RICE UNIVERSITY

Spectroscopy of III-V Narrow Gap Semiconductor Quantum Wells

by

Diane C. Larrabee

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE

Doctor of Philosophy

APPROVED, THESIS COMMITTEE:

[Signatures of committee members]

Junichiro Kono, Chairman
Associate Professor of Electrical and Computer Engineering

Daniel M. Mittleman
Associate Professor of Electrical and Computer Engineering

Frank K. Tittel
J.S. Abercrombie Professor in Electrical and Computer Engineering and Professor of Bioengineering

Houston, Texas
January, 2006
INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.
Abstract

Spectroscopy of III-V Narrow Gap Semiconductor Quantum Wells

by

Diane C. Larrabee

The narrow gap III-V semiconductors, InAs/AlSb/GaSb and InSb, exhibit an array of extreme physical properties, from the lightest effective mass and largest nonparabolicity of III-V semiconductors to heterostructure conduction band offsets ranging from -0.15 to +2.0 eV. In this work, I present three spectroscopic techniques which exploit these unusual properties to provide new insight into the physics of these materials.

First, my measurement of cyclotron resonance in InAs/AlSb and InSb/AlInSb quantum wells was the first spectroscopic application of a new laser, the THz quantum cascade laser. The physical properties mentioned above put these materials into an experimentally accessible range, and InAs's high room temperature mobility and low temperature carrier density enabled us to explore a large temperature range. Previous investigations of other materials in limited temperature ranges had suggested what we confirmed: the cyclotron resonance effective mass increases with temperature, contrary to theoretical expectations.
Second, we applied time resolved cyclotron resonance to InSb quantum wells for the first time. Because of InSb's large effective g-factor and nonparabolicity, time resolved cyclotron resonance enabled us to monitor the carrier relaxation and recombination from each Landau- and Zeeman-quantized state directly in time. This unprecedented level of detail could be extended to longer times to probe spin-flip relaxation, a significant parasitic process in quantum computation.

Finally, I measured intersubband absorption in narrow InAs/AlSb quantum wells with widths from 10.5 to 1.8 nm. I observed the highest energy intersubband resonance in InAs/AlSb quantum wells: 650 meV at 77 K in a 1.8 nm well. I also performed detailed measurements of the temperature dependence of intersubband absorption and confirmed the correlation between the integrated intensity of intersubband absorption and the carrier distribution inferred from Shubnikov-de Haas and Hall measurements. Because of InAs/AlSb intersubband transitions' large accessible energy and temperature robustness, they are ideal candidates for resonant nonlinear optics. In particular, I discuss the potential of InAs/AlSb double quantum wells as a compact, room temperature, and coherent THz source. Such a source could revolutionize chemical sensing by providing convenient access to the strong fundamental vibrational fingerprints which all molecules have in the THz, potentially transforming applications from medicine to the military.
Acknowledgments

I am grateful to my advisor, Dr. Jun Kono, for all the advice and support he has given me during the last five years. He gave me the opportunity to work on several different experiments and had the patience to let me learn many experimental techniques. He is remarkably good at asking the right questions. He is a strong role model, especially as a teacher and a writer.

Thanks to my thesis committee members, Dr. Daniel Mittleman and Dr. Frank Tittel (again!), for taking the time to serve on my committee.

This thesis contains the efforts of many people, all of whom were a pleasure to work with:

**CW Cyclotron Resonance:** The sample was grown by Prof. Sasa’s group at Osaka Institute of Technology (see below). The THz quantum cascade laser (QCL) was fabricated by Giacomo Scalari, Lassaad Ajili, and Jerome Faist at the University of Neuchâtel, Switzerland, and Harvey Beere, Giles Davies, Edmund Linfield, and David Ritchie at the University of Cambridge, U. K. The modeling is being done by Gary Sanders and Chris Stanton at the University of Florida. Thanks to Dr. Tittel for initiating the collaboration with Dr. Faist. Many thanks to Anatoliy Kosterev, Yury Bakhirkin, Stephen So, and Adrian Barkan at Rice for help with setting up the QCL.

**Time Resolved Cyclotron Resonance:** The sample was grown by Seokjae Chung and Mike Santos at the University of Oklahoma. The experiment was per-
formed at the Stanford Free Electron Laser Center (FEL) with Doug King, Giti Khodaparast, and Dr. Kono. The Pidgeon-Brown calculation was by Giti Khodaparast. Thanks to the many people who kept the FEL happy: Alan Schwettman, Todd Smith, Richard Swent, Takuji Kimura, George Marcus, Bill Armstrong, and Don Keegan.

**THz Generation in Antimonide-based QWs:** The samples were grown by Masato Nakai, Furukawa Masashi, M. Karasaki, Yoji Nakajima, and K. Ueda in the group of Profs. Shigehiko Sasa and Masataka Inoue at the Osaka Institute of Technology, Japan. The MIR QCLs were lent by Claire Gmachl, then at Bell Labs. The bandstructure calculations were by Konstantin Kolokolov and the many-body absorption calculations were by Jiangzhong Li, both in the group of Cun-Zheng Ning at NASA Ames Research Center. The difference frequency generation calculations, both edge-coupled and ultrafast, were by Alexey Belyanin of Texas A&M University. Giti Khodaparast and Jun Tang helped with the early spectroscopy work.

At Rice, I would like to thank Dr. Kono’s group, in particular Gordana Ostojic, who encouraged me to join in the first place; Chanjuan Sun, who made our time repairing the CPA/OPA enjoyable; and all the bright-eyed undergrads I’ve worked with. Thanks also to Michael Dye, Dick Chronister, Lam Yu, and Clayton Simien for helping with electronics, and to Dwight Dear for teaching me to machine.

I am very grateful to Giti Khodaparast, now at Virginia Tech. She got me started in the lab beginning in the days before I knew what an O-ring was. She taught me
the importance of working long hours, figuring out which details are the important ones, and eating good meals. She is one of the most generous people I know, and I feel fortunate to have worked with her.

I am very grateful to my boyfriend, Lam Yu, for his humor and his unfailing support and encouragement. And thanks most of all to my parents, who fed me a good meal every Sunday and were always ready to listen.
Contents

Abstract ii
Acknowledgments iv
List of Illustrations xi
List of Tables xx

1 Introduction 1

1.1 Narrow Gap Semiconductors ........................................ 1

1.2 All-Optical Intersubband THz Generation .......................... 3

1.2.1 The THz Gap .................................................. 3

1.2.2 Why InAs/AlSb QWs? ........................................ 6

2 Quantum Well Band Structure Calculations 9

2.1 k·p Method for Bulk Band Structure Calculation .................. 9

2.1.1 Löwdin Perturbation Theory .................................. 11

2.1.2 k·p Hamiltonian ............................................. 13

2.1.3 Hamiltonian Matrix in the Luttinger-Kohn Representation .... 15

2.1.4 Small Gap Approximation .................................... 16

2.1.5 4-band Approximation ........................................ 21

2.1.6 Measurement of Key Parameters ............................... 23

2.2 Quantum Well Band Structure Calculations ....................... 24

2.2.1 Envelope Function Approximation ............................ 25
2.2.2 Kronig-Penney Model ........................................ 28
2.2.3 Ben Daniel-Duke Model ...................................... 30
2.2.4 Kane Model with Confinement ............................... 31
2.2.5 Doped QWs: Self-consistent Solution of the Quantum Well Potential ........................................ 33
2.2.6 Strained QWs: the Bir-Pikus Hamiltonian ............... 34

3 CW Cyclotron Resonance ........................................ 37

3.1 Origin of Landau Levels ....................................... 39
3.2 Review of the Temperature Dependence of the Effective Mass .......... 42
3.3 Experiment and Results ........................................ 46
3.4 Analysis .......................................................... 48
  3.4.1 Bastard Model for $m^*(T)$ .................................. 48
  3.4.2 Detailed Model for $m^*(T)$ .................................. 52

4 Time-Resolved Cyclotron Resonance .......................... 55

4.1 Review of Relaxation Measurements in Landau-Quantized Systems . 57
4.2 Experiment and Results ......................................... 60
4.3 Pidgeon-Brown Model .......................................... 63
4.4 Simulation ......................................................... 66
4.5 Further experiments ............................................. 70

5 Intersubband Transitions ........................................ 72

5.1 Sample Growth and Characterization ........................... 72
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1.1</td>
<td>Carrier Distribution</td>
<td>73</td>
</tr>
<tr>
<td>5.2</td>
<td>Intrasubband Absorption in Single QWs</td>
<td>75</td>
</tr>
<tr>
<td>5.2.1</td>
<td>Selection Rules and Experimental Geometry</td>
<td>75</td>
</tr>
<tr>
<td>5.2.2</td>
<td>Temperature Dependence</td>
<td>78</td>
</tr>
<tr>
<td>5.2.3</td>
<td>High Energy</td>
<td>81</td>
</tr>
<tr>
<td>5.2.4</td>
<td>No Rashba Splitting</td>
<td>83</td>
</tr>
<tr>
<td>5.2.5</td>
<td>GaSb Substrates</td>
<td>84</td>
</tr>
<tr>
<td>5.2.6</td>
<td>Integrated Absorption</td>
<td>85</td>
</tr>
<tr>
<td>5.3</td>
<td>Double QWs for Difference Frequency Generation</td>
<td>88</td>
</tr>
<tr>
<td>5.4</td>
<td>Modeling of Many-Body Effects</td>
<td>92</td>
</tr>
<tr>
<td>5.4.1</td>
<td>Single-particle Broadening Mechanisms</td>
<td>92</td>
</tr>
<tr>
<td>5.4.2</td>
<td>Intersubband Semiconductor Bloch Equations</td>
<td>93</td>
</tr>
<tr>
<td>5.4.3</td>
<td>Many-body Effects</td>
<td>96</td>
</tr>
<tr>
<td>5.4.4</td>
<td>Temperature Dependence</td>
<td>97</td>
</tr>
<tr>
<td>6</td>
<td>QCL Difference Frequency Generation</td>
<td>100</td>
</tr>
<tr>
<td>6.1</td>
<td>Second Order Nonlinear Susceptibility ($\chi^{(2)}$)</td>
<td>100</td>
</tr>
<tr>
<td>6.2</td>
<td>Resonant Nonlinear Optics in QWs</td>
<td>104</td>
</tr>
<tr>
<td>6.3</td>
<td>Phase Matching</td>
<td>107</td>
</tr>
<tr>
<td>6.4</td>
<td>Power in the Single-Bounce Geometry</td>
<td>114</td>
</tr>
<tr>
<td>6.5</td>
<td>Power in the Non-Phase Matched Edge Coupled Geometry</td>
<td>116</td>
</tr>
<tr>
<td>6.6</td>
<td>Experiment</td>
<td>118</td>
</tr>
</tbody>
</table>
6.7 Assessment ............................................................................. 123

7 Coherence Oscillations ......................................................... 126

7.1 Density Matrix ................................................................. 127

7.2 Review: Coherence Oscillations in QWs ....................... 128

7.3 Proposal of Few-Cycle THz Pulse via Resonant Mixing in QWs . 130

7.4 Experiment ................................................................. 132

8 Conclusions ........................................................................... 139

8.1 Future Work ................................................................. 140

Bibliography ............................................................................. 143
Illustrations

1.1 Band alignments of the 6.1 Å semiconductors. .......................... 2

1.2 Conduction band profile of the active region of a THz QCL. From [1]. 5

1.3 Designs for an intersubband laser (left) and difference frequency
mixer (right). ................................................................. 7

2.1 The full band structure of InAs, calculated using the nonlocal
pseudopotential method. From [2]. ........................................ 12

2.2 Dispersion relations for GaAs, InAs, and InSb. The black lines
include the energy-dependent effective mass (eq. 2.16), and the grey
lines are the parabolic approximation (eq. 2.19). This figure clearly
illustrates that nonparabolicity increases as the band gap decreases.
The plots are extended beyond their range of validity ($k \sim \pm 0.05$ [3])
in order to show their overall behavior more clearly. .................. 20

2.3 Diagram of the Kronig-Penny superlattice. ............................. 28

2.4 $\cos(qL)$ vs. energy (eq. 2.41) for (left) GaAs/AlGaAs and (right)
InAs/AlSb superlattices, as described in the text. The horizontal
lines $\cos(qL) = \pm 1$ are also plotted. Note the different horizontal
and vertical scales. .......................................................... 30
2.5 Left: calculated self-consistent potential of a 10 nm InAs/AlSb QW with \( n = 1.84 \times 10^{12} \) cm\(^{-2}\). Right: lowest few conduction band wavefunctions. .......................................................... 35

3.1 Density of states of a semiconductor in 3D (left) and 2D (right).
Dashed lines are without magnetic field, solid lines are with magnetic field. .......................................................... 42

3.2 Temperature dependence of the Hall density and mobility of a 10 nm InAs/AlSb multiple QW sample (0209042T). ....................... 47

3.3 Transmission as a function of magnetic field at several temperatures for InAs/AlSb quantum wells (sample 0209042T). The quantum cascade laser wavelength is 64 \( \mu \)m (4.7 THz). ....................... 49

3.4 (a) Resonance field as a function of photon frequency at 1.5 K for InAs/AlSb quantum wells (sample 0209042T). The straight line has a slope of 1.51 T/THz, corresponding to an effective mass of 0.042 \( m_0 \). (b) Cyclotron mass versus temperature at 4.7 THz for the InAs/AlSb quantum wells. The experimental mass (triangles) increases with increasing temperature, whereas the theoretical masses (open and filled circles) for two possible CR transitions show the opposite behavior. ....................... 50

3.5 Experimental and calculated \( m^* \) vs. \( T \). ....................... 54
4.1 Schematic of the TRCR measurement, showing the interband pump and inter-Landau-level probe (not to scale). ........................................ 56

4.2 Picosecond TRCR spectra of photocreated electrons in InSb QWs. The photoinduced transmission change is plotted as a function of time delay and magnetic field. The photoinduced transmission change is defined as $(T_0 - T)/T_0$, where $T_0$ is the transmission value before the arrival of the NIR pump pulse. The wavelength of the FIR radiation was 42 μm and the sample temperature was 1.5 K. .......................... 62

4.3 (a) Transmission change of the FIR probe pulse as a function of magnetic field at different time delays. The wavelength of the probe was 42 μm, the sample temperature was 1.5 K, and the NIR fluence was ~1.3 mJ/cm². (b) Average cyclotron mass vs. time delay. (c) Carrier density ($n$) vs. time delay. (d) Average momentum scattering time ($t_p$) vs. time delay. .......................... 63

4.4 Conduction band Landau levels calculated using a modified Pidgeon-Brown model, which takes into account confinement and strain effects. Black (gray) lines are for the first (second) subband. Solid (dashed) lines are spin up (down) states. The arrows show the possible CR transitions for 42 μm radiation. .......................... 65
4.5  (a) Simulated TRCR spectra for the traces in fig. 4.3, based on the energy levels shown in fig. 4.4. Simulation successfully reproduces the main observed features arising from nonparabolicity. (b) Cartoon at left (right) shows a Fermi distribution at short (long) time delay, when the carrier population is large and hot (small and cold). (c) Simulated carrier temperature as a function of time delay. .......................... 67

4.6  Diagram of Landau levels (top) and the TRCR signal expected at long time delays when spin-flip transitions are allowed (middle) and forbidden (bottom). ................................. 71

5.1  Three common experimental geometries for the observation of ISBT. Counterclockwise from top left, the Brewster angle geometry, the trapezoid geometry, and the multi-bounce parallelogram geometry (this work). ................................................. 78

5.2  The intersubband absorption spectrum of a 7.4 nm InAs/AlSb QW (sample 0201181T) as a function of temperature. Curves vertically offset for clarity. ................................. 79

5.3  Temperature dependence of the intersubband absorption peak energy (left) and full-width at half maximum (right) for InAs/AlSb QWs with densities around $1 \times 10^{13}$ cm$^{-2}$ per well. ................................. 80
5.4 The absorption spectrum of the highest energy $E_{12}$ intersubband resonance, in a 2.1 nm InAs/AlSb QW (sample 0403081T), plotted as $p$-polarized (intersubband active) and $s$-polarized (intersubband inactive) absorption. ........................................ 82

5.5 The ISBT FWHM vs. well width. The FWHM goes like $(\text{width})^{-1.1}$. The FWHM increases sharply for well widths less than 3 nm. .... 82

5.6 Intersubband absorption spectra of three samples with native interface asymmetry. The splitting in the sample on the left (0205171T) was tentatively attributed to Rashba splitting until two nearly identical samples (center, 0311122T, and right, 0211062T) showed no such splitting. Sample 0311122T has almost two orders of magnitude fewer carriers compared with the other samples, hence the smaller integrated absorption. ........................................ 84
5.7 Left: Intersubband absorption coefficient of two samples measured in the Brewster geometry at 77 K: red: 20 periods of 10 nm QWs on a GaAs substrate with a density of $3.5 \times 10^{12}$ cm$^{-2}$ per well (0211012T); blue: 20 periods of 7 nm QWs grown on a GaSb substrate with a density of $5 \times 10^{12}$ cm$^{-2}$ per well (0411102T).

Right: Intersubband absorption coefficient of sample 0211012T in the multibounce geometry at 77 K. Note the different vertical scale. The resonance is wider in the multibounce geometry because the absorption is saturated.

5.8 The calculated lowest energy levels and wavefunctions of a typical double QW structure designed for intersubband difference frequency generation. Note the break in the vertical axis.

5.9 Intersubband absorption spectra at 77 K of seven ADQW samples designed for THz difference frequency generation. Each graph is labeled with the sample name, total density in cm$^{-2}$, and number of ADQW periods. Sample 0404231T is Si-doped in the wide well but still shows clearly resolved THz splitting.

5.10 High energy intersubband absorption spectra at 77 K of three ADQW samples with high densities. In sample 0404191T, the Fermi energy is above $E_3$ and no $E_{12}$ or $E_{13}$ resonance is observed.
5.11 Calculated dispersion of the first three subbands of a 10 nm InAs/AlSb QW. The horizontal line is the Fermi energy for an electron density of $1.84 \times 10^{12}$ cm$^{-2}$. The “parabolic” line has constant effective mass.

5.12 Calculated ISBT spectra for 10 nm InAs/AlSb QW at 4.5 K with a density of $8 \times 10^{11}$ cm$^{-2}$. The effects of many-body interactions on the intersubband absorption are shown singly and in combination. In (e), the dotted line is the single particle spectrum. From [4].

5.13 Nonparabolic dispersion relation and absorption coefficient at 0 K (left) and 300 K (right). The arrows indicate the single-particle ISBT energies.

6.1 Energy levels used in calculation of $\chi^{(2)}$.

6.2 The two different schemes for incorporating intersubband nonlinear mixing regions into QCLs used in ref. [5]. Later work found that the scheme on the right, with the mixing region and the active region in the same QWs, gave better modal overlap and higher mixing efficiency [6].
6.3 Top: intersubband absorption of the phase matched SHG sample, showing the slightly off-resonance SHG-active intersubband resonances and the phase matching resonance. From [7]. Bottom: schematic diagram of the dispersion due to the phase matching resonance. ................................. 110

6.4 Top: Schematic diagram of a material dispersion which achieves phase matched DFG for a small difference frequency. Bottom: The associated absorption coefficient. .................................. 112

6.5 The angled multi-bounce phase matching geometry. ................. 113

6.6 Diagram of the pump and difference frequency angles for a phase matched interaction in the single-bounce trapezoid geometry. ...... 114

6.7 Diagram of the edge coupled geometry showing the pump polarization direction. ................................................. 118

6.8 Schematic diagram of the QCL circuit. ................................. 119

6.9 Diagram of setup for non-phase matched edge coupled DFG. ....... 120

6.10 The 250 kHz pulse train of the lasers is chopped at 200 Hz for detection by the bolometer: schematic diagram. ................. 121

7.1 Intersubband energy levels and pump pulse spectrum for ultrafast excitation of DFG. ............................................... 126

7.2 The experimental setup for DFG with the OPA pump. ............... 134
7.3 Setup for intersubband linear absorption spectrum measurement  
with the OPA. ................................................................. 137

7.4 Setup for intersubband absorption saturation measurement. ....... 138

8.1 Experimental setup for optically detected cyclotron resonance with a 
THz QCL. From [8]. .......................................................... 141

8.2 Calculated intersubband absorbance for InAs/AlSb QWs, all with a 
density of $1 \times 10^{12}$ cm$^{-2}$ per well, from [9]. Dotted lines: single 
particle spectrum; dashed lines: various many-body effects; solid 
lines: full many-body calculation. ....................................... 142
Tables

6.1 Calculation of DFG power with QCL pumps in single-bounce geometry. 116

6.2 Typical time average of the power product for a fixed average pump
    power. Various combinations of lasers. In arbitrary units. 124

6.3 \( \exp(-\alpha_{THz}l) \) calculated for GaAs at the THz wavelengths and
    sample path lengths used in this experiment. 125

7.1 Samples used for DFG with the OPA. 133
Chapter 1

Introduction

1.1 Narrow Gap Semiconductors

The narrow gap III-V semiconductor heterostructures are a relatively new group of materials. The binary compounds are InAs, GaSb, and AlSb, all nearly lattice matched with lattice constants around 6.1 Å, and InSb with a lattice constant of 6.48 Å. Although InAs/GaSb heterostructures have been studied for decades [10], it is only since the early 90's that high quality InAs/AlSb and InSb/AlInSb QWs have been available. InAs in particular has been studied extensively for electrical device applications, such as high electron mobility transistors, but more remains to be learned about these materials' optical properties.

The narrow gap semiconductors have a variety of unusual characteristics. Due to their narrow band gaps compared to GaAs ($E_g = 1.52$ eV in GaAs, 0.417 eV in InAs, and 0.235 eV in InSb), these materials have small effective masses, leading to high mobility [11], and large nonparabolicity. In addition, their strong spin-orbit coupling leads to a large effective $g$-factor.

6.1 Å semiconductor heterostructures also have unusual band alignments, as shown in fig. 1.1. InAs/GaSb is semimetallic, and both electrons and holes can be confined in separate layers at the same time. InAs/AlSb is a semiconducting
quantum well (QW). It is type II, meaning that the confined electron states are in the well while the hole states are in the barrier. InSb/AlInSb, on the other hand, is a type I QW, with both electrons and holes confined in the well.

This thesis is divided into seven sections. Chapter 2 describes $k\cdot p$ semiconductor bandstructure calculations and several QW modeling techniques which will be applied throughout the thesis. Chapter 3 describes cyclotron resonance measurements of InAs/AlSb QWs, which was the first spectroscopic application of THz quantum cascade lasers (QCLs). These cyclotron resonance measurements cover temperatures from 4 to 300 K, and our collaborators are working to explain our results using $k\cdot p$ theory. Chapter 4 describes time-resolved cyclotron resonance measurements of InSb/AlInSb QWs. This unusual technique takes advantage of the large nonparabolicity in InSb and allows us to monitor charge relaxation and recombination in the
Landau-quantized system directly in time.

The rest of the thesis is concerned with THz generation via optical pumping of intersubband resonant nonlinearities in InAs/AlSb QWs. In the following section, we will give the motivation for this work.

1.2 All-Optical Intersubband THz Generation

1.2.1 The THz Gap

The terahertz (THz) region of the electromagnetic spectrum, from 1 to $10 \times 10^{12}$ Hz (300-30 μm; 4-40 meV), is rich with possibilities for communications, chemical sensing, astronomy, and condensed matter spectroscopy [12, 13]. But, this spectral range is currently under-utilized due to a lack of convenient sources and detectors. Approaching the THz from the high-energy side, optical sources are limited by the difficulty of creating materials with a small enough energy level separation. Thermal population of those levels is also an obstacle, since $k_B \times 300 \text{ K} = 26 \text{ meV} = 6 \text{ THz}$.

A variety of coherent THz sources exist. Molecular gas lasers have many lines in the far-infrared. Broadly tunable or broadband THz sources include the free electron laser [14–16]; Auston switches [17]; and p-type germanium laser, which operates by stimulated cyclotron resonance emission in crossed electric and magnetic fields [18, 19]. These sources are cumbersome, however, requiring significant infrastructure or cryogenic cooling. Compact, table-top sources include optical rectification of fs pulses at semiconductor surfaces [20] and in periodically poled lithium niobate
[21, 22], polariton excitation by angled fs beams in ferroelectrics [23], photomixing of visible diode lasers in low temperature grown GaAs [24], and THz optical parametric oscillators [25].

