumes that a 2D state exists in the emitter accumulation region, one failure mechanism considered was lack of such a state at low magnetic fields. The experimental evidence disputes

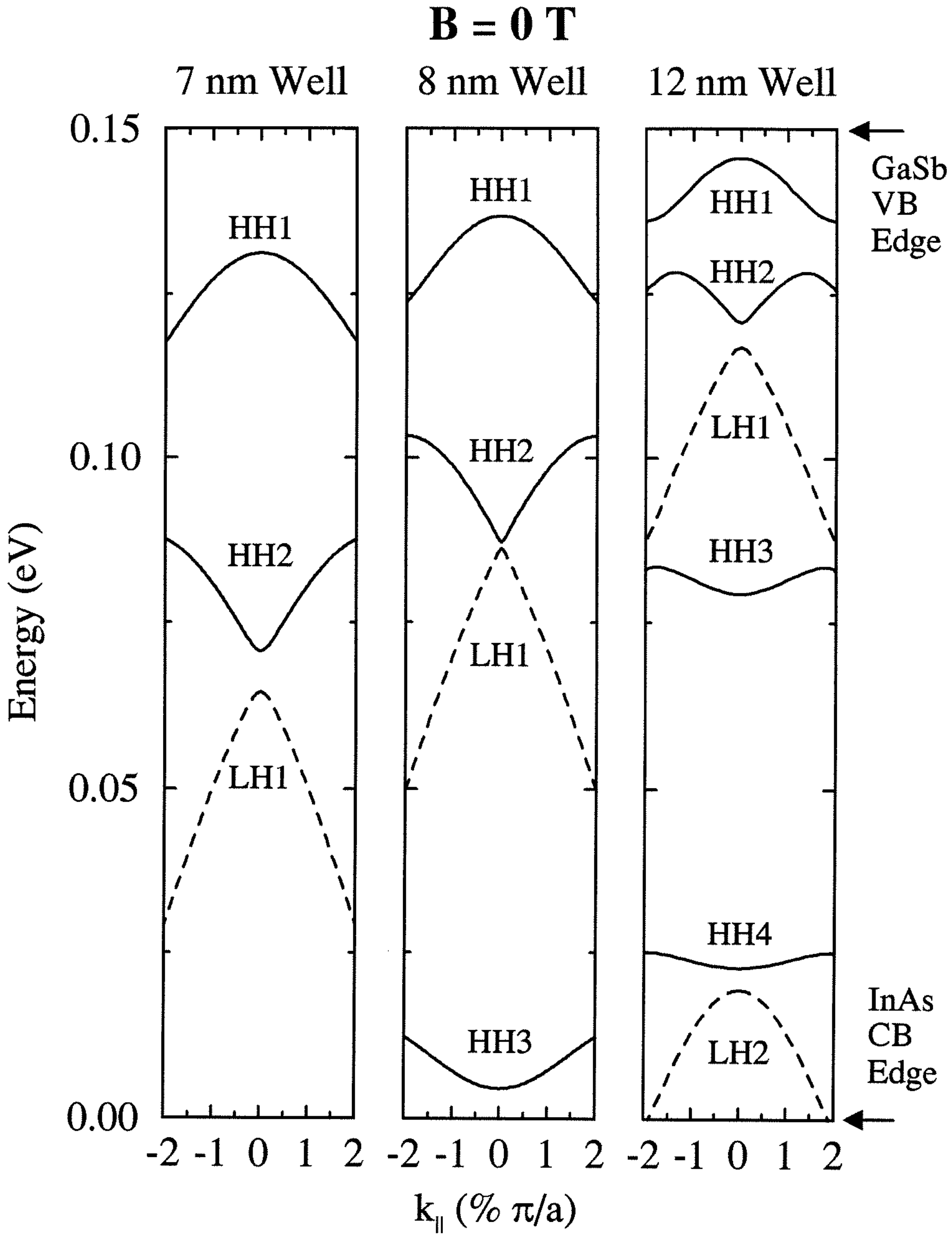


Figure 3.14: Theoretical zero-field well subband structure for RIT devices having GaSb well widths of 7.0, 8.0, and 12.0 nm, respectively, in the energy range between the InAs conduction band edge and GaSb valence band edge. Based upon the calculated quantum mechanical transmission coefficients (not shown), the majority of the resonant current tunnels through the HH2 and LH1 subbands.

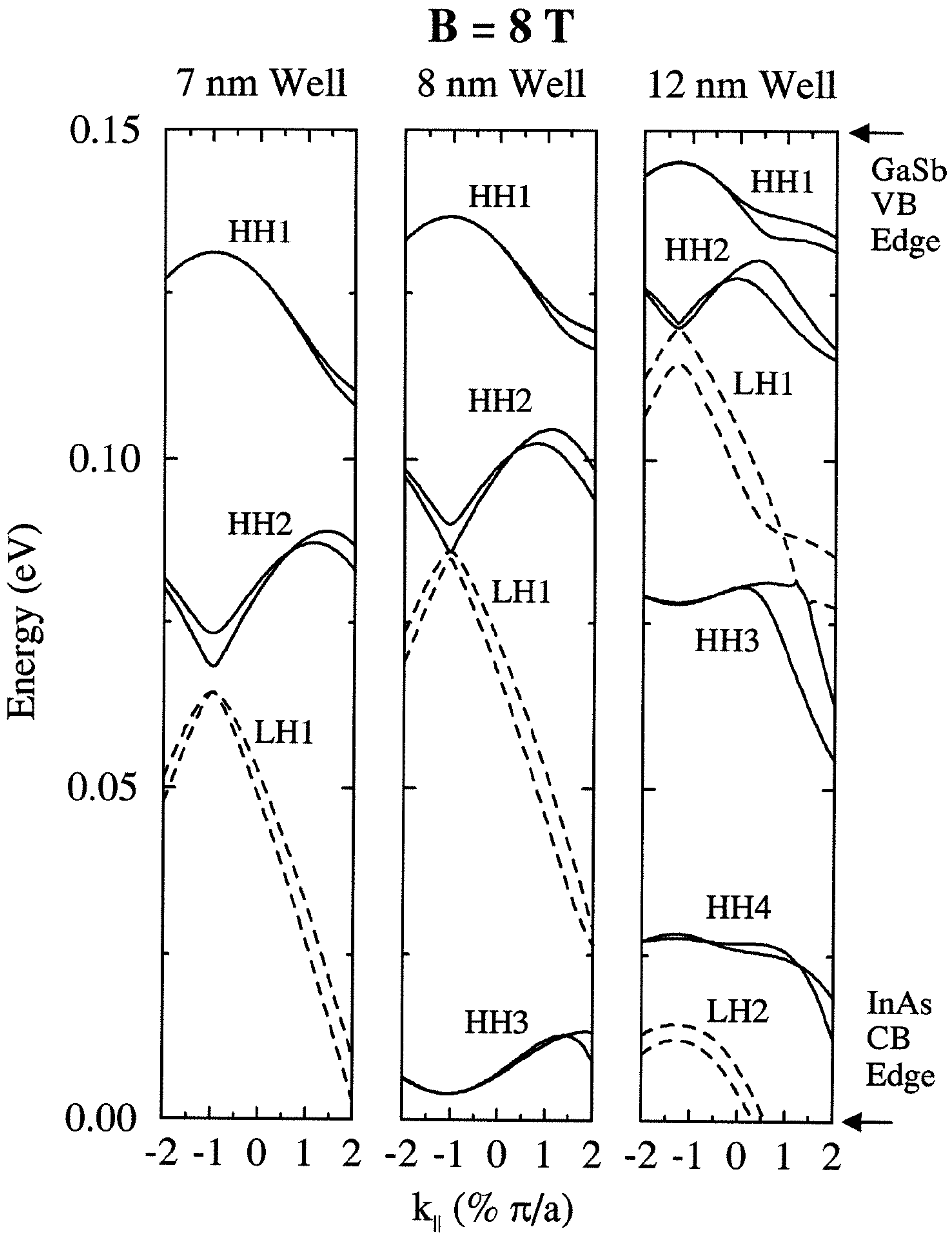


Figure 3.15: Theoretical well subband structure for RIT devices having GaSb well widths of 7.0, 8.0, and 12.0 nm, respectively, at a magnetic field of 8.0 tesla. The predicted shift of the well subbands is shown, along with the additional features induced by the applied field.

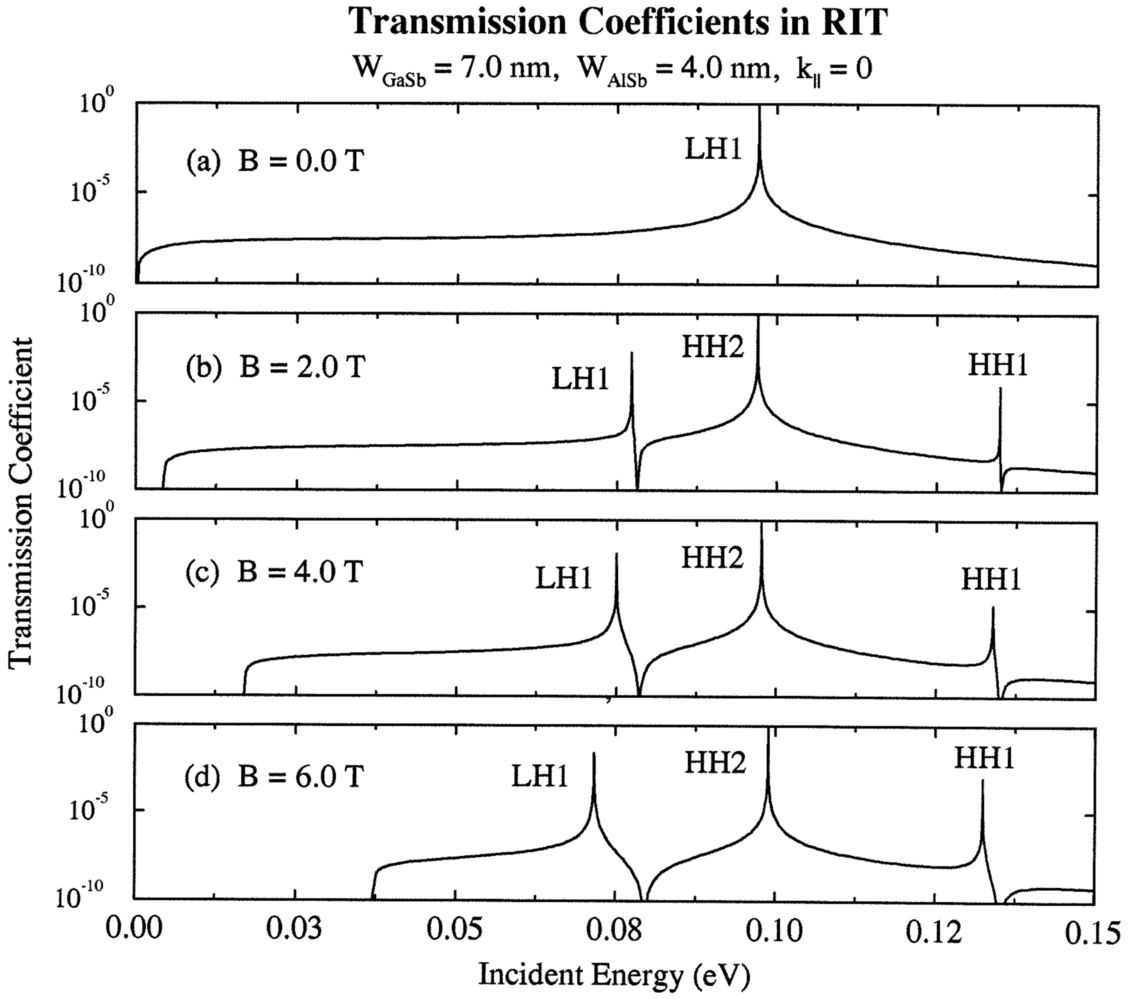


Figure 3.16: Calculated transmission curves for a RIT having a 7.0 nm-wide GaSb well, at zero, 2.0, 4.0, and 6.0 tesla in the RMTS geometry. At non-zero fields, additional resonances attributable to HH1 and HH2 states are present in addition to the LH1 resonance found at $B = 0$.

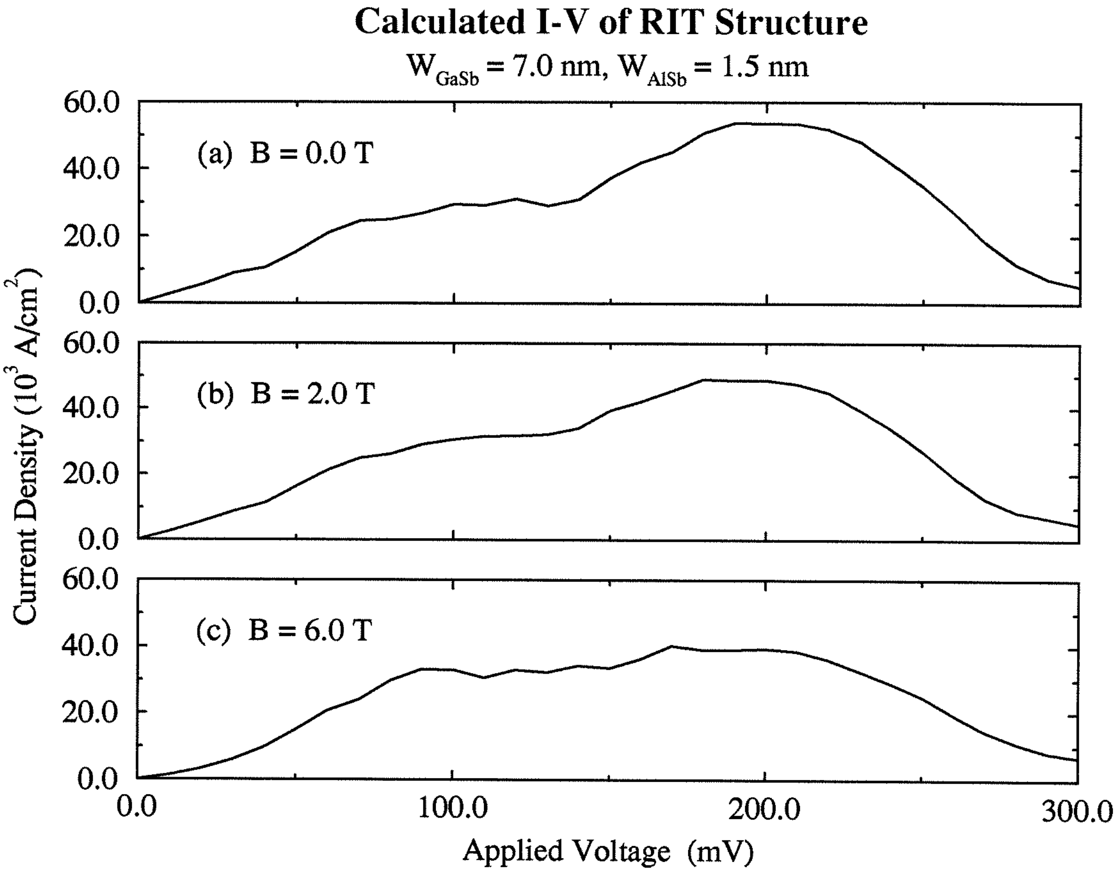


Figure 3.17: Calculated current density-voltage characteristics for a 7.0 nm-wide p-well RIT in zero, 2.0, and 6.0 tesla fields. This calculation was done for T = 77 K, and for a device having thinner barriers than those investigated (primarily to speed the calculation), but the results bear a strong resemblance to the experimental data.

this possibility, however. The InAs/AlSb double barrier exhibited no signs of such an effect, even though it had identical electrodes to the RIT structures. Also, the additional band bending induced by the transfer of charge between the GaSb well and InAs electrodes of a RIT should actually *increase* the likelihood and degree of quantization in the InAs emitter in comparison to the InAs/AlSb double barrier. Lastly, the mechanism by which such an abrupt transition may be attributed to a change in dimensionality of the emitter electrons remains unclear. It is similarly unlikely that this phenomenon is due to a classical cyclotron effect. For the n = 0 state and the InAs bulk mass, a 50 G change in field strength at 5 T (all that is necessary to distinguish the onset of the additional I-V peak) results in only a 0.13 Å change in cyclotron radius, and a 2.3 μeV alteration of the cyclotron energy. Furthermore, since this failure of the semi-classical RMTS model occurs over a large range of magnetic fields, a theory evoking the field-dependent cyclotron motion of the carriers cannot wholly explain the experimental data. The exact cause of this behavior is still unknown.