One remarkable new single-color THz source is the quantum cascade laser (QCL). First demonstrated in the midinfrared [26], this unipolar, electrically driven laser is based on transitions between quantized levels in narrow quantum well structures (Fig. 1.2) [1]. Design improvements have led to lasing at wavelengths as long as 140 μm (2.1 THz) [27], or 159 μm (1.9 THz) in a magnetic field [28], with powers as high as 15 mW at 10 K [29]. Despite the remarkable achievements of THz QCLs, they are subject to intrinsic limitations including free carrier absorption, which is proportional to the square of the emission wavelength, and reduction of population inversion due to electron-electron scattering. Both of these limitations are more significant at higher carrier densities. Therefore, it is worthwhile to investigate optically driven THz sources. These have fewer free carriers because they do not require heavily doped contact layers and because it is possible to create carriers only in selected regions by a judicious choice of pumping wavelengths. Fewer heterostructure layers and fewer free carriers may lead to compact devices with superior THz performance.
Figure 1.2: Conduction band profile of the active region of a THz QCL. From [1].

Several optically pumped THz sources have been demonstrated. Gauthier-Lafaye et al. created a three-level laser within the conduction band of GaAs/Al$_{0.35}$Ga$_{0.65}$As asymmetric coupled double quantum wells [30]. It is pumped by a CO$_2$ laser and emits at 15.5 µm with a peak power of 0.4 W. Sirtori et al. demonstrated difference frequency mixing of CO$_2$ lasers in GaAs/Al$_{0.33}$Ga$_{0.67}$As symmetric coupled double quantum wells, achieving up to 0.7 nW emission at 60 µm [31]. One significant improvement in the convenience of these sources would be solid-state pumping, for example by a near-infrared (NIR) diode laser. But this is not possible in GaAs/Al$_x$Ga$_{1-x}$As quantum wells, in which the conduction band (CB) offset is only about 1 eV for $x = 1$ [32]. The maximum possible intra-well pumping energy is less than this value. So in order to use a NIR pump, we must use a material system with a larger CB offset.
1.2.2 Why InAs/AlSb QWs?

We have chosen to work with the 6.1 Å semiconductors because of the favorable characteristics of intersubband resonances in general and these materials in particular. The conduction band offsets in this material system are enormous: as much as 2 eV between InAs and AlSb, compared with around 0.5 eV in GaAs/AlGaAs. This large band offset allows us to confine high energy subbands in InAs/AlSb QWs. In addition, the intersubband resonance energy can be tuned within nearly the entire range of the band offset, giving us great flexibility in designing all-intraband devices. Another favorable feature of InAs/AlSb QWs is their large intraband dipole matrix elements. These arise both from the strong confinement, which concentrates the wavefunction in the well, and the small InAs effective mass, which allows a wider well to be grown for a given intersubband transition energy. The large dipole matrix element increases the strength of linear and nonlinear intersubband optical interactions. Intersubband transitions are also robust at high temperatures and low mobilities due to the collective nature of the resonance.

It has been shown that the properties of narrow gap semiconductors listed above, along with a reduced optical phonon scattering rate, have a favorable impact on QCL performance [33]. In fact, QCLs based on 6.1 Å semiconductors have been demonstrated in the mid-infrared from 4.5 to 14 µm [34–37]. Much of the work on these lasers aims to operate them at shorter wavelengths, taking advantage of the large conduction band offsets possible in this material system [38].
Figure 1.3: Designs for an intersubband laser (left) and difference frequency mixer (right).

For optically pumped THz generation, we considered two schemes, shown in fig. 1.3. The right hand diagram shows a difference frequency generation structure. The quantum wells are engineered to have intersubband resonance energies close to two pump energies. This creates a large second-order nonlinear optical susceptibility, which enables the production of light at the difference energy of the two pumps. Because of the large CB offset between InAs and AlSb, it may be possible to create resonance energies in the near-infrared, where convenient diode laser pumps are available. Because the intersubband resonance is fairly robust with increasing temperature, it is possible to use this structure at room temperature. Figure 1.3 (left) shows a design for an intersubband THz laser [39]. Because the intersubband pump energy is large, it is possible to locate the lasing levels well above the Fermi level and avoid thermal population of the lasing level. One possible drawback of this design is interband absorption in the GaSb layers, if the pump photon energy exceeds its band gap (0.8 eV).

The second half of the thesis describes efforts to create a THz difference frequency
mixer based on intersubband resonances in InAs/AlSb double QWs. Chapter 5 presents measurements of intersubband absorption in single and double quantum wells. This chapter also outlines our collaborators' theoretical treatment of many-body effects on the intersubband resonance energy and lineshape. Chapter 6 presents an effort to observe difference frequency mixing of MIR QCLs. Two schemes were considered: the single-bounce technique used by Sirtori et al. [31], and a non phase matched edge-coupled geometry, simulated by a collaborator. Chapter 7 presents another attempt based on excitation with an ultrafast MIR pump, and a simulation by a collaborator. Chapter 8 presents conclusions and future work.
Chapter 2

Quantum Well Band Structure Calculations

Predicting the optical properties of a quantum well (QW) requires modeling its band structure [40]. Once the QW’s energy levels and wavefunctions have been calculated, parameters such as optical transition energies and probabilities can be computed and compared with experiment. The first half of this chapter describes the 8-band $k\cdot p$ method for calculating bulk band structures. The second half treats several methods for calculating QW band structures. These calculations underlie the modeling of all of the experiments described in this thesis. Although I did not do any of the elaborate band structure calculations myself, understanding where they come from is important for interpreting their results.

These calculations deal entirely with the one-electron picture. As we will see in chapters 3 and 4, this picture adequately describes cyclotron resonance in narrow gap QWs, in which Kohn’s theorem [41] is satisfied. However, many-body effects play a major role in the intersubband resonances presented in chapter 5 [42]. Calculation of many-body effects will be described in sec. 5.4.3.

2.1 $k\cdot p$ Method for Bulk Band Structure Calculation

There are many methods for calculating band structure, for example the pseudopotential and tight binding methods [11]. However, most of these methods are compu-
tationally intense and do not easily lead to physical insight. The $\mathbf{k}\cdot\mathbf{p}$ method, on the other hand, is computed by a matrix diagonalization and takes as its inputs physically relevant parameters such as energy gaps and effective masses. Consequently, the $\mathbf{k}\cdot\mathbf{p}$ method has become a popular band structure calculation method.

The $\mathbf{k}\cdot\mathbf{p}$ method is a way to solve the one-electron Schrödinger equation for the eigenvalues and eigenfunctions of the bands in a crystal. In principle, this is a very complicated problem, since there are a large number of occupied bands and an infinite number of unoccupied bands, and all of the bands interact with each other via the Hamiltonian. However, the strength of the interaction between any two bands is inversely proportional to their energy separation. In other words, only bands which are close together in energy (and have the right symmetries) will interact strongly. So, the $\mathbf{k}\cdot\mathbf{p}$ method takes a modified perturbative approach: the exact interactions within the set of close together bands are renormalized (shifted in energy) by the interactions with remote bands as a second order perturbation. The larger the number of exact bands, the more accurate and the more difficult the calculation.

The $\mathbf{k}\cdot\mathbf{p}$ method evolved gradually into the powerful technique it is today. Bouckaert et al. [43] in 1936 were perhaps the first to formulate it; Seitz [44] used it in 1940 to derive the effective mass. In 1950, Schockley [45] extended it to incorporate degenerate bands, and in 1955, Dresselhaus et al. [46] added the spin-orbit interaction. Cardona and Pollack [47], [48] showed in 1966 that by including 30 bands exactly in the $\mathbf{k}\cdot\mathbf{p}$ method, they could accurately describe the entire Brillouin zone. Several
modern groups favor this global approach [49], [50]. However, the most popular variety of \( k \cdot p \) method today is the 8-band formulation, reviewed by Kane in 1966 [51]. In the next section, we will outline this approach.

### 2.1.1 Löwdin Perturbation Theory

In the Löwdin perturbation approach to the \( k \cdot p \) method, the bands are divided into two groups, \( A \) and \( B \) [40], [51]. Group \( A \) consists of the bands which are close together in energy and therefore interact strongly. These interactions are treated exactly. Then, the interactions between group \( A \) and all the other, remote bands (group \( B \)) are added as a second-order perturbation. The renormalized interactions of the group \( A \) bands are:

\[
U_{ij}^A = h_{ij} + \sum_{\beta}^B \frac{h'_{i\alpha} h'_{\beta j}}{h_{ii} - h_{\beta\beta}}
\]  

(2.1)

where \( h \) is the interaction matrix element, \( h = h^0 + h' \), \( i, j \) are in \( A \), and \( \beta \) is in \( B \). The eigenfunctions and eigenvalues of the unperturbed Hamiltonian, \( h^0 \), are assumed to be known. The Löwdin perturbation method is valid when:

\[
|h_{ij}| \ll |E_i - E_\beta|,
\]

(2.2)

that is, the matrix elements between the strongly interacting bands are much smaller than the energy differences between the strongly interacting and the remote bands.

Kane’s study of the \( k \cdot p \) method was spurred by experiments on InSb [52], in which the low temperature band gap is 0.235 eV and the split-off energy is 0.81 eV.
Figure 2.1: The full band structure of InAs, calculated using the nonlocal pseudopotential method. From [2].

For InSb, then, a sensible choice for group A is the conduction, heavy hole, light hole, and split-off bands (CB, HH, LH, and SO). In fact, this is the case for all III-V zincblende semiconductors: the separations between the conduction, heavy hole, light hole, and split-off bands at the Gamma point are only about 1-3 eV, whereas all the other bands are at least 5 eV away. Figure 2.1 shows the full band structure of InAs, calculated using the nonlocal pseudopotential method. The red box highlights the group A bands around the Γ point.

This 8-band $k \cdot p$ method (4 bands × 2 spins) is a favorable compromise between accuracy and computational difficulty. Eight bands are enough to describe the dispersion accurately in only a small region of the Brillouin zone (about one quarter [32]) around some central wavevector $k_0$. In III-V semiconductors, $k_0$ is almost always chosen to be the direct gap. Since the optical properties of a semiconductor typi-
cally depend only on low-energy excitations of electrons and holes, which stay near
the band edges, the 8-band model works well for describing the optical properties of
semiconductors.

2.1.2 k·p Hamiltonian

The following description of the 8-band k·p method is due to Kane [51] and Bastard
[3] unless otherwise noted. Schrödinger's equation for a single electron in a periodic
potential \( V(\mathbf{r}) \) is:

\[
\left\{ \frac{p^2}{2m_0} + V(\mathbf{r}) + \frac{\hbar}{4m_0^2c^2} [\nabla V \times \mathbf{p}] \cdot \sigma \right\} \psi = E\psi, \tag{2.3}
\]

where \( \sigma \) is the Pauli spin matrix, \( m_0 \) is the electron mass, the third term is the
spin-orbit energy, and \( E \) is the energy of the electron. The solutions to this equation
are the Bloch functions

\[
\psi = e^{i\mathbf{k} \cdot \mathbf{r}} u_{nk}(\mathbf{r}), \tag{2.4}
\]

where \( u_{nk}(\mathbf{r}) \) is a function with the periodicity of the lattice, \( n \) is the band index,
and \( \mathbf{k} \) is restricted to the first Brillouin zone. Substituting the Bloch functions into
the Schrödinger equation gives

\[
\left\{ \frac{p^2}{2m_0} + V(\mathbf{r}) + \frac{\hbar}{4m_0^2c^2} (\sigma \times \nabla V) \cdot \mathbf{p} + \frac{\hbar^2 k^2}{2m_0} + \frac{\hbar \mathbf{k}}{m_0} \cdot (\mathbf{p} + \frac{\hbar}{4m_0c^2} \sigma \times \nabla V) \right\} u_{nk} = E_{nk} u_{nk}, \tag{2.5}
\]
The set of $u_{nk}(r)$ can be written in terms of the complete set of Bloch functions at $k=0$, $u_{n0}(r)$ [53]:

$$u_{nk} = \sum_m c_m(k) u_{m0}. \quad (2.6)$$

Substituting equation 2.6 into equation 2.5, multiplying by $u_{n0}^*(r)$ on both sides, and integrating over the unit cell gives:

$$\sum_m \left\{ \left( E_{n0} - E_{mk} + \frac{\hbar^2 k^2}{2m_0} \right) \delta_{nm} + \frac{\hbar k}{m_0} \cdot p_{nm} \right. + \frac{\hbar}{4m_0 c^2} \left( [\nabla V \times p] \cdot \sigma + [\nabla V \times k] \cdot \sigma \right)_{nm} \left. \right\} c_m(k) = 0 \quad (2.7)$$

where the matrix elements of an operator $O$ are given by

$$O_{nm} = \int_{\text{unit cell}} u_{n0}^*(r) O u_{m0}(r) dr. \quad (2.8)$$

This is the eigenvalue equation for the point $k$ written in the $k = k_0 = 0$ representation. The last term, the $k$-dependent spin-orbit coupling, is an order of magnitude or so smaller than the other spin-orbit term [40] and is neglected.

What we have done here is write out a new Hamiltonian with the same energies as the usual single-electron Schrödinger equation but different eigenfunctions: $u_{nk}(r)$ instead of $\psi$ [54]. One advantage of this approach is that the new Hamiltonian and its eigenfunctions have the same periodicity as the lattice. So, we can solve the eigenvalue equation over only one unit cell (e.g., 8 electrons) rather than the whole crystal ($\sim 10^{23}$ electrons). Another advantage of using the $u_{n0}(r)$ as the basis vectors is that the $k$-dependent part of the Hamiltonian is zero at $k=0$. Therefore, the $k$-dependent terms will be small at the wavevectors of interest, near $k=0$, and we can...
apply perturbation theory to solve for the energies and wavefunctions.

2.1.3 Hamiltonian Matrix in the Luttinger-Kohn Representation

So, what is a good set of basis states $u_{n0}(x)$? A popular choice is the Luttinger-Kohn representation [53] because the spin-orbit interaction is diagonal in this basis. This basis is also known as the $|J, m_J\rangle$ basis. $J$ is the total angular momentum and $m_J$ is the component of $J$ along $k$. The Luttinger-Kohn basis states are:

$$
\begin{align*}
|J, m_J\rangle & \quad \psi & \quad H(k = 0) \\
CB & S \uparrow & E_g \\
& S \downarrow & E_g \\
HH & |3/2, 3/2\rangle & \left( \frac{X+iY}{\sqrt{2}} \right) \uparrow & 0 \\
& |3/2, -3/2\rangle & -\frac{1}{\sqrt{2}}(X-iY) \downarrow & 0 \quad (2.9) \\
& |3/2, -3/2\rangle & -\frac{1}{\sqrt{6}}(X+iY) \uparrow -\sqrt{\frac{2}{3}}Z \downarrow & 0 \\
& |3/2, -1/2\rangle & -\frac{1}{\sqrt{6}}(X-iY) \uparrow -\sqrt{\frac{2}{3}}Z \downarrow & 0 \\
LH & |3/2, 1/2\rangle & -\sqrt{\frac{2}{3}}Z \uparrow +\frac{1}{\sqrt{6}}(X+iY) \downarrow & 0 \\
& |3/2, -1/2\rangle & -\frac{1}{\sqrt{6}}(X-iY) \uparrow -\sqrt{\frac{2}{3}}Z \downarrow & 0 \\
SO & |1/2, 1/2\rangle & \frac{1}{\sqrt{3}}(X+iY) \downarrow +\frac{1}{\sqrt{3}}Z \uparrow & -\Delta \\
& |1/2, -1/2\rangle & -\frac{1}{\sqrt{3}}Z \downarrow +\frac{1}{\sqrt{6}}(X-iY) \uparrow & -\Delta
\end{align*}
$$

where $|S\rangle$, $|X\rangle$, $|Y\rangle$, and $|Z\rangle$ are the lowest spherical harmonics: $Y_{00}$, $Y_{10}$, and $Y_{1\pm 1}$.

The spherical harmonics are closely related to the band wavefunctions because the conduction band edge has the symmetry of an $s$-like atomic state $(|S\rangle)$ and the HH, LH, and SO edges are $p$-like $(|X\rangle, |Y\rangle, and |Z\rangle)$. The $\uparrow$ and $\downarrow$ states in each band are degenerate. A major effect of spin-orbit interaction is to split the 6-fold degenerate $J=3/2$ $p$-like states into the 4-fold degenerate (at $k=0$) $m_J = 3/2$ states (LH and
HH) and the 2-fold degenerate $m_j = 1/2$ states (SO).

When we apply Löwdin perturbation theory to solve this problem, we take $h' = \frac{\hbar}{m_0} \mathbf{k} \cdot \mathbf{p}$. The resulting $8 \times 8$ Hamiltonian matrix includes momentum and spin orbit matrix elements between states in group $A$ to first order. Momentum matrix elements between states in groups $A$ and $B$ are included as a second order perturbation, which renormalizes the interactions between the group $A$ states. Spin-orbit interaction between groups $A$ and $B$ as well as all $\mathbf{k}$-dependent spin-orbit interactions are typically neglected. This $8 \times 8$ matrix can be solved numerically. However, in order to get analytical expressions, we will consider two approximations.

### 2.1.4 Small Gap Approximation

The small gap approximation is a two-step simplification of the full 8-band treatment. First, the interactions between group $A$ states only are solved exactly. Then, the interactions with remote bands are added by first-order perturbation. The first step leads to a block-diagonal Hamiltonian:

$$H = \begin{pmatrix} H & 0 \\ 0 & H \end{pmatrix}$$

$$\mathbf{H} = \begin{pmatrix} H & 0 & 0 & Pk \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -\Delta & \sqrt{2}\Delta/3 \\ Pk & 0 & \sqrt{2}\Delta/3 & -2\Delta/3 \end{pmatrix}$$ (2.10)
where the sequence of eigenvectors is \( \text{CB}, \text{HH}, \text{LH}, \text{SO} \), and the constants are:

\[
P = -\frac{i\hbar}{m} \langle S | p_z | X \rangle = -\frac{i\hbar}{m} \langle S | p_y | Y \rangle = -\frac{i\hbar}{m} \langle S | p_z | Z \rangle \tag{2.12}
\]

\[
\Delta = \frac{3\hbar}{4m^2c^2} \langle X | \frac{\partial V}{\partial x} p_y - \frac{\partial V}{\partial y} p_x | Y \rangle. \tag{2.13}
\]

In writing the Hamiltonian matrix, we have taken advantage of the bands’ symmetry properties, which reduce the number of terms significantly because many interactions are forbidden by symmetry. We have also simplified the expression by choosing \( \mathbf{k} \parallel \mathbf{J} \), which decouples the heavy hole and light particle (CB, LH, SO) states. We are free to do this because, as we will see, the dispersion is isotropic in this treatment.

Now we can see why this method is called “\( \mathbf{k} \cdot \mathbf{p} \)”: the \( \mathbf{k} \cdot \mathbf{p} \) term in the Hamiltonian is the one which mixes the bands. (The spin-orbit term also cause mixing, of course, but historically this term was added later.)

The momentum matrix elements can be calculated using the symmetrized wavefunctions at \( k = (0,0,0) \) in the nearly-free electron model for zincblende crystals [55]:

\[
S = \sqrt{8} \cos \frac{2\pi x}{a_0} \cos \frac{2\pi y}{a_0} \cos \frac{2\pi z}{a_0}
\]

\[
X = \sqrt{8} \sin \frac{2\pi x}{a_0} \cos \frac{2\pi y}{a_0} \cos \frac{2\pi z}{a_0}
\]

\[
Y = \sqrt{8} \cos \frac{2\pi x}{a_0} \sin \frac{2\pi y}{a_0} \cos \frac{2\pi z}{a_0}
\]

\[
Z = \sqrt{8} \cos \frac{2\pi x}{a_0} \cos \frac{2\pi y}{a_0} \sin \frac{2\pi z}{a_0}
\] \tag{2.14}

The resulting matrix element is:

\[
\langle S | p_z | X \rangle = \frac{2\pi\hbar}{a_0}. \tag{2.15}
\]
This matrix element is usually expressed in terms of energy as \( E_P = \frac{2m_0 P^2}{\hbar^2} \). \( E_P \) has a value of 19-29 eV in the semiconductors considered in this thesis.

The secular equation of this Hamiltonian is

\[
E' = 0
\]

\[
E'(E' - E_g)(E' + \Delta) - k^2 P^2(E' + 2\Delta/3) = 0
\]

where

\[
E' = E - \frac{\hbar^2 k^2}{2m_0}
\]

By expanding in small values of \( k \), this leads to the following parabolic dispersion relations (energy vs. wavevector):

\[
E_{CB}(k) = E_g + \frac{\hbar^2 k^2}{2m_0} + \frac{k^2 P^2(E_g + 2\Delta/3)}{E_g(E_g + \Delta)}
\]

\[
E_{HH}(k) = \frac{\hbar^2 k^2}{2m_0}
\]

\[
E_{LH}(k) = \frac{\hbar^2 k^2}{2m_0} - \frac{2k^2 P^2}{3E_g}
\]

\[
E_{SO}(k) = -\Delta + \frac{\hbar^2 k^2}{2m_0} \frac{k^2 P^2}{3(E_g + \Delta)}
\]

Note that the dispersion relations are isotropic in \( k \). This is a consequence of ignoring remote band effects. The effective masses at the band edge, \( k=0 \), are

\[
\frac{1}{m_{CB}^*} = \frac{1}{m_0} + \frac{4P^2}{3E_g} + \frac{2P^2}{3(E_g + \Delta)}
\]

\[
\frac{1}{m_{HH}^*} = \frac{1}{m_0}
\]

\[
\frac{1}{m_{LH}^*} = \frac{1}{m_0} - \frac{4P^2}{3E_g}
\]
\[
\frac{1}{m_{SO}} = \frac{1}{m_0} - \frac{2P^2}{3(E_g + \Delta)}
\]

$1/m_0$ is the free electron contribution to the mass. For light masses, such as InAs ($m_{CB}^* = 0.026m_0$), this term is not very important.

The terms in the effective mass expressions indicate the interactions between the bands [3]. For example, the conduction band contains contributions from both the light hole and split-off bands, since it interacts with both of them. The light hole band has a term which is equal in magnitude and opposite in sign to the CB's interaction with the LH. Most strikingly, the HH band is dispersionless. This is a consequence of the fact that it does not interact with the other three bands. When we add remote band interactions, we will find a more realistic HH dispersion.

The dispersion relations given by eq. 2.18 are nonparabolic, meaning that the effective mass is a function of energy or, equivalently, $k$. From eq. 2.16, the energy-dependent CB effective mass is:

\[
\frac{1}{m_{CB}^*} = \frac{2P^2}{3} \left( \frac{2}{E' + E_g} + \frac{1}{E' + E_g + \Delta} \right)
\]

(2.20)

This equation shows that in the conduction band, the effective mass increases with increasing energy. This result is plotted in figure 2.2 for GaAs, InAs, and AlSb, along with the parabolic approximation. It is clear from the figure that nonparabolicity is stronger for materials with a smaller band gap. When the conduction and valence bands are closer together in energy, there is more mixing between positive and negative mass bands. Nonparabolicity plays a significant role in all the experiments in this thesis.
Figure 2.2: Dispersion relations for GaAs, InAs, and InSb. The black lines include the energy-dependent effective mass (eq. 2.16), and the grey lines are the parabolic approximation (eq. 2.19). This figure clearly illustrates that nonparabolicity increases as the band gap decreases. The plots are extended beyond their range of validity ($k \sim \pm 0.05$ [3]) in order to show their overall behavior more clearly.

Another striking effect in narrow gap semiconductors is the spin-orbit coupling. For GaAs, inclusion of spin-orbit coupling in the $\mathbf{k} \cdot \mathbf{p}$ Hamiltonian leads to a small correction to the $P^2$ calculated without spin-orbit coupling. But for InAs, the correction is about 10%, and for InSb, it is 50% [54].

The second step in the small gap approximation is to use first order perturbation to add remote band interactions to eq. 2.19. The result for the conduction band is an isotropic effective mass [40], [32]:

$$\frac{1}{m_e^*} = \frac{1}{m_0} + \frac{2F}{m_0} + \frac{2F^2(E_g + 2\Delta/3)}{\hbar^2 E_g(E_g + \Delta)} \quad (2.21)$$

The Kane $F$ parameter describes interactions between the conduction band and
remote bands:

\[ F = \frac{\hbar^2}{m^2} \sum_j \frac{|\langle X | p_x | u_j \rangle|^2}{(E_{VB} - E_j)} + \frac{P^2}{E_{CB} - E_{VB}} \]  

(2.22)

The summation is over all states \( u_j \) which have the symmetry \( \Gamma_1 \). According to group theory, interactions with all other symmetry types are zero. The prime indicates that the summation does not include states in group \( A \). \( E_{CB} \) and \( E_{VB} \) are the energies of the conduction band edge and the average valence band edge in group \( A \). Remote band interactions with the conduction band will be important when we discuss the temperature dependence of the conduction band effective mass in section 3.4.2.