3.5 Conclusions

The complex nature of the interband tunneling mechanism was shown to lead to a complicated, and indirect, relationship between the observed effects of the magnetic field and the parallel subband dispersions of the quasi-bound states in the well. As interband transport at a given bias involves a broad spectrum of states in both the k_x and k_y directions, transport through several subbands concurrently, and the resultant effect of the field upon the peak position involves changes in the number of states tunneling, their respective energies, and their quantum mechanical transmission probabilities, the observed peak shift cannot be uniquely assigned to the well dispersion at a field-dependent wavevector, as it can for intraband transport. However, multiband theoretical calculations provide strong evidence that the peak and shoulder positions above B_{crit} correspond to the HH2 and LH1 subbands, respectively.

The data presented raise interesting questions as to the nature of transport in interband tunneling devices. The abrupt, catastrophic changes observed at the critical field are not understandable within the semi-classical theory invoked to explain RMTS experiments. Furthermore, simple explanations for deviations from this model fail to satisfy, or are disproved by, the experimental data. Further observations are warranted at higher magnetic fields and additional device geometries, both to study the peak shifts further, and to also possibly lend further insight into the nature of the observed critical magnetic field.

Bibliography

[1] R. K. Hayden *et al.*, Phys. Rev. Lett. **66**, 1749 (1991).

[2] J. P. Eisenstein, T. J. Gramila, L. N. Pfeiffer, and K. W. West, Phys. Rev. B **44**, 6511 (1991).

[3] S. Y. Lin *et al.*, Appl. Phys. Lett. **60**, 601 (1992).

[4] U. Gennser *et al.*, Phys. Rev. Lett. **67**, 3828 (1991).

[5] T. Osada, N. Miura, and L. Eaves, Solid State Commun. **81**, 1019 (1992).

[6] D. A. Collins *et al.*, Appl. Phys. Lett. **57**, 683 (1990).

[7] D. H. Chow *et al.*, in *Quantum-Well and Superlattice Physics III*, proceedings of the SPIE **1283** (Bellingham, WA, 1990).

[8] D. A. Collins *et al.*, in *Resonant Tunneling in Semiconductors: Physics and Applications*, edited by L. L. Chang, E. E. Mendez, and C. Tejedor (Plenum, New York, 1991).

[9] E. E. Mendez, in *Resonant Tunneling in Semiconductors: Physics and Applications*, *ibid.*

[10] E. E. Mendez, J. Nocera, and W. I. Wang, Phys. Rev. B **45**, 3910 (1992).

[11] E. E. Mendez, Surf. Sci. **267**, 370 (1992).

[12] T. Takamasu *et al.*, Surf. Sci. **263**, 217 (1992).

[13] D. A. Collins, Ph.D. Thesis, California Institute of Technology, 1993.

[13] V. J. Goldman, D. C. Tsui, and J. E. Cunningham, Phys. Rev. Lett. **58**, 1256 (1987).

[14] T. C. L. G. Sollner, Phys. Rev. Lett. **59**, 1622 (1987).

[15] Y. X. Liu, D. Z.-Y. Ting, and T. C. McGill, unpublished.

Chapter 4

Interband Tunneling Through Landau Levels

4.1 Introduction

The transport properties of quantum heterostructure devices are greatly altered by the application of magnetic fields at arbitrary angles to the epitaxial layers. Unlike the RMTS experiments detailed in Chapter 3, in this orientation the energy spectra, wavefunctions, and density of states are dramatically altered. The field component along the growth direction induces additional in-plane quantization, resulting in the formation of Landau levels. Experimentally, the effect of the field and the existence of Landau levels is observed through one of two distinct phenomena, either from the direct observation of individual peaks (corresponding to single Landau levels) in the I-V characteristic, or via Shubnikov-de Haas-like oscillations in one or more of the transport variables. While much of the historical study of low dimensional transport in semiconductor devices has involved the investigation of Shubnikov-de Haas (SdH) oscillations in 2D carrier gases,[1] these techniques were characteristically in-plane transport measurements, often using Si MOS technology for precise, adjustable control of the 2D carrier confinement and population. Only recently have perpendicular quantum transport studies observed Landau level tunneling and SdH oscillations. Such observation of magnetotunneling has furthered the understanding of quasi-2D states in heterostructure devices by revealing the relative energies of well subbands, their effective masses, and the coupling between well and electrode states.

As with the RMTS technique (c.f., Chapter 3), a vast majority of the observations of resonant tunneling through Landau levels have been in Type I unipolar DBH diodes. While many of these studies report individual Landau-level resonance peaks in I-V [2, 3, 12, 5] or conductance [6], still others observe SdH oscillations of a variety of transport variables, including current [7], conductance [8, 5], d^2I/dV^2 [10] and dI/dB [9]. The only reports of experiments in the InAs/GaSb/AlSb interband system have been for an n-well RIT [11], where conduction band-like Landau levels are formed in the well, and the barrierless InAs/GaSb/InAs quantum well-like structure [13]. This chapter presents experimental results of resonant Landau level tunneling in p-well RIT's having well widths of 7.0 nm and 11.9 nm in magnetic fields as high as 8 tesla aligned parallel to the tunnel current. Qualitatively different results are obtained for these samples with valence band-like Landau levels. The remainder of this chapter details these results and their interpretation. In Section 4.2 we examine the physics of resonant tunneling in this regime, and develop a simple theory for the explanation of the experimental results. The measurements are presented and analyzed in Section 4.3. A summary and conclusions are found in Section 4.4.

---

[1] The discovery of the quantum Hall effect [1] was made during the experimental study of SdH phenomena in Si MOS accumulation regions.

4.2 Theory

4.2.1 Landau Levels in Heterostructures

The theoretical treatment of Landau levels in heterostructures is simplified by the orientation of the magnetic field relative to the epitaxial layers. With the magnetic field applied solely along the growth direction, its effect is seen only in the growth plane. Consequently, the single-carrier Hamiltonian is separable into contributions parallel and perpendicular to the layers,

$$H_0 = H_\perp + H_{||}, \tag{4.1}$$

where, in the envelope function approximation, $H_{||}$ operates on the in-plane component of the single-particle wavefunction, and $H_\perp$ acts solely on the portion of the wavefunction dependent upon the growth direction, z, of the heterostructure. The band structure problem $H_0\psi = E(k)\psi$ therefore simplifies through a separation of variables. The single-particle envelope function in the z-direction, $H_\perp$, is written as,

$$H_\perp = \frac{p_z^2}{2m_\star} + V(z), \tag{4.2}$$

where $V(z)$ represents the variation in potential energy due to the heterointerfaces. In the transverse direction, $H_{||}$ incorporates the normal magnetic field through the in-plane vector potential, $\vec{A}$,

$$H_{||} = \frac{1}{2m^\star}\left(\vec{p}_{||} + e\vec{A}\right)^2 + g^\star \frac{\hbar e B}{m^\star}. \tag{4.3}$$

The spin contribution is ignored as negligible for materials with small band gaps, and subsequently small effective masses. The contributions of $H_\perp$ are identical to the zero-field problem, and the solution of $H_{||}$, in the effective mass approximation, becomes identical to the bulk problem discussed in Chapter 2. The energy structure for carriers in a simple quantum well therefore is of the form,

$$E(q,l) = E_q + (l + \frac{1}{2})\hbar\omega_c, \tag{4.4}$$

where q and l are the quantum well subband and Landau level indices, respectively, E_q is the zero-field, $k_{||} = 0$ energy of the qth quantum well subband, and $\hbar\omega_c = eB/m^\star$ is the well cyclotron energy.

Unfortunately, the effective mass approximation does not provide a good model for the complex parallel dispersions that exist in many quantum well problems, especially those involving hole-like states. To consider the problem in general, deviations from the effective mass approximation, $E \propto k^2$, must be developed. An arbitrarily complex scalar Hamiltonian may be expanded in a Taylor series in $\overline{k}$ about $\overline{k} = 0$ for an arbitrary magnetic field,

$$H(\overline{k}) = \sum_{\alpha=0}^{\infty} C_\alpha \overline{k}^\alpha, \tag{4.5}$$

where the C_α are the Taylor expansion coefficients, $\overline{k}$ is the generalized crystal momentum operator defined by $\overline{k_i} = (p_i + eA_i)/\hbar$, and $\vec{A}$ is the vector potential. With such a Hamiltonian, the Landau levels are no longer evenly spaced in energy and the eigenstates of the effective mass problem are not stationary states of the generalized Hamiltonian. As motivated in the derivation of Landau levels in the effective mass approximation (c.f. Chapter 2), raising ($a^\dagger$) and lowering (a) operators may be defined in terms of the in-plane components of $\overline{k}$,

$$a^\dagger = -\frac{1}{\sqrt{2}s}(\overline{k}_x + i\overline{k}_y) \tag{4.6}$$

$$a = -\frac{1}{\sqrt{2}s}(\overline{k}_x - i\overline{k}_y), \tag{4.7}$$

where $1/s = \sqrt{\hbar/eB}$ is the magnetic length. These operators applied to individual wavefunctions relate a given Landau level indexed by l to its adjacent neighbors with indices $l+1$ and $l-1$,

$$a^{\dagger}\,\psi_l = \sqrt{l+1}\,\psi_{l+1} \tag{4.8}$$

$$a\,\psi_l = \sqrt{l}\,\psi_{l-1} \tag{4.9}$$

$$a^{\dagger}a\,\psi_l = l\,\psi_l \tag{4.10}$$

Whereas the effective mass Hamiltonian may be rewritten in terms of the number operator, $\hat{N} = a^{\dagger}a = \overline{k}_x^2 + \overline{k}_y^2$, thereby leaving the state vector unaltered, arbitrary powers of $\overline{k}$, which may be reexpressed as various powers of $a^{\dagger}$ and a, $(a^{\dagger})^{\alpha}(a)^{\beta}$, thereby raise and lower Landau level wavefunctions to other indices, and allow for coupling between Landau levels. Solutions of the Schrödinger Equation will thus, in general, involve linear combinations of many Landau level basis functions. If the effective mass Hamiltonian is only slightly perturbed with additional powers of $\overline{k}$, solutions result which are primarily composed of a single Landau level, but which retain non-zero contributions from other Landau indices as well.

The simplest model for which deviations from the effective mass approximation may be investigated is the two-band model due to Kane [15]. Calculating the conduction and light hole bands simultaneously, while ignoring the contributions of other states, the model is first presented in the absence of a magnetic field. The Kane theory is, in fact, one of a number of realistic band structure models that may be derived from the more general $\vec{k}\cdot\vec{p}$ single-particle Hamiltonian. For a given band, indexed by m, we can express the wave function as the product of a plane wave factor and a Bloch function, $u_m(\vec{k},\vec{r})$, such that $\psi_m(\vec{k},\vec{r}) = e^{i\vec{k}\cdot\vec{r}}\,u_m(\vec{k},\vec{r})$. Substituting this form into the general Hamiltonian in Equation 4.3 results in the so-called $\vec{k}\cdot\vec{p}$ Hamiltonian,

$$\left[\frac{P^2}{2m_e} + \frac{\hbar\vec{k}\cdot\vec{p}}{m_e} + \frac{\hbar^2k^2}{2m_e} + V(\vec{r})\right] u_m(\vec{k},\vec{r}) = E_m(\vec{k})\,u_m(\vec{k},\vec{r}), \tag{4.11}$$

where $\vec{p}$ is the momentum operator and $E_m(\vec{k})$ is the energy eigenvalue for the

m^{th} band. Abstractly, this Hamiltonian requires solutions expanded in terms of all zone center Bloch functions indexed by m. Typically, an infinite number of such Bloch states are implicitly included in Equation 4.11, although the Hamiltonian may be block diagonalized to incorporate only the bands of interest. The details of the approximations which go into the adaptation of the $\vec{k} \cdot \vec{p}$ Hamiltonian to a finite problem are detailed in Appendix B. Following Fasolino and Altarelli [16], the in-plane eigensystem, $H_{||}\psi = E_{||}(k)\psi$, for the 2-band system consisting of the lowest conduction band and the light hole valence band may be written as,

$$H_{||}\psi = \begin{bmatrix} E_c & P(\overline{k}_x + i\overline{k}_y) \\ P(\overline{k}_x - i\overline{k}_y) & E_v \end{bmatrix} \begin{pmatrix} \psi_c \\ \psi_v \end{pmatrix} = E_{||}(k) \begin{pmatrix} \psi_c \\ \psi_v \end{pmatrix}, \tag{4.12}$$

where E_c, E_v, ψ_c, and ψ_v are the conduction and valence band edges and basis states, respectively. The parameter P is the Kane interband matrix element of the momentum operator, $iP \equiv \langle s|p_x|x\rangle$, where $|s\rangle$ and $|x\rangle$ are the spherically symmetric s-like and x-symmetric p-like zone-center states. The convenient consequence of the 2D Hamiltonian in this representation is the one-to-one mapping between the $\overline{k}_x$ and $\overline{k}_y$ dependence of $H_{||}$ and the raising and lowering operators, $a^\dagger$ and a. Therefore, for non-zero magnetic field $H_{||}$ may be expressed in terms of $a^\dagger$ and a rather than the non-commuting quantities of $\overline{k}_x$ and $\overline{k}_y$. Rewriting $H_{||}$ as,

$$H_{||} = \begin{bmatrix} E_c & -\sqrt{2}Psa^\dagger \\ -\sqrt{2}Psa & E_v \end{bmatrix}, \tag{4.13}$$

we find that eigenstates of the Hamiltonian may be written in this basis as,

$$\psi = \begin{pmatrix} c_1\, h_n \\ c_2\, h_{n-1} \end{pmatrix}, \tag{4.14}$$

where h_n and h_{n-1} are the harmonic oscillator wavefunctions for the $l = n$ and $l = n - 1$ Landau levels, and the constants c_1 and c_2 determine the fractional contribution of the conduction band and valence band basis states, respectively.

Here n is a parameter of the wavefunction in which the Landau level index, l, of each component may be expressed. Note that the constant c_2 must equal 0 for $n = 0$ as there is no $l = -1$ harmonic oscillator wavefunction. What is important to note is the unique nature of the eigenstates. In this model, the stationary states of the Hamiltonian involve more than one Landau level, but each band has a distinct, single Landau index. Solutions of this form are not unique to the 2-band model. In the 8-band $\vec{k} \cdot \vec{p}$ formalism (see Appendix B for discussion), the eigenvector in the Kramers basis set may be written as,

$$\psi = \left(c_1\, h_n,\ c_3\, h_{n-1},\ c_5\, h_{n+1},\ c_7\, h_{n+1},\ c_2\, h_{n+1},\ c_4\, h_{n+2},\ c_6\, h_n,\ c_8\, h_n\right), \tag{4.15}$$

where the odd-c components correspond to spin-up states and the even-c components to spin-down, and the vector is ordered so as to simplify the resulting Hamiltonian. The association between the basis components, $|1\rangle$ through $|8\rangle$, and the various physical bands is given by,

$$\begin{aligned}
|1\rangle &= |CB\!\uparrow\rangle &&= |\tfrac{1}{2}\,\tfrac{1}{2}\rangle_{CB} \\
|3\rangle &= |HH\!\uparrow\rangle &&= |\tfrac{3}{2}\,\tfrac{3}{2}\rangle \\
|5\rangle &= |LH\!\uparrow\rangle &&= |\tfrac{3}{2}\,-\tfrac{1}{2}\rangle \\
|7\rangle &= |SO\!\uparrow\rangle &&= |\tfrac{1}{2}\,-\tfrac{1}{2}\rangle_{VB} \\
|2\rangle &= |CB\!\downarrow\rangle &&= |\tfrac{1}{2}\,-\tfrac{1}{2}\rangle_{CB} \\
|4\rangle &= |HH\!\downarrow\rangle &&= |\tfrac{3}{2}\,-\tfrac{3}{2}\rangle \\
|6\rangle &= |LH\!\downarrow\rangle &&= |\tfrac{3}{2}\,\tfrac{1}{2}\rangle \\
|8\rangle &= |SO\!\downarrow\rangle &&= |\tfrac{1}{2}\,\tfrac{1}{2}\rangle_{VB},
\end{aligned} \tag{4.16}$$

where CB, HH, LH, and SO refer to the conduction band, heavy hole, light hole, and split-off bands, respectively, and ↑ and ↓ signify spin up and down states; in addition, the $|j\, m_j\rangle$ angular momentum state definitions are shown. While many authors [16, 17, 18] have demonstrated these interdigitated Landau ladders in multiband models, none has proposed a physically-intuitive model for their existence; all have merely used them to simplify the resulting calculation. The underlying physics itself, however, is important to the proper understanding of Landau-level tunneling in interband systems.

The Landau indices of components of the two- and eight-band eigenstates are not as arbitrarily related as they might first appear. Each eigenstate in fact represents a state for which the total angular momentum projected along the z-axis is a constant. To verify this assertion, it is important to consider the additional angular momentum induced in the z-direction by the magnetic field through the semi-classical cyclotron orbits in which the carriers travel. In general, the angular momentum $\vec{L}$ is defined as,

$$\vec{L} = \vec{r} \times \vec{p}. \tag{4.17}$$

To calculate the component of angular momentum along the applied field, consider the nature of semi-classical Landau orbits. As these orbits are circular, the components of $\vec{r}$ and $\vec{p}$ projected into the plane must be perpendicular. Therefore, ignoring the components of $\vec{r}$ and $\vec{p}$ parallel to $\vec{B}$, the magnitude of the cross product is the product of the vector lengths, and $\vec{L}$ is either parallel or anti-parallel to $\vec{B}$. As discussed in Chapter 2, the semi-classical orbital radii in real- and crystal momentum-space are given by,

$$R = \frac{1}{s}\sqrt{l + \frac{1}{2}}, \tag{4.18}$$

and,

$$K = s\sqrt{l + \frac{1}{2}}, \tag{4.19}$$

where $1/s$ is again the magnetic length, $\sqrt{\hbar/eB}$. Thus, the magnitude of the cyclotron orbital angular momentum is given by $|\vec{r} \times \vec{p}| = \hbar RK = (l + \frac{1}{2})\hbar$. To determine whether $\vec{L} \cdot \vec{B} = \pm 1$, one must examine the Lorentz force on a carrier,

$$\vec{F} = \frac{q}{m}\vec{p} \times \vec{B} = \dot{\vec{p}} = \hbar\dot{\vec{k}}. \tag{4.20}$$

Thus, for an electron with charge $q = -e$,

$$\dot{k}_x = -\frac{eBk}{m} \tag{4.21}$$

$$\dot{k}_y = \frac{eBk}{m}, \tag{4.22}$$

and the rotation is clockwise with respect to the field for negatively charged carriers. The angular momentum vector is consequently parallel to $\vec{B}$, and the z-component of the cross product $\vec{r} \times \vec{p}$ is positive.

The field induces a positive angular momentum along the z-axis (for electrons), and this additional macroscopic angular momentum component must be conserved in addition to the projection of the orbital angular momentum of the band and the carrier's intrinsic spin. As the macroscopic cyclotron orbital angular momentum is proportional to l, the conservation of all three angular momenta corresponds to the conservation of $f \equiv l + m_j$, where m_j is the projection of the sum of orbital and spin angular momenta along the z-direction. As all carriers in the Landau level indexed by l have the projected angular momentum $(l + \frac{1}{2})\hbar$ in the z-direction, no additional hybridization of the states due to the quantum mechanical addition of angular momentum must occur. Implicitly the eigenvalues of L_z are wholly specified, and the eigenvalues of L^2 are complicated by the rigors of angular momentum addition. For the current problem, the total angular momentum after the addition of the cyclotron motion is irrelevant, and will be ignored. The basis of the Hamiltonian may be relabeled with the Landau index, in addition to j and

m_j, as $|n : j\, m_j\rangle$, becoming,

$$\begin{aligned}
|1\rangle &= |CB\uparrow\rangle &&= |n : \tfrac{1}{2}\,\tfrac{1}{2}\rangle_{CB} \\
|3\rangle &= |HH\uparrow\rangle &&= |n-1 : \tfrac{3}{2}\,\tfrac{3}{2}\rangle \\
|5\rangle &= |LH\uparrow\rangle &&= |n+1 : \tfrac{3}{2}\,-\tfrac{1}{2}\rangle \\
|7\rangle &= |SO\uparrow\rangle &&= |n+1 : \tfrac{1}{2}\,-\tfrac{1}{2}\rangle_{VB} \\
|2\rangle &= |CB\downarrow\rangle &&= |n+1 : \tfrac{1}{2}\,-\tfrac{1}{2}\rangle_{CB} \\
|4\rangle &= |HH\downarrow\rangle &&= |n+2 : \tfrac{3}{2}\,-\tfrac{3}{2}\rangle \\
|6\rangle &= |LH\downarrow\rangle &&= |n : \tfrac{3}{2}\,\tfrac{1}{2}\rangle \\
|8\rangle &= |SO\downarrow\rangle &&= |n : \tfrac{1}{2}\,\tfrac{1}{2}\rangle_{VB},
\end{aligned} \tag{4.23}$$

where n is again a parameter of the eigenstate relating the Landau level index of each band to the state as a whole. For all components of the eigenvector, the sum $l + m_j$ is constant and equal to $n + \frac{1}{2}$. Thus, in the absence of scattering and other inelastic processes, tunneling may only occur between states of equal $l + m_j$, in the current approximation to the full single-particle Hamiltonian. Perturbations from this idealized model will in turn couple the states to additional Landau levels in each band.

4.2.2 Tunneling Through Landau Levels

The physics of Landau level tunneling in double barrier heterostructures depends critically upon the dimensionality of carriers in the emitter of the device. If the source of carriers tunneling into the well from the emitter is simply the bulk, 3D electrode, then 3D-2D tunneling occurs in the absence of a magnetic field, and 1D-0D tunneling with a B-field applied along the growth direction. Consequently, the well states define a more limited subset of energies and momenta, and transport properties therefore reflect the well states alone. In contrast, if carriers from the emitter tunnel from a quantized notch state, either 2D-2D ($B = 0$) or 0D-0D ($|\vec{B}| >$

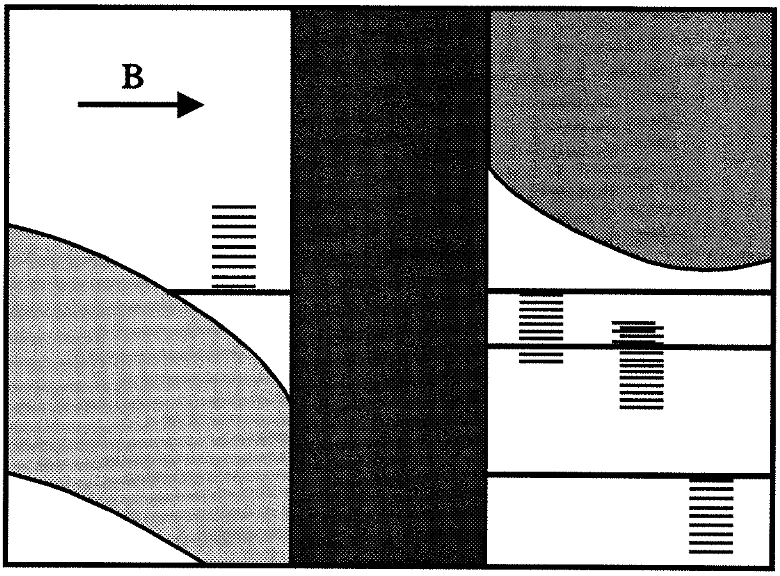


Figure 4.1: Emitter and well regions of a p-well RIT in a magnetic field perpendicular to the heterointerfaces. In this geometry, Landau levels are formed, some of which may deviate significantly from the evenly-spaced ladders predicted by effective mass theory, as shown in the well. Because of the intrinsic angular momenta of the well subbands, and the additional angular momentum provided by the magnetic field, carriers tunneling from the emitter subband to the well are expected to conserve the quantity $l + m_j$, or the total projection of angular momentum along the field.

$0, \vec{B} \,||\, \vec{J}$) tunneling occurs, and the properties of both the well and emitter states equally affect the transport characteristics. Unlike many experiments [6, 8, 9], there is evidence to support the existence of a 2D accumulation region in the interband RIT structures that have been studied, as detailed in Chapter 3. The existence of such a notch level is consistent with the large charge transfer expected between InAs and GaSb. Thus, only the specifics of 0D-0D Landau level tunneling will be developed.

Figure 4.1 shows the effect a magnetic field perpendicular to the heterointerfaces has upon the energy states in both the 2D emitter accumulation region and the well of an interband RIT device. Non-ideal Landau level behavior is indicated to highlight the potential complexity of the GaSb valence band-like well states. Each emitter and well quasi-bound state has a ladder of Landau levels associated with it, and consequently, the criteria for elastic resonant tunneling between such states is modified. While energy must still be conserved, the conservation of $k_{||}$ is alternatively stated as a condition on the initial and final state Landau level indices. Typically, as there is a direct mapping from Landau levels to $k_{||}$, this criterion is expressed as the conservation of Landau level index, $\Delta l = 0$ [7, 3, 11]. While this selection rule is valid for n-type Type I unipolar devices, as studied in References. [7] and [3], it is not so in general, in light of the additional angular momentum contributed by the cyclotron orbits of carriers in Landau levels. For n-type Type I devices, the general tunneling criteria $\Delta(l + m_j) = 0$, dictated by the conservation of the z-projection of angular momentum, is equivalent to the simple rule $\Delta l = 0$, as the initial and final states are inherently identical angular momentum states. In interband devices, where both conduction and valence bands are involved as initial and final states, and for hole tunneling in Type I structures, for which states of various m_j are simultaneously involved in transport, the m_j of initial and final states may differ in the presence of a magnetic field, and the full selection rule $\Delta(l + m_j) = 0$ must be applied. This conclusion is not inconsistent with the results of Mendez *et al.* [11], who, in their study of the I-V data for a GaSb/AlSb/InAs/AlSb/GaSb n-well RIT, cite the lack of additional ladders of peaks, corresponding to inter-Landau level coupling, as evidence that $\Delta l = 0$ is valid in interband systems. The general $\Delta(l + m_j) = 0$ rule allows a conduction band state of given l to couple to only select valence band states, which should lead to behavior such as that observed in their data.

While many studies observe Landau levels directly through a field-dependent ladder of peaks [2, 3, 12, 5] in I-V characteristics, still others observe oscillatory Shubnikov-de Haas behavior of either current or conductance [8, 7, 5]. Referring to Figure 4.1, the origin of SdH-like oscillations in perpendicular transport through RTD's may be argued for zero temperature in the effective mass approximation. With the field applied normal to the epitaxial layers, Landau levels form, and the Fermi level in the emitter lies at the energy of the highest occupied Landau level. As the field is increased, the spacing between the levels, $\hbar\omega_c$, and the maximum carrier occupation of each level, eB/h, both increase. The Fermi energy thus rises linearly with the applied field as the Landau level to which it is fixed increases in energy. When the last carrier falls from the uppermost Landau level l into the $l-1$ state, due to the increased occupation density of the lower states, the Fermi level discontinuously jumps down to the next rung of the Landau ladder. This shift in the Fermi level occurs whenever the total number of states in the lowest L Landau levels identically equals the total 2D charge density in the emitter notch state, $LeB/h = N_E^{2D}$. Thus, for fixed bias, and therefore fixed N_E^{2D}, the fields at which E_F jumps to a new level are periodic in $1/B$ with a fundamental frequency defined by $1/B_f \equiv e/(hN_E^{2D})$. In 2D transverse transport measurements, the resistivity shows an equivalent periodic behavior as scattering increases dramatically whenever the Fermi energy jumps down to a fully populated level. In perpendicular transport through an RTD, however, resonant elastic transport involves both well and emitter states and the SdH periodicity arises from the alternate alignment and misalignment of emitter and well states conserving energy and $l + m_j$. For a constant bias, at magnetic fields for which the emitter state at the Fermi energy and a well state of identical $l + m_j$ have equal energies, conductance is maximized. At intermediate fields, the tunnel current is decreased, and at sufficiently low temperature should fall to zero when no states meet the elastic criteria. However, at more elevated temperatures, Landau level broadening results in non-zero tunneling conductance minima. The periodicity of Landau level tunneling as a function of the applied field is therefore not determined from the oscillation of the Fermi energy in the emitter alone, but rather by the structure of both the emitter and well Landau level ladders.

The fields for which the tunneling conductance is maximized may be derived from simple considerations of energy and longitudinal momentum conservation. First, consider resonant tunneling in a RIT p-well device obeying the simplified selection rule, $\Delta l = 0$. In the effective mass approximation, the energies of the emitter and well states are,

$$E = E_E + \left(l + \tfrac{1}{2}\right) \hbar\omega_c^E, \qquad \text{(Emitter)} \tag{4.24}$$

and,

$$E = E_W - \left(l + \tfrac{1}{2}\right) \hbar\omega_c^W, \qquad \text{(Well)} \tag{4.25}$$

respectively, where E_E and E_W are the emitter and well $B = 0$, $k_{\|} = 0$ band edges, and $\hbar\omega_c^E$ and $\hbar\omega_c^W$ are the emitter and well cyclotron energies, respectively. For tunneling processes conserving both energy and Landau index, these energies are equal, and conductance maxima occur for,

$$\Delta E \equiv E_W - E_E = \left(l_F + \frac{1}{2}\right) \frac{e\hbar B}{\mu^\star}, \tag{4.26}$$

where l_F is the index of the emitter Landau level in which the Fermi level resides, and $\mu^\star$ is the reduced mass defined by $1/\mu^\star \equiv 1/m_E^\star + 1/m_W^\star$. Therefore, the inverse field at which the conductance is maximized by transport through Landau level l is given by,

$$\frac{1}{B_l} = \frac{e\hbar}{\Delta E \mu^\star} \left(l + \frac{1}{2}\right), \tag{4.27}$$

and the fundamental frequency of the SdH conductance oscillations is,

$$\begin{aligned} \frac{1}{B_f} &= \frac{1}{B_{l+1}} - \frac{1}{B_l}, && (4.28) \\ &= \frac{e\hbar}{\Delta E \mu^\star}. && (4.29) \end{aligned}$$

The fundamental field thus depends upon the mass and band edge of both emitter and well states.

The situation is more complex when the full z-component of angular momentum is conserved such that $\Delta(l+m_j)=0$. For the RIT structures studied, the incident carriers all lie in the conduction band (CB) of the InAs emitter, and therefore correspond to angular momentum states of $|\frac{1}{2}\pm\frac{1}{2}\rangle$ (the split-off bands will be ignored for the remainder of the discussion, and all $|\frac{1}{2}\pm\frac{1}{2}\rangle$ states will represent conduction band states and not split-off valence band states). In the absence of an applied field, at $k_{\|}=0$ the $|\frac{1}{2}\,\frac{1}{2}\rangle$ state in InAs only couples to the light hole (LH) $|\frac{3}{2}\,\frac{1}{2}\rangle$ state in the GaSb well. It is forbidden for a tunneling electron to end up in a heavy hole (HH) $|\frac{3}{2}\pm\frac{3}{2}\rangle$ state. However, for $k_{\|}\neq 0$, the LH and HH states mix, becoming mixed angular momentum states. The LH and HH labels then, in general, refer to energy bands and not specific angular momentum eigenstates. The $|HH\uparrow\rangle$ and $|HH\downarrow\rangle$ bands at non-zero $k_{\|}$ may be written as linear combinations of several of the zone center basis states, such that, for $\beta_{\uparrow}>\beta_{\downarrow}$, and $\alpha\gg\beta_{\uparrow},\beta_{\downarrow}$ for $|k_{\|}|\ll\frac{2\pi}{a}$,

$$|HH\uparrow\rangle=\alpha|\frac{3}{2}\,\frac{3}{2}\rangle+\beta_{\uparrow}|\frac{3}{2}\,-\frac{1}{2}\rangle+\beta_{\downarrow}|\frac{3}{2}\,\frac{1}{2}\rangle, \tag{4.30}$$

and, $\delta_{\downarrow}>\delta_{\uparrow}$, and $\gamma\gg\delta_{\uparrow},\delta_{\downarrow}$ for $|k_{\|}|\ll\frac{2\pi}{a}$,

$$|HH\downarrow\rangle=\gamma|\frac{3}{2}\,-\frac{3}{2}\rangle+\delta_{\downarrow}|\frac{3}{2}\,\frac{1}{2}\rangle+\delta_{\uparrow}|\frac{3}{2}\,-\frac{1}{2}\rangle. \tag{4.31}$$

Thus, for a non-zero $k_{\|}$, CB states in the InAs emitter of a RIT can couple to the HH band through the LH-like $m_j=\pm\frac{1}{2}$ components. With the magnetic field on, tunneling electrons conserve $l+m_j$ and CB states may couple to either LH-like ($m_j=\pm\frac{1}{2}$) or HH-like ($m_j=\pm\frac{3}{2}$) states with a change in Landau index $\Delta l=-1,0,+1,+2$ for the $|\frac{3}{2}\,\frac{3}{2}\rangle$, $|\frac{3}{2}\,\frac{1}{2}\rangle$, $|\frac{3}{2}\,-\frac{1}{2}\rangle$, and $|\frac{3}{2}\,-\frac{3}{2}\rangle$ states, which in turn correspond to the primary components of the HH↑, LH↓, LH↑, and HH↓ bands, respectively. Writing the energies in the emitter and well as,

$$E = E_E + \left(l + \tfrac{1}{2}\right) \hbar\omega_c^E, \qquad \text{(Emitter)} \tag{4.32}$$

and,

$$E = E_W - \left(l + \Delta l + \tfrac{1}{2}\right) \hbar\omega_c^W, \qquad \text{(Well)} \tag{4.33}$$

the magnetic fields corresponding to conductance maxima of Shubnikov-de Haas oscillations are given by,

$$\frac{1}{B_l} = \frac{e\hbar}{\Delta E \mu^\star}\left(l + \frac{1}{2}\right) + \frac{e\hbar}{\Delta E m_W^\star}\Delta l, \tag{4.34}$$

and the fundamental field of the oscillations is defined by,

$$\frac{1}{B_f} \equiv \frac{1}{B_{l+1}} - \frac{1}{B_l} = \frac{e\hbar}{\Delta E \mu^\star}. \tag{4.35}$$

Thus, the frequency of the Shubnikov-de Haas oscillations remains the same under the selection rule $\Delta(l + m_j) = 0$, but the phase is altered by an amount, $e\hbar\Delta l/\Delta E m_W^\star$, dependent upon the change in Landau index (determined by the final hole subband eigenstate), the effective mass of the hole subband, and the zero-field energy difference of the well and emitter subband edges.