### 2.1.5 4-band Approximation

For hole energies which are small compared to the band gap and split-off energy, \( \Delta \), the \( 8 \times 8 \) L"owdin matrix can be approximated by a \( 4 \times 4 \) matrix for the heavy hole and light hole states. Dresselhaus et al. [46] showed that the eigenenergies in this case are

\[ E(k) = A k^2 \pm \sqrt{B^2 k^4 + C^2 (k_x^2 k_y^2 + k_y^2 k_z^2 + k_z^2 k_x^2)}^{1/2}, \]  

(2.23)

where the \( - \) is for heavy hole states and the \( + \) is for light hole states. The constants are:

\[
\begin{align*}
A &= \frac{1}{3} (L + 2M) + \frac{\hbar^2}{2m_0} \\
B &= \frac{1}{3} (L - M) \\
C^2 &= \frac{1}{3} [N^2 - (L - M)^2]
\end{align*}
\]  

(2.24)
where [51]

\[
\begin{align*}
L &= F + 2G \\
M &= H_1 + H_2 \\
N &= F - G + H_1 - H_2 \\
G &= \frac{\hbar^2}{2m^2} \sum_j^{\Gamma_{15}} \frac{|\langle X|p_j|u_j}\rangle|^2}{(E_{VB} - E_j)} \\
H_1 &= \frac{\hbar^2}{m^2} \sum_j^{\Gamma_{15}} \frac{|\langle X|p_j|u_j}\rangle|^2}{(E_{VB} - E_j)} \\
H_2 &= \frac{\hbar^2}{m^2} \sum_j^{\Gamma_{25}} \frac{|\langle X|p_j|u_j}\rangle|^2}{(E_{VB} - E_j)}
\end{align*}
\]

(2.25)

Again, group theory has been used to reduce the number of parameters.

The remote band matrix elements are often expressed in terms of the Luttinger parameters, \( \gamma \) [55]:

\[
\begin{align*}
(\hbar^2/2m_0)\gamma_1 &= -A \\
(\hbar^2/2m_0)\gamma_2 &= -B/2 \\
(\hbar^2/2m_0)\gamma_3 &= [(B^2/4) + (C^2/12)]^{1/2}
\end{align*}
\]

(2.26)

The Luttinger parameters are a mathematically succinct expression of the momentum matrix elements between states in group A and B in the Luttinger-Kohn basis.

Equation 2.23 can be differentiated to find the effective masses in different direc-
tions. The results are [32]:

\[
\left( \frac{m_0}{m_{HH}} \right)^{\gamma_1 - 2\gamma_2} = \left( \frac{m_0}{m_{HH}} \right)^{[110]} = \frac{1}{2}(2\gamma_1 - 2\gamma_2 - 3\gamma_3) \\
\left( \frac{m_0}{m_{HH}} \right)^{[111]} = \frac{1}{2}(2\gamma_1 - 2\gamma_2 - 3\gamma_3)
\]

(2.27)

\[
\left( \frac{m_0}{m_{\ell H}} \right)^{\gamma_1 + 2\gamma_2} = \left( \frac{m_0}{m_{\ell H}} \right)^{[110]} = \frac{1}{2}(2\gamma_1 + 2\gamma_2 + 3\gamma_3) \\
\left( \frac{m_0}{m_{\ell H}} \right)^{[111]} = \gamma_1 + 2\gamma_3
\]

(2.28)

Equations 2.23 and 2.27-2.28 describe hole mass anisotropy, or warping.

In the 4-band approximation, the Löwdin Hamiltonian for the split-off band can also be solved as a 2 \times 2 matrix. In this approach, the isotropic split-off mass is [55], [32]:

\[
\left( \frac{m_0}{m_{SO}} \right) = \gamma_1 - \frac{E_P\Delta_{SO}}{3E_g(E_g + \Delta_{SO})}
\]

(2.29)

2.1.6 Measurement of Key Parameters

One of the strengths of the k \cdot p method is that its calculations depend on only a few parameters, most of which are easy to measure from experiment. The band gap and split-off energies can be measured from optical absorption with excitonic corrections. Effective masses can be measured by cyclotron resonance with a correction for band filling and the nonresonant polaron energy [32]. The P parameter can be calculated
from eq. 2.19 or the more accurate eq. 2.21. Unfortunately both $P$ and $F$ depend on remote band effects, which cannot be measured directly. The Zeeman splitting factor $g^*$, however, depends less on remote band interactions than $m^*$ does. So, $g^*$ has been used in addition to the other quantities to compile consistent sets of band parameters for each III-V semiconductor [32].

One important question is, how good is the $\mathbf{k}\cdot\mathbf{p}$ method? The main limit is that the $\mathbf{k}\cdot\mathbf{p}$ method fails when the energy of the band becomes comparable to the energy gap with some other band which is not included in set $A$. In GaAs, for example, the $X$ point conduction band minimum is about 0.5 eV above the $\Gamma$ point conduction band minimum, so the 8-band $\mathbf{k}\cdot\mathbf{p}$ method is good only up to about 0.3 eV in the conduction band, or $k \sim 0.05$ [3]. This limit is a natural consequence of the limited number of bands upon which the calculation is based. However, this method accurate describes complicated effects, such as hole dispersion anti-crossings in superlattices, which occur within the range of validity.

### 2.2 Quantum Well Band Structure Calculations

In the rest of the chapter, we will discuss QW band structure calculation methods based on the envelope function approximation. After describing the envelope function approximation, we will consider three analytical applications of this approximation, in order of increasing complexity: the Kronig-Penney model, the Ben Daniel-Duke model, and the Kane model with confinement. These models apply only to flat band
systems, *i.e.* undoped wells. To include charge transfer effects, which modify the
QW potential and shift the energy levels, we must self-consistently solve for the
electrostatic potential and wavefunctions. Strain also affects the QW energy levels
and is included in the self-consistent calculations.

2.2.1 Envelope Function Approximation

One problem that is frequently encountered in semiconductor physics is to find the
effect of a perturbing potential, such as an impurity, a magnetic field, or a quantum
well potential, on a known band structure [40], [3]. If this perturbing potential varies
slowly on the scale of the crystal’s periodic potential, then the problem can be treated
in the envelope function (or effective mass) approximation. The main result of the
envelope function approximation is to subsume all of the atomic-scale details of the
system into effective parameters: the interband matrix elements, the band offsets (for
QWs), and the effective masses. These parameters define a Schrödinger equation for
the large-scale envelope function.

The envelope function approximation assumes that the band structure of the
unperturbed (*i.e.* bulk) crystal has already been solved. So, the energies and
wavefunctions are the solutions of

\[ H_0 \psi_{nk}(r) = E_n(k) \psi_{nk}(r). \]  

(2.30)

The dispersion relation for non-degenerate bands can be written

\[ E_n(k) = E_n(0) + \sum_{\alpha, \beta} \frac{\hbar^2}{2m^*} \left( \frac{1}{m^*} \right)_{\alpha, \beta} k_{\alpha \beta}; \quad \alpha, \beta = x, y, z \]  

(2.31)
where the elements of the effective mass tensor $m^*$ are given by

$$\frac{1}{m^*_\alpha\beta} = \frac{1}{m_0\delta_{\alpha\beta}} + \frac{2}{m_0^2} \sum_{m \neq n} \frac{\pi_{nm} \pi_{nm}}{E_{n0} - E_{m0}}. \tag{2.32}$$

$\pi$ is all the $k$ dependent parts of the unperturbed Hamiltonian $H_0$; the matrix elements $\pi_{nm}$ are given by eqs. 2.5 and 2.8.

The problem is to find the solution of

$$[H_0 + U(r)]\psi(r) = E\psi(r), \tag{2.33}$$

where $U(r)$ is the slowly varying potential. The wavefunction can be written as a product of a rapidly varying periodic function $u_{n0}(r)$ and a slowly varying envelope function $f$:

$$\psi_n = f_n u_{n0}(r). \tag{2.34}$$

Then, the envelope function $f$ is the solution of

$$\left[ E_{n0} + \sum_{\alpha,\beta} \frac{\hbar^2}{2m^*} \left( \frac{1}{m^*} \right)_{\alpha,\beta} \left( -i \frac{\partial}{\partial x_{\alpha}} \right) \left( -i \frac{\partial}{\partial x_{\beta}} \right) + U(r) \right] f(r) = Ef(r). \tag{2.35}$$

Here it is clear that all of the atomic-scale features of the problem have explicitly disappeared. They are included only implicitly, via the effective mass.

The band edge effective mass (that is, the effective mass at $k = 0$) is one of the most important semiconductor parameters because it wraps up the effects of the periodic crystal potential into a single number. So, an electron in an external field responds like a free particle with a mass of $m^*$. This is why holes are assigned negative mass: since holes have positive charge but negative mass, they can be treated just like electrons, which have negative charge and positive mass.
When the envelope function approximation is applied to quantum wells, several more assumptions are necessary besides that of the slowly varying perturbing potential \[3\]. We assume that the Bloch functions are the same in both the well, \(W\), and barrier, \(B\), layers:

\[ u_{nk}^{W}(\mathbf{r}) = u_{nk}^{B}(\mathbf{r}) \tag{2.36} \]

Then the total wavefunction is the sum of the wavefunctions in each material:

\[ \psi = \sum_{W,B} f_{[W,B]}^{[W,B]} u_{nk}^{W}(\mathbf{r}) \tag{2.37} \]

This approximation is fairly good as long as the energy stays close to the same \(k_0\) edge in each layer. (This approximation breaks down in, for example, narrow GaAs/AlGaAs \(011\) wells, where the \(\Gamma\) band edge becomes degenerate with the \(X\) band edge.) Likewise, we assume that \(k_\perp\) is the same in both layers, so that the envelope function is separable:

\[ f_{[W,B]}^{[W,B]}(r_\perp, z) \propto e^{ik_\perp r_\perp} f_{[W,B]}^{[W,B]}(z) \tag{2.38} \]

We also assume that the QW potential is abrupt and step-like. This assumption ignores interface effects, such as bond types and interface roughness, which in reality give rise to a more complicated potential. However, the abrupt potential assumption is justified because the wavefunction has a small overlap with the interfaces. And, although this assumption may appear to conflict with the original assumption of a slowly varying perturbing potential, in fact the overall QW potential is still much larger than the atomic spacing.
2.2.2 Kronig-Penney Model

The Kronig-Penney model is one of the simplest ways to calculate the band structure of a superlattice [40], [3]. It treats the one-dimensional periodic potential of the superlattice quantum mechanically. The parameters required by the Kronig-Penney model are the widths of the well and barrier, the barrier potential height, and the band edge effective masses in the well and barrier (see fig. 2.4).

The Kronig-Penny model is based on the propagation matrix approach. Each period $L$ of the superlattice has an envelope function which looks like:

$$f(z) = \begin{cases} A_0 e^{ikz} + B_0 e^{-ikz} & -w \leq z \leq 0 \\ C_1 e^{ik_1(z-b)} + D_1 e^{-ik_1(z-b)} & 0 \leq z \leq b \\ A_1 e^{ik(z-L)} + B_1 e^{-ik(z-L)} & b \leq z \leq L \end{cases}$$

(2.39)

The boundary conditions that the wavefunction must satisfy are continuity of:

$$f \text{ and } \frac{\partial f}{\partial z}.$$  

(2.40)
Using these boundary conditions, a propagation matrix can be defined for the transition from layer to layer. For a regular periodic structure such as a superlattice, this leads to the following determinantal equation \([3, 40]\):

\[
\cos(qL) = \cos(k_W w) \cosh(\kappa_B b) - \frac{1}{2} \left( -\xi + \frac{1}{\xi} \right) \sinh(k_W w) \sinh(\kappa_B b) \quad (2.41)
\]

where \(\xi = \frac{\kappa_B}{k_W}\)

\[
k_W = \sqrt{\frac{2m^*_W}{\hbar^2} (E + V_B)}
\]

\[
\kappa_B = \sqrt{\frac{-2m^*_B}{\hbar^2} E} = -ik_B
\]

\(q\) is the superlattice wavevector. Eq. 2.41 can be solved graphically using the condition \(\cos(qL) \leq 1\) to give the miniband energies. In the limit of independent QWs (infinitely thick or high barriers), eq. 2.41 reduces to:

\[
\cos(k_W w) - \frac{1}{2} \left( -\xi + \frac{1}{\xi} \right) \sin(k_W w) = 0. \quad (2.42)
\]

Fig. 2.4 shows eq. 2.41 plotted as a function of energy above the well’s conduction band edge for two situations. The left figure is calculated for a GaAs/AlGaAs superlattice with well widths of 10 nm, barrier widths of 5 nm, an effective mass of 0.067 \(m_0\), and a CB offset of 0.3 eV. The allowed minibands—those regions for which \(\cos(qL) \leq 1\)—are 10’s of meV wide. The right figure is calculated for an InAs/AlSb superlattice, characteristic of all of the InAs/AlSb QWs considered in this thesis. In this calculation, the well widths are 10 nm, the barrier widths are 10 nm, the effective mass is 0.026 \(m_0\), and the CB offset is 2 eV. In this case, the lowest few “minibands” are only 10’s of \(\mu\)eV wide, much narrower than the nonparabolicity-induced level
Figure 2.4: \( \cos(qL) \) vs. energy (eq. 2.41) for (left) GaAs/AlGaAs and (right) InAs/AlSb superlattices, as described in the text. The horizontal lines \( \cos(qL) = \pm 1 \) are also plotted. Note the different horizontal and vertical scales.

broadening or even the observed intersubband absorption linewidth (sec. 5.2.2). Because the wells are so deep, the QWs in this superlattice are effectively independent.

2.2.3 Ben Daniel-Duke Model

The Ben Daniel-Duke model takes the same basic approach as the Kronig-Penny model, but uses the Bastard boundary conditions [56], [3]:

\[
f \quad \text{and} \quad \frac{1}{m^*(z)} \frac{\partial f}{\partial z} \quad (2.43)
\]

are continuous at the QW interfaces. The second boundary condition conserves probability current between the well and the barrier. In other words, the Bastard boundary condition accounts for the fact that the electrons are moving in the periodic potential, not in vacuum. The solution for a single quantum well looks just like the
Kronig-Penny case (eq. 2.42), but with

$$\tilde{\xi} = \frac{\kappa_B m^*_W}{k_W m^*_B}$$  \hspace{1cm} (2.44)

The in-plane effective mass in the CB increases with increasing energy level, n.

For small $k_\perp$ and $m^*_W \sim m^*_B$, the $k_\perp$ dispersion is approximately parabolic. The $k_\perp$ effective mass, $m^*_n$, is given by

$$\frac{1}{m^*_n} = \frac{1}{m^*_W} [1 - P_B(E_n)] + \frac{1}{m^*_B} P_B(E_n),$$  \hspace{1cm} (2.45)

where $P_B(E_n)$ is the probability of finding the electron in the barrier when it is in the n state. For $m^*_W > m^*_B$, $m^*_n$ increases with increasing $n$ because as the effective barrier height decreases, the wavefunction penetration into the barrier increases. This effect adds to the mass increase with increasing energy due to nonparabolicity.

### 2.2.4 Kane Model with Confinement

The model of a superlattice can be extended from a single band, as in the previous section, to include multiple Kane-like bands. The simplest case is zero in-plane wavevector: $k_\perp = 0$. Then $m_J$ is a good quantum number because $k \parallel J \parallel z$, and the HH states are decoupled from the other three bands. (The two spin states in each band are degenerate.) This model neglects interactions with remote bands, so the dispersion relations are the same as eq. 2.16. The QW bound states are solutions of
eq. 2.42 but with an energy- and position-dependent effective mass included via [3]:

\[
\dot{\xi} = \frac{k_w m'^*_B(E)}{m'^*_W(E)} \frac{m'^*_B}{k_B} \tag{2.46}
\]

\[
\frac{1}{m^*(E, z)} = \frac{2p^2}{3} \left[ \frac{E + E_W - V_H(z)}{E + E_W + \Delta_W - V_{SO}(z)} + \frac{1}{E' + E_W + \frac{2\Delta_W}{3}} \right] \tag{2.47}
\]

\[
(E' - V_{CB})(E' - V_{CB} + E_B)(E' - V_{CB} + E_B + \Delta_B) = \hbar^2 k_B^2 \frac{p^2}{k_B} \left( E' - V_{CB} + E_B + \frac{2\Delta_B}{3} \right) \tag{2.48}
\]

Here, \( V_{CB}, V_H, \) and \( V_{SO}, \) are the band offsets at each band edge, \( E_W \) and \( E_B \) are the band gaps in the well and barrier, and \( \Delta_W \) and \( \Delta_B \) are the split-off energies in the well and barrier.

For \( k_\perp \neq 0, \) the band structure becomes more complicated because the heavy hole states couple with the light particle states. However, there is still an analytic solution to this four-band problem. The QW conduction band bound states are solutions of [57]:

\[
\cos(k_W w) + \frac{1}{2} \sin(k_W w) \left[ \frac{k_B m'^*_W}{m'^*_B} - \frac{k_W m'^*_B}{m'^*_W} \frac{m'^*_B}{k_B k_W} - \frac{k_\perp^2}{k_W k_B} \left( \frac{1}{m_W} - \frac{1}{m_B} \right)^2 \right] = 0 \tag{2.50}
\]

where

\[
k_W^2 = \frac{2m'^*_CB}{\hbar^2} \frac{E'(E' + E_W)(E' + E_W + \Delta_W)(E_W + 2\Delta_W/3)}{E_W(E_W + \Delta_W)(E' + E_W + 2\Delta_W/3)} - k_\perp^2 \tag{2.51}
\]

\[
k_B^2 = k_\perp^2 - \frac{2m'^*_CB}{\hbar^2} \frac{(E' - V_{CB})(E' - V_{CB} + E_B)(E' - V_{CB} + E_B + \Delta_B)(E_W + 2\Delta_W/3)}{E_W(E_W + \Delta_W)(E' - V_{CB} + E_B + 2\Delta_B/3)} \tag{2.52}
\]

\( m^* \) is the same as in eq. 2.47. This 4-band model will be used in section 3.4.1 as a simple way to calculate the Landau levels of a QW.
2.2.5 Doped QWs: Self-consistent Solution of the Quantum Well Potential

A realistic treatment of doped quantum wells must include the effects of band bending caused by the local charge distribution of the well [40]. For example, in a well which is modulation-doped with donors a few nm into the barrier, the positive charges of the ionized donors plus the negative charges of the electrons in the well combine to produce an electric field pointing into the well on both sides. This field causes the center of the well to curve up relative to the built-in (no charges) potential. The band bending in the samples considered in this thesis is on the order of 10’s of meV.

Mathematically, the QW potential can be found by a self-consistent perturbative approach. Band bending can be treated in the effective mass theory (eq. 2.35), where the QW’s potential is, for electrons:

\[ U(z) = V_{\text{built-in}}(z) + |e|V_{\text{appl}} z + V_H(z). \]  

(2.53)

Here \( V_{\text{built-in}} \) is the potential profile of the QW before charge transfer, \( V_H(z) = -|e|\phi(z) \) is the Hartree potential (sec. 5.4.3), and \( V_{\text{appl}} \) is an applied electric field which we will set to zero. The electrostatic potential \( \phi \) must satisfy Poisson’s equation:

\[ \varepsilon \nabla^2 \phi = -c(z), \]  

(2.54)

where \( \varepsilon \) is the dielectric constant and the charge distribution \( c(z) = -|e|[-\rho^A(z) + \rho^e(z) + \rho^A(z) - \rho^D(z)] \) is the sum of contributions from free electrons and holes.
and ionized donors and acceptors. For an n-type well, the 2D electron density, $\rho^e$, depends on the envelope function as:

$$\rho^e(z) = \sum_n |f_n(z)|^2 \rho^e_n$$  \hspace{1cm} (2.55)

where the summation is over the occupied subbands. The 2D electron density in the $n^{th}$ subband, $\rho^e_n$, is:

$$\rho^e_n = \frac{2}{\text{area}} \sum_{\mathbf{k}_\perp} \frac{1}{1 + e^{(E_n(\mathbf{k}_\perp) - E_F)/k_B T}}$$  \hspace{1cm} (2.56)

$$= \frac{k_B T m^*_{CB}}{\pi \hbar^2} \ln(1 + e^{(E_F - E_n(0))/k_B T})$$

Here $k_B$ is the Boltzmann constant, and the dispersion relation of $k_\perp$ is parabolic. The Fermi level is set by the overall charge neutrality condition.

Now, we solve the coupled Schrödinger and Poisson equations for the envelope wavefunction $f(z)$ and the electrostatic potential $\phi(z)$ self-consistently. Figure 2.5 shows the calculated envelope functions and self-consistent potential for an InAs/AlSb QW [58]. The well width is 10 nm, the electron density is $1.84 \times 10^{12}$ cm$^{-2}$, and the temperature is 0 K. Only the lowest subband is populated. The well potential is bent up by 40 meV due to charge transfer.

### 2.2.6 Strained QWs: the Bir-Pikus Hamiltonian

Like band bending, strain also causes energy level shifts in quantum wells. Even when no external strain is applied to the QW sample, strain can arise from lattice mismatch between well and barrier materials. The lattice mismatch in narrow gap
Figure 2.5 : Left: calculated self-consistent potential of a 10 nm InAs/AlSb QW with $n = 1.84 \times 10^{12}$ cm$^{-2}$. Right: lowest few conduction band wavefunctions.
quantum wells is non-negligible: 1.3% between InAs and AlSb, 0.5% between InSb and Al_{0.96}In_{0.91}Sb. The enormous (\sim 8\%) lattice mismatch between the narrow gap active region and the GaAs substrate (for most of the samples in this thesis) is assumed to be compensated by the buffer layers.

Strain affects QWs in complicated ways. For example, InSb wells with AlInSb barriers have biaxial compressive strain, which is compressive in the QW plane and tensile along the growth axis. This strain increases the band gap of InSb and shifts the light hole energy down. In contrast, the InAs/AlSb QWs considered in this thesis are under biaxial tensile strain due to pseudomorphic growth on AlSb. In the calculation shown in fig. 2.5, strain raises the light hole band above the heavy hole band by 100 meV. Strain also reduces the band gap from 417 meV to 190 meV.

Mathematically, strain is incorporated in the bandstructure via the 8-band Bir-Pikus Hamiltonian. The strain Hamiltonian consists of terms that are anisotropic in $k^2$, the coefficients of which are the Luttinger coefficients, and terms that are functions of various components of the strain tensor. [40]
Chapter 3

CW Cyclotron Resonance

As we have seen in chapter 2, the band edge effective mass, $m^*_0$, is one of the most important parameters for making predictions about the optical properties of semiconductor heterostructures. But how do we measure the band edge mass? The most straightforward and accurate method is cyclotron resonance (CR). We detect the resonance by applying a magnetic field and measuring the transmission of light through the sample while we vary the wavelength of the incident light or the magnetic field. The cyclotron mass is given by the position of the resonance according to $m^*_CR = eB/\omega_c$, where $\omega_c$ is the frequency of the light and $B$ is the magnetic field. In a parabolic semiconductor, the band edge effective mass is found by subtracting the nonresonant polaron enhancement (a few percent) from the cyclotron mass [59] (sec. 3.2). So, the band edge mass is also known as the bare mass. In a nonparabolic semiconductor, the cyclotron mass increases with increasing energy. Since free-carrier CR exists only at finite carrier densities, the band edge mass must be found from a bandstructure calculation incorporating the measured cyclotron mass and density.

One poorly understood aspect of the effective mass is its dependence on temperature. Naïvely, we expect that the temperature dependence of the mass will be dominated by the temperature dependence of the band gap. The band gap energy decreases with increasing temperature due to lattice dilation and phonon interac-
tion [60, 61]. The temperature dependence of the band gap is described well by the empirical Varshni formula [32],

\[ E_g(T) = E_g(0) - \frac{\alpha T^2}{T + \beta}, \]

(3.1)

where \( \alpha \) and \( \beta \) are material-dependent parameters. We expect that the conduction band mass should decrease with temperature as well due to increased coupling between the conduction and valence bands (eq. 2.21). However, cyclotron resonance measurements of the effective mass do not confirm this expectation. Several theoretical explanations have been proposed.