4.3 Experiment

4.3.1 Measurement Technique

Two p-well RIT samples were investigated in magnetic fields as great as 8 tesla aligned parallel to the tunnel current. These samples, having well widths of 7.0 nm and 11.9 nm, respectively, are the same devices as samples A and C in Chapter 3. The specifics of the device geometries were summarized in Table 3.1. A 16.0 nm well InAs/AlSb double barrier RTD was investigated as well, to contrast the behavior of otherwise-identical Type I and Type II broken-gap structures. The details of device fabrication and nominal zero-field characteristics are summarized in Section 3.2.2. The alignment of the sample relative to the field was verified to within 0.5° by an adjacent Hall probe. All measurements detailed in this chapter were made at a temperature of 4.5 K. Several experiments of transport at ≈ 1.6 K were conducted, but no significant deviations were observed from the data taken at higher temperature. All conductances were derived from the I-V data by fourth-order nine-point Savitzky-Golay [19] numerical differentiation. The I-V data was taken in 1 mV steps, and discrete conductance values were obtained at 5 mV intervals.

4.3.2 Current-Voltage Characteristics

Figure 4.2 shows the I-V characteristics of the 11.9 nm-wide well sample at 2 and 6 tesla, respectively. The behavior of the 7.0 nm well device was similar. While some change to the I-V curve is noticeable, especially at 6 tesla, the most profound changes in behavior are noticeable in the conductance at a fixed voltage for magnetic fields between these two values. For $B < 2$ T, little effect was observed directly in the I-V curve, except for a subtle shift of all features to higher biases as the field strength was increased. Above 6 tesla, the NDR region narrows rapidly, and disappears entirely by 8 T, presumably due to the effects of magnetoresistance in the contact layers. As the back contacts of these devices were not on the opposite face of the wafer, but instead surrounded the perimeter of each device, current from the back contact to the mesas experiences a maximum magnetoresistance in this geometry. This field-dependent resistance was unavoidable, but does not appear to significantly impact the experimental results. The features seen in the NDR region in Figure 4.2b, labeled A through D, may be the signature of Landau level formation at this field, although their presence in the NDR region limits the degree of confidence which may be assigned to them. The entire structure of the

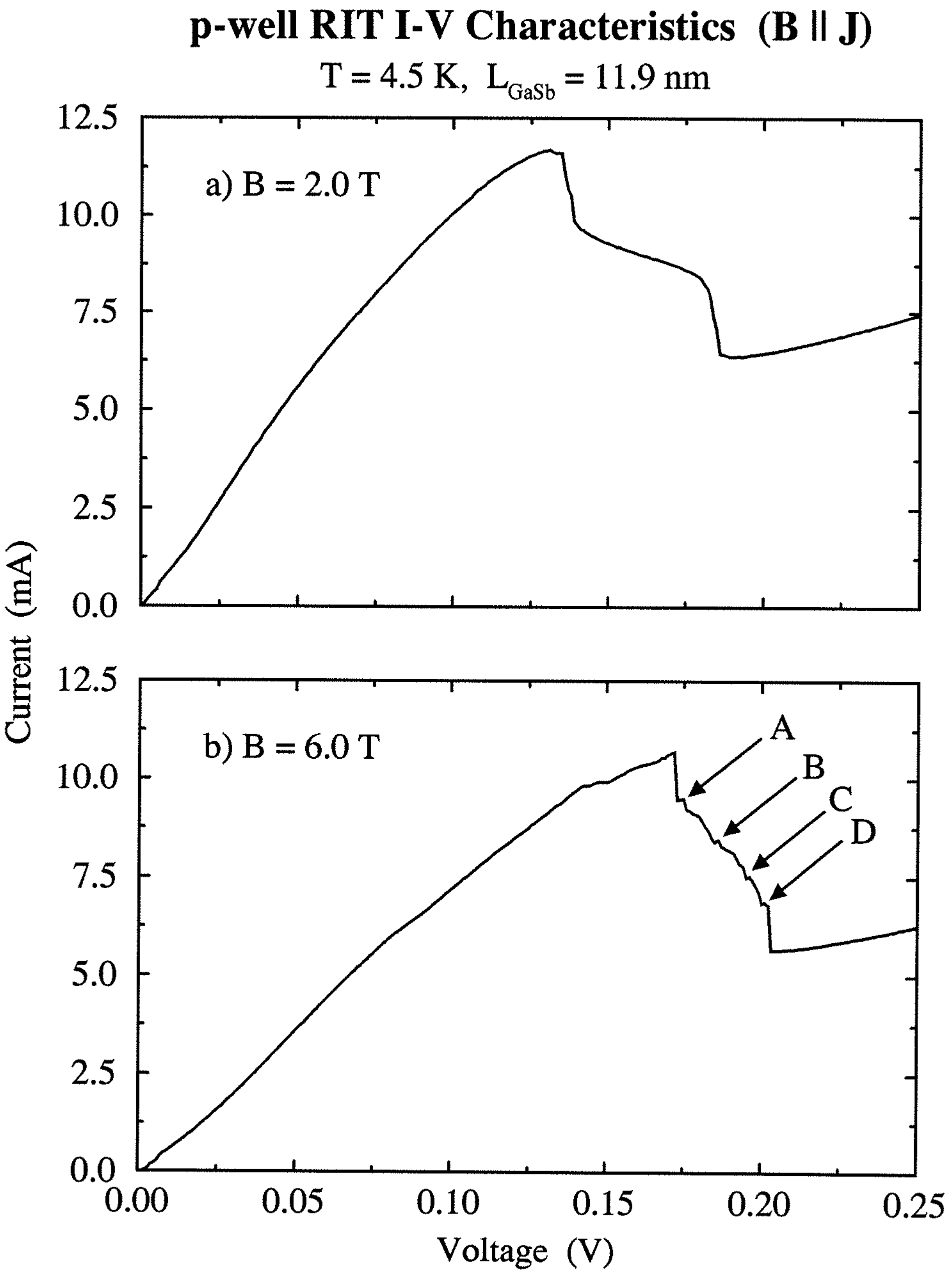


Figure 4.2: Representative current-voltage characteristics of the 11.9 nm well RIT device at 2 and 6 tesla. These fields represent the extremal values for which Shubnikov-de Haas oscillations of the conductance were observed. The uniform spacing between the features labeled A through D on the $B = 6$ T I-V may be signatures of Landau levels.

NDR region is thought to be dominated by circuit instabilities due to the negative differential RTD resistance [20, 21]. However, the regular spacing of these small peaks (the A-B and B-C spacing is 10 mV, while the C-D spacing is 5 mV) is suggestive of a Landau ladder with $\hbar\omega_c = 10$ mV. In such an interpretation, peak D would correspond to the band edge of the quasi-bound state, and peaks A, B, and C to the $l = 2$, 1, and 0 Landau levels, respectively. This interpretation is consistent with the presumed structure of the highest energy valence band state in the GaSb well (the HH1 state), as the higher Landau levels should be accessible at lower voltage, due to the negative dispersion. With sufficient level broadening, the band edge might be experimentally accessible, and the assignment of peak D to such is consistent with the abrupt onset of the valley current at slightly higher biases. A 10 mV cyclotron energy at 6.0 tesla is consistent with an effective mass of $0.07\,m_e$, which is a reasonable number for the HH1 state (the 8-band $\vec{k}\cdot\vec{p}$ model of Liu *et al.* [14] predicts a value of $\approx 0.12\,m_e$). Unfortunately, the B-dependent width of the NDR region and the related change in effective experimental resolution do not allow the field dependence of these peaks to be observed with enough accuracy to unequivocally attribute them to individual Landau levels.

4.3.3 Oscillatory Tunneling Conductance

Without additional features in the experimental I-V characteristics, it was necessary to look at the tunneling conductance, dI/dV, to observe the effect of the field on quantum interband transport. Figures 4.3 and 4.4 plot the conductance of the 7.0 nm- and 11.9 nm-wide well samples, respectively, against inverse magnetic field at fixed biases of 25, 50, 75, and 100 mV. For clarity, arbitrary offsets have been added to the data. In both samples, the peak-to-peak magnitudes of the oscillations are no greater than 10% of the local average conductance at any given field. Superimposed on the 7.0 nm well data, however, are a strong field-dependent

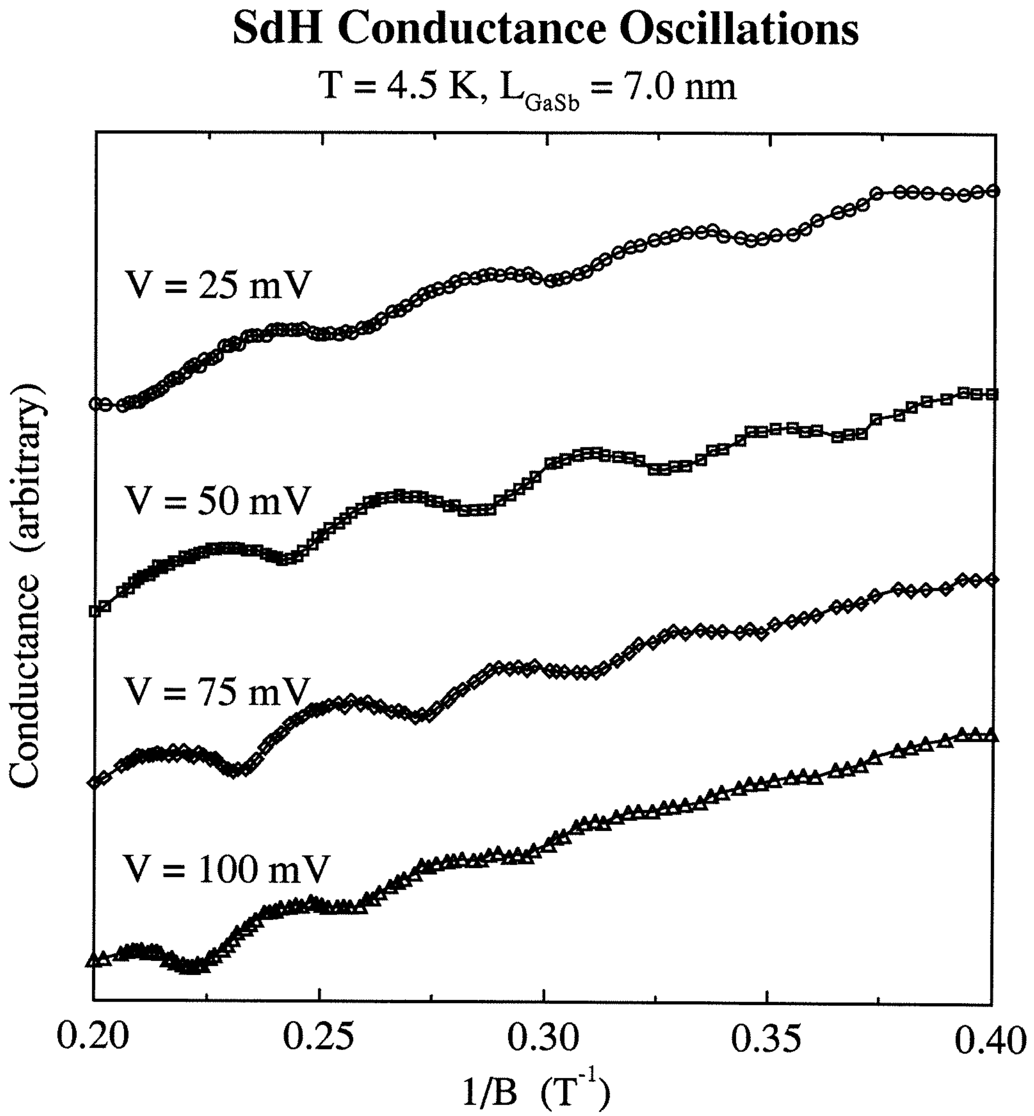


Figure 4.3: Tunneling conductance, dI/dV, of the 7.0 nm well RIT diode as a function of inverse magnetic field, at biases of 25, 50, 75, and 100 mV, respectively. An arbitrary offset has been added to each curve for clarity. The peak-to-peak oscillation amplitude is no greater than roughly 10% of the local average conductance at any given field.

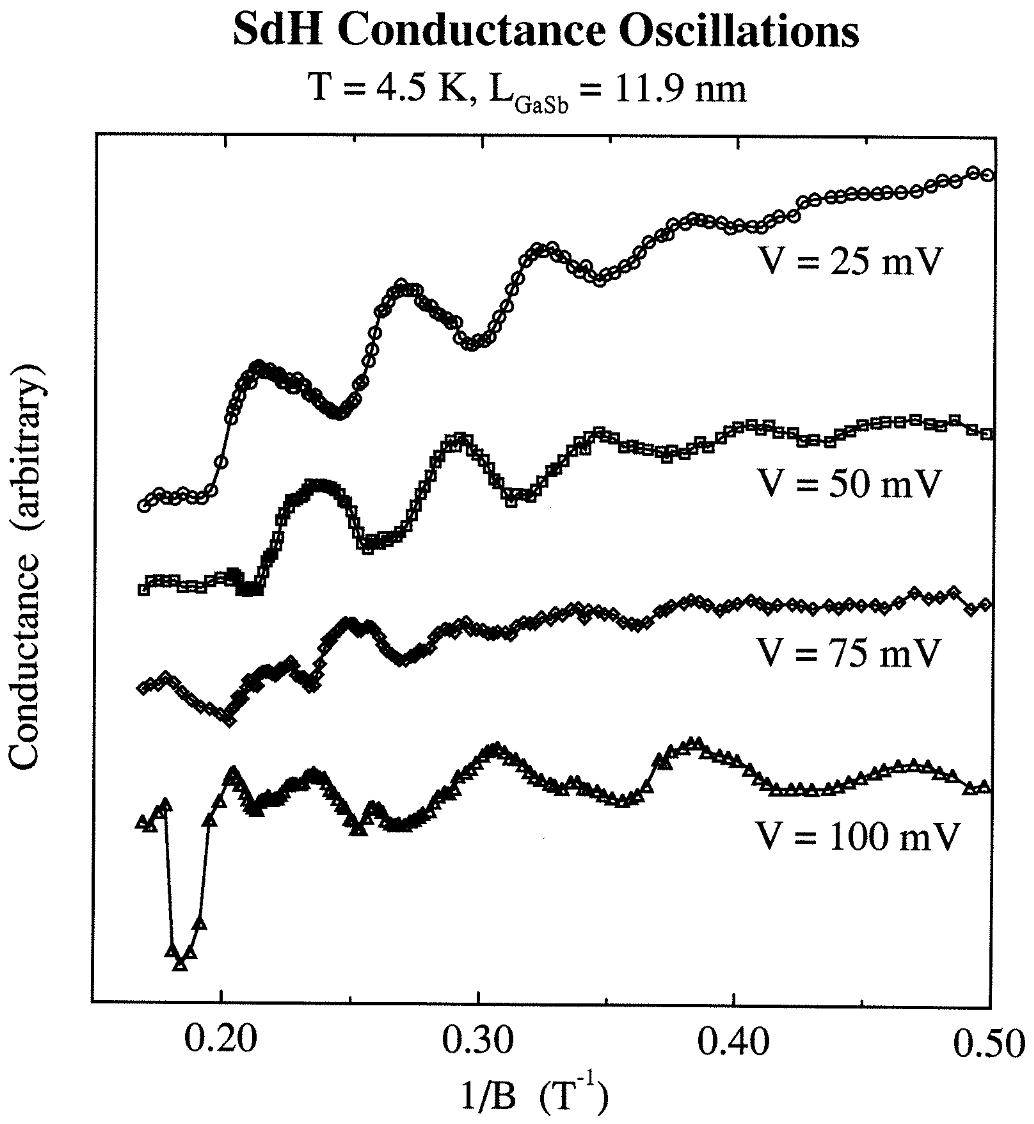


Figure 4.4: Tunneling conductance, dI/dV, of the 11.9 nm well RIT diode as a function of inverse magnetic field, at biases of 25, 50, 75, and 100 mV, respectively. An arbitrary offset has been added to each curve for clarity. The peak-to-peak oscillation amplitude varies considerably but is no greater than roughly 10% of the local average conductance at any given field.

monotonic background variations in conductance which cause the apparent magnitude of the oscillations to appear less significant. In contrast, there is some field-dependent background conductance at 25 mV and 50 mV in the 11.9 nm well sample, but it is of much less consequence. At higher fields (lower $1/B$), the density of states in both the well and the emitter become highly non-linear, as the Landau levels separate and become more distinct. As this occurs, the conditions for resonant, elastic tunneling become more restrictive, leading to a lower net tunneling conductance. However, when proper Landau levels providing resonant energy- and angular momentum-conserving current paths align at a given field, the conductance increases, causing the observed Shubnikov-de Haas-like conductance oscillations. Thus, the magnitude of such oscillations is expected to be greater at larger fields, when the difference between the on- and off-resonance current should be maximized. As seen in both figures, this behavior is observed in the experimental data. At sufficiently low temperatures, the magnitude of these oscillations should increase, as the level broadening decreases. Qualitatively little improvement was observed for either sample at a temperature of 1.6 K, the lowest temperature achievable in our dewar. Oscillatory behavior was observed, to varying degrees, at all biases lower than the peak voltage in each sample. These oscillations were evident for magnetic fields in the range of roughly 2 to 6 tesla. Below that field, presumably the level broadening due to scattering processes [22] is of the same order as the cyclotron energy, giving a lower limit on the energy resolution of the experiments. Above 6 tesla, the inverse field spacing between maxima requires substantially higher fields for the observation of additional oscillations. SdH-like oscillations were not observed in the InAs/AlSb double barrier sample at voltages both above and below the current peak, in all magnetic fields between 0 and 8 tesla.