Attempts to explain the temperature dependence of the effective mass raise fundamental questions about the validity of \( \mathbf{k} \cdot \mathbf{p} \) theory at finite temperature. In this chapter, we present a CR measurement and a theory of the temperature dependence of the effective mass (\( m^*(T) \)) in InAs/AlSb QWs. The first section describes Landau levels; the second section reviews the varied experimental and theoretical results for \( m^*(T) \) in the literature. Next comes a description of our CR experiment, which was the first spectroscopic application of a THz quantum cascade laser. Finally, we present our collaborators' theoretical modeling, which attributes the observed \( m^*(T) \) to a combination of effects including polaron enhancement and temperature dependence of the \( F \) parameter.
3.1 Origin of Landau Levels

When a semiconductor is placed in a strong magnetic field, the energies of free carriers become quantized into Landau levels. In a semi-classical picture, electrons spiral around the magnetic field at the cyclotron frequency

\[ \omega_c = \frac{eB}{m^*}. \tag{3.2} \]

The magnitude of the cyclotron frequency can be understood by equating the centripetal acceleration of the electron to the Lorentz force: \( m^* \omega_c^2 r = e \omega_c r B \), where \( r \) is the radius of the motion. What is meant by a “strong” magnetic field? It is one for which a electron can complete at least one period of its orbit before being scattered. That is, \( \omega_c \geq 1/\tau \), where the momentum scattering time \( \tau \) is related to the mobility \( \mu \) by \( \mu = e\tau/m^* \). (Note that the momentum scattering time is in general different from the quantum lifetime [62] or the CR lifetime.) The energy of an electron in a strong magnetic field is quantized; its stationary states can be thought of as those for which the circumference of its orbit is composed of an integer number of wavelengths. This condition leads to the energy levels \( E = n\hbar \omega_c \) for \( n = 1, 2, 3, \ldots \). In fact, the following quantum mechanical treatment will show that the correct orbital energy spacing is

\[ E = \left( n + \frac{1}{2} \right) \hbar \omega_c \text{ for } n = 0, 1, 2, \ldots \tag{3.3} \]

These states are the Landau levels. (Here \( n \) is the Landau level index, not the band index as in chapter 2.)
Now we will treat the Landau level structure of a QW quantum mechanically [3]. Consider a QW in a magnetic field $\mathbf{B}$ along the growth direction, $\hat{z}$. The Lorentz force on a charge is $q \mathbf{v} \times \mathbf{B}$. Choose the gauge $\mathbf{A} = [0, Bx, 0]$. Then, the conduction band Landau Hamiltonian is [55]:

$$H = \frac{1}{2m_C^*} \left[ p_y + eA_y \right]^2 + \frac{p_z^2}{2m_C^*} + U(z) + g^* \mu_B \sigma \cdot \mathbf{B}$$

(3.4)

$$= \frac{1}{2m_C^*} \left[ -i\hbar \frac{\partial}{\partial y} + eBx \right]^2 + \frac{p_z^2}{2m_C^*} + U(z) + g^* \mu_B \sigma \cdot \mathbf{B}.$$  

Here $U(z)$ is the QW confinement potential (eq. 2.53). $g^*$ is the effective $g$-factor, which equals 2 for electrons in a vacuum. However, in a semiconductor, spin-orbit coupling can change the $g$-factor significantly, leading to wide variation between materials. For example, $g^* \sim 0.44$ in GaAs and -50 in InSb [55].

The wavefunction of the Landau Hamiltonian is separable into orbital and spin parts. The eigenvalues of the spin part are $\pm \frac{1}{2} g^* \mu_B B$. The orbital wavefunction is separable into three parts:

$$\psi = \frac{e^{i(k_y y)}}{\sqrt{L_y}} f(z) \phi(x)$$

(3.5)

The $z$ part is the solution of

$$H_z f_m(z) = \left[ \frac{p_z^2}{2m_C^*} + V_{QW}(z) \right] f(z) = E_m f_m(z).$$

(3.6)

The solutions of this equation are just the eigenvalues and envelope functions of the QW at $B = 0$. The $x$ part of the orbital wavefunction is the solution of

$$H_x \phi_n(x) = \left\{ \frac{p_x^2}{2m_C^*} + \frac{1}{2m_C^*} \left[ \hbar k_y + cBx \right]^2 \right\} \phi_n(x) = E_n \phi_n(x)$$

(3.7)
This is the harmonic oscillator equation with frequency $\omega_c = eB/m^*$ and center position $x_0 = -\hbar c/eB$. The eigenfunctions are the Hermite polynomials, and the eigenvalues are given by eq. 3.3. Note that the Landau level energy is independent of $k_y$.

The total energy of the Landau-quantized system is

$$E = E_m + \left( n + \frac{1}{2} \right) \hbar \omega_c \pm \frac{1}{2} g^* \mu_B B \quad (3.8)$$

So, each QW state $E_m$ has associated with it a ladder of Landau levels separated by energy $\hbar \omega_c$, and each Landau level is spin-split. Each of the terms in eq. 3.8 is energy dependent due to nonparabolicity. However, for a 10 nm InAs QW such as the one in sec. 3.3, approximate values are $E_m = 200$ meV, $\hbar \omega_c = 20$ meV at 7 T, and $g \mu_B B = 5$ meV at 7 T.

The degeneracy of the orbital states of a Landau level can be found from the condition that the center of the harmonic oscillator must be inside the sample: $0 < x_0 < L_x$, or $0 < k_y < (eB/h)L_x$. Since the density of $k_y$ values is $L_y/2\pi$, the degeneracy per area of each Landau level is $eB/h$. In a two-dimensional system, the degeneracy leads immediately to the definition of the filling factor $\nu = neB/h$, where $n$ is the areal density of carriers. The filling factor describes the fraction of Landau levels that are full; for example, a filling factor of 3.5, as shown in fig. 3.1, indicates that the lowest three Landau levels are completely filled and the fourth is half-full.

Landau quantization leads to qualitatively different behavior for two-dimensional
and three-dimensional systems [63]. A QW in a strong magnetic field is effectively zero-dimensional: carriers are quantized in the $z$ direction by the QW potential and in $x$ and $y$ by the magnetic field. The result for the density of states is a series of spikes, as shown in figure 3.1. Each spike has a degeneracy of $eB/h$ (neglecting spin-splitting). On the other hand, in a bulk system in a strong magnetic field, carriers’ $z$ motion is not confined. The density of states versus Fermi energy has sharp features that correspond to adding another Landau level, but unlike the QW case, the density of states is never zero except for $E_F < \frac{1}{2} \hbar \omega_c$.

### 3.2 Review of the Temperature Dependence of the Effective Mass

In this section, we will describe all of the factors which can affect the temperature dependence of the effective mass. Then, we will review previous experiments on
$m^*(T)$ in the conduction band of III-V semiconductors.

The temperature dependence of the effective mass is most strongly influenced by
the temperature dependence of the band structure. With increasing temperature,
the band gap decreases. A smaller band gap means more interband interaction and
so a lighter conduction band mass. In strongly nonparabolic semiconductors, this
trend is counteracted by thermal effects on the carrier distribution. With increasing
temperature, higher Landau levels are thermally populated. The carrier density
may also increase with increasing temperature, as in our InAs/AlSb QWs, further
contributing to the population of higher Landau levels. Due to nonparabolicity,
higher Landau levels have higher masses. So, thermal effects on the population in a
nonparabolic semiconductor act to increase the CR mass.

The polaron effect renormalizes the effective mass, making it a few percent heavier
than the bare band edge mass in III-V semiconductors. Polaron interaction, between
electrons and phonons, arises in polar crystals due to the macroscopic electric fields
created by long-wavelength phonons. Polaron interaction is characterized by the
Fröhlich coupling constant, which is 0.07 in bulk GaAs and 0.02 in bulk InSb [64]. The
temperature dependence of the polaron effect arises due to the strong temperature
dependence of the phonon population. The polaron mass enhancement peaks at
intermediate temperature, for example, around 80 K in a GaAs heterojunction with
$n = 5 \times 10^{11} \text{ cm}^{-2}$ [65].

Localization due to interface fluctuations has been shown to decrease the mass
in very low density ($\leq 10^8 \text{ cm}^{-2}$) QWs [66]. This effect weakens with increasing temperature as electrons gain enough thermal energy to overcome the localization potentials. This result is in contrast to the usual expectation for an impurity potential, which leads to a finite CR frequency at zero magnetic field and higher mass.

Finally, electron-electron interactions become important when carriers of different masses are present [41]. Coulomb interaction renormalizes the Landau levels, coupling spin-split levels into a single resonance for most conditions in GaAs [67–70] and reducing the spin splitting in InAs QWs near $\nu = 1$ [71,72].

There are not very many measurements of $m^*(T)$, and they cover different regimes of density, magnetic field, dimensionality, and material. In bulk GaAs, $m^*(T)$ is constant to within 1% from 4 K to room temperature [73,74]. Batke et al. [73] measured $m^*(T)$ at 3.9, 6.5, and 13.6 T. At the higher two fields, they were able to resolve and assign transitions from different Landau levels, and each transition had a constant CR mass. The low field data was fit well with a model that included the thermal population of different levels and constant CR mass. Batke et al. treated the CR mass in a two band model:

$$m_{CR}^* = m^* + \frac{2\hbar e}{E_g^*} B(N + 1) \quad (3.9)$$

where $m^*$ is the bare effective mass (nearly equal to the polaron mass), $N$ is the Landau index, and $E_g^*$ is an effective band gap which describes nonparabolicity [75]. However, they were unable to account for the observed CR mass with this model, using the measured temperature dependence of the magnetopolaron mass [76] and
the temperature dependence of the band gap. They speculated that the discrepancy arises due to an unaccounted for increase in the polaron effect with temperature.

A series of studies on GaAs heterojunctions with \( n \sim 10^{11} \text{ cm}^{-2} \) at \( \nu \sim 1 \) have shown that electron-electron interaction is important in this regime. Brummel et al. [74] observed a drop in the CR mass of 2% from around 80 K to 4 K. They attributed this drop to screening of the polaron enhancement at low temperature due to thermal narrowing of the Landau levels and increase in the density of states. However, subsequent calculations of the dynamic screening by Xiaoguang et al. [65] showed that this mechanism cannot explain the observed size of \( m^*(T) \). Hu et al. performed the measurement again in a gated heterostructure at slightly lower densities (3 & 7 \( \times 10^{10} \text{ cm}^{-2} \)). They also observed a drop in the effective mass with decreasing temperature, and the drop was more pronounced at the higher density. They were able to explain this observation in terms of Coulomb interaction between partially filled Landau levels.

Measurements of \( m^*(T) \) in InAs heterostructures have until now been restricted to ultra-high pulsed magnetic fields (\( \sim 100 \text{ T} \)). The advantage of going to high magnetic field is that CR can be observed even in low mobility samples, such as InAs/GaSb superlattices [77]. High magnetic fields also clearly resolve the spin-splitting in high mobility InAs/AlSb QWs (2.6 \( \times 10^5 \text{ cm}^2/\text{Vs} \)), enabling Arimoto et al.'s [71] interesting observation of hysteresis in the Landau level population with magnetic field sweep direction. Both of these high field experiments were in the
regime $\nu \sim 1$, where electron-electron interactions are predicted to renormalize the Landau level energies [69, 72]. Indeed, both experiments showed a reduction in the splitting between the $(0, \uparrow -1, \uparrow)$ and $(0, \downarrow -1, \downarrow)$ transitions at around 100 K. The sizes of the reductions were consistent with the renormalization effect. However, this many-body interaction complicates the interpretation of the temperature dependence of the band structure in these experiments.

3.3 Experiment and Results

In order to elucidate the temperature dependence of the band structure, we measured cyclotron resonance in InAs/AlSb QWs [78]. This heterostructure is well-suited to cyclotron resonance measurements over a wide temperature range because it maintains a high carrier density at 4 K (see sec. 5.1.1) and a moderate mobility at room temperature.

The sample was grown by molecular beam epitaxy at Osaka Institute of Technology (number 0209042T). It consists of 20 periods of (10 nm/10 nm) InAs/AlSb quantum wells, with a total electron density of $1.5 \times 10^{12}$ cm$^{-2}$ and a mobility of 73,000 cm$^2$/Vs at 4 K. Shubnikov-de Haas measurements of this sample give a density of $9.7 \times 10^{11}$ cm$^{-2}$, indicating that about 60% of the electrons are located in a single well (see sec. 5.1.1). The temperature dependence of the total density and mobility are shown in fig. 3.2.

The light sources were GaAs/AlGaAs quantum cascade lasers (QCLs) operating
Figure 3.2: Temperature dependence of the Hall density and mobility of a 10 nm
InAs/AlSb multiple QW sample (0209042T).

at 4.7 THz (64 μm), 3.5 THz (86 μm), and 2.3 THz (127 μm) with a maximum
continuous wave power of about 4 mW [1]. The lasers were operated at 135 Hz
with a duty cycle of 25%. The light was collimated and focused onto the sample by
parabolic mirrors. The sample was placed in a superconducting magnet with cold
and room temperature z-cut quartz windows (f/2.4), and the transmitted light was
collected with a parabolic mirror and detected with a liquid-helium-cooled silicon
bolometer. The entire beam path was purged with dry nitrogen. The short-term
wavelength drift was \~20 MHz over 30 s, measured by beating with a THz gas
laser [79]. Any long-term drift was unnoticeably small during the CR measurement.
The short-term intensity fluctuations were \~0.5% over 1 s; long-term drift was
dominated by humidity fluctuations in the beam path.
Figure 3.3 shows the transmission of the sample at 4.7 THz as a function of magnetic field from 1.5 to 300 K. Figure 3.4(a) shows the frequency dependence of the CR at 1.5 K, which corresponds to a cyclotron mass of $m_{eR}^c = 0.042m_0$. Figure 3.4(b) shows the temperature dependence of the cyclotron mass in fig. 3.3.

The cyclotron mass in fig. 3.4(b) increases with increasing temperature. This behavior is the opposite of that expected from the band gap decrease with increasing temperature. Many of the effects discussed in sec. 3.2 are potential culprits. Only interface fluctuation localization can be ruled out, since the lateral size of the interface fluctuations in this sample, estimated to be 50 Å from intersubband absorption measurements [80], is smaller than the cyclotron radius, $l = \sqrt{\hbar/eB} = 240$ Å.

The observed CR lineshape is asymmetric, particular at low temperature. One mechanism which could give rise to asymmetry in this nonparabolic system is cyclotron resonance from multiple Landau levels with different populations (see sec. section: modelTRCR for a full discussion of this issue). However, the lineshape of the observed resonance is fit better by a single Lorentzian (one CR transition) than by two, so it is not clear that this mechanism is active.

3.4 Analysis

3.4.1 Bastard Model for $m^*(T)$

In order to qualitatively explore the effects of band structure and nonparabolicity on $m^*(T)$, we performed a simple calculation of the Landau level structure including
Figure 3.3: Transmission as a function of magnetic field at several temperatures for InAs/AlSb quantum wells (sample 0209042T). The quantum cascade laser wavelength is 64 μm (4.7 THz).
Figure 3.4: (a) Resonance field as a function of photon frequency at 1.5 K for InAs/AlSb quantum wells (sample 0209042T). The straight line has a slope of 1.51 T/THz, corresponding to an effective mass of 0.042 $m_0$. (b) Cyclotron mass versus temperature at 4.7 THz for the InAs/AlSb quantum wells. The experimental mass (triangles) increases with increasing temperature, whereas the theoretical masses (open and filled circles) for two possible CR transitions show the opposite behavior.
dispersion using the 4-band model in sec. 2.2.4 [57]. A magnetic field applied in the
growth direction quantizes the transverse wavevector according to [81]:

\[
\frac{\hbar^2 k_\perp^2}{2m^*} = \left( n + \frac{1}{2} \right) \hbar \omega_c
\]  

(3.10)

so

\[
k_\perp^2 = (2n + 1)eB/\hbar.
\]  

(3.11)

Then for any combination of magnetic field and Landau index, the allowed QW
energy levels can be calculated using eq. (2.50)-(2.52). The energy dependence of
the effective g-factor is not automatically included in this model; it was found using
a separate Pidgeon-Brown calculation (sec. 4.3). Furthermore, this model neglects
band bending in these doped QWs, which shifts the zero-field QW levels up in energy
by several 10's of meV. However, this shift has only a small effect on the cyclotron
mass. Finally, temperature was incorporated via the temperature dependence of \( E_g \),
which is described by the empirical Varshni formula (eq. 3.1).

The results of this simple calculation are shown in fig. 3.4(b). The low temper-
ature density indicates that the observed CR is predominantly the \((2, \uparrow) \rightarrow (3, \uparrow)\)
transition at low temperature. Here \((n, s)\) labels the \( n \) Landau level and the \( s \) spin
state. At higher temperatures, the density and thermal excitation increase, so the
CR shifts to include other transitions such as \((2, \downarrow) \rightarrow (3, \uparrow)\). Although higher levels
have higher masses, every level's mass decreases strongly with increasing tempera-
ture. Mass increase due to nonparabolicity alone is unable to qualitatively explain
our results. In the next section we describe a much more sophisticated theory.
3.4.2 Detailed Model for $m^*(T)$

Our collaborators Sanders and Stanton [82] are modeling the temperature dependence of the cyclotron resonance mass by allowing the $F$ parameter (eq. (2.22)) to vary with temperature. They employ a generalized 8-band Pidgeon-Brown model [83] that includes the QW confinement potential and the wavefunction's $k_z$ dependence [84]. The model includes strain and uses position dependent material parameters, which are important considerations for these QWs since the lattice mismatch between InAs and AlSb is not negligible (1.3 %). The carrier distribution is calculated using the Fermi-Dirac distribution, and the CR spectra are calculated using Fermi's golden rule.

In the conventional $k \cdot p$ model, the band gap, lattice constant, and carrier density are assumed to vary with temperature; the interband matrix elements and band offsets are assumed to be constant. Since the conventional treatment does not fit the observed temperature dependence of the CR effective mass, this model takes $F$ as a new temperature dependent parameter. $E_g$ varies with temperature according to the empirical Varshni formula (eq. (3.1)). The polaron enhancement of the mass is about 1%. The temperature dependence of the polaron enhancement, however, is already included in the Varshni formula, so it is not treated separately. The lattice constant, and consequently the strain, varies according to:

$$a_0(\text{AlGaSb}) = 6.1157 + 3.66 \times 10^{-5}(T - 300), \quad (3.12)$$

where the lattice constant is in Å and the temperature is in K. The QWs are assumed
to be pseudomorphically strained with the AlGaSb buffer. The electron density was measured throughout the temperature range of interest by Hall measurements on a different section of the sample.

In order to fit the observed temperature variation of the effective masses, the $F$ parameter is allowed to vary with temperature. The momentum matrix element $P$, the split-off energy $\Delta$, and the band offsets remain temperature independent, as in the usual $\mathbf{k}\cdot\mathbf{p}$ model. The $F$ parameter is expected to be temperature dependent because of the temperature-dependent energy term in its denominator (eq. 2.22). It is the easiest parameter to treat mathematically, since it appears in the Hamiltonian only in the CB-CB Landau terms.

Figure 3.5 shows the preliminary results of this model. The model cannot reproduce the observed increase of the effective mass with temperature. This is probably because the model does not actually incorporate the density increase yet. Adding this effect should increase the mass at higher temperature, leading to the right trend.
Figure 3.5: Experimental and calculated $m^*$ vs. $T$. 
Chapter 4

Time-Resolved Cyclotron Resonance

As we saw in the previous chapter, two-dimensional electron systems can provide a unique spectroscopic environment that is not achievable in atomic or molecular systems. Landau and Zeeman quantization lead to an externally tunable, nearly-equally-spaced energy ladder. Various orbital, spin and combined resonances have been spectroscopically investigated in Landau-quantized systems [85], especially in 2D, providing significant insight into the low-energy quantum dynamics of charge and spin carriers in the frequency domain.

However, there has been very limited success in directly probing carrier dynamics in these ladder systems in the time domain [86–88] with ultrashort and/or intense far-infrared (FIR) pulses. In particular, to our knowledge, there have been no studies on spin relaxation in a 2D system in high magnetic fields, although the corresponding 3D situation has been studied theoretically [89]. Understanding different spin relaxation mechanisms in confined structures in the presence of spin-orbit coupling is crucial for developing novel spin-based devices [90].

In this chapter, we report results of a picosecond time-resolved cyclotron resonance (TRCR) study of photogenerated electrons in undoped InSb QWs [91]. TRCR is a recently developed spectroscopic method for studying intraband carrier relaxation in a magnetic field [87]. An ultrashort near-infrared (NIR) pulse creates tran-
Figure 4.1: Schematic of the TRCR measurement, showing the interband pump and inter-Landau-level probe (not to scale).

Persistent carriers, and the subsequent dynamics are probed by a delayed FIR pulse, as shown in fig. 4.1. This technique allows us to monitor the effective mass of relaxing carriers as a function of time. InSb is an ideal material for TRCR studies. Its small effective mass and large $g$-factor result in large Landau and Zeeman splittings; its strongly energy-dependent mass and $g$-factor provide a natural mechanism for CR to expose the electron energy distribution, which evolves in time; and its strong spin-orbit coupling makes it an ideal system in which to compare different spin-relaxation mechanisms. Our data unambiguously demonstrated that the average electron cyclotron mass decreases as the electrons relax towards the band edge, as expected from band structure considerations.

In this chapter, we first ask, what can TRCR reveal about electron relaxation
which other techniques can not? Then, we present our experimental results and simulation. Finally, we discuss possible further applications of TRCR.

4.1 Review of Relaxation Measurements in Landau-Quantized Systems

Many techniques exist to study electron relaxation in Landau-quantized semiconductors. A simple, "brute force" method is CR saturation spectroscopy, which yields Landau level lifetimes on the order of 10 ps in an InAs/GaSb QW [92] or 10s of ps in an InAs/AlGaSb QW [93]. A more direct measure of the electron relaxation time is FIR degenerate pump-probe spectroscopy. This technique gives a Landau level lifetime of about 50 ps in InAs/AlSb QWs [88]. Both of these techniques are sensitive to magnetophonon resonances, or the phonon bottleneck. The inter-Landau level relaxation rate is enhanced by about an order of magnitude whenever \( \omega_{CR} = n\omega_{LO} \) [67] or \( \omega_{LO} = n\omega_{CR} \).

Kohn's theorem states that many-body effects are important in Landau-quantized systems containing carriers with different masses, for example both electrons and holes. Chemla et al. have performed a series of elegant studies of magnetoeexciton dynamics using degenerate four-wave mixing [94, 95]. They find, indeed, that Coulomb correlations between the magnetoeexcitons and with the electron collective modes dominate the observed dynamics.

Several other techniques can also measure the spin relaxation time in Landau
quantized systems. Spin relaxation time is the relaxation time between Zeeman-split levels, which depends on the spin-flip time. All of these methods are based on creating a non-equilibrium spin population and monitoring its decay. In spin-flip Raman spectroscopy, the upper spin state of the lowest Landau level is populated by a CO₂ laser pulse. The spin relaxation time in bulk InSb is 50-130 ns, probed by a second optical pulse [96, 97]. Another way to create a non-equilibrium spin population is by sweeping the magnetic field very quickly. Using this technique, a spin relaxation rate of 1-0.2 μs has been deduced in InAs/AlSb QWs [71]. Finally, circularly polarized interband pumping has been used to create a non-equilibrium spin population. Polarization-dependent photoluminescence is a sensitive measure of the decay, giving in GaAs QWs a spin relaxation time of 5 ns [98].

TRCR is unique because it is two-color, while previous time-resolved [88] and nonlinear [92, 93] FIR work on Landau levels in 2D systems was single-color. Because CR occurs only across the Fermi energy, the samples in the previous studies had to have carriers present at low temperature. Depending on where the Fermi energy lies, there are only one or two spin-split transitions which can be excited and probed by one color, even in the highly nonparabolic systems studied. This restriction limits the range of excited carrier temperatures which can be created and monitored, and does not allow a thorough study of the carrier distribution. In contrast, the two-color nature of our experiment allows us to create much hotter carriers and then probe their distribution over all magnetically accessible energy levels. Two-color TRCR
gives much more information about the time evolution of the Fermi distribution in these Landau-quantized multi-level systems.