Of particular note in each sample are the oscillatory characteristics as a function

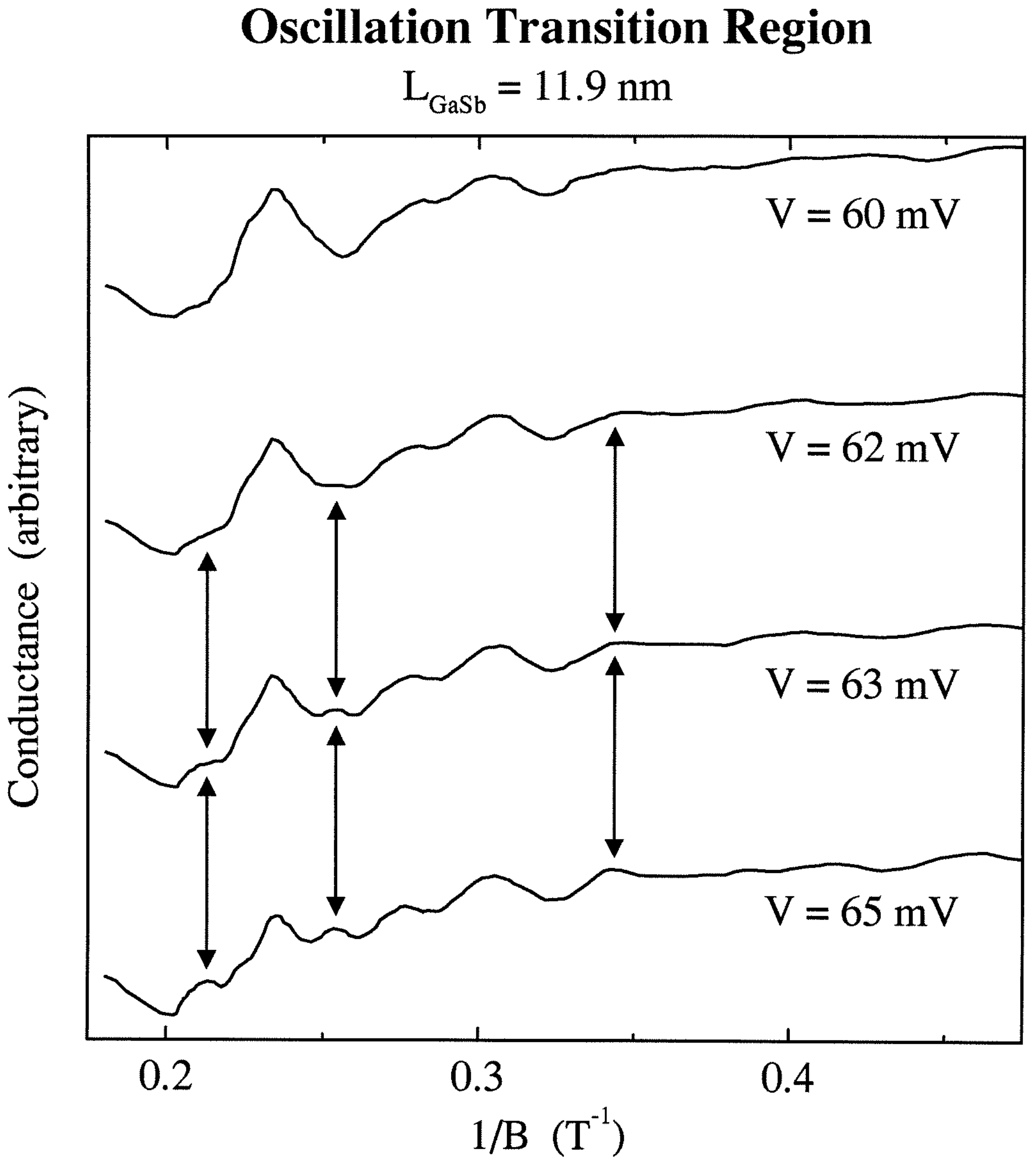


Figure 4.5: Observed conductance values of the 11.9 nm-wide device plotted against inverse magnetic field at biases of 60 mV, 62 mV, 63 mV, and 65 mV, respectively. In this voltage range, there is a noticeable transition in oscillation behavior. Additional peaks, highlighted by arrows, appear and grow at higher biases. Especially sharp are the distinctions between the data taken at 62 mV and 63 mV bias. With the additional 1 mV of bias come many additional peaks, observable at even relatively low magnetic fields.

of applied bias. While Figures 4.3 and 4.4 show the behavior at only four voltages, many of the general trends are contained in this data. In both samples, the phase of the SdH-like oscillations is seen to vary with applied voltage. This trend is not surprising, as changes in voltage alter the value of ΔE in Equation 4.34, causing a voltage-dependent phase shift. Also, the quality of the oscillations deteriorates at higher biases, especially as the applied voltage nears the I-V peak. This dependence is especially noticeable in the wider-well sample. Whereas the oscillation peaks are discernible in the narrow well RIT at voltages up to 110 mV, they persist to even higher voltages in the 11.9 nm well sample, although their form becomes rather complex at biases above 90 mV. The 100 mV data in Figure 4.4 show the non-ideal conductance oscillation behavior in the wide-well device, having two or more distinct frequencies and associated amplitudes. The anomalous oscillations at 75 mV in the 11.9 nm well are part of a transition region in which the conductance oscillations all but die away and then reappear at higher biases with different frequency. The nature of the conductance oscillations in the transition region is shown in Figure 4.5. While the amplitude of the oscillation maxima seen at 60 mV lessons, additional peaks, as shown by the arrows, appear as the bias is increased to 65 mV. In the 1 mV bias increase from 62 mV to 63 mV, many of the peaks rapidly become obvious. For voltages between 75 mV and 90 mV, well-discerned, single-frequency periodicity is observed in the wide-well data. The added complexities at high bias should be expected in the wider-well sample; the quasi-bound states clump together deeper in the well, increasing their interaction and complicating transport at these voltages relative to the more-diffuse subbands of the narrow well.

4.3.4 Fan Diagrams

A compact, efficient method of summarizing the Shubnikov-de Haas oscillatory phenomena is the use of fan diagrams. These diagrams are a common means of displaying the field-dependence of numerous transport features simultaneously. Consider maxima occurring at inverse fields predicted by Equation 4.34. At a fixed bias, the positions of the conductance maxima in $1/B$ should maintain a linear relationship when plotted against a sequential series of integers. In the absence of a change in Landau index ($\Delta l = 0$), when these integers are chosen such that the x-intercept of the fan line lies between 0 and -1 (ideally at $-\frac{1}{2}$), they in turn represent the Landau indices corresponding to each of the maxima. With such a calibration, the physical parameters (ΔE and $\mu^\star$) of the system may be extracted from the fan line slope, and multiple lines may be displayed in a fan-like pattern to show the data for a variety of applied biases. When $\Delta l \neq 0$, a distinction must be made between the emitter and well Landau indices. If, as in Equation 4.34, the emitter and well states are indexed by l and $l + \Delta l$, respectively, the maxima are found at,

$$\frac{1}{B} = \frac{e\hbar}{\Delta E \mu^\star}\left(l + \frac{1}{2} + \Delta l \frac{\mu^\star}{m_W^\star}\right). \tag{4.36}$$

In general, the resulting fan lines are not strictly plotted against either Landau level index. The Δl-dependent phase term can be no larger than Δl itself; in that limit ($m_W^\star \ll m_E^\star$) the fan diagram ordinate is the well Landau index. For $m_W^\star \gg m_E^\star$ ($\mu^\star/m_W^\star \to 0$), the diagram ordinate is the emitter Landau index. Without *a priori* knowledge of the masses, the Landau indices of the conductance maxima cannot be properly calibrated when $\Delta l \neq 0$. However, the phase itself provides an important clue as to the biases at which transport through the well switches from one subband to another; an abrupt change in Landau phase should be observed when such a change in conduction channels occurs. Furthermore, the bias-dependence of the slopes may still be used to estimate ΔE and $\mu^\star$.

Figures 4.6 and 4.7 show the fan diagrams of the narrow and wide well samples, respectively. For each bias, the set of conductance maxima are plotted in addition to a linear least-squares fit of the resulting fan line. A representative set of biases is shown for each sample, although space and labeling difficulties limit the displayed fan lines to coarse bias steps. Fan lines in addition to those shown were considered, the parameters of which will be used in later analyses. Each fan diagram is divided into regions delineated by thick lines and labeled by Roman numerals. The fan lines in each region have similar Landau index phase, γ, or x-axis intercept. There are abrupt changes in phase between these regions, which in turn are related to changes in Landau level index during tunneling. As resonant transport shifts between well subbands, there occurs a corresponding change in Δl, and, ultimately, the observed γ. Thus, if in one bias region the incident carriers tunnel into well states in the LH1 band, and consequently couple primarily to the $|\frac{3}{2} \pm \frac{1}{2}\rangle$ hole components, they should have $\Delta l = 0$ or $+1$, depending on the incident spin state. This change in Landau index introduces a phase shift of $(e\hbar/\Delta E\mu^\star)\Delta l$ into the inverse fields predicted by Equation 4.34. When, at slightly higher biases, the incident electrons instead couple to a well state having a different mass or Δl (or both), the subsequent change in Landau phase is observed in the fan diagram. Strictly speaking, the differing phases for opposite incoming spin states should cause two sets of oscillations, of the same frequency, but of different phases. Such behavior is not seen at most biases, and it therefore is likely that the well mass, $m^\star_W$, and the reduced mass, $\mu^\star$, are nearly identical, causing the phase difference between spin states to be an integral multiple of the overall periodicity.

In general, all of the fan lines, for both samples, are linear and of a single period, with greater variance at higher bias. This trend is consistent with the effective mass approximation, and Landau levels equally spaced in energy. Deviations from

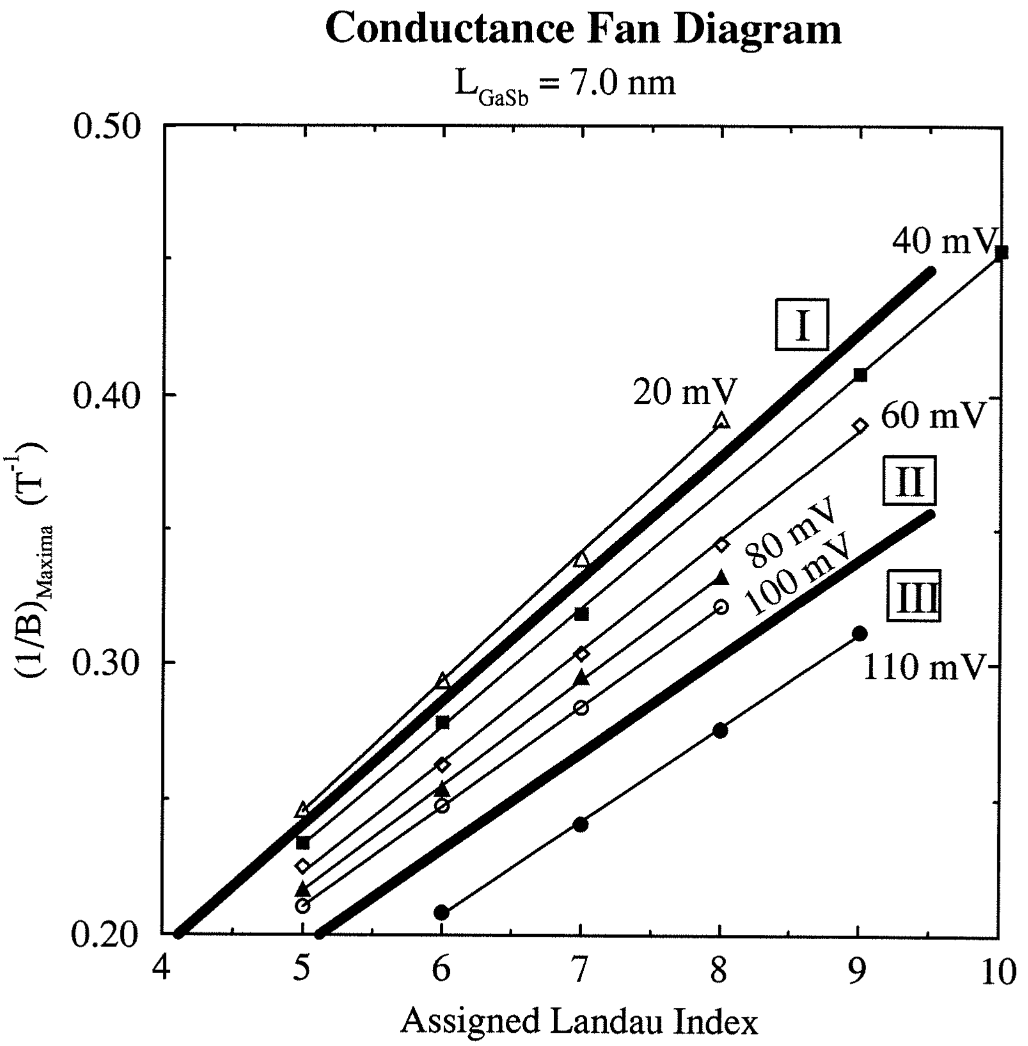


Figure 4.6: Fan diagram showing the inverse fields at which maxima occur in the 7.0 nm well RIT conductance oscillations, plotted versus assigned Landau level index. The ordinate value is assigned such that the fan line of a given bias intersects the x-axis between 0 and -1, calibrating the individual maxima to well Landau indices in the limit of $\omega_W^\star = \omega_E^\star$. Thick lines separate regions of differing Landau-index phase.

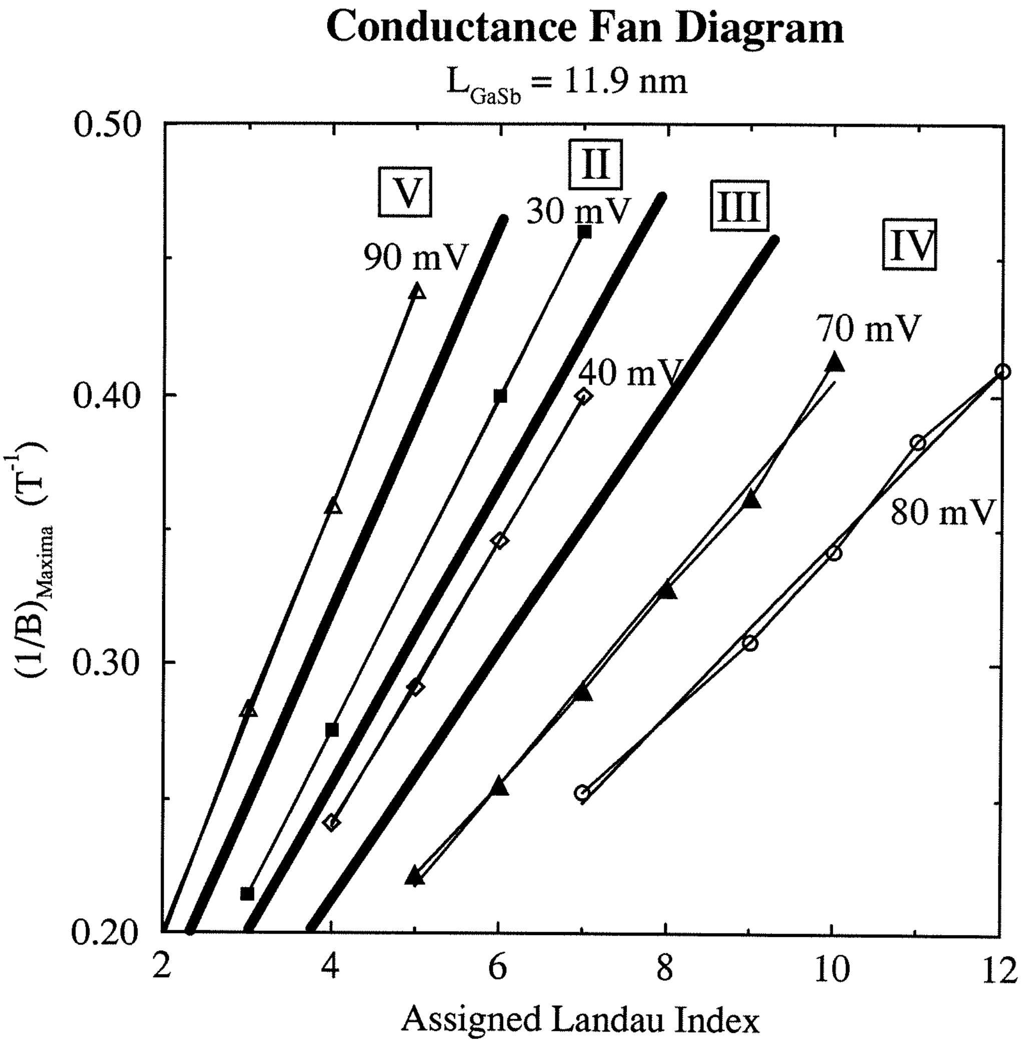


Figure 4.7: Fan diagram showing the inverse fields at which maxima occur in the 11.9 nm well RIT conductance oscillations, plotted versus assigned Landau level index. The ordinate value is assigned such that the fan line of a given bias intersects the x-axis between 0 and -1, calibrating the individual maxima to well Landau indices in the limit of $\omega_W^\star = \omega_E^\star$. Thick lines separate regions of differing Landau-index phase.

ideal Landau level behavior are expected at higher biases and in the wider well sample, where interactions between bands deep in the GaSb well result in strongly nonparabolic dispersions. Such behavior is somewhat evident in the 70 mV and 80 mV fan lines in region IV of the wide well diagram. The high bias cutoff of fan data in both samples is attributable to non-ideal, multi-periodic conductance oscillations and decreasing oscillatory amplitude. The narrow well fan lines in Figure 4.6 are well-ordered and evenly spaced as a function of bias, although there is a large shift in fan position between biases of 100 mV and 110 mV. In contrast, the data for the wider, 11.9 nm well are less ideal. There is a very abrupt shift in slope between the 80 mV and 90 mV fan lines, causing the latter to more nearly approximate the 30 mV data. In addition, although they could not all be shown, the biases intermediate to 40 mV and 70 mV in region III have slightly decreasing slopes as a function of bias, while the slopes from the 15 mV to 30 mV data in region II actually increase. The added complexity of transport in the wide well sample is also evidenced by the larger number fan diagram regions and observed shifts in Landau phase.

Figures 4.8 and 4.9 plot the fan line slope and Landau phase γ at 5 mV bias increments for the narrow and wide well fan diagrams, respectively. The dashed vertical lines, and corresponding Roman numerals, denote the various regions of similar Landau phase. The narrow well data shows very distinct and abrupt changes in both slope and phase between each of these regions. The bias dependence of these slopes is useful for the determination of the well state mass and energy. As indicated in Equation 4.34, the slope M of the fan lines should be related to these parameters as $M = e\hbar/\mu^{\star}\Delta E$. The voltage across the device enters this equality through the energy difference, $\Delta E = \Delta E_0 - \alpha\, eV$, where $\Delta E_0 \equiv E_W - E_E$ is the zero-bias energy difference of the $B = 0$, $k = 0$ band edges of the well and emitter states, and α is the fraction of the total voltage dropped between the the emitter

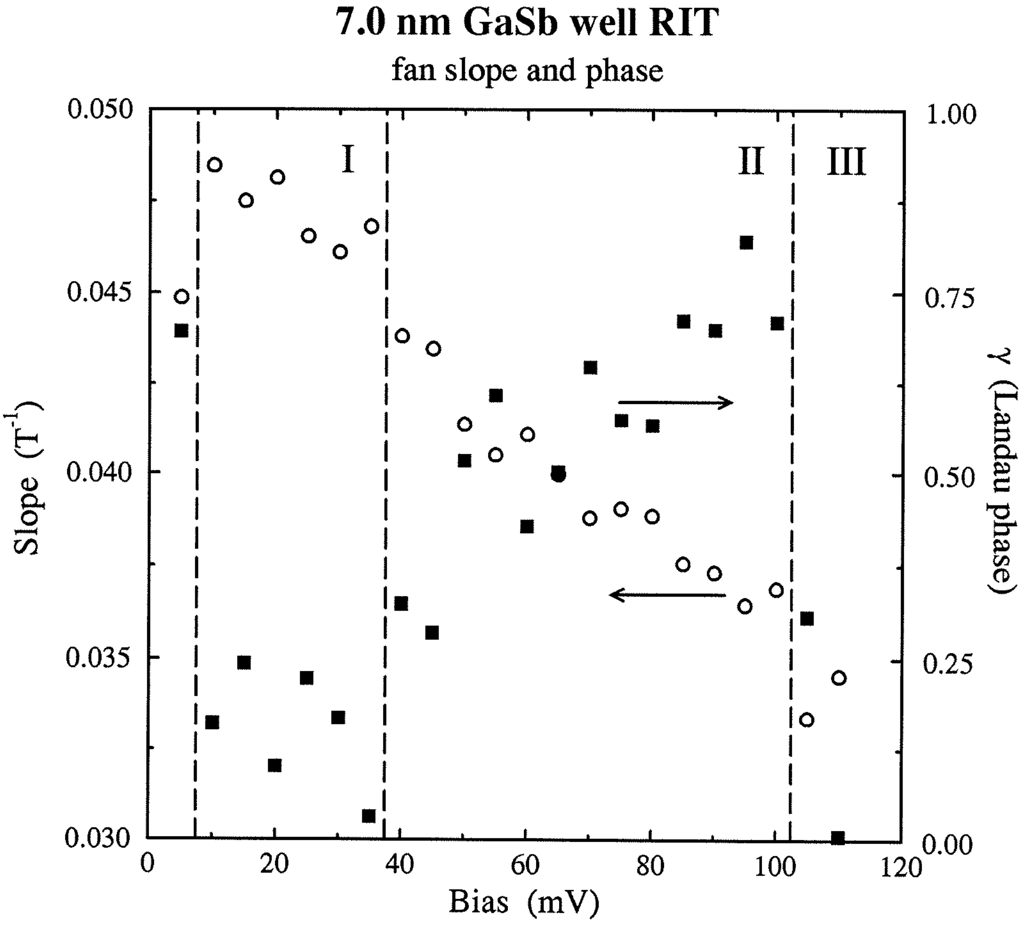


Figure 4.8: The slope and Landau index phase (γ) for fan lines of the 7.0 nm well oscillation data plotted against applied bias at 5 mV increments. The vertical dashed lines separate regions between which large, discontinuous changes in Landau phase occur.

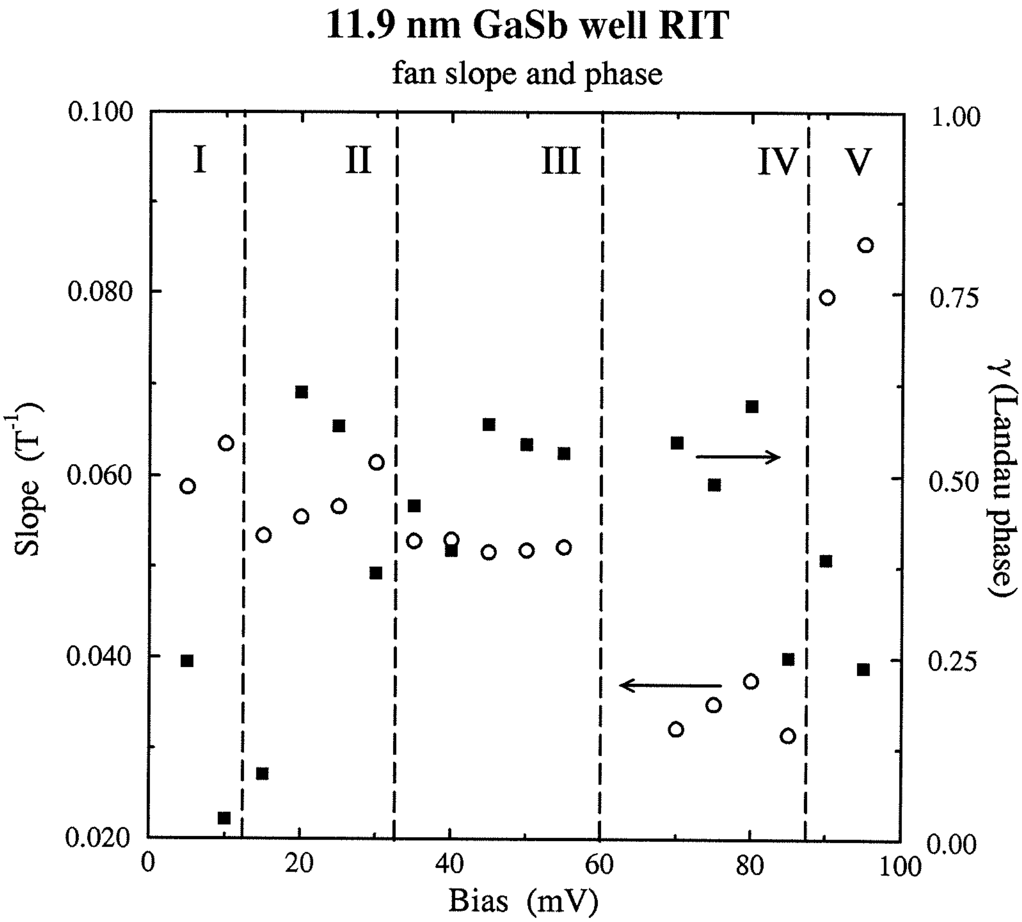


Figure 4.9: The slope and Landau index phase (γ) for fan lines of the 11.9 nm well oscillation data plotted against applied bias at 5 mV increments. The vertical dashed lines separate regions between which large, discontinuous changes in fan slope occur. Unlike the 7.0 nm well data, these discontinuities in slope are not matched by simultaneous shifts in Landau index phase. The phase shifts appear offset by up to 5 mV from the discontinuities in fan slope, perhaps because the transitions between subbands are less abrupt in this sample.

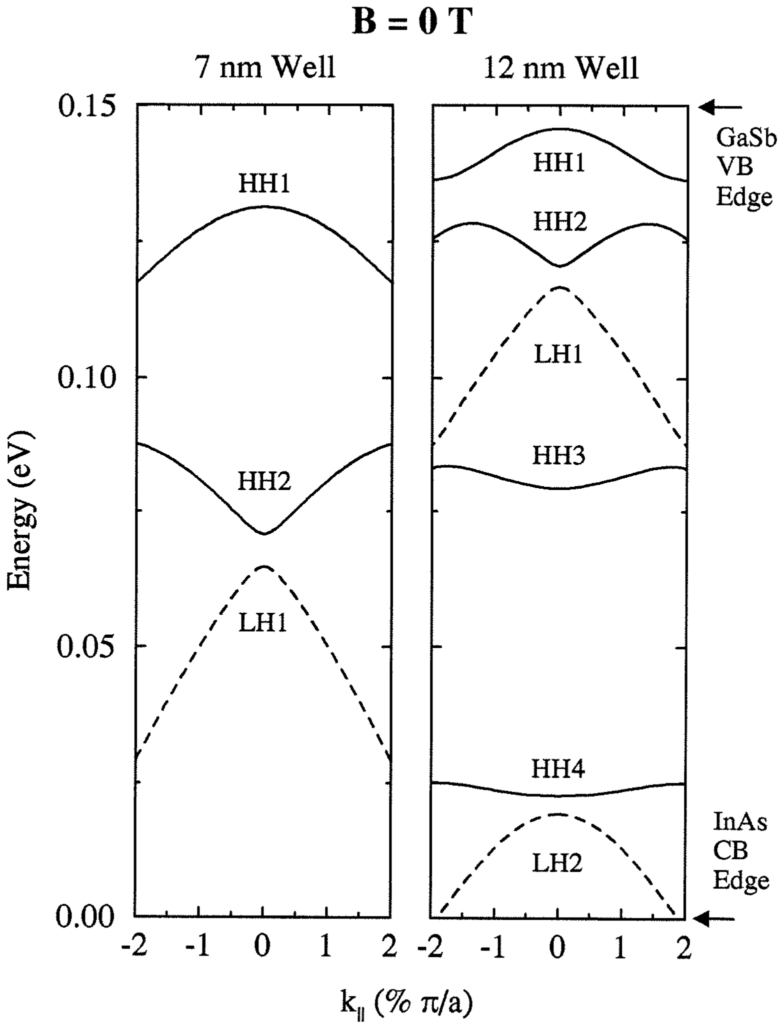


Figure 4.10: Theoretical [14] zero-field well subband structure for RIT devices having GaSb well widths of 7.0 and 11.9 nm, in the energy range between the InAs conduction band edge and the GaSb valence band edge.

state and the well. The parameter α is expected to be less than 0.5 as the potential drop from the outer emitter to the center of the well should be one half the total voltage in a symmetric device. As the bias is increased, the emitter notch state falls lower into the accumulation region, and while α is not strictly constant, it is estimated from band bending calculations to be be roughly 0.35 for the voltage range in question. This accuracy of this estimate, and the lack of a voltage dependence to α, is no worse than the other assumptions and approximations that will be made in this analysis. The overall voltage dependence of the fan line slope is therefore inversely proportional to the applied bias. Rewriting the inverse slope, $1/M$, as a function of the applied voltage yields an expected linear dependence,

$$\frac{1}{M} = \frac{\mu^\star}{e\hbar}\left[\Delta E_0 - 0.35\, eV\right], \tag{4.37}$$

having a slope,

$$\beta \equiv \partial(1/M)/\partial V = -0.35\, \mu^\star/\hbar \tag{4.38}$$

and an offset,

$$(1/M)_0 = \mu^\star \Delta E_0/e\hbar. \tag{4.39}$$

Fitting the slopes in region I of the narrow well diagram gives estimates of $\beta = 36 \pm 13$ T/V and $(1/M)_0 = 20.0 \pm 0.3$ T. Likewise, in region II, fitting gives $\beta = 74 \pm 5$ T/V and $(1/M)_0 = 20.0 \pm 0.4$ T. In both regions, the voltage-dependence of the inverse slopes is positive, whereas Equation 4.37 predicts they should be negative. This fundamental difference can only be explained by the existence of positive dispersions for these valence band states. In such a case, $\mu^\star$ is no longer the reduced mass of the two bands, but rather the related quantity $1/m_E^\star - 1/m_W^\star$. When $m_W^\star < m_E^\star$, this quantity is negative, and leads to the observed behavior of the fan line slopes. Because both masses are implicitly involved in determining the voltage dependence of the fan line slopes, the well mass cannot be analytically determined from the values obtained in fitting the data to Equation 4.37. However, if the emitter state mass is assumed to be identical to the InAs bulk mass, then details of the well subbands may be extracted from these fits. In region I with these assumptions, the data imply $m_W^* = 0.008\,m_e$ and $\Delta E_0 = -200$ meV. Similarly, in region II, these assumptions yield $m_W^* = 0.012\,m_e$ and $\Delta E = -95$ meV. Neither of these sets of values is reasonable, however. The bulk energy difference between the InAs conduction band and GaSb valence band edges is only 150 meV, and band bending and quantization in the heterosystem reduced that further. Thus, the large energy differences, while still possible, are less plausible. Additionally, since the well mass differs substantially from the reduced mass in both of these cases, two sets of conductance oscillations of identical frequency, but offset by a phase factor, would be expected from the two possible spin states, but were not seen. Thus, it appears more likely that the in-plane mass of the emitter band is rather large, relative to the bulk value, such that the well mass is roughly equal to the magnitude of the reduced mass. This assumption would give masses of $0.012\,m_e$ and $0.024\,m_e$, as well as energy offsets of -135 meV and -45 meV, respectively, for the well states in regions I and II. For comparison, the zero-field band structure predicted by Liu *et al* for these devices is shown in Figure 4.10. While theory predicts a positive dispersion for the HH2 subband, the exact relationship between the theory and experiment is uncertain. First, only one band is predicted to have positive dispersion, and yet a single band cannot account for the behavior in both regions I and II of the fan diagram, especially since these together span nearly 100 mV of bias. It is possible that the HH1 band has a positive dispersion for small k as well, yet the cause of this deviation from the theory is unclear.