In addition, two-color TRCR can be applied to samples in which there are no carriers present at low temperature. Therefore this technique is uniquely positioned to study the spin relaxation dynamics between the lowest two energy states. These states are (1,0,-1/2) and (1,0,+1/2); the levels are labelled by \( n, N, s \), where \( n = 1, 2 \) (subband index), \( N = 0, 1, 2, \ldots \) (Landau index), and \( s = \pm 1/2 \) (spin index). A very slow spin-flip relaxation is observed in a closely related system, semiconductor quantum dots (QDs). Due to confinement, the QD dispersion relation is quantized in all directions, so that sense, a QD is just like a 2DEG in a perpendicular magnetic field. Very long spin relaxation times have been observed in QDs, for example, more than 20 ms in a single GaInAs dot [99]. Long spin relaxation lifetimes are expected in QDs because the spin relaxation mechanisms which act in two or three dimensions are all restricted by the lack of dispersion in a QD. In fact, the spin relaxation lifetimes in quantum dots are significantly longer than the interband recombination times [100]. This long spin lifetime makes QDs an attractive tool for quantum computing applications, an active area of research. Elucidating spin relaxation mechanisms in QDs is one motivation for measuring TRCR.
4.2 Experiment and Results

The sample was an undoped multiple QW structure, containing 25 periods of 35 nm InSb wells separated by 50 nm thick Al$_{0.08}$In$_{0.92}$Sb barriers, grown by molecular beam epitaxy on a semi-insulating (001) GaAs substrate. The sample growth procedures and basic sample characteristics have been described by Dai et al. [101]. Because of the lattice mismatch between the well and the barrier materials (~0.5%), the wells were under compressive strain. In order to eliminate any effects of photoexcited carriers in the GaAs substrate, we etched away the substrate using bromine and methanol with the sample mounted on a sapphire window using low temperature epoxy. We coated the sample surface with semitransparent NiCr (~20 nm) in order to reduce unwanted Fabry-Perot interference effects [102].

The experimental setup for the two-color TRCR spectroscopy of this work was similar to that in [87,103]. The FIR source was the Stanford Picosecond Free Electron Laser (FEL), which was tunable from 3 to 80 μm and emitted a transform-limited and diffraction-limited pulsed beam with pulse durations between 600 fs to 2 ps and energies as high as 1 μJ. The NIR source was a Spectra Physics Tsunami Ti:Sapphire laser seeding a regenerative amplifier (Positive Light Spitfire) operating at 800 nm. The FIR and NIR beams were combined by a Pellicle plate and then focused onto the sample, which was placed in an 8 Tesla Oxford Instrument Spectromag 4000 with sapphire cold windows and polypropylene room temperature windows. The measurements were performed in the Faraday geometry at three FIR wavelengths (46, 42 and
38.5 μm) either by varying the time delay at a fixed magnetic field or by varying the magnetic field at a fixed time delay. The FEL produced millisecond macropulses at 10 Hz, which contained picosecond micropulses at 11.8 MHz. The Ti:Sapphire oscillator was locked at the 7th harmonic of the FEL micropulse repetition rate, i.e., 82.6 MHz, and our synchronization electronics allowed us to obtain one NIR pulse from the regenerative amplifier for every FIR macropulse. A liquid-helium-cooled Ge:Ga photoconductive detector was used to collect the FIR beam transmitted through the sample.

Figure 4.2 shows typical TRCR data. The photoinduced transmission change of the FIR probe pulse is plotted as a function of magnetic field (B) and time delay in the ranges of 0-8 Tesla and 0-600 ps. The photoinduced transmission change is defined as (T₀ − T)/T₀, where T₀ is the transmission value before the arrival of the NIR pump pulse. The wavelength of the FIR radiation was 42 μm and the sample temperature was 1.5 K.

We can concentrate on particular aspects of the data by looking at slices along either B or delay. Typical B-scans are shown in fig. 4.3(a) for six different fixed delays. These traces show the smooth evolution of CR from 25 ps to 1.3 ns, a time scale much shorter than the interband lifetime. At 25 ps after excitation, the CR line is significantly broadened to higher B, i.e., to heavier effective mass. As time progresses, the electrons relax towards the band edge, resulting in a lighter mass at 1.3 ns. The average cyclotron mass, carrier density, and average scattering time,
Figure 4.2: Picosecond TRCR spectra of photocreated electrons in InSb QWs. The photoinduced transmission change is plotted as a function of time delay and magnetic field. The photoinduced transmission change is defined as \((T_0 - T)/T_0\), where \(T_0\) is the transmission value before the arrival of the NIR pump pulse. The wavelength of the FIR radiation was 42 \(\mu\)m and the sample temperature was 1.5 K.

obtained through Lorentzian fits to the traces in (a), are plotted vs. delay in Figs. 4.3 (b-d). The CR lineshape was fitted by calculating the complex conductivity in the Drude model [104]:

\[
\sigma_{\pm}(B) = \frac{n e^2}{z_0 m^*_{CR}} \frac{1}{1/\tau_{CR} + i(\omega \pm eB/m^*_{CR})}. \tag{4.1}
\]

Here, \(n\) is the carrier density, \(z_0\) is the thickness of the 2DEG, and \(\tau_{CR}\) is the cyclotron resonance scattering time. The resonances observed are fitted best as the sum of two to four Lorentzians, which is consistent with our Landau level calculations below. The scattering time is kept the same for all of these single-particle resonances at a particular delay. By doing so we are assuming that the same broadening mechanisms act on each transition at the same density. We note that the \(B\) width corresponding to the finite spectral width of the FEL pulse \((\Delta B \approx 0.087 \, \text{T})\), using \(m^* = 0.025 m_0\)
Figure 4.3: (a) Transmission change of the FIR probe pulse as a function of magnetic field at different time delays. The wavelength of the probe was 42 µm, the sample temperature was 1.5 K, and the NIR fluence was \( \sim 1.3 \text{ mJ/cm}^2 \). (b) Average cyclotron mass vs. time delay. (c) Carrier density \( (n) \) vs. time delay. (d) Average momentum scattering time \( (t_p) \) vs. time delay.

is much smaller than the width of any of the resonances we observe.

The wavelength dependence of the resonance position is consistent with CR. At a delay of 1.4, 1.3, and 1.3 ns, the resonance positions of 46, 42, and 38.5 µm probe light are 4.9, 5.1 and 6.1 T, respectively. These resonance positions correspond to an effective mass of 0.021 +/- 0.002 \( m_0 \).

### 4.3 Pidgeon-Brown Model

To better understand the TRCR spectra, we calculated energy levels in the QW structure. We first used an eight-band self-consistent model in the absence of \( B \) to calculate the subband energies in the conduction and valence bands by taking
into account the energy dependence of the electron and hole masses in the envelope function formalism (see ch. 2). We also included strain effects according to [105]. For the 35 nm InSb / 50 nm Al_{0.05}In_{0.21}Sb QW structure, we found three subbands in the conduction band with energies 12, 42, and 83 meV, assuming the well and barrier band gaps to be 249 meV and 416 meV, respectively, and the conduction band offset to be 62% of the band gap difference [106]. The first three heavy hole (HH) subband energies were calculated to be 1, 4, and 9.5 meV. The strain in this system gives rise to a large splitting between the HH and light hole (LH) bands of 30 meV. For the structure studied here, two LH subbands are expected at 7 and 25 meV below the top of the LH confinement potential.

The Landau levels of the QW are calculated using a modified Pidgeon-Brown model [83]. The Pidgeon-Brown model is an 8-band $k\cdot p$ model which includes an applied magnetic field via the envelope function approximation. The Hamiltonian is basically eq. 2.7, with magnetic field incorporated by redefining $k$ as $p + eA$ and adding a Zeeman term. The resulting secular equation is two $4 \times 4$ matrices which can be solved numerically at a particular magnetic field and Landau index to give the energies of the two spin states. By mapping these energies over a range of magnetic fields and Landau indices, one can extract the energy dependence of the effective mass and $g$-factor. To apply the Pidgeon-Brown model to a QW, the bulk energy gap is replaced by an effective band gap which is the energy separation between the lowest conduction and highest valence subbands.
Figure 4.4: Conduction band Landau levels calculated using a modified Pidgeon-Brown model, which takes into account confinement and strain effects. Black (gray) lines are for the first (second) subband. Solid (dashed) lines are spin up (down) states. The arrows show the possible CR transitions for 42 μm radiation.

As an aside, the $\mathbf{k} \cdot \mathbf{p}$ method with magnetic field has different names depending on the number of bands included in the calculation. Two bands is the Bowers-Yafet model [107], four bands is the Luttinger model [108], six bands is the Suzuki-Hensel model [84], and eight bands is of course the Pidgeon-Brown model [83].

Several calculated low-lying conduction band Landau levels are shown in fig. 4.4, where the black (gray) lines originate from the first (second) subband and solid (dashed) lines represent spin up (down) states. This calculation takes into account the strong energy dependence of the $g$-factor and of the effective mass, which cause different transitions to occur at different magnetic fields.

We have neglected transitions in the hole bands. Because we used linearly polarized FIR light, our measurement is sensitive to hole CR as well as electron CR.
And in fact, the light hole band edge effective mass, $0.016m_0$, is close to the electron effective mass, $0.0139m_0$, so it is possible that LH CR overlaps electron CR. However, our analysis in terms of electron CR alone is self-consistent.

The calculated electron energy levels in fig. 4.4 strongly suggest that the resonances we observe consist of multiple transitions. The allowed transitions are such that $\Delta n = 0$, $\Delta N = 1$, and $\Delta s = 0$. The electric field vector of our FIR probe lay in the QW plane and thus did not couple with the intersubband transition ($\Delta n = 1$). Electric-dipole-excited electron spin resonance ($\Delta n$, $\Delta N$, $\Delta s$) = (0,0,1) occurs at fields above 8 T for the wavelengths of the experiment [109,110]. Finally, combined resonance ($\Delta n$, $\Delta N$, $\Delta s$) = (0,1,1) is not allowed in the Faraday geometry [111]. In fig. 4.4, the allowed CR transitions for 42 $\mu$m probe light in the first and second subbands are indicated by arrows. Because the individual CR transitions are too broad to resolve separately, however, we do not directly observe spin- or Landau-level-resolved CR.

### 4.4 Simulation

Because different transitions occur at different magnetic fields, it is possible to combine the data and the calculation in fig. 4.4 to extract the time evolution of the energy of the excited carrier population. Our detailed simulation of the TRCR spectra obtained for 42 $\mu$m probe light is shown in fig. 4.5. The fixed parameters in the simulation were the calculated $B$ positions of the CR transitions and, at each time
Figure 4.5: (a) Simulated TRCR spectra for the traces in fig. 4.3, based on the energy levels shown in fig. 4.4. Simulation successfully reproduces the main observed features arising from nonparabolicity. (b) Cartoon at left (right) shows a Fermi distribution at short (long) time delay, when the carrier population is large and hot (small and cold). (c) Simulated carrier temperature as a function of time delay.

delay, the carrier density and scattering time, taken from fig. 2. The fitted carrier density accounts for recombination. The variable parameter was the temperature of the Fermi-Dirac distribution of the excited carriers. We neglected continuum states above the barrier energy and many-body effects such as band gap renormalization. We took into account the reduction in CR-active population that occurs when the upper level of a transition is populated.

The results accurately reproduced the observed features, which arise from the strong conduction band nonparabolicity. Namely, the initially excited carriers, having a high carrier temperature, populate some states with high Landau indices, where the cyclotron masses are higher than that of the lowest electron Landau level. This is the main cause for the significant initial broadening (inhomogeneous broadening).
As time progresses, the electrons relax towards the lowest Landau level, and, as a result, the number of transitions contributing to the observed linewidth decreases with time. In addition, the linewidth of individual CR transitions decreases with time. This indicates that each energy level is initially broadened due to the high scattering rate in the high density and high temperature regime (homogeneous broadening). As time progresses, both the electron temperature and density decrease, making individual transitions narrower. The simulation was not representative at very short time delays (< 200 ps) because it did not include the dynamics of highly-excited carriers relaxing into the well.

Although the individual Landau level transitions are too broad to resolve, our simulation, Landau level calculation, and fittings are consistent with each other. Hence, our single-particle analysis is valid and Coulomb collapse of the nonparabolicity-induced CR splitting is not present. This result contrasts with GaAs, in which the nonparabolicity splitting is so small that it is overwhelmed by many-body effects (Coulomb collapse). Only a single resonance is observed in most experimental situations [67–69, 72, 112–114]. However, in InAs, nonparabolicity-induced CR splittings can be observed easily [111, 115–117]. InSb is even more nonparabolic, so we are confident in our multi-line analysis.

Based on the simulated carrier temperatures, we have calculated the LO phonon scattering times. We used

\[ \frac{1}{\tau} = \frac{1}{\tau_{LO}} e^{-\hbar \omega_{LO}/k_B T}, \]  

(4.2)
where $\tau$ is the scattering time due to LO phonon emission and $1/\tau_{LO}$ is a material parameter given in 2D by [118]

$$
\frac{e^2 \sqrt{2m^* \hbar \omega_{LO}}}{8 \hbar^2} \left( \frac{1}{\epsilon_\infty} - \frac{1}{\epsilon_0} \right)
$$

(4.3)

For InSb, $\tau_{LO}$ in 2D $\sim$400 fs [119]. We calculated a scattering time due to LO phonons of 5 ps at 25 ps delay time, increasing smoothly to 61 ps at 1300 ps delay. The scattering time due to LO phonons is almost two orders of magnitude larger than the scattering times deduced from the fitting, indicating that some other scattering mechanism is dominant. While there are a number of possible decay mechanisms, the most important one, especially in narrow gap semiconductors at high densities (which is the case for the InSb QWs in the current study), is Auger relaxation [120]. Our recent TRCR work on bulk InSb [119] clearly showed that this is the main recombination mechanism, and it is consistent with the observation that the scattering time is roughly inversely proportional to the density [119,120].

In summary, we performed picosecond two-color (NIR and FIR) TRCR spectroscopy on undoped InSb/AlInSb quantum wells. We monitored the dynamics of FIR transmission while we varied the magnetic field and the time delay between the NIR and FIR pulses. The strong nonparabolicity of the InSb conduction band allowed us to study the relaxation dynamics of quasi-2D electrons in a quantizing magnetic field in unprecedented detail.
4.5 Further experiments

Our results suggest several further experiments on carrier dynamics in nonparabolic Landau-quantized systems. One line of inquiry is phonon resonances. We expect a faster relaxation time when the probe is resonant with the phonon energy divided by an integer [67, 88, 92] or an integer times the phonon energy. It is possible for us to investigate this intriguing regime using the equipment available to us ($\lambda_{FIR} = 26 \mu m$), but we have not yet done so. If we could go to higher fields, around 15 T, we could also see what happens when the phonon energy is resonant with the spin splitting.

In addition, TRCR may be useful for studying spin relaxation near the quantum limit at low densities and in high magnetic fields, where relaxation slow-down and even quenching are predicted. Experimentally, we may be able to do this simply by measuring at longer time delays, as shown in fig. 4.6.

One difficulty that spin relaxation experiments may face, however, is that the spin relaxation time may turn out to be much longer than the interband relaxation time, as it is in quantum dots. In this case, the CR signal, which is due to free carriers in the conduction band, would disappear well before the spin relaxation occurred. One possible solution to this problem is circularly polarized interband excitation of an $n$-type quantum well. In this configuration, the spin polarization of the photocreated carriers may be transferred to the degenerate population, as has been demonstrated in GaAs QWs [121]. This mechanism would allow CR detection of the spin relaxation
Figure 4.6: Diagram of Landau levels (top) and the TRCR signal expected at long time delays when spin-flip transitions are allowed (middle) and forbidden (bottom). at long delay times.
Chapter 5

Intersubband Transitions

In this chapter, I will summarize the important results about intersubband transitions (ISBTs) in my MS thesis [122]. In addition, I will present new observations of intersubband transitions in very narrow QWs, in QWs grown on GaSb substrates, and in asymmetric double QWs. These advances are important components of an intersubband-based THz generator.

5.1 Sample Growth and Characterization

The samples were grown by molecular beam epitaxy (MBE) at Osaka Institute of Technology. The substrate was GaAs, and the buffer structure consisted of GaAs (300 nm) / AlAs (10 nm) / AlSb (100 nm) / GaSb (300 nm) / AlSb (1000 nm) / [AlSb/GaSb (6 nm / 6 nm)]×15 / AlGaSb (200 nm). The purpose of the thick buffer region is to compensate for the lattice mismatch between GaAs, with a lattice constant of 5.65 Å, and the 6.1 Å quantum wells. The active region consisted of 20 periods of InAs (x nm) / AlSb (10nm), where x = 1.2 to 10.5 nm. The active region was capped by 10 nm of GaSb or InAs to prevent oxidation of the AlSb. Most samples were not intentionally doped, but had a total electron density of $3 \times 10^{10}$ to $4.5 \times 10^{14}$ cm$^{-2}$ at 77 K. Some samples were doped in the well with Si. These had total electron densities of $8.7 \times 10^{12}$ to $2.7 \times 10^{15}$ cm$^{-2}$ at 77 K. The wells were
under tensile strain due to the 1.3% lattice mismatch between InAs and AlSb.

In addition to single quantum wells, we grew more than a dozen double quantum well (DQW) structures, both symmetric and asymmetric. They were grown on GaAs substrates with the buffer region described above. The symmetric DQWs were described in a pervious work [122]. The asymmetric DQWs (ADQWs) consisted of 10 or 20 periods of InAs (12.4, 13.6, or 14.2 nm) / AlSb (0.9 nm) / InAs (5.4 or 5.8 nm), separated by 20 nm AlSb barriers. The undoped ADQWs had densities of $1.5 \times 10^{12}$ to $3.8 \times 10^{13} \text{ cm}^{-2}$ at 77 K. Several of the ADQWs were Si doped in the wide well; these had densities of $2.2 \times 10^{13}$ to $2.7 \times 10^{15} \text{ cm}^{-2}$ at 77 K.

We also grew several samples on GaSb substrates. These substrates are attractive because they are lattice matched to the heterostructure. So, the buffer region consists of only GaSb (200 nm) / AlGaSb (200 nm), and GaSb-substrate samples are quick to grow. However, because semi-insulating GaSb substrates are not available, they cannot be used for transport measurements or far-infrared generation.

### 5.1.1 Carrier Distribution

In the not intentionally doped QWs, the carrier distribution is not necessarily uniform, which affects the observed intersubband absorption. We deduced the carrier distribution by comparing Hall measurements of the density, which are sensitive to the total carrier density in the sample, to Shubnikov-de Haas measurements of the density, which give the density of a single quantum well [122]. It turns out that
the carrier distribution depends on how soon the sample is grown after the start of new MBE sources. At the start of the MBE run, all of the QWs in the sample are uniformly populated. By the time 25 or 30 samples have been grown, only one QW is populated.

The most likely sources of electrons in these wells are bulk [123] and interface [124,125] donors and Fermi level pinning at the surface of the GaSb cap layer [126], $\sim500$ meV above the InAs conduction band. We attribute the large variation in total electron density to variation in the density of volatile bulk donors, which are initially present in the MBE source but are removed rapidly during the run. So, the early grown samples are dominated by bulk donors, while the late grown samples are populated primarily by Fermi level pinning at the GaSb surface.

Several steps can be taken to improve the uniformity of multiple quantum well samples. First, growth should wait until the density of bulk donors is low. Second, the QWs should be doped with Si to control the density. Finally, the band bending can be minimized by capping the structure with InAs instead of GaSb. In the InAs cap layer, the Fermi energy is pinned only 80 meV above the InAs conduction band [127], causing less band bending and hence greater uniformity in the energy band profile and carrier distribution.
5.2 Intersubband Absorption in Single QWs

5.2.1 Selection Rules and Experimental Geometry

We detect the intersubband resonance via its polarization selection rule. Light is absorbed or emitted by the intersubband resonance only when it is polarized in the growth direction, \( z \) \cite{75,128}. Semi-classically, the polarization selection rule arises because the strong \( z \) direction confinement makes the electrons oscillate in the \( z \) direction much faster than in \( x \) and \( y \). This confinement creates dipoles oriented along the \( z \) direction which can radiate and absorb only radiation polarized along \( z \).

Quantum mechanically, the intersubband transition rate between initial and final states depends on Fermi’s golden rule \cite{75}:

\[
W = \frac{4\pi}{\hbar} | < \psi_f | H | \psi_i > |^2 \frac{2\pi}{\hbar} \delta(E_f - E_i - \hbar\omega)
\]  

(5.1)

Here the dipole interaction Hamiltonian is

\[
H = -qE_0 \hat{\mathbf{e}} \cdot \mathbf{r}
\]  

(5.2)

where \( E_0 \) is the magnitude of the electric field and \( \hat{\mathbf{e}} \) is a unit vector in the polarization direction. The matrix element can be written as

\[
< \psi | \hat{\mathbf{e}} \cdot \mathbf{r} | \psi' > = \hat{\mathbf{e}} \cdot < u_u | r | u_{u'} > < f_n | f_{n'} > + \hat{\mathbf{e}} \cdot < u_{u'} | u_u > < f_{n'} | r | f_n >
\]  

(5.3)

where we have separated the wavefunction into Bloch function \((u)\) and envelope function \((f)\) components. The momentum matrix element of the Bloch function is nonzero only for interband transitions. The overlap integral of the Bloch functions
is zero for interband transitions and one for intraband transitions. So, the first term describes interband transitions and the second describes intraband (intersubband) transitions. The dipole matrix element of the envelope function is

\[
< f_n | \hat{d} \cdot \mathbf{r} | f_{n'} > = \int d^3r f_n^*(z) e^{-ik'_\perp \cdot r'_\perp} [\hat{e}_x x + \hat{e}_y y + \hat{e}_z z] f_{n'}(z) e^{ik_\perp \cdot r_\perp} \tag{5.4}
\]

The \(x\) and \(y\) components are zero except when \(k'_\perp = k_\perp\) and \(n' = n\), that is, when \(\omega = 0\). So, only Drude absorption is possible for light polarized in the QW plane. Intersubband absorption occurs only for light polarized along the growth direction, \(z\). This polarization selection rule determines the geometry of intersband measurements and devices.

The simplest way to couple light to the ISBT is by putting the sample at an angle to the incident beam, as shown in fig. 5.1. However, because the \(z\) component of the beam is small in this geometry, the ISBT signal is weak. In order to couple light to the intersubband resonance more strongly, we use the multibounce parallelogram geometry. The sample is polished at 45° and the light is incident normal to the polished face. In this geometry, \(s\)-polarized light cannot be absorbed by the intersubband resonance but half of the \(p\)-polarized light can be absorbed. A similar strong coupling geometry is the single-bounce trapezoid, which we used in some difference frequency generation experiments.

We coated the QW surface with 100 nm of gold. This treatment is expected to make light beams electric field have an antinode rather than a node at the QW surface, increasing the coupling to the QWs [129]. We have compared coated and
uncoated QWs. We observed the same absorption intensity (unlike Warburton et al. [42]) and a small shift in energy (2 meV higher when uncoated). All of the samples discussed here are coated.

The intersubband absorption was measured with a Fourier transform infrared spectrometer. We used a Jasco FTIR-660 with a spectral range of 15,000-100 cm$^{-1}$ (666 nm 100 μm). The sample was placed in a helium-flow cryostat, a wire grid polarizer was placed before the sample, and the beam path was purged with nitrogen. In order to detect the resonance, we took the ratio of the $p$-polarized light to the $s$-polarized light (intersubband active to inactive). In order to cancel out background differences in the signals of $s$- and $p$-polarized light, we also took the $s/p$ ratio of an empty aperture in the same shape and position of the sample. So, the final spectrum was

$$T = \frac{T_{p} / T_{s}}{T_{p} / T_{s}}_{\text{sample}} / T_{p} / T_{s} \text{reference}. \quad (5.5)$$

The absorption coefficient of the sample, $\alpha$, was calculated as

$$\alpha = \frac{-\ln(T) + R}{L}, \quad (5.6)$$

where $R$ is a sample- and temperature-dependent constant chosen to make $\alpha$ equal zero away from the absorption, and $L$ is the total length of well material that the infrared beam passes through [75].
Figure 5.1: Three common experimental geometries for the observation of ISBT. Counterclockwise from top left, the Brewster angle geometry, the trapezoid geometry, and the multi-bounce parallelogram geometry (this work).

5.2.2 Temperature Dependence

One feature of intersubband resonances which makes them attractive for device applications is their weak temperature dependence compared to interband transitions. This is a result of the collective nature of intersubband resonances, which will be discussed in sec. 5.4. Figure 5.2 shows the temperature dependence of the intersubband absorption spectrum of sample 0201181T, with 20 periods of 7.4 nm QWs and a total carrier density of $1.7 \times 10^{13}$ cm$^{-2}$ at 77 K. The resonance redshifts and broadens slightly with increasing temperature.

Figure 5.3 shows the temperature dependence of the intersubband absorption energy and full-width at half maximum (FWHM) for wells from 5 to 10 nm wide [130]. These samples all have a density of around $1 \times 10^{13}$ cm$^{-2}$ per well, so only the lowest
Figure 5.2: The intersubband absorption spectrum of a 7.4 nm InAs/AlSb QW (sample 0201181T) as a function of temperature. Curves vertically offset for clarity.