The wide well fan data display even greater complexity. The Landau phase changes discontinuously at four different biases, although, unlike the narrow well device, the discontinuities in slope do not, in general, occur at the same biases. In region IV, the data for the bias range of 60 – 70 mV is omitted due to the extended change in behavior shown in Figure 4.5. In all but region III, the slope increases as a function of bias, as predicted by Equation 4.37 for negative dispersion bands. In region III, β is statistically indistinguishable from 0, indicating that $\mu^\star$ is essentially infinite. Transport in this voltage regime is thus likely due to a positive dispersion state of mass nearly identical to the emitter mass, such that $1/m_E^\star - 1/m_W^\star \approx 0$. Least-squares fits to the inverse slope in region II give $\beta = -154 \pm 29$ T, and $(1/M)_0 = 21 \pm 0.7$ T. If we again assume $m_W^\star \ll m_E^\star$, as we did for the narrow well data, these fit parameters correspond to a mass of roughly $0.051\, m_e$ and energy difference ΔE_0 of 47 meV. In region IV, the data does not appear to behave linearly, and the effects of the transition observed between 60 mV and 70 mV may introduce nonidealities at higher biases. Referring to the theoretical band structure in Figure 4.10, transport in region II is most likely through the LH1 well state, while region III corresponds to the HH2 band. Since the HH2 band has both positive and negative dispersion at small $k_{||}$ (low Landau level index), the transition occurring at 60 mV, and the subsequent nonideal (but negative dispersion-like) fan behavior at higher biases may be experimental evidence of the change in sign of the local differential mass. Or, more likely, as the HH1 and HH2 bands are only 10 meV apart, transport in the 60 mV to 70 mV region may involve mixed transport through different Landau levels of each hole subband. As the wider well data indicates, the closer the well subbands approach each other, the more complex and varied transport through the device becomes.

4.4 Conclusions

Shubnikov-de Haas oscillations of the tunneling conductance through RIT diodes provide an indirect means of exploring the subband structure of the GaSb well and observing 0D-0D quantum transport behavior in a unique hybrid system.

The observation of distinct and abrupt changes in the Landau phase of fan lines as a function of applied bias provide evidence that the Landau level index l is not conserved in quantum transport involving states of differing z components of angular momentum. The complete selection rule for 0D-0D Landau tunneling, $\Delta(l + m_j) = 0$, reflects the conservation of total longitudinal angular momentum, including the macroscopic contribution of cyclotron orbits induced by an external magnetic field. Through the study of the bias dependence of fan line slopes for regions of similar Landau phase, the mass and band edge of states in the well may be determined, though only in the limit of $m_W^\star \ll m_E^\star$. In general, only the reduced mass of the emitter and well states can be determined experimentally. The former limit, however, may be justified by the lack of separate, out-of-phase oscillation sequences for the spin-up and spin-down components of each subband. In both samples, the anomalous bias-dependence of the fan line slopes is indicative of a positive, low-mass well dispersion for the range of Landau indices experimentally accessible. The additional complexity of the structure in the wide well was observed in variety of nonideal behaviors, including a 5 mV-wide region where oscillations having several phases coexist.

This study, as one of the first complete investigations of magnetotransport in Type II broken-gap resonant tunneling diodes, demonstrates the novel nature of transport in the InAs/AlSb/GaSb system. The evidence for the non-conservation of Landau index in Shubnikov-de Haas oscillations of resonant, elastic tunneling conductance is unique, and has implications for transport in other quantum transport systems for which states of differing longitudinal angular momentum coexist (e.g., hole tunneling in Type I heterostructures). Only very recently have experimental observations of Landau level mixing in unipolar hole tunneling been reported [23]. The data contained in this chapter represent the first evidence of Landau level mixing in interband tunneling, and the first observation of the conservation of total projected angular momentum through phase shifts of Shubnikov-de Haas oscillations. Additional ultra-low temperature ($< 4.2K$) studies of this and other systems appear warranted in light of the data presented in this chapter. Ultimately, if evidence of Landau level tunneling remains limited to the conductance, additional measurements with a tunneling spectrometer, in which the conductance is measured directly and with greater accuracy, may be appropriate.

Bibliography

[1] K. von Klitzing, Rev. Mod. Phys. **58**, 519 (1986).

[2] C. E. T. Gonçalves de Silva and E. E. Mendez, Phys. Rev. B **38**, 3994 (1988).

[3] M. L. Leadbeater, F. W. Sheard, and L. Eaves, Semicond. Sci. Technol. **6**, 1021 (1991).

[4] C. H. Yang, M. J. Yang, and Y. C. Kao, Phys. Rev. B **40**, 6272 (1989).

[5] E. E. Mendez, Surf. Sci. **267**, 370 (1992).

[6] E. E. Mendez, L. Esaki, and W. I. Wang, Phys. Rev. B **33**, 2893 (1986).

[7] S. Ben Amor *et al.*, Appl. Phys. Lett. **54**, 1908 (1989).

[8] V. J. Goldman, D. C. Tsui, and J. E. Cunningham, Phys. Rev. B **35**, 9387 (1987).

[9] M. L. Leadbeater *et al.*, Solid-State Elec. **31**, 707 (1988).

[10] L. Eaves *et al.*, Appl. Phys. Lett. **52**, 212 (1988).

[11] E. E. Mendez, H. Ohno, L. Esaki, and W. I. Wang, Phys. Rev. B **43**, 5196 (1992).

[12] C. H. Yang, M. J. Yang, and Y. C. Kao, Phys. Rev. B **40**, 6272 (1989).

[13] Takamasu *et al.*, Surf. Sci. **263**, 217 (1992).

[14] Y. X. Liu, D. Z.-Y. Ting, and T. C. McGill, unpublished.

[15] E. O. Kane, in *Semiconductors and Semimetals* **1**, edited by R. K. Willardson and A. C. Beer (Academic Press, New York, 1966).

[16] A. Fasolino and M. Altarelli, Surf. Sci. **142**, 322 (1984).

[17] D. Smith, private communication.

[18] J. M. Luttinger, Phys. Rev **102**, 1030 (1956).

[19] A. Savitzky and M. J. E. Golay, Anal. Chem. **36**, 1627 (1964); See also W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C*, 2^{cd} edition, (Cambridge University Press, Cambridge, 1992), pp. 650–655.

[20] V. J. Goldman, D. C. Tsui, and J. E. Cunningham, Phys. Rev. Lett. **58**, 1256 (1987).

[21] T. C. L. G. Sollner, Phys. Rev. Lett. **59**, 1622 (1987).

[22] T. Ando, A. B. Fowler, and F. Stern, Rev. Mod. Phys. **54**, 437 (1982).

[23] A. Zaslavsky *et al.*, Surf. Sci. **305**, 307 (1994).

Appendix A

Magnet History and Operation

A.1 Introduction

Well, if you are reading this appendix, you are probably planning on using the 8 tesla superconducting magnet. Since you are, and I was the individual responsible for the original purchase, I feel obligated to provide you with a certain amount of historical background and a few heuristic operating guidelines. For details of the magnet design, and the "official" factory instructions, you should refer to the documentation provided by Janis [1, 2, 3]. After reading this appendix, I hope you understand the strengths and weaknesses of the magnet design, the potential problems you may encounter, and the proper operation of the magnet and associated cryogenic dewar. Let us both hope that you need not repeat the same mistakes, and encounter the same design flaws, that I have.

The magnet itself is a split-coil Nb/Ti alloy superconducting solenoid (from American Magnetics), and is mounted in a D2-style Janis dewar with optical access both perpendicular and parallel to the field axis. The details of the design are summarized in Reference [1]. The magnet specifications, taken from this document as well as determined empirically in routine operation, are summarized below:

Rated Central Field at 4.5 K	8.00 T
Rated Current (at 4.5 K)	73.1 Amps
Field-to-Current Ratio	0.1094 T/A
Homogeneity	0.1% (1 cm DSV)
Inductance	35.6 Henries
Max recommended charging voltage, $B < 5$ T	2.5 V
Max recommended charging voltage, $B > 5$ T	1.5 V
Total magnet resistance, 300 K	21.7 Ω
Total magnet resistance, 77 K	16.6 Ω
Total LHe chamber volume (w/out magnet)	$\approx$ 26 liters
Estimated 100% LHe volume	$\approx$ 20 liters
Inner-most windows	ZnS
Outer windows	ZnS or Quartz

Table A.1: Magnet design specifications

A.2 History

The magnet apparently has a long history in the McGill group dating back long before it was ever purchased. It was a phantom piece of equipment that appeared in many proposals, but which never found a student sponsor. Its true justification came when the work with the interband InAs/AlSb/GaSb devices progressed to a stage that magnetotransport measurements were seen as important. I believe that you will be amused at exactly how many former students return for a visit to be amazed that a magnet was actually bought and used. Unfortunately, there were many pitfalls associated with the purchase of the magnet itself, the details of which you should be aware in case the problems arise yet again. In general, I think it fair to state that we (namely, Tom and I) are not impressed with either Janis or American Magnetics, the two vendors which assembled the system.

While there were many of the problems which might be considered typical for a large equipment purchase, such as shipping delays, inappropriately-wired electronics, and poor quality control at the factory, the biggest single problem encountered with the dewar was with the inner window seals around the sample tube. The dewar was specified to have ZnS windows so that the system may be used for magneto-optical experiments on far-IR detectors. Janis however, does not, apparently, have a good track record with the installation of specialized windows. Problems arose with a previous purchase of a dewar with ZnSe windows as well. Basically, two flaws were found with the windows when the dewar arrived at Caltech. First, they were absolutely filthy; so much so, in fact, that they were nearly opaque in one direction. While the outer windows may be removed and cleaned, or replaced, in a straight-forward manner, the inner-most set is not designed for easy access, and is hermetically sealed to maintain vacuum isolation of the sample tube (in principle, at least, see more later). Because ZnS is very soft, we were loath to clean them ourselves, lest we scratch the surfaces. When the second problem with the dewar was discovered, a leak in the sample tube, the whole magnet was sent back to the factory for repair, and while there, the inner windows were cleaned, and scratched, by Janis personnel. The leak in the sample tube was rather extreme; when pumped down to a few hundred microns, it would return to atmosphere in about 15 or 20 minutes. The problem, it turned out, was the inner window seals themselves, and the complications introduced by ZnS. When the magnet returned, it was cooled down and first-field was achieved on June 15, 1992. Unfortunately, by August of that year, there was again a leak in the sample tube, and the whole dewar was sent back to Janis yet again for repair of the inner window seals. The entire system was not fully operational until January of 1993, fully two years after it was first ordered. The inner window seals appear fine to date, but they may cause problems in the future, and should be the first suspect should a leak in the sample tube develop.

In addition to the windows, there are several other design flaws that I have encountered of which you should be made aware. Some of these were simply due to naiveté on my part when the system was specified, but others are true engineering problems. First, a thermometer of some sort should have been directly mounted onto the top of the magnet coil, so that the temperature of the magnet could be directly monitored during the cool-down process. The danger is that if the magnet is not completely pre-cooled to 77 K, the subsequent cooling step to 4.2 K will require vast quantities of LHe. I have determined empirically that the resistance of the coil when at 77 K is 16.6 Ω, and this fact may be used in lieu of a thermometer during the magnet pre-cool. The meter used to display the magnet current is only accurate to a tenth of an Ampere, which represents an uncertainty in the field greater than the designed magnet homogeneity. This problem is especially bad at low fields, and is compounded by the fact that the meter also seems to have an offset associated with it. As a work-around, I used a microvolt meter to directly monitor the current shunt voltage, and get a more accurate reading of the field strength. This procedure worked well, and is recommended in all further operation. The computer control aspect of the system is admittedly crude, but also seems plagued with offsets and drift. Consequently, the field could not be reliably set or read remotely, and all attempts at full computer control were abandoned as the delays associated with the inner windows developed. Lastly, the current reversal switch never seems to work (although you will probably never need to use it) because of the offsets in the built-in current ammeter. The reversing switch refuses to operate, and chimes an annoying warning, if it thinks there is current still present in the coils. This feature is good, but because of the offset in the meter (which never was fully compensated), the reversal switch always "sees" current in the magnet, even when the power supply is off. Lastly, the evacuation valve for the vacuum jacket is positioned at such an angle that the LN_2 access vent nearest it interferes with connection to the turbo pump. As a result, I rigged together an extension tube for the pump. It should be stored in the toolbox near the magnet. If it cannot be found, another must be acquired before the dewar can be used. None of these are major problems, and so long as the inner windows hold, I think you will find the system (as I left it, at least), to be quite functional and straightforward to learn to use.

A.3 Operation

The exact cool-down procedure recommended by Janis is found in Reference [1]. Incidentally, as these documents are often referred to during magnet operation, a set is kept in the lab with the magnet. To satisfy proper instruction manual archiving, however, an additional copy should be located in the manual storage file cabinets.

The **most** important concept to keep in mind when operating the magnet is *safety.* There are two main safety concerns: electrical shock, and cryogenic liquid handling. While the magnet operates at high currents (upwards of 75 Amps), the voltages involved are low (< 5 V), and therefore there is usually not much of a hazard. However, if the current path is disrupted for whatever reason, the massive inductance of the coil will cause immense voltages across (most likely) the input terminals on top of the dewar. Thus, you should always be careful around the magnet when it is charged, even if the power supply is off (I find that the menacing sound of the power supply is more than enough to remind me of the dangers involved, but in persistent current mode, where the power supply is off, the calm and quiet belie the possible dangers). The danger from cryogenic liquids is real as well. When filling the dewar with LHe, both the magnet dewar and the LHe dewar may have cryogenic liquids under high pressure, and consequently, there is a very real danger of frostbite if the flow out of either becomes uncontrolled. In addition, there is an asphyxiation hazard, although with the proper ventilation, this should not be a concern. Should the magnet ever quench (go normal during operation), all 20+ liters of LHe will be dumped out the release valve on the side of the dewar in a matter of a minute, and it is best to leave the room. The power supply controller should drop the current to zero when a quench is detected, but there is a danger that very large voltages will cause the current in the magnet to discharge to a convenient nearby potential. Try to not make yourself the lightning rod. Also, the risk of asphyxiation and frostbite is greatly increased in the presence of that much LHe angrily venting from the side of the dewar. On the one occasion when a quench occurred as I was using the system the room was subsequently cooled to the low 60° F-range. The entire experience was quite impressive. It should be noted that there are occasions in which you will find it necessary to clamp off the release valve for the LHe chamber. Often this is done to increase the He gas overpressure in the chamber if there are flow problems through the capillary to the sample tube. If you do clamp off the release valve, which, although I have done so, I should not recommend, be always on the ready to yank the clamp off in the case of a quench. When the quench did occur as I was using the magnet, the valve was clamped off, and if I had not released it, the dewar may very well have been a twisted shard of metal at this moment. In general, it is best not to fiddle with the safety features of the equipment. Lastly, beware of the effect the field has on external items. Always keep loose tools away from the magnet while in operation. The opaque plexiglass window guards were installed to keep metal objects accelerated by the field from breaking the exterior windows, in addition to keeping the sample free of external illumination.

The cool-down procedure, detailed in Reference [1], is summarized below.

1. Evacuate vacuum jacket (using turbo pump).

2. Evacuate helium reservoir (using mech pump attached to Tri-clamp connector off of the release valve. Open needle valve and evacuate sample tube as well).

3. Back-fill helium reservoir and sample tube with He gas and seal. Close needle valve to isolate the two, and maintain He pressure to both.

4. Fill nitrogen reservoir with LN_2.

5. Fill helium reservoir with LN_2. Use same transfer line for both LN_2 fills. The copper tube at the end is designed to fit in the seal for the initial LHe fill port.

6. As the LN_2 is added to the LHe reservoir, there will at first be a rapid and violent boil-off, and subsequent high pressure in the LHe chamber. It is

easiest to relieve this by keeping the release valve open with a spare gasket or such jammed between the two faces of the spring-loaded valve. Also, at this stage, continue He gas overpressure to the sample tube, but discontinue such to the LHe chamber once the LN_2 creates a suitable overpressure.

7. Monitor the coil resistance across the input terminals. The power supply should be off, but the controller, and current reversal switch need to be on due to the active circuitry in the energy absorber module. Once the magnet reaches 77 K, fill for a little while longer, and then turn off LN_2 flow.

8. At this point the dewar can sit overnight, with positive He gas pressure to both the sample tube and the LHe reservoir. Usually, I performed all steps to this point the night prior to the actual use of the dewar.

9. Siphon off the LN_2 from the LHe reservoir (usually the next morning) by inserting the LHe transfer tube down the initial fill port, and applying an overpressure of He gas. Continue roughly 5 minutes beyond the time the last LN_2 flows from the tube.

10. Pump down LHe reservoir. It should reach pressures down in the hundreds of microns. If it does not, and only goes as low as 1 torr (or greater), there is still LN_2 in the chamber, and He gas should be flowed in, and then pumped on, sequentially, until the pressure lowers well below 1 torr.