As expected from quantum mechanics, the ISBT energy increases with decreasing well width from 212 meV (5.84 μm) for 10 nm wells to 385 meV (3.22 μm) for 5 nm wells at 10 K. For a given well width, the ISBT peak position redshifts with increasing temperature. The amount of redshift from 10 to 300 K increases with decreasing well width from 6.2 meV for 10 nm to 12.0 meV for 5 nm. Much smaller redshifts have been reported for ISBTs in GaAs/AlGaAs quantum wells [131–133]. The ISBT linewidth has a weak temperature dependence, increasing from 10 K to 300 K by a factor of less than 2. The low temperature linewidths are relatively small, ranging from 11 to 23 meV (2.8 to 5.8 THz), indicating that the interface type and smoothness are well controlled. The ISBT linewidth does not vary systematically with well width.
Figure 5.3: Temperature dependence of the intersubband absorption peak energy (left) and full-width at half maximum (right) for InAs/AlSb QWs with densities around $1 \times 10^{13}$ cm$^{-2}$ per well.
5.2.3 High Energy

For near-infrared pumping of intersubband-based devices, it is necessary to have near-infrared intersubband resonances. We have reported the highest energy $E_{12}$ intersubband absorption in an InAs/AlSb QW: 670 meV (1.85 $\mu$m) in a 2.1 nm QW at 77 K (sample 0403081T). Othani et al. reported the next highest energy ISBT in InAs/AlSb: 626 meV (1.98 $\mu$m) with FWHM = 60 meV in 2.7 nm Si-doped QWs [34]. We found that the key to observing this high-energy resonance was heavy doping. Our sample had a density of $9 \times 10^{12}$ cm$^{-2}$ per well at 77 K. Fig. 5.4 shows the absorption spectrum of this resonance, plotted as $p$-polarized (intersubband active) and $s$-polarized (intersubband inactive) absorption. The FWHM of this resonance is about 120 meV. The oscillations in the $s$-polarized spectrum with a period of about 100 meV are believed to be due to multiple reflections within the MBE-grown region. The increased absorption above 775 meV in both polarizations is due to interband absorption in the GaSb buffer layer.

The ISBT FWHM increases sharply for well widths less than 3 nm, as shown in fig. 5.5. This result is not unexpected, since the broadening of the ISBT absorption is strongly influenced by interface roughness scattering [134]. The narrower the well, the stronger the interface roughness broadening. For this reason, it may be worth exploring using higher transitions, such as $E_{14}$, in wider wells to achieve near-infrared resonances with narrow linewidths.

$N$-type intentional doping of InAs/AlSb QWs must be made by doping in the
Figure 5.4: The absorption spectrum of the highest energy $E_{12}$ intersubband resonance, in a 2.1 nm InAs/AlSb QW (sample 0403081T), plotted as $p$-polarized (intersubband active) and $s$-polarized (intersubband inactive) absorption.

Figure 5.5: The ISBT FWHM vs. well width. The FWHM goes like $(\text{width})^{-1.1}$. The FWHM increases sharply for well widths less than 3 nm.
wells, not the barriers, but well-doping does not broaden the intersubband resonance much. Well-doping is necessary because the best n-type donor for the AlSb barrier is Te, but this element contaminates later MBE growths. As expected, doping in the well reduces the 2DEG mobility compared to remote doping. It has also been shown that well doping increases the ISBT linewidth [133]. However, the ISBT linewidth is influenced primarily by collective effects and scattering off interface roughness; ionized impurities are less important [42,134]. We have not systematically studied the effects of doping on our ISBTs, but it does not dramatically increase the linewidth.

5.2.4 No Rashba Splitting

In a previous work [122], we reported results that were suggestive of Rashba splitting in InAs/AlSb QWs. However, further studies have shown that this effect is not present. InAs/AlSb QWs can be grown with two different types of interfaces, depending on the growth sequence: InSb-like or AlAs-like. These two types of interface layers are expected to have different effects on the electrostatic potential in the QW. So, a QW grown with a different interface on each side experiences an asymmetric electric field; this is an example of native interface asymmetry [135]. The electrons in the well feel the asymmetric electric field as an effective magnetic field, which lifts the spin-degeneracy of the subbands. We observed a split intersubband resonance in a 20-period 10.5 nm QW sample with native interface asymmetry (0205171T), as shown in fig. 5.6 (left). However, two 10 nm QW samples with native inter-
Figure 5.6: Intersubband absorption spectra of three samples with native interface asymmetry. The splitting in the sample on the left (0205171T) was tentatively attributed to Rashba splitting until two nearly identical samples (center, 0311122T, and right, 0211062T) showed no such splitting. Sample 0311122T has almost two orders of magnitude fewer carriers compared with the other samples, hence the smaller integrated absorption.

face asymmetry (0211062T and 0311122T) showed no such splitting. We conclude that the original observation was not Rashba splitting but perhaps a change in QW thickness during the growth.

5.2.5 GaSb Substrates

GaSb substrates are attractive candidates for InAs/AlSb QW growth because unlike GaAs, they are lattice-matched to the QWs, so a thick buffer structure is not required. We have observed intersubband absorption in 7 and 10 nm QWs grown on p-type GaSb substrates. Because of strong free-carrier absorption in the substrate,
we found that these samples could not be measured in the usual multibounce geometry (fig. 5.1), unlike the InAs/AlSb QWs of Prevot et al. on n-GaSb substrates [136]. Instead, the Brewster angle geometry was used. This geometry results in a much weaker absorption coefficient than the multibounce geometry, as shown in fig. 5.7. For this reason, intersubband absorption in QWs on GaSb substrates was observed only in wells which were heavily doped.

The origin of this reduced absorption in the Brewster geometry is unclear. The calculation of the absorption coefficient takes into account the number of passes of the beam through the active region and the beam’s propagation direction inside the sample. The reduction may be due to reduced electric field overlap in the Brewster geometry compared with the coated, multibounce geometry [129].

### 5.2.6 Integrated Absorption

The area under the intersubband absorption curve is proportional to the square of the dipole matrix element and to the carrier density. Sirtori et al. showed that the integrated absorption is:

\[
\ln 10 \int \log T(\omega) d\omega = \frac{\pi \hbar \rho_s N n_p e^2 \sin^2 \theta}{2 m_0 \varepsilon_0 \varepsilon_n \cos \theta} f_{0j} \tag{5.7}
\]

where \(T\) is defined in eq. 5.5, \(\rho_s\) is the 2D density, \(N\) is the number of QW periods, \(n_p\) is the number of passes the light makes through the QWs, \(\theta\) is the angle the beam makes with the growth direction, and \(n\) is the refractive index. The oscillator
Figure 5.7: Left: Intersubband absorption coefficient of two samples measured in the Brewster geometry at 77 K: red: 20 periods of 10 nm QWs on a GaAs substrate with a density of $3.5 \times 10^{12}$ cm$^{-2}$ per well (0211012T); blue: 20 periods of 7 nm QWs grown on a GaSb substrate with a density of $5 \times 10^{12}$ cm$^{-2}$ per well (0411102T). Right: Intersubband absorption coefficient of sample 0211012T in the multibounce geometry at 77 K. Note the different vertical scale. The resonance is wider in the multibounce geometry because the absorption is saturated.

Strength, $f_{0j}$, is taken to be

$$f_{nn'} = \frac{2m_0\omega_{nn'}}{\hbar} |\langle n|z|n' \rangle|^2.$$  \hspace{1cm} (5.8)

This expression is different from the more usual definition of the oscillator strength in which $m_0$ is replaced by $m^*$. The usual oscillator strength obeys the well-known $f$-sum rule: $\sum_{n'} f'_{nn'} = 1$, where the sum is over all transitions. Sirtori et al.’s expression is more general because it applies even when the effective mass changes with position.

Note also that the expression for the integrated absorption does not include the QW thickness. So, this expression avoids the ambiguity inherent in describing a 2D system by a 3D absorption coefficient [75].

For a rough estimate of the dipole matrix element of an InAs/AlSb QW, the QW
can be treated as infinitely deep and parabolic [75]. The envelope functions are then

$$\varphi_n(z) = \sqrt{\frac{2}{w}} \sin \left( \frac{n\pi z}{w} \right), \quad (5.9)$$

where $w$ is the QW thickness. The dipole matrix element is

$$| <1|z|2> | = \frac{8\hbar}{3\omega m^* \omega_{12}}. \quad (5.10)$$

For a 10 nm InAs QW, taking $\hbar \omega_{12} = 205$ meV and $m^* = 0.035 m_0$, this expression gives $z_{12} = 2.8$ nm.

I calculated the intersubband dipole matrix element from integrated absorption spectra and found good agreement with the theoretical expectation. I considered four 10 nm QWs for which we have both Hall and Shubnikov-de Haas density measurements. The samples all had only one or two wells populated [137]. The average dipole matrix element calculated from the density and integrated absorption was 2.6 nm, in good agreement with the theoretical expectation [138].

In principle, the integrated absorption allows us to measure the densities of samples grown on conducting substrates, such as p-GaSb, for which Hall measurements may be less reliable. However, as discussed in sec. 5.2.5, ISBT in QWs on p-GaSb substrates had to be measured in the Brewster angle geometry. There is a systematic reduction in the absorption coefficient for this geometry, regardless of the substrate material, as shown in fig. 5.7. This reduction means that the carrier density deduced from the integrated absorption in the Brewster angle geometry is a few hundred times smaller than the actual density. Further study is necessary to resolve this discrepancy.
5.3 Double QWs for Difference Frequency Generation

We have measured intersubband absorption with THz splitting in asymmetric double QWs (ADQWs) designed for intersubband difference frequency generation. The calculated lowest energy levels and wavefunctions of a typical design are shown in fig. 5.8. Although our ultimate goal is difference frequency mixing of near-infrared pumps, this preliminary structure was designed for mid-infrared pumps in order to take advantage of the narrow resonances in wider wells. This structure has large intersubband dipole matrix elements of \( z_{12} = 1.7 \) nm, \( z_{13} = 18 \) nm, and \( z_{23} = 44 \) nm. These large matrix elements will contribute to a strong nonlinear susceptibility (sec. 6.1).

Figure 5.9 shows the intersubband absorption spectra of low- and medium-density ADQWs. Samples 0306272T and 0311142T have the structure shown in fig. 5.8; the rest have slightly narrower wells. All of these samples show a clearly resolved splitting of 30-100 meV between the \( E_{12} \) and \( E_{13} \) intersubband transitions. Several samples also show other small resonances of indeterminate origin. One ADQW sample which was doped in the wide well (0306271T; not shown) did not display splitting of the intersubband absorption. However, the other doped samples, 0404231T and 0404191T (fig. 5.10), showed clear splitting, indicating strong inter-well coherence despite the Si impurities.

Figure 5.10 shows the intersubband absorption spectra of the three highest density ADQWs. In sample 0404191T, the Fermi energy is above \( E_3 \), so no \( E_{12} \) or \( E_{13} \)
Figure 5.8: The calculated lowest energy levels and wavefunctions of a typical double QW structure designed for intersubband difference frequency generation. Note the break in the vertical axis.
Figure 5.9: Intersubband absorption spectra at 77 K of seven ADQW samples designed for THz difference frequency generation. Each graph is labeled with the sample name, total density in cm$^{-2}$, and number of ADQW periods. Sample 0404231T is Si-doped in the wide well but still shows clearly resolved THz splitting.
Figure 5.10: High energy intersubband absorption spectra at 77 K of three ADQW samples with high densities. In sample 0404191T, the Fermi energy is above $E_3$ and no $E_{12}$ or $E_{13}$ resonance is observed. In general, the nonlinear susceptibility increases with increasing density. However, this huge density actually reduces the nonlinear susceptibility between levels 1-3 because the susceptibility depends on the population difference, which is negligible when all three levels are fully populated (sec. 6.1). In the other two samples, the Fermi energy is just below $E_2$, so thermal population of $E_2$ and $E_3$ permits absorption at $E_{34}$ and $E_{24}$. 
5.4 Modeling of Many-Body Effects

5.4.1 Single-particle Broadening Mechanisms

In GaAs, which is not very nonparabolic, the intersubband absorption linewidth can be attributed primarily to single particle broadening effects. Homogeneous broadening mechanisms include ionized impurity scattering and phonon scattering. Interface roughness scattering can also behave as a homogeneous broadening mechanism when the in-plane size of the fluctuations is less than the characteristic wavelength of the electrons, $\lambda_F$. This is the case in our InAs/AlSb QWs, where the IFR fluctuations are $\sim 43 \text{ Å}$ [80] and $\lambda_F \sim 0.5 \text{ μm}$. However, when the fluctuations are much larger, they cause inhomogeneous broadening, along with well-to-well width fluctuations and nonparabolicity.

In strongly nonparabolic semiconductors such as InAs, nonparabolicity broadening completely dominates the single-particle spectrum. For example, since the lowest subband has a lighter mass than the second subband, electrons of different wavevectors see different intersubband energy separations, as shown in fig. 5.11. In a 10 nm QW at 0 K with a density of $1.84 \times 10^{12} \text{ cm}^{-2}$, the subband separation is 208 meV at $k = 0$ and 166 meV at $k_F$. Given this spread of single-particle transition energies, the expected width of the intersubband absorption is $208 \times 166 = 42 \text{ meV}$. But in fact, the observed absorption is less than half as broad. The resonance behaves so differently from the single-particle calculation due to the strong effect of many-body interactions.
Figure 5.11: Calculated dispersion of the first three subbands of a 10 nm InAs/AlSb QW. The horizontal line is the Fermi energy for an electron density of $1.84 \times 10^{12}$ cm$^{-2}$. The "parabolic" line has constant effective mass.

### 5.4.2 Intersubband Semiconductor Bloch Equations

Many-body interactions play a critical role in intersubband absorption. In order to incorporate them properly into the model, our collaborators at NASA Ames Research Center calculated the intersubband absorption coefficient using the semiconductor Bloch equations (SBEs) [4, 80].

The SBEs [139] describe the interaction of a semiconductor with light through the dipole interaction. They provide the most general formalism to incorporate coherent and incoherent light-semiconductor interactions by taking into account band dispersions, Coulomb interactions, and various decoherence and relaxation processes. They are an extension of the Bloch equations, originally derived to describe the dynamics of a magnetic moment in a time-dependent magnetic field, and the optical
Bloch equations, which describe the dynamics of an electric dipole in an electric field. The SBEs have been successfully applied to the study of interband dynamics in many semiconductor systems, but this is one of the first cases in which they have been applied to intersubband dynamics. Important parameters in the Bloch equations are $T_1$, the population relaxation time, $T_2$, the spin relaxation time, and $\Omega$, the Rabi frequency, which is related to the moments oscillation frequency. Li \textit{et al.} use the reduced single particle density matrix:

\begin{equation}
\begin{pmatrix}
\langle a_k^+ a_k \rangle & \langle b_{-k} a_k \rangle \\
\langle a_k^+ b_{-k}^+ \rangle & \langle b_{-k}^+ b_{-k} \rangle
\end{pmatrix} = \begin{pmatrix}
n_k^c & P_k \\
P_k^* & n_k^v
\end{pmatrix}.
\end{equation}

Here $a^+, a$ and $b^+, b$ are the electron and hole creation and annihilation operators, respectively, $P_k$ is the electron-hole transition amplitude, and $n_k^c, n_k^v$ are the conduction and valence band occupation probabilities. The Hamiltonian contains a single-particle energy, Coulomb interaction between carriers, and dipole interaction between carriers and light. They use the Heisenberg equation of motion to derive the coupled equations of motion for $P_k$ and $n_k^c, n_k^v$. The resulting Hartree-Fock interband semiconductor Bloch equations are:

\begin{align}
\frac{\partial P_k}{\partial t} &= -i(e_{e_2,k} + e_{e_1,k})P_k - i(n_{e_2,k} + n_{e_1,k} - 1)\omega_{R,k} + \frac{\partial P_k}{\partial t} \bigg|_{\text{coll}} \\
\frac{\partial n_{e_2,k}}{\partial t} &= -2\text{Im}(\omega_{R,k}P_k^*) + \frac{\partial n_{e_2,k}}{\partial t} \bigg|_{\text{coll}} \\
\frac{\partial n_{e_1,k}}{\partial t} &= -2\text{Im}(\omega_{R,k}P_k^*) + \frac{\partial n_{e_1,k}}{\partial t} \bigg|_{\text{coll}}
\end{align}

where

\begin{equation}
\hbar e_{e_2,k} = \hbar E_{e_2,k} - \sum_q V_{|\mu,q|} n_{e_2,q}
\end{equation}
\[ \omega_{R,k} = \frac{1}{\hbar} \left[ d_{21} \varepsilon + \sum_{q \neq k} V_{k\rightarrow q} P_q \right]. \] (5.14)

In these equations, \( d \) is the dipole matrix element between final and initial states. \( e \) is the renormalized single-particle energy, equal to the single-particle energy plus the exchange self-energy (a many-body effect described later). \( \omega_R \) is the generalized Rabi frequency. It describes the fact that the system reacts to an effective field that is the sum of the applied field and the dipole field of all the excited electron-hole pairs. The \( \text{Im}(\omega_{R,k} P_q^* \text{e}^{-i\omega t}) \) terms describe the generation of electrons and holes by light. The collision term on the polarization represents dephasing, like \( T_2 \), and the collision terms on the population represent decay, like \( T_1 \). The SBEs also include phase space filling. The SBEs apply to light intensities which are strong enough to create enough excited electrons that there is significant Coulomb interaction between carriers. If there were no Coulomb interaction, which would happen if the excitation intensity were very low, then \( V \) would go to zero and the generalized Rabi frequency would be the usual Rabi frequency. Unlike the optical Bloch equations, the semiconductor Bloch equations include dispersion via \( k \).

For the intersubband semiconductor Bloch equations, the electron and hole levels are replaced by two lowest CB energy levels. In this case the dispersion relation of the lower level now has the same curvature direction as the upper level, rather than opposite as in the interband case. The semiconductor Bloch equations are usually applied to interband phenomena; this is one of their first applications to intersubband resonances [140].
5.4.3 Many-body Effects

The four types of Coulomb interaction can be classified according to whether they are direct or exchange, and static or dynamic. We have already discussed the direct, static Coulomb interaction. This causes band bending and screening and is included in the band structure calculation via the heterostructure potential, according to the Hartree approximation (sec. 2.2.5). The exchange, static Coulomb interaction is the self-energy correction [11]. This interaction has the effect of lowering the potential at the electrons, thus lowering the subband energies. This change affects the lowest subband the most, leading to a blueshift of the ISBT [75].

The direct, dynamic Coulomb interaction is known as the depolarization shift, or resonant screening. Allen et al. [141] showed that the ISBT could not be described by static many-body effects alone. They found that they had to take into account the screening of the applied electric field by the carriers, which causes the effective field to be smaller than the applied field. So, the absorption occurs at a higher applied frequency. The frequency shift is

\[ \Delta \omega = \frac{2e^2 \rho_s}{\varepsilon \varepsilon_0 E_{\text{12}}} S, \]  

where \( S \) is an integral related to the wavefunction overlap. We can see from this expression that the depolarization shift is larger for higher densities and lower inter-subband energies (wider wells).

The exchange, dynamic Coulomb interaction is the excitonic or vertex correction. It is the interaction between an excited electron and the remaining electrons. Ando
showed that it partially cancels the depolarization shift [142]. The dynamic many-body effects correspond to a correction to the frequency-dependent conductivity of the 2DEG.

Figure 5.12 shows the effects of the many-body interactions one at a time. Panel (a) is the single-particle spectrum, which is broad due to nonparabolicity. Adding the self energy shifts the resonance to higher energy and broadens it slightly, as does the depolarization shift. The vertex correction causes a strong redshift and a dramatic narrowing of the resonance. The vertex correction turns out to be the dominant effect in this regime. However, the interplay of many-body effects is sensitive to carrier density and well width. In general, for a fixed density, the vertex correction is more important in narrower wells and the depolarization effect is more important in wider wells.

5.4.4 Temperature Dependence

The temperature dependence of the ISBT is incorporated through the carrier density, band filling, and nonparabolicity. The temperature dependence of phonon scattering is not included. This simple treatment reproduces the experimental energy shift and linewidth change well. At low temperatures, the carrier density is low and the carriers are all at the lowest possible energies. This gives the minimum nonparabolicity broadening and a sharp absorption line (fig. 5.13 left). At room temperature, the carrier density is about three times higher and the carriers are thermally excited to
Figure 5.12: Calculated ISBT spectra for 10 nm InAs/AlSb QW at 4.5 K with a density of $8 \times 10^{11}$ cm$^{-2}$. The effects of many-body interactions on the intersubband absorption are shown singly and in combination. In (e), the dotted line is the single particle spectrum. From [4].
Figure 5.13: Nonparabolic dispersion relation and absorption coefficient at 0 K (left) and 300 K (right). The arrows indicate the single-particle ISBT energies.

higher energies in the subband. This redistribution strongly increases the nonparabolicity broadening, shifting the resonance to lower energy and broadening it (fig. 5.13 right). Because GaAs has smaller nonparabolicity, ISBT in GaAs/AlGaAs QWs show smaller redshifts [131–133].
Chapter 6
QCL Difference Frequency Generation

6.1 Second Order Nonlinear Susceptibility \((\chi^{(2)})\)

When an electromagnetic field is applied to a material, the electrons in the material respond by oscillating. These oscillating charges emit their own field which interferes with the incident field. If the incident field is strong enough, the electrons will be taken out of the bottom, parabolic part of the potential well that they feel. In this nonlinear regime, the electron's oscillation includes different frequency components. If the material is noncentrosymmetric, that is, if the potential in one direction is different from the potential in the opposite direction, then nonlinear processes can occur in second order. Second order processes involve three photons, for example

- **Second Harmonic Generation (SHG)**: \(\omega_3 = \omega_1 + \omega_1\)
- **Sum Frequency Generation (SFG)**: \(\omega_3 = \omega_1 + \omega_2\)
- **Difference Frequency Generation (DFG)**: \(\omega_1 = \omega_3 - \omega_2\)

Examples of noncentrosymmetric materials are \(\text{LiNbO}_3\) and \(\text{Beta-TaB}_2\text{O}_4\) (BB0), both of which have large second order nonlinear susceptibilities \(\sim 10^{-3}\) m/V. Another way to create an asymmetric potential is in a quantum well, either by biasing it or by varying the composition asymmetrically.

As we will see at the end of this section, the requirement of an asymmetric
potential can also be thought of in terms of allowed transitions. The second order susceptibility is strongest when all three resonant transitions are allowed, that is, have nonzero dipole matrix elements. In symmetric QWs, intersubband transitions between next-nearest neighbors are forbidden by parity, but in asymmetric QWs, all transitions are allowed.

In fact, ISBTs can have quite large nonlinear susceptibilities because of their large dipole matrix elements (sec. 5.2.6), typically on the order of the width of the QW (that is, 10's of Å) [143]. For comparison, the dipole matrix element in bulk material is only on the order of the atomic separation.

Our intersubband system is nearly in resonance with the incident fields, that is, the detuning of the lasers is less than the broadening of the resonance. In this case, the susceptibility should be derived in the density matrix formalism in order to properly account for the resonance broadening mechanisms.

The density matrix equation of motion is:

$$\dot{\rho}_{nm} = -\frac{i}{\hbar} [\hat{H}, \rho]_{nm} - \gamma_{nm}(\rho_{nm} - \rho_{eq})$$

where $\rho_{nm}$ are the elements of the density matrix, $n$ and $m$ label the energy levels of the system, $\gamma$ is a phenomenological broadening rate, and $\rho_{eq}$ is the equilibrium population (which will be taken as zero for the off-diagonal elements). The density matrix equation can be solved perturbatively, where the perturbation part of the Hamiltonian is the electric dipole interaction:

$$\hat{V} = -\hat{\mu} \cdot \mathbf{E}(t) = -q\mathbf{r} \cdot \mathbf{E}(t).$$

(6.2)
In order to find the susceptibility from the density matrix, we use, for first order interactions,

\[
< \mu(t) > = \text{tr}(\rho^{(1)} \tilde{\mu}) = \sum_p < \mu(\omega_p) > e^{-i\omega_p t}
\]  

(6.3)

Here \(< \mu(t) >\) is the expectation value of the dipole moment, \(\rho^{(1)}\) is the first order correction to the density matrix, and the right hand side of the equation is the Fourier transform of the dipole moment. The material polarization is given by (in Gaussian units)

\[
P^{(1)}(\omega_p) = N < \tilde{\mu}(\omega_p) > = \varepsilon_0 \chi^{(1)}(\omega_p) E(\omega_p).
\]  

(6.4)

\(N\) is the density of carriers in our semiconductor system. Using these relations and the first order perturbation expression for the density matrix, the linear susceptibility is found to be

\[
\chi^{(1)}_{ij}(\omega_p) = \frac{N}{\varepsilon_0 \hbar} \sum_{nm} (\rho^{(0)}_{nm} - \rho^{(0)}_{mn}) \frac{\mu^i_{mn} \mu^j_{nm}}{(\omega_{nm} - \omega_p) - i\gamma_{nm}}
\]  

(6.5)

where \(i, j\) are Cartesian coordinates. From this expression, we can see that if the population of two levels is the same, then there is no first-order interaction between them. For the common case when all of the population of the system is in the ground state, \(a\), the linear susceptibility can be approximated by

\[
\chi^{(1)}_{ij}(\omega_p) = \frac{N}{\varepsilon_0 \hbar} \frac{\mu^i_{an} \mu^j_{na}}{(\omega_{na} - \omega_p) - i\gamma_{na}},
\]  

(6.6)

where we have dropped a nonresonant term. The dipole matrix elements in the susceptibility come from the off-diagonal (coherence) terms in the density matrix.
The linear susceptibility, $\chi^{(1)}(\omega)$, is related to the refractive index of the material, $n(\omega)$, according to

$$n(\omega) = \sqrt{\varepsilon(\omega)} = \sqrt{1 + 4\pi\chi^{(1)}(\omega)}$$

$$\simeq 1 + 2\pi\chi^{(1)}(\omega) \text{ for } 4\pi\chi^{(1)} \ll 1,$$

where $\varepsilon(\omega)$ is the dielectric constant. The imaginary part of the refractive index gives the absorption coefficient:

$$\alpha = 2n''\omega/c = 4\pi\chi^{(1)''}\omega/c.$$  \hspace{1cm} (6.8)

The derivation of the second order susceptibility, which is the one we want for difference frequency generation, goes along the same lines as the first order susceptibility. The second order perturbation term in the density matrix equation of motion is calculated. It is converted into susceptibility via

$$\mathbf{P}^{(2)}(\omega_p + \omega_q) = N < \bar{\mu}(\omega_p + \omega_q) >$$

$$\mathbf{P}_i^{(2)}(\omega_p + \omega_q) = \sum_{jk} \sum_{pq} \chi^{(2)}_{ijk}(\omega_p + \omega_q, \omega_p, \omega_q) E_j(\omega_p) E_k(\omega_q)$$

The second order susceptibility has quite a few terms. But in our experimental situation, with all three levels nearly in resonance and only the ground state populated, $\chi^{(2)}$ has only one term:

$$\chi^{(2)}_{ijk}(\omega_3, \omega_2, \omega_1) = \frac{N}{2\varepsilon_0\hbar^2} \frac{\mu_{iak} \mu_{jck} \mu_{bka}}{[(\omega_{ca} - \omega_3) - i\gamma_{ca}] [(\omega_{ba} - \omega_2) - i\gamma_{ba}]}$$

where the levels are assigned in fig. 6.1. For intersubband resonances, the dipole matrix elements have only a $z$ component, so $\chi^{(2)}_{zzz}$ is the only nonzero term in $\chi^{(2)}$. 
Since $\mu_{ab} = qz_{ab}$, the susceptibility becomes

$$\chi^{(2)}(\omega_3, \omega_2, \omega_1) = \frac{Ng^3}{2\varepsilon_0 \hbar^2} \frac{z_{ac}z_{cb}z_{ba}}{(\omega_{ca} - \omega_3 - i\gamma_{ca})(\omega_{ba} - \omega_2 - i\gamma_{ba})}$$  \hspace{1cm} (6.12)

### 6.2 Resonant Nonlinear Optics in QWs

The first measurement of the large $\chi^{(2)}$ in intersubband resonances was by Fejer et al. in multiple GaAs/AlGaAs QW [144]. This sample was biased to provide the necessary asymmetry. Although the ISBTs were only singly resonant, a large $\chi^{(2)}$ of 28 nm/V was observed for SHG of CO$_2$ laser light.

The next advancement was to step QWs, in which the asymmetry is provided by the changing composition of the well. The Orsay group used this design in GaAs/Al$_{1-x}$Ga$_x$As/Al$_{1-y}$Ga$_y$As step QWs to achieve two equally spaced ISBTs [145, 146]. This double resonance with SHG enhanced $\chi^{(2)}$ to about 740 nm/V. The Bell Labs group extended this SHG work to AlInAs/GaInAs asymmetric double QWs (ADQWs) [147]. In these doubly resonant systems, they observed a maximum $\chi^{(2)}$ of 75 nm/V.

In order to get to shorter pump wavelengths, such as near-infrared diode wavelengths, it was necessary to change to a material system with a larger band off-
set. The Stanford group published several works on SHG in strained InP/InGaAs ADQWs [148].

Both the Bell Labs and Stanford groups also investigated DFG in ADQWs. The Bell Labs group used CO\textsubscript{2} laser pumps to generate far-infrared output around 60 μm with a $\chi^{(2)}$ of about 1 μm/V [31]. The Stanford group used two beams from a ns optical parametric oscillator at around 1.9 and 2.4 μm to generate difference wavelengths around 10 μm [148]. Their measured $\chi^{(2)}$ was 12 nm/V.

One inherent problem with resonant nonlinear mixing is resonant linear absorption. Vodopoyanov et al. designed a structure which minimizes this problem while maintaining good SHG output power [7]. They demonstrated SHG of 8 μm light in an InGaAs/InAlAs ADQW structure with fundamental and second harmonic ISBTs about 25 meV off resonance. $\chi^{(2)}$ in this configuration was only 9 nm/V. But, they achieved SHG conversion efficiencies as high as 1% thanks to low absorption loss and two innovative phase matching schemes which will be described in detail in sec. 6.3.

The latest trend in intersubband nonlinear optics is monolithically integrating a nonlinear structure with a quantum cascade laser. This approach eliminates three problems which have prevented externally pumped intersubband nonlinear devices from becoming commonly used [6]. First, efficient pump coupling to the nonlinear region is achieved by integrating the nonlinear region with the pump laser waveguide. Second, by proper design, all of the fields and intersubband transitions can be resonant with each other. Third, true phase matching can be achieved by judicious
choice of waveguide modes and width. Finally, resonant linear absorption can be overcome by gain [149].

The first demonstration of intersubband nonlinear mixing inside a QCL was by Owschimikow et al. [5]. They demonstrated SHG and SFG of 7.1 and 9.5 μm fundamentals. They used two types of intersubband nonlinear region: one region which contained both the QC lasing structure and nonlinear intersubband resonances, and one region with intersubband resonances solely for nonlinear generation (fig. 6.2). Further work showed that the former approach works better [6]. The same collaboration has also demonstrated SHG in a self-mode-locked QCL, with increased SHG power [150], and in a single mode distributed-feedback QCL [151]. They have achieved exact phase matching by adjusting the laser ridge width in a structure with a zeroth order fundamental transverse mode and a second order second harmonic transverse mode [152]. Tuning the ridge width causes slight changes to the modal refractive indices. In this structure, they achieved an external linear-to-nonlinear conversion efficiency of 36 mW/W². Finally, this collaboration has demonstrated third harmonic generation at 3.7 μm in a 11.1 μm QCL in which the lasing transition served as one of the three intersubband resonances [153].

As an aside, another group has demonstrated SFG and SHG in QCLs grown on < 111 > oriented GaAs substrates [154,155]. This technique utilizes the bulk nonlinearity of the substrate—not an intersubband-based nonlinearity, but intriguing nonetheless.
Figure 6.2: The two different schemes for incorporating intersubband nonlinear mixing regions into QCLs used in ref. [5]. Later work found that the scheme on the right, with the mixing region and the active region in the same QWs, gave better modal overlap and higher mixing efficiency [6].

Resonant nonlinearities integrated in a QCL waveguide are a promising technique for nonlinear frequency conversion. This method may allow the efficient production of wavelengths which are not otherwise available.

6.3 Phase Matching

When we calculate the power of a nonlinear process, the degree of phase matching must be taken into consideration. Qualitatively, an interaction is phase matched if all the waves in the process remain in the same relative phase throughout the medium. A phase matched interaction is the most efficient type of nonlinear conversion. When an interaction is not phase matched, the waves get out of phase as they propagate, and eventually the power in the generated wave goes back into the pump wave. The phase mismatch, $\Delta k = 2\pi/l_c$, is defined in terms of the coherence length, $l_c$, which is how far the waves can propagate before they get out of phase.

The expression for the phase mismatch, $\Delta k$, can be found by considering the
coupled wave equations. The wave equation in the absence of absorption is [156]:

\[- \nabla^2 \mathbf{E}_n + \frac{\varepsilon^{(1)}(\omega_n)}{c^2} \frac{\partial^2 \mathbf{E}_n}{\partial t^2} = -\frac{4\pi}{c^2} \frac{\partial^2 \mathbf{P}^{NL}_n}{\partial t^2}\]

(6.13)

where \(\varepsilon^{(1)}\) is the (linear) dielectric tensor. For a DFG process in which \(\omega_1 = \omega_3 - \omega_2\), the amplitude of the \(E_1\) wave depends on propagation length \(z\) like \(E_1 \sim e^{ik_1z}\), where the wavevector is

\[k_1 = n_i\omega_1/c\]

(6.14)

and the refractive index is

\[n_i = \sqrt{\varepsilon^{(1)}(\omega_i)}\].

(6.15)

The nonlinear source term on the right hand side of the wave equation goes like \(\mathbf{P}^{NL}_1 \sim e^{i(k_3-k_2)z}\). Plugging these expressions into the wave equation and making the slowly varying amplitude approximation,

\[\left| \frac{d^2 \mathbf{E}_3}{dz^2} \right| \ll \left| k_3 \frac{d \mathbf{E}_3}{dz} \right|,\]

(6.16)

gives

\[\frac{d^2 \mathbf{E}_3}{dz^2} \sim e^{i(k_3-k_1-k_2)z}\].

(6.17)

The phase mismatch for DFG is \(\Delta k = k_3 - k_1 - k_2\).

The phase matching condition, \(\Delta k = 0\), is difficult to achieve due to normal dispersion. Away from absorption resonances, the index of refraction of a material increases with frequency: \(n_3 > n_2 > n_1\) for \(\omega_3 > \omega_2 > \omega_1\). For DFG, the phase
matching condition is

\[ k_3 = k_1 + k_2 \]  \hspace{1cm} (6.18)

\[ n_3 \omega_3 = n_1 \omega_1 + n_2 \omega_2 \]  \hspace{1cm} (6.19)

But since \( \omega_3 = \omega_2 + \omega_1 \) and \( n_3 > n_2, n_1 \), the phase matching condition can not be satisfied in materials with normal dispersion. The same holds true for other nonlinear processes.

The most common way to achieve phase matching takes advantage of birefringence, that is, different refractive indices for different polarizations in a crystal. By properly choosing the polarization of the pump(s) and the angle of the crystal, it is possible to achieve phase matching for a wide range of frequencies. This phase matching technique is not applicable to intersubband-based nonlinear mixers because III-V zincblende semiconductors are not birefringent.

Another technique for phase matching takes advantage of the anomalous dispersion which occurs near an absorption resonance. On the high frequency side of the resonance, the index decreases with increasing frequency. Vodopayanov et al. applied this technique to SHG in an intersubband-based nonlinear device [7]. The phase matching absorption feature was an intersubband resonance in a specially-designed multi-QW grown on both sides of the SHG active region (see fig. 6.3). Using this design, the achieved phase matched SHG in edge-coupled waveguides up to 200 \( \mu m \) long.

We considered using anomalous dispersion for phase matching in our DFG exper-
Figure 6.3: Top: intersubband absorption of the phase matched SHG sample, showing the slightly off-resonance SHG-active intersubband resonances and the phase matching resonance. From [7]. Bottom: schematic diagram of the dispersion due to the phase matching resonance.
iment based on a phonon resonance. However, this technique turns out to be much less appropriate for DFG than for SHG. In SHG, the phase matching condition is met when \( n_\omega = n_{2\omega} \), which can happen with a resonance around \( 1.5\omega \). For DFG in the technologically useful configuration where the difference frequency is quite a bit smaller than either of the pump frequencies, the refractive index at the difference frequency must be quite a bit larger than the refractive index at the pump frequencies in order to achieve phase matching. Unfortunately, a large refractive index only happens near the peak of an absorption (see fig. 6.4). In this case, the phase matched difference frequency is strongly absorbed. If the pump wavelengths are close to the difference frequency wavelength, then phase matching can be achieved further from the absorption resonance. However, there are limited applications for making light at a wavelength only slightly longer than the pumps.

Vodopayanov et al. also demonstrated a novel phase matching technique in a multi-bounce waveguide. In the multi-bounce geometry (fig. 5.1), the length of one bounce through the active region is so short that the interaction is effectively phase matched. But, as the fundamental and second harmonic propagate through the thick substrate, they get out of phase. This work showed that it is possible to get the beams back into the same relative phase at each bounce by adjusting the angle of incidence of the pump beam, as shown in fig. 6.5. For SHG, the phase matching condition in this geometry is:

\[
2k_\omega L = 2k_{2\omega} L + 2m\pi
\]  

(6.20)
Figure 6.4: Top: Schematic diagram of a material dispersion which achieves phase matched DFG for a small difference frequency. Bottom: The associated absorption coefficient.

where $L$ is the length traveled through the substrate between interactions with the active region and $m$ is an integer. The conversion efficiency of this process was five times lower than for their anomalous dispersion waveguide geometry due to a smaller intersubband-active polarization component. However, the output power was higher in the multibounce geometry due to better pump coupling.

We considered using the angled multi-bounce technique for phase matching DFG. For DFG, the phase matching condition is

\[ k_3L = k_1L + 2j\pi = k_2L + 2m\pi, \]  

(6.21)

where $j$ and $m$ are integers. Calculations show that this technique should work for DFG. However, we decided not to pursue it since we needed more than the factor of five or so enhancement of difference frequency output that this technique can provide.
Figure 6.5: The angled multi-bounce phase matching geometry.

The phase matching condition also changes the direction of the emitted difference frequency radiation relative to the pump direction in a bounce geometry, as shown in fig. 6.6. The wavevector of the difference frequency must simultaneously satisfy the phase matching and energy conservation conditions. The phase matching condition means that the x component of the wavevector (that is, the intersubband active part, propagating in the QW plane) must satisfy

$$k_{1x} = k_{3x} - k_{2x} = \left(\frac{\omega_3 n_3}{c} - \frac{\omega_2 n_2}{c}\right) \sin \theta_0$$  \hspace{1cm} (6.22)

where $\theta_0$ is the angle the pump beams make with the growth direction inside the sample. Energy conservation requires that $\omega_1 = \omega_3 - \omega_2$. As a result, the difference frequency beam emerges at a smaller angle from the growth direction, $\theta_1$, given by

$$\sin \theta_1 = \frac{k_{2x}}{k_2} = \frac{(\omega_3 n_3 - \omega_1 n_1)}{\omega_2 n_2} \sin \theta_0$$  \hspace{1cm} (6.23)
Figure 6.6: Diagram of the pump and difference frequency angles for a phase matched interaction in the single-bounce trapezoid geometry.

For my experimental setup described in ch. 7, which is a single-bounce trapezoid polished at 45°, θ₀ = 57.5°, E₁₂ and E₁₃ = 121 to 168 meV, E₂₃ = 14 meV, and θ₁ = 49.2°. Although the change in propagation angle is small inside the sample, refraction at the exit face of the sample means that the difference frequency’s propagation direction is different from the pumps’ direction outside the sample by θ₂ = 29°.

6.4 Power in the Single-Bounce Geometry

For DFG in a single-bounce geometry, I used the expression of Sirtori et al., which describes a setup nearly identical to mine. The pumps are focussed onto the sample, which is polished in a single-bounce trapezoid geometry, and the emitted light is collected and focussed onto a bolometer. The difference frequency power is:

\[
P(\omega₃) = \frac{2n₃(1 - R₃) \sin^6 \theta₁ L² P(\omega₁) P(\omega₂) |\chi(2)|²}{\pi\varepsilon₀ n₁ n₂ c^5 (r₁² + r₂²)} \times \frac{\exp[-(1/2)(\alpha₁ + \alpha₂ + \alpha₃)(L/\cos \theta₁)]}{(\alpha₁ + \alpha₂)²} \times \int_{0}^{\gamma_{max}} \left[ \frac{2J₁(\beta)}{\beta} \right]^2 \sin \gamma d\gamma
\]

(6.24)
where

\[ n_\omega = \text{refractive index at frequency } \omega \]

\[ R_3 = \text{reflection coefficient at } \omega_3 \]

\[ L = \text{active layer thickness} \]

\[ P = \text{power} \]

\[ r_1, r_2 = \text{radii of pump beams} \]

\[ \alpha_\omega = \text{absorption coefficient at frequency } \omega \]

\[ \gamma_{\text{max}} = \text{half the angle subtended by the detector} \]

\[ \beta = k_3 a \sin \gamma \]

\[ 1/a = 1/r_1 + 1/r_2 \]

The second factor describes absorption; the third factor describes diffraction of the DFG beam by the aperture formed by the pump spot. \( \chi_2 \) is given by eq. 6.12. There is no phase matching factor in this expression because the coherence length is greater than the interaction length in a single bounce.

To compute the power I expect from my setup, I extrapolated from the results of Sirtori et al., as shown in table 6.1. However, I never actually tried this experimental configuration with a sensible combination of pump wavelengths and long-pass filter. By the time I had gotten those issues sorted out, we had moved on to a different geometry which was expected to give much more difference frequency power.
Table 6.1: Calculation of DFG power with QCL pumps in single-bounce geometry.

<table>
<thead>
<tr>
<th></th>
<th>Sirtori et al.</th>
<th>This work</th>
</tr>
</thead>
<tbody>
<tr>
<td>L (nm)</td>
<td>775</td>
<td>280</td>
</tr>
<tr>
<td>lc (μm)</td>
<td>100</td>
<td>10</td>
</tr>
<tr>
<td>p_{peak}(ω_1)p_{peak}(ω_2) (mW^2)</td>
<td>10,000</td>
<td>810</td>
</tr>
<tr>
<td>χ^{(2)} (m/V)</td>
<td>1×10^{-6} [157]</td>
<td>4×10^{-6}</td>
</tr>
<tr>
<td>duty cycle</td>
<td>0.5</td>
<td>0.02</td>
</tr>
<tr>
<td>P_{average}(ω_3) (pW)</td>
<td>700</td>
<td>5</td>
</tr>
</tbody>
</table>

6.5 Power in the Non-Phase Matched Edge Coupled Geometry

A significant limitation of the single-bounce geometry is the short interaction length of the pumps with the active region. In order to get a longer interaction length, it is necessary to switch to a different geometry, such as edge coupling, as shown in fig. 6.7. Our collaborator at Texas A&M, Alexey Belyanin, showed that even for a non-phase matched interaction such as ours, this waveguide geometry can have reasonable conversion efficiency if the length of the waveguide is chosen properly.

The difference frequency power is calculated from an ordinary differential equation for the THz field, formed by combining the wave equation with the waveguide mode boundary conditions. The resulting THz power, including pump absorption
and detuning, is

\[ P_1 = \frac{128\pi^5 P_p P_s \Sigma |\chi^{(2)}|^2 FT}{\lambda^2 c n_1 n_2 n_3 \varepsilon_0} \]  

(6.25)

where

\[ \Sigma = \left( \frac{L}{t_{\text{pump}} t_{THz}} \right)^2 \]

= 1/interaction area

\[ P = \text{pump power} \]

\[ F = \frac{1}{\Delta k^2 + (\alpha_{THz} - 2\alpha_p)^2} \left\{ e^{-4\alpha_p l} + e^{-2\alpha l} - 2e^{(-2\alpha_p - \alpha)l} \cos \Delta kl \right\} \]

= absorption factor

\[ l = \text{sample length} \]

\[ T = \text{product of transmission coefficients at all three wavelengths} \]

\[ \chi^{(2)} = \frac{e^3 N z_{12} z_{13} z_{23}}{\varepsilon_0} D \]

\[ N = \text{volume density of electrons in the wells} \]

\[ D = \frac{\sqrt{|\Gamma_{21}|^2 + |\Gamma_{31}|^2}}{\sqrt{|\Gamma_{22}|^2 |\Gamma_{21}|^2 |\Gamma_{31}|^2}} \]

= detuning factor

\[ |\Gamma_{21}|^2 = \gamma_{21}^2 + \Delta \omega_{21}^2 \]

\[ \gamma_{21} = \text{FWHM of } E_{12} \]

\[ \Delta \omega_{21} = \text{detuning of the 1-2 pump frequency from the } E_{12} \text{ resonance} \]

Here, the absorption coefficient is an average over the entire beam: it is the absorption coefficient of eq. 5.6 reduced by the ratio of the active region area to the beam
area. The maximum sample length for efficient non-phase matched conversion is given by $1/\alpha_3$, $1/\alpha_p$, or $1/\Delta k$, whichever is bigger. The next section will describe the experimental setup for the non-phase matched edge coupled geometry and then discuss the DFG power expected.

### 6.6 Experiment

The pump sources were mid-infrared QCLs operating at 7.9 and 10.4 $\mu$m, on loan from Claire Gmachl. They were mounted on the cold finger of a liquid nitrogen cryostat. The QCLs were In soldered to Cu blocks which also served as the positive contact of the lasers. The copper blocks were electrically insulated but thermally connected to the copper laser holders by a thin layer of thermal grease and a glass coverslip soldered to the holder with In. The QCLs were fixed on the laser holder via stainless steel screws which also provided the electrical connection. Each QCL was connected via $\sim 30$ cm of stripline to a Directed Energy PCO-7120 laser diode driver (30 A, 1 MHz), as shown schematically in fig. 6.8. Each PCO current-amplified the pulse from a home-built pulse generator with pulse width variable from 15 to 80 ns. The peak current for each laser was controlled by a DC voltage. The timing between
Figure 6.8: Schematic diagram of the QCL circuit.

The pulse generators was controlled by a SRS DG535 4-channel digital delay/pulse generator with a timing jitter of 5 ps.

The experimental setup for non-phase matched edge coupled DFG is shown in fig. 6.9. The MIR light was output from the cryostat through a 2" diameter, 5 mm thick ZnSe window, anti-reflection (AR) coated from 8 to 12 μm. Each laser's output was collimated by a 0.5" focal length, 0.5" diameter AR coated ZnSe lens. About 65% of the output of each laser was collected. Because the QCL beam diverges ~60° in the growth direction but only ~30° in the QW plane, the collimated beam is elliptical. Fortunately, since the beam size is larger in the growth direction, it is also possible to focus the beam more tightly in that direction, leading to better overlap with the edge-couple ADQW sample. The collimated beams were combined by a 1" diameter ZnSe 50-50 beamsplitter and focussed onto the sample using a 1" diameter, 2.5" focal length ZnSe AR-coated lens. The transmitted and emitted light was collected and focused by two parabolic mirrors onto a Si bolometer.

A bolometer is a very sensitive type of far-infrared detector. It works by measur-
ing the change in resistance of a material (in this case, Si) due to absorption of light. For FIR detection, the detection element must be cooled; ours was operated at 4.2 K. The noise-equivalent power of our bolometer is 2 pW at 200 Hz. The trade-off for this high sensitivity at long wavelengths is that the bolometer is also slow.

One challenge of this experiment is the different timescales of the lasers and the detector. The QCLs overheat with pulses longer than 1 μs and require a duty cycle of less than a few percent. So, they were operated with 80 ns pulses at 250 kHz. The bolometer, on the other hand, responds optimally only up to 200 Hz; its response falls to 92% at 400 Hz. So, the bolometer is effectively an integrating detector. In order to combine the disparate timescales of the source and detector, the 250 kHz QCL pulse train must be chopped at a few hundred Hz, creating a micropulse-macropulse train as shown in fig. 6.10. A mechanical chopper does not work for
Figure 6.10: The 250 kHz pulse train of the lasers is chopped at 200 Hz for detection by the bolometer: schematic diagram.

the low frequency chopping because then the bolometer signal is dominated by the temperature difference between the chopper blade and the optical background. In order to avoid this problem, the low frequency chopping was performed electrically, via the pulse generators. However, it is still possible that the bolometer could detect the temperature change of the lasers on the ms timescale. If DFG were detected, a careful control measurement would have to be made to rule out this source of noise.

Spatial overlap of the two MIR beams was achieved using a pinhole. Timing overlap was monitored by a Boston Electronics PEM-L-3 room temperature MCT detector with < 0.5 ns rise time. The peak powers at the sample, measured with the room temperature MCT detector, were 130 mW for the 7.9 μm laser and 74 mW for the 10.4 μm laser. The focal spot diameter was about 70 μm, so the peak intensities were 3.4 kW/cm² and 1.9 kW/cm².

The sample substrate was thinned to ~ 120 μm and cleaved to a length of ~ 150 μm with optically smooth facets. The sample was mounted on a 150 μm pinhole with the active region in the center to facilitate the positioning and orientation of the
sample. Although the pinhole is significantly larger than the focal spot, the pump beams are probably strongly diffracted by the edge of the sample. This effect was not taken into account in the calculation (sec. 6.5), which assumed plane waves inside the sample slab.