11. Backfill with He gas, and then begin LHe transfer. Start cooling down the transfer tube first, while preparing to open the seal to the initial fill port. When LHe is coming out of the transfer line, insert the tube into the fill port. Jam open the release valve to minimize the overpressure in the LHe chamber, and transfer the LHe until the dewar is full. Note that you may have to pressurize the LHe dewar with He gas to continue the transfer beyond

the first few minutes. The LHe chamber must be a *minimum* of 65% full for the magnet to be used without fear of quenching.

12. When the LHe chamber is full, apply He gas overpressure, and open needle valve to cool sample to desired temperature. Open valve to manifold of the vapor-cooled leads (if not opened earlier). Allow leads to cool before current is applied. These must be kept cold by a flow of He, or they will vaporize at high currents!

13. Let entire system sit for 15-30 minutes before charging the magnet. (Be sure to let the power supply warm up at this time.)

14. You are ready to magnetize!!!! Be sure to keep the LN_2 reservoir at no less than 80%, or else the LHe evaporation rates will be excessive.

Aside from safety, the most important thing to do during cool-down is *not* allow atmospheric gases into any portion of the dewar that is cold. Usually it is sufficient to keep a 1-3 PSI He gas overpressure to both the LHe reservoir and the sample tube. If you do not, N_2 ice can form in the LHe chamber, possibly causing a rupture. And the capillary tube between the LHe reservoir and the sample tube may freeze. If this happens, the whole dewar must be warmed up to clear the ice, and allow proper sample temperature control.

There are very few maintenance issues with the system, so long as it is used properly. If you have to change windows, you will need the big yellow monster crane that the group owns. It takes about half a dozen strongish people to setup, and needs to fit on either side of the optical table where the dewar is mounted. There are two chain hoists associated with this crane. The dewar (and stand) must first be lifted off the table and onto the floor before the cryostat can be lifted from the vacuum jacket. Attach one hoist to either side of the dewar with the eyelet fasteners on the top. Disconnect the seating screws on the underside of the main O-ring flange, and lift the inner guts of the dewar out by evenly lifting both sides with the chain hoists. There is only a few microns of tolerance between the upper-most reaches of the crane, and the height to which you have to lift the cryostat. Try and minimize the height of your connections between the eyelets and the hoists (we had to bypass the hooks and attach the eyelets to the chain directly). Once the dewar is out, the second layer of windows is easily accessible. It can be harrowing to see the cryostat floating so precariously above the ground, but I have personally gone through this procedure at least 20 times (due to the window problems) without a major incident. The only other maintenance that I have had to perform was to sand down the copper contacts for the power supply connection to the magnet. These get wet and have large amounts of current going through them and therefore corrode very quickly. Sanding both sides of the contacts down can dramatically improve the magnet operation. The series resistance due to the corrosion can adversely change the charging characteristics of the magnet. I had to do this when I began to notice large offset voltages were necessary to charge the magnet.

Well, good luck, and enjoy your new cryogenic toy!

Bibliography

[1] "Operating Instructions for Superconducting Magnet Systems," Janis Research Company. (This document contains specific design parameters of our dewar in addition to generic operational guidelines for the magnet cryostat and the *Supervaritemp* variable temperature insert. It is a must-read before attempting to cool down the dewar.)

[2] "Operating Instructions for Supplied Ancillary Equipment," Janis Research Company. (This is basically a collection of manuals for various pieces of equipment bought with the dewar and magnet. Many are from American Magnetics, which made the superconducting coil.)

[3] M. N. Jirmanus, "Introduction to Laboratory Cryogenics," Janis Research Company. (A good overview of cryogenic laboratory technique.)

Appendix B

Bulk $\vec{k} \cdot \vec{p}$ Theory Incorporating Magnetic Fields

B.1 Introduction

As was motivated in Chapter 2, the effect of a magnetic field upon the physics of carriers in bulk semiconductors fundamentally occurs in k-space, where the vector potential, $\vec{A}$, describing the field, modifies the Hamiltonian. Aside from spin-dependent potential energy terms, the field enters the general single-particle Hamiltonian through the kinetic energy term. The momentum operator $\vec{p}$ becomes $\vec{p} - q\vec{A}$ in the presence of a magnetic field, and the vector $q\vec{A}/\hbar$ may be viewed as a B- and position-dependent vector in k-space. Because the field enters the Hamiltonian naturally in k-space, solutions of the Hamiltonian are best attempted using $\vec{k} \cdot \vec{p}$, rather than real-space methods such as tight binding. Because this thesis makes reference to the $\vec{k} \cdot \vec{p}$ formalism, and the specifics of 2- and 8-band models derived from it, the goals of this chapter are to motivate this theory in bulk, show how the zero-field Hamiltonian is changed in the presence of a magnetic field, and lend a degree of intuition to the predicted results.

B.2 Bulk $\vec{k} \cdot \vec{p}$ Theory

We will first consider the zero-field, bulk Hamiltonian, and derive the formalism on which $\vec{k} \cdot \vec{p}$ solution techniques are based. We start with the time-independent single particle Hamiltonian,

$$H\,\Psi = \left[\frac{p^2}{2m} + V(\vec{r})\right]\,\Psi = E\Psi, \tag{B.1}$$

where m is the free electron mass, $\vec{p}$ and $\vec{r}$ are the momentum and position operators, respectively, and $V(\vec{r})$ is the bulk crystal potential. We consider solutions taking the form of a plane wave multiplied by a crystal-symmetric Bloch function, $u_{n\vec{k}}(\vec{r})$, having the periodicity of the lattice:

$$\Psi = e^{i\vec{k}\cdot\vec{r}}\,u_{n\vec{k}}(\vec{r}). \tag{B.2}$$

Substitution of this form of the wavefunction into the general Hamiltonian generates the $\vec{k} \cdot \vec{p}$ Hamiltonian for the $u_{n\vec{k}}(\vec{r})$,

$$\left(\frac{p^2}{2m} + \frac{\hbar}{m}(\vec{k}\cdot\vec{p}) + \frac{\hbar^2 k^2}{2m}\right)\,u_{n\vec{k}}(\vec{r}) + V(\vec{r})\,u_{n\vec{k}}(\vec{r}) = E_n\,u_{n\vec{k}}(\vec{r}). \tag{B.3}$$

In principle, this is an infinite-dimensional system, indexed (through n) by the bands to which the $u_{n\vec{k}}(\vec{r})$ correspond. To solve Equation B.3, we must first pick a basis. It is traditional to chose the orthonormal zone center Bloch functions, $u_{n0}(\vec{r})$, as the basis in which to express the $u_{n\vec{k}}(\vec{r})$. The choice of this basis set is convenient only when band structure near a $k = 0$ minimum is relevant. To expand about a different $\vec{k}_0$, the Bloch functions at that reciprocal lattice vector should be chosen. If the general $u_{n\vec{k}}(\vec{r})$ are written as,

$$u_{n\vec{k}}(\vec{r}) = \sum_{n'} C_{n'n}(\vec{k})\,u_{n'0}(\vec{r}), \tag{B.4}$$

and we note that the $u_{n'0}(\vec{r})$ satisfy,

$$\left(\frac{p^2}{2m} + V(\vec{r})\right)\,u_{n'0}(\vec{r}) = E_{n0}\,u_{n'0}(\vec{r}), \tag{B.5}$$

then substitution of the expansion into Equation B.3 yields the relationship,

$$\sum_{n'} C_{n'n}(\vec{k}) \left[E_{n0}\, u_{n'0}(\vec{r}) + \frac{\hbar(\vec{k}\cdot\vec{p})}{m}\, u_{n'0}(\vec{r}) + \frac{\hbar^2 k^2}{2m}\, u_{n'0}(\vec{r}) \right] = \sum_{n'} E_n(\vec{k})\, C_{n'n}(\vec{k})\, u_{n'0}(\vec{r}). \tag{B.6}$$

If we multiply both sides of this expression by $u^{\star}_{n'0}(\vec{r})$ and integrate over the unit cell, we end up with the C-number Hamiltonian,

$$\sum_{n'} \left[\left\{ E_{n0} + \frac{\hbar^2 k^2}{2m} \right\} \delta_{nn'} + \frac{\hbar \vec{k}}{m} \cdot \vec{p}_{nn'} \right] C_{nn'} = E_n(\vec{k})\, C_{nn'}, \tag{B.7}$$

where the parameters $\vec{p}_{nn'}$ are defined by,

$$\vec{p}_{nn'} = \int_{\Omega} u^{\star}_{n'0}(\vec{r})\, \vec{p}\, u_{n'0}(\vec{r})\, d\vec{r}. \tag{B.8}$$

This Hamiltonian, H_{∞}, is still infinite in extent, but for most problems only a few bands are of primary importance. Taking only the lowest order symmetries of the crystal, we consider states s, x, y, and z having spherical, x-, y-, and z-like symmetry, respectively. For these zone-center basis states, H_{∞} in matrix form becomes,

$$H_{\infty} = \begin{array}{c|cccc|c} & s & x & y & z & \\ \hline s & E_s + \frac{\hbar^2 k^2}{2m} & \frac{\hbar}{m}\vec{k}\cdot\vec{p}_{sx} & \frac{\hbar}{m}\vec{k}\cdot\vec{p}_{sy} & \frac{\hbar}{m}\vec{k}\cdot\vec{p}_{sz} & \\ x & \frac{\hbar}{m}\vec{k}\cdot\vec{p}_{xs} & E_x + \frac{\hbar^2 k^2}{2m} & \frac{\hbar}{m}\vec{k}\cdot\vec{p}_{xy} & \frac{\hbar}{m}\vec{k}\cdot\vec{p}_{xz} & \\ y & \frac{\hbar}{m}\vec{k}\cdot\vec{p}_{ys} & \frac{\hbar}{m}\vec{k}\cdot\vec{p}_{yx} & E_y + \frac{\hbar^2 k^2}{2m} & \frac{\hbar}{m}\vec{k}\cdot\vec{p}_{yz} & \\ z & \frac{\hbar}{m}\vec{k}\cdot\vec{p}_{zs} & \frac{\hbar}{m}\vec{k}\cdot\vec{p}_{zx} & \frac{\hbar}{m}\vec{k}\cdot\vec{p}_{zy} & E_z + \frac{\hbar^2 k^2}{2m} & \\ \hline & & & & & \end{array} \tag{B.9}$$

where extending in the matrix all around these entries are the contributions of other bands. If the magnitude of the off-diagonal terms $\hbar/m(\vec{k}\cdot\vec{p}_{nn'})$ are small however (valid for this expansion for all $\vec{k}$ near zone center), the effects of other bands may be treated as a perturbation. If, in general, we divide the states into two sets, A and B, such that set A contains all the sets in which we are interested (s, x, y, and z), and set B contains all additional states, then the perturbation theory due to Löwdin [1] may be used to iteratively remove the interactions between states in the two sets. In essence, this perturbation theory takes the infinite Hamiltonian and block-diagonalizes it in terms of the states for which there is interest. Formally, matrix elements of the renormalized Hamiltonian are related to those of H_∞ by,

$$H'_{ij} = (H_\infty)_{ij} + \sum_{\beta \in B} \frac{(H_\infty)_{i\beta}\,(H_\infty)_{\beta j}}{E_i - (H_\infty)_{\beta\beta}}. \qquad \text{(B.10)}$$

All coupling between A and B states is removed in the procedure, and a finite Hamiltonian results. The infinite sums from the perturbative calculation are typically treated as parameters to the theory, and fit to experimental data. The resulting 4×4 matrix for the $A \in \{s, x, y, z\}$ system is the 4-band $\vec{k} \cdot \vec{p}$ Hamiltonian for the states closest to the band gap of the crystal.

When spin-orbit coupling is added, the more complicated Kramers basis (in which the spin-orbit terms are diagonal) must be used as a basis for H_∞, and greater complexity arises. The Kramers basis, expressed as linear combinations of s, x, y, and z, may be written as,

u_i	$\lvert j, m\rangle$	ψ_{jm}	Band
1	$\lvert\frac{1}{2}, \frac{1}{2}\rangle_S$	$\lvert s\uparrow\rangle$	CB$\uparrow$
3	$\lvert\frac{3}{2}, \frac{3}{2}\rangle$	$\frac{1}{\sqrt{2}}\lvert(x+iy)\uparrow\rangle$	HH$\uparrow$
5	$\lvert\frac{3}{2}, -\frac{1}{2}\rangle$	$\frac{1}{\sqrt{6}}[\lvert(x-iy)\uparrow\rangle + 2\lvert z\downarrow\rangle]$	LH$\downarrow$
7	$\lvert\frac{1}{2}, -\frac{1}{2}\rangle_P$	$\frac{i}{\sqrt{3}}[-\lvert(x-iy)\uparrow\rangle + \lvert z\downarrow\rangle]$	SO$\downarrow$,
2	$\lvert\frac{1}{2}, -\frac{1}{2}\rangle_S$	$i\lvert s\downarrow\rangle$	CB$\downarrow$
4	$\lvert\frac{3}{2}, -\frac{3}{2}\rangle$	$\frac{i}{\sqrt{2}}\lvert(x-iy)\downarrow\rangle$	HH$\downarrow$
6	$\lvert\frac{3}{2}, \frac{1}{2}\rangle$	$\frac{i}{\sqrt{6}}[-\lvert(x+iy)\downarrow\rangle + 2\lvert z\uparrow\rangle]$	LH$\uparrow$
8	$\lvert\frac{1}{2}, \frac{1}{2}\rangle_P$	$\frac{1}{\sqrt{3}}[\lvert(x+iy)\downarrow\rangle + \lvert z\uparrow\rangle]$	SO$\uparrow$

where the corresponding bands are indicated, and ↑ and ↓ refer to spin up and down, respectively. There are several sources in which the full derivation of the 8-band $\vec{k}\cdot\vec{p}$ Hamiltonian is given [2, 3, 4]. The details which remain in the calculation are the exact specification of the infinite sums as parameters, and the reduction of these to a minimal set through consideration of the symmetry properties of the crystal.

B.3 Incorporation of the Magnetic Field

The classic seminal work on the incorporation of magnetic fields into $\vec{k}\cdot\vec{p}$ theory is due to Luttinger [4]. This paper is a must-read for anyone considering the calculation of realistic band structure in a magnetic field. In the zero-field, bulk Luttinger Hamiltonian, the $\vec{k}\cdot\vec{p}$ matrix elements are dependent upon the Luttinger parameters, $\gamma_1, \gamma_2, \gamma_3, \kappa,$ and q, which are in turn related to the infinite sums resulting from the perturbative renormalization of H_∞. It should be noted that these parameters are *not* identical to those used by Kane [2], and the differences are the source of great consternation to those learning $\vec{k}\cdot\vec{p}$ theory. For a magnetic field applied along one of the principle axes of the crystal, the problem reduces to that of introducing the field dependence into the general bulk Hamiltonian. For a field at an arbitrary angle with respect to the crystal, however, more rigor must be introduced to properly treat the angular momentum states of the various bands in a new basis aligned with the magnetic field (see Luttinger [4]). In either case, raising and lowering operators, defined as in Chapter 2, may be introduced in terms $\overline{k}_x$ and $\overline{k}_y$. If $a^\dagger$ and a are introduced, and the anisotropy terms proportional to $(\gamma_3 - \gamma_2)a^\dagger$ (and similarly for a) are neglected (which is a reasonable assumption for the Γ-valleys of most III-V compounds), the eigenvector of the 8-band $\vec{k}\cdot\vec{p}$ Hamiltonian takes the form,

$$\psi = \left(c_1\, h_n,\ c_3\, h_{n-1},\ c_5\, h_{n+1},\ c_7\, h_{n+1},\ c_2\, h_{n+1},\ c_4\, h_{n+2},\ c_6\, h_n,\ c_8\, h_n\right), \qquad \text{(B.11)}$$

where the h_l are simple quantum harmonic oscillator wavefunctions. The resulting eigenstate contains a mixture of Landau levels indexed by l, and is, in fact, a state of constant projected angular momentum along the axis of the magnetic field. As shown in Chapter 4, the orbital angular momentum attributable to the Landau orbits is proportional to the Landau index l, and the sum of l and m_j, the z-component of the total band angular momentum, is constant for the above eigenstates of the 8-band Hamiltonian. We therefore expect a selection rule $\Delta(l + m_j) = 0$ for resonant, elastic tunneling in magnetic fields applied perpendicular to the heterointerfaces. The incorporation of the bulk 8-band $\vec{k} \cdot \vec{p}$ Hamiltonian in a magnetic field into heterostructure transport calculations is the work of Liu *et al.* [5].

Bibliography

[1] P. Löwdin, J. Chem. Phys. **19**, 1396 (1951).

[2] E. O. Kane, in *Semiconductors and Semimetals* **1**, edited by R. K. Willardson and A. C. Beer (Academic Press, New York, 1966).

[3] G. Wu, Ph.D. Thesis, California Institute of Technology, 1988.

[4] J. M. Luttinger, Phys. Rev. **102**, 1030 (1956).

[5] Y. X. Liu, D. Z.-Y. Ting, and T. C. McGill, unpublished.