The choice of pump wavelengths is complicated by several factors. The DFG wavelength must avoid the phonon absorption bands in the sample. There must be a sample with intersubband resonances near the pump wavelengths; the best one for this setup was 0404231T. And, there must exist a suitable long-pass filter. Because the pump power is so much stronger than the predicted DFG power, it is necessary to have a long-pass filter with an optical density (OD) of 6 or 7 for the MIR pumps. Apart from a custom dielectric filter, the only appropriate filter material for the QCLs and samples available to us is sapphire. At 4 K, sapphire has a 7.9 \( \mu \text{m} \) absorption coefficient of about 52 cm\(^{-1} \) for the extraordinary polarization. The difference wavelength between our 7.9 and 10.4 \( \mu \text{m} \) pumps is 33 \( \mu \text{m} \). Sapphire's 4 K absorption coefficient at 33 \( \mu \text{m} \) is about 4.6 cm\(^{-1} \). So, I used 2.85 mm of sapphire as a 4 K window inside the bolometer to get an OD of 6.4 for the pumps and 0.6 for the DFG. The bolometer also had a room temperature white polyethylene window and a 4 K white polyethylene window to cut down on blackbody radiation.

The predicted average DFG power for this configuration was 700 pW.
6.7 Assessment

The biggest question in this work is, why was I unable to observe DFG? The predicted DFG power was more than two orders of magnitude bigger than the noise level of the detector, so it should have been observable. It is also possible that the simulation does not adequately represent the system; for example, it does not take into account the detailed profile of the electric field at the active region. Since DFG is a second-order process, small errors in one parameter can lead to large errors in the final result.

Another important question is, what could I change about the experiment in order to detect DFG?

Faster detector? Would a faster detector give a better collection efficiency by being better matched to the short laser pulses? No. The most sensitive fast FIR detector, a Ga:Ge photoconductive detector (4-120 μm), has a time constant of 5 kHz. However, its detectivity is low enough that it has no advantage over the bolometer.

Thicker active region: The DFG power goes as the square of the number of ADQW periods in both the single bounds and waveguide geometries. So, growing 60 ADQWs instead of 20 would give 10 times more DFG power.

More powerful laser: The DFG power goes as the product of the pump powers, so it is tempting to use a more powerful MIR source such as a CO₂ laser. However, it is necessary to limit the average power on the sample to avoid thermal damage.
Table 6.2: Typical time average of the power product for a fixed average pump power. Various combinations of lasers. In arbitrary units.

<table>
<thead>
<tr>
<th>Lasers</th>
<th>Average $P_{DFG}$ (a. u.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.9 and 10.4 μm QCL (15-80 ns pulse)</td>
<td>1</td>
</tr>
<tr>
<td>QCL + CW CO$_2$</td>
<td>0.003</td>
</tr>
<tr>
<td>QCL + RF CO$_2$ (50 μs pulse)</td>
<td>0.3</td>
</tr>
<tr>
<td>RF CO$_2$ + RF CO$_2$</td>
<td>5</td>
</tr>
<tr>
<td>CW CO$_2$ + CW CO$_2$</td>
<td>50</td>
</tr>
<tr>
<td>QCL + Q-switched CO$_2$ (15 ns pulse)</td>
<td>100-200</td>
</tr>
<tr>
<td>Q-switched CO$_2$ + Q-switched CO$_2$</td>
<td>17,000</td>
</tr>
</tbody>
</table>

This requirement means that, for a given total power on the sample, the DFG power is proportional to the peak power of the pumps. When the average pump power on the samples is limited to 0.2 W$^2$ [31], then the average power product (proportional to the DFG power) varies as shown in table 6.2. These values were calculated from typical values of the pulse width and repetition rate of CW, RF, and Q-switched CO$_2$ lasers. From this table, we can see that the important parameters for a high average power product are a high duty cycle and good timing matching between the two pumps.

**Narrower intersubband resonances:** Although improving InAs/AlSb QW material quality is a gradual process, it is interesting to consider what difference a dramatic reduction in intersubband linewidth would make to DFG. If the intersub-
Table 6.3: \(\exp(-\alpha_{THz}l)\) calculated for GaAs at the THz wavelengths and sample path lengths used in this experiment.

<table>
<thead>
<tr>
<th>THz Wavelength</th>
<th>(\alpha (cm^{-1}))</th>
<th>(l = 2) mm</th>
<th>(l = 150) (\mu m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.8 (\mu m) - 10.4 (\mu m = 29) (\mu m)</td>
<td>230</td>
<td>(1\times10^{-20})</td>
<td>0.03</td>
</tr>
<tr>
<td>7.9 (\mu m) - 10.4 (\mu m = 33) (\mu m)</td>
<td>38</td>
<td>(5\times10^{-4})</td>
<td>0.6</td>
</tr>
<tr>
<td>89 (\mu m) (ch. 7)</td>
<td>4</td>
<td>0.4</td>
<td>0.9</td>
</tr>
</tbody>
</table>

band FWHM were 4-5 meV, as is now commonplace in GaAs/AlGaAs QWs, then \(\chi^{(2)}\) would be 20 times higher and the calculated DFG power would be almost 400 times higher in the non phase matched edge-coupled geometry.

**Watch out for THz absorption:** The THz absorption coefficient is very sensitive to the exact wavelength in the regime of this experiment, close to the 35 meV (35 \(\mu m\)) GaAs phonon absorption band. The absorption factor, \(\exp(-\alpha_{THz}l)\), for several THz wavelengths considered in this thesis is give in table 6.3. Free carrier absorption also affects the THz absorption coefficient and is stronger for longer wavelengths. However, since the free carriers are localized in the narrow active region, the DFG beam’s interaction with them is much weaker than its interaction with the GaAs substrate.
Chapter 7

Coherence Oscillations

This chapter describes our second effort to observe THz DFG in ADQWs. As the pump source, we use an ultrafast MIR pulse, as shown in fig. 7.1. Simplistically, this interaction can still be considered difference frequency mixing, since the ultrafast is spectrally broad enough that it contains both $\omega_{12}$ and $\omega_{13}$. However, since the pulse length is comparable to the intersubband coherence time, it is more illuminating to describe this interaction in the language of coherence oscillations and the mathematics of the Bloch equations.

This chapter begins with a description of the density matrix formalism, then reviews some recent work in the busy field of intersubband coherence oscillation emission. Next comes a description of our experimental setup and finally the calculation which encouraged us to try it.

![Diagram](image)

Figure 7.1: Intersubband energy levels and pump pulse spectrum for ultrafast excitation of DFG.
7.1 Density Matrix

The density matrix formalism describes quantum mechanical systems in which the wavefunction is not completely known [158]. Often, this uncertainty in the wavefunction arises from having a large number of particles, in which case it is impossible to know the wavefunction of each particle exactly. The density matrix formalism takes a statistical mechanical average over this large number of particles. Then, the density matrix operator can be treated as a quantum mechanical operator. Note that this statistical mechanical uncertainty is distinct from quantum mechanical uncertainty, in which the observables of the system may be unknown even if the wavefunction is known exactly.

The density matrix operator for a pure state \( |\psi(t)\rangle \) is defined as:

\[
\hat{\rho}(t) = |\psi(t)\rangle \langle \psi(t)|
\]  

(7.1)

A matrix representation of the density operator may be found by expansion of the wavefunction in terms of a basis set:

\[
|\psi(t)\rangle = \sum_n c_n |n\rangle
\]  

(7.2)

\[
\hat{\rho}(t) = \sum_n \sum_m c_n^* c_m |n\rangle \langle m|
\]  

(7.3)

\[
= \sum_n \sum_m |n\rangle \rho_{nm} \langle m|
\]  

(7.4)

The elements of the density matrix, \( \rho_{nm} \), can be considered the ensemble average of a system of \( N \) particles (\( N \) large) which are all prepared in the same state, as far as
possible given the limited knowledge of the system:

\[ \rho_{nm}(t) = \frac{1}{N} \sum_{s=1}^{N} c_m^s(t) c_n(t) \]  \hspace{1cm} (7.5)

The diagonal elements of the density matrix give the population of each state and always add up to 1. The off-diagonal elements describe the coherent interactions between states which give rise to radiation. The dipole moment of a transition \( nn' \) is given by

\[ p(t) = 2q z_{nn'} \mathcal{R}(\rho_{nn'}(t)) \]  \hspace{1cm} (7.6)

The emitted electric field is proportional to the time derivative of the dipole moment:

\[ E \propto \frac{\partial p}{\partial t} \]  \hspace{1cm} (7.7)

The expectation value of an operator \( \hat{O} \) is given by

\[ \langle \hat{O} \rangle = \sum_n \sum_m \rho_{nm} O_{nm} = \text{Tr}[\rho \hat{O}] \]  \hspace{1cm} (7.8)

The time evolution of the density operator is described by the Liouville equation:

\[ \frac{\partial \hat{\rho}}{\partial t} = -\frac{i}{\hbar} [\hat{H}, \hat{\rho}] \]  \hspace{1cm} (7.9)

These density matrix equations are the basis for descriptions of coherent processes in a variety of systems, from nuclear magnetic resonance to coherent optical processes in QWs.

### 7.2 Review: Coherence Oscillations in QWs

THz emission from coherence oscillations in QWs was observed as early as 1992. Roskos et al. used fs interband excitation in a gated GaAs/AlGaAs ADQW, tuned
to create carriers in just one well [159]. When the splitting between the overlapping conduction subbands was small, these charges oscillated back and forth between wells. The THz wave emitted by this time-varying polarization was directly detected by a photoconductive antenna. The frequency of the oscillation was equal to the splitting energy. Plancken et al. observed THz emission from a gated GaAs/AlGaAs single QW [160]. They attributed the THz frequency oscillations to coherence between the HH and LH subbands.

These and other demonstrations have inspired many theoretical studies of quantum control of coherent dynamics in QWs. For example, M. S. C. Luo et al. calculated and demonstrated how a combination of two pulses could enhance or eliminate the coherence oscillations in a DQW structure, depending on the time delay and relative phase of the second pulse [161]. Krause et al. used a genetic algorithm to optimize the pump wavepacket or QW structure to achieve a desired THz emission waveform [162]. These are just two examples of the many proposals to manipulate coherence oscillations.

A wealth of sophisticated coherent control experiments have been performed recently, enabled by ever higher sample quality and shorter laser pulse lengths. For example, C. W. Luo et al. induced Rabi oscillations on the MIR E_{12} transition in a single GaAs QW and measured the emitted transient using electro-optic sampling in ZnTe [163]. By changing the strength of the driving field (i.e., the Rabi frequency), they could control the behavior of the emitted field. More recently, Dynes
et al. demonstrated Stark splitting in a three-level InGaAs/AlInAs ADQW [164]. By strongly driving the 1-2 transition, they induced a splitting in level 2 which was observed as a splitting in the 2-3 intersubband absorption. Simulation of these results gave insight into the intersubband relaxation and dephasing processes. Indeed, that is the goal of many coherence oscillation experiments: to better understand coherent dynamics in semiconductors.

7.3 Proposal of Few-Cycle THz Pulse via Resonant Mixing in QWs

Pestov et al. have calculated the mid- or far-infrared emission from QWs due to coherence oscillations induced by an ultrafast pump [165]. The authors' interest in this scheme is primarily as a novel way to generate ultrashort M/FIR pulses. They show that the coherence oscillations set up by the pump pulse are periodic, with period $T = 4\pi/\sqrt{2\Omega_0^2 + \omega_2^2}$, where $\Omega_0$ is the Rabi frequency. By selecting the correct combination of pulse amplitude and length, the system can be returned to its original state at the end of the pulse, ending the emitted M/FIR pulse. Our interest in this scheme is simply as a THz generation mechanism using the ultrafast MIR optical parametric amplifier source and ADQW samples available to us.

The calculation is based on the Bloch equations. Unlike the previous chapter, in which the density matrix equations of motion were solved perturbatively for an effectively steady state interaction, here the field strength is so strong that the per-
turbative situation is inadequate. Instead, the density matrix equations must be solved exactly. The calculation is based on a three-level scheme, shown in fig. 7.1. The pump pulse energy should lie between the resonance E_{12} and E_{13}, and the pump pulse spectrum should be broader than E_{23}. The density matrix equations of motion are written in the rotating wave approximation as:

\[
\frac{d\sigma_{11}}{dt} = -2\Omega(t)\mathcal{S}(\alpha)
\]

\[
\frac{d\sigma_s}{dt} + \left\{ \gamma + \frac{i\omega_{32}}{2} \right\}\sigma_s = i\frac{\Omega(t)}{2}\left\{ \frac{1}{2}(3\sigma_{11} - 1) - \rho_{32} \right\}
\]

\[
\frac{d\rho_{32}}{dt} + \left\{ \gamma_{32} + i\omega_{32} \right\}\rho_{32} = -i\Omega(t)\sigma_s
\]

(7.10)

where \( \rho_{21} = -\sigma_s^*(t)e^{-i\omega t} \), \( \rho_{31}(t) = \sigma_s(t)e^{-i\omega t} \), \( \rho_{nn}(t) = \sigma_{nn}(t) \), \( \sigma_{22}(t) = \sigma_{33}(t) = \frac{1}{2}\{1 - \sigma_{11}(t)\} \), the Rabi frequency is \( \Omega(t) = zeE/\hbar \), and \( \gamma \) is the relaxation rate from level 2 to 1 and from level 3 to 1. Several simplifying assumptions have also been made: \( \omega = (\omega_{12} + \omega_{13})/2 \); \( d_{12} = d_{13} \) and the dipole moment is real, so the 1-2 and 1-3 Rabi frequencies are equal and real; \( \gamma_{12} = \gamma_{13} = \gamma \); and \( \gamma_{23} = 0 \) during the excitation pulse. The generated M/FIR field is

\[
E(t) = -n^+\frac{\pi\rho_sN_{QW}}{c\varepsilon^{3/2}}\frac{\partial \rho}{\partial t}
\]

(7.11)

where \( n^+ \) is a unit vector in the plane of incidence, \( N_{QW} \) is the number of QWs, \( \varepsilon \) is the dielectric constant in the material, and \( \rho \) is the intersubband dipole moment of one electron (related to \( \rho \) via eq. 7.6). The solution of these coupled equations depends on the pump pulse shape. Calculations based on our experimental parameters, described in the next section, give an average THz power of 1 nW.
7.4 Experiment

The pump source was a Light Conversion TOPAS optical parametric amplifier (OPA). The OPA was pumped by a Clark-MXR CPA 2010 regeneratively amplified Ti:S laser. The CPA emitted transform limited 150 fs pulses centered at 780 nm with 1 mJ/pulse at a repetition rate of 1 kHz. The OPA signal (1.1 - 1.6 μm) and idler (2.7 - 1.5 μm) beams had a combined energy of ~ 200 μJ. For MIR output, the signal and idler beams were difference frequency mixed in a GaSe crystal to provide wavelengths from 4 to 22 μm. The maximum MIR pulse energy was 5.2 μJ at 7.0 μm, decreasing smoothly to, for example, 2.3 μJ at 12 μm.

Three ADQW samples were appropriate for this experiment, as shown in table 7.1. The main condition is that $E_{23}$ should be as small as the spectral width of the OPA pulse, approximately 15 meV. The samples that met this condition are shown in the following table. They all had intersubband resonances spaced 14 meV = 89 μm apart (in addition to a third resonance farther away). Besides matching the pump laser spectrum, this difference energy is well below the phonon band of GaAs (around 35 meV), allowing good transmission of the difference frequency. Each sample had a moderate number of wells, a high density around $1 \times 10^{12}$ cm$^{-2}$ per period, and a pump wavelength long enough to be adequately blocked by sapphire.

The experimental setup is shown in fig. 7.2. The MIR beam from the OPA was filtered by a 2.5 μm dielectric longpass filter to block the signal and idler and a piece of SI-GaAs, oriented at Brewster's angle, to absorb any remaining CPA pump
Table 7.1: Samples used for DFG with the OPA.

<table>
<thead>
<tr>
<th></th>
<th>0306272T</th>
<th>0310081T</th>
<th>0404211T</th>
</tr>
</thead>
<tbody>
<tr>
<td>(N_{QW})</td>
<td>10</td>
<td>10</td>
<td>20</td>
</tr>
<tr>
<td>(\rho_s) per well (cm(^{-2}))</td>
<td>(1.4 \times 10^{12})</td>
<td>(1 \times 10^{12})</td>
<td>(2 \times 10^{12})</td>
</tr>
<tr>
<td>((\omega_{13} + \omega_{12})/2) (eV)</td>
<td>135</td>
<td>161</td>
<td>128</td>
</tr>
<tr>
<td>((\omega_{13} + \omega_{12})/2) ((\mu)m)</td>
<td>9.2</td>
<td>7.7</td>
<td>9.7</td>
</tr>
</tbody>
</table>

light. The beam was focused onto the sample with a 2.5" focal length AR coated ZnSe lens. The focal spot diameter was 180 \(\mu\)m. The sample was in the single-bounce trapezoid geometry. The QW surface was coated with gold and the edges of the sample were polished at 45°. The sample was 0.5 mm thick, 1.3 mm wide on the QW side, and 0.3 mm wide on the substrate side. After polishing, the linear absorption spectrum was measured with the FTIR to confirm that no degradation of the intersubband absorption had taken place. The sample was clamped between polished Al pieces. Since the MIR beam was polarized parallel to the optical table, the sample was oriented with its polished edges perpendicular to the table so that the beam polarization was intersubband-active. The emitted THz light was to be collimated and focused by polyethylene lenses onto a bolometer. The plastic lenses and bolometer were mounted on a rail which could be rotated to find the exact emission angle (see sec. 6.3). The THz beam path, about 30 cm long, was enclosed and purged with nitrogen to reduce atmospheric absorption.

First I tried to observe THz emission. But when that did not work, I decided to
Figure 7.2: The experimental setup for DFG with the OPA pump.
back up a step and confirm that the MIR beam really was interacting with the QWs. First, I tried to use the OPA to measure the linear intersubband absorption spectrum, as shown in fig. 7.4. I put a grating spectrometer and liquid nitrogen cooled MCT detector after the sample to select ~ 2 meV sections of the spectrum. Because the overall intersubband absorption was broader than the OPA spectrum, I set the OPA at several different wavelengths to cover the entire intersubband spectrum. Unfortunately, changing the OPA wavelength also changes the steering slightly due to rotation of the nonlinear crystals in the OPA. I believe this steering change is the reason why I was unable to make any reproducible measurement of the intersubband linear absorption spectrum. In addition, the beam intensity may have been too high, saturating the absorption.

My second approach to confirm the interaction of the pump beam with the QWs was to try to measure absorption saturation of the intersubband resonance at a single wavelength. Absorption saturation is an unambiguous sign of an absorption resonance [93]. Absorption saturation should be easy to observe with a high power laser such as the OPA and a strong linear absorption such as 40% in the ADQW intersubband resonance. However, I was unable to make systematic measurements of the absorption saturation. I varied the intensity on the sample while keeping the intensity on the detector approximately constant by using a set of nearly neutral density filters make of gold films on SI-GaAs. I also varied the intensity by moving the sample along the beam path away from the focal spot. As a reference signal, I
tried both a non-resonant wavelength and rotating the sample to the intersubband-inactive position. Preliminary measurements indicate that the saturation intensity is probably around 7 MW/cm$^{-2}$ (peak intensity), which is the intensity of the unfocused beam about 0.5 m from the OPA. This value is consistent with other absorption saturation measurements of intersubband transitions with ultrafast lasers [166].
Figure 7.3: Setup for intersubband linear absorption spectrum measurement with the OPA.
Figure 7.4: Setup for intersubband absorption saturation measurement.
Chapter 8
Conclusions

In summary, we have presented three spectroscopic studies of narrow gap quantum wells. First, we performed temperature-dependent cyclotron resonance on InAs/AlSb QWs using a THz QCL. Besides being the first spectroscopic application of a THz QCL, this work may enable a new theoretical understanding of the temperature dependence of the \( \mathbf{k} \cdot \mathbf{p} \) bandstructure model.

Second, we performed time-resolved cyclotron resonance in InSb/AlInSb QWs using a free electron laser. This novel spectroscopic technique, coupled with a highly nonparabolic semiconductor such as InSb, allows us to directly monitor carrier cooling and relaxation in a Landau-quantized system. This technique may enable the study of spin-flip relaxation mechanisms, which are currently of interest to the quantum computation community.

Third, we studied intersubband absorption in InAs/AlSb QWs in a wide range of experimental parameters, such as well width, temperature, and density. We reported the highest energy \( E_{12} \) transition to date in this system, 670 meV (1.85 \( \mu \)m) in a 1.8 nm QW. And, our measurements enabled our theoretician collaborators to develop a microscopic theory of many-body effects, which strongly influence the resonance energy and absorption lineshape of the collective intersubband resonance.

Finally, we attempted THz generation via difference frequency mixing in asym-
metric coupled double quantum wells with resonant subbands. We mixed MIR QCLs in a non phase matched edge coupled geometry, and we mixed a single ultrafast MIR pulse in a single-bounce geometry. But, we were unable to observe THz generation.

8.1 Future Work

Each of these projects suggests future work. One of the most unusual things about the THz QCL is its small size compared to other THz sources. So, our group has proposed putting the THz QCL in the bore of a windowless magnet in order to perform, for example, optically detected cyclotron resonance spectroscopy [8], as shown in fig. 8.1. Putting the THz source inside the magnet avoids problems related to windows, such as absorption and reflection.

Time resolved cyclotron resonance could be extended to longer time delays and doped QWs in order to fully explore spin relaxation (fig. 4.6).

The physical understanding of collective effects in intersubband transitions is still incomplete. Our collaborators have simulated the transition of the intersubband resonance from vertex-correction dominated in narrow wells (Fermi edge singularity) to depolarization dominated in wide wells (intersubband plasmon). In between, they predict that the competition of the two effects leads to single particle-like behavior. In nonparabolic semiconductors such as InAs, the transition is accompanied by dramatic changes in the intersubband absorption energy and lineshape, as shown in fig. 8.2. A systematic study of constant-density samples test this simulation and lead to better
Figure 8.1: Experimental setup for optically detected cyclotron resonance with a THz QCL. From [8].
Figure 8.2: Calculated intersubband absorbance for InAs/AlSb QWs, all with a density of $1 \times 10^{12}$ cm$^{-2}$ per well, from [9]. Dotted lines: single particle spectrum; dashed lines: various many-body effects; solid lines: full many-body calculation.

understanding of the physics of intersubband resonances.

THz generation by ultrafast pumping still seems feasible. It should be tried again.
Bibliography


[38] D. Barate, R. Teissier, Y. Wang, and A. N. Baranov, “Short wavelength intersubband emission from InAs/AlSb quantum cascade structures,” Appl. Phys.


[65] W. Xiaoguang, F. M. Peeters, and J. T. Devreese, “Temperature dependence


[71] H. Arimoto, N. Miura, and R. A. Stradling, “Effects of nonequilibrium electron distribution and electron-electron interaction observed in spin-split cyclotron


[111] B. D. McCombe, S. G. Bishop, and R. Kaplan, “Combined Resonance and
Electron $g$ Values in InSb,” Phys. Rev. Lett. 18, 748 (1967).


[137] The number of wells populated was taken to be $\frac{2}{3}n_{Hall}/n_{SdH}$ instead of simply $n_{Hall}/n_{SdH}$. This adjustment ensures that sample 0212062T, in which only one
QW was grown, has only one QW populated. The extra density in the Hall measurement may from parallel conduction in the buffer layers, although the growers say that the GaSb, AlGaSb, and AlSb in the buffer are $p$-type.

[138] For samples in which the intersubband absorption measurement was saturated, that is $T_{\text{min}} < 0.5$, the intersubband matrix element calculated from the integrated absorption was about 40% smaller than for unsaturated measurements. As an aside, it is not clear how $T_{\text{min}}$ can be less than 0.5, since only half of the amplitude of the incident $p$-polarized beam is in the $z$ direction. $T_{\text{min}} < 0.5$ implies that some of the polarization which lies in the QW plane is also being absorbed by the intersubband resonance, which is contrary to the well-established polarization selection rule. However, others have also observed $T_{\text{min}} < 0.5$ in the multibounce geometry.


[157] Note: this is the $\chi^{(2)}$ calculated from the measured DFG power, not the theoretical $\chi^{(2)}$. Comparing this measured $\chi^{(2)}$ to my calculated $\chi^{(2)}$ takes into account the possible collection error in the Bell Labs experiment. So, the difference frequency power in the table is the power I can expect at the correct collection angle.